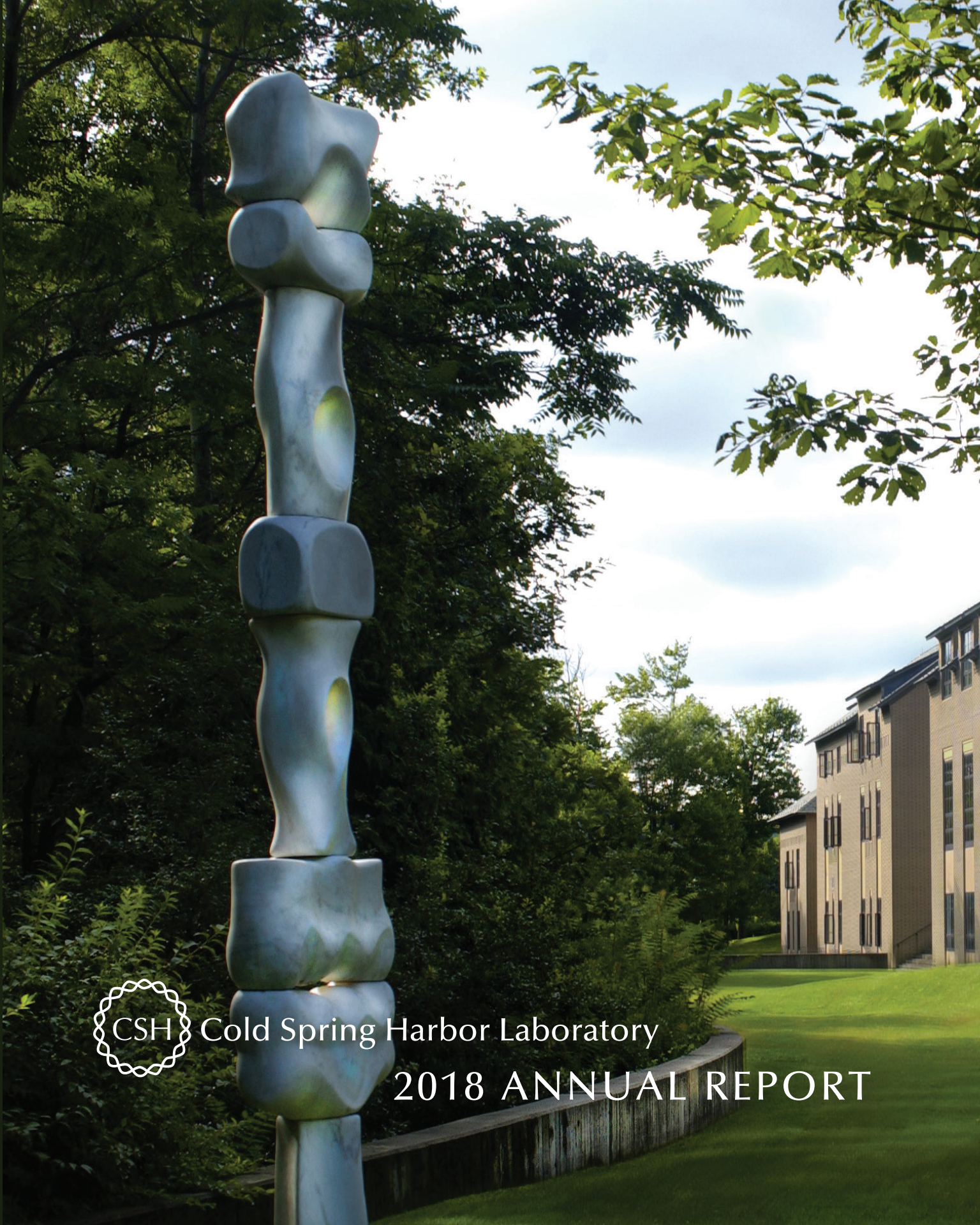
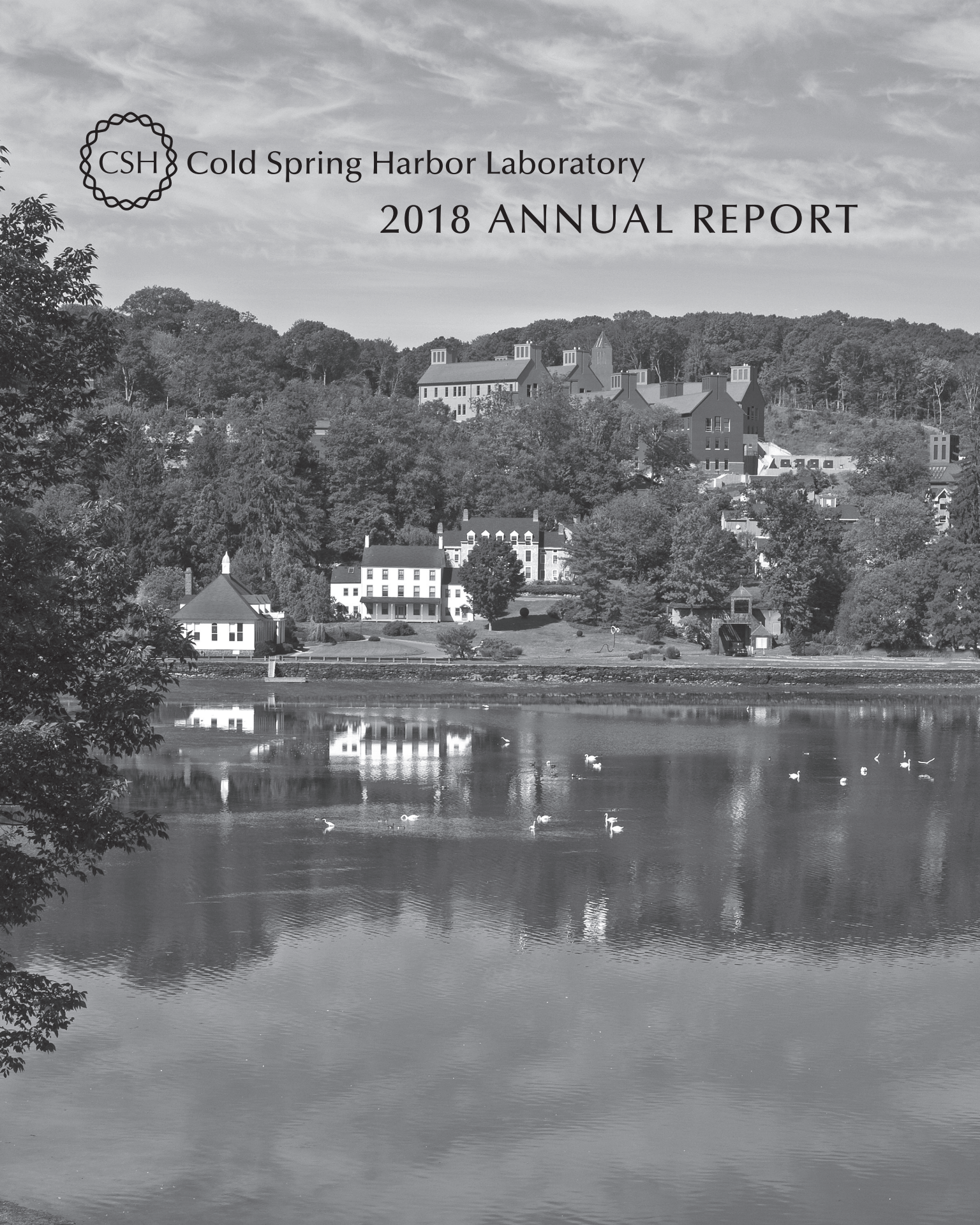


Cold Spring Harbor Laboratory

2018 ANNUAL REPORT


CSH
Cold Spring Harbor Laboratory
2018 ANNUAL REPORT

ANNUAL REPORT 2018

Cold Spring Harbor Laboratory
One Bungtown Road
Cold Spring Harbor, New York 11724

www.cshl.edu

Managing Editors Philip Renna, Dagnia Zeidlickis
Production Editor Kathleen Bubbeo
Copy Editors Dorothy Brown, Joanne Dolce
Production Manager Denise Weiss
Nonscientific Photography Constance Brukin, Gina Motisi
Cover Designer Margot Bennett

Front cover: Photography by Phil Renna

Contents

Officers of the Corporation

Marilyn H. Simons, Ph.D., *Chair*
Charles I. Cogut, *Vice-Chair*
Paul J. Taubman, *Vice-Chair*

Elizabeth McCaul, *Treasurer*
Robert W. Lourie, *Secretary*

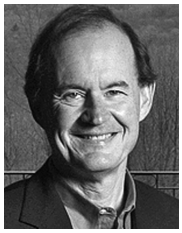
Bruce W. Stillman, Ph.D., *President and CEO*
John P. Tuke, *Chief Operating Officer*

Board of Trustees

Individual Trustees

Lalit R. Bahl, Ph.D.
Senior Research Scientist, Renaissance Technologies Corp.

Joanne Berger-Sweeney, Ph.D.
President, Professor in Neuroscience, Trinity College



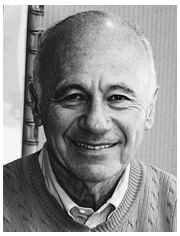
David Boies
Chairman, Boies, Schiller & Flexner LLP

Michael R. Botchan, Ph.D.
Richard and Rhoda Goldman Distinguished Professor of Biochemistry, University of California, Berkeley

Michele Munn Celestino
Laurel Hollow, New York

Charles I. Cogut
Senior M&A Counsel, Simpson Thacher & Bartlett LLP

Elaine Fuchs, Ph.D.
Rebecca C. Lancefield Professor/HHMI Investigator, Rockefeller University/ Howard Hughes Medical Institute



Leo A. Guthart
Founder and CEO, Topspin Partners

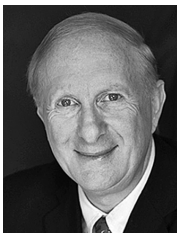
Jeffrey E. Kelter
Senior Partner & Chief Executive Officer, KTR Capital Partners

Laurie J. Landeau, V.M.D.
General Manager, Listowel, Inc.

Robert D. Lindsay
Co-Managing Partner, Lindsay Goldberg

Robert W. Lourie, Ph.D.
Head of Future Research, Renaissance Technologies Corp.

Elizabeth McCaul
Partner-in-Charge, New York Office, Promontory Financial Group, LLC

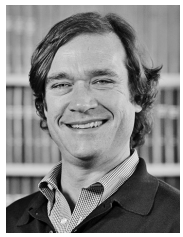


Howard L. Morgan, Ph.D.
Partner, First Round Capital; Director, Idealab

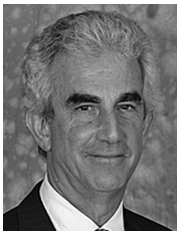
Jamie C. Nicholls
New York, New York

John C. Phelan
Managing Partner, MSD Capital, L.P.

Bruce Ratner
Chairman, Forest City New York



Geoffrey Robertson
Director of Business Assistance, Vermont Sustainable Jobs Fund



George Sard
Chairman and CEO, Sard Verbinnen & Co.

Thomas A. Saunders III
CEO, Ivor & Co., LLC

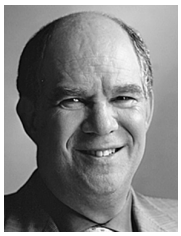
Charles L. Sawyers, M.D.
Investigator, HHMI; Chair, Human Oncology & Pathogenesis Program, MSKCC

Douglas Schloss
President and CEO, Rexford Management

Marilyn H. Simons, Ph.D.
President, The Simons Foundation

Bruce W. Stillman, Ph.D.
President and Chief Executive Officer, Cold Spring Harbor Laboratory

Individual Trustees *(continued)*

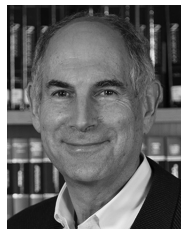


James M. Stone, Ph.D.
Chairman, The Plymouth Rock Company

Karel Svoboda, Ph.D.
Janelia Senior Group Leader, HHMI Janelia Research Campus

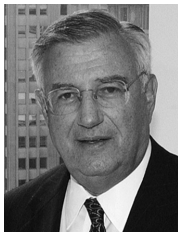
Paul J. Taubman
Chairman and CEO, PJT Partners Inc.

Edward Travaglianti
President, TD Bank Long Island



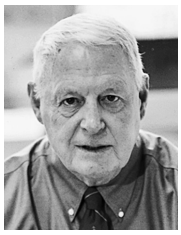
Stuart Weisbrod, Ph.D.
Chief Investment Officer, Iguana Healthcare Partners

George Yancopoulos, M.D., Ph.D.
Chief Scientific Officer & President, Regeneron Pharmaceuticals, Inc.



Roy J. Zuckerberg
Senior Director, Goldman Sachs & Co.

Honorary Trustees



Bayard Clarkson, M.D.
Memorial Sloan Kettering Cancer Center

Charles F. Dolan
Cablevision Systems Corp.

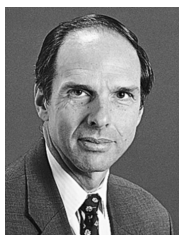
Helen Ann Dolan
Cablevision Systems Corp.

Lola N. Grace
Founder, Middle East Children's Institute

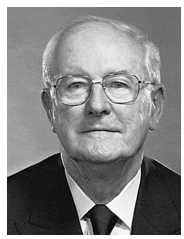
David H. Koch
Koch Industries, Inc.

Mary D. Lindsay
Laurel Hollow, New York

Nancy Abeles Marks
Nancy Marks Interiors and Carl Marks & Co., Inc.



Eduardo G. Mestre
Vice Chairman, Evercore Partners

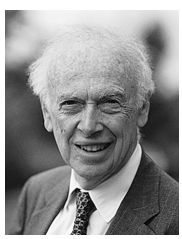


William R. Miller
New York, New York

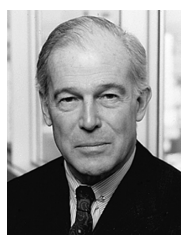
William S. Robertson
Chairman, Robertson Foundation for Government

Alan Seligson
NAK International

James H. Simons, Ph.D.
Chairman, The Simons Foundation



James D. Watson, Ph.D.
Chancellor Emeritus, Cold Spring Harbor Laboratory



Henry Wendt III
Friday Harbor, Washington

CSHL Scientific Advisory Council

The Scientific Advisory Council (SAC) is an external advisory group that advises the senior management of Cold Spring Harbor Laboratory (CSHL) on matters pertaining to science (both current and future), including the development of a research strategy to maintain CSHL as a world leader. The SAC is a nine-member Council including a Chair of Council who is an individual known for scientific breadth and a detailed understanding of research management at the senior management level. The other eight members are world leaders in their respective fields and as such are able to provide advice on the different research areas of the Laboratory.

Frederick W. Alt, Ph.D. (Chair)
HHMI Investigator/Harvard University, Professor & Co-Chief, Division of Molecular Medicine, Children's Hospital Boston

Angelika Amon, Ph.D.
HHMI Investigator/Massachusetts Institute of Technology; Professor, Biology

Joanne Chory, Ph.D.
HHMI Investigator/The Salk Institute for Biological Studies; Professor and Director, Plant Biology Laboratory

Carol W. Greider, Ph.D.
Johns Hopkins School of Medicine, Professor and Director of Molecular Biology and Genetics

Eve Marder, Ph.D.
Brandeis University, Professor of Neuroscience, Biology Department and Volen Center

Markus Meister, Ph.D.
California Institute of Technology, Professor, Molecular and Cellular Biology

Steven Salzberg, Ph.D.
Johns Hopkins University School of Medicine, Professor of Biomedical Engineering, Computer Science, and Biostatistics

Kevan M. Shokat, Ph.D.
HHMI Investigator/University of California, San Francisco; University of California, Berkeley, Professor and Chair, Cellular and Molecular Pharmacology, and Professor of Chemistry

Max S. Wicha, M.D.
University of Michigan, Professor and Director, University of Michigan Comprehensive Cancer Center

Governance

The Laboratory is governed by a Board of Trustees of up to 35 members that meets three or four times a year. Authority to act for the Board of Trustees between meetings is vested in the Executive Committee of the Board of Trustees. The Executive Committee is composed of the Officers of the Board and any other members who may be elected to the Executive Committee by the Board of Trustees. Additional standing and ad hoc committees are appointed by the Board of Trustees to provide guidance and advice in specific areas of the Laboratory's operations.

Representation on the Board of Trustees itself is divided between business and community leaders and scientists from major educational and research institutions.

The Laboratory is chartered as an educational and research institution by the Board of Regents of the Education Department of the State of New York. It is authorized to operate a graduate program under the name "Cold Spring Harbor Laboratory, Watson School of Biological Sciences" and thereat to confer the degrees of Doctor of Philosophy (Ph.D.), Master of Science (M.S.), and Doctor of Science (Sc.D.), Honorary.

It is designated as a "public charity" under Section 501(c)(3) of the Internal Revenue Code.

Committees of the Board 2018

Executive

Marilyn Simons, Chair
Casey Cogut
Robert Lourie
Elizabeth McCaul
Howard Morgan
Bruce Stillman
James Stone
Paul Taubman
Edward Travaglianti
Roy Zuckerberg

Development

Howard Morgan, Chair
David Boies
Casey Cogut
Kristina Davison
Jeff Kelter
Laurie Landeau
Robert Lourie
George Sard
Bruce Stillman
Paul Taubman
John Tuke
Roy Zuckerberg

Commercial Relations

James Stone, Chair
Casey Cogut
Rob Martienssen
Bruce Stillman
John Tuke
Andrew Whiteley

Audit

Leo Guthart, Chair
Elizabeth McCaul
Douglas Schloss
Edward Travaglianti

Investment

John Phelan, Chair
Jamie Nicholls
Thomas Purcell
Remy Trafelet
Edward Travaglianti
John Tuke

Finance

Elizabeth McCaul, Chair
Jamie Nicholls
Lari Russo
Douglas Schloss
Marilyn Simons
Edward Travaglianti
John Tuke

Academic Affairs

James Stone, Chair
Lalit Bahl
Michael Botchan
Elaine Fuchs
Laurie Landeau
Robert Lourie
Bruce Stillman
Karel Svoboda

Business Development

Bryan (Hunt) Lawrence
Frank Brosens
William Spire
James Stone
Paul Taubman
Stuart Weisbrod
Andrew Whiteley

Nominating

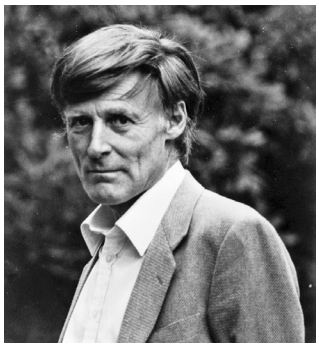
Paul Taubman, Chair
Casey Cogut
Leo Guthart
Robert Lindsay
Elizabeth McCaul
Bruce Ratner
Thomas Saunders
Douglas Schloss
Bruce Stillman
George Yancopoulos
Roy Zuckerberg

DNA Learning Center

Laurie Landeau, Chair
Casey Cogut
Kristina Davison
Jeffrey Kelter
David Micklos
Ed Rover
Marilyn Simons
Bruce Stillman

Robertson Research Fund (Board of Trustees)

Walter Goldschmidts
William Gridley
Leo Guthart
Andrew McKibben
Walter Meier
Geoff Robertson
Marilyn Simons
Bruce Stillman
Edward Travaglianti



John Hugh Cairns (1922–2018)

John Hugh Cairns who died on November 12, 2018, just a few days short of his 96th birthday, was the first scientist to use autoradiography to visualize replicating DNA. He holds a special place in the history of Cold Spring Harbor Laboratory as the first director of the newly reorganized institution. Later John Cairns became interested in DNA repair, mutations, and cancer and suggested provocative theories such as "adaptive mutagenesis" and immortal DNA strands in stem cells.

Cairns was born on November 22, 1922, in the Bishop's Room in the Master's Lodge of Balliol College. Cairns' mother was the youngest of the seven daughters of A.L. Smith, the Master of Balliol, and so was born in his maternal grandfather's house. John's father, Sir Hugh Cairns, was Australian and after training at Oxford became England's preeminent neurosurgeon.

John Cairns began by following in his father's footsteps—studying medicine at Oxford and being admitted to Balliol. In 1947, he joined the Department of Clinical Pathology at the Radcliffe Infirmary, Oxford. There he published what he regarded as his first scientific paper, a study of the origin of patients' penicillin-resistant staphylococci. More importantly, it was there that he met and married Elfie, a marriage that lasted more than 70 years.

Although the war was over, doctors were still being drafted; however, Cairns was able to take up a studentship offered by the Colonial Office to learn about viruses in the British colonies. In 1950, the Cairns family moved to Melbourne where John took up his studentship in Macfarlane Burnet's laboratory. In 1952, the Cairns moved to the Virus Research Institute in Entebbe, Uganda. His two years there were not productive, perhaps best illustrated by the fact that the Institute "… was known locally as Hangover Hill (which, in hard-drinking East Africa, was something of an accolade)" (JH Cairns, unpubl.).

In 1954, the Cairns returned to Canberra, and John joined the Australian National University, working on the influenza virus under the chairmanship of the famous Australian virologist Frank Fenner. In 1957, he was awarded a Rockefeller Foundation fellowship to learn about cell culture in Renato Dulbecco's laboratory at Caltech. There he shared a house with Howard Temin, Matt Meselson, and Jan Drake. What an eye-opener it must have been, to move from a conventional microbiology laboratory to one working at the leading edge of molecular biology! Weekends in the desert with Max and Manny Delbrück completed the experience.

After four months, John returned to Canberra and continued working on virus replication. He used autoradiography to analyze the replication of vaccinia virus in tissue culture cells, research that stood him good stead for the later work for which he is remembered best.

John became eligible for a sabbatical year in 1960 and, perhaps influenced by his experience at Caltech, spent it with Al Hershey at Cold Spring Harbor. John used DNA from a bacterial virus called bacteriophage T2 to carry out a simple experiment that for the first time revealed an intact DNA molecule. John came to the not surprising conclusion that there was no need to assume that the T2 DNA molecule "was anything other than an uncomplicated double helix" (Cairns 1961). In hindsight, it seems remarkable that such an experiment needed to be done.

Like so many visitors to Cold Spring Harbor, John was captivated by the atmosphere of the Laboratory where "… just about everyone at Cold Spring Harbor was embroiled in the new science of molecular biology" (JH Cairns, unpubl.), and that there were endless discussions and arguments about every aspect of the science. But the year was over all too quickly and the family returned to Canberra.

Back at the Australian National University, John continued studying DNA molecules and chromosomes using autoradiography, the technique in which DNA was labeled with the then newly developed radioisotopes in thymidine, a precursor to DNA synthesis, and visualized on a thin film after many months of exposure. The result was a series of elegant papers in which simple experiments—labeling DNA with H3 thymidine, carefully extracting the DNA, and preparing autoradiographs—were followed by elegant analysis and interpretation. His most significant study was of *Escherichia coli* chromosomes, showing that the DNA in bacteria duplicates at a fork and that there appeared to be two forks per molecule that started from a single origin.

John's participation in the 1962 Cold Spring Harbor Laboratory Symposium on Quantitative Biology was notable not so much for his presentation but because of a conversation he had with Paul Margolin, a scientist in the Biological Laboratory. John was rather taken aback by Margolin's suggestion that John apply for the position of director of a new entity about to come into existence, the Cold Spring Harbor Laboratory of Quantitative Biology.

At the time John was unaware of the turmoil at Cold Spring Harbor. There were two institutions on the campus, the Carnegie Institute of Washington's Department of Genetics and the Biological Laboratory. Milislav Demerec was the director of both, but on his retirement, Carnegie decided to close the genetics department, taking with them their considerable endowment support for science. This heightened the difficulties of the Biological Laboratory, which was already in desperate financial straits. After considerable machinations and maneuvering, a new institution, Cold Spring Harbor Laboratory (CSHL), was formed by the fusion of the Biological Laboratory with the remnants of the Carnegie Department of Genetics. Unfortunately, this did not improve the financial situation; the new institution would be close to bankruptcy.

In late 1962, an official invitation arrived from Ed Tatum, chair of the CSHL Board of Trustees and later to become John's nemesis. Still ignorant of the deplorable state of affairs at Cold Spring Harbor, John accepted and arrived in New York in May 1963. Those who know Cold Spring Harbor Laboratory today will have difficulty in appreciating the CSHL of the 1960s. John found that there was only enough cash on hand to pay the staff wages for two weeks. The National Science Foundation funded a review of the physical plant that concluded that "If the Laboratory disregards the accumulated problems of maintenance, it can survive about one year. If it attempts to maintain its buildings in the manner to which they should have been accustomed, it will go out of business almost at once" (emphasis in original).

The Laboratory was saved by the income from the sales of the volumes of the 1961 and 1962 "Cold Spring Harbor Symposia on Quantitative Biology," and John increased the number of meetings and courses, which provided a steady and substantial revenue stream. Of the Symposia John organized, the most notable was held in 1966 on *The Genetic Code*; intellectually stimulating and financially rewarding, it brought in $107,000. Grant money increased as well, but staff left CSHL, taking grants with them, and by 1967, there were only three research groups at CSHL. John had had enough of struggling to find funds, with little or no help from the trustees, especially Ed Tatum, the chair. In 1966, John announced his intention to step down but it was two years

before he could do so and return to research. Jim Watson, who was a trustee, prepared a report on the state of the Laboratory for the trustees, and on presenting it, Watson was invited to become director. He accepted, taking up his appointment in 1968 while John acted as Associate Director.

Despite all the obstacles, John had begun to turn things around and left the Laboratory with a balanced budget and a small cash reserve. As Watson wrote in his first annual report: "… let me express the strong debt which everyone connected with Cold Spring Harbor owes to John Cairns. … he inherited a terrible mess … But with much devotion, intelligence, and great indulgence of his free time, he again made the Laboratory a going concern. Without his achievement, I would have never taken over as director." Watson should have added praise for Elfie and the children, who underwent trials and tribulations as they helped John.

Watson acted rapidly on John's behalf and in 1969 John received an American Cancer Society Lifetime Award. John returned to the laboratory and with his assistant Paula De Lucia carried out his most significant research project. There was increasing evidence that the DNA polymerase discovered by Arthur Kornberg (for which he was awarded a Nobel Prize) was not the polymerase that replicated DNA. John and Paula mutagenized *E. coli* and assayed colonies for their DNA polymerase activity, searching for a colony that had no detectable DNA polymerase activity but could still grow. Paula had to test 3,486 colonies before she determined that the 3487th colony had no detectable DNA polymerase activity but could grow. The mutant was called *polA*, pronounced *Paula.* It was later shown that the Kornberg polymerase was a DNA repair enzyme, and it was Arthur's son Tom who isolated the polymerase that actually replicates the bacterial DNA.

In 1973, Michael Stoker invited John to run the Imperial Cancer Research Fund (ICRF) unit at Burtonhole Lane, Mill Hill, London. Much as he had done at CSHL, John transformed what had been a rather backwater outpost of the mother ship in Lincoln's Inn Fields into a dynamic laboratory, focusing on DNA mutations and repair and on developmental biology. Notable members of the laboratory included Julian Gross, Brigid Hogan, David Ish-Horowicz, Tomas Lindahl, and Jonathan Slack. John's return to research was clearly a happy one, and with Leona Samson, he discovered a new DNA repair pathway in *E. coli.* He also proposed the "immortal strand" hypothesis—that is, that stem cells could protect against accumulating mutations if, upon chromosome segregation, during mitosis the stem cell always inherited the old DNA template for every chromosome, thereby the possibility of accumulating mutations in stem cells.

Alas, John's pleasure at being back in the laboratory was short-lived. Britain's Medical Research Council (MRC) held the lease to the Burtonhole Lane building that would expire in 1986, and the MRC made plain that it would not be renewed. Discussions took place about what should be done and John resigned, hoping that this would force the appointment of a younger head of the laboratory who could keep the Burtonhole Lane group together. Alas, it was not to be and Cairns' highly successful team was broken up.

While at ICRF, John had become interested in the broader public health issues relating to cancer, and in 1978 he published an important and successful book, *Cancer: Science and Society*. In 1980, Howard Hiatt persuaded John to join the Harvard School of Public Health, and the Cairns family, yet again, moved across the Atlantic. And, once again, John found that he had moved to a position of financial insecurity. Nevertheless, he spent 10 years at Harvard, mainly reading in the Countway Library and training for marathons. (In 1982, he won his age group in the Boston Marathon.) It was at Harvard in 1988 that John published what proved to be his most controversial research. This seemed to show that populations of nongrowing cells under selection that is not in itself mutagenic preferentially accumulate the appropriate mutations without at the same time accumulating unselected mutations. That is, the bacteria were able to direct mutations, counter to the well-accepted Darwinian natural selection of preexisting mutations. This aroused a storm of criticism that he was denying the role of natural selection and invoking Lamarckism.

In 1991, John and Elfie moved back to the United Kingdom, to Wilcote, a lovely, very small village outside Oxford.

Although now fully retired, John was welcomed to the University's Clinical Trial Service Unit by Richard Peto, Richard Doll, and Rory Collins, providing him with an academic affiliation. John worked on his book *Matters of Life and Death: Perspectives on Public Health, Molecular Biology, Cancer, and the Prospects for the Human Race,* which was published in 1997, and revisited his earlier interests of mutational mechanisms and stem cell kinetics. They moved to London in 2006 but returned to Oxford in 2013. John's health deteriorated during 2018 and he died that November.

John was a remarkable man. He had an incisive and analytical mind; he was witty and a raconteur; and he wrote with style and an economy of words that may have owed something to Al Hershey. CSHL owes a great debt to John Cairns and his family. It is no exaggeration to assert that the newborn Cold Spring Harbor Laboratory might have expired within one or two years had not John persevered with what must have seemed an impossible and thankless task. That he did so is why the Laboratory is here today.

Jan A. Witkowski

Reference

Cairns J. 1961. An estimate of the length of the DNA molecule of T2 bacteriophage by autoradiography. *J Mol Biol* **3:** 756–761.

PRESIDENT'S REPORT

As we advance toward the middle of the twenty-first century, humanity faces an existential challenge: figuring out how to feed the world's rapidly growing population in the face of climate change and the increasingly limited availability of key nutrients and suitable land for farming. We need solutions that are local, national, and global to increase food production in spite of these threats to our livelihood, and to do so on what is an ever-smaller amount of suitable land. Cold Spring Harbor Laboratory's (CSHL) biologists pursuing fundamental discovery research are critical to these solutions.

The Challenges Defined

It's important to start with some facts that precisely define the challenge. By 2050, the world's population is expected to reach almost 10 billion, from 7 billion today. According to the United Nations, nearly 821 million people suffered from chronic food deprivation in 2017, up from 804 million in 2016. Those numbers are expected to continue to rise. Meanwhile, climate change and related extreme events have been a major cause of food shortages and damage to the world's ecosystems.

On top of this upheaval, overfarming and other factors have degraded much of the world's soil, leaving it too salty for agricultural use. Nitrogen, an essential nutrient for crops that is at once dangerously scarce in its availability as a fertilizer and is also detrimental to the environment and to humans when not managed efficiently, has been polluting our land, air, and water. Phosphorus, another crucial nutrient for plants and for our survival, is simultaneously in short supply where it is needed and harming the environment where it is found in excess.

Of course, there are political, infrastructural, economic, and social factors contributing to the globe's food shortages. The pressures are coming from multiple points. But science—and specifically, the kind of fundamental discovery research that is conducted at CSHL—may very well be our best hope for meeting and beating these challenges.

Biological Roots of Agriculture

Plant biology research at CSHL explores fundamental mechanisms in plant development and genetics, and this kind of discovery research underlies all innovations that have the potential to combat nutrient shortages and poor adaptation of plants to changing environments and to improve crop productivity to feed the growing population. CSHL scientists have made and are continuing to make pioneering discoveries that help explain how plants adapt to different and changing environments, and how we can influence their ability to do so; they are exploring the possibilities of adapting nontraditional, wild, or "orphan" crops for large-scale agriculture, expanding our selection of major food sources; and they are making impressive headway in the quest to increase yields of existing staple crops. Their experimentation can lead to remarkable results and real-world impacts.

Groundbreaking discoveries in genetics and plant biology have been taking place in Cold Spring Harbor for well over a century. Charles Davenport, funded by the Carnegie Institution, established a study center for genetics here in 1904 that soon produced truly impactful results. Botanist George Harrison Shull's fundamental work here culminated in a 1908 paper, "The Composition of a Field of Maize," which led to the use of heterosis or "hybrid vigor"—the tendency of a crossbred individual to show qualities far superior to both parents—in plant breeding. When later

introduced to farmers, this technology dramatically increased maize crop yield for both animal and human consumption, turning corn into this country's most valuable crop.

In the 1940s, Barbara McClintock discovered that genetic information is not stationary by identifying DNA elements in the genome of corn that she called "transposable elements." These elements could jump around the genome—and when landing near a gene, they take control of that gene by affecting the surrounding genetic material. McClintock's discoveries laid the groundwork for understanding how genes can be controlled in plants, and she received the 1983 Nobel Prize for Physiology or Medicine for her pioneering work.

Shull and McClintock's research was foundational to the field of plant biology and revolutionary in its day. Today, it is thrilling to see the next great plant science revolution taking place at the Laboratory. We have begun to understand and map the genomic sequences of plants and to use those insights and knowledge to make crops and other plants heartier, more adaptable, and more robust—all at an unprecedented rate of discovery. Indeed, a second green revolution is beginning.

The Yield of Genomic Data

We are now in the age of genome biology—and that includes the sequencing of plant genomes. The first plant genome to be sequenced was that of the popular model plant organism, *Arabidopsis thaliana*, and was accomplished with substantial input from CSHL scientists Dick McCombie and Rob Martienssen. More recently, genome sequencing of many plant genomes, particularly the staple crops like sorghum, maize, wheat, and rice that are critical to food security in countries across the globe, has helped plan new approaches to crop improvement.

Exemplifying CSHL's open and collaborative spirit, Dick McCombie and plant biologist Doreen Ware are committed to generating genome sequences and building genomic databases that can be utilized by scientists anywhere, in the interest of improving the yield and range of modern plants. For example, McCombie has been involved in a multinational program to analyze the genome of an important food legume crop that is grown in the semi-arid regions of the world.

By comparing the genomes of plants of the same lineage, Ware is part of a large international consortium gaining key insights into gene conservation and loss, gene regulation, and gene expression of these plants, including studies that compare 13 varieties of rice. In other work, the sequencing of the genome of sorghum plants and identifying mutations in a gene that regulates hormone production have made it possible to triple the number of grains each plant produces. Also, in 2018, genome analysis of *Arabidopsis* uncovered a set of gene regulators that are key to understanding nitrogen metabolism.

Discovering biological principles related to how plants reproduce, how they develop, and how they control their shape and yield of flowers and fruit is a core component of the plant biology group at CSHL. Rob Martienssen has done leading-edge work in plants' "secondary" genetic code—the epigenome. This work by Martienssen, who was mentored by and worked closely with McClintock for his first two years at CSHL, starting in 1989, has far-reaching implications for plant breeding, bioenergy, and human health. In 2018, he was honored with the Barbara McClintock Prize for Plant Genetics and Genome Studies at the Maize Genetics Conference in Saint-Malo, France. This award recognized decades of work piecing together an understanding of how the several interacting layers of epigenetic signals in plants function, and how they work together to regulate gene activity.

For example, by studying modifications in the DNA of cloned palm oil plants, Martienssen and colleagues in the United States, Malaysia, and Indonesia developed a simple and inexpensive DNA test that will greatly improve the productivity of farms that produce palm oil, thereby reducing the need for farmers to acquire more land by invading nearby rain forests. In addition to important, fundamental discoveries about how genes are controlled in plants, Martienssen has

identified methods that can generate plant embryos without fertilization, thereby limiting the variegation that occurs with classical plant breeding.

This research is helping us understand how to increase crop yields and make plants more adaptable to changing environmental conditions. In other projects, Martienssen is pioneering research into using a diminutive but plentiful aquatic plant called duckweed as a sustainable source of biofuel. This contributes to tackling another major challenge of our time—humanity's dependence on fossil fuels.

Improvements in yield of major crops also come from discovery research aimed at understanding how stem cells in the growing tips of plants, called meristems, develop into various tissues. David Jackson, who together with Martienssen has built up the CSHL plant biology group, studies how the shoot meristem cells differentiate into the inflorescences that produce flowers and ultimately seeds, which in maize are the kernels on the cob. Jackson has shown that regulatory proteins can move between cells in the plant, an unexpected finding.

Jackson has also studied the function of signaling proteins that, when manipulated, can increase the size of the meristem, ultimately resulting in more rows of maize kernels on the cob, hence increasing yield per plant. His insights into stem cell biology and cell-to-cell signaling come from the powerful genetic tools in maize, as well as his studies on the model plant *Arabidopsis*. The results of his work are readily translated into crop species because of the conserved pathways that determine plant development.

Nurturing New Technologies

It is a testament to the Laboratory's sustained reputation as a deeply supportive and welcoming institution that so many of our scientists come here and stay here as a result not only of the science being done, but of the relationships they form with their colleagues, mentors, and students. Martienssen admired and built upon the seminal work of McClintock, and in turn, he mentored now-Professor Zach Lippman when Lippman first came to CSHL as a graduate student in 2000.

Lippman studies the genes that are responsible for the number and location of inflorescences, or flower-producing branches, on tomato plants, and he has used the CRISPR-Cas9 gene editing tool to make extraordinary advances in fine-tuning plant architecture and flower production. Although mechanistic discovery science is the bedrock of CSHL's scientific program, Lippman's work is a shining example of how the work we do here to understand plant biology can lead—and lead quickly—to extraordinary translational benefits.

Lippman has studied the genetics of hybrid vigor, following up on Shull's ideas of more than a century ago. Lippman has shown that the levels of hormones in a plant and the exact makeup or ratio of multiple genes that determine these hormone levels contribute to hybrid vigor, or heterosis. In addition, Lippman, in a collaboration with Jackson, has used CRISPR-Cas9 gene editing to make a series of mutations in the DNA that controls the expression of genes, called a promoter. By making many such promoter mutations in genes that control expression of hormones, plants can be created with specific characteristics like fruit number and size. Other promoter mutations could help plants better adjust to growing environments in specific locations, thereby increasing the potential geographic area in which crops can grow.

CSHL has licensed this Promoter Fine Tuning technology and other inventions of our plant biology group to the plant breeding industry. Whereas commercial production of soybean, corn, wheat, tomato, and other crops has historically been constrained by the availability of just a few standard seed varieties (varieties that may not be ideally suited to the growing conditions where they are planted), the new technology will eventually offer farmers the ability to purchase seeds customized to the specific growing conditions of their fields. Seeds for plants that are better suited to the currently available growing locations and conditions will mean increases in plant yields or

the growth of crops that are tuned to use water and nitrogen more efficiently. And as the climate continues to change, we can much more swiftly and easily adapt crop planting to the growing conditions that arise.

Indeed, thanks to fundamental discovery research in this genome age, the rate of discovery in plant biology has increased and new technologies can be implemented relatively quickly. As the leap from discovery science to field trials has become increasingly less expensive, the process has also been democratized. Small companies and developing nations are able to use technologies to augment the traditional, large multinational plant industry offerings—with potentially transformative positive implications for future food security.

Lippman has also used CRISPR technology to tweak the genes of the groundcherry, a small and historically wild relative of the tomato, with the hope of rendering it suitable for larger-scale agriculture. His work illustrates the possibility that new strains of traditionally wild crops—yet untapped food sources—could be bred relatively quickly for large-scale crop farming.

The most recent faculty member in the productive and influential CSHL plant biology program, Ullas Pedmale, is helping us understand how plants modulate their own growth and development by sensing the environment around them. By studying how plants use information about both biotic and abiotic aspects of their environment—temperature, light quality and quantity, herbivores, pathogens, water availability, and more—to direct their own growth, he is investigating ways to develop crops that can adapt efficiently to unfavorable environments, like those that will inevitably arise from our changing climate, without decreases in yield.

A Way Forward

The challenge of growing crops to feed billions of people, on a total amount of suitable land that is rapidly diminishing, is formidable. But the basic research I have described, and the research that we will continue to pursue, gives us reason to believe that we can help plants adapt to conditions like hotter temperatures, soil salinity, infections, and water shortages. Since the institution's inception, CSHL and its scientists have dedicated themselves to the fundamental discovery research that is essential to tackling some of the world's most complex and intimidating problems. For the past 129 years, CSHL scientists have made major advances in many research areas, but particularly notable has been the impact of CSHL in plant biology. Thus, as challenges of population growth and climate change face us, we look toward the new green revolution to help address such issues.

Bruce Stillman, Ph.D., F.R.S.
President and Chief Executive Officer

Highlights of the Year

Research

Hundreds of scientists working in Cold Spring Harbor Laboratory's 50-plus laboratory groups contributed to research that in 2018 was published in the world's major research journals. Their efforts reflect the full spectrum of this institution's scientific activity in Cancer, Neuroscience, Plant Biology, Quantitative Biology, and Genomics. The following is a sampling of the year's important findings.

Doubling the Number of Grains in Sorghum

D. Ware

A simple genetic modification can triple the grain number of sorghum, a drought-tolerant plant that is an important source of food, animal feed, and biofuel. By lowering the level of a key hormone, the team (led by Doreen Ware in collaboration with her USDA colleague Zhangou Xin of Lubbock, Texas) generated more flowers and more seeds.

Sorghum's grains are produced in clusters of flowers that develop from a branched structure at the top of the plant called a panicle. Each panicle can produce hundreds of flowers. There are two types of flowers, and usually only one of these, known as the sessile spikelet, is fertile. The other flower type, called pedicellate spikelets, does not make seeds. In modified plants, however, *both* sessile and pedicellate spikelets produced seeds, doubling each plant's grain number.

By sequencing the genomes of the modified plants, Ware's team found key mutations affected a gene that regulates hormone production. Plants carrying the mutation produce abnormally low levels of a development-regulating hormone called jasmonic acid, particularly during flower development. Through subsequent experiments, the team learned that jasmonic acid prevents pedicellate spikelets from producing seeds.

The team now wants to understand whether the strategy can be applied to increase grain production in related plants that are vital in the global food supply, such as rice, corn, and wheat.

Unique Communication Strategy in Pathway Controlling Plant Growth

D. Jackson

CSHL plant geneticists identified a receptor protein on stem cells involved in plant development that can issue different instructions about how to grow depending on what peptide activates it. The findings by David Jackson and colleagues have implications for efforts to boost yields of food crops such as corn and rice.

Plant growth and development depend on structures called meristems—reservoirs that contain stem cells. When prompted by peptide signals, stem cells develop into any of the plant's organs—roots, leaves, or flowers, for example. These signals work like a key (the peptide) fitting into a lock (the receptor protein) on the surface of a cell. The lock opens, stimulating a messenger inside the cell. The messenger carries instructions for the cell to grow into a leaf or flower cell or even stop growing altogether. Conventionally, one or more peptides fit into each receptor to stimulate a single type of messenger.

In 2001, Jackson and colleagues discovered that a receptor protein called FEA2 can stimulate one of two distinct messengers into the cell depending on which of two peptides, ZmCLE7 or ZmFCP1, switches it on. Receptors such as FEA2 that stimulate more than one messenger are rare. This is the first one discovered that plays a role in crop production.

Jackson and his team believe that FEA2 is bound to two different co-receptors, each of which acts as the "lock" for one of the two peptide "keys." Future research will explore how the two different peptide signals are translated by FEA2 into distinct messenger outputs.

A Way to Make Prostate Cancer Cells Run Out of Energy and Die

Cells lacking the tumor-suppressor protein PTEN—a feature of many cancers—have been determined by Lloyd Trotman to be vulnerable to drugs that impair their energy-producing mitochondria. Unlike normal cells, cells without PTEN seem driven to preserve their mitochondria at all costs. Trotman and colleagues found that when such cells are treated with certain mitochondrial inhibitors, they consume vast quantities of glucose to fuel these efforts. As a result, they quickly run out of energy and die.

L. Trotman

Some mitochondrial inhibitors, including the diabetes medication metformin, are being tested for their ability to treat cancer. Trotman's findings suggest that such drugs have the potential to eliminate cancer cells at doses that leave healthy cells intact, particularly in cells that have compromised PTEN function. "The hope is that carefully timed administration of these drugs can generate a much better window of selective killing," Trotman says.

Two related compounds emerged from a screen performed by the team. Both killed cells missing PTEN and another tumor suppressor, p53. Loss of these together is common among men with advanced prostate cancer that is also highly metastatic, a particularly lethal form of prostate cancer.

The two drugs had little effect on nearly identical cells with functional PTEN. One, rotenone, is a known mitochondrial inhibitor. Trotman established that the second compound, deguelin, shuts down mitochondrial function just as well in cells with PTEN as it does in cells that lack it. This raised the question of why normal cells were able to tolerate the toxic compound.

The answer has to do with how cells use glucose. Cells without PTEN take in glucose from their environment to generate the energy-rich molecule ATP, which they import into mitochondria to keep them intact. But, mitochondria are supposed to *generate* ATP for the rest of the cell, not consume it. For the cells lacking PTEN, unless there is an endless supply of glucose, they quickly use up the sugar and die.

All cells need glucose, but cells with an unusual need for it, like cancer cells lacking PTEN, are especially vulnerable to its availability. This means it could be critical to administer mitochondrial inhibitors to cancer patients when their blood sugar is low, Trotman says.

From Many Mice, Unexpected Variety in Decision-Making Strategies

When it comes to sample sizes in experiments to understand decision-making, a CSHL team found that testing more subjects in more trials is not only better, but necessary, to truly grasp what an individual is thinking.

Anne Churchland and colleagues presented about three dozen mice with a task: Watch blinking lights and determine the frequency of the flashing, or stimulus. "If they are correct, the mice learn that if it's a low-rate stimulus they should move to the left port to get a reward; and if it's high-rate, they go the other direction to get the reward," the neuroscientist explains.

A. Churchland

The scientists assumed that the rate of the flash was all the mice would pay attention to, and set up the experiment to ensure that the mice had no other source of information to distract them. But, unexpectedly they found that some mice simply estimated the brightness of the complete flashing experience.

"If you think about it, a high-rate stimulus is 15 flashes per second. The overall amount of photons emitted will be higher than if it's a low-rate stimulus, flashing seven times per second. So, some mice figured out that they don't actually have to count the flashes," said Churchland. "The mice could just measure the total number of photons coming from the LED panels that we used. They found this extra piece of information that we hadn't realized was there."

If the researchers had just looked at a few mice in this experiment, they would have missed the fact that some of them were more clever. The team included this finding in their equations and set themselves up to learn more about each individual mouse's ability to deduce.

"If we understand why an individual is making a decision (i.e., what they're considering), we can better understand the mechanisms in the brain supporting that decision," Churchland said. "It's critical to thoroughly understand what the animal's strategy is and to be in a position to distinguish different decision-making strategies."

How Pancreatic Cancer Spreads after Surgery

Surgery is usually not an option for pancreatic cancer patients whose primary tumor has metastasized. But doctors have been puzzled by the poor outcome of patients whose tumor seems confined to the pancreas at the time of diagnosis and thus qualify for surgery. In many such patients, the liver appears cancer-free. Yet within 2 years, most of these patients develop lethal metastatic cancer, often in the liver.

D. Fearon

Doug Fearon discovered how the cancer spreads in patients whose tumor is successfully removed. After surgery, patients experience a 2-week period during which their immune system is depleted as a result of a surge in postoperative stress hormone (cortisol) levels. With killer T-cell levels sagging, isolated, dormant cancer cells that have already traveled to the liver via the bloodstream begin to grow or metastasize.

This postoperative period, suggests Fearon, "offers a window during which efforts might be made to keep cortisol levels down and T cells strong so the patient's own immune system can kill the cancer cells that have made their way to the liver but until this point have been dormant."

Fearon's team explains that dormant cancer cells are already in the liver well before patients have their primary tumor removed. They are carried there by the bloodstream, having been shed by the primary tumor. The immune system can kill most—but not the dormant—non-growing cancer cells that are deposited in the liver.

The immune system seeks and destroys cancer cells by sensing proteins called MHC I present on the outer membrane of the cancer cells. Fearon's team found that the cancer cells that have been lying dormant in the liver of pancreatic cancer patients do not express these proteins, so killer T cells cannot find them. In situations such as postoperative surgical stress, in which T cells in the liver are depleted, the dormant cancer cells reexpress MHC I and begin to divide, becoming seeds of metastatic lesions. Even though the MCH I molecule is expressed, the stress-induced cortisol blocks the T-cell response, allowing the growing metastatic cells to escape T-cell killing.

How a Sleeping Cancer Awakens and Metastasizes

Even after successful cancer treatment, dormant, nondividing cancer cells that previously detached from the original tumor may exist elsewhere in the body. If awakened, these cells can grow into metastatic tumors. Studying metastasis to the lungs, Mikala Egeblad's laboratory identified signals accompanying inflammation that can awaken dormant cancer cells.

M. Egeblad

The team demonstrated that sustained lung inflammation, including that caused by tobacco smoke exposure, can cause dormant breast and prostate cancer cells that have traveled to the lungs to awaken and metastasize in the lungs.

With colleagues, they demonstrated a way to block the signal that awakened the dormant cancer cells, a concept that could prevent cancer recurrence. The team showed that sustained lung inflammation induced common white blood cells called neutrophils to awaken nearby dormant cancer cells.

Neutrophils normally kill invaders like bacteria and yeast. They can expel their DNA into the space beyond the cell membrane. Laced with toxic enzymes, this expelled DNA forms net-like traps (called neutrophil extracellular traps, or NETs) that can kill a pathogen.

Egeblad showed that sustained lung inflammation causes the formation of NETs in the area around dormant cancer cells. Two enzymes in the NETs interact with a protein in tissue called laminin. In sequence, the enzymes make cuts in laminin proteins and change their shape, exposing a new surface (called an epitope).

When recognized by dormant cancer cells nearby, the epitope spurs signaling that awakens the cancer cells. The Egeblad team created an antibody to block the epitope, and, in mice, this prevented the reawakening of dormant cancer cells. Work has begun to optimize the antibody, with the hope of conducting trials in humans.

Organoid Profiling Personalizes Treatments for Pancreatic Cancer

D. Tuveson

A team of researchers led by CSHL Cancer Center Director David Tuveson demonstrated that patient-derived organoids, hollow spheres of cells cultured directly from a patient's tumor, can quickly and accurately predict how patients with pancreatic cancer respond to a variety of treatments.

Pancreatic cancer is one of the deadliest cancers. Currently, surgical removal of the cancerous tissue is the only effective treatment—but because the disease progresses so quickly, only 15% of patients are eligible for the procedure. Surgery-ineligible patients can be treated with chemotherapy, but patient response is highly varied, and there is no good method to determine which treatment is best for any given patient.

For several years, the Tuveson laboratory has been honing organoid technology. Taking only 3–6 weeks to grow, a major advantage of organoids is that they can be derived from tumors using tiny needle biopsies. Tuveson's team grew organoids from 66 pancreatic cancer patients and tested the sensitivity of each sample to five standard-of-care chemotherapy drugs. The team also measured the gene expression patterns in these organoids and compared the gene expression patterns to the drug sensitivity of the five different chemotherapies. They found that three "signatures" of gene activity in the organoids correctly identified patients who had responded well to these drugs, using tumor samples from several Canadian clinical trials. "The signatures are promising and may enable physicians to choose the best initial chemotherapy treatment for pancreatic cancer patients," Tuveson says. Tuveson also said that such "pharmaco-typing" approaches may apply to other cancer types.

Tuveson and his team plan to further refine the gene signatures and to test in prospective clinical trials the ability of these signatures found in organoids to predict the responses of pancreatic cancer patients.

A New Type of Lung Cancer

C. Vakoc

Researchers discovered a new kind of small-cell lung cancer (SCLC), paving the way for developing personalized medicine approaches to target it. About 10%–15% of all lung cancers are SCLC, a cancer without a specific treatment that often spreads early.

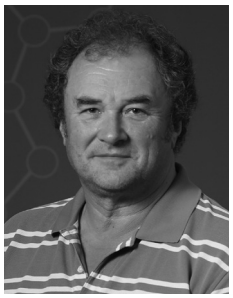
Analysis of gene activity in human SCLC tumors revealed an unexpected activity pattern in ~20% of samples. Christopher Vakoc and his team found a paucity of neuroendocrine markers in pulmonary neuroendocrine cells, a cell type thought to be the source of SCLC.

To further characterize this minority of cells, Vakoc and colleagues used a method they developed in 2015 that employs the gene-editing tool CRISPR to screen for specific proteins that are critical to the growth of various human cancer cell lines. They found that a transcription factor called POU2F3 is expressed exclusively in the minority of

SCLC tumors with low levels of neuroendocrine markers. Developing drugs that specifically target the function of POU2F3 may be particularly effective in the subset of patients with tumors that express high levels of this transcription factor.

Vakoc's team is now looking to do preclinical tests in mice to test compounds that target POU2F3.

Toward an Improved Wilson's Disease Drug



N. Tonks

In collaboration with DepYmed Inc., a CSHL spinout company, Nicholas Tonks and his team conducted promising preclinical experiments on a compound to treat Wilson's disease and other disorders in which levels of copper in the body are elevated.

Wilson's disease, affecting 1 in 30,000 people, is a severe inherited disorder that leads to profound liver and neurological damage. It is caused by mutations in a gene that encodes an enzyme critical in the excretion of excess copper from cells and organs. Copper is obtained mainly through the diet. Although essential in bodily function, it can be toxic when it accumulates. Normally, amounts of copper are precisely regulated both at the cellular level and in the body as a whole. In Wilson's patients, abnormal copper buildup begins in the liver. Copper toxicity can lead to liver enlargement, hepatitis, cirrhosis, and liver failure.

The team's research confirms that DPM-1001, a small molecule, robustly reduces copper levels in cells grown in culture that were sampled from Wilson's disease patients, as well as systemically in a mouse model of the disease.

The team showed that DPM-1001 could be taken as a pill, and is "exquisitely specific" for copper. Current de-coppering agents tend to affect levels of other metals in addition to copper—an undesirable feature in a drug for an illness like Wilson's.

In a mouse model of Wilson's disease, DPM-1001 ameliorated associated liver complications. This was accompanied by dramatic lowering of tissue copper levels and reduced disease symptoms. Optimization work on the compound continues in his laboratory in collaboration with DepYmed Inc.

Using CRISPR to Bring Orphan Crops to Market

Z. Lippman

Most people have never tasted a groundcherry—the small, sweet relative of the tomato. That's because the groundcherry plant, with long, straggly branches and sporadically ripening fruit, is unsuitable for large-scale agriculture. Zach Lippman used the gene-editing tool CRISPR to engineer groundcherries that maintain compact, manageable stems and that produce larger, more abundant fruit.

Lippman and colleagues described the effects of changes to three groundcherry genes. One change reins in production of a hormone that regulates flowering, making plants more compact and producing fruit in clusters. A change to another flowering hormone leads to denser fruit production: Plants with a CRISPR-generated mutation in this gene produced up to 50% more fruit than the unmodified plant. The third change boosts the number of seedy sections within each fruit, increasing its overall size.

The team is working on other aspects of the plant to make the groundcherry practical for large-scale production. The current success demonstrates that it is possible. In less than 2 years, they achieved improvements that would have taken far longer using traditional breeding practices. The implications are much broader. With gene editing, other wild plants or orphan crops may be brought into agricultural production.

Research Faculty

Awards

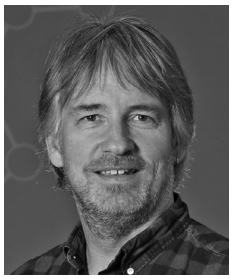
Professor Adrian Krainer was awarded the prestigious Breakthrough Prize at the celebrity-studded, televised ceremony known as the "Oscars of Science." The November ceremony included

A. Krainer

Emma Larson, a little girl who participated in the clinical trial to test the lifesaving spinal muscular atrophy (SMA) treatment drug that Krainer co-developed. Krainer received numerous additional accolades throughout the year for his ability to dive deeply into the basic biology of RNA splicing and then work collaboratively with industry partners to apply this knowledge to develop a drug.

Krainer came to the Lab more than 30 years ago, as the first CSHL Fellow. He had just received his Ph.D. at Harvard, and CSHL gave him the extraordinary opportunity to run his own, curiosity-driven lab. Programs to support early-career scientists in pursuing independent research continue here today.

In addition to winning the 2018 Breakthrough Prize, Krainer was named a National Academy of Inventors (NAI) 2018 Fellow, honored for his work on Spinraza®, the first FDA-approved treatment for SMA. NAI Fellows are nominated by their peers; those chosen for the honor demonstrate a highly prolific spirit of innovation in creating or facilitating outstanding inventions that have made a tangible impact on the quality of life, economic development, and welfare of society.



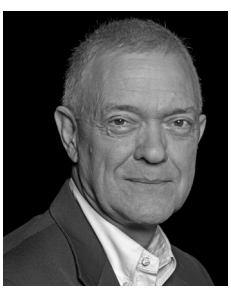
D. Jackson

CSHL had two American Association for the Advancement of Science (AAAS) 2018 Fellows. Professor David Jackson was honored in the field of Agriculture, Food and Renewable Resources for his discoveries of the genes and signals that regulate stem cell behavior in plants, thereby affecting plant architecture and yield.

In the field of Biological Sciences, Jan A. Witkowski, Watson School of Biological Sciences professor, was awarded for his seminal role in advancing science through his leadership of the Banbury Center. CSHL's think tank for science, Banbury Center is recognized by AAAS for holding discussion meetings that review key issues in biology and medicine.

Professor Leemor Joshua-Tor was honored by the American Society for Biochemistry and Molecular Biology's 2018 Mildred Cohn Award in Biological Chemistry. The award is named for the pioneering biochemist who developed powerful techniques for understanding how molecules behave in the body. The highly prestigious Cohn Award recognizes innovative scientists who have made substantial advances in understanding biological chemistry. Accepting the award, her lecture focused on the structure of a molecule that enables DNA to be accurately replicated as each new cell is born. It is one of countless projects that Joshua-Tor has tackled since joining CSHL's faculty in 1995.

Professor Rob Martienssen received the 2018 Barbara McClintock Prize for Plant Genetics and Genome Studies. The Barbara McClintock Prize for Plant Genetics and Genome Studies was created to memorialize the contributions of Dr. McClintock. Her 1983 Nobel Prize in Physiology or Medicine was awarded for her discovery and characterization of transposable genetic elements that she determined could move within the genome. The McClintock prize recognizes the most outstanding plant geneticists of the present era.



J. Witkowski

L. Joshua-Tor

R. Martienssen

Martienssen was recognized for his pioneering contributions to epigenetic mechanisms of gene regulation and inheritance; his stellar work with transposons, DNA methylation, and histone modification, which has linked these with chromatin remodeling and RNA interference; his efforts revealing unifying mechanisms that underlie transcriptional and posttranscriptional silencing; demonstrating the existence of plant cell type–specific small RNAs and their capacity to alter imprinting by moving to adjacent cell types; and bringing diverse approaches to bear on identifying many genes that control classical genetic traits.

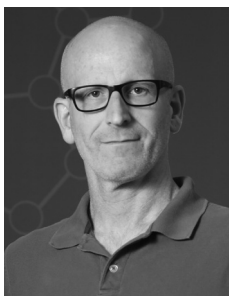
Professor Zachary Lippman was selected as a National Finalist in Life Sciences for the 2018 Blavatnik National Awards. The Blavatnik National Awards honor outstanding scientists under the age of 42 in the fields of Life Sciences, Chemistry, and Physical Sciences & Engineering. Lippman's research focuses on the genes that determine when, where, and how many flowers are produced on a plant, using tomatoes as a model system. Employing a combination of genetic, genomic, and molecular approaches, his team is developing new strategies for improving crop yields.

Blavatnik Award

Professor Anthony Zador was named a Gill Symposium Transformative Investigator for his work on MAPseq. The prize honors researchers who have made exceptional contributions to cellular or molecular neuroscience. MAPseq (multiplexed analysis of projections by sequencing) is a revolutionary brain-mapping method developed by Zador and his team. The tool barcodes thousands of neurons and determines their wiring patterns at single-neuron resolution.

Professor Nicholas Tonks was awarded the American Society for Biochemistry and Molecular Biology's (ASBMB) 2019 Earl and Thressa Stadtman Distinguished Scientist Award. Tonks' research has largely focused on the protein tyrosine phosphatase (PTP) family of enzymes, the "inaugural member" of which, PTP1B, he discovered in the late 1980s. In recent years, Tonks has obtained promising results aimed at devising new therapeutic strategies for diseases such as breast cancer, diabetes, Parkinson's, and Alzheimer's. This award, given to established scientists for their outstanding achievements in basic research, is issued once every two years.

Associate Professor Molly Hammell was awarded the Chan Zuckerberg Initiative (CZI) Ben Barres Early Career Acceleration Award for her proposed work on amyotrophic lateral sclerosis, better known as ALS or Lou Gehrig's disease. Hammell proposes to develop machine-learning software that would systematically identify genetic factors and molecular mechanisms that lead to motor neuron death. She will focus on transposable elements—viral-like genomic parasites that normally lie dormant in the genome—that are implicated in multiple diseases, including ALS. The award is a part of CZI's Neurodegeneration Challenge Network, which connects researchers who are studying neurodegenerative diseases and encourages a cross-disease perspective.


A. Zador


N. Tonks


M. Hammell

F. Albeanu and
A. Koulakov

M. Egeblad receiving the Suffrage
Science award

Associate Professor Florin Albeanu and Professor Alexei Koulakov received the National Institutes of Health (NIH) Director's Transformative Research Award for an innovative neuroscience research project on the olfactory system, one of the basic senses that is still quite mysterious. The project will study how the brain interprets odors, an aspect of neuronal processing. Together, Albeanu and Koulakov are building a general framework that will help standardize the study of olfactory receptors and how they broadcast odor-related information brain-wide. By gathering data in the Albeanu laboratory and applying mathematical methods used in artificial intelligence in the Koulakov laboratory, they hope to be able to predict olfactory "rules" based on the chemical structure of the odor. The NIH Director's Transformative Research Award provides funding for innovative research that is considered "high-risk" but has the potential to create or overturn fundamental paradigms of science.

Associate Professor Mikala Egeblad was honored with a Suffrage Science award at the Academy of Medical Sciences, London. The award, in the form of heirloom jewelry, honors women in science with the aim of encouraging more women to enter science, stick with it, and reach senior leadership roles. Rather than having committees select the winners, each winner selects who will get the heirloom jewelry next. Egeblad's cancer research laboratory focuses on understanding the network of immune cells, blood vessels, chemical signals, and support structures that make up what is known as the tumor microenvironment.

T. Engel

Assistant Professor Tatiana Engel is helping build computational tools for data collected specifically from the brain. She was awarded a Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative grant from the National Institutes of Health. The goal of the federal government's BRAIN Initiative is to accelerate the development and application of innovative technologies to produce a revolutionary new dynamic picture of the brain. Engel will build mathematical models of decision-making activity in two different areas of the brain. Engel will work with Associate Professor Anne Churchland on this project.

C. dos Santos

Assistant Professor Camila dos Santos was awarded the Pershing Square Sohn Prize for Young Investigators in Cancer Research. This is the fifth year the Alliance has awarded the prize to promising early-career, New York City–area scientists—this year, to six others in addition to Dr. dos Santos. She aims to find ways to prevent breast cancer by examining the epigenome, a system of molecular marks that change the way genes are expressed without changing the DNA itself. Her research builds on a correlation that has been observed for nearly 100 years: that a full-term pregnancy early in a woman's life dramatically reduces

her risk of developing breast cancer later in life. By analyzing the epigenome in animal models of breast cancer, Dr. dos Santos's group has already established that pregnancy changes the epigenome of breast cells known as mammary epithelial cells.

CSHL Fellow Lingbo Zhang aims to extinguish myelodysplastic syndromes (MDSs) and was recognized by the Edward P. Evans Foundation with an EvansMDS Young Investigator Award. This highly competitive award supports Zhang's translational research on MDS. Zhang is one of the first researchers to receive this award, created to support the development of the next generation of research leaders who will blaze a path toward cures for MDSs.

L. Zhang

To assess a potential therapeutic compound that Zhang identified and move the project toward clinical trials for MDSs, he is currently collaborating with a medicinal chemist at Northwell Health, with which CSHL formed a strategic affiliation in 2015. Using CRISPR gene editing to rapidly and precisely make changes to the genome, Zhang has screened MDS cells for a list of vulnerabilities, which he has since narrowed down to those that could be most effectively targeted with drugs.

CSHL Fellow Semir Beyaz received three different honors for his research, which looks at how dietary choices impact the body's immunity against cancer. Beyaz first received the Jeffrey Modell Prize, given by the Jeffrey Modell Foundation and the Harvard Committee of Immunology to a Harvard graduate student for excellence in his or her graduate career and dissertation.

S. Beyaz

He also received the Turkish American Scientists & Scholars Association (TASSA) Aziz Sancar Award, given to young Turkish researchers and scholars in STEM fields.

Additionally, Beyaz received the Sabri Ulker Early-Career Award from the Organizing Committee of Cell Symposia: Translational Immunometabolism for an abstract submission.

New Hires/Promotions

2018 brought a new CSHL Fellow, Semir Beyaz, and Assistant Professor, Tobias Janowitz, to pursue new, whole-organism approaches to cancer, obesity, and nutrition. Central to their efforts will be the historic 1953 Demerec Laboratory, under renovation and scheduled to open in 2019. Alexander Dobin, who had been a postdoc at CSHL since 2008, also joined as Assistant Professor to further investigate cancer genomics.

Plant biologist Professor Zach Lippman began his CSHL career as a Ph.D. student; he returned to join the faculty after a postdoc in Israel. This year he was named a Howard Hughes Medical Institute (HHMI) Investigator. CSHL is proud of this addition to our existing HHMI Investigators, Professors Rob Martienssen and Leemor Joshua-Tor.

Molly Hammell and Dan Levy were promoted to Associate Professor.

S. Beyaz

T. Janowitz

A. Dobin

Z. Lippman

Education Highlights

Meetings & Courses Program

CSHL Meetings this year attracted 7,000 participants from more than 50 countries to the main campus. The 83rd Cold Spring Harbor Symposium, Brains and Behavior: Order and Disorder in the Nervous System, explored how fundamental brain research and technologies are translating to improved brain health and treatments for psychiatric and neurological disorders. The symposium was supported by the Tianqiao & Chrissy Chen Institute. Single Biomolecules and Nutrient Signaling were new additions to the meetings program. The Evolving Concept of Mitochondria: From Symbiotic Origins to Therapeutic Opportunities was the topic of the 10th meeting of the Genentech Center, History of Molecular Biology & Biotechnology series.

The Cold Spring Harbor Asia (CSHA) conference program drew 3,500 scientists to symposia, meetings, and Banbury-style discussions designed for scientists from the Asia/Pacific region. CSHL was pleased to sign a new partnership agreement with Suzhou Industrial Park to both extend and expand the commitment to operating meetings and courses on the Dushu Lake campus for another decade.

Covering a diverse array of topics in molecular biology, neurobiology, structural studies, and bioinformatics, 700 instructors, lecturers and assistants have come to teach at CSHL from universities, medical schools, research institutes, and companies around the world. In 2018, 650 trainees—advanced graduate students, postdocs, and faculty—attended courses lasting from one to three weeks. The latest addition to the program is the course on Cryoelectron Microscopy.

The Courses program relies on grants and foundation support, including major support from the Helmsley Charitable Trust, the Howard Hughes Medical Institute, the National Institutes of Health, and the National Science Foundation. The Courses also benefit from the loan of equipment, reagents, and technical support from many companies, whose support is indispensable to ensure that the program remains cutting-edge.

Banbury Center

This year's 18 meetings demonstrated two of Banbury's core strengths: bridging interdisciplinary divides and hosting discussions at the frontiers of science and technology. Discussions covered the legal and ethical implications for the growing neuromonitoring and neuromodulatory devices

Banbury Conference Center meeting

market, expected to top $3 billion by 2020; explored the use of DNA to store data; considered comprehensive and community approaches to HIV prevention; considered new questions about aging from the long-living bat; and took aim at the opioid epidemic from a practical, policy standpoint.

The Evolving Phenomenon of Direct-to-Consumer Neuroscience meeting produced a January 2019 publication in *Science* suggesting ways to provide systematic support for regulatory agencies, funding bodies, and a public that is thirsty for knowledge about the efficacy of consumer-targeted neurotechnology products.

A second paper resulting from the 2016 Banbury Meeting Diagnostic Tests for Lyme Disease: a Reassessment and Pathways Forward was published in *Clinical Infectious Diseases*, marking a major advance in the ability to diagnose and treat Lyme disease sooner.

Signals of Trust in Science Communication gathered a diverse group of eminent scientists, journal editors, science writers, and communications experts to debate complex issues surrounding trustworthiness in science. A Twitter campaign associated with the meeting allowed users from across the world to share their perspectives on the issue.

Sponsored by CSHL, Increasing Gender Diversity in Biosciences convened to identify practical solutions to better recruit, promote, and support women in science. A report from the meeting has been accepted for publication in the prestigious journal *Science* in 2019.

Banbury continues to attract financial support from across sectors, with almost half drawn from not-for-profit organizations.

Watson School of Biological Sciences

CSHL welcomed its 20th incoming class of Ph.D. students from the United States, Canada, China, France, Hong Kong, India, Pakistan, and Turkey. A goal of the program is to graduate students faster than students in comparable Ph.D.-granting institutions and position them to secure excellent jobs early in their careers. As of this year, 105 CSHL Ph.D. graduates are now thriving in the world.

Twenty-seven graduates have secured tenure-track faculty positions and are receiving federal grants and publishing papers as independent researchers. One alumnus, Dr. Zachary Lippman, is a professor at CSHL and this year became the first CSHL graduate to be named a Howard Hughes Medical Institute Investigator. The School's alumni have also moved into influential positions in administration, publishing, consulting, and industry—with one of this year's graduates taking a data scientist position. During the year, scientific papers published by students appeared in major journals, bringing the cumulative total to 400.

WSBS entering Class of 2018

At the 2018 graduation ceremony, five students were awarded Ph.D. degrees. Drs. David and Leon Botstein were awarded honorary degrees. David Botstein is a prominent geneticist, CSHL course instructor, and CSHL Trustee from 2003 to 2013. Leon Botstein was the youngest person appointed as college president in American history and currently is the President of Bard College and the music director of the American Symphony Orchestra.

From June through August, 20 undergraduates from around the world performed advanced research in the laboratory of a CSHL faculty member. In its 59th year, the annual immersive experience called the Undergraduate Research Program reaped intellectual as well as social rewards for the participants. The equally innovative Partners for the Future program brought gifted local high school students to CSHL labs for hands-on research experience.

DNA Learning Center

Just as a universal product code (UPC "barcode") uses a unique set of bars to identify each consumer product, a DNA barcode is a unique set of DNA "letters" that identifies each living thing.

Students and faculty swabbing snakes to collect microbiomes

Over the last eight years, the DNALC has developed simplified biochemistry and bioinformatics analysis that puts DNA barcode research within reach of students and citizen scientists. Funding from the Thompson Family Foundation, the Laurie Landeau Foundation, the Simons Foundation and the Pinkerton Foundation supported barcode projects by 582 students from 80 high schools across New York City and Long Island this year.

The DNALC's barcoding infrastructure also supports course-based *undergraduate* research experiences (CUREs), which have been shown to increase graduation rates and student retention in STEM disciplines. In October, DNALC began a new 5-year project with funding from the National Science Foundation (NSF) program Improving Undergraduate STEM Education (IUSE). In addition to popularizing our existing barcoding platform for student projects, the initiative will develop affordable technology for metabarcoding—which uses next-generation sequencing to identify the variety of microbes, fish, or other living things in an environmental sample. The project will train and mentor 80 college faculty as they implement CUREs in a variety of classes and school settings. Among key collaborators on the project are City Tech, where we will open a new DNALC in 2020, and James Madison University (JMU), whose barcoding CURE serves 500 students per semester.

During the academic year, 21,176 students conducted labs at Dolan DNA Learning Center, DNALC West and Harlem DNA Lab; and 1,347 students attended week-long summer camps. An additional 7,905 students conducted in-school labs led by DNALC staff, and 1,983 used footlocker kits. Also, this year, 5.6 million visitors accessed DNALC's suite of multimedia resources online, including 4 million visits to DNALC websites, 911,378 views of YouTube videos, and 661,726 downloads of smartphone/tablet apps, the *3D Brain*, *Weed to Wonder*, and *Gene Screen*.

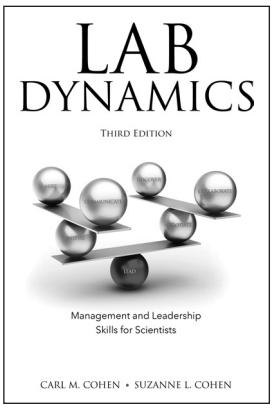
The DNALC has been a partner of the Breakthrough Prize competition, delivering a state-of-the-art science laboratory to the high school of the Junior Breakthrough Challenge winner. In 2016, Hillary Diane Andales entered the Junior Breakthrough Challenge, which asks high school students to make a video about their favorite concept in science or math. Among the thousands of submissions from all over the world, Hillary's video was voted the most popular, with a whopping 40,000 Facebook "likes."

Although the video did not win the competition's grand prize, it did earn Hillary's Philippine Science High School a new $100,000 laboratory classroom designed and equipped by the DNALC. In 2018 she submitted another video and won a $250,000 scholarship and *another* $100,000 DNALC-designed laboratory for her school.

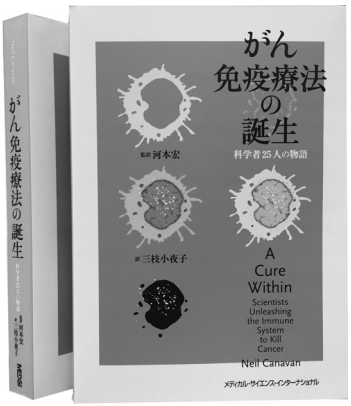
Cold Spring Harbor Laboratory Press

The Cold Spring Harbor Laboratory Press provides scientists worldwide with authoritative, affordable, and up-to-date information to further their research and aid in their career development. The CSHL Press continues to respond creatively and pragmatically to the opportunities offered by both online and print distribution and to the changing landscape of funding and business models in the sharing of scientific information.

The seven established research and review subscription journals continued to publish highly valued content and maintain or expand their international audience. *Genes & Development* and *Genome Research* remained at the top of their disciplines among primary research journals. *CSH Molecular Case Studies* (*MCS*), a newer open access title, is gaining ground among authors who have obtained results from the application of precision medicine techniques to one patient or a small cohort. Launched in April 2018, *Life Science Alliance* (*LSA*) is a new open access journal owned and published jointly by the Laboratory, the European Molecular Biology Organization (EMBO), and Rockefeller University. Using a unique cross-publisher "cascade" publishing model, *LSA* is a publication channel for manuscripts referred by nine



Lab Dynamics, 3E



A Cure Within, Japanese translation

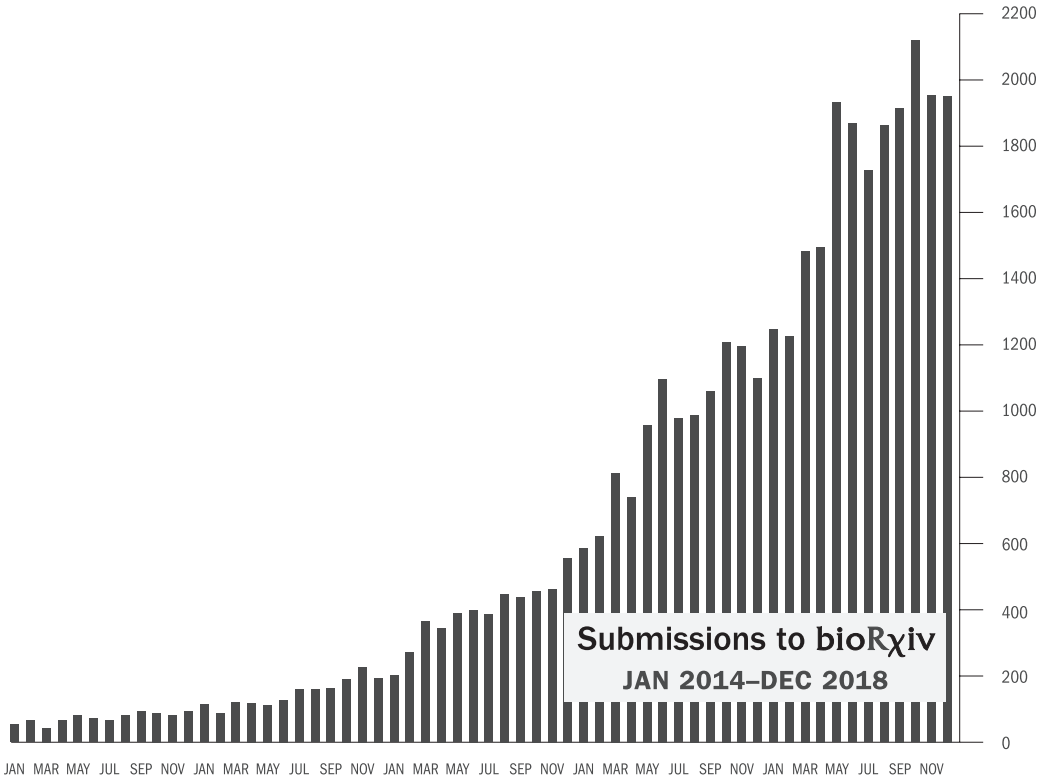
of the highly selective journals published by the three partner organizations.

The CSHL Press published 15 new book titles in 2018. Ten were themed collections of articles published in the *Perspectives in Biology* and *Perspectives in Medicine* journals. This publishing model, a successful marriage of serial online and print book publishing, successfully engages thought leaders as editors and distinguished investigators as authors. The handbook *Lab Dynamics* was published in a third edition that for the first time assists principal investigators in the challenging responsibilities of interviewing prospective team candidates and assessing performance.

Published in 2017, *A Cure Within: Scientists Unleashing the Immune System to Kill Cancer* by Neil Canavan continued to be a best seller, with a Japanese translation complete and a Chinese translation in progress—and gained added impetus with the award of the 2018 Nobel Prize in Physiology or Medicine to two scientists featured in the book, James Allison and Tasuko Honjo.

bioRxiv, the Preprint Server for Biology

A preprint is a research manuscript distributed by its authors before certification by peer review and publication by a journal. The Laboratory's preprint server, bioRxiv, turned five in 2018 and, with major support from the Chan Zuckerberg Initiative (CZI), continued its rapid



Growth in bioRxiv submissions

growth, doubling in size with 20,000 new submissions from more than 100 countries. Each month, papers on the server are read more than 4.5 million times and discussed in social networks and dedicated preprint assessment sites. The server is clearly accelerating science: 70% of bioRxiv manuscripts are shared for community evaluation as long as two years (median six months) before journal publication. Thirty journals now enable authors to simultaneously submit a manuscript and post it on bioRxiv, and 130 journals will accept automatic submission of preprints for editorial consideration. In 2018, an additional CZI grant made possible the conversion of all bioRxiv papers into XML format so they can be read in a web browser without downloading. bioRxiv continues to transform the way biologists communicate their science.

Marilyn Simons, Bruce Stillman, and Jamie C. Nicholls

Board of Trustees

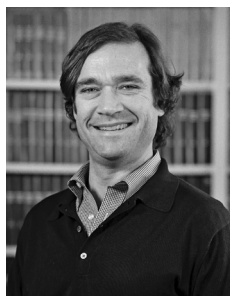
Dr. Marilyn Simons, president and co-founder of the Simons Foundation with her husband Dr. Jim Simons, CSHL honorary trustee, was elected chairman of the CSHL Board of Trustees in November. Marilyn was previously the vice-chairman of the Board, and led the successful 125th Anniversary Capital Campaign. She takes the baton from Jamie C. Nicholls, who served as chairman from 2010 to 2018. In recognition of Nicholls' unprecedented efforts to ensure the institution's financial strength and stability, as evidenced by the doubling of the endowment to more than $640 million during her tenure, Jamie was named the Laboratory's first-ever lifetime trustee.

New trustees elected in 2018 include Bruce Ratner, chairman of Forest City, New York; Dr. Stuart Weisbrod, former CSHL postdoctoral fellow and current chief investment officer of Iguana Healthcare Partners; Dr. Karel Svoboda, former CSHL professor and current Howard Hughes Medical Institute (HHMI) investigator and group leader at HHMI Janelia Research Campus; Dr. Elaine Fuchs, HHMI investigator and professor at the Rockefeller University; and Geoffrey Robertson, director of Business Assistance at the Vermont Sustainable Jobs Fund—who is the grandson of Charles and Marie Robertson, who in 1973 generously provided the initial gift to seed the CSHL endowment.

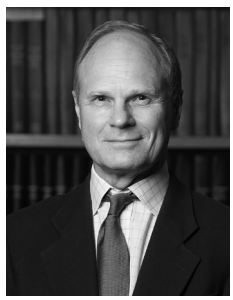
John P. Tuke was appointed chief operating officer in November 2018. He had been the chief financial officer of the Hotchkiss School for more than 19 years, overseeing a broad range of financial, operational, risk management, and strategic activities.

K. Svoboda

E. Fuchs

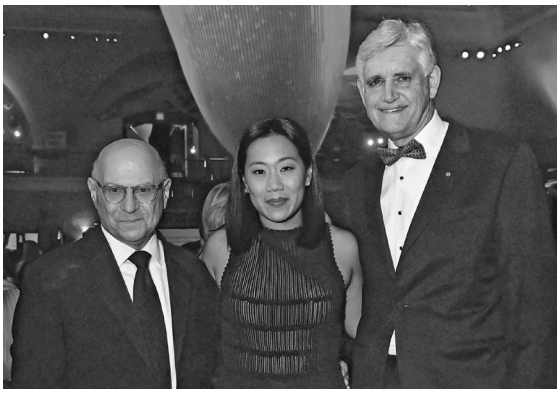


G. Robertson



J. Tuke

W.D. Ayres

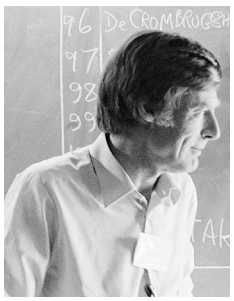


Larry Norton, Priscilla Chan, and Bruce Stillman

John succeeded W. Dillaway Ayres, who helped lead this organization with distinction for 20 years, setting the highest of standards—but, more importantly, was a leader in maintaining the collegial and entrepreneurial culture that is integral to the Lab's success.

The power of the CSHL Board of Trustees was demonstrated at the institution's record-breaking Double Helix Medals Dinner, which raised $4 million honoring Mark Zuckerberg and Dr. Priscilla Chan of the Chan Zuckerberg Initiative and renowned oncologist Dr. Larry Norton. Through other events and fund-raising initiatives, trustees and friends of CSHL rallied to raise a total of >$7 million of unrestricted money that supports the Lab's most innovative research initiatives. The Helix Society, a planned-giving initiative, continues to grow in members and dollars, becoming a vital source to guarantee the Lab's future.

CSHL Director John Cairns



J. Cairns

The Laboratory fondly remembers John Cairns, CSHL Director from 1963 to 1968, who passed away in late 2018. Cairns is known for pioneering work in DNA autoradiography. A physician and molecular biologist trained at Oxford, Cairns saw the Laboratory as "Arcadia," an unspoiled, harmonious wilderness, a mythical place with scientific ideal. He was dedicated to the Lab's survival during stressful financial times in the 1960s. See the special tribute to him in this publication.

Business Development & Technology Transfer

A. Whiteley

In April, CSHL announced the appointment of Andrew Whiteley as the new vice president of Business Development and Technology Transfer. He had served as executive in residence at CSHL since November 2015. Andrew succeeds Teri Willey, who served in this position since 2013.

This "innovation" team are business agents for CSHL and its scientists, helping to bring discoveries made at the Lab to the public through intellectual property licensing, industry collaborative research and new ventures. Highlights of 2018 included:

- Monetizing the assets from the development of the 2016 FDA-approved spinal muscular atrophy drug Spinraza®, which added $82 million to the endowment and contributed royalty streams that added $8 million to the Science Fund for reinvestment into research here.
- More than $3.58 million in licensing revenue and $630 million in sponsored research funding was received under agreements negotiated and managed by this team.

Library and Archives

The Message

All educational and scholarly programs are now part of the Center for Humanities Studies of Molecular Biology, an interdisciplinary destination for historians, scholars, researchers, and artists to take advantage of the rich history found in the original materials in our archives for their humanities projects. The Center is also home to all Library and Archives–hosted events related to modern biology.

In August, the Center hosted *New York Times* columnist and science writer Carl Zimmer, who spoke about his new book, *She Has Her Mother's Laugh*. In October, the Center hosted a play about a major event in modern science history. The new play, *The Message*, written by Keith Burridge, directed by Hal Brooks, and starring Broadway actors Brad Cover, Dominic Cuskern, and Rachel Botchan, was based on the discovery of messenger RNA (and written after Burridge's extensive research in our collections).

The Center has hosted two international meetings: a workshop on Historical Research on Model Organisms in Biology, and the annual History of Science meeting The Evolving Concept of Mitochondria. The meeting on mitochondria was attended by more than 150 scientists. Videos of the speakers' talks, along with their presentation slides and more, are available on the meeting website at http://library.cshl.edu/Meetings/Mitochondria. The workshop on model organisms gathered 21 scientists and historians from around the world to discuss the historical rise of select species from the wide array of experimental organisms to the vaunted status of model organisms and to formulate a suggested path forward for historical study of this aspect of the history of biology.

Sydney Brenner Research Scholarship recipient Matthew Cobb presented a talk on his research from the Archives about the working relationship between two Nobel laureates, Francis Crick and Sydney Brenner.

Based on an idea from Jim Watson to create a book about the major players of the Human Genome Project (HGP), editors Mila Pollock, Jan Witkowski, and Dick McCombie published *Faces of the Genome*. It consists of pencil sketch portraits of scientists involved in the HGP by famed Australian artist Lewis Miller. The original sketches are housed in the Archives. The book also contains biographical sketches of the scientists, written by their colleagues, and an introduction by Watson.

Infrastructure

Made possible with funds raised by the 125th Anniversary Capital Campaign, reconstruction of the Demerec Laboratory took place during the entire year, with scheduled completion in 2019. This major construction project brought a unique set of challenges, including the need to create a modern laboratory building within a historically sensitive building. The project's location at the center of campus near the main entrance added additional considerations to ensure unimpeded normal campus operations, including a busy Meetings & Courses Program year.

The long-overdue renovation of Dolan Hall was completed and fully operational for the second half of 2018. Guest rooms were renovated and improved to meet modern expectations, resulting in a modern boutique-style property with high-quality—albeit small—guest rooms.

The Laboratory also undertook a number of infrastructure renewal projects, including:

- *North Chiller Plant: chiller replacement.* Two aging chillers supplying the Hillside Laboratories were replaced with a larger, more energy-efficient chiller. This new high-efficiency unit offers a far greater service life and operates at a far lower cost than the preexisting chillers.

- *South Chiller Plant: emergency generator replacement.* The diesel emergency generator serving the Marks, James, and Freeman Laboratories suffered a catastrophic failure, requiring emergency replacement.
- *James Laboratory: walkway replacement.* The elevated walkway entrance to the James Laboratory had degraded over time, becoming unsafe for use. This walkway was reconstructed to current standards.
- *Tiffany House: retaining wall.* The Laboratory had been managing drainage and water issues with the Tiffany House for a number of years. A major project was undertaken to construct a new retaining wall behind the house and to create sufficient drainage to protect the property.
- *Housing Improvements.* The Laboratory continued its program of modernizing and improving its housing stock, with a number of renovations in the Robertson, Davenport, and Moore's Hill Road properties.

The Laboratory continued its ongoing program of researching and implementing energy conservation projects. These include both equipment and lighting retrofits with the intention of providing improved lighting and comfort at a lower cost and a reduced carbon footprint.

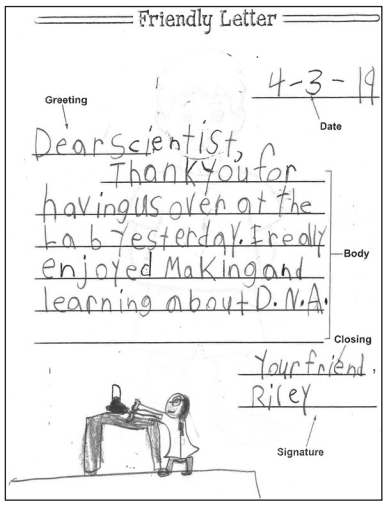
Community Outreach

The Public Affairs Department works closely with faculty, students, and employees across the Lab to create opportunities for the public to engage with the institution. This includes public tours, lectures, and talks, concerts, and various channels of regular communication through digital and printed materials available at cshl.edu. The Lab's digital platform was upgraded in 2018, making information about CSHL more easily accessible to audiences across the globe. An example of the department's success in digital content creation and distribution is *Base Pairs*, the three-season podcast series about the power of genetic information, which was selected as a finalist in the 2018 Webby Awards.

Base Pairs with Brian Stallard and Andrea Alfano

The Lab's public tour program attracted nearly 1,400 visitors to the Bungtown Road campus from spring through fall. Fifty-four group tours and 19 weekend public tours were guided by our guide team of 15 Ph.D. students.

CSHL remains active in our local schools, as faculty, students, and employees volunteer time to share our science and personal career stories with students and their families. In addition to the varied volunteer opportunities, the Lab annually hosts the Cold Spring Harbor School District's first graders and their parents for a hands-on science fair. This year's experimental stations shared knowledge ranging from brain biology to states of matter to DNA codes and chromatography. All of these activities are developed and delivered by Watson School graduate students and DNA Learning Center teachers.


Friendly Letter

4-3-19

Greeting

Date

Dear Scientist,

Thank You for having us over at the lab yesterday. I really enjoyed Making and learning about D.N.A.

Body

Closing

Your friend,

Riley

Signature

Thanks for the science fair

In addition to a series of scientific lectures that invites the community to the Lab's Grace Auditorium, a more casual event called Cocktails & Chromosomes is gaining a significant local following. With neighborhood bars as the

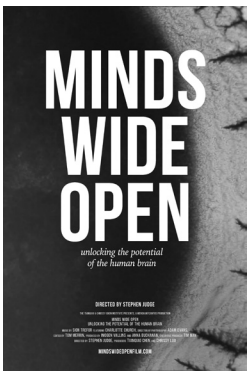
venue, this series of talks is strictly for nonscientists. CSHL faculty presenters make a special effort to avoid scientific jargon and simply share their curiosity about the wonder of their science with our neighbors.

CSHL Public Presentations

February 11: Screening and discussion at Cinema Arts Centre in Huntington Co.; *Food Evolution*. Panelists: **David Jackson, Ph.D.,** professor, CSHL; **Zachary Lippman, Ph.D.,** professor, CSHL; **Ullas Pedmale, Ph.D.**, assistant professor, CSHL; **Doreen Ware, Ph.D.**, molecular biologist, USDA ARS, adjunct associate professor, CSHL; co-presented by Cold Spring Harbor Laboratory & Science Advocacy of Long Island.

April 23: Public lecture, Energy from Thin Air: Basic Research to Biofuels. **Rob Martienssen, Ph.D.**, professor, CSHL; **Frank O'Keefe**, founder & CEO, Infinitree.

May 29: East coast film premiere of Tianqiao & Chrissy Chen Institute's (TCCI) documentary, followed by Q&A; *Minds Wide Open: Unlocking the Potential of the Human Brain*. TCCI was the exclusive supporting partner of the 2018 Cold Spring Harbor Symposium.



Minds Wide Open

June 12: Public lecture, Metastasis & Immunity: How Immune Cells Can Help Cancer Spread or Stop It in Its Tracks; co-sponsored by CSHL, US Trust, Northwell Health, and St. Johnland Nursing Center. **Mikala Egeblad, Ph.D.**, associate professor, CSHL; **Sylvia Adams, M.D.,** associate professor, Department of Medicine, director of Clinical Research, Breast Cancer Disease Management Group, NYU Medical Center.

June 20: *Cocktails & Chromosomes* at Six Harbors Brewing Co. **Jessica Tollkuhn, Ph.D.**, assistant professor, Cold Spring Harbor Laboratory.

July 31: Screening and discussion at Cinema Arts Centre in Huntington Co.; *The Most Unknown*. Panelists: **Bruce Stillman, Ph.D.,** professor, president & CEO, CSHL; **Scott McLennan, Ph.D.,** professor, Stony Brook University; **Lisa Miller, Ph.D.**, biophysical chemist, Brookhaven National Laboratory; co-presented by Cold Spring Harbor Laboratory & Cinema Arts Centre as a Science on Screen event.

October 5: Play by **Keith Burridge**. *The Message: A Play about the Discovery of Messenger RNA*. Panel including **Wally Gilbert, Ph.D.**; **Matthew Meselson, Ph.D.**; **James D. Watson, Ph.D.**

October 23: Public lecture, Genetic Privacy: Friend or Foe? 2018 Lorraine Grace lectureship on societal issues of biomedical research. **Yaniv Erlich, Ph.D.,** chief science officer, MyHeritage, associate professor of Computer Science, Columbia University [leave of absence], adjunct core member, New York Genome Center.

Cocktails & Chromosomes

December 5: *Cocktails & Chromosomes* at Six Harbors Brewing Co. **Tatiana Engel, Ph.D.**, assistant professor, Cold Spring Harbor Laboratory.

CSHL Public Concerts

April 20: Tomer Gewirtzman, piano
May 4: Yoonah Kim, clarinet
May 18: Naomi Louisa O'Connell, mezzo soprano

Naomi Louisa O'Connell

Employees celebrating Top Long Island Workplaces win

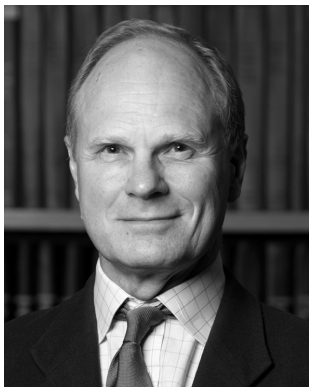
September 14: Argus Quartet, string quartet
September 28: Dominic Cheli, piano
October 12: SooBeen Lee, violin

Looking Forward

CSHL received a special honor this year from our own employees, winning Top Long Island Workplaces 2018. *Newsday* surveyed employees across Long Island on the quality of their workplaces. The anonymous employee surveys were analyzed by a third party. As part of the large company category (more than 500 employees), CSHL shared the honor with only 14 other institutions. Thank you to all who contribute to the success of this institution and the productive environment of all of our campuses. With your support, the future is bright.

Bruce Stillman, Ph.D., F.R.S.
President and Chief Executive Officer

CHIEF OPERATING OFFICER'S REPORT



J. Tuke

This is my first contribution to the *Annual Report*, having joined CSHL in November. I succeed Dill Ayres, who helped lead this organization with distinction for 20 years. Dill set high standards, but, more importantly, along with Bruce Stillman, he was a leader in maintaining the collegial and entrepreneurial culture that is integral to the Lab's success.

In 2018, the Lab's operating results included a strong performance in third-party revenues, and with expenses remaining in line with budgets, generated a surplus of $0.3 million after taking some appropriate reserves. We budget conservatively and maintain fiscal discipline throughout the year, which allows flexibility to pursue research and hiring opportunities in the very dynamic environment in which we operate. Revenues totaled $177.7 million, which included $108.7 million of external grant funding, $41.5 million of nonresearch-division support, and $27.5 million of contributions from the endowment, annual fund, and royalties. Growing the endowment, especially unrestricted funds, remains a strategic priority of the Board as federal and private grant funding becomes increasingly competitive. Operating expenses are dominated by investment in people, with compensation and benefits totaling $90 million. Infrastructure expenses were at $8.2 million. These are relatively fixed costs, required so as to not disrupt our research activities, and we strive to recognize this reality by securing and growing stable funding sources.

Endowment balances grew from $530 million to $595 million during 2018, thanks to proceeds from the Spinraza® monetization and gifts, and net of the operating draw and investment returns. The 4.4% loss in 2018 outperformed our policy benchmark (5.7%) as financial markets ended the year in turmoil with the fourth quarter wiping out a 3.3% gain that had been made before its start. Endowments are managed for long-term horizons, and the trailing 10-year return through 2018 is 7.5%. The Board-approved spending policy is to appropriate a prudent 4.5% of the trailing 3-year average endowment value for operations.

After an expansion of both human capital and infrastructure over the last decade, there was no change in either in 2018. Employee count remained essentially flat, and there was no increase in our five campuses' aggregate facility footprint. A comprehensive $29 million renovation of the Demerec Laboratory is underway. Built in 1953, it is where four Nobel laureates worked. Reopening in summer 2019, it was funded by gifts and support from New York state. It demonstrates the investment needed to keep our labs state-of-the-art. Bittersweet news for the Lab was that after 38 years of dedicated and thoughtful leadership as vice president of Facilities, Art Brings announced his retirement. A well-crafted master plan and the magnificent Hillside complex are among Art's most enduring legacies.

CSHL prides itself on an administrative structure that is flat, lean, and collaborative. The chief operating officer's role is to orchestrate the resources and activities of 11 departments, whose mission is to efficiently support our research and the five auxiliary, nonresearch divisions: the Meetings & Courses Program, the DNA Learning Center, the Banbury Center, the CSHL Press, and the Watson School of Biological Sciences. These units excel in programming linked to the core research and, in the aggregate, provide net financial support to the Lab.

I highlight two of the 11 departments as unique to an organization like CSHL and as being integral to our enterprise model. The Office of Sponsored Programs (OSP) works closely with our scientists on all aspects of the grant process, from advising on applications to helping build the budgets and assisting with the ongoing stewardship of the grants. The leverage provided by OSP

allows investigators to focus their time and energy in the labs. The Business Development and Technology Transfer team also works closely with our scientists to evaluate the commercial potential of their research. They manage the intricate process of building networks of possible funders and collaborators, as well as protect the Lab's intellectual property.

I came to CSHL because of its leadership, culture, and mission, and I am thrilled to be part of this community. CSHL operates in a highly competitive, complex, and dynamic space. It has the vision, resources, and resiliency to continue to adapt and succeed as a world-class institution.

John P. Tuke
Chief Operating Officer

Long-Term Service

Back row (*left to right*): Jeffrey Klaverweiden, Paul Edwards, Ming Wang, Dill Ayres, David Spector, Danny Jusino; *center row* (*left to right*): Patricia Brady, Lifang Zhang, Margot Bennett, Lari Russo, Bruce Stillman; *front row* (*left to right*) Margaret Falkowski, Frank Bowdren, Michael Wigler, James Watson, Christopher McEvoy.

The following employees celebrated milestone anniversaries in 2018:

50 years	James Watson
40 years	Christopher McEvoy, Michael Wigler
35 years	Margaret Falkowski, Danny Jusino
30 years	Margot Bennett, Frank Bowdren, Arne Stenlund
25 years	Deborah Aufiero, Patricia Brady, Wayne Pav, David Stewart
20 years	Dill Ayres, Drew Comer, Paul Edwards, Louis Hunter, Jeffrey Klaverweiden, Amanda McBrien, Tom McIlvaine, Susana Roman, Lari Russo, Ming Wang, Lifang Zhang

(*Left to right*) Eileen Earl, Bruce Stillman, Wayne Hamilton, David Spector, Manuel Ramones

15 years	Eileen Earl, Corrisa Farmer, Wayne Hamilton, Alex Koulakov, Mary Lamont, Erin McKechnie, Partha Mitra, Manuel Ramones, Krystyna Rzonca


RESEARCH

CANCER: GENE REGULATION AND CELL PROLIFERATION

Camila dos Santos' laboratory studies the epigenetic regulation of normal and malignant mammary gland development, with an emphasis on the alterations brought by pregnancy. Significant changes mark the pre- and postpubescence mammary developmental stages, but those associated with pregnancy have the greatest effect on cellular function, tissue reorganization, and breast cancer susceptibility. Her group has recently found that mammary glands react differently to a second pregnancy than they do to the first one, with associated changes in DNA methylation. These findings suggested that pregnancy changes the state of mammary cells, and these may permanently alter how they react to the next pregnancy. In addition, the dos Santos lab is exploring how the pregnancy-induced epigenetic changes might influence cell transformation and the risk of breast cancer. This research uses genomic and computational approaches to define the pre- and postpregnancy mammary epigenome. An additional objective of the dos Santos laboratory is to use functional genomics to discover novel transcriptional regulators that modulate mammary stem-cell self-renewal, lineage specification, and cell transformation. The long-term objective of Camila's group is to improve our understanding of the mammary epigenome during normal development and use this information to gain insight into new preventive and curative strategies to target breast cancer.

In **Leemor Joshua-Tor**'s lab, researchers study the molecular basis of nucleic acid regulatory processes using the tools of structural biology and biochemistry. One such regulatory process is RNA interference (RNAi), in which a small double-stranded RNA triggers gene silencing. Joshua-Tor and her team offered critical insight when they solved the crystal structure of the Argonaute protein and identified it as the long-sought Slicer. They then went on to explore the mechanism of the slicing event. The structure of human Argonaute 2 (hAgo2) bound to a microRNA (miRNA) guide allowed Joshua-Tor and her colleagues to understand how messenger RNA (mRNA) is cleaved during RNAi. This year, members of the Joshua-Tor lab explored the function of a very similar protein, called Argonaute 1, that has no slicing ability although it is almost identical in structure to the slicing hAgo2. Using biochemical methods and mutational analysis, they were able to identify key parts of the protein that are required for slicing activity. The lab also studies the generation of PIWI-interacting RNAs (piRNAs), which serve to protect the genome of germ cells. With colleagues in the Hannon lab, Joshua-Tor's team also determined the structure and function of Zucchini, a key nuclease in the initial generation of piRNAs in fruit flies. In other work, the lab is exploring the mechanisms of heterochromatin formation and gene silencing through the study of a protein complex called RNA-induced initiation of transcriptional gene silencing (RITS). Joshua-Tor is also well known for her work on the E1 helicase enzyme, which acts to unwind DNA strands during the DNA replication process.

Adrian Krainer's lab studies the mechanisms of RNA splicing, the ways in which they go awry in disease, and the means by which faulty splicing can be corrected. In particular, they study splicing in spinal muscular atrophy (SMA), a neuromuscular disease that is the leading genetic cause of death in infants. In SMA, a gene called *SMN2* is spliced incorrectly, making it only partially functional. The Krainer lab found a way to correct this defect using a powerful therapeutic approach. It is possible to stimulate SMN protein production by altering mRNA splicing through the introduction into cells of chemically modified pieces of RNA called antisense oligonucleotides (ASOs). Following extensive work with ASOs in mouse models of SMA, one such molecule, known as nusinersen or Spinraza®, was taken to the clinic, and at the end of 2016 it became the

first FDA-approved drug to treat SMA, by injection into the fluid surrounding the spinal cord. The Krainer lab is currently using this approach for the study of other diseases caused by splicing defects, including familial dysautonomia. In addition, they are applying antisense technology to stabilize mRNAs that are destroyed by a process called nonsense-mediated mRNA decay (NMD), both to learn about the underlying mechanisms and develop new therapies. The Krainer lab has also worked to shed light on the role of splicing proteins in cancer. They found that the splicing factor SRSF1 functions as an oncogene, and they characterized the splicing changes it elicits when overexpressed in the context of breast cancer; several of these changes contribute to various aspects of cancer progression. Finally, the lab continues to study fundamental mechanisms of splicing and its regulation, and they identified novel ways in which the U1 small nuclear RNA (snRNA) can recognize natural 5′ splice sites that deviate from the consensus.

David L. Spector's laboratory is focused on characterizing long noncoding RNAs (lncRNAs) that show altered levels of expression in breast cancer progression and during embryonic stem-cell differentiation. A major focus of their efforts has been on Malat1 lncRNA, which is one of the most abundant lncRNAs. The Spector lab previously identified a novel mechanism of 3′-end processing of this RNA. More recent studies have revealed that increased levels of Malat1 lncRNA impact breast cancer progression and metastasis. Knockout or ASO knockdown of Malat1 results in the differentiation of mammary tumors and a significant reduction in metastasis. Studies are currently under way to elucidate the mechanism of action of this abundant nuclear retained lncRNA and implement innovative therapeutic approaches that can impact its function in vivo. In addition, they have identified other lncRNAs, termed mammary tumor-associated RNAs, that are up-regulated in breast tumors, and they are currently assessing the function of these lncRNAs using 3D tumor organoids as well as mouse models.

A second area of study in the Spector lab is based on their earlier discovery of an increase in random autosomal monoallelic gene expression dependent on the differentiation of mouse embryonic stem cells to neural progenitor cells. These data support a model in which stochastic gene regulation during differentiation results in monoallelic gene expression, and for some genes, the cell is able to compensate transcriptionally to maintain the required transcriptional output of these genes. Therefore, random monoallelic gene expression exemplifies the stochastic and plastic nature of gene expression in single cells. Ongoing studies are examining the relationship of monoallelic gene expression to lineage commitment.

Arne Stenlund and colleagues have obtained a detailed understanding of processes required for initiation of DNA replication from the papillomavirus, using this system to gain a general biochemical understanding applicable in other systems. Papillomaviruses are a large viral family that induces cell proliferation at the site of infection, usually giving rise to benign tumors. But certain types of human papillomaviruses (HPVs) generate tumors that progress toward malignancy. Among these are HPVs that cause most cervical cancers. Members of the Stenlund lab also pursue studies aimed at developing an effective small-molecule inhibitor of HPVs that might someday be used by women who do not receive the preventive anti-HPV vaccine now available or those already infected with HPV who would not be helped by the vaccine.

Bruce Stillman's lab studies the process by which DNA is copied within cells before they divide in two. Working with yeast and human cells, Stillman and colleagues have identified many of the cellular proteins that function at the DNA replication fork during the S phase, the portion of the cell-division cycle when DNA synthesis occurs. Among these proteins are those that facilitate the assembly of chromatin, the protein–DNA complexes that form the chromosomes. Current research focuses on the mechanism that initiates the entire process of DNA replication in eukaryotic cells. At the heart of this mechanism is a protein that binds to "start" sites on the

chromosomes, called the origin recognition complex (ORC). The Stillman lab is part of an ongoing collaboration that determined the cryo-electron microscopy (EM) structure of ORC proteins in complex with a group of proteins, called a helicase, that unwind DNA during replication. These images offer molecular insights into how the helicase is loaded onto DNA. Stillman's research also focuses on the process by which duplicated chromosomes are segregated during mitosis. They found ORC at centrosomes and centromeres, structures that orchestrate chromosome separation during mitosis. At centromeres, ORC subunits monitor the attachment of duplicated chromosomes to the mitotic spindle that pulls the chromosomes apart when they are correctly aligned. Stillman's team has discovered that mutations in the Orc1 protein alter the ability of ORC to regulate both DNA replication and centrosome duplication. These mutations have been linked to Meier–Gorlin syndrome, a condition that results in people with extreme dwarfism and small brain size but normal intelligence.

Cancer can be understood as a disease of dysfunctional gene expression control. Research in **Chris Vakoc**'s lab investigates how transcription factors and chromatin regulators cooperate to control gene expression and maintain the cancer cell state. This work makes extensive use of genetic screens to reveal cancer-specific functions for transcriptional regulators, as well as genomic and biochemical approaches to identify molecular mechanisms. One theme that has emerged from their efforts is that blood cancers are often vulnerable to targeting transcriptional coactivators, such as BRD4 and the SWI/SNF chromatin remodeling complex. Vakoc's team showed that chemical inhibition of BRD4 shows therapeutic effects in mouse models of leukemia, a finding that has motivated ongoing clinical trials in human leukemia patients. The Vakoc lab has also developed a CRISPR-Cas9 screening approach that can reveal individual protein domains that sustain cancer cells. Their lab is now deploying this technology in a diverse array of human cancers to reveal therapeutic opportunities and basic mechanisms of cancer gene control.

UNDERSTANDING THE EPIGENETIC REGULATION OF NORMAL AND MALIGNANT MAMMARY GLAND DEVELOPMENT

C. dos Santos C. Chen M. Feigman
M. Ciccone M. Moss
S. Cyrill S.T. Yang

To Define a Role for BPTF Inhibition in Blocking Breast Cancer Progression

C. Chen, S.T. Yang, M. Ciccone

During this past year, we focused on defining the effects of BPTF inhibition on controlling the growth of mouse mammary tumors and on the survival of human breast cancer lines and patient-derived breast cancer organoids.

To date, we have established two transgenic mouse lines of mammary oncogenesis models to evaluate BPTF roles in cancer development.

a. $KRT5^{CRE\text{-}ERT2}BPTF^{fl/fl}BRCA1^{fl/fl}p53^{null}$ and $KRT5^{CRE\text{-}ERT2}BPTF^{WT}BRCA1^{fl/fl}p53^{null}$—In this model, deletion of BPTF and BRCA1 or BRCA1 alone ($BPTF^{WT}$) is driven by a tamoxifen (TAM)-inducible CRE-recombinase under the control of KRT5 promoter. Previous characterization of this model indicated that $KRT5^{CRE\text{-}ERT2}BPTF^{WT}BRCA1^{fl/fl}p53^{null}$ female mice will develop hormone-negative mammary tumors within three months from TAM treatment. We are currently monitoring tumor development in a cohort of $BPTF^{null}BRCA1^{null}p53^{null}$ and $BPTF^{WT}BRCA1^{null}p53^{null}$ female mice.
b. $KRT5^{CRE\text{-}ERT2}BPTF^{fl/fl}$MMTV-PyMT—In this model, BPTF deletion is driven by a TAM-inducible CRE-recombinase under the control of KRT5 promoter in a classical transgenic mouse model of luminal-like mammary oncogenesis (MMTV-PyMT). Preliminary tumor monitoring experiments suggest that BPTF deletion delays the onset and progression of MMTV-PyMT mammary tumors. We are currently characterizing epithelial cell autonomous and non–cell autonomous signals that were brought by BPTF deletion and impact with the development of MMTV-PyMT mammary tumors.

In addition, studies utilizing CRISPR-Cas9 genomic editing and small molecule inhibitor to target BPTF demonstrated that loss of BPTF function blocks cell cycle progression and induces apoptosis of human breast cancer cell lines, suggesting a role for BPTF on maintaining breast cancer progression and survival. Interestingly, our studies have also identified cell lines that are resistant to the effects of BPTF loss of function, providing an excellent platform to identify cancer-specific susceptibility to BPTF inhibition. We are currently developing CRISPR-Cas9-inducible stable cell lines to define a role for BPTF loss of function in the context of in vivo tumor/metastatic potential of human cancer cells and identify molecular pathways altered by BPTF loss of function.

Investigation of the Epigenetic Modifications Brought About by Pregnancy

M. Moss, C. Chen, M. Ciccone, C. dos Santos

Our previous study established that postpregnancy mammary epithelial cells are epigenetically distinct from their prepregnancy counterparts. More recently, we have characterized dynamics of several histone marks, including H3K27ac, H3K9me3, H3K4me1, and H3K4me3, as well as gene expression before, during, and after pregnancy. We found that histone marks of active transcription indicate that pregnancy augments the activity of distal regulatory regions, also known as enhancers, supporting the notion that the enhancer landscape of mammary epithelial cells is altered by pregnancy cycles.

In identifying these changes to the enhancer landscape, we have defined a set of enhancers that become active after pregnancy, which we have named pregnancy-induced elements (PIEs). We demonstrated

that these PIEs influence gene expression dynamics across different developmental stages in the mammary gland, including a more robust up-regulation of genes associated with milk production in a second pregnancy cycle, a finding that supports a role for epigenomic changes in controlling mammary epithelial cells' ready response to pregnancy signals.

We have recently optimized an organoid culture system to recapitulate the pregnancy-induced development in vitro, a strategy that has shown postpregnancy mammary organoids engaging in differentiation and gene expression changes more rapidly than organoid culture derived from prepregnancy mammary glands. These findings further support the idea that postpregnancy mammary epithelial cells (MECs) carry an epigenetic memory of prior pregnancies, in addition to providing a platform to dissect the contribution of specific gene sets and enhancer regions to gene regulation across pregnancy cycles.

Defining the Molecular Basis of Pregnancy-Induced Breast Cancer Protection

M. Feigman, M. Moss, C. Chen, C. dos Santos

For nearly 100 years, population studies have consistently and definitively found that a full-term pregnancy early in life dramatically reduces the lifetime incidence of breast cancer. Although research has shown increased breast cancer risk for roughly 5–10 yr after, there is a long-term reduction of risk of breast cancer for women completing a full-term pregnancy before the age of 30 yr. A similar risk decrease following pregnancy has been observed in mice, in which completion of a pregnancy cycle dampens the frequency of mammary tumor development. Given our findings demonstrating alteration to molecular programs brought by pregnancy to MECs and the longevity of cancer preventive effects in rodents and humans, it is possible that such effects have an epigenetic basis.

To characterize the effects of a pregnancy-induced epigenome in response to oncogenic stress, we established a transgenic mouse strain (CAGMYC) in which overexpression of the oncogene *cMYC*, a known inducer of mammary tumor development, is driven in a doxycycline-dependent manner. Using this transgenic mouse strain, we found that the postpregnancy epigenome was incompatible with *cMYC* overexpression, blocking the activation of cMYC-downstream signals and their progression to oncogenesis. Moreover, utilizing mammary fat pad transplantation and organoid cultures, we determined that cell-autonomous features of postpregnancy MECs in response to cMYC overexpression represent the basis for blocking their initial progression into oncogenesis. We are currently preparing these findings for publication.

Characterization of Factors That Block Breast Cancer Development and Progression

S. Cyrill, S.T. Yang, C. Chen, M. Ciccone

Our analysis of the effects of the postpregnancy epigenome in blocking cMYC overexpression oncogenic signals yielded several candidates as the drivers of pregnancy-induced breast cancer protection.

Among many, we found that Epha2, a tyrosine receptor kinase, and Tbx3, a transcription factor, both activated by cMYC-driven signals in prepregnancy CAGMYC MECs, failed to accumulate H3K27ac signals and cMYC DNA occupancy in postpregnancy CAGMYC MECs. Interestingly, deletion of Epha2 and TBX3 impacted tumor development and mammary gland development and therefore represents a mechanism by which postpregnancy CAGMYC MECs resist oncogenesis. To investigate this hypothesis, we have generated an array of reagents to study gain and loss of function of Epha2 and TBX3 in mammary epithelial cells to validate their role in preventing mammary tumor development.

In a complementary effort, we have also focused our efforts on the histone methyltransferase, Kmt2c (Mll3), an enzyme responsible for the deposition of single methyl groups at lysine 4 of histone 3 (H3K4me1), a modification that has been associated with poised gene regulation. Given that our findings showed that H3K4me1 levels are dynamically regulated throughout pregnancy cycles, we hypothesize that targeting Mll3 could alter the pregnancy-induced epigenome of MECs and their predisposition to cancer development. Thus, we generated murine mammary epithelial cell lines bearing CRISPR-induced mutations in the catalytic domain of Mll3 in order to identify guide RNAs that efficiently target Mll3 for future use in organoid cultures. Mll3 mutant cell lines ($n = 2$) were then transplanted into the fat pad of prepubescent mice and analyzed for their proliferation state during

mouse pregnancy and postlactation stage (involution). Preliminary histological analysis revealed mammary glands transplanted with Mll3 mutant cell lines suggested aberrant involution with a substantial increase in the number of ductal structures compared to those injected with Mll3 wild-type cell lines. We are currently setting up follow-up analysis to determine whether such histological abnormalities are accompanied by alteration of MEC epigenome and gene regulation.

The Effects of Pregnancy on Reprogramming the Immune Environment of Mammary Epithelial Cells

M. Feigman, M. Moss

Previously, our lab found that pregnancy induces a dramatic, long-lasting reorganization of the mammary epithelial epigenome. We have found that many of these epigenetic modifications drive expression of genes associated with immune cell communication. Specifically, we discovered CD1d overexpression in breast epithelium after pregnancy. Given that CD1d is a known marker for antigen presentation, we believe that alteration of CD1d levels after pregnancy and during malignant transformation may govern cancer susceptibility in mammary epithelial cells. To address this hypothesis, we have crossed transgenic mouse lines lacking CD1d expression (CD1d knockout [KO] mice) with our mouse model of pregnancy-induced breast cancer protection (CAGMYC) to probe for a role of CD1d KO MECs in decreasing mammary malignant development. Supporting our hypothesis, we found that CD1d KO cells develop into malignant lesions in response to *c-MYC* overexpression regardless of parity.

Given the role of CD1d in recruitment and maturation of immune cells, we characterized the immune composition of postpregnancy mammary glands. Using single-cell RNA-Seq and flow cytometry analysis, we found that the mammary immune population dramatically changes after pregnancy in mice, which included the presence of a new population of iNKT cells after pregnancy. We are currently utilizing a series of transgenic mice and transplantation assays to investigate whether there is a tangible link between increased expression of CD1d and presence of iNKT subpopulations after pregnancy and their role in blocking oncogenesis.

To Understand the Effect of an Immune Response to Infections on Pregnancy-Induced Mammary Gland Development and Breast Cancer

S. Cyrill, C. dos Santos

In prior studies, pregnancy-induced breast cancer protection was studied in rodents in which pregnancy signals were mimicked by the implantation of hormone pellets (pseudopregnancy). Thus, we tested whether pseudopregnancy would block the oncogenic effect of *cMYC* overexpression on MECs. We found that post-pseudopregnancy CAGMYC female mice developed urethral obstructions, leading to the development of urinary tract infections (UTIs), and that their mammary glands displayed tissue hyperplasia, suggesting that pregnancy-induced oncoprotection had been blocked.

Further analysis demonstrated that post-pseudopregnancy, non-cMYC-overexpressing mice harboring UTIs also displayed aberrant mammary histology marked by expansion of epithelial compartment. This result suggests that, independently of oncogenic signals, systemic alterations brought by immune response to infection affects mammary gland biology. We are currently using a series of approaches to investigate the immune environment and MEC epigenome from mice bearing UTI-like infections, in addition to screening for systemic signals that, in response to UTI infection, alter mammary gland tissue homeostasis and cancer predisposition.

To Define the Epigenomic Landscape of Organ-Specific Breast Cancer Metastasis

S. Cyrill, C. dos Santos [in collaboration with L. Van Aelst and M. Egeblad, CSHL]

The most aggressive breast cancers (stage IV) are characterized by metastasis. The heterogeneity that develops within the primary and metastatic tumors is largely effected by interactions with the organ-specific microenvironments, making the disease more aggressive and difficult to treat. However, not much is known about the organ-specific dependencies of breast cancer metastasis. In collaboration with the Van Aelst lab and Egeblad lab at CSHL, we will use murine models of metastatic breast cancer to study the role of the host microenvironment in metastasis,

with particular emphasis on epigenetic regulation and immunomodulation of these invasive cancer cells. To date, we have developed murine breast cancer cell lines (4T1) to express fluorescent proteins, sfGFP or mCherry, on their nuclear surface by tethering them to the nuclear envelope protein, SunI. The engineered cells were transplanted into the mammary fat pad of Balb/C mice and found to rapidly develop tumors that become aggressively metastatic. We are currently profiling the epigenome (ATAC-seq) and gene expression patterns (RNA-Seq) of fluorescently tagged 4T1 cells isolated from metastatic lesions present in the lungs and spleen of tumor-bearing Balb/C female mice, with the goal of defining specific epigenomic signatures underlying their organ-specific metastasis preferences.

PUBLICATIONS

Fein MR, He XY, Almeida AS, Bruzas E, Pommier A, Eberhardt A, Fearon DT, Wilkinson JE, dos Santos CO, Egeblad M. 2018. Cancer cell CCR2 orchestrates suppression of the adaptive immune response. bioRxiv doi:10.1101/390187

In Press

Feigman M, Moss M, Chen C, Cyrill SL, Ciccone FM, Frey W, Yang ST, Wilkison EJ, dos Santos CO. 2019. Pregnancy reprograms the enhancer landscape of mammary epithelial cells and alters the response to cMYC driven oncogenesis. bioRxiv doi:10.1101/642330

Kinberger SE, Ycas PD, Johnson JA, Chen C, Ciccone MF, Woo RWL, Urick AK, Zahid H, Shi K, Aihara H, et al. 2019. Selectivity, ligand deconstruction, and cellular activity analysis of a BPTF bromodomain inhibitor. *Org Biomol Chem* **17:** 2020–2027.

Slepicka P, Cyrill SL, dos Santos CO. 2019. Pregnancy and breast cancer: pathways to understand risk and prevention. *Trends Mol Med* (in press).

STRUCTURAL BIOLOGY OF NUCLEIC ACID REGULATORY PATHWAYS

L. Joshua-Tor B. Bibel, E. Elkayam, A. Epstein, C. Faehnle, S. Goldsmith, J. Ipsaro, M. Jaremko, J. Jung, K. Meze, K. On, J. Scheuring, R. Tseng

We study the molecular basis of nucleic acid regulatory processes—RNAi and DNA replication in particular. We use the tools of structural biology, biochemistry, and biophysics to study proteins and protein complexes associated with these processes to elucidate how they work. X-ray crystallography, cryo-electron microscopy, and other structural techniques enable us to obtain the three-dimensional structures of these molecular machines. Biochemistry, biophysics, and molecular biology allow us to study properties that can be correlated to their function and biology.

Mechanisms of RNAi and Noncoding RNAs

RNAi has made an enormous impact on biology in a very short period of time. Not only are we still discovering new cellular pathways for the regulation of gene expression involving these pathways, but RNAi has become an extraordinarily useful and simple tool for gene silencing. Almost from its beginnings, people have used genetics, biochemistry, molecular biology, and bioinformatics to study the mechanism of RNAi and related pathways. To get a true mechanistic understanding of these pathways, however, we must understand how the components of the RNAi machinery work at a molecular level.

Molecular Basis of piRNA Silencing

J. Ipsaro [in collaboration with A. Haase, NIH; M. Hammell, CSHL]

The stability of germline genomes is critical to the survival of a species. If unchecked, mobile genetic elements (transposons) are able to integrate into distant genomic sites and disrupt genome integrity. To protect against this, a germline-specific RNAi pathway has evolved to specifically and robustly repress transposon expression.

At the center of this critical defense pathway are Piwi proteins and their associated small RNAs (piRNAs). piRNAs are highly variable in sequence, which affords them the ability to silence many of the molecular targets that transposon transcripts present. Nonetheless, the diversity of Piwi-piRNAs is restricted by their preference of a uridine in the 5′-most position. The basis of this "1U" bias, particularly whether it is established by a Piwi protein or earlier in piRNA biogenesis, was unknown.

We found that the 1U bias of Piwi-piRNAs is established by sequential discrimination against all nucleotides other than U—first during piRNA biogenesis and then by interaction with Piwi's specificity loop. Additionally, sequence preferences during piRNA processing also limit uridine occurrences across the piRNA body. Together, these processes and the resulting 1U bias may modulate the efficacy of transposon silencing by piRNAs and provide a means for purifying selection in the ongoing arms race between germline genomes and their mobile genetic parasites.

ELTA: An Enzymatic Method to Label Free or Protein-Conjugated ADP-Ribose Monomer and Polymer

E. Elkayam [in collaboration with A. Leung, Johns Hopkins]

ADP ribosylation, the attachment of one or more ADP-ribose groups onto proteins, is a therapeutically important protein modification. The attached ADP-ribose monomer or polymer, known as PAR, modulates the activities of the modified substrates and/or their binding affinity to other proteins. However, there has been a woeful lack of tools to properly investigate these important modifications and their potential functions. In collaboration with

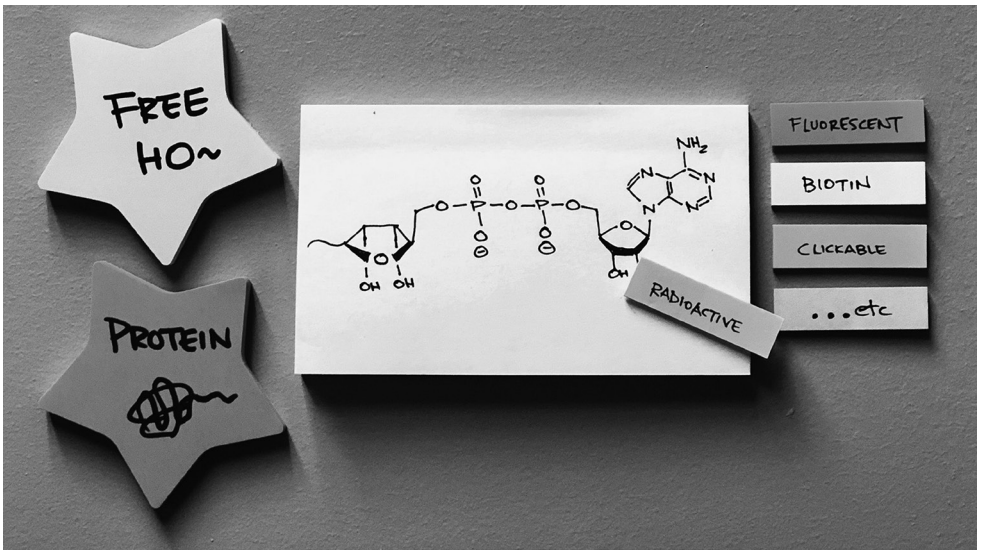


Figure 1. ELTA (enzymatic labeling of terminal ADP-ribose) is a method to label PAR modifications with radioactive, fluorescent, or otherwise tagged ADP. This enables simplified detection of PAR modifications, affinity measurements, assessment of PAR lengths, and enrichment of femtomole ADP-ribosylated peptides from complex mixtures.

Anthony Leung from Johns Hopkins University, we developed a simple, efficient, and versatile enzymatic labeling method, named enzymatic labeling of terminal ADP-ribose (ELTA) (Fig. 1), to label free or protein-conjugated ADP-ribose monomers and polymers using the enzyme OAS1 and dATP. ELTA can be coupled with a diverse range of chemical analogs of dATP (radioactive, fluorescent, biotin-tagged, clickable functional groups, and more) for various applications such as fluorescence-based biophysical measurement and biotin/click chemistry–based enrichment. Of particular importance, ELTA provides a timely tool to directly assess PAR length and distribution following treatment with FDA-approved PARP inhibitors that are now being administered for several cancer indications and others that are being tested. In addition, ELTA could be extended in the future to label other ADP-ribose derivatives with a free 2′-OH group such as the ADP ribosylation of the antibiotic rifamycin and the recently discovered modification of DNA ADP ribosylation.

PUBLICATIONS

Krishnan N, Bonham CA, Rus IA, Kumar Shrestha O, Gauss CM, Haque A, Tocilj A, Joshua-Tor L, Tonks NK. 2018. Harnessing insulin- and leptin-induced oxidation of PTP1B for therapeutic development. *Nature Comm* **9:** 283. doi:10.1038/s41467-017-02252-2

Li J, Wu C, Wang W, He Y, Elkayam E, Joshua-Tor L, Hammond PT. 2018. Structurally modulated co-delivery of siRNA and Argonaute 2 for enhanced RNA interference. *Proc Natl Acad Sci* **115:** E2696–E2705. doi:10.1073/pnas.1719565115

On KF, Jaremko M, Stillman B, Joshua-Tor L. 2018. A structural view of the initiators for chromosome replication. *Curr Opin Struct Biol* **53:** 131–139. doi:10.1016/j.sbi.2018.08.003

Zhang T, Wu Y-C, Mullane P, Ji YJ, Li H, He L, Arora A, Hwang H-Y, Alessi AF, Niaki AG, et al. 2018. FUS regulates activity of microRNA-mediated gene silencing. *Mol Cell* **69:** 787–801. doi:10.1016/j.molec.2018.02.001

In Press

Ando Y, Elkayam E, McPherson RL, Dasovich M, Cheng S-J, Voorneveld J, Filippov DV, Ong S-E, Joshua-Tor L, Leung AKL. 2019. ELTA: enzymatic labeling of terminal ADP-ribose. *Mol Cell* **73:** 845–856.e5. doi:10.1016/j.molec.2018.12.022

Stein C, Genzor P, Mitra S, Elchert AR, L. Benner L, Sobti S, Ipsaro JJ, Hammell M, Joshua-Tor L, Haase AD. 2019. Decoding the 5′ nucleotide bias of PIWI-interacting RNAs (piRNAs). *Nature Comm* **10:** 828. doi:10.1038/s41467-019-08803-z

RNA SPLICING

A.R. Krainer M. Daley, V. Kergourlay, Y.J. Kim, J. Layne, K-T. Lin, Y.H. Liu, W.K. Ma, L. Manche, L. Marasco, M. Rahman, J. Scharner, D. Segovia, N. Sivetz, L. Wan, M.S. Wong, Q. Zhang

Mechanisms of Constitutive and Alternative Pre-mRNA Splicing

RNA splicing is required for expression of most eukaryotic protein-coding genes. The spliceosome selects authentic splice sites with very high fidelity, relying on limited sequence information present throughout introns and exons. In humans, >90% of genes are expressed via alternative splicing, giving rise to multiple protein isoforms. The choice of alternative splice sites is commonly regulated to alter gene expression, either tissue-specifically or in response to a developmental program or to signaling pathways. The fact that multiple protein isoforms can be expressed from individual genes demonstrates that the classical "one gene–one enzyme" paradigm is no longer valid and provides an explanation for the unexpectedly small number of protein-coding genes uncovered by genome-sequencing projects.

Both constitutive and alternative splicing mechanisms involve numerous protein components, as well as five noncoding RNA components that are part of small nuclear ribonucleoprotein (snRNP) particles. The work in our lab focuses on the identification and molecular characterization of protein factors and sequence elements that are necessary for the catalysis and fidelity of splicing and/or for the regulation of alternative splice-site selection. We are interested in how the spliceosome correctly identifies the exons on pre-mRNA and how certain point mutations in either exon or intron sequences cause aberrant splicing, leading to various human genetic diseases. Related areas of interest include the remodeling of mRNP architecture as a consequence of splicing, which influences downstream events, such as nonsense-mediated mRNA decay (NMD); the various roles of alternative splicing misregulation in cancer; and the development of effective methods—particularly antisense technology—to correct defective splicing or modulate alternative splicing or for gene/allele-specific inhibition of NMD, in a disease context. A summary of some of our recently published studies is provided below.

Experimental and Computational Analyses of Splicing

Mutations at the 5′ splice site frequently cause defective splicing and disease, because of interference with the initial recognition of the exon–intron boundary by U1 snRNP, a key component of the spliceosome. In collaboration with J. Kinney (CSHL), we developed a massively parallel splicing assay (MPSA) in human cells to quantify the activity of all 32,768 unique 5′ splice site sequences (NNN/GYNNNN) in three different three-exon minigene contexts (*SMN1*, *BRCA2*, and *IKBKAP*). Our results revealed that although splicing efficiency is mostly governed by the 5′ splice site sequence, there are substantial differences in this efficiency across gene contexts. The rank order among 5′ splice site sequences is substantially similar in these contexts, but each context imposes an activity threshold. For example, the *IKBKAP* context is much more stringent than the other two, allowing only a small subset of strong 5′ splice sites to function. This restrictive context was relaxed by strengthening the upstream 3′ splice site, reflecting exon-definition interactions. Among other uses, these MPSA measurements facilitate the prediction of 5′ splice site sequence variants that are likely to cause aberrant splicing. Thus, our approach provides a framework to assess potential pathogenic variants in the human genome and streamline the development of splicing-corrective therapies.

Percent spliced-in (PSI) values are commonly used to report alternative pre-mRNA splicing changes. Previous PSI-detection tools were limited to specific alternative splicing events and were evaluated by in

silico RNA-Seq data. We developed PSI-Sigma, a computational tool that uses a new PSI index, and we employed actual (nonsimulated) RNA-Seq data from spliced synthetic genes (RNA Sequins) to benchmark its performance (i.e., precision, recall, false-positive rate, and correlation) in comparison with three leading tools (rMATS, SUPPA2, and Whippet). PSI-Sigma outperformed these tools, especially in the case of alternative splicing events with multiple alternative exons and intron-retention events. We also evaluated its performance in long-read RNA-Seq analysis, by sequencing a mixture of human RNAs and RNA Sequins with nanopore long-read sequencers. PSI-Sigma is available at https://github.com/wososa/PSI-Sigma.

Splicing Alterations in Genetic Diseases and Cancer, and Antisense Therapeutics

Familial dysautonomia (FD), or Riley–Day syndrome, is a rare inherited neurodegenerative disorder caused by a point mutation in the *IKBKAP* gene that results in defective splicing of its pre-mRNA. The mutation weakens the 5′ splice site of exon 20, causing this exon to be partially skipped and thereby introducing a premature termination codon. Although detailed FD pathogenesis mechanisms are not yet clear, correcting the splicing defect in the relevant tissue(s), and thus restoring normal expression levels of the full-length IKAP protein, could be therapeutic. Splice-switching antisense oligonucleotides (ASOs) can be effective targeted therapeutics for neurodegenerative diseases such as nusinersen (Spinraza®), an approved drug for spinal muscular atrophy (SMA), which we previously developed in collaboration with Ionis Pharmaceuticals. Using a two-step screen for splice-switching ASOs targeting *IKBKAP* exon 20 or the adjoining intronic regions, we identified a lead ASO that fully restores exon 20 splicing in FD patient fibroblasts (in collaboration with Ionis Pharmaceuticals). We also characterized the corresponding *cis*-acting regulatory sequences that control splicing of exon 20. When administered into a transgenic FD mouse model, the lead ASO promoted expression of full-length human *IKBKAP* mRNA and IKAP protein levels in several tissues tested, including the central nervous system. These findings provided insights into the mechanisms of *IKBKAP* exon 20 recognition and preclinical proof of concept for an ASO-based targeted therapy for FD.

Although motor-neuron degeneration is the hallmark of SMA, accumulating evidence indicates that it is a multisystem disorder, particularly in its severe forms. Several studies delineated structural and functional cardiac abnormalities in SMA patients and mouse models, yet the abnormalities have been primarily attributed to autonomic dysfunction. A collaborative study with Y. Hua's lab (Soochow University, China) and Ionis Pharmaceuticals showed that cardiomyocytes from a severe SMA mouse model undergo G_0/G_1 cell cycle arrest and enhanced apoptosis during postnatal development. Microarray and real-time RT-PCR analyses revealed that a set of genes associated with the cell cycle and apoptosis are dysregulated in newborn pups. Of particular interest, the *Birc5* gene, which encodes Survivin, an essential protein for heart development, was down-regulated as early as presymptomatic postnatal day 0. Cultured cardiomyocytes depleted of SMN recapitulated the gene-expression changes, including down-regulation of Survivin and abnormal cell cycle progression, and overexpression of Survivin rescued the cell cycle defect. Finally, increasing SMN in SMA mice with a therapeutic ASO improved heart pathology and rescued expression of deregulated genes. These data indicate that the cardiac malfunction of the severe SMA mouse model is mainly a cell-autonomous defect, caused by widespread gene deregulation in heart tissue, particularly of *Birc5*, resulting in developmental abnormalities through cell cycle arrest and apoptosis.

Pre-mRNA splicing can contribute to the switch of cell identity that occurs in carcinogenesis. We analyzed a large collection of RNA-Seq data sets and found that splicing changes in genes coding for hepatocyte-specific enzymes, such as *AFMID* (arylformamidase, involved in tryptophan metabolism, and hence NAD^+ production) and *KHK* (ketohexokinase), are associated with the survival and relapse of hepatocellular carcinoma (HCC) patients. We found that the switch of *AFMID* isoforms is an early event in HCC development and is associated with driver mutations in *TP53* (p53 tumor suppressor) and *ARID1A* (AT-rich interaction domain 1A, a SWI/SNF family member). Interestingly, this switch of *AFMID* isoforms is human-specific and not detectable in other species, including primates. We further showed that overexpression of the full-length *AFMID* isoform in HepG2 cells results in a higher NAD^+ level, lower DNA-damage response, and slower cell growth. Thus,

our integrative analysis uncovered a mechanistic link between alternative-splicing switches, de novo NAD^+ biosynthesis, driver mutations, and HCC recurrence.

Splice-switching ASOs are promising therapeutic tools to target various genetic diseases, including cancer. However, in vivo delivery of ASOs to orthotopic tumors in cancer mouse models or to certain target tissues remains challenging. A viable solution already in use is receptor-mediated uptake of ASOs via tissue-specific receptors. For example, the asialoglycoprotein receptor (ASGP-R) is exclusively expressed in hepatocytes. Triantennary *N*-acetylgalactosamine (GalNAc) (GN3)-conjugated ASOs bind to the receptor and are efficiently internalized by endocytosis, enhancing ASO potency in the liver. In collaboration with Ionis Pharmaceuticals, we explored the use of GalNAc-mediated targeting to deliver therapeutic splice-switching ASOs to cancer cells that ectopically express ASGP-R, both in vitro and in tumor mouse models. We found that ectopic expression of the major isoform ASGP-R1 H1a is sufficient to promote uptake and increase GN3-ASO potency to various degrees in four of five tested cancer cell lines. We also showed that cell type–specific glycosylation of the receptor does not affect its activity. Finally, we showed that in vivo, GN3-conjugated ASOs specifically target subcutaneous xenograft tumors that ectopically express ASGP-R1 and modulate splicing significantly more strongly than unconjugated ASOs. Our work shows that GN3 targeting is a useful tool for proof-of-principle studies in orthotopic cancer models, until endogenous receptors are identified and exploited for efficiently targeting cancer cells.

Mechanistic Links between Splicing and NMD

We previously reported that the splicing factor SRSF1 stimulates NMD, a quality-control mechanism that degrades mRNAs with premature termination codons (PTCs). We have now addressed the underlying molecular mechanism and found that transcript-bound SRSF1 increases the binding of the key NMD RNA helicase UPF1 to mRNAs while in (or associated with) the nucleus, bypassing UPF2 recruitment and promoting NMD. SRSF1 promotes NMD when positioned downstream of a PTC, which resembles the mode of action of exon junction complex (EJC) and NMD factors. Moreover, splicing and/or EJC deposition increase the effect of SRSF1 on NMD. Finally, we showed that SRSF1 enhances NMD of natural PTC-containing endogenous transcripts that result from various alternative-splicing events. Our findings revealed an alternative mechanism for UPF1 recruitment, uncovering an additional connection between splicing and NMD. SRSF1's role in the mRNA's journey from splicing to decay has broad implications for gene-expression regulation and genetic diseases.

PUBLICATIONS

Aznarez I, Nomakuchi TT, Tetenbaum-Novatt J, Rahman MA, Fregoso O, Rees H, Krainer AR. 2018. Mechanism of nonsense-mediated mRNA decay stimulation by splicing factor SRSF1. *Cell Rep* **23:** 2186–2198.

Lin KT, Ma WK, Scharner J, Liu YR, Krainer AR. 2018. A human-specific switch of alternatively spliced AFMID isoforms contributes to TP53 mutations and tumor recurrence in hepatocellular carcinoma. *Genome Res* **28:** 275–284.

Sheng L, Wan B, Feng P, Sun J, Rigo F, Bennett CF, Akerman M, Krainer AR, Hua Y. 2018. Downregulation of Survivin contributes to cell-cycle arrest during postnatal cardiac development in a severe spinal muscular atrophy mouse model. *Hum Mol Genet* **27:** 486–498.

Sinha R, Kim YJ, Nomakuchi T, Sahashi K, Hua Y, Rigo F, Bennett CF, Krainer AR. 2018. Antisense oligonucleotides correct the familial dysautonomia splicing defect in IKBKAP transgenic mice. *Nucleic Acids Res* **46:** 4833–4844.

Wong MS, Kinney JB, Krainer AR. 2018. Quantitative activity profile and context dependence of all human 5′ splice sites. *Mol Cell* **71:** 1012–1026.

In Press

Bennett CF, Krainer AR, Cleveland DW. 2019. Antisense oligonucleotide therapies for neurodegenerative diseases. *Annu Rev Neurosci* **42:** 385–406.

Feng H, Bao S, Rahman MA, Weyn-Vanhentenryck SM, Khan A, Wong J, Shah A, Flynn ED, Krainer AR, Zhang C. 2019. Modeling RNA-binding protein specificity in vivo by precisely registering protein–RNA crosslink sites. *Mol Cell* **20:** 1189–1204.

Lin KT, Krainer AR. 2019. PSI-Sigma: a comprehensive splicing-detection method for short-read and long-read RNA-seq analysis. *Bioinformatics* doi:10.1093/bioinformatics/btz438

Scharner J, Qi S, Rigo F, Bennett CF, Krainer AR. 2019. Delivery of GalNAc-conjugated splice-switching ASOs to non-hepatic cells through ectopic expression of asialoglycoprotein receptor. *Mol Ther Nucleic Acids* **16:** 313–325.

Wang E, Lu SX, Pastore A, Chen X, Imig J, Chun-Wei Lee S, Hockemeyer K, Ghebrechristos YE, Yoshimi A, Inoue D, et al. 2019. Targeting an RNA-binding protein network in acute myeloid leukemia. *Cancer Cell* **35:** 369–384.

REGULATION OF GENE EXPRESSION

D.L. Spector D. Aggarwal, G. Arun, B. Balasooriya, S. Barrera, S. Bhatia, L. Brine, K-C. Chang, R. Hazra, B. Liu, P. Naik, S. Russo, W. Xu, A. Yu

Most cellular processes can trace their beginnings to the nucleus, where a gene is activated resulting in the production of an RNA molecule—some of which encode proteins, whereas others are functional in the form of RNA. Although much biochemical information is available regarding many of the factors involved in gene expression, the spatial and temporal parameters that influence gene expression and the role of noncoding RNAs in regulating this multifaceted process are just beginning to be elucidated. Over the past year our research has continued to focus on identifying and characterizing the role of long noncoding RNAs (lncRNAs) in cancer progression and/or differentiation. In addition, we have been examining the role of lineage commitment in establishing random autosomal monoallelic gene expression and the role of lncRNAs in lineage commitment. An overview of some of our accomplishments over the past year follows.

Identification of lncRNAs Involved in Breast Cancer Progression

S. Bhatia, K-C. Chang, B. Liu, D. Aggarwal, W. Xu, A. Yu, S. Russo, L. Brine [in collaboration with F. Rigo, R. MacLeod, and C. Frank Bennett, Ionis Pharmaceuticals]

Large-scale genome-wide studies have indicated that thousands of RNAs that lack protein-coding capacity are transcribed from mammalian genomes. A subset of these noncoding RNAs are >200 nt in length and are referred to as long noncoding RNAs (lncRNAs). With breast cancer being the most frequent malignancy in women worldwide, we aim to identify lncRNAs that play roles in breast tumor progression and evaluate their mechanism of action and potential as therapeutic targets. We have previously reported on our studies focused on the lncRNA *MALAT1* as a potential therapeutic target impacting breast cancer progression. In addition to *MALAT1*, we have previously identified a number of potentially oncogenic lncRNAs, termed mammary tumor associated RNAs (MaTARs), in mouse models of human breast cancer. Our results suggest that these lncRNAs are likely important drivers of mammary tumor progression and represent promising new therapeutic targets.

MaTAR25 is an ~2,000-nt nuclear enriched and chromatin-associated lncRNA. *MaTAR25* is overexpressed in mammary tumors in the MMTV-PyMT (luminal B subtype), MMTV-NEU-NDL (HER2 subtype), and in triple-negative mammary cancer (TNMC) cells compared to normal mammary epithelial cells. Using CRISPR-Cas9 genome editing, *MaTAR25* knockout (KO) clones were generated in mouse TNMC 4T1 cells. When comparing *MaTAR25* KO cells to control 4T1 cells, we found a significant decrease in cell proliferation (–50%), cell motility (–40%), and invasion (–45%). Ectopic expression of *MaTAR25* in 4T1 *MaTAR25* KO cells rescued the KO phenotypes, indicating that *MaTAR25* lncRNA plays an important role in these processes. Injection of *MaTAR25* 4T1 KO cells into the mammary fat pad or tail vein of BALB/c mice resulted in a decrease in both tumor growth (–56%) and the number of lung metastatic nodules (–62%) compared to the 4T1 control group. Furthermore, antisense oligonucleotide (ASO)-mediated knockdown of *MaTAR25* in the MMTV-Neu-NDL mouse model resulted in a decrease in tumor growth (–59%) compared to the scrambled ASO control group. By comparing differentially expressed genes from *MaTAR25* KO cells with *MaTAR25*-targeted genes from chromatin isolation by RNA purification (ChIRP) analysis, we identified a total 446 overlapping genes. One of the top genes identified is *Tensin1*, the protein product of which positively regulates cell migration. By qRT-PCR and western blot analyses, we found that the expression level of Tensin1 is lower in *MaTAR25* KO cells than in control cells, and overexpression of *MaTAR25* in *MaTAR25* KO cells can also

up-regulate the level of Tensin1 in these cells. In addition, overexpression of Tensin1 in *MaTAR25* KO cells can rescue the cell viability phenotype. These results suggest that *MaTAR25* regulates expression of Tensin1 through binding to its DNA sequence, and Tensin1 is one of the major downstream targets of *MaTAR25* regulating tumor cell progression. Analysis of synteny between the mouse and human genomes, combined with TCGA RNA sequencing data from individuals with breast cancer, has allowed us to identify the human ortholog of *MaTAR25* (*hMaTAR25*), which has been confirmed by rescue experiments. Together, our data suggest that *hMaTAR25* may have diagnostic and/or therapeutic potential as a novel drug target in breast cancer progression.

MaTAR42 was found to be up-regulated in MMTV-PyMT and MMTV-Neu-NDL mammary tumors compared to normal mammary epithelial cells. The *MaTAR42* gene is located on mouse chromosome 4, with a conserved potential human ortholog identified on human chromosome 9. RACE, qRT-PCR, and northern blot analyses verified that the major transcript expressed in MMTV-PyMT mouse mammary tumor cells is a 2,821-nt single-exon transcript with low protein coding potential as indicated by computational prediction. qRT-PCR verified that the potential MaTAR42 human ortholog is up-regulated in several breast tumor–derived organoids compared to adjacent normal breast organoids. Cellular fractionation followed by qRT-PCR confirmed that *MaTAR42* is predominantly located in the nucleus. *MaTAR42* is up-regulated when cells are grown in suspension, and it is down-regulated to a basal level when cells are grown attached to a substrate. Analysis of the *MaTAR42* promoter, which is conserved in mammals, suggests a possible regulation by the GATA signaling pathway. Antisense oligonucleotide–mediated knockdown of *MaTAR42* in MMTV-PyMT tumor organoids resulted in a significant reduction in organoid size and cell viability. Further analysis in MMTV-PyMT cell lines confirmed that the loss of viability was not due to a cell proliferation defect, but rather to a loss of adhesion to the Matrigel extracellular matrix (ECM) substrate. These results suggest a potential role of *MaTAR42* in cellular mechanisms sensing the loss of cell–ECM contacts and further suggest a possible role of *MaTAR42* in colonization after extravasation during tumor metastasis, which is the major future direction of the project.

Human Breast Tumor Organoid Project

G. Arun, S. Bhatia, S. Russo, S. Barrera, P. Naik
[in collaboration with K. Kostroff, Northwell Health]

Recent studies have demonstrated that various tumor types can be propagated in the laboratory as 3D tumor organoids. These organoids retain many of the structural and functional complexities of the tumor and can be used to study tumor composition, heterogeneity, and development ex vivo. Tumor organoids, therefore, provide a very innovative and unique platform to study cancer, as they can recapitulate many aspects of the disease with high fidelity. As such they represent an excellent system for identifying new therapeutic targets and for drug development and screening in a patient-specific manner. Our goal has been to develop a human breast tumor organoid biobank that can be used to study the role of lncRNAs in breast cancer progression and establish a therapeutic platform for rapid screening of cancer-relevant lncRNAs.

Thus far, we have established 50 breast tumor organoid models along with adjacent or distal normal organoids from a subset of these patients. Additionally, we have established normal mammary organoids from individuals who have undergone reductive mammoplasty. We have performed DNA sequencing to identify potential mutations from a panel of 143 pan-cancer driver genes. We currently have mutation data on 23 of the patient-derived tumor organoids. Additionally, to test whether the ex vivo–derived tumor organoids retain the mutation profile of the original tumor, we have sequenced seven tumor tissue and tumor organoid pairs. These data show that the majority of the mutations of the original tumor were retained in the cultured organoid system. Furthermore, we find that ~36% of the sequenced tumor organoids have oncogenic mutations in the gene *PIK3CA*, whereas 14% have mutations in cadherin genes (*CDH1*). Interestingly, all of the samples with a *CDH1* mutation belong to the organoids derived from lobular carcinomas and have a loosely organized morphology in culture. RNA sequencing has been performed from 35 breast tumor organoids and 21 normal organoids. The data show very high patient specificity in the organoid profiles, in which the normal and tumor organoids from the same patient cluster closely together—providing strong evidence for studying breast tumors in a patient-specific manner. We currently have ongoing projects that are using this patient-specific approach to identify and

study lncRNAs up-regulated in the luminal A and triple-negative breast cancer subtypes.

Probing the Role of a Highly Expressed Long Noncoding RNA in Hepatocellular Carcinoma

A. Yu [in collaboration with C. Berasain, University of Navarra, Pamplona, Spain; F. Rigo, Ionis Pharmaceuticals]

Hepatocellular carcinoma (HCC), the most common type of liver malignancy, is one of the most lethal forms of cancer. HCC is not diagnosed until late stages and has a poor five-year survival rate of <14%. Excluding liver transplantation, the current standard of care for HCC is treatment with sorafenib, a multikinase inhibitor that targets Raf, receptor tyrosine kinases, and platelet-derived growth factor receptor, which extends median survival time from 7.9 mo to 10.7 mo. This modest gain emphasizes the urgent need to identify new and effective therapeutic targets for HCC. Genome-wide analyses such as the ENCODE (Encyclopedia of DNA Elements) project have revealed that most of the genome is transcribed, even though <2% of the genome encodes proteins. Thousands of transcripts >200 nt in length (lncRNAs) are expressed in a tissue-specific manner and undergo changes in expression level during cellular differentiation and in cancers. lncRNAs have been implicated in numerous molecular functions including modulating transcriptional patterns, regulating protein activities, serving structural or organizational roles, altering RNA processing events, and serving as precursors to small RNAs. One lncRNA (lnc05) is of particular interest in that it is enriched in mouse hepatocellular carcinoma cell lines and it is conserved in humans. Lnc05 is up-regulated in HCC cells as well as during hepatocarcinogenesis, as compared to normal mouse hepatocyte cell lines. Knockout of lnc05 in HepA1-6 cells using CRISPR-Cas9 genome editing results in a 23% increased doubling time or a 47% reduction in proliferation. In addition, weighted-gene correlation network analysis (WGCNA) places lnc05 in a module that is enriched with cell cycle components. Moreover, knockdown of lnc05 expression using antisense oligonucleotides (ASOs) decreases proliferation, as assessed by expression levels of Ki67, a proliferation marker, and colony formation assays. RNA antisense pulldown assays identified nucleolysin TIAR, a translational repressor, as a potential binder of lnc05. Modulating lnc05 levels also affects one of TIAR's downstream targets, c-Myc, in a positively correlated fashion. We have identified a 71-nt hairpin within lnc05 that harbors four TIAR binding sites. These appear as distinct bands when assayed for gel shifts and can be successfully competed out using a cold probe. Thus, we believe lnc05 may be sequestering TIAR in order to allow for c-Myc translation. In summary, our data suggest that lnc05 may be targeted as a therapy for HCC and/or have diagnostic potential in liver cancer.

The Role of the Long Noncoding RNA, *Platr4*, in Lineage Commitment

R. Hazra, L. Brine

lncRNAs are differentially expressed in a development-specific manner across tissues, suggesting regulatory roles in cell fate decision and differentiation. We have identified the functional role of a novel ESC-specific lncRNA, *Platr4* (pluripotency-associated transcript 4), in cell fate determination. *Platr4* is a 1,035-nt poly(A)$^+$ transcript comprising two exons. Cellular fractionation of mouse ESCs indicates that *Platr4* is mainly present in the nuclear fraction and is associated with chromatin. Using the CRISPR-Cas9 genome-editing system, we have generated mESC lines (V6.5 and AB2.2) with the deletion of the *Platr4* promoter resulting in a significant depletion (homozygous deletion, *Platr4*$^{-/-}$) of the *Platr4* transcript (up to 99%) as measured by qRT-PCR and single-molecule RNA-FISH analysis. Deletion of *Platr4* in mESCs did not affect cell cycle kinetics or pluripotency, whereas we identified abnormalities in the spontaneous contraction of embryoid bodies (EBs) compared to control cells, which was further confirmed by a decreased level of the major contractile protein transcripts—myosin (*myh7b*) and troponin (*Tnnt3*)—during ESC differentiation. Further, gene set enrichment analysis has predicted that *Platr4* depletion in ESCs affects mesendoderm (ME) lineages, which was further confirmed by measurements determining the level of the ME transcripts, Sox17, Foxa2, Bra(T), and Eomes. In addition, iRegulon analysis (using differentially expressed genes from RNA-Seq data) predicted the Tead transcription factor, which is one of the known co-activators of ME specification. Together, these findings indicate that *Platr4* is

an important lncRNA regulator of ME specification during mammalian development.

Random Autosomal Monoallelic Gene Expression and Differentiation

B. Balasooriya

Random autosomal monoallelic (RAM) gene expression describes the transcription from only one of two homologous alleles of a particular gene in a diploid cell. RAM gene expression may partly explain the variable penetrance of disease-associated mutations, and hence this outcome of gene expression warrants extensive investigation in terms of how it is initiated, maintained, and regulated among a clonal population of cells. In this regard, over the past year we set out to investigate the spatiotemporal factors that are involved in the establishment and maintenance of RAM gene expression. An allele-specific live-cell imaging system was developed using differentially labeled alleles for two different monoallelically expressed genes: *Cth* and *Ttc4* using CRISPR-Cas9-D10A technology. First, we developed doxycycline (Dox)-inducible mCherry-fused PP7 coat protein (mCherry-PP7CP) and EYFP-fused MS2 coat protein (EYFP-MS2CP) expressing diploid F_1 mESCs. EYFP-MS2CP and mCherry-PP7CP sequences were independently cloned into the 3′ end of the *P*TRE3GS element in the Tet-One inducible system (TakaRa®). SNP-specific guide RNAs (gRNAs) were used to independently knock in these expression cassettes into the 3′ UTR of CAST/EiJ and C57BL/6J alleles of the *Rosa26* locus, and then single cell–derived double knock-in clones were established. Dox-inducible dosage-dependent EYFP and mCherry expression in these clones was confirmed by immunofluorescence microscopy and immunoblotting. The knock-in clones contain low-background fluorescent signals because of the modest activity of the *P*hPGK promoter. This will be an advantage in image analysis, as significant background noise reduction will not be necessary. Second, 24 repeats of a PP7 stem-loop (PP7-SL) cassette were knocked into the 3′ UTR of the CAST/EiJ allele and 24 repeats of an MS2 stem-loop (MS2-SL) cassette into the 3′ UTR of the C57BL/6J allele of the *Cth* and *Ttc4* genes using allele-specific SNPs. To date, single cell–derived EYFP- and mCherry-inducible MS2 and PP7 stem-loop double knock-in F_1 mESC clones have been established for both the *Cth* and *Ttc4* genes. The nascent *Cth* and *Ttc4* transcripts were visualized at 36 h post-Dox addition in single cell–derived clones. Using spinning disk confocal microscopy, one "green dot" and one "red dot" were detected per nucleus in 50%–70% of cells in the F1 ESC clones. Interestingly, when biallelically expressed, the two dots were observed in very close proximity to each other. Moreover, one "green dot" was observed per nucleus in F1 mESC-derived neural progenitor cells (NPCs), indicating monoallelic expression. These preliminary results demonstrate the inactivation of the CAST/EiJ allele of the *Cth* and *Ttc4* genes in NPCs, confirming the successful generation of the allele-specific imaging system.

PUBLICATIONS

Arun G, Diermeier SD, Spector DL. 2018. Therapeutic targeting of long non-coding RNAs in cancer. *Trends Mol Med* **24:** 257–277.

Huang Y-H, Klingbeil O, He X-Y, Wu XS, Arun G, Lu B, Somerville TDD, Milazzo JP, Wilkinson JE, Demerdash OE, et al. 2018. POU2F3 is a master regulator of a tuft cell–like variant of small cell lung cancer. *Genes Dev* **32:** 915–928.

In Press

Abulwerdi FA, Xu W, Ageeli A, Yonkunas M, Arun G, Nam H, Schneekloth Jr JS, Dayie TK, Spector D, Baird N, Le Grice SFJ. 2019. Selective small-molecule targeting of a triple helix encoded by the long non-coding RNA, *MALAT1*. *ACS Chem Biol* **14:** 223–235.

Arun G, Spector DL. 2019. *MALAT1* long non-coding RNA and breast cancer. *RNA Biol* **16:** 860–863. doi:10.1080/15476286.2019.1592072

MOLECULAR BIOLOGY OF PAPILLOMAVIRUSES

A. Stenlund M. Kornaj

The papillomaviruses are a group of viruses that infect and transform the basal epithelium, inducing proliferation of the cells at the site of infection. The resulting tumors (warts) are, in most cases, benign and will usually regress after some time, but certain types of human papillomaviruses (HPVs) give rise to tumors that are prone to progress toward malignancy—especially, frequently, cervical carcinoma. Indeed, HPV infection appears to be a necessary cause of invasive cervical carcinoma and thus represents one of the few firmly established links between viral infections and the development of cancer.

An impediment to the study of papillomaviruses has been the inability to define simple in vitro cell culture systems for analysis of the viral life cycle. These viruses normally require specialized differentiating cells that only with difficulty can be generated in cell culture. However, for a bovine papillomavirus (BPV-1), a convenient cell culture system exists in which viral gene expression, oncogenic transformation, and viral DNA replication can be studied. Thus, BPV has become a useful model for these aspects of the viral life cycle. The DNA replication properties of the papillomaviruses show some unique and interesting characteristics. As part of their normal life cycle, these viruses can exist in a state of latency, which is characterized by maintenance of the viral DNA as a multicopy plasmid in infected cells. The copy number of the viral DNA is tightly controlled, and the viral DNA is stably inherited under these conditions. Papillomaviruses, therefore, provide a unique opportunity to study plasmid replication in mammalian cells. In addition, the viral DNA replication machinery represents one of the most promising targets for antiviral therapy.

In previous years, we have reported the characterization of the papillomavirus replicon and the identification of the viral components that are required for viral DNA replication. In recent years, we have directed our attention toward the biochemical events that are associated with initiation of DNA replication. We are studying the biochemical properties of the viral E1 and E2 proteins and how these two proteins interact with the viral origin of DNA replication and with the cellular replication machinery to generate initiation complexes. Our studies show that the E1 protein has all the characteristics of an initiator protein, including ori recognition, DNA-dependent ATPase activity, and DNA helicase activity. The transcription factor E2, whose precise function has remained more elusive, appears to serve largely as a loading factor for E1. Through direct physical interactions with both E1 and the ori, E2 provides sequence specificity for the formation of the initiation complex.

We are currently attempting to elucidate how the E1 and E2 proteins orchestrate the precise biochemical events that precede initiation of DNA replication at the viral ori. These events include binding of the initiator to the ori, the initial opening of the DNA duplex (melting), and the assembly and loading of the E1 replicative helicase at the replication fork. Our studies so far indicate that these activities are generated in an ordered process that involves the sequential assembly of E1 molecules on the ori. This sequential assembly generates different complexes with different properties that, in turn, recognize ori, destabilize the double helix, and function as the replicative DNA helicase.

The papillomavirus E1 protein is a representative of a subgroup of hexameric helicases. However, E1 is not a dedicated helicase, as it also serves several additional functions in the viral life cycle. E1 takes part in initiation of viral DNA replication in several different ways: The E1 DNA-binding domain (DBD) provides sequence-specific binding of E1 to the viral origin of replication (ori). A particular form of E1, a DT, can then generate local melting of the viral ori. Once the DNA is melted locally, an E1 DH is formed on the melted DNA and unwinds the DNA in front of the replication fork.

The use of a single polypeptide to provide multiple functions presents many challenges. In addition to accommodating multiple different activities in one polypeptide, mechanisms have to exist that allow switching between the different activities. Furthermore, mechanisms have to exist to prevent the different activities from interfering with each other. For example, in the DT to DH transition, E1 switches from site-specific

DNA binding and complex formation dictated by E1-binding sites in the ori to formation of hexamers dictated by the oligomerization domain and binding to single-stranded DNA (ssDNA). How this feat is accomplished is not understood.

The Formation of Functional E1-DNA Complexes That Can Serve to Initiate Viral DNA Replication Involves the Use of Multiple Different DNA-Binding Activities Present in E1

We have previously shown that the E1 protein supports two distinct DNA-binding activities. One of these is the E1 DBD, which binds to specific sites in the origin of DNA replication and directs binding of E1 as a head-to-head dimer. A second DNA-binding activity consists of a β-hairpin structure in the helicase domain, which directs non-sequence-specific DNA binding. These two elements are required to form the functional DT and DH complexes. In addition to the DT and DH complexes that form on the origin of DNA replication, E1 can form two sequence-independent complexes on double-stranded DNA. These complexes, a trimer in the presence of ADP and a hexamer that depends on ATP hydrolysis, provide information about the intrinsic ability of E1 to oligomerize and are thought to relate to the formation of the DT and DH, respectively.

We wanted to determine whether the formation of the trimer and hexamer also relies on the same two DNA-binding elements that are important for formation of the DT and DH. We therefore generated mutations in the DBD and β-hairpin in the helicase domain, in residues involved in DNA contacts, and tested these for trimer and hexamer formation. This result was surprising in several ways. Although the β-hairpin was required for trimer formation, the result of mutating this structure was not complete loss of complex formation; instead, in the absence of a functional β-hairpin, E1 formed a complex containing two E1 molecules. This result indicates that the trimer, in fact, is a mixed complex in which two E1 molecules are bound in one way, whereas the third molecule is bound in a different way, and only the binding of the third molecule relies on the β-hairpin structure. This result explains a number of earlier observations that indicated the E1 molecules in the trimer are bound to the DNA in a staggered fashion based on the effects of probe length on complex formation.

The results of mutating the DBD were also unexpected. The DNA-binding surface of the DBD consists of two structural elements, the DNA-binding loop and DNA-binding helix, both of which are absolutely required for binding of E1 as a sequence-specific dimer. Surprisingly, although the DNA-binding helix did not contribute to trimer formation, one DNA-contacting residue in the DNA-binding loop was required for trimer formation. This result shows that the individual components of the DNA-binding surface of the DBD are used independently for the formation of E1 complexes.

Because the β-hairpin and DBD are not sufficient for trimer formation, we wanted to examine E1 for additional DNA-binding activities. Although E1 is well studied, functions have not been assigned to all parts of the protein and, especially, the amino-terminal ~150 residues remain mysterious. A part of this sequence is involved in nuclear import and export, but these sequences account for only a fraction of the amino-terminal domain. Given the parsimony that usually characterizes viral proteins, it is likely that additional functions are present in the amino-terminal domain. We identified a DNA-binding activity present in the amino-terminal domain by generating fragments from this domain and expressing these as GST fusion proteins in *Escherichia coli*. In this manner we could pinpoint a nonspecific DNA-binding activity located between residues 102 and 120. Point mutations that disrupted the DNA-binding activity of this peptide failed to form the DT complex, and instead resulted in formation of a DH complex, consistent with a role for this peptide in DT formation.

In addition to its DNA-binding activity, this peptide has an additional function in trimer and DT formation. The 102-120 peptide directly interacts with the oligomerization domain of the E1 protein and prevents oligomerization that is required for hexamer formation. This interaction prevents E1 from forming the hexamer and DH complexes. Hydrolysis of ATP by the E1 helicase domain releases the 102-120 peptide from the oligomerization domain and, instead, allows the 102-120 peptide to interact with a short peptide from the amino terminus where it is sequestered. As a result, as a consequence of ATP hydrolysis, E1 switches from forming a DT complex to forming a DH ring complex.

To clarify how the different DNA-binding activities in E1 combine to form the different E1 complexes important for the function of E1, we tested mutants in the different DNA-binding peptides for complex formation. We found that a total of five DNA-binding activities, in different combinations, are involved in E1 complex formation. Forming the trimer complex requires a combination of the DNA-binding loop in the DBD, β-hairpin, and $E1_{102\text{-}120}$ peptide. Formation of the hexamer complex requires a combination of the β-hairpin and both the DNA-binding loop and DNA-binding helix in the E1 DBD, which, under these conditions, bind DNA without sequence specificity. Forming the DT complex requires four of these DNA-binding activities. The β-hairpin, $E1_{102\text{-}120}$ peptide and the sequence-specific DNA-binding activity present in the E1 DBD, as well as the non-sequence-specific DNA binding by the DNA-binding loop in the E1 DBD, all play a role in DT formation. Finally, forming the DH complex requires all five DNA-binding activities.

Our data show that the five DNA-binding activities are not all exposed and accessible at the same time. For example, the DNA-binding activity of the $E1_{102\text{-}120}$ peptide is not detectable in the context of the intact amino-terminal domain. However, successive deletions from the amino terminus gradually increase the DNA-binding activity of this fragment. These results show that the DNA-binding activity present in $E1_{70\text{-}120}$ is hidden or sequestered in the presence of the intact amino terminus. Similarly, the E1 DBD can be inactivated through an interaction with a short peptide in the amino-terminal domain. These two interactions are mutually exclusive, resulting in the exposure of either one or the other of the two DNA-binding activities. The toggling between the two different DNA-binding activities allows E1 to transition from one kind of complex (i.e., the DT) to another kind of complex (i.e., the DH). Although we do not understand the exact mechanism of the switch, it is clear that it is ATP binding and hydrolysis by E1 that causes the toggle to change its position.

DNA REPLICATION AND CHROMATIN INHERITANCE

B. Stillman A. Amendolara, K. Bhalla, H-C. Chou, M. Hossain, Y. Hu, S. Nandi, F. Nasrin, M. Quyang, J. Shapp, Y-J. Sheu, M. Tramantano

For many years we have focused on understanding the process of DNA replication and chromatin inheritance by studying the genome of the budding yeast *Saccharomyces cerevisiae*. While that continues, we have increasingly begun comparing the inheritance of DNA and chromatin in human cells, which, particularly for the initiation of DNA replication, appears to be quite different from the yeast paradigm. In the budding yeast, the initiator protein, the origin recognition complex (ORC), is a six-subunit protein that is stable throughout the cell division cycle and remains bound to origins of DNA replication. Its initiator partner protein Cdc6, which is highly related to the large Orc1 subunit, is synthesized during late G_2 phase and then binds to ORC following mitotic exit. Cdc6 cooperates with ORC to assemble pre-replicative complexes (pre-RCs) in G_1 phase, but it is then degraded as cells transition from G_1 to S phase (see Fig. 1). The pre-RC consists of ORC and a double hexamer of the Mcm2-7 DNA helicase component that is loaded onto origin DNA by ORC and Cdc6, in cooperation with the Cdt1 protein that binds each Mcm2-7 hexamer prior to its loading onto origins. Cdt1 is removed as each Mcm2-7 hexamer is loaded onto origin-bound ORC.

In contrast, the scenario in human cells is far more complicated. First of all, the six-subunit ORC is not stable during the cell division cycle. ORC1 is degraded at the G_1 to S phase transition by a SKP2-dependent ubiquitin ligase (SCF^{SKP2}), is resynthesized in late G_2 phase, and, following nuclear envelope breakdown, ORC1 binds to the condensed mitotic chromosomes in prophase. The other ORC subunits remain in the nucleosol, not bound to chromosomes. Upon mitotic exit, the other ORC subunits bind to ORC1 on DNA; however, it is not known where on DNA ORC binds and how this assembly happens. Second, the human CDC6 protein is degraded as cells exit from mitosis and is resynthesized in mid-G_1 phase if cells commit to enter into another round of cell division (Fig. 1),

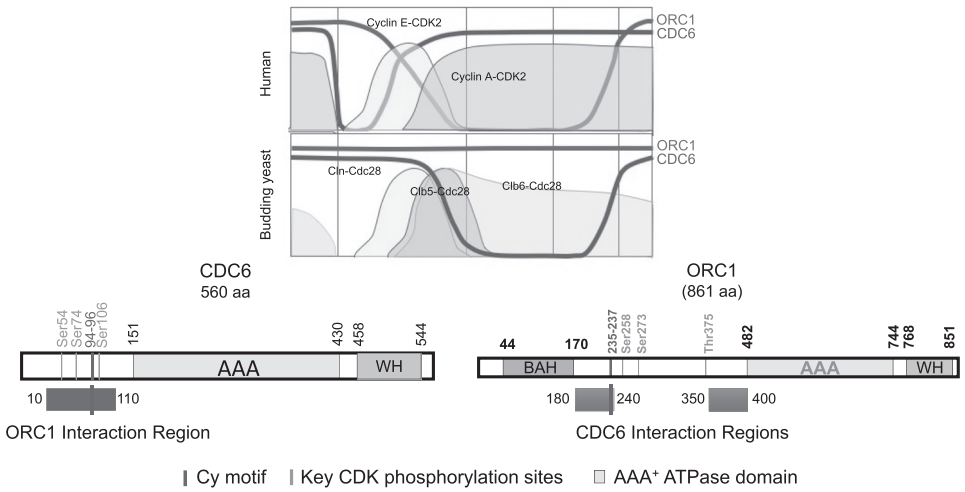


Figure 1. The *top* panel compares the relative levels of ORC1 and CDC6, as well as the cyclin-CDKs that control DNA replication in both the budding yeast *S. cerevisiae* and human cells. These differences contribute significantly to the varied mechanism and control of the initiation of DNA replication between yeast and human cells. The *bottom* panel shows maps of human ORC1 and CDC6 and the location of their interaction sites, with the Cy motifs shown in red.

but when and how it binds to ORC is not known. The CDT1 protein is synthesized during G_2 phase and is bound to the Geminin protein that prevents its interaction with the MCM2-7 hexamer. Geminin is degraded as cells transition from metaphase to anaphase in mitosis and the released CDT1 is inherited into the next G_1 phase. After it functions to load the MCM2-7 proteins during G_1, CDT1 is degraded during S phase in a process that depends on the DNA polymerase accessory protein PCNA. We have previously shown that the human cell MCM2-7 proteins bind chromatin during mid-G_1 phase, but it was not known how the pre-RC is assembled in human cells and where on DNA this complex is assembled.

In previous reports, we demonstrated that human ORC binds to the retinoblastoma (Rb) protein and to the histone H3 methyltransferase SUV39H1, and all three are required for silencing of gene expression at promoters that are regulated by the E2F1-DP1 transcription factors, notably the promoter for the *CCNE1* gene encoding cyclin E. The cyclin D-CDK4/6 kinase phosphorylates Rb and initiates the derepression of *CCNE1* expression, but cyclin E-CDK2 kinase bound to CDC6 is required for full derepression, demonstrating that CDC6 antagonizes ORC1 repression of gene expression in a feedback mechanism that ensures commitment to cell proliferation.

Because of the involvement of ORC and CDC6 in both pre-RC assembly and the control of gene expression in the G_1 phase, we have studied the interaction between ORC and CDC6 and have uncovered multiple new roles for a short linear peptide motif present in both ORC1 and CDC6 that was previously shown to regulate cyclin-dependent kinase activities during the cell division cycle.

Multiple interactions between ORC and CDC6. Using purified ORC subunit and CDC6 proteins, we demonstrated that CDC6 binds both ORC1 and ORC2. This was consistent with the X-ray and cryo-EM structure of ORC and the model of how CDC6 would interact that was determined in collaboration with Leemor Joshua-Tor's laboratory and reported last year. We mapped the regions within ORC1 that bind CDC6 and found two independent binding domains (Fig. 1): domain ORC1-1 (amino acids 180–240) and domain ORC1-2 (aa 350–400). The first domain, ORC1-1, contains a cyclin-binding (Cy) motif that has a consensus sequence of R/KxL that is required for ORC1 to bind cyclin A-CDK2, but is dispensable for the interaction between ORC1 and cyclin E-CDK2. Although mutation of the Cy motif in ORC1 prevented cyclin A-CDK2 from binding, it did not prevent ORC1 from binding to CDC6. The second domain, ORC1-2, is related to an intrinsically disordered, lysine-rich sequence in the yeast Orc1 protein that interacts with DNA when ORC is first recruited to the origin DNA, but not later when the first Mcm2-7 hexamer is loaded. Strikingly, the ORC1-1 and ORC1-2 domains interact with each other, and this self-interaction requires the Cy motif in the ORC1-1 region and a basic patch of amino acids in the ORC1-2 region. The region within CDC6 that binds ORC1 is amino acids 10–110, and, interestingly, this domain also contains a Cy motif that is essential for CDC6 to bind to ORC1. This Cy motif in CDC6 also mediates binding of CDC6 to both cyclin E-CDK2 and cyclin A-CDK2 kinases.

The involvement of the Cy motifs in ORC1 and CDC6 is interesting because cyclin A-CDK2 kinase, which binds the Cy motif in both ORC1 and CDC6, prevents ORC interacting with CDC6 by phosphorylating ORC1. In contrast, cyclin E-CDK2 kinase, which also binds to ORC1, but in a different region than cyclin A-CDK2, cannot block the interaction between these two proteins—rather, cyclin E-CDK2 kinase enhances the interaction between ORC and CDC6. Cyclin A-CDK2 kinase is active from S phase through the metaphase to anaphase transition in mitosis, whereas cyclin E-CDK2 kinase is only active during a short window in G_1 phase, after commitment to cell division. ORC1 and CDC6 are both present in mitotic cells before metaphase, as is cyclin A-CDK2, and thus ORC1 and CDC6 should not interact during this period, and studies in synchronized human cells confirmed that ORC1 and CDC6 do not interact in mitosis. However, when the ORC1 Cy motif was mutated, the mutant ORC1 now bound to CDC6 in mitosis. Because ORC, CDC6, CDT1, and Mcm2-7 are all present during mitosis, we have uncovered a new regulatory step to prevent premature pre-RC assembly during mitosis that involves cyclin A-CDK2 blocking ORC-CDC6 interaction. It is known that Geminin, which binds CDT1, also prevents pre-RC assembly at this time.

When cells progress from metaphase to anaphase, Geminin is degraded by APC^{CDC20}, whereas CDC6 and cyclin A are degraded by APC^{CHD1} upon mitotic exit—both processes requiring the anaphase-promoting complex-cyclosome (APC/C) ubiquitin ligase. In the absence of CDC6 and cyclin A in early S phase,

the pre-RC assembly process is halted, but during this period ORC1 bound to Rb and SUV39H1 to repress E2F1-DP1-dependent gene transcription. Upon commitment to cell division by cyclin D-CDK4/6 kinase activity, cyclin E and CDC6 are resynthesized. We have shown, using highly synchronized human HeLa cells, that when CDC6 reappears in the middle of G_1 phase, it binds to ORC1 in a complex that includes other ORC subunits, but not Rb. This is the time when the pre-RC can form, loading MCM2-7 proteins onto the chromatin. It was shown many years ago by Ron Laskey's laboratory that overexpression of both CDC6 and cyclin E greatly shortens G_1 phase and promotes early DNA replication, a mutagenic event. Indeed, both cyclin E and CDC6 are oncoproteins and are overexpressed in many cancers.

Because cyclin A-CDK2 phosphorylates ORC1 and this phosphorylation blocks binding to CDC6, ORC1 must be dephosphorylated upon mitotic exit and during G_1 phase to allow binding to CDC6 when it is resynthesized. We noticed that ORC1 interacted with protein phosphatase 1 (PP1) and found a PP1 binding motif in ORC1, similar to the RVxF binding motif found in other PP1 substrates. Mutation of this motif prevented PP1 from binding to ORC1, and removal of phosphate groups from ORC1 in early G_1 phase was partially impaired. Expression of the mutant ORC1 that cannot bind PP1 in human cells caused a dominant inhibition of cell proliferation, suggesting that removal of the phosphate groups on ORC1 is critical for its activity in G_1 phase.

Mutation of the ORC1 Cy motif also caused a dominant-negative phenotype by inhibiting cell proliferation, but interestingly it also prevented the degradation of ORC1 as cells transitioned from G_1 into S phase. The cells had clear defects in progression through S phase and mitosis, suggesting that in human cells, ORC1 must be degraded for normal cell cycle progression.

Previously we have reported that degradation of ORC1 is mediated by the SKP2-dependent ubiquitin E3 ligase. We have now shown that cyclin A-CDK2 binds to SKP2 and that cyclin A-CDK2, by binding to ORC1, recruits the SKP2 ligase to the ORC1 substrate. This explains why the ORC1 Cy mutant is not degraded when cyclin A-CDK2 becomes active during the G_1 to S phase transition.

The Cy motif, a short amino acid region in proteins that contains within it the (R/K)xL amino acids, was discovered as a motif in cyclin-CDK inhibitors, such as $p27^{KIP1}$, and cyclin-CDK substrates like ORC1 and CDC6. But our recent studies described here show that the Cy motif has multiple additional functions in the cell division cycle, including functions that do not involve cyclin-CDK. One function of the Cy motif is that it acts as a ligand docking motif because it is required in CDC6 for this protein to bind ORC1 in the absence of cyclin-CDK. Interestingly, the ORC1 Cy motif is not required for interaction with CDC6, but is involved in an ORC1 intramolecular interaction. The ORC1 Cy motif is also required for binding cyclin A-CDK2 kinase at two stages during the cell division cycle, once in mitosis to prevent CDC6 from binding ORC1, thereby blocking premature pre-RC assembly, and later at the G_1 to S phase transition to promote ORC1 degradation. Thus, the ORC1 and CDC6 Cy motifs harbor three separate activities: functioning as a ligand docking motif for ORC1/CDC6 protein–protein interactions, as a motif that accelerates the enzymatic phosphorylation of ORC1 by cyclin A-CDK2 kinase, and as part of a degron to promote the destruction of ORC1 (see Fig. 1).

We have recently found many other proteins that interact with ORC1, some in an ORC1 Cy motif-dependent manner that does not involve cyclin-CDK. Therefore, the Cy motif in ORC1 and CDC6 proteins, despite being derived from a common ancestor and displaying related sequences, have evolved separate functions for their respective Cy motifs. The ORC1 Cy motif is conserved in vertebrates but not in budding yeast. As noted above and in Figure 1, budding yeast has a very different ORC cycle than the ORC cycle in human cells, and we suggest that these differences are mediated in large part by the acquisition of the Cy motif in ORC and CDC6 that functions in different contexts as a protein–protein interaction ligand, as an accelerator of enzyme catalysis, and as a component of a protein degron.

PUBLICATIONS

On KF, Jaremko M, Stillman B, Joshua-Tor L. 2018. A structural view of the initiators for chromosome replication. *Curr Opin Struct Biol* **53:** 131–139.

Stillman B. 2018. Histone modifications: insights into their influence on gene expression *Cell* **175:** 6–9.

Tramentozzi E, Ferraro P, Hossain M, Stillman B, Bianchi V, Pontarin G. 2018. The dNTP triphosphohydrolase activity of SAMHD1 persists during S-phase when the enzyme is phosphorylated at T592. *Cell Cycle* **17:** 1102–1114.

TRANSCRIPTIONAL ADDICTION IN DIVERSE TUMOR LINEAGES

C.R. Vakoc L. Benjamin, Y. Gao, Y. Huang, S. Hur, O. Klingbeil, C. Lopez-Cleary, B. Lu, D. Maia-Silva, J. Milazzo, S. Polyanskaya, D. Skopelitis, T. Somerville, M. Sroka, Y. Tarumoto, Y. Wei, X. Wu, Z. Yang, T. Yoshimoto

Broad changes in gene expression are critical for cells to undergo malignant transformation. As a consequence, cancer cells are vulnerable to perturbations of transcription, which includes targeting of DNA-binding transcription factors (TFs) and chromatin regulatory machineries. Our laboratory has taken a genetic screening approach to identify transcriptional addictions in cancer, which is an effort to expose basic regulatory mechanisms and opportunities for therapeutic intervention. Research in our lab has diversified in recent years as we have transitioned from a singular focus on the BRD4 pathway in leukemia to a broader effort to expose dependencies and understand lineage transitions in four different cancers. This change was enabled by critical collaborations and through our use of domain-focused CRISPR screening, a technique for probing the essentiality of protein domains in cancer cells (Shi et al., *Nat Biotechnol 33:* 661 [2015]).

A TFIID Perturbation That Targets Myb and Suppresses Acute Myeloid Leukemia

Y. Xu

We previously introduced a strategy for disabling cancer oncogenes by targeting of transcriptional coactivators (Zuber et al., *Nature 478:* 24 [2011]). However, we now appreciate that such perturbations can produce pleiotropic biological effects on normal tissues. To overcome this issue, we have pursued RNA interference (RNAi) screening of diverse coactivator subunits in normal and malignant myeloid cells, which identified the TFIID subunit TAF12 as a selective dependency in leukemia. We traced this dependency to a direct interaction between the TAF12/TAF4 histone-fold heterodimer and the TF MYB. We showed that ectopic expression of the TAF4 histone-fold fragment can efficiently squelch TAF12 in cells, suppress MYB, and regress leukemia in mice. This study highlights how an oncogenic TF can be neutralized by targeting a general coactivator complex (Xu et al. 2018).

Lineage and Enhancer Reprogramming in Pancreatic Ductal Adenocarcinoma (PDA)

D. Maia-Silva, T. Somerville, S. Hur [in collaboration with D. Tuveson's laboratory, CSHL]

The aberrant expression of squamous lineage markers in PDA is known to correlate with poor clinical outcomes. However, the functional role of this putative trans-differentiation event in PDA pathogenesis is unclear. We have shown that expression of the TF TP63 is sufficient to install the enhancer landscape and transcriptional signature of the squamous lineage in human PDA cells. In addition, we showed that TP63-driven enhancer reprogramming promotes aggressive tumor phenotypes, including enhanced cell motility and invasion and an accelerated growth of primary PDA tumors and metastases in vivo. Taken together, our study validates the functional significance of squamous trans-differentiation in PDA, and revealed TP63-based reprogramming of PDA cells as an experimental system for investigating mechanisms and vulnerabilities linked to this aberrant cell fate. (Somerville et al. 2018). In ongoing work, we are investigating the origin of TP63+ squamous cells in PDA and how these cells interact with the tumor stroma. In addition, we are investigating the function of other TFs that become aberrantly expressed in PDA, which includes the BED-family zinc finger protein ZBED2 and the ciliogenesis master regulator FOXJ1.

A POU2F3-Addicted Tuft Cell Variant of Small-Cell Lung Cancer

Y. Huang, X. Wu

Small-cell lung cancer (SCLC) is widely considered to be a tumor of pulmonary neuroendocrine cells; however, a variant form of this disease, which lacks neuroendocrine features, has been described. We applied

domain-focused CRISPR screening to human cancer cell lines to identify the TF POU2F3 as a powerful dependency in a subset of SCLC lines. An analysis of human SCLC specimens revealed that POU2F3 is expressed exclusively in variant SCLC tumors that lack expression of neuroendocrine markers and instead express markers of a chemosensory lineage known as tuft cells. Using chromatin- and RNA-profiling experiments, we obtained evidence that POU2F3 is a master regulator of tuft cell identity in a variant form of SCLC. These findings revealed POU2F3 as a cell identity determinant and a dependency in a tuft cell variant of SCLC, which may reflect a previously unrecognized cell of origin or a trans-differentiation event in this disease (Huang et al. 2018).

Because POU2F3-expressing tuft cells exist in the normal mouse lung, we hypothesize that this cell type serves as a novel cell-of-origin in this malignancy. To address this, we are inactivating the tumor suppressor genes *Rb1* and *Trp53* in the normal tuft cell lineage of the mouse lung using a Cre knockin at the *Pou2f3* locus, and animals will be characterized to determine whether tumors emerge from this lineage that resemble the human disease we previously identified. In a parallel effort, we are using established cell lines and patient-derived tumor specimens to identify existing cancer drugs that exert exceptional responses in one of the three lineages of SCLC. Finally, we are also investigating strategies to directly block the function of POU2F3 using antisense oligonucleotides or by targeting its coactivator p300.

ZFP64 Addiction in Leukemia Is Imposed by the *MLL* Oncogene Promoter Sequence

B. Lu

The mixed lineage leukemia (*MLL*) gene is recurrently altered in leukemia by chromosomal translocations to produce oncoproteins comprising the MLL amino terminus fused to the carboxyl terminus of a partner protein. Although the fusion partners of MLL are diverse, the resulting oncoprotein is invariably expressed from the endogenous *MLL* promoter. We used domain-focused CRISPR screening to identify ZFP64 as an essential TF in leukemia cell lines harboring *MLL* translocations. Using epigenomic, biochemical, and genetic-rescue experiments, we found that the critical function of ZFP64 in leukemia cells is to bind the *MLL* promoter and maintain MLL oncoprotein expression. The remarkable specificity of ZFP64 for *MLL* was accounted for by an exceptional density of ZFP64 motifs embedded within the *MLL* promoter. In addition to identifying a vulnerability in high-risk leukemia, these findings showed how a sequence anomaly of an oncogene promoter can impose a powerful transcriptional addiction in cancer (Lu et al. 2018).

Disabling the MEF2C Oncoprotein by Targeting of Salt-Inducible Kinases

Y. Tarumoto

Expression of the lineage-specific TF MEF2C is directly activated by the MLL fusion oncoprotein in leukemia, which renders leukemia cells addicted to this TF. However, strategies to target MEF2C have yet to be identified. We have used a domain-focused CRISPR screen to reveal an essential role for LKB1 and its salt-inducible kinase (SIK) effectors (SIK3, in a partially redundant manner with SIK2) to maintain MEF2C function in acute myeloid leukemia (AML). A key phosphorylation substrate of SIK3 in this context is histone deacetylase 4, a repressive cofactor of MEF2C. Consequently, targeting of LKB1 or SIK3 diminishes histone acetylation at MEF2C-bound enhancers, and deprives leukemia cells of the transcriptional output of this essential TF. We also found that MEF2C-dependent leukemias are sensitive to on-target chemical inhibition of SIK activity. This study revealed a chemical strategy to block MEF2C function in AML, highlighting how an oncogenic TF can be disabled by targeting of upstream kinases (Tarumoto et al. 2018). In unpublished work, we have taken SIK inhibitors into leukemia mouse models, and showed single-agent activity and extension of animal survival through an on-target mechanism.

Novel Dependencies in MLL-Rearranged Leukemia

S. Polyanskaya, Y. Wei, Z. Yang

CRISPR screening efforts in the lab continue to uncover remarkable transcriptional and signaling vulnerabilities in MLL-rearranged leukemia. In addition to the targets above, we have also begun investigating

a nuclear serine phosphatase SCP4, the Fanconi anemia DNA repair complex, and the myo-inositol transporter SLC5A3 in this disease. Ongoing mechanistic studies will reveal the underlying role of these factors in this context.

Epigenetic Vulnerabilities in Rhabdomyosarcoma

C. Lopez-Cleary, M. Sroka, T. Yoshimoto

Alveolar rhabdomyosarcoma (aRMS) is a rare muscle cancer that affects primarily children and adolescents. Although the disease bears a low overall mutational burden, >60% of aRMS patients harbor the t(2;13)(q35;q14) translocation that leads to expression of the PAX3-FOXO1 fusion oncoprotein, in which the DNA binding domain of PAX3 is linked with the transactivation domain of FOXO1. Fusion-positive tumors are the most aggressive, with a 4-year overall survival rate of 8% for metastatic tumors. Numerous studies showed that aRMS cancers are dependent on the presence of the PAX3-FOXO1 chimera, and withdrawal causes cell death or differentiation. However, no therapies targeting the fusion protein exist to date, partly because of challenges associated with designing drugs that target TFs. The goal of our ongoing efforts is to elucidate the entire PAX3-FOXO1 fusion oncoprotein pathway in aRMS. To this end, we are characterizing how aRMS cells respond to fusion inactivation, as well as the molecular signatures and dynamics of each response. We have developed an unbiased, reporter-based, fluorescence-activated cell sorting (FACS)-assisted CRISPR screening strategy to identify genes that phenocopy fusion inactivation dependent on knockout. These efforts will impact our understanding of the basic biology of fusion-positive aRMS. Further, the findings might allow the development of therapeutics that indirectly silence the pathway by interfering with factors that cooperate with the fusion oncoprotein in the sustenance of the transformed state.

In the past year, we have also initiated efforts to reveal how we can amplify the anticancer effects of existing epigenetic drugs, chemotherapeutic agents, and kinase inhibitors in aRMS. These efforts rely on the use of domain-focused CRISPR screening, in which we seek to expose synthetic lethality between genetic knockouts and exposure to low concentrations of cancer drugs.

Next-Generation CRISPR Screening Strategies

O. Klingbeil, D. Skopelitis, Y. Wei [in collaboration with O. El Demerdash, Genetic Screening Core]

Genetic redundancy is pervasive in metazoan species, and conceals important gene functions when we are performing large-scale CRISPR knockout screens in cancer cells. Over the past two years, we have optimized novel lentiviral vectors that allow for efficient coexpression of two independent single-guide RNAs (sgRNAs) for the efficient generation of cells deficient in specific gene pairs. Using this approach, we are now producing CRISPR sgRNA libraries that allow for cotargeting of homologous domains of various chromatin (e.g., methyltransferases) and signaling proteins (e.g., kinases). We will begin deploying these libraries in diverse cancer models to gain a deeper understanding of essential cancer-promoting transcriptional pathways. One area of focus will be to map redundant chromatin regulator dependencies in the distinct lineages of SCLC (neuroendocrine vs. tuft cell) and pancreatic cancer (progenitor vs. squamous) in search of synthetic lethality with undruggable oncogenes, tumor suppressors, and lineage master regulators.

We have performed numerous genetic screens in tissue culture to discover epigenetic vulnerabilities in cancer; however, this strategy is unable to reveal genes required for cancer cells to engage the tissue microenvironment. We have recently adapted our domain-focused CRISPR strategy to leukemia cells grown in mice. Importantly, we can now interrogate thousands of genetic knockouts in a single cohort of animals, which we are now deploying to uncover a new class of genes that are dispensable for leukemia growth in culture but are critical for disease expansion in vivo. One such gene is MGAT1, an acetylglucosaminyltransferase that functions in the N-linked glycosylation pathway, which we are now studying on a mechanistic level. Our novel screening strategy forms the basis of a long-term effort to comprehensively elucidate the genetic programs that allow leukemia cells to evade the immune system, form productive interactions with stromal components, and evade targeted therapies.

PUBLICATIONS

Banito A, Li X, Laporte AN, Roe JS, Sanchez-Vega F, Huang CH, Dancsok AR, Hatzi K, Chen CC, Tschaharganeh DF, et al. 2018.

The SS18-SSX oncoprotein hijacks KDM2B-PRC1.1 to drive synovial sarcoma. *Cancer Cell* **33:** 527–541.

Bhagwat AS, Lu B, Vakoc CR. 2018. Enhancer dysfunction in leukemia. *Blood* **131:** 1795–1804.

Bharathy N, Berlow NE, Wang E, Abraham J, Settelmeyer TP, Hooper JE, Svalina MN, Ishikawa Y, Zientek K, Bajwa Z, et al. 2018. The HDAC3-SMARCA4-miR-27a axis promotes expression of the *PAX3:FOXO1* fusion oncogene in rhabdomyosarcoma. *Sci Signal* **11:** eaau7632. doi:10.1126/scisignal.aau7632

Brien GL, Remillard D, Shi J, Hemming ML, Chabon J, Wynne K, Dillon ET, Cagney G, Van Mierlo G, Baltissen MP, et al. 2018. Targeted degradation of BRD9 reverses oncogenic gene expression in synovial sarcoma. *Elife* **7:** e41305.

Hörmann A, Hopfgartner B, Köcher T, Corcokovic M, Krammer T, Reiser C, Bader G, Shi J, Ehrenhöfer K, Wöhrle S, et al. 2018. RIOK1 kinase activity is required for cell survival irrespective of *MTAP* status. *Oncotarget* **9:** 28625–28637.

Huang YH, Klingbeil O, He XY, Wu XS, Arun G, Lu B, Somerville TDD, Milazzo JP, Wilkinson JE, Demerdash OE, et al. 2018. POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. *Genes Dev* **32:** 915–928.

Lu B, Klingbeil O, Tarumoto Y, Somerville TDD, Huang YH, Wei Y, Wai DC, Low JKK, Milazzo JP, Wu XS, et al. 2018. A transcription factor addiction in leukemia imposed by the MLL promoter sequence. *Cancer Cell* **34:** 970–981.

Skucha A, Ebner J, Schmöllerl J, Roth M, Eder T, César-Razquin A, Stukalov A, Vittori S, Muhar M, Lu B, et al. 2018. MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity. *Nat Commun* **9:** 1983.

Somerville TDD, Xu Y, Miyabayashi K, Tiriac H, Cleary CR, Maia-Silva D, Milazzo JP, Tuveson DA, Vakoc CR. 2018. TP63-mediated enhancer reprogramming drives the squamous subtype of pancreatic ductal adenocarcinoma. *Cell Rep* **25:** 1741–1755.e7.

Tarumoto Y, Lu B, Somerville TDD, Huang YH, Milazzo JP, Wu XS, Klingbeil O, El Demerdash O, Shi J, Vakoc CR. 2018. LKB1, salt-inducible kinases, and MEF2C are linked dependencies in acute myeloid leukemia. *Mol Cell* **69:** 1017–1027.

Tiriac H, Belleau P, Engle DD, Plenker D, Deschênes A, Somerville T, Froeling FEM, Burkhart RA, Denroche RE, Jang GH, et al. 2018. Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. *Cancer Discov* **8:** 1112–1129.

Vakoc CR, Kentsis A. 2018. Disabling an oncogenic transcription factor by targeting of control kinases. *Oncotarget* **9:** 32276–32277.

Wai DCC, Szyszka TN, Campbell AE, Kwong C, Wilkinson-White LE, Silva APG, Low JKK, Kwan AH, Gamsjaeger R, Chalmers JD, et al. 2018. The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *J Biol Chem* **293:** 7160–7175.

Xu Y, Milazzo JP, Somerville TDD, Tarumoto Y, Huang YH, Ostrander EL, Wilkinson JE, Challen GA, Vakoc CR. 2018. A TFIID-SAGA perturbation that targets MYB and suppresses acute myeloid leukemia. *Cancer Cell* **33:** 13–28.

Zafra MP, Schatoff EM, Katti A, Foronda M, Breinig M, Schweitzer AY, Simon A, Han T, Goswami S, Montgomery E, et al. 2018. Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol* **36:** 888–893.

CANCER: GENETICS

Christopher Hammell's lab is interested in understanding gene regulatory processes that give rise to robust phenotypes associated with normal development in animals (specifically, how the timing of developmental processes is controlled), as well as the alterations in these pathways that give rise to diseases such as cancer (as in the alterations in mitogenic pathways in melanoma). Hammell and colleagues approach this elemental problem by using a variety of model organisms and patient-derived cancer cell lines. To directly identify the components that function in controlling normal developmental timing, they use the small nematode *Caenorhabditis elegans*, applying forward and reverse genetic approaches. In contrast to the extreme robustness of cell-fate lineage in *C. elegans*, in which specification of developmental programs is hardwired, mutations that alter conserved signaling pathways in melanoma create relatively plastic developmental landscapes that allow these lesions to become aggressive tumors. Notably, the gene regulatory architecture of melanoma cells allows them to acquire resistance to therapeutic agents. Hammell's team is interested in epigenetic mechanisms that contribute to resistance, specifically dramatic changes in gene expression patterns and intracellular signaling pathways. They are performing high-throughput screens to identify cellular factors that allow these rewiring events to occur, with the idea that these components would make ideal therapeutic targets to complement existing clinical strategies.

The **Je H. Lee** lab studies how cells interact with their microenvironment to regulate gene expression during development. Single-cell heterogeneity in gene expression can result from spatial differences in cell–cell and cell–extracellular matrix interactions. Such differences contribute to stochastic evolution of tumor cells, as well as morphogenesis, during normal development. However, the spatial control of gene expression in complex tissues, embryos, or tumors remains poorly understood because most genome-wide studies sample bulk tissues or dissociated single cells.

They have recently developed a method to sequence RNA molecules directly within single cells and tissues using subcellular resolution imaging, and they have shown subtle differences in cell–cell/extracellular matrix (ECM) signaling and gene expression genome-wide in situ. By clustering transcripts into functionally or morphologically discrete regions, they find many unique spatial markers and signaling pathways. The Lee lab focuses on the role of noncoding RNA in chromatin remodeling and tumor progression using mouse and organoid models of human cancer. They use in situ sequencing, cell lineage tracing, and single-cell profiling to understand how noncoding RNA affects tumor cell evolution in their native context. The long-term goal is to develop better tumor classification tools and anticancer therapeutics using their understanding of the tumor microenvironment.

Alea Mills is studying genetic pathways important in cancer, aging, and autism, identifying the genetic players and determining how aberrations in their functions culminate in human disease. Through innovative use of a technique called "chromosome engineering," the Mills group discovered that one of the most common genetic alterations in autism—deletion of a 27-gene cluster on chromosome 16—causes autism-like features in mice. These autism-like movement impairments can be identified just days after birth, suggesting that these features could be used to diagnose autism. The Mills group has also used chromosome engineering to identify a tumor suppressor gene that had eluded investigators for three decades. The gene, called *Chd5*, was shown to regulate an extensive cancer-preventing network. This year, the Mills lab uncovered how *Chd5* acts as a tumor suppressor: It binds to a protein found within chromatin to turn specific genes on or off, halting cancer progression. The epigenetic role of *Chd5* in development, cancer, and stem-cell maintenance is currently being investigated. The Mills lab is also studying p63 proteins, which regulate development, tumorigenesis, cellular senescence, and aging in vivo. They succeeded in

halting the growth of malignant tumors by turning on production of one of the proteins encoded by the *p63* gene, called TAp63. TAp63 also exerts other protective effects. This year, the Mills lab generated a mouse model that allowed them to find that TAp63 is required to prevent a genetic disorder, known as EEC (ectrodactyly-ectodermal dysplasia cleft lip/palate syndrome), which is characterized by a cleft palate and major deformities of the skin and limbs in infants. In addition, they recently discovered that a different version of *p63*, called ΔNp63, reprograms stem cells of the skin to cause carcinoma development—the most prevalent form of human cancer. Modulation of these proteins may offer new ways to treat human malignancies in the future.

Human cancers show a diverse array of genomic gains and losses that alter the dosage of hundreds of genes at once. About 90% of solid tumors display whole-chromosome aneuploidy, whereas many tumors with diploid karyotypes nonetheless harbor segmental or arm-length aneuploidies that also result in significant gene copy number alterations. Despite the prevalence of aneuploidy in cancer, its functional consequences for cell physiology remain poorly understood. The work of **Jason Sheltzer** and colleagues has shown the existence of several surprising phenotypes that are shared among cells with different chromosomal imbalances. They showed that aneuploidy can function as a novel source of genomic instability, as aneuploid cells tend to display elevated levels of mutation, mitotic recombination, and chromosome loss. They have also identified a transcriptional signature of aneuploidy that is associated with cellular stress and slow proliferation and is found in aneuploid primary and cancer cells across a host of organisms. More recently, they have investigated the link between aneuploidy and cellular transformation. Using a series of genetically matched euploid and aneuploid cell lines, they have shown that aneuploidy can paradoxically function as a barrier to tumor growth. They are currently continuing their investigation of the role in aneuploidy in cancer. They are also applying CRISPR-Cas9-mediated genome engineering to develop novel mouse models for exploring the impact of gene dosage alterations on tumor development in vivo.

Although aneuploidy is a ubiquitous feature of human tumors, it occurs rarely in somatic cells. Thus, differences between aneuploid and euploid cells may represent crucial therapeutic vulnerabilities in cancer. By identifying phenotypes that are shared among tumors with different aneuploidies, the Sheltzer lab hopes to discover pathways that can be manipulated to selectively eliminate aneuploid cells or block aneuploidy's non–cell autonomous effects. Drugs that target these pathways may have broad utility against a wide range of aneuploid cancers while showing minimal toxicity in euploid tissue.

Michael Wigler's work provides a new paradigm for understanding and exploring human disease. The Wigler lab studies human cancer and the contribution of new mutation to genetic disorders. The cancer effort (with James Hicks, Alex Krasnitz, and Lloyd Trotman) focuses on breast and prostate cancers. It involves collaborative clinical studies to discover mutational patterns predicting treatment response and outcome and the development of diagnostics to detect cancer cells in bodily fluids such as blood and urine. The major tools are single-cell DNA and RNA analysis. The single-cell methods, which are in development, are also being applied to problems in neurobiology (with Josh Huang and Pavel Osten) to characterize neuronal subtypes, somatic mutation, and monoallelic expression. The Wigler lab's genetic efforts are in collaboration with Ivan Iossifov and Dan Levy, and this team focuses on determining the role of new mutations in pediatric disorders. In a large-scale population sequencing project with W. Richard McCombie and the Genome Sequencing Center at Washington University in St. Louis, and supported by the Simons Foundation, the team has proven the contribution of this mechanism to autism. The work further suggests a relationship between the mutational targets in autism and the process of neuroplasticity that lies at the heart of learning. Smaller-scale population studies of congenital heart disease and pediatric cancer (collaborations with scientists at Columbia University and Memorial Sloan Kettering Cancer Center, respectively) also point to new mutation as a causal factor in these disorders.

GENETIC DISSECTION OF GENES THAT MODULATE TEMPORAL GENE REGULATION

C.M. Hammell C. Aguirre-Chen H. Ouyang
K. Hills-Muckey N. Stec
B. Kinney J. Wang

The long-standing questions that my lab aims to understand are how the precise timing of developmental gene regulation is established in the most basic sense and how these patterns of gene expression are maintained in highly variable environments. These efforts originated through the genetic dissection of the gene regulatory pathways that establish temporal cell fate specification in *Caenorhabditis elegans* and have historically been centered on how the activities of specific microRNAs (miRNAs) impact these processes. Although initially centered on integrating how miRNA-mediated gene regulation is integrated in temporal transitions, our current work has expanded to two major themes: understanding (1) how oscillatory gene expression (both protein coding and regulatory RNAs) is established during postembryonic development and (2) how a novel and conserved family of proteins modulates translational output during development.

Development of a Long-Term Microfluidics Imaging Platform to Characterize Gene Expression Patterns in Developing *C. elegans* Larvae

N. Stec, K. Hills-Muckey, B. Kinney

One of the discoveries from our laboratory was the identification of a number of regulatory components that generate cyclical patterns of transcription that are coupled to the growth cycles of developing *C. elegans* larvae. In the intervening years, we have leveraged these initial findings to identify additional components of this system. The characterization of these new genes (including those encoding the transcription factors LIN-42, BLMP-1, ELT-3, and NHR-23) is through the use of a variety of transcriptional reporters (GFP-pest) whose expression is driven by the upstream regulatory regions of putative target genes. The transparency of our model organism enables us to use these reporters to visualize changes in target gene expression that are due to alterations in this normal, conserved gene regulatory circuit. One of the unique features of cyclical transcriptional expression patterns that distinguishes them from simple on/off types of gene regulation is that specific features of their patterns can be modulated independently: Specifically, aspects of periodicity, amplitude, and phase can vary across different target genes. We have typically relied on ensemble measurements to characterize these features. This is accomplished by monitoring a large number of animals over short intervals and is not continuous—essentially, we take snapshots of populations over the course of development. Because ensemble measurements tend to "average out" some of the fine details of these patterns, our lab has struggled to properly visualize these dynamics to a level at which we could move past the descriptive. As our intention ultimately would be to model these features mathematically, we needed to develop a novel approach to these types of specific measurements.

To this end, our lab applied for and received additional funding to develop a novel microfluidics platform that will enable us to monitor changes of gene expression in individual animals over the entire time of postembryonic development. In essence, this system uses small, two-layered microfluidic chambers in which the bottom layer constrains the developing animal within a defined location (Fig. 1). Importantly, this layer enables growth media and a food source to flow through this chamber, enabling growth and development. The upper chamber can be compressed through separately controlled pressure changes. This enables the specimen in the lower chamber to be immobilized temporarily. During this immobilization period, we image the fluorescent signal (in all *z*-stacks) of the various GFP-pest transcriptional reporters at cellular resolution. Once the imaging is complete, the

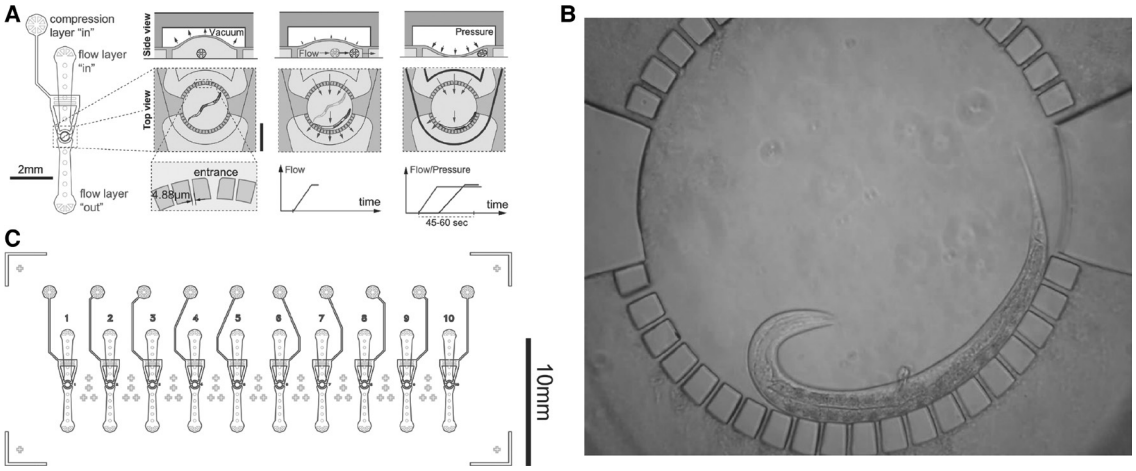


Figure 1. A novel microfluidics platform that enables rapid and reversible immobilization and long-term imaging of developing *C. elegans* larvae. (*A*) A schematic of a single well of the microfluidic device. When animals are to be imaged, the flow is increased in the bottom chamber, causing the animal to move to the bottom of the enclosure. Once the animal has been repositioned in the bottom chamber, pressure is increased to hold the animal in place. During this procedure, the animal is imaged with a microscope, capturing images of the animal with differential interference contrast imaging (for animal anatomy) and then in a particular fluorescent channer to measure changes in gene expression. Once this imaging is done, the animal is released to resume normal movement and development. (*B*) An image showing an L2-staged animal in a single well of a chip. (*C*) A schematic of an entire chip, demonstrating the compact organization of the design.

pressure in the top chamber is released, allowing the animal to move, forage, and develop normally. This procedure can be multiplexed for up to 10 animals (10 chambers) per experiment and is gentle enough to enable these specimens to be imaged throughout all of postembryonic development.

Mutations in BLMP-1 and ELT-3 Function Redundantly to Control the Amplitude and Duration of Transcription for Their Targets

miRNAs function in a dose-dependent manner to control the expression of mRNA targets and therefore enforce stage-specific developmental transitions. As we had already shown that most heterochronic miRNA genes are expressed in pulsatile patterns throughout development, we hypothesized that there may be distinct transcriptional factors that control the output of miRNA genes by modulating any or all of the three elements of pulsatile transcription (periodicity, amplitude, and/or phase). We had already published evidence that this is the case for LIN-42, the *C. elegans* ortholog of *Period* that dampens the amplitude of cyclical miRNA transcriptional pulses. In the preceding year, we had identified *blmp-1* and *elt-3* as genes that antagonize *lin-42* to maintain normal miRNA expression levels. In addition to a variety of other molecular measurements, we subjected *blmp-1;elt-3* animals to our new microfluidics apparatus. In these experiments, we monitored the expression of a number of transcriptional reporters of BLMP-1/ETL-3 targets that had been identified through ChIP-seq analysis. These efforts demonstrated that *blmp-1;elt-3* mutations primarily alter the amplitude and duration of their targets (see the ZK180.5 and *mlt-10* transcriptional profiles in Fig. 2). Importantly, this alteration in transcriptional patterns of BLMP-1 and ELT-3 targets is similar.

These experiments suggest to us that, unlike core transcriptional components of the developmental clock that work episodically at a single point in the transcriptional cycle, BLMP-1 and ELT-3 function redundantly throughout larval development. This model is also consistent with two additional observations that we have recently made regarding BLMP-1's and ELT-3's mechanism of modulating miRNA (and other target) transcription. First, BLMP-1 and ELT-3 are expressed continuously throughout each larval stage, suggesting that they are bound and regulating target gene expression both before and after active transcription. Second,

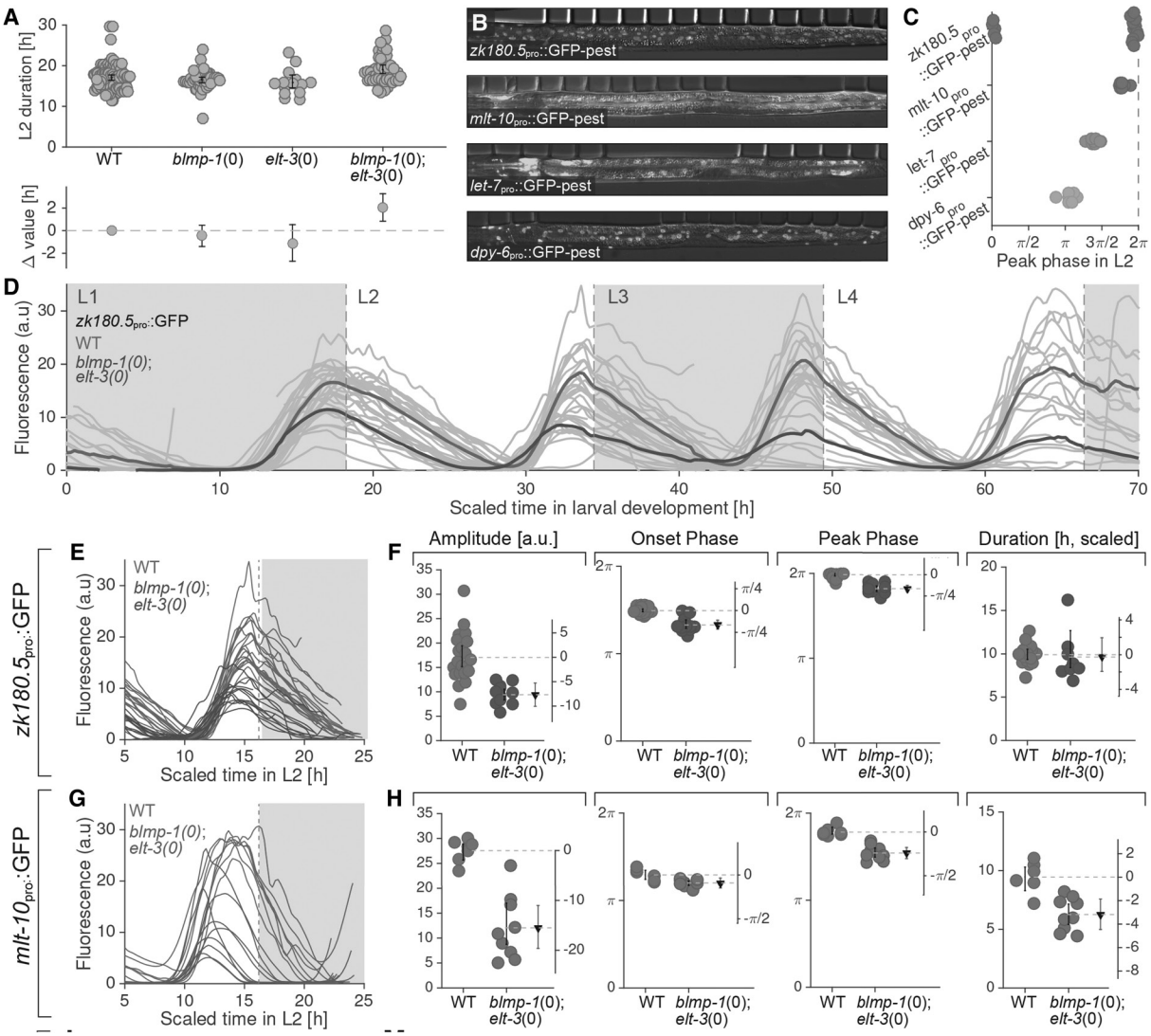


Figure 2. *blmp-1(0);elt-3(0)* mutations alter transcriptional output by modulating amplitude and pulse duration. (*A*) *blmp-1(0);elt-3(0)* animals do not alter the periodicity of the molting cycle. (*B*) Representative pictomicrographs of various transcriptional GFP-pest reporters that were analyzed in the microfluidics device. All of the reporters used in these experiments were expressed primarily in the hypodermis. (*C*) A graphical representation of where the peak phase of each reporter occurs within each larval cycle. (*D*) Whole-animal, developmental expression profiles for the *ZK180.5::GFP-pest* reporter in wild-type or *blmp-1(0);elt-3(0)* double mutants. Note that in each larval stage, the peak expression levels observed in *blmp-1(0);elt-3(0)* animals are always lower than those measured in wild-type animals. (*E*) A detailed view of the L2 expression patterns of the *ZK180.5::GFP-pest* reporter in wild-type and *blmp-1(0);elt-3(0)* animals. Note that the time of transcriptional onset is nearly identical in both genotypes and that it is primarily the amplitude and pulse duration of the expression pattern that is altered in *blmp-1(0);elt-3(0)* animals. (*F*) A detailed quantification of the specific cyclical transcriptional parameters of the *Zk180.5::GFP-pest* reporter in more than 10 animals of each genotype. (*G,H*) An identical analysis of the expression parameters of the *mlt-10::GFP-pest* reporter.

comparison of BLMP-1 and ELT-3 genomic target sites and ATAC-seq data that demarcates open chromatin regions indicates that BLMP-1 and ELT-3 are statistically associated with regions of open local chromatin structure. This may indicate that both of these proteins function as pioneer transcription factors that facilitate the binding of core developmental clock transcription factors. We aim to integrate these findings with other projects in the lab regarding the activity of the nuclear hormone receptor NHR-23.

Dissecting the Conserved Molecular Roles of the "Prion-Like" Proteins, PQN-59/UBAP2/2L, That Amplify Translational Output

H. Ouyang, J. Wang, N. Stec

Dynamic gene expression in development requires the coordination of a number of processes that converge to regulate the abundance of RNAs and proteins. Although most of our lab is focused on understanding how the precise timing of transcriptional events mediates robustness in cell fate specification, we also are interested in how protein translation affects these decisions. This project is, again, focused on one of the genes we identified on our original *lin-4* suppressor screens. Among the suppressors that correct lin-4(lf) phenotypes without altering *lin-4* expression levels is the conserved "prion-like" protein, PQN-59. In last year's report, we had presented molecular evidence that PQN-59 and its human orthologs, UBAP2 and UBAP2L, function in translation. By performing a proteomic analysis of the factors that interact with PQN-59 in vivo, we were able to infer two kinds of potentially functional interactions between PQN-59. The first was to be expected: PQN-59 interacts with an overlapping set of similar proteins whose orthologs also interact with UBAP2L (Fig. 3A). Of these, a

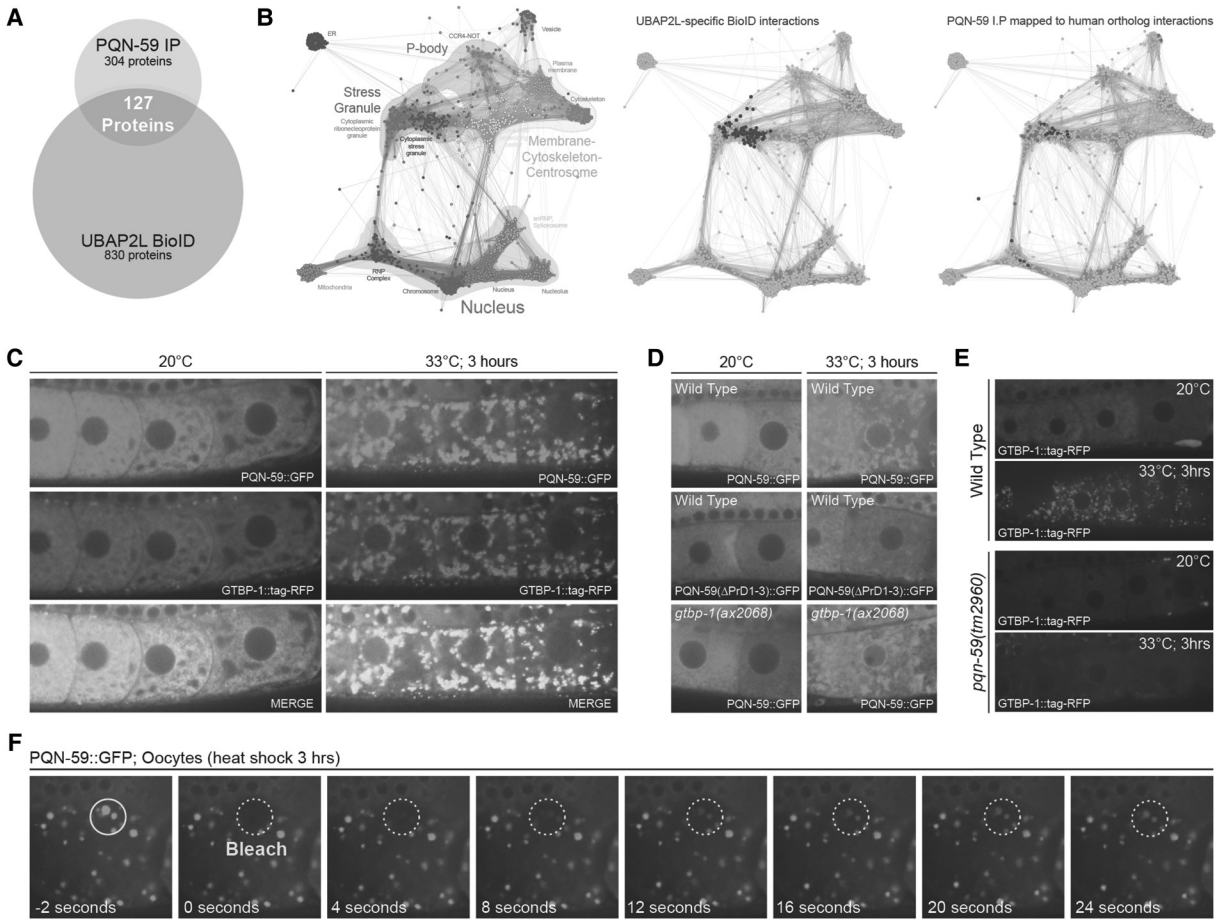


Figure 3. PQN-59 interacts with stress-granule components and is needed for their full assembly. (*A*) A comparison of the overlap between proteins that coprecipitate with PQN-59 in worm lysates and proteins that associate with UBAP2L in human cells. (*B*) A pictogram demonstrating the human RNP complexes in cells and the overlapping sets of proteins that cluster in these maps around UBAP2L and PQN-59. (*C*) Micrographs demonstrating that PQN-59 and the core stress-granule component, GTBP-1, colocalize to the cytoplasm in stress and nonstress conditions. (*D*) Localization of PQN-59 to stress granules requires the prion-like domains (PdD1-3), but it does not require the expression of GTBP1. (*E*) Deletion of *pqn-59* prevents the localization of GTBP-1 to stress granules. (*F*) PQN-59 condensates are "liquid-like" coacervates as measured by FRAP.

large number of binding partners in both species are associated directly with the process of translation. In addition to that, PQN-59 interacts with a number of stress-granule components implicated in cytoplasmic, membrane-less compartments associated with stalled translational intermediates. In a comparison between the interactors of UBAP2L, PQN-59 exhibits a number of overlapping binding partners with its human orthologs (Fig. 3B).

Primary among the stress-granule-associated proteins that interact with both PQN-59 and UBAP2L is the central stress-granule component GTBP-1 (G3BP1 in humans). Using CRISPR-tagged translational reporters of these proteins, we were able to demonstrate that heat-shocking animals led to the relocalization of these proteins from their normal diffuse distribution to concentrated droplets for each protein. Importantly, and consistent with their biochemical interactions, they are colocalized in these liquid droplets (Fig. 3C). We then used genetic mutations in the *pqn-59* or *gtbp-1* genes to probe the mechanistic relationship between these genes and the normal localization of their products. First, we demonstrated that the "prion-like" domains of PQN-59 are required for its relocalization during heat shock (Fig. 3D). Second, using a null mutation of *gtbp-1*, we demonstrated that PQN-59 could relocalize to these "granule-like" structures without GTBP-1 expression (Fig. 3D). In contrast, relocalization of GTBP-1 to these structures requires PQN-59 expression (Fig. 3E). Finally, we demonstrated that the PQN-59 granules resemble liquid-like droplets in that they can fuse with each other in vivo and recover rapidly after photobleaching (FRAP) assays (Fig. 3F).

We are currently focusing on the construction of *C. elegans* strains and human-derived cell lines where we can rapidly deplete these proteins in vivo in a ligand-dependent manner. These reagents will be used to further dissect the roles of these proteins in the process of normal translation.

SUBCELLULAR RECONSTRUCTION OF THE TRANSCRIPTOME DURING DEVELOPMENT

J. Lee D. Furth S. Weinmann
D. Ghosh X. Yuan
X. Wang

Our laboratory is developing single-cell, genomics, and imaging technologies as well as experimental approaches to investigate how subcellular localization of mRNAs establishes asymmetric cellular patterning using *Drosophila* as a model. We are also evolving our platform to reconstruct somatic mutations for detecting early cancer or premalignant lesions.

Long RNA Variant Sequencing in *Drosophila* Ommatidia

D. Furth, X. Wang

Most cell types are asymmetric in subcellular composition, shape, or function. Mechanisms that regulate cell asymmetry include the distribution of polarity proteins as well as subcellular localization of mRNAs. Classical approaches have shown the importance of RNA localization in ascidian eggs, *Xenopus* oocytes, *Drosophila* embryos, and *Saccharomyces cerevisiae*. In fibroblasts, mRNAs coding for actin enable localized translation at the leading edge. In neurons, differential distribution of mRNAs to dendrites enables localized translation of synaptic mediators. However, the full extent to which mRNA localization regulates cellular processes is not known.

Epithelial cells form a sheet of connected cells over the underlying connective tissue. The apical side faces the external environment and contains specialized structures such as microvilli or cilia that mediate absorption or planar movement of particles, respectively. These processes are established by protein complexes that define apical, subapical, and basolateral surfaces of the epithelium and regulate asymmetric trafficking of cargo content to the surface. For example, pair-rule genes localize apically in epithelial cells at the *Drosophila* blastoderm stage. These mRNAs are packaged into particles that are transported to the apical surface along microtubule tracks. Later in development, wingless (*wg*) mRNA localizes apically, and this polarized distribution is essential for optimal Wg activity. These studies suggest that mRNA subcellular localization contributes to epithelial cell structure and function. Yet, it remains to be determined whether this is simply a general phenomenon or one regulated by an intricate network of *cis*-acting zip codes, RNA-binding proteins, and delivery systems.

Modern tools enable imaging of RNAs in live cells, as well as identification of *cis*-acting motifs and RNA-binding proteins. These approaches have shown that *cis*-acting RNA motifs in the 3′ untranslated region (UTR) play an important role in mRNA localization in a cell type– or tissue-specific manner. For example, short and long 3′ UTR isoforms of BDNF mRNA are differentially localized to the soma and dendrites of hippocampal neurons, and the long 3′ UTR isoform is required for proper dendritic morphology and synaptic plasticity. Because alternative 3′ UTR or polyadenylation is widespread, an unbiased method for localizing RNAs or their isoforms with subcellular resolution across multiple tissue types could be key to elucidating the role of 3′ UTRs in development.

Unbiased next-generation sequencing (NGS) methods enable one to examine multiple tissue regions using microdissection, laser-activated mRNA capture, or array-based mRNA capture. Although these approaches could attain high spatial resolution by increasing the sampling frequency, the number of sampled regions increases exponentially, making such approaches cost-prohibitive for subcellular RNA localization studies. Alternatively, single-molecule fluorescence in situ hybridization (smFISH) can examine multiple genes selected based on existing annotations or single-cell analysis; however, it is not possible to interrogate the whole transcriptome, variants, and isoforms in an unbiased manner using hybridization-based methods.

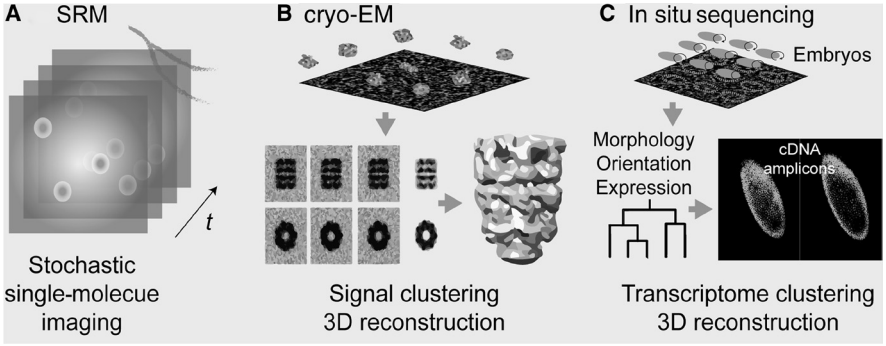


Figure 1. High-resolution spatially resolved 3D structures can be generated from a large array of similar protein complexes (cryo-EM) or biological patterns (FISSEQ), overcoming sparseness in data.

In contrast, fluorescent in situ sequencing (FISSEQ) utilizes randomly primed reverse transcription of the whole transcriptome followed by in situ cDNA sequencing to visualize gene expression or splicing across intact cells in an unbiased manner (Lee et al., *Science 343:* 1360 [2014]); however, FISSEQ suffers from low sensitivity and short read lengths. The low sensitivity is due to the overwhelming amount of ribosomal RNAs (rRNAs), whereas the short sequencing read length is due to the number of time-consuming imaging cycles. As a result, it could take up to a month to produce 30-base reads across a field of approximately 1000 cells with subcellular resolution. In addition, the short read length makes it challenging to map alternatively spliced transcripts in situ. Over the past year, we have solved these problems using *Drosophila* as a model system.

To compensate for the low sensitivity, we utilize highly ordered *Drosophila* ommatidia. The concept is similar to that of cryo-EM, in which low-resolution images from thousands of molecules are averaged for tomographic reconstruction (Fig. 1). *Drosophila* retinas expressing UAS-catenin-GFP and eye-specific GMR-GAL4 are used to segment individual cells. Retinas are dissected, fixed, and mounted with the apical surface facing the coverslip. The cDNA amplicon library is generated in situ following tissue permeabilization (Fig. 2), and we designed reverse transcription primers to enrich for cytoplasmic mRNAs. To sequence cDNA amplicons within the *Drosophila* retina, our bidirectional sequencing by ligation (SBL) uses a two-color coding scheme so that two bases are sequenced per imaging cycle, reducing the imaging time by one-half (Fig. 3). To do so, we developed a terminator cleavage chemistry for bidirectional SBL using off-the-shelf

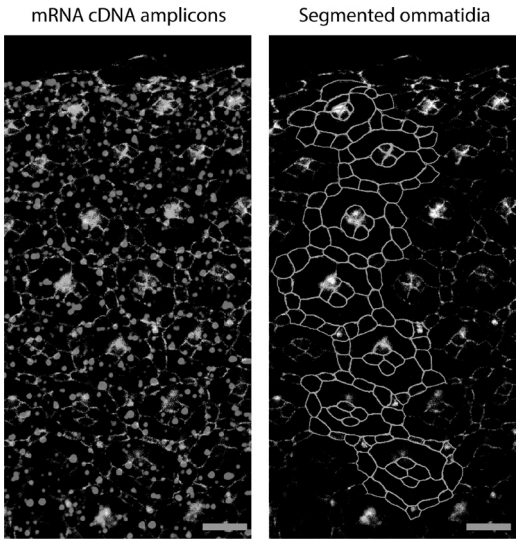


Figure 2. Dissected *Drosophila* retina mounted on glass, ommatidia, eye-specific catenin-GFP, and 12-base in situ RNA sequencing of the mRNA transcriptome.

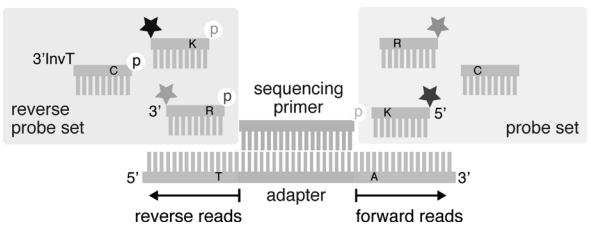


Figure 3. New bidirectional paired-end ribonucleotide-inosine cleavage *k*-mer ligation (PRICKLi) chemistry enables one to sequence two bases at each imaging cycle using off-the-shelf probes and enzymes.

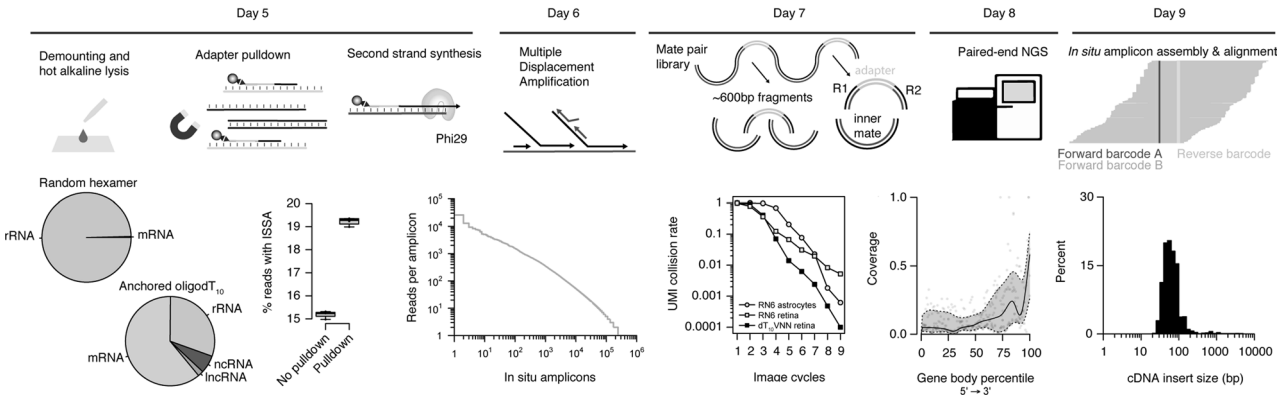


Figure 4. FISSEQ sequences short (six imaging cycles) barcodes from each amplicon in a day. After imaging, the sample is collected, lysed, and used for mated-pair Illumina sequencing. The contig length of cDNA amplicons is as long as 4.7 kb, and the amplicons contain unique barcodes for mapping them to specific locations within the image.

oligonucleotides and enzymes, requiring only four oligonucleotides. To increase the read length, we sequence the first 12 bases from each amplicon in situ, which serve as a random barcode. This is then followed by hybridization-based capture of the amplicons and mated-pair sequencing library construction (Fig. 4). The whole process is completed in 3–4 d, instead of months. Currently, we are generating mated-pair sequences using paired-end 150-base sequencing from the cDNA amplicons as long as 10 kb, and each mated-pair read is placed within a mesh representation of the *Drosophila* retina (Fig. 5). In the near future, we will transition from Illumina short reads to Oxford nanopore long reads, and this will enable us to reconstruct de novo single-nucleotide variants (SNVs), hypervariable somatic mutations (i.e., T-cell receptors), and RNA editing loci in situ. We are now developing software to perform 3D tomography of mRNAs, 3′ UTRs, and putative *cis*-acting motifs in the representative ommatidium.

In summary, we achieved a major feat in spatial functional genomics and obtained long RNA sequencing reads in situ. Our solution enables one to reconstruct the transcriptome, including variants, isoforms, and regulatory motifs, across repetitive biological patterns. We are applying our platform to study *Drosophila* development from blastoderms to wing imaginal discs to adult whole brains. Furthermore, we are analyzing asymmetric embryonic stem

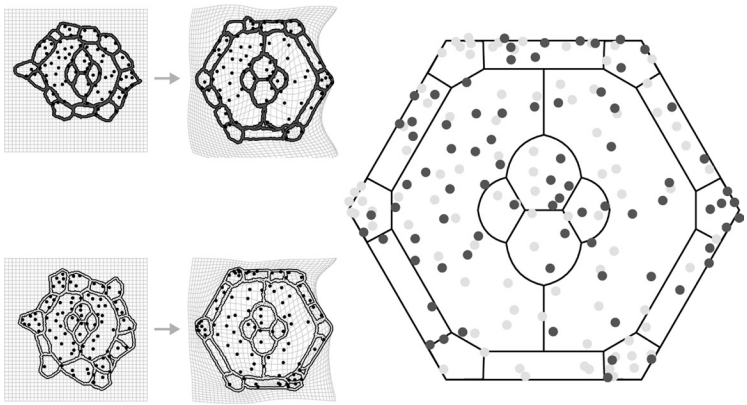


Figure 5. The cDNA sequences (poly(A′)) span the whole 3′ UTR as well as some of the coding region, enabling one to detect localized patterns of alternative 3′ UTR usage or polyadenylation with subcellular resolution. Preliminary analysis indicates that FISSEQ reads are similar to those of CLIP-seq, from which RNA–protein interactions can be inferred with ~10-nt resolution.

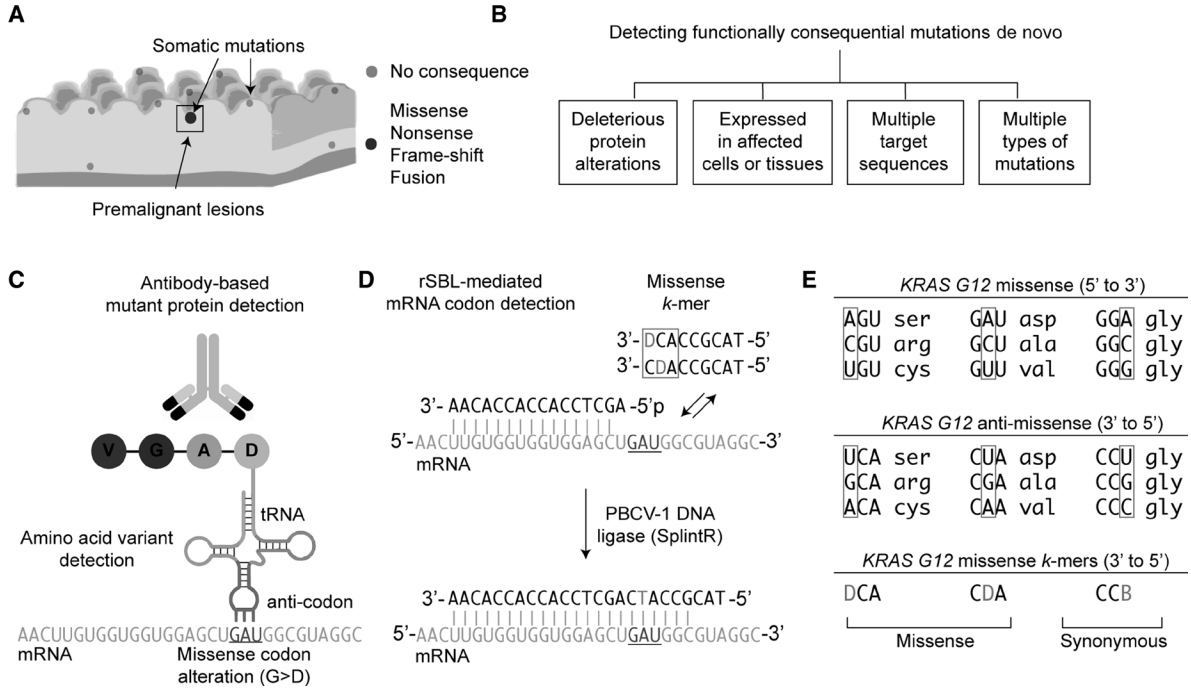


Figure 6. Sequence-agnostic labeling of mRNA or single cells containing functionally deleterious mutations. (*A*) Nonsynonymous mutations are rare but accumulate with aging. (*B*) Basic criteria for assessing de novo deleterious mutations. (*C*) Mutation-specific antibodies essentially read changes in individual codons; however, mutation-specific antibodies are difficult to generate. (*D*) Our approach uses pseudo-random *k*-mers and sequencing by ligation (SBL) to label transcripts containing nonsynonymous mutations in situ without using NGS.

cell (ESC) patterning in mouse blastocysts to investigate whether polarized *cis*-acting motifs are involved in mammalian development. In addition, we have sequenced all possible cadherin isoforms in the *Drosophila* brain, enabling us to identify cadherin isoforms necessary for cellular compartmentalization during development.

Somatic Mutation and Short RNA Variant Detection in Cancer

D. Furth, D. Ghosh, S. Weinmann, C. Yuan

Cells in the human body accumulate hundreds of de novo mutations over the span of a lifetime. Although most mutations are harmless, deleterious variants eventually emerge and contribute to tumorigenesis (Fig. 6A). Therefore, premalignant lesion screening requires detecting functional mutations in rare cells scattered among a large number of normal cells (Fig. 6B). To detect coding mutations within single cells, antibodies (Fig. 6C) can be used. However, antibodies specific for coding mutations are difficult to produce. Alternatively, high-throughput DNA sequencing can profile any number of genetic alterations, but deep sequencing is required to detect rare cells containing sequence variants, whose single-cell sensitivity is limited by the sequencing cost. For early disease detection,

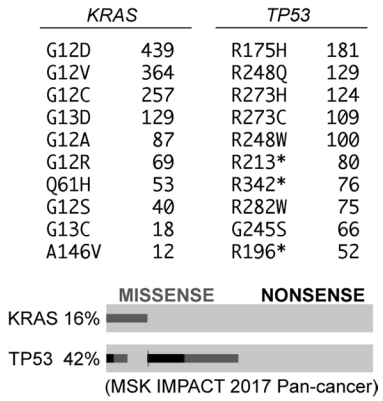


Figure 7. If one were to interrogate all missense or nonsense mutations in *KRAS* or *TP53*, >50% of all premalignant lesions could be detected across various human cancers. For proof of concept, we are testing 10 codons from each, which are found in 23% of all human cancers.

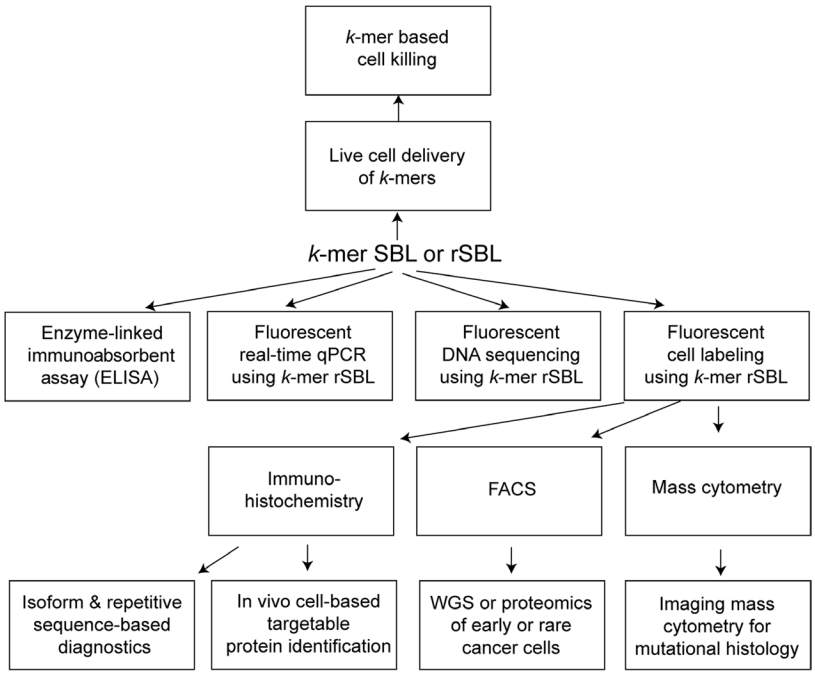


Figure 8. Our method for detecting and labeling nucleic acids or single cells is both broad and generic, potentially enabling multiple applications from qPCR, ELISA, FACS, NGS, single-cell genomics, and potentially cytotoxic cell therapies. This is because our approach and reagents enable molecular or single-cell labeling or quantification without the need for separate NGS or computational analysis. This chart describes our current effort in demonstrating feasibility across different applications.

a new method is needed in which a large number of functionally relevant de novo mutations are profiled with single-cell resolution using low-cost assays. In addition, if one could label, isolate, or visualize functional somatic mutations within single cells from tissue sections, it could increase the diagnostic specificity.

To label DNA, RNA, or single cells that contain functional mutations, we developed SBL using pseudo-degenerate probes (Fig. 6D). For any codon mutation, one can categorize variant sequences by their functional class (i.e., missense, nonsense). SBL can then be performed using *k*-mers specific for functional point mutation to label only those DNA or RNA molecules that are deleterious. By utilizing mixed bases, most codon mutations only require two probes, and our method works directly on RNA as well as DNA (Fig. 6E). We have now quantified the sensitivity, specificity, and read length of our method and designed *k*-mers against 20 nonsynonymous codon mutations (*KRAS* and *TP53*) found in up to 23% of all human cancers (Fig. 7). In addition, we have performed key feasibility experiments and generated a set of preliminary data (Fig. 8). Currently, our goal is to demonstrate its application in low-cost liquid biopsy and mutation histology. In addition, we are using *k*-mer SBL to label single cells in situ post-Cas9-induced mutagenesis, enabling us to dissect the biochemical property of individual amino acids in vivo.

GENETIC AND EPIGENETIC BASIS OF STEM CELLS AND CANCER

A. Mills S. Balinth C. Ballon L. Banks Y. Chang M. Fisher D. Johnson P. Shrestha S. Sun C. Wu

Our group is focused on determining the genetic/epigenetic basis of cancer and neurodevelopmental syndromes. We have discovered genes impacting these conditions, have determined how the encoded proteins work in normal cells, and have determined how their deregulation contributes to disease. These findings have had a major impact and have affected how clinicians analyze and treat patients with these syndromes.

Major discoveries:

- Identifying *p63* as a gene affecting development, aging, and cancer
- Defining the genetic basis of autism
- Discovering *CHD5* as a gene that prevents cancer
- Determining that Chd5 loss causes male infertility

p63 in Development, Aging, and Cancer

We discovered *p63*, a gene related to *p53*— which encodes a tumor suppressor defective in more than half of all human cancers. Although p63 is similar to p53, its function was not at all clear. We found that p63 deficiency leads to premature aging, as p63 is needed for stem cell renewal. Indeed, lack of p63 causes curvature of the spine, hair loss, and severe skin lesions. Yet there is a fine balance, as an excess of one version of p63 (ΔNp63α) causes carcinoma—the most prevalent type of human cancer. In contrast, we found that a different version of p63 (TAp63) prevents cancer. Our work showing that TAp63 inhibits tumor growth, even when p53 is absent, was surprising: It had previously been assumed that p53 was needed to prevent cancer. We showed that TAp63 can do the job alone.

We discovered that p63 was needed for development: Its loss causes malformations of the limbs, skin, and palate. This finding led others to interrogate p63 and to reveal that its mutation causes seven different human syndromes involving birth defects affecting the limbs, skin, and palate. By generating mouse models for one of these syndromes—ectrodactyly, ectodermal dysplasia, clefting (EEC) syndrome—we found the reason why some children with EEC syndrome have symptoms that are very severe and even life threatening, whereas other children with EEC—even those in the same family with the same p63 mutation—have symptoms that are barely noticeable. Within the past year we published our collaborative studies implicating p63 in limb and cartilage development (Hwang et al. 2018; Lu et al. 2019). We are currently working to understand how p63 regulates stem cell biology and how its perturbation leads to cancer.

Revealing the Genetic Basis of Autism

We found that one of the most common genetic alterations in autism—deletion of a 27-gene cluster on chromosome 16—causes autism-like features. By generating mouse models of autism, we provided the first functional evidence that inheriting fewer copies of these genes leads to features resembling those used to diagnose children with autism. First, "autism" mice had behaviors such as higher activity, resistance to change in environment, sleeping problems, and repetitive/restrictive behavior; each of these features is similar to clinical criteria used to diagnose autism in humans. Second, these mice had distinctive architectural changes in the brain that were detectable by MRI. Our work provides functional evidence for the genetic basis of autism that was not previously appreciated. Our recent collaborative work reveals that genes deleted in autism affect specific regions of the brain and that this leads to sleeping deficits (Lu et al. 2018, 2019). We believe that these models will be invaluable for pinpointing the genes responsible for autism. They could also be used for designing ways to diagnose children with autism and potentially to alleviate some of the detrimental symptoms associated with this syndrome.

CHD5, a New Tumor Suppressor

We discovered *CHD5* as a tumor suppressor mapping to human 1p36—a region of our genomes frequently deleted in cancer. 1p36 deletions occur in many different types of human tumors, including those of the epithelia, brain, and blood. Although this suggested that a cancer-suppressing gene resided in this region, the gene responsible was unknown. By generating mice with deletions and duplications of the genomic region corresponding to 1p36 using chromosome engineering technology—a strategy with which we can generate precise chromosome rearrangements in the mouse—we identified a region of the genome with potent tumor suppressive activity. Using genetic and molecular approaches, we discovered *CHD5* as the tumor suppressor gene in the region and found that its product turns on a network of tumor suppressors. In addition, we found that *CHD5* is frequently deleted in human glioma. Chromosome engineering proved such a powerful way to identify cancer genes that we also used it for studying neurodevelopmental syndromes, including schizophrenia and autism. This technology was essential for several collaborative studies focused on copy number variations responsible for autism.

We continue to focus on defining the role of CHD5 in chromatin dynamics and deciphering how dysregulation of CHD5 and the pathways it regulates leads to disease. We found that Chd5 uses its plant homeodomains to bind histone 3, and that this is essential for tumor suppression. Our work paved the way for further discoveries, and CHD5 is now known to be mutated in human cancers of the breast, ovary, and prostate, as well as in melanoma, glioma, and neuroblastoma. Furthermore, recent reports indicate that patients with high levels of CHD5 have much better overall survival than those with low levels. We found that Chd5 is essential for packaging DNA, and that loss of Chd5 causes improperly packaged DNA that is prone to DNA damage. Intriguingly, Chd5's absence is particularly important during the process of sperm maturation—an event in which the DNA is first unpackaged and then repackaged using an elaborate series of steps—and that deficiency of Chd5 causes male infertility. We discovered that Chd5 is expressed highly in neurons, and that Chd5 plays a pivotal role in the brain, suggesting that inappropriate DNA packaging contributes to neurodevelopmental syndromes such as autism. We reported that, like CHD5, other members of the CHD family regulate chromatin and are implicated in cancer. Within the past year, we discovered that Chd5 regulates a ribosome biogenesis switch that dictates neuronal cell fate and that Chd5 deficiency leads to an excessive number of astrocytes and too few neurons. We are currently delving deeper into the mechanisms whereby Chd5-mediated regulation of chromatin affects gene expression cascades regulating neuronal stem cells and how deregulation of this process sets the stage for neurodevelopmental syndromes and cancer.

PUBLICATIONS

Hwang D-W, Jaganathan A, Shrestha P, Jin Y, El-Amine N, Wang S, Hammell M, Mills AA. 2018. Chromatin-mediated translational control is essential for neural cell fate specification. *Life Sci Alliance* **1:** e201700016. doi:10.26508/lsa.201700016

Lu HC, Mills AA, Tian D. 2018. Altered synaptic transmission and maturation of hippocampal CA1 neurons in a mouse model of human chr16p11.2 microdeletion. *J Neurophysiol* **119:** 1005–1018. doi:10.1152/jn.00306.2017

In Press

Lu HC, Pollack H, Lefante JJ, Mills AA, Tian D. 2019. Altered sleep architecture, rapid eye movement sleep, and neural oscillation in a mouse model of human chromosome 16p11.2 microdeletion. Sleep **42:** pii zsy253. doi:10.1093/sleep/zsy253

CANCER AND HUMAN GENETICS

M. Wigler J. Alexander, P. Andrews, C. Danyko, A. Gruet, I. Hakker, J. Kendall, S. Li, B. Ma, S. Park, C. Podszus, M. Riggs, M. Ronemus, J. Rosenbaum, A. Stepansky, D. Trimboli, Z. Wang, M. Wroten, B. Yamrom, C. Ye, C. Yoon, Z. Yu

Introduction

Our lab works in three areas: genomics, cancer, and autism. The first area is largely a collaborative effort with Kenny Ye from Albert Einstein College of Medicine and Dan Levy from CSHL; the second with Alex Krasnitz (CSHL) and Levy; and the third with Ivan Iossifov (CSHL), Levy, and Ye. The unifying themes are identifying genomic variations that characterize disorders and then using the variations to understand disorders and to improve treatments. To accomplish this, we develop molecular and computational tools for genomics.

Genomics

We continue to build genomic tools. The most active areas during the past year have been error-free sequence variant detection, single-cell sequencing of genomes and transcripts, measuring biallelic transcription bias, haplo-assembly of genomes, and measuring transmission bias from parents to their children. We will describe these tools and relate them to other ongoing efforts.

(1) We are motivated to monitor the progress of cancer patients by measuring their residual disease (MRD). Nucleotide sequencing is among the most powerful detection methods available and, with care, can be used for accurate quantitation. In the cancer context, we achieve this by using tumor-specific variants that distinguish the DNA or RNA of the cancer cells from the patient's normal cells. DNA sequencing is error-prone (one error per 104 positions), which limits its sensitivity for this application, and polymerase chain reaction (PCR) distortion limits its use for quantitation. Therefore, we developed methods for error-correcting sequences that lower error rate by two orders of magnitude and eliminate PCR distortion. These methods are based on the concept and techniques for altering original template nucleic acid molecules with "varietal tags" and obtaining many independent reads of original templates. The concept, value, and application of varietal tagging has been described in many past years' annual reports. We call our present method MASQ (multiplex accurate and sensitive quantitation). Developed by Zihua Wand in the Wigler group and Andrea Moffit in the Levy group, MASQ can trace up to 50 tumor variants at a time to a depth of one part per million each. We are applying MASQ to measuring neoplastic load (see later sections).

(2) We were the first to use single-cell methods for characterizing cancer (9). Over the years we made incremental progress with the methodology, but this past year, because of the efforts of Siran Li, we have taken a quantum leap forward. Using what we call BAG technology, we embalm single cells or nuclei into beads of polyacrylamide gel (BAG) and copolymerize our choice of their molecular contents: DNA, RNA, protein, or any selection thereof. We tag the captured molecules in individual beads by "pool and split" technology (12). BAG technology is cheap, deep, flexible, scalable from low (100) to high (>10,000) cell input, uses sampling without loss, and has very low levels of cross-contamination. The sequence produced by BAG is error-correcting because we also add varietal tags to each of the captured molecules. We use this technology to measure copy number in cancers, detect cancer cells in the body, monitor the host response to cancer by clarifying the cell identities in mixed cell populations, and determine variable expression patterns in mixed cell populations (see the section on Autism).

(3) Some of our research explores how genetic variation alters gene expression. To do this accurately, we use variation in single-nucleotide polymorphisms (SNPs) to distinguish alleles and use techniques for unbiased allele mapping to a reference genome. Other unbiased methods exist, but we have data available to us that the other methods do not provide. We have a whole-genome sequence (WGS) from each person, enabling us to create a reference genome for that individual. We apply this method to detect *cis*-acting variation that alters transcription, loss of parental imprinting, and sporadic allele silencing.

(4) Correct haplo-assembly is sometimes needed in genomics and genetics. This task was discussed in the *2017 Annual Report*, along with two techniques: MuSeq and HaHa. We will not repeat a description here, except to say that we have extended the bench techniques and algorithms for MuSeq haplo-assembly. We have demonstrated the method works on two different 5-kb genome segments, one including the problematic HLA-B region. Improvements on the bench techniques should soon allow the method to be used on any length of genome that can be copied by long-read polymerases in vitro. Applications of the method are contemplated for resolving problems of haplotyping generally, de novo assembly of new organisms, and precise structural determination for complex de novo rearrangements.

(5) Finally, we have developed tools for measuring transmission bias from parents to their children. These tools have the potential to investigate the contribution of genetic mechanisms to phenotypes. Our first tool was A2DS (16). Our group has since developed a second, PHSS (parental haplo-sharing by siblings). Whereas A2DS measures ancestral variation shared between families, PHSS measures recent and ancestral variation shared within a family. In particular, we measure genomic sharing between siblings—separately for each parent. Half the autosomal genome of a child derives from each parent. However, two siblings do not typically share precisely the same half from each parent, as each parent has two haplotypes. The null expectation is that a sibling pair (excluding monozygotic [MZ] twins) will share half of their parental inheritance with a statistical variation determined by meiotic block lengths. If, however, concordant affected siblings share causative alleles from one parent, the expectation is that they will on average share more than half of that parent's haplotypes. If the parent transmits a highly penetrant allele, then the expectation for discordant siblings is that they will share less than half. We have made these notions statistically more precise by relating the observed sharing of haplotypes, H, to the underlying net sharing at Q causal loci. The method uses loci that are heterozygous in one parent and homozygous in another to correctly phase the offspring at those loci. We apply PHSS to reach some startling conclusions in the analysis of autism (see the section on Autism).

Cancer

Until recently most academic cancer research has focused on genetic causes and treatments. For some patients—those at low risk of recurrence—surgery suffices as treatment. For others, therapies are ineffective or lose their effectiveness. In the absence of predictably effective therapies, clinicians need to know when intervention has succeeded, when to change therapy, and how to spare patients costly and possibly harmful intervention when not needed. With the goal of providing clinical guidance, we have been learning how to apply molecular and genomic tools to assess the risk of relapse, to detect it early, and to gauge the patient's response to therapy. These methods might be used someday for early detection of cancer.

To accomplish these goals we have used three classes of technologies: SMASH (15) and TreSeq (unpublished, but see the *2017 Annual Report*) to measure copy number variation, MASQ for error-free sequence detection, and single-cell sequencing for identifying cell type.

(1) The methods for copy number measurement (SMASH and TreSeq) were described in previous years' reports and are still being used in a collaboration on breast cancer with Anders Zetterberg of the Karolinska Institut, Sweden (6). We have recently shown that sampling at multiple sites within a tumor is needed for correct risk assessment. We have also established that proper risk assessment

requires knowledge of the breast cancer type. Many copy number events in estrogen receptor (ER)-positive cancers is a danger sign, whereas in ER-negative cancers this is expected and increases risk only if the numbers are very high.

(2) The detection of tumor signature in blood is central to carrying out our idea of measuring response to therapy and early detection of relapse. We use the cancer signature obtained by comparing DNA in the primary to host DNA to find single-nucleotide variants (SNVs) that are specific to the neoplasm. For solid cancer, we then use MASQ, as described in the Genomics section, to monitor the presence of that signature in cell-free DNA (see 3, 5, 14) and in circulating cells (see 7, 8, 10). At present, we are having more success with cell-free DNA than circulating cells in breast (a collaboration with Northwell) and pancreatic cancer (a collaboration with Boston Deaconess Hospital, through Lakshmi Muthuswamy and Senthil Muthuswamy). We have also had success accurately measuring residual leukemia in a collaboration with Northwell (through Steve Allen and Jonathan Kolitz).

(3) We have used single-cell techniques in the past to detect tumor heterogeneity (9) and to evaluate prostate cancer risk (1). We are continuing the latter in collaboration with Herbert Lepor at NYU. Other applications of single-cell methods include monitoring the host response to cancer by clarifying cell identities in the mixed-cell populations found at the tumor site and characterizing the rare epithelial cells in circulation. The latter has broader application beyond cancer: It might pertain to determining sites of organ inflammation in patients with undiagnosed organ damage.

Autism

Autism is a neurodevelopmental disorder characterized by cognitive and social deficits that occurs in 2% of children, with about a 4:1 gender bias (boys to girls). The evidence of a genetic etiology of autism is very strong, and our lab has been characterizing the genetic landscape of the disorder since 2004 in an effort funded largely by the Simons Foundation. We have determined that there are high-risk and low-risk mate pairs. Only the high-risk mate pairs can generate multiplex families (at least two children affected), whereas the low-risk mate pairs generate about half of the simplex families (only one child affected) (17). An excess of de novo mutations is found almost entirely in the simplex families.

Our effort over the past year has been mainly (but not exclusively) in three areas: evidence for contribution of a de novo noncoding mutation, a pilot study to evaluate causation using RNA expression, and analysis of transmission in multiplex and simplex families. The analysis has been based on data from three collections: the Simons Simplex Collection (SSC) (4), the Autism Genetic Resource Exchange (AGRE), and SPARK (13). We use WGS prepared at the New York Genome Center (NYGC), and SNP chip data from Regeneron, all made at the behest of the Simons Foundation.

(1) Within simplex families, Ivan Iossifov found significantly more de novo mutation in introns in the affected than in their unaffected siblings. We estimate, with some extrapolation, that these noncoding mutations contribute to ~7% of the autism in the SSC. We add together estimates of the excess from all de novo sources—namely, copy number mutation (5%), missense mutations (12%), potentially truncating mutations (a.k.a. LGD, 9%), noncoding intronic mutations (7%) (11), and estimates by others for noncoding regulatory mutations (7%) (2). Thus, ~40% of simplex autism has a contribution from de novo mutation. Because simplex families are about half low-risk, and because excess de novo mutations are not found in high-risk families (work in progress), we conclude that de novo mutation contributes to the diagnosis of autism in 80% of the offspring from low-risk mate pairs. This percentage is very close to predictions we made in 2007 (17).

(2) Driven by our results with noncoding mutations, we initiated a study of gene expression. The SSC contains cultured lymphoid cells (CLCs) from virtually all members of the collection. We wanted to confirm that some of the de novo noncoding mutations altered expression of the *cis*-allele. We conducted a pilot study of 50 families, a collaboration with the NYGC. The main tool was measurement of expression levels and relative allele expression (see Genomics section). This study is

now more than half complete, enough for the following assessments: We achieved uniformity over technical replicates; we achieved uniformity over biological replicates; and we observed excellent balance for allele expression—when there is gene imbalance, it most often follows a Mendelian pattern of transmission. Some of the exceptions to Mendelian transmission are expected; we generally observe proper parental imprinting. In addition, we find examples in which de novo mutation alters gene expression: We find examples of loss of imprinting, and—importantly—we find many examples of unexplained sporadic allele suppression (see below). We should therefore be able to measure contributions to autism from these latter sources.

In a collaboration with Kristin Baldwin (Scripps, La Jolla), we showed that CLCs from the SSC could be used to derive stem cells (induced pluripotent stem cells [iPSCs]) that could be differentiated into neural lineages (NLs). Analyzing the RNAs of CLCs, iPSCs, and NLs, we determined that biased allele expression is largely conserved between lineages, and that most (70%) candidate autism genes are expressed in both types. Thus, analysis of the RNA profiling from CLCs should be relevant to neural lineages, and hence to autism. In any event, we can use CLC-to-NL conversion for validation.

In the course of our pilot study, we found evidence of unexplained sporadic allele silencing. The evidence was that some expression bias was not seen in parents, in identical twins, nor in the one-quarter of the genome of a sibling pair that is identical. There are too many of these events to be explained by de novo mutation in the zygote. Preliminary analysis indicates that some occur prior to X-chromosome inactivation (the 100- to 200-cell embryonic stage). They are sporadic in that there is little overlap in the silenced loci between different sibling pairs. This sporadic unexplained silencing may have a trivial technical explanation, but it might represent a random epigenetic source of expressive irregularity explaining variable penetrance, and thus be of great importance for understanding disease severity.

(3) We believe we have a good explanation for autism in many low-risk families—namely, de novo mutation affecting critical genes needed for normal neurodevelopment. But what explains autism in high-risk families? We developed several genomic tools for exploring transmission from parents. Results from one of these tools, A2DS, were previously described and published (16)—namely, that distant ancestral alleles are shared between unrelated children with autism. We had an additional (but statistically weak) result from that method—namely, that fathers seemed to be the preferred carrier of these ancestral alleles. This year we confirmed all the results of A2DS from AGRE and SSC with results from SPARK, a newly available set of multiplex and simplex families collected by the Simons Foundation.

Theories differ with respect to the nature of the transmitted variants. Given the very reasonable assumption that genetic variation is transmitted equally to males and females, it is striking that females have a lower incidence of autism than males. In multiplex families, following two prior children with autism, the next male born has a nearly 50% chance of having autism (42/86), whereas the chance of autism in the next-born female is about 25% (17/79) (17). We confirmed these statistics with SPARK. In simple models of genetic causation, it follows that more females than males carrying causative loci would be asymptomatic.

So, the result of preferential transmission from fathers, seen with A2DS, struck us as surprising. We desired a more definitive result, and so we developed a method called PHSS (see Genomics section), which does not depend on distant ancestry but directly examines shared loci between siblings. In preliminary results, we have confirmed sharing in concordant siblings—mostly coming from the father's haplotypes. Based on our results, the probability that mothers transmit more than fathers to concordant siblings is at most 0.05. We observe something similar for discordant siblings: They avoid sharing alleles coming from the father more than they avoid sharing alleles coming from the mother. From our results we also conclude that risk alleles are of high penetrance and number a little less than one per family in families at high risk of autism.

The father's role is quite a surprise, as it flies in the face of the lower autism incidence in girls. There are four possible explanations we can consider: Girls with an autism genotype die in utero; men have sperm dominated by mutation or by epigenetic modulation;

fathers with borderline autistic features have a selective mating advantage; and antigenic incompatibility between mother and father results in maternal–fetal incompatibility that sometimes damages the fetus. This last is the hypothesis we prefer, for reasons too numerous to explain here. If we are correct, then some autism might be preventable by immunological intervention, just as we intervene to prevent hydrops fetalis and GALD (gestational alloimmune liver disease).

REFERENCES

1. Alexander J, Kendall J, McIndoo J, Rodgers L, Aboukhalil R, Levy D, Stepansky A, Sun G, Chobardjiev L, Riggs M, et al. 2017. Utility of single cell genomics in diagnostic evaluation of prostate cancer. *Cancer Res* **78:** 348–358.
2. An JY, Lin K, Zhu L, Werling DM, Dong S, Brand H, Wang HZ, Zhao X, Schwartz GB, Collins RL, et al. 2018. Genome-wide de novo risk score implicates promoter variation in autism spectrum disorder. *Science* **362**: eaat6576. doi:10.1126/science.aat6576
3. Coombes RC, Page K, Salari R, Hastings RK, Armstrong A, Ahmed S, Ali S, Cleator S, Kenny L, Stebbing J, et al. 2019. Personalized detection of circulating tumor DNA antedates breast cancer metastatic recurrence. *Clinical Cancer Research* **18:** 3663.
4. Fischbach GD, Lord C. 2010. The Simons Simplex Collection (SSC): a resource for identification of autism genetic risk factors. *Neuron* **68:** 192–195. doi:10.1016/j.neuron.2010.10.006
5. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, Cheang M, Osin P, Nerurkar A, Kozarewa I, et al. 2015. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci Transl Med* **7:** 302.
6. Hicks J, Krasnitz A, Lakshmi B, Navin N, Riggs M, Leibu E, Esposito D, Alexander J, Troge J, Grubor V, et al. 2006. Novel patterns of genomic rearrangement and their association with survival in breast cancer. *Genome Res* **16:** 1465–1479.
7. Kalinich M, Bhan I, Kwan TT, Miyamoto DT, Javaid S, LiCausi JA, Milner JD, Hong X, Goyal L, Sil S, et al. 2017. An RNA-based signature enables high specificity detection of circulating tumor cells in hepatocellular carcinoma. *Proc Natl Acad Sci* **114:** 1123–1128.
8. Kuske A, Gorges TM, Tennstedt P, Tiebel A-K, Pompe R, Preißer F, Prues S, Mazel M, Markou A, Lianidou E, et al. 2016. Improved detection of circulating tumor cells in non-metastatic high-risk prostate cancer patients. *Sci Rep* **6:** 39736.
9. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, et al. 2011. Tumor evolution inferred by single cell sequencing. *Nature* **472:** 90–94.
10. Ozkumur E, Shah AM, Ciciliano JC, Emmink BL, Miyamoto DT, Brachtel E, Yu M, Chen P-I, Morgan B, Trautwein J, et al. 2013. Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells. *Sci Transl Med* **5:** 179.
11. Ronemus M, Iossifov I, Levy D, Wigler M. 2014. The role of de novo mutation in the genetics of autism spectrum disorders. *Nat Rev Genet* **15:** 133–141.
12. Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, et al. 2018. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360:** 176–182.
13. The SPARK Consortium. 2018. SPARK: a US Cohort of 50,000 families to accelerate autism research. *Neuron* **97:** 488–493. doi:10.1016/j.neuron.2018.01.015
14. Tie J, Wang Y, Tomasetti C, Li L, Springer S, Kinde I, Silliman N, Tacey M, Wong H-L, Christie M, et al. 2016. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci Transl Med* **8:** 346.
15. Wang Z, Andrews P, Kendall J, Ma B, Hakker I, Rodgers L, Ronemus M, Wigler M, Levy D. 2016. SMASH, a fragmentation and sequencing method for genomic copy number analysis. *Genome Res* **26:** 1–8.
16. Ye K, Iossifov I, Levy D, Yamrom B, Buja A, Krieger A, Wigler M. 2017. Measuring shared variants in cohorts of discordant siblings with applications to autism. *Proc Natl Acad Sci* **114:** 7073–7076.
17. Zhao X, Leotta A, Qiu S, Kustanovich V, Lajonchere C, Geschwind DH, Law K, Law L, Qui S, Lord C, et al. 2007. A unified genetic theory for sporadic and inherited autism. *Proc Natl Acad Sci* **104:** 12831–12836.

CANCER: SIGNAL TRANSDUCTION

Mikala Egeblad and colleagues study cancer and, in particular, the microenvironment in which the cancer cells arise and live. Solid tumors are abnormally organized tissues that contain not only cancer cells but also various other stromal cell types and an extracellular matrix, and these latter components constitute the microenvironment. Communications between the different components of the tumor influence its growth, its response to therapy, and its ability to metastasize. Among the tumor-associated stromal cells, the lab's main focus is on myeloid-derived immune cells, a diverse group of cells that can enhance angiogenesis and metastasis and can suppress the cytotoxic immune response against tumors. Egeblad is interested in how different types of myeloid cells are recruited to tumors and how their behaviors—for example, their physical interactions with cancer cells and other immune cells—influence cancer progression, including metastasis. The Egeblad lab studies the importance of the myeloid cells using mouse models of breast and pancreatic cancer and real-time imaging of cells in tumors in live mice. This enables them to follow the behaviors of and the interactions between cancer and myeloid cells in tumors during progression or treatment. This technique was instrumental when the lab showed that cancer drug therapy can be boosted by altering components of the tumor microenvironments, specifically reducing either matrix metalloproteinases (enzymes secreted by myeloid cells) or chemokine receptors (signal receptors on myeloid cells). Most recently, the Egeblad lab has shown that when a specific type of myeloid cell, called a neutrophil, is activated during inflammation, it can awaken sleeping cancer to cause cancer recurrence. The neutrophils do so by forming so-called neutrophil extracellular traps, structures of extracellular DNA, and these alter the extracellular matrix surrounding the sleeping cancer cells to provide a wake-up signal.

Douglas Fearon's laboratory studies the interaction between cancer and the immune system. Their underlying premise is that the tumor microenvironment is immune suppressive because cancer cells elicit responses characteristic of wound healing and tissue regeneration. This approach has led to the finding that activated fibroblasts in the tumor stroma mediate immune suppression in several mouse models of cancer, including the autochthonous model of pancreatic ductal adenocarcinoma of the Tuveson lab. Their understanding of the basis of immune suppression is evolving, but they know that it involves the production of the chemokine CXCL12 by the fibroblastic stromal cells, binding of this CXCL12 by pancreatic cancer cells, and exclusion of T cells from the vicinity of the cancer cells. T-cell exclusion, which also occurs in several types of human adenocarcinomas, causes antagonists of T-cell checkpoints to be ineffective, despite the presence of cancer-specific $CD8^+$ T cells. This immune suppression is interrupted by administering AMD3100, an inhibitor of CXCR4, the receptor for CXCL12, which leads to the rapid accumulation of T cells among cancer cells, thereby uncovering the efficacy of anti-PD-L1 and eliminating cancer cells. Because human pancreatic cancer has certain immunological characteristics of the mouse model, a phase 1 clinical trial of AMD3100 in patients with pancreatic cancer was initiated in 2015. Some of their next steps are to determine the biological process that causes cancer cells to express nonmutated, shared antigens and the means by which dormant metastases escape immune elimination.

Darryl Pappin's lab develops chemical and computational methods for analysis of proteins and peptides. These are fundamental tools for proteomics, and they are vital in many fields of biological investigation. Proteins and peptides are typically analyzed via mass spectrometry, a method that involves fragmenting samples by colliding them with gas atoms in a vacuum. Masses of the resulting fragments are measured, and computer algorithms match the results with known or predicted molecules whose amino acid sequences are either known or inferred. The Pappin lab has

developed search engines for mass spectrometry data that enable investigators to sift hundreds of thousands of experimental spectra at a time for database matches. They also seek to reduce sample complexity via an approach they call chemical sorting. This includes the use of chelation to enrich phosphopeptides from the total peptide pool and the use of specific affinity-tagged small-molecule inhibitors to segregate classes of kinases or phosphatases for more specific mass spectroscopic analysis.

Despite their large variety of genetic abnormalities, cancer cells have been found to be extremely sensitive to the reversal of certain mutations. **Raffaella Sordella** and colleagues study why cells in certain cancers are responsive to the inhibition of one particular gene or gene product. Why, for instance, do non-small-cell lung cancer (NSCLC) cells that have a particular mutation in the epidermal growth factor receptor (EGFR) respond dramatically to its inhibition by the drug Tarceva®? This occurs in 15%–20% of patients, the great majority of whom, within 1–3 years, develop resistance. Various mutations have been implicated in about half of resistant patients. Sordella and colleagues have discovered a new resistance mechanism in a subpopulation of NSCLC cells that are intrinsically resistant to Tarceva. These tumor cells were observed to secrete elevated amounts of a growth factor called transforming growth factor-β (TGF-β), which in turn increases secretion of interleukin-6 (IL-6), an immune signaling molecule. Significantly, these effects were independent of the EGFR pathway. The team therefore hypothesizes that inflammation is one of the factors that can render a tumor cell resistant to treatment with Tarceva. In other work, Sordella collaborates with the Krainer lab to study whether alternative splicing has a role in the failure of p53-mediated senescence to halt oncogenesis in certain lung cancers.

Nicholas Tonks and colleagues study a family of enzymes called protein tyrosine phosphatases, or PTPs, which remove phosphate groups from proteins and other signaling molecules, such as lipids, in cells. Disruption of PTP function is a cause of major human diseases, and several of the PTPs are potential therapeutic targets for such diseases. Tonks' group seeks to fully characterize the PTP family, understanding how PTP activity is controlled and how PTPs modify signaling pathways. In addition, they are working to determine how those pathways are abrogated in serious illnesses, including cancer, diabetes, and Parkinson's disease. The overall goal is to identify new targets and strategies for therapeutic intervention in human disease. Tonks and colleagues have defined new roles for PTPs in regulating signaling events in breast cancer, identifying three PTPs as novel potential tumor suppressors. They have characterized the regulation of PTP1B by reversible oxidation, showing that it is regulated by covalent modification of the active site by hydrogen sulfide (H_2S) under conditions of endoplasmic reticulum (ER) stress that are linked to protein folding–related pathologies, such as Parkinson's and Alzheimer's. In addition, they have generated recombinant antibodies that selectively recognize the oxidized conformation of PTP1B; these antibodies display the ability to promote insulin signaling in cells and suggest novel approaches to therapy for diabetes. Finally, they have also discovered a novel mechanism for allosteric regulation of PTP1B activity, offering the possibility of developing small-molecule drugs that could inhibit the phosphatase and thereby modulate signaling by insulin and the oncoprotein tyrosine kinase HER2, potentially offering new ways to treat insulin resistance in type-2 diabetes and breast cancer.

Lloyd Trotman's recent research path begins at his discovery some years ago that the loss of a single copy of a master tumor-suppressing gene called PTEN is sufficient to permit tumors to develop in animal models of prostate cancer. His team later found that complete loss of PTEN paradoxically triggers senescence, an arrested state that delays or blocks cancer development in affected cells. These findings explained why many patients only display partial loss of this tumor suppressor when diagnosed with prostate cancer. Now the team is researching ways to restore the PTEN protein levels in these patients. This therapeutic approach could slow disease progression and thus greatly reduce the need for surgical removal of the prostate or similar drastic interventions that

carry the risks of incontinence and impotence. Their second approach to combat prostate cancer is to model the lethal metastatic disease in genetically engineered mice. They are developing a novel approach that allows for quick generation and visualization of metastatic disease. The efficacy of existing and novel late-stage therapies, such as antihormonal therapy, can then be tested and optimized in these animals. At the same time, the Trotman lab is exploring the genome alterations associated with metastatic disease and resistance to therapy. To this end, they use single- and multicell genome-sequencing techniques developed at CSHL by Dr. Wigler and Dr. Hicks.

David Tuveson's laboratory uses murine and human models of pancreatic cancer to explore the fundamental biology of malignancy and thereby identify new diagnostic and treatment strategies. The lab's approaches run the gamut from designing new model systems of disease to developing new therapeutic and diagnostic approaches for rapid evaluation in preclinical and clinical settings. The lab's studies make use of organoid cultures—three-dimensional cultures of normal or cancerous epithelia—as ex vivo models to probe cancer biology. Current projects in the lab explore changes in redox metabolism associated with pancreatic cancer tumorigenesis, dissect signaling by the Ras oncogene, discover new biomarkers of early pancreas cancer, and identify mechanisms of cross talk between pancreatic cancer cells and the tumor stroma. Novel treatment approaches suggested by these studies are then tested by performing therapeutic experiments in mouse models. To dissect molecular changes associated with pancreatic tumorigenesis, the Tuveson lab has generated a large collection of human patient-derived organoid models. By measuring the therapeutic sensitivities of patient-derived organoids, the lab is working to identify novel strategies to treat patients as well as markers of therapeutic response. The Tuveson laboratory maintains strong links to clinical research, and the ultimate goal is confirmation of preclinical findings in early-phase trials. Collectively, the lab's bench-to-bedside approach is codified as the "Cancer Therapeutics Initiative," and this initiative will provide these same approaches to the entire CSHL cancer community.

Dr. Tuveson serves as Director of the Cold Spring Harbor Laboratory Cancer Center and the Chief Scientist for the Lustgarten Foundation.

Linda Van Aelst's lab studies how aberrations in intracellular signaling involving enzymes called small GTPases can result in disease. They are particularly interested in Ras and Rho GTPases, which help control cellular growth, differentiation, and morphogenesis. Alterations affecting Ras and Rho functions are involved in cancer and various neurodevelopmental disorders. Van Aelst's team has extended its prior study of mutations in a Rho-linked gene called *oligophrenin-1* (*OPHN1*), part of an effort to connect the genetic abnormalities associated with mental retardation to biological processes that establish and modify the function of neuronal circuits. In addition to a role for *OPHN1* in activity-driven glutamatergic synapse development, lab members have obtained evidence that *OPHN1* has a critical role in mediating mGluR-LTD (long-term depression), a form of long-term synaptic plasticity, in CA1 hippocampal neurons. Their findings provide novel insight not only into the mechanism and function of mGluR-dependent LTD, but also into the cellular basis by which mutations in *OPHN1* could contribute to the cognitive deficits observed in patients. Defects in cortical neurogenesis have been associated with cerebral malformations and disorders of cortical organization. The Van Aelst team discovered that interfering with the function of the Rho activator DOCK7 in neuronal progenitors in embryonic cerebral cortices results in an increase in the number of proliferating neuronal progenitors and defects in the genesis of neurons. In an extension of these studies, the Van Aelst team this year showed that DOCK7 has a central regulatory role in the process that determines how and when a radial glial cell progenitor "decides" to either proliferate (i.e., make more progenitor cells like itself) or give rise to cells that will mature, or "differentiate," into pyramidal neurons. These lines of research provide novel insight into mechanisms that coordinate the maintenance of the neural progenitor pool and neurogenesis.

THE TUMOR MICROENVIRONMENT AND CANCER PROGRESSION: HOW HIJACKING HOST CELLS HELPS TUMORS SPREAD

M. Egeblad J. Albrengues, E. Bružas, J. Curtis, J. Dassler Plenker, C. Evans, X. He, L. Maiorino, D. Ng, L. Puckett, M. Shields, L. Sun

Solid tumors are aberrant tissues. Like organs, solid tumors are composed of cancer cells and stroma. The stroma is the supportive framework of the organs and includes the extracellular matrix (ECM) as well as immune cells, fibroblasts, adipocytes, and cells of the vascular system. Interactions between epithelium and stroma are essential for normal organ development as well as for tumor progression. In solid tumors, the stroma is also known as the tumor microenvironment.

We study how the tumor microenvironment influences cancer cells in the context of tumor initiation, growth, drug resistance, and metastasis. We use mouse models of breast and pancreatic cancer together with real-time spinning disk confocal and multiphoton microscopy in living mice, known as intravital imaging. This allows us to study how cancer cell proliferation, migration, and survival are influenced by stromal components in real time.

Neutrophil Extracellular Traps Produced during Inflammation Awaken Dormant Cancer Cells

J. Albrengues, M. Shields, D. Ng, E. Bružas, L. Maiorino, J. Dassler Plenker [in collaboration with V. Kuettner]

Most cancer patients die, not from their original tumor, but of cancer that recurs after metastasizing to a different tissue. Preventing recurrence is therefore a major opportunity to increase cancer patient survival. Recurrence often occurs after a prolonged period of dormancy, a stage in which residual cancer is present but clinically undetectable. Disseminated cancer cells can remain dormant for years, even decades, before recurring, or "awakening," as metastatic cancer. T cells and natural killer cells can eliminate disseminated cancer cells as they start proliferating, preventing them from reaching clinically detectable levels. However, little is known about the cues that drive metastatic recurrence after a long period of clinically undetectable, dormant cancer.

In breast cancer survivors, elevated plasma levels of C-reactive protein, a nonspecific marker of chronic inflammation, are associated with reduced disease-free survival. Smoking is a well established inducer of chronic lung inflammation, and two recent large pooled analysis studies showed that current smoking or prior heavy smoking was significantly associated with an elevated risk of breast cancer recurrence and death due to breast cancer. Whether smoking can cause metastasis through induction of inflammation is not clear, but in a mouse model, experimental lung metastasis was previously shown to be increased twofold by tobacco smoke exposure.

Neutrophils are one of the major types of inflammatory cells. In an experimental model of inflammation, it was recently shown that neutrophils are a critical cell type in cancer cell awakening, but how neutrophils awaken cancer remained unclear. We previously showed that neutrophils can promote metastasis through the formation of neutrophil extracellular traps (NETs)—scaffolds of chromatin with associated cytotoxic enzymes and proteases that are released into the extracellular space, where their physiological function is to trap microorganisms. NETs are generated through a signaling process that involves chromatin decondensation and disintegration of the nuclear membrane. Contents from the neutrophil's secretory granules—including neutrophil elastase (NE), cathepsin G (CG), and matrix metalloproteinase 9 (MMP9)—associate with the decondensed chromatin. Finally, the plasma membrane ruptures, and the protease-associated chromatin fibers are released into the extracellular space.

We found that sustained experimental lung inflammation caused by tobacco smoke exposure or nasal instillation of lipopolysaccharide (LPS) converted

disseminated, dormant cancer cells to aggressively growing metastases in mice. To observe how inflammation influenced the dormant cancer cells, we used our intravital lung imaging. After induction of inflammation, we observed NET-like structures around the dormant cancer cells, and this was associated with initiation of proliferation. Excitingly, treatment with NET-digesting, DNase-I-coated nanoparticles significantly reduced awakening of the cancer after lung inflammation. Thus, NETs were required for awakening. To determine how NETs induced awakening, we developed an in vitro assay. We found that NETs promoted awakening by enabling two NET-associated proteases, NE and MMP9, to sequentially cleave laminin. We propose that the DNA of the NETs concentrates the neutrophil proteases at their substrate. This is likely critical because only sequential proteolytic remodeling of laminin, first by NE and then by MMP9, generated a laminin epitope that induced awakening.

We next determined that NET-cleaved laminin contained an integrin-activating epitope that induced proliferation of dormant cancer cells. This knowledge allowed us to generate blocking antibodies against NET-remodeled laminin, and these antibodies prevented tobacco smoke exposure and LPS from inducing awakening in mice. Preventing dormant cancer cells from awakening is a major opportunity to prolong patient survival, and our data implicate NETs as critical mediators of inflammation-induced metastatic cancer recurrence.

Chronic inflammation and smoking are risk factors in metastatic recurrence. Inflammation is a complex process involving dramatic changes in cellular composition, ECM remodeling, and cytokine production. However, we have now determined that neutrophils recruited during inflammation are key initiators of awakening of dormant cancer, and neutrophils act through the formation of NETs. We envision that NETs and their downstream effectors eventually can be targeted to reduce the risk of cancer recurrence.

Cancer Cell Chemokine Receptor CCR2 Orchestrates Suppression from the Adaptive Immune Response

X. He, L. Puckett, E. Bružas, J. Curtis

The immune system is very efficient at eliminating pathogens that can cause harm to the organism. The immune system also has the potential to eliminate neoplastic cells. The concept of "immune surveillance" was first described more than 50 years ago, and refers to the ability of immune cells to detect tumor cells and destroy them. T cells, part of the adaptive immune system, are critical for tumor immune surveillance. Immune surveillance may lead to a period in which cancer cells are kept in check by the immune system, and the tumor neither expands nor regresses. Eventually, tumors develop means to escape immune control. Tumors have multiple mechanisms of escaping immune control, including cancer cell–intrinsic changes that alter how the cancer cell is recognized by the immune system and extrinsic changes that suppress immune cell activities. As examples of intrinsic changes, cancer cells can decrease the surface expression of major histocompatibility complex (MHC) class I, making them effectively invisible to T lymphocytes. Another mechanism is increased expression of programmed cell death ligand 1 (PD-L1) on the cancer cells. PD-L1 binds the PD-1 receptor on activated T cells, leading to protection against T-cell-mediated killing. Current immunotherapies that target this so-called "immune checkpoint" have led to long-lasting regression in several cancers. Extrinsic mechanisms of immune escape include the downregulation of costimulatory molecules (e.g., CD86) on antigen-presenting cells, the secretion of cytokines that directly inhibit cytotoxic T lymphocytes (CTLs), and the promotion of regulatory T-cell infiltration. In contrast, infiltration of $CD103^+$ dendritic cells (DCs) has emerged as a mechanism by which tumors may be kept under immune control: $CD103^+$ DCs are highly efficient at acquiring and processing exogenous antigens, and they present the antigens on MHC class I molecules directly to $CD8^+$ CTLs.

Chemokines, or chemotactic cytokines, have critical roles in mediating recruitment of immune cells to sites of inflammation and tumors. For example, the C-C chemokine ligand 2 (CCL2) recruits CC chemokine receptor (CCR2)-expressing immune cells to tumors. The primary role of CCR2 in cancer has therefore been considered to be the regulation of immune cell infiltration, and we previously showed that CCL2 causes recruitment of CCR2-expressing monocytes to tumors after treatment with chemotherapy, and these newly recruited monocytes inhibited the chemotherapy response. CCL2/CCR2-mediated recruitment of

CCR2$^+$ inflammatory monocytes to the lung has also been shown to promote breast cancer extravasation and metastasis in mice. Furthermore, elevated levels of CCL2 in tumors and serum are associated with advanced disease and poor prognosis in breast carcinoma patients. These findings have sparked interest in targeting the CCR2 pathway for therapeutic benefit in breast cancer.

However, it is not just immune cells that express CCR2; breast cancer cells also express CCR2. Indeed, CCR2 is up-regulated on breast cancer cells compared to normal breast epithelium. In vitro, cancer cell expression of CCR2 has been proposed to promote tumor growth through increased motility and invasion, cell proliferation, and cell survival. In vivo, the potential role(s) of CCR2 signaling in cancer cells have, however, not been well studied, largely because they were thought to be minor compared to the roles of CCR2 in myeloid cells.

To test the function of CCR2 in breast cancer cells, we used orthotopic transplantation of MMTV-PyMT breast cancer cells as our breast cancer mouse model. We found that *Ccr2* deletion in cancer cells led to reduced tumor growth and twofold longer survival. The longer survival was accompanied by multiple alterations associated with better immune control: increased infiltration and activation of CTLs and CD103$^+$ cross-presenting DCs, as well as up-regulation of MHC class I and down-regulation of checkpoint regulator PD-L1 on the cancer cells. Pharmacological inhibition of CCR2 increased cancer cell sensitivity to CTLs and enabled the cancer cells to induce DC maturation toward the CD103$^+$ subtype. The combination of all these changes likely results in more effective immune surveillance of the *Ccr2*$^{-/-}$ cancer cells and reduced growth of tumors derived from these cells. Indeed, the *Ccr2*$^{-/-}$ cancer cells were not growth-restricted in *Batf3*$^{-/-}$ mice lacking the CD103$^+$ DC subtype or in nude mice lacking CTLs. Thus, breast cancer CCR2 plays a central role in inhibiting immune surveillance.

Our results establish a novel role for CCR2 signaling in cancer cells in orchestration of the suppression of the immune response. These new data, together with our previous findings regarding the role of CCR2 in recruitment of monocytes that protect tumors against chemotherapy, make CCR2 an attractive target in combination with both chemotherapy and immunotherapy.

Lysyl Oxidases Inhibit Metastasis of Pancreatic Cancer

M. Shields, L. Sun, J. Albrengues [in collaboration with V.M. Weaver, UCSF]

Tumors are often discovered as nodules that are harder than the surrounding tissue. This reflects a stiffening in the ECM and changes in the ECM architecture. The interstitial ECM consists of macromolecules such as fibrillar collagens, fibronectin, and proteoglycans. Type I collagen is the major fibrillar collagen in many tissues, and it forms a scaffold that provides stability. Type I collagen also has signaling functions mediated by, for example, integrins. The synthesis, as well as the proteolytic remodeling, of the fibrillar type I collagen increases in many tumors, including breast and pancreatic tumors. Collagen cross-linking in mouse models of mammary carcinoma delays tumor onset and slows tumor progression. Furthermore, collagen architectural structure becomes abnormal with progression of solid tumors: Whereas collagen fibers are curly and oriented in parallel to normal or hyperplastic epithelium, there is a progressive change in the fibers so that they are straighter and mostly perpendicular to the tumor border in the late stages. Lysyl oxidases are a family of five enzymes—lysyl oxidase (LOX) and lysyl-oxidase-like (LOXL) 1–4—that can cross-link collagen, thereby generating linearized fibers. Lysyl oxidases have been shown to promote cancer progression and metastasis in mouse models of breast cancer by generating linearized collagen fibers and enhancing integrin signaling. Consistently, inhibiting lysyl oxidases reduces tumor progression and metastasis in several mouse models of breast cancer.

Pancreatic cancer has a very high incidence of metastasis: About 80% of patients present with metastasis and most of the remaining patients develop metastasis within 2 years of diagnosis. Pancreatic cancer also has a very pronounced tumor microenvironment featuring deposition of type I collagen. Recent reports have shown that reducing the stromal response, either by inhibiting paracrine hedgehog signaling or genetically ablating collagen-producing myofibroblasts, resulted in more aggressive, undifferentiated tumors. Nevertheless, it is not well understood how the collagen-rich ECM affects pancreatic cancer progression. Gene expression analyses indicate that high expression of lysyl oxidase family members correlates with worse outcomes in breast and pancreatic cancer

patients. Despite these findings, a recent clinical trial in pancreatic cancers showed no survival benefit when chemotherapy was combined with targeting of LOXL2 using a blocking antibody.

Lysyl oxidases are recognized for their ability to cross-link ECM proteins, but it has largely been forgotten that LOX was first described as an "antioncogene of *ras*," and its expression was shown to inhibit Ras-mediated transformation and tumor growth in NIH 3T3 fibroblasts. Ras mutations are very rare in breast cancer but present in >90% of pancreatic tumors. Collagen content is also different, with pancreatic tumors generally having a more collagen-rich tumor microenvironment than breast cancer.

We set out to test whether prometastatic effects of lysyl oxidases on ECM cross-linking or antitumorigenic effects mediated by inhibition of Ras signaling would prevail in orthotopic pancreatic tumor models driven by oncogenic Ras. We found that Lox and Loxl2, the two most highly expressed lysyl oxidases, had tumor-suppressive functions in pancreatic cancer, independent of collagen cross-linking. Using intravital imaging, we could observe that dependent on either genetic or pharmacological targeting of lysyl oxidases, the pancreatic cancer cells increased their invasion along linearized collagen fibers. LOX inhibition significantly increased distant metastasis. We discovered that Lox and Loxl2 inhibition, specifically in cancer cells with mutant Ras, activated focal adhesion kinase (FAK) signaling, and this activation was independent of the effects of lysyl oxidases on collagen cross-linking. Importantly, FAK signaling was required for metastasis. If the mutant Ras was downregulated in the cancer cells, then LOX inhibition no longer had an effect on FAK signaling. These results suggest that the inhibitory effects of LOX on pancreatic cancer metastasis are dependent on signaling from Ras and FAK.

Lysyl oxidases, through their prometastatic effect, were attractive targets in pancreatic cancer; however, targeting LOXL2 was ineffective in a clinical trial. Our preclinical work confirms that targeting of lysyl oxidases will not benefit pancreatic cancer patients because of the activation of the proinvasive FAK signaling pathway leading to increased invasion and metastasis. Our results, therefore, strongly caution against inhibiting lysyl oxidases in cancers driven by mutant Ras.

PUBLICATIONS

Albrengues J, Shields MA, Ng D, Park CG, Ambrico A, Poindexter M, Upadhyay P, Uyeminami D, Pommier A, Küttner V, et al. 2018. *Science* **361:** eaao4227. doi:10.1126/science.aao4227

Bružas E, Egeblad M. 2018. Spotlight. Bone talk: activated osteoblasts promote lung cancer growth. *Trends Mol Med* **24:** 237–239. doi:10.1016/j.molmed.2018.01.007

Huang Y-H, Wu X, Klingbeil O, He X-Y, Arun G, Lu B, Somerville TDD, Milazzo JP, Wilkinson JE, Demerdash OE, et al. 2018. POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. *Genes Dev* **32:** 915–928.

Pommier A, Anaparthy N, Memos N, Kelley Z, Gouronnec A, Yan R, Auffray C, Albrengues J, Egeblad M, Iacobuzio-Donahue CA, et al. 2018. Unresolved endoplasmic reticulum stress engenders immune-resistant, latent pancreatic cancer metastases. *Science* **360:** eaao4908. doi:10.1126/science.aao4908

THE INTERACTION BETWEEN THE IMMUNE SYSTEM AND PANCREATIC CANCER

D.T. Fearon J. Li P. Moresco Z. Wang R. Yan M. Yao

The focus of the Fearon lab continues to be the interaction between the host immune system and pancreatic ductal adenocarcinoma (PDA). This past year, we published our discovery that the immune system selects "silent," or "dormant," metastases in a mouse model of this problem (Pommier et al. 2018). The majority of patients with PDA develop metastatic disease after resection of their primary tumor, indicating that they harbored dormant metastases at the time of surgery. Pommier et al. (2018) found that livers from patients and mice with PDA harbor single disseminated cancer cells (DCCs) lacking expression of cytokeratin 19 (CK19) and major histocompatibility complex class I (MHCI) (Pommier et al. 2018). A mouse model was created to determine how these DCCs develop. Intraportal injection of immunogenic PDA cells into preimmunized mice seeded livers only with single, nonreplicating DCCs that were $CK19^{-}$ and $MHCI^{-}$. The DCCs showed an endoplasmic reticulum (ER) stress response, but paradoxically lacked both inositol-requiring enzyme 1α activation and expression of the spliced form of transcription factor XBP1 (XBP1s). Inducible expression of XBP1s in DCCs, in combination with T-cell depletion, stimulated the outgrowth of macrometastatic lesions that expressed CK19 and MHCI. Thus, unresolved ER stress enables DCCs to escape immunity and establish latent metastases in this mouse model of PDA and, possibly, in patients with this cancer. This study complements a collaboration with Mikala Egeblad in which her lab showed that neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in a model of breast cancer (Albrengues et al. 2018).

Finally, following up on our study in 2016 of the systemic effects of the cytokine IL-6 on the impaired and immune suppressive metabolic response to cachexia in models of cachexia, the wasting syndrome suffered by many patients with PDA and other cancers, we collaborated with the Cantley lab at Weill Cornell Medicine to show that the drug fenofibrate, which is an agonist of the PPARα, modifies the IL-6-induced suppression of

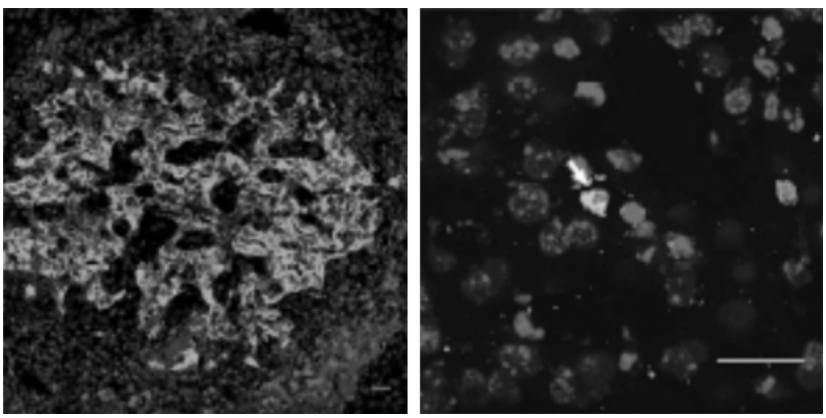


Figure 1. Confocal photomicrographs of two types of pancreatic cancer metastases in the liver that have been reproduced in an animal model. On the *left* is a growing metastatic lesion with many cancer cells (green) that exclude the cancer-specific T cells (red). On the *right* is a single, nongrowing, dormant pancreatic cancer cell (also green) and a nearby, irrelevant T cell (red). The former would be clinically evident and will eventually cause the demise of the animal, whereas the latter is not detectable by standard clinical tests and may start growing at some future time, especially in a host with an impaired immune system. The arrow identifies a dormant pancreatic cancer cell in the liver of a patient with cancer.

the expression of this transcription factor that may be associated with cancer cachexia (Goncalves et al. 2018). This effect ameliorates the muscle wasting and immune suppression that occurs with cachexia.

PUBLICATIONS

Albrengues J, Shields MA, Ng D, Park CG, Ambrico A, Poindexter ME, Upadhyay P, Uyeminami DL, Pommier A, Küttner V, et al. 2018. Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. *Science* **361:** eaao4227. doi:10.1126/science.aao4227

Goncalves MD, Hwang S-K, Pauli C, Murphy CJ, Cheng Z, Hopkins BD, Wu D, Loughran RM, Emerling BM, Zhang G, et al. 2018. Fenofibrate prevents skeletal muscle loss in mice with lung cancer. *Proc Natl Acad Sci* **115:** E743–E752.

Pommier A, Anaparthy N, Memos N, Kelley ZL, Gouronnec A, Yan R, Auffray C, Albrengues J, Egeblad M, Iacobuzio-Donahue CA, et al. 2018. Unresolved endoplasmic reticulum stress engenders immune-resistant, latent pancreatic cancer metastases. *Science* **360:** eaao4908. doi:10.1126/science.aao4908

MASS SPECTROMETRY

D. Pappin A. Makarenko K. Rivera J. Snider
K. Marquant C. Shaw J. Wilson

Mitochondrial Complex I Inhibitors Expose a Vulnerability for Selective Killing of Pten-Null Cells

D.J. Pappin [in collaboration with A. Naguib, G. Mathew, K. Watrud, A. Ambrico, T. Herzka, I.C. Salas, M.F. Lee, N. El-Amine, W. Zheng, and L.C. Trotman, CSHL; C.R. Reczek and N.S. Chandel, Northwestern Medical School; M.E. Di Francesco and J.R. Marszalek, University of Texas]

A hallmark of advanced prostate cancer (PC) is the concomitant loss of PTEN and p53 function. To selectively eliminate such cells, the Trotman laboratory screened cytotoxic compounds on $Pten^{-/-}$,$Trp53^{-/-}$ fibroblasts, and their Pten-WT reference. Highly selective killing of Pten-null cells could be achieved by deguelin, a natural insecticide. Deguelin eliminates Pten-deficient cells through inhibition of mitochondrial complex I (CI). Five hundred-fold higher drug doses were needed to obtain the same killing of Pten-WT cells, even though deguelin blocks their electron transport chain equally well. Selectivity arises because mitochondria of Pten-null cells consume ATP through complex V, instead of producing it. The resulting glucose dependency can be exploited to selectively kill Pten-null cells with clinically relevant CI inhibitors, especially if they are lipophilic. In vivo, deguelin suppressed disease in a genetically engineered mouse model of metastatic PC. The data thus introduce a vulnerability for highly selective targeting of incurable PC with inhibitors of CI.

DPM-1001 Decreased Copper Levels and Ameliorated Deficits in a Mouse Model of Wilson's Disease

K. Rivera, D.J. Pappin [in collaboration with N. Krishnan, C. Felice, and N.K. Tonks, CSHL]

The levels of copper, an essential element in living organisms, are under tight homeostatic control. Inactivating mutations in ATP7B, a P-type Cu-ATPase that functions in copper excretion, promote aberrant accumulation of the metal, primarily in the liver and brain. This condition underlies Wilson's disease, a severe autosomal recessive disorder characterized by profound hepatic and neurological deficits. Current treatment regimens rely on the use of broad-specificity metal chelators as "decoppering" agents; however, there are side effects that limit their effectiveness. In a collaboration with the Pappin laboratory, the Tonks laboratory characterized DPM-1001 {methyl 4-[7-hydroxy-10,13-dimethyl-3-({4-[(pyridin-2-ylmethyl)amino]butyl}amino)hexadeca hydro-1H-cyclopenta[a]phenanthren-17-yl] pentanoate} as a potent and highly selective chelator of copper that is orally bioavailable. Treatment of cell models, including fibroblasts derived from Wilson's disease patients, eliminated adverse effects previously associated with copper accumulation. Furthermore, treatment of the toxic milk mouse model of Wilson's disease with DPM-1001 lowered the levels of copper in the liver and brain, removing excess copper by excretion in the feces while ameliorating symptoms associated with the disease. The data suggest that it may be worthwhile to investigate DPM-1001 further as a new therapeutic agent for the treatment of Wilson's disease, with potential for application in other indications associated with elevated copper, including cancer and neurodegenerative diseases.

Proteome Modifications on Tomato under Extremely High Light-Induced Stress

K. Rivera, D. Pappin [in collaboration with D. Parrine, B.S. Wu, B. Muhammad, X. Zhao, and M. Lefsrud, McGill University]

Abiotic stress reduces photosynthetic yield and plant growth, negatively impacting global crop production, and is a major constraint faced by agriculture. However, the knowledge about the impact on plants of extremely high irradiance is limited. In a collaboration

between the Pappin and Lefsrud laboratories, the first in-depth proteomics analysis of plants treated with a stepped gradient of extremely intense light was performed. The proteomics results showed the presence of specific defense responses to each level of light intensity, with a possible involvement of PsbH, Psb28, PsbR, and PsbS proteins.

PUBLICATIONS

Krishnan N, Felice C, Rivera K, Pappin DJ, Tonks NK. 2018. DPM-1001 decreased copper levels and ameliorated deficits in a mouse model of Wilson's disease. *Genes Dev* **32:** 944–952. doi:10.1101/gad.314658.118

Melnikov SV, Manakongtreecheep K, Rivera KD, Makarenko A, Pappin DJ, Söll D. 2018. Muller's ratchet and ribosome degeneration in the obligate intracellular parasites *Microsporidia*. *Int J Mol Sci* **19:** E4125. doi:10.3390/ijms19124125

Naguib A, Mathew G, Reczek CR, Watrud K, Ambrico A, Herzka T, Salas IC, Lee MF, El-Amine N, Zheng W, et al. 2018. Mitochondrial complex I inhibitors expose a vulnerability for selective killing of Pten-null cells. *Cell Rep* **23:** 58–67. doi:10.1016/j.celrep.2018.03.032

Parrine D, Wu BS, Muhammad B, Rivera K, Pappin D, Zhao X, Lefsrud M. 2018. Proteome modifications on tomato under extreme high light induced-stress. *Proteome Sci* **16:** 20. doi:10.1186/s12953-018-0148-2

TUMOR CELL–INTRINSIC PHENOTYPIC PLASTICITY DRIVES ADAPTIVE CELLULAR REPROGRAMMING AND DRUG RESISTANCE

R. Sordella D. Chaluts M. Miller N. Sachan
A. Gray D. Pal P. Safaroic Tepes
N. Korimerla P. Pisterzi N. Shirole

The tremendous advances in genomics we witnessed in the past decades have led to the identification of specific "driver" mutations and the development of targeted small-molecule inhibitors to block their functions by us and others. There was much excitement when these specific inhibitors were employed in the clinic, as they showed initial remarkable efficacies. For lung cancer, this has been the case for erlotinib, gefitinib, afatinib, and osimertinib for mutant EGFR, vemurafenib and dabrafenib for mutant BRAF, and crizotinib, ceritinib, and alectinib for ALK and/or ROS1 gene rearrangements. These discoveries have revolutionized the management and treatment of cancers and have improved the quality of life for millions of patients worldwide. Unfortunately, their clinical success has been limited by the occurrence of resistance. Ultimately in the majority of the cases, these therapies failed to significantly prolong the lives of individuals with cancer, as acquired resistance to these drugs invariably develops despite initial responses. Many mechanisms of resistance to targeted therapy have been identified by us and others, and they can be generally grouped into three main categories: (1) molecular changes in the driver oncogene as a consequence of the occurrence of secondary mutations, (2) activation of critical signaling pathway(s) in a parallel or downstream fashion, and (3) rewiring of the cellular signaling network to drive pro-survival signaling through a different signaling pathway.

In the past years we have identified a novel class of resistance that encompasses histological transformation from one cell lineage, such as epithelial, to another. We dubbed cells that have morphed into these different cell states drug-resistant mesenchymal (DMR) cells. This population of tumor cells that never completely responds to therapy likely functions as a transition state culminating eventually in a drug-resistant tumor (acquired resistance). These cells originate from a stochastic/epigenetic mechanism centered on the methylation and consequent silencing of a specific locus containing the miRNA MiR335. Importantly, we found that the occurrence of DMR cells is not restricted to the acquisition of resistance to one specific drug but that DMR cells are preexisting in the majority of tumors we have analyzed and are contributing to resistance to almost all treatments. Hence, understanding the mechanisms involved in the survival of this subpopulation of residual tumor cells is critical for identifying successful cancer therapies. If critical signaling pathways can be targeted in a patient in an up-front manner, the emergence of resistance may be delayed or prevented entirely, offering a promising approach to combat the heterogeneity and adaptiveness of most cancers.

Over the years our studies have identified numerous potential Achilles' heels in DMR cells that can be targeted by drugs. These include the AXL, IL-6, and NF-κB pathways, which can be exploited therapeutically.

Interestingly, we also found that DMR cells are characterized by intrinsic defects in their DNA double-strand break (DSB) repair capability. This resulted in an increased accumulation of DNA copy number alterations, genetic diversity of cancer cell populations, and improved adaptability to drug treatment. Because these cells are naturally occurring and can be generated by a stochastic/epigenetic program, our findings suggest that the interconversion between different cell states can promote intratumor genetic heterogeneity, spur the tumor's evolution, and hence increase the tumor's fitness. This is important because it argues that the transition into a drug-tolerant/mesenchymal state could provide a mechanism that allows a small subpopulation of tumor cells to withstand an initial destructive attack of drug to enable their survival until they accumulate more permanent resistance mechanisms. We are now revisiting the data from the original EGFR Tki clinical trial to try to understand if increased genetic instability could be a

signature that might be used to predict response to therapy and time to relapse.

Despite initial frustrations over the inevitable emergence of resistance to targeted kinase inhibitors, these new insights, by providing a novel explanation of how tumors can become drug-resistant, might provide hope for the development of more durably effective treatment strategies.

One of the signaling pathways we found to be constitutively activated in DMR cells was the TGF-β–IL-6 axis. Although we showed the importance of the TGF-β–IL-6 axis in mediating acquired resistance, the detailed mechanistic aspects of how TGF-β induces the expression of IL-6 were not understood. In a recent study, we found that TGF-β inhibition of ZC3H12A was essential in regulating IL-6 expression. ZC3H12A is a bifunctional endoribonuclease and deubiquitinase protein that controls the stability and expression of multiple cytokines. We demonstrated that TGF-β's direct inhibition of ZC3H12A transcription regulated the expression of IL-6 via a bimodal mechanism encompassing increased mRNA stability and mRNA transcription. In the latter case, this occurred by releasing the activation of nuclear factor kappa B (NF-κB), a sequence-specific transcription factor that is best known as a major regulator of inflammatory and innate immune responses. From a biological standpoint, regulation of IL-6 by ZC3H12A was dispensable for TGF-β-induced EMT, but it was instead required for TGF-β-induced drug resistance and metastatic colonization of cancer cells. These findings open novel opportunities for the treatment of non-small-cell lung cancer (NSCLC). At the same time, the increased activation of NF-κB also suggests that the inhibition of MCPIP1 by TGF-β could be important for the expression of other cytokines and, in general, in orchestrating the dialog of cancer cells with the surrounding tumor microenvironment.

In addition to the TGF-β–IL-6 axis, other signaling components are altered in DMR cells. In particular, DMR cells undergo a switch in their receptor tyrosine kinase repertoire. In addition to expressing EGFR, they also express other RTKs such as PDGFR.

Since the discovery of the first receptor tyrosine kinase (RTK) more than 25 years ago, many members of the receptor tyrosine kinase family have been shown to play critical roles in a multiplicity of biological processes such as cell proliferation, motility, migration, differentiation, survival, and differentiation.

The ability to interpret extracellular stimuli with robust and precise cellular responses relies on signaling specificity. Factors contributing to protein interaction specificity such as colocalization, co-expression, docking motifs, and scaffold or adaptor proteins have been shown to ensure that a specific set of signaling cues will be triggered only when a defined RTK is activated.

Yet RKTs, despite activating similar signaling effector proteins and being localized in the same cellular compartment, have distinct biological activities. Hence a fundamental question arises: How can different RTKs prompt distinct cellular responses such as proliferation versus differentiation while activating highly similar components of the cellular signaling network? For example, if both the EGF and PDGF receptors activate similar signaling cascades, how does EGFR lead to cell proliferation and the PDGF receptor promote increased motility and extracellular matrix invasion in the same cells?

To address this conundrum, it has been proposed that the integration of packets of information within cellular signaling networks is important. The intensity and duration of signaling, as well their contextual activation within signaling networks, have been proposed to be at the root of RTKs' specificity.

Although not extensively studied in RTKs, these macromolecular-level elements defining signaling specificity have been shown to be complemented by the ability of a given kinase domain to recognize and phosphorylate a specific peptide within the protein substrate. Although many studies have investigated substrate specificity in soluble tyrosine kinase, only a handful have been focused on RTKs. By using a positional scanning peptide library screening approach, we found that PDGFR and EGFR have profound substrate specificity differences.

Peptide specificity arises due to the electrostatic, hydrogen, and hydrophobic interactions between the peptide and catalytic domain of the kinases. Detailed analysis of the catalytic domain provided further insight into these interactions and, in principle, can explain the differences in substrate specificity of these kinases.

We found that the differences we identified in our in vitro assays impinge upon important in vivo regulatory functions by changing the kinase activity, protein stability, and interactions of preferred substrates with other signaling proteins. In particular, we observed that PDGFR, unlike EGFR, can phosphorylate the

suppressor of cytokine signaling 3 (SOCS3), a key negative regulator of the IL-6/STAT3 signaling pathway, at a specific tyrosine residue (Y166). As a result, we found SOCS3 to be rapidly degraded. Consistent with these findings, cells in which both the IL-6 and the PDGF axis were activated have similar features of cells with constitutive IL-6/STAT3 activation. These cells have distinct changes in their morphology and behaviors and acquire pro-metastatic features.

The cross talk between these signaling axes centered around the regulation of SOCS3 and could explain some of the contextual activity of STAT3 and some of the features of DMR cells. In general, our findings could be of particular interest because they could support the intriguing possibility that the gene-duplicated events at the root of the vertebrate branch points generated a class of kinases with insulated biological networks regulating novel signaling networks important for some aspect specific for the vertebrate physiology. Of note, PDGFR appeared much later in evolution compared with EGFR during vertebrate development. These novel regulatory networks might then have been hijacked by cancer cells during the acquisition of resistance to drug treatment.

PROTEIN TYROSINE PHOSPHATASES AND THE CONTROL OF SIGNAL TRANSDUCTION

N.K. Tonks C. Bonham, Y. Cen, C. Felice, C. Gauss, Q. Hu, Z. Qian, D. Song, P. Venkataramani, S. Zhang, I. Zubiete Franco

As cells encounter stimuli, such as growth factors, cytokines, and hormones, receptors on the cell surface modulate the activities of protein kinases and phosphatases. The functions of these enzymes, which promote the addition and removal of phosphate groups, are coordinated in signal transduction pathways to mediate the cellular response to the environmental stimuli. These pathways are of fundamental importance to control of cell function, and their disruption frequently underlies major human diseases. Consequently, the ability to modulate such signal transduction pathways selectively with drugs holds enormous therapeutic potential. In the area of tyrosine phosphorylation–dependent signal transduction, drug discovery efforts to date have emphasized the protein tyrosine kinases (PTKs). Although there have been spectacular successes, challenges remain, including the acquisition of drug resistance. Considering the reversibility of protein tyrosine phosphorylation, there is the potential to manipulate signal transduction pathways at the level of both PTKs and protein tyrosine phosphatases (PTPs). Although the PTPs have been garnering attention as potential therapeutic targets, they remain largely an untapped resource for drug development. The long-term objectives of the work of the Tonks lab are to characterize the structure, modes of regulation, and physiological function of members of the protein tyrosine phosphatase family of enzymes. Through basic research to understand the mechanism of action and function of PTPs, the Tonks lab is trying to devise creative new approaches to exploiting these enzymes as targets for therapeutic intervention in major human diseases, including diabetes, obesity, and cancer.

During the last year we were joined by Yuxin Cen, a graduate student from the Stony Brook MCB program.

Allosteric Inhibitors of PTP1B

PTP1B is a major regulator of the signaling pathways initiated by insulin, which controls glucose uptake and metabolism, and leptin, which controls appetite. Gene targeting studies demonstrated that PTP1B-null mice are healthy, display enhanced insulin sensitivity, do not develop type 2 diabetes, and are resistant to obesity when fed with a high-fat diet. This generated considerable interest in PTP1B as a therapeutic target for treatment of diabetes and obesity. Major programs in industry focused on developing small-molecule inhibitors of PTP1B, but they have been frustrated by technical challenges arising from the chemical properties of the PTP active site. In particular, although it is possible to generate potent, selective, and reversible active site–directed inhibitors, the tendency for such molecules to be highly charged, such as pTyr-substrate-mimetics, presents problems with respect to their oral bioavailability and limits their drug development potential. As a result, industry views PTP1B, and other PTPs, as challenging. Consequently, innovative strategies are required, such as can be produced in an academic setting, to generate inhibitors of this highly validated target that may be readily exploited for drug development. One focus of the lab continues to be novel approaches to exploiting PTP1B as a therapeutic target in various human diseases.

In addition to its role in down-regulating insulin and leptin-induced signaling, PTP1B plays a positive role in promoting signaling events downstream of the oncoprotein tyrosine kinase HER2; consequently, small-molecule inhibitors of PTP1B may also represent novel therapeutics for treatment of tumorigenesis and malignancy associated with elevated HER2. Previously, we demonstrated that Trodusquemine/MSI-1436 is an allosteric inhibitor of PTP1B, which attenuates HER2-dependent tumorigenesis and abrogates metastasis

in the NDL2 mouse model of breast cancer. This led to us testing MSI-1436 in a Phase 1 clinical trial in metastatic breast cancer patients (ClinicalTrials.gov: NCT02524951), in collaboration with the Northwell Health Monter Cancer Center. In this small trial, involving five patients, we observed that the compound was well tolerated. Our current focus remains on examining the effect of MSI-1436 on tyrosine phosphorylation in HER2-positive breast cancer cells, such as BT474, to define its mechanism of action. Furthermore, a major problem exists with current therapeutic strategies in that 70% of HER2-positive breast cancer patients display de novo resistance to the frontline targeted therapy Herceptin (Trastuzumab), and the remaining 30% that do respond initially acquire resistance in ~2 yr. Consequently, the identification of alternative or combinatorial targets for therapeutic intervention is desperately needed. We are working to define how MSI-1436 may be able to overcome the de novo and acquired resistance to Herceptin that is observed in HER2-positive cancer patients.

An important aspect of our approach is the application of PTP "substrate trapping" mutants, a technology that we developed several years ago. These mutants are capable of forming stable complexes with target substrates in vivo, allowing us to examine direct cellular targets and protein tyrosine phosphorylation events regulated through the action(s) of specific PTPs. We are integrating our "substrate trapping" technology with recently described proximity-dependent biotin identification (BioID) labeling strategies. By coupling with quantitative, high-throughput mass spectrometry–based proteomics and phospho-proteomics, we have established and validated a single workflow capable of identifying direct substrates of PTP1B and their respective sites of tyrosine dephosphorylation. We are applying this approach to cell-based models of HER2-positive breast cancer that are sensitive to Herceptin (BT474) and those that exhibit de novo resistance to the drug (JIMT1) to identify direct substrates and signaling axes influenced by the activities of PTP1B. Our mass spectrometry data have uniquely identified novel, tyrosine-phosphorylated protein targets of PTP1B among the sensitive (54) and resistant (18) cell populations, most notably those with established roles in HER2-positive breast cancer pathophysiology and in therapeutic resistance to Herceptin, respectively. Global pathway analysis of the identified candidates has uncovered targets involved in cell migration, lipid biosynthesis, and metabolic function—commonly dysregulated cellular processes observed in HER2-positive and metastatic breast cancers. We have now validated approximately 20 candidates by immunoblot analysis and are continuing our analysis of changes in tyrosine phosphorylation levels upon treatment with MSI-1436/Trodusquemine.

A New Approach to Copper Chelation

Although MSI-1436 demonstrates efficacy in an injectable format, like many PTP1B inhibitors identified to date it also shows limited oral bioavailability. In 2018, we published two papers describing the identification and characterization of an orally bioavailable analog of MSI-1436 that is a potent and specific inhibitor of PTP1B. This compound, termed DPM-1001, exhibits a unique specificity and high affinity for copper that enhances its potency as an inhibitor of PTP1B. This has focused our attention on the ability of DPM-1001 to function as a copper chelator. Initially, we examined Wilson's disease, in which inactivating mutations in ATP7B, a P-type Cu-ATPase that functions in copper excretion, promote aberrant accumulation of the metal, primarily in liver and brain. We have demonstrated that treatment of cell models, including fibroblasts derived from Wilson's disease patients, eliminated adverse effects associated with copper accumulation. Furthermore, treatment of the toxic milk mouse model of Wilson's disease with DPM-1001 lowered the levels of copper in liver and brain, removing excess copper by excretion in the feces, while ameliorating symptoms associated with the disease (Fig. 1). These data suggest that DPM-1001 may form the basis for a new therapeutic approach to Wilson's disease.

Recently, there have been reports about the potential application of copper suppression as a cancer therapy, including the observation that copper exerts a stimulatory effect on MEK and is required for oncogenic BRAF signaling in melanoma. With this in mind, we have tested DPM-1001 on a panel of triple-negative breast cancer (TNBC) cells and observed that it was cytotoxic at low micromolar concentrations. Most importantly, we have identified an analog, DPM-1003, which has the same chemical composition as DPM-1001, but in which copper chelation is markedly impaired; unlike DPM-1001, DPM-1003 was not cytotoxic in the TNBC cells.

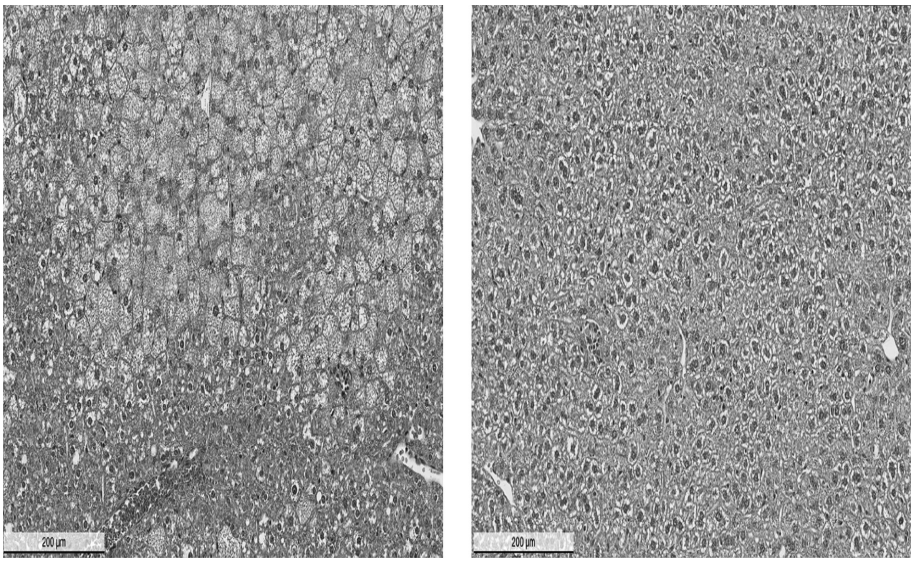


Figure 1. Enlarged size, irregular shape, and arrangement of cells in liver tissue are characteristic features of Wilson's disease (*left*). Treatment with DPM-1001 reversed these pathologies (*right*).

Interestingly, we observed that the levels of copper in these TNBC cells were elevated compared to controls, and the effects of DPM-1001 coincided with suppression of those levels of copper. In contrast, copper levels were unchanged after treatment with DPM-1003. Furthermore, we tested the compounds in tumor xenograft assays in which MDA-MB-231 cells were implanted in mammary fat pads and large, palpable tumors were allowed to form before commencing administration of the compounds. In contrast to treatment with saline or DPM-1003, DPM-1001 abrogated tumor growth. We are extending this analysis into additional cancer cell lines; our data indicate that copper chelation may offer a new approach to treatment of a wide variety of different cancers.

Exploiting Redox Regulation for Therapeutic Development

In addition to developing allosteric inhibitors of PTP1B, we have taken the approach of trying to harness a physiological mechanism for redox regulation of PTP function that reflects a new tier of control of tyrosine phosphorylation–dependent signaling. Previously, we observed that the activity of PTP1B is attenuated by reversible oxidation of an essential cysteinyl residue at the active site of the enzyme. The architecture of the PTP active site is such that this essential cysteinyl residue displays unique properties that favor its role as a nucleophile in catalysis but also render it prone to oxidation. Insulin stimulation of mammalian cells leads to enhanced and localized production of intracellular H_2O_2, which causes reversible oxidation of PTP1B and inhibition of its enzymatic activity; this, in turn, concomitantly promotes the signaling response to insulin. We have shown that mild oxidation of PTP1B, such as occurs in response to insulin, results in profound conformational changes in the active site of the enzyme that transiently inhibit substrate binding and catalysis. These structural changes are reversible and the enzyme can be reduced back to its active state. Therefore, reversible oxidation of PTP1B in response to insulin provides a mechanism for fine-tuning the signaling response to the hormone. Previously, we used phage display to identify conformation-sensor antibodies, such as scFv45, that recognize the reversibly oxidized form of PTP1B (PTP1B-OX) selectively and stabilize this inactive state, inhibiting its reactivation by reducing agent and thereby inhibiting phosphatase activity. We demonstrated that expression of these antibodies in cells enhanced insulin-induced signal transduction. These data provide proof of concept that stabilization of the oxidized, inactive form of PTP1B with appropriate therapeutic molecules may offer a novel

paradigm for phosphatase drug development. We developed this concept further in a paper that was published in 2018.

Initially, we focused on defining the molecular basis for recognition of PTP1B-OX by the conformation-sensor antibody scFv45. A mass spectrometry analysis of the ability of scFv45 to protect PTP1B-OX from proteases highlighted the importance of a loop composed of residues 36-46. Furthermore, investigation of the specificity of scFv45 for PTP1B-OX, compared to its closest relative TCPTP, demonstrated the importance of basic residues within this loop. We generated mutants in which we exchanged residues between PTP1B and TCPTP. Whereas scFv45 did not bind TCPTP-OX, it did bind TCPTP-mut2 (TCPTP-mut2 F39L/E41K/R43K); introduction of these three residues from PTP1B into TCPTP was sufficient to promote binding between scFv45 and TCPTP-OX. Conversely, binding between scFv45 and PTP1B-mut2 (PTP1B-mut2 L37F/K39E/K41R), which contains those three residues from TCPTP, was abrogated. Finally, in collaboration with Leemor Joshua-Tor's lab, we generated a crystal structure of the conformation-sensor antibody scFv45, and modeling of the complex revealed an important role for acidic residues in the complementarity-determining regions of scFv45 and these same basic residues in PTP1B at the binding interface.

This analysis laid the foundation for examining whether small molecules could mimic the effects of the PTP1B-OX-directed conformation-sensor antibodies. We developed an assay that allowed us to screen a library of compounds for small molecules that would stabilize the oxidized form of PTP1B by inhibiting its reduction and reactivation. This pilot screen revealed two such hits, which displayed selectivity for PTP1B-OX over the reduced form of the enzyme and did not affect TCPTP. Testing analogs, we identified chelerythrine, which displays improved potency for PTP1B-OX, while also not affecting TCPTP. Using isothermal titration calorimetry (ITC), we demonstrated target engagement through *direct* binding of chelerythrine to PTP1B-CASA, a mutant that adopts the reversibly oxidized conformation of PTP1B in a stable manner. The data indicated a binding constant that was consistent with the IC_{50} for stabilization of PTP1B-OX. In contrast, we did not detect saturable binding to the mut2 form of PTP1B-CASA, in which the basic residues important for scFv45 binding have been disrupted, suggesting an overlapping binding site for scFv45 and chelerythrine. Treatment of HEK293T and hepatic stellate cells with chelerythrine led to enhanced insulin and leptin signaling, respectively, and this effect was abrogated by expression of catalase, or NOX inhibitors, which suppress H_2O_2 levels. Furthermore, treatment of high-fat diet (HFD)-fed C57Bl6/J mice with chelerythrine resulted in weight loss, whereas there was no effect of treatment with saline or protopine, an inactive analog of chelerythrine. We observed a 3% decrease in body weight in HFD-fed mice treated with chelerythrine, which plateaued after 14 d; however, there was no effect in mice fed a normal chow diet. Treatment with chelerythrine also improved glucose tolerance and insulin sensitivity, coincident with enhanced phosphorylation of the insulin receptor β subunit and AKT in liver. In addition, chelerythrine led to enhanced leptin signaling in the hypothalamus. Overall, these data are consistent with a mechanism in which chelerythrine, like scFv45, stabilizes PTP1B-OX to enhance insulin and leptin signaling. This work illustrates a novel paradigm for inhibiting the signaling function of PTP1B, which focuses on the critical pool of PTP1B that is responsible for regulation of insulin and leptin signaling—the pool of the enzyme that is acutely regulated following insulin and leptin stimulation. By focusing on this pool of PTP1B, the approach may minimize the potential for complications arising from inhibition of the native enzyme as a whole. Overall, our data suggest that this approach may represent the basis for a novel strategy for therapeutic intervention in diabetes and obesity.

Developing Inhibitors of the Protein Kinases PIM and DYRK

In a long-standing collaboration with the laboratories of Darryl Pappin and Leemor Joshua-Tor, we have purified and characterized CSH-4044, a small-molecule natural product that we isolated from fermented wheat germ extract. We demonstrated that CSH-4044 has a unique structure and a unique specificity for PIM and DYRK family kinases. PIMs and DYRKs have been implicated in a wide variety of hematological and epithelial tumors, with the expectation that inhibitors of these kinases may have broad therapeutic utility. In collaboration with Elad Elkayam and Leemor Joshua-Tor, the crystal structure of the kinase PIM-1

was determined in a complex with CSH-4044 at 1.95-Å resolution. CSH-4044 binds to the ATP binding site of PIM-1 in two alternate conformations because of the symmetrical nature of the inhibitor. We identified several specific interactions between the inhibitor and the protein that would help to guide the synthesis of analogs of the compound. A formal collaboration was established with Vichem Chemie to support a medicinal chemistry program required for optimizing CSH-4044. A synthetic route was established and approximately 180 analogs were produced. We generated inhibitors that were ATP-competitive and displayed both improved potency and selectivity relative to CSH-4044. We identified compounds that inhibited DYRK preferentially, inhibited PIM preferentially, or inhibited both. Specificity was confirmed by profiling against a panel of 140 kinases and predictive ADME analysis confirmed drug-like properties.

Now, we are focusing on the DYRKs as our targets because we feel they offer the best opportunities for therapeutic development. The DYRKs belong to the CMGC family of protein kinases, which includes cyclin-dependent kinases, MAP kinases, and glycogen synthase kinases. Newly translated DYRKs undergo intramolecular autophosphorylation on a single Tyr residue in the activation loop, but the mature proteins recognize Ser and Thr residues in their target substrates. There are five DYRKs (1A, 1B, 2, 3, and 4); although 1A and 1B feature nuclear localization signals, all of these enzymes show both nuclear and cytoplasmic localization that vary according to conditions. These are exciting enzymes that have been implicated in the etiology of several major diseases. The *DYRK1A* gene is probably the best characterized because it is located at the Down syndrome critical locus, is overexpressed in trisomy 21, and has been implicated in the neurodegeneration and cancer susceptibility of Down syndrome patients. It has been implicated as a therapeutic target in many cancers, including EGF-dependent glioblastoma. DYRK1B is also overexpressed in several cancers. Importantly, there are currently no DYRK-directed drugs; although inhibitory compounds, such as INDY and the natural product Harmine, have received attention, they display off-target effects that negate any therapeutic potential. Consequently, this provides an opportunity for exploiting our DYRK-directed inhibitors as cancer therapeutics.

We are developing further two of the inhibitors from the SAR program we conducted with Vichem (IC50 20-50nM). Although it was a challenging target, Elad determined the crystal structure of DYRK1A in a complex with one of these at 2.4-Å resolution. By combining sequence analysis, structural insights and biochemical assays, we have generated point mutants of DYRK1A in which catalytic function was maintained, but in which affinity for the inhibitors was markedly attenuated—from nanomolar to micromolar. Currently, we are expressing these mutants in cancer cells in order to establish that the effects of our small-molecule drug candidates are due to "on-target" inhibition of DYRK, rather than "off-target" effects. We have demonstrated that our compounds were inhibitory to growth of glioblastoma cell lines in culture. Furthermore, they inhibited neurosphere formation by U87MG glioblastoma cells, showing improved effects compared to INDY, an established inhibitor of DYRK. In addition, we are collaborating with Chris Vakoc, who used a kinase domain–focused CRISPR screen to examine dependencies in AML cell models. Interestingly, he identified two AML cells, HEL and SET-2, which displayed a dependency on DYRK1A. We have now demonstrated that these cells were sensitive to our DYRK1A inhibitors; in contrast, MOLM13 and K562 cells, which were not dependent on DYRK1A, were approximately fivefold less sensitive to the inhibitor. Going forward, we have established a collaboration between our team at CSHL and Dr. Yousef Al-Abed and his group at the Feinstein Institute. We plan to optimize our compounds further to both improve bioavailability and improve potency and selectivity. In addition, we will examine the possibility of exploiting unique structural features of the DYRKs for generation of new inhibitor classes.

PUBLICATIONS

Krishnan N, Konidaris KF, Gasser G, Tonks NK. 2018. A potent, selective and orally bioavailable inhibitor of the protein tyrosine phosphatase PTP1B improves insulin and leptin signaling in animal models. *J Biol Chem* **293:** 1517–1525.

Krishnan N, Bonham CA, Rus IA, Shrestha OK, Gauss CM, Haque A, Tocilj A, Joshua-Tor L, Tonks NK. 2018. Harnessing insulin- and leptin-induced oxidation of PTP1B for therapeutic development. *Nat Commun* **9:** 283.

Krishnan N, Felice C, Rivera K, Pappin DJ, Tonks NK. 2018. DPM-1001 decreased copper levels and ameliorated deficits in a mouse model of Wilsons' disease. *Genes Dev* **32:** 944–952.

Zhang J, Wang Z, Zhang S, Chen Y, Xiong X, Li X, Tonks NK, Fan G. 2018. Spatial regulation of signaling by the coordinated action of the protein tyrosine kinases MET and FER. *Cell Signalling* **50:** 100–110.

MECHANISMS AND MODELING OF METASTASIS

L. Trotman A. Ambrico I. Casanova-Salas H. Cox A. Jhaveri M. Lee G. Mathew D. Nowak M. Swamynathan M. Walsh K. Watrud

Resistance to therapy of metastatic prostate cancer (CRPC) is responsible for the annual deaths of some 30,000 U.S. men. Although there is considerable progress in development of improved antihormone therapy for treatment of metastatic disease, this standard-of-care approach will invariably fail at some point.

Our focus is to understand the mechanisms driving human prostate cancer in its most lethal form—metastatic disease. We have studied the human genetics behind the transition from indolent to lethal metastatic prostate cancer and combined it with viral transgene delivery into prostate. With this approach, we have succeeded in generating a unique, fast, and faithful mouse model for metastatic prostate cancer. We have termed this system RapidCaP, as it allows us to generate any genetically mutant mouse prostate cancer with a very accelerated time frame compared to breeding-based approaches. Now, we use RapidCaP for analysis and therapy of metastatic disease, and we use human genomics analysis for discovery of new candidate drivers of metastasis.

At the same time, we aim to better understand how the PTEN tumor suppressor works. This has given us unique insights into how the process of endocytosis is intimately associated with the tumor suppressor function of PTEN, allowing us to redefine this pathway.

Genomics of Lethal Human Prostate Cancer

G. Mathew, H. Cox, I. Casanova-Salas, A. Jhaveri, K. Watrud, A. Ambrico [in collaboration with S. Hall, M. Vira, C. Metz, O. Yaskiv, O. Rodriguez, and M. Ryan, Northwell Health; L. Kollath and C. Morrissey, University of Washington, Seattle]

Next-generation sequencing techniques have provided the ability to incorporate cutting-edge genomic profiling in understanding, prognosis, and treatment of various tumor types. However, heterogeneity of the prostate tissue during tumorigenesis makes it difficult to conduct exhaustive transcriptome analyses. Spontaneous genetic changes arising in prostate cancer (PC) are a crucial imprint of this variability. Therefore, one of our aims is to shed light on key drivers for metastatic progression utilizing genomic information from prostate cancer patients. The project goal was to obtain an extensive copy number landscape of visceral and bone metastases from 10 patients who had consented to subject their bodies to rapid autopsy (RAP) after death from PC. Samples were obtained from the University of Washington, Seattle, and Northwell Health, New York. We have successfully established a pipeline for processing of frozen tissue samples for single-nucleus sequencing to determine copy number alterations (CNAs) at the single-cell level. Recurrent CNA involving cancer genes has emerged as the primary driver of lethal metastatic PC, whereas recurrent missense mutations are infrequent. After processing all bone metastases, we analyzed matched visceral metastases (liver, lung, lymph node, etc.) from each of the 10 patients. A normal muscle sample from each patient was obtained and processed as a baseline control for CNA.

Based on the data collected from 2914 cells in 31 metastatic sites of 10 patients, we first answer these general questions on the metastatic landscape seen at single-cell resolution.

1. What is the CNA-based clonality of metastasis (within tumor site and between sites)?
2. Can we infer fitness of clones based on representation (within a site and/or between sites of a patient)?
3. Are there recurrent CNAs that have been missed by bulk sequencing of metastatic PC?

Given our expertise and the emergence of PTEN deletion as the most prominent feature of lethal metastatic PC, we place special emphasis on the below questions.

1. What genes are most significantly co-deleted with PTEN at the single-cell level?
2. Does loss of PTEN dominate clonality as expected from a strong driver event?

These data are complemented by our analysis of tumors from primary PC patients using the same approach. Samples from these early patients are collected through our collaboration with clinician scientists at Northwell Health. Analysis of genome-wide DNA and RNA alteration in primary and metastatic RapidCaP samples is used for cross-species prioritization of results. Based on our preliminary results, this project allows us to discover novel markers of metastasis and new drivers of the lethal disease that have escaped our notice based on bulk sequencing analysis.

Blood-Based Monitoring of Prostate Cancer

I. Casanova-Salas, G. Mathew, H. Cox, A. Jhaveri, M. Swamynathan [in collaboration with S. Hall, M. Vira, C. Metz, O. Yaskiv, O. Rodriguez, and M. Ryan, Northwell Health; L. Kollath and C. Morrissey, University of Washington, Seattle]

We have developed a method to infer CNAs from exosomal DNA (exoDNA) in patient blood. Because CNAs are the major driver behind metastatic PC, we use this low-cost approach to discover new biology that can be translated into tests that will help patients. First, we find that the whole genome is represented in exoDNA, whereas overt CNA events are mostly detectable only in metastatic patients. These results are consistent with the notion that exoDNA from normal cells far outnumbers that from a tumor site until there is overt metastasis. Therefore, we explore other parameters in these samples to define biomarkers and new biology that allows us to get insights into the nature of the tumor in a patient. Based on our preliminary results, our approach can lead to tests that are directly applicable to validation studies in patients who are under active surveillance for signs of PC progression.

The PHLPP2 Phosphatase Protects MYC and Is a Target for Prevention of Prostate Cancer Progression

D.G. Nowak, K. Watrud, I. Casanova-Salas, A. Ambrico, M. Swamynathan [in collaboration with C-H. Pan, CSHL; K.C. Katsenelson and A.C. Newton, University of California, San Diego; J.E. Wilkinson, University of Michigan, Ann Arbor]

Metastatic PC commonly presents with targeted, biallelic mutation of PTEN and TP53 tumor suppressor genes. In contrast, however, most candidate tumor suppressors are part of large recurrent hemizygous deletions, such as the common chromosome 16q deletion, which involves the AKT-suppressing phosphatase PHLPP2.

Using RapidCaP, a genetically engineered mouse model of Pten-Trp53-mutant metastatic PC, we found that complete loss of *Phlpp2* effectively blocks prostate tumor growth and progression to otherwise lethal metastasis. We show that Phlpp2 activates Myc, a key driver of PC and metastasis. Mechanistically, Phlpp2 dephosphorylates the Thr-58 site of Myc, thus directly increasing MYC stability. Finally, we show that small-molecule inhibitors of PHLPP2 can suppress MYC and cause cell death. Our findings reveal how PTEN-deficient tumors can thrive in the absence of AKT activation, driven by PHLPP2 stabilization of MYC. They also suggest that the frequent hemizygous deletions on chromosome 16q present a druggable vulnerability for targeting the MYC protein through PHLPP2 phosphatase inhibitors.

Selective Killing of PTEN-Deficient Cancer Cells

G. Mathew, K. Watrud, A. Ambrico, M.F. Lee [in collaboration with D. Pappin, A. Naguib, N. El-Amine, W. Zheng, T. Herzka, I.C. Salas, CSHL; E. Di Francesco and J. Marszalek, MD Anderson Cancer Center; C.R. Reczek and N.S. Chandel, Northwestern Medical School]

A hallmark of advanced PC is the concomitant loss of PTEN and p53 function. Although there is progress in the development of improved drugs for treatment of metastatic disease, the vast majority of patients are in one type of treatment approach: antihormone therapy. This approach goes back to discoveries made in the 1940s and has improved the condition of most men with the disease. However, it is not a cure: Emergence of androgen resistance is invariably observed, even after 70 years of improvements. We are interested in searching for alternatives that specifically target the PTEN-deficient cancer cells.

The genes for p53 and PTEN are most frequently inactivated in lethal PC. Their diminished activity provides a growth, proliferative, and anti-apoptotic advantage. In mouse PC models, loss of p53 does not result in malignancy, yet co-deletion of *Pten* and *Trp53* genes causes prostatic adenocarcinoma, illustrating the need to combat the cooperative power of these combined genetic lesions. To selectively eliminate

such cells, we screened cytotoxic compounds on $Pten^{-/-};Trp53^{-/-}$ fibroblasts and their Pten-wild type (wt) reference. Highly selective killing of Pten-null cells can be achieved by deguelin, a natural insecticide. Deguelin eliminates Pten-deficient cells through inhibition of mitochondrial complex I (CI). Five-hundred-fold-higher drug doses are needed to obtain the same killing of Pten-wt cells, even though deguelin blocks their electron transport chain equally well. Selectivity arises because mitochondria of Pten-null cells consume ATP through complex V, instead of producing it. The resulting glucose dependency can be exploited to selectively kill Pten-null cells with clinically relevant CI inhibitors, especially if they are lipophilic. In vivo, deguelin suppressed disease in our genetically engineered mouse model for metastatic PC. Our data thus introduce a vulnerability of highly selective targeting of incurable PC with inhibitors of CI. Now, through collaboration with clinicians at Northwell Health and in the United Kingdom, we are testing whether the popular antidiabetic drug metformin shows disease-slowing effects in PC patients who routinely take it.

PUBLICATIONS

Albrengues J, Shields MA, Ng D, Park CG, Ambrico A, Poindexter ME, Upadhyay P, Uyeminami DL, Pommier A, Küttner V, et al. 2018. Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. *Science* **361:** eaao4227. doi:10.1126/science.aao4227

Alexander J, Kendall J, McIndoo J, Rodgers L, Aboukhalil R, Levy D, Stepansky A, Sun G, Chobardjiev L, Riggs M, et al. 2018. Utility of single-cell genomics in diagnostic evaluation of prostate cancer. *Cancer Res* **78:** 348–358. doi:10.1158/0008-5472.CAN-17-1138

Naguib A, Mathew G, Reczek CR, Watrud K, Ambrico A, Herzka T, Salas IC, Lee MF, El-Amine N, Zheng W, et al. 2018. Mitochondrial complex I inhibitors expose a vulnerability for selective killing of Pten-null cells. *Cell Rep* **23:** 58–67. doi:10.1016/j.celrep.2018.03.032

CANCER MEDICINE

D. Tuveson

L. Afinowicz, B. Alagesan, P. Alagesan, L. Baker, G. Biffi, E. Brosnan, D. Cheng, A. Deschenes, E. Elyada, D. Engle, F. Froeling, L. Guerriero, A. Handly-Santana, C-I. Hwang, M. Kheir Gouda, R. Koike, E. Lee, M. Lucito, M. Ma, K. Miyabayashi, S. Moran, T. Oni, Y. Park, H. Patel, D. Plenker, C. Schoepfer, B. Spielman, J. Thalappillil, H. Tiriac, C. Tonelli, M. Yao, G. Yordanov, K. Yu

Our laboratory investigates pancreatic ductal adenocarcinoma (PDAC), the primary form of pancreatic cancer and the third leading cause of cancer-related deaths in the United States. More specifically, we aim to generate insights into the molecular underpinnings of PDAC, which could inform novel strategies to detect and treat this currently incurable cancer. As part of this goal, we are committed to developing tractable model systems that accurately mimic PDAC disease biology. We have more than 20 years of experience developing murine models of PDAC. More recently, we have developed three-dimensional organoid cultures as ex vivo models of PDAC biology. Organoids have enabled us to gain new insights into the factors driving PDAC development and ways to improve patient care. By examining genes up- or down-regulated in tumor organoids derived from our mouse models of PDAC compared to organoids derived from normal, healthy murine pancreata, we have identified candidate oncogenes and tumor suppressors, respectively, which we are now following up on. Using an organoid-fibroblast co-culture system, which we developed, we have begun to disentangle the signal cascades that mediate patterning of the PDAC microenvironment (Biffi et al. 2019). We have also begun to identify other cell types present in the PDAC microenvironment and are working to incorporate those into our co-culture models. Finally, we have begun a deep analysis of the human patient-derived PDAC organoid collection that we have generated. Through this analysis, we have identified signatures of genes whose expressions predict patient response to chemotherapy agents (Tiriac et al. 2018). We are now developing clinical trials to refine those signatures and evaluate whether they can be used prospectively to stratify each patient into the most optimal treatment.

Delineating Signaling Cascades that Pattern PDAC Stroma

This work was done in collaboration with J. Preall, A. Dobin, and M. Hammell (CSHL) and P. Robson (The Jackson Laboratory).

PDAC tumors typically contain a dense desmoplastic stroma that includes numerous noncancer cell types such as cancer-associated fibroblasts (CAFs). The tumor stroma was once hypothesized to perform a solely tumor-supportive role. However, recent research from our lab and others challenges that notion. We have identified two distinct types of CAFs present in the PDAC stroma: CAFs that express myofibroblast markers, which we termed myofibroblastic cancer-associated fibroblasts (myCAFs), and CAFs that express inflammatory cytokines, which we termed inflammatory cancer-associated fibroblasts (iCAFs). We observed that myCAFs tend to be located in close proximity to the cancer cells, whereas iCAFs tend to be more distally located.

We sought to better understand signaling pathways involved in patterning these distinct cell types. Because transforming growth factor beta (TGF-β) signaling was previously shown to promote a myofibroblastic phenotype, we hypothesized that TGF-β signaling might promote myCAF formation. In support, we identified TGF-β target genes up-regulated in the myCAFs relative to the iCAFs. Furthermore, higher levels of the active, phosphorylated SMAD2 and SMAD3, key effectors of TGF-β signaling, were observed in myCAFs. Notably, when CAFs were cultured in vitro, addition of TGF-β to culture medium promoted the myCAF phenotype while antagonizing the iCAF phenotype, suggesting that the signaling pathways needed for iCAF and myCAF formation help to enforce either cell state.

Next, we turned our attention to the signaling pathways required in iCAFs. We hypothesized that NF-κB signaling might play a role in iCAF formation because this pathway was previously shown to induce an inflammatory profile in CAFs. We found that cancer cells sorted from our mouse model of PDAC and cancer organoids derived from this model produce the NF-κB signaling ligands IL1α and TNFα. In contrast, CAFs produce the corresponding receptors for these ligands. Moreover, we have uncovered an IL1α-driven signaling cascade, involving activation of LIF and JAK/STAT signaling, responsible for up-regulating inflammatory genes in iCAFs.

Because iCAFs secrete inflammatory cytokines, such as IL6, they likely serve to promote tumor growth through paracrine interactions. On the other hand, myCAFs are responsible for deposition of the desmoplastic stroma that prevents therapies from effectively being delivered to PDAC cells. Accordingly, when we treated our mouse model of PDAC with an inhibitor of JAK signaling, we observed depletion of the iCAF population, coupled with increased desmoplasia. The results of this study were published in October in *Cancer Discovery* (Biffi et al. 2019). We are now researching whether therapeutic combinations of JAK inhibition with agents that would target desmoplasia production would more effectively treat PDAC tumors. In addition, we are using single-cell RNA-sequencing approaches to catalog all of the cells present in the PDAC microenvironment. Excitingly, we have identified a third subtype of CAFs, which express major histocompatibility complex class II molecules, normally found only on professional antigen-presenting cells. We are now working to better understand the role of these cells in the PDAC stroma. In the future, we will also better examine the other cells present in the PDAC microenvironment.

Defining Molecular Signatures of Chemotherapy Resistance

This work was done in collaboration with A. Krasnitz and C. Vakoc (CSHL); S. Gallinger (Toronto); R. Moffitt, E. Li, and J.M. Buscaglia (Stony Brook); W. Nealon (Northwell); L. Martello-Rooney (SUNY Downstate); C.L. Wolfgang (Johns Hopkins); and B. Wolpin (Dana-Farber).

Although our genetically engineered mouse model has been a source of many significant breakthroughs, this model is not able to mimic the significant inter- and intratumoral genetic heterogeneity found in human patients. Therefore, we previously developed human organoid cultures as ex vivo models of human PDAC. We have now amassed a collection of more than 200 human patient-derived PDAC organoids (PDOs), including PDOs generated from the limiting material present in a fine-needle biopsy (Fig. 1A). To better characterize our collection, we performed whole-exome sequencing on 65 PDOs. In addition, we performed whole-genome sequencing (WGS) on a subset of PDOs and their matched primary tumor ($n = 13$). In 11 of the 13 cases, the PDO cultures completely recapitulated the PDAC mutational profile found in the primary specimen, although mutations were sometimes difficult to call in primary tumor specimens in which cancer cells are outnumbered by noncancerous stromal cells (Fig. 1B). RNA sequencing (RNA-Seq) was performed on 44 PDOs and 11 organoids derived from normal, noncancerous tissue. Gene set enrichment analysis (GSEA) of the differentially expressed genes in human PDAC relative to normal organoids indicated pathways dysregulated in pancreatic cancer cells, including MYC and E2F targets, the G2-M checkpoint, and pathways involved in metabolism. Previous studies have established that there are two main subtypes of PDAC tumors: a classical subtype and a basal-like subtype. Based on analysis of our RNA-Seq data, 70% of the PDO cultures are the classical subtype (31/44) and 30% are basal-like (13/44)—a notable finding as there are very few available cell line models of the classical PDAC subtype.

To better understand the therapeutic sensitivity profile of each PDO, we developed a method for performing high-throughput, therapeutic screens on PDOs. In our assay, known numbers of cells from PDOs are plated under organoid conditions in 384-well plates and allowed to re-form organoids. Organoids are then exposed to varying concentrations of chemotherapeutic agents or targeted therapies, and viability is assessed 3–5 days later. We can use these data to generate therapeutic response curves and assess the response of each organoid to a given therapy. Using this assay, which we call "pharmacotyping," we observed a striking degree of heterogeneity in the response of our PDOs to standard-of-care chemotherapies (Fig. 1C,D). By comparing pharmacotyping data with RNA-Seq data, we identified genes whose expression levels in PDOs best correlate with chemotherapy response. Moreover, we were able to show that our gene signatures were able to predict patient response in two different cohorts of patients (Fig. 1E). We expect these gene signatures will be a valuable clinical tool, allowing oncologists to

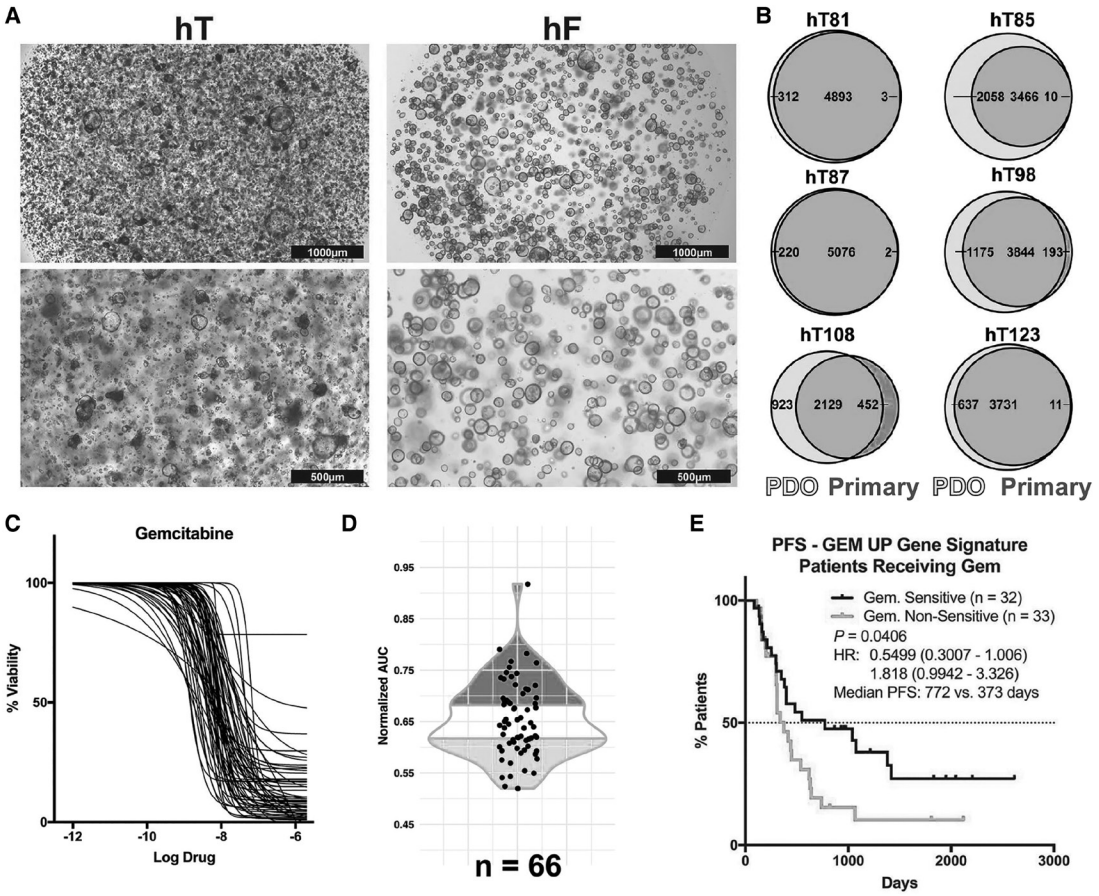


Figure 1. Analysis of human patient-derived organoids (PDOs) reveals gene signatures of chemotherapy response. (*A*) Bright-field microscopy images of organoid cultures generated from human PDAC tumors (hT) or fine-needle biopsies of human PDAC lesions (hF). Scale bars, 1000 or 500 µm, as indicated. (*B*) Representative Venn diagrams showing the concordance of somatic single-nucleotide variants found in primary tumor specimens and matched PDO cultures following germline variant removal. (*C*) Dose–response curves of individual PDO cultures treated with gemcitabine (Gem.). Drug concentration is Log-transformed mol/L. (*D*) Normalized area-under-the-curve (AUC) values showing the sensitivity of each PDO to Gem., calculated from the dose–response curves in *C*. High AUC values (the dark gray portion of the plot) indicate the most Gem.-resistant organoids, whereas low AUC values (the light gray portion of the plot) indicate the most Gem.-sensitive organoids. (*E*) Kaplan–Meier analysis of progression-free survival (PFS) in patients who received single-agent Gem. and were predicted to be either "Gem. sensitive" or "Gem. nonsensitive" based on the application of our gemcitabine-sensitivity prediction signature to RNA-sequencing data from each patient's primary tumor. Patients were part of the International Cancer Genome Consortium-Canada cohort.

identify which therapies a given patient is most likely to respond to. We published the results of our findings in Tiriac et al. (2018). We are now working to use bioinformatic methods to better refine our gene signatures and applying similar methods to identify signatures of response to targeted therapies. We have also forged numerous clinical collaborations, which will allow us to retrospectively test whether our signatures are predicative of patient response in additional cohorts of patients. Finally, we have begun to set up a clinical trial to test the effectiveness of our signatures.

PUBLICATIONS

Tiriac H, Belleau P, Engle DD, Plenker D, Deschenes A, Somerville TDD, Froeling FEM, Burkhart RA, Denroche RE, Jang GH, et al. 2018. Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. *Cancer Discov* **8:** 1112–1129.

In Press

Biffi G, Oni TE, Spielman B, Hao Y, Elyada E, Park Y, Preall J, Tuveson DA. 2019. IL1-induced JAK/STAT signaling is antagonized by TGFβ to shape CAF heterogeneity in pancreatic ductal adenocarcinoma. *Cancer Discov* **9:** 282–301.

Ras AND Rho GTPases AND THE CONTROL OF SIGNAL TRANSDUCTION

L. Van Aelst F. Boato J. He
J. Freedman Y. Tai
N.B. Gallo M. Wang

Research in our laboratory is focused on the role of Ras and Rho GTPase family members in signal transduction. Ras and Rho family members play key roles in cellular activities controlling cell growth, differentiation, and morphogenesis. Alterations in Ras and Rho functions have been implicated in cancer as well as brain/mental disorders; the latter include intellectual disability (ID), autism, schizophrenia, epilepsy, and mood disorders. Our interests lie in understanding how defects in Ras- and Rho-linked proteins contribute to the development of these diseases/disorders. Toward this end, my lab has continued to define the functions of selected GTPases, their regulators, and effectors, using animal models of cancer and neurodevelopmental/neurological disorders. Below are highlighted our key projects.

Multifunctional Role of the X-Linked Intellectual Disability (XLID) Protein Oligophrenin-1 in Cognition and Behavior

Oligophrenin-1 (*OPHN1*), which encodes a Rho-GTPase activating protein, was the first identified Rho-linked ID gene. It was initially identified by the analysis of a balanced translocation t(X;12) observed in a female patient with mild ID. Subsequent studies revealed the presence of *OPHN1* mutations in families with a syndromic form of ID, with affected individuals commonly showing mild to severe ID and behavioral problems. Most of the *OPHN1* mutations identified to date were shown, or predicted, to result in loss of OPHN1 function; however, the pathophysiological role of *OPHN1* has remained poorly defined. To gain insight into how mutations in *OPHN1* could lead to the cognitive deficits, we initially focused on defining OPHN1's role in hippocampal plasticity associated with learning and memory. Our studies unveiled multiple roles for OPHN1 at hippocampal CA1 synapses. In addition to an essential role in controlling activity-driven glutamatergic synapse development, we found that temporal regulation of OPHN1 translation plays a critical role in mGluR-dependent long-term depression (LTD), a form of plasticity linked to drug addiction and cognitive disorders that is dependent on rapid new protein synthesis. We subsequently showed that OPHN1's involvement in mGluR-LTD and its ability to control glutamatergic synapse development are independent of each other. Whereas the latter requires OPHN1's Rho-GAP activity and association with Homer 1b/c proteins, the former is dependent on OPHN1's interaction with Endophilin A2/3. Together, our findings provided first insights into how mutations in *OPHN1* could contribute to the cognitive deficits in individuals with *OPHN1* mutations.

Apart from the above-described role for OPHN1 in hippocampal plasticity and learning, we more recently discovered that it also plays a critical role in moderating stress-induced depressive-like behaviors in a learned helplessness (LH) model of depression. In this model, animals are exposed to an unpredictable and uncontrollable stressor and subsequently evaluated for their coping capabilities (i.e., resilient vs. helpless/depression-like behavior) in a testing session that consists of an aversive but escapable stressor. We found that *Ophn1*-deficient mice ($Ophn1^{flox/Y;}$ β-$actin^{Cre-/+}$) subjected to the LH procedure showed a marked increase in helpless behavior compared to control littermates. Importantly, *Ophn1*-deficient mice did not display altered motor activity and were capable of learning and performing the avoidance task. Dependent on further examination of the brain regions and neuronal cell types involved, we remarkably found that *Ophn1* deficiency exclusively in parvalbumin (PV) interneurons in the prelimbic area of the medial prefrontal cortex (PL-mPFC) is sufficient to induce helpless behavior. We then investigated how loss of OPHN1 affects the function of PL-mPFC

PV interneurons, leading to maladaptive behavioral responses to stress. Interestingly, we uncovered that *Ophn1* deficiency dampens the excitatory synapses onto mPFC PV interneurons, leading to less active PV interneurons and consequently increased activity in neighboring pyramidal neurons (PyNs). Together, these findings uncovered the importance of the XLID protein OPHN1 in the establishment of resilience to stress and its involvement in the regulation of mPFC PV interneuron function.

Axo-Axonic Innervation of Neocortical PyNs by GABAergic Chandelier Cells Requires AnkyrinG-Associated L1CAM

Proper assembly and functioning of cortical circuits relies on the formation of specific synaptic connections between excitatory PyNs and different types of GABAergic interneurons. Among the various cortical interneuron subtypes, chandelier cells (ChCs), in particular, have a powerful influence over the output of excitatory PyNs because of their unique morphology and the type of connections they make. Specifically, ChCs possess a very distinctive axonal arbor with multiple arrays of short vertical sets of cartridges, each harboring a string of synaptic boutons. This unique architecture enables a single ChC to couple to a large population of PyNs. Furthermore, ChC cartridges make exclusive contact with the axon initial segment (AIS) of PyNs, which is the most excitable part of a neuron in which action potentials are initiated. Importantly, aberrant ChC/PyN AIS innervation has been reported in several disease states associated with altered cortical excitability, including schizophrenia, epilepsy, and autism spectrum disorder. Despite the importance of ChCs, very little is known about the mechanisms governing ChC structure and connectivity. To date, the only molecules implicated in neocortical ChC morphogenesis are the atypical Rac activator DOCK7 and receptor tyrosine kinase ErbB4. In particular, we uncovered that silencing of DOCK7 in ChCs via a novel vMGE-directed in utero electroporation (IUE) approach markedly decreases ChC cartridge bouton size and density, and intriguingly does so by modulating the activity of ErbB4. Notably, though, we found that DOCK7- and ErbB4-depleted ChCs still make contact with PyN AISs, indicating that other molecules must regulate ChC/PyN axo-axonic innervation.

To identify the molecular factors required for neocortical ChC/PyN AIS innervation, we initiated an in vivo RNA interference (RNAi) screen of PyN-expressed axonal cell adhesion molecules (CAMs) and select Ephs/ephrins. More specifically, we devised a strategy taking advantage of IUE and RNAi to individually knock down these molecules in neocortical PyNs while concurrently labeling ChCs using the recently generated *Nkx2.1-CreER* mouse line, which enables tamoxifen-dependent ChC red fluorescent protein labeling via Nkx2.1-driven CreER expression. Strikingly, of all the candidates tested, we found the panaxonally expressed CAM L1CAM to be the only molecule required for neocortical ChC/PyN AIS innervation as knockdown of PyN L1CAM, but none of the other screened candidates significantly reduced PyN AIS innervation by ChCs. In line with this, we observed the number of VGAT and gephyrin puncta at the AIS, but not along the somatodendritic compartment, to be concomitantly reduced in L1CAM-depleted PyNs, indicating that PyN L1CAM selectively regulates ChC/PyN AIS synaptic innervation and not the subcellular targeting of other PyN subcellular domains by other interneuron subtypes. Furthermore, we showed that L1CAM is required during both the establishment and maintenance of neocortical ChC/PyN AIS innervation. Finally, we provided evidence that anchoring of L1CAM at the AIS by the ankyrin-G/βIV-spectrin AIS cytoskeletal complex is essential for ChC subcellular innervation of PyN AISs. Taken together, our findings identify L1CAM as the only PyN-expressed CAM known to date to regulate axo-axonic innervation of PyNs by ChCs in the neocortex.

TGF-β/Smad Signaling through DOCK4 Facilitates Lung Adenocarcinoma Metastasis

The cytokine transforming growth factor beta (TGF-β) plays an important, albeit complex, role in epithelial tumorigenesis. During early stages of tumorigenesis, TGF-β typically functions as a tumor suppressor. At later stages, however, it can act as a potent promoter of multiple events driving the metastatic process, which comprises local motility/invasion, entry of cancer cells into the blood stream (intravasation), exit from the blood vessels (extravasation), and colonization of distant organs. The relevance of TGF-β signaling for

disease progression has been particularly recognized in tumors in which cancer cells retain the core TGF-β signaling components, as is frequently the case in breast and lung cancers. However, a major remaining challenge is the identification of TGF-β target genes that drive specific events during metastasis, especially because TGF-β modulates gene expression in a highly cell- and context-specific manner. Although some progress has been made in the context of breast cancer metastasis, the genes and mechanisms that mediate the prometastatic effects of TGF-β in lung adenocarcinoma (ADC) remain largely unknown.

To identify molecular mechanisms that mediate the prometastatic effects of TGF-β in lung ADC, we took a candidate gene approach and started by scrutinizing members of the DOCK180-related protein superfamily, which emerged as a distinct class of Rac and/or Cdc42 guanine nucleotide exchange factors (GEFs). We found that in lung ADC cells, expression of DOCK4, but not other DOCK180-family members, is rapidly and robustly induced by TGF-β in a Smad-dependent manner. Subsequently, we uncovered that DOCK4 is a direct TGF-β/Smad target gene, and, importantly, high DOCK4 expression correlates with activated TGF-β signaling and poor prognosis in human lung ADC. These findings prompted us to assess the role of DOCK4 in mediating the prometastatic effects of TGF-β in lung ADC in vivo in a mouse model. Remarkably, we found that DOCK4 induction is essential for TGF-β-driven lung ADC metastasis. Specifically, we observed that blockade of TGF-β-mediated DOCK4 induction attenuates the ability of lung ADC cells to extravasate into distant organ sites, resulting in a marked reduction in metastatic burden. At a cellular level, our evidence supports a model in which TGF-β-induced DOCK4 facilitates extravasation by stimulating lung ADC cell protrusive activity, motility, and invasion, without promoting epithelial-to-mesenchymal transition (EMT) and, intriguingly, that it does so by driving Rac1 activation. So far, Rac1 has only been linked to TGF-β via a noncanonical pathway. Thus, our findings identified the atypical Rac1 activator DOCK4 as a novel key component of the TGF-β/Smad pathway that promotes lung ADC cell extravasation and metastasis.

Recently, we expanded this line of research toward identifying genes that mediate lung ADC organ-specific metastases with a particular focus on genes that mediate colonization of specific organs. To this end, we established a multiple-organ metastasis model system using mice injected intracardially with lung ADC cell populations derived from primary tumors arising in $Kras^{G12D}/p53^{-/-}$ mice. With this model system, we initiated an in vivo RNAi screen designed to assess the possible involvement of a select set of genes in modifying the potential of lung ADC cells to metastasize to specific host organs. These studies were performed in collaboration with Kenneth Chang (CSHL). Interestingly, we found that silencing of this gene set resulted in a significant decrease in occurrence of brain metastases and an increase in the occurrence of bone metastases. These data suggest that one or more genes in this set modulate(s) lung ADC metastatic cell homing to and/or colonization of the brain and bone. We are currently investigating which of these genes, individually or in combination, modulate organ-specific homing and/or colonization of lung ADC cells in the brain and bone. These studies have the potential to identify novel mediator(s) and suppressor(s) of lung ADC metastases to brain and bone.

PUBLICATIONS

Van Aelst L, Downward J, McCormick F, eds. 2018. *Ras and cancer in the 21st century*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

In Press

Tai Y, Gallo NB, Wang M, Yu J-R, Van Aelst L. 2019. Axo-axonic innervation of neocortical pyramidal neurons by GABAergic chandelier cells requires AnkyrinG-associated L1CAM. *Neuron* **102:** 338–372.

NEUROSCIENCE

How does the brain encode stimuli from the outside world, within and across sensory modalities, to generate specific perceptions that trigger complex behaviors? How is the brain shaped by sensory experience, and what modifications occur in neuronal circuits that allow us to learn and remember? These are the questions guiding the work of **Florin Albeanu**, who is using the olfactory bulb and olfactory cortex of mice as the subject of his current studies. Airborne chemicals, translated into neuronal signals by specific receptors in the nose, are sent directly to the olfactory bulb. Advances in optical imaging and optogenetics, combined with electrophysiological recordings, enable Albeanu and colleagues to monitor and/or alter patterns of activity at unprecedented synaptic and millisecond resolution, in real time, as animals are engaged in various behaviors. For survival, rodents need to identify the smells of objects of interest such as food, mates, and predators, across their recurring appearances in the surroundings, despite apparent variations in their features. Furthermore, animals aptly extract relevant information about environment across different sensory modalities, combining olfactory, visual, or auditory cues. By recording neuronal activity in the input and output layers of the olfactory bulb, as well as feedback from olfactory cortical areas and neuromodulatory signals, Albeanu and his team aim to understand computations the bulb performs and how this information is decoded deeper in the brain. They have recently published evidence suggesting that the mouse olfactory bulb is not merely a relay station between the nose and cortex, as many have supposed. Using optogenetic tools and a novel patterned illumination technique, they discovered that there are many more information output channels leaving the olfactory bulb for the cortex than there are inputs received from the nose. They are currently investigating how this diversity of bulb outputs is generated, as well as how downstream areas, such as the piriform and parietal cortexes, make use of such information during behaviors.

The study of decision-making provides a window into the family of brain functions that constitute cognition. It intervenes between perception and action and can link one to the other. Although much is known about sensory processing and motor control, much less is known about the circuitry connecting them. Some of the most interesting circuits are those that make it possible to deliberate among different interpretations of sensory information before making a choice about what to do. **Anne Churchland**'s lab investigates the neural machinery underlying decision-making. Lab members use carefully designed paradigms that encourage experimental subjects to deliberate over incoming sensory evidence before making a decision. Recent results show that rats and humans have a statistically similar decision-making ability. To connect this behavior to its underlying neural circuitry, the researchers measure electrophysiological responses of cortical neurons in rodents as they perform designated tasks. The lab's current focus is on the parietal cortex, which appears to be at the midpoint between sensory processing and motor planning. Churchland and colleagues also use theoretical models of varying complexity to further constrain how observed neural responses might drive behavior. This approach generates insights into sensory processing, motor planning, and complex cognitive function.

The brain's activity is in constant motion: It ebbs and flows in big waves when we are in a deep slumber, turns into small ripples when we reawaken, and flows in orchestrated streams when we perceive, decide, and remember. These complex dynamics are driven by intricate networks of microscopic interactions between hundreds of thousands of neurons, and thus are only vaguely seen in spike trains of single neurons. Fortunately, recent advances in recording techniques enable us to monitor the activity of large neural populations in behaving animals, offering the opportunity to investigate how dynamic variations of collective neural-activity states translate into behavior.

To gain insights from these large-scale recordings, **Tatiana Engel** and colleagues develop and apply computational methods for discovering collective neural dynamics from sparse, high-dimensional spike-train data. They also develop models and theory to explain how collective neural dynamics support specific network computations and how these dynamics are constrained by biophysical properties of neural circuits. In these endeavors, they use and extend tools and ideas from diverse fields, including statistical mechanics, machine learning, dynamical systems theory, and information theory. Their work benefits from close collaborations with experimental neuroscience laboratories that are collecting neurophysiological data in animals engaged in sophisticated tasks, such as attention, decision-making, and learning.

Hiro Furukawa's lab studies receptor molecules involved in neurotransmission. Its members mainly focus on the structure and function of NMDA (*N*-methyl-D-aspartate) receptors—ion channels that mediate excitatory transmission. Dysfunctional NMDA receptors cause neurological disorders and diseases, including Alzheimer's disease, Parkinson's disease, schizophrenia, depression, and stroke-related ischemic injuries. The Furukawa lab is working to solve the three-dimensional structure of the very large NMDA receptor by dividing it into several domains. They seek to understand the pharmacological specificity of neurotransmitter ligands and allosteric modulators in different subtypes of NMDA receptors at the molecular level. Toward this end, they use cutting-edge techniques in X-ray crystallography to obtain crystal structures of the NMDA receptor domains and validate structure-based functional hypotheses by a combination of biophysical techniques, including electrophysiology, fluorescence analysis, isothermal titration calorimetry, and analytical centrifugation. Crystal structures of NMDA receptors serve as a blueprint for creating and improving the design of therapeutic compounds with minimal side effects for treating neurological disorders and diseases. During the last several years, the team discovered and mapped several regulatory sites in specific classes of NMDA receptors—progress that now opens the way to the development of a new potential class of drugs to modulate the receptor activity.

Josh Huang and colleagues study the assembly and function of neural circuits in the neocortex of the mouse. The neocortex consists of a constellation of functional areas that form a representational map of the external (sensory, social) and internal (visceral, emotional) world. These areas are strategically interconnected into elaborate information processing networks that guide behavior. The group's overarching hypothesis is that, at the cellular level, cortical processing streams and output channels are mediated by a large set of distinct glutamatergic pyramidal neuron types, and functional neural ensembles are regulated by a diverse set of GABAergic interneuron types. Understanding cortical circuit organization requires comprehensive knowledge of these basic cellular components. The Huang lab uses state-of-the-art genetic approaches to systematically target cell types and facilitate the application of a full set of modern techniques for exploring cortical circuits. Among GABAergic interneurons, the chandelier cell is one of the most distinctive cell types that controls pyramidal neuron firing at the axon initial segment. Huang and colleagues are studying the developmental specification, activity-dependent circuit integration, and functional connectivity of chandelier cells—an entry point toward understanding a local circuit module. Regarding pyramidal neurons, they are systematically characterizing the developmental origin, axon projection pattern, and input connectivity of multiple classes of pyramidal neuron types, focusing on the forelimb motor cortex. They combine a range of approaches that include genetic and viral engineering, genetic fate mapping, gene expression profiling, cellular imaging, electrophysiology, and behavior analysis. Recently, they began to integrate their studies in the context of the motor cortex control of forelimb movements.

Adam Kepecs and colleagues are interested in identifying the neurobiological principles underlying cognition and decision-making. They use a reductionist approach, distilling behavioral questions

to quantitative behavioral tasks for rats and mice that enable the monitoring and manipulation of neural circuits supporting behavior. Using state-of-the-art electrophysiological techniques, they first seek to establish the neural correlates of behavior and then use molecular and optogenetic manipulations to systematically dissect the underlying neural circuits. Given the complexity of animal behavior and the dynamics of neural networks that produce it, their studies require quantitative analysis and make regular use of computational models. The team also has begun to incorporate human psychophysics to validate its behavioral observations in rodents by linking them with analogous behaviors in human subjects. Currently, the team's research encompasses study of (1) neural basis of decision confidence, (2) the division of labor among cell types in the prefrontal cortex, (3) how the cholinergic system supports learning and attention, and (4) social decisions that rely on stereotyped circuits. A unifying theme is the use of precisely timed cell-type and pathway–specific perturbations to effect gain and loss of function for specific behavioral abilities. This year, the Kepecs lab was able to link foraging decisions—the choice between staying or going—to a neural circuit and specific cell types in the prefrontal cortex. In other work, they identified a class of inhibitory neurons that specializes in inhibiting other inhibitory neurons in the cerebral cortex and conveys information about rewards and punishment. Through manipulations of genetically and anatomically defined neuronal elements, the team hopes to identify fundamental principles of neural circuit function that will be useful for developing therapies for diseases such as schizophrenia, Alzheimer's disease, and autism spectrum disorder.

Alexei Koulakov and colleagues are trying to determine the mathematical rules by which the brain assembles itself, with particular focus on the formation of sensory circuits such as those involved in visual perception and olfaction. The visual system of the mouse was chosen for study in part because its components, in neuroanatomical terms, are well understood. What is not known is how projections are generated that lead from the eye through the thalamus and into the visual cortex, how an individual's experience influences the configuration of the network, and what parameters for the process are set by genetic factors. Even less is known about the assembly of the neural net within the mouse olfactory system, which, in the end, enables the individual to distinguish one smell from another with astonishing specificity and to remember such distinctions over time. These are among the challenges that engage Koulakov and his team.

Understanding the link between neural circuits and behavior has been the focus of research in **Bo Li**'s lab. They are particularly interested in studying the synaptic and circuit mechanisms underlying reward processing, attention, and learning and memory, as well as synaptic and circuit dysfunctions responsible for maladaptive behaviors that are related to major mental disorders. They integrate in vitro and in vivo electrophysiology, imaging, molecular, genetic, optogenetic, and chemogenetic techniques to probe and manipulate the function of specific neural circuits—with a focus on the fear and reward circuits—in the rodent brain and determine how these circuits participate in adaptive or maladaptive behavioral responses in various tasks.

Partha Mitra is interested in understanding intelligent machines, which are products of biological evolution (particularly animal brains), with the basic hypothesis that common underlying principles may govern these "wet" intelligent machines and the "dry" intelligent machines that are transforming the present economy. Dr. Mitra initiated the idea of brain-wide mesoscale circuit mapping, and his laboratory is involved in performing such mapping in the mouse (http://mouse.brainarchitecture.org) and the marmoset (in collaboration with Japanese and Australian scientists at the RIKEN Brain Science Institute and Monash University).

Dr. Mitra spent 10 years as a member of the theory department at Bell Laboratories and holds a visiting professorship at IIT Madras, where he is helping establish the Center for Computational Brain Research. He has an active theoretical research program in machine learning and control

theory, wherein he is using tools from statistical physics to analyze the performance of distributed/networked algorithms in the "thermodynamic" limit of many variables.

Pavel Osten's lab works on identification and analysis of brain regions, neural circuits, and connectivity pathways that are disrupted in genetic mouse models of autism and schizophrenia. Osten hypothesizes that (1) systematic comparison of multiple genetic mouse models will allow determination of overlaps in pathology—neural circuit endophenotypes—responsible for the manifestation of neuropsychiatric disorders, and (2) neural circuit–based classification of autism and schizophrenia will provide key circuit targets for detailed mechanistic studies and therapeutic development. Osten and colleagues have developed the first systematic approach to the study of neural circuits in mouse models of psychiatric diseases, based on a pipeline of anatomical and functional methods for analysis of mouse brain circuits. An important part of this pipeline is high-throughput microscopy for whole-mouse brain imaging, called serial two-photon (STP) tomography. This year, they used this method to describe the first whole-brain activation map representing social behavior in normal mice. They are currently focusing on using this approach to study brain activation changes in two mouse models of autism: the 16p11.2 df/+ mouse model, which shows an increased propensity to seizures and hyperactivity, and the CNTNAP2 knockout mouse model, which shows abnormal social behavior.

Stephen Shea's lab studies the neural circuitry underlying social communication and decisions. They use natural social communication behavior in mice as a model to understand circuits and processes that are evolutionarily conserved and therefore shared broadly across species, likely contributing to disorders such as autism. Shea and colleagues have examined how emotion and arousal enable mice, via their olfactory systems, to store memories of other individuals and related social signals. The team has exploited the intimate relationship between memory and emotion to effectively create memories in anesthetized mice, allowing them unprecedented access to neurobiological processes that typically only occur during behavior. The lab has been making a detailed analysis of the changes in neural connections that underlie odor memory. The team is particularly focused on an enigmatic cell type (granule cells or GCs) that has long been hypothesized to be crucial for memories but has resisted direct study. They have developed methods for recording, giving them the first glimpse of the dynamics of these cells while the animal is learning an odor. The results show unexpectedly complex population dynamics among the GCs, which were independently predicted by a model of odor learning developed in Alexei Koulakov's lab. The two labs are collaborating to discern how GC population activity gets integrated by olfactory bulb output neurons and to pinpoint the synaptic circuit that underlies this form of learning. In parallel, another member of the lab is using imaging techniques to determine how memories are stored among broad neuronal ensembles at a different level of the system. Recently, the lab made a key breakthrough, developing the ability to record from GCs in awake animals and discovering that their activity is dramatically modulated by state of consciousness. Finally, the Shea lab completed a series of studies of a different form of social recognition: auditory recognition of pup vocalizations by their mothers. Through this research, they have shown that a mouse model of Rett syndrome shows deficits in communication and learning not unlike those in human patients. Grants from the Simons and Whitehall Foundations are allowing the lab to extend this work by directly linking these deficits to the action of the gene *MeCP2* in the auditory cortex.

Jessica Tollkuhn's lab seeks to understand how transient events during brain development exert lasting effects on gene expression, circuit function, and, ultimately, behavior. They study how sex-specific neural circuits in rodents are established and modulated by the gonadal hormones estrogen and testosterone. The cognate receptors for these hormones are nuclear receptor transcription factors, which orchestrate modification of local chromatin environment and thus exert

long-term effects on gene expression. However, the genes regulated by these receptors, as well as the specific mechanisms they use, remain poorly understood in the brain. This is in part because the extraordinary cellular heterogeneity of the brain complicates analysis of the small subpopulations of neurons that mediate sex-specific behaviors.

Having recently identified sex differences in both gene expression and chromatin in brain regions known to regulate sex-specific behaviors, the Tollkuhn lab is now working to understand how hormones generate these molecular sex differences during development through the use of biochemical, genomic, and behavioral analyses. They have developed a method that permits genome-wide analysis of histone modifications or DNA methylation in genetically defined populations of neurons. They hypothesize that these epigenetic data, combined with gene expression profiling, define the molecular signature of the critical period for sexual differentiation of the brain. Their goal is to provide a mechanistic link between the transcriptional effects of hormone signaling during development and the consequential social behaviors displayed in adulthood.

Anthony Zador and colleagues study how brain circuitry gives rise to complex behavior. Work in the lab is focused on two main areas. First, they ask how the cortex processes sound, how that processing is modulated by attention, and how it is disrupted in neuropsychiatric disorders such as autism. Recently, the lab found that when a rat makes a decision about a sound, the information needed to make the decision is passed to a particular subset of neurons in the auditory cortex wherein axons project to a structure called the striatum. In the second major line of work in the Zador lab, they are developing new methods for determining the complete wiring instructions of the mouse brain at single-neuron resolution, which they term the "connectome." In contrast to previous methods, which make use of microscopy, these methods exploit high-throughput DNA sequencing. Because the costs of DNA sequencing are plummeting so rapidly, these methods have the potential to yield the complete wiring diagram of an entire brain for just thousands of dollars.

UNDERSTANDING THE NEURONAL BASIS OF INTERNAL MODELS OF THE WORLD AND MAPPING THE ODOR SPACE

F. Albeanu W. Bast S. Choi P. Gupta
H. Chae D. Cowan D. Hernandez Trejo
Y. Chen M. Davis P. Villar

The focus of our research group is twofold: (1) understanding the logic of odor space and olfactory neuronal representations underlying olfactory perception, and (2) understanding how brain-wide neuronal circuits learn the statistics of the world (internal models) and solve fundamental sensorimotor challenges in closed-loop behaviors. We investigate how the brain encodes and interprets inputs from environment as a function of context and prior expectations to implement meaningful behaviors. We focus on neuronal circuits in the olfactory stream (olfactory bulb, olfactory cortex, and striatum), as well as on inputs to sensory circuits from the motor cortex and association areas, in wild-type mice and models of psychiatric disorders.

Understanding the Neuronal Substrates of Internal Models of the World

Our perception of the world is a closed-loop process that relates sensory inputs to motor actions, which in turn shape the sensory inputs we sample. As a result of this reciprocal relationship between sensory inputs and motor actions, the brain learns to predict the sensory consequences of its actions by running internal simulations (internal models). Using novel closed-loop behavioral tasks in rodents, in combination with ultrasound, optical and electrical monitoring of activity, and optogenetic manipulations, we aim to identify the neuronal substrates of internal models across the brain.

During behavior, sensation and action operate in a closed loop. Movements shape sensory input, and sensory inputs guide motor commands: *Where one looks determines what one sees.* Through experience, the brain learns the reciprocal relationship between sensory inputs and movements to build internal models that predict the sensory consequences of upcoming actions (*sensorimotor predictions*). Comparing internal sensory predictions to actual sensory observations generates prediction errors that can be minimized by learning increasingly accurate models of the world. This exchange of sensory inputs and egocentric expectations is at the core of *active perception*. Experimental investigation of this idea has been sparse and split between behavioral interrogation of sensory-guided, precise motor control in primates (visuomotor adaptation tasks) and the search for neuronal substrates of sensory predictions in rodents via simpler running-based closed-loop behaviors.

To study *internal models* at both the behavioral and the circuit level, we developed a novel behavioral task in which head-fixed mice are trained to steer the left–right location of an odor source by controlling a lightweight lever with their forepaws. In this manner, we (1) link a precise motor action to well-defined sensory expectations (odor location) and (2) subsequently violate the learnt expectations via online feedback perturbations in trained animals. Expert mice (6 out of 6 trained mice, training period <2 wk, >90% accuracy, 400–800 trials/session) showed precise movements that were locked to the instantaneous odor feedback during normal closed-loop coupling. However, when sensory feedback was transiently interrupted (halting of odor source) or distorted (displacement of odor source or change in movement gain; Figs. 1 and 2), movements were initially guided by each animal's learned internal model and further quickly adapted (within a few sniffs in single trials) in accordance with the instantaneous sensory error. We are currently probing activity in olfactory and motor cortex and the olfactory striatum to understand the sensorimotor transformations that enable this behavior.

Neuronal Substrates of Olfactory Perception

Olfactory perception is critically dependent on responses of a large cohort of olfactory receptors (ORs) to an even larger set of odorants. Progress in olfaction has been substantially impeded by failure in the field

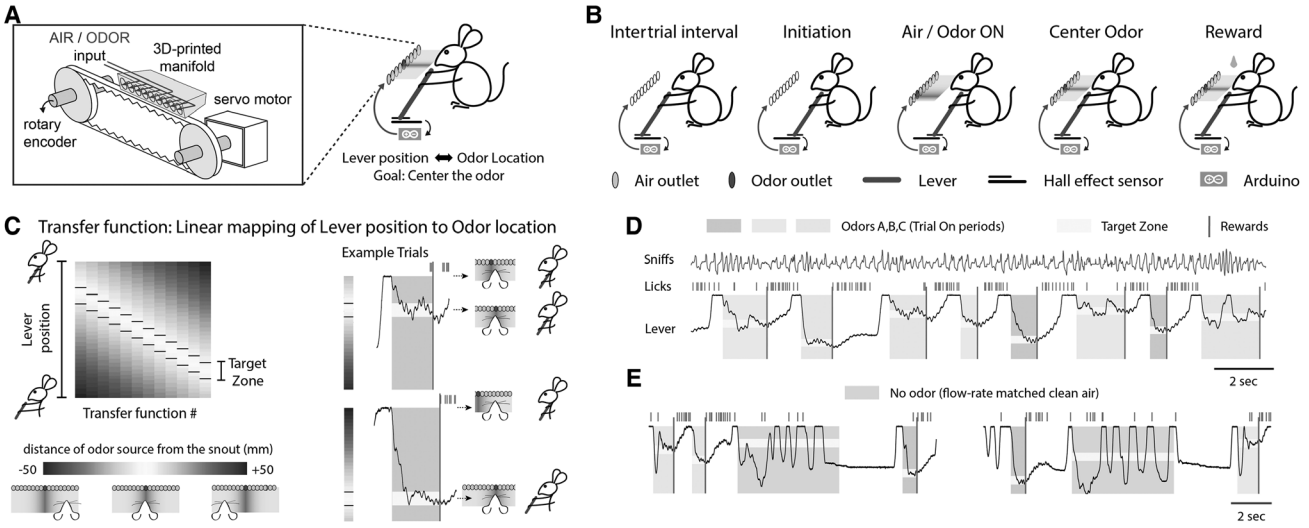


Figure 1. A novel method for fast, closed-loop control of odor source location. (*A*) Schematic of the system: 1D movement of the lever (read by a Hall effect sensor) is transformed to left–right displacement of an odor source. The odor location is manipulated by displacing a 3D-printed manifold affixed to a timing belt servo system. The manifold comprises a central odor outlet and 16 air outlets on either side. Flow rate (0.2 L/min) is matched across all outlets. To obtain rewards, animals are required to align the odor outlet to their snout. (*B*) Trial structure: mice initiate trials by retracting and holding the lever to activate air/odor flow and place the odor source at a fixed starting location. Mice can then steer and hold the odor, centered on their snout, by bringing the lever within a narrow target zone. Maintaining target zone hold for ~300 msec triggers a reward and terminates the trial. (*C*) (*Left*) Different transfer functions map lever position to odor source displacement. Colors indicate relative lateral distance from the snout. Black lines demarcate the target zone—that is, the set of lever positions that place the odor <3 mm from the snout. All transfer functions have the same gain. Unit displacement of the lever results in unit displacement of odor, but the odor source location at the trial start differs across trials. Using different transfer functions across trials ensures that reward availability is not associated with any specific lever position. (*Right*) Two example trials with different transfer functions and corresponding lever trajectories in time (black). A gray bar indicates trial-on period. Yellow demarcates the target zone. Schematics show odor locations at the trial start and in the target zone. Green ticks show a water reward; red ticks show licks. (*D*) Example behavioral trace from an expert mouse showing that the animal reliably centers the odor in each trial despite varying target zone locations (yellow bands). (*E*) Lever movements are unstructured in catch trials in which olfactory feedback is not provided, confirming that animals rely on closed-loop odor feedback to guide their movements during this task.

to accumulate information about the OR/odorant binding affinity matrices in vivo with sufficiently high throughput and to relate them across different conditions in a predictive manner. Our approach relates the molecular identity of ORs to their in vivo responses and to downstream activity and connectivity patterns for hundreds of odorants and thousands of individual neurons per brain (Fig. 3). To this end, we use in vivo functional imaging of large odor panels in conjunction with fluorescent in situ RNA sequencing (FISSEQ), a novel technology developed outside neuroscience, and MAPseq, which was developed using CSHL and modern machine learning techniques. We aim to predict the perceptual quality of an odorant based solely on its molecular structure, a long-sought-after dream in olfaction. In vision and other senses, substantial progress was achieved by understanding the features of the stimulus space that are represented by the brain. The realization that color perception is based on three types of cone photoreceptors enabled the invention of cameras and displays that faithfully reproduce any natural stimulus by mixing a basis set of just three lights. In the case of smell, we lack any comparable conceptual understanding. We do not understand what properties of odorants lead to particular percepts and how these properties are represented in the neuronal activity. Our approach is ideally suited to provide the answer, by first building a framework for relating the physical and neuronal spaces, and further connecting them to odor perception using modern machine learning techniques (deep neural networks [DNNs]). Altogether, we aim to identify the analogs of the red–green–blue basis set of odor perception. (This work was done in collaboration with the Zador and Koulakov labs.)

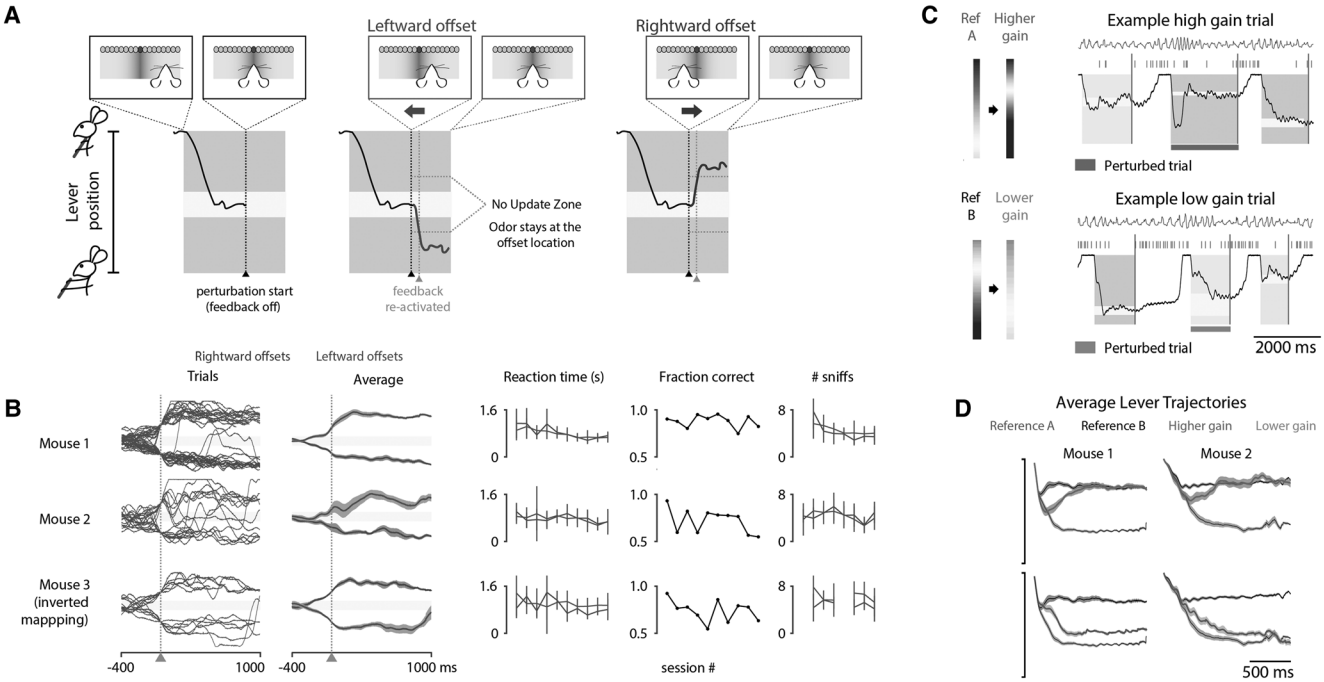


Figure 2. Feedback perturbations trigger corrective movements that reveal the animal's internal model. (*A*) Schematic of within-trial perturbations of odor location and expected corrective movements (red, leftward odor offset; blue, rightward odor offset). This example illustrates the scenario for a mouse that was trained to move the odor from left → right via forward motion of the lever (away from the body). The vertical black dotted line indicates the time of odor displacement (perturbation start). No update in odor location is provided until the corrective movement crosses the threshold (2× the width of the target zone). This ensures that observed corrective movements purely reflect the animal's internal model and are not refined via instantaneous odor feedback. The vertical orange dotted line marks the time of reactivation of closed-loop feedback. (*B*) (*Left*) Distinct trial-averaged and single-trial corrective movements in three mice, sorted by leftward (red) and rightward (blue) odor offsets. Mice 1 and 2 were trained to move the odor left → right by forward motion of the lever, whereas mouse 3 was trained on the opposite (forward lever movement = right → left odor displacement). Note that corrective movements for left and right odor offsets diverge well before closed-loop feedback is reactivated. (*Right*) Summary plots showing session-averaged reaction times (time from offset start to threshold crossing), success rates (fraction of trials in which the corrective response was in the correct direction), and number of sniffs during perturbation period (offset start to feedback reactivation). (*C*) Schematic of gain-change perturbations. (*Top*) Example trial with higher (3×) lever gain compared to reference. As a result, centering the odor requires smaller lever displacement (target zone closer to the body) despite the same odor start location in perturbed and reference trials. Note how the animal initially overshoots and steers the lever to where the target zone would be given normal gain (1×) and then quickly corrects to successfully center the odor, despite the higher gain. (*Bottom*) Example trial with lower gain (0.4×) compared to reference. Here, the target zone is lower than expected. Thus, the animal initially stalls the lever (undershoots), but then quickly corrects. (*D*) Average lever trajectories during gain perturbation trials in comparison to reference control trials showing consistent overshoots upon higher-gain and undershoots upon lower-gain perturbations.

Understanding the Relationship between Olfactory Perceptual Discriminability and Glomerular Response Features

For rodents, the ability to recognize and discriminate particular combinations of volatile compounds is essential for their survival. Mice can easily report the difference between weak, similar odors in rich sensory scenes, even when stronger odorants fluctuate in the background. To date, the neural mechanisms underlying such behavior remain unknown.

To understand the neural basis of odor discrimination, we measured and manipulated the activity of the input nodes of the olfactory system, the glomeruli. By using wide-field optical imaging in conjunction with odor stimulation, we tracked the position of glomeruli

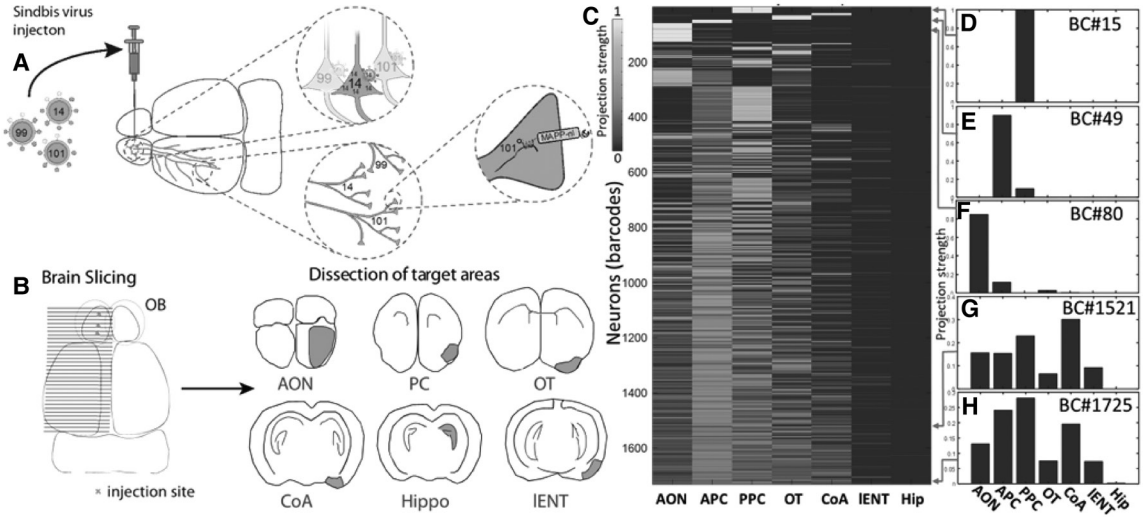


Figure 3. MAPseq experiments reveal biases in OB projection patterns. (*A,B*) Schematics of injection site, slicing, and LASER microdissection of tissue from the six major OB target areas. (*C*) Heatmaps of projection patterns of 1,732 barcodes (BCs) from four mice. Projection density is color-coded. BCs are clustered according to similarity20. Clusters are sorted according to the specificity of their projections: specific on *top* and broad at the *bottom*. (*D–H*) Individual BC projection patterns range from specific (BCs #15, #29, and #80) to broad (BCs #1521 and #1725).

and quantified their odor response properties; this allowed us to define different sets of affine and nonaffine glomeruli with a variable number of components. We aim to determine the relationship between the discriminability of olfactory stimuli and the similarity of glomerular odor response profiles. We additionally quantified the discriminability of the stimuli with the degree of overlap between different sets of glomeruli, as well as the physical separation of glomeruli on the bulb surface.

Toward this end, and to assess the specificity of photo-stimulation, we express red-activatable channelrhodopsin1 (ReaChR) in all mature olfactory sensory neurons and GCaMP6f in the OB output neurons. We use DMD-based patterned illumination to selectively stimulate combinations of glomeruli on the dorsal surface of the bulb with subglomerular resolution (~10 μm) and high temporal precision (3 msec) in awake, head-fixed mice. Prior to optogenetic stimulation, using a large odor panel (up to approximately 100 stimuli), we identified the exact locations of glomeruli, revealing their shapes and response tuning to the odors sampled. We further create glomerular light patterns of known odor response similarity (within the range of our panel) and project specific glomerular inputs. In a two-alternative forced-choice discrimination task, we systematically relate the similarity of these light patterns to the perceived difference between them. Further using a novel strategy to decouple patterned photo-stimulation and two-photon imaging across different axial planes, we are monitoring the responses of mitral and tufted (M/T) cells in the deeper layers of the bulb.

We are further implementing strategies that will enable noninvasive, functional dissection of neuronal networks, with cellular resolution, in behaving animals. This will be brought about via a closed-loop strategy involving real-time control of activity of select neurons with simultaneous monitoring of the concomitant effects of these manipulations on neuronal outputs within the circuit and elsewhere in the brain. Briefly, we are employing digital holography methods via spatial light modulators (SLMs) to optogenetically control neurons of interest at the single-cell level and DMD-based methods to control cell type–specific populations across large brain regions (Fig. 4). This allows us to both replicate and systematically manipulate stimulus-evoked activity patterns in a circuit. We are simultaneously employing two-photon calcium imaging and electrophysiology within the same and different brain regions (olfactory bulb vs. olfactory cortex) to dissect how the alteration of select circuit elements, or their specific properties, affects the output of the network. This closed-loop approach will make it possible to determine the spatiotemporal

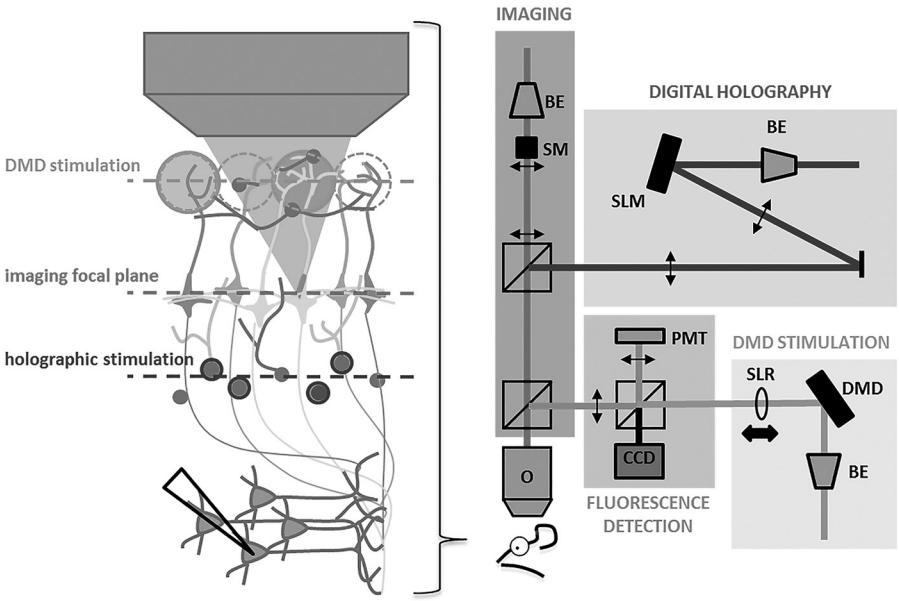


Figure 4. Combined imaging and photo-stimulation. (*Left*) Experimental configuration, combining scanning two-photon imaging (red), DMD photo-stimulation (blue), and holographic photo-stimulation (brown). (*Right*) Schematic of the experimental microscope. DMD stimulation is used to create spatiotemporal light patterns on the surface (<100 µm). Digital holography is used to photo-stimulate deeper (<500 µm) in the brain with cellular resolution. Calcium activity is monitored in an independent optical plane by two-photon imaging and by electrodes in downstream brain regions. (BE) Beam expander, (SM) scan mirrors, (O) objective, (PMT) photo-multiplier, (SLR) camera lens, (DMD) digital micromirror device, (SLM) spatial light modulator, (CCD) charge-coupled device.

integration rules within the bulb and the olfactory cortex, investigate the relevance of spike-time codes, and reveal underlying decoding schemes.

Task-Related Representations in the Corticobulbar Feedback

Sensory systems format information about objects in the environment into neural representations used by the brain depending on context and prior experience. Sensory representations emerge from the interplay between feedforward inputs, as well as local and feedback signals across brain areas. Given that the output neurons of the main olfactory bulb—the first processing station of olfactory input—are richly modulated by task contingencies, we asked whether the feedback from piriform cortex provides an underlying substrate. To this end, we engage mice in a serial reversal learning task, in which a stimulus (odorant or sound) carries distinct motivational value according to different contingencies. In parallel, we monitored the activity of GCaMP5-labeled corticobulbar feedback axons. We report the existence of two largely independent populations of axon terminals, responsive in sound versus odor trials, and spanning diverse temporal profiles across both response polarities. Correlation and dimensionality reduction population analysis identified sets of choice-selective and choice-independent boutons active in both odor and sound trials (Fig. 5). Sound responses occurred mainly during reward (hit) and error (false alarm) trials, further suggesting that the corticobulbar feedback carries information on action outcome. Optogenetic inactivation of cortical feedback axons locally in the bulb substantially impaired the task performance compared to controls for both odor and sound trials. These effects were dominated by increased error rates (false alarms) and persisted beyond the light-on trials. In ongoing experiments, we are monitoring the cortical feedback and OB outputs while varying the size of reward, the interval between sensory cues and the time when reward is available, and the identity of odor–sound pairs.

Taken together, our results suggest that cortical feedback conveys information related to task contingencies such as learning rules and reward and is well positioned to flexibly control the bulb output during olfactory behaviors.

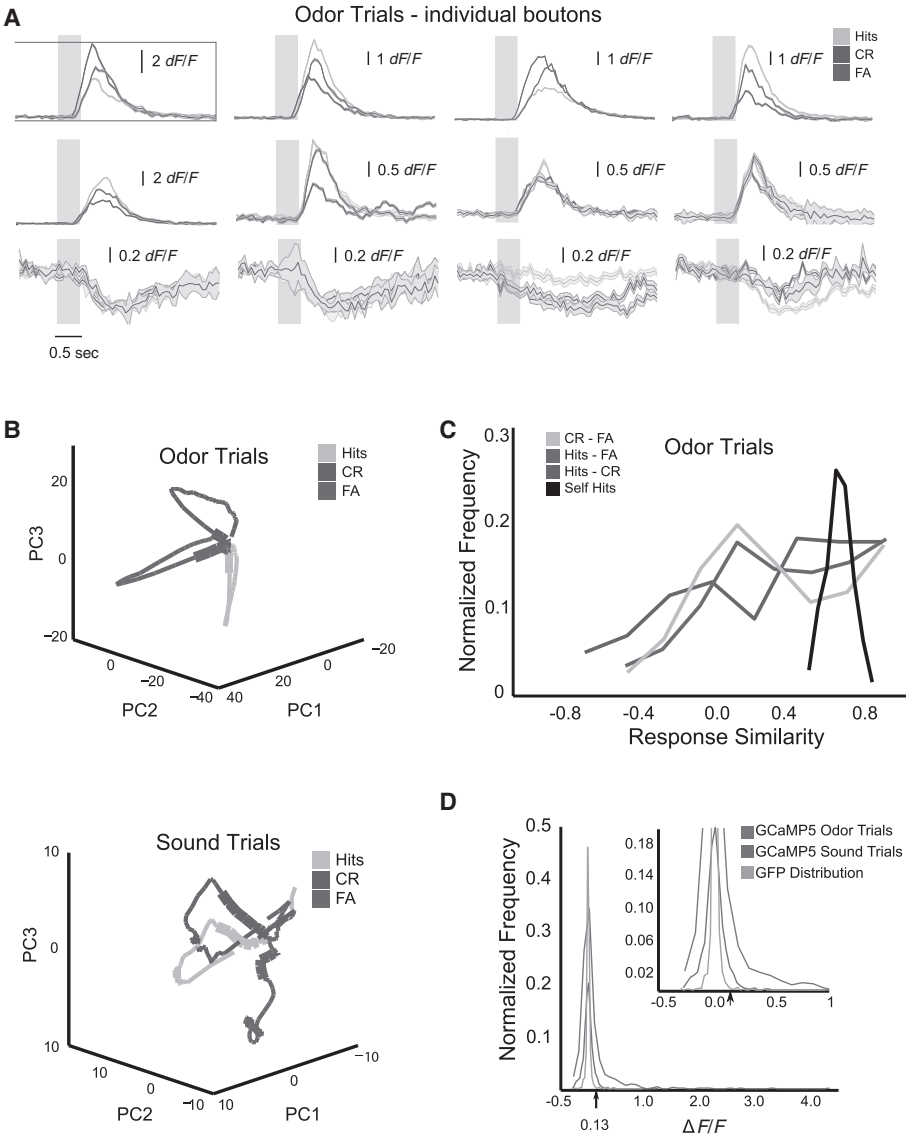


Figure 5. Cortical feedback responses to odor stimuli cluster differently across different task conditions. (*A*) Excitatory and inhibitory trial-averaged responses from individual feedback boutons exemplifying differential modulation across task conditions. (Thick line) average response, (shaded area) ± SEM, (gray bar) stimulus (350 msec). (*B*) Ensemble responses of boutons across trial outcomes projected onto the first three principal components (64% of variance). Each trajectory represents the bouton population activity over time for odor (*top*) and sound trials (*bottom*). (Origin marks) the start of the trial, (thick lines) the stimulus period. (*C*) Histograms of pairwise response similarity for odor trials across conditions. (Colored lines) Uncentered correlation between responses of single boutons for different trial outcomes, (black line) distribution of self-correlations between two halves of hit trials resampled by bootstrapping 1,000 times. (*D*) Histograms of average Δ*F*/*F* for GCaMP responses versus GFP. Each point is the average fluorescence of a 600-msec window starting at the stimulus onset. An arrow denotes the 99th percentile of the GFP distribution, used as the significance threshold for GCaMP responses for further analysis.

Two Parallel Feedforward and Feedback Pathways for Olfactory Information Processing

Understanding the interplay between feedforward and feedback neuronal signals across interconnected brain areas is essential for unveiling the computations they perform. Across the brain, specialized long-range circuits support different processing streams. These broadcast information ranging from multiple features of sensory stimuli, decision variables, and inner states to substrates for planning and execution of motor actions.

To date, the logic of information flow within the early mammalian olfactory system remains poorly understood. It is not known whether different projection neurons carry different signals to particular areas, and to what degree the feedback from such target brain areas to the sensory periphery is specific to the inputs they receive. We find that the two classes of olfactory bulb outputs, the mitral and tufted cells (MTCs), which innervate distinct sets of higher brain areas including piriform cortex, PC, and anterior olfactory nucleus (AON), respectively, are in turn specifically regulated by differential negative feedback from these areas. Specifically, bulbar feedback from these cortical areas modulates differentially the response amplitude, sparseness, and pairwise correlations of odor representations at the level of mitral versus tufted cell ensembles. Moreover, the effect of these feedback signals is precise, acting proportionally to the strength of feedforward drive and in an odor–cell pair-specific manner. Furthermore, we find that concentration-invariant odor-specific representations emerge already in the bulb outputs. Surprisingly, tufted ensembles largely outperform mitral cells in decoding stimulus identity with concentration invariance. Lack of cortical feedback predominantly impairs this odor-decoding performance in mitral cells while largely sparing tufted cell ensembles. These results identify two parallel feedforward–feedback loops in the early olfactory system and indicate that they have specialized roles in odor processing (Fig. 6).

Within this scenario, we are currently testing the hypothesis that cortical targets of TCs (AON, olfactory tubercle) are poised to infer concentration-invariant odor identity, whereas the cortical targets of MCs such as the PC are more suited to solving odor-related tasks of a different nature (i.e., contextual learning, flexible rule assignment, generating internal representations of the environment, etc.). In summary, we describe two parallel feedforward and feedback streams of odor processing in the mammalian brain with differential inhibitory organization and neural representation for odor identity and concentration, further providing a testable framework during olfactory behaviors that go beyond sensory decoding.

Mapping Odor Space onto Neuronal Representations in the Olfactory System

Unlike many other sensory systems, low-dimensional metrics for characterizing stimuli have remained elusive for olfaction, and it is unclear what features of chemical stimuli are represented by neurons. We would like to relate neural activity in the early olfactory system of mice to the physical–chemical properties of odorants. The elementary stimulus features encoded by the olfactory system remain poorly understood. We examined the relationship between 1,666 physical–chemical descriptors of odors and the activity of olfactory bulb inputs as well as outputs in awake mice. Glomerular and M/T cell responses were sparse and locally heterogeneous, with only a coarse dependence of glomerular positions on physical–chemical properties. Odor features represented by ensembles of M/T cells were overlapping but distinct from those represented in glomeruli, consistent with extensive interplay between feedforward and feedback inputs to the bulb. This reformatting was well described as a rotation in odor space. The descriptors accounted for a small fraction in response variance, and the similarity of odors in physical–chemical space was a poor predictor of similarity in neuronal representations. Our results suggest that commonly used physical–chemical properties are not systematically represented in bulbar activity and encourage further search for better descriptors of odor space (Fig. 7). (This work was done in collaboration with the Koulakov lab.)

Other Collaborative Projects with CSHL Groups

Huang: Functionally mapping the motor cortex via DLP-based patterned stimulation.

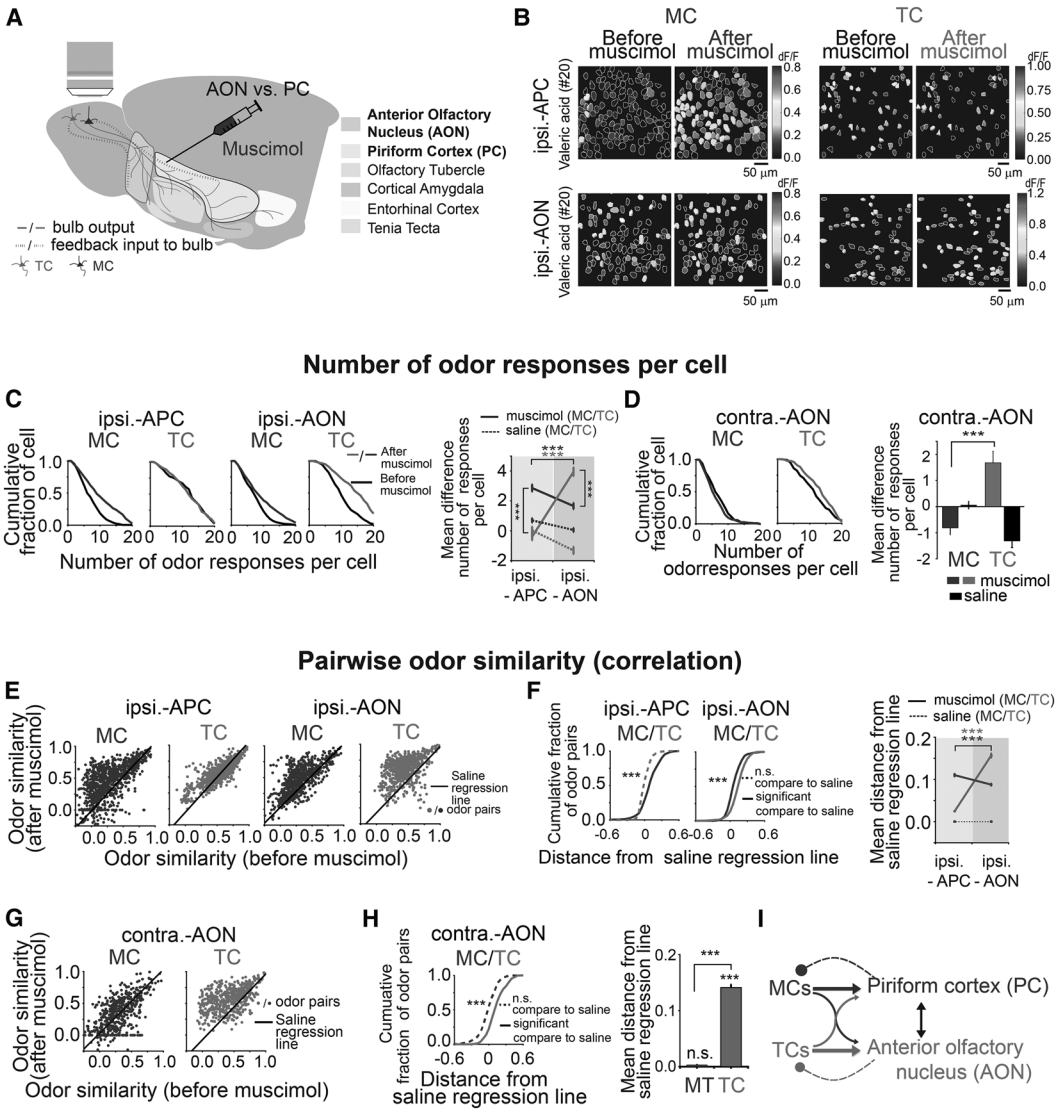


Figure 6. Two long-range feedforward–feedback functional loops in the olfactory system. (*A*) Schematic of the experimental framework: monitoring the activity of the olfactory bulb (OB) outputs, the mitral cells (MC), and tufted cells (TC) in awake head-fixed mice via multiphoton imaging while suppressing the feedback from the anterior olfactory nucleus (AON) and anterior piriform cortex (APC). (*B*) Color maps showing average fluorescence change in response of MCs and TCs to valeric acid (1:100 dilution) in example fields of view before (*left*) and after (*right*) muscimol injection in the APC (*top*) and AON (*bottom*). (*C,D*) (*Left*) Number of odor responses per cell before and after muscimol injection in the ipsi-APC, ipsi-AON, and contra-AON. (*Right*) Quantification of the average number of responses across conditions (before vs. after muscimol injection). (*E,G*) Pairwise odor similarity across cells before and after suppression of cortical feedback; each dot corresponds to one odor–odor pair. (*F,H*) Cumulative distributions of distances from saline regression line of pairwise odor similarity across conditions (before vs. after muscimol injection). (*I*) Cartoon representations of specific feedforward–feedback loops engaging the mitral cells and APC and the tufted cells and the AON.

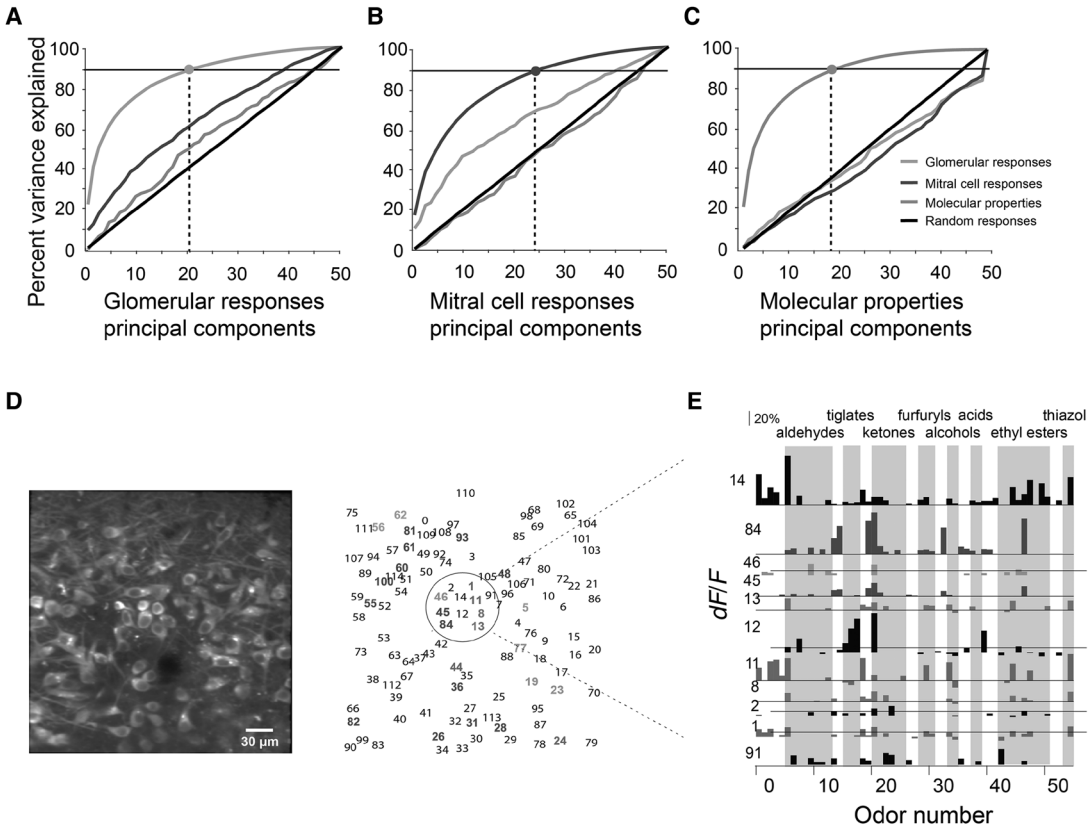


Figure 7. Diversity of odorant representations in the glomeruli and mitral cells, independent of molecular properties. (*A*–*C*) Results of principal component analysis (PCA) for (*A*) glomerular cell responses, (*B*) mitral cell responses, and (*C*) molecular properties ($n = 871$ glomeruli, $n = 639$ mitral cells, $n = 1{,}320$ properties). Percent of variance explained is shown as a function of the number of included principal components (PCs). (*A*) Percent variance explained for glomerular cells' odor responses (green), mitral cells' odor responses (blue), molecular property strength vectors (gray), and random data controls (black) shown as a function of the number of included glomerular response principal components. (*B,C*) Percent variance explained of glomerular cells' odor responses (green), mitral cells' odor responses (blue), molecular property strength vectors (gray), and random data controls (black) shown as a function of the number of included mitral cell response principal components (*A*) and of molecular property principal components (*B*). Note that neural responses of both glomeruli and olfactory bulb outputs (mitral cells) are poorly tuned to the physical–chemical properties analyzed and instead reflect odor features that are not well captured by these molecular properties commonly used in computational chemistry and by previous studies in olfaction. (*D*) (*Left*) Average resting fluorescence multiphoton image in the mitral cell layer. (*Right*) Numbers indicate the relative positions of 112 mitral cell bodies in the imaged field of view. (*E*) An example of odor responses of mitral cell bodies within an arbitrarily picked 75-µm-diameter region circled in *D* sorted according to functional chemical groups (aldehydes, tiglates, ketones, furfuryls, alcohols, acids, ethyl esters, thiazoles). Note that the neighboring neurons (enclosed in the circle in *D*) have diverse, as well as similar, odor tuning (e.g., cells 45 and 84 vs. cell 12). We probed the local diversity further, by rearranging the odor response spectra (ORS) such as to group odors in terms of chemical classes (ketones, acids, etc.). Mitral cells responding to many different classes of chemicals can be found even within a small region of glomerular size.

Koulakov and Zador: Sequencing the olfactory bulb—bringing the gap between glomerular odors responses and odor receptor sequences by identifying the molecular identity of glomeruli.

Li: Monitoring neuronal activity in punishment and reward neuronal circuits during behavior.

Osten: Developing light-sheet-based approaches for fast optical reconstruction of neuronal circuits.

Kepecs: Modulating VIP neurons by reward and punishment assessed using fiber photometry.

Zador: Optical monitoring and manipulation of neuronal activity in genetically and anatomically defined cortical circuits in animal models of cognition.

PUBLICATIONS

Albeanu DF, Provost AC, Agarwal P, Soucy E, Murthy VN. 2018. Olfactory marker protein (OMP) regulates formation and refinement of the olfactory glomerular map. *Nat Commun* **9:** 5073.

In Press

Chae HG, Kepple DR, Bast WG, Murthy VN, Koulakov A, Albeanu DF. 2019. Mosaic representations of odors in input and output layers of the mouse olfactory bulb. *Nat Neurosci* (in press).

INVESTIGATING NEURAL CIRCUITS FOR MULTISENSORY DECISION-MAKING

A. Churchland G. Bekheet, L. Chartarifsky, S. Gluf, M. Kaufman, A. Khanai, S. Musall, F. Najafi, S. Pisupati, J. Roach, X. Sun, A. Urai

Making use of sensory information requires more than simply relaying incoming signals from the sensory organs. It requires interpreting information, classifying it, drawing inferences, and ultimately using the context of behavioral goals to make a decision about its meaning. A decision is a commitment to a particular choice or proposition at the expense of competing alternatives. In some situations, decisions involve integration of evidence—that is, they combine multiple pieces of information from the environment or from memory. These decisions can provide a framework in which to investigate complex cognitive processes and open a window into higher brain function.

Although previous experiments have begun to reveal how neural systems combine evidence to make decisions, they have left a critical gap in our understanding. Specifically, very little is currently known about the neural mechanisms that make it possible to combine information from multiple sensory modalities for decisions. The gap is apparent even though it is clear from behavioral observations that neural systems can combine multisensory information: When parsing speech in a crowded room, for example, the listener makes use of both auditory information (the speaker's vocal sounds) and visual information (the speaker's lip movements). Understanding the neural mechanisms of multisensory decisions is critical for two reasons. First, it is essential for a complete understanding of sensory perception because real-world stimuli rarely activate a single sense in isolation. Therefore, understanding how the brain interprets incoming information requires understanding how the brain merges information from different senses. Second, it is likely of clinical importance as several developmental abnormalities appear to be related to difficulties in integrating sensory information. For example, abnormalities in multisensory processing are a hallmark of subjects with autism spectrum disorder. Impairments in multisensory processing are also observed in subjects with a collection of sensory abnormalities known together as sensory processing disorder and may also be evident in patients with Rett syndrome and dyslexia. Understanding the neural mechanisms of multisensory integration could inform treatment of those conditions. Our long-term goal is to understand how the brain can make decisions that integrate inputs from our multiple senses, stored memories, and innate impulses. Our projects in previous years constituted the very first steps toward achieving this goal: We developed a multisensory behavior that could be precisely quantified in both humans and rodents (Sheppard et al., *J Vis 13:* 4 [2013]; Raposo et al., *J Neurosci 32:* 3726 [2012]). In addition, we built on the foundation by measuring the responses of neurons in rodents engaged in the behavior, affording insight into the neural circuits underlying multisensory decisions (Raposo et al., *Nat Neurosci 17:* 1784 [2014]).

Understanding Lapses in Perceptual Decision-Making

L. Chartarifsky, S. Pisupati, A. Khanal

During perceptual decisions, subjects often display a constant rate of errors independent of evidence strength, referred to as "lapses." Their proper treatment is crucial for accurate estimation of perceptual parameters; however, they are often treated as a nuisance arising from motor errors or inattention. Based on new results from the last year, we are proposing a new explanation—that lapses can reflect a dynamic form of exploration. We demonstrated that uncertainty about sensory stimuli modulates the probability of lapses in rats. These effects cannot be accounted for by traditional models in the field and are instead concisely explained by a normative model of uncertainty-guided

exploration. Further, we showed that changing the reward associated with one of the decisions selectively affects the lapses associated with that decision in uncertain conditions, while leaving unchanged very easy "sure-bet" decisions, as predicted by the model. Finally, we demonstrated that muscimol inactivations of secondary motor cortex and posterior striatum affect lapses across modalities. Together, our results suggest a novel value-based account for lapses, and that far from being a nuisance, lapses are informative about individual animals' exploration–exploitation trade-off.

A Novel Device to Support Electrophysiological Recordings during Ethological Behavior

G. Bekheet [in collaboration with A. Juavinett, University of California, San Diego]

The starting point for this project was a major problem we saw in the field: The field has recently begun to benefit from Neuropixel probes, a unique biosensor with the ability to measure neural activity from far more neurons than was previously possible. These probes are being released to the community and afford the opportunity to make high-quality recordings of hundreds of neurons simultaneously spanning multiples brain areas. However, these probes are essentially unusable to scientists studying freely moving or ethological behaviors, which together constitute a critical component of neuroscience research. Further, Neuropixels were not designed to allow reuse after chronic implantation. This is a devastating problem for any researcher who needs chronically implanted probes in any experimental setup or animal.

We devised a solution to move the field forward: we designed a new device, the apparatus to mount individual electrodes (AMIE). The AMIE can be used in conjunction with Neuropixel probes. The AMIE holds the Neuropixel probe safely and securely in place for weeks until the experiment is finished, at which point it can be explanted and reused in another animal. This transformative technology will bring Neuropixel probes to an entirely new community who have, until now, been excluded from the opportunity to benefit from Neuropixel probes.

The paper is currently on bioRxiv (Juavinett et al. 2018), where it has generated considerable interest. Critically, we have made all the materials related to this device available to the community: The technical drawings, the methodological instructions, the photographs, and supporting code will, together, allow any researcher to rapidly adopt this new technology and begin to benefit from Neuropixel probes.

Population Dynamics of Neurons during Decision-Making

This work was done in collaboration with F. Najafi (University of Pennsylvania).

The starting point of this project is that mathematical models of decision-making have long relied on inhibitory neurons, but the role of such neurons has never been tested in vivo. This stands in stark contrast to the neural circuits field in which the responses of inhibitory neurons have been carefully studied, but mainly during passive conditions. Our approach joins the study of decision-making with the field of neural circuits: during two-choice perceptual decisions, we used two-photon imaging to measure neural activity in large neural populations. We ran these experiments in transgenic mice in which we could distinguish excitatory from inhibitory neurons. Surprisingly, inhibitory population activity predicted the animal's trial-by-trial choice as accurately as excitatory population activity. The ability to predict choice emerged jointly in the two populations during learning.

Our experimental observations, combined with our powerful network simulations, argue that inhibitory neurons form subnetworks with excitatory neurons according to their functional properties, conferring stability and robustness. This conclusion is in stark contrast to leading decision-making models. Our discovery of this architecture in mouse decision-making structures suggests that it may reflect a canonical computation that is relevant to many behaviors. This work is under review and is on bioRxiv (Najafi et al. 2018).

Brainwide Macrocircuits That Support Decision-Making

S. Musall, S. Gluf, M. Kaufman [in collaboration with R. Sun, Princeton]

Cognition and action are typically studied separately and are assumed to recruit largely nonoverlapping neural structures. Little is known about how cognitive and action processes interact, especially for actions beyond simple instructed movements used in laboratory behaviors. Here, we tracked a vast array of movements

in mice making decisions about auditory/visual stimuli, measured neural activity across the entire dorsal cortex, and developed a model to connect single-trial neural activity to movement and cognitive variables. Neural activity reflected both kinds of variables, but was dominated by movements. Using new methods, we then partitioned average neural activity into cognitive and movement components. This revealed that neurons with similar average responses could reflect utterly different combinations of cognitive and movement variables. Taken together, our observations argue that cognitive functions and movements can be tightly intertwined, and that during cognition, movements are a much higher priority than previously believed. This work is on bioRxiv (Musall et al. 2018) and was presented at the Society for Neuroscience Annual Meeting, at which it was selected for the Allen Institute's "What's hot in neuroscience in 2018?"

International Brain Laboratory

A. Urai

I cofounded this organization and secured funding in 2017 from the Wellcome Trust and the Simons Collaboration on the Global Brain. The International Brain Laboratory is a virtual laboratory, unifying a group of 21 experimental and theoretical neuroscience groups distributed across the world to understand the neural computations supporting decision making. During the past year, we presented our first data at the Society for Neuroscience Meeting and devoted considerable resources to building electrophysiology rigs.

PUBLICATIONS

Juavinett AL, Bekheet G, Churchland AK. 2018. Chronically-implanted Neuropixels probes enable high yield recordings in freely moving mice. bioRxiv doi:101101/406074

Musall S, Kaufman MT, Gluf S, Churchland AK. 2018. Movement-related activity dominates cortex during sensory-guided decision making. bioRxiv doi:10.1101/308288

Najafi F, Elsayed GF, Pnevmatikakis E, Cunningham J, Churchland AK. 2018. Inhibitory and excitatory populations in parietal cortex are equally selective for decision outcome in both novices and experts. bioRxiv doi:10.1101/354340

DYNAMICS OF CORTICAL NETWORKS FOR PERCEPTION AND COGNITION

T. Engel M. Genkin Y. Shi

Core brain functions—perception, attention, decision-making—emerge from complex patterns of neural activity coordinated within local microcircuits and across brain regions. This multiscale activity is driven by intricate networks of microscopic interactions between millions of neurons, with dynamics on timescales down to milliseconds. Yet, our current understanding of brain functions is largely defined by conventional single-neuron recordings and is grounded in a classical view that single neurons "represent" behaviorally relevant parameters in average firing rates. Within the last decade, rapid advances in massively parallel neural recording technologies enabled monitoring activity from hundreds of neurons simultaneously, presenting the opportunity to investigate how activity is orchestrated across large neural populations to drive behavior. Now, progress is mainly limited by the availability of computational methods to interpret high-dimensional, high-resolution neural activity data, and by the lack of theories linking dynamic features of neural activity to circuit computations and behavior.

The long-term goal of our research is to understand how coordinated activity in distributed neural circuitry gives rise to cognitive functions. To achieve this goal, we develop theory and computational methods, and apply them to analyze large-scale neural activity recordings from our experimental collaborators. Taking a step beyond the classical single-neuron level of understanding, we investigate a neural population dynamics hypothesis. In this hypothesis, behaviorally relevant parameters are encoded in a collective activity state of the entire population and computations are realized through orchestrated changes of this neural population state. In the last year, our research focused on modeling dynamic activity states of cortical populations during attention and decision-making. First, we developed network models that reveal how dynamic patterns of neural activity in the primate visual cortex are controlled during selective attention. Second, we developed a novel computational framework for discovering population dynamics from large-scale neural activity recordings, which we will apply to uncover how decision computations are realized by neural populations within local microcircuits and across cortical areas.

Theory of Correlated Fluctuations in Cortical Networks

Y. Shi [in collaboration with T. Moore, Stanford University; A. Thiele, Newcastle University]

The primate neocortex is the most evolved and specialized of all brain structures. Neocortical activity fluctuates endogenously, with much variability shared among neurons. These co-fluctuations are generally characterized as correlations between pairs of neurons, termed noise correlations. Noise correlations depend on anatomical dimensions, such as cortical depth and lateral distance, and they are also dynamically influenced by behavioral states, in particular, during spatial attention. Noise correlations echo the population activity on broader spatial and temporal scales, but the dynamics of population-wide fluctuations and their relationship to cortical functions remain unknown.

The goal of this project was to develop a theory that relates noise correlations to spatiotemporal dynamics of the underlying population activity. The theory integrates vast experimental data on noise correlations with our recent discovery that population activity in single cortical columns spontaneously transitions between phases of vigorous (On) and faint (Off) spiking, synchronously across cortical layers. These On–Off dynamics reflect global changes in cortical state associated with arousal, and they are also modulated at a local scale during selective attention. By creating a network model of interacting cortical columns, we aimed to elucidate how attentional inputs restructure spatiotemporal On–Off dynamics, leading to attention-related changes in noise correlations.

We developed a biophysically inspired network model of cortical columns with spatially structured

connectivity. In the model, single-column dynamics replicate spontaneous On–Off transitions, as observed in recordings from the visual cortex. The spatial connectivity correlates On–Off dynamics across columns. First, we found that the single-column model accurately predicts the broad distribution of attention-related changes in noise correlations within single columns, indicating that they largely arise from the On–Off dynamics. Second, the model of interacting columns mechanistically explains differences in attention-related changes in noise correlations at different lateral distances: Changes in lateral noise correlations largely arise from changes in the correlation length of the On–Off dynamics across columns. The model predicts that at short lateral distances, noise correlation changes decrease or increase with distance, when the correlation length increases or decreases, respectively. We validated model predictions using multielectrode array recordings of spiking activity from all layers of the primate visual cortex. This work provides a unifying framework that links network mechanisms shaping noise correlations to spatiotemporal population activity and underlying cortical circuit structure.

The future work will focus on analyzing and modeling neural activity recorded simultaneously from multiple columns within the same and across multiple cortical areas. We aim to elucidate how spontaneous activity is coordinated across cortical areas and to develop network models of how synchronous activity propagation across areas leads to conscious perception.

Understanding Decision Circuits through Dynamical Model Discovery

M. Genkin [in collaboration with K. Shenoy, Stanford; A. Churchland, CSHL]

Recently, massively parallel technologies enabled activity recordings from many neurons simultaneously, offering the opportunity to investigate how activity is orchestrated across neural populations to drive behavior. To reveal dynamic features in these large-scale data sets, computational methods are needed that can uncover neural population dynamics and identify how individual neurons contribute to the population activity. Existing methods rely on fitting ad hoc parametric models to data, which often leads to ambiguous model comparisons and estimation biases, limiting the potential of these methods for scientific discovery. To push these limits, we aimed to develop a broadly applicable, nonparametric inference framework for discovering population dynamics directly from the data without a priori model assumptions. Our nonparametric methods explore the entire space of all possible dynamics in search of the model consistent with the data, leading to a conceptual shift from model fitting to model discovery. Our framework reconstructs population dynamics on single trials and reveals heterogeneous contributions of single neurons to circuit-level computations.

In the first year of this project, we developed and implemented the computational framework for discovering population dynamics from neural activity recordings. Our framework extends latent dynamical models to a general form, where the population dynamics are governed by arbitrary dynamical-systems equations in which driving forces are directly optimized. Activity of each neuron is related to the instantaneous population state via unique firing-rate functions, which allow for heterogeneity of single-neuron responses. We developed a gradient-descent optimization, which traverses the space of dynamical models of increasing complexity. While more complex models better fit the training data, they typically exhibit features not present in the ground-truth dynamics. To control for model complexity, we adapted several regularization techniques, including cross-validation with early stopping, data augmentation, and stochastic gradient descent. We validated our framework by demonstrating that it accurately recovers ground-truth population dynamics and firing-rate functions on synthetically generated data sets. The salient features of the population dynamics can be reliably recovered from low amounts of data (as few as ~200–300 spikes). Our framework is also scalable to high amounts of data (~150,000 spikes) from many neurons with diverse firing-rate functions. This powerful tool can be used to reveal mechanisms of neural computations on single trials and is broadly applicable to different behaviors and neural circuits in the brain.

In the future, we will apply our methods to examine large-scale physiological recordings during decision-making to reveal how neural activity is coordinated to drive decisions and how functional heterogeneity of single-neuron responses aligns with anatomical organization of decision-making circuits. Specifically, we will investigate how two fundamental elements of decision-making circuits—excitatory versus inhibitory

cell types and laminar organization—contribute to functional heterogeneity of single-neuron responses.

PUBLICATIONS

Van Kempen J, Boyd M, Engel TA, Gieselmann MA, Thiele A. 2018. Coordination of cortical state within and between areas V1 and V4 of the Macaque. In *Abstracts presented at the Society for Neuroscience 48th Annual Meeting*, San Diego, CA.

Zeraati R, Engel TA, Levina A. 2018. Critical avalanches in a spatially structured model of cortical On–Off dynamics. In *Abstracts presented at the Bernstein Conference*, Berlin, Germany.

In Press

Genkin M, Engel TA. 2019. Non-parametric discovery of population dynamics from large-scale neural activity recordings. In *Abstracts presented at the American Physical Society Meeting*, Boston, MA (in press).

DISSECTING FUNCTIONS OF NEURONAL ION CHANNELS BY CRYO-EM AND ELECTROPHYSIOLOGY

H. Furukawa E. Chou J. Syrjanen
K. Michalski J.X. Wang
N. Simorowski

The research in the Furukawa lab attempts to establish a molecular basis for cellular signal transductions involved in neurotransmission and neuroplasticity in the mammalian brain with a scope to develop therapeutic compounds for treatment of neurological diseases and disorders, including schizophrenia, depression, stroke, and Alzheimer's disease. To achieve our goals, we conduct structural and functional studies on cell surface receptors and ion channels that regulate intracellular calcium signaling dependent on stimulation by voltage and/or neurotransmitters. Those ion channels regulate strength of neurotransmission, the fundamental process for neuronal communication. Dysfunction of the ion channels studied in our group is highly implicated in neurological disorders and diseases noted above. In the last decade, the Furukawa lab has been working on a class of ligand-gated ion channels called *N*-methyl-D-aspartate (NMDA) receptors. The abnormal activation of NMDA receptors is caused by a number of factors, including excessive transmission of neurotransmitters and point mutations in the ion channels that alter their functional properties. To unravel the molecular basis for normal and abnormal NMDA receptor activities, we use structural biology techniques, including X-ray crystallography and single-particle cryo-electron microscopy (cryo-EM), to determine three-dimensional (3D) atomic structures of the intact NMDA receptors, or fragments of them, and test structure-based mechanistic hypotheses by site-directed mutagenesis in combination with biochemical and biophysical techniques such as electrophysiology. In 2018, we advanced our understanding of how splice variants of NMDA receptors are capable of changing pH-sensitive inhibitory responses by a combination of cutting-edge cryo-EM at Cold Spring Harbor Laboratory and electrophysiology.

Proton Inhibition of NMDA Receptors Is Controlled by Alternative Splicing

Alternative splicing of the GluN1 subunit occurs both within the carboxy-terminal end (Exons 21 and 22) and the amino-terminal domain (ATD) at the extracellular region (Exon 5; Paoletti et al., *Nat Rev Neurosci 14:* 383 [2013]). A common way to modulate protein function is alternative splicing of the mRNA, and numerous examples of altered function through alternative splicing exist among a wide variety of ion channel families, including Shaker (Hoshi et al., *Neuron 7:* 547 [1991]) and BK (Chen et al., *J Biol Chem 280:* 33599 [2005]) potassium channels, voltage-gated sodium (Farmer et al., *PLoS One 7:* e41750 [2012]) and calcium (Castiglioni et al., *J Physiol 576:* 119 [2006]) channels, acid-sensing sodium channels (Bassler et al., *J Biol Chem 276:* 33782 [2001]), and transient receptor potential channels (Gracheva et al., *Nature 476:* 88 [2011]; Zhou et al., *Nat Commun 4:* 2399 [2013]), as well as iGluRs (Lerma et al., *Physiol Rev 81:* 971 [2001]; Schiffer et al., *Neuron 19:* 1141 [1997]; Palmer et al., *Pharmacol Rev 57:* 253 [2005]; Penn et al., *Neuron 76:* 503 [2012]; Sommer et al., *Science 249:* 1580 [1990]). One of the hallmark functions in NMDA receptors is that the splice variants of the GluN1 subunit containing the 21 amino acid residues encoded by Exon 5 in the ATD have significantly reduced sensitivity to protons and faster deactivation speed (Fig. 1; Adams et al., *Mol Genet Metab 113:* 161 [2014]; Rumbaugh et al., *J Neurophysiol 83:* 1300 [2000]; Traynelis et al., *Science 268:* 873 [1995]; Vance et al., *J Physiol 590:* 3857 [2012]), indicating that those splice variants contribute to excitatory postsynaptic current and elicit calcium signaling differently, both under normal conditions and in the acidified environment created locally by high-frequency neuronal firing, stroke, and seizure (Kaku et al., *Science 260:*

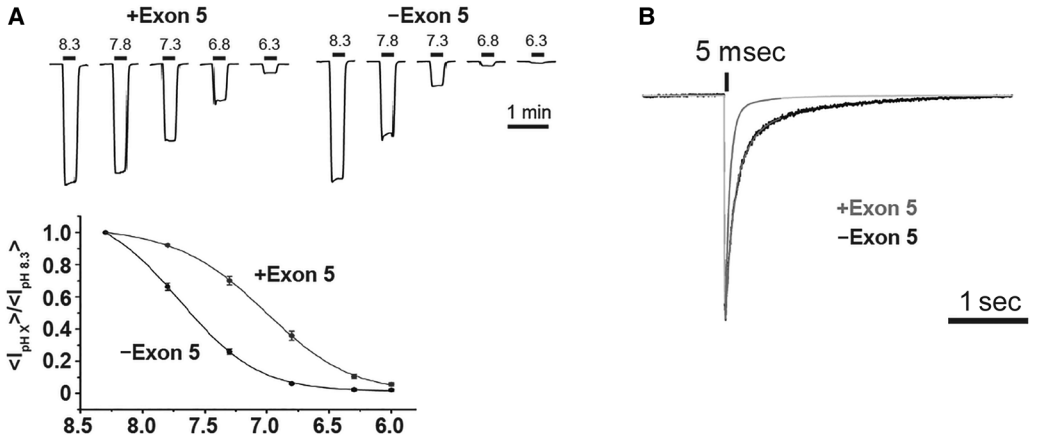
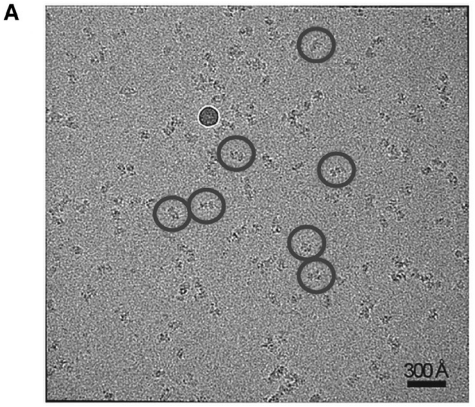


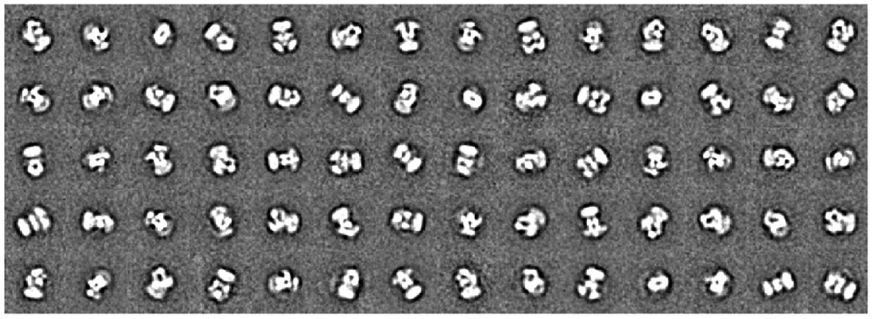
Figure 1. Exon 5 in GluN1 changes ion channel properties in *N*-methyl-D-aspartate (NMDA) receptors. Two-electrode voltage clamp (TEVC) recording from complementary RNA (cRNA)-injected *Xenopus* oocytes (*A*) and whole-cell patch clamp recording from DNA transfected HEK293 cells (*B*) of the GluN1-4a-GluN2B (–Exon 5) and GluN1-4b-GluN2B (+Exon 5) NMDA receptors showing reduced proton sensitivity and faster deactivation speed elicited by Exon 5, respectively. Error bars in *A* represent ±SD for data obtained from 15 and 17 different oocytes for the –Exon 5 and +Exon 5, respectively.

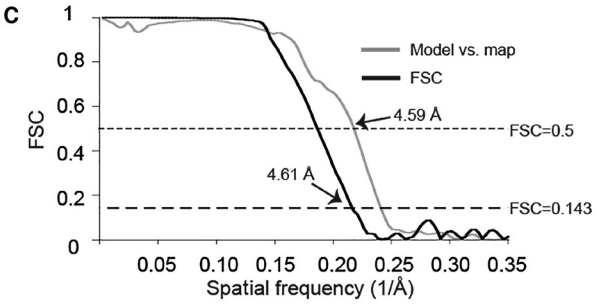
1516 [1993]). The Exon 5 sequence is conserved throughout vertebrate GluN1 subunits as far back as cartilaginous fishes and hemichordates, which suggests that the location and composition of Exon 5 is evolutionarily advantageous for the maintenance of proper neurological functions. Despite the importance of this classical paradigm in basic neuroscience and pathology, the molecular mechanism underlying functional alteration by Exon 5 has been enigmatic because of the lack of a structure of the intact NMDA receptor with Exon 5. Thus, we decided to determine the structure of GluN1b-GluN2B NMDA receptors that contain the Exon 5–encoded motif.

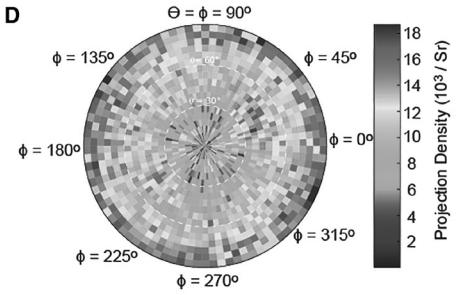
Cryo-EM Structure Locates Exon 5–Encoded Motif at Domain and Subunit Interfaces

To define the structure and location of the Exon 5–encoded motif, we performed single-particle cryo-EM on the intact rat GluN1-GluN2B NMDA receptor containing the Exon 5 sequence (GluN1b-GluN2B$_{EMX}$) (Figs. 2 and 3) and prepared the samples with glycine and glutamate bound in the ligand-binding domains (LBDs), ifenprodil in the ATD, and MK-801 in the transmembrane domain (TMD) ion channel. This modified construct formed the GluN1-GluN2B NMDA receptors that retained functional properties, including activation by glycine and glutamate and inhibition by an allosteric modulator (ifenprodil or zinc), antagonists (5,7 dichlorokynurenic acid or D-AP5), or channel blockers (magnesium and MK-801; Fig. 3). 3D classification resulted in one structure that resolves to ~4.6 Å resolution as estimated by Fourier shell correlation (FSC) using the 0.143 cutoff (Fig. 2), enabling us to visualize density for bulky aromatic residues and Asn-linked glycosylation and permitting building of molecular models with confidence. The cryo-EM structure clearly showed the heterotetrameric assembly of the GluN1b-GluN2B NMDA receptors. The bilobed structures of the GluN2B ATD, the GluN1 LBD, and the GluN2B LBD in the current cryo-EM structure are in the closed conformation, occupied by ifenprodil, glycine, and glutamate, consistent with crystal structures of the isolated domains (Jespersen et al., *Neuron 81: 366* [2014]; Karakas et al., *Nature 475:* 249 [2011]). At the TMD, the ion channel pore is blocked by MK-801, as illustrated by the cryo-EM density. The cryo-EM structure shows ordered density for eight residues (GluN1b Tyr204-Lys211) encoded by Exon 5 at the ATD–LBD interdomain interface, which "caps" the GluN1b-GluN2B subunit interface of the LBD through interactions between GluN1b Lys190 and GluN2B Tyr507 and between GluN1b Lys211 and GluN1b Asp786 (Fig. 4B). The regions surrounding GluN1b Lys190, GluN2B Lys211, and GlwuN2B Tyr507 are devoid of missense variants









E

Data Collection/Processing		Model Statistics	
Microscope	Titan Krios	Number of atoms	
Voltage (kV)	300	Protein	20,885
Defocus range (μm)	-1.5 to -3.0	NAG	280
Exposure time (sec)	21		
Dose (e^-/Å^2/sec)	4.67	R.M.S. deviations	
Number of frames	70	Bond lengths (Å)	0.01
Pixel size (Å/pixel)	1.31	Bond angles (°)	0.83
Number of particles	73,790		
Resolution (Å)	4.6	Ramachandran	
		Favored (%)	90.54
		Allowed (%)	9.27
		Disallowed (%)	0.19

Figure 2. Single-particle analysis of the GluN1b-GluN2B. (*A*) Representative motion-corrected electron micrograph at 22,500 magnification. Particles with different orientations are highlighted by green circles. (*B*) Representative 2D class averages. (*C*) Fourier shell correlation (FSC) curves for masked, and model versus cryo-electron microscopy (EM) map. (*D*) Orientation plot showing the distribution of Euler angles assigned to all particles contributing to 3D reconstruction with an occupancy of at least 80%. (*E*) Data collection details and model statistics for the current structure.

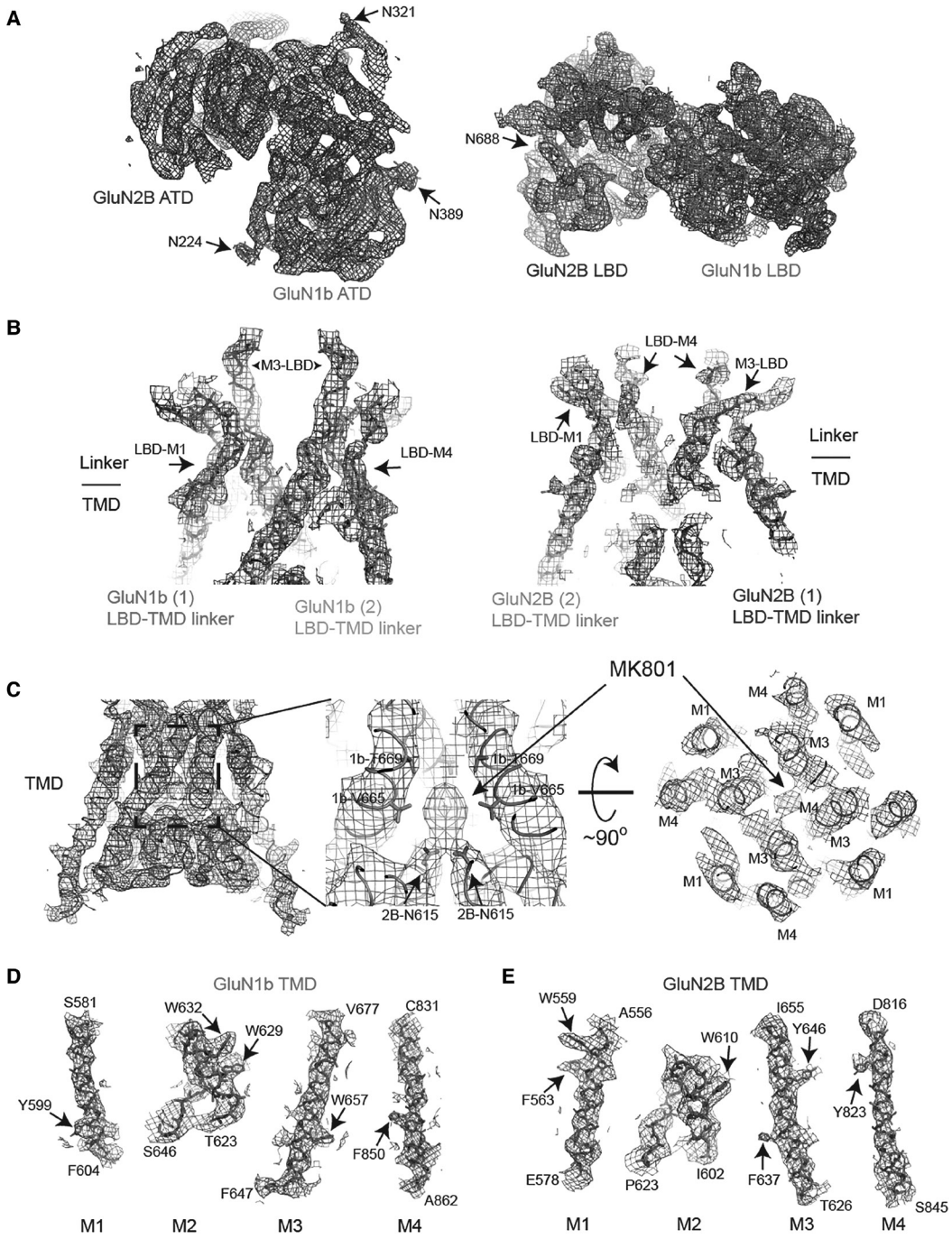


Figure 3. Representative cryo-EM density of the GluN1b-GluN2B. (*A*) The extracellular domains, amino-terminal domain (ATD) (*left*) and LBD (*right*) from the GluN1b and GluN2B subunits. Clear densities for Asn-linked glycosylation for GluN1b Asn224, Asn321, and Asn389 and GluN2B Asn688 are observed. (*B*) All of the LBD-TMD linkers from the GluN1b and GluN2B subunits are visible. (*C*) The TMD viewed from the "side" and the "cytoplasm." The density in the middle of the ion channel surrounded by the M3 helices likely represents MK-801 (arrows). Cryo-EM density was especially strong for GluN2B Asn615, which appears to be directly interacting with MK-801. (*D*,*E*) In contrast to previous structures, our current cryo-EM structure has excellent density at the TMD, from which we are even able to model a number of amino acid side chains.

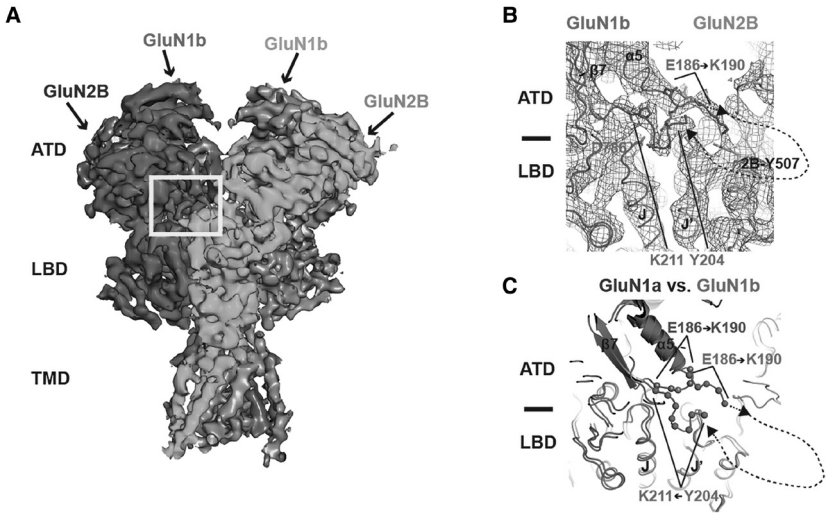


Figure 4. Cryo-electron microscopy (cryo-EM) structure identifies the Exon 5–encoded motif at domain and subunit interfaces. (*A*) Cryo-EM structure of the intact GluN1b-GluN2B NMDA receptor at ~4.6 Å. (*B*) Highlighted region from *A,* illustrating cryo-EM density for residues of the Exon 5 motif at the ATD–LBD interface. (*C*) Overlay of models from the GluN1a-GluN2B and GluN1b-GluN2B intact NMDA receptors. Note the large shift in position of GluN1 Lys190 between the GluN1a and GluN1b isoforms.

in humans, suggesting they play important roles in protein function and may be under strong selection (Lek et al., *Nature 536:* 285 [2016]; Ogden et al., *PLoS Genet 13:* e1006536 [2017]). Remarkably, the region between GluN1b Glu186 and Lys190, which is not encoded by Exon 5, has an altered architecture compared to that in the GluN1a-GluN2B NMDA receptor lacking Exon 5 (Fig. 4C; Karakas and Furukawa, *Science 344:* 992 [2014]). Specifically, those five residues are extended toward the ATD–LBD linker of the GluN2B subunit rather than turning back to GluN1 β7, as in GluN1a, thereby creating the GluN1b-specific contacts with the GluN2B subunit, including those between GluN1b Lys190 and GluN2B Tyr507 and between GluN1b Lys211 and GluN2B Tyr507 (Fig. 4B). The structure of the region immediately surrounding the Exon 5 motif is similar to that of the allosterically inhibited receptor (PDB code: 4PE5), the nonactive receptor (PDB code: 5FXH and 5FXI), and the active receptor (PDB code: 5FXG); therefore, the pattern of interactions between the Exon 5 motif and other domains is likely independent of the functional states of the receptors (Karakas and Furukawa, *Science 344:* 992 [2014]; Tajima et al., *Nature 534:* 63 [2016]).

Interactions between the Exon 5 Motif and the LBDs Control Proton Sensitivity

Based on the current cryo-EM structure, we explored the molecular mechanism for reduced proton sensitivity by Exon 5. We conducted mutagenesis in the context of the full-length GluN1-4b-GluN2B NMDA receptor, including disrupting interactions between the Exon 5 motif and the GluN1 LBD and/or between the Exon 5 motif and the GluN2B LBD, charge reversal of the three negatively charged residues encoded by Exon 5, and truncations of the disordered region (Fig. 5A–C). Potency of proton inhibition was measured by detecting whole-cell macroscopic currents at various pH points using two-electrode voltage clamp (TEVC) in *Xenopus laevis* oocytes. Disrupting the interaction between the Exon 5 motif and the GluN1 LBD (by mutating GluN1-4b Lys190Ala or GluN1-4b Asp786Ala) or between the Exon 5 motif and the GluN2B LBD (by mutating GluN1-4b Lys211Ala or GluN2B Tyr507Ala) increased the IC_{50} values for proton inhibition from pH 7.04 to pH ~7.3 and ~7.4, respectively, but did not restore proton sensitivity to the level observed for the Exon 5–lacking GluN1-4a-GluN2B NMDA receptor (IC_{50} = pH 7.72) (Fig. 5A).

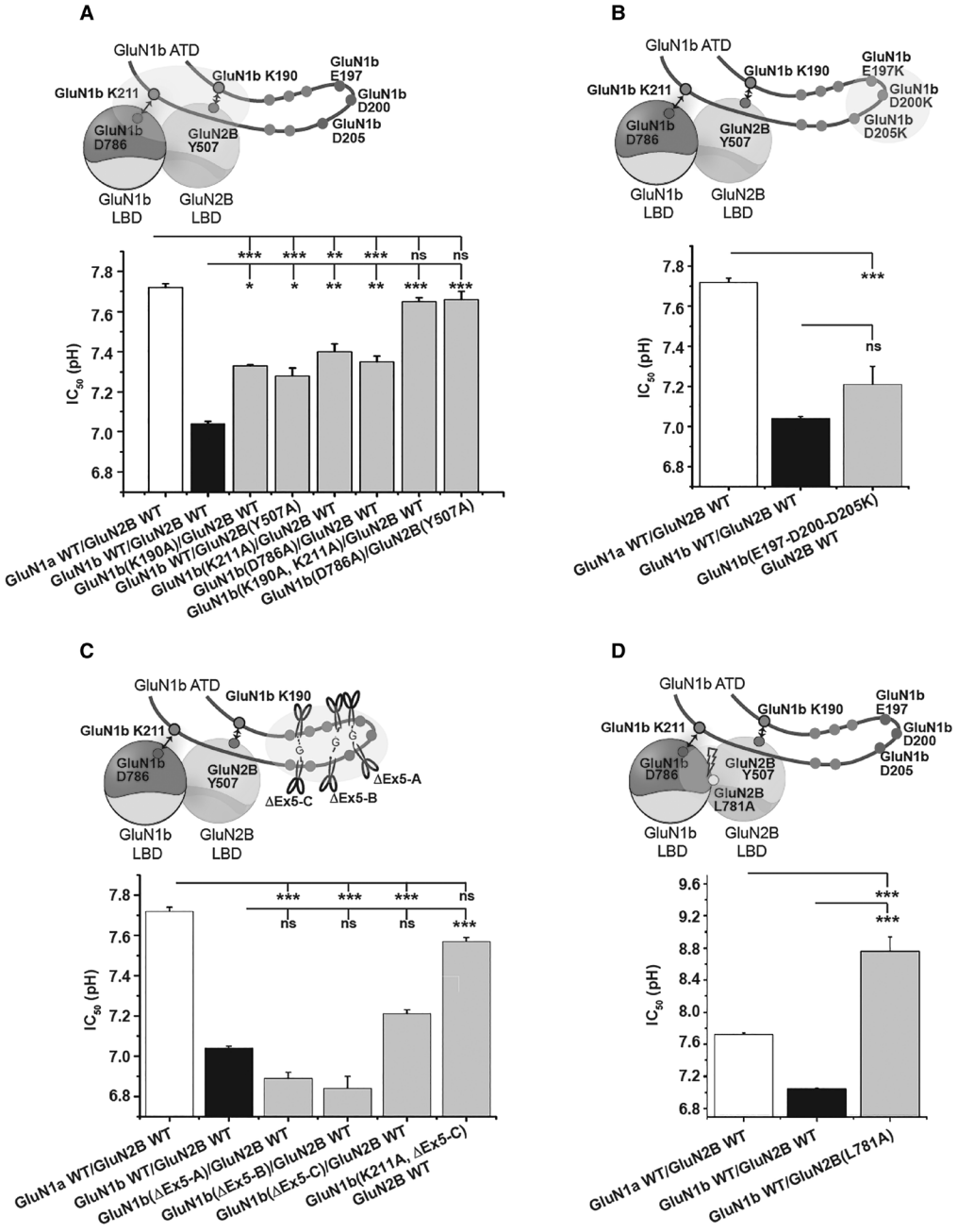


Figure 5. Interactions between the Exon 5 motif and the LBDs control proton sensitivity. Here, we used four types of mutations in the context of TEVC experiments to explore several potential effects of the Exon 5 motif: (*A*) disruption of the direct Exon 5–LBD interactions, (*B*) reversal of negative charges to positives, (*C*) alteration of local structure of the Exon 5 motif by truncations, and (*D*) disruption of the heterodimeric interface between the GluN1 and GluN2B LBDs. Shown in gray lines and two-tone spheres are the Exon 5 motif and the GluN1b-GluN2B LBD heterodimers, respectively. The lightning bolt and scissors indicate disruption of interactions and truncations, respectively. The areas of engineered mutations are highlighted by transparent ovals. IC_{50} values of proton inhibition for the wild-type GluN1-4a-GluN2B and GluN1-4b-GluN2B NMDA receptors and the mutant GluN1-4b-GluN2B NMDA receptors were estimated by recording macroscopic current at various pHs by TEVC from at least five oocytes per construct. (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$, (ns) not significant, as determined by one-way analysis of variance (ANOVA) with Tukey's test. Bar graphs represent mean ± standard error of mean (SEM). All of the TEVC recordings look similar to the ones in Figure 1A.

When we disrupted both of the two interdomain interactions above with the GluN1-4b-Lys190Ala/Lys211Ala mutations, and similarly with GluN1-4b-Asp786Ala/GluN2B-Tyr507Ala mutations, the IC_{50} values for these respective mutant combinations shifted to pH 7.65 and 7.66, which is statistically indistinguishable from the wild-type GluN1-4a-GluN2B NMDA receptor (Fig. 5A). Contrary to the large effect observed by disrupting the Exon 5–LBD interactions, charge reversal of negatively charged residues (Fig. 5B) and truncations of the disordered region (Fig. 5B) had little or no effect on receptor proton sensitivity. The effect of truncation is observed only in the most extreme construct (ΔExon5-C) together with GluN1-4b-Lys211Ala. These results imply that the Exon 5 motif reduces proton sensitivity of the ion channel by affecting the subunit interface between the GluN1b LBD and the GluN2B LBD. To further test this hypothesis, we disrupted the GluN1b-GluN2B LBD heterodimer interface with the GluN2B Leu781Ala mutation and observed a significant increase in proton sensitivity (Fig. 5D). Overall, these results agree with the concept that the GluN1-GluN2B LBD dimer interface is a critical region for controlling proton sensitivity, and the Exon 5 motif reduces proton sensitivity by stabilizing the GluN1-GluN2B LBD dimer interface by interacting with both the GluN1 LBD and the GluN2B LBD. Disrupting the similar LBD subunit interface in the GluN1-GluN2A NMDA receptors was also shown to result in increased proton sensitivity (Gielen et al., *Neuron 57:* 80 [2008]), indicating that the GluN1-GluN2 LBD interface is the common locus for controlling the proton sensitivity in both GluN2A- and GluN2B-containing NMDA receptors. Taken together, our structural and functional data show that the ATD–LBD interface is an important locus for regulating ion channel function, including proton sensitivity and deactivation rate in NMDA receptors. The Exon 5 motif acts as a ligand naturally incorporated into the primary sequence in order to control the GluN1-GluN2B LBD intersubunit interface, which in turn regulates ion channel activity at the TMD.

PUBLICATIONS

Hansen KB, Yi F, Perszyk RE, Furukawa H, Wollmuth LP, Gibb AJ, Traynelis SF. 2018. Structure, function, and allosteric modulation of NMDA receptors. *J Gen Physiol* **150:** 1081–1105.

Regan MC, Grant T, McDaniel MJ, Karakas E, Zhang J, Traynelis SF, Grigorieff N, Furukawa H. 2018. Structural mechanism of functional modulation by gene splicing in NMDA receptors. *Neuron* **98:** 521–529.e3.

Wu Q, Regan M, Furukawa H, Jackson D. 2018. Role of heterotrimeric Gα proteins in maize development and enhancement of agronomic traits. *PLoS Genet* **14:** e1007374.

In Press

Regan MC, Zhu Z, Yuan H, Myers SJ, Menaldino DS, Tahirovic YA, Liotta DC, Traynelis SF, Furukawa H. 2019. Structural elements of a pH-sensitive inhibitor binding site in NMDA receptors. *Nat Commun* **10:** 321. doi:10.1038/S41467-019-08291-1

Wang JX, Furukawa H. 2019. Dissecting diverse functions of NMDA receptors by structural biology. *Curr Opin Struct Biol* **54:** 34–42. doi:10.1016/j.sbi.2018.12.009

ASSEMBLY AND ORGANIZATION OF CORTICAL CIRCUITS UNDERLYING MOTOR CONTROL

J. Huang X. An, S.W. Galbavy, J. Hatfield, D. Huilgol, K. Kelly, G. Kim, H. Kondo, Yi. Li, K. Matho, H. Mohan, A. Paul, Y. Qian, R. Raudales, P. Wu

The neocortex contains a massive interacting network of multidozen functional areas that integrates sensory information with emotional drive and internal goals to guide intelligent behavior. Despite its immense cellular complexity, fundamental aspects of cortical organization (e.g., major interareal processing streams, output channels, and local circuit templates) are stereotypic among individuals of the same species and conserved across mammals. These basic network scaffolds and circuit modules are implemented by a large set of cardinal types of glutamatergic projection neurons (forming myriad processing streams and output channels) and GABAergic interneurons (forming local connectivity motifs), which are reliably generated through developmental programs rooted in the genome. These self-evident first principles provide the basis for a genetic dissection of cortical assembly and organization through systematic screening and fate mapping of its major cellular components—a powerful approach that has been successfully applied to neural circuits in *Drosophila* and the vertebrate retina and spinal cord. The overarching theme of our research is to understand general principles underlying the assembly and organization of cortical circuitry by integrating multifaceted studies of cortical neuron types in the functional context of motor control in the mouse. To achieve this, we combine genetic, genomic, anatomical, physiological, and behavioral approaches. A unique aspect of our research program is aiming to link cell types to circuit and behavior on the one hand and to developmental and evolutionary trajectory on the other.

Cell Type Genetic Tools

Specific and systematic experimental access to neuronal subpopulations is prerequisite to explore circuit organization, function, and assembly. We are leading a Center for Mouse Brain Cell Atlas in the BRAIN Initiative Cell Census Network (BICCN). We continue to systematically generate cell type tools targeting forebrain projection neurons, including cortical, striatal, and thalamic projection neurons. These cell type tools will greatly accelerate studying the cortical–striatal–thalamic system, the most prominent network in the mammalian brain, that mediates a wide range of sensory, motor, emotional, and cognitive functions.

Developmental Lineage Mechanisms Underlying Striatal Compartments and Circuit Organization

The circuitry of the striatum is characterized by two organizational plans: the division into striosome and matrix compartments, thought to mediate evaluation and action, and the direct and indirect pathways, thought to promote or suppress behavior. The developmental origins of these organizations and their developmental relationships are unknown, leaving a conceptual gap in understanding the cortico-basal ganglia system. Through genetic fate mapping, we show that striosome-matrix compartmentalization arises from a lineage program embedded in lateral ganglionic eminence radial glial progenitors that mediate neurogenesis through two distinct types of intermediate progenitors (IPs). The early phase of this program produces striosomal spiny projection neurons (SPNs) through fate-restricted apical intermediate progenitors (aIPSs) with limited capacity; the late phase produces matrix SPNs through fate-restricted basal intermediate progenitors (bIPMs) with expanded capacity. Notably, direct and indirect pathway SPNs arise within both aIPS and bIPM pools, suggesting

that striosome-matrix architecture is the fundamental organizational plan of basal ganglia circuitry organization.

Genetic Single-Neuron Anatomy Reveals Fine Granularity of Cortical Axo-Axonic Cells

Parsing diverse nerve cells into biological types is necessary for understanding neural circuit organization. Morphology is an intuitive criterion for neuronal classification and a proxy of connectivity, but morphological diversity and variability often preclude resolving the granularity of neuron types. Combining genetic labeling with high-resolution, large-volume light microscopy, we established a single-neuron anatomy platform that resolves, registers, and quantifies complete neuron morphologies in the mouse brain. We discovered that cortical axo-axonic cells (AACs), a cardinal GABAergic interneuron type that controls pyramidal neuron (PyN) spiking at axon initial segment, consist of multiple subtypes distinguished by highly laminar-specific soma position, dendritic, and axonal arborization patterns. Whereas the laminar arrangements of AAC dendrites reflect differential recruitment by input streams, the laminar distribution and local geometry of AAC axons enables differential innervation of PyN ensembles. This platform will facilitate genetic targeted, high-resolution, and scalable single-neuron anatomy in the mouse brain.

Activity-Dependent Chandelier Cell Elimination Shapes Binocularity in Primary Visual Cortex

In mammalian primary visual cortex (V1), binocular responses that integrate the left–right visual scene derive from convergent ipsi- and contralateral retinal-geniculate inputs, as well as trans-callosal projections between the two hemispheres; the underlying developmental mechanisms remain incompletely understood. Using genetic methods in mice, we found that during the second postnatal week before eye opening, retinal and callosal activities drive massive apoptosis of GABAergic chandelier cells (ChCs, which control PyN firing at axon initial segment) at V1 binocular region. Blockade of ChC elimination results in a contralateral-dominated V1 and deficient binocular vision. As correlated activities within and between retinas convey organization of the visual field, their regulation of ChC density through the trans-callosal pathway may prime a nascent binocular territory for subsequent experience-driven tuning.

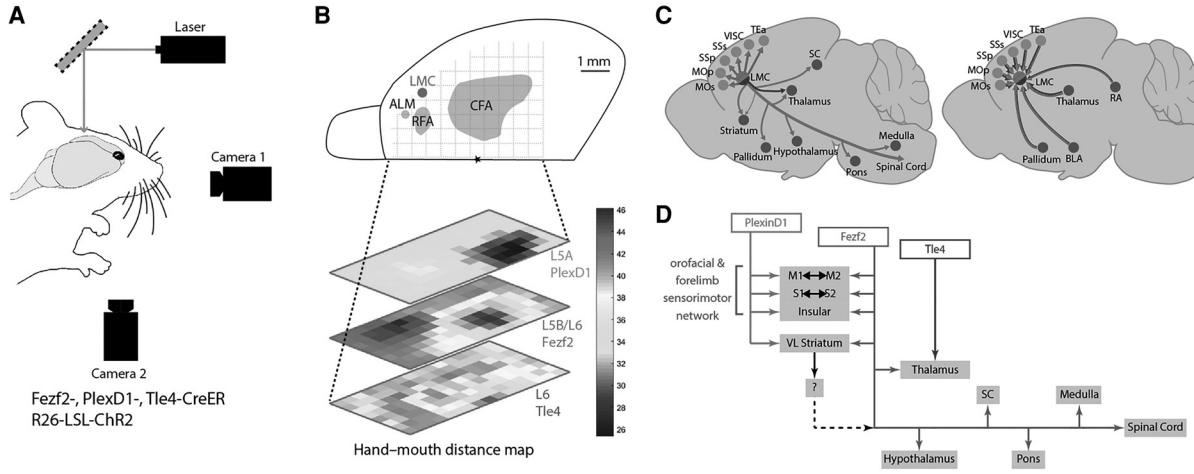


Figure 1. (*A*) Optogenetic activation screen of PyN subpopulations in forelimb and orofacial movements. (*B*) PlexD1, Fezf2, Tle4 PyNs construct different and related sets of cortical "motor maps." The lateral motor cortex (LMC) is indicated in relation to other known motor areas. (*C*) The output and input connectivity patterns of PlexD1, Fezf2, and Tle4 PyNs in LMC revealed by anterograde and retrograde viral tracing. (*D*) Summary schematic showing that the three PyN types in LMC form a cortico-striato-thalamic subnetwork that emphasizes the forelimb–orofacial sensorimotor system, with subcortical output mainly from Fezf2 neurons.

Cortical Circuits Controlling Complex Movement and Food Manipulation

In rodents and primates, food handling and manipulation using synergistic hand–mouth maneuvering is necessary for feeding and requires online sensorimotor coordination, as well as dexterous motor control; the underlying neural circuits are not understood. We performed a systematic optogenetic screen of PyN projection types and cortical areas that induce forelimb and orofacial movements in the mouse. We define a rostral forelimb orofacial (RFO) area (centered around anterior 1.7 mm and lateral 2.3 mm), in which activation of Fezf2 cortico-fugal PyNs (PyNFezf2) and PlexinD-1cortico-striatal PyNs (PyNPlexD1) induce highly coordinated forelimb and orofacial movements resembling feeding. We further reveal a caudal forelimb orofacial (CFO) area (centered around anterior 0.3 mm and lateral 3 mm) in which activation of Tle4 cortico-thalamic PyNs (PyNTle4) induce similar action, an effect dependent on functional RFO. Antero- and retrograde tracing from these PyN types in RFO and CFO depict a highly connected cortical network involving primary and secondary sensory and motor areas of the forelimb and orofacial regions. This cortical network is embedded in the cortico-striatal-thalamic system, with outputs to numerous subcortical targets from midbrain, hypothalamus, to pons and spinal cord. Chemo- and optogenetic inactivation of RFO and/or CFO PyNs impairs hand maneuvering and hand-mouth synergy during food handing, manipulation, and eating in head-fixed and free-moving mice. These results begin to implicate specific neuron types, cortical circuits, and brain systems in sensorimotor coordination for object manipulation.

PUBLICATIONS

Crow M, Paul A, Ballouz S, Huang ZJ, Gillis J. 2018. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat Commun* **9:** 884.

He M, Huang ZJ. 2018. Genetic approaches to access cell types in mammalian nervous systems. *Curr Opin Neurobiol* **50:** 109–118.

Kelly SM, Raudales R, He M, Lee, JH, Kim Y, Gibb LG, Wu P, Matho K, Osten P, Graybiel, AM, et al. 2018. Radial glial lineage progression and differential intermediate progenitor amplification underlie striatal compartments and circuit organization. *Neuron* **99:** 345–361.

Paul A, Huang ZJ. 2018. Single-cell RNA sequencing of fluorescently labeled mouse neurons using manual sorting and double in vitro transcription with absolute counts sequencing (DIVA-seq). *J Vis Exp* (140), e58690. doi:10.3791/58690

In Press

Huang ZJ, Paul A. 2019. Diversity of GABAergic neurons and diversification of communication modules in cortical networks. bioRxiv doi:10.1101/490797

Wang X, Tuccarione J, Jiang S, Yin F, Wang B-S, Wang D, Jia Y, Jia X, Li Y, Yang T, et al. 2019. Genetic single neuron anatomy reveals fine granularity of cortical axo-axonic cells. *Cell Rep* **26:** 3145–3159. doi:10.1016/j.cellrep.2019.02.040

NEUROBIOLOGY OF COGNITION AND DECISION-MAKING

A. Kepecs

A. Arora, D. Bonda, M. Bosc, Q. Chevy, E. Demir, E. Gibson, T.S. Gouvea, Q.A. Li, S. Li, P. Masset, T. Ott, T. Pinkhasov, L. Ramirez Sanchez, S. Ren, K. Schmack, A. Siebels, S. Starosta, J.F. Sturgill, M. Wulf, W. Zhong

Our long-term goal is to reverse engineer the computational and neurobiological processes underlying cognition and decision-making and apply these insights to biological psychiatry. We start with quantifiable behavioral tasks, for both humans and rodents, that enable us to examine the computational principles responsible for cognition. We probe the neural basis of these processes in rodents using state-of-the-art electrophysiological, imaging, and optogenetic techniques to establish the underlying neural processes. Given the complexity of behavior and the dynamics of the neural networks producing it, we also develop new algorithms for analyzing data and models to help us interpret them. By understanding how brains accomplish cognitive tasks, often beyond the capacity of current machine learning algorithms, we also expect to uncover new computational principles. Our goal is to build a bridge from our animal studies to psychiatric disorders and, by identifying the neural processes underlying behavior in "humanized" mouse models of cognition, provide insights into what goes awry in the brain during mental illness. Ultimately we hope these insights will lead to the development of novel therapeutic strategies for psychiatric disorders such as addiction, major depression, schizophrenia, and autism spectrum disorder.

Confidence in Rats, Humans, Brains, and Statistics

This work was done in collaboration with M. Lagler and T. Klausberger (Medizinische Universität Wien), A. Lak (UCL, University of Oxford), and J. Hirokawa (Doshisha University).

Every decision we make is accompanied by a sense of confidence about its likely outcome. This sense informs subsequent behavior, such as investing more—whether time, effort, or money—when the payoff is more certain. Conversely, the pathological misevaluation of confidence contributes to a wide range of neuropsychiatric conditions, including anxiety, obsessive–compulsive disorder, and addiction. Our long-term goal is to understand how the brain implements confidence judgments and acts on these.

We have developed a set of behavioral tasks and a theoretical framework that rigorously translates the psychological concept of confidence into a formally defined decision variable. Using this approach we identified orbitofrontal cortex (OFC) neurons that encode the confidence associated with a perceptual decision. Further, we have mathematically derived a mathematical framework for decision confidence from first principles of statistics (Hangya et al., *Neural Comput 28:* 1 [2016]). We showed that key properties of statistical decision confidence match human self-reported confidence (Sanders et al., *Neuron 90:* 499 [2016]), providing a deep link between objective and subjective notions of confidence.

Previously we showed that the firing of many OFC neurons encodes statistical confidence about olfactory-discrimination decisions, and OFC inactivation specifically impairs rats' ability to optimally invest time waiting for reward. However, a neural representation of abstract confidence should not just (i) reflect a confidence computation, as we have shown, but also (ii) predict multiple confidence-guided behaviors and (iii) be independent of the source of information used to make a choice (i.e., independent of sensory modality). Using a new task design, we now have been able to show that single OFC neurons encode statistical confidence and predict two confidence-guided behaviors (trial-by-trial time investment and confidence-guided updating) and do so irrespective of whether the sensory discrimination was olfactory or auditory. Therefore OFC appears to contain a modality-general representation of confidence that could provide an information source–independent probability estimate, useful for confidence-driven adaptive behaviors, such as learning and time investment.

Confidence or uncertainty has long been suggested to modulate the degree of learning. We identified a novel form of reinforcement learning during perceptual decisions that depends on the confidence of past sensory judgments. We showed that these outcome-dependent biases depend on the strength of past sensory evidence, suggesting that they are consequences of confidence-guided updating of choice strategy. We illustrate that this form of choice updating is a widespread behavioral phenomenon that can be observed across various perceptual decision-making paradigms in mice, rats, and humans. This trial-to-trial choice bias was also present in different sensory modalities and transferred across modalities in an interleaved auditory/olfactory choice task. To explain these observations, we have formulated a class of reinforcement learning models that compute prediction errors scaled by decision confidence and produce confidence-guided updating of choice bias.

Neurons in orbitofrontal cortex, like in other regions of frontal cortex, display baffling complexity, responding to a mixture of sensory, motor, and other variables. We developed a new approach to understand the representational content and architectural logic of higher-order cortical areas. We found that discrete groups of OFC neurons encode distinct decision variables and these categorical representations map directly onto decision variables, such as reward size, decision confidence, and integrated value, in a choice model explaining our task. This suggests that like sensory neurons, frontal neurons form a sparse and overcomplete population representation aligned to the natural statistics of the world—in this case spanning the space of decision variables required for optimal behavior.

Beyond our rodent work, we are now in a position to use these quantitative measures of decision confidence in humans, through collaborations, with neuroimaging and genetic approaches. Moreover, by showing that a single confidence measure is applicable to humans and rodents, our results strengthen the case for using the rat as a model system for translational studies in cognition and psychiatry.

Cell Type–Specific Cortical Architecture

This work was done in collaboration with Z.J. Huang and J. Tucciarone (CSHL), Z. Szadai and B. Rozsa (Hungarian Academy of Sciences), and A. Nectow (Princeton).

Reverse engineering cortical algorithms requires understanding the contributions of a multitude of different cortical neuron types, each with unique properties and connectivity. Thus, to identify neural circuit implementations of computations among diverse cortical elements, we must couple sophisticated tasks with targeted, high-throughput monitoring and manipulation of neural circuits at cell type resolution. Because there is no consensus definition of neural types, we use complementary techniques based on genetic and projection targeting with optogenetics-assisted cell type identification. We continued our efforts to expose the cell type–specific logic of cortex by focusing on a few inhibitory and projection cell types.

Chandelier cells (ChCs) constitute perhaps the most unique GABAergic interneurons in cortex. They specialize in innervating the axon initial segment of excitatory pyramidal neurons, the site for action potential generation, yet it remains unclear whether they function to inhibit their targets. Taking advantage of a genetic and viral approach developed in the Huang lab, we are able to target a subgroup of ChCs in PFC. By combining electrical stimulation at the basolateral amygdala (BLA) and extracellular recordings in PFC, we are aiming to determine whether ChCs inhibit or excite their BLA-projecting pyramidal neurons. We are also developing a behavioral paradigm for testing approach–avoidance conflict in head-fixed mice that engage these circuits. We aim to identify ChCs' impact on their local circuit and evaluate when they are recruited during behavior.

We have recently identified a disinhibitory cortical circuit motif that appears to be a conduit for fast neuromodulatory action in the cortex. This circuit is controlled by a class of inhibitory interneurons that express vasoactive intestinal polypeptide (VIP) and inhibit other interneurons, thereby disinhibiting a subpopulation of principal neurons. Functionally, we showed that VIP interneurons in the auditory cortex are recruited in response to specific reinforcement signals such as reward and punishment.

To explore the generality of these observations across cortex regions, we are collaborating with B. Rozsa (KOKI, Hungarian Academy of Sciences), using a state-of-the-art 3D random-access AOD two-photon imaging system to record the sparse VIP population across large regions. We found most VIP neurons are activated by reward and punishment across multiple cortical regions, suggesting that their behavioral recruitment has a cortex-wide function in reinforcement learning.

We have also begun to record VIP neurons in a cued-outcome task to determine what aspects of reinforcement they signal. Do they respond simply to the delivery of primary reinforcers, reinforcement prediction errors, or reinforcement surprise? We found that cue responses emerge with learning and are proportional to outcome value, suggesting a more sophisticated role in reinforcement learning. By establishing the detailed circumstances under which VIP interneurons are recruited and identifying their generality and circuit mechanisms, we expect that these studies will reveal fundamental principles about cortical microcircuits that are applicable across cortical regions.

To understand OFC neuron types, we use two complementary approaches. First, we target specific projection neurons using retrograde viruses and use optogenetic stimulation to identify these in electrophysiological recordings. Using this technique, we have found that OFC projections to ventral striatum show a characteristic response in which negative value signals are sustained throughout the intertrial interval to the beginning of the subsequent trial. Second, in collaboration with the Klausberger laboratory, we use juxtacellular labeling to target neurons based on their functional response profiles. Once labeled, neurons are subjected to detailed ex vivo analysis of the axonal projection patterns. Using these techniques, we have begun to record and identify OFC neurons that specifically signal the confidence-dependent time investments. We expect that the combination of juxtacellular and optogenetically identified extracellular recordings will enable us to reverse engineer the cell type–specific circuit logic of orbitofrontal cortex.

Computational Logic of Neuromodulation

This work was done in collaboration with B. Hangya (Hungarian Academy of Sciences) and Y. Li (Peking University).

Neuromodulators constitute central brain systems with cell bodies located in deep brain areas that project across large areas of the brain, providing broadcast signals to reconfigure circuits through a unique set of neurotransmitters, such as acetylcholine (ACh), dopamine (DA), serotonin (5HT), and norepinephrine (NE). These neuromodulators have been implicated in a broad range of behavioral functions, many overlapping across modulators, such as learning (ACh, DA), arousal (NE, ACh), and impulsivity (5HT, DA). Based on these differences, each neuromodulator has been suggested to have a distinct computational function. Specifically, DA has been well studied across a range of behaviors; its responses are well described by a simple computational principle—reward prediction error—that explains a broad set of observations. Less is known about NE, 5HT, or ACh, largely because of the technical challenges of recording these. Nevertheless, current evidence suggests that these neuromodulators have many more overlapping functions than initially thought.

Basal forebrain (BF) cholinergic neurons constitute a major neuromodulatory system implicated in normal cognition functions, including learning, memory, and attention. Cognitive deficits in Alzheimer's disease, Parkinson's dementia, age-related dementia, and healthy aging are correlated with the extent of deterioration of BF cholinergic neurons. Cholinergic projections densely innervate neocortex and release acetylcholine, which is thought to regulate arousal, attention, and learning.

We recorded optogenetically identified cholinergic neurons using optogenetic identification in mice performing an auditory detection task requiring sustained attention. We found cholinergic neurons responded to reward and punishment. This reinforcement response invites comparison to dopamine neurons, for which a key conceptual advancement was that they compute reward prediction error—the difference between reward expectation (as informed by a predictive cue) and the reward received. Therefore, we used fiber photometry to measure bulk GCaMP signal in cholinergic neurons in a cued probabilistic outcome task. After the mice learned the task, cholinergic neurons responded to reward-predicting cues, and the neurons' reward responses were diminished by cued expectations, similar to dopamine neurons. These results reveal that the cholinergic system broadcasts a rapid and precisely timed reinforcement signal that could support fast cortical activation and plasticity. The basal forebrain also contains long-range GABAergic projections to the cortex, but these are poorly understood. Using optogenetic identification of PV^+ NB neurons, we found that unlike cholinergic neurons, these neurons encode the expected value of the cue and predict the reaction time of the mice on a trial-to-trial basis. We expect that these studies

will reveal fundamental principles about the parallel GABAergic and that cholinergic projection subsystems in the basal forebrain differentially control cortical processing.

We are also investigating the roles of multiple neuromodulators in simple learning tasks to understand how similar they are. We are further probing the roles of dopamine in foraging, impulsivity, and motor behaviors to understand how its functions generalize across behaviors.

Computational and Circuit Psychiatry: Hallucinations, Impulsivity, and Social Cognition

This work was done in collaboration with R. Axel (Columbia University).

Our behavioral research is aimed at determining the computational principles of cognition. The starting point of our studies is the observation that behavior is often described using folk psychological categories, even though almost everyone agrees that the neural architecture that supports behavior performs computational functions. Hence, there is a large gap between the well established molecular and neural circuit level descriptions of the brain and much less developed computational descriptions of behavior produced by the brain. We try to fill this gap using two main sources of insight. First, we use psychophysical and behavioral economic approaches to design tasks and collect large data sets. Second, we use machine learning and theoretical neuroscience models in an attempt to reverse engineer the algorithms. Our studies have mainly focused on decision confidence in rats, mice, and humans, and we are beginning to expand into impulsivity and cognitive control and computational phenotyping of human behavior, including psychiatric populations.

Hallucinations, a core symptom of schizophrenia, can be operationalized as false percepts that are experienced with the same certainty as veridical percepts. Based on this insight we have developed a behavioral task to quantitatively measure hallucination-like percepts in mice. We trained mice to perform a psychometric auditory detection task with a time investment–based confidence report. Using computational modeling, we found that choice behavior closely followed the predictions of an ideal observer model, suggesting that mice indeed reported their perception and their confidence. We are currently testing the translational validity of this behavioral test by reproducing previous results and setting up patient testing. This mouse model will enable a circuit-level description of the causal link between dopamine and schizophrenia symptoms.

Impulsivity is a behavioral trait present in many psychiatric disorders that significantly increases the risk of suicide, violence, and criminal behavior. Whether a particular decision made too early is impulsive is difficult to determine, because misevaluation of expected outcome or misestimation of time could lead to similar consequences. Therefore, as an initial step toward understanding the underlying neural circuits, we sought to develop a task that isolates the contribution of impulsivity to individual choices and separates it from reward valuation. We devised a novel behavioral paradigm that enables us to parse out the contribution of these pathways to reward valuation and motor inhibition. Our goal is to elucidate the contribution of neural circuits that are involved in impulsivity, particularly the anterior cingulate cortex's control of the serotonergic and dopaminergic systems, using photometry recordings and optogenetic manipulation. Ultimately, we hope our circuit-based understanding of impulsivity will contribute to the design of circuit-specific treatments for impulsivity disorders.

Social behavior is integral to animals' survival and reproduction; social deficits are at the heart of cognitive disorders such as autism spectrum disorder that have proven profoundly difficult to study in model organisms. We would like to understand how social information is represented, computed, and used by mice. In rodents, a main source of information for social decision-making and reward valuation is the chemosensory system. One component of these scents is darcin, a mouse urinary protein that is sufficient to induce innate attraction in sexually receptive female mice and also acts as an unconditioned stimulus in associative learning paradigms. We have identified a neural circuit that extends from the vomeronasal organ to the medial amygdala and mediates the innate response to darcin. Genetic silencing of either accessory olfactory bulb or the medial amygdala eliminates the innate attraction to darcin. Optical reactivation of darcin-activated medial amygdala neurons elicits attraction behavior.

We are also interested in understanding the rules that mice use to choose partners. For this purpose, we have developed a psychophysical social behavior task, the "social carousel," inspired by perceptual psychophysics and game theoretic traditions that have been instrumental in studying other facets of cognition. Our task enables reliable, quantitative, and high-throughput analysis of social interactions in mice. Subjects can choose to engage in extended social interactions with the caged mice at the expense of delaying the water reward. This task allows us to infer the "social value" of a mouse based on the trade-off between social interactions and appetitive rewards. In addition, this task is compatible with our electrophysiological studies, because it is devised for precise stimulus delivery and reproducible behavioral contingencies.

Neurotechnologies: Viral Targeting, Behavioral Language, and Nanophotonics

This work was done in collaboration with A. Mohanty, M.A. Tadayon, and M. Lipson (Columbia University).

Progress in neuroscience relies on continual technique development and improvement. Part of our effort is devoted to developing, improving, and adopting instrumentation that enables the study of neural circuits and behavior. For instance, we developed an open-source behavioral control system (Bpod) that is used by more than 60 laboratories around the world. Building on this, we are also developing a formal behavioral description language that is hardware-independent and brings rigor and reproducibility to complex behavioral tasks. On the biological side, we also developed a new viral complementation strategy that enables tropism-free retrograde viral delivery for targeting long-range projections. We collaborate to develop and test a new class of nanophotonics silicon probes for high-density optical.

We continue to work on a receptor complementation strategy to achieve efficient, tropism-free retrograde targeting of neurons. Retrogradely transported neurotropic viruses enable researchers to target neurons based on their long-range projections and have thus become indispensable tools for linking neural connectivity with function. A major limitation of viral techniques is that they rely on cell type–specific molecules for uptake and transport and, as a result, may fail to infect neurons that do not express the requisite complement of surface receptors (viral tropism). To overcome this problem, we designed AAV constructs to express the coxsackievirus and adenovirus receptor (CAR) and thereby potentiate canine adenovirus type 2 (CAV-2) infection in candidate projection neurons. Enhancement of CAR expression greatly increased retrograde labeling rates with CAV-2 in multiple long-range projecting neural circuits in both mice and rats, providing a robust method for high-efficiency, tropism-free retrograde labeling. The CAR/CAV-2 system will facilitate efficient retrograde targeting for functional analysis of neural circuits. We are currently developing variants that will enable systemic delivery of receptors.

The ability to activate neural populations using optogenetics has revolutionized the study of neural circuits. However, this is currently primarily done by using a single fiber to flood light into a large volume of the brain. We demonstrated the first active optical switch to excite individual neurons at 473 nm, which also enables us to control multiple independent beams for deep brain neural stimulation. Using a silicon nitride waveguide platform, the Lipson group developed a cascaded Mach–Zehnder interferometer (MZI) network located outside the brain to direct light to eight different grating emitters located at the tip of the neural probe. To demonstrate the potential of the platform and provide bidirectional neural interfacing for localized and specific stimulation and recording in deep brain area in vivo, we package the device with tungsten wires placed near the output gratings with a resolution of <10 microns. We report fast (up to 200 Hz), specific, and independent control of single-neuron activity across cortical laminar and hippocampus in anesthetized mice. The demonstrated depth-specific manipulation approaches single-cell resolution, and this device could control neural activity in vivo independently across beams and with high spatiotemporal resolution.

PUBLICATIONS

Chevy Q, Kepecs A. 2018. When acetylcholine unlocks feedback inhibition in cortex. *Neuron* **97:** 481–484.

Li SJ, Vaughan A, Sturgill JF, Kepecs A. 2018. A viral receptor complementation strategy to overcome CAV-2 tropism for efficient retrograde targeting of neurons. *Neuron* **98:** 905–917.e5.

Mohanty A, Li Q, Tadayon MA, Bhatt GR, Cardenas J, Miller SA, Kepecs A, Lipson M. 2018. A reconfigurable nanophotonics platform for sub-millisecond, deep brain neural stimulation. arXiv:1805.11663

Ott T, Masset P, Kepecs A. 2018. The neurobiology of confidence: from beliefs to neurons. *Cold Spring Harb Symp Quant Biol* **83:** 038794.

In Press

Fishell G, Kepecs A. 2020. Interneuron types as attractors and controllers. *Annu Rev Neurosci* doi:10.1146/annurev-neuro-070918-050421

Hirokawa J, Vaughan A, Masset P, Ott T, Kepecs A. 2019. Structured and cell-type-specific encoding of decision variables in orbitofrontal cortex. *Nature* (in press).

Lak A, Okun M, Moss M, Gurnani H, Farrell K, Wells MJ, Reddy CB, Kepecs A, Harris KD, Carandini M. 2019. Neural basis of learning guided by sensory confidence and reward value. bioRxiv doi:10.1101/411413

MATHEMATICAL PRINCIPLES OF NEURAL COMPUTATION

A.A. Koulakov B. Baserdem D. Kepple N. Tran
H. Giaffar S.A. Shuvaev

Our laboratory works on theories of neural computation. Our overall strategy is to use methods developed in mathematics, physics, machine learning, computer science, and statistics to build experimentally testable models of neural networks and their function. In most cases, we base our theories on what is known about particular biological systems; however, given that the principles of brain function remain unclear, in many cases, we resort to building computational theories. This means that we formulate the problems solved by the brain in a mathematically rigorous fashion and hypothesize how an engineer would solve the problem, given biological and experimental constraints. We then use these solutions to form experimentally testable predictions. Testing these predictions in collaboration with our experimental colleagues helps us refute or refine our theories. For example, we are interested in understanding how connectivity is established in the brain. We have proposed several theories that may determine the rules of making connections between neurons based on a limited set of instructions contained in the genome. These theories address several levels of organization, including computational, biological, engineering, and evolutionary. Our theories may explain the differences between connectivities in normal and abnormal brain circuits. We are also interested in understanding the principles of perceptual invariance—that is, how can sensory systems represent objects in the environment despite substantial variations in intensity and background. Visual percepts, for example, retain basic features, such as perceived shape and color composition, despite variable luminance, spectral composition, scale, and position of the stimuli. Although we study the question of perceptual invariance in application to well-defined problems, we believe that the principles that we will uncover may generalize across sensory modalities. Finally, we are pursuing the question of how modern theories of machine learning and artificial intelligence can apply to brain function. Although reinforcement learning, deep learning, long short-term memory networks, etc., are successful in solving a variety of artificial intelligence problems, their mapping onto brain circuits remains unclear. We attempt to bring these systems closer to satisfying the constraints imposed by biology. We hope that the convergence of theoretical constructs and their biological underpinning will help us learn more about the brain function.

DeepNose: Using Artificial Neural Networks to Represent the Space of Odorants

N. Tran, D. Kepple, S.A. Shuvaev, A.A. Koulakov

The olfactory system employs an ensemble of odorant receptors (ORs) to sense odorants and to derive olfactory percepts. We trained artificial neural networks to represent the chemical space of odorants and used that representation to predict human olfactory percepts. We hypothesized that ORs may be considered 3D spatial filters that extract molecular features and can be trained using conventional machine learning methods. First, we trained an autoencoder, called DeepNose, to deduce a low-dimensional representation of odorant molecules, which were represented by their 3D spatial structure. Next, we tested the ability of DeepNose features in predicting physical properties and odorant percepts based on 3D molecular structure alone. We found that despite the lack of human expertise, DeepNose features led to perceptual predictions of comparable accuracy to molecular descriptors often used in computational chemistry. We propose that the DeepNose network can extract de novo chemical features predictive of various bioactivities and can help us understand the factors influencing the composition of the OR ensemble (Tran et al. 2018).

A Computational Framework for Converting High-Throughput DNA Sequencing Data into Neural Circuit Connectivity

D.D. Ferrante, A. Koulakov [in collaboration with J. Dubnau, A. Zador, M. Oyibo, C. Cao, H. Zhan, CSHL; L. Enquist, Princeton]

There is growing interest in determining the connectivity of neural circuits—the "connectome"—at single-neuron resolution. Most approaches to circuit mapping rely on either microscopy or physiology, but these approaches have very limited throughput. We have recently proposed BOINC (Barcoding of Individual Neuronal Connectivity), a radically different approach to connectivity mapping based on high-throughput DNA sequencing. Here we describe the set of computational algorithms that serve to convert sequencing data into neural connectivity. We developed and applied our computational pipeline to the results of proof-of-principle experiments illustrating an implementation of BOINC based on pseudorabies virus (PRV). PRV is capable of traversing individual synapses and carrying DNA barcodes from one cell to another. Using this high-throughput sequencing data, we obtain 456×486 connectivity matrix between putative neurons. An inexpensive high-throughput technique for establishing circuit connectivity at single-neuron resolution would represent a major advance in neuroscience (Oyibo et al. 2018).

Primacy Model and the Evolution of the Olfactory Receptor Repertoire

H. Giaffar, A.A. Koulakov [in collaboration with D. Rinberg, NYU]

For many animals, the neural activity in early olfactory circuits during a single sniff cycle contains sufficient information for fine odor discrimination. Although much is known about the transformations of neural representations in early olfactory circuits, exactly how odorant-evoked activity in the main olfactory bulb shapes the perception of odors remains largely unknown. In olfaction, odorant identity is generally conserved over a wide range of conditions, including concentration. We present a theory of identity assignment in the olfactory system that accounts for this invariance with respect to stimulus intensity. We suggest that the identities of relatively few high-affinity olfactory receptor types determine an odorant's perceived identity. This set of high-affinity receptors is defined as the primary set, and the coding model based on their responses is called the primacy theory. In this study, we explore the impact that primacy coding may have on the evolution of the ensemble of olfactory receptors. A primacy coding mechanism predicts the arrangement of different receptor types in a low-dimensional structure that we call a primacy hull. We present several statistical analyses that can detect the presence of this structure, allowing the predictions of the primacy model to be tested experimentally (Giaffar et al. 2018).

Computational Algorithms and Neural Circuitry for Compressed Sensing in the Mammalian Main Olfactory Bulb

D. Kepple, H.S. Demmer, A.A. Koulakov [in collaboration with S.D. Shea, B.N. Cazakoff, D. Eckmeier, CSHL]

A major challenge for many sensory systems is the representation of stimuli that vary along many dimensions. This problem is particularly acute for chemosensory systems because they require sensitivity to a large number of molecular features. Here we use a combination of computational modeling and in vivo electrophysiological data to propose a solution for this problem in the circuitry of the mammalian main olfactory bulb. We model the input to the olfactory bulb as an array of chemical features that, because of the vast size of chemical feature space, is sparsely occupied. We propose that this sparseness enables compression of the chemical feature array by broadly tuned odorant receptors. Reconstruction of stimuli is then achieved by a supernumerary network of inhibitory granule cells. The main olfactory bulb may therefore implement a compressed sensing algorithm that presents several advantages. First, we demonstrate that a model of synaptic interactions between the granule cells and the mitral cells that constitute the output of the olfactory bulb can store a highly efficient representation of odors by competitively selecting a sparse basis set of "expert" granule cells. Second, we further show that this model network can simultaneously learn separable representations of each component of an odor mixture without exposure to those components in isolation. Third,

our model is capable of independent and odor-specific adaptation, which could be used by the olfactory system to perform background subtraction or sensitively compare a sample odor with an internal expectation. This model makes specific predictions about the dynamics of granule cell activity during learning. Using in vivo electrophysiological recordings, we corroborate these predictions in an experimental paradigm that stimulates memorization of odorants (Kepple et al. 2018).

Spatial Geometry of Stem Cell Proliferation in the Adult Hippocampus

The modes of stem cell divisions (e.g., symmetric vs. asymmetric) can have a profound impact on the number of progeny and tissue growth, repair, and function. This is particularly relevant for adult neural stem cells, because stem cell–derived neurons affect cognitive and mental states, resistance to stress and disease, and response to therapies. In this study, we show that although dividing stem cells in the adult hippocampus display a certain bias toward paired distribution (which could imply the prevalence of symmetric divisions), this bias already exists in the distribution of the general population of stem cells and may be responsible for the perceived occurrence of symmetric stem cell divisions. Remarkably, the bias in the distribution of stem cells decreases with age. Our results argue that the preexisting bias in stem cell distribution may affect current assumptions regarding stem cell division and fate as well as conjectures on the prospects of brain repair and rejuvenation.

PUBLICATIONS

Giaffar H, Rinberg D, Koulakov AA. 2018. Primacy model and the evolution of the olfactory receptor repertoire. bioRxiv doi:10.1101/255661

Kepple D, Cazakoff BN, Demmer HS, Eckmeier D, Shea SD, Koulakov AA. 2018. Computational algorithms and neural circuitry for compressed sensing in the mammalian main olfactory bulb. bioRxiv doi:10.1101/339689

Mineyeva OA, Enikolopov G, Koulakov AA. 2018. Spatial geometry of stem cell proliferation in the adult hippocampus. *Sci Rep* **8:** 3444.

Mineyeva O, Koulakov A, Enikolopov G. 2018. Neural stem cell spacing questions their self-renewal. *Aging* **10:** 1793–1794.

Oyibo H, Cao C, Ferrante DD, Zhan H, Koulakov A, Enquist L, Dubnau J, Zador A. 2018. A computational framework for converting high-throughput DNA sequencing data into neural circuit connectivity. bioRxiv doi:10.1101/244079

Tran N, Kepple D, Shuvaev SA, Koulakov AA. 2018. DeepNose: using artificial neural networks to represent the space of odorants. bioRxiv doi:10.1101/464735

THE FUNCTION AND PLASTICITY OF CENTRAL SYNAPSES IN ADAPTIVE AND MALADAPTIVE BEHAVIORS RELATED TO PSYCHIATRIC DISORDERS

B. Li C. Bizingre, A. Bouhuis, S. Boyle, C. Bravo-Rivera, H. Deng, A. Furlan, J. Giovanniello, O. Gschwend, W. Guan, T. Russo, X. Xiao, T. Yang, K. Yu, X. Zhang

The focus of research in our laboratory has been to understand the link between neural circuits and behaviors. We are particularly interested in studying the synaptic and circuit mechanisms underlying aspects of motivated behaviors, such as attention, motivation, and learning and memory, as well as synaptic and circuit dysfunctions that may underlie the pathophysiology of mental disorders, including anxiety disorders, depression, autism, and drug addiction. We integrate in vitro and in vivo electrophysiology, imaging, molecular, genetic, optogenetic, and chemogenetic methodologies to probe and manipulate the function of specific neural circuits in the rodent brain and to determine their roles in adaptive or maladaptive responses in various behavioral paradigms. We are currently undertaking the following major lines of research.

The Role of the Amygdala Circuitry in Motivated Behaviors

Our previous studies demonstrate that the central amygdala (CeA) has a key role in learning and expression of defensive responses to threats. In particular, our studies indicate that somatostatin-expressing (SOM^+) neurons in the lateral division of the central amygdala (CeL) are essential for the acquisition and recall of fear memories. Furthermore, we show that another major class of CeL neurons, the protein kinase C-δ-expressing ($PKC\text{-}\delta^+$) neurons, is essential for the synaptic plasticity underlying learning in the lateral amygdala, as it is required for lateral amygdala neurons to respond to unconditioned stimulus, and moreover carries information about the unconditioned stimulus (US) to instruct learning.

In a more recent study (Ahrens et al. 2018), we demonstrate that enhanced excitatory synaptic inputs onto SOM^+ CeL neurons and the resulting reduction in inhibition onto downstream SOM^+ neurons in the bed nucleus of the stria terminalis (BNST) plays an important role in the generation of anxiety-related behaviors. Notably, our results indicate that an increase in dynorphin signaling in SOM^+ CeA neurons mediates the paradoxical reduction in inhibition onto SOM^+ BNST neurons, and that the consequent enhanced activity of SOM^+ BNST neurons is both necessary and sufficient to drive the elevated anxiety. Our results unravel previously unknown circuit and cellular processes in the central extended amygdala that can cause maladaptive anxiety.

In another study (Zhang and Li 2018b), we found that associative learning driven by reward and punishment profoundly alters population responses in the basolateral amygdala (BLA), reducing noise correlations and transforming the representations of conditioned stimuli (CSs) to resemble the valence-specific representations of USs. This transformation is accompanied by the emergence of prevalent inhibitory CS and US responses, and by the plasticity of CS responses in individual BLA neurons. During reversal learning wherein the expected valences are reversed, BLA population CS representations are remapped onto ensembles representing the opposite valences and predict the switching in valence-specific behaviors. Our results reveal how signals predictive of opposing valences in the BLA evolve during learning, and how these signals are updated during reversal learning to guide flexible behaviors.

The Basal Ganglia Circuit in Motivated Behaviors

The basal ganglia, a group of subcortical nuclei, play a crucial role in motivated behaviors. Recently we

showed that neurons in the habenula-projecting globus pallidus (GPh), an output of the basal ganglia, are essential for evaluating action outcomes. Our current study addresses the roles of another basal ganglia output, the ventral pallidum, in motivated behaviors.

Reward-seeking and punishment-avoidance behaviors are driven by positive and negative motivational processes. The ventral pallidum (VP) is critical for invigorating reward seeking and may also be involved in punishment avoidance, but how the VP contributes to such opposing behavioral actions remains unclear. By using electrophysiological recording combined with optogenetics in behaving mice, in a recent study (Bravo-Rivera et al. 2018) we show that the VP contains two valence-specific populations—positive-valence neurons (PVNs) and negative-valence neurons (NVNs), which are GABAergic and glutamatergic, respectively. These two populations bidirectionally and oppositely encode positive and negative valences or values, and are differentially modulated by an animal's internal states. Furthermore, optogenetic manipulations reveal that PVNs and NVNs control reward seeking and punishment avoidance, respectively, at least in part through their projections to the lateral habenula. Together, our results suggest that the coordinated activities between PVNs and NVNs orchestrate the generation of incentive and aversive saliences that trigger or promote valence-specific behaviors.

Circuit Mechanisms of the Distinct Phenotypes of a 16p11.2 Deletion Model of Autism

Previous studies and preliminary results in our lab indicate that mice carrying a deletion corresponding to the human 16p11.2 microdeletion (the "16p11.2 mice") have behavioral changes including (1) increased locomotor activity and repetitive behaviors, (2) cognitive deficit, and (3) enhanced anxiety. These behavioral changes, which are related to the symptoms of humans carrying the homologous deletion, suggest potential dysfunctions in the basal ganglia, a group of nuclei in the brain that have been implicated in motor function and cognition, as well as emotion and anxiety-related behaviors. Notably, recent studies have identified cellular and structural changes in the basal ganglia of the 16p11.2 mice, which may underlie both the motor and the cognitive deficits in these mice. Furthermore, the changes in the basal ganglia may also contribute to the anxiety phenotypes of these mice.

On the basis of all these findings, including our own, we are currently testing the hypothesis that impaired function in different nuclei of the basal ganglia causes the divergent behavioral changes in the 16p11.2 mice, including cognitive deficit, increased locomotor activity, and enhanced anxiety (Giovanniello et al. 2018). We are testing this hypothesis using a combination of state-of-the-art neuroscience technologies, including electrophysiology, imaging, optogenetics, chemogenetics, and novel behavioral techniques. This line of research, which addresses the roles of the basal ganglia in distinct behavioral changes in the 16p11.2 deficiency model of autism, will establish a framework for research into the mechanisms underlying the core symptoms and comorbid symptoms of autism spectrum disorders, and will guide the development of novel and effective therapeutics.

PUBLICATIONS

Ahrens S, Wu M, Hwang G-R, Li H, Penzo MA, Tollkuhn J, Li B. 2018. A central extended amygdala circuit that modulates anxiety. *J Neurosci* **38:** 5567–5583.

Bravo-Rivera C, Stephenson-Jones M, Furlan A, Fernandes-Henriques C, Zhang X, Xiao X, Yang T, Li B. 2018. Genetically-distinct ventral pallidal neurons drive the motivation for reward approach and punishment avoidance through projections to the lateral habenula. *Society for Neuroscience Annual Meeting*. Program No. 109.02. 2018 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2018. Online. Oral presentation in a Nanosymposium.

Giovanniello J, Ahrens S., Li B. 2018. Sex-specific behavioral and neuronal circuit disruptions in a 16p11.2 microdeletion model of autism. *Society for Neuroscience Annual Meeting*. Program No. 103.01. 2018 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2018. Online. Oral presentation in a Nanosymposium.

Li B. 2019. Central amygdala cells for learning and expressing aversive emotional memories. Invited review. *Curr Opin Behav Sci* **26:** 40–45.

Schiff H, Bouhuis AL, Yu K, Penzo MA, Li H, He M, Li B. 2018. An insula-central amygdala circuit for guiding tastant-reinforced choice behavior. *J Neurosci* **38:** 1418–1429.

Zhang X, Li B. 2018a. A pathway to avoiding threats? *Neuron* **100:** 780–782.

Zhang X, Li B. 2018b. Population coding of valence in the basolateral amygdala. *Nat Commun* **9:** 5195. doi:10.1038/s41467-018-07679-9

THE STUDY OF INTELLIGENT MACHINES

P. Mitra S. Banerjee, M. Bota, D. Ferrante, A. Field-Pollatou, S. Georgiou, J. Hsu, B. Huo, C. Lee, X. Li, L. Lodato, I. Mallardi, J. O'Rourke, S. Savoia, A. Tolpygo, F. Xu

Mesoscale Circuit Mapping in the Mouse Brain

The Mouse Brain Architecture (MBA) project aims to develop a brain-wide connectivity map of the adult male C57BL/6 mouse brain at the mesoscopic scale. The approach has been to inject neuronal tracers in a grid of points that form an unbiased sampling of one hemisphere in the mouse brain; there is a sister project at RIKEN studying marmosets. The goal is to bring about a new conceptual synthesis with a long-term impact on neuroscientific understanding and computational tools. Dr. Mitra originally proposed such a mapping strategy and defined the mesoscale as the transitional length scale from a microscopic scale at which individual variation is prominent to a macroscopic scale at which one can see species-typical patterns. The mesoscale roughly corresponds to the scale of interest in classical neuroanatomical reference atlases.

Experiment

The current focus of the project is to complete injection coverage using AAV-retro (AAVr) custom viral constructs produced by the UPenn Vector Core. AAVr operates using fluorescent labels that do not require a costly histochemical step, like CTB (previously utilized in the project), and achieves twice the experimental efficiency through both red and green color differentiation and labeling. We make computer-guided injections of neuronal tracers, targeting predetermined sites on a 3D grid covering the entire volume of the left hemisphere, with a minimum of one injection per cubic millimeter of brain. So far, 270 AAVr injections have been placed.

Data Analysis
Brain to atlas registration

This work was done in collaboration with M. Miller, D. Tward, and B. Lee (Johns Hopkins University).

In a continuing collaboration with Michael Miller at Johns Hopkins University, we are improving previously adopted techniques from differential geometry (large-deviation diffeomorphic metric mapping) for registering brains to a reference atlas. Progress in 2018 included treatment of cross-modal registration across multiple data modalities (Lee et al. 2018b), as well as assessment of the metric distortions due to the histological processing using the determinant of the Jacobian of the transformation between the atlas coordinate system and the individual brain (Lee et al. 2018a).

Semantic segmentation and topological skeletonization

This work was done in collaboration with S. Das (Indian Institute of Technology, Madras) and Y. Wang, S. Wang, and D. Wang (Ohio State University).

We improved methods to detect nuclei in whole-brain volumetric data with high precision (>95%) (Pahariya et al. 2018). Work was commenced on semantic segmentation of the images using machine learning methods to detect axonal and dendritic processes, as well as to carry out semantic segmentation of the images using supervised machine learning approaches. We also advanced work on topological skeletonization of individual neurons as well as groups of neurons labeled by tracer injections, using methods taken from computational topology (discrete Morse theory) (Wang et al. 2018).

Web portal

In 2018, the Mitra laboratory released a new data portal for the MBA project (http://mouse.brainarchitecture.org). The new interface involves navigation through data sets using 3D cursor selection of injection site or a hierarchical tree-based injection site search. The portal also now contains data from mesoscale circuit mapping of the marmoset brain (see next section).

Mesoscale Circuit Mapping in the Marmoset Brain

This work was done in collaboration with Y. Takahashi, M. Lin, J. Nagashima, M. Hanada, and B. Huo (Laboratory at the RIKEN Center for Brain Science); M.G.P. Rosa (Monash University); and P.R. Martin (University of Sydney).

In parallel with the MBA project, since 2014, Dr. Mitra has led the laboratory at RIKEN Center for Brain Science as a part of Japan's Brain/MINDS project. The laboratory was directed toward mapping the mesoscale connectivity in the marmoset brain. The common marmoset (*Callithrix jacchus*) is an emerging primate model system of widespread interest seeing rapid growth compared with the macaque because of faster generation times (allowing more facile genetic manipulations), a smaller brain with fewer cortical folds (allowing for easier access to the cortical surface), and behavioral traits such as a rich social vocal communication repertoire.

A high-throughput neurohistological and computational pipeline was developed for brain-wide mesoscale connectivity mapping of the marmoset brain (Lin et al. 2018). As of 2018, 190 injections of anterograde and retrograde tracers have been placed in 49 brains. Among them, 37 data sets were digitized to enable production of a rich data set that is currently being subjected to computational analysis using techniques developed in the mouse. A dedicated cross-modal registration pipeline was developed in collaboration with Michael Miller's group at Johns Hopkins University and applied both to the multimodal marmoset data sets (Lee et al. 2018a). The marmoset connectivity data set is leading to specific new discoveries about marmoset brain architecture. In 2018 this included two results obtained in collaboration with Marcello Rosa's Laboratory at Monash University (Majka et al. 2019) and Paul Martin's Laboratory at the University of Sydney (Huo et al. 2018).

BICCN Data Core, Spatial Transcriptomics

This work was done in collaboration with M. Miller (Johns Hopkins University); Y. Wang (Ohio State University); M. Sivaprakasan, K. Ram, and S. Das (IIT Madras); and E. Macosko and A. Regev (Broad Institute/MIT).

In 2017, Dr. Mitra started the Data Core component of a U19 award as part of the NIH Brain Initiative Cell Census Network (BICCN) led by Josh Huang (CSHL) and Paola Arlotta (Harvard); similarly, Dr. Mitra started the Data Core component for a U01 award to Pavel Osten. Part of the work in the data core is to centralize project data from different investigators for analysis, web presentation, and for transfer to NIH-designated web archives. Work in the data core can be broken down into (1) data ownership, stewardship, and metadata management; (2) storage and infrastructure management; and (3) portal development and tool enhancements.

The Mitra laboratory is collaborating with Evan Macosko's laboratory at the Broad Institute to perform spatially localized mapping of single-nucleus RNA sequencing (snRNAseq)-based cell type analysis in the mouse brain, as a part of the U19 award. A standard experimental procedure has been established along with a strategic plan to cover the entire brain. Brains with tissue punched for snRNAseq were registered in 3D and to the reference brain using an expectation maximization variant of the diffeomorphic geometry algorithm.

Virtual Pathologist Project

This work was done in collaboration with A. Kepecs (CSHL); J. Crawford, M. Conforti, and D. Savant (Northwell); B. Gallas and Q. Gong (FDA); and C. Abbey (UCSB).

With exponential growth in AI/machine learning, there is significant interest in computer-aided diagnoses (CADs). The Virtual Pathologist Project was designed as a system for acquiring extensive behavioral data from pathologists as they read a physical slide on a microscope, which constitutes their standard working environment. The system for capturing this data is based on the FDA's Evaluation Environment for Digital and Analog Pathology (eeDAP) platform running on a CSHL workstation. The goals of the project are to better understand how pathologists query a vast amount of data available to them on slides containing thin tissue sections stained according to the standards

of the indicated medical exam (breast tissue assessment, GI tissue assessment, etc.), and then evaluate machine learning approaches using the same information as image annotations.

The eeDAP system has been modified to accommodate various extensions needed for this project. A laboratory information management system (LIMS) was modified from the MBA project to organize the output data into a hierarchical database. Both the eeDAP system and LIMS are being prototyped at the Mitra lab.

Understanding Modern Machine Learning: Interpolating Learners

This work was done in collaboration with M. Belkin (OSU) and D. Hsu (Columbia University).

The field of machine learning has undergone exponential growth in recent years, driven largely by artificial neural network algorithms with many feedforward layers (deep nets). Although commercially successful, there is not yet a good theoretical understanding of why these algorithms work. A particularly puzzling aspect of these algorithms is the usage of very large numbers of parameters—often exceeding the number of data points. According to classical statistics, this should lead to overfitting and poor generalization. However, there are important examples in which the generalization is quite good, even though the training data is fit exactly. How this happens is a topic of great theoretical interest and activity.

In a collaborative paper with Mikhail Belkin (OSU) and Daniel Hsu (Columbia University) presented in the 2018 NeurIPS meeting, Dr. Mitra has helped provide an answer to this puzzle by proposing an algorithm (weighted interpolating nearest neighbors [wiNN]) that exhibits the same behavior as do deep nets in terms of successful overparametrization. This algorithm fits the training data exactly (i.e., interpolates), but also shows good generalization properties. It achieves the Bayes risk (the theoretical limit) in the large-sample limit and does so in a min–max optimal manner. Additionally, the algorithm explains the existence of adversarial examples for deep nets.

This theoretical work also provides hints as to why the deep net–based approach is fundamentally different from how biological brains work. It is now appreciated that biological brains learn with much smaller data sets than the deep nets (i.e., have much stronger prior structure). Modern supervised ML approaches such as deep nets may simply be sophisticated data interpolation techniques, whereas biological brains have superior abilities to generalize because of their prior structure present even before learning takes place (in the form of a genetic developmental program). Dr. Mitra's experimental research program is predicated on the premise that much of this prior structure is provided by the mesoscale circuit architecture of vertebrate brains—which constrains the potential synapses between neurons. In the future, the experimental mesoscale data set gathered in mouse and marmoset brains will be examined in light of the ongoing theoretical machine learning work.

PUBLICATIONS

Allegra N, Bamich B, Mitra P, Sire C. 2018. Phase transitions in distributed control systems with multiplicative noise. *J Statist Mech: Theory Exp* **2018:** 013405.

Belkin M, Hsu DJ, Mitra P. 2018. Overfitting or perfect fitting? Risk bounds for classification and regression rules that interpolate. arXiv:1806.05161

Huo B-X, Zeater N, Lin MK, Takahashi YS, Hanada M, Nagashima J, Lee BC, Grünert U, Miller MI, Rosa MGP, et al. 2018. Continuity between koniocellular layers of dorsal lateral geniculate and inferior pulvinar nuclei in common marmosets. bioRxiv doi:10.1101/315598

Lee BC, Lin MK, Fu Y, Hata J, Miller MI, Mitra PP. 2018a. Multimodal cross registration and quantification of metric distortions in whole brain histology of marmoset using diffeomorphic mappings. arXiv:1805.04975

Lee, BC, Tward DJ, Mitra PP, Miller MI. 2018b. On variational solutions for whole brain serial-section histology using a Sobolev prior in the computational anatomy random orbit model. *PLoS Comp Biol* **14:** e1006610. doi:10.1371/journal.pcbi.1006610

Mitra PP. 2018. Fast convergence for stochastic and distributed gradient descent in the interpolation limit. In *2018 26th European Signal Processing Conference (EUSIPCO)*. IEEE, Piscataway, NJ.

Pahariya G, Das S, Jayakumar J, Banerjee S, Vangala V, Ram K, Mitra PP. 2018. High precision automated detection of labeled nuclei in terabyte-scale whole-brain volumetric image data of mouse. bioRxiv doi:10.1101/252247

Wang S, Li X, Mitra PP, Wang Y. 2018. Topological skeletonization and tree-summarization of neurons using discrete Morse theory. arXiv:1805.04997

In Press

Lin MK, Takahashi YS, Huo B-X, Hanada M, Nagashima J, Hata J, Tolpygo AS, Ram K, Lee BC, Miller MI, et al. 2019. A high-throughput neurohistological pipeline for brain-wide mesoscale connectivity mapping of the common marmoset. *eLife* **8:** 72. doi:10.7554/eLife.40042

Majka P, Rosa MGP, Bai S, Chan JM, Huo B-X, Jermakow N, Lin MK, Takahashi YS, Wolkowicz IH, Worthy KH, et al. 2019. Unidirectional monosynaptic connections from auditory areas to the primary visual cortex in the marmoset monkey. *Brain Struct Funct* doi:10.1007/s00429-018-1764-4

MAPPING THE BRAIN STRUCTURE AT CELLULAR RESOLUTION AND SUPER-RESOLUTION

P. Osten A. Bludova C. Elowsky R. Munoz Castaneda R. Palaniswamy X. Qi
R. Drewes J. Mizrachi A. Narasimhan J. Palmer J. Taranda

Mapping the Brain at Cellular Resolution and Super-Resolution

How are brain circuits built to achieve the vast array of mammalian behaviors, both innate and learned, and emotions? How can we distinguish normal individual-to-individual variation from abnormal brain features that may underlie the risk for neurodevelopmental and psychiatric disorders in the human brain?

Recently, several automated light microscopy methods capable of imaging mouse brains at high spatial resolution have been developed for analysis of neuronal morphology and long-range connectivity, including serial two-photon tomography (STPT) (Kelly et al. 2018; Turcan et al. 2018; Mandelbaum et al. 2019) and the oblique light-sheet tomography (OLST) developed in our laboratory (Narasimhan et al., bioRxiv 132423 [2017]). The technology development focus in our laboratory in 2018 to 2019 has been on building upon our OLST instrument by enabling the study of brain anatomy in a way that combines whole-brain connectivity analysis of OLST v1.0 with synaptic-level super-resolution analysis in the same tissue. In addition to this new effort, we are continuing our mouse brain cell type atlasing funded by the BRAIN Initiative, extending these methods to other species such as the prairie vole (collaboration with Steven Phelps at the University of Texas, Austin) and deer mouse *Peromyscus maniculatus* (collaboration with Hopi Hoekstra at Harvard), two rodent species with monogamous behaviors.

Research Program 1. Volumetric super-resolution imaging in the mouse brain

Judith Mizrachi, SBU graduate student, leads the project on developing a super-resolution capacity for the OLST v2.0 instrument built by Xiaoli Qi, a computational science analyst in our lab. Although the second-generation OLST instrument is capable of high-resolution imaging (e.g., at $0.4 \times 0.4 \times 1.0$ microns across the entire mouse brain), the work of Ms. Mizrachi aims to use a computational approach, called super-resolution optical fluctuation (SOFI) imaging, that uses temporal correlations of independently fluctuating "blinking" fluorophores, such as Alexa dyes conjugated to secondary antibodies, for the calculation of super-resolved images from a recorded image time series (Figs. 1 and 2). We expect that this method, once fully implemented and integrated in the OLST instrument, will enable highly novel studies of the relationships between local super-resolution analyses of, for example, the postsynaptic distribution of glutamate AMPA receptors, and whole-brain tracing of long-range connectivity of the same cells.

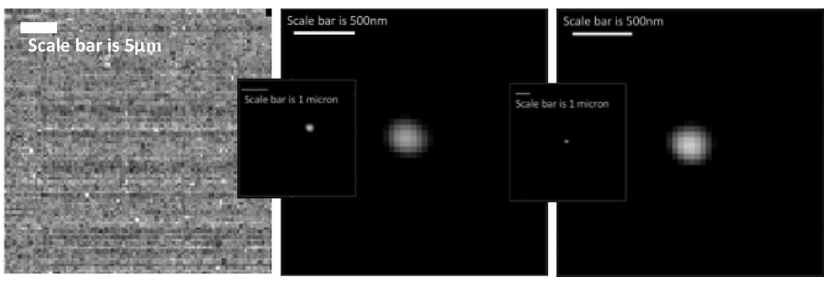


Figure 1. OLST-SOFI of single Alexa-488-conjugated secondary antibodies diluted in agarose. (*Left*) No apparent fluorescent signal in the single FOV image (100×100 pixels) acquired with OLST. (*Middle, right*) SOFI eighth-order processing of several such time series resulted in appearance of fluorescence spots of similar size and brightness, likely single Alexa-488-Ab molecules.

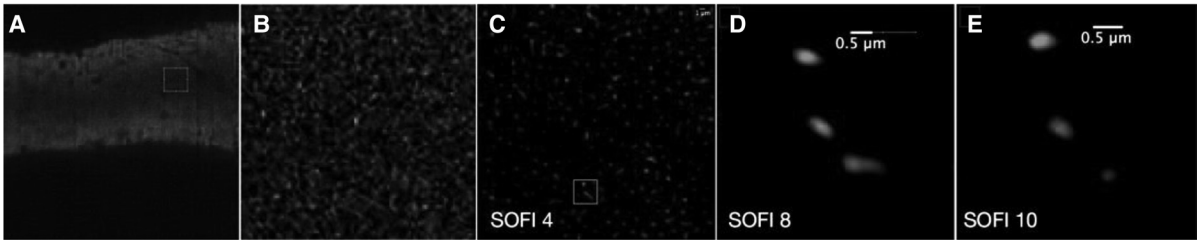


Figure 2. OLST-SOFI example 3: anti-homer1 immunolabeling. Brain tissue was labeled with primary antibodies against homer1A and secondary antibodies conjugated with Alexa-568. (*A*) Low-magnification image of cortical tissue with labeled neuropil and outlined 100 × 100-pixel area that was used for SOFI processing. (*B*) Single FOV image showing only minimal structural information in the labeling. Already (*C*) fourth-order SOFI of only 500 series stack revealed what appears as postsynaptic structures in the neuropil in the mouse cortex, and this signal is further enhanced with (*D*) eighth- and (*E*) 10th-order SOFI.

Research Program 2. Cell type–based understanding of the mouse brain

Our U01 BRAIN Initiative project "Collaboratory for atlasing cell type anatomy in the female and male mouse brain" focuses on building a neuronal and glial-based cell type atlas of the mouse brain, including mapping the distribution and ratios of brain cell types and their wiring into neuronal circuits that underlie the vast diversity of mammalian behaviors. To date we have mapped the brain-wide distribution of more than 20 cell types in the male and female brains, comprising 200 whole-brain data sets, and we also imaged and are in the process of tracing more than 10,000 pyramidal neuron morphologies distributed across all cortical areas (Fig. 3). These data sets offer an unprecedented wealth of quantitative information about the brain-wide distribution of neuronal and glia cell types, representing a highly unique resource for the neuroscience community that is accessible via our web portal at http://mouse.brainarchitecture.org/ost/—which was built in a collaboration with Partha Mitra's lab at CSHL. The U01 funding from the BRAIN Initiative will allow us to map and atlas additionally more than 50 cell types, resulting in a detailed collection of cell type data for most known cell types in the brain to date.

Research Program 3: The structure and function of neural circuits involved in social bond formation

From parental bonds to lasting relationships, social attachments are central to human experience and well-being. The neurobiology of social attachment is a logical and imperative focus of animal research, but common lab models do not form social bonds. To address this question in relevant animal models, we are collaborating with the lab of Steven Phelps at the University of Texas, Austin, to study the differences in social behaviors and underlying brain activations in monogamous prairie voles and polygamous meadow voles and with the lab of Hopi Hoekstra at Harvard to study these questions in monogamous *Peromyscus maniculatus* and polygamous *Peromyscus polionotus*, two species of deer mice. As these rodent species have not been studied with modern tools of systems neuroscience before, we start with developing a pipeline of computational methods for analysis of brain anatomy and function. As this work progresses, we hope that by mapping the structure–function relationships in these related rodent species, we will be able to obtain a detailed understanding of the brain regions and circuits mediating these markedly distinct and evolutionarily highly significant social behaviors.

Deciphering Mouse Models of Human Neurodevelopmental and Psychiatric Disorders

Identifying vulnerable circuits underlying severe neurodevelopmental and psychiatric disorders has been a major motivation for the development of the whole-brain imaging and analysis methods described above. In the first application of these methods to clinically relevant questions, graduate student Eric Szelenyi mapped the distribution of X-chromosome inactivation (XCI) in normal wild-type mice and in a mouse model of the neurodevelopmental fragile X syndrome, demonstrating nonrandom skewing that favors more cells with maternal X chromosome active and

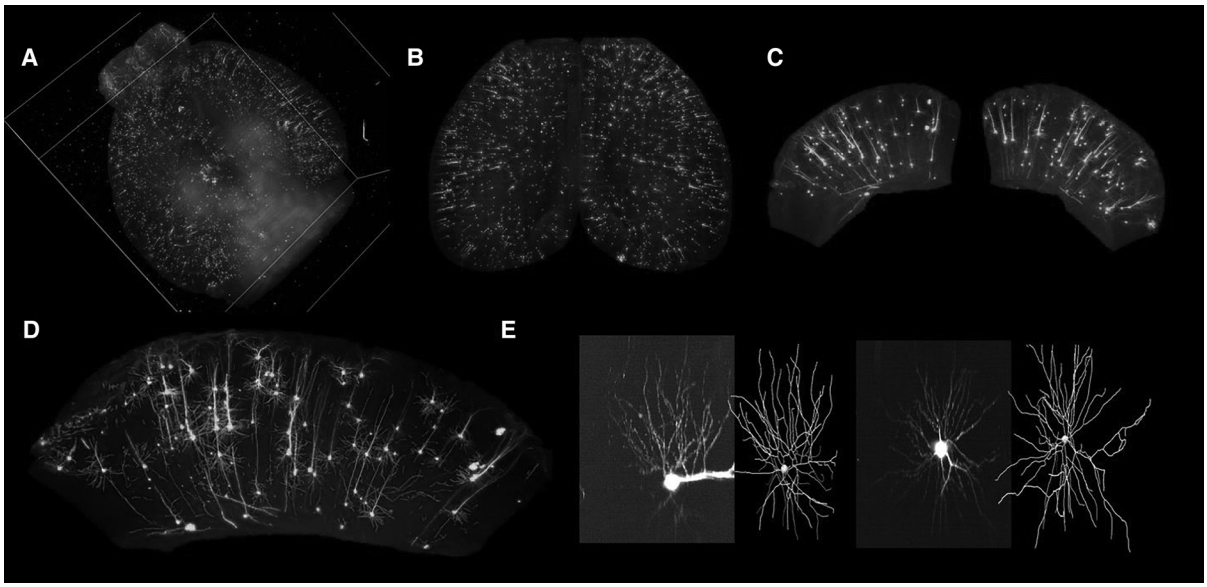


Figure 3. Analysis of cortical neuron morphologies by OLST. (*A*) 3D view of the entire Emx1-Cre mouse brain imaged by OLST with cortical pyramidal neuron labeling achieved by systemic (tail vein) delivery at Cre-dependent AAV9-DIO-GFP in Emx1-Cre mice. Note that the sparsity of the labeling is because AAV9 crosses the blood–brain barrier in the adult mouse at low efficiency, resulting in only ~1000–2000 cortical pyramidal neurons labeled in the entire brain at the chosen virus titer. (*B*) Top view of computationally "dissected" cortical hemispheres. (*C*) Coronal view of maximum projection of the computationally dissected left and right primary motor cortex (MOp) selected for region-specific analyses. (*D*) Convolutional network-based prediction of the GFP-labeled neuronal morphologies in the right MOp (note that we are continuing to train this convolutional network to further improve the accuracy). (*E*) Skeletonization of selected basal somatodendritic morphologies.

stochastic within-individual variability that further modifies possible phenotypes of X-linker neurodevelopmental disorders. These data, in conjunction with detailed behavioral studies, demonstrate that even modest XCI skewing can significantly impact phenotypic penetrance, providing an etiological insight into the source of phenotypic variability seen in heterozygous females with X-linked disorders.

In a second study focused on neurodevelopmental conditions, Julian Taranda identified a remarkable condition of an incomplete penetrance of structural, epileptiform, and behavioral phenotypes in a mouse model of 16p11.2 deletion and recurrent copy number variations (CNVs) linked to developmental delays, intellectual disability, autism, and childhood seizures. Approximately one-half of the genetically identical 16p11.2 deletion (16p11.2 del/+) mice showed a number of prominent phenotypes, including spontaneous epileptiform episodes of cortical activity, increased propensity to convulsant-evoked seizures, pronounced volume reductions in cortical areas correlating with increased convulsant-evoked local cortical activity, disrupted sleep, hyperactivity, and increased repetitive behaviors. In contrast, the remaining 16p11.2 del/+ mice showed only moderate brain volume changes and hyperactivity, but otherwise appeared phenotypically normal. The phenotypic discordance was observed in isogenic 16p11.2 del/+ mice within the same litter and across multiple generations and parents, which implies that the clustered phenotypes share a common mechanism of origin and their divergence reflects a bifurcating developmental choice occurring stochastically in the presence of the CNV, without the requirement for secondary mutations or environmental factors. These results also call for a reevaluation of the traditional two-cohort, mutant versus wild-type comparisons used in genetic mouse studies to date, as sensitive assays capable of measuring deficits in individual mutant mice are likely to reveal novel and clinically relevant phenotypes related to incomplete penetrance seen in all genetic neurodevelopmental disorders.

Summary

Motivated by the gap in knowledge presented by difficulties in quantifying cellular phenotypes across the entire mammalian brain, our lab has pioneered a series of high-throughput and high-resolution methods to enable new discoveries relevant to brain circuit functions, gender dimorphism, social behavior, brain evolution, and abnormal neurodevelopment caused by human genetic risk factors.

PUBLICATIONS

Kelly SM, Raudales R, He M, Lee JH, Kim Y, Gibb LG, Wu P, Matho K, Osten P, Graybiel AM, et al. 2018. Radial glial lineage progression and differential intermediate progenitor amplification underlie striatal compartments and circuit organization. *Neuron* **99:** 345–361 e344.

Turcan S, Makarov V, Taranda J, Wang Y, Fabius AWM, Wu W, Zheng Y, El-Amine N, Haddock S, Nanjangud G, et al. 2018. Mutant-IDH1-dependent chromatin state reprogramming, reversibility, and persistence. *Nat Genet* **50:** 62–72.

In Press

Mandelbaum G, Taranda J, Haynes TM, Hochbaum DR, Huang KW, Hyun M, Umadevi Venkataraju K, Straub C, Wang W, Robertson K, et al. 2019. Distinct cortical-thalamic-striatal circuits through the parafascicular nucleus. *Neuron* **102:** 636–652 e637.

NEURAL CIRCUITRY FOR SOCIAL COMMUNICATION

S. Shea A. Corona A. Nowlan
R. Dvorkin D. Rupert
C. Kelahan L. Shen

The broad goal of our laboratory's research is to understand how the brain detects and interprets sensory stimuli to guide flexible behavior. We are particularly interested in how neural activity and plasticity in olfactory and auditory brain circuits facilitate communication and social behavior. We are revealing neural mechanisms that allow organisms to detect and recognize familiar individuals, gather information about their identity and social status, and select appropriate behaviors. The smells and sounds emitted by a mouse during a social encounter convey a remarkable amount of information to its partners in the encounter. For example, they can signal the mouse's sex, genetic identity, reproductive state, levels of distress, or sexual interest. Their friends can even determine what food they recently ate and whether it was good. As you might imagine, proper interpretation of social signals is indispensable for survival and mating success.

This research program has two major goals. First, we want to identify the fundamental principles that govern how the brain adaptively controls complex behavior. Natural social behaviors are well suited for this purpose because the brain is exquisitely adapted through evolution to resolve and integrate social cues and link them to powerful behavioral responses. Second, we hope to pinpoint and repair neural circuitry defects that impair appropriate use of social information. Difficulty with social perception and cognition are core features of the autism spectrum disorders (ASDs). For example, patients may have trouble perceiving and interpreting communication gestures such as speech, facial expressions, and "body language." This broad feature is also evident in many mice that carry genetic variants identified in human ASD populations. Therefore, if we can ascertain the neural circuit substrates of social behavior in normal mice, we can make and test predictions for how the circuitry is affected in the mouse models. The results are likely to tell us more about the synaptic modifications that occur in human autism.

In recent years, we have focused on understanding the neural activity and circuitry in primary sensory brain areas that support adaptive behavior. We are now moving deeper into the brain to ascertain where the sensory data from those regions are collected and integrated into hormonal and electrical signals that promote appropriate behavioral choices. We are particularly interested in how all stages of neural processing are flexible according to experience and behavioral state.

Auditory Plasticity Sharpens Vocal Perception during Parental Learning

A. Corona, L. Shen

For several decades, we have understood that mice are continuously "speaking" (or vocalizing) to one another in a "language" that we have only just begun to understand. This vocal behavior has become a subject of renewed interest in recent years, partly because of the potential to understand its genetic basis and apply genetically encoded tools to dissect the underlying neural circuits. Many distinct vocalization types are produced by males, females, juveniles, and adults in a variety of behavioral situations. We would like to better understand the perceptual significance of these calls to the mice and how they are used to guide behavioral choices. Each vocalization type is a distinct message with an intended receiver and an intended response. Therefore, to be a reliable communication signal, the receiver must possess the neural circuitry to distinguish between different messages and link each of them to the appropriate behavioral response set. Identifying, monitoring, and perturbing this circuitry and its plasticity across development and life events is a major goal of our research program.

One form of vocalization for which we have a solid understanding of both the message and appropriate response is the ultrasonic distress vocalization (USV). Young mice, before vision and full mobility,

will occasionally become separated from the nest. This is stressful for them, and they will therefore call out to their mother with a very high frequency squeak. New mothers develop sensitivity to these cries and respond by moving toward their source to retrieve the pup. Moreover, females who have never given birth do not innately show approach responses to pups or their calls; however, they can learn to perform the behavior with experience. We refer to this group as "surrogates." Accurate maternal learning by mothers and surrogates is suspected to involve rewiring (or "plasticity") in the auditory cortex.

How do mice distinguish between vocalization types? They are constantly surrounded by many sounds, some of which may share pitch or timing features with important vocal signals. To which features do they attend? How much tolerance do they have for variability in those features? And do these limits change after maternal experience? Graduate student Luqun Shen is developing a sensitive, high-throughput method that we are using to identify those perceptual category boundaries to compare them to the features that define selective neural responses to USVs and monitor how both of these are altered over time as an animal gains maternal experience. His initial results show that in behavioral assays, over the first few days of maternal experience, mothers become increasingly more selectively attracted to pup calls as compared to band-limited noise that matches the temporal amplitude envelope of the calls. He is now recording responses to these stimuli from the auditory cortex of naive and maternally experienced mice.

The vast majority of studies of parental behavior in mice have focused entirely on females and maternal behavior. Nevertheless, graduate student Alberto Corona has shown in preliminary data that, under certain narrower conditions, males will show paternal behavior. Alberto is very interested in identifying contextual factors that favor paternal behavior. We hypothesize that vocal and olfactory cues emitted by the mother may be very important for encouraging paternal behavior. He is also interested in identifying brain regions in males that overlap with circuits that govern maternal behavior and brain regions that are uniquely activated during paternal behavior. Therefore, Alberto is performing behavioral experiments and whole-brain imaging (in collaboration with P. Osten at CSHL) on male and female mice to screen for paternal behavior-specific neural activity. He is particularly focused on regions that supply input to the medial preoptic area (MPOA), because it has been identified as a major hub for controlling parental behaviors. Indeed, Alberto finds that males, like females, show elevated expression of an activity-related gene in MPOA after interacting with pups. He has also identified the locus coeruleus (LC) as an important input to MPOA (see below). Therefore, he is planning to perform functional experiments on this pathway to directly assess its role in parental behavior.

Auditory Plasticity and Parental Behavior Are Impaired in a Mouse Model of Rett Syndrome

K. Krishnan, B. Lau, D. Rupert

Consistent with our objective of identifying impairments in neural circuitry that underlie social communication difficulties in ASD, postdoctoral fellows Billy Lau and Keerthi Krishnan led a collaboration with CSHL professor Dr. Josh Huang on vocal perception in a mouse model of Rett syndrome. This neurodevelopmental disorder is caused by loss of function of a gene called *Mecp2*. Female mice that possess only a single copy of *Mecp2* are not able to develop proficiency at gathering pups. We previously showed that this likely happens because MECP2 (the protein product of the gene) plays a critical role in maintenance and plasticity of the auditory cortex by acting on inhibitory interneurons.

Based on changes in the pattern of expression of certain molecular markers that we observed after maternal experience, we predicted that a network of inhibitory neurons expressing a protein called parvalbumin (PV) are central to auditory cortical plasticity. Notably, deletion of *Mecp2* only in this minority of neurons is sufficient to disrupt pup care. Therefore, Billy made neuronal recordings in awake behaving females of both genotypes that differed in their maternal experience. The data show that when a normal adult female mouse is first exposed to pups, her auditory cortex becomes "disinhibited" (i.e., there is suppression of the inhibitory network). Interestingly, we found that the removal of inhibition was selective for deep layers of the cortex, preferentially acted on late components of inhibition, and only affected responses to behaviorally meaningful calls, as opposed to synthetic tones. In contrast to typical mice, we observed no disinhibition

in *Mecp2*-deficient mice. The results of this study are consistent with our model that MECP2 regulates plasticity in adults and juveniles through its effects on inhibitory interneurons. We intend to submit a manuscript reporting these findings in spring 2019. Graduate student Deborah Rupert is continuing with behavioral and electrophysiological studies that will genetically dissect the contributions of several different classes of interneurons, as well as excitatory neurons.

Multisensory Integration in Maternal Retrieval Behavior

A. Nowlan, C. Kelahan

In addition to vocal signals, pup retrieval behavior depends heavily on olfactory cues from the pups. Surprisingly, these odors modulate auditory processing by mothers; however, the functional significance of this modulation is not well understood. One possibility is that pup odors serve as a contextual cue that elevates the behavioral significance of vocalizations, yet little is known about the neural pathways that integrate these two senses. It is also possible that olfaction and audition work together to trigger long-term synaptic modification in the auditory cortex so that mothers are more attentive to USVs. Together with technician Clancy Kelahan, graduate student Ally Nowlan has discovered a novel projection from the basal amygdala (BA) to the auditory cortex. This pathway appears to be well positioned to convey pup odors to the auditory cortex from the olfactory amygdala. By using newly developed optical methods for monitoring the activity of specific neural circuits in deep brain regions of freely behaving mice, Ally found that neurons in the BA that project to the auditory cortex are indeed active during pup retrieval. These cells show consistently elevated activity during olfactory investigation and search, and they directly respond to odors. These data suggest that the BA may provide the auditory cortex with access to odor information during parental behavior.

Neural Activity Signaling Emotion, Arousal, and Reward during Social Encounters

R. Dvorkin

Organisms are constantly being bombarded by an overwhelming number and variety of stimuli from all of their senses. Therefore, one of the greatest challenges faced by the nervous system is to make sensible and efficient choices about which stimuli to attend and remember. Decades of evidence have established that this calculation is achieved, in large part, with the contribution of a class of neurochemicals that enable neuronal communication and are collectively referred to as "neuromodulators." Neurotransmitters of this type generally do not participate in fast, moment-to-moment communication between neurons. Instead, their levels fluctuate more slowly and modify the function of larger groups of neurons or circuits. One important function of neuromodulators is to alert the brain to the possibility that something noteworthy is happening in the environment. Similarly, they may also broadcast a signal that the animal encountered a reward. This is likely to motivate the animal to examine the events leading up to that reward so that it may build a model of the actions and stimuli that predict positive outcomes. Of course, the other side of this coin is that the animal also wants to keep track of recent and impending negative events. In that sense, neuromodulators can inform the brain prospectively that something interesting is about to happen or retrospectively that something interesting recently happened. We are specifically interested in a small nucleus in the brainstem called locus coeruleus (LC), which sends neuromodulatory signals throughout the brain by releasing the chemical noradrenaline (NA). Evidence strongly suggests that in artificial, operant learning tasks, the activity of LC neurons both prospectively influences behavior and retrospectively signals recent stimuli of behavioral interest. However, very little is known about how these neurons participate in unstructured social interaction.

Postdoctoral fellow Roman Dvorkin is making recordings of individual neurons during actual social encounters between mice, including courtship and parental interactions. His goal is to monitor the activity of noradrenergic neurons in LC during these encounters. These neurons are likely responsive to rewarding social signals and may modulate coding of sensory data. Understanding the context-dependent activity patterns in LC is therefore critical to developing models for how this structure affects behavior. Roman has succeeded in chronically observing LC neural activity, day after day, as the mouse has repeated social interactions using two independent methods. He has made electrical recordings of individual LC neurons,

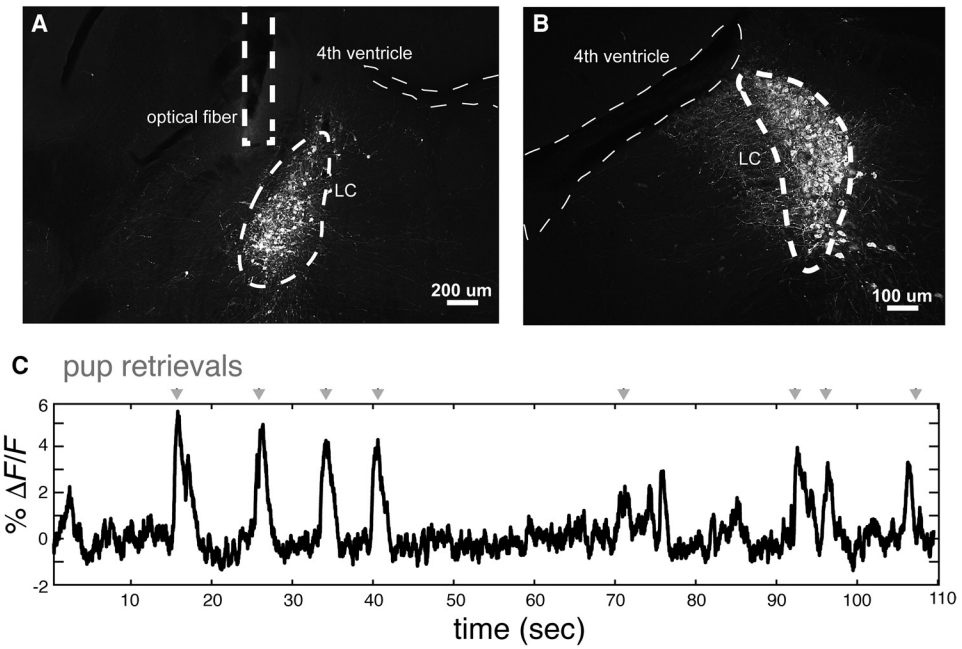


Figure 1. Phasic bursts of noradrenaline are released by locus coeruleus (LC) each time a maternal caregiver retrieves a pup. (*A*,*B*) Fluorescent photomicrographs showing specific labeling of noradrenergic neurons in LC with the genetically encoded activity sensor GCaMP6. Labeling was achieved by injecting a Cre-recombinase-dependent adeno-associated virus (AAV) into a mouse expressing Cre in dopamine β hydroxylase-positive neurons. The implanted optical fiber for measuring GCaMP6 signals is visible in *A*. (*C*) A plot of GCaMP6 fluorescence intensity in terms of Δ*F*/*F* in an actively retrieving maternal caregiver. Note that there is a sharp peak in the signal associated with each retrieval event as denoted by the gray arrowheads.

and also used optical techniques to image patterns of firing among the genetically defined noradrenergic cell class. His most dramatic and exciting result is that each time a maternal caregiver retrieves or interacts with a pup, there is a large, precisely timed burst of firing in LC neurons just before she makes contact with the pup (Fig. 1). This suggests that the release of a bolus of the powerful neuromodulator NA may help motivate and reinforce accurate maternal behavior, and it underscores the profound emotional content of interactions with offspring.

Dynamics of Neuronal Inhibition during Olfactory Learning

B. Cazakoff [in collaboration with D. Kepple and A. Koulakov, CSHL]

Behavioral state and previous experience can dramatically change sensory responses. Rather than a one-way street, sensory perception is probably better viewed as an ongoing and dynamic interaction between the brain and its environment. This interaction can allow sensory activity to be optimized in accordance with behavioral demands, and it can help the brain converge to a more accurate and efficient representation of complex stimuli. For the past several years, we have been addressing the contribution to this flexibility of a crucial group of inhibitory interneurons in the main olfactory bulb called granule cells (GCs). Our studies combine extensive collection of experimental data from GCs in awake and anesthetized mice and theoretical modeling of GC dynamics during learning in collaboration with the research group led by our CSHL colleague, Dr. Alexei Koulakov.

In 2018, we submitted a manuscript reporting the results of our collaboration with the Koulakov laboratory developing a model of GCs and their synaptic interactions with the mitral cells (MCs) that convey odor information to deeper brain structures. We speculate that the interaction between these circuits enables the main olfactory bulb (MOB) to implement an algorithm for "compressed sensing" that endows

the structure with an ability to compress and decompress its representations of many-dimensional stimuli. Compressing sparse multidimensional stimuli would allow the brain to most efficiently represent a vast and diverse chemical space with a limited number of receptors. This compressed sensing scheme has the added benefit that the MOB is capable of learning the identity of components in a mixture without ever experiencing the components in isolation. The GCs are absolutely central to this computation, and our model makes unexpected predictions about their activity during odor learning. We, therefore, compared these predictions with experimental data from GCs collected during a learning paradigm. We found that the empirical observations matched closely with the patterns predicted by our model. Thus, they corroborate our speculation that the MOB refines an implementation of a compressed sensing mechanism through a learning-dependent process.

PUBLICATIONS

Kepple D, Cazakoff BN, Demmer HS, Eckmeier D, Shea SD, Koulakov AA. 2018. Computational algorithms and neural circuitry for compressed sensing in the mammalian main olfactory bulb. bioRxiv doi:10.1101/339689

HORMONAL REGULATION OF GENE EXPRESSION IN THE BRAIN

J. Tollkuhn R. Bronstein B. Gegenhuber M. Wu

Our lab seeks to understand the mechanisms that shape and regulate sex differences in the brain. Females and males differ in many behaviors and are differentially affected by mental health disorders, but the distinct developmental trajectories that give rise to these sex differences remain poorly understood. Much of our knowledge about the cellular and molecular differences between the sexes in the mammalian brain has been obtained through studies of the hormonal regulation of the differentiation and function of neural circuits underlying innate, sex-typical behaviors and physiology in rodents. Paradoxically, estrogen is required to both feminize and masculinize the brain. Males undergo a transient perinatal testosterone surge, and this circulating testosterone is converted to estradiol (the most abundant endogenous estrogen) locally in the brain. Treating females with estradiol at birth irreversibly masculinizes both adult behaviors and gene expression patterns, suggesting that perinatal estrogen directs gene regulatory events that organize persistent sex differences in the brain. The receptors for estrogen, estrogen receptor alpha (ERα) and estrogen receptor beta (ERβ), are nuclear receptor transcription factors that are recruited to DNA in the presence of hormone. Although the regulatory strategies used by estrogen receptors are well understood in the context of breast cancer, they remain obscure in the brain. Defining the gene programs regulated by estrogen in the developing and adult brain is the focus of our current research program. Our central hypothesis is that hormone-responsive genes underlie sex differences in the incidence and etiology of psychiatric and neurological diseases.

Identification of Sex-Specific Gene Programs in the Brain

To understand how estrogen regulates gene expression in the brain, we have performed transcriptomic profiling of two interconnected sexually dimorphic brain regions that regulate social behaviors: the bed nucleus of the stria terminalis (BNST) and the medial amygdala (MeA). To circumvent the heterogeneity of the mammalian brain, we are performing our RNA-sequencing (RNA-Seq) experiments specifically in ERα+ neurons. This specificity is achieved through use of the RiboTag mouse allele, which permits Cre-defined tagging of ribosomes, followed by purification of the affiliated translating mRNAs. We identified novel sex-biased genes in mouse pups. To dissociate sex differences in gene expression from sex differences in circulating gonadal hormones in adult animals, we have gonadectomized animals of both sexes and treated them with either estradiol or vehicle for 4 hours to identify genes that are rapidly induced by estradiol treatment. We have validated several brain-specific targets of ERα by in situ hybridization, as well as select coactivators and corepressors that have been previously shown to interact with ERα in nonneuronal tissues. Surprisingly, this histological validation has revealed genes that show striking induction by estradiol in one sex more than the other. We are now working to validate direct interactions between ERα and sex-biased cofactors through the use of rapid immunoprecipitation mass spectrometry (RIME).

Sex Differences in the Epigenome

Our lab is particularly interested in identifying the gene regulatory strategies used by ERα in neurons. Chromatin immunoprecipitation with sequencing (ChIP-seq) data in cell lines and nonneural tissues has revealed that ERα primarily occupies distal regulatory elements that do not contain a canonical estrogen response element, and ERα is frequently recruited to DNA through other cell type–specific transcription factors. As a first step in identifying such factors in neurons, we performed an assay for transposase-accessible chromatin with sequencing (ATAC-seq) to reveal genomic regions that become accessible in response to acute estradiol treatment. Motif analysis of estradiol-opened regions can be used to infer the identity of

transcription factors that occupy these opened regions. To gain access to chromatin from our limited populations of ERα neurons in the BNST and MeA, we are using a protocol termed isolation of nuclei tagged in specific cell types (INTACT). This approach uses a Cre-inducible nuclear envelope tag to permit purification of chromatin from genetically defined cell types.

From these experiments, we have identified NFI and Foxo motifs as significantly increased in genomic regions that are open in the presence of estradiol compared with a vehicle control. We previously identified members of these families as showing increased expression in male BNST compared with female, suggesting that these factors are causal for sex differences in gene regulation in the brain. We are now working to determine the genomic occupancy patterns of these factors in BNST through the use of cleavage under targets and release under nuclease (CUT&RUN), a recently developed method that provides results similar to ChIP-seq, but with 1000-fold fewer cells required. Combined with the RNA-Seq experiments described above, these results are the first examples of sex differences in the expression of transcription regulators in the brain that are not steroid hormone receptors.

Estrogen Receptor Genomic Binding in the Mouse Brain

Previously, the transcription factor ERα has been found necessary for the programming and display of sex-typical social behaviors in mice. We aim to understand how ERα regulates the expression of genes involved in these behaviors and, ultimately, whether dysregulation of sex-specific transcriptional programs contributes to the pathology of neurodevelopmental disease. Because of technical limitations of ChIP-seq, the genomic targets of ERα in the brain remain unknown. To answer this question, we are applying the recently described CUT&RUN method to nuclei purified from brain tissue. CUT&RUN uses a Protein A-micrococcal nuclease fusion protein to cleave DNA around transcription factor–binding sites in situ. In addition, we are combining CUT&RUN with the INTACT technique, which uses a Cre-dependent nuclear envelope tag, to study ERα genomic binding in genetically defined neuron populations. Thus far, we have found that CUT&RUN can capture well established ERα-binding sites in the MCF7 cell line with fewer than 1% of the number of cells required for ChIP-Seq. Currently, we are determining estrogen- and sex-dependent ERα binding sites in two brain areas crucial to sex-typical social behaviors—the BNST and MeA. For the first time, we will identify the *cis*-regulatory elements, gene targets, and interacting transcription factors of ERα in the mouse brain. We anticipate that our approach may be broadly used to determine transcription factor–binding sites in defined neuron populations throughout the brain.

PUBLICATIONS

Ahrens S, Wu MV, Furlan A, Hwang GR, Paik R, Li H, Penzo MA, Tollkuhn J, Li B. 2018. A central extended amygdala circuit that modulates anxiety. *J Neurosci* **38:** 5567–5583.

Manoli DS, Tollkuhn J. 2018. Gene regulatory mechanisms underlying sex differences in brain development and psychiatric disease. *Ann NY Acad Sci* **1420:** 26–45.

CIRCUITS UNDERLYING DECISIONS IN AUDITORY CORTEX

A. Zador V. Aguillon, B. Burbach, X. Chen, M. Dworjan, E. Fong, A. Funamizu, S. Ghosh, G. Henry, L. Huang, C. Krasniak, Y. Li, S. Lu, A. Reid, C. Stoneking, Y-C. Sun, A. Vaughan, W. Wadolowski, L. Yuan, H. Zhan, A. Zhang

Our laboratory is interested in how neural circuits underlie cognition, with a focus on the cortical circuits underlying auditory processing and decision-making. To address these questions, we use a combination of molecular, electrophysiological, imaging, and behavioral approaches. In addition, we have developed a method that allows us to use high-throughput DNA sequencing to determine the wiring diagram of the brain at single-neuron resolution rapidly and efficiently, at low cost.

Correlating Gene Expression to Long-Range Neuronal Connectivity

X. Chen, Y-C. Sun

Neurons in the cerebral cortex communicate with neurons in other brain areas through diverse long-range projections. Correlating projections with other neuronal properties, such as the myriad transcriptomic cell types recently found using single-cell RNA-Seq, is crucial in understanding the organization and function of neuronal circuits. We have previously developed BARseq, a high-throughput method for mapping long-range neuronal projections using in situ sequencing. We are currently exploring two strategies to correlate high-throughput neuroanatomical data with gene expression. The first strategy utilizes single-cell RNA-Seq to read out both endogenous gene expression and barcodes at cellular resolution. We have performed proof-of-principle experiments to demonstrate simultaneous sequencing of transcriptome and barcodes and are currently optimizing cell dissociation protocols to increase throughput. The second strategy involves combining BARseq with in situ gene detection methods, including targeted in situ sequencing and fluorescence in situ hybridization (FISH). Using this strategy, we were able to capture the known relationship between gene expression and long-range projections of major classes of cortical projection neurons. Beyond the major classes, we identified additional gene correlates of neuronal projections, but these genes do not necessarily distinguish transcriptomic types of neurons. Our results suggest that long-range projections and transcriptomic types may be organized orthogonally in the cortex beyond the major classes.

Neural Encoding and Decoding in Perceptual Decision-Making

A. Funamizu

Neurons in the auditory cortex represent ("encode") information by spike activity, and this activity is "decoded" by downstream areas. Here we compare the encoding and decoding of information by neurons in the auditory cortex. We trained head-fixed mice on an auditory frequency discrimination task (based on Marbach and Zador, bioRxiv doi:10.1011.073783 [2016]), in which either (i) the stimulus probability or (ii) the reward size for category A and category B trials changes in blocks. We then imaged populations of neurons in auditory cortex using two-photon calcium imaging.

In the task, both the stimulus probability and reward amount biased mice toward choices associated with high-probability stimuli or large reward, respectively. Our proposed reinforcement learning model suggested that mice integrated the auditory stimulus and prior (reward expectation) to optimize the behavior. The auditory cortical neurons increased the sound responses when the preferred stimulus was associated with a high stimulus probability or large reward amount, suggesting the neural encoding of prior. In contrast, the sound decoding of neurons was stable irrespective of varied outcomes or stimulus probabilities, partly because the prior encoding did not change

the sound categorization threshold in the auditory cortical neurons. Our results propose a role of sensory cortex in which the downstream areas reliably read out sensory information from the prior modulated neurons.

Corticostriatal Plasticity Underlying Learning and Reversal of Auditory–Motor Association in Mice

S. Ghosh

Animals use complex sensory cues from their environment to make a variety of decisions in their lives. Such behavior requires integration of sensory discrimination, decision-making, and appropriate motor actions. We used an auditory discrimination task to understand the brain circuits involved in such decision-making, and how these circuits evolve during learning of sensorimotor associations. Previous studies from our lab have shown that the connections between auditory cortex and the auditory striatum in rats are instrumental for an animal to perform this task (Znamenskiy and Zador, *Nature 497:* 482 [2013]). Moreover, learning of this discriminatory task results in formation of a memory of the learned association in this circuit (Xiong et al., *Nature 521:* 348 [2015]). These findings suggest that parts of the striatum receiving predominantly sensory inputs might be involved in promoting contralateral movements. We tested this model in a reversal paradigm using the "tonecloud" task, in which an animal initially trained to form a specific auditory–motor association was then forced to reverse its association to obtain reward successfully. We find that mice can indeed learn to reverse these associations, taking comparable training times and reaching similar performances in both training epochs. We then investigated the pattern of plasticity in this circuit following the reversal and found that reversing the association does not result in a simple reversal of the memory trace. In fact, we observe a strong persistence of the synaptic plasticity pattern that only reflects the initial association. Our results are consistent with a model in which forming new memories does not erase previous ones. These results suggest that the sensory striatum might not be simply transducing sensory information into a contralateral motor output, and it raises the question of how task-related context information is integrated in this circuit.

Improvements to the Barcode Carrier Scheme

G. Henry

Efficient axonal transport of barcode RNA is a central problem in mapping neuronal projections by DNA sequencing. Both the MAPseq and BARseq methods rely on a carrier protein called MAPPnl, which utilizes a short localization domain found in the Neurexin protein to target the phage nLambda RNA binding protein to axonal terminals. Because barcode RNAs contain the boxB sequence bound by nLambda, the carrier protein mediates anterograde barcode transport. In the past year we improved the efficiency of this system by screening a panel of alternate second-generation carrier proteins. To do this we replaced the MAPP fusion protein with a number of protein sequences known to target axons. These included Tau, Neurexin, SNAP25, Syntaxin1A, Synapsin, Synaptophysin, Synaptotagmin, VAMP1, VAMP2, SV2A, and the Ras Gap palmitoylation signal. A MAPseq assay was used to measure the ability of each candidate to enhance both projection frequency and strength. Compared to the control MAPPnl, VAMP2 shows a fivefold enhancement in projection frequency and a 1.8-fold to 2.0-fold enhancement in projection strength.

Both carrier protein and barcode RNA expression are mediated by Sindbis virus, a well described single-stranded RNA alphavirus. Though this system permits extremely high levels of polycistronic expression and large payloads, it is incompatible with the ubiquitous Cre-driver scheme that is the most popular means of achieving cell type–specific expression in mice because the recombinase acts on DNA. A solution to this problem involves the development of a barcode carrier scheme in adeno-associated virus (AAV). Because extremely high levels of expression cannot be achieved with a DNA virus, we have designed and tested a scheme that in principle should generate highly stable barcode RNA. To do this, we have inserted a barcode cassette into the intronic sequence of tRNAs that are known to be spliced (Arg, Ile, Leu, Ser, and Tyr). Removal of the barcoded intron in the nucleus generates a circularized tricRNA that has been shown by others to enter the cytoplasm. Detectable levels of barcoded intron are obtained from $tRNA^{Arg}$, and we are currently generating viruses that co-express the VAMP2nl carrier protein.

High-Throughput Mapping of Mesoscale Connectomes in Individual Mice

L. Huang

Brain function is determined by connectivity among brain areas, and disruption of this connectivity leads to neuropsychiatric disorders. Understanding connectivity is essential to modern neuroscience, but mesoscale connectivity atlases are currently slow and expensive to generate, exist for few model systems, and require pooling across many brains. To overcome the difficulties above, we developed a novel method, muMAPseq (multisource multiplexed analysis of projections by sequencing), which leverages barcoding and high-throughput sequencing to generate atlases from single animals in a few weeks and at low cost. We apply muMAPseq to tracing the neocortical connectome of individual mice and demonstrate high reproducibility and accuracy. Applying muMAPseq to the mutant BTBR mouse strain, which lacks a corpus callosum, we recapitulate its known connectopathies, indicating the potential of muMAPseq in comparative connectomics. muMAPseq allows individual laboratories to generate atlases tailored to individuals, disease models, and new model species and will facilitate quantitative comparative connectomics, permitting examination of how age, sex, environment, genetics, and species affect neuronal wiring.

The International Brain Laboratory

C. Krasniak

A key challenge in modern neuroscience is understanding how neural systems work together as a whole to support adaptive behavior. Adaptive behavior requires processing sensory information with focused attention, reaching decisions, acting, and learning from the results of those actions. These require the brain to combine a vast array of information from prior experience, current sensory stimuli, and internal and environmental contexts. These computations involve dynamic interactions between millions of neurons within local circuits and across many brain regions. Understanding these processes is a problem with a scale and complexity that far exceed what can be tackled by any single laboratory and that demands computational theory to be interwoven with experimental design and analysis in a manner not yet achieved. To overcome these challenges, we have created a virtual laboratory, unifying a group of 21 highly experienced neuroscience groups distributed across the world. We are using a single standardized visual decision-making task in labs across the world to ensure our experiments are repeatable. As a first step, we are recording throughout as much of the mouse brain as possible in behaving mice to address how local circuits and teams of distal areas combine to form a single behavior.

Pairing Gene Expression and Spatial Location to Explore Whole-Brain Patterns of Transcription

S. Lu [in collaboration with J. Gillis, CSHL]

What is the relationship between gene expression and brain areas defined by conventional neuroanatomy? To study this relationship, we are using data from a recently developed spatial sequencing approach—spatial transcriptomics—in conjunction with the Allen Brain Atlas adult mouse in situ hybridization data. Combining these data will provide a foundation for understanding how gene expression relates to neuronal connectivity.

Mapping Neural Projection In Situ by BARseq

L. Yuan, X. Chen, Y-C. Sun

A neural projection map in individual brains provides an anatomical basis to study a wide range of topics in neuroscience like behavior and neurological disorders. MAPseq is a novel projection mapping technique, in which individual neurons are barcoded with unique nucleotide sequences, and both soma location and axonal projections can be identified by barcode sequencing. Thus, MAPseq enables high-throughput mapping of neural projections in various brain regions with single-cell resolution within a single brain.

However, MAPseq utilizes bulk sequencing for barcode readout; thus, it has limited spatial resolution for both somatic and axonal barcodes. Because the brain is highly heterogeneous, neurons can be further divided into different groups based on soma locations and projection patterns even within the same brain region. Therefore, it is crucial to capture and study these fine-scale cell-specific projections. To preserve the soma

location during barcode sequencing, BARseq was developed recently to directly sequence somatic barcode on individual brain sections. Combining both techniques—soma location mapping by BARseq and projection mapping by MAPseq—mapping of the mouse primary auditory cortex was successfully achieved. The projection map is aligned with previous findings, and more importantly it provided novel insights into projection pattern on the population level (Chen et al. 2018a). However, this approach still has spatial limitation in projection sites so that it is not able to resolve projection preference to different groups or functional regions within a brain region. To this end, we are currently optimizing BARseq to sequence axonal barcode in situ.

At present, we are tackling two major challenges for axonal BARseq: limited detection sensitivity and low signal-to-noise ratio. The current optimized protocol enables sequencing and base-calling dendritic/axonal barcode for more than 10 bases. Individual dendritic/axonal barcodes are also able to be matched with corresponding barcoded soma, and partial neural morphology can be visualized with single-cell resolution in a single brain. Next, we will focus on increasing the sensitivity of individual barcode detection. We predict that these improvements can make axonal BARseq a powerful and practical tool for high-resolution projection mapping for scientific discoveries.

Mapping Synaptic Connectivity In Situ by Proximity Ligation Assay (PLA)-Dependent BARseq

H. Zhan

We have developed a novel method we call SYNPLA (synaptic proximity ligation assay), a synapse-specific, high-throughput, and potentially brain-wide method capable of detecting circuit-specific learning-induced synaptic plasticity. PLA is a highly sensitive, high-throughput, and specific biochemical method that reliably detects the close (<40-nm) juxtaposition of two proteins in situ. The key advantages over traditional immunostaining are that the absolute requirement for a pair of proteins confers very high specificity, and the amplification renders each positive signal very bright and therefore easily distinguishable from the background. We have demonstrated that PLA can specifically detect the interaction of NRXN (Neurexin) and NLGN (Neuroligin) across the synaptic cleft. Our current work focuses on increasing the sensitivity and specificity of this approach.

Comparing Roles of Primary Visual and Auditory Cortices in Sensory Decisions

A. Zhang

Sensory information of different modalities is processed and extracted through distinct but analogous series of hierarchically organized brain areas. Visual information travels from the retina to the lateral geniculate nucleus of the thalamus and then on to primary visual cortex (V1). Auditory information travels similarly from the cochlea, through the medial geniculate nucleus of the thalamus, to primary auditory cortex (A1). Feature-encoding principles are analogous at these early stages of processing, and perceptual decision-making work has begun to elucidate the mechanisms by which such features can be associated with appropriate motor outputs. However, because of varied stimulus complexities, task demands, and animal models, it has been impossible to directly compare the mechanisms and pathways that underlie visual-based associations to those that underlie auditory-based associations.

To overcome these challenges, we have developed a novel visual discrimination task that asks animals to judge the dominant spatial location of the stimulus. This task is directly analogous to an existing auditory discrimination task used in the lab that asks animals to judge the dominant frequency of an auditory stimulus. Animals acquire the visual task quickly and to high levels of performance, and more detailed behavioral experiments show that subjects consistently use a shortcut to solve the task. Their strategy allows us to divide the presented stimulus into "used" and "unused" portions. We use tetrode recordings to investigate the information carried by neurons in V1 in these animals, and we find preliminarily that there exist differences in the coding statistics of subgroups of neurons depending on whether they preferentially respond to the used or unused stimulus portions. We continue to investigate how stimulus use or behavioral context shapes stimulus encoding in the early visual system in this task and in a modified task in which the animal's strategy is more constrained. We are also in the process of answering our central question of how primary

sensory areas of different modalities contribute within a common perceptual decision-making paradigm. To do this, we are using optogenetic techniques in V1 similar to those previously used in A1 to compare the contributions of these two primary sensory areas and their projections to sensory decision-making.

PUBLICATIONS

Chen X, Zhan H, Kebschull JM, Sun Y-C, Zador AM. 2018a. High-throughput mapping of long-range neuronal projection using in situ sequencing. bioRxiv doi:10.1101/294637

Chen X, Sun YC, Church GM, Lee JH, Zador AM. 2018b. Efficient in situ barcode sequencing using padlock probe-based BaristaSeq. *Nucleic Acids Res* **46:** e22.

Han Y, Kebschull JM, Campbell RAA, Cowan D. Imhof F, Zador AM, Mrsic-Flogel TD. 2018. The logic of single-cell projections from visual cortex. *Nature* **556:** 51–55.

Kebschull JM, Zabor AM. 2018. Cellular barcoding: lineage tracing, screening and beyond. *Nat Meth* **15:** 871–879.

In Press

Shuvaev SA, Baserdem B, Zador AM, Koulakov AA. 2019. Network closing using DNA barcodes. *Proc Natl Acad Sci* **116:** 9610–9615.

Zador AM. 2019. A critique of pure learning: what artificial neural networks can learn from animal brains. *Nat Comm* (in press).

PLANT BIOLOGY

David Jackson and colleagues study genes and signals that regulate plant growth and architecture. They are investigating a unique way in which plant cells communicate by transporting regulatory proteins via small channels called plasmodesmata. These channels, which direct the flow of nutrients and signals through growing tissues, are regulated during development. The team discovered a gene encoding a chaperonin, *CCT8*, that controls the transport of a transcription factor called SHOOTMERISTEMLESS (STM) between cells in the plant stem-cell niche, or meristem. STM is critical for stem-cell maintenance, and studies of the *CCT8* gene indicate that movement of STM between cells is required for this function. The lab also continues to identify other genes that control plant architecture through effects on stem-cell maintenance and identity, and their work has implications for crop yields. Recent examples include discovery of a subunit of a heterotrimeric G protein that is conserved throughout animals and plants, and their studies indicate that this gene controls stem-cell proliferation. They have found that in plants, the G protein interacts with a completely different class of receptors than in animals. Their discovery helps to explain how signaling from diverse receptors is achieved in plants. This year, they also showed that weak mutations in one of the receptor proteins can enhance seed production in maize, which could lead to yield increases. Separately, the lab has characterized system-wide networks of gene expression using next-generation profiling and chromatin immunoprecipitation methods that have led to many new hypotheses regarding developmental networks controlling inflorescence development. They are also developing a collection of maize lines that can drive expression of any reporter or experimental gene in any tissue type—tools of great interest to maize researchers that are being made available to the broader scientific community, enabling experiments never before possible in crop plants.

Zachary Lippman's research focuses on the processes of flowering and flower production in plants, which are major contributors to reproductive success and crop yield. Specifically, Lippman's research program integrates development, genetics, genomics, and gene editing to explore the mechanisms that determine how plant stem cells become shoots and flowers. The lab takes advantage of extensive natural and mutant variation in inflorescence production and architecture in tomato and related nightshade species (e.g., potato, pepper, ground cherry) to explore how differences in these processes explain the remarkable diversity in the architectures of flower-bearing shoots (inflorescences) observed in nature and agriculture. Recent discoveries regarding the genes and networks underlying this diversity have led to broader questions about the significance of genomic structural variation, gene redundancy, and epistasis in development, domestication, and breeding. Based on their fundamental discoveries, the group is developing and applying innovative concepts and tools for crop improvement.

Epigenetic mechanisms of gene regulation—chemical and conformational changes to DNA and the chromatin that bundles it—have had an important impact on genome organization and inheritance and cell fate. These mechanisms are conserved in eukaryotes and provide an additional layer of information superimposed on the genetic code. **Robert Martienssen**, a pioneer in the study of epigenetics, investigates mechanisms involved in gene regulation and stem-cell fate in yeast and model plants, including *Arabidopsis* and maize. He and his colleagues have shed light on a phenomenon called position-effect variegation, caused by inactivation of a gene positioned near densely packed chromosomal material called heterochromatin. They have discovered that small RNA molecules arising from repeating genetic sequences program that heterochromatin. Martienssen and colleagues have described a remarkable process by which "companion cells" to

sperm in plant pollen grains provide them with instructions that protect sperm DNA from transposon damage. They found that some of these instructions, or epigenetic marks, could be inherited in the next generation. These marks, and the small RNAs responsible for guiding them, can sense the number of chromosomes inherited from pollen and may allow *Arabidopsis*, a flowering plant, to produce egg cells without meiosis—an important step toward a long-time goal of plant breeding: generating clonal offspring to perpetuate hybrid vigor. The lab has also shown that when RNA polymerase II has transcribed a stretch of DNA, the RNA interference mechanism causes the enzyme to release its hold on the DNA and fall away. This allows the replication fork to progress smoothly and the DNA strands to be copied; histone-modifying proteins, which follow right along, establish heterochromatin. Martienssen's group also continues to work on problems related to the creation of plant-based biofuels. As part of a collaborative project to generate a high-quality full genome map of the oil palm plant, Martienssen and his colleagues identified a transposon whose modification controls the yield of oil palm trees. This discovery will increase yields and should lessen the environmental burden of oil palm production, which often threatens already endangered rainforest lands.

Plants and animals interact with their environment. Because plants are incapable of moving around, they are sensitive to their surrounding environment and modify their development according to external signals. Plants face variability in growth conditions—temperature, light quality and quantity, herbivores, pathogens, water availability, etc. Yet, plants respond to these biotic and abiotic factors and survive substantial fluctuations in their environment. Plants also must balance the range of potential threats and benefits confronting them and should make appropriate decisions on resource allocation. Such adaptability is essential given the sessile nature of the plants. The mechanisms that underlie this adaptability likely involve complex signaling to generate the appropriate response. In some adaptive responses (e.g., when the plants have to cope with climate change and increased competition for light), there is a decrease in productivity (yield, biomass) as the plant relocates resources to better adapt.

Ullas Pedmale's lab seeks to determine the mechanisms behind how a plant perceives and successfully adapts to its environment. They also aim to understand how a plant must integrate intrinsic and extrinsic cues and "decide" how best to respond to environmental cues. Understanding how plants deal with and respond to a multitude of environmental signals could help to develop crops that cope with unfavorable growth conditions without significant changes in yield.

DEVELOPMENTAL BIOLOGY—STEM CELL SIGNALING AND CROP ARCHITECTURE

D. Jackson G. Carver, H. Claeys, E. Demesa-Arevalo, C. Fugina, M. Kitagawa, L. Liu, A. McKane, T. Skopelitis, E. Smith, M. Venezia, Y. Wen, Q. Wu, F. Xu, X. Xu

Our research asks how the growth of plants is controlled, with the ultimate goal of improving crop yields. We identify genes, signals, and pathways that regulate plant architecture and development. Like all organisms, plants grow and develop by carefully controlling the passage of signals between cells. We are interested in discovering the signals that carry key information, how they are transmitted, and how they function. A major focus has been identification of genes that control stem cell signaling. In the past year we reported a new way to increase kernel number in maize by tweaking expression of signaling proteins called G proteins in stem cells. We also identified a new way in which plants control branching, through use of enzymes that function in sugar metabolism, but our research suggests these enzymes play a more important role in the cell nucleus. In other studies, we continue to expand our use of CRISPR genome editing to uncover new gene functions as well as to increase allelic diversity to improve maize yield traits.

The Control of Meristem Size in Maize

Q. Wu, F. Xu [in collaboration with B. Il Je, Busan University; H. Furukawa, CSHL; B. Yang, Danforth Center; R. Meeley, DuPont Crop Genetics]

All plant organs derive from populations of stem cells called meristems. These stem cells have two purposes: to divide and maintain themselves and to give rise to daughter cells, which will differentiate into plant organs. Consequently, meristems must precisely control the size of the stem cell niche via a network of positive- and negative-feedback signals. A loss of function in a negative regulator of stem cell fate results in an enlarged or fasciated meristem phenotype and a dramatic alteration in the morphology of the maize ear and tassel. Maize is an excellent genetic model system for these studies because of a large collection of developmental mutants and a diverse genome. Our lab uses genetics to identify key regulators of stem cell homeostasis and meristem size. Two previously cloned mutants, *fasciated ear2* and *thick-tassel dwarf1*, encode orthologs of the *Arabidopsis thaliana* genes *CLAVATA1* and *CLAVATA2*, indicating the well-known *CLAVATA-WUSCHEL* regulatory feedback loop is conserved in monocot crops. However, little else is known about the control of this important developmental process in maize. Here, we describe progress in identifying additional genes contributing to stem cell homeostasis.

A common class of proteins that signal directly downstream of cell surface receptors is the heterotrimeric G proteins, consisting of Gα, Gβ, and Gγ subunits. Our previous work found that the maize *COMPACT PLANT2* (*CT2*) gene, which encodes the α subunit of a heterotrimeric GTPase (Gα), functions in the CLAVATA pathway to control meristem size through its interaction with the FEA2 receptor. To further study the mechanism of G-protein signaling in meristem development, we knocked out the sole *Gβ* gene of maize, *ZmGB1,* using CRISPR-Cas9. Using a functional translational fusion of *ZmGB1* and *YFP* under the control of its native promoter, we localized the protein to both cytosol and plasma membrane throughout the meristem, suggesting that it acts in receptor complexes, consistent with our previous findings for Gα. However, to our surprise, we found that the *Zmgb1^{cri}* null mutant plants died at an early stage of seedling development, with overaccumulation of H_2O_2 and salicylic acid, constitutive activation of MAP kinases, and up-regulation of *PATHOGENESIS-RELATED* immune marker genes. These results suggest that *ZmGB1* mutation causes autoimmune symptoms. We therefore crossed the *Zmgb1^{cri}* null alleles into different lines to see if the lethal phenotype can be suppressed, and indeed found that it was partially suppressed in a CML103 background. In this lethality-suppressed background, the *Zmgb1^{cri}* mutants

had significantly larger shoot apical meristems (SAMs) and fasciated ear and tassel meristems, suggesting that ZmGB1 plays an important role in meristem development. We mapped the suppressor of *Zmgb1* lethality in CML103 and identified a disease resistance (*R*) gene as a candidate, suggesting that *Gβ* acts as an immune sensor, or "guardee," in maize—unlike in *Arabidopsis*, where *Gβ* mutants are viable.

Concurrently, by map-based cloning of a new fasciated mutant, *fea*-GN183,* by chance we identified a second allele of *Zmgb1* with a single amino acid change in a conserved residue of one of the WD40 domains. Yeast-three-hybrid experiments revealed that the Zmgb1^{fea*183} protein was impaired in forming a complex with the alpha subunit, CT2, and the gamma subunit, ZmRGG1. *Zmgb1^{fea*183}* mutants had fasciated ears and thick tassels and were dwarf with enlarged SAMs. They also developed necrotic lesions, reminiscent of the *Zmgb1^{cri}* null mutants, and failed to complement *Zmgb1^{cri}* mutants, suggesting that they are allelic. The fact that the *Zmgb1^{fea*183}* alleles were viable allowed double-mutant analysis, which suggested that ZmGB1 functions in the same pathway as CT2, and *fea2* mutants were epistatic to *Zmgb1*, so ZmGB1 and CT2 might function together downstream of FEA2. In contrast, *Zmgb1^{fea*183}; fea3* double-mutant meristems were massively enlarged, suggesting that ZmGB1 functions in a pathway parallel to FEA3. Further studies will focus on understanding the roles of Gβ in maize development and immune responses.

We also continue to study the mechanism of action of *FEA3,* which encodes a predicted leucine-rich repeat receptor-like protein related to *FEA2*. FEA3 is of particular interest because it is expressed in the organizing center of the SAM and in leaf primordia, and expression of maize WUSCHEL, a marker for the stem cell niche organizing cells, spreads downward in *fea3* mutants, which is strikingly different from its response in the known *CLAVATA* stem cell mutants.

To further understand the role of FEA3 in stem cell signaling, we are using immunoprecipitation-mass spectrometry (IP-MS) with FEA3 tagged plants. Our preliminary analysis found some candidate interactors, which we are now validating by co-IP experiments in *Nicotiana benthamiana*. We are also using CRISPR to knock out FEA3 homologs, to explore whether they have related functions. We already obtained different CRISPR alleles for four FEA3 homologs; however, no obvious phenotype was observed. By reanalyzing the phylogeny based on new annotations in the improved maize B73 v4 genome, we identified an additional *FEA3* homolog, which we have also knocked out by CRISPR-Cas9 to overcome genetic redundancy and are screening for phenotypes.

Genetic Redundancy in Circuits Controlling Meristem Development

L. Liu, E. Demesa-Arevalo, F. Xu, T. Skopelitis [in collaboration with B. Il Je, Busan University; M. Bartlett, University of Massachusetts Amherst; Z. Nimchuck, University of North Carolina at Chapel Hill; B. Yang, Danforth Center; Z. Lippman, CSHL]

The CLE (CLAVATA3/Endosperm surrounding region-related) peptides are fundamental players in meristem maintenance in plants, acting as mobile signals that establish feedback circuit signaling between stem cell differentiation and division. Disruption of this pathway generates overproliferation in meristems, or fasciation. In our laboratory, we have described different fasciated mutants encoding leucine-rich repeat receptor-like kinases or receptor-like proteins (LRR-RLKs or RLPs); however, the signals perceived by many of these receptors remain elusive. Genetic evidence in maize suggests a divergence in signaling pathways controlling meristem size. The putative orthologs in maize for CLAVATA1 (THICK TASSEL DWARF1, TD1), CLAVATA2 (FASCIATED EAR2, FEA2), and FASCIATED EAR3 (FEA3) have synergistic effects, suggesting that additional pathways control meristem size. Forty-nine CLE peptide genes have been identified in maize, suggesting either specialization or redundancy in these ligands.

To analyze the role of CLE peptide ligands in maize meristem regulation and their involvement in redundant circuits, we are generating CRISPR knockouts. We first analyzed expression patterns from publicly available data sets combined with our transcriptional profiles from meristematic tissues and identified 31 candidates expressed in meristems. CRISPR frameshift mutations have been induced in 27 of the 31 CLEs, and mutants from the same subgroup are being crossed together to evaluate their functional redundancy. To identify *Zm*CLE peptides that regulate meristem size in a redundant way, we performed RNA-Seq in dissected inflorescence meristems from mutants in the maize CLV3 ortholog, *Zmcle7^{cr}*. Only two CLE genes, *Zm*FCP1 and *Zm*CLE7, were

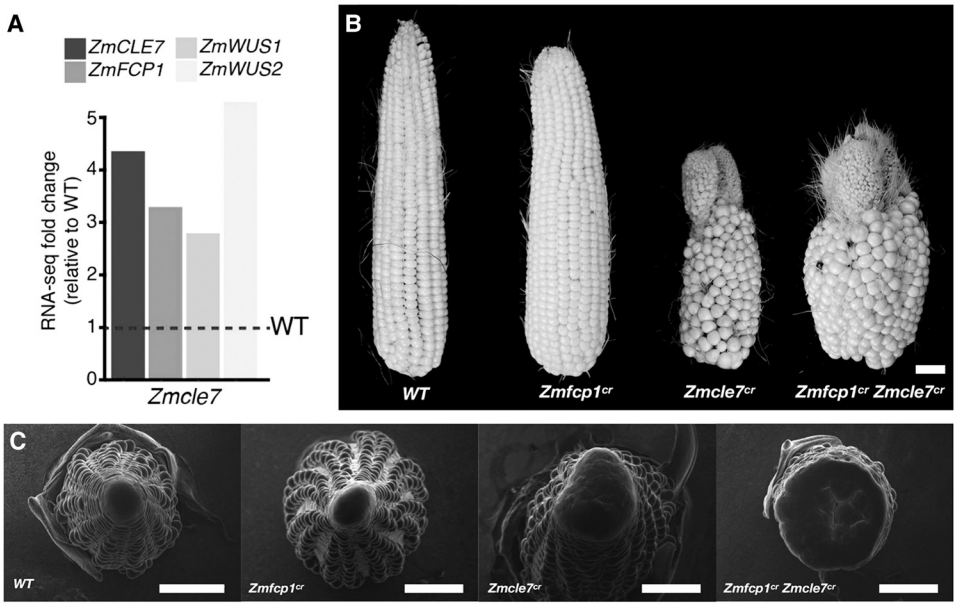


Figure 1. (*A*) *ZmCLE7*, *ZmFCP1*, *ZmWUS1*, and *ZmWUS2* expression was significantly up-regulated in *Zmcle7*^{cr} mutants. (*B*) CLE mutants exhibit redundancy; *Zmfcp1*^{cr} mutants are weak, but strongly enhance *Zmcle7*^{cr} phenotype in double mutants, also shown by scanning electron microscopy (*C*). Scale bars, 1 cm (*B*); 500 µm (*C*).

significantly up-regulated (Fig. 1A). We therefore asked if ZmFCP1 exerts a compensatory function in a *Zmcle7*^{cr} genetic background. *Zmfcp1*^{cr} mutants had a weak phenotype; however, in double *Zmcle7*^{cr} *Zmfcp1*^{cr} mutants the fasciation was dramatically enhanced, supporting our hypothesis that *Zm*FCP1 partially compensates for loss of *Zmcle7*^{cr} (Fig. 1B,C). However, the mechanism of compensation is distinct from that observed in *Arabidopsis* and tomato, suggesting that different mechanisms underlie the evolution of genetic buffering in stem cell pathways.

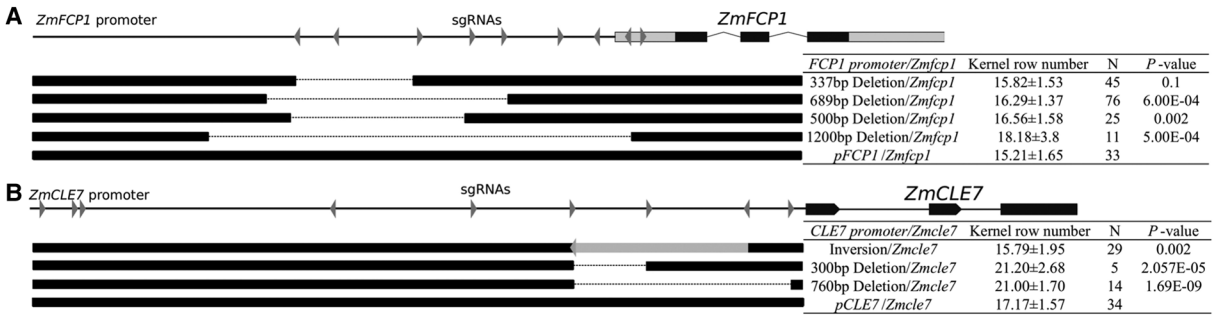
Redundancy in stem cell signaling is also seen on the receptor side, as *Arabidopsis clv1* mutant phenotypes are enhanced by mutations in the related *BARELY ANY MERISTEM 1*, *2*, and *3* (*BAM1*, *BAM2*, and *BAM3*) receptors. Maize has seven *BAM*-like genes and a single *CLAVATA1* gene (*TD1*), and to characterize their function and redundancy, higher-order CRISPR mutants are being made. The genetic interactions between different *BAM* genes and CLEs will allow us to dissect additional signaling pathways for meristem maintenance in maize.

We previously found that weak alleles of fasciated ear mutants can improve maize yield traits, such as kernel row number, by increasing meristem size and number of primordia, while maintaining structural integrity of the meristem. We found that *fea3* weak allele hybrids also enhance overall yield in lab strains of maize. These results are particularly exciting because in our previous studies of weak *fea2* alleles, we found an increase in kernel row number but no overall increase in ear weight (due to a compensatory reduction in kernel size). Therefore, the newly identified *FEA3* signaling pathway could be used to develop new alleles for crop improvement. As *Zmfcp1* and *Zmcle7* single mutants are also fasciated, we gene-edited their promoters using CRISPR to create weak alleles to ask if they could also enhance yield traits. The promoters (-2 kb) of *ZmFCP1* and *ZmCLE7* were targeted by multiplex sgRNAs, and a variety of edited haplotypes were obtained (Fig. 2). Some promoter weak alleles showed a significantly enlarged but nonfasciated ear, indicating potential utility of these favorable alleles in maize breeding.

Control of Shoot Branching and Determinacy

H. Claeys, E. Demesa-Arevalo, X. Xu, Y. Wen, T. Skopelitis
[in collaboration with B. Yang, Danforth Center]

The *RAMOSA* (*RA*) genes function to impose determinacy on axillary meristem growth; consequently, *ra* loss-of-function mutants (*ra1, ra2,* and *ra3*) have



FCP1 promoter/Zmfcp1	Kernel row number	N	*P*-value
337bp Deletion/*Zmfcp1*	15.82±1.53	45	0.1
689bp Deletion/*Zmfcp1*	16.29±1.37	76	6.00E-04
500bp Deletion/*Zmfcp1*	16.56±1.58	25	0.002
1200bp Deletion/*Zmfcp1*	18.18±3.8	11	5.00E-04
pFCP1 /Zmfcp1	15.21±1.65	33	

CLE7 promoter/Zmcle7	Kernel row number	N	*P*-value
Inversion/*Zmcle7*	15.79±1.95	29	0.002
300bp Deletion/*Zmcle7*	21.20±2.68	5	2.057E-05
760bp Deletion/*Zmcle7*	21.00±1.70	14	1.69E-09
pCLE7 /Zmcle7	17.17±1.57	34	

Figure 2. CRISPR mutagenesis and phenotype of promoters of *ZmFCP1* (*A*) and *ZmCLE7* (*B*).

more highly branched inflorescences. *RA3* encodes a trehalose phosphate phosphatase, which catalyzes the conversion of trehalose-6-phosphate (T6P) to trehalose. T6P is an important regulatory metabolite that connects sucrose levels, and thus the sugar status, to plant growth and development, but its mode of action is still unclear. *RA3* is expressed at the base of axillary inflorescence meristems and localizes to distinct puncta in both nuclear and cytoplasmic compartments, suggesting that its effect on development may not be simply metabolic. These data support the hypothesis that *RA* genes may serve as mediators of signals, maybe a sugar signal, originating at the boundary domain and regulating determinacy. *RA3* itself may have a direct gene regulatory function, because it is nuclear and affects the expression of specific downstream genes.

We aim to identify factors that act in the same pathway by screening for enhancers of the *ra3* phenotype. Typically, *ra3* mutants have three to eight ectopic branches at the base of the ear. We mutagenized *ra3* and looked for plants that have more branches and/or branches in the upper part of the ear. So far, four independent alleles of *TPP4* (*TREHALOSE-6-PHOSPHATE PHOSPHATASE*), an *RA3* paralog, were identified, and we confirmed that *TPP4* is the causative gene using additional CRISPR-Cas9-generated alleles. *TPP4* was expressed in the same domain as *RA3*, and upon mutation of *ra3* its expression was up-regulated, suggesting that it acts as a redundant backup to compensate for loss of *RA3*. All EMS-induced alleles contained single amino acid substitutions, and some of the resulting mutant proteins still had considerable enzymatic activity, despite all having similar phenotypic strength. Therefore there is no straightforward relationship between TPP activity and phenotype, and additional regulatory functions of TPP4 may be important. Using CRISPR-Cas9, we also generated knockouts of *TPP12*, a more distant family member that is also expressed in developing inflorescences—but unlike *TPP4*, mutating *TPP12* did not enhance the RA3 phenotype, showing functional divergence within the TPP family.

Another *ra3* enhancer mutant was mapped to a gene encoding an RNA-binding protein that functions in inflorescence development. We have made maize lines carrying functional YFP fusions of this gene and will use these to further characterize its role in meristem determinacy, and its relationship with RA3, by looking at protein–protein and protein–RNA interactions. More recently, we mapped a third enhancer mutant to *IDS1 (INDETERMINATE SPIKELET1). ids1* mutants have more floral meristems, and *IDS1* expression was significantly increased in *ra3* mutants. We are testing if the enhancement of RA3 by IDS1 may be due to their physical interaction. For the other *ra3* enhancer mutants, we are currently at various stages of mapping and confirming the causal genes.

To further probe RA3 functions, we also developed a specific polyclonal antibody. This has been used to find interactors using immunoprecipitation and mass spectroscopy (IP-MS), and to probe its subcellular localization (Fig. 3). Using different nuclear markers, we are defining the nature of the RA3 nuclear speckles.

We also wish to understand the contribution of the enzymatic function of RA3 to its biological mechanism. We achieved partial complementation of *ra3* mutants using a catalytically dead mutant version of RA3, supporting a moonlighting hypothesis, and are now refining this experiment by mutating endogenous *RA3* and *TPP4* loci to catalytically dead versions using CRISPR-Cas9-mediated base editing. We are also using CRISPR-Cas9 to generate mutants in trehalose

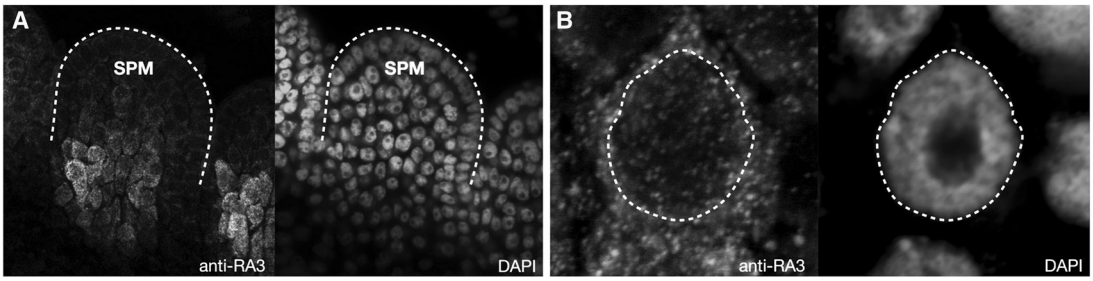


Figure 3. RA3 is localized in cells subtending the spikelet pair meristems (SPMs), highlighted with dashed lines (*A*). RA3 is localized in cytoplasmic and nuclear speckles; nucleus is outlined by a dashed line (*B*).

phosphate synthases (TPSs), which catalyze the step preceding RA3 in the trehalose metabolic pathway. If accumulation of T6P causes the *ra3* phenotype, then combining *ra3* with *tps* mutants should alleviate the phenotype, giving important insights into its mechanism.

Last, we are attempting to knock out the entire TPP gene family in *Arabidopsis* using CRISPR. So far, mutation of two RA3 homologs, *TPPI* and *TPPJ*, did not reveal any obvious phenotype, and we are now mutating additional family members to overcome the likely redundancy in this gene family.

Natural Variation in Inflorescence Architecture

H. Claeys

Maize inflorescence architecture has been a target for extensive selection by breeders since domestication, and the maize genome is highly diverse; hence, different maize inbreds vary greatly in these traits. The genetic basis underlying this diversity is largely unknown, but is of great interest for both fundamental and applied science. To identify natural variation relevant to inflorescence traits, we screened for inbred backgrounds that can enhance or suppress the phenotypes of different mutants. We focused on the 25 nested association mapping (NAM) founder inbreds, because they were selected to capture the diversity of maize germplasm, and because of the available genetic tools. We crossed these 25 inbreds to our collection of mutants (in a B73 background) and screened the F2s for mutants with suppressed or enhanced phenotypes.

fea2-0 was strongly enhanced in the NC350 background, and we mapped this enhancement to a single major-effect locus on chromosome 5 using both bulked segregant analysis and crosses to recombinant inbred lines (RILs). Interestingly, we used B73-NC350 heterogeneous inbred families (HIFs) to show that the NC350 allele of *TD1* positively affects kernel row number (KRN), demonstrating its potential usefulness in breeding. Fine mapping of the NC350 enhancer locus led us to a small region containing only approximately five genes, with a prominent candidate *THICK TASSEL DWARF1* (*TD1*), a known meristem regulator. Consequently, a *td1* loss-of-function mutant cannot rescue the enhanced phenotype. There are no obvious causal changes in the coding sequences of the two alleles of *TD1*, but we found regulatory changes that may explain the enhanced phenotype; the B73 allele of *TD1* is up-regulated in a *fea2* mutant background, whereas the NC350 allele is insensitive to loss of *FEA2*. Our results reveal changes in the wiring of regulatory networks controlling compensation in meristem size between ecotypes. To test this candidate, we generated alleles with mutations in the *TD1* promoter through CRISPR-Cas9-mediated promoter editing, and are asking if they can recapitulate the behavior of the NC350 allele.

The Effects of Drought on Early Inflorescence Development

H. Claeys [in collaboration with H. Cline and B. Meeley, DuPont Crop Genetics; E. Vollbrecht, Iowa State University; S. Hake, USDA-UC Berkeley; J. Dinneny, Stanford University; A. Eveland, Danforth Center]

Drought stress is one of the major environmental factors limiting maize yield. Some progress has been made in studying how drought affects grain filling during the later stages of reproductive development, but almost

nothing is known about how drought affects early inflorescence development, where the number of grains is determined. Yield is significantly affected when maize is subjected to drought at this stage; therefore, in collaboration with DuPont-Pioneer, we subjected plants to drought at different stages of early inflorescence development to assess how drought interacts with development. Severe drought was applied when ears were ~1 cm long, allowing us to dissect different meristem types from these ears to assess how they respond to drought. In a second experiment, plants were drought-stressed when ears were ~1 mm long, to study how drought affects early ear development, and in a third trial, drought was applied at an even earlier plant growth stage, when tassels were ~1 mm long. Because the plants were smaller, water loss through evaporation was slower, and stress levels were generally lower, allowing us to assess early responses to mild drought. From all these experiments, we collected developing ears or tassels for transcriptome analysis. We found evidence for a large-scale reprogramming in each meristem type—consisting of, on the one hand, a common core of up-regulated genes that are involved in stress resistance and down-regulated genes associated with cell proliferation, reflecting lower meristem activity, and, on the other hand, meristem type–specific developmental processes being affected. Of note were effects on gibberellin and carbohydrate metabolism, hinting at possible signaling mechanisms underlying these changes in meristem activity. The third experiment in particular yielded interesting data on very early reprogramming of metabolism in response to drought.

Mechanism of Active Transport of Transcription Factors through Plasmodesmata

M. Kitagawa, T. Skopelitis

In plants, some transcription factors (TFs) are actively and selectively transported between cells to specify cell fates. These TFs are transported through plasmodesmata (PD), membrane-lined channels traversing the cell wall. To this date, however, the mechanism underlying the active and selective transport of TFs through PD has been largely unknown. Previously, we established a system for evaluating the capacity of the active transport of TFs in *Arabidopsis* seedlings using a mobile homeodomain TF, KNOTTED1 (KN1). Using this system, we isolated mutants in which the active transport of TFs may be defective. One such mutant encodes an aspartyl-transfer RNA synthetase (AspRS), an enzyme that attaches L-aspartate onto its transfer RNA (tRNA) to function in translation. Interestingly, we found the enlarged shoot apical meristem (SAM) showing disorganized arrangement of cells and fasciation of stems in this mutant, suggesting this AspRS regulates SAM development. Based on the phenotypes, we named this gene *OKINA KUKI* (*OKI1*, Japanese for big stem). OKI1 is a mitochondria-localized AspRS, and in *oki1* mutants, the expression domains of CLAVATA3 (CLV3) and WUSCHEL (WUS), which form a negative-feedback loop to maintain the stem cell population in the SAM, were expanded concomitantly with the enlarged SAM phenotype. Our double-mutant analysis indicated that *wus* and *clv3* are epistatic to *oki1* in meristem size control. To address the possible mechanism by which *oki1* mutants cause meristem disruption, we reasoned that a block in mitochondrial function by partial loss of an essential translation factor might lead to redox imbalance, which is known to impact meristem size. Indeed, reactive oxygen species (ROS) staining indicated that superoxide and hydrogen peroxide were up-regulated in *oki1* meristems, suggesting that redox imbalance may cause the increases in meristem size in *oki1*.

We also isolated two new mutants, *rb31-7* and *mk5-140*, by the screen described above, that are mutated in the same gene, encoding ribosomal RNA processing protein (RRP) 44A. This protein is a subunit of the RNA exosome complex and functions in degradation and processing of RNAs. We found that the mutants had smaller SAMs and fusions between leaf and stem, phenotypes that are also observed in plants where transport of the *Arabidopsis* KN1 homolog SHOOTMERISTEMLESS (STM) is restricted. In addition, double *rrp44;stm* mutants had enhanced fusions between leaves or petioles (Fig. 4). RNA-immunoprecipitation RT-PCR assays indicate that RRP44A binds to *STM* mRNA, suggesting RRP44A controls *STM* mRNA trafficking between cells through direct binding to its mRNA. Furthermore, we found that RRP44A colocalized with plasmodesmata (Fig. 4). Therefore, PD may be a platform where RRP44A interacts with *STM* mRNA to mediate its trafficking.

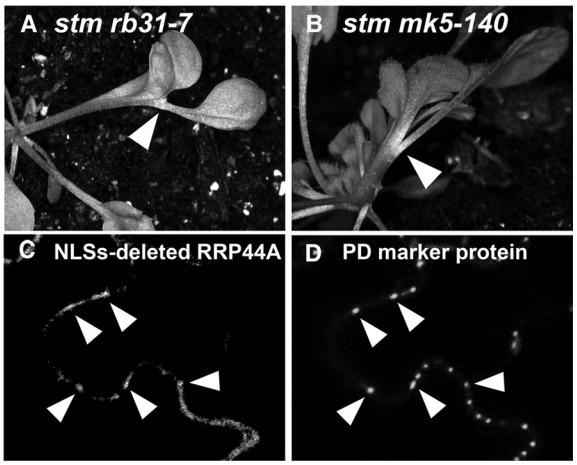


Figure 4. Fusion phenotype of *stm rrp44a* double mutants, and PD colocalization of RRP44A. (*A,B*) Double *rb31-7 stm* (*A*) and *mk5-140 stm* (*B*) mutants show prominent leaf-petiole fusion phenotypes (arrowheads). (*C,D*) Nuclear localization sequences (NLSs)-deleted RRP44A fused a fluorescent protein (*C*) colocalized with a PD marker protein (*D*), shown by arrowheads.

Functional Annotation of the Maize Genome by ChIP-Seq and FACS

X. Xu, D. Jackson [in collaboration with C. Ortiz Ramirez and K. Birnbaum, Center for Genomics and Systems Biology, New York University; X. Wang, L. Wang, J. Drenkow, T. Gingeras, and D. Ware, CSHL]

To develop functional annotation of the maize genome, we are conducting genome-wide transcription factor (TF) binding analysis by chromatin immunoprecipitation-sequencing (ChIP-Seq) and expression profiling of cell types by fluorescence-activated cell sorting (FACS). These are important goals of the MaizeCODE project, an initial analysis of functional elements in the maize genome.

To perform genome-wide TF binding site analysis, we generated fluorescent protein (FP) tagged transgenic lines for diverse TFs that function in different aspects of maize development. These TFs belong to several different families, such as the MADS box, TUNICATE1A, functioning in maize domestication; the homeodomain TF, WUSCHEL1 (WUS1), functioning in meristem maintenance; and the GATA TF, TASSELSHEATH1 (TSH1), functioning in bract suppression. To overcome limited tissue availability for conducting ChIP-Seq, we crossed these lines into a double mutant, *branched silkless*;*Tunicate* (*bd*;*Tu*), that transforms the maize ear into a "cauliflower" with overproliferating meristems. The tissues have been collected and the ChIP-seq libraries are under preparation.

To profile cell type–specific expression by FACS, we generated tissue-specific promoter FP lines to isolate cell types or domains for RNA-Seq, small RNA, DNA methylation, and ATAC-Seq analysis. The promoter lines cover several different types of cells or tissues in the meristem, such as the organizing center (p*ZmWUSCHEL1*-mRFP1), lateral organ zone (p*ZmYABBY14*-tagRFPt), and epidermal zone (p*ZmHOMEODOMAIN-LEUCINE ZIPPER IV6*-tagRFPt). RNA-Seq of FACS-sorted cells has been completed, and preliminary analysis of YABBY14 FACS-sorted cells found that all known *YABBY* genes were significantly enriched, indicating that the FACS experiment was successful. Other FACS profiling data are being analyzed, and will be available through the CyVerse MaizeCODE project portal as a public resource.

New Insights into Maize Ear Development Using Single Cell (sc)RNA-Seq

X. Xu, L. Liu, D. Jackson [in collaboration with M. Crow, L. Wang, D. Ware, J. Preall, and J. Gillis, CSHL]

Productivity of maize depends on the inflorescences, and development of the ear requires a programmed series of meristem fate decisions between different cell populations. A fundamental understanding of development requires insight into the full diversity of cell types and developmental domains and the gene networks required to specify them. However, current studies have classified cell types and domains mainly by morphology and signatures derived from bulk RNA sequencing. Major insights into genes required to specify cell types and developmental domains have come from classical genetics; however, this approach is limited by genetic redundancy and pleiotropy. Single-cell transcriptome profiling of heterogeneous tissues, such as the maize ear inflorescence, can provide high-resolution windows to understand genome-wide transcriptional signatures of specific cell types and identify new developmental domains in a quantitative and comprehensive manner.

We isolated single cells from developing maize ear primordia by a new rapid protoplasting method, and used the high-throughput 10X Genomics Chromium™ platform to profile the transcriptomes of more than 10,000 individual cells. Profiling detected on average 2938 transcripts per cell, and in total we

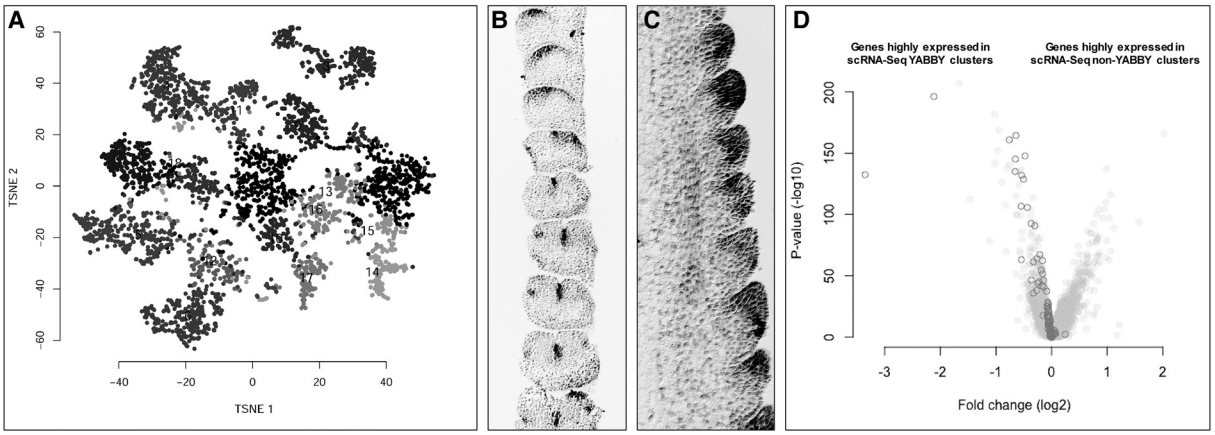


Figure 5. (*A*) Eighteen groups of cells are shown in t-distributed stochastic neighbor embedding (tSNE) plots. (*B,C*) mRNA in situ hybridization of selected marker genes from different cell clusters show signal in meristem branching sites (*B*) or meristem tips (*C*). (*D*) The majority of enriched transcripts from *YABBY* FACS RNA-seq (black circles) showed concordant differential expression with genes highly expressed in YABBY clusters in our scRNA-seq analysis (gray dots on the *left* branch).

detected transcripts from 28,254 genes. Graph-based clustering partitioned cells based on their transcriptomes into 18 groups, which we visualized using t-distributed stochastic neighbor embedding (tSNE) plots (Fig. 5A). Many of the groups were clearly marked by known marker genes, such as an L1/epidermal layer group, marked by *LIPID TRANSFER PROTEIN* and *OUTER CELL LAYER* genes; an L2 layer meristem group, marked by *KNOTTED1*; a lateral primordium group, marked by *YABBY* genes; and a vasculature group, marked by *RAN BINDING PROTEIN2*. Each of these groups contains an additional 22–554 new cell type– or domain-enriched genes that are novel candidates for markers or developmental regulators. To validate the accuracy of clustering and determine where each cell population was located in the maize ear inflorescence, we selected marker genes within each group to perform mRNA in situ hybridization. The results revealed that our scRNA-seq method identified specific and spatially restricted markers for each group. Strikingly, we also identified novel markers for specific developmental domains, such as meristem branching sites and meristem tips (Fig. 5B,C). In addition, a large majority of high-confidence marker genes identified from our FACS assay of promoter lines, such as *pYABBY14-tagRFPt*, also showed concordant differential expression patterns in our scRNA-seq analysis (Fig. 5D).

Collectively, we demonstrated that scRNA-seq is a powerful tool to predict genome-wide gene expression domains for tens of thousands of maize ear inflorescence cells and identified many new candidate regulators of maize inflorescence development.

PUBLICATIONS

Je BI, Xu F, Wu Q, Liu L, Meeley R, Gallagher JP, Corcilius L, Payne RJ, Bartlett ME, Jackson D. 2018. The CLAVATA receptor FASCIATED EAR2 responds to distinct CLE peptides by signaling through two downstream effectors. *eLife* doi:10.7554/eLife.35673

Liu L, Li C, Song S, Teo Z, Shen L, Wang Y, Jackson D, Yu H. 2018. FTIP-dependent STM trafficking regulates shoot meristem development in arabidopsis. *Cell Rep* **23:** 1879–1890.

Ortiz-Ramírez C, Demesa-Arevalo E, Xu X, Jackson D, Birnbaum KD. 2018. An efficient cell sorting protocol for maize protoplasts. *Curr Protoc Plant Biol* **3**: e20072.

Wu Q, Jackson D. 2018. Detection of MAPK3/6 phosphorylation during hypersensitive response (HR)-associated programmed cell death in plants. *Methods Mol Biol* **1743:** 153–161.

Wu Q, Xu F, Jackson D. 2018. All together now, a magical mystery tour of the maize shoot meristem. *Curr Opin Plant Biol* **45:** 26–35.

Wu Q, Regan M, Furukawa H, Jackson D. 2018. Role of heterotrimeric G proteins in maize development and enhancement of agronomic traits. *PLoS Genet* **14:** e1007374.

Xu F, Jackson D. 2018. Learning from CIK plants. *Nat Plants* **4:** 195–196. doi:10.1038/s41477-018-0125-x

In Press

Claeys H, Vi SL, Xu X, Satoh-Nagasawa N, Eveland AL, Goldshmidt A, Feil R, Beggs GA, Sakai H, Brennan RG, Lunn JE, Jackson D. 2019. Control of meristem determinacy by trehalose-6-phosphate phosphatases is uncoupled from enzymatic activity. *Nat Plants* **5:** 352–357.

Kitagawa M, Jackson D. 2019. Control of meristem size. *Ann Rev Plant Biol* **70:** 269–291.

STEM CELL PRODUCTION, MATURATION, AND REPRODUCTIVE SUCCESS IN PLANTS

Z.B. Lippman J. Dalrymple, A. Hendelman, B. Hendrick, E. Joshua, J. Kim, A. Krainer, C-T. Kwon, Z. Lemmon, S. Qiao, G. Robitaille, D. Rodriguez-Leal, S. Soyk, X. Wang

Plant reproductive success in nature and crop yield in agriculture rely on flowers, which are the foundation for fruit and seed production. Flowers develop from inflorescences—reproductive branching systems that originate from leaf-producing vegetative shoots when environmental and endogenous signals induce small populations of stem cells at the tips of shoots known as meristems to transition into reproductive states. The number of inflorescences produced on a plant, as well as how many branches and flowers form on each inflorescence, can vary dramatically both within and between species. At the center of this diversity lie two critical processes of stem-cell regulation, which are the focus of research in the lab: (i) maturation, during which stem cells transition from a vegetative to a reproductive growth program, and (ii) proliferation, which controls stem-cell population size. Our research program integrates development, genetics, genomics, and gene editing to explore mechanisms of meristem maturation and maintenance and their relationship to plant shape, flowering, and flower production. We take advantage of extensive natural and mutant variation in inflorescence production and architecture in tomato and related nightshades to explore how differences in these processes explain the remarkable diversity in shoot architecture. Recent discoveries on these topics have led us to broader questions about the significance of genomic structural variation, gene redundancy, and epistasis in development, domestication, and breeding.

Redundancy and Compensation among Ligand-Receptor Interactions in Stem-Cell Homeostasis

C-T. Kwon, D. Rodriguez-Leal, A. Hendelman

The CLAVATA (CLV) network is a key regulatory pathway for plant stem-cell proliferation, which is also highly associated with crop domestication and improvement. Ligand-receptor relationships are the foundation of CLV signaling, and it has been shown that binding of the processed CLV3 small signaling peptide to the leucine-rich repeat (LRR) receptor kinase, CLV1, is the core mechanism underlying the CLV network and its control of stem-cell homeostasis and shoot meristem size. Previously, we found that tomato *SlCLE9*, a paralog of tomato *SlCLV3*, acts as a compensator during meristem proliferation. Along with ligand redundancy, the tomato *SlCLV1* receptor has four paralogous LRR receptors, encoded by the *BARELY ANY MERISTEM* (*SlBAM*) genes.

Based on findings in *Arabidopsis* showing receptor compensation among *CLV1* and three *BAM* homologs, we tested if these related receptors in tomato are similarly regulated by transcriptional compensation. Unlike in *Arabidopsis*, transcriptome profiling showed none of the four *SlBAMs* were significantly up-regulated in any of the single-receptor mutants. To further understand functional redundancy of four *SlBAMs*, we used CRISPR-Cas9 to mutate all four tomato *SlBAM* genes. Homozygous null mutations for three genes (*SlBAM1*, *SlBAM3*, and *SlBAM4*) revealed no impact on stem-cell proliferation of meristem size individually, consistent with findings in *Arabidopsis*. However, meristem and floral organ phenotypes of *Arabidopsis clv1* mutants are enhanced stepwise by the *bam* mutants. Interestingly, we observed that tomato *slclv1 slbam1 slbam4* triple mutants were indistinguishable from *slclv1* single mutants, indicating that receptor compensation between *SlCLV1* and *SlBAMs* is wired differently than *Arabidopsis*. It is possible that other *SlBAMs* or receptors could contribute to CLV signaling. Thus, we are now generating *slclv1 slbam1/2/3/4* quintuple mutants, and also higher-order mutants with additional LRR receptors we have targeted by CRISPR-Cas9 to determine the extent to which receptor compensation in tomato has diverged from *Arabidopsis*.

In related work, we are investigating the origin of peptide compensation in tomato and related

Solanaceae species. In collaboration with M. Bartlett (UMass-Amherst) we have analyzed the syntenic chromosomal blocks that carry *SlCLV3* and *SlCLE9* through eudicot evolution, and found both blocks are in the order Solanales, but only *SlCLV3*-like genes are found in syntenic blocks outside the Solanales. This indicates *SlCLV3* and *SlCLE9* duplicated in the lineage leading to the Solanales. Interestingly, *SlCLE9*-like genes in the Solanales showed a dynamic evolution whereby the ortholog of this compensator in different species has been duplicated or deleted. For example, no *SlCLE9*-like genes are found in eggplant or potato, but two fragments of *SlCLE9*-like genes are found in pepper. To decipher the evolution and functional relationship of both peptides in the Solanales, we are now generating knockout mutants of *SlCLV3* and *SlCLE9* orthologs in multiple species. Combined phenotypic and transcriptome analyses of each mutant in each species should help reveal whether and how peptide compensation has evolved in the control of stem-cell proliferation in diverse plants.

Structural Variant Landscapes in Tomato and Their Association with Phenotypic Diversity

X. Wang, S. Soyk, J. Kim

Extensive phenotypic and genetic variation underlies economically important traits in crop plants, including tomato. Linking genotype to phenotype remains difficult, and one of the greatest challenges is identification and characterization of sequence variants underlying quantitative variation. Despite several examples of significant roles of structural variants (SVs) in crop domestication and agronomic traits, knowledge of SV diversity is incomplete as a result of difficulty in their detection using short-read (Illumina) sequencing technology. However, recent improvements in long-read sequencing technology have provided a powerful tool to identify the large and often complicated SVs that are undoubtedly widespread in diverse genotypes. To construct a high-quality tomato panSV-genome and elucidate how SVs shape plant genomes and contribute to domestication and improvement, we are collaborating with M. Schatz (Johns Hopkins) and F. Sedlazek (Baylor College of Medicine), and using Oxford Nanopore (ONT) technology to generate long-read sequencing data for a diverse panel tomato germplasm. Wild and domesticated genotypes were chosen using an algorithm that captures the majority of the predicted, although ill-defined, SV diversity from a large resequenced (Illumina) core collection, which we helped establish. In the past year, more than 60 tomato accessions, including wild ancestors, early domesticates, and modern cultivated accessions, have been sequenced to at least 40× coverage with mean read length of 10–30 kbp. The Nanopore long reads were aligned to the tomato reference genome to identify SVs. Our SV calls for the first 12 accessions to reveal thousands of SVs, and the number of identified SVs for each accession varies substantially, ranging from 25,000 to 45,000 variations each. The identified types of SVs include deletions, insertions, translocations, duplications, and inversions, with deletions and insertions being the most predominant SVs. Furthermore, our analyses show that many identified SVs are specific to individual samples (i.e., represent likely rare alleles), but a substantial number of SVs are shared among the 12 accessions.

With more than 800 existing resequenced genomes, we identified SVs using three SV analysis pipelines and then integrated the results using our consensus algorithm, SURVIVOR. Because of the high false-positive rate of using short-read data, stringent selection criteria were applied for filtering to identify SVs with putative functions. To determine the functional significance and molecular impact of gene-associated SVs, we aim to profile the RNA expressions from different major tissues of the top 20 accessions that capture the most SVs and investigate both the global and specific impact of SVs on gene expression. Analyses of our first set of RNA-sequencing (RNA-Seq) data of vegetative (leaf) tissue and publicly available RNA-Seq data of fruit pericarp tissue show clear correlation of several verified duplication SVs with increased gene expression. Several such duplication SVs have dramatically increased allele frequencies during tomato domestication and improvement, perhaps indicating active selection by breeders. Selected genes associated with the duplications are being functionally validated. One of the examples is a duplication containing a tomato fruit weight quantitative trait locus (QTL) gene. Previous reports indicated that a promoter variant might have caused increased gene expression and, thus, larger fruit. However, our identification of a duplication carrying this gene and QTL region suggests that the duplication could be causative. Functional characterization using CRISPR-Cas9 and a dosage

analysis is under way. With additional RNA-Seq data from more tissues and high-confidence SV calls from our long-read sequencing, the effects of different classes of SVs on gene expression will be investigated on a genome-wide scale.

Cryptic Genetic Variation Impacting Development, Domestication, and Crop Improvement

S. Soyk, Z. Lemmon

Over the last century, plant breeders have had remarkable success in developing highly productive crops. However, improvements in agriculture are plateauing. Great enthusiasm and high expectations currently exist that genome editing will overcome this problem and revolutionize plant breeding and crop improvement. Yet, it is largely ignored that engineering beneficial genetic variation in diverse genetic backgrounds can result in undesirable outcomes caused by interactions with preexisting secondary mutations, which breeders are often blind to. Such hidden epistasis caused by so-called "cryptic variation" must, then, be overcome through further laborious and time-consuming breeding. Surprisingly, how such neutralization of epistasis is achieved has never been studied, and this lack of understanding continues to put breeders at a disadvantage as they strive to attain predictable outcomes in agriculture.

An enlightening example of such cryptic variation involves the classical "jointless" trait of tomato. The jointless mutation was discovered 75 years ago and promised to bring immediate and widespread improvements in harvesting practices because it prevents fruit drop, similar to domestication mutations that eliminated seed shattering in cereals. However, breeding for this highly desirable trait in tomato continues to be extremely challenging because of a cryptic variant in a closely related flower development gene that arose during domestication and reached near fixation in modern cultivars. Thus, attempts to create new jointless varieties—using either natural mutations or genome editing—almost always fail because mutations in both flowering genes cause excessive flower production but low fruit yields as a result of sterility. Remarkably, we found that soon after the jointless trait was discovered, breeders exploited new cryptic variation that restored normal flowering and fertility—but only in a small collection of tomato genotypes developed for open-field production in Florida. Using both quantitative and molecular genetics approaches, we identified the cryptic natural mutation that drove the suppression of negative epistasis on fruit yield. We discovered how one dominant genetic locus drove restoration of fertility caused by a recently evolved gene duplication, which itself is a cryptic variant, and harbors the ancestral cryptic variant responsible for the undesirable epistasis. Mechanistically, we show how this duplication increased gene copy number to overcome a partial loss of functional transcripts because of missplicing from the ancestral variant. Significantly, we then translated our findings to agriculture by showing that modern breeding lines carrying the duplication are protected from negative epistasis when using CRISPR-Cas9 genome editing to engineer jointless mutations. This has allowed us to devise multiple solutions to predictably engineer the jointless trait into any genetic background by genome editing (Fig. 1). Our study provides a compelling demonstration of how a thorough identification and characterization of cryptic genetic variation in both plants and animals will be essential to capitalize on genome editing to revolutionize breeding and agriculture.

Dissecting *cis*-Regulatory Control of Gene Expression and Quantitative Trait Variation

X. Wang, D. Rodriguez-Leal [in collaboration with L. Aguirre, WSBS]

Previous work from our lab has shown that a wide range of quantitative phenotypic variations in plant can be achieved through altering the expression of key developmental genes by engineering various alleles of gene promoters. Using CRISPR-Cas9 to mutagenize *cis*-regulatory regions, we showed that a collection of promoter alleles with various levels of gene expression changes could be created rapidly and efficiently. Our initial use of this tool to characterize the promoter of the key plant stem-cell regulator gene *SlCLV3* in tomato has suggested a complex architecture of *cis*-regulatory regions, transcriptional regulation, and quantitative trait variation. Although quantitative trait variations were successfully achieved, genetic understanding of the *cis*-regulatory control on gene expression and phenotypic variation remains unclear. To dissect the complex architecture, additional

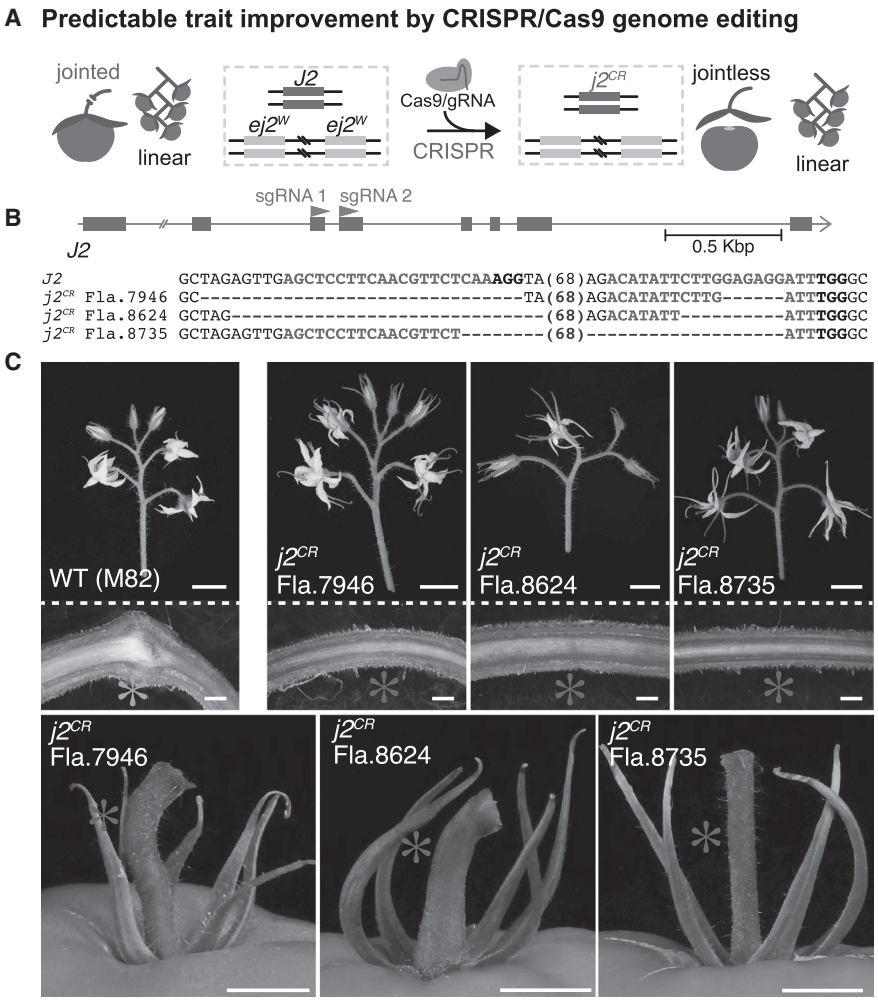


Figure 1. Predictable improvement of agronomic traits by genome editing mediated by cryptic variation. (*A*) The *sb3* duplication allows predictable breeding for the jointless trait. (*B*) Targeting of *J2* was performed with two single-guide RNAs (red arrowheads). Sequences of CRISPR-Cas9 engineered $j2^{CR}$ null alleles in three different *sb3* large-fruited breeding lines are shown. sgRNA targets and protospacer-adjacent motifs (PAMs) are indicated in red and bold font, respectively. Deletions are indicated by blue dashes, and sequence gap length (bp) is shown in parentheses. (*C*) $j2^{CR}$ mutations in three breeding lines containing the *sb3* tandem duplication result in desirable jointless pedicels on flowers and fruits and produce normal unbranched inflorescences. Lignified cells were stained using phloroglucinol to show loss of the abscission zone. Scale bars, 1 cm in *A*, and the top and bottom panels of *C*; 1 mm in the middle panel of *C*. Green and red asterisks mark the presence and absence of a pedicel abscission zone, respectively.

SlCLV3 promoter alleles from our previous mutagenesis were isolated, and their characterization reveals a higher resolution of continuum of the quantitative trait variation. However, the diversity and complexity of the promoter alleles make it hard to directly assign individual *cis*-regulatory elements (CREs) within the promoter their specific functions and contributions to gene expression and phenotypic variation, which impedes our understanding of the *cis*-regulatory control of critical developmental genes. To investigate the associations between individual *cis*-regulatory regions with phenotypic variation and quantitative, spatial, and temporal gene expression, promoter alleles mutating individual smaller evolutionarily conserved regions of the native *SlCLV3* gene were generated using CRISPR-Cas9, and their effects on trait variation and gene expression are currently being evaluated. Preliminary results show that deletions of these smaller *cis*-regulatory regions either have weak or no phenotypic effects, indicating that multiple CREs contribute to

the regulation of *SlCLV3* expression with different effect strengths. To dissect the interactions between regions within the promoter, we are aiming to generate promoter alleles combining deletions of different regions. Interestingly, one *SlCLV3* promoter allele that fortuitously provided two *cis*-regulatory regions deleted together displayed an enhanced phenotype greater than an additive effect of two individual regions, suggesting redundancy and epistasis exist between *cis*-regulatory regions. With the newly generated "micro" deletion promoter alleles and their combinations, we are well positioned to dissect the functions of specific *cis*-regulatory regions, as well as the modules in which they function to regulate *SlCLV3* expression and quantitative trait variation.

Genetic and Epigenetic Landscapes of *cis*-Regulatory Regions

A. Hendelman

Recently, our lab showed that creating a range of mutations in *cis*-regulatory regions can expose a range of phenotypic variation. This has opened the door to address new questions on how quantitative variation is controlled, and dissect the roles of specific CREs in transcriptional activity. Transcriptional regulation is controlled by transcription factors (TFs) binding to specific CREs, found nearby and also far from gene bodies. In addition, chromatin structure and epigenetic factors (e.g., DNA and histone methylation) collectively influence transcriptional regulation as accessibility of CREs is needed for transcriptional activation and repression. The complexity of transcriptional control in plants is poorly understood, and our ability to predict and functionally interrogate gene regulatory regions is still restricted.

Our focus is on meristems, in which genomic data as described above are lacking. We have begun using ATAC-seq and whole-genome sequencing to expose the open chromatin regions. ATAC-seq libraries were produced from meristems and leaf samples, which revealed ~70,000 accessible regions and a similar distribution of open chromatin for both tissues. More than half were localized in proximal promoter regions (up to 2 kb from TSS), 20% within gene bodies, and 20% surrounding transcription end sites. As 60% of the tomato genome comprises of transposons and related repeats, 70% of accessible regions were in such sequences. Interestingly, 40% of the accessible regions in meristems were unique, and RNA-Seq data showed that many of the corresponding genes were expressed in this tissue. However, not all genes with accessible regions in their promoters were expressed, and it has been shown that open chromatin is not always associated with transcription complexes, but can also be occupied by repressor elements. To build a higher chromatin resolution on regulatory states, we have included DNA methylation and histone modifications associated with gene transcription or repression. Finally, because essential regulatory regions in gene promoters are more likely to be evolutionarily conserved, we are exploring whether accessible regions are maintained among related Solanaceae species. We are now testing the functional significance of our findings using CRISPR-Cas9 against candidate regulatory regions.

Rapid Improvement of Productivity Traits in Orphan Crops by Genome Editing

Z. Lemmon, J. Dalrymple, S. Soyk, D. Rodriguez-Leal

Over the last three decades, numerous studies on crop domestication have revealed the genes and mutations that were critical for improving growth and productivity. Prominent examples include mutations that reduced branching in maize, prevented seed shattering in rice, and increased fruit size in tomato, to name just a few. Such alleles arose randomly over thousands of years, making domestication and improvement of crops a slow and laborious process. Now, with the rapid increase of genome editing technologies, there is great hope and anticipation that yield barriers can be broken by creating targeted mutations in the major staple crops. However, potentially more impactful will be to advance breeding of dozens of orphan crops, which are incredibly important in regional agriculture and have the potential to make a major impact on food security. Yet, most orphan crops have not been fully domesticated and, therefore, suffer from poor productivity.

By uniting developmental genetics with genomics and CRISPR-Cas9 gene editing, we have collaborated with J. Van Eck (Boyce Thompson Institute) to rapidly improve multiple domestication traits in an orphan fruit crop. *Physalis pruinosa*, commonly known as "ground cherry," is a distant relative of tomato cultivated primarily in Central and South America for

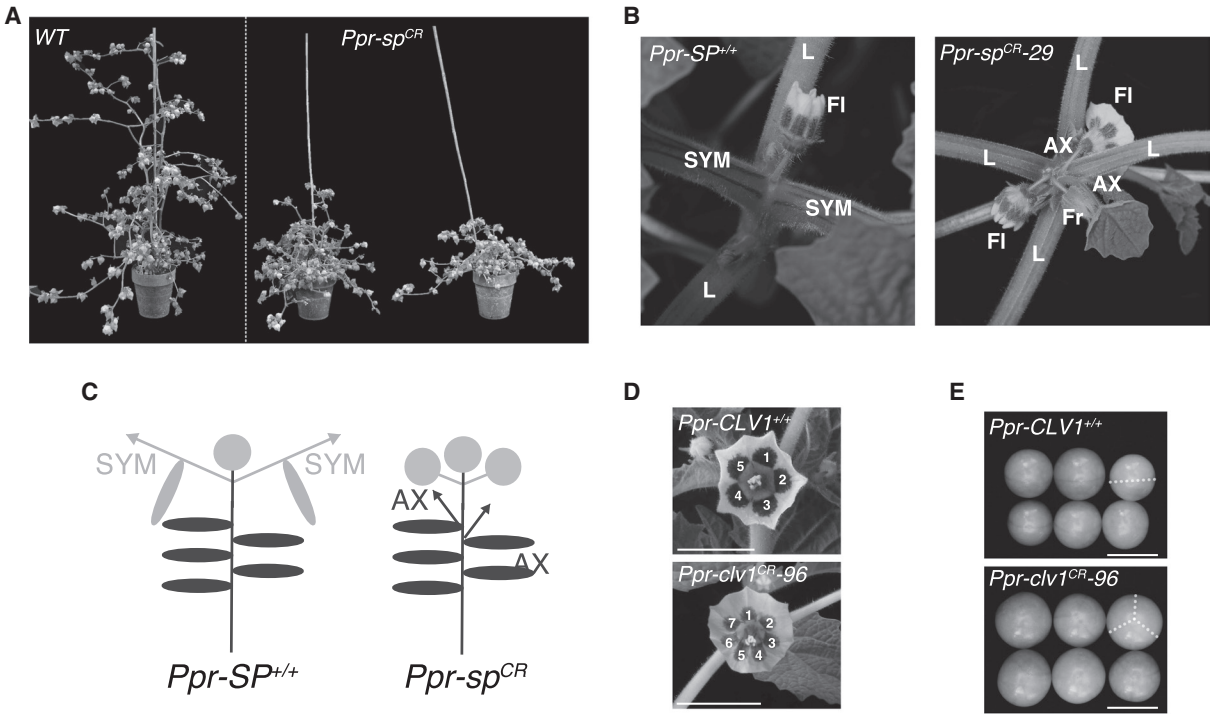


Figure 2. CRISPR-Cas9 targeting of domestication and improvement-related genes in *Physalis pruinosa*. (*A*) *Ppr-sp*^{CR} stable homozygous null T_1 plants showing compact growth compared to the wild-type (*WT*) plants. (*B,C*) Shoot apices of *Ppr-SP*$^{+/+}$ control and *Ppr-sp*^{CR} T_1 plants, showing immediate termination of sympodial shoots into flowers in *Ppr-sp*^{CR}. Vegetative growth continues from axillary shoots. (L) leaf, (SYM) sympodial shoot, (AX) axillary shoot, (Fl) flower, (Fr) fruit. (*D,E*) Flowers and fruits of *Ppr-clv1*^{CR} T_0 plants showing increased petal and locule numbers compared to the *Ppr-CLV1*$^{+/+}$ control. Petals are numbered and demarcation of locules is indicated with dotted lines on one representative fruit for each genotype.

its sweet and nutritious berries. Despite the potential of ground cherry and related *Physalis* orphan crops to become widespread, high-value fruit crops and nutraceuticals, expanding productivity and cultivation has been hindered by its wild sprawling growth habit and tiny fruit. We used ground cherry to test the hypothesis that knowledge from model crops could be translated with CRISPR-Cas9 to create novel allelic diversity and breeding germplasm in related orphan crops. To improve key domestication traits in ground cherry, we captured the gene space using Illumina RNA and DNA sequencing, defined clear orthologs of key tomato developmental and productivity genes through phylogenetic analyses, developed a transformation protocol, showed CRISPR-Cas9 is functional and efficient, and engineered mutations in three major productivity genes. In less than two years, we succeeded in making ground cherry plants more compact, productive, and bear larger fruits (Fig. 2). Our study provides a first demonstration and framework for precision breeding of orphan crops and further illustrates how the imminent implementation of genome editing technologies in both major and orphan crops can provide complementary approaches toward sustainable agriculture and food security.

PUBLICATIONS

Huang X, Tang L, Yu Y, Dalrymple J, Lippman ZB, Xu C. 2018. Control of flowering and inflorescence architecture in tomato by synergistic interactions between ALOG transcription factors. *J Genet Genomics* **45:** 557–560.

Lemmon Z, Reem N, Dalrymple J, Soyk, S, Swartwood K, Rodriguez-Leal D, Van Eck J, Lippman ZB. 2018. Rapid improvement of domestication traits in an orphan crop by genome editing. *Nat Plants* **4:** 766–770.

In Press

Alonge M, Soyk S, Ramakrishnan S, Wang X, Goodwin S, Sedlazek FJ, Lippman ZB, Schatz M. 2019. Fast and accurate reference-guided scaffolding of draft genomes. bioRxiv doi:10.1101/519637

EPIGENETIC INHERITANCE IN PLANTS AND FISSION YEAST

R. Martienssen

C. Alves
E. Anderson
B. Berube
S. Bhattacharjee
F. Borges
W. Dahl
J. Dolata
E. Ernst
M. Gutbrod
H.S. Kim
S-C. Lee
J-S. Parent
U. Ramu
M. Regulski
J. Ren
B. Roche
A. Schorn
A. Shimada
J. Simorowski
J. Steinberg

Plants and fission yeast provide excellent models for epigenetic mechanisms of transposon regulation, heterochromatic silencing, and gene imprinting—important both for plant breeding and for human health. We are investigating the role of RNA interference (RNAi) in heterochromatic silencing in the fission yeast *Schizosaccharomyces pombe* and in the model plant *Arabidopsis thaliana*, as well as continuing our work on *Zea mays*. In fission yeast, we have found that RNAi promotes DNA replication and repair, as well as histone modification required for centromere function. In quiescence, RNAi becomes essential, because it is required for release of RNA polymerase. In the mouse, LTR retrotransposons are controlled by small RNA derived from tRNA that match the highly conserved primer binding site, in which retrotransposons are primed for reverse transcription. Remarkably, plant retrotransposons are similarly targeted by a microRNA (miRNA) in pollen. This miRNA triggers transposon small RNA that sense chromosome dosage and cause triploid seeds to abort. Along with DNA methylation, retrotransposon RNAi is required for centromere function and histone modification. And as in fission yeast, histone modification represses recombination near maize and *Arabidopsis* centromeres. We continue to develop duckweeds for biofuel by sequencing the genomes of several species and by developing an efficient transformation system in the clonally propagated aquatic macrophyte *Lemna minor*.

This year we said good-bye to postdocs Filipe Borges, Jean-Sebastien Parent, and Jie Ren, who obtained faculty positions at INRA Versailles in Paris, Agriculture Canada in Ottawa, and the Chinese Academy of Sciences in Beijing, respectively. We welcomed postdocs Cristiane Alves and Jakub Dolata and M.D./Ph.D. student Joshua Steinberg.

RNA Polymerase I Is Regulated by RNAi and Long Noncoding RNA in Cellular Quiescence

B. Roche, R. Martienssen [in collaboration with B. Arcangioli, Institut Pasteur, Paris]

Most cells in nature, and in the human body, are present in a nondividing state (G_0). Cellular quiescence is an important G_0 state characterized by its reversibility and metabolic activity, found, for example, in stem cells and memory lymphocytes. To establish and maintain quiescence, cells undergo a major transcriptional reprogramming. However, this phenomenon is still largely unexplored at the molecular level, and little is known about the mechanisms involved in maintaining viability in quiescent states. Because cellular quiescence is in many aspects an epigenetic transition, we hypothesized that specific epigenetic pathways would be involved. Indeed, we have found that several of these pathways are rewired in cellular quiescence and become essential specifically in G_0. We have designed a strategy to identify G_0-specific suppressors of these new functions and found that RNA interference acquires a novel, and essential, nucleolar function in G_0 cells, in which it regulates the epigenetic state of rDNA repeats by regulating RNA polymerase I transcription. Mutants in key RNAi factors, such as Dicer and Argonaute, display a major loss of viability specifically in quiescent cells because of an overaccumulation of H3K9 methylation on rDNA repeats. Strikingly, this phenomenon can be suppressed by specific mutants in the RNA polymerase I holoenzyme itself. Dicer catalytic activity is required, showing that specific target RNAs are likely to mediate this function. We have recently discovered a novel group of long noncoding RNAs that are strongly up-regulated in Dicer-deficient G_0

cells. In particular, we have determined that at least one of these noncoding RNAs is necessary for accumulating H3K9 methylation at rDNA, thus appearing to be the first *trans*-acting functional homolog of the pRNA in human cells, a classical RNA-mediated rDNA silencing factor.

tRNA Fragments as Regulators of Retrotransposons

A. Schorn, M. Gutbrod, R. Martienssen

tRNA fragments (tRFs) are a class of small, regulatory RNAs with diverse functions. The 3′-derived tRFs perfectly match long terminal repeat (LTR) retroelements, which use the 3′ end of tRNAs to prime reverse transcription. We have recently shown that tRFs target LTR retroviruses and retrotransposons for the RNAi pathway and also inhibit mobility by blocking reverse transcription. The highly conserved tRNA primer-binding site (PBS) in LTR retroelements is a unique target for 3′-tRFs to recognize and block abundant but diverse LTR retrotransposons that become transcriptionally active during epigenetic reprogramming in development and disease. 3′-tRFs are processed from full-length tRNAs under so far unknown conditions and potentially protect many cell types. tRFs appear to be an ancient link between RNAi, transposons, and genome stability.

Transposon-Derived Small RNAs Triggered by miR845 Mediate Genome Dosage Response in *Arabidopsis*

F. Borges, J-S. Parent, R. Martienssen [in collaboration with C. Kohler, Linnean Institute of Agricultural Sciences, Uppsala University]

One hundred years ago at Cold Spring Harbor, Albert Blakeslee discovered that newly arising tetraploid plants behaved like a new species, in that backcrosses to diploid progenitors resulted in reproductive isolation via triploid seed abortion. Later work in maize demonstrated that imbalance of imprinted genes in the endosperm was likely responsible. Chromosome dosage has substantial effects on reproductive isolation in both plants and animals, but the underlying mechanisms have remained obscure. We previously postulated that small RNAs in pollen could regulate dosage response by targeting transposons neighboring imprinted genes in the seed. We have now shown that a conserved microRNA in plants, miR845, targets the $tRNA^{Met}$ PBS of LTR retrotransposons in *Arabidopsis* pollen and triggers the accumulation of 21–22-nt small RNAs in a dose-dependent fashion via RNA polymerase IV. We show that these epigenetically activated small interfering RNAs (easiRNAs) mediate dosage response in crosses between diploid seed parents and tetraploid pollen parents (the "triploid block"), and that natural variation for miR845 may account for "endosperm balance," allowing the formation of triploid seeds. Targeting of the PBS with small RNA is a common mechanism for transposon control in mammals and plants and provides a uniquely sensitive means to monitor chromosome dosage and imprinting in the developing seed. In his last year as director of Cold Spring Harbor, Blakeslee hired Barbara McClintock, who also worked with chromosome dosage and discovered transposable elements.

RNA Interference Promotes *Arabidopsis* Chromosome Segregation in the Absence of DNA Methylation

A. Shimada, R. Martienssen [in collaboration with D. Grimanelli, Institut de Recherche pour le Développement, Montpellier; T. Kakutani, National Institute of Genetics, Mishima, Japan]

Small RNAs and DNA methylation have pivotal roles in transposon silencing. In *Arabidopsis*, 21-nt easiRNA arise from transposons in mutants defective in DNA methylation. This class of small RNAs is generated in somatic cells by RDR6 (RNA-dependent RNA polymerase), but its biological significance is unclear. We have shown that mutants that lose both DNA methylation and small RNA from pericentromeric regions, *rdr1 rdr2 rdr6 ddm1*, suffer from pleiotropic developmental defects. Fertility defects can be traced to the centromeric region of chromosome 5 and can be rescued by an ectopic increase in DNA methylation at the pericentromeric retrotransposon *ATHILA5*. Overexpression of *ATHILA5* did not impact developmental phenotypes, suggesting that proteins encoded by *ATHILA5* are not the cause of sickness. We observed that RDR6-dependent easiRNAs are required for

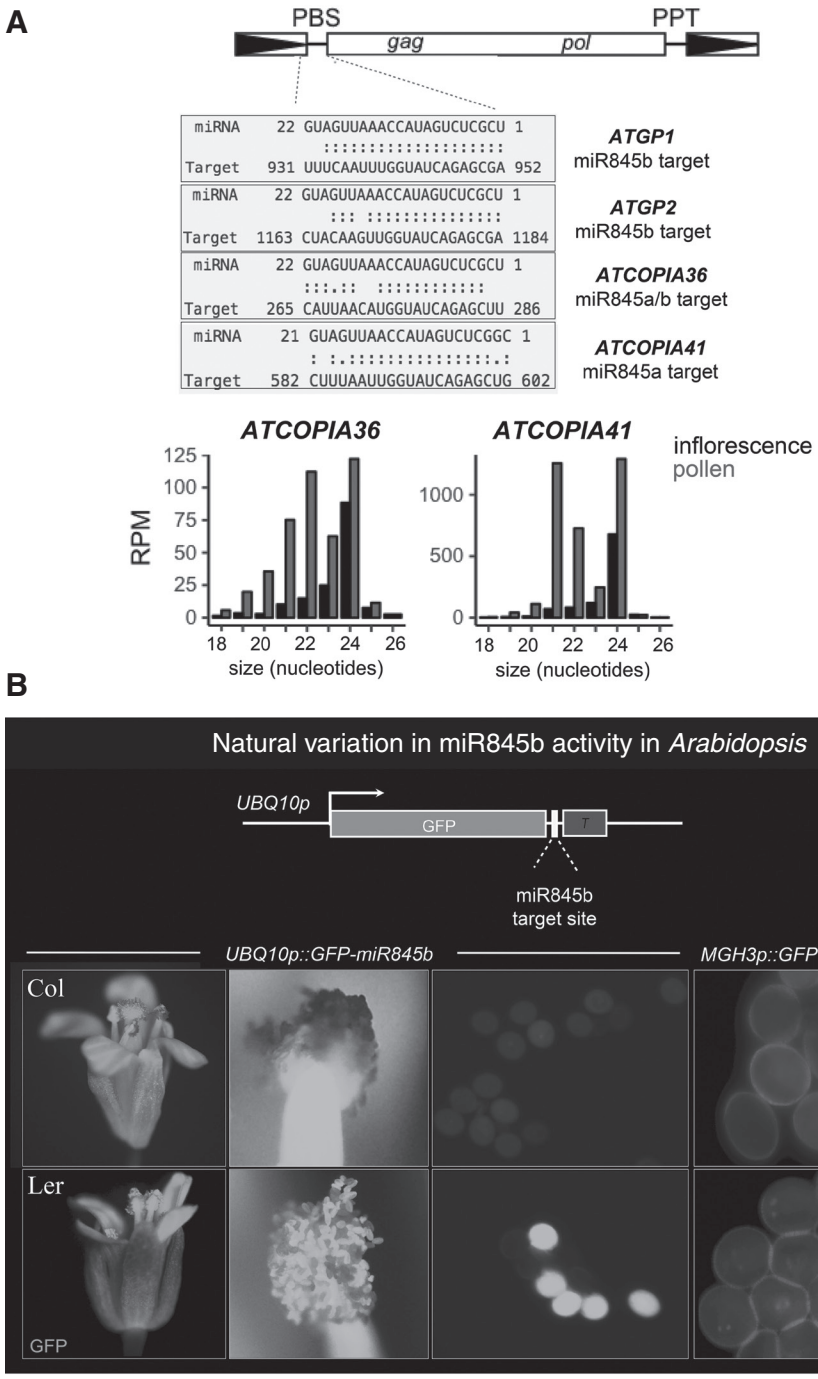


Figure 1. Pollen microRNA miR845 targets the primer-binding site of retrotransposons and is polymorphic among natural *Arabidopsis* accessions. (*A*) miR845 targets the tRNA primer binding site of *COPIA* and *GYPSY* LTR retrotransposons. These transposons give rise to 21–22-nt epigenetically activated small RNA in pollen (red histograms), but not in somatic tissues (black histograms). Levels of miR845 levels vary among natural *Arabidopsis* accessions, including Landsberg *erecta*, from which it is absent. (*B*) A GFP reporter gene, equipped with a miR845 target site, was introduced into transgenic plants in the Columbia (Col) and Landsberg *erecta* (Ler) genetic backgrounds. GFP expression was observed in all tissues in Landsberg *erecta*, including pollen and sperm, but was absent in Columbia pollen because of the activity of miR845.

the incorporation of histone H3 into pericentromeric heterochromatin and H3K9 methylation, and that loss of easiRNAs results in diminished chromocenters and severe mitotic mis-segregation of chromosome 5. Importantly, reformation of heterochromatin at the *ATHILA5* locus by artificial siRNAs is sufficient to rescue the chromosome mis-segregation defect. We propose a model in which RDR6-dependent easiRNAs compensate for loss of DNA methylation, ensuring centromeric structure and genome integrity.

DDM1 Cooperates with easiRNA to Control Transposons via Chromatin Remodeling

S.C. Lee, E. Ernst, J-S. Parent, R. Martienssen [in collaboration with D. Grimanelli, Institut de Recherche pour le Développement, Montpellier]

Heterochromatin is largely comprised of transposons and repeats, which contain epigenetic signatures such as DNA methylation and histone modifications. Histones H3.1 and H3.3 are two major H3 histone variants deposited in heterochromatin and euchromatin, respectively. We found that DDM1 is involved in the deposition of H3.1 in *Arabidopsis*. H3.1 incorporation was markedly compromised in *ddm1,* whereas H3.3 accumulation was enhanced in heterochromatin and DNA methylation and H3K9me2 marks were greatly reduced. H3K27me3 is normally a euchromatic silencing mark for genes but spreads to heterochromatin in *ddm1.* As a consequence of improper regulation of DNA methylation and histone modification in *ddm1* mutants, a large number of transposons are transcriptionally activated, resulting in accumulation of easiRNA by RDR6. We found *ddm1rdr6* displayed higher levels of H3K27me3 and lower levels of H3K9me2 than did *ddm1* single mutants, suggesting easiRNAs may play a role in histone modification when DNA methylation levels are low. *EVADE* is a Copia LTR retrotransposon that is activated in *ddm1* plants and can be epigenetically inherited by wild-type plants in an active form. In wild-type backgrounds, *EVADE* is maternally silenced but remains active when inherited through the male (pollen). Our data suggest *EVADE* easiRNAs are absent in pollen but present in flowers of *ddm1* plants, consistent with maternal silencing. Thus, DDM1 and RDR6-dependent easiRNAs play crucial roles in shaping heterochromatin to maintain genome stability.

Epigenetic Activation of Meiotic Recombination near *Arabidopsis thaliana* Centromeres via Loss of H3K9me2 and Non-CG DNA Methylation

C. Underwood, J. Simorowski, E. Ernst, R. Martienssen [in collaboration with I. Henderson, Department of Plant Sciences, Cambridge University]

Eukaryotic centromeres contain the kinetochore, which connects chromosomes to the spindle, allowing segregation, flanked by pericentromeric heterochromatin. During meiosis, centromeres are suppressed for interhomolog crossover, as recombination in these regions can cause chromosome mis-segregation and aneuploidy. Plant centromeres are surrounded by transposon-dense pericentromeric heterochromatin that is epigenetically silenced by histone 3 lysine 9 dimethylation (H3K9me2), and DNA methylation in CG and non-CG sequence contexts. We have found that disruption of *A. thaliana* H3K9me2 and non-CG DNA methylation pathways (e.g., via mutation of the H3K9 methyltransferase genes, or the CHG DNA methyltransferase gene *CMT3*) increases meiotic recombination in proximity to the centromeres. Using immunocytological detection of MLH1 foci and genotyping by sequencing, we observed that H3K9me2 and non-CG DNA methylation pathway mutants have increased pericentromeric crossovers and likely involve contributions from both the interfering and noninterfering crossover repair pathways. We also showed that meiotic DNA double-strand breaks (DSBs) increase in H3K9me2/non-CG mutants within the pericentromeres, but also in *met1* mutants, coincident with reduced nucleosome occupancy and gain of transcription and H3K4me3. However, the frequency of crossover in these regions was reduced in *met1* mutants. Therefore, although DNA methylation represses meiotic DSB, H3K9me2 and non-CG DNA methylation repress crossover formation in plant pericentromeric heterochromatin. Our results may account for the selection of enhancer trap *Dissociation* (*Ds*) transposons into the *CMT3* gene by recombination with proximal transposon launchpads.

Single Pollen Sequencing Reveals Recombination Control by H3K9me2 and Non-CG Methylation in Maize

B. Berube, M. Regulski, R. Martienssen [in collaboration with D. Grimanelli, Institut de Recherche pour le Développement, Montpellier]

Meiotic recombination is a fundamental evolutionary driver and an indispensable tool for agricultural breeding. Advances in next-generation sequencing, coupled with genotyping by sequencing approaches, have allowed for the development of large-scale, population-level assessments of crossover frequency and localization, as well as the construction of high-density linkage maps for quantitative trait mapping. Despite these advances, unbiased, genome-wide study of recombination in meiotic mutants remains difficult. The computational complexity of identifying crossovers in hybrid populations, constraints on the availability of plant material, and variable contributions of both the maternal and paternal genomes all pose significant challenges. We have developed a single-cell genomic DNA sequencing method to sequence the genome of individual maize pollen grains. Sequencing of pollen grains from hybrid populations, coupled with statistical modeling of recombination breakpoints, allows for genome-wide, high-resolution mapping of crossover intervals. Using this approach, the distortion of patterns of recombination in DNA methyltransferase mutants is being characterized.

Genome-Wide Analysis of *Arabidopsis* Replication and DNA Methylation

F. Borges, C. LeBlanc [in collaboration with M. Vaughn, University of Texas at Austin; W. Thompson and L. Hanley-Bowdoin, North Carolina State University]

Although replication timing programs have been described for yeast and animal systems, much less is known about the temporal regulation of plant DNA replication or its relationship to genome sequence and chromatin structure. We used the thymidine analog, 5-ethynyl-2′-deoxyuridine, in combination with flow sorting and Repli-Seq to describe, at high resolution, the genome-wide replication timing program for *A. thaliana* Col-0 suspension cells. We identified genomic regions that replicate predominantly during early, mid, and late S phase and correlated these regions with genomic features and with data for chromatin state, accessibility, and long-distance interaction. *Arabidopsis* chromosome arms tend to replicate early, whereas pericentromeric regions replicate late. Early- and mid-replicating regions are gene-rich and predominantly euchromatic, whereas late regions are rich in transposable elements and primarily heterochromatic. The distribution of chromatin states and DNA (cytosine) methylation is complex, with each replication time corresponding to a mixture of states. During DNA replication, 5-methylcytosines in symmetric contexts (mCG) are guided for methylation by hemimethyl C on parental DNA strands. But in plant genomes, replication of non-CG methylation needs modified histones and small RNA guides. By sequencing bisulfite-treated EdU-labeled nuclei sorted by cell cycle stage, we have determined genome-wide methylation in *Arabidopsis* cell cultures. We find mCG becomes transiently asymmetric during late S phase but is rapidly restored, whereas mCHG (in which H = A, T, or C) remains asymmetric and is substantially reduced. mCHH is almost absent from dividing cells, accompanied by a shift from 24-nt to 21-nt siRNA and a loss of H3K9me2. In arrested cells, mCHH is restored, but only at loci that retained 24-nt siRNAs. Comparisons with methylomes and small RNA profiles in pollen cell types and in cell culture indicate that germline reprogramming reflects cell cycle stage, as does loss of non-CG methylation in plants micropropagated as clones.

The *Karma* of Clones: Genomics and Epigenomics of *Lemnaceae*

E. Ernst, R. Martienssen [in collaboration with T. Michael, J. Craig Venter Institute, La Jolla; E. Lam, Rutgers University]

Genomics has greatly impacted breeding in domesticated crops, through genome- and marker-assisted selection, as well as finding genes underlying key traits. But crops with only minimal domestication, such as oil palm and duckweed, present challenges to this approach, especially when they are propagated asexually as clones. Although clones have the advantage of potentially fixing hybrid vigor, lack of germline passage, in which epigenetic reprogramming occurs,

can lead to epigenetic variation. Advances in single-molecule genomic sequencing and epigenomic profiling have enabled the rapid generation and assembly of large and highly repetitive plant genomes. In oil palm, these technologies have enabled the identification of key genetic traits underlying yield, as well as epigenetic traits arising from loss of *Karma* transposon DNA methylation in tissue culture. We are now exploring potential implications for natural variation, genetic modification, and next-generation biofuels in Lemnaceae, aquatic plants with rapid clonal growth habit. We have employed long single-molecule sequencing from PacBio and Oxford Nanopore to complete the first long-read chromosomal assemblies of the *Lemna gibba* 7742 and *Lemna minor* 8627 genomes, as well as *Spirodela polyrhiza* in collaboration with our colleagues, along with updated gene annotations informed by full-length cDNA sequences from the ONT MinION. We have found that all three genomes lack CHH methylation and have reduced 24-nt siRNA, despite having the requisite Dicer and DNA methyltransferase genes. We are investigating the possibility that loss of RdDM reflects continuous clonal growth.

PUBLICATIONS

Borges F, Parent JS, van Ex F, Wolff P, Martínez G, Köhler C, Martienssen RA. 2018. Transposon-derived small RNAs triggered by miR845 mediate genome dosage response in *Arabidopsis*. *Nat Genet* **50:** 186–192. doi:10.1038/s41588-017-0032-5

Choi K, Zhao X, Tock AJ, Lambing C, Underwood CJ, Hardcastle TJ, Serra H, Kim J, Cho HS, Kim J, et al. 2018. Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. *Genome Res* **32:** 532–546. doi:10.1101/gr.225599.117

Concia L, Brooks AM, Wheeler E, Zynda GJ, Wear EE, LeBlanc C, Song J, Lee TJ, Pascuzzi PE, Martienssen RA, Vaughn MW, Thompson WF, Hanley-Bowdoin L. 2018. Genome-wide analysis of the *Arabidopsis* Replication Timing Program. *Plant Physiol* **176:** 2166–2185. doi:10.1104/pp.17.01537

Ikeda Y, Nishihama R, Yamaoka S, Arteaga-Vazquez MA, Aguilar-Cruz A, Grimanelli D, Pogorelcnik R, Martienssen RA, Yamato KT, Kohchi T, et al. 2018. Loss of CG methylation in *Marchantia polymorpha* causes disorganization of cell division and reveals unique DNA methylation regulatory mechanisms of non-CG methylation. *Plant Cell Physiol* **59:** 2421–2431. doi:10.1093/pcp/pcy161

Schorn AJ, Martienssen R. 2018. Tie-break: host and retrotransposons play tRNA. *Trends Cell Biol* **28:** 793–806. doi:10.1016/j.tcb.2018.05.006

Underwood CJ, Choi K, Lambing C, Zhao X, Serra H, Borges F, Simorowski J, Ernst E, Jacob Y, Henderson IR, Martienssen RA. 2018. Epigenetic activation of meiotic recombination near *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Res* **32:** 519–531. doi:10.1101/gr.227116.117

MOLECULAR SIGNALING EVENTS UNDERLYING ENVIRONMENTAL CONTROL OF PLANT GROWTH

U. Pedmale M. Biggiani A. Matheus
J. Hart S. Sankaranarayanan
Y. Hu O. Spassibojko
L.N. Lindbäck

A fundamental question in biology, which remains unanswered, is how the environment of the organism regulates its growth and development. Both plants and animals interact with their environment; however, plants grow postembryonically as they are incapable of moving around. Unlike animals, plants do not have specific organs that either see or hear various stimuli, yet, plants are sensitive to their surrounding environment and modify their growth according to various external and internal signals. Plants regularly face variability in growth conditions—temperature, light quality and quantity, herbivores, pathogens, water availability, etc. Plants respond to these biotic and abiotic factors and survive substantial fluctuations in their environment. Plants also must balance the range of potential threats and benefits confronting them and should make appropriate decisions on resource allocation. Remarkably, lacking a brain, plants can successfully integrate various cues and make appropriate decisions about growth. Such adaptability is essential to the sessile nature of the plants. In some adaptive responses—for example, when the plants have to cope with climate change and increased competition for light—there is a decrease in productivity (yield, biomass) as the plant reallocates resources to adapt better.

The goal of our laboratory is to determine the mechanisms behind how a plant perceives and successfully adapts to its environment. We also aim to understand how a plant must integrate intrinsic and extrinsic cues and "decide" how best to respond to environmental cues. Understanding how plants deal with and respond to a multitude of environmental signals could help to develop crops that cope with unfavorable growth conditions without significant changes in yield. Our laboratory primarily studies the effect of light environment on plant growth and development. Light is among the most relevant environmental signals because it not only drives photosynthesis but also provides critical information about the local growth environment as well as seasonal time. Light is perceived by a complex array of photoreceptors, which include phytochromes (PHYA-E), cryptochromes (CRY1-2), phototropins (PHOT1-2), the zietlupe family (FKF1, LKP2, and ZTL), and UVR8. Plants have developed various adaptive responses to interpret and utilize light directionality, quantity, and quality. In vegetational shading, when plants are under the shade of another plant, they perceive a decrease in the ratio of red to far-red light (R:FR) due to absorption of red light by chlorophyll and reflection of far-red light by the neighboring foliage. Simultaneously, there are also decreases in blue light and the available photosynthetically active radiation (PAR).

We focus on blue light–absorbing CRYs. Apart from being an excellent genetic and molecular tool to tease out the complexities of growth and adaptation, there are still many open questions about the molecular function of CRYs in plants. Understanding the role of CRYs is not only important for agriculture, but it also has an impact on human health, which could make this field appealing to diverse funding agencies. CRYs regulate growth and development and provide circadian entrainment to both plants and animals. In metazoans, disruption of CRY activity is linked to cancer, altered behavior, magnetoreception, and metabolism. Therefore, understanding CRY function in plants is not only important as they are integral to growth of plants but also can have an impact on human health.

Decoding the Nature of Long-Distance Communication in Plants: The Case of Shoot Control of Root Growth

During shading, many aerial organs elongate rapidly, whereas the root growth is reduced with the delay

in the emergence of the lateral roots. Roots not only serve as a mechanical anchor but play a vital role in the well-being of the entire plant. Therefore, a robust and well-developed root system is required for healthy plant growth. As one can imagine, there is a negative cycle occurring during shading; shoot-perceived shade leads to reduced root growth, which in turn is unable to support the shoot, leading to unproductive plants. However, this phenomenon is an excellent model to understand growth at a systems level because of the different growth phenotypes observed in the various organs of the same plant, as well as enabling exploration of the nature of the interorgan and long-distance communication that is used to signal when a distant organ is exposed to an adverse environment. Unfortunately, and surprisingly, not much is known about the mechanisms that underlie reduced root growth seen during shading.

During shading, newly synthesized auxin hormone in the cotyledons is required for the hypocotyl growth. First, we measured free auxin levels in the dissected root and shoot of seedlings exposed to low R:FR. As expected we saw an increase in free auxin in the shoot and interestingly a 10-fold accumulation in the root. However, we failed to see the activity of various auxin reporters in the root compared to the hypocotyl. High auxin levels are known to stimulate lateral root growth, though, intriguingly, we did not detect growth of lateral roots, but lateral root primordia were present. These results indicate that although auxin accumulation occurred, auxin signaling or perception was blocked in the root during shading. We also performed broad time course transcriptomic analysis from excised cotyledons, hypocotyl, and roots from *Arabidopsis* seedlings exposed to shade and nonshading conditions. Our transcriptomic analysis detected induction of auxin-responsive genes in the hypocotyl but not in the root. Interestingly, genes induced during abiotic and biotic stress responses were significantly up-regulated in the root. To further test whether the roots in shaded plants activated stress responses to slow down their growth, we also employed various reporters that are activated during stress or defense against pathogens. Furthermore, we have identified genes that are specifically expressed in the roots under shade (Fig. 1), which can now be exploited further as a tool to uncover the signal originating from the shoots. Therefore, our results suggest that stress responses in the root are likely inhibiting its growth. Furthermore, it is known that the plant can either grow or defend, and it is probable that in shaded plants, by turning on the defense or stress response in the roots, resources are diverted to the shoot organs, which then compete for light. However, the molecular determinant that switches a plant between these two states of defense/stress and growth is unknown and is of immense interest to biologists. We are taking several approaches to identify this molecular switch.

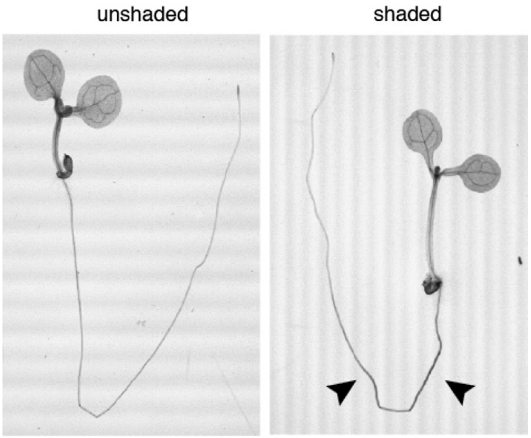


Figure 1. *Arabidopsis* seedlings expressing a gene promoter that is specifically activated in the roots of plants grown under shade. A GUS histochemical marker (marked by arrowheads) is shown to be expressed in the shade under the control of this promoter.

The Role of RNA-Binding Proteins in Cryptochrome-Mediated Signaling

We also identified many RNA-binding proteins that co-purified with CRY2. Recent studies reported co-purification of RNA-binding proteins with human CRY1/2. However, the significance of CRYs associating with RNA-binding proteins is not known. Interestingly, the CRY2 nuclear speckles resemble those formed by pre-mRNA splicing factors, SR proteins, and other RNA-binding proteins in plants and animals, indicating that CRYs likely have a role in RNA metabolism. This observation may provide mechanistic insights into posttranscriptional control, known to be essential for animal and plant circadian biology, and into control of alternative splicing in plants by light. We are currently focusing on two unknown proteins that are hypothesized to bind to modified

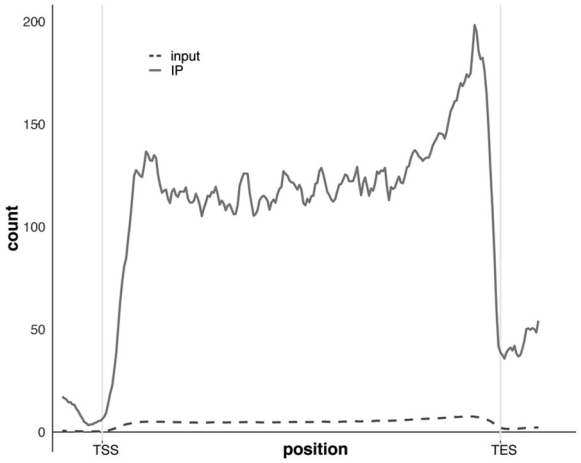


Figure 2. Genome-wide metagene analysis showing enrichment of a specific RNA modification around the transcriptional end site (TES) of genes compared to the rest of the gene body. Enrichment of a specific RNA modification (IP) was conducted using modification-specific pull. (TSS) Transcription start site.

RNA; their mutants resemble *cry2* mutant plants, indicating an epistatic relationship between them. Furthermore, similar to our experiments in plants, we are also studying their orthologs in animals to discover whether they can also interact directly with mammalian CRY2. We have started profiling RNA modification in different mutant backgrounds using genome-wide approaches (Fig. 2). RNA modifications are emerging as important regulators of various cellular processes not limited to protein translation, mRNA degradation, alternative splicing, and nuclear export. We are generating loss-of-function mutants of these RNA-binding proteins to understand their impact on plant growth, circadian rhythms, alternative splicing, and other physiological responses.

Molecular Determinants of CRY2 Protein Signaling and Stability

CRYs were first identified in plants and then discovered in animals. CRY2 protein accumulates in the dark and in vegetational shade and is readily degraded by the 26S proteasome under high intensities of BL. Therefore, it is obvious that the CRY2 protein level and activity are tightly regulated by light in order to ensure proper signaling and response. However, the signaling events from the photoactivated CRYs to growth and development programs are not known. In animals and plants, CRY protein levels and activity are tightly modulated to influence signaling outcome. Therefore, to elucidate the CRY signaling pathway, our laboratory has purified CRY2-containing protein complexes from *Arabidopsis thaliana* seedlings exposed to low-intensity blue light, which is typically encountered under shading. We identified CRY2-associated proteins by tandem affinity purification and mass spectrometry. Interestingly, the orthologs of some of the CRY2-associated proteins were also present in CRY protein complexes purified from human cells. This indicates that there could be a similar signaling mechanism in these two different evolutionary lineages.

Two of the highly enriched proteins in the CRY2-associated protein complex were deubiquitinases called UBP12 and UBP13. Deubiquitinases remove the ubiquitin protein covalently bound to a target protein. We found out that CRY2 and these deubiquitinases can interact directly in the nucleus of the cell. We hypothesize that CRY2 is ubiquitinated continuously, but, in certain situations, it recruits deubiquitinases to protect itself from proteasomal degradation such that downstream signaling can proceed. Next, we tested CRY2 protein levels in the deubiquitinase mutants and in plants in which they are overexpressed. Surprisingly, we found that CRY2 protein levels were very high in the deubiquitinase mutant and lower when overexpressed (Fig. 3). This matches with the physiological response exhibited by the seedling stem length of these genetic backgrounds. The deubiquitinase mutant seedling had a short hypocotyl when compared to the wild-type and the *cry2* mutant, and the overexpression line had a longer hypocotyl similar to *cry2* (Fig. 3). This observation indicates that the deubiquitinase-CRY2 protein complex likely modifies a protein that affects CRY2 protein levels like a E3 ubiquitin ligase. Efforts are under way to identify ubiquitylated residues in the CRY2 protein and to evaluate the effect of catalytically dead deubiquitinases in plants. Using seedlings expressing catalytically dead deubiquitinase, we have identified several candidate E3 ligases that may affect CRY2 protein levels. Unlike animals, substrates for the large number (approximately 50) of plant deubiquitinases remain unidentified, except for histones. In parallel, to identify the E3 ubiquitin ligase responsible for targeting plant CRY2 for degradation, we are undertaking a forward genetic

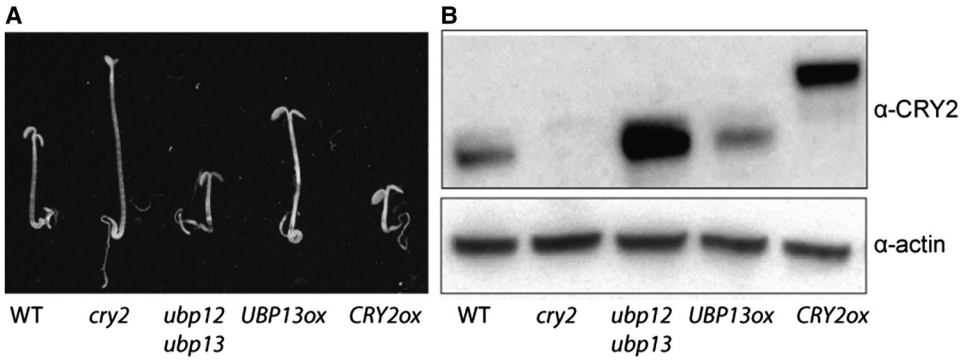


Figure 3. (*A*) Hypocotyl phenotype of seedlings of indicated genotypes grown in blue light for 4 d and (*B*) the CRY2 protein levels determined by immunoblotting using its specific antibody.

screen. Thus, we are presented with a unique opportunity to study how deubiquitinases participate in the CRY2 signaling pathway and also their role in plant growth and development. In conclusion, our findings will provide novel insight into the regulation of CRYs by reversible ubiquitination, as well as the role of deubiquitinases in plant growth and development, which is largely unknown so far.

GENOMICS

There has been a growing appreciation in recent years that gene function is frequently context-dependent, with a large part of that context provided by the activities of other genes. But trying to understand how genes interact to produce function is a hugely complicated problem, and one that appears likely to become more so as genomic information becomes more detailed. **Jesse Gillis** and colleagues are computational biologists who are challenging an oft-taken approach to the problem in which the functions of genes are interpreted in the context of networks derived from gene association data. Such networks consist of millions of interactions across thousands of genes, derived from protein-binding assays, RNA coexpression analysis, and other sources. Historically, many attempts to understand gene function through networks have leveraged a biological principle known as "guilt by association." It suggests that genes with related functions tend to share properties (e.g., physical interactions). In the past decade, this approach has been scaled up for application to large gene networks, becoming a favored way to grapple with the complex interdependencies of gene functions in the face of floods of genomics and proteomics data. Gillis' work centers on identifying the limits of the approach and making fundamental improvements to its operation, as well as applying those improvements to neuropsychiatric gene network data.

Thomas Gingeras and colleagues study where and how functional information is stored in genomes. These efforts help explain the biological and clinical effects of disease-causing gene mutations in humans and other organisms. Gingeras is a leader of the ENCODE (ENCyclopedia of DNA Elements) and the mouseENCODE and modENCODE (model genome ENCODE) projects of the National Institutes of Health. His research has altered our understanding of the traditional boundaries of genes, revealing that almost the entire lengths of genomes in organisms ranging from bacteria to humans can be transcribed into RNA (pervasive transcription) and most RNA products made by a cell are not destined to be translated into proteins (noncoding RNAs, or ncRNAs). In fact, ncRNAs are proving to be involved in a variety of other important biological functions. Some have been shown to be critical components in the pre- and posttranscriptional and translational processes, as scaffolds on which large protein complexes are assembled and as extracellular signals. The initial studies that led to these observations have been extended to cover the entire human genome.

The insights of **W. Richard McCombie** and colleagues have led to the introduction and optimization of novel methods of high-throughput genome sequencing. His team has made it possible to catalog variation among individual organisms in a way that would have been unthinkable 10 years ago. They have brought online a new generation of Illumina sequencers and optimized their function to a level at which eight to 10 trillion DNA bases can be sequenced in a month. McCombie's team has been involved in international efforts culminating in genome sequences for maize, rice, and bread wheat—three of the world's most important food crops. They have also had an important role in projects to sequence the flowering plant *Arabidopsis thaliana* (the first plant genome sequence) and the fission yeast *Schizosaccharomyces pombe,* as well as the human genome and other important genomes. McCombie's group is currently involved in several important projects to resequence genes in patient samples that are of special interest to human health, including *DISC1* (a strong candidate gene for schizophrenia), looking for genetic variants implicated in bipolar illness and major recurrent depression. They are also looking for genes that contribute to cancer progression using whole-genome sequencing or a method called exome sequencing, which they developed with Greg Hannon, to look at mutations in the regions of the genome that code for proteins.

Using multidisciplinary approaches that combine computational analysis, modeling, and prediction with experimental verification, **Doreen Ware**'s lab seeks a deeper understanding of the evolution of genome sequences in plants and their implications for agricultural improvement. By looking comparatively across the genomes of plants in the same lineage, they seek answers to the following questions: How are genes conserved and lost over time? What are the fates of duplicated genes? What is the impact of structural variation on phenotypic variation? Ware's team also studies gene regulation in plants, focusing on gene regulatory networks, targeting transcription factors, and microRNA genes with the objective of understanding how these parts of the plant genome work together in determining spatial and temporal expression of genes. The lab had an important role in the project to produce a haplotype map reference genome of maize, spearheading the most comprehensive analysis of the crop yet. This has provided important information on the variation of the reference genome, as well as comparative data showing changes in the genome acquired through domestication and breeding. They have devoted special attention to examining diversity within maize, grape, and tomato, aiming to accelerate the development of strategies to introduce new germplasm, which is needed to meet demands of increasing population and a changing environment. The lab also has brought fully sequenced genomes into an integrated data framework to enhance the power of their comparative studies. This past year, Ware was named as principal investigator for the National Science Foundation–funded Gramene project, a comparative genomics resource for agriculturally important crops and models to support sustainable food and fuel production. Ware, as principal investigator for plants, has also helped lead an effort funded by the Department of Energy to create—out of many separate streams of biological information—a single, integrated cyber-knowledgebase for plants and microbial life.

GENE EXPRESSION, FUNCTION, AND META-ANALYSIS

J. Gillis S. Ballouz S. Fischer S. Lu
M. Crow B. Harris M. Shah

Research in the Gillis laboratory involves two interwoven elements: improving the interpretability of network analysis and characterizing transcriptional data in the brain. These topics form a naturally complementary unit because the complexity of the brain as a system means that it is essential that the methods for analyzing it yield clear and precise signals. A dominant interest within computational biology is the analysis of gene networks to provide insight into diverse levels of functional activity, typically starting with regulatory interactions and moving up to more diffuse associations important for understanding systemic dynamics. Gene associations (of various sorts) are believed to encode functional interaction, and this interaction is frequently shown to be able to substantially predict gene function. This approach, commonly called "guilt by association," is embedded in everything from prioritization of de novo variants to uncovering novel regulatory interactions or mechanisms of disease. Although black box style network analyses are common, explaining the basis of how and why methods work is more rarely attempted. In the Gillis laboratory, we are developing network-based methods and software that improve both the sophistication and breadth of data available for determining how genes interact to produce function, particularly focusing on how genes interact to cause organismal or cell phenotypes. Broadly, our research can be divided into methods development and our own research applications, often carried out in collaboration with other laboratories to test computational predictions experimentally. In addition to Jesse Gillis, the members of the Gillis laboratory are postdoctoral fellows Sara Ballouz, Maggie Crow, and Stephan Fischer; graduate students Shaina Lu and Benjamin Harris; and computational science developer Manthan Shah.

Understanding Drivers of Differential Expression

Differential expression (DE) is commonly used to explore molecular mechanisms of biological conditions. Although many studies report significant results between their groups of interest, the degree to which results are specific to the question at hand is not generally assessed, potentially leading to inaccurate interpretation. This could be particularly problematic for meta-analysis in which replicability across data sets is taken as strong evidence for the existence of a specific, biologically relevant signal, but which instead may arise from recurrence of generic processes. To address this, we developed an approach to predict DE based on an analysis of more than 600 studies. A predictor based on empirical prior probability of DE performs very well at this task (mean area under the ROC curve [AUROC] ~0.8), indicating that a large fraction of DE hit lists is nonspecific. In contrast, predictors based on attributes such as gene function, mutation rates, or network features perform poorly. Genes associated with sex, the extracellular matrix, the immune system, and stress responses are prominent within the "DE prior." In a series of control studies we show that these patterns reflect shared biology rather than technical artifacts or ascertainment biases. Finally, we demonstrate the application of the DE prior to data interpretation in three use cases: (1) breast cancer subtyping, (2) single-cell genomics of pancreatic islet cells, and (3) meta-analysis of lung adenocarcinoma and renal transplant rejection transcriptomics. In all cases, we find hallmarks of generic differential expression, highlighting the need for nuanced interpretation of gene phenotypic associations.

Signatures of Transcriptomic Individuality Revealed by the Nine-Banded Armadillo

A delicate balance between robustness and stochasticity is present within developmental programs in mammals. The outcome is individuality, or phenotypic uniqueness. One way to assess purely epigenetic variability leading to different phenotypes is to look at the variation of phenotypes across identical genomes, such as twins. But because twinning occurs rarely

and randomly in humans, there are still uncertainties associated with the contributions of "nature, nurture, or noise." In recent work, we exploit the obligate polyembrony of the wild nine-banded armadillo (*Dasypus novemcinctus*), which is programmed to produce monozygotic quadruplets, to test for "epigenetic signatures" of individuality.

First, we sequenced the blood transcriptomes of five litters of armadillo quadruplets to generally assess and quantify epigenetic variation, using gene expression as a proxy for epigenetic marks. Then, we looked longitudinally across a year to see if this variability is itself stable and unique to genetic background or a common signature of armadillos. Unsurprisingly, we find a "sibling effect" (i.e., within-litter transcriptional similarity is always significantly higher than between-litter similarity). But remarkably, we find weak but significant "individuality" signatures of siblings. Mostly, these signatures are unique to each armadillo litter, implying distinct processes being affected. Although expression level of individual genes alone provides a significant signal ($p < 0.05$), allele-specific expression patterns exhibit much greater individuality ($p < 1E\text{-}3$), allowing armadillos to be uniquely identified relative to their siblings over time.

Consensus Genomes

A major intent behind the original sequencing of the human genome was to provide a reference for future analysis, and this has been wildly successful. The human reference genome has shaped methods and data across modern genomics, as it functions as the backbone for all genomic data and databases. It provides a scaffold in genome assembly to a background for read sequence alignment. However, there are some issues with the current reference—in particular, its nearness to a "type specimen." In recent work with the Dobin laboratory, we asked whether updating the current reference genome to a consensus genome—a genome that contains variation representing the most common alleles of a population—would have a strong impact on current practices such as DNA alignment, RNA-Seq alignment, and variant calling. We first assessed the current reference, and we show it is a good model of a human but not a general reference. We then built a series of population-specific consensus genomes and benchmarked their use in place of the current reference. We suggest that switching to a consensus reference offers important advantages over the current reference with few disadvantages.

Cross-Kingdom Transcriptomics

Coexpression networks describe shared expression profiles between genes. Because shared expression patterns often reflect shared function, coexpression networks describe the functional relationships between genes. In a large-scale meta-analysis, members of the Gillis laboratory have been constructing coexpression networks covering species from across the tree of life. By carefully aggregating preexisting expression data, we have developed "gold standard" networks describing the relationships between genes as well as developing tools to interpret these networks for targeted biological insight. A particular focus of our recent efforts has been the use of these networks to understand molecular functions in single-cell data. In collaboration with the Jackson laboratory, we have been assessing single-cell maize data and are working to characterize functions underlying cell type variation and shared in the independent evolution of multicellularity in plants and animals.

Exploring Cellular Heterogeneity in X-Linked Inflammatory Function

Cellular heterogeneity is a critical feature of both healthy function and disease. Recent advances in single-cell RNA-Seq have led to numerous insights in both areas, highlighting how cell types arise through the activities of coexpressed sets of genes as well as how variability across cells can have a strong prognostic significance in cancer. Understanding how individual cells contribute to organismal phenotype is often inherently limited in its degree of precision and control; variation measured between individuals rarely targets cell type–specific effects. To directly assay the impact of cell heterogeneity in disease, we have been conducting research to exploit a system with an inherent internal control, in the form of X-inactivation status within female carriers of an X-linked disorder. Which X-chromosome is inactivated varies from cell to cell in women in a cell lineage–specific fashion, leading to nonrandom trends within tissues, defined by the local

environment. The contribution of individual cells to the phenotype is clearly nonlinear because female carriers do not generally exhibit, for example, half as strong a phenotype as males for the X-linked disorder and are typically much less symptomatic. However, female carriers of X-linked disorders can demonstrate skewed or nonskewed X chromosome inactivation (XCI) and this tends to increase with age. When the wild-type allele is on the inactivated X chromosome (Xi), the carrier can become symptomatic if wild-type protein function drops below critical thresholds. X-linked chronic granulomatous disease (X-CGD) exhibits nonskewed XCI of the CYBB gene. However, some X-CGD carriers exhibit skewed XCI favoring inactivation of wild-type allele, with increased susceptibility to infections and autoimmunity as CYBB-encoded neutrophil oxidative burst activity decreases below 20% of wild type. By careful study of these individuals and the cellular contribution to their specific disease phenotype, our research works to disentangle the link between heterogeneous cellular phenotypes and health.

RNA-Seq Alignment

A major computational challenge in RNA sequencing is the quantification of gene expression. At present, the most commonly used approach consists of mapping RNA-Seq reads to a reference, followed by calculation of the transcript or gene expression. Many tools are available for RNA-Seq alignment and expression quantification, with comparative value being hard to establish. Benchmarking assessments often highlight methods' good performance, with occasional outlier tools, but are focused on either model data or fail to detail variation in performance. This leaves us to ask, what is the most meaningful way to assess different alignment choices? And importantly, where is there room for progress? In this work, we explore the answers to these two questions by performing an exhaustive assessment of the STAR aligner in collaboration with Alex Dobin and Tom Gingeras. We assess STAR's performance across a range of alignment parameters using common metrics and then on biologically focused tasks. We find technical metrics such as fraction mapping or expression profile correlation are uninformative, capturing properties unlikely to have any role in biological discovery. Surprisingly, we find that changes in alignment parameters within a wide range have little impact on both technical and biological performance. Yet, when performance finally does break, it happens in difficult regions, such as X-Y paralogs and *MHC* genes. We believe improved reporting by developers will help establish where results are likely to be robust or fragile, providing a better baseline to establish where methodological progress can still occur.

Expression Outliers in Rare Disease

In characterizing a disease, it is common to search for dysfunctional genes by assaying the transcriptome. The resulting differentially expressed genes are typically assessed for shared features, such as functional annotation or coexpression. Although useful, the reliability of these "systems" method is far less clear. To better understand shared disease signals, we assess their replicability by first looking at gene-level recurrence and then pathway-level recurrence along with coexpression signals across the six pedigrees of a rare homogeneous X-linked disorder, TAF1 syndrome. We find most differentially expressed genes are not recurrent between pedigrees, and functional enrichment is largely unique. We do find two highly recurrent "functional outliers" (CACNA1I and IGFBP3), genes acting atypically with respect to coexpression and therefore missed by standard systems assessment. We show this occurs in reanalysis of public Huntington's disease, Parkinson's disease, and schizophrenia data. Our results suggest a significant role for genes missed in systems approaches.

Learning Cell Identity

The exceptional diversity of neurons has been appreciated from the time of Ramón y Cajal. His depictions of cells that varied wildly in size, shape, and connectivity provided foundational guidance in the field of neuroscience. The precise extent of this diversity has been debated since that time, with each new technology indicating novel facets of neuronal identity: from morphology, to electrical activity, and now single-cell transcriptomics. When faced with a transformative technology, like single-cell RNA-Seq, how are we to make sense of the varied and distinct results of papers attempting to reshape the landscape

of molecular neuroscience? One approach is to build on a firm bedrock of our prior knowledge; comparison across studies to detect replicability is another. Our ongoing work with the Huang laboratory combines these two ideas in a naturally complementary way: By assessing clusters of cells for their overlap across studies with respect to the known functional properties of the genes that describe them, we characterize not just the overlap in transcriptional signal but also its functional implications. These innovations, and our method of cross-laboratory study design, are critical means of defining the overall state of neuroscientific knowledge as to cell identity. We first measured the replicability of neuronal identity by comparing more than 13,000 individual scRNA-Seq transcriptomes, then assessed cross-data-set evidence for novel pyramidal neuron and cortical interneuron subtypes identified by scRNA-Seq. We found that 24 of 45 cortical interneuron subtypes and 10 of 48 pyramidal neuron subtypes have evidence of replication in at least one other study, and we provide lists of candidate marker genes. Across tasks we found that any large set of variably expressed genes can identify equivalent cell types across data sets with high accuracy, indicating many of the transcriptional changes defining cell identity are pervasive and easily detected.

Epigenetic Meta-Analysis

Chromatin accessibility provides an important window into the regulation of gene expression. Recently, the assay of transposase accessible chromatin with sequencing (ATAC-seq) was developed to profile genome-wide chromatin accessibility. Although pipelines for implementation have been developed on an ad hoc basis for the analysis of individual data sets, there has been little comparative or aggregate evaluation. This is critical both to determine appropriate methodologies, controls, and efficacy, as well as to determine the global biological landscape of chromatin accessibility across diverse conditions. One major technical problem to address is that the counts of ATAC-seq reads underlying each peak vary substantially within a single sample and also between samples. Such variation makes the comparison to determine presence and absence of peaks (i.e., the open and closed state of chromatin) more difficult and less statistically well grounded. In this work we analyzed 197 ATAC-seq data samples from 13 studies to evaluate the robustness of results, as well as their specificity across studies. We find that peaks are promiscuously identified, with approximately 34,000 peaks per sample on average. These substantially overlap with transcription start sites (TSS), with 11,000 genes on average overlapping with the called peaks. Among those genes, 447 genes have peaks at their TSS from all 197 samples we analyzed. We evaluate the properties of these genes in detail, including mean expression across a diverse corpus of data, coexpression between this set and other genes, and functional enrichment. Finally, we suggest a novel approach to evaluate the robustness of peak signals and sensitivity by bootstrapping reads and recalling peaks for each resampling, whose calls are then aggregated. This yields peak calls that are highly robust to variation in noise as a source of peaks within data itself. We applied this novel data to our own ATAC-seq conditional experiment, with a reduction of likely spurious peaks to improve specificity.

PUBLICATIONS

Ballouz S, Dobin A, Gingeras TR, Gillis J. 2018. The fractured landscape of RNA-seq alignment: the default in our STARs. *Nucl Acids Res* **46:** 5125–5138. doi:10.1093/nar/gky325

Crow M, Gillis J. 2018. Co-expression in single-cell analysis: saving grace or original sin? *Trends Genet* **34:** 823–831. doi:10.1016/j.tig.2018.07.007

Crow M, Paul A, Ballouz S, Huang ZJ, Gillis J. 2018. Characterizing the replicability of cell types defined by single cell RNA-sequencing data using MetaNeighbor. *Nat Commun* **9:** 884. doi:10.1038/s41467-018-03282-0

Kalish BT, Cheadle L, Hrvatin S, Aurel Nagy M, Rivera S, Crow M, Gillis J, Kirchner R, Greenberg ME. 2018. Single-cell transcriptomics of the developing lateral geniculate nucleus reveals insights into circuit assembly and refinement. *Proc Natl Acad Sci* **115:** E1051–E1060. doi:10.1073/pnas.1717871115

In Press

Crow M, Gillis J. 2019. Single cell RNA-sequencing: replicability of cell types. *Curr Opin Neurobiol* **56:** 69–77. doi:10.1016/j.conb.2018.12.002

Crow M, Lim N, Ballouz S, Pavlidis P, Gillis J. 2019. Predictability of human differential gene expression. *Proc Natl Acad Sci* **116:** 6491–6500. doi:10.1073/pnas.1802973116

GENOME ORGANIZATION, REGULATION, AND FUNCTIONAL ROLES OF NONCODING RNAs

T.R. Gingeras C. Danyko G. Nechooshtan
C. Davis S. Patel
A. Dobin D. Yunusov
J. Drenkow

Summary of Contributions to Phase 3 of the ENCODE Project

C. Danyko, A. Dobin, J. Drenkow [in collaboration with C. Zaleski, A. Scavelli, and former members who worked on ENCODE Phase 3]

A total of nearly 6,000 new experiments (4,834 human, 1,158 mouse) were performed to define and annotate diverse classes of functional elements in the human and mouse genomes. Our contributions were in the area of developing a catalog of the types of long (>200-nt) and short (<200-nt) RNAs and identifying the transcript start sites (TSSs), splice sites, and potential *cis*-regulatory regions. Collectively, the ENCODE Encyclopedias encompass data from more than 1,000 different cell and tissue contexts, of which the vast majority are primary. Across multiple data types, the increase in the scale of experimental data provided new purchase on questions of genome organization and function and catalyzed new capabilities for deriving biological insights and principles. In addition to the accomplishments mentioned previously, we identified core gene sets and regulatory features underlying classical histopathological features, and we enhanced annotation of active but previously intractable genomic elements such as *Alu* repeats. Additionally, for this phase of ENCODE, we developed a new approach for profiling of 5′-capped RNA species called RNA annotation and mapping of promoters for the analysis of gene expression (RAMPAGE) (Batut et al., *Genome Res 23:* 169 [2013]) and extensively applied it to primary cells. Unlike RNA-Seq, RAMPAGE can (i) position TSSs with single-nucleotide resolution; (ii) generate accurate quantitative and reproducible measurements of promoter-specific RNA expression; and (iii) connect precise 5′ transcription initiation sites with specific alternative splice isoforms, thus providing a heretofore unavailable connection between promoter regulation and spliced products over long genomic intervals. RAMPAGE also enables annotation of previously intractable classes of RNA transcripts emanating from repetitive elements such as Alu regions.

ENCODE Tissue Expression (EN-TEx) Project

C. Danyko, A. Dobin, J. Drenkow

The main goals of the EN-TEx project have been to determine the genome sequences for four individuals (two male, two female) and to augment these DNA sequences (DNA-seq) with information concerning the transcriptional profiles (RNA-Seq and RAMPAGE), DNAse hypersensitive regions (DNAse-seq), chromatin modification profiles (ChIP-seq), and 3D chromosomal interacting maps (HiC-seq) donor. To accomplish these goals, approximately 20 tissues per donor were analyzed with each of these assays (except for DNA-seq and HiC-seq). These data have been collected to construct four highly detailed precision phased diploid genomes and to correlate the genetic variations found in each genome with the tissue transcriptional, epigenetic, and chromosomal 3D profiles.

Precision Genome Sequences: Multiple sequencing technologies and analysis software were employed for each genome. Short-read Illumina genome sequence data for each of the four donors was collected at a depth of >100× coverage. This approach was used to determine at high accuracy the single-nucleotide sequence variations (SNVs) and short insertion and deletions (indels). Long-read sequencing was carried out for the purpose of the detection of larger indels and to provide biallelic identification for each genome using PacBio (PB) and Oxford Nanopore Technologies (ONT). In addition, full chromosome

length phasing of the genomes and the identification of chromosomal interactions were determined using HiC-seq analysis and 10X analyses. Unique and common SNVs and structural variations for each genome were determined for each allele of each genome. An average of 3.9 million SNVs and 5.5 hundred thousand short SVs (<50 nt) were detected using Illumina sequencing and 23.7 thousand long SVs (50–10 kb) were detected by PB technology. The use of each sequencing technology revealed inherent biases. Only PB technologies could detect tandem expansion and contraction sequences, whereas the error rate of the PB in the detection of SNV was 100-fold that of the Illumina technology.

Reference versus Precision Mapping of the Transcriptional Landscape: Of the current approximately 19,000 and approximately 43,600 protein-coding and -noncoding genes, respectively, a total of approximately 15,000 and approximately 5,000 protein-coding and -noncoding genes were expressed (TPM 1–10) in at least one of the 80 tissues obtained from the four donors. Only a handful (25–50) of the expressed genes were seen to be expressed in either the reference (hg38) or the diploid precision genomes, indicating that there are a few errors in either genome sequence assemblies. The remainder of the 15,000–20,000 expressed genes were highly correlated in their expression levels when mapped to either the reference or diploid genomes. A total of 10,000–80,000 more reads were mapped to the diploid genome. Thus, as expected, more multiple and fewer unique reads mapped to diploid genomes, indicating that the differences in sequence between the versions of the genome were distributed across the genome rather than being collected in large regions. Extreme differences in expression after mapping the RNA-Seq results to the reference versus precision versions of the genomes (e.g., no expression mapped in mapping to the reference genome compared to detected expression in the precision genome and vice versa) for each of 60,000 coding and noncoding genes was observed for 728 genes. These results are likely to have been caused by errors in either the genome assemblies or the mappings of the RNA-Seq data.

Effects on Gene Expression When SVs Overlap Functional Elements of the Genome: Considering the cases of complete or partial (at least 1-bp) overlap with any 22 classes of functional elements, negative selection is observed for all classes except for ENSEMBL CTCF binding sites in all four donors, thus indicating that functional elements are not often the sites of SV location. The expression levels of 87 genes and 75 promoters are affected significantly in two of the donors when these genes and promoters are covered completely by an SV. Affected genes of this group with TPM of >5 show no decrease in expression. Similarly, affected promoters with TPM of >5 all appear not to have their expression affected. Analyses of these deleted genes and promoters indicate that these involve heterozygous deletions, leaving one functional allele. Because these are biallelically expressed genes in genomes with no SVs, these results suggest that compensatory mechanisms are contributing to maintaining the levels of expression for the heterozygous genes. The effects of SVs on epigenetic and methylation marks are under analysis.

MaizeCode: Construction of an Encyclopedia of DNA Elements of the Maize Genome

The overarching goal of the MaizeCode project is to produce empirical data sets for the identification of biochemically active and biologically functional elements encoded in the genome sequences of four maize lines (B73, NC350, W22, and Til11). A comprehensive catalog of these elements has been used to construct a critical link between the genotypes of each of the lines with phenotype in this classical plant system. The project began with the genomic sequencing of the NC350 and Til11 lines. These data have been used to construct a plant/maize genomics database with associated metadata, workflows, and analysis pipelines for the project and community. These are being disseminated using CyVerse, an NSF-funded platform and therefore accessible to all researchers to use in an unrestricted fashion, to process their data through and/or to reproduce our results and update them as desired. These data will serve as a basis for further annotation efforts in "wild-type" backgrounds and ideally as a centralized genomics living repository for additional maize high-throughput assays. Further curation of this repository will allow community-wide standardization for plant genomics data, ease their integration into analysis pipelines (e.g., to update older data to newer versions of assemblies), and/or expand upon current analyses to incorporate novel features and data types.

Our laboratory serves as the principal coordination center for five CSHL laboratories and laboratories from Johns Hopkins and NYU to achieve these goals. In addition, we have undertaken the transcriptional cataloging and genomic localization for five tissues from each of these maize as well as single-cell transcriptional analyses for the root of B73. Currently, a total of 141 of the proposed 175 (80.5%) assays have been completed. A total of 46 long RNA-Seq reactions each were conducted on fluorescence-activated cell sorting (FACS)-sorted root tip protoplast cells. The average number cells analyzed presample is 10^2–5×10^3, and 15 RNA-Seq reactions were completed on long RNAs on FACS-sorted ear protoplast cells and tissues. The average number of cells analyzed presample ranged from 10^2 to 10^4 of the cells. As part of the RNA-Seq pipeline, metadata including QC metrics and sample evaluations were collected based on the B73.v4 assembly and annotations. The v1.2 draft NC350 genome is available on the CyVerse and MaizeCode portals: http://www.maizecode.org/data/ and http://datacommons.cyverse.org/browse/iplant/home/shared/maizecode/NC350/assembly/hybrid/Zm-NC350-MAIZECODE-1.2.fasta.

Sorting and Processing of Various Types of RNAs in the Extracellular Environment

G. Nechooshtan, S. Patel, A. Prakash

Extracellular RNA processing agent

G. Nechooshtan

RNA is not only found within the boundaries of cells. We had turned our attention to the identification of a collection of RNAs that undergo site-specific processing in the extracellular milieu. Examples of these RNAs include tRNAs and YRNAs. One of our first goals with this project is to identify and to isolate the factors involved in this processing activity. Following proteomic and biochemical work, we have identified RNase1 as a possible candidate for this activity.

Previous work at our lab and several other labs has shown that the extracellular environment is rich in small noncoding RNAs. Notably, most of these molecules are processed from larger precursor RNAs. This research project aims to identify the mechanism leading to the specific processing of these RNAs. Following proteomic and biochemical work, we have identified RNase1 as a possible candidate for this activity, and decided to generate a modified cell line carrying an inactivated *RNASE1* gene using a CRISPR-Cas9 approach.

During 2018, we characterized the role of RNase1 in RNA processing in the cell, in extracellular vesicles, and in the nonvesicular extracellular RNA fraction, by RNA sequencing and further validation methods. Our findings indicate that although RNase1 has minimal effect on the processing of cellular and extracellular vesicle small RNAs, it has a crucial role in processing of nonvesicular extracellular small RNA. This effect is most noticeable for tRNAs and YRNAs. Our findings further indicate that mature tRNAs and YRNAs are released from cells as full-length precursors and are subsequently processed by RNase1 outside cells to yield tRNA halves and distinct YRNA fragments, similar to those identified in several biofluids.

tRNA halves have been shown to be processed from mature tRNAs inside cells by angiogenin in response to various stresses. However, they are found outside cells regardless of stress and our results show that RNase1, rather than angiogenin, is responsible for their formation in the extracellular environment.

Our findings suggest RNase1 as a strong candidate for formation of extracellular tRNA halves and distinct YRNA fragments in vivo and pave the way to biochemical studies of the role of these molecules.

Biological importance of extracellular vesicles and their RNA cargo

D. Yunusov

Extracellular vesicles (EVs) have recently gained attention as RNA and protein delivery agents that can be either intercepted for diagnostic purposes or artificially designed or modified for highly specific, targeted treatments. Despite ample diagnostic and therapeutic potential of EVs, a lot remains unknown about such important aspects of EV biology as the degree of selectivity of packaging of RNA and protein molecules into EVs and phenotypic changes elicited by those RNA and protein molecules in EV-receiving cells. The ongoing postdoctoral research project aims to understand the specific role of EV RNAs in the phenotypic changes in EV-receiving cells.

In vivo samples of EVs are collected from biological fluids. EVs in such samples originate from a large number of cells from various, usually highly heterogeneous tissues. For that reason, although permitting studies of their molecular cargoes, complex and heterogeneous populations of EVs from biological fluids are of little

use for dissection of phenotypic changes in the EV-receiving cells. At the same time, large sample volumes (liters) are required to isolate even minute quantities of EVs. For that reason and to reduce heterogeneity of collected populations of EVs, it is most convenient to isolate them from culture media conditioned by EV-producing cell line(s) in vitro. It is also very convenient to first study EV- and EV RNA-associated phenotypes if EV-receiving cells are represented by a homogeneous population of cells that are cultured in vitro under tightly controlled conditions.

The major challenge faced by any study designed in this way is in vivo validation of biological relevance of results obtained. In our current project, we took advantage of the in vitro system whose results can be easily extrapolated and tested in a potentially unlimited number of biological replicates in vivo. Our in vitro system represents the simplest in vivo situation—when an organism has only two interacting cell types. This in vitro system consists of EV-producing mouse embryonic stem cells (mESCs) and EV-receiving trophoblast stem cells (mTSCs). These cells represent the only two cell types that exist (and hence can interact with each other) in vivo in a mouse embryo on day 3.5 of development—inner cell mass (ICM) cells and trophectoderm cells, respectively.

The activities during the last year of the current project mostly involved the development of robust and cost-efficient protocols for EV and EV RNA isolation from mESCs and mTSCs, as well as of pipelines for data analysis. During this year, we calibrated and optimized the culture conditions for mESCs and mTSCs, established carefully controlled EV and EV RNA isolation protocols for mESCs (EV-producing cells) and mTSCs (control EV-producing cells), sequenced mESC EV RNA contents (results are currently being analyzed), and designed and are currently testing EV treatment and EV RNA transfection protocols and initial phenotype screening assays for mTSCs (EV-receiving cells). We also developed a pipeline to analyze short RNA sequencing data (short RNAs constitute the majority of EV RNAs) and long RNA sequencing data to test transcriptional changes in mTSCs in response to mESC EVs and EV RNAs.

The current goals of the project are (1) test the hypothesis that mESC EVs (possibly via EV RNAs) are capable of maintaining the undifferentiated state of mTSCs, (2) screen for alternative or parallel phenotypic responses of mTSCs to mESC EVs and EV RNAs, (3) identify specific RNA(s) responsible for such phenotypic changes, and (4) validate in vitro results through in vivo experiments in preimplantation mouse blastocysts.

EVs and EV small RNA (sRNA) as an initiator of epithelial-to-mesenchymal transition

An epithelial-to-mesenchymal transition (EMT) has been first described in embryonic development. EMT has been shown to be associated with other physiological and pathological processes such as fibrosis, wound healing, and cancer. EMT facilitates cell disassociation from the primary tumor and intravasation into blood and lymphatic circulations and leads to invasion of cancerous cells into surrounding or distant organs.

It is known that virtually all living cells secrete EVs into the extracellular space. EVs are endocytic or plasma membrane vesicles that include exosomes, microvesicles, and apoptotic bodies. EVs contain proteins, lipid, DNA, and RNA. The cargoes of EVs represent enriched members of the RNA landscape and are depleted of large members of the RNA landscape present within the cells of origin. EVs may be an important factor in determining whether their cargo plays a role in the development, spread, and prognosis of the cancer. This preliminary study evaluated whether EVs and EV sRNA have any role in EMT using normal human epithelial cell line MCF10A.

These human epithelial cells have been chosen because they recapitulate the developmental EMT characteristics. This model system has been instrumental in identification of several EMT factors that were necessary for mammary gland epithelial cell EMT and also for breast cancer cell invasion and metastasis formation.

In this preliminary study, EVs were isolated from two breast cancer cell lines and one leukemic cell line, K562. We used breast carcinoma MCF7 and metastatic MDA-MB-231 cell lines. The protein content of isolated EVs was measured using the BCA method. MCF10A cells were treated with 50 μg protein equivalent of EVs from all three cancer cell lines. MCF10A cells treated with or without TGF-β were used as the positive and negative control, respectively. The MCF10A cells were monitored, and images were taken daily for up to 72 h of treatment. After 72 h of treatment, EMT markers were measured using RT-PCR assay.

MDA-MB-231 EV (M31 EV)-treated MCF10A cells up-regulated mesenchymal marker FN1 approximately

50-fold, M7 approximately fivefold, and K5 approximately 15-fold compared to nontreated cells. There was a slight difference in NCAD level (mesenchymal marker) and ECAD level between all the cancer EV-treated MCF10A cells.

MDA-MB-231 EVs (M31 EV) induce morphological changes comparable to TGF-β-treated MCF10A cells, whereas M7 EV- and K5 EV-induced morphological changes were intermediate to those of TGF-β-treated cells. The combined set of data suggests that metastatic MDA-MB0231 cells' EVs might have a higher amount of EMT-inducing factor(s) compared to MCF7 cells' and K562 cells' EVs. Changes in RNA level of EMT marker further suggest that MCF10A cells are not fully transitioned to mesenchymal cells rather than that they show a partial/intermediate EMT type phenotype. This partial EMT transition phenotype could be attributed to the shorter duration of EV treatment (72 h compared to 12–14 d treatment), the single dose of EV treatment, and also the presence of the EMT inhibitor factor within EV cargoes. Overall, this study suggests that EVs have a role in EMT, and this model system could be used to identify EMT factors related to EVs.

PUBLICATIONS

Ballouz S, Dobin A, Gingeras TR, Gillis J. 2018. The fractured landscape of RNA-seq alignment: the fault in our STARs. *Nucleic Acids Res* **46:** 5125–5138.

Li K, Rodosthenous RS, Kashanchi F, Gingeras T, Gould SJ, Kuo LS, Kurre P, Lee H, Leonard JN, Liu H, et al. 2018. Advances, challenges, and opportunities in extracellular RNA biology: insights from the NIH exRNA Strategic Workshop. *JCI Insight* **3:** e98942. doi:10.1172/jci.98942

In Press

Zhang Q, Chao TC, Patil VS, Qin Y, Tiwari SK, Chiou J, Dobin A, Tsai CM, Li Z, Dang J, et al. 2019. The long noncoding RNA *ROCKI* regulates inflammatory gene expression. *EMBO J* doi:10.15252/embj.2018100041

GENOME ANALYSIS

W.R. McCombie Y. Berstein, B. Dong, E. Ghiban, S. Goodwin, S. Iyer, M. Kramer, J. Lihm, S. Mavruk Eskipehlivan, O. Mendivil Ramos, S. Muller, A. Qui, S. Teplin, R. Wappel

In 2018, we continued optimizing use of long-read sequencing platforms and applying that capability to both human genetics (primarily cancer) and plant genetics. We have additionally continued with several collaborations using more standard Illumina sequencing.

Our most significant advances for the year were the publication of the results of our study comparing long-read and short-read sequencing of the genome of the breast cancer cell line SK-BR-3 (this with Pacific Biosciences long-read technology) and the further optimization and use of the Oxford Nanopore PromethION long-read instrument. We then used the PromethION to sequence SK-BR-3 and several organoids derived from breast tumors. As the publication of the comparison paper shows, there are significant weaknesses in standard short-read sequencing of genomes. The short-read sequencing missed more than 10,000 structural variations in that one genome based on our analysis and comparison to long-read sequencing. These are only recently becoming well understood.

We are focusing on using the new long-read capabilities to explore the genomes of plants and animals: in plants, to understand genomes ranging from those of common food species to those of great evolutionary interest and, in animals, to analyze genomes of interest to understand cancer and aging.

CANCER GENETICS

Application of Long-Read Technologies to Probe the Genomic Architecture of Tumor Organoids

S. Goodwin, J. Lihm, M. Kramer, R. Wappel, W.R. McCombie [in collaboration with G. Arun and D. Spector, CSHL; K. Kostroff, Northwell Health; M. Schatz, R. Sherman, I. Lee, and W. Timp, Johns Hopkins University; F. Sedlazeck, Baylor College of Medicine]

Organoids are thought to effectively recapitulate tumor characteristics in a controlled, reproducible manner, leading them to become important tools for modeling cancer and personalized cancer treatment. By isolating tumor cells from the surrounding normal tissue, homogeneous cell cultures can be used for any array of assays that would not be possible from primary tumor tissue. The ability to generate a large amount of homogeneous material is particularly important for the study of large-scale structural variations. Long-read sequencing has been shown to be superior to short-read methods in detecting large-scale structural variation and filling gaps in difficult-to-sequence regions of genomes. However, long-read sequencing methods currently require a large amount of DNA—more than is available from a typical biopsy. Organoid culture provides sufficient material to carry out these analyses.

In 2018, we completed high-depth sequencing of two breast tumor–derived organoids and the SKBR3 breast cancer model cell line. Lower coverage sequencing was also carried out on a matched normal tissue–derived organoid; however, poor growth performance limited the available material. A second matched normal organoid underwent sparse sequencing for copy number variation (CNV) profiling, and we determined it was not truly normal tissue. Sequencing was carried out via PacBio, Oxford Nanopore, and 10X Genomics methods. The yield/coverage for each method is found in Table 1.

In 2019, additional sequencing work will be carried out on tumor and SKBR3 samples so that each sample will have the following data sets: PacBio, Oxford

Table 1. Data types available for each sample

Sample	PacBio coverage	ONT coverage	10X coverage	Illumina MethylSeq (Gb)
51T	55	58	12	5.6
48T	63	58	NA	
SKBR3	75	73	12	5.6
51N	NA	12	36	NA

(NA) Not available.

Nanopore, 10X Genomics, and Illumina MethylSeq. These data will be compared to determine the optimal combination of data sets to achieve high coverage, high assembly contiguity, high (true) structural variation (SV) detection rate, and low cost. Additionally, the Oxford Nanopore methylation calling will be compared to the Illumina-based methylation calling to assess its performance and to explore the relationship between global methylation patterns and SVs. An additional two colon-derived organoids, one primary tumor and one matched metastasis, will be sequenced. These will be provided by Dr. Semir Beyaz (CSHL).

Quality Assurance of Organoid Cancer Models

W.R. McCombie, S. Goodwin, M. Kramer, S. Muller [in collaboration with D. Tuveson, D. Spector, P. Sridevi, D. Plenker, G. Arun, and S. Bhatia, CSHL; V. Corbo, University of Verona]

Organoids derived from tumor tissue are an important resource for modeling cancer treatment response. However, generation of these models can be challenging and can be compromised by the growth of normal cells in culture. We previously described our analysis pipeline to identify driver mutations in a capture panel of 140 genes known to be involved in cancer. In 2018, we continued to analyze organoid models generated by our collaborators using an updated pipeline for assessment of cancer drivers. To identify organoids derived from tumor types that did not show clear drivers in our 140-gene panel capture, we have now incorporated low-coverage (~0.5×) whole-genome sequencing to identify CNVs indicative of the tumor compared to the matched blood. Our CNV pipeline comprises running both CNVNator and Gingko and filtering events found in the matched blood. Then we identify variant calls found by both tools to increase confidence. In general, this low-coverage "skim" approach allows us to view gross changes across chromosomes.

We have now processed 229 panel captures (including the models and matched blood samples) and have assessed 62 samples for CNVs. The models have been derived from an array of cancers including colon, pancreas, lung, liver, and breast. Thus far, 104 organoids have shown clear drivers, and several additional samples have been classified as provisional pending further review. In particular, we have noted difficulty in identifying drivers for the breast organoid samples, although we are considering additional sequencing of either the whole exome or an expanded set of genes to improve classification of these models.

Genomic Analysis of the Ethnic Disparity in Colon Cancer

M. Kramer, S. Goodwin, R. Wappel, S. Muller, W.R. McCombie [in collaboration with X. Wang, X. Yu, E. Li, and J. Williams, Stony Brook University; M. Salifu, L. Martello-Rooney, S. Vignesh, and M. Stewart, SUNY Downstate]

Recent studies have shown genetic differences in African–American (AA) colon cancer drivers compared to those in Caucasian–Americans (CAs). This is critical information, as colon cancer is more likely to lead to poor prognosis and death when diagnosed in AAs. We continued our collaboration with Stony Brook and Downstate to generate linked genomic, transcriptomic, and epigenetic data sets for the study of colon cancer in minorities, with a slight shift of focus in 2018 to increase sample number for the RNA-Seq set. Initial RNA-Seq analysis of a small sample indicated possible enrichment of immune signatures in AA samples, which will be explored in the larger data set in 2019. Also, for the reduced representation bisulfite sequencing (RRBS), we noted that CpG targeting was reduced in tumor samples compared to adjacent normal tissue. Higher coverage would be necessary for further analysis; therefore, the methylation analysis strategy was switched to microarray profiling to address cost limitations. The samples completed for this pilot are outlined in Table 2.

Long-Read Sequencing of Bats as a Model for Cancer Resistance Mechanisms

O. Mendivil Ramos, M. Kramer, S. Goodwin, R. Wappel, W.R. McCombie [in collaboration with N. Simmons and S. Oppenheim, AMNH; M. Schatz, Johns Hopkins University]

Not all animals age at the same rate; neither do they get cancer at the same rate. Interestingly, mammalian species have evolved a remarkable diversity of aging rates spanning from 2 to 211 years. This fact implies that there is an extraordinary diversity in mechanisms of aging, which include cancer-resistance mechanisms.

Within the mammalian species, the bat clade is prominent because of their long life spans given their small body size. Thus, the bat clade holds promise in comparative studies for studying aging and cancer-resistance mechanisms.

In collaboration with the AMNH, we have sequenced two new bat species, *Artibeus jamaicensis* and

Table 2. Sequencing of tumor-normal pairs at Cold Spring Harbor Laboratory

		Before 3/1/18		After 3/1/18		Total	
		non-AA	AA	non-AA	AA	non-AA	AA
Stony Brook	Matched tumor-normal pair exomes sequenced at CSHL	6	7	2	0	8	7
	Matched tumor-normal pair RNA-Seq samples at CSHL	4	9	22	0	26	9
Downstate	Matched tumor-normal pair exomes sequenced at CSHL	0	3	0	2	0	5
	Matched tumor-normal pair RNA-Seq samples at CSHL	0	3	0	18	0	21
Matched tumor-normal RRBS samples sequenced		6	7	0	0	6	7
Total exomes, both sites		**6**	**10**	**2**	**2**	**8**	**12**
Total RNA-Seq, both sites		**4**	**12**	**22**	**18**	**26**	**30**

Pterotus mesoamericanus, which are above the median age in years of all species of bats noted so far and have noncarnivorous diets in comparison with other bats. We are using Oxford Nanopore Technologies (ONT) long-read sequencing along with short-read sequencing to provide long-read error correction. We have achieved a coverage of 107× ONT and ~30× Illumina for *Artibeus jamaicensis* and 56× ONT and ~30× Illumina for *Pterotus mesoamericanus.* In collaboration with JHU, we are in the process of obtaining a final assembly, and in parallel we are constructing a pipeline for comparative genomics to analyze the assembly.

In addition to a high-quality assembly, we anticipate an important improvement in genome annotation and detection of large- and small-scale genome variation that could provide clues to positively selected genes or gene families involved in cancer resistance, which may not have been observed before.

PLANT GENOME ANALYSIS

Exploring the Wheat Epigenome

W.R. McCombie, S. Goodwin, E. Ghiban [in collaboration with R. Martienssen, M. Regulski, J. Han, C. LeBlanc, M. Donoghue, CSHL; L. Gardiner, Earlham Institute; A. Hall and J. Kenny, University of Liverpool; K. Mayer and M. Pfeifer, Helmholtz-Zentrum Muenchen; M. Bevan, John Innes Centre Institute]

Elucidation of the regulatory mechanisms of wheat will create a valuable community resource and provide genetic clues that may allow growth of wheat in new habitats, as well as improving crop quality. In 2018, we continued our collaboration as part of the ERA-CAPS project to sequence both long and short RNAs from multiple elite wheat varieties including CS42 and Paragon. We previously reported on the high-coverage bisulfite sequencing of the Paragon strain. The addition of expression data will allow us to correlate the regulatory effects of the epigenetic marks that were found. We have begun library preparation and sequencing for both long- and short-RNA profiling of multiple lines, including stable hexaploids, triploids, and parents, as well as tissue-specific assays of CS42, which will be analyzed by collaborators at the Earlham Institute. We plan to complete the additional transcriptome sequencing and analysis and refine the integration of epigenetic and expression data in 2019.

Toward an Encyclopedia of Maize Genomics with the MaizeCODE Project

W.R McCombie, S. Goodwin, M. Kramer, E. Ghiban [in collaboration with T. Gingeras, C. Danyko, D. Jackson, R. Martienssen, M. Regulski, D. Micklos, M. Schatz, D. Ware, CSHL; Ken Birnbaum, NYU; Hank Bass, Florida State]

Maize is an important crop species with a complex genomic structure. Although many varieties have been sequenced using short-read technologies, the application of long-read sequencing can provide insight into key structural variants and repetitive elements in these genomes. Our previous work with NC350 combined ~55× coverage of Pacific Biosciences long-read data with 10× Chromium linked short reads to produce a high-quality assembly with a scaffold N50 of >1 Mb. In 2018, we have generated an additional >50× coverage of NC350 with long reads from the Oxford Nanopore platform and have also obtained HiC data from Phase Genomics to further refine the assembly into chromosomes. Furthermore, we generated >50× Oxford Nanopore long-read coverage of the teosinte (Til11) maize precursor. Continuing our work with the Gingeras lab, we have sequenced transcriptome libraries for long-, RAMPAGE, and short RNA-Seq of B73, W22, Til 11, and FACS-sorted ears and submitted them to the NCBI SRA. In 2019, our collaborators in the Schatz lab will finalize the combined multiplatform assembly of NC350 and begin assembly of

teosinte, and the group will finalize the RNA-Seq and ChIP experiments for all species.

PROMETHION OPTIMIZATION AND SEQUENCING

Sequencing and Assembly of Tomato Genomes with Oxford Nanopore Technology

S. Goodwin, R. Wappel, W.R. McCombie [in collaboration with S. Soyk, X. Wang, Z. Lemmon, Z. Lippman, CSHL; F. Sedlazeck, Baylor College of Medicine; M. Alonge and M. Schatz, Johns Hopkins University]

In 2018 the McCombie lab, in collaboration with the next-generation sequencing core, received an Oxford Nanopore PromethION device. This device is capable of running 24 PromethION cells in parallel—with the expectation that it will be able to run 48 in parallel in the future. Each cell has 3000 sequencing pores and can generate many gigabytes of data per run. This instrument was acquired to support our ongoing collaboration with Northwell Health to develop improved methods of long-read sequence of organoids. However, there has been substantial interest in the instrument from many groups both within and outside CSHL. In 2018, one of the primary users of the instrument was Dr. Zach Lippman. His project aims to carry out whole-genome sequencing (WGS) on 100 tomato varieties to determine the structural and evolutionary driving factors driving important phenotypes. In 2018, >3 Tb of data from 45 cells was generated for the tomato project. This large number of cells has given us the opportunity to optimize and develop pipelines for improved yield and read lengths. Over this time, mean yield increased 14% and total DNA input decreased 90%. Average read length remained constant at ~8.5 kb. Other projects that have made use of the technology are WGS sequencing of two bat species, organoid sequencing, human WGS for the Entex project from Dr. Thomas Gingeras, maize and teosinte as part of the MaizeCODE project (also with Dr. Gingeras), and a grass project in collaboration with Dr. Adam Siepel. Although none of these other projects have had the depth of work dedicated to them like the tomato project, they have nevertheless provided us with an opportunity to validate our pipelines on varied samples. One of our main findings is that there is no one method that works for all samples. Care must be taken to ensure the highest sample quality. When possible, samples should be sequenced very soon after extraction without an intermediary freeze step. If freezing is unavoidable, fragmentation parameters may need to be optimized. To achieve optimal fragmentation, pipelines employing the Diagenode megaruptor (which can consistently fragment DNA up to 75 kb) and the Agilent Femtopulse (which can accurately size DNA up to 165 kb) have been developed. This involves a preliminary QC step on the Femtopulse instrument to assess the length of the unfragmented DNA. Based on these results and the anticipated condition of the DNA, an aliquot sample is fragmented to achieve an average fragment length between 30 and 50 kb. We have found that although larger read lengths are possible, this range gives better total yield while maintaining sufficient read length. After fragmentation the sample is again run with the Femtopulse to ensure proper sizing. This process is repeated if the sample is not properly sized.

In 2019, we anticipate significant interest in the use of this device. Our short-term goal is to optimize the instrument such that every WGS cell from all samples yields at least 80 b of data with a read length of N50 = 15–20 kb. We anticipate increasing this goal over the course of 2019 to ensure we are achieving the highest-quality nanopore sequencing. One problem we have found is the substantial amount of raw data generated per cell—up to 2 Tb in many cases. If we continue at the current rate of sequencing (around 10 cells a month) we will generate 240 Tb of data in 2019 or ~25% of the total storage available to all McCombie users, the NGS core, and our collaborators. We anticipate sequencing up to 30 cells a month. This level of data generation is unsustainable and will require a better storage solution to be implemented in early 2019.

PUBLICATIONS

Nattestad M, Goodwin S, Ng K, Baslan T, Sedlazeck FJ, Rescheneder P, Garvin T, Fang H, Gurtowski J, Hutton E, et al. 2018. Complex rearrangements and oncogene amplifications revealed by long-read DNA and RNA sequencing of a breast cancer cell line. *Genome Res* **28:** 1126–1135. doi:10.1101/gr.231100.117

Ryan NM, Lihm J, Kramer M, McCarthy S, Morris SW, Arnau-Soler A, Davies G, Duff B, Ghiban E, Hayward C, et al. 2018. DNA sequence-level analyses reveal potential phenotypic modifiers in a large family with psychiatric disorders. *Mol Psychiatry* **23:** 2254–2265. doi:10.1038/s41380-018-0087-4

Tamura K, Yu J, Hata T, Suenaga M, Shindo K, MacGregor-Das A, Borges M, Wolfgang CL, Weiss MJ, He J, et al. 2018. Mutations in the pancreatic secretory enzymes CPA1 and CPB1 are associated with pancreatic cancer. *Proc Natl Acad Sci* **115:** 4767–4772. doi:10.1073/pnas.1720588115

PLANT GENOME ORGANIZATION, EVOLUTION, AND FUNCTION

D. Ware K. Chougule, N. Gladman, F. Hu, Y. Jiao, V. Kumar, S. Kumari, Y.K. Lee, Z. Lu, I. Meijs, D. Muna, A. Olson, M. Regulski, J. Stein, M.K. Tello-Ruiz, J. Thomason, P. Van Buren, B. Wang, L. Wang, X. Wang, S. Wei, L. Zhang

The Ware lab has two primary goals: (1) understanding plant genome function in agriculturally important crop plants; and (2) development of tools, data sources, and resources for the genomics research community.

PLANT GENOME RESEARCH

In the last decade, the sequencing and annotation of complete plant genomes has helped us understand plant function and evolution, as well as how to alter economically important traits. Efforts in many disparate disciplines are required to generate reference genomes. The work at the Ware lab often starts with experimental biologists who generate the raw sequence data. Next, computational biologists and bioinformaticians kick off a series of computational steps to interpret the raw data. The process of interpretation involves the assembly of raw sequence reads into overlapping segments (contigs), which are combined to create a scaffold. This scaffold, in turn, discerns the position, relative order, and orientation of contigs within the chromosomes. The next step is annotation, the discovery and description of genes and other functional elements, and homologies (evolutionary relationships) with other genomes. This information must be faithfully communicated and visualized in web-based platforms such as Gramene.

All of these activities are rapidly evolving with the advent of new sequencing technologies, algorithms, and data-handling requirements. For example, high-depth, low-cost sequencing of RNA transcripts is providing a vast stream of new evidence that informs genome annotation; this, in turn, has spurred the development of new software for modeling and performing genome annotation. Low-cost sequencing has also made it possible to ask whole new classes of questions, moving beyond the generation of single references for individual species and supporting the development of multispecies representation as a pangenome. Ongoing projects within the maize, rice, *Arabidopsis*, sorghum, and grape research communities are now sequencing hundreds or thousands of genotypic backgrounds, chosen from carefully constructed populations, wild populations, and breeding germplasms in each species. Information about genetic variation is helping scientists to understand the genetic basis of phenotypic traits and the origins of domesticated crop species. New technologies are also driving research in epigenetics, the study of heritable variation not attributable to changes in the underlying genome sequence. Epigenetic mechanisms include modification of DNA by methylation and various forms of histone modification, which can cause changes in gene expression and other phenotypes. Both types of modifications can be studied using new sequencing technologies and analytical methods.

Updating the Maize B73 Reference Genome and Annotation Using Long Sequence Read Technologies

Y. Jiao, B. Wang, M. Regulski, A. Olson, J. Stein, S. Kumari [in collaboration with The Maize Genome Sequencing Consortium; R. McCombie and S. Goodwin, CSHL; K. Guill and M. McMullen, Missouri University, USDA ARS; J. Ross-Ibarra, UC Davis; K. Dawe, University of Georgia; D. Rank, P. Peluso, and T. Clark, Pacific Biosciences; A. Hastie, BioNano Genomics]

Maize is an important feed and fuel crop, as well as a model system in developmental genetics. A complete and accurate reference genome is imperative for sustained progress in understanding the genetic basis of trait variation and crop improvement. The Ware lab

has played a leading role in the development and stewardship of the maize reference genome for more than a decade. The 2009 release of the B73 reference sequence was a milestone in plant genomics research because of the unprecedented size and complexity of the maize genome. Through several updates, this foundational resource has remained the principal genome reference for the maize research community. Yet, it continues to be a work in progress, with gaps and misassemblies that have defied available sequencing technologies, especially over the highly repetitive regions that are also the most dynamic and rapidly evolving.

This year marks a new milestone with the release of an entirely new reference assembly, designated B73 RefGenV4. In collaboration with Pacific Biosciences, we sequenced maize to ~65× coverage using single-molecule real-time (SMRT) technology. These exceptionally long reads, typically >10 kb, were able to span not only genes but also the extensive intergenic and repetitive regions rich in transposons. Using an optimized correction and assembly pipeline, we built a de novo assembly consisting of 2,958 contigs and totaling 2.10 Gb, with more than half of this length in contigs of >1.2 Mb. This represents a 52-fold improvement in assembly contiguity compared with the previous bacterial artificial chromosome (BAC)-based maize assembly, whereas the nucleotide agreement was maintained at >99.9%. Optical maps of the maize genome, developed at BioNano Genomics, enabled the scaffolding of these contigs into 10 chromosomes, thereby placing >99% of the maize genes. Many improvements in gene order and orientation were found. The new sequence fills in many intergenic regions that harbor transposons and regulatory regions influencing gene expression. In addition, the comparative optical mapping of two other inbreds revealed a prevalence of deletions in the low gene density region and maize lineage-specific genes (Jiao et al., *Nature 456:* 524 [2017]).

Uncertainties about the complete structure of mRNA transcripts, particularly with respect to alternatively spliced isoforms, can be a limiting factor for research in the system. In addition to working on the reference genome sequence, we are using the same single-molecule sequencing technology to investigate the maize transcriptome. For this work, we have sampled full-length cDNAs from six tissues of the maize inbred line B73. These were barcoded, pooled, size-fractionated (<1 kb, 1–2 kb, 2–3 kb, 3–5 kb, 4–6 kb, and 5–10 kb), and sequenced on the PacBio RS II platform with P6-C4 chemistry. We were able to capture 111,151 transcripts, representing ~70% of the genes annotated in the maize RefGenV3 genome assembly. A large proportion of transcripts (57%) are novel. We were able to validate ~90% of the transcript splice-site junctions within high-depth short reads generated from the matched tissues. In addition, we identified a large number of novel long noncoding RNAs (lncRNAs) and fusion transcripts. Our results show that the characterization of the maize B73 transcriptome is far from complete, and maize gene expression is more complex than previously thought (Wang et al., *Nat Commun 7:* 11708 [2016]).

A Comparative Transcriptional Landscape of Maize and Sorghum Obtained by Single-Molecule Sequencing

B. Wang, M. Regulski, A. Olson, D. Ware [in collaboration with R. McCombie and S. Goodwin, CSHL; E. Tseng, Pacific Biosciences]

Maize and sorghum are both important crops with similar overall plant architectures, but they have key differences, especially in regard to their inflorescences. To better understand these two organisms at the molecular level, we compared the expression profiles of both protein-coding and -noncoding transcripts in 11 matched tissues using single-molecule, long-read, deep RNA sequencing. This comparative analysis revealed large numbers of novel isoforms in both species (Fig. 1). Evolutionarily young genes were likely to be generated in reproductive tissues and usually had fewer isoforms than old genes. We also observed similarities and differences in alternative splicing patterns and activities—both among tissues and between species. The maize subgenomes A and B exhibited no bias in isoform generation; however, genes in the B genome were more highly expressed in pollen tissue, whereas genes in the A genome were more highly expressed in endosperm. We also identified a number of splicing events conserved between maize and sorghum. In addition, we generated comprehensive and high-resolution maps of poly(A) sites, revealing similarities and differences in mRNA cleavage between the two species. Overall, our results reveal considerable splicing and expression diversity between sorghum and maize, well beyond what was reported in previous studies, likely

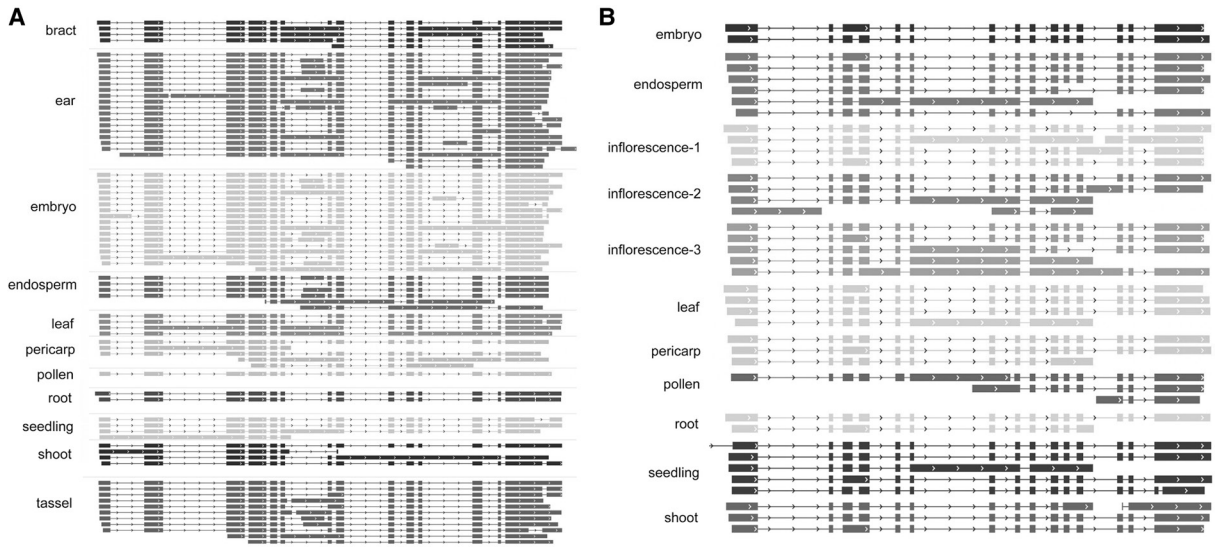


Figure 1. Different isoforms in maize and sorghum. (*A*) Different isoforms among different tissues in maize. (*B*) Different isoforms among different tissues in sorghum.

reflecting the differences in architecture between these two species (Wang et al. 2018a).

Gramene: Comparative Genomic Resource for Plants

K. Chougule, Y. Jiao, V. Kumar, S. Kumari, Y.K. Lee, A. Olson, J. Stein, M.K. Tello-Ruiz, B. Wang, S. Wei, L. Zhang
[in collaboration with P. Jaiswal, Oregon State University; P. Kersey, I. Papatheodorou, and R. Petryszak, EMBL-European Bioinformatics Institute; L. Stein, Ontario Institute of Cancer Research; C. Taylor, American Society of Plant Biologists; R. Wing, University of Arizona]

The Gramene project provides online reference resources for plant genomes and curated pathways to aid functional genomics research in crops and model plant species. Our website (www.gramene.org) facilitates studies of gene function by combining genome and pathway annotation with experimental data and cross-species comparisons. In other words, the data and tools in Gramene enable plant researchers to use knowledge about gene function in one species to predict gene function in other species. Drawing these connections facilitates translational research in plant development and physiology that influences economically important traits—for example, grain development, flowering time, drought tolerance, and resistance to diseases. In the past year, the project accomplished several major milestones, culminating in our 59th data release (October 2018) since the inception of the project, with a significantly enhanced search user interface and back-end functions. The Gramene team develops comparative genomics databases in collaboration with the Ensembl Plants project at the European Bioinformatics Institute (EMBL-EBI) and collaborates closely with the EBI's Expression Atlas project to provide manually curated, quality-controlled, and analyzed transcriptomic data. We continue to host genome and pathway annotations via (1) the Ensembl genome browser and (2) the Plant Reactome pathways portal.

We consolidated our dedicated maize genome browser resource (maize4.gramene.org) into the prototype of a maize pangenome browser. Although Gramene's main site (www.gramene.org) is committed to supporting community-recognized annotations for a single reference genome per species, this resource features three maize assemblies: B73 *Zea mays* RefGenV4, *Z. mays* W22, and *Z. mays* PHP and the ancestral teosinte TIL11 assembly. To facilitate comparative and functional genomics research in maize, this resource also hosts a subset of key species, including the dicot model *Arabidopsis thaliana* and the monocot model *Oryza sativa Japonica*. Featured annotations will be used to build phylogenetic gene trees and define orthologous and paralogous relationships using the Ensembl Compara gene tree method. These results will yield new insights into the taxonomic origin of genes and patterns of duplication,

movement, and loss influenced by genome architecture. The maize V4 assembly and annotations are also available from the main Gramene genome browser.

Pathogen Immunity Genes in Wild Related Species of Rice

J. Stein, K. Chougule, S. Wei [in collaboration with R. Wing, University of Arizona; The International Oryza Map (I-OMAP) Consortium]

Disease pathogens, such as rice blast (*Magnaporthe oryzae*), severely impact rice production and pose an increasing threat as climate change alters the geographical range of pests in the future. Breeding for natural host resistance is a proven strategy but is limited by sources of variation. Wild relatives of rice, collected from around the world, provide a reservoir of resistance genes that can be transferred to cultivated rice by introgressive breeding. Taking advantage of a 13-species set of I-OMAP consortium reference genomes in OGE (Oryza Genome Evolution) Gramene (www.oge.gramene.org) that spans the Oryzeae tribe (including wild species in the *Oryza* and *Leersia* genera), we discovered more than 5,400 NLR (nucleotide-binding domain and leucine-rich repeat-containing) genes in 28 families (Stein et al. 2018). The rapid diversification of complex haplotypes by gene expansion and loss is typical of NLR genes, contributing to disease adaptation. Applying phylogenetic reconciliation methods to gene trees in these 28 families, we found a 10-fold increase in duplication rates in lineages leading to both Asian and African cultivated rice, consistent with selection for resistance traits before domestication. Most NLR genes were positionally clustered, often forming complex arrangements of distantly related genes (Fig. 2). Yet, clear orthologous relationships and evidence of conserved underlying haplotype structures could be drawn, even in the most distantly related (~17 MY) species of *Leersia*. In rice and other plants, disease resistance is sometimes conferred by the action of two neighboring NLR genes whose products function as heterodimers. Examining all possible combinations of adjacent NLR pairs, we found that those reminiscent of functionally coupled NLR genes and, therefore, belonging to heterogeneous families and configured in a head-to-head (divergently transcribed) arrangement are significantly more prevalent than expected by chance and are also more likely to be conserved across species than other arrangements. This finding suggests that the evolution of haplotype structure may be constrained by underlying regulatory and functional interactions

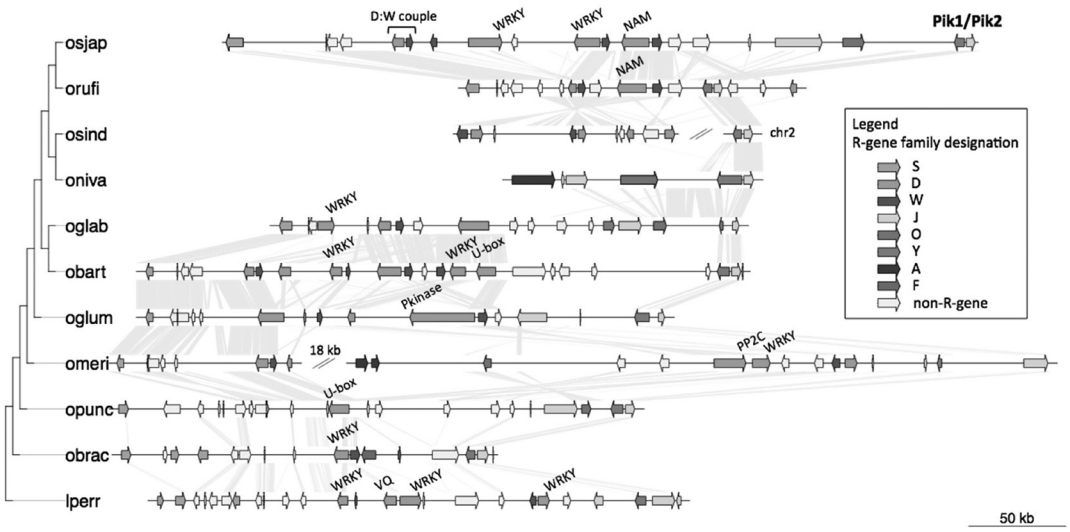


Figure 2. Complex evolution of an NLR gene cluster in the *Oryza* genus adjacent to the *Pik1*/*Pik2* disease resistance locus, which confers resistance to rice blast fungus (*Magnaporthe grisea*). Color coding of NLR gene families reveals underlying conservation of haplotype structure and prevalence of putatively functionally coupled heterogeneous gene pairs. Diversification of pathogen recognition specificity is hypothesized to be conferred by noncanonical domains, such as WRKY, which function as integrated decoy domains. (Figure adapted from Stein et al. 2018.)

between such putative coupled NLR genes. Furthermore, we found a greater prevalence of putative integrated decoy domains among such pairs, which are thought to function in pathogen recognition by mimicking host targets of pathogenicity factors. Striking variation in domain structure suggests that swapping of various decoy domains contributes to the evolution of haplotype diversity and resistance specificity. This study has opened a treasure trove of potentially novel resistance functions that may help in the future development and sustainability of rice.

PLANT GENETICS AND SYSTEMS BIOLOGY

The global challenges confronting agricultural security are falling into sharper relief: declining water for irrigation; surging pest pressures due to longer and hotter growing seasons; degrading arable land; increasing population; and long-term geographical adjustments brought about through climate change. Overcoming these strata of obstacles necessitates nimble and reliable approaches. Predictive genetics of desirable traits in concert with rapid germplasm conversion has become the norm since high-throughput sequencing has reached cost-effectiveness and genome editing and transformation techniques continually improve; pangenomes and new crop genomes are available for molecular investigation and comparative genomics, strengthening and accelerating the output of researchers and hastening the fruits of such labor into the hands of producers.

Although the predictive genomic paradigms are still being optimized, they are showing improved reliability depending on the desired trait. However, there are limitations when sought-after changes are focused around heretofore uncharacterized molecular mechanisms and gene regulatory networks acting in narrow spatiotemporal windows. It is crucial for plant science investigators to continue the molecular dissection of pathways controlling beneficial agronomic traits like flower fertility, inflorescence architecture, root formation, microbiome interaction, and nutrient use efficiency. These research areas have noted quantitative improvements in crop yield and sustainability. Additionally, characterization across numerous plant species can yield a more unified systems biology model that can be effectively applied to numerous agricultural challenges.

Dissection of Gene Regulatory Networks Associated with Nitrogen Use Efficiency

L. Zhang, A. Olson, V. Kumar, S. Kumari, K. Chougule [in collaboration with A-M. Bagman, A. Gaudinier, and S. Brady, UC Davis; M. Frank and B. Shen, DuPont-Pioneer, Inc.]

Nitrogen (N) is an essential micronutrient for plants. Maximizing nitrogen use efficiency (NUE) in plants is critical to the increase in crop production and reduction of negative impacts on the environment due to seasonal excess of runoff fertilizer. To explore the gene regulatory network that controls these processes, we have used a gene-centered approach to characterize transcription factors that regulate genes involved in nitrogen uptake, transport, and metabolism in the model plant *Arabidopsis*. The outcome of the screen was the identification of 23 novel transcription factors (Gaudinier et al. 2018). To prioritize which of these genes are more likely to have an impact on NUE and in turn impact biomass, root development, and time to flower, we developed NECorr (Liseron-Monfils et al. 2018), a model that combines network topology and expression data to rank genes and their interactions in a given tissue or condition. Using the prioritized gene list, we identified 26 loss-of-function mutants; amazingly, 25 of the genes perturbed showed a root or flowering time phenotype. With the demonstration of efficacy in the model, we are now extending it to agronomically important species such as maize and sorghum.

Developmental Networks Controlling Inflorescence Architecture and Grain Yield in Grasses

Y. Jiao, Y.K. Lee, S. Kumari, N. Gladman [in collaboration with Z. Xin and J. Burke, USDA ARS]

The objective of this work is to integrate genetics and genomics data sets to find molecular networks that influence a variety of agronomic traits to improve the morphology (architecture) of grass inflorescences (flowers), with a specific focus on *Zea mays* and more recently *Sorghum bicolor* [L.] Moench—a top five global crop in terms of dedicated acreage. Because inflorescences bear the fruits and grains that we eat, either through direct consumption or via animal feed, the genetic and regulatory factors that govern their

development are obviously relevant to important agronomic traits such as grain yield and harvesting capability. Sorghum has recently emerged as a potentially potent bioenergy crop in addition to its role for human consumption in sub-Saharan Africa and other global regions. Because of its drought tolerance and tendency to be grown on marginal land, sorghum can also be used as a predictive genomics model for identifying water- and nutrient-use efficiency gene candidates that could be implemented in other broadleaf crops.

The number of grains per panicle is a developmental trait contributing to overall sorghum yield. Sorghum flowers comprise one fertile (sessile) and two sterile (pedicellate) spikelets (Fig. 3). Only the sessile spikelet is fertile and will produce seed. Using a publicly available EMS (ethyl methanesulfonate) population, we identified independent multiseeded (*msd*) mutants that manifest both fertile sessile and pedicellate spikelets throughout the inflorescence. A detailed dissection of developmental stages of wild

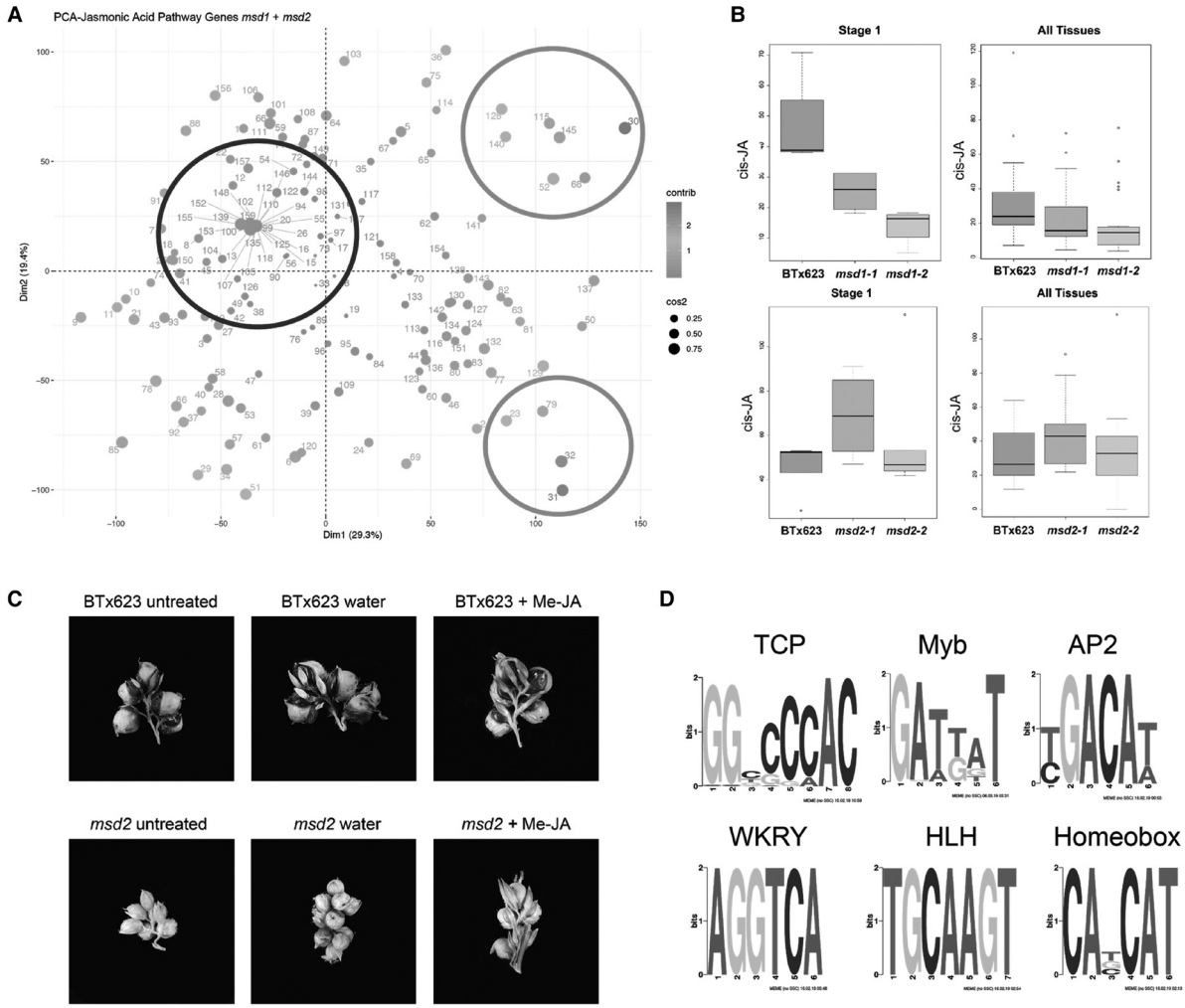


Figure 3. (*A*) PCA analysis of *msd1* and *msd2* transcriptomic profiles of jasmonic acid (JA) pathway genes. Red circles highlighting variable genes are enriched for signaling proteins (JAZ), whereas the blue circle is enriched for final steps in JA biosynthesis. (*B*) Lipidomics profile *cis*-JA in *msd1* and *msd2* developing panicles (stage 1 and cumulative levels in all tissues). (*C*) Phenotypic rescue of floral fertility with exogenous application of JA during meristem development. (*D*) Enriched transcription factor binding motifs in MSD1 DAP-seq peaks. Only the TCP motif is identified in peaks surrounding the transcriptional start site. Additional developmental and environmental response motifs were identified when enrichment analysis was applied to all peaks, which includes putative MSD1-targeted enhancer regions.

type and *msd1* revealed that pedicellate spikelets in wild type do not have floral organs, including ovary, stigma, filament, or anther, whereas the *msd1* mutants generate intact floral organ in both sessile and pedicellate spikelet. Using a bulk segregant analysis of F_2 individuals, we determined that the *msd1* mutations are located within a TCP transcriptional factor. The six causal SNPs found in *msd1* are highly conserved across grass species. The *TCP* gene was found to be differentially expressed during inflorescence development within a narrow spatiotemporal region. To characterize the gene networks associated with pedicellate spikelet fertility, we generated whole-genome expression profiling data of floral tissues at four different inflorescence development stages in both wild type and *msd1*. Preliminary analyses suggested *MSD1* may impact programmed cell death signaling in pedicellate spikelets in wild type by modulating hormone pathways. To unbiasedly probe for MSD1 regulation targets, DNA affinity purification sequencing (DAP-seq) was performed using bacterially expressed GST-MSD1 proteins. The successive peak-calling analysis determined that MSD1 binds near the transcriptional start site of other putatively regulatory genes, including other developmental transcription factors, signaling cascade proteins, and ribosomal proteins; these targets have been confirmed to be acted on by TCPs in other plant species. Curiously, the DAP-seq analysis showed that there is no direct regulation of JA biosynthetic pathway genes by MSD1, indicating more complex methods of regulatory network alteration in the mutants through different downstream effectors and feedback loops.

CYBER INFRASTRUCTURE PROJECTS

The success of the Human Genome Project, completed at the start of the twenty-first century, inspired numerous large-scale sequencing projects such as the 1000 Genomes Project to catalog human genetic variations, 1000 Plant Genomes to sequence expressed genes of 1,000 different plant species, 1001 Genomes to discover whole-genome sequence variations in more than 1,001 strains of the reference plant model *Arabidopsis thaliana*, 100K Pathogen Genomes aimed at 100,000 infectious microbes, and Genome 10K targeting 10,000 vertebrate species. However, deriving biological function and meaningful predictive models from sequences has continued to be a challenge. As a result, the scale and complexity of genomics research have advanced from studying a few genomic sequences of an organism to characterizing genome variations, gene expressions, biological pathways, and phenotypes for several thousands of organisms and their complex communities.

This has necessitated the availability of a wide array of computational tools that can process complex and heterogeneous data sets in an interoperable manner and sophisticated workflows that can seamlessly integrate these tools and their results at different stages of the analysis. Moreover, researchers often working collaboratively on large and complex systems need to be able to easily discover and integrate the analyses and results of their peers while simultaneously sharing their own results in a reproducible manner. Sustained efforts to lower the barriers to meet these monumental challenges have come to fruition in the form of a number of cyberinfrastructure platforms that adhere to the FAIR (findable, accessible, interoperable, reusable) guiding principles for research data stewardship. For more than a decade, the Ware lab has been invested in the growth and success of several such community initiatives since their inception. The rest of this section briefly discusses the specific projects that the Ware lab is currently engaged in and provides guidance and resources for, particularly to champion the cause of the plant genomics community.

CyVerse (Formerly The iPlant Collaborative): Transforming Science through Data-Driven Discovery

L. Wang, Z. Lu, K. Chougule, X. Wang, P. Van Buren, J. Stein, D. Ware [in collaboration with CSHL, employing more than 100 staff and headquartered at the University of Arizona, principal investigators (PIs) A. Parker, D. Ware, N. Merchant, M. Vaughn, and E. Lyons; dozens of collaborators located at more than 20 institutions]

The U.S. National Science Foundation (NSF) established the iPlant Collaborative in 2008 to develop cyberinfrastructure for life sciences research and democratize access to U.S. supercomputing capabilities. In 2015, iPlant was rebranded as CyVerse to emphasize an expanded mission to serve all life sciences.

For the Ware lab, work in the last year culminated in the development of a workflow management system, SciApps, for supporting automation

of complicated analysis. SciApps uses CyVerse Data Store for managing user data. The CyVerse Data Store was designed as a distributed storage system for hosting data on the cloud, providing convenient access. Currently, the CyVerse Data Store consists of two major storage systems with fast internet connections, one at the University of Arizona (UA) and another one at the Texas Advanced Computing Center (TACC). SciApps also uses TACC for computing through integration with the AGAVE science API. For each analysis, input data are staged from the CyVerse Data Store to the TACC clusters and results are archived back to the Data Store once completed.

As another major contribution to the CyVerse project, the Ware lab has developed and released several automated scientific workflows covering MAKER-based annotation, small genome assembly, GWASs, RNA-Seq analysis, methylation profiling, and bulk segregation analysis. The RNA-Seq workflows have been used to actively process the MaizeCODE expression data sets, and automation of data analysis and metadata processing is achieved by calling both SciApps API and the CyVerse API.

CyVerse is supported by NSF grants DBI-0735191, DBI-1265383, and DBI-1743442.

SciApps: A Cloud-Based Reproducible Workflow Platform

L. Wang, Z. Lu, P. Van-Buren, D. Ware [in collaboration with the CyVerse project]

As part of the CyVerse project effort, the Ware lab has developed SciApps, a cloud-based automated reproducible workflow platform. It is designed to manage distributed and edge computing assets, including both academic and commercial cloud assets. For supporting the FAIR data principles, SciApps assigns a unique workflow ID to each data set analyzed, and the workflow ID is attached as metadata of raw data and derived data like genome browser tracks (Findable); all raw data, derived data, and metadata are retrievable through the unique workflow ID (Accessible); and workflows and metadata are available as JSON files and extractable through the use of ontologies and standards (Interoperable/Reusable). In addition to FAIR data principles, SciApps tracks data provenance through analysis steps, recording results in standardized formats, and providing access to scripts, runs, and results. As a workflow manager, SciApps provides users with a way to track analysis and ensures reproducibility across clouds via the underlying technology, Docker/Singularity, which is adapted from established community efforts, such as BioConda and BioContainer.

For creating a workflow, each analysis job is submitted, recorded, and accessed through the SciApps web portal. Part or all of the series of recorded jobs can be saved as reproducible, sharable workflows for future execution with the original or modified inputs and parameters. The platform is designed to automate the execution of modular Agave apps and support executing workflows on either local or remote clusters or the combination of both. For executing workflows on U.S. supercomputing centers, users need to have a CyVerse account and upload their data to the CyVerse Data Store, a cloud storage built on the Integrated Rule-Oriented Data System (iRODS). All intermediate and final results of the analysis workflow are archived back to the Data Store once done to easily reproduce any one step of the analysis workflow.

Besides derived data, each SciApps workflow also captures the complete computational metadata to reproduce the entire analysis with one click and also provides direct links to input data and associated experimental metadata residing inside the CyVerse Data Store. Therefore, SciApps workflows can be used to share the complete analysis and input data among users or to the public—for example, for releasing the MaizeCODE data sets.

SciApps is supported by NSF grant DBI-1265383, and USDA-ARS (1907-21000-030-00D).

MaizeCODE: An Initial Analysis of Maize Function Elements

L. Wang, Z. Lu, K. Chougule, X. Wang, P. Van Buren, M.K. Tello-Ruiz, J. Stein, D. Ware [in collaboration with CSHL, principal investigators (PIs) T. Gingeras, K. Birnbaum, D. Jackson, R. Martienssen, W.R. McCombie, D. Micklos, M. Schatz, and D. Ware]

MaizeCODE, a project for an initial analysis of functional elements in the maize genome, has assayed five tissues of four maize genomes (B73, NC350, W22, TIL11) for RNA-Seq, ChIP-seq, Rampage, small RNA, and MNase (outside collaboration). MaizeCODE is

committed to open-access and reproducible science based on the FAIR data principles, supporting various ways to access data. First, all raw data are available through CyVerse Data Store, and the user can bulk download all data sets through iCommands (command line) or CyberDuck (GUI). Second, all ground-level analysis results of MaizeCODE data are available through SciApps (https://www.sciapps.org/?page_id=dataWorkflows&data_item=MaizeCODE) for integrative analysis. For example, differential expression analysis between tissues can be done in a few minutes because all gene quantification results are staged and chained together by the SciApps workflows. Third, peaks and signals from the ground-level analysis are available on a Genome Browser, JBrowser. To increase data discoverability, by calling CyVerse API, SciApps workflow IDs are added as metadata of the raw data in CyVerse Data Store during the automated analysis process (through SciApps API) so that the users can retrieve the entire analysis when browsing through the CyVerse Data Store or Data Commons. On SciApps.org, the workflow diagram provides a direct link to the raw data via CyVerse Data Commons so that users can check experimental metadata associated with each data set or workflow. On the Genome Browser side, SciApps workflow IDs are attached to each browser track for bringing up the specific workflow in a web-based iframe when users click on a track. In this way, the user can check all parameters that have been used to generate the track data, QC report, and metadata of the raw data or reproduce the entire analysis on TACC with a click.

In addition, all raw data have been submitted to the NCBI short read archive (SRA) using the SRA submission pipeline in the CyVerse Discovery Environment (DE). Through the submission process, all experimental metadata are stored in the iRODS-based Data Store of CyVerse and used for automating the ground-level analysis on SciApps. SciApps organizes both replicates (and controls if available) of each assay as one experiment (or a workflow with the unique ID), which represents an entity that chains raw data, analysis results, experimental metadata, and computational metadata together. The ground-level analysis includes quality control (QC), alignment to the reference genome, filtering, quantification (e.g., for gene expression), and peak calling (if needed). In summary, SciApps provides both a graphical user interface (GUI) and a RESTful API for users to check QC results, process new data, and reproduce existing analysis on the TACC cloud.

MaizeCODE is supported by an NSF grant IOS 1445025.

KBase: Department of Energy Systems Biology Knowledgebase

V. Kumar, S. Kumari, Z. Lu, J. Thomason [in collaboration with the Department of Energy (DOE) National laboratories and led by PI A. Arkin, Lawrence Berkeley National Laboratory (LBNL); co-PIs C. Henry, Argonne National Laboratory (ANL) and R. Cottingham of Oak Ridge National Laboratory (ORNL). As Plants Science Lead for KBase, D. Ware continues to informally serve as a co-PI on the project.]

The Systems Biology Knowledgebase (KBase, http://kbase.us) is a free, open-source platform for systems biology of plants, microbes, and their communities at scales ranging from biomolecular to ecological. The users can collaboratively generate, test, compare, and share hypotheses about biological functions; perform large analyses on scalable computing infrastructure; and combine experimental evidence and conclusions to model plant and microbial physiology and community dynamics.

The KBase platform has extensible analytical capabilities that currently include (meta)genome assembly, annotation, comparative genomics, transcriptomics, and metabolic modeling; a web-based user interface that supports building, sharing, and publishing reproducible and well-annotated analyses with integrated data; and a software development kit that enables the community to add functionality to the system. The Ware lab has been engaged in the development and maintenance of a number of analysis tools and data resources that enable the plant science community to gain insight into the evolution of genes and genomes, profile transcriptomes, perform genome functional modeling with metabolic networks, and identify differential expression between tissues, developmental stages, environmental conditions, and genetic backgrounds. These capabilities are directly relevant to important DOE research targets such as optimizing biomass production in biofuel feedstocks.

KBase was conceived from the beginning as a knowledgebase that would bring together relevant computational systems biology tools and data for microbes, plants, and interactions between the two. To

support this, KBase's data model integrates reference data and shared user data for both microbes and plants so that results for one organism can be applied to others. In further support of the knowledgebase this year, the KBase team introduced new platform features such as search, generics, and a relation engine. The search was extended to include public user data and shared analyses in addition to reference data. Also, the search UI was revamped to provide the search results views at different granularity levels. The introduction of support for generics data types in KBase makes it possible for the KBase apps to interoperably operate on objects with similar informational structures. For example, many transcriptomic, proteomic, and metabolomic data come in the form of a generic expression matrix data type. KBase also created an early prototype of a relation engine that allows queries on data connected across the system through various relationship types. For example, the users can query for gene candidates that can potentially perform a particular reaction of interest in a genome of interest. These features will ultimately evolve into knowledge-discovery features, enabling KBase to propose new hypotheses by making connections across the system.

In addition to the platform development, KBase invested significant resources in external engagements, such as co-development with JGI for improved JGI search and implementation of JGI metagenome tools and genome homology service in KBase, and roadmap development with Environmental Molecular Sciences Laboratory (EMSL) targeted on EMSL tools in metabolomics. Also, user working groups (UWGs) were identified to help external research teams bring their data, tools, and workflows to KBase in the areas of metabolism, microbiome, and functional genomics. The Ware lab is engaged in the functional genomics UWG, which focuses on the integration of tools and development of science stories related to gene and regulatory functions in plants and microbes—in particular, linking genes to function and exploring the impact of genetic variation on whole genome phenotypes.

KBase is actively engaging the external community to help us improve our tools and workflows for plant science, including support for large-scale reads upload and analysis, plant genome annotation, functional genomic clustering and enrichment, physiological modeling and variation, and trait-based modeling analysis. The KBase team published its first platform paper in *Nature Biotechnology* in July 2018 (Arkin et al. 2018). The Ware lab has actively engaged the community through various channels such as webinars, presentations, talks, posters, and demonstrations during major community events such as Plant Biology, Plant and Animal Genomes, Genome Informatics, and Biology of Genomes. We welcome the community's feedback on our current tools and the new functionality we should add to the platform. We also invite the community members to share their scientific analyses using KBase Narratives.

PUBLICATIONS

Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, et al. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* **36:** 566–569. doi:10.1038/nbt.4163

Chougule KM, Wang L, Stein JC, Wang X, Devisetty UK, Klein RR, Ware D. 2018. Improved RNA-seq workflows using CyVerse cyberinfrastructure. *Curr Prot Bioinform* **63:** e53. doi:10.1002/cpbi.53

Gaudinier A, Rodriguez-Medina J, Zhang L, Olson A, Liseron-Monfils C, Bagman AM, Foret J, Abbitt S, Tang M, Li B, et al. 2018. Transcriptional regulation of nitrogen-associated metabolism and growth. *Nature* **563:** 259–264.

Harper L, Campbell J, Cannon EKS, Jung S, Main D, Poelchau M, Walls R, Andorf C, Arnaud E, Berardini T, et al. 2018. AgBioData consortium recommendations for sustainable genomics and genetics databases for agriculture. *Database* **2018:** bay088. doi:10.1093/database/bay088

Jiao Y, Burow G, Gladman N, Acosta-Martinez V, Chen J, Burke J, Ware D, Xin Z. 2018a. Efficient identification of causal mutations through sequencing of bulked F_2 from two allelic bloomless mutants of *Sorghum bicolor*. *Front Plant Sci* **8:** 2267. doi:10.3389/fpls.2017.02267

Jiao Y, Lee YK, Gladman N, Chopra R, Christensen SA, Regulski M, Burow G, Hayes C, Burke J, Ware D, Xin Z. 2018b. MSD1 regulates Pedicellate spikelet fertility in sorghum through the jasmonic acid pathway. *Nat Comm* **9:** 822. doi:10.1038/s41467-018-03238-4

Liseron-Monfils C, Olson A, Ware D. 2018. NECorr, a tool to rank gene importance in biological processes using molecular networks and transcriptome data. bioRxiv doi:10.1101/326868

Majoros WH, Holt C, Campbell MS, Ware D, Yandell M, Reddy TE. 2018. Predicting gene structure changes resulting from genetic variants via exon definition features. *Bioinformatics* **25:** 1–8. doi:10.1093/bioinformatics/bty324

Springer NM, Anderson SN, Andorf CM, Ahern KR, Bai F, Barad O, Barbazuk WB, Bass HW, Baruch K, Ben-Zvi G, et al. 2018. The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat Genet* **50:** 1282–1288.

Stein JC, Yu Y, Copetti D, Zwickl DJ, Zhang L, Zhang C, Chougule K, Gao D, Iwata A, Goicoechea JL, et al. 2018. Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus *Oryza*. *Nat Genet* **50:** 285–296. doi:10.1038/s41588-018-0040-0

Sun S, Zhou Y, Chen J, Shi J, Zhao H, Zhao H, Song W, Zhang M, Cui Y, Dong X, et al. 2018. Extensive intraspecific gene order and gene structural variations between Mo17 and other maize

genomes. *Nat Genet* **50:** 1289–1295. doi:10.1038/s41588-018-0182-0

Wang B, Regulski M, Tseng E, Goodwin S, Olson A, McCombie R, Ware D. 2018a. A comparative transcriptional landscape of maize and sorghum obtained by single-molecule sequencing. *Genome Res* **28:** 921–932. doi:10.1101/gr.227462.117

Wang L, Lu Z, Van Buren P, Ware D. 2018b.SciApps: a cloud-based platform for reproducible bioinformatics workflows. *Bioinformatics* **34:** 3917–3920.

Wang L, Lu Z, Van Buren P, Ware D. 2018c. SciApps: a bioinformatics workflow platform powered by XSEDE and CyVerse. Association for Computing Machinery Digital Library. https://dl.acm.org/citation.cfm?id=3219101

Wittmeyer K, Cui J, Chatterjee D, Lee TF, Tan Q, Xue W, Jiao Y, Wang PH, Gaffoor I, Ware D, et al. 2018. The dominant and poorly penetrant phenotypes of maize *Unstable factor for orange1* are caused by DNA methylation changes at a linked transposon. *Plant Cell* **30:** 3006–3023. doi:10.1105/tpc.18.00546

QUANTITATIVE BIOLOGY

Gurinder S. "Mickey" Atwal and colleagues are applying insights from the physical and computational sciences to the study of population genetics and human disease. The Atwal lab has modeled the process by which genetic variants, or alleles, have evolved in the last 100,000 years of human history. This has recently led to surprising insights about the role of *p53*, a master tumor suppressor gene, in female fertility and furthered our understanding of how complex gene networks evolve. The lab has analyzed the comparative genomics and physical organization of cancer-related genes and their role in mediating tumorigenesis across numerous tissue types. Recently, they have begun to focus efforts on understanding cancer genome evolution on shorter timescales by analyzing nucleotide sequences from single cells.

Human development requires the regulated activity of thousands of genes in hundreds of distinct cell types throughout life. One requirement for this process is that each cell must contain an intact, functional genome free of mutations. One type of mutation can arise from the activation of transposable elements (TEs). These viral-like parasites lay dormant within our genomes, but have the capacity to hop into new genomic locations, causing mutations as they break the surrounding DNA sequence. **Molly Hammel**'s lab has discovered mounting evidence implicating transposon activity in a host of human diseases, with particular evidence for TE activation in neurodegenerative diseases: amyotrophic lateral sclerosis (ALS) and fronto-temporal dementia (FTD).

Ivan Iossifov's lab focuses on the development of new methods and tools for genomic sequence analysis and for building and using molecular networks, and they apply these methods and tools to specific biomedical problems. They study the genetics of common diseases in humans using two main tools: next-generation sequencing and molecular networks representing functional relationships among genetic loci. These approaches in combination enable the kind of large-scale studies necessary for furthering our understanding of the complex etiology of disorders such as autism, bipolar disorder, and cancer.

Justin Kinney completed his Ph.D. in Physics at Princeton University in 2008 and began his term as a Quantitative Biology Fellow in 2010. His research focuses on developing next-generation DNA sequencing as a tool for dissecting the structure and function of large macromolecular complexes. Of particular interest to his lab is the biophysical basis of transcriptional regulation—how simple interactions between proteins and DNA allow promoters and enhancers to modulate genes in response to physiological signals.

In 2010, Kinney and colleagues published a paper demonstrating Sort-Seq, a novel sequencing-based method that can measure the functional activity of hundreds of thousands of slightly mutated versions of a specific DNA sequence of interest. Using a novel information-theoretic analysis of the resulting data, Kinney et al. were able to quantitatively measure, in living cells, the protein–DNA and protein–protein interactions controlling messenger RNA (mRNA) transcription at a chosen promoter. Kinney continues to develop this approach using a combination of theory, computation, and experiment. From a biological standpoint, Sort-Seq allows researchers to investigate important but previously inaccessible biological systems. Kinney's lab is currently using Sort-Seq to address open problems in transcriptional regulation, DNA replication, and immunology. These experiments also present new challenges for the field of machine learning, and a substantial fraction of Kinney's efforts are devoted to addressing the theoretical and computational problems relevant to the analysis of Sort-Seq data.

Alexander Krasnitz and colleagues develop mathematical and statistical tools to investigate population structure of cells constituting a malignant tumor and reconstruct evolutionary processes leading to that structure. These tools are designed to make optimal use of emerging molecular technologies—chief among them high-throughput genomic profiling of multiple individual cells harvested from a tumor. By analyzing these profiles, Krasnitz derives novel molecular measures of malignancy, such as the number of aggressive clones in a tumor, the invasive capacity of each clone, and the amount of cancer-related genetic alteration sustained by clonal cells. Krasnitz and colleagues collaborate closely with clinical oncologists to explore the utility of such measures for earlier detection of cancer, more accurate patient outcome prediction and risk assessment, and better-informed choice of treatment options.

There is increasing evidence that rare and unique mutations have a significant role in the etiology of many diseases such as autism, congenital heart disease, and cancer. **Dan Levy**'s group develops algorithms to identify these mutations from large, high-throughput data sets comprising thousands of nuclear families. After previously working with high-resolution comparative genome hybridization (CGH) arrays, Levy's group now uses targeted sequence data. Levy has developed methods for identifying de novo mutations (i.e., those seen in a child but not in his or her parents) by simultaneously genotyping the entire family; the team is currently focused on building algorithms to detect copy number variants and multiscale genomic rearrangements. Although their copy number methods are based on "read" density, there are classes of mutations that require analysis at the level of the read. Thus, they are developing algorithms to identify insertions, deletions, inversions, transpositions, and other complex events. Other projects in the Levy lab include analysis of single-cell RNA, phylogenetic reconstruction from sparse data sets, and disentangling haplotypes from sperm and subgenomic sequence data.

The **David McCandlish** lab develops computational and mathematical tools to analyze and exploit data from high-throughput functional assays. The current focus of the lab is on analyzing data from so-called "deep mutational scanning" experiments. These experiments simultaneously determine, for a single protein, the functional effects of thousands of mutations. By aggregating information across the proteins assayed using this technique, they seek to develop data-driven insights into basic protein biology, improved models of molecular evolution, and more accurate methods for predicting the functional effects of mutations in human genome sequences.

Critically, these data also show that the functional effects of mutations often depend on which other mutations are present in the sequence. They are developing new techniques in statistics and machine learning to infer and interpret the complex patterns of genetic interaction observed in these experiments. Their ultimate goal is to be able to model these sequence–function relationships with sufficient accuracy to guide the construction of a new generation of designed enzymes and drugs and to be able to predict the evolution of antigenic and drug-resistant phenotypes in rapidly evolving microbial pathogens.

Michael Schatz is a computational biologist and an expert at large-scale computational examination of DNA sequencing data, including the alignment, assembly, and analysis of next-generation sequencing reads. These methods have been used to reconstruct the genomes of previously unsequenced organisms, probe sequence variations, and explore a host of biological features across the tree of life. Recent improvements in sequencing technologies are challenging our capacity to store and analyze the huge volume of DNA sequence data being generated. Consequently, the Schatz lab is particularly interested in capitalizing on the latest advances in distributed and parallel computing, especially cloud computing technologies, to advance the state of the art in bioinformatics and genomics. In a recent breakthrough, the Schatz lab was able to create a hybrid software-based solution to eliminate errors in so-called third-generation sequencing. This makes it remarkably easier to compile, align, and analyze full genome sequences.

Modern genomic technologies make it relatively easy to generate rich data sets describing genome sequences, RNA expression, chromatin states, and many other aspects of the storage, transmission, and expression of genetic information. For many problems in genetics today, the limiting step is no longer in data generation, but in integrating, interpreting, and understanding the available data. Addressing these challenges requires expertise both in the practical arts of data analysis and the theoretical underpinnings of statistics, computer science, genetics, and evolutionary biology.

Adam Siepel's group focuses on a diverse collection of research questions in this interdisciplinary area. Over the years, their research has touched on topics including the identification of recombinant strains of HIV, the discovery of new human genes, the characterization of conserved regulatory elements in mammalian genomes, and the estimation of the times in early human history when major population groups first diverged. A general theme in their work is the development of precise mathematical models for the complex processes by which genomes evolve over time, and the use of these models, together with techniques from computer science and statistics, both to peer into the past and to address questions of practical importance for human health. Recently, they have increasingly concentrated on research at the interface of population genomics and phylogenetics, with a particular focus on humans and the great apes. They also have an active research program in computational modeling and analysis of transcriptional regulation in mammals and *Drosophila*, in close collaboration with Prof. John Lis at Cornell University.

QUANTITATIVE BIOLOGY

G.S. Atwal J. Carter
P. Gilbo
K. Grigaityte
R. Utama

Fueled by data generated from recent technological developments in DNA sequencing, the Atwal lab primarily develops mathematical models and computational methods to tackle problems in cancer genomics, immunology, and biological machine learning. Recently we have focused efforts on understanding the cellular heterogeneity and evolution of the tumor microenvironment and how we may leverage this knowledge to inform effective cancer immunotherapeutics in breast cancers.

This year we have continued to leverage advances in microfluidic-based single-cell sequencing platforms to study both the human T-cell receptor repertoire in peripheral blood and the tumor microenvironment of breast tumors. A major milestone has been the completion of work generating and analyzing the largest collection of single T cells in collaboration with Juno Therapeutics, in which we uncovered universal clonal distributions and common clones across individuals. We have also been following up on the investigation of ectopic expression of germline genes, transcribing genetic loci exclusively expressed in testes/ovaries and in various tumors. Preliminary results have also identified ectopic expression of germline piwi genes in samples of glioblastoma multiforme extracted from the Cancer Genome Atlas. However, the landscape of ectopically expressed germline genes is unknown, and their functional impact on cancer development remains elusive. We have continued an integrative study that sought to determine the ectopic expression of germline transcripts in glioblastoma multiforme and breast cancer, elucidating the predictive value of germline molecular programs as diagnostic markers and immunogenic therapeutic targets. We have continued to collaborate with Betsy Barnes at the Feinstein Institute to test candidate peptides as vaccines in mouse models of triple-negative breast cancer.

Single-Cell Sequencing of the Adaptive Immune System

K. Grigaityte, J. Carter [in collaboration with Juno Therapeutics (now Celgene)]

A diverse T-cell repertoire is a critical component of the adaptive immune system, providing protection against invading pathogens and neoplastic changes, relying on the recognition of foreign antigens and neoantigen peptides by T-cell receptors (TCRs). However, the statistical properties of the αβ T-cell repertoire in healthy individuals have remained poorly characterized, in large part because of the laborious task of sequencing single T cells in a high-throughput fashion.

Over the last year, we continued to explore a high-throughput data set of full-length, high-quality, paired αβ sequences ($n = 205{,}950$) from peripheral blood samples of five healthy individuals (three males, two females, ages 33–69) acquired through a recently developed microfluidic method of single-cell RNA sequencing in emulsion droplets. The sequenced T cells were further stratified into $CD4^+$ ($n = 73{,}495$) and $CD8^+$ ($n = 30{,}321$) subtypes, based on paired sequence tags introduced by labeling with DNA-conjugated antibodies. Previously, we showed that the paired αβ TCR repertoire cannot necessarily be directly inferred by observing the repertoire of one chain alone. That is, although single-chain bulk sequencing provides valuable information regarding the TCR repertoire, it is not equivalent to the paired repertoire and does not fully capture TCR diversity. This is alarming because almost all previous studies on T-cell repertoire, including clinical cancer research, made their conclusions using β chain sequencing alone. In addition, we demonstrated that the T-cell clone size distribution of every subject is a double power law. We suggest that this mixture of two power laws comes from two generation processes of TCRs, such as memory versus naive T cells and proliferated versus nonproliferated T cells. Interestingly, $CD8^+$ T-cell population contains a higher fraction

(0.1) of lower exponent power law than that of $CD4^+$ T-cell population (0.01). We also investigated TCRs that are shared between our data set and a publicly available database of TCRs with known antigen specificity. We compared these public sequences with the rest of our data set and found that the public clones tend to reside at the tail of the power law, are more likely to be generated, and are closer to germline DNA compared to nonpublic receptors.

The analyses we have done so far were only on the T-cell repertoire from a single time point. Over the past year we collected PBMC samples monthly from three individuals and collected longitudinal samples before and after a flu shot. Once sequenced and analyzed, the data will be invaluable for understanding T-cell repertoire changes over time. In addition, this study could be particularly critical for developing successful immunotherapeutic approaches and correctly assessing patient response to treatment.

Bayesian Inference of Power Law Distributions

K. Grigaityte

Power law distributions abound in empirical data. Some examples in which power laws appear are geography (city population sizes), literature (word usage), biology (T-cell clone sizes), networks (connections per node), and statistical physics (order parameters in phase transitions). The widespread and enigmatic appearance of these long-tailed distributions has generated much research and debate into their origins and their detection. A number of differing theories have been proposed to explain the generation of power law behavior, and their study still provides an active rich area of interest. On the other hand, accurate detection of actual power law distributions in both simulated and real-world data remains challenging, especially in the low sample limit, precisely because of the long-tailed nature of the distributions. Currently, the most accurate method for fitting the exponent of power laws is via a maximum likelihood approach. Although efficient and accurate, this method does not provide uncertainty of the inferred exponent value and does not perform well with low sample sizes.

We have developed a novel Bayesian inference methodology, implemented in an open-source Python package, that overcomes limitations posed by maximum likelihood algorithms. In addition, our approach is able to infer a mixture of power law. Specifically, we utilize Markov chain Monte Carlo sampling (Metropolis–Hastings) to sample the posterior distribution to obtain point estimates and uncertainty. We have also analytically derived an uninformative (Jeffreys) prior for both continuous and discrete power laws.

We compared our algorithm to previously used linear regression and maximum likelihood methods with simulated distributions and showed that the Bayesian inference method is superior, especially in the low sample limit. Analysis on real molecular data revealed the existence of a mixture of two power laws (~2 and ~4.5) in the T-cell clone size distribution obtained from high-throughput single-cell sequencing.

Cell-Nonautonomous Interactions during Nonimmune Stromal Progression in the Breast Tumor Microenvironment

R. Utama [in collaboration with the Lee laboratory at City of Hope]

The breast tumor microenvironment of primary and metastatic sites is a complex milieu of differing cell populations, consisting of tumor cells and the surrounding stroma. Despite recent progress in delineating the immune component of the stroma, the genomic expression landscape of the nonimmune stroma (NIS) population and its role in mediating cancer progression and informing effective therapies are not well understood. Here we obtained 52 cell-sorted NIS and epithelial tissue samples across 37 patients from (i) normal breast, (ii) normal breast adjacent to primary tumor, (iii) primary tumor, and (iv) metastatic tumor sites. Deep RNA-Seq revealed diverging gene expression profiles as the NIS evolves from normal to metastatic tumor tissue, with intrapatient normal-primary variation comparable to interpatient variation. Significant expression changes between normal and adjacent normal tissue support the notion of a cancer field effect, but extended out to the NIS. Most differentially expressed protein-coding genes and lncRNAs were found to be associated with pattern formation, embryogenesis, and the epithelial–mesenchymal transition. We validated the protein expression changes of a novel candidate gene, *C2orf88*, by immunohistochemistry staining of representative

tissues. Significant mutual information between epithelial ligand and NIS receptor gene expression, across primary and metastatic tissue, suggests a unidirectional model of molecular signaling between the two tissues. Furthermore, survival analyses of 827 luminal breast tumor samples demonstrated the predictive power of the NIS gene expression to inform clinical outcomes. Together, these results highlight the evolution of NIS gene expression in breast tumors and suggest novel therapeutic strategies targeting the microenvironment.

Single-Cell Profiling of the Stromal Tissue in Breast Cancer

R. Utama [in collaboration with the Lee laboratory at City of Hope]

In conjunction with our efforts to characterize the transcriptional landscape of the breast tumor microenvironment via bulk sequencing, we also investigated the cellular phenotypic heterogeneity of stromal tissue via single-cell RNA-Seq. Library preparation of breast tumor biopsy samples was conducted on a 10X Genomics platform by the Lee group at City of Hope (CoH). Previously designed protocols were implemented for tissue acquisition, cell digestion, isolation, sorting, and barcoding with unique molecular identifiers. We encountered and resolved technical challenges such as low single-cell viability and collagen binding of stromal fibroblasts. Sequencing of the various samples was run at both CoH and CSHL using a NextSeq 500 Illumina machine. Downstream bioinformatics analysis was conducted using Cellranger and Seurat pipelines for four luminal and one triple-negative breast cancer patients from whom we obtained both primary tumor and draining lymph node samples. Clustering of the single cells was done using a graph-based methodology on the top 10 principal components and then projected onto tSNE plots. Differential expression analyses revealed molecular subtypes of cells in differing proportions across samples.

We successfully captured a sufficient number of NIS cells ($N \sim 200$) from only one patient. The NIS clusters were found to be stratified further into cancer-associated fibroblasts (CAFs), pericytes, and endothelial cells based on canonical markers such as *FAP, RGS5,* and *PECAM1*. Preliminary analyses suggested that the CAF cluster is robustly split into two molecular subtypes, one of which is associated with a myofibroblast (myCAF) phenotype, characterized by overexpression of ACTA2. Analyses of a triple-negative sample surprisingly revealed the existence multiple subtypes of breast cancer, such as basal-like and LAR subtypes. Finally, we compared the T-cell receptor clonality between a tumor and its draining lymph node and found identical clones highly expanded in both sites, suggesting direct communication between the two sites.

Functional Roles for αβ T-Cell Receptor Pairing in Cell Lineage and Antigen Specificities

J. Carter [in collaboration with Juno Therapeutics]

Although structural studies of individual T-cell receptors (TCRs) have revealed important roles for both the α and β chain in directing MHC and antigen recognition, repertoire-level immunogenomic analyses have historically examined the β chain in isolation. Consequently, the amount of useful information about TCR function encoded within αβ pairings remains largely unknown. To determine whether information about αβ pairings can meaningfully increase our understanding of TCR repertoire function, we analyzed paired TCR sequences from nearly 100,000 unique CD4 and CD8 T cells captured using two different high-throughput, single-cell sequencing approaches. Our results demonstrate little overlap in the CD4 and CD8 repertoires, with shared TCR sequences possessing significantly shorter CDR3 sequences with higher generation probabilities. We further utilized tools from information theory and machine learning to show that although α and β chains are only weakly associated with lineage, αβ pairings appear to synergistically drive TCR–MHC interactions. Interestingly, we found Vαβ gene pairings to be the TCR feature most informative of T-cell lineage, thereby supporting the existence of germline-encoded TCR–MHC interaction motifs. Finally, annotating our TCR pairs using a database of sequences with known antigen specificity, we found that up to one-third of the cells had differing antigen specificities for each of the chains, suggesting that αβ pairing is required for the accurate inference of repertoire functionality. Together, these findings provide biological insight into the functional implications of αβ pairing.

Inference of Allelic Inclusion Rates in the Human T-Cell Receptor Repertoire

J. Carter

A small population of αβ T cells are known to express more than one unique T-cell receptor (TCR). However, limitations in the requisite single-cell sequencing technologies have precluded a comprehensive understanding of such T-cell allelic inclusion. To address this, we utilized recent advances in high-throughput single-cell sequencing to examine the peripheral T-cell repertoires from eight healthy individuals. We then developed a Bayesian inference model to reliably estimate the true rate of α and β TCR allelic inclusion from differing emulsion-barcoding single-cell sequencing strategies. In brief, our generative model attempts to estimate repertoire allelic inclusion rates by capturing three major components of the experimental process: (i) the number of cells loaded into each droplet, (ii) the rate of allelic inclusion, and (iii) the probability that an expressed allele is not experimentally observed. We experimentally validated allelic inclusion rates both at the mRNA level using 96-well plate–based single-cell sequencing and at the protein level using flow cytometry, demonstrating a high concordance among all three measures. Furthermore, we generated a novel, high-confidence database composed of more than 51,000 allelic inclusion TCR sequences derived directly from single cells. Our analyses reveal that T-cell allelic inclusion likely contributes unique sequences to the human TCR repertoire and potentially plays an important functional role in the healthy adaptive immune response. Together, these findings provide a valuable new resource for the study of T-cell allelic inclusion and suggest that allelic inclusion T cells represent a substantial component of the human TCR repertoire.

Transcriptomic Model of Immune Checkpoint Therapy Response

P. Gilbo

We investigated the ability to create a predictive model for immune checkpoint therapy using machine learning and differential analyses of the whole transcriptome of tumor core biopsies. The underlying hypothesis was that infiltrating T cells respond to aberrant expression in addition to neoantigens stemming from genomic single-nucleotide mutations. We initially focused our efforts on RNA-Seq data from metastatic melanoma patients prior to anti-CTLA4 therapy. The task was to develop a computational classifier that would accurately predict which patients would respond (tumor regression) to immunotherapy within a specified time frame. After pseudoalignment of the raw FASTQ files, a novel random feature algorithm was developed to capture informative low-dimensional representations of the high-dimensional ($K \sim 20{,}000$) input data from the low number of samples ($N \sim 40$). Preliminary investigation of the classifier performance, using k-fold cross-validation, revealed significant nonrandom sensitivity and specificity (AUROC ~ 0.6). Ongoing investigations include development of a classifier incorporating both genomic and transcriptomic data from core biopsies.

PUBLICATIONS

Carter JA, Preall JB, Grigaityte K, Goldfless SJ, Briggs AW, Vigneault F, Atwa GS. 2018. T-cell receptor alpha-beta chain pairing is associated with CD4 and CD8 lineage specification. bioRxiv doi:10.1101/293852

In Press

Utama R, Bastian A, Sadagopan N, Jin Y, Antoniou E, Gao Q, Huang Y, Gopalakrishna-Pillai S, Lee P, Atwal G. 2019. Cell non-autonomous interactions during non-immune stromal progression in the breast tumor microenvironment. bioRxiv doi:10.1101/5401122

INTEGRATING GENOMIC DATA SETS TO UNDERSTAND GENE REGULATION IN DEVELOPMENT AND DISEASE

M. Hammell Y. Hao, Y-J. Ho, Y. Jin, K. O'Neill, N. Rozhkov, R. Shaw, O. Tam, M-K. Yip

Human development requires the regulated expression of thousands of genes in hundreds of distinct cell types throughout life. One requirement for this process is that each cell must contain an intact, functional genome. Furthermore, there is a set of rules for how the genes in our genomes coordinate their activities, and these rules are encoded within gene regulatory networks. Many human diseases occur when these basic processes are altered—either through altering the genome itself (as in the mutations seen in cancers) or through altering the way in which genes interact with each other. The focus of the M. Hammell lab is to understand how mutations in our genomes lead to both alterations in the function of the mutated gene itself and the repercussions of these alterations on the hundreds of other neighboring genes within the network. To this end, the Hammell lab uses computational algorithms to integrate multiple types of genomic and transcriptomic sequencing data into models of cellular function.

One particular focus of our lab is to apply these novel statistical analysis methods to better understand how transposable elements are controlled in animal cells. Transposons are viral-like parasites (Bourque et al. 2018) that lie dormant within our genomes, but have the capacity to hop into new genomic locations, causing mutations as they break the surrounding DNA sequence. Mounting evidence has implicated transposon activity in a host of human diseases, with particular evidence for activation in several neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). By combining the power of systems-level, high-throughput data analysis with patient-derived ALS tissue samples, the lab aims to better understand how these viral-like parasites contribute to cell death and disease.

Endogenous Retroviral-Like Elements May Contribute to Neurodegeneration

Y. Jin, N. Rozhkov, R. Shaw, O. Tam

TDP-43 is an RNA-binding protein that is known to control proper processing of many RNA targets in neurons. Mutation of TDP-43 has been associated with a variety of neurodegenerative diseases including ALS, frontotemporal lobar degeneration (FTLD), and Alzheimer's disease (AD). However, the normal function of TDP-43 in neuronal development and maintenance has not been fully characterized, and few of its mRNA targets have been definitively associated with the neurodegenerative diseases that result from loss of TDP-43 function. In collaboration with the Dubnau lab at Stony Brook University, the group has explored the novel hypothesis that TDP-43 normally plays a large and hitherto uncharacterized role in regulating the expression of transposable elements (TEs). TEs are retroviral-like elements encoded within our genomes whose unregulated expression leads to genetic instability as well as cellular toxicity. In collaboration with the Dubnau lab at Stony Brook, we have shown that TE transcripts are elevated upon expression of mutant, aggregate-prone forms of human TDP-43 in the fly brain, and that neurodegeneration results from aberrant expression of this hTDP-43 protein. In previous publications, members of the Hammell lab have shown that TDP-43 binds widely to TE transcripts in mammals, and that TDP-43 binding to TEs is lost in human patients diagnosed with FTLD, a disease characterized by TDP-43 proteinopathy. Although these studies support a role for TDP-43 in regulating TE expression, our future goals are centered on determining the role that dysregulation of TEs plays in neurodegenerative disease. Ongoing research in the lab is focused on three main areas: (1) improving our ability to detect active transposons in ALS and FTLD patient

samples through software innovation, (2) examining the degree to which transposons and other retroviral-like elements are interfering with cellular function in ALS and FTLD patient tissues, and (3) understanding the basic biology of how TDP-43 interacts with the general transposon control machinery.

In a collaboration with a large ALS patient sequencing consortium under way at the New York Genome Center (NYGC), members of the Hammell lab have integrated targeted genotyping data with deep expression profiling for hundreds of ALS patients and controls. Analysis of these patient profiles has revealed that elevated TE expression does occur in the cortical regions for a substantial fraction of ALS patients. Correlated gene markers revealed that those patients with elevated TE levels also show dysregulation of other targets of TDP-43, suggesting that TDP-43 protein dysfunction is leading to TE desilencing in a subset of ALS patients. Collaborations with ALS Consortium members has already led to two publications (Conlon et al. 2018; Nicolas et al. 2018). The lab was also recently awarded a large grant from the Chan Zuckerberg Initiative to follow up these results in a larger set of ALS patient tissues and to see if related processes are occurring in other neurodegenerative diseases. Finally, efforts are also under way to examine the exact molecular mechanisms by which TDP-43 contributes to TE silencing.

Algorithms to Enable Better Recovery of Repeat Regions in Genomic Studies

Y. Jin, K. O'Neill, O. Tam

Transposable elements, viral-like sequences in our genomes, have been historically difficult to study because of their highly repetitive nature. Nearly half of the human genome is composed of TE-derived sequences, with millions of copies of TEs scattered throughout the chromosomes. Although most of these copies are nonfunctional, thousands of TEs retain the ability to mobilize and create new copies of themselves elsewhere in the genome. The difficulty lies in differentiating these active TEs from the millions of other harmless copies with nearly identical sequences. This presents both a technical challenge for experimentally isolating TE-derived sequences from the genomes of cells as well as a computational challenge for determining where each TE copy originates in the genome of a particular sample. Members of our lab have recently developed novel statistical inference methods to solve the computational challenge of analyzing TE expression in sequencing studies (Jin et al., *Bioinformatics 31:* 3593 [2015]; Jin and Hammell 2018; O'Neill et al. 2018). All of these algorithms, which together form the TEtoolkit, use maximum likelihood frameworks to statistically infer the correct originating locus of sequencing reads that map ambiguously to many related genomic regions. These algorithms have been used to examine the basic mechanisms of controlling TE expression and activity in germline tissues in many collaborative studies. Ongoing efforts will establish optimized analysis protocols for many different types of TE studies (genome resequencing studies, chromatin association studies, etc.). In addition, several members of the lab are developing optimized protocols for isolating and identifying novel TE insertion sites in the genomes of individual cells, as active TEs provide one source of genetic mutation that can occur somatically, in adult cells. Together, these efforts will provide the tools with which to determine the extent of TE activity in patient samples.

The most recently published software program in the TEtoolkit is TEsmall (O'Neill et al. 2018). TEsmall is specifically designed for improving the analysis of small RNA sequencing reads that map to repetitive genomic loci. MicroRNAs (miRNAs) are small 22-nt RNAs that act to inhibit the expression of mRNA target genes through direct binding to mRNA targets and are the most abundant class of small RNAs in most cell types. Although miRNAs typically dominate small RNA transcriptomes, many other classes are present, including tRNAs, snoRNAs, snRNAs, Y-RNAs, piRNAs, and siRNAs. Interactions between processing machinery and targeting networks of these various small RNA classes remain unclear, largely because these small RNAs are typically analyzed separately. TEsmall allows for the integrated analysis of all small RNA classes in a single integrated workflow that includes allocation of reads from repetitive regions.

To test the TEsmall pipeline, we applied TEsmall to small RNA libraries generated from melanoma cells responding to targeted inhibitors of the MAPK pathway. Targeted oncogene inhibitors have emerged as a way to tailor cancer therapies to the particular mutations present in a given tumor. Although these targeted strategies are typically effective for short intervals, the emergence of resistance is extremely common, limiting the effectiveness of single-agent

therapeutics and driving the need for a better understanding of resistance mechanisms. Using TEsmall, we identified several microRNAs and other small RNA classes that are enriched in inhibitor-resistant melanoma cells in multiple melanoma cell lines and may be able to serve as markers of resistant populations more generally.

SAKE: Analysis Software for Single-Cell RNA-Seq Data Sets

Y-J. Ho

The Cancer Genome Atlas (TCGA) was a large-scale cancer profiling project that aimed to sequence the genetic mutations and expression profiles for hundreds of patients in dozens of different cancers. Rather than finding a few genetic mutations that explain most cancers, these TCGA studies uncovered thousands of mutations with surprisingly little overlap from patient to patient. This makes the task of designing targeted therapies to treat an individual patient's set of cancer-causing mutations a difficult one. Although the genetic mutations showed little pattern of recurrence among patients, the gene expression data did show clear expression patterns—or cancer molecular subtypes. Moreover, expression subtypes are often predictive of patient survival rates, likelihood to metastasize, and response to targeted therapies. For the melanoma cancer samples that our lab generates, lab members have also been able to show that these molecular subtypes persist in 2D cultured cell lines, in 3D cultured organoid cells growing in Matrigel, in 3D tumors grown in a mouse xenograft model, and from tumor samples taken directly from patients. However, for all of these samples, the transcriptomes came from bulk tissue or cell populations whose behavior and phenotypes are generally an average over a very large population of tumor cells. This ignores the heterogeneity that exists both between different cell types within a tissue and between different individual cells within a population. This problem is especially important in the context of human cancers, which are continually evolving to develop invasive properties as well as resistance to therapeutics. To address the problems of cellular heterogeneity, the field of biology needs better tools, in terms of both the technology for single-cell sequencing assays and the statistical methods for analyzing data from single-cell assays.

Although our technological ability to generate single-cell data sets has improved considerably over the last few years, our ability to analyze these data sets with algorithms that can robustly detect expression patterns in the presence of high noise and sparse sampling is lagging behind. Members of our lab began by thoroughly testing dozens of statistical models for expression clustering to find a method that would be accurate, robust to noise, and computationally efficient for hundreds of libraries. Using thousands of iterative randomizations and a set of gold standard cell types, we have shown that an adapted version of nonnegative matrix factorization (NMF) shows the most robust performance and is computationally efficient enough to rapidly analyze data from hundreds to thousands of cells. Furthermore, this NMF-based clustering method provides quantitative confidence estimates on the number of clusters present as well as the assignment of any given cell to a cluster, providing a means of assessing the accuracy of our methods as compared to other analysis strategies. We have built a user-friendly web browser application to provide this adapted NMF analysis package to the scientific community, which we have entitled SAKE for single-cell analysis and klustering evaluation (Ho et al. 2018).

PUBLICATIONS

Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, Imbeault M, Izsvák Z, Levin HL, Macfarlan TS, et al. 2018. Ten things you should know about transposable elements. *Genome Biol* **19:** 199.

Conlon EG, Fagegaltier D, Agius P, Davis-Porada J, Gregory J, Hubbard I, Kang K, Kim D, The New York Genome Center ALS Consortium, Phatnani H, et al. 2018. Unexpected similarities between C9ORF72 and sporadic forms of ALS/FTD suggest a common disease mechanism. *Elife* **7:** e37754. doi:10.7554/eLife.37754.001

Ho YJ, Anaparthy N, Molik D, Mathew G, Aicher T, Patel A, Hicks J, Hammell MG. 2018. Single-cell RNA-seq analysis identifies markers of resistance to targeted BRAF inhibitors in melanoma cell populations. *Genome Res* **28:** 1353–1363.

Hwang DW, Jaganathan A, Shrestha P, Jin Y, El-Amine N, Wang SH, Hammell M, Mills AA. 2018. Chromatin-mediated translational control is essential for neural cell fate specification. *Life Sci Alliance* **1:** e201700016.

Jin Y, Hammell M. 2018. Analysis of RNA-seq data using TEtranscripts. *Methods Mol Biol* **1751:** 153–167.

Nicolas A, Kenna KP, Renton AE, Ticozzi N, Faghri F, Chia R, Dominov JA, Kenna BJ, Nalls MA, Keagle P, et al. 2018. Genome-wide analyses identify *KIF5A* as a novel ALS gene. *Neuron* **97:** 1268–1283.

O'Neill K, Liao WW, Patel A, Hammell MG. 2018. TEsmall identifies small RNAs associated with targeted inhibitor resistance in melanoma. *Front Genet* **9:** 461.

GENETIC VARIANTS LINKED TO AUTISM TRAITS

I. Iossifov Y-h. Lee A. Munoz
S. Marks B. Yamrom

Overview

In 2018, my lab included the following members: Adriana Munoz, Boris Yamrom, Steven Marks, and Yoon-ha Lee. The bulk of our work was in analyzing the large data set of whole-genome sequencing (WGS) data generated from approximately 2400 of the Simons Simplex Collection (SSC) families and approximately 900 families from the Autism Genetic Resource Exchange (AGRE), a collection of families that have multiple children with autism. We also started the analysis of the whole-exome sequencing (WES) data from the growing SPARK collection: data for approximately 5,000 of the SPARK families have been released, and SPARK is expected to grow to around 50,000 families in a couple of years. In addition, we started a pilot project to explore the potential of RNA sequencing (RNA-Seq) in family collections like SSC. These data are a rich resource that we use in numerous projects.

Below, I list the abstracts of four projects that are in submission or near-submission status and that show our current efforts in studying the role of de novo noncoding variants, rare structural rearrangements, and common variants in autism's etiology.

De Novo Disruption of Introns Contributes to Autism

A. Munoz, B. Yamrom, Y-h. Lee, S. Marks, I. Iossifov
[in collaboration with S. Yoon, P. Andrews, Z. Wang, D. Levy, and M. Wigler, CSHL; C. Reeves and L. Winterkorn, New York Genome Center; A.M. Krieger and A. Buja, University of Pennsylvania; K. Pradhan, Albert Einstein College of Medicine, Montefiore Medical Center; K.K. Baldwin, Scripps Research Institute]

De novo mutation contributes to the incidence of autism, as it does for many other developmental disorders. Most of the previous single-gene de novo mutations were found by WES and, hence, were restricted to coding mutations. Here, we report on findings from the whole-genome sequence of both simplex families from the SSC and multiplex families from the AGRE. Using the SSC, we find de novo disruptions within introns contribute significantly (p value ~0.0006) to simplex autism as evidenced by increased incidence of both short indels and larger copy number variations (CNVs). We estimate this contribution to be on the order of one-third the contribution from coding mutations and similar in magnitude to the recently reported incidence of promoter mutations. On studying multiplex families from the AGRE, we find that many samples have excessive genomic drift in culture. After removing those samples, we find no statistical evidence that de novo mutation contributes to multiplex autism. These results support the prediction of our unified hypothesis of autism that de novo mutations of high penetrance contribute to simplex but not multiplex autism.

Damaging De Novo Mutations Diminish Motor Skills in Children on the Autism Spectrum

I. Iossifov [in collaboration with A. Buja and A.M. Krieger, University of Pennsylvania; C. Lord, Weill Cornell Medical College; N. Volfovsky and A.E. Lash, Simons Foundation; M. Wigler, CSHL]

In individuals with autism spectrum disorder (ASD), de novo mutations have previously been shown to be significantly correlated with lower IQ, but not with the core characteristics of ASD: deficits in social communication and interaction and restricted interests and repetitive patterns of behavior. We extend these findings by showing in the SSC that damaging de novo mutations in ASD individuals are also significantly and convincingly correlated with measures of impaired motor skills. This correlation is not explained

by a correlation between IQ and motor skills. We find that IQ and motor skills are distinctly associated with damaging mutations, and, in particular, motor skills are a more sensitive indicator of mutational severity than is IQ, as judged by mutational type and target gene. We use this finding to propose a combined classification of phenotypic severity: mild (little impairment of either), moderate (impairment mainly to motor skills), and severe (impairment of both IQ and motor skills).

MUMdex: MUM-Based Structural Variation Detection

I. Iossifov [in collaboration with S. Marks, P.A. Andrews, J. Kendall, Z. Wang, D. Levy, and M. Wigler, CSHL; L. Muthuswam, New York Genome Center]

Standard genome sequence alignment tools primarily designed to find one alignment per read have difficulty detecting inversion, translocation, and large insertion and deletion events. Moreover, dedicated split-read alignment methods that depend only on the reference genome may misidentify or find too many potential split-read alignments because of flaws in the reference genome.

We introduce MUMdex, a maximal unique match (MUM)-based genomic analysis software package consisting of a sequence aligner to the reference genome, a storage-indexing format, and analysis software. Discordant reference alignments of MUMs are especially suitable for identifying inversion, translocation, and large indel differences in unique regions. Extracted population databases are used as filters for flaws in the reference genome. We describe the concepts underlying MUM-based analysis, the software implementation, and its usage.

We show via simulation that the MUMdex aligner and alignment format are able to correctly detect and record genomic events. We characterize alignment performance and output file sizes for human whole-genome data and compare with Bowtie 2 and the BAM format. Preliminary results show the practicality of the analysis approach by detecting de novo mutation candidates in human whole-genome DNA sequence data from 510 families. We provide a population database of events from these families for use by others.

A Platform for Access and Analysis of Genetic Variants in Phenotype-Rich Family Collections

Y-h. Lee, B. Yamrom, S. Marks, I. Iossifov [in collaboration with M. Cokol, Axcella, Boston; A. Nenkova, University of Pennsylvania; L. Chorbadjiev, SeqPipe Ltd., Sofia]

WES, a technique that enables the inexpensive identification of genetic variants in the gene-encoding regions of the genomes of thousands of people, is quickly transforming human genetics. Particularly successful are the numerous studies that used WES in large collections of families to study the genetic architectures of human disorders with strong detrimental effect in fecundity, including autism, intellectual disability, schizophrenia, epilepsy, and congenital heart disease. These studies identified large numbers of genetic variants segregating in the families or arising de novo in children, gathered detailed phenotypic measurements of the studied individuals, and used the complex data sets to develop models of genotype and phenotype relationships.

There is an enormous amount of work that needs to follow the early success in the genetics of such complex disorders in order to develop effective treatment and early diagnostic strategies. A variety of future research projects will study in detail the effects of hundreds of genetic variants and genes at molecular, cellular, and organismic levels. Such projects will greatly benefit from the accumulated family WES data sets, but their large size and complex structure create a major obstacle for their efficient use. Here we present the GPF (genotype and phenotype in families) system, which manages such data sets and has an intuitive interface that makes it possible for the wider scientific community to benefit from the new collections.

RNA-Seq of SSC

We are finalizing our analysis of the whole-genome data from approximately 2,400 of the SSC families. The major result of that effort is the estimate of the contribution of the de novo noncoding variants. Specifically, we observed a significantly increased rate of de novo intronic variants in affected children compared with their unaffected siblings when we restrict the rate observation to the autism genes previously

implicated by WES. The increase in the rate is consistent with a contribution of ~5% of de novo intronic indels to the autism diagnosis in SSC (see the "De novo Disruption of Introns Contributes to Autism" section above). We do not observe a similar increase in the rate of de novo intronic substitutions, but it is expected that the size of the study is insufficient to detect that signal given the much higher rate of background noise for substitutions. Nevertheless, we also expect that de novo intronic substitutions have a contribution, and we guess that is likely of similar magnitude to the contribution of the de novo intronic indels. As others have reported an increased rate of de novo mutation in affected versus unaffected children within the control regions of the intergenic space, we expect that the contribution of noncoding de novo mutation is close to 15%, perhaps only slightly less than the contribution from de novo coding mutation.

Despite the large contribution of the noncoding variants, we have no good purely analytic method to distinguish the specific causal sequence variation from the many random ones. We proposed to address that through study of the RNA. We expect that for the majority of the causal de novo noncoding variants, the immediate effect would be on the expression of nearby genes, and such changes in expression can be detected through RNA-Seq by comparing the expression of the affected gene allele to that of the unaffected allele, a method called allele-specific expression (ASE). In the last year we initiated pilot experiments to test the feasibility of this approach. This is a collaborative effort, including several groups at CSHL (Wigler and Levy labs), several groups at the New York Genome Center (Hemali Phatnani's and Tuuli Lappalainen's groups and Tom Maniatis), and Kristen Baldwin from the Scripps Research Institute.

We have access to Epstein–Barr virus (EBV) immortalized lymphoblastoids (LCLs) for all the individuals of the SSC. One of our pilots addressed the question of whether LCLs are a good source to study ASE. In collaboration with Kristen Baldwin, we transformed an LCL into an induced pluripotent stem cell (iPSC) and further down neuronal lineages. We then generated and sequenced RNA libraries from the original LCLs, the iPSCs, and the derived neurons. The analysis of these data is ongoing, but we have already made a few useful observations. First, the transformation process worked successfully. Second, the derived neurons express nearly 90% of the autism genes identified by exome sequencing, whereas the LCLs express ~70% of these genes. Third, when a gene is expressed in both cell types, the ASE is preserved.

In an additional pilot, we performed RNA-Seq from LCLs of six of the SSC families—two of which had identical affected twins and an unaffected sibling, and four of which had one affected and one unaffected child. Among the goals of this pilot were tuning our bench protocols and analytical tools and estimating the noise sources (like *trans*-regulatory and epigenetic effects) that would decrease the power of detecting ASE. It appears from the pilot that such noise sources are not negligible but are manageable.

PUBLICATIONS

Buja A, Volfovsky N, Krieger AM, Lord C, Lash AE, Wigler M, Iossifov, I. 2018. Damaging de novo mutations diminish motor skills in children on the autism spectrum. *Proc Natl Acad Sci* **115:** E1859–E1866.

Castel SE, Cervera A, Mohammadi P, Aguet F, Reverter F, Wolman A, Guigo R, Iossifov I, Vasileva A, Lappalainen T. 2018. Modified penetrance of coding variants by *cis*-regulatory variation contributes to disease risk. *Nat Genet* **50:** 1327–1334.

GENE REGULATION, BIOPHYSICS, AND MACHINE LEARNING

J.B. Kinney A. Ayaz T.L. Forcier A. Tareen
W-C. Chen A. Posfai M.S. Wong

My research career began in theoretical physics, but early in graduate school I was drawn to biology by the immense variety of open problems and by the possibility of testing theoretical ideas with experiments. At first I pursued dry lab research focused on machine learning methods for analyzing large but noisy biological data sets. Then, in my last year of graduate school, I became captivated by the possibility of using ultra-high-throughput DNA sequencing to quantitatively study the biophysical mechanisms of transcriptional regulation. To pursue this vision, I proposed and carried out wet lab experiments that culminated in Sort-Seq, the first massively parallel reporter assay (MPRA) for studies in living cells. As an independent investigator, I have continued pursuing a tightly knit combination of experiment, computation, and theory.

Measuring *cis*-Regulatory Energetics in Living Cells Using Allelic Manifolds

Gene expression in all organisms is controlled by cooperative interactions between DNA-bound transcription factors (TFs), but quantitatively measuring TF–DNA and TF–TF interactions remains difficult. In Forcier et al. (2018), we introduced a strategy for precisely measuring the Gibbs free energy of such interactions in living cells. This strategy centers on the measurement and modeling of "allelic manifolds," a multidimensional generalization of the classical genetics concept of allelic series. Allelic manifolds are measured using reporter assays performed on strategically designed *cis*-regulatory sequences. Quantitative biophysical models are then fit to the resulting data. We used this strategy to study regulation by two *Escherichia coli* TFs, CRP and σ^{70} RNA polymerase. Doing so, we consistently obtained energetic measurements precise to ~0.1 kcal/mol. We also obtained multiple results that deviate from the prior literature. Our strategy is compatible with MPRAs in both prokaryotes and eukaryotes and is expected to be both highly scalable and broadly applicable.

Quantitative Activity Profile and Context Dependence of All Human 5′ Splice Sites

Pre-mRNA splicing is an essential step in the expression of most human genes. Mutations at the 5′ splice site (5′ss) frequently cause defective splicing and disease because of interference with the initial recognition of the exon–intron boundary by U1 small nuclear ribonucleoprotein (snRNP), a component of the spliceosome. In Wong et al. (2018), our lab and Adrian Krainer's lab together developed a massively parallel splicing assay (MPSA) in human cells and used this assay to quantify the activity of all 32,768 unique 5′ss sequences (having the form NNN/GYNNNN) in three different disease-relevant gene contexts. Our results revealed that, although splicing efficiency is mostly governed by the 5′ss sequence, there are substantial differences in this efficiency across gene contexts. Among other uses, these MPSA measurements were found to facilitate the prediction of 5′ss sequence variants that are likely to cause aberrant splicing. This study thus provides a framework to assess potential pathogenic variants in the human genome and streamline the development of splicing-corrective therapies.

Density Estimation on Small Data Sets

How might a smooth probability distribution be estimated, with accurately quantified uncertainty, from a limited amount of sampled data? In Chen et al. (2018), we described a field-theoretic approach that addresses this problem in one dimension remarkably well. In particular, this approach provides an exact nonparametric Bayesian posterior without relying on tunable parameters or large-data approximations. Among other results, we found that non-Gaussian constraints, which require a nonperturbative treatment, often play a major role in reducing distribution uncertainty. In addition to our theoretical findings, we developed a fast and user-friendly software implementation of this method

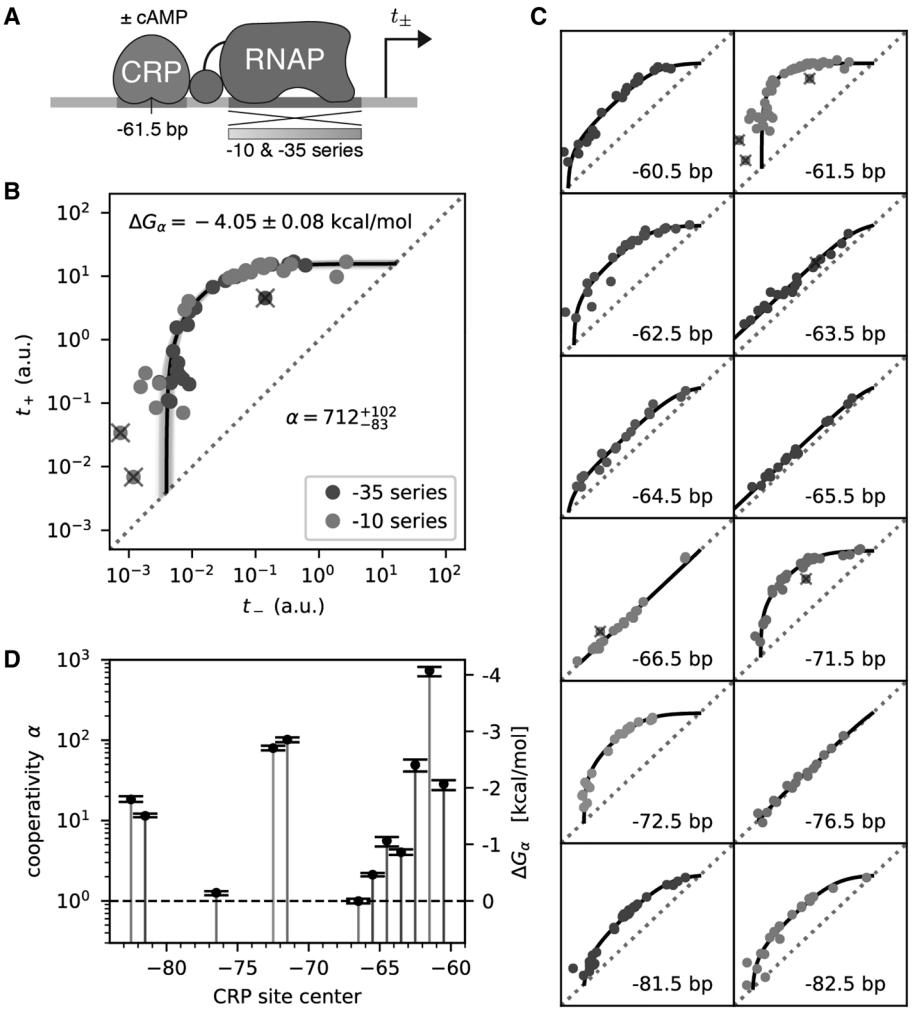


Figure 1. Precision measurement of class I CRP-RNAP interactions. (*A*) t_+ and t_- were measured for promoters containing a CRP binding site centered at –61.5 bp relative to the transcription start site. The RNAP sites of these promoters were mutagenized in either their –10 or –35 regions, generating two allelic series. t_+ and t_- correspond to expression measurements made in the presence and absence, respectively, of cAMP. (*B*) Data obtained for 47 variant promoters having the architecture shown in *A*. Three data points designated as outliers are indicated by Xs. The allelic manifold that best fits the 44 nonoutlier points is shown in black; 100 plausible manifolds, estimated from bootstrap-resampled data points, are shown in gray. The resulting values for α and $\Delta G_\alpha = k_B T \log \alpha$ are also shown, with 68% confidence intervals indicated. (*C*) Allelic manifolds obtained for promoters with CRP binding sites centered at a variety of class I positions. (*D*) Inferred values for the cooperativity factor α and corresponding Gibbs free energy ΔG_α for the 12 different promoter architectures assayed in *C*. Error bars indicate 68% confidence intervals. (Adapted from Fig. 5 of Forcier et al. 2018.)

called Statistics Using Field Theory (SUFTware). The Kinney lab is currently adapting this approach for the analysis of survival data. Our hope is that, by enabling the inference of smooth survival curves as well as estimates of their uncertainty without requiring large-data approximations, we will be able to increase the statistical power of both preclinical studies and small clinical trials.

PUBLICATIONS

Belliveau NM, Barnes SL, Ireland WT, Jones DJ, Sweredoski MJ, Moradian A, Hess S, Kinney JB, Phillips R. 2018. Systematic approach for dissecting the molecular mechanisms of transcriptional regulation in bacteria. *Proc Natl Acad Sci* **115:** E4796–E4805. doi:10.1073/pnas.1722055115

Chen W, Tareen A, Kinney JB. 2018. Density estimation on small data sets. *Phys Rev Lett* **121:** 160605. doi:10.1103/PhysRevLett.121.160605

Forcier T, Ayaz A, Gill MS, Jones D, Phillips R, Kinney JB. 2018. Measuring *cis*-regulatory energetics in living cells using allelic manifolds. *eLife* **7:** e40618. doi:10.7554/eLife.40618

Lu B, Klingbeil O, Tarumoto Y, Somerville TDD, Huang Y-H, Wei Y, Wai DC, Low JKK, Milazzo JP, Wu XS, et al. 2018. A transcription factor addiction in leukemia imposed by the MLL promoter sequence. *Cancer Cell* **34:** 970–981. doi:10.1016/j.ccell.2018.10.015

Posfai A, Zhou J, Plotkin JB, Kinney JB, McCandlish DM. 2018. Selection for protein stability enriches for epistatic interactions. *Genes* **9:** E423. doi:10.3390/genes9090423

Sachan N, Miller M, Shirole NH, Senturk S, Corbo V, Kinney JB, Sordella R. 2018. Inducible CRISPR-based genome editing for the characterization of cancer genes. In *Genome editing and engineering: from TALENs, ZFNs and CRISPRs to molecular surgery* (ed. Appasani K, Church GM), p. 337. Cambridge University Press, Cambridge.

Wong MS, Kinney JB, Krainer AR. 2018. Quantitative activity profile and context dependence of all human 5′ splice sites. *Mol Cell* **71:** 1012–1026.e3. doi:10.1016/j.molcel.2018.07.033

In Press

Adams RM, Kinney JB, Walczak AM, Mora T. 2019. Epistasis in a fitness landscape defined by antibody-antigen binding free energy. *Cell Syst* **8:** 86–93.e3. doi:10.1016/j.cels.2018.12.004

Barnes SL, Belliveau NM, Ireland WT, Kinney JB, Phillips R. 2019. Mapping DNA sequence to transcription factor binding energy in vivo. *PLoS Comp Biol* **15:** e1006226. doi:10.1037/journal.pcbi.1006226

Kinney JB, McCandlish DM. 2019. Massively parallel assays and quantitative sequence–function relationships. *Annu Rev Genom Hum Genet* doi:10.1146/annurev-genom-083118-014845

Tareen A, Kinney JB. 2019. Logomaker: beautiful sequence logos in Python. bioRxiv doi:10.1101/635029

COMPUTATIONAL GENOMICS OF CANCER

A. Krasnitz B. Becerra P. Belleau N. Ranade

Research in our group is focused on in silico cancer genomics. In the last several years there has been explosive growth in the volume as well as quality, variety, and detail of cancer-related genomic data available for analysis. This development came about through rapid accumulation of public data sets generated by individual laboratories worldwide, through large-scale cancer genome projects such as The Cancer Genome Atlas (TCGA), and with the advent of new experimental methodologies, especially next-generation sequencing and single-cell genomics. We see our goal as channeling this flood of data into a number of clinically relevant applications. These include discovery of genomic markers for clinical outcome and molecular classification of cancer; elucidating the clonal structure of tumors and its relation to progression, invasion, metastasis, and response to treatment; and pinpointing and prioritizing targets for functional analysis. Our work is done in close coordination with experimental studies performed by the Wigler, Tuveson, Spector, Stillman, and Fearon laboratories at CSHL.

Computational Framework for Single-Cell Genomics

Investigation of single-cell genomes and transcriptomes is the focus of a massive research effort worldwide. In application to cancer, this line of research has revealed the genomic complexity of the disease and the presence of multiple genealogically related cell populations in a tumor. Detailed knowledge of the clonal structure of a cancer potentially is of high clinical value: Multiplicity of clones or of lesions in most advanced clones is a possible measure of progression; spatial pattern of clone dispersal in a tumor may signal elevated propensity to invade; and lesions observed in individual clones but not in the bulk tissue may point to targets for therapy. DNA copy number profiling of cells from sparse sequencing is an accurate, economically feasible technological approach to the study of cancer subpopulation structure. Novel multiplex sequencing techniques, developed, among others, by the Wigler lab at CSHL, permit simultaneous sequencing of hundreds of single-cell DNA specimens and their subsequent copy number profiling at up to 50 kb resolution. Optimal use of this data form for robust reconstruction of cancer cell phylogenies is a challenging computational problem calling for new and robust informatic and statistical tools.

We responded to this challenge by developing a computational pipeline for single-cell genomics. The pipeline, which has now been released for public use, comprises two major modules—one for deriving integer-valued copy number profiles of individual cells and one for establishing genealogical relations among the cells in a sample and identification of clones. The input into the pipeline consists of cell-specific sets of sequencing reads. These are first aligned with the genome. The sequencing read density as a function of genomic position is then used to derive integer-valued DNA copy number profiles for each cell. Prior to any further processing, these profiles are examined for evidence of extensive DNA damage or degradation, and profiles derived from damaged DNA are filtered out. Next, each of the remaining profiles is reduced to a set of copy number change points in which, for each change point, the genomic interval of its likely location and the sign of the change are specified.

From this point on, we combine change-point reduced copy number profiles of the cells constituting the sample, with the ultimate goal of reconstructing their phylogeny. To this end, we first derive a minimal set of features to account for all the change points observed in the sample and construct a table to indicate, for each cell genome, the presence or absence of each feature. Pairwise similarity of cell genomes is quantified based on the number of features shared by the pair. We then examine the resultant similarity matrix for the presence of clones (i.e., groups of cells whose high number of shared features is far in excess of what is expected at random and can only be explained by their descent from a recent

common ancestor). Such clones may in turn contain subclones of even more closely related cells.

Single-cell data as interpreted by the pipeline must be presented to cancer biologists and clinicians in a comprehensible way in order to have impact on clinical outcome. In practice, this means that the data should be viewable in an organized way, with the viewer retaining the ability to change the organization. The end user must be able to see and navigate the phylogenetic organization. Moreover, the data and their interpretation must be viewed in the context of other parameters, such as anatomical sites along with their pathological assessment. Such examination of the data in aggregate offers the best chance to reveal the critical properties of the sample relevant to clinical assessment. An integrated user interface, capable of handling specialized forms of data that arise in single-cell research, is therefore essential. With this necessity in mind, we have created and made public the Single-Cell Genome Viewer (SCGV). The viewer displays multiple single-cell genomic profiles in the chromosomal order, with colors encoding the sign and magnitude of copy number aberrations. The cells on display are ordered as leaves on a tree, reflecting the structure of the population being sampled. In particular, clonal and subclonal identities of the cells are indicated in the appropriate tracks. Other tracks provide information on the cell ploidy, the quality of the cell genomic data, and, importantly, the anatomic origin of the cell. Elements of the data can be examined in greater detail by a combination of selection and zooming in. For example, any number of cells can be selected and the corresponding copy number profiles plotted in a separate interface. A link to the UCSC genome browser is available for any genomic region displayed by the viewer to help put the findings in a broader genomic context.

In developing our single-cell genomics software, we took great care to ensure its user friendliness and thereby increase its value to the research community. Both major modules of the single-cell genomics pipeline and SCGV can be easily installed, with any major operating system, using Anaconda, a popular software distribution and management platform. Future releases of our tools will include parallel computing options for enhanced performance and further improvements in portability achieved by use of container technology such as Docker™.

Organoid Models of Pancreatic Cancer: Faithful and Predictive

Pancreatic cancer is one of the least curable malignancies, often diagnosed at an advanced stage and with <30% of patients surviving the first year following diagnosis. Nearly all patients require pharmacological treatment, most commonly chemotherapy. Given the unfavorable prognosis, and the narrow time window it leaves for therapeutic intervention, it is vital that pharmacological regimen be optimized individually for each patient, accounting for the molecular profile of the disease.

With the ultimate goal of improving patient stratification, we collaborated with the Tuveson laboratory at CSHL, which recently has achieved a major breakthrough in the ability to consistently generate realistic, three-dimensional in vitro models of pancreatic cancer, termed patient-derived organoids (PDOs). These cultures can be grown either from surgical specimens or from fine-needle tissue biopsies and both from primary tumors and metastatic lesions of pancreatic ductal adenocarcinoma (PDAC). Once grown, PDOs provide a platform for massively parallel screening of pharmacological agents. The goal of our collaboration was to examine the correspondence between the molecular properties of the PDOs and their PDAC tissues of origin and to determine the relevance of drug responses in the PDOs to those observed in the clinic of PDAC.

To this end, we performed extensive molecular profiling of the PDO cultures in the library, including whole-exome sequencing (WES), RNA-Seq, and deep DNA sequencing for a panel of 500 PDAC-associated genes. To examine how faithfully PDOs reflect the somatic mutation spectrum of the tumor of origin, we performed whole-genome sequencing (WGS) of the PDO, the matching primary PDAC tissue from a surgical specimen, and matching normal tissue from the donor for 20 patient cases. We observed a high degree of concordance for all classes of somatic variants (substitutions, insertions, deletions) occurring at high variant allele frequencies (VAFs) in the PDOs. At the same time, the PDOs made it possible to discover variants at low VAFs, rendered undetectable by the low neoplastic cellularity in the primary tumors. A broader examination of all genomic data from the PDO library demonstrated the presence in the PDO genomes of mutations in the same driver genes, and at highly similar overall frequencies, as has been

previously observed in PDAC patient cohorts. An important example is that of *KRAS*, which was found mutated in 96% of PDOs in the library—highly similar to the rate of >90% found in PDAC patient data.

We further determined that PDO cultures recapitulate molecular phenotypes of PDAC as found in the human hosts. In particular, most PDOs in our library can be assigned, with a high degree of confidence, either to the previously described classical or basal molecular subtypes based on their mRNA expression profiles. Accordingly, unsupervised clustering of PDO cultures by mRNA expression robustly produces a partition into two classes—one of which is dominated by the classical and the other by the basal profiles. Thus, PDO cultures retain phenotypic diversity of the disease in the human hosts, unlike PDAC cell lines, which overwhelmingly are basal.

To better understand how organoid sensitivities to chemotherapeutic agents correspond to patient sensitivities, we conducted a screen of PDO responses to five standard-of-care cytotoxic agents for PDAC, using the area under the drug response curve (AUC) as a measure of response. Next, we addressed the questions of (a) whether transcriptional profiles of PDOs are predictive of the therapeutic responses observed in this screen and (b) whether these transcriptional signatures of response are applicable to PDAC patient outcomes. To this end, we identified, separately for each of the five agents, a small subset of genes whose mRNA expressions most strongly correlated with the AUC values for the agent in the PDO library. For each PDO, we then summed the z-scores of the genes in the drug-specific subset, to form a single predictive molecular sensitivity score (MSS) for each PDO's sensitivity to each agent. We found these predictive scores to be strongly correlated with the corresponding AUC values.

To see how well our PDO-derived MSSs would apply to human patients, we used the gemcitabine-sensitivity genes defined in the PDOs to compute MSSs from 95 patients who received adjuvant treatment containing gemcitabine, one of the five agents in the PDO therapeutic screen. For this analysis, we used RNA-Seq data from resected PDAC specimens. The 50% of the patients with the highest MSS values experienced significantly longer median PFS than the low-scoring 50% of the cohort (608 vs. 442 days; Cox regression, $p = 0.046$). Moreover, the MSS values were not associated with survival benefit in the 38 patients who received no adjuvant treatment, demonstrating that the score is predictive only in the setting of the chemotherapy used to calculate it.

Finally, we examined the MSS performance using molecular and clinical data from the COMPASS trial, for which patients with advanced, unresectable PDACs were recruited, and each was assigned a chemotherapy regimen (m-FOLFIRINOX or gemcitabine-paclitaxel). We focused on the 41 patients in the trial who received FOLFIRINOX. We calculated MSS values for oxaliplatin, a component of FOLFIRINOX, for each patient. Patient MSS values were then compared with the corresponding tumor responses to treatment as measured by changes in the tumor volume. Patients whose MSS predicted they would be sensitive to oxaliplatin exhibited significantly better response to FOLFIRINOX than those deemed nonsensitive ($r = -0.4$, $p = 0.008$). In summary, our results indicate the prognostic utility of PDO-derived MSSs for cytotoxic treatment outcomes in patients with advanced, as well as with localized, PDAC.

PanCanAtlas: Stemness and Immune Response across Multiple Tumor Types

TCGA is a massive, comprehensive data repository for cancer research. More than 30 cancer types are currently represented in the atlas, with hundreds of patient cases per type. For each tumor somatic mutation, DNA copy number variation, DNA methylation, and mRNA and micro-RNA profiles are available, along with clinical annotation for the patient. By combining multiple data sets, TCGA consortium generated integrated molecular portraits of a number of cancer types. PanCanAtlas, a successor project to TCGA, sought to take data integration one step further and across multiple cancer types. Our group joined this effort in June 2015, with two focal areas of interest. The first of these is concerned with pan-cancer characterization of immune response to tumors with a particular focus on the role of cancer-specific antigens, including neoantigens and cancer testis antigens (CTAs) (i.e., products of genes with amino acid sequence altered by the disease and those that are never presented to the immune system in normal tissues). Expression of such genes in tumors elicits an immune response—in particular, by cytotoxic T cells—and therefore occupies a central place in the immune landscape of cancer. As part of cancer-specific antigen identification, we collaborated with Seven Bridges Genomics to determine HLA class I

A, B, and C alleles for nearly 9,000 TCGA patients with available RNA sequencing data.

We also were interested in quantifying the presence of stem-cell-like cell populations in tumors. In particular, loss of differentiation has long been recognized as a key feature of cancer pathology. On a cellular level, this loss of differentiation is accompanied by acquisition of stem-cell-like properties. In an analysis spanning multiple cancer types, we found that genes associated with loss of differentiation in tumors frequently occur in published transcriptional signatures of embryonic stem cells.

PUBLICATIONS

Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omberg L, Gevaert O, et al. 2018. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* **173:** 338–354.e15.

Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, et al. 2018. The immune landscape of cancer. *Immunity* **48:** 812–830.e14.

Tiriac H, Belleau P, Engle DD, Plenker D, Deschênes A, Somerville TDD, Froeling FEM, Burkhart RA, Denroche RE, Jang GH, et al. 2018. Organoid profiling identifies common responses to chemotherapy in pancreatic cancer. *Cancer Discov* **8:** 1112–1129.

COMPUTATIONAL GENETICS

D. Levy A. Moffitt C. Wunderlich

The data-rich environment at CSHL generates a wealth of opportunities for the application of mathematics and computation to further our understanding of biology and genetics. The primary activities of our lab are algorithm and protocol development, data analysis, and genetic theory, with a focus on sporadic human diseases such as autism, congenital heart disease, and cancer.

Quantitative-Sensitive Detection

Measuring genomic variants at very low frequency is important in many applications, but especially relevant in measuring residual disease in cancer. Present-day short-read sequencers generate hundreds of millions of high-quality sequence reads with error rates of <1 per 100 bases. These errors determine the lower limits of variant detection: Using standard sequencing, it is impossible to distinguish a variant frequency of <1% from machine error. We would like to measure variants at frequencies of 0.0001% or one part in a million.

Together with Michael Wigler, Zihua Wang, and Andrea Moffitt, we developed a protocol that adds a unique sequence identifier (or "varietal tag") to the initial template molecule. Because sequence error is sparse and independent of the template molecule, reads with the same varietal tag are unlikely to have the same sequence errors. Taking a sequence consensus from all reads with the same tag corrects for sparse error. We found that systematic errors have a sequence context–specific signature, and by analyzing multiple high-depth tagged experiments, we are able to model the error of consensus reads. Some sequence contexts (such as CpG) have a high background rate that has nothing to do with machine error. However, other contexts are very stable, with background error rates for consensus sequences lower than one part in a million. The MASQ (multiplex accurate sensitive quantitation) protocol enables the simultaneous measurement of variants from up to 50 loci and millions of templates per locus. We developed an informatics pipeline that selects variants with low background error rates, determines the optimal reagents for the protocol, and designs the necessary sequencing primers.

With Mona Spector, Alex Krasnitz, and Joan Alexander, we have applied this method to measure residual disease in patients treated for acute myeloid leukemia (AML). Using MASQ, we measured tumor-specific variants in blood samples taken after treatment. In some patients, the residual load was detectable at levels observable by traditional sequencing. In other patients, we were able to measure rates as low as 1 part in 100,000.

We are presently testing the utility of MASQ in the context of solid tumors, in which circulating tumor cells and cell-free tumor DNA provide targets for detection. Analyzing DNA in the cell-free component presents a unique challenge with DNA fragments that are short and few in number.

Mutational Sequencing

The latest third-generation sequencing platforms like PacBio and Oxford Nanopore generate long-read information that is important in high-quality genome assemblies. Long reads are especially useful when the genome studied is diploid and heterozygous variants are too far apart to phase by short reads. Unfortunately, compared to the present generation of short-read sequencers, these long-read platforms are expensive and error-prone.

We previously proposed a method for turning short-read sequencers into virtual long-read sequencers by embedding a unique molecular identity throughout each template molecule by random mutation. Theoretical computations suggested that we could then count templates by counting unique mutation patterns and that we could assemble very long templates by connecting reads with overlapping mutation patterns.

Recently, we implemented this idea in practice, using incomplete bisulfite conversion as the mechanism for introducing mutations. Loosely speaking, sodium bisulfite converts a C to T in a DNA template,

and, by tuning our rate of conversion to 50%, we label each template molecule with a unique and dense mutational signature of C-to-T conversions. Clustering reads with the same conversion pattern enabled accurate count and long-range assembly of initial template molecules from short-read sequence data. Using a PstI representation, we demonstrated that muSeq improves copy number measurement and significantly reduces sporadic sequencing error. Using a cDNA library, we demonstrated long-range assembly of template molecules up to 4 kb in length.

However, both counting and cDNA assembly required first mapping reads to a reference genome. This limits the utility of muSeq to well-sequencing organisms, and even in those cases, we would like to avoid reference bias when determining insertion/deletion polymorphism, splice junctions, and assembling complex genomic regions rich in polymorphism, such as the HLA locus.

For these reasons, we are liberating muSeq from the reference genome. Together with Siran Li, we developed a protocol and informatics for targeted and phased de novo muSeq assembly. At present, we can generate haplotype-level assemblies for target regions up to 5 kb in length. In place of mapping to a reference, the mutated template molecules are assembled using a De Bruijn graph. We developed custom assembly methods to augment this graph with read-pair information, resulting in many mutated template assemblies that span the full length of the target region. In lieu of a reference genome, we use an unconverted sequence library to "correct" mutations in the template assemblies. We then partition the corrected templates into two (or more) haplotypes. Finally, we compute a consensus over each haplotype, correcting any residual mutations and PCR errors.

The present protocol and informatics can assemble hundreds of mutated template molecules and generate high-quality haplotype-phased assemblies for regions up to 5 kb. The protocol is simple, robust, and, requiring fewer than a million Illumina reads, very cheap. Although changes in technology for long-read sequencers may alter the equation, at present, de novo muSeq offers an accessible and inexpensive alternative for targeted applications.

PUBLICATIONS

Alexander J, Kendall J, Mclndoo J, Rodgers L, Aboukhalil R, Levy D, Stepansky A, Sun G, Chorbardjiev L, Riggs M, et al. 2018. Utility of single-cell genomics in diagnostic evaluation of prostate cancer. *Cancer Res* **78:** 348–358. doi:10.1158/0008-5472.CAN-17-1138

Kumar V, Rosenbaum J, Wang Z, Forcier T, Ronemus M, Wigler M, Levy D. 2018. Partial bisulfite conversion for unique template sequencing. *Nucl Acids Res* **46:** e10. doi:10.1093/nar/gkx1054

PREDICTING EFFECTS OF MUTATIONS FROM HIGH-THROUGHPUT DATA

D. McCandlish A. Posfai J. Zhou

Understanding the relationship between the DNA sequence of an organism's genome and the measurable characteristics of that organism is one of the fundamental goals of biology. Recent progress in high-throughput experimental techniques now allows us to measure the effects on a cellular or molecular level of thousands to millions of changes to the DNA sequence in a single experiment. In the McCandlish lab, we are focused on developing new computational and mathematical techniques for making sense of this wealth of data. Our ultimate goals are to be able to predict the pathogenicity of mutations observed in human genome sequences, understand the evolution of drug resistance and immune escape, and help to construct highly optimized enzymes for biotechnology applications.

An important challenge in predicting the effects of mutations is that the effect of any given mutation may depend on which other mutations are already present, a phenomenon known as genetic interaction or epistasis. Our group is particularly interested in developing techniques to quantify and better understand the form and causes of these genetic interactions, with the dual goals of improving our ability to predict the effects of combinations of mutations and to understand the influence that these interactions have on the process of biological evolution.

Inference of Genetic Interactions

Because of the large number of possible mutations to any given gene and the far larger number of possible combinations of these mutations, the form that genetic interactions can take can be extremely complicated. Nonetheless, because these mutations are embedded within very specific biochemical, cell biological, and developmental mechanisms, we can hope that these mechanisms will confer some type of order or simple structure on the form of the genetic interactions. For instance, it turns out that the effects of mutations that change the amino acid sequence of a protein have approximately additive effects on the free energy of folding for that protein. Moreover, basic results in statistical mechanics show that the probability of the protein being found in a folded state is determined as a specific nonlinear function of the free energy of folding. It has been hypothesized that most protein-coding mutations confer their effects via disruption to protein folding, in which case the interactions between mutations should take a very particular form that is determined by the nonlinear mapping between folding energy and the probability of folding.

To test this hypothesis, in collaboration with Jakub Otwinowski and Joshua Plotkin at the University of Pennsylvania, we have developed a statistical technique for inferring models of genetic interaction that generalizes the form of interaction hypothesized for proteins. The statistical model assumes that mutations have additive effects on some latent underlying trait (e.g., the free energy of folding in the example above), which is transformed by some nonlinear function into the quantity that we actually observe (e.g., the activity of the protein). We then learn optimal parameters for both the effects of mutations on the latent underlying trait and the shape of the nonlinear function. We call these models "global epistasis" models because the interactions in the model are all due to one shared nonlinearity rather than fine-scale interactions between specific pairs of sites. Having applied this model to high-throughput mutagenesis data from several proteins, we find that the model fits well, consistent with the hypothesis that the effects of most mutations occur via their influence on thermodynamic stability.

Despite the strong performance of global epistasis models for most mutations, some of the most interesting mutations, such as those that change substrate specificity, display more complex patterns of interaction that are not well fit by these models. We are therefore also working to develop techniques to analyze these more complex interactions. In our group, Juannan Zhou has been developing a very general class of models to understand these interactions using

a statistical technique known as Gaussian process regression. In the course of developing these models, we have begun a collaboration with Justin Kinney, Adrian Krainer, and their joint postdoc Mandy Wong to apply these methods to data from a high-throughput splicing assay that they recently developed. We are currently working to validate the predictions of this modeling with low-throughput experiments and are investigating possible mechanisms to explain the apparent functionality of certain noncanonical splicing motifs that have been identified via this analysis.

Finally, postdoc Anna Posfai (joint with the Justin Kinney laboratory), who is a mathematician by training, has been working on techniques to allow better interpretation of the results of this modeling. In particular, it is frequently the case in models fit to data from high-throughput experiments that many different sets of model parameters will give identical predictions. She is working on mathematical techniques to resolve this problem in order to unambiguously identify the specific positions in a genetic sequence that have the largest influence and/or strongest interactions.

The Role of Genetic Interactions in Molecular Evolution

Although our current focus is on developing techniques to infer the form of genetic interactions from high-throughput data, we are also interested in the implications of these interactions for other biological processes and, in particular, for understanding patterns of genetic divergence between different species. In a collaboration between Anna Posfai, Juannan Zhou, Joshua Plotkin, Justin Kinney, and myself, we conducted a theoretical investigation into the influence of genetic interactions for thermodynamic stability on patterns of long-term protein evolution. This modeling revealed a surprising result—namely, that even if there is only a small amount of epistasis for thermodynamic stability, this epistatic contribution to folding will typically be essential in evolved proteins. That is, despite the small contribution of epistasis to folding stability overall, the observed protein sequences would nonetheless fail to fold correctly in the absence of these interactions.

Although the project on long-term protein evolution was based on a specific model of genetic interaction, another important question is whether we can understand the influence of arbitrary patterns of genetic interaction on molecular evolution. I have been engaged for some time in developing mathematical techniques to address this problem, and this year published another manuscript in this longer series of articles. In McCandlish (2018), I describe a natural neighborhood structure on the space of possible sequences that is induced by natural selection operating in the presence of genetic interactions. More intuitively, although it is clear that two sequences are similar if they share most of their letters, this does not necessarily mean that it is easy to evolve from one to another because, for example, the intermediate sequences may be nonfunctional. This manuscript defines the set of sequences that are particularly easy to evolve to from any given sequence once the influences of natural selection, mutation, and genetic drift are taken into account and shows that by using this definition, we can provide intuitive explanations for several previously existing formal results.

Influence of Mutational Biases on Molecular Adaptation

For many applications such as predicting immune escape and drug resistance mutations, we are interested not only in which mutations can potentially confer these phenotypic changes, but also which mutations are most likely to contribute to the emergence of resistance in a clinical or natural setting. In this context, an important fact is that different mutations occur at substantially different rates. For instance, in many genomes, including those of mammals and birds, mutations occur at highly elevated rates at sites where a cytosine nucleotide is found immediately 5′ to a guanine nucleotide, known as CpG sites. One might hypothesize that if a mutation at such a site is beneficial it might be particularly likely to contribute to adaptation because of its high rate of occurrence. However, more detailed mathematical analysis shows that this effect will only be observed in some circumstances and not others. In a long-standing collaboration with Arlin Stoltzfus at the National Institute of Standards and Technology, we have been working to document the prevalence of this influence of mutational biases on the outcome of molecular adaptation in nature. This year we collaborated with Jay Storz at the University of Nebraska to show that mutations at CpG sites are highly enriched among the set of mutations that confer increased oxygen affinity to

hemoglobin proteins of high-altitude birds, which is a well-studied system for understanding molecular adaptation. These results add to the growing evidence that the influence of mutational biases on molecular adaptation is in fact a widespread phenomenon.

PUBLICATIONS

McCandlish DM. 2018 Long-term evolution on complex fitness landscapes when mutation is weak. *Heredity* **121:** 449–465.

Otwinowski J, McCandlish DM, Plotkin JB. 2018. Inferring the shape of global epistasis. *Proc Natl Acad Sci* **115:** E7550–E7558.

Posfai A, Zhou J, Plotkin JB, Kinney JB, McCandlish DM. 2018. Selection for protein stability enriches for epistatic interactions. *Genes* **9:** 423.

In Press

Storz JF, Natarajan C, Witt CC, McCandlish DM, Stoltzfus A. 2019. The role of mutation bias in adaptive molecular evolution: insights from convergent changes in protein function. *Philos Trans R Soc Lond B Biol Sci* **374:** 20180238.

COMPUTATIONAL SEQUENCE ANALYSIS

M. Schatz

Our laboratory develops novel computational analysis methods to study the structure and function of genomes, especially genomes with medical or agricultural significance. These include methods for assembling sequence data into complete genomes, aligning sequences to discover variations or measure transcription levels, and methods for mining these data to discover relationships between an organism's genome and its traits. We then apply these methods to study several biological systems, to discover mutations associated with human diseases such as cancer and autism, or to reconstruct the genomes of important agricultural crops and biofuels. For this work, Schatz was named a 2018 Highly Cited Researcher by Clarivate Analytics for publishing multiple highly cited papers that rank in the top 1% by citations in the Web of Science.

A central theme of our current research is to develop novel computational algorithms to analyze the data from new sequencing technologies, especially long-read single-molecule sequencing data from Pacific Biosciences or Oxford Nanopore as well as long-range mapping data from 10X Genomics or Hi-C-based assays. We have found these technologies can reveal substantially more variants in human genomes or other samples compared to short-read alternatives. Here we highlight three demonstrations of this work: (1) a comprehensive review article discussing all of the ways in which these technologies can be used to analyze novel genomes (Sedlazeck et al. 2018b); (2) a description of our long-read mapping and structural variation detection algorithms that can be used to find complex variants (Sedlazeck et al. 2018a); and (3) an application of these methods to study the important SK-BR-3 breast cancer cell line where we detect nearly 20,000 structural variations in the cancer that are undetectable using alternative approaches (Nattestad et al. 2018). This last work was published as the cover story for the August 2018 issue of *Genome Research* and was highlighted with news-and-views articles in both *Nature* and *Science Translational Medicine*. Relatedly, this year we published two reports of using Oxford Nanopore long-read sequencing to rapidly sequence and diagnose clinical infections of Klebsiella (Tamma et al. 2018; Simner et al. 2018).

Other major works this year include the publication of several major genome studies: the publication of the agriculturally important sugarcane genome (Zhang et al. 2018), the parrot genome as a model system for speech and cognition (Wirthlin et al. 2018), the bumblebee genome to support our analysis and understanding of colony collapse disorder (Kent et al. 2018), and an analysis of hybrid fish (Dennenmoser et al., in press). We also coauthored a novel method for identifying de novo mutations from high throughput sequencing data (Gómez-Romera et al. 2018) and coauthored a study on the role of circular RNA in gastric cancer (Wang et al., in press). Finally, we coauthored a publication describing KBase, a systems biology knowledgebase for bioenergy-related genomics research (Arkin et al. 2018).

Piercing the Dark Matter: Bioinformatics of Long-Range Sequencing and Mapping

Several new genomics technologies have become available that offer long-read sequencing or long-range mapping with higher throughput and higher resolution analysis than ever before. These long-range technologies are rapidly advancing the field with improved reference genomes, more comprehensive variant identification, and more complete views of transcriptomes and epigenomes. However, they also require new bioinformatics approaches to take full advantage of their unique characteristics while overcoming their complex errors and modalities. Here, we discuss several of the most important applications of the new technologies, focusing on both the currently available bioinformatics tools and opportunities for future research.

Accurate Detection of Complex Structural Variations Using Single-Molecule Sequencing

Structural variations are the greatest source of genetic variation, but they remain poorly understood

because of technological limitations. Single-molecule long-read sequencing has the potential to dramatically advance the field, although high error rates are a challenge with existing methods. Addressing this need, we introduce open-source methods for long-read alignment (NGMLR; https://github.com/philres/ngmlr) and structural variant identification (Sniffles; https://github.com/fritzsedlazeck/Sniffles) that provide unprecedented sensitivity and precision for variant detection, even in repeat-rich regions and for complex nested events that can have substantial effects on human health. In several long-read data sets, including healthy and cancerous human genomes, we discovered thousands of novel variants and categorized systematic errors in short-read approaches. NGMLR and Sniffles can automatically filter false events and operate on low-coverage data, thereby reducing the high costs that have hindered the application of long reads in clinical and research settings.

Complex Rearrangements and Oncogene Amplifications Revealed by Long-Read DNA and RNA Sequencing of a Breast Cancer Cell Line

The SK-BR-3 cell line is one of the most important models for HER2$^+$ breast cancers, which affect one in five breast cancer patients. SK-BR-3 is known to be highly rearranged, although much of the variation is in complex and repetitive regions that may be underreported. Addressing this, we sequenced SK-BR-3 using long-read single-molecule sequencing from Pacific Biosciences, and developed one of the most detailed maps of structural variations (SVs) in a cancer genome available with nearly 20,000 variants present—most of which were missed by short-read sequencing. Surrounding the important ERBB2 oncogene (also known as HER2), we discovered a complex sequence of nested duplications and translocations, suggesting a punctuated progression. Full-length transcriptome sequencing further revealed several novel gene fusions within the nested genomic variants. Combining long-read genome and transcriptome sequencing enables an in-depth analysis of how SVs disrupt the genome and sheds new light on the complex mechanisms involved in cancer genome evolution.

PUBLICATIONS

Arkin AP, Cottingham RW, Maslov S, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, et al. 2018. KBase: the United States Department of Energy systems biology knowledgebase. *Nat Biotechnol* **36:** 566–569. doi:10.1038/nbt.4163

Gómez-Romera L, Palacios-Flores K, Reyes J, García D, Boege M, Dávila G, Flores M, Schatz, MC Palacios, R. 2018. Precise detection of de novo single nucleotide variants in human genomes. *Proc Natl Acad Sci* **115:** 5516–5521. doi:10.1073/pnas.1802244115

Kent CF, Dey A, Patel H, Tsvetkov N, Tiwari T, MacPhail VJ, Gobeil Y, Harpur BA, Gurtowski J, Schatz MC, et al. 2018. Conservation genomics of the declining North American bumblebee *Bombus terricola* reveals inbreeding and selection on immune genes. *Front Genet* **9:** 316. doi:10.3389/fgene.2018.00316

Nattestad M, Goodwin S, Ng K, Baslan T, Sedlazeck FJ, Rescheneder P, Garvin T, Fang H, Gurtowski J, Hutton E, et al. 2018. Complex rearrangements and oncogene amplifications revealed by long-read DNA and RNA sequencing of a breast cancer cell line. *Gen Res* **28:** 1126–1135. doi:10.1101/gr.231100.117

Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, Schatz, MC. 2018a. Accurate detection of complex structural variations using single molecule sequencing. *Nat Methods* **15:** 461–468. doi:10.1038/s41592-018-0001-7

Sedlazeck FJ, Lee H, Darby CA, Schatz MC. 2018b. Piercing the dark matter: bioinformatics of long-range sequencing and mapping. *Nat Rev Genet* **19:** 329–346. doi:10.1038/s41576-018-0003-4

Simner PJ, Antar AA, Hai S, Gurtowski J, Tamma PD, Rock C, Opene BNA, Tekle T, Carroll KC, Schatz MC, et al. 2018. Antibiotic pressure on the acquisition and loss of antibiotic resistance genes in *Klebsiella pneumoniae*. *J Antimicrob Chemother* **73:** 1796–1803. doi:10.1093/jac/dky121

Tamma PD, Fan Y, Bergman Y, Pertea G, Kazmi A, Lewis S, Carroll KC, Schatz MC, Timp W, Simner PJ. 2018. Applying rapid whole genome sequencing to predict phenotypic antimicrobial susceptibility testing results among carbapenem-resistant *Klebsiella pneumoniae* clinical isolates. *Antimicrob Agents Chemother* **63:** e01923. doi:10.1128/AAC.01923-18

Wirthlin M, Lima NCB, Guedes RLM, Soares AER, Almeida LGP, Cavaleiro NP, Loss de Morais G, Chaves AV, Howard JT, Teixeira MM, et al. 2018. Parrot genomes and the evolution of heightened longevity and cognition. *Curr Biol* **28:** 4001–4008. doi:10.1016/j.cub.2018.10.050

Zhang J, Zhang X, Tang H, Zhang Q, Hua X, Ma X, Zhu F, Jones T, Xhu X, et al. 2018. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nat Genet* **50:** 1565–1593. doi:10.1038/s41588-018-0237-2

In Press

Dennenmoser S, Sedlazeck FJ, Schatz MC, Altmüller J, Zytnicki M, Nolte AW. 2019. Genome-wide patterns of transposon proliferation in an evolutionary young hybrid fish. *Mol Ecol* (in press). doi:10.1111/mec.14969

Wang Z, Ma K, Pitts S, Cheng Y, Liu X, Ke X, Kovaka S, Ashktorab H, Smoot DT, Schatz M, et al. 2019. Novel circular RNA NF1 acts as a molecular sponge, promoting gastric cancer by absorbing miR-16. *Endocr Relat Cancer* (in press). doi:10.1530/ERC-18-0478

POPULATION GENETICS AND TRANSCRIPTIONAL REGULATION

A. Siepel A. Blumberg, N. Dukler, B. Gulko, H. Hijazi, Y. Huang, M. Hubisz, E. Hutton, A. Platts, R. Ramani, A. Xue, Y. Zhao

For the past several years, our research has focused in two major areas: human population genetics and transcriptional regulation in humans and *Drosophila*. The research in population genetics is done either with publicly available genomic sequence data or with a variety of collaborators, whereas most of the work on transcriptional regulation is done with our collaborators Charles Danko and John Lis at Cornell University. We also have smaller collaborative projects on topics ranging from molecular evolution of micro-RNAs in *Drosophila* (with Eric Lai, Memorial Sloan Kettering Cancer Center), to prediction of the fitness consequences of mutations in rice, maize, and other crops (with Michael Purugganan, New York University, and Ed Buckler, Cornell University), to studying the speciation process of recently diverged *Sporophila* songbirds (with John "Irby" Lovette, Cornell University, and Ilan Gronau, Herzliya Interdisciplinary Center, Israel). We focus on theoretical and computational research and do not generate our own data, but we often work closely with experimental collaborators on projects that have substantial experimental as well as computational components. We are broadly interested in molecular evolution, population genetics, and gene regulation, as well as in machine learning, probabilistic modeling, and Bayesian statistics, and our research projects cut a broad swath across these diverse areas. Our research group is highly interdisciplinary, with members trained in computer science, mathematics, physics, genetics, and biochemistry, among other areas. The size of the group is stable at present, with one Ph.D. student having recently graduated (Brad Gulko), a new postdoctoral associate having recently joined (Alexander Xue), and another Ph.D. student having graduated but stayed on as a postdoctoral associate (Noah Dukler).

Below we describe recent progress in three main research areas.

Reconstruction of Demographic History from Complete Genome Sequences

Our research group has a long-standing interest in reconstructing the demographic history of complex, structured populations from DNA sequence data. We developed the first method for inference of human population sizes, divergence times, and gene flow between populations that explicitly models the genealogical relationships among individuals and is efficient enough for genome-wide use. We used this method, *G-PhoCS,* to estimate the origin of one of the earliest branching extant human populations, the San hunter-gatherers of Southern Africa, and more recently in analyses on dogs, wild canids, and six species of *Sporophila* birds. Another method developed by our group, *ARGweaver,* generalizes *G-PhoCS* by capturing the manner in which recombination alters genealogies along the genome sequence. In a joint analysis, we applied *G-PhoCS* and *ARGweaver* to detect significant evidence of gene flow from modern humans into the Altai Neanderthal genome sequence, in the opposite direction and much earlier than previously reported. In addition, we used *ARGweaver* to date this human-to-Neanderthal introgression event at ~100,000 years ago, suggesting an earlier migration of modern humans out of Africa than indicated by most current estimates.

ARGweaver is especially powerful for identifying very early introgression events, yet is currently limited in that it naively assumes a prior distribution based on a single randomly mating population of constant size. Melissa Hubisz has recently implemented a demography-aware version of *ARGweaver* that can sample ancestral recombination graphs (ARGs) under an arbitrary prespecified demographic model, which may include multiple populations, population size changes, and migration events. This approach presents a powerful and flexible way to detect introgressed regions,

with or without samples from the introgressing population. We have been using this method to analyze ancient hominin genomes, including Neanderthals, Denisovans, and other unsequenced ancient hominins (M Hubisz and A Siepel, in prep.). In collaborative studies, we have also applied *ARGweaver* to study introgression and selective sweeps in birds (H Hijazi, L Campagna, I Gronau, et al., in prep.).

In addition, we have continued to improve the *ARGweaver* software, adding many new features since the original publication. *ARGweaver* can now integrate over haplotype phase, thread ancient genomes, read in VCF files, take genotype uncertainty into account, and it now works under an extension of the original sequentially Markov coalescent (SMC) model known as the SMC′. Melissa has also made the powerful but complex *ARGweaver* program more accessible to users of software for population genetic analysis by authoring a detailed user guide, with an accompanying github repository, published as part of a book on *Statistical Population Genetics* (Hubisz and Siepel, in press).

Analysis of Natural Selection on Regulatory Sequences in the Human Genome

We also have a long-standing interest in characterizing the influence of natural selection on DNA sequences, particularly in noncoding regions of the genome. A few years ago, we developed a probabilistic model and inference method, called INSIGHT, which makes use of joint patterns of divergence and polymorphism to shed light on recent natural selection. We have used INSIGHT to show that natural selection has profoundly influenced transcription factor binding sites across the genome during the past five million years of evolution, with major contributions both to adaptive changes in humans and to weakly deleterious variants currently segregating in human populations. Afterward, we realized that the INSIGHT model could also be used to produce "fitness consequences" (fitCons) scores across the entire human genome. Using high-throughput data from the ENCODE project, we partitioned the genome into classes of sites having characteristic functional genomic "fingerprints" in a given cell type and then used INSIGHT to calculate a fitCons score for each fingerprint (Gulko et al., *Nat Genet 47:* 276 [2015]). Finally, we plotted these scores along the genome sequence. These fitCons scores turn out to be remarkably powerful for identifying unannotated regulatory elements in the human genome.

A major shortcoming of fitCons is that it does not scale up for use with large numbers of functional genomic covariates. We developed an alternative approach that bypasses the need for clustering genomic sites and instead assumes a linear-logistic relationship between covariates along the genome and the parameters of the INSIGHT model (Huang et al., *Nat Genet 49:* 618 [2017]). This method, called LINSIGHT, is extremely fast and scalable, enabling it to exploit the "big data" available in modern genomics. We have shown that LINSIGHT outperforms the best available methods in identifying human noncoding variants associated with inherited diseases (including fitCons). In addition, we have applied LINSIGHT to an atlas of human enhancers and shown that the fitness consequences at enhancers depend on cell type, tissue specificity, and constraints at associated promoters.

In parallel, Brad Gulko has recently devised a powerful alternative approach to the fitCons clustering problem, which both scales well and avoids the linearity assumptions of LINSIGHT. His new algorithm, called fitCons2, builds a decision tree by repeatedly splitting classes of genomic sites in a manner that is guaranteed to increase a global measure of the "information" associated with natural selection. This approach allows us to consider dozens of genomic features both individually and in complex combinations. We have now applied fitCons2 to all of the data from Roadmap Epigenomics, considering nine epigenomics features across 115 cell types. A side benefit of this approach is that it allows us to measure the genome-wide "information" about function associated with these epigenomic features (Gulko and Siepel 2018). We found that several epigenomic features yield more information in combination than they do individually. In addition, we found that the entropy in human genetic variation predominantly reflects a balance between mutation and neutral drift. Our cell type–specific FitCons scores reveal relationships among cell types and suggest that around 8% of nucleotide sites are constrained by natural selection.

In addition, Yi-fei Huang in the group has been developing a deep-learning method that both allows for arbitrarily complex relationships among genomic features, and makes use of population genetic theory to estimate allele-specific selection coefficients at every nucleotide in the human genome. This ap-

proach, called linear allele-specific selection inference (LASSIE), unifies methods for deleterious variant prediction with methods for inferring distributions of fitness effects. We applied LASSIE to 51 high-coverage genome sequences annotated with 33 genomic features, and constructed a map of allele-specific selection coefficients across protein-coding sequences in the human genome (Huang and Siepel 2018). This map is generally consistent with previous inferences of bulk distribution of fitness effects, but reveals pervasive weak negative selection against synonymous mutations. In addition, the estimated selection coefficients are highly predictive of inherited pathogenic variants and cancer-driver mutations, outperforming state-of-the-art variant prioritization methods. By constraining our estimated model with ultrahigh coverage ExAC exome-sequencing data, we identified 1,118 genes under unusually strong negative selection, which tend to be exclusively expressed in the central nervous system or associated with autism spectrum disorder, as well as 773 genes under unusually weak selection, which tend to be associated with metabolism. This combination of classical population genetic theory with modern machine-learning and large-scale genomic data is a powerful paradigm for the study of both human evolution and disease.

Transcriptional Regulation and Its Evolution in Primates

For several years, our research program in transcriptional regulation has focused on developing new methods for interpreting the rich nascent RNA sequencing data generated using the powerful global run-on and sequencing (GRO-seq) protocol or its higher-resolution successor, PRO-seq. These methods isolate and sequence newly transcribed RNAs, revealing genome-wide locations of engaged polymerases. It has gradually become clear that an unanticipated benefit of both GRO-seq and PRO-seq is that they are uniquely well suited for detecting so-called enhancer RNAs (or eRNAs), and consequently, for identifying active enhancers and other regulatory elements in mammalian cells.

In our latest work in this area (Danko et al. 2018), we made use of PRO-seq to carry out the first comparative study of nascent transcription in primates. We mapped actively transcribing RNA polymerases in resting and activated $CD4^+$ T cells in multiple human, chimpanzee, and rhesus macaque individuals, with rodents as outgroups. This approach allowed us to directly measure active transcription separately from posttranscriptional processes. We observed general conservation in coding and noncoding transcription, punctuated by numerous differences between species, particularly at distal enhancers and noncoding RNAs. Among other findings, we observed that transcription factor binding sites are a primary determinant of transcriptional differences between species, that rates of evolutionary change are strongly correlated with long-range chromatin interactions, and that adaptive nucleotide substitutions are associated with lineage-specific transcription. We also found that genes regulated by larger numbers of enhancers are more frequently transcribed at evolutionarily stable levels, despite reduced conservation at individual enhancers. At one locus, *SGPP2*, we were able to experimentally confirm that multiple substitutions contribute to human-specific transcription by altering NF-κB binding sites in the human genome. Collectively, our findings suggest a pervasive role for evolutionary compensation across ensembles of enhancers that jointly regulate target genes.

Finally, Noah Dukler has developed a more formal modeling approach to address the problem of identifying gains and losses of regulatory elements across an entire genome. Noah has implemented a probabilistic framework for simultaneously calling peaks in ChIP-seq data across species and modeling the gain and loss of regulatory elements along the branches of a general phylogeny (N Dukler, Y-F Huang, A Siepel, et al., in prep). This approach, called epi-Phylo, combines established statistical methods for calling with a hidden Markov model for identifying gains and losses on a phylogeny. It can be easily generalized to consider any sequence-based readout of epigenomic function and can identify epigenomic conservation as well as turnover.

PUBLICATIONS

Danko CG, Choate LA, Marks BA, Rice EJ, Wang Z, Chu T, Martins AL, Dukler N, Coonrod SA, Tait Wojno ED, et al. 2018. Dynamic evolution of regulatory element ensembles in primate $CD4^+$ T cells. *Nat Ecol Evol* **2:** 537–548.

Fang H, Huang YF, Radhakrishnan A, Siepel A, Lyon GJ, Schatz MC. 2018. Scikit-Ribo enables accurate estimation and robust modeling of translation dynamics at codon resolution. *Cell Syst* **6:** 180–191.

Gulko B, Siepel A. 2019. An evolutionary framework for measuring epigenomic information and estimating cell-type-specific fitness consequences. *Nat Genet* **51:** 335–342.

Huang, YF, Siepel A. 2018. Estimation of allele-specific fitness effects across human protein-coding sequences and implications for disease. bioRxiv doi:10.1101/441337

Mohammed J, Flynt AS, Panzarino AM, Mondal MMH, DeCruz M, Siepel A, Lai EC. 2018. Deep experimental profiling of microRNA diversity, deployment, and evolution across the *Drosophila* genus. *Genome Res* **28:** 52–65.

Ramani R, Krumholz K, Huang YF, Siepel A. 2018. PhastWeb: a web interface for evolutionary conservation scoring of multiple sequence alignments using phastCons and phyloP. *Bioinformatics.* doi:10.1093/bioinformatics/bty966

In Press

Hubisz MJ, Siepel AC. 2019. ARGweaver. In *Statistical population genetics* (ed. Dutheil J), Chap. 11. Springer, New York.

COLD SPRING HARBOR LABORATORY FELLOWS

In 1986, Cold Spring Harbor Laboratory began a Fellows Program to encourage independent research by outstanding young scientists who, during their graduate studies, displayed exceptional promise of becoming leading scientists of the future. The purpose of this program is to provide an opportunity for these young scientists to work independently at the Laboratory for a period of 3 years on projects of their own choosing. Fellows are provided with a salary, research support, and technical assistance so that they can accomplish their goals free of distraction. The interactions among research groups at the Laboratory and the program of courses and meetings on diverse topics in biology contribute to a research environment that is ideal for scientific innovation by these Fellows. The CSHL Fellows Program has been tremendously successful and has served as a paradigm for several analogous programs at other institutions, most recently a Fellows Program sponsored by the National Institutes of Health.

The success of the program is apparent from the list of distinguished alumni. Most notably, Carol Greider—recipient of the 2009 Nobel Prize in Physiology or Medicine for her work on telomerase and telomere function—joined the Fellows Program in 1998. After completing her fellowship, Carol was on the CSHL faculty for 9 years, and she is currently the Daniel Nathans Professor and Director of Molecular Biology and Genetics at Johns Hopkins University School of Medicine.

The first CSHL Fellow, Adrian Krainer (1986), is currently a Professor at the Laboratory, as are Chris Vakoc (2008) and Florin Albeanu (2008), currently holding Professor and Associate Professor positions at CSHL, respectively. Scott Lowe (1995) is a Howard Hughes Medical Institute (HHMI) Investigator. After nearly 15 years on the CSHL faculty, he took on a Professorship at Memorial Sloan Kettering Cancer Center in New York City. Marja Timmermans (1998) was a member of the CSHL faculty for more than 17 years and recently accepted the Humboldt Professorship at the University of Tübingen. Eric Richards (1989) currently is the Vice President of Research and Senior Scientist at the Boyce Thompson Institute for Plant Research at Cornell University; David Barford (1991) is a Fellow of the Royal Society and Professor of Molecular Biology at the Institute of Cancer Research in London; Ueli Grossniklaus (1994) is Professor at the Institute of Plant Biology, University of Zürich, Switzerland; and Térence Strick (2000) left at the end of his fellowship to become a Group Leader at the Institut Jacques Monod in Paris. Lee Henry (2000) joined HHMI's Janelia Farm in Ashburn and joined a project headed by Thomas Südhof. Ira Hall (2004) is an Associate Professor and Associate Director of the Genome Institute at Washington University. Patrick Paddison, who had joined the Fellows Program in 2004, currently is an Associate Member at the Fred Hutchinson Cancer Research Center in Seattle, Washington.

Lingbo Zhang has been a Fellow at the Laboratory since 2013. He joined us from Harvey Lodish's laboratory at the Whitehead Institute of the Massachusetts Institute of Technology, where he studied the regulation of erythroid progenitor cell self-renewal. As a CSHL Fellow, Lingbo is conducting genetic and small-molecule screens to discover novel regulators of normal and aberrant stem cell biology. **Jason Sheltzer** has been a CSHL Fellow since 2015 after completing his graduate work in Angelika Amon's laboratory at MIT. His research focuses on studies of aneuploidy and how it impacts cancer progression. **Semir Beyaz** has been a CSHL Fellow since 2017 after completing his graduate work with Stuart Orkin at Harvard University. He studies how dietary fat intake alters intestinal stem cells, the immune system, and cancer.

ENVIRONMENT–GENE INTERACTION IN CANCER AND IMMUNITY

S. Beyaz C. Chung H. Mou
I. Ergin K. Papciak
O. Eskiocak B. Yueh
M. Fein

Cells respond and adapt to the signals that they receive from their environment. Environmental factors such as nutrients affect cellular states by altering cell state–specific gene expression or metabolic programs. Established in early 2018, the Beyaz laboratory investigates the causal cellular and molecular mechanisms that link nutrition to organismal health and disease. For example, diets that lead to obesity, such as high-fat diets, are significant environmental risk factors that influence cancer incidence and progression in several tissues. Our studies interrogate the functional consequences of diet and obesity on cancer initiation, progression, and response to therapy. We dissect tumor-intrinsic and tumor-extrinsic mechanisms to uncover mechanistic links that can be therapeutically exploited for the treatment of cancer. The current focus in the laboratory is to decipher the causal molecular and cellular mechanisms that link obesity to increased risk of cancer in two different organs: intestine and uterus. We study how dietary and metabolic alterations influence cancer risk through their effects on stem cell biology and immunity in the intestine and uterus by interrogating changes in gene expression, epigenetic state, metabolism, and microbiome. We also develop innovative models to determine the significance of interactions between diverse cell types, including stem cells and immune cells within these organs, in modulating cancer risk. We ultimately aim to build a comparative blueprint of cancer risk in all relevant tissue types in response to diet and obesity and develop a comprehensive understanding of environment–gene interactions in cancer and immunity.

Dietary Regulation of Stem Cell Biology and Cancer

Diets that lead to obesity, such as a high-fat diet (HFD), are considered significant environmental risk factors that influence cancer incidence in several tissues, including intestine. However, the causal molecular and cellular mechanisms through which a pro-obesity HFD affects intestinal tumorigenesis are not well understood. $Lgr5^+$ intestinal stem cells (ISCs) drive the rapid renewal of the intestinal epithelium and remodel intestinal composition in response to diet-induced cues. Importantly, lineage tracing studies demonstrated that ISCs serve as the cell of origin for intestinal cancers. We previously reported that a pro-obesity lard-based HFD increases cancer incidence in the intestine, in part through enhancing ISC and progenitor cell function by activating the lipid-sensing transcription factor PPAR-d, inducing in them a premalignant state with increased vulnerability to undergoing oncogenic transformation. This study revealed a causal stem cell-intrinsic mechanism that links diet-induced obesity to increased intestinal tumor formation.

In the past year, to explore how diverse dietary fatty acids influence ISC activity in an unbiased manner, we developed an ex vivo fatty acid screening assay in intestinal organoids, three-dimensional (3D) epithelial structures that are grown in a lab dish using defined factors. These organoids maintain the structures from normal intestine, including stem cells and differentiated cells in culture, and therefore allow a reliable comparison of functional and phenotypic changes in epithelial cells in response to diverse fatty acids. Our phenotypic screen in both mouse and human intestinal organoids using all dietary fatty acids revealed that omega-6 fatty acid treatment led to increased stem cell activity and reduced differentiation. Mechanistically, we found that omega-6 fatty acids are metabolized to lipid mediators in ISCs, and signaling through their receptor is both necessary and sufficient to mediate the stem cell–enhancing effects of omega-6 fatty acids. Finally, we uncovered the epigenetic mechanisms involving transcription factors that modulate the enhanced intestinal stem

cell state in response to omega-6 fatty acids. We performed similar studies in other epithelial organoids, including prostate, breast, and endometrium organoids. We found that endometrial organoids exhibit functional and molecular alterations in response to fatty acids that are similar to intestinal organoids. Our ongoing studies aim to comparatively assess the molecular mechanisms that govern the tissue specificity of fatty acid–induced alterations in epithelial stem cell activity.

Dietary Regulation of Cancer Immunity

In addition to contributing cancer risk, obesity is associated with immune dysfunction and susceptibility to infections. However, little is known about how obesity or pro-obesity diets affect cancer immunity. Over the past year, we have dissected the mechanisms through which diet-induced obesity impacts cancer immune recognition, response, and immune cell migration pathways.

Immune recognition

The intestinal epithelium serves as the interface between the dietary intake of nutrients, commensal microbes, and immune cells. Intestinal tumorigenesis in this dynamic interface is significantly influenced by the cross talk between cancer cells, immune cells, and microbes. Although cancers develop several strategies to evade the immune system, little is known about how diet-induced obesity impacts cancer immune recognition throughout intestinal tumorigenesis. T cells of the adaptive immune system recognize antigens in the context of major histocompatibility complex (MHC) molecules and play a critical role in cancer immune surveillance. MHC-II-mediated activation of $CD4^+$ T cells can engage multiple mechanisms that contribute to immunity against tumors. Although MHC-II expression and function are usually considered to be restricted to professional antigen-presenting cells like dendritic cells, several studies demonstrated that intestinal epithelial cells express high levels of MHC-II and are able to capture, process, and present antigens to $CD4^+$ T cells. We find that, at steady state, ISCs express high levels of MHC-II protein on their cell surface, which is significantly down-regulated in response to HFD-induced obesity. Down-regulation of immune recognition molecules is one of the key strategies that cancer cells use to evade immune-mediated clearance. Indeed, several human cancers down-regulate MHC-II expression to evade antitumor immune responses, and lower MHC-II expression in tumors correlates with poor survival. To test whether HFD-mediated down-regulation of MHC-II in ISCs enhances intestinal tumorigenesis, we utilized an in vivo orthotopic syngeneic colon transplantation assay in mice that we recently pioneered. Using Lgr5-Cre $APC^{L/L}$ mice that have been on a purified control diet or HFD, we sorted MHC-II^+ APC-null or MHC-II^- APC-null premalignant ISCs by flow cytometry after deletion of the tumor suppressor gene *Apc* by tamoxifen administration. We then transplanted these premalignant cells into the distal colon of syngeneic immune-competent or immune-deficient hosts. We found that reduced MHC-II expression in premalignant ISCs leads to increased tumor initiation rate in vivo in immune-competent hosts but not immune-deficient hosts. Moreover, we demonstrated that ISC-specific genetic ablation of MHC-II in engineered *Apc*-mediated intestinal tumor models led to increased tumor burden in a cell autonomous manner.

The intestinal microbiome plays a significant role in regulating intestinal immunity. Because dietary perturbations are among the major external factors shaping the intestinal microbiome, we asked whether HFD-induced alterations in the microbiome influence MHC-II expression in ISCs and the intestinal epithelium. Consistent with previous findings, HFD-induced obesity led to microbial dysbiosis with reduced bacterial diversity. To determine whether the microbiome is involved in regulation of epithelial MHC-II levels, we treated mice with broad-spectrum antibiotics, which ablated bacterial diversity and massively altered community composition. Notably, antibiotic treatment was accompanied by decreased MHC-II expression in ISCs and the intestinal epithelium, comparable to that observed in HFD. Among the bacterial genera most strongly ablated under HFD conditions and most strongly correlating with MHC-II levels was *Helicobacter*. Indeed, mice harboring *Helicobacter* species had significantly higher MHC-II expression in ISCs compared to mice lacking these species. Mechanistically,

we found that pattern recognition receptor and JAK/STAT signaling regulate MHC-II expression in ISCs. Our ongoing studies are aiming to determine the significance of epithelial MHC-II expression in the context of tumor progression and responsiveness to cancer immunotherapy.

Immune response

To further explore how diet-induced obesity impacts immunity and contributes to cancer risk in the intestine, we performed single-cell RNA sequencing of intestinal immune cells in response to diet-induced obesity. We found that HFD leads to dampened cytotoxicity in intestinal sentinel T cells. To test whether fatty acid constituents of the HFD are sufficient to drive the impaired cytotoxicity, we pretreated T cells with diverse fatty acids and assayed for cytotoxicity and effector molecule production ex vivo. Interestingly, fatty acid–treated T cells exhibited defective tumor killing capacity and effector molecule production. We found that the transcriptional programs regulating T-cell activation, proliferation, and cytotoxicity were down-regulated in response to fatty acid treatment. Our ongoing studies are investigating the transcriptional, epigenetic, and signaling pathways linking fatty acid metabolism to T-cell dysfunction.

Immune migration

Cancer immune surveillance relies on the ability of immune cells to infiltrate target tissues to recognize and respond to tumor. We asked whether diet-induced obesity influences immune cell recruitment dynamics to intestinal epithelium. We found that abundance of intestinal tissue–resident T cells was significantly reduced in response to HFD. Our in vitro studies suggest that fatty acid constituents of the HFD lead to increased T-cell egress, in part because of activation of PPAR. To determine the mechanisms, we established ex vivo intestinal organoid–T cell co-culture to assess T-cell epithelial migration dynamics in response to dietary fatty acids and PPAR activation. We have utilized in vivo genetic models to test the necessity and sufficiency of PPAR and fatty acid utilization in this process. Finally, we are developing in vivo models to assess diet-induced alterations in T-cell migration dynamics in the context of intestinal cancer.

Integrative Analysis of Cancer Risk in Response to Diet and Obesity

Our data suggest that a lard-based HFD leads to increased tumorigenicity in the intestine through several orthogonal mechanisms involving tumor-initiating stem cells, immune cells, and microbiome. To identify how diverse dietary inputs and patterns perturb the cellular and molecular networks associated with cancer risk, we developed novel dietary models and assessed the interactions between intestinal epithelial cells (including stem cells), immune cells, and microbes over time. We performed gene expression, metagenomics, and metabolomics analyses and integrated the data to discover multidimensional features and define potential mediators that contribute to intestinal tumorigenesis. Moreover, we are performing comparative analyses of different tissue types that exhibit increased cancer risk in the context of obesity. Our goal is to build a unifying model for understanding how diet and obesity influence cancer risk at the molecular, cellular, tissue, and organismal level.

Development of Organoid Models to Study Endometrial Cancer

Endometrial cancer is one of the most common gynecologic malignancies and is strongly associated with obesity. However, little is known about the causal cellular and molecular mechanisms that contribute to endometrial cancer risk. There are several types of endometrial cancers that are classified based on histological features. Type-1 endometrial cancers involve low-grade endometrioid endometrial cancers. These are the most common endometrial cancers and are strongly associated with obesity. Type-2 endometrial cancers involve high-grade serous endometrial cancer, clear cell endometrial cancer, and uterine carcinosarcoma. We have successfully developed patient-derived organoid models for all endometrial cancer types as well as normal endometrium. We are characterizing these organoids using histology, genomics, transcriptomics, and epigenomics tools. Moreover, we optimized co-culture conditions for growing endometrial organoids with immune cells to model and

study tumor-immune interactions. We aim to utilize these models to better understand endometrial tumorigenesis and define mechanisms that can be exploited therapeutically.

PUBLICATIONS

Biton M, Haber AL, Rogel N, Burgin G, Beyaz S, Schnell A, Ashenberg O, Su CW, Smillie C, Shekhar K, et al. 2018. T helper cell cytokines modulate intestinal stem cell renewal and differentiation. *Cell* **175:** 1307–1320.e22.

Michelet X, Dyck L, Hogan A, Loftus RM, Duquette D, Wei K, Beyaz S, Tavakkoli A, Foley C, Donnelly R, et al. 2018. Metabolic reprogramming of natural killer cells in obesity limits antitumor responses. *Nature Immunology* **19:** 1330–1340.

In Press

Beyaz S, et al. 2019. Dietary suppression of MHC-II expression in intestinal stem cells enhances intestinal tumorigenesis. *Cell Stem Cell* 2019 *in press*.

ANEUPLOIDY AND CANCER DEVELOPMENT

J. Sheltzer A. Chait, Z. Galluzzo, V. Girish, C. Giuliano, K. John, A. Lin, L. Liu, D. Lukow, A. Palladino, C. Scaduto, J. Smith, A. Vasudevan

The Sheltzer lab applies a variety of techniques, including chromosome engineering, CRISPR mutagenesis, and single-cell analysis, to address fundamental questions in cancer biology. We are particularly interested in exploring the role of gene dosage imbalances in cancer development and progression. Additionally, we apply bioinformatic approaches to investigate genomic alterations associated with cancer patient outcomes in order to identify better biomarkers for treatment and prognosis.

Genetic Investigation of Cancer Drug Targets

A. Chait, Z. Galluzzo, C. Giuliano, K. John, A. Lin, L. Liu, A. Palladino

Substantial progress has been made in the treatment of certain malignancies by targeting cancer "addictions," or genetic dependencies, that encode proteins required for the survival and/or proliferation of cancer cells. Therapeutic agents that block the function of a cancer dependency—like the kinase inhibitor lapatinib in $Her2^+$ breast cancer—can trigger apoptosis and durable tumor regression. Discovering and characterizing druggable cancer dependencies is a key goal of preclinical research.

We are using CRISPR-Cas9 to identify genetic addictions in different cancer types. While screening several cancer cell lines, we discovered that many genes previously reported to be both cancer-essential and the target of anticancer drugs are actually dispensable for cancer growth. For instance, we found that MELK (maternal embryonic leucine zipper kinase), a putative addiction in multiple cancer types, could be eliminated using CRISPR without any detectable loss in cancer cell fitness (Fig. 1). Additionally, we showed that OTS167, a small-molecule inhibitor of MELK undergoing phase II clinical trials, continued to kill MELK-knockout cancer cells with no decrease in potency. This indicated that an anticancer agent had entered clinical trials in human patients because of flawed preclinical data and based on an incorrect understanding of that drug's mechanism of action.

We are working to discover how preclinical cancer research can be conducted in a robust and reproducible manner to prevent similarly flawed results from driving the treatment of human patients. Moreover, we are applying a variety of genetic and biochemical approaches to uncover the true mechanisms of action of potent anticancer drugs that have been mischaracterized.

Chromosomal Instability and Aneuploidy in Cancer

D. Lukow, A. Vasudevan, C. Scaduto

Human cancers show a diverse array of genomic gains and losses that alter the dosage of hundreds or thousands of genes at once. The prevalence of aneuploidy in cancer—first noted more than 100 years ago—has led to a widespread belief that genomic imbalances play a crucial role in tumor development. Indeed, in

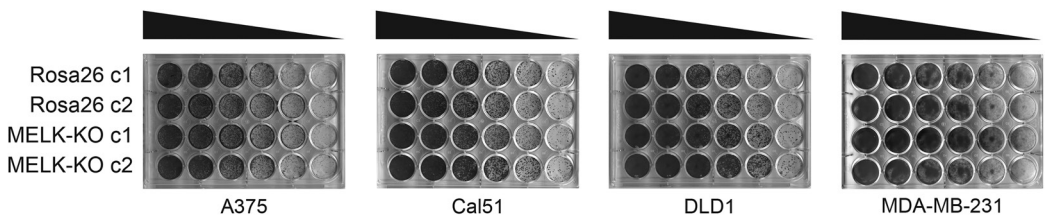


Figure 1. Knocking out MELK with CRISPR-Cas9 fails to affect cancer cell proliferation.

the early twentieth century, Theodor Boveri speculated that abnormal karyotypes altered the balance between pro- and antiproliferative cellular signals and were therefore sufficient to induce transformation. "Boveri's hypothesis" has motivated decades of research into the origins and consequences of aneuploidy, but the precise relationship between abnormal karyotypes and tumorigenesis remains unclear.

We are developing novel models of aneuploidy to explore the impact of genome dosage alterations on tumor development and progression. Using a variety of techniques, including CRISPR-Cas9, microcell-mediated chromosome transfer, and small-molecule mitotic accelerants, we are changing chromosome copy number in human cells. We can then study how these aneuploidies impact a number of cancer-related phenotypes, including metastasis, chemotherapy resistance, and cell cycle progression.

We have characterized a series of isogenic colon cancer cell lines that harbor single extra chromosomes and found that these aneuploidies show significant tumor-suppressive properties (Sheltzer et al., *Cancer Cell 31:* 1 [2017]). Additionally, we have worked to understand how these aneuploidies affect metastasis. We found that adding a single extra copy of chromosome 5 caused a partial epithelial–mesenchymal transition. These cells strongly down-regulated E-cadherin, Epcam, and Claudin-7 and showed increased motility and invasive behavior. These patterns were not observed in cells harboring several additional aneuploidies, suggesting that the phenotype is caused by the increased dosage of a gene or genes found on chromosome 5. We are applying a variety of genetic approaches to identify these key factors.

More generally, we hypothesize that aneuploidy is commonly detrimental under "normal" growth conditions. That is, when a cell is grown in rich media with an adequate supply of nutrients and growth factors, aneuploidy is disfavored. However, in stressful environments, unique karyotypes may exist that confer an environment-specific growth advantage. To test this, we have treated cancer cells with Mps1 inhibitors to generate populations of cells with random aneuploidies. We have observed that pretreatment with Mps1 inhibitors speeds the evolution of drug resistance in cells exposed to various chemotherapy agents. In the case of one drug, vemurafenib, this resistance consistently co-occurs along with the gain of chromosomes 11 and 18. We speculate that similar aneuploidy patterns may exist for other drugs or environments—that is, an "optimal" karyotype can be found to maximize growth potential in each condition. We believe that these results may explain the close association that we have previously documented between aneuploidy and poor prognosis in cancer.

Discovery and Characterization of Genes Affecting Patient Survival in Cancer

J. Sheltzer, J. Smith

Cancers that arise from the same tissue can show vast differences in clinical behavior. For instance, among individuals diagnosed with early-stage colorectal cancer, ~60% of patients will be cured by surgery alone, whereas the remaining 40% will experience a recurrence that is frequently fatal. Biomarkers that can successfully differentiate between patients with benign and aggressive cancers could lead to improved risk prediction, better clinical management, and a decrease in dangerous and unnecessary overtreatment.

To gain a global understanding of the genomic features in a primary tumor that influences cancer prognosis, we are collecting and analyzing molecular profiles from tens of thousands of patients with known clinical outcomes. These data have revealed that mutations in almost all cancer driver genes contain remarkably little information on patient prognosis. However, copy number alterations in these same driver genes harbor significant prognostic power. Focal copy number alterations (CNAs) are associated with worse outcomes than broad alterations, and CNAs in many driver genes remain prognostic when controlling for stage, grade, *TP53* status, and total aneuploidy. The biological importance of cancer copy number alterations is unexpected—and a topic of continuing investigation.

PUBLICATIONS

Giuliano CJ, Lin Al, Smith JC, Palladino AC, Sheltzer JM. 2018. MELK expression correlates with tumor mitotic activity but is not required for cancer growth. *eLife* **7:** e32838. doi:10.7554/eLife.32838

Smith JC, Sheltzer JM. 2018. Systematic identification of mutations and copy number alterations associated with cancer patient prognosis. *eLife* **7:** e39217. doi:10.7554/eLife.39217

TARGETING SELF-RENEWAL PATHWAY AND METABOLIC DEPENDENCY IN HEMATOPOIETIC MALIGNANCIES: FROM BASIC DISCOVERY TO PRECLINICAL DEVELOPMENT OF INNOVATIVE THERAPEUTICS

L. Zhang J. Korossy G. Mazzeo A. Seidner D. Tsang Hu E. Wong
B. Li W. Mei G. Trivedi K. Vishnu L. Zhu

Dr. Zhang is a principal investigator and CSHL fellow at Cold Spring Harbor Laboratory who directs his independent lab in the investigation of the mechanism that governs the balance of self-renewal and differentiation in hematopoietic stem and progenitor cells. The Zhang laboratory employs both CRISPR-Cas functional genomic and forward chemical genomic approaches to identify novel regulators that modulate this process and utilize genetically engineered murine models and human patient sample culture systems to validate the efficacy of therapeutic approaches that target these novel regulators. Through collaborations with medicinal chemists and physicians, the long-term goal of our research is to develop novel therapeutic strategies and cures for hematological malignancies, including myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML). For our research, we have been awarded the NIH Research Evaluation and Commercialization Hub (REACH) Feasibility Award, the NIH REACH Proof of Concept Award, and the Edward P. Evans Foundation EvansMDS Young Investigator Award.

Thus far, we have identified a G-protein-coupled receptor (GPCR) as a novel regulator of early erythroid progenitor self-renewal and a drug target for treatment-refractory MDS. In collaboration with medicinal chemist Yousef Al-Abed at Northwell Health System, we have developed novel chemical entities as selective antagonists for this GPCR with improved drug metabolism and pharmacokinetic (DMPK) properties. In collaboration with Yousef Al-Abed and physician Johnson Liu at Northwell Health System, we have characterized drug efficacy, DMPK properties, and toxicology properties for lead compounds. We are working on these investigational new drug (IND)-enabling studies and preparing Phase 1 clinical trials for treatment-refractory MDS at Northwell Health System. The novel therapy will benefit not only MDS patients but also cancer patients who are undergoing chemotherapy and radiation therapy. The project has been supported by CSHL, the NIH, Northwell Health System, and the Edward P. Evans Foundation.

We have also determined that pyridoxal kinase (PDXK), an enzyme that catalyzes the formation of the bioactive form of vitamin B_6—pyridoxal phosphate (PLP)—was selectively required for leukemia cell proliferation relative to other normal cell types. We have established a novel liquid chromatography-mass spectrometry (LC-MS) method to monitor intracellular PLP and found that PLP-dependent generation of polyamine and the PLP-dependent GOT2 pathway, which produces nucleotides, selectively support AML cell proliferation. Our work has identified the vitamin B_6 pathway as a pharmacologically actionable dependency in AML. We are collaborating with medicinal chemist Ouathek Ouerfelli at Memorial Sloan Kettering Cancer Center to modify PDXK inhibitors for better drug DMPK properties, and we expect to extend this work to a collaboration with the Tri-Institutional Therapeutics Discovery Institute and Takeda Pharmaceutical Company to provide lead compounds for clinical development.

Identifying and Targeting Novel Self-Renewal Pathways in Refractory Myelodysplastic Syndromes

MDSs represent the most common cause of acquired bone marrow failure in adults. Peripheral blood cytopenias resulting from progressive bone marrow failure are a major manifestation of MDSs and are predictive of poor prognosis. Current therapeutic options for MDSs are very limited. Only a small proportion of MDS patients benefit from standard erythropoietin (EPO) treatment, and the recently approved compound lenalidomide only targets the 5% of MDS patients who carry a chromosome 5q

deletion. In addition, many of the initial responders do not have a long-term response. Currently, the only option for patients who do not respond to EPO and lenalidomide is red blood cell transfusion, which exposes patients to insufficient correction of anemia, alloimmunization, and organ failure secondary to iron overload. Therefore, there is an unmet and urgent demand for novel therapeutics that target treatment-refractory MDSs.

Burst-forming unit-erythroid (BFU-E) and colony-forming unit-erythroid (CFU-E) are two progenitor cell types of the erythroid lineage. BFU-E is an earlier progenitor capable of extensive self-renewal and expansion, whereas the later progenitor CFU-E is only able to undergo approximately four to five divisions to generate approximately 30 cells. Clinically, bone marrow early erythroid progenitor cell levels predict responsiveness of MDS patients to EPO treatment. MDS patients with early erythroid progenitor numbers similar to normal individuals respond to erythropoietin, whereas patients with inadequate early erythroid progenitor fail to respond. Currently, however, our understanding of the molecular mechanisms underlying BFU-E self-renewal, especially druggable regulators and corresponding small chemical compounds targeting these regulators, is extremely limited, precluding us from targeting this process to treat MDSs.

We have recently established novel methods to purify BFU-Es (defined as lineage Ter-119-CD16/CD32$^-$ Sca-1$^-$ CD41$^-$ c-Kit$^+$ CD71/Cd24a$^{10\%\ low}$ cells), which allowed us to identify drug targets and small chemical compounds that boost their expansion. A chemical screening effort we recently performed for druggable regulators of BFU-E expansion unexpectedly revealed a GPCR pathway as a novel regulator of BFU-E self-renewal. Consistent with this, pharmacologic inhibition of this pathway improved erythropoiesis in a mouse model of MDS and in MDS patient samples. Through application of chemical and genomic approaches to these purified populations, we identified multiple druggable GPCRs that are differentially expressed during BFU-E self-renewal and tested drugs against each. Of the compounds tested, two closely related antagonists targeting the same GPCR triggered BFU-E expansion. Although the role of this GPCR has been extensively evaluated in neural physiology, the role of this receptor and its signaling intermediates in hematopoiesis is not known.

We have demonstrated that pharmacological inhibition of this GPCR with nanomolar concentration of selective antagonists triggered primary early erythroid progenitor expansion. Importantly, in a genetically engineered conditional knock-in Mx1-Cre Srsf2 P95H/WT MDS model that faithfully recapitulates essential features of human MDS, injection of GPCR-selective antagonist completely corrected anemia of MDS, exhibited a sustained and long-term therapeutic efficacy, and extended survival of MDS mice to levels comparable to wild-type littermate controls. Together, these data suggest that this GPCR pathway promotes BFU-E self-renewal by up-regulating the expression of genes important for the maintenance of BFU-E progenitor status, and that inhibition of this pathway may improve hematopoiesis across multiple subtypes of MDS.

Through collaboration with medicinal chemist Yousef Al-Abed at the Feinstein Institute, we have further performed medicinal chemistry modifications on selective antagonists of this GPCR to reduce their potential side effects. We have demonstrated that multiple lead compounds exhibited dramatic reduction of side effects and demonstrating both in vitro and in vivo efficacies. With human MDS patient samples, we have further showed that GPCR-selective antagonists significantly improved erythroid progenitor expansion in samples isolated from MDS patients. These preclinical developments have been supported by the NIH Research Evaluation and Commercialization Hub (REACH) Feasibility Award, the Edward P. Evans Foundation EvansMDS Young Investigator Award, and the NIH REACH Proof of Concept Award.

Identifying and Targeting Vitamin B_6 Addiction in Acute Myeloid Leukemia

AML is one of the most devastating blood cancers, and it affects around one million people and results in 147,000 deaths per year worldwide. AML is characterized by the abnormal production of myeloid lineage of blood cells and the rapid growth of abnormal leukemia blasts in bone marrow and peripheral blood. The symptoms of AML include shortness of breath, bleeding, and increased risk of infection. AML is a very aggressive cancer, it progresses rapidly, and AML patients become fatal within weeks to months.

Currently, therapeutic options for AML are very limited. Only ~35% of AML patients less than 60 years old and 10% of those more than 60 years old benefit from standard chemotherapy. Older AML patients whose health cannot afford intensive chemotherapy only have a survival of 5–10 mo. Thus, there is an urgent and unmet medical need to develop novel therapeutics for AML patients.

AML cells rely on altered metabolism to support their abnormal proliferation. The clinical success of isocitrate dehydrogenase (IDH) inhibitors in treating AML harboring IDH mutations highlights the therapeutic potential of targeting deregulated metabolism in cancer cells. However, the vast majority of oncogenic drivers alter cellular metabolism through indirect mechanisms, and, as such, the metabolic regulators that sustain tumorigenesis cannot be inferred from genomic analyses alone. Combining the results of our metabolism-focused single-guide RNA (sgRNA) library screening and the analysis of publicly available functional genomic data sets, we found sgRNAs targeting PDXK (pyridoxal kinase), an enzyme that catalyzes the formation of the bioactive form of vitamin B_6, PLP, were selectively required for leukemia cell proliferation relative to many other normal and cancer cell types. Indeed, an analysis of publicly available genome-wide CRISPR screens for genes essentially indicate that PDXK has a requirement similar to that of Bcl-2, which is a target of a recently FDA-approved drug for AML, and is preferentially required in human leukemia cells compared to other cell types.

For validation, we generated individual sgRNAs and short hairpin RNAs (shRNAs) that target PDXK and performed competition assays in multiple murine and human AML cell lines. Consistent with CRISPR-Cas9 screening results, knockout of Pdxk using individual sgRNAs inhibited proliferation of Nras(G12D)/MLL-AF9 leukemia cells. The validity of these findings was reinforced by shRNA-mediated knockdown of Pdxk as an orthogonal approach. Knockdown of Pdxk inhibited the proliferation of multiple additional mouse AML cell lines produced by different oncogenic events, as well as a panel of immortalized human cell lines. Of note, PDXK depletion did not trigger leukemia cell differentiation but instead reduced cell cycle progression and produced an increase in apoptosis. Accordingly, gene set enrichment analysis of RNA-Seq data obtained following PDXK sgRNAs revealed a significant down-regulation of genes associated with "cell cycle progression," "DNA replication," and "nucleotide metabolism." Collectively, these results establish PDXK as a metabolic vulnerability in AML.

The PDXK product PLP is a cofactor for multiple enzymes involved in amino acid, nucleic acid, and lipid metabolism and is a biomarker of PDXK activity. Although plasma PLP can be readily measured, no reliable assays currently detect intracellular PLP. To determine whether PDXK inhibition affected PLP levels in leukemia cells, we developed a high-performance liquid chromatography-mass spectrometry (HPLC-MS)-based method to compare intracellular PLP levels in leukemic cells in the presence or absence of PDXK inhibition. Genetic inhibition of PDXK dramatically decreased PLP levels in mouse and human leukemic cells. The availability of this pharmacodynamic marker of PDXK inhibition enabled us to relate the leukemia-selective dependence of PDXK to its role in PLP production and vitamin B_6 metabolism. Hence, reexpression of a wild-type PDXK cDNA but not a kinase dead PDXK mutant (D235A) was able to rescue the proliferative defects and PLP depletion produced by a PDXK sgRNA that targets the intronic region of the human gene. These results validate the PDXK kinase activity as a therapeutic target.

We also validated PDXK in vivo using genetic and pharmacologic approaches. First, we used mouse AML lines expressing a reverse tetracycline transactivator (rtTA) and a tetracycline-responsive element (TRE) promoter, in which shRNA expression is induced by doxycycline. In these cells, Nras(G12D) was co-expressed with luciferase to allow for monitoring of disease progression with bioluminescence imaging. Leukemia cells were transduced with shRNAs targeting a control gene or Pdxk and were transplanted into sublethally irradiated recipient mice. Induction of Pdxk shRNAs by doxycycline treatment significantly delayed disease progression and extended overall animal survival. We also tested the small chemical compound PDXK inhibitor 4′-O-methylpyridoxine. Treatment of mice bearing Nras(G12D)/MLL-AF9 leukemia with 4′-O-methylpyridoxine significantly delayed disease progression and produced a survival advantage.

In summary, by performing a focused CRISPR-Cas9 screen to uncover metabolic vulnerabilities in AML, we have shown that PDXK, an enzyme that

catalyzes the formation of the bioactive form of vitamin B_6, PLP, was selectively required for leukemia cell proliferation relative to other normal cell types. We have established a novel LC-MS method to monitor intracellular PLP, and found that PLP-dependent generation of polyamine and the PLP-dependent GOT2 pathway, which produces nucleotides, selectively supports AML cell proliferation. Our work has identified the vitamin B_6 pathway as a pharmacologically actionable dependency in AML. We are collaborating with medicinal chemist Ouathek Ouerfelli at Memorial Sloan Kettering Cancer Center to modify PDXK inhibitors for better drug DMPK properties, and we expect to extend this work to collaboration with the Tri-Institutional Therapeutics Discovery Institute and Takeda Pharmaceutical Company to provide lead compounds for clinical development.

Identification of the GR-JARID2 Axis as a Negative Feedback Loop for Early Erythroid Progenitor Self-Renewal

Erythropoiesis is a multistage developmental process that results in erythrocyte production. EPO is a crucial regulator of this process, triggering survival and proliferation of the late erythroid progenitor CFU-E, and has thus been widely used for the treatment of anemias. In bone marrow failure disorders such as MDS and cancer chemotherapy and radiation therapy patients, however, insufficient amounts of CFU-Es limit EPO-dependent erythrocyte output. EPO treatment is therefore not able to correct this type of anemia. Treatment of EPO-resistant anemia therefore requires increased production of CFU-E progenitors from the most immature committed erythroid progenitor cells, the BFU-Es. Therefore, targeting BFU-Es to expand and increase CFU-E output is crucial for sustained erythropoiesis and treatment of anemias in these conditions, and a systemic-level functional analysis to better understand molecular mechanisms underlying erythropoiesis at the BFU-E stage holds promise in the development of novel therapeutics for EPO-refractory anemias associated with cancer chemotherapy and radiation therapy.

We integrated glucocorticoid-induced differential gene expression profiling, glucocorticoid receptor (GR) chromatin immunoprecipitation sequencing (ChIP-seq), and shRNA functional genomic screening to provide the first systemic-level functional understanding of GR-dependent regulation of erythropoiesis at the BFU-E stage. We showed that groups of genes involved in negative regulation of stem cell differentiation and apoptosis are essential for BFU-E expansion, providing the first link between the previously observed contribution of cellular-level phenotypes, including differentiation and apoptosis, to BFU-E expansion and their molecular-level explanations. We identified multiple novel regulators of erythropoiesis at the BFU-E stage and classified these novel regulators into two categories: those that are direct GR target genes, such as Zfp36l2, a known BFU-E self-renewal regulator, and those that are secondary to GR, such as *Ddit*, a gene that functions downstream of HIF1α, which is a co-regulator of GR that modulates BFU-E self-renewal.

We and others previously showed that GR target genes such as *ZFP36L2* and cofactors including HIF1α and PPARα positively regulated BFU-E expansion. However, because GR triggered limited expansion of BFU-E, it is not known how GR balanced BFU-E expansion and restricted it to a limited level. Molecularly, it is also not known whether there are direct GR transcriptional target genes to function as negative feedback loops of this process. Here, we identified JARID2 as a negative regulator of BFU-E expansion and provided the first molecular clue of this negative feedback loop. GR binds to a genomic region near Jarid2 and up-regulates its expression. The up-regulated JARID2 then negatively regulates BFU-E expansion to restrict GR's effect to a limited level. Therefore, the axis consisting of GR-JARID2 functions as a negative feedback loop to restrict the erythroid cell to a limited expansion. GR triggers the up-regulation of both pro-expansion genes such as *Zfp36l2* and anti-expansion genes such as *Jarid2* to balance erythroid expansion. The identification of JARID2 as a novel negative regulator of BFU-E expansion also suggested a potential approach to enhance BFU-E expansion to treat anemia through genetically or pharmacologically inhibiting JARID2.

In conclusion, our research investigates the mechanism that governs the balance of self-renewal and differentiation in hematopoietic stem and progenitor cells. Through both genetic and chemical functional genomics, our research identified a GPCR as a novel druggable regulator of self-renewal of an early erythroid progenitor and uncovered the metabolic

enzyme PDXK as a novel dependency in AML. We have demonstrated that pharmacological inhibition of this GPCR and PDXK overcame the refractoriness of MDS and AML, respectively, to current therapies. We have further modified chemical structures of selective antagonists for the GPCR and PDXK for improved DMPK properties. Our research provided selective antagonists and their corresponding toxicology and safety pharmacology profiles. These results will form the foundation of IND application or clinical developments of lead compounds for MDS and AML, respectively. These novel therapeutics will potentially revolutionize clinical management of hematopoietic malignancies including MDS and AML.

PUBLICATIONS

In Press

Al-Abed Y, Zhang L. 2019. Compositions and methods for treating anemia. Patent filed.

Chen C-C, Li B, Morris IV JP, Chen C, Li X, Millman SE, Mayle A, Ho Y-J, Liu H, Shah H, et al. 2019a. Vitamin B_6 addiction in acute myeloid leukemia. Under revision

Chen C, Galeev R, Flygare J, Zhang L. 2019b. Functional genomics identify JARID2 as a negative regulator of early erythroid progenitor self-renewal. Under revision

Lowe S, Zhang L, Chen CC. 2019. Compositions and methods for treating acute myeloid leukemia. Patent Pub. No. WO/2019/108559A1

Trivedi G, Inoue D, Chen C, Bitner L, Chung YR, Taylor J, Gönen M, Wess J, Abdel-Wahab O, Zhang L. 2019. A GPCR regulates self-renewal of early erythroid BFU-E progenitor. Under revision

AUTHOR INDEX

WATSON SCHOOL OF BIOLOGICAL SCIENCES

ADMINISTRATION

Alexander Gann, Ph.D., WSBS Professor and Dean
Monn Monn Myat, Ph.D., Associate Dean
Alyson Kass-Eisler, Ph.D., Postdoctoral Program Officer and Curriculum Director
Kimberly Creteur, M.Ed., M.S.Ed., Admissions and Recruitment Manager
Kimberley Geer, Administrative Assistant

EXECUTIVE COMMITTEE

Chair
Alexander Gann

Members
Camila dos Santos
John Inglis
Alexander Krasnitz
Bo Li
David L. Spector (Director of Research)
Anthony Zador

Student Representatives
Emilis Bružas, WSBS
Zhe Qian, SBU

Secretary
Monn Monn Myat

ADMISSIONS COMMITTEE

Chair
Adrian Krainer

Members
Alexander Gann (ex officio)
Bo Li
Zachary Lippman
W. Richard McCombie
Stephen Shea
Adam Siepel
Nicholas Tonks
Christopher Vakoc
Linda Van Aelst

Secretary
Kimberly Creteur

QUALIFYING EXAM COMMITTEE

Chair
Linda Van Aelst

Members
Molly Hammell
W. Richard McCombie
Stephen Shea
Christopher Vakoc

EXTERNAL ADVISORY COMMITTEE

Keith Yamamoto (Chair)
Executive Vice Dean, School of Medicine
University of California, San Francisco

Victor Corces
Professor, Department of Biology
Emory University
Professor, Howard Hughes Medical Institute

Gail Mandel
Senior Scientist, Vollum Institute
Oregon Health and Science University
Investigator, Howard Hughes Medical Institute

Barbara Meyer
Professor of Genetics and Development
University of California, Berkeley
Investigator, Howard Hughes Medical Institute

Frank Solomon
Professor, Department of Biology and Center for Cancer Research
Massachusetts Institute of Technology

WATSON SCHOOL OF BIOLOGICAL SCIENCES DEAN'S REPORT

The 14th WSBS Graduation

On April 29, 2018, we celebrated the Watson School's 15th graduation ceremony. Five students were awarded Ph.D. degrees: Fred Marbach and Onyekachi Odoemene from the Entering Class of 2011, Yu-Jui (Ray) Ho and Paul Masset from the Entering Class of 2012, and Daniel Kepple from the Entering Class of 2013. Drs. David and Leon Botstein, who are brothers, were awarded honorary degrees and gave co-commencement addresses.

David Botstein is a prominent geneticist whose advocacy for gene mapping was crucial in laying the groundwork for the Human Genome Project. After a long academic career at Harvard, University of Michigan, MIT, Stanford, and Princeton, and as Vice President of Science at Genentech, David is currently the chief scientific officer of Calico (California Life Company), a biotechnology company focused on longevity and health. David co-taught the Advanced Bacterial Genetics Course at Cold Spring Harbor Laboratory in the 1970s, where he also did research while on sabbatical in 1974–1975 and was a member of the Board of Trustees from 2003 to 2013. He urged the graduates to "Take every opportunity to learn, teach, and apply the fundamental principles." Adding, "No discovery or experiment is complete until it becomes knowledge for others to build on, even if it is preliminary, and don't skimp on describing your methods."

Leon Botstein, the youngest college president in American history, is currently the president of Bard College and the music director of the American Symphony Orchestra. He was honored as a remarkable academic leader and an outstanding educator. In encouraging the students to seek knowledge within and beyond science, he said, "I have never met or known of a great scientist who has ever been really a narrow specialist, immune from a deep engagement with matters such as literature, art, politics, and history."

2018 WSBS DOCTORAL RECIPIENTS

Student	Thesis advisor	Academic mentor	Current position
Lital Chartarifsky	Anne Churchland	John Inglis	Biotechnology and Business Development Coordinator, Cold Spring Harbor Laboratory
Talitha Forcier	Justin Kinney	Nicholas Tonks	Postdoctoral Fellow, Cold Spring Harbor Laboratory (Advisor: Molly Hammell)
Daniel Kepple	Alexei Koulakov	David Stewart	Senior Machine Learning Engineer, Samsung Artificial Intelligence Center, New York
Paul Masset	Adam Kepecs	Jan Witkowski	Postdoctoral Fellow, Harvard University (Advisors: Venkatesh Murthy and Naoshige Uchida)

2018 Graduates: (*Left to right*) WSBS Dean Alexander Gann; CSHL President Bruce Stillman; CSHL Board of Trustees Chair Jamie Nicholls; Paul Masset; Daniel Kepple; honorary degree recipient, David Botstein; Yu-Ji (Ray) Ho; CSHL Chancellor Emeritus James Watson; honorary degree recipient, Leon Botstein

2018 THESIS DISSERTATION DEFENSES

ENTERING CLASS OF 2012

Talitha Forcier, July 31, 2018

Precision measurement of cis*-regulatory energetics in living cells.*

Thesis Examining Committee

Chairperson:	**Adam Siepel**
Research Mentor:	**Justin Kinney**
Academic Mentor:	**Nicholas Tonks**
Committee Member:	**Molly Hammell**
External Examiner:	**Bryce Nickels,** *Rutgers University*

Paul Masset, April 18, 2018

Building confidence: Behavioral and neural signatures for metacognition.

Thesis Examining Committee

Chairperson:	**Stephen Shea**
Research Mentor:	**Adam Kepecs**
Academic Mentor:	**Jan Witkowski**
Committee Member:	**Alexei Koulakov**
Committee Member:	**Anthony Zador**
External Examiner:	**Michael Shadlen,** *Columbia University Medical Center*

ENTERING CLASS OF 2013

Lital Chartarifsky, October 12, 2018

Visual, auditory, and multisensory decision-making: Behavior and neural circuits.

Thesis Examining Committee

Chairperson:	**Stephen Shea**
Research Mentor:	**Anne Churchland**
Academic Mentor:	**John Inglis**
Committee Member:	**Anthony Zador**
External Examiner:	**Heather Read,** *University of Connecticut*

Daniel Kepple, April 24, 2018

An exploration of olfactory spaces.

Thesis Examining Committee

Chairperson:	**Florin Albeanu**
Research Mentor:	**Alexei Koulakov**
Academic Mentor:	**David Stewart**
Committee Member:	**Stephen Shea**
External Examiner:	**Joel Mainland,** *Monell Chemical Senses Center*

DOCTORAL THESIS RESEARCH

Student	Academic mentor	Research mentor	Thesis research
ENTERING CLASS OF 2013			
Giorgia Battistoni *Starr Centennial Scholar* *Boehringer Ingelheim Fonds Fellow*	Christopher Hammell	Gregory Hannon	One-carbon metabolism and DNA methylome dynamics in pancreatic cancer.
Sanchari Ghosh *Charles A. Dana Fellow*	Josh Dubnau	Anthony Zador	Role of corticostriatal plasticity in learning an auditory discrimination task.
Michael Gutbrod *Bristol-Myers Squibb Fellow*	Zachary Lippman	Robert Martienssen	Small RNA and the RNAi pathway in transposable element regulation and differentiation in the preimplantation embryo.
Laura Maiorino *George A. and Marjorie H. Anderson Fellow* *Boehringer Ingelheim Fonds Fellow*	Nicholas Tonks	Mikala Egeblad	Understanding the role of the epithelial–mesenchymal plasticity in pancreatic cancer metastasis.
Georgi Yordanov *Leslie C. Quick, Jr. Fellow* *Boehringer Ingelheim Fonds Fellow*	Leemor Joshua-Tor	David Tuveson	Role of c-Myc in pancreatic cancer.
ENTERING CLASS OF 2014			
Emilis Bružas *Starr Centennial Scholar*	Alea Mills	Mikala Egeblad	Investigation of mechanisms responsible for reawakening and chemoresistance in a breast cancer dormancy model.
Hamza Giaffar *Robert and Teresa Lindsay Fellow*	Jan Witkowski	Alexei Koulakov	The primacy model of olfactory coding.
Jacqueline Giovanniello *NIH Predoctoral Trainee*	Bruce Stillman	Bo Li	Disruption of central amygdala fear circuit in a 16p11.2 microdeletion model of autism.
Elizabeth Hutton *Elizabeth Sloan Livingston Fellow*	Molly Hammell	Adam Siepel	Functional variant prediction in noncoding regions.
Sashank Pisupati *Cashin Fellow*	Stephen Shea	Anne Churchland	Dissecting the circuits and mechanisms that support optimal multisensory integration in rodents.
Colin Stoneking *NIH Predoctoral Trainee*	Zachary Lippman	Anthony Zador	Neuronal mechanisms enabling decision-making to be learned.
Jue Xiang Wang *George A. and Marjorie H. Anderson Fellow* *Boehringer Ingelheim Fonds Fellow*	Mikala Egeblad	Hiro Furukawa	Impact of subunit composition and de novo mutations on NMDA receptor structure, channel function, and interactions.
Anqi Zhang *Starr Centennial Scholar*	Bo Li	Anthony Zador	From corticostriatal plasticity to a common pathway
ENTERING CLASS OF 2015			
Benjamin Berube *National Science Foundation Fellow* *NIH Predoctoral Trainee* *Elisabeth Sloan Livingston Fellow*	Zachary Lippman	Robert Martienssen	A single-cell assessment of germline epigenetic heterogeneity.
Kristina Grigaityte *Farish-Gerry Fellow*	John Inglis	Mickey Atwal	Computational analyses of high-throughput single-cell T-cell receptor sequences in health and disease.
Matt Lee *David H. Koch Fellow*	Nicholas Tonks	Lloyd Trotman	The road to metastasis: defining the initial stages of prostate cancer progression.
Katarina Meze *Leslie C. Quick, Jr. Fellow*	Jay Lee	Leemor Joshua-Tor	Structural and functional studies of RNA regulatory mechanisms mediated by Lin28.
Alexandra Nowlan *Genentech Fellow* *George A. and Marjorie H. Anderson Fellow*	Jessica Tollkuhn	Stephen Shea	Multisensory experience dependent plasticity: network dynamics in auditory processing following parturition.

(*continued*)

DOCTORAL THESIS RESEARCH (*continued*)

Student	Academic mentor	Research mentor	Thesis research
Sofya Polyanskaya *Starr Centennial Scholar*	Alexander Krasnitz	Christopher Vakoc	Identification of fusion oncoprotein codependencies in cancer.
Ngoc Tran *Samuel Freeman Fellow*	Leemor Joshua-Tor	Alexei Koulakov	Predicting bioactivity using a data-driven representation of three-dimensional chemical structures.
ENTERING CLASS OF 2016			
Brianna Bibel *National Science Foundation Fellow*	Hiro Furukawa	Leemor Joshua-Tor	Structural and functional studies of phosphorylation-mediated regulation of the RNAi effector *Argonaute.*
Alberto Corona *NIH Predoctoral Trainee* *Hearst Foundation Fellow*	David Jackson	Stephen Shea	Identification of neural circuitry underlying paternal behaviors.
David Johnson *National Science Foundation Fellow* *Hearst Foundation Scholar*	Zachary Lippman	Alea Mills	Elucidating the role of BRPF1 in human glioblastoma multiforme.
Christopher Krasniak *NIH Predoctoral Trainee*	Jan Witkowski	Anthony Zador	The role of cholinergic input to visual cortex in mouse spatial visual attention.
Shaina Lu *Edward & Martha Gerry Fellow*	Leemor Joshua-Tor	Anthony Zador	Development of a high-throughput pipeline to study the relationship of neuron projections and gene expression underlying mouse models of neuropsychiatric disorders.
Kathryn O'Neill *National Science Foundation Fellow* *NIH Predoctoral Trainee*	Camila dos Santos	Molly Hammell	Investigations into TDP-43-mediated effects on sRNA biology.
Luqun Shen *Edward & Martha Gerry Fellow*	David Stewart	Stephen Shea	Predicting bioactivity using a data-driven representation of three-dimensional chemical structures.
Olya Spassibojko *David & Fanny Luke Fellow*	Jessica Tollkuhn	Ullas Pedmale	Molecular determinants controlling cryptochrome light signal transduction.
Martyna Sroka *George A. and Marjorie H. Anderson Fellow*	Molly Hammell	Christopher Vakoc	Molecular dissection of the PAX3-FOXO1 fusion oncoprotein pathway in rhabdomyosarcoma.
Ran Yan *George A. and Marjorie H. Anderson Fellow*	David Tuveson	Douglas Fearon	Identification of endogenous antigen-specific T cells in pancreatic cancer metastasis.
Chengxiang (Charlie) Yuan *A*STAR Fellow*	Nicholas Tonks	Jay Lee	Linking the cell cycle and developmental fate specification.
ENTERING CLASS OF 2017			
Lyndsey Aguirre *Hearst Foundation Fellow*	Ullas Pedmale	Zachary Lippman	Decoding *cis*-regulatory control of quantitative trait variation in tomato.
Sara Boyle *NIH Predoctoral Trainee*	Jessica Tollkuhn	Bo Li	Can the central amygdala's interaction with midbrain dopamine areas control motivated behavior?
Jordan (Bruno) Gegenhuber *Charles A. Dana Fellow* *John & Amy Phelan Fellow* *NIH Predoctoral Trainee*	John Inglis	Jessica Tollkuhn	Gene regulatory mechanisms underlying brain organization by perinatal estradiol.
Benjamin Harris *Crick-Clay Fellow*	W. Richard McCombie	Jesse Gillis	Robust analysis of single-cell RNA sequencing data to study development and emerging model organisms.
Yuzhao (Richard) Hu *George A. and Marjorie H. Anderson Fellow*	Justin Kinney	Ullas Pedmale	Characterizing cryptochrome 2 downstream targets in *Arabidopsis.*
Dennis Maharjan *John & Amy Phelan Fellow*	Florin Albeanu	Adam Kepecs	Investigating the neural basis of impulsive choice and impulsive action.

(*continued*)

DOCTORAL THESIS RESEARCH (*continued*)

Student	Academic mentor	Research mentor	Thesis research
Diogo Maia e Silva *Robert and Teresa Lindsay Fellow*	Bruce Stillman	Christopher Vakoc	Defining the molecular origins of squamous pancreatic ductal adenocarcinoma.
Cole Wunderlich *NIH Predoctoral Trainee*	Adam Siepel	Dan Levy	Single-cell RNA sequencing analysis of neurons derived from autistic patients and their healthy siblings.

Teaching Award

At this year's graduation ceremony, the School awarded its 14th Winship Herr Award for Excellence in Teaching to Dr. Zachary Lippman, the lead instructor of the Specialized Disciplines Course in Genetics. Zach, who also won the award in 2013, was chosen by the first-year students for this award. Zach was not present at the ceremony, but he delivered his remarks via Skype. Here is what just one student said in their nomination:

Zachary Lippman attends via Skype

"All the problem sets were very well constructed, encouraging relevant literature searches. Zach gave us useful, specific feedback after each problem set that helped learning the module's contents and correcting our misconceptions. His extensive biological knowledge transpired in his vivid illustrations of core genetic concepts. He was very dedicated, often staying after class to clarify questions that arose in lectures or problem sets. Overall, Zach was a superb instructor that managed to transmit his enthusiasm and thorough scholarship, putting together a very enjoyable and instructive course."

Faculty Changes

Two new faculty members joined the Watson School in 2018: Alexander Dobin and Tobias Janowitz. A new fellow also joined CSHL in 2018, Semir Beyaz.

Alexander earned his Ph.D. in physics at the University of Minnesota followed by a postdoctoral position in Tom Gingeras' lab at CSHL. His expertise is in the application of computational and statistical skills to genomics. His lab is developing new ways to analyze DNA sequence data in the hope of fully annotating specific traits in the human genome.

Tobias earned his M.D. and Ph.D. at the University of Cambridge. Before coming to CSHL, he was a clinician scientist at the Cancer Research UK Cambridge Institute. He is interested in understanding cancer as a systemic condition—one that does not just affect a particular tissue or organ, but rather the entire organism. His aim is to discover how the body's response to a tumor can be used to improve cancer treatment.

Semir earned his Ph.D. at Harvard University, where he studied the relationship between diet and the incidence of cancer. His lab at CSHL studies how dietary choices impact the body's natural immunity against cancer, specifically in the immune recognition and response pathways.

Alexander, Tobias, and Semir have already participated in WSBS activities, including giving Research Topics talks to the first-year students. Tobias and Semir also lectured in the Specialized Disciplines Course on Cancer. We look forward to their growing participation as members of the faculty.

Admissions 2018

The School received a very high-quality pool of more than 200 applications for the Entering Class of 2018 and is indebted to its Admissions Committee, which reviewed, interviewed, and selected candidates for our doctoral program. The Admissions Committee for the 2018 entering class comprised Adrian Krainer (Chair), Bo Li, Zachary Lippman, W. Richard McCombie, Stephen Shea, Adam Siepel, Nicholas Tonks, Christopher Vakoc, Linda Van Aelst, and myself (ex officio).

Entering Class of 2018

On August 20, 2018, the WSBS welcomed the 20th incoming class, consisting of 11 new students: King Hei (Teri) Cheng, Danielle Ciren, Marie Dussauze, Ilgin Ergin, Connor Fitzpatrick, Amritha Varshini Hanasoge Somasundara, Asad Aziz Lakhani, Ziyi Mo, Alexa Pagliaro, Jenelys Ruiz, and Jonathan Werner.

ENTERING CLASS OF 2018

King Hei "Teri" Cheng, University of Edinburgh, United Kingdom: B.S. in Biological Sciences (Molecular Genetics) (2018)
Academic Mentor: Adrian Krainer

Danielle Ciren, Queen's University, Canada: B.S. with Honours in Biology (2018)
Academic Mentor: Ullas Pedmale

Marie Dussauze, Ecole Normale Supérieure Paris-Saclay, France: M.S. in Life Sciences, Neuroscience (2018); University of Versailles Saint Quentin: B.S. in Physiology (2015)
Academic Mentor: Stephen Shea

Ilgin Ergin, Bilkent University, Turkey: B.S. in Molecular Biology and Genetics (2018)
Academic Mentor: Thomas Gingeras

Connor Fitzpatrick, Stony Brook University, New York: B.S. in Biochemistry (2015)
Academic Mentor: Robert Martienssen

Amritha Varshini Hanasoge Somasundara, University of Pennsylvania, Philadelphia: M.S. in Biotechnology, Molecular Biology (2015); Sri Jayachamarajendra College of Engineering, India: B.E. in Biotechnology (2013)
Academic Mentor: Jason Sheltzer

Asad Aziz Lakhani, University of Cambridge, United Kingdom: B.A. with Honors in Biochemistry (2018); CSHL Undergraduate Research Program: Egeblad Lab (2017)
Academic Mentor: David Tuveson

Ziyi Mo, New York University, Abu Dhabi: B.S. in Biology (2018)
Academic Mentor: David McCandlish

Alexa Pagliaro, Wellesley College, Dover, Delaware: B.A. in Neuroscience and Spanish (2018)
Academic Mentor: John Inglis

Jenelys Ruiz, University of Puerto Rico, Rio Piedras: B.S. in Molecular and Cellular Biology (2018)
Academic Mentor: Molly Hammell

Jonathan Werner, University of Maryland, Baltimore County: B.S. in Bioinformatics and Computational Biology (2018); NSF Graduate Research Fellowship Program Award (2018)
Academic Mentor: Adam Siepel

Academic Mentoring

The Watson School takes great pride in the mentoring that it offers its students. One example is our two-tiered mentoring approach, whereby each student chooses both an academic and a research mentor. The academic mentor is a critical advisor during the intensive coursework of the first term, during their rotations, and when identifying a suitable research mentor. Furthermore, the academic mentor continues to follow them throughout their doctoral experience, often serving as important advocates for the students. Entering students select, by mutual agreement, a member of the research or nonresearch faculty to serve as their academic mentor. This program continues

2018 Entering Class: (*Top row from left to right*) Asad Aziz Lakhani, Jenelys Ruiz, Amritha Varshini Hanasoge Somasundara, Marie Dussauze, Jonathan Werner. (*Bottom row from left to right*) Ziyi Mo, Danielle Ciren, Connor Fitzpatrick, Ilgin Ergin, Alexa Pagliaro, King Hei (Teri) Cheng.

to receive much support from the faculty who volunteer to be academic mentors, and it has rightfully become a vital ingredient in our success. The following are the Academic Mentors for the Entering Class of 2018:

STUDENT	MENTOR
King Hei (Teri) Cheng	Adrian Krainer
Danielle Ciren	Ullas Pedmale
Marie Dussauze	Steve Shea
Ilgin Ergin	Tom Gingeras
Connor Fitzpatrick	Rob Martienssen
Amritha Varshini Hanasoge Somasundara	Jason Sheltzer
Asad Aziz Lakhani	David Tuveson
Ziyi Mo	David McCandlish
Alexa Pagliaro	John Inglis
Jenelys Ruiz	Molly Hammell
Jonathan Werner	Adam Siepel

Recruiting Efforts

This year, we once again focused primarily on targeted visits to graduate fairs and minority conferences to identify prospective students for the School. In addition to these visits, the WSBS sends information to top undergraduate science departments around the world. A multiprogram booklet, incorporating the graduate, undergraduate, and postdoctoral programs, was updated for this recruitment season. Additionally, e-mails were sent to personalized contacts and an electronic mailing list of more than 50,000 individuals who receive information from the Cold Spring Harbor Laboratory Press or have attended Meetings or Courses at the Lab. We are grateful to these departments for sharing this contact list. We received 487 applications for the Entering Class of 2019—more than double that for 2018—and it appears that many outstanding candidates have

WATSON SCHOOL OF BIOLOGICAL SCIENCES 2018 RECRUITMENT SCHEDULE

Event	Location	Date
Scientista Symposium, Recruiting Fair	New York, New York	April 13–15
American Association of Cancer Research, Annual Meeting	Chicago, Illinois	April 14–18
Medgar Evers College, Information Session	Brooklyn, New York	April 16
LaGuardia Community College, STEM Research Panel and Careers Event	Long Island City, New York	April 20
California Forum for Diversity in Graduate Education, Graduate School Fair	Stockton, California	April 21
Caldwell University, 2nd Annual Research and Creative Arts Day	Caldwell, New Jersey	April 25
York College of The City University of New York, Undergraduate Research Day	Jamaica, New York	May 3
University of California, Riverside, STEM-California Alliance for Minority Participation (CAMP), Pathway to Graduate School and Research Seminar	Riverside, California	June 8
NIH Graduate and Professional School Fair, Annual Meeting	Bethesda, Maryland	July 18
University at Buffalo McNair Research Conference, Graduate School Fair	Niagara Falls, New York	July 19–21
University of Maryland, Baltimore County, Information Session	Baltimore, Maryland	September 24
University of Maryland, Baltimore County, Meyerhoff Scholars Program, Graduate School Fair	Baltimore, Maryland	September 24
University of Notre Dame, Information Session	Notre Dame, Indiana	September 27
Big 10+ Graduate School Expo, Graduate School Fair	West Lafayette, Indiana	September 30–October 1
Rose-Hulman Institute of Technology, Graduate School Fair	Terre Haute, Indiana	October 1
Community College Day, Resource Fair	San Antonio, Texas	October 10
Hunter College MARC and MBRS/RISE, Information Session	New York, New York	October 10
Society for Advancement of Chicanos and Native Americans in Science (SACNAS), National Conference	San Antonio, Texas	October 11–13
Vanderbilt University, SIM Research Match, Information Fair	Nashville, Tennessee	October 15
American Society for Human Genetics, Annual Meeting	San Diego, California	October 16–20
Vanderbilt University, Scientific Immersion and Mentorship, Information Session	Nashville, Tennessee	October 16
University of Tennessee, Information Session	Knoxville, Tennessee	October 18
Vassar College, Information Session	Poughkeepsie, New York	October 24
University of Notre Dame, Fall Undergraduate Research Fair, College of Science Research Fair	Notre Dame, Indiana	October 25
California Forum for Diversity in Graduate Education, Graduate School Fair	San Diego, California	October 27
Metropolitan Association of Colleges and University Biologists (MACUB) at Queensborough Community College, Annual Meeting	Bayside, New York	October 27
San Diego State University, Women in Science Society, Information Session	San Diego, California	October 30
University of California, San Diego, Information Session	San Diego, California	November 1
Society for Neuroscience Annual Meeting, Graduate School Fair	San Diego, California	November 3–7
Queensborough Community College, Internship & Cooperative Education Forum	Bayside, New York	November 7
Mount Holyoke College, Information Session	South Hadley, Massachusetts	November 8
Annual Biomedical Research Conference for Minority Students (ABRCMS), National Conference	Indianapolis, Indiana	November 14–17
American Society for Cell Biology, Annual Meeting	San Diego, California	December 8–12

once again applied to the program. The large increase in applications is likely due to many factors, but one major change this year was the elimination of the application fee.

Students from Other Institutions

WSBS students account for approximately half of the total graduate student population here at CSHL; the other half comprises visiting graduate students from other universities who have decided to conduct some or all of their thesis research in CSHL faculty members' laboratories. A large fraction of students are from Stony Brook University (SBU), via a program established between CSHL and SBU more than 40 years ago. Over the years, we have built relationships with other institutions around the world, enabling their students to conduct research here at CSHL. Currently, we have visiting students from institutions in Austria, China, the Netherlands, Russia,

Spain, and the United States. The Watson School provides a contact person for the students and maintains relationships with the administrators from their home institutions. These students are fully integrated into the CSHL community and receive all the necessary assistance as they navigate the complexities of performing doctoral research away from their home institutions. The following students, listed in the box below, joined us from SBU this year:

STUDENT	CSHL RESEARCH MENTOR	SBU PROGRAM
Disha Aggarwal	David Spector	Genetics
Seamus Balinth	Alea Mills	Molecular and Cellular Biology
Yuxin Cen	Nicholas Tonks	Molecular and Cellular Biology
Cristian Lopez-Cleary	Christopher Vakoc	Genetics
Margaret Shevik	Mikala Egeblad	Pharmacological Sciences
Joshua Steinberg	Robert Martienssen	Genetics
Gaurang Trivedi	Lingbo Zhang	Genetics
Zhezhen Yu	Michael Wigler	Molecular and Cellular Biology

Graduate Student Symposium

Each year, the students participate in two Graduate Student Symposia held at the Laboratory's Genome Research Center in Woodbury: one in May, the other in October. Each Symposium consists of senior students giving short talks, while coffee breaks and lunch provide opportunities for more informal interactions. The prize for best talk for the May session was awarded to Judith Mizrachi (SBU, Osten lab), and for the October session, it was awarded to Benjamin Berube (WSBS, Martienssen lab) and Luqun Shen (WSBS, Shea lab). We are grateful to the two student chairs—Sashank Pisupati (WSBS) and Padmina Shrestha (SBU)—for their hard work and to WSBS's Kim Geer for helping to make the Symposium a great success.

Graduate Student and Postdoctoral Fellow Departures

With each year come not only new arrivals, but also departures. The following graduate students and postdoctoral fellows left the Laboratory during 2018:

POSTDOCTORAL FELLOWS

Sandra Ahrens	Byoung Il Je	Luis Mejia, Jr.	Jae Seok Roe
Federico Carnevale	Ashley Juavinett	Shinichi Nakamuta	Elena Rozhkova
Cristina Chen	Matthew Kaufman	Dawid Nowak	Juergen Scharner
Luana Guerriero	Virginie Kergourlay	Jean-Sebastien Parent	Mario Shields
Aman Husbands	Young Koung Lee	Michael Regan	Nami Tajima
Chang-Il Hwang	Zachary Lemmon	Daniel Rodriguez-Leal	John Wilson

GRADUATE STUDENTS

Brinda Alagesan	Jesse Levine	Justin Snider
Daniel Kepple	Lei Sheng	Yali Xu

Executive Committee

The School's Executive Committee, in its monthly meetings, provides year-round direction for the School and its students through its invaluable policy recommendations. I wish to thank faculty members Camila dos Santos, John Inglis, Alexander Krasnitz, Bo Li, David Spector, and Anthony Zador for their service in 2018. I would also like to thank student representatives Qian Zhe (SBU) and Emilis Bružas (WSBS), who contributed to discussions and provided useful suggestions and feedback from their colleagues.

The Watson School Continues to Benefit from Generous Benefactors

We are extremely grateful for the generous donors, whose one-time gifts or continued support made our 2018 programs possible, including the Annenberg Foundation, Lita Annenberg Hazen Foundation, Bristol-Myers Squibb Corporation, Mr. and Mrs. Landon Clay, Lester Crown, the Dana Foundation, Henriette and Norris Darrell, the Samuel Freeman Charitable Trust, the William Stamps Farish Fund, Mr. and Mrs. Alan Goldberg, Florence Gould Foundation, Gonzalo Río Arronte Foundation, William Randolph Hearst Foundation, Dr. Nancy Hopkins, Annette Kade Charitable Trust, Dr. James Karam, Mr. David H. Koch, Mr. and Mrs. Robert D. Lindsay and Family, Mr. and Mrs. David Luke III, Mr. and Mrs. William R. Miller, Estate of Edward L. Palmer, the Patrina Foundation, Mr. and Mrs. John C. Phelan, the Quick Family, Mr. and Mrs. Thomas A. Saunders III, Estate of Elisabeth Sloan Livingston, Joan Smith and Jason Sheltzer, the Starr Foundation, the Roy J. Zuckerberg Family Foundation, the Ainslie Foundation, and anonymous donors.

We are also grateful for our endowed lectureships: the John P. and Rita M. Cleary Visiting Lectureship, the George W. Cutting Lectureship, the William Stamps Farish Lectureship, the Martha F. Gerry Visiting Lectureship, the Edward H. Gerry Visiting Lectureship, the Edward H. and Martha F. Gerry Lectureship, the Susan T. and Charles E. Harris Visiting Lectureship, the Klingenstein Lectureship, the Mary D. Lindsay Lectureship, the Pfizer Lectureship, the George B. Rathmann Lectureship, the Seraph Foundation Visiting Lectureship, the Sigi Ziering Lectureship, the Daniel E. Koshland Visiting Lectureship, the Michel David-Weill Visiting Lectureship, and the Fairchild Martindale Visiting Lectureship.

We are also very fortunate to hold a prestigious Ruth L. Krischstein National Research Service Award Predoctoral Training Grant from the National Institutes of Health, National Institute of General Medical Sciences, which was competitively renewed for an additional five years in 2017.

Student and Alumni Achievements

To date, 105 students have received their Ph.D. degree from the WSBS. Twenty-eight graduates have now secured faculty positions. Twelve of them have already been promoted to Associate Professor, and one is a full professor. The following graduates have also moved into influential positions in administration, publishing, consulting, industry, and medicine:

- **Onyekachi Odoemene** joined Booz Allen Hamilton as a data scientist.
- **Ahmet Denli** is an associate editor at *Genome Research*.
- **Paloma Guzzardo** joined Regeneron Pharmaceuticals as manager in R&D Planning, Strategy, and Operations.
- **Colin Malone** co-founded and is the head of Biology at VNV NewCo.
- **Beth Nakasone** matched with the Fred Hutchinson Cancer Research Center as a Hematology/Oncology fellow, a role she will begin in July 2019.
- **Daniel Kepple** joined the Samsung Artificial Intelligence Center as a senior machine learning engineer.
- **Kaja Wasik** and **Stephane Castel** co-founded a biotech company called Variant Bio, based in New York.

In 2018, our current students and alumni were successful in receiving the following prestigious awards and fellowships:

- WSBS graduate **Justus Kebschull** was named the 2018 Larry Katz Memorial Lecturer, Neuronal Circuits meeting, CSHL.
- WSBS students **Kathryn O'Neill** and **Jonathan Werner** were awarded National Science Foundation Graduate Research Fellowships.
- WSBS student **Laura Maiorino** received a JSMF Postdoctoral Fellowship Award in Understanding Dynamic and Multi-Scale Systems.

2018 WSBS STUDENT (CURRENT OR PREVIOUS) PUBLICATIONS

Albrengues J, Shields MA, Ng D, Park CG, Ambrico A, Poindexter ME, Upadhyay P, Uyeminami DL, Pommier A, Küttner V, **Bružas E**, **Maiorino L**, et al. 2018. Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice. *Science* **361**: eaao4227.

Alexander J, Kendall J, McIndoo J, Rodgers L, **Aboukhalil R**, Levy D, Stepansky A, Sun G, Chobardjiev L, Riggs M, et al. 2018. Utility of single-cell genomics in diagnostic evaluation of prostate cancer. *Cancer Res* **78:** 348–358.

Aznarez I, Nomakuchi TT, Tetenbaum-Novatt J, Rahman MA, **Fregoso O**, Rees H, Krainer AR. 2018. Mechanism of nonsense-mediated mRNA decay stimulation by splicing factor SRSF1. *Cell Rep* **23:** 2186–2198.

Banito A, Li X, Laporte AN, Roe JS, Sanchez-Vega F, Huang CH, Dancsok AR, Hatzi K, Chen CC, Tschaharganeh DF, Chandwani R, **Tasdemir N**, et al. 2018. The SS18-SSX oncoprotein hijacks KDM2B-PRC1.1 to drive synovial sarcoma. *Cancer Cell* **33:** 527–541.

Bružas E, Egeblad M. 2018. Bone talk: activated osteoblasts promote lung cancer growth. *Trends Mol Med* **24**: 237–239.

Choi K, Zhao X, Tock AJ, Lambing C, **Underwood CJ**, Hardcastle TJ, Serra H, Kim J, Cho HS, Kim J, et al. 2018. Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. *Genome Res* **28:** 532–546.

Forcier TL, Ayaz A, Gill MS, Jones D, Phillips R, Kinney JB. 2018. Measuring *cis*-regulatory energetics in living cells using allelic manifolds. *Elife* **pii:** e40618.

Han Y, **Kebschull JM**, Campbell RAA, Cowan D, Imhof F, Zador AM, Mrsic-Flogel TD. 2018. The logic of single-cell projections from visual cortex. *Nature* **556:** 51–56.

Ho YJ*, Anaparthy N,* Molik D, Mathew G, Aicher T, Patel A, Hicks J, Gale Hammell M. 2018. Single-cell RNA-seq analysis identifies markers of resistance to targeted BRAF inhibitors in melanoma cell populations. *Genome Res* **38:** 10143–10155.

Kebschull JM, Zador AM. 2018. Cellular barcoding: lineage tracing, screening and beyond. *Nat Methods* **15:** 871–879.

Knott SRV, **Wagenblast E**, Khan S, Kim SY, Soto M, Wagner M, Turgeon MO, Fish L, Erard N, Gable AL, et al. 2018. Asparagine bioavailability governs metastasis in a model of breast cancer. *Nature* **556:** 135.

Naguib A, Mathew G, Reczek C.R, Watrud K, Ambrico A, Herzka T, Salas IC, **Lee MF**, El-Amine N, Zheng W, et al. 2018. Mitochondrial complex I inhibitors expose a vulnerability for selective killing of Pten-Null cells. *Cell Rep* **23:** 58–67.

Nattestad M, Goodwin S, Ng K, Baslan T, Sedlazeck F, Rescheneder P, **Garvin T**, Fang H, Gurtowski J, **Hutton E**, et al. 2018. Complex rearrangements and oncogene amplifications revealed by long-read DNA and RNA sequencing of a breast cancer cell line. *Genome Res* **28:** 1126–1135.

Odoemene O, **Pisupati S**, Nguyen H, Churchland AK. 2018. Visual evidence accumulation guides decision-making in unrestrained mice. *J Neurosci* **38:** 10143–10155.

O'Neill K, Liao WW, Patel A, Hammell MG. 2018. TEsmall identifies small RNAs associated with targeted inhibitor resistance in melanoma. *Front Genet* **9:** 461.

Pommier A, Anaparthy N, Memos N, Kelley ZL, Gouronnec A, **Yan R**, Auffray C, Albrengues J, Egeblad M, Iacobuzio-Donahue CA, et al. 2018. Unresolved endoplasmic reticulum stress engenders immune-resistant, latent pancreatic cancer metastases. *Science* **360:** eaao4908.

Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, **Nattestad M**, von Haeseler A, Schatz MC. 2018. Accurate detection of complex structural variations using single-molecule sequencing. *Nat Methods* **15:** 461–468.

Somerville TDD, Xu Y, Miyabayashi K, Tiriac H, Cleary CR, **Maia-Silva D**, Milazzo JP, Tuveson DA, Vakoc CR. 2018. TP63-mediated enhancer reprogramming drives the squamous subtype of pancreatic ductal adenocarcinoma. *Cell Rep* **25:** 1741–1755.

Sonzogni O, Haynes J, Seifried LA, Kamel YM, Huang K, BeGora MD, Yeung FA, Robert-Tissot C, Heng YJ, Yuan X, Wulf GM, Kron KJ, **Wagenblast E**, et al. 2018. Reporters to mark and eliminate basal or luminal epithelial cells in culture and in vivo. *PLoS Biol* **16:** e2004049.

Underwood CJ, Choi K, Lambing C, Zhao X, Serra H, Borges F, Simorowski J, Ernst E, Jacob Y, Henderson IR, Martienssen RA. 2018. Epigenetic activation of meiotic recombination near *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Res* **28:** 519–531.

*Authors contributed equally to the work. **Boldface indicates Watson School student.**

WSBS GRADUATES IN FACULTY POSITIONS (IN ORDER OF COMPLETION)

Name	Faculty Position
Amy Caudy	Associate Professor, University of Toronto, Canada
Ira Hall	Associate Professor, Washington University in St. Louis, Missouri
Niraj Tolia	Chief, Host–Pathogen Interactions and Structural Vaccinology Section Laboratory of Malaria Immunology and Vaccinology, National Institutes of Health, Bethesda, Maryland
Patrick Paddison	Associate Member, Fred Hutchinson Cancer Research Center, Seattle, Washington
Elizabeth Bartom (nee Thomas)	Assistant Professor, Northwestern University, Illinois
Michelle Heck (nee Cilia)	Research Molecular Biologist, U.S. Department of Agriculture, and Adjunct Assistant Professor, Cornell University, Ithaca, New York
Zachary Lippman	Professor, HHMI Investigator/Cold Spring Harbor Laboratory
Ji-Joon Song	Associate Professor, Korea Advanced Institute of Science and Technology (KAIST), South Korea
Elena Ezhkova	Associate Professor, Mount Sinai School of Medicine, New York, New York
Masafumi Muratani	Associate Professor, University of Tsukuba, Japan
Marco Mangone	Associate Professor, Arizona State University, Phoenix
Elizabeth Murchison	Reader, Cambridge University, United Kingdom
Hiroki Asari	Group leader, EMBL Monterotondo, Rome
François Bolduc	Associate Professor, University of Alberta, Edmonton, Canada
Wei Wei	Associate Professor, University of Chicago, Illinois
Christopher Harvey	Associate Professor, Harvard University, Boston, Massachusetts
Tomas Hromadka	Group Leader, Institute of Neuroimmunology, Slovak Academy of Sciences, Slovakia
Monica Dus	Assistant Professor, University of Michigan, Ann Arbor
Daniel Chitwood	Assistant Professor, Michigan State University, East Lansing
Jeremy Wilsuz	Assistant Professor, University of Pennsylvania, Philadelphia
Oliver Fregoso	Assistant Professor, University of California, Los Angeles
Amy Leung	Assistant Research Professor, City of Hope's Beckman Research Institute, Duarte, California
Hiroshi Makino	Assistant Professor, Nanyang Technological University, Singapore
Katherine McJunkin	Stadtman Tenure Track Investigator, National Institutes of Health, Bethesda, Maryland
Yaniv Erlich	Assistant Professor, New York Genome Center, Columbia University, New York
Michael Pautler	Research Scientist, Vineland Research and Innovation Centre, Vineland Station, Ontario, Canada
Wee Siong Goh	Junior Principal Investigator/GIS Fellow, Genome Institute of Singapore

WSBS GRADUATES IN NONACADEMIC POSITIONS (IN ORDER OF COMPLETION)

Student	Current Position
Emiliano Rial-Verde	Vice President, Food & Ingredients Strategy, Bunge Limited, New York
Ahmet M. Denli	Associate Editor, Genome Research, Cold Spring Harbor Laboratory
Rebecca Ewald	International Business Leader, Ventana Medical Systems/Roche, Tucson, Arizona
Catherine Seiler (nee Cormier)	Senior Manager Biosample Operations, Kaleido Biosciences, Lexington, Massachusetts
Darren Burgess	Senior Editor, *Nature Reviews Genetics*, United Kingdom
Rebecca Bish-Cornelissen	Scientific Director, The Mark Foundation for Cancer Research, New York, New York
Angelique Girard	Director of Finance and Administration, Amplitude Studios, Paris, France
Allison Blum	Account Supervisor, LifeSci Public Relations, LLC, New York, New York
Keisha John	Associate Dean for Diversity and Inclusion, University of Virginia, Charlottesville
Colin Malone	Co-Founder and Head of Biology at VNV NewCo, New York, New York
Oliver Tam	Computational Science Analyst, Cold Spring Harbor Laboratory
Amy Rappaport	Senior Scientist, Gristone Oncology, Emeryville, California
Frederick Rollins	Engagement Manager, LEK Consulting, Boston, Massachusetts
Patrick Finigan	Senior Specialist, Regulatory Affairs CMC Biologics, Merck, Bridgewater, New Jersey

(*continued*)

WSBS GRADUATES IN NONACADEMIC POSITIONS (IN ORDER OF COMPLETION) (*continued*)

Student	Current Position
Elizabeth Nakasone	Resident Physician, UCLA Health, Los Angeles, California
Maria Pineda	Co-Founder, CEO, Envisagenics, New York, New York
Felix Schlesinger	Bioinformatics Scientist, Illumina, Inc., San Diego, California
Paloma Guzzardo	Associate Manager, R&D Planning, Strategy and Operations, Regeneron Pharmaceuticals, Eastview, New York
Saya Ebbesen	Associate Director, Medical + Scientific Strategy at BluPrint Oncology, London, United Kingdom
Joshua Sanders	Founder and C.E.O., Sanworks, L.L.C., Stony Brook, New York
Zinaida Perova	Bioimaging Data Scientist, European Bioinformatics Institute (EMBL-EBI), Cambridge, United Kingdom
Katie Liberatore	Research Scientist, Calyxt, Minneapolis, Minnesota
Kaja Wasik	Co-Founder, Gencove & Variant Bio, New York, New York
Stephane Castel	Co-Founder at Variant Bio and Senior Research Fellow at New York Genome Center, New York
Mitchell Bekritsky	Senior Bioinformatics Scientist, Illumina, Inc., Cambridge, United Kingdom
Sang-Geol Koh	Scientist and Entrepreneur, (Mind), South Korea
Susann Weissmueller	Strategic Partnering Associate, Roche, Switzerland
Ian Peikon	Lead Scientist, Kallyope, New York, New York
Cinthya Zepeda Mendoza	Laboratory Genetics and Genomics Fellow, Mayo Clinic, Rochester, Minnesota
Jack Walleshauser	Research Scientist at Plexxikon Inc., Berkeley, California
Colleen Carlston	Medical student, University of California, San Francisco
Lisa Krug	Scientist, Kallyope, New York, New York
Robert Aboukhalil	Senior Bioinformatics Software Engineer, Genapsys, Redwood City, California
Tyler Garvin	Head of Operations, Stockwell AI Inc., Oakland, California
Anja Hohmann	Senior Scientist, KSQ Therapeutics, Boston, Massachusetts
Charles Underwood	Scientist, KeyGene, Wageningen, Netherlands
Matt Koh	Natural Language Processing Research Scientist, Bloomberg LP, New York, New York
Brittany Cazakoff	Law student, Stanford University, California
Annabel Romero Hernandez	Associate Scientist, Regeneron Pharmaceuticals, Tarrytown, New York
Maria Nattestad	Scientific Visualization Lead, DNA Nexus and Founder, OMGenomics, California
Onyekachi Odoemene	Data Scientist, Booz Allen Hamilton, Maryland
Daniel Kepple	Senior Machine Learning Engineer, Samsung Artificial Intelligence Center, New York
Lital Chartarifsky	Biotechnology and Business Development Coordinator, Cold Spring Harbor Laboratory
Charles Kopec	Associate Professional Specialist, Princeton University, New Jersey (Advisor: Dr. Carlos Brody)
Claudia Feierstein	Research Associate, Champalimaud Neuroscience Programme, Portugal
Gowan Tervo	Research Specialist, Janelia Farms Research Campus, Loudoun County, Virginia
Shu-Ling Chiu	Research Associate, Johns Hopkins University School of Medicine, Baltimore, Maryland (Advisor: Dr. Richard Huganir)
Shraddha Pai	Postdoctoral Fellow, Centre for Addiction and Mental Health, Toronto, Ontario, Canada (Advisor: Gary Bader)
Galen Collins	Postdoctoral Fellow, Harvard Medical School, Boston, Massachusetts (Advisor: Dr. Alfred Goldberg)
David Simpson	Postdoctoral Fellow, Stanford University, California (Advisor: Dr. Alejandro Sweet-Cordero)
Claudio Scuoppo	Instructor, Columbia University, New York (Advisor: Riccardo Dalla-Favera)
Kyle Honegger	Postdoctoral Fellow, Harvard University, Cambridge, Massachusetts (Advisor: Benjamin de Bivort)
Zhenxun Wang	Research Fellow, Genome Institute of Singapore (Advisor: Bing Lim)
Eyal Gruntman	Postdoctoral Associate, Janelia Farm, Ashburn, Virginia (Advisor: Michael Reiser)
Megan Hogan (nee Bodnar)	Senior Research Scientist, New York University Langone Medical Center, New York
Petr Znamenskiy	Senior Research Associate, Sainsbury Wellcome Centre, London, United Kingdom (Advisor: Thomas Mrsic-Flogel)
Yevgeniy Plavskin	Postdoctoral Fellow, New York University, New York (Advisor: Mark Siegal)
Hassana Oyibo	Postdoctoral Fellow, FMI, Basel, Switzerland (Advisor: Georg Keller)

(*continued*)

WSBS GRADUATES IN NONACADEMIC POSITIONS (IN ORDER OF COMPLETION) (*continued*)

Student	Current Position
Marek Kudla	Postdoctoral Fellow, University of California, Berkeley (Advisor: Nick Ingolia)
Melanie Eckersley-Maslin	Postdoctoral Fellow, Cambridge University, United Kingdom (Advisor: Wolf Reik)
Dario Bressan	Research Associate, CRUK, Cambridge (Advisor: Gregory Hannon)
Philippe Batut	Postdoctoral Fellow, Princeton University, New Jersey (Advisor: Michael Levine)
Ozlem Aksoy (nee Mert)	Postdoctoral Fellow, University of California, San Francisco (Advisor: William Weiss)
Nilgun Tasdemir	Postdoctoral Associate, University of Pittsburgh, Pennsylvania (Advisors: Steffi Osterreich and Nancy Davidson)
Elvin Wagenblast	Postdoctoral Fellow, University of Toronto, Ontario, Canada (Advisor: John Dick)
Kristen Delevich	Postdoctoral Fellow, University of Berkeley, California (Advisor: Linda Wilbrecht)
Silvia Fenoglio	Postdoctoral Fellow, Massachusetts Institute of Technology, Cambridge (Advisor: Michael Hemann)
Arkarup Bandyopadhyay	Postdoctoral Fellow, New York University, New York (Advisor: Michael Long)
Justus Kebschull	Postdoctoral Fellow, Stanford University, California (Advisor: Liqun Luo)
Joaquina Delas Vives	Postdoctoral Fellow, Francis Crick Institute, United Kingdom (Advisor: James Briscoe)
Abram Santana	Senior Research Scientist, Columbia University, New York, New York (Advisor: Christine Chio); Postdoctoral Fellow, Harvard University, Boston, Massachusetts (Advisor: Joan Brugge, August 2018)
Fred Marbach	Postdoctoral Fellow, Sainsbury Wellcome Centre, United Kingdom (Advisor: Marcus Stephenson-Jones)
Yu-Jui (Ray) Ho	Computational Biologist, Memorial Sloan Kettering Cancer Center, New York, New York (Advisor: Scott Lowe)
Paul Masset	Postdoctoral Fellow, Harvard University, Boston, Massachusetts (Advisor: Venkatesh Murthy and Naoshige Uchida)
Talitha Forcier	Postdoctoral Fellow, Cold Spring Harbor Laboratory (Advisor: Molly Hammell)

Prizes for the best posters by a postdoctoral fellow and by a graduate student were awarded at the Laboratory's annual In-House Symposium held in November 2018. The poster session provides a forum for the postdoctoral fellows and students to show off their research and gives the entire scientific community a chance to come together and share ideas beyond the walls of their individual laboratories. This year, Nicholas Gallo, a Stony Brook student from Linda Van Aelst's laboratory, and Xiaoli Wu, a Stony Brook student from Chris Vakoc's laboratory, shared the graduate student prize. The postdoctoral prize was shared by Giulia Biffi from David Tuveson's laboratory and Xueyan He from Mikala Egeblad's laboratory.

Alexander Gann
WSBS Professor and Dean

SPRING CURRICULUM

TOPICS IN BIOLOGY

Each year, invited instructors offer week-long courses at the Banbury Conference Center exploring specialized topics outside the expertise of the Cold Spring Harbor Laboratory faculty. These courses include morning and evening lectures, as well as afternoon sessions during which students read assigned papers or work on problem sets and presentations. In spring 2018, there were two courses: Microbial Pathogenesis and Evolution.

Microbial Pathogenesis

February 11–17	**Attended by the entering classes of 2014 and 2015**
INSTRUCTOR	**Stanley Maloy,** San Diego State University
VISITING LECTURERS	**Kelly Doran,** University of Colorado **Yiping Han,** Columbia University **Linda Kenney,** University of Illinois, Chicago
TEACHING FELLOW	**Jeffrey Maloy,** University of California, Los Angeles

Understanding bacterial pathogenesis demands a detailed knowledge of the host response as well as the pathogen itself. Each of these perspectives provides potential strategies for solving important clinical problems. To elucidate these distinct aspects of microbial pathogenesis requires an interdisciplinary approach, integrating the fields of microbiology, cell biology, immunology, biochemistry, genetics, and genomics.

This course focused on mechanisms of microbial pathogenesis and the host response. How do microbes adhere to host cells? How do environmental cues direct the response of microbial pathogens? How do microbial pathogens modulate host cells to expedite virulence? How does the host respond to microbial pathogens? How do new microbial pathogens evolve? How can we thwart microbial pathogens? The course integrated lectures, directed readings of research papers, seminars, and student presentations on various aspects of microbial pathogenesis.

Evolution

February 18–24 **Attended by the entering classes of 2016 and 2017**

INSTRUCTOR **Nipam Patel,** University of California, Berkeley

VISITING LECTURERS **John Huelsenbeck,** University of California, Berkeley
Katya Mack, University of California, Berkeley
Charles Marshall, University of California, Berkeley

TEACHING FELLOW **Rachel Thayer,** University of California, Berkeley

The field of evolutionary biology touches upon all other areas of the biological sciences, because every form of life and every biological process represents an ongoing evolutionary "experiment." Our aim in this course was both to discuss our understanding of the mechanisms of evolution and to explore how evolutionary data can be used to further our understanding of various biological problems.

The course began with a discussion of the diversity of organisms that currently exist and methods for understanding the evolutionary relationships among these organisms. It then went on to study how paleontological data are collected and used to understand the history of life on earth. How DNA sequence data can be used to understand the evolutionary history of organisms, genes, and genomes was also examined. Within this molecular and genetic framework, the focus shifted to the mechanisms of evolutionary change and how variation within populations leads to the evolution of new species. Finally, the course also discussed how morphological changes are brought about through evolutionary changes in development and gene regulation. The hope was that this course would provide the student with a general overview of evolutionary biology and how to use evolutionary data to gain further insight into all manner of biological problems. The course included a class favorite field trip: a curated tour at the American Museum of Natural History in New York City.

SPECIAL COURSES

Optical Methods

January 7–9 **Attended by the entering class of 2017**

INSTRUCTOR **Florin Albeanu,** CSHL

TEACHING FELLOWS **Walter Bast,** CSHL
Priyanka Gupta, CSHL

Optical imaging techniques are widely used in all areas of modern biological research. Our aim for this course was to give students an introduction into widely used basic and advanced optical methods. Given the experimental nature of the topic, a central aim of our course was to offer

students a practical hands-on experience. This included both the use of commercially available systems and, more importantly, a primer on custom building and adapting optical setups to address specific biological needs.

Teaching Experience at the Dolan DNA Learning Center

Entering Class of 2017

DIRECTOR **David A. Micklos**

INSTRUCTORS **Amanda McBrien** (Lead)
Elna Gottlieb
Erin McKechnie
Bruce Nash
Sharon Pepenella

As science plays an increasing role in society, there is also an increasing need for biologists to educate nonscientists of all ages about biology. The WSBS doctoral program offers its students the opportunity to teach in the Laboratory's Dolan DNA Learning Center, where students teach laboratory courses to high school and middle school students. In so doing, they learn how to communicate with nonbiologists and to inspire and educate creative young minds. The teaching module entailed pairs of students teaching one morning or afternoon a week for 12 weeks. In the initial weeks, the Dolan DNA Learning Center instructors taught the Watson School students the didactic process. It was not until the fifth week that the graduate students taught on their own. At the end of the 12 weeks, the students were very excited about their teaching experience.

Laboratory Rotations

Entering Class of 2017

The most important element of a doctoral education is learning to perform independent research. After the fall term courses, students participate in laboratory rotations; these provide students and faculty the opportunity to get to know each other and to explore possibilities for doctoral thesis research. At the end of each rotation, students make short presentations of their studies to the other students and their rotation advisors. These talks give students an opportunity to share their laboratory experiences and to practice giving a scientific presentation. This year, 20 WSBS faculty members served as rotation mentors, some mentoring more than one student.

ROTATION MENTORS

Mickey Atwal	Molly Hammell	Dan Levy	Jason Sheltzer
Anne Churchland	Josh Huang	Bo Li	Adam Siepel
Mikala Egeblad	Dave Jackson	Zachary Lippman	Jessica Tollkuhn
Jesse Gillis	Adam Kepecs	Ullas Pedmale	Chris Vakoc
Tom Gingeras	Jay Lee	Stephen Shea	Linda Van Aelst

FALL CURRICULUM

Entering Class of 2018

The students started the semester by attending boot camps in Molecular, Cellular, and Quantitative Biology as an introduction to the techniques and terminology that they encounter in subsequent courses. The Molecular and Cellular Biology boot camp featured seven lectures from faculty members Hiro Furukawa, Dick McCombie, Pavel Osten, and Lloyd Trotman and from Associate Dean Monn Monn Myat. The Quantitative Biology boot camp lectures were given by Mickey Atwal and Justin Kinney.

CORE COURSES

The Leslie C. Quick, Jr. Core Course on Scientific Reasoning and Logic

INSTRUCTORS	**Linda Van Aelst (Lead)**	**Leemor Joshua-Tor**
	Alexander Gann	**Bo Li**
	Christopher Hammell	
GUEST LECTURERS	**Hiro Furukawa**	**Adrian Krainer**
	Z. Josh Huang	**Robert Martienssen**
	Justin Kinney	
TEACHING ASSISTANTS	**Sonali Bhattacharjee**	**Dhananjay Huilgol**
	Cristina Chen	**Jon Ipsaro**

In this core course, which forms the heart of the curriculum, students (1) acquired a broad base of knowledge about the biological sciences, (2) learned the scientific method, and (3) learned how to think critically. The initial four to five modules are on a different general theme; in each, students read an assigned set of research articles, and at the end of the module, they provided written answers to a problem set that guides them through several of the articles.

Twice weekly, students attended lectures related to the module's topic that included concepts and fundamental information as well as experimental methods. The students met among themselves to discuss the assigned papers not covered by the problem set. Each week, students spent an evening discussing the assigned articles with faculty. In the final module of the course, students participated in a mock study section in which real National Institutes of Health R01 grants are reviewed and critiqued. This allowed students to evaluate the questions before the answers were known, to evaluate routes toward discovery before knowing where they will end, and to make critical judgments about how to proceed in the face of an uncertain outcome.

In 2018, the following were the module topics for this course:

Topic	Instructor(s)
Gene Expression	Alex Gann
Gene Regulatory Logic and the Construction of Multicellular Organisms: Insights from Flies, Plants, and Worms	Christopher Hammell
The Brain: Wiring, Plasticity, and Maladaptation	Bo Li
Macromolecular Structure and Function	Leemor Joshua-Tor
Study Section	Linda Van Aelst

The Darrell Core Course on Scientific Exposition and Ethics

The 2018 Scientific Exposition and Ethics (SEE) core course was again separated into three distinct sections covering writing, oral communication, and ethics. It was taught by Lead instructor David Jackson, along with Sydney Gary, Molly Hammell, and Charla Lambert. As usual, the course hosted expert guest lecturers who covered topics including scientific presentations, ethical and legal responsibilities of scientists, DNA profiling and postconviction appeals, and reproductive genetics.

INSTRUCTORS	**David Jackson (Lead)** **Sydney Gary** **Molly Hammell** **Charla Lambert**
GUEST LECTURERS	**Diane Esposito** **Alyson Kass-Eisler** **Ullas Pedmale** **Rachel Rubino** **Richard Sever**
VISITING LECTURERS	**Keith Baggerly,** MD Anderson Cancer Center **Lydia Franco-Hodges,** Alan Alda Center for Communicating Science **Susan Friedman,** The Innocence Project **Avner Hershlag,** North Shore University Hospital

This core course offered instruction in the fundamental elements of scientific exposition—writing skills and public speaking—and ethics. The ability to communicate effectively and to appreciate the intricacies of ethical issues are essential skills for biologists; both subjects were taught in a series of example-based lectures and discussion groups. Writing skills included the fundamentals of modern scientific English and the organization and preparation of papers, research abstracts, and grant applications. Oral presentation skills were taught by instructors with different modes of presentation. Together with instructors, students critiqued formal seminar presentations at the Laboratory. Instruction and discussions about ethics included the ethical implications of biological discovery on society as well as the nature and boundaries of ethical behavior of scientists and their rights and responsibilities. A primary objective of the course was that students consider exposition and ethics integral aspects of scientific research.

Research Topics

ORGANIZERS **Kimberley Geer**
Alyson Kass-Eisler

This core course provided students with an in-depth introduction to the fields of research that the Laboratory scientists investigate. Students and faculty attended a weekly Research Topics seminar at which faculty members and CSHL fellows presented their current research topics and methods of investigation each Wednesday evening over dinner. The students learned how to approach important problems in biology. These seminars, together with the annual fall Laboratory In-House Symposium, provided students with a basis for selecting laboratories in which to do rotations.

SPECIALIZED DISCIPLINES COURSES

The students in the Entering Class of 2018 took a total of four Specialized Disciplines courses this fall: Quantitative Biology, Genetics and Genomics, Cancer, and Systems Neuroscience.

Quantitative Biology

INSTRUCTOR **Justin Kinney (Lead)**

GUEST LECTURERS **Mickey Atwal**
Tatiana Engel
Alexander Krasnitz
Dan Levy
David McCandlish
Adam Siepel

This course was given throughout the semester. Quantitative reasoning is a powerful tool to uncover and characterize biological principles ranging from the molecular scale all the way to the ecological. With the advent of high-throughput technologies in genomics and neuroscience, it has become increasingly necessary for biological researchers to be able to analyze and interpret high-dimensional data and frame biological hypotheses mathematically. To this end, this course aimed to equip the students with a basic training in computer programing, modern statistical methods, and physical biology. By the end of the course, students not only were able to answer many of the statistical questions that arise in data analyses, but also had become familiar with the more complex techniques used by fellow computational biologists. Topics covered included probabilities, statistical fluctuations, Bayesian inference, significance testing, fluctuations, diffusion, information theory, neural signal processing, dimensional reduction, Monte Carlo methods, population genetics, and DNA sequence analyses. A common theme throughout the course was the use of probabilistic and Bayesian approaches.

Genetics and Genomics

September 5–28

INSTRUCTOR	**Ullas Pedmale (Lead)**
GUEST LECTURER	**Zachary Lippman**
VISITING LECTURER	**Guillaume Lettre,** University of Montreal
TEACHING ASSISTANT	**Benjamin Roche**

This course placed modern genetics and genomics into the context of classical genetics. History, technique, and perspective of genetic inference were described around four levels of analysis: forward genetics, natural genetic variation, gene interaction, and genomics. Emphasis was placed on integrating classical with modern questions of genetic analysis: How are genes mapped and "cloned"? How do gene mutations help to define biological processes? How are more complex traits genetically dissected into their component parts? What concepts and techniques are used to organize genes into pathways and networks? What defines a gene and what gene variation exists in natural populations? What are the functional consequences of gene variation, and how is it detected? How are genomes organized and coordinately regulated? How can genomic information be cataloged, organized, and mined? These questions and concepts were fleshed out using examples from the literature.

Cancer

October 4–26

INSTRUCTORS	**Mikala Egeblad (Co-lead)** **Christopher Vakoc (Co-lead)**
GUEST LECTURERS	**Semir Beyaz** **Tobias Janowitz** **Robert Maki** **Jason Sheltzer** **David Tuveson**
TEACHING ASSISTANT	**Lindsey Baker**

Cancer represents an increasing cause of morbidity and mortality throughout the world as health advances continue to extend the life spans of our populations. Our basic understanding of cancer has increased considerably since 1971, when U.S. President Richard Nixon initiated the "War on Cancer." Specific hypotheses developed from our knowledge of cancer biology are being tested in increasingly complex model systems ranging from cell culture to genetically engineered mouse models, and such investigations should prove invaluable in discovering new methodologies for the detection, management, and treatment of cancer in humans. Importantly, our ability to translate our knowledge of cancer biology into a health benefit for patients is now starting to take form.

At the conclusion of this course, students were able to elaborate an understanding of cancer as a pathobiological process that invades our bodies without offering any known benefit to the host, discuss how cancer progresses, and contemplate how to expand on the methods currently used to treat cancer. Students were also able to design tractable methods to investigate fundamental aspects of cancer biology and became familiar with translational approaches to defeating cancer. Topics covered in this course included hallmarks of cancer, tumor progression, the cancer genome, microenvironment, tumor immunology, metastasis, and approaches to treating cancer, including targeted therapy.

Systems Neuroscience

October 23–November 17

INSTRUCTORS	**Stephen Shea (Lead)** **Florin Albeanu**
GUEST LECTURERS	**Adam Kepecs** **Jessica Tollkuhn**
TEACHING ASSISTANT	**Priyanka Gupta**

This course provided an overview of key aspects of neuroscience, with a focus on learning and plasticity from its cellular basis, through development, to systems and behavior. Both experimental and theoretical viewpoints were explored. There were three main components to the class: lectures, a problem set, and paper presentations.

POSTDOCTORAL PROGRAM

PROGRAM DIRECTOR **Nicholas Tonks**

PROGRAM ADMINISTRATOR **Alyson Kass-Eisler**

An important measure of our postdoctoral program's success is the ability of postdoctoral fellows to secure positions after they complete their training. Recently, our fellows accepted positions at AbbVie, Illinois; Beijing Institute of Genomics, Chinese Academy of Sciences; Case Western Reserve University; Columbia University; Envisagenics; Kings College London, United Kingdom; Queen's University, Ireland; Ohio State University; Ottawa Research and Development Centre, Canada; Penn State University; Pfizer; Sanofi; Salk Institute for Biological Studies; University of California, Davis; University of California, San Diego; and University of Chicago.

Postdoctoral Liaison Committee

The Postdoctoral Liaison Committee (PDLC), which is an elected group of postdoctoral fellows who communicate information and ideas between the administration and the postdoctoral community, continues to enhance CSHL's postdoctoral experience. The PDLC is essentially the voice of the community, and it holds regular meetings with Dr. Bruce Stillman, CSHL President. The current PDLC members are Leah Banks, Debarati Ghosh, Nicholas Gladman, Prabhadevi Venkataramani, and Min Yao. The PDLC hosted a retreat on September 21 at the Banbury Conference Center. The agenda included research talks from CSHL postdocs; a mock chalk talk; a panel discussion with former postdocs who are now in nonacademic positions; and an R-course for beginners.

The PDLC, along with the CSHL Library, continued the very successful Perspectives on Science Careers series that was started last year. Recent guest speakers included David Kolosic, Regulatory Information Specialist and Business Developer, Asphalion, Spain, and CSHL alumnus Dr. Fatih Mercan, Scully, Scott, Murphy & Presser, who spoke about his transition to a career in intellectual property (IP) law. The PDLC worked with the Office of Sponsored Programs (OSP) to put together a panel of CSHL fellowship recipients who shared their insights on applying for their individual fellowships. Dhananjay Huilgol (NARSAD), Bor-Shuen Wang (F32), and Roger Tseng (NSF) each participated.

The PDLC also oversees and distributes funds provided by Dr. Stillman to two career development groups: the Career Development Program and the Bioscience Enterprise Club. These groups are primarily composed of postdoctoral fellows, but they also include graduate students. Today's postdoctoral fellows face a number of challenges, including a very difficult and competitive job market. CSHL endeavors to prepare postdocs to be competitive for the scarce number of jobs available. It is increasingly becoming CSHL's role to introduce the diversity of career opportunities available and to provide the tools postdocs need to prepare for these positions. As a result, a number of events were organized with the assistance of the PDLC and career development groups.

Career Development Program

The Career Development Program (CDP) provides programming geared toward careers in academia. The *Coffee Chat* lecture series, at which postdocs are provided career insights within an informal and interpersonal format, was initiated by faculty member Ullas Pedmale and was coorganized by postdoc Leah Banks. The 2018 sessions were hosted by faculty members Alea Mills and Anthony Zador and CSHL Fellow Jason Sheltzer.

The CDP has also been bolstering connections with local colleges and universities in order to provide teaching opportunities for interested CSHL postdocs. Informal partnerships have now been established with Adelphi University, Molloy College, and Long Island University. In addition, selected postdoctoral fellows continue to participate in the Watson School as tutors—either one-on-one or in the classroom setting—thus providing CSHL postdocs with valuable experience in teaching and mentoring.

Bioscience Enterprise Club

The Bioscience Enterprise Club (BEC) disseminates information about nonacademic careers to the CSHL postdoc community. Topic areas include biotechnology, intellectual property, law, regulatory affairs, and venture capitalism. The Club hosted a number of guest speakers recently. Dr. Bernadette So (New York University) spoke about preparing résumés for private sector job applications and also held a résumé feedback session. Dr. Geneviève Smith (from the Insight Data Science fellowship program in New York City) talked about careers in data science, health data, and artificial intelligence. Dr. Michael Gilman (CEO and Director, Obsidian Therapeutics; Co-founder, X4; Board Chair, Arrakis Therapeutics, and a CSHL alumnus) talked about entrepreneurship and gave advice on how to transition to a career in biotech. Drs. Lisa A. Napolione, Kurt Schilling, and Klodjan Stafa (from the Estée Lauder Company) also shared their experience in transitioning to careers in industry as part of a panel discussion and networking event. Mr. William Doyle discussed his career journey that led him to the position of Chairman of the Board of Novocure and shared his insights into venture capital, investment banking, and the best way to transition from bench science to roles outside of academia. BEC along with a new CSHL club, the Consulting and Investing Club (COIN), hosted seven attorneys from Dilworth & Barrese, a distinguished law firm on Long Island that has previously hosted internships for CSHL postdocs. They discussed their own career paths into IP law as well as opportunities and requirements associated with transitioning into this career. Dr. Christine Ardito-Abraham, a former postdoc at CSHL, talked about her path to becoming an oncology Medical Science Liaison at Amgen.

BEC also organized a trip to Regeneron Pharmaceuticals enabling postdocs to participate in the Regeneron Science to Medicine Forum. With the assistance of the WSBS's Associate Dean, Monn Monn Myat, a visit to Google's New York City offices was organized to give students and postdocs a taste of what working at Google is like and the current employment opportunities. A guided tour was led by former Undergraduate Research Program participant and current Google employee Paul Baranay.

Finally, BEC and the PDLC sponsored two teams of students and postdocs to participate in the Tufts New England Case Competition. One team, composed of graduate students Katie Meze, Deborah Rupert, and Jue Wang and postdocs Sonali Bhattacharjee and Johanna Syrjanen, won first place, beating out teams from Harvard, MIT, and Tufts.

Social Events

This year, the Laboratory has paid special attention to the social needs of the postdoctoral community. The vast majority of CSHL postdocs are not from Long Island and do not have a built-in social network. Alyson Kass-Eisler, our Program Administrator, is working with the PDLC to identify a group of postdoc "buddies" to have lunch with a new postdoc shortly after that postdoc starts at the Laboratory. This will give new postdocs an immediate peer and point of contact. We have also been hosting a monthly postdoc lunch table where postdocs can join other postdocs at a designated place and time. A postdoc BBQ was held at the beginning of summer and was a huge

success. The PDLC recently sent the community a survey to gauge the kinds of social events they would be most interested in attending.

National Postdoc Appreciation Week

Annual ice cream socials are held on CSHL's campus to celebrate National Postdoc Appreciation week. These events provide a great opportunity for the community to join together for some fun, but also to network and learn about CSHL's ongoing programs.

The Science Alliance

All CSHL postdoctoral fellows and graduate students are enrolled in a special initiative of the New York Academy of Science (NYAS), the Science Alliance. The Science Alliance for graduate students and postdoctoral fellows is a consortium of universities, teaching hospitals, and independent research facilities in the New York City metro area that have formed a partnership with NYAS. The Alliance's aim is to provide career and professional development mentoring for postdoctoral fellows and graduate students in science and engineering, through a series of live events and a dedicated web portal. In addition, the Science Alliance offers graduate students and postdoctoral fellows the opportunity to network with their peers across institutions and with key leaders in academia and industry.

The Science Alliance meetings, courses, and workshops this year discussed the following topics: Know Your Rights: Legal Literacy for Scientists; STEM Exchange: Research and Career Symposium–PepsiCo's R&D; Using Informational Interviews to Find Your Dream Job (Webinar); Stories Behind the CV: A Career Exploration Town Hall (Webinar); Beating the STEM Burnout (Webinar); Immigration Info Session for Advanced Degree Holders in STEM (Webinar); Using Informational Interviews to Find Your Dream Job; How to Build Your All Star Resume and LinkedIn Profile; Thinking About Teaching; Scientists Teaching Science Online Course; Grantsmanship for Students and Postdocs: Pathways to Individual Fellowship; and Introductory Coding for Researchers.

WiSE and DIAS

There are two groups on campus dedicated to promoting diversity, inclusion, and equity in science. The Women in Science and Engineering (WiSE) group was established a few years ago to create a strong and collaborative support system for women scientists at CSHL and beyond. To address challenges disproportionally affecting women in STEM, WiSE provides a platform for professional development and empowerment through mentorship, career planning, community outreach, and educational opportunities. They recently held a #WomenInSTEM Wikipedia edit-a-thon to create and update the Wikipedia pages of prominent female scientists. Along with BEC, WiSE hosted a workshop on "How to Negotiate for What You Need" by Dr. Sandra Masur, Director of the Office for Women's Careers at the Icahn School of Medicine at Mount Sinai.

The CSHL DIAS (Diversity Initiative for the Advancement of STEM) is an on-campus organization of underrepresented minority (URM) scientists. DIAS is broadly interested in raising awareness and inclusivity at CSHL, hosting on-campus seminars by prominent URM speakers, and providing outreach to nearby community colleges. For their inaugural event, they hosted Dr. Medeva Ghee, Executive Director of the Leadership Alliance, a consortium of 35 institutions that are committed to supporting underrepresented minorities in higher education and leadership positions.

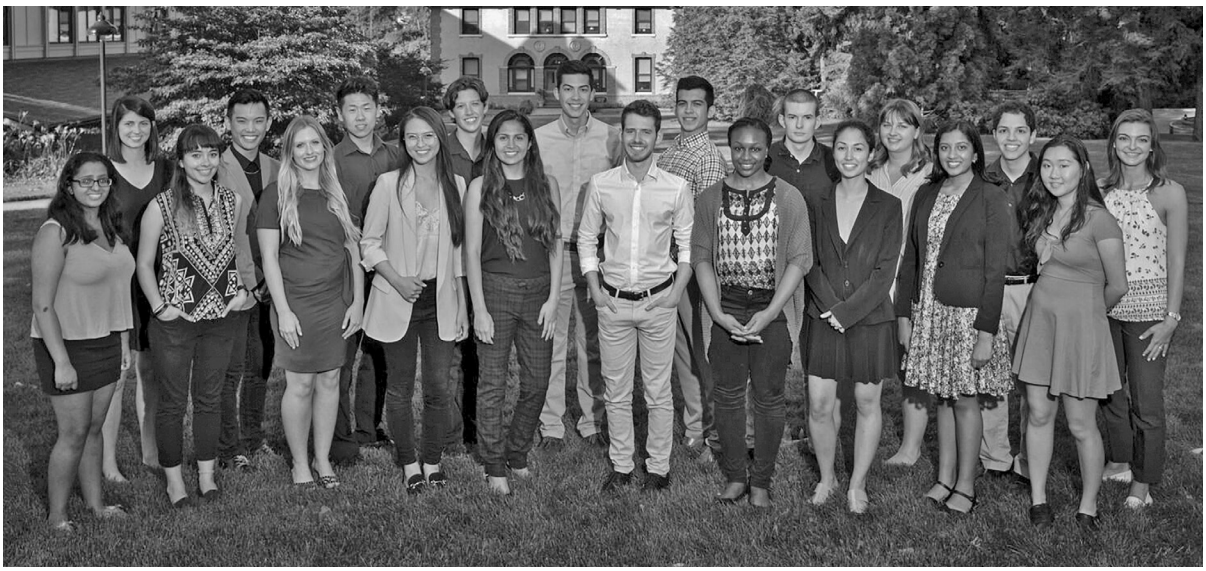
UNDERGRADUATE RESEARCH PROGRAM

PROGRAM DIRECTORS **Jesse Gillis**
Christopher Hammell

PROGRAM ADMINISTRATOR **Kimberly Creteur**

Established almost 60 years ago, the CSHL Undergraduate Research Program (URP) provides undergraduates from around the world with hands-on undergraduate research training in biology. The 10-week program begins the first week of June. Several activities are implemented to ensure that URP participants transition smoothly into the Laboratory community and research. For example, during the first week, the students attend various orientations and receive a guided historical tour of campus and all the facilities and resources available to them. The URPs work, live, eat, and play among CSHL scientists and have a very busy academic and social calendar throughout the remaining 9 weeks of the summer. The students receive training in scientific research, science communication, career preparation, and bioinformatics and computational biology, all while interacting socially with fellow program participants and members of the CSHL community in formal and informal activities. Some of the 2018 activities included the Director's Tea, a pizza party with Dr. and Mrs. Watson, dinner with Dr. and Mrs. Stillman, volleyball games, designing the URP T-shirt, a Broadway show, a scavenger hunt, and the ever-famous URP versus PI volleyball match.

The students' scientific development is the most important component of the program. At the beginning of the summer, the URP students write an abstract and present a talk on their proposed research. The URP participants work alongside scientists and become increasingly independent throughout the summer. Concluding the program in August, the URP students prepare a final report and present their results in a 15-minute talk at the URP Symposium. As in previous years, the program directors and faculty mentors were highly impressed with the accomplishments of the URP students.



2018 Undergraduate Research Program Participants

The following 20 students, selected from 672 applicants, took part in the 2018 program:

Chimsom Agbim
Advisor: **Trotman Laboratory**
Funding: **William Townsend Porter Foundation Scholar**
Exosomes as therapeutic biomarkers in prostate cancer.

Anisha Babu
Advisor: **Albeanu Laboratory**
Funding: **National Science Foundation Scholar**
Revolving odor delivery machine.

Basheer Becerra
Advisor: **Krasnitz Laboratory**
Funding: **National Science Foundation Scholar**
A bioinformatics pipeline for copy-number feature extraction used for predicting tumor organoid chemosensitivity.

Gavriela Carver
Advisor: **Jackson Laboratory**
Funding: **Former URP Fund Scholar/30th Anniversary URP Scholar**
Investigating the roles of trehalose-6-phosphate phosphatases in plant development.

George Chen
Advisor: **Li Laboratory**
Funding: **Libby Fellowship/William Shakespeare Fellowship**
Monitoring home-cage behavior of 16p11.2$^{+/-}$ mice to determine their viability as a model for autism in humans.

Kevin Chen
Advisor: **Stillman Laboratory**
Funding: **Robert H.P. Olney Fellow/Garfield Fellowship**
Role of ORC4 and ORC2 in ORC origin sequence specificity.

Sterling Evans
Advisor: **Martienssen Laboratory**
Funding: **Katya H. Davey Fellowship**
Understanding easiRNA in *Arabidopsis thaliana* pollen.

Itai Levin
Advisor: **Zador Laboratory**
Funding: **National Science Foundation Scholar**
Mapping neuronal projections from the mouse periaqueductal gray.

Matheo Morales
Advisor: **Spector Laboratory**
Funding: **National Science Foundation Scholar**
Characterization of long noncoding RNA lnc10 in neurodevelopment.

Noelle Ozimek
Advisor: **Osten Laboratory**
Funding: **National Science Foundation Scholar**
C-fos screening for cellular resolution mapping of behaviorally evoked whole-brain activation in APPSWE mouse model.

Matthew Peacey
Advisor: **Pedmale Laboratory**
Funding: **James D. Watson Undergraduate Scholar**
Investigating interaction of CRY2 with MOS1 and FVE in the cryptochrome signaling pathway.

Daniel Quintero
Advisor: **Churchland Laboratory**
Funding: **National Science Foundation Scholar**
Looming and receding stimuli influence innate defensive behaviors and neuron populations.

Tess Rinaldo
Advisor: **Shea Laboratory**
Funding: **National Science Foundation Scholar**
Analyzing excitatory and inhibitory neuronal response in the basal amygdala during the learning of maternal retrieval behavior.

Nicole Sivetz
Advisor: **Krainer Laboratory**
Funding: **Alfred L. Goldberg Fellowship/Joan Redmond Read Fellowship**
Inhibition of nonsense-mediated mRNA decay of the *CFTR* gene using splice-switching antisense oligonucleotides.

Lauren Stiene
Advisor: **Egeblad Laboratory**
Funding: **Von Stade Fellowship/Burroughs Wellcome Fellowship**
Evaluating intratumoral clonal heterogeneity in breast cancer by multicolor lineage tracing.

Maya Talukdar
Advisor: **Gillis Laboratory**
Funding: **National Science Foundation Scholar**
Shaping our understanding of transcriptional network rewiring via V-shaped relationships.

Amelia Tian
Advisor: **Tollkuhn Laboratory**
Funding: **National Science Foundation Scholar**
Validation of hormone-regulated differential gene expression in mice and voles.

Macy Vollbrecht
Advisor: **dos Santos Laboratory**
Funding: **James D. Watson Fellow/Dorcas Cummings Scholar**
Investigating how the immune system affects the development, proliferation, and transcriptome of mammary epithelial cells.

Magdalene Walters
Advisor: **McCandlish Laboratory**
Funding: **University of Notre Dame URP Scholar**
Computational detection and characterization of epistatic interactions in influenza A hemagglutinin protein through surveillance sequencing and deep mutational scanning data.

Julia Wang
Advisor: **Engel Laboratory**
Funding: **National Science Foundation Scholar**
Coordination of cortical state between frontal and visual cortex during spatial attention.

SUMMER RESEARCH INTERNSHIP FOR MEDICAL STUDENTS

PROGRAM ADMINISTRATOR	**Jessica Gotterer**
PROGRAM DIRECTOR	**Priya Sridevi**
PROGRAM COORDINATOR	**Rachel Desrochers**

In the past, CSHL has offered incoming second-year medical students from the Donald and Barbara Zucker School of Medicine at Hofstra/Northwell the opportunity to experience the latest in basic and translational research. In 2018, CSHL piloted SRIMS, in which students not only participate in a summer research project but are also mentored by one of CSHL's outstanding faculty as well as the Program Director, Priya Sridevi, Ph.D. In addition, students participate in a welcome lunch with President and CEO Bruce Stillman, Ph.D., attend weekly discussions with the Program Director, get hands-on experience in programming workshops, receive training in scientific communication, attend targeted lecture series, and are trained in the Responsible Conduct of Research. This 8- to 10-week program gives students a glimpse into the world of research, focusing on cancer, neuroscience, and the genetics of human disease.

The following student took part in the 2018 program:

STUDENT	CSHL MENTOR
Danielle Soberman	Jason Sheltzer

PARTNERS FOR THE FUTURE

PROGRAM DIRECTOR **David Jackson**

PROGRAM ADMINISTRATOR **Lynn Carmen**

The Partners for the Future Program, established by Dr. James Watson in 1990, provides an opportunity for talented Long Island high school students to have hands-on experience in biomedical research at Cold Spring Harbor Laboratory. This highly competitive program is open to Long Island high school students in their junior year. Each high school science chairperson may nominate up to three students. The top candidates are interviewed by CSHL scientists. Students selected for the program are paired with a scientist mentor and spend a minimum of 10 hours per week, September through March of their senior year, conducting original research. At the conclusion, the students present their projects to an enthusiastic audience of the students, scientific mentors and colleagues, CSHL administrators, parents, and teachers. Although the students learn a great deal about modern biology and state-of-the-art research techniques, the main advantage of the program is that they are exposed to day-to-day life in a lab. Interacting with scientists and support staff, the students learn to define and pursue a research goal while solving problems that may occur along the way.

The following 2018–2019 Partners for the Future were chosen from among 50 nominations:

NAME	HIGH SCHOOL	LABORATORY	MENTOR
Brooke Bier	Syosset High School	Semir Beyaz	Semir Beyaz
Michael Biggiani	Oyster Bay High School	Ullas Pedmale	Olya Spassibojko
Gabriel Chan	Commack High School	Tatiana Engel	Tatiana Engel
Cindy Cheng	Hicksville High School	Robert Martienssen	Hyun Soo
Sinar Deniz	Commack High School	Robert Martienssen	Sonali Bhattacharjee
Ariana Fang	Massapequa High School	Jonathan Preall	Jonathan Preall
Rochelle Kris	Harborfields High School	Lloyd Trotman	Kaitlin Watrud
Jessica Layne	Oyster Bay High School	Adrian Krainer	Adrian Krainer
Kaitlin Long	Kellenberg High School	Jason Sheltzer	Jason Sheltzer
Joseph Massa	Cold Spring Harbor High School	Tatiana Engel	Mikhail Genkin
Jack McCooey	Friends Academy	Robert Martienssen	Benjamin Roche
Pragati Muthukumar	Commack High School	Doreen Ware	Doreen Ware
Noa Sussman	Cold Spring Harbor High School	Jessica Tollkuhn	Jessica Tollkuhn
Matthew Venezia	Lynbrook High School	David Jackson	Fang Xu
Nathaniel Wang	Northport High School	Bo Li	Xian Zhang

MEETINGS & COURSES PROGRAM

ACADEMIC AFFAIRS

The Cold Spring Harbor Laboratory meetings and courses program attracts scientists from all over the world to communicate, learn, and critique the latest ideas and approaches in the biological sciences. The meetings and courses program at the Laboratory attracted strong attendance in 2018, with 7,300 meeting participants and more than 1,400 course participants (trainees, teaching, and support faculty). The Cold Spring Harbor Asia program included 17 conferences and attracted more than 3,500 participants, bringing the anticipated year-end total for both United States– and China-based programs to almost 12,000.

The Laboratory held more than 30 academic meetings this year, which brought together scientists from around the world to discuss their latest research. The spring meeting season culminated in the 83rd Cold Spring Harbor Symposium, which focused on Brains and Behavior: Order and Disorder in the Nervous System, addressed differing themes in brain research, and explored the tremendous recent progress in neuroscience and technologies and how these advances may be used to improve brain health and address psychiatric and neurological disorders. The Symposium attracted almost 300 participants, including notable scientists such as David Anderson, Hugo Bellen, Antonello Bonci, Edward Boyden, Yang Dan, Karl Deisseroth, Ricardo Dolmetsch, Frances Edwards, Joshua Gordon, Richard Huganir, Anatol Kreitzer, Andres Lozano, Robert Malenka, Helen Mayberg, P. Read Montague, Clifford Saper, Amita Sehgal, Karel Svoboda, Beth Stevens, Kevin Tracey, Kay Tye, Flora Vaccarino, Mehmet Fatih Yanik, and Huda Zoghbi, to name but a few. Dissemination includes the proceedings of the Symposium, published each year by the CSHL Press, and videotaped interviews with leading speakers conducted by editors and journalists attending the Symposium now available on our Leading Strand YouTube channel. The Symposium therefore reaches a much wider audience nationally and internationally than can possibly attend.

Cold Spring Harbor Laboratory meetings are unique in that organizers are encouraged to select talks from abstracts that are submitted on the basis of scientific merit, ensuring that the latest findings will be presented and that young scientists will have the chance to describe their work. The year 2018 saw the continuation of many successful annual and biennial meetings as well as the introduction of several new meetings, including Single Biomolecules and Nutrient Signaling. The CSHL Genentech Center Conferences on the History of Molecular Biology and Biotechnology this year addressed The Evolving Concept of Mitochondria: From Symbiotic Origins to Therapeutic Opportunities. The program featured many notable speakers, including Alexey Amunts, Navdeep Chandel, Marcia Haigis, Carla Koehler, Nils-Göran Larsson, Jennifer Lippincott-Schwartz, Agnes Rötig, L'institut Imagine, Franc, Jared Rutter, Bruce Spiegelman, Anu Suomalainen, John Walker, Douglas Wallace, Richard Youle, and Massimo Zeviani. Partial support for individual meetings is provided by grants from the National Institutes of Health, the National Science Foundation, other foundations, and companies. Core support for the meetings program is provided by the Laboratory's Corporate Sponsor Program.

The Courses program covers a diverse array of topics in molecular biology, neurobiology, structural studies, and bioinformatics. Instructors update their course curricula annually, invite new speakers who bring a fresh perspective, and introduce new techniques and experimental approaches based on student feedback and progress in the field. New techniques—for example, genome editing using tools such as CRISPR or super-resolution microscopy—are introduced as methodologies develop and evolve. We strongly encourage each course to include the latest technical and conceptual developments in their respective fields. Instructors, lecturers, and assistants come from universities, medical schools, research institutes, and companies around the world to teach at Cold Spring Harbor. Students include advanced graduate students, postdoctoral trainees, and principal investigators and senior scientists from around the world. Their excellence and

dedication make the course program work so uniquely well. We would especially like to thank Drs. Benjamin Allen, Tiago Branco, Sean Cutler, William Furey, Karmella Haynes, Mustafa Khokha, Tuuli Lappalainen, Karen Liu, Alexander McPherson, Katalin Medzihradszky, Amy Ralston, and David Veesler, whose exemplary teaching and leadership of their respective courses have benefitted so many young scientists. We should especially single out for thanks Drs. Furey and McPherson, who have taught the X-ray crystallography course for more than two decades.

Grants from a variety of sources support the courses. The core support grants provided through the Helmsley Charitable Trust and Howard Hughes Medical Institute are critical to our course program. The courses are further supported by multiple awards from the National Institutes of Health and the National Science Foundation, and additional support for individual courses is provided by various foundations and other sources. The courses also depend on equipment and reagents that are loaned or donated by a large number of companies—partnerships that are invaluable in ensuring that the courses offer training in the latest technologies.

Now in its ninth year of operation, the Cold Spring Harbor Asia (CSHA) program, under the direction of Dr. Maoyen Chi, is headquartered at the Suzhou Dushu Lake Conference Center in the Suzhou Innovation Park high-technology suburb. In 2018, 17 scientific conferences and two summer schools were held in Suzhou. CSHA's scientific program is designed for scientists from the Asia/Pacific region, who make up more than 75% of attendance, and include symposia and meetings, training workshops, and occasional Banbury-style discussion meetings. This program is described in more detail in a separate Annual Report.

Special events included the first Double Helix Day event on Forensics & DNA in February, several bioentrepreneur networking events, and numerous local area one- and two-day retreats—including several affiliated with the Feinstein Institute and Northwell Health. Although distinct from our regular academic program, these events attract significant numbers of leaders and individuals associated with biomedicine and bio-business from the tri-state area and beyond.

The meetings and courses program staff comprises a diverse team of talented professionals who handle the complexities of database design, programming, web and multimedia design, educational grants management, marketing and recruitment, conference and course administration, audiovisual and digital design services, and other activities. We said goodbye to Nicholas Moore and Samuel Stewart, after many years of devoted service to CSHL. We also welcomed several new staff in 2018, including Ira Russo into the Audio Visual Department, who are already bringing a high level of professionalism to their positions.

David Stewart
Executive Director,
Meetings & Courses Program
President, Cold Spring Harbor Asia

Academic guidance

Terri Grodzicker
Dean of Academic Affairs

CSH ASIA SUMMARY OF CONFERENCES

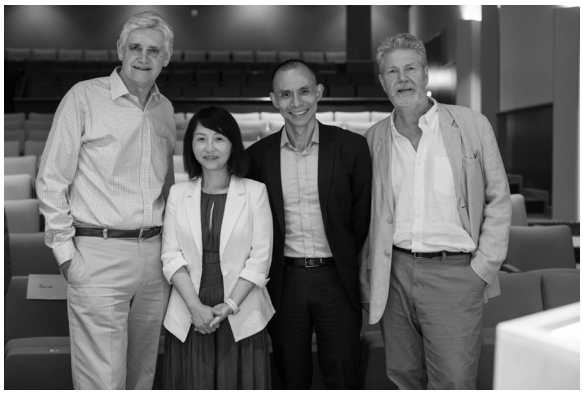
Dates	*Title*	*Organizers*
March 26–30	Cancer and Metabolism	Navdeep Chandel, Hozumi Motohashi, (Hongyuan) Rob Yang, Shimin Zhao
April 9–13	Ubiquitin Family, Autophagy, and Diseases	Ivan Dikic, Jianping Jin, Xiaobo Qiu, Hong Zhang
April 16–20	Chromatin, Epigenetics, and Transcription	Genevieve Almouzni, Hiroyuki Sasaki, Yang Shi, Bing Zhu
April 23–27	Genome Editing: All Things Considered	Jens Boch, Jin-Soo Kim, Wensheng Wei
May 7–11	Stem Cell Crossroads	Duanqing Pei, Austin Smith, Yasuhiro Yamada
June 11–15	DNA Metabolism, Genomic Stability, and Human Disease	Antony Carr, Daochun Kong, Binghui Shen, Patrick Sung, Stephen West
June 26–July 7	Summer School: Building and Mining Brain Cell Atlases and Connectomes	Giorgio Ascoli, Z. Josh Huang, Qingming Luo, Partha Mitra
June 26–July 19	Summer School: Computational and Cognitive Neuroscience	Eric DeWitt, Zachary Mainen, John Murray, Xiao-Jing Wang, Si Wu
September 3–7	Frontiers in Computational Biology and Bioinformatics	Steven Brenner, Keith Dunker, Julian Gough, Luhua Lai, Yunlong Liu
September 10–14	Frontiers of Immunology in Health and Disease	Xuetao Cao, Richard Flavell, Fabienne Mackay, Tadatsugu Taniguichi
September 17–21	Autism and Neurodevelopment Disorders—From Genetic Discoveries to Interventions	Yong-Hui Jiang, Zilong Qiu, Jonathan Sebat, Toru Takumi
September 25–28	Latest Advances in Development and Function of Neuronal Circuits	Frank Bradke, Kazuo Emoto, Hailan Hu, Alcino Silva
October 8–12	Systems Biology of Gene Regulation and Genome Editing	Xiaole Shirley Liu, Huck-Hui Ng, Len Pennacchio, Bing Ren
October 15–19	Advances in Optical Imaging of Living Cells and Organisms: Focus on the Brain	Guoqiang Bi, Haruhiko Bito, Arthur Konnerth, Yi Zuo
October 22–26	Telomeres and Telomerase	Alison Bertuch, Eric Gilson, Ming Lei, Roger Reddel, Jerry Shay
October 29–November 2	RNA Biology	Ling-Ling Chen, Xiang-Dong Fu, Masatoshi Hagiwara, Adrian Krainer
November 5–9	Frontiers in Single Cell Genomics	John Marioni, Nicholas Navin, Fuchou Tang
November 12–16	Biology of Extracellular Vesicles	Raghu Kalluri, Graca Raposo, Li Yu
November 26–30	Iron, Reactive Oxygen Species, and Ferroptosis in Life, Death, and Disease	Quan Chen, Marcus Conrad, Xuejun Jiang, Brent Stockwell
December 3–7	Novel Insights in Glia Function and Dysfunction	Shumin Duan, Helmut Kettenmann, Brian MacVicar, Mami Noda, Long-Jun Wu

83RD COLD SPRING HARBOR LABORATORY SYMPOSIUM ON QUANTITATIVE BIOLOGY

Brains and Behavior: Order and Disorder in the Nervous System

May 30–June 4 294 Participants

ARRANGED BY **David Stewart and Bruce Stillman,** Cold Spring Harbor Laboratory



B. Stillman, Y. Li, D. Tan, D. Stewart

The 83rd Cold Spring Harbor Symposium focused on Brains and Behavior: Order and Disorder in the Nervous System. Previous Cold Spring Harbor Symposia on Function and Dysfunction in the Nervous System in 1996 and Cognition in 2014 explored differing themes in brain research, and the 2018 Symposium built on these successes in exploring the tremendous recent progress in neuroscience and technologies and how these advances may be used to improve brain health and address psychiatric and neurological disorders. Topics addressed at the Symposium included cellular basis of neurodegeneration, memory systems and aging, neural circuit of anxiety, addicted brain, circuits and treatment frontiers, sleep and sleep disorders, hormonal regulation and mood disorders, restoring vision, corticostriatal circuits in perception and action, deep brain stimulation for psychiatry, computational psychiatry, development of social cognition, brain–machine interface, and new techniques for human brain manipulation.

The Symposium attracted more than 290 participants and provided an extraordinary 5-day synthesis of current understanding in the field. Opening night talks setting the scene for later sessions included Mehmet Yanik on engineering brain activity patterns for therapeutics of disorders, Daphna Shohamy on how memory guides value-based decisions, Daniel Wolpert on probabilistic models of sensorimotor control and decision, and Antonello Bonci on synaptic plasticity,

H. Zoghbi interviewed by J. Witkowski

D. Bambah, M.A. Devineni

A. Kepecs, Z. Mainen

T. Pinkhasov, D. Rupert

optogenetics, and a novel treatment against cocaine use disorders. P. Read Montague delivered a fascinating Dorcas Cummings Lecture on "Connecting Mind and Brain in a Computational Age" for the Laboratory's friends and neighbors. Rising to the challenging task of condensing more than 50 talks over the prior five days, Adam Kepecs provided a masterly summary of the state of the field at the conclusion of the Symposium. Interviews with leading scientists captured during the Symposium provide a snapshot of the state of current research and are available on the CSHL Leading Strand channel (https://www.youtube.com/user/LeadingStrand).

Thank you to the Tianqiao & Chrissy Chen Institute for major support for this Symposium.

PROGRAM

Introduction
B. Stillman, *Cold Spring Harbor Laboratory*

Psychiatric Neuroscience
Chairperson: F. Edwards, *University College London, United Kingdom*

Sleep and Neuromodulation
Chairperson: L-H. Tsai, *Massachusetts Institute of Technology, Cambridge*

Decisions and Motor Control/Genetics
Chairperson: A. Seghal, *HHMI/University of Pennsylvania, Philadelphia*

Rapid Fire Talks: Chen Fellows
Chairpersons: Y. Li, *Tianqiao & Chrissy Chen Institute, Menlo Park, California;* D. Stewart, *Cold Spring Harbor Laboratory*

Neurological Health
Chairperson: Y. Dan, *University of California, Berkeley*

Neurodevelopment and Synapses
Chairperson: D. Anderson, *HHMI/California Institute of Technology, Pasadena*

Social Behavior
Chairperson: B. Stevens, *Children's Hospital/Harvard Medical School, Boston, Massachusetts*

K. Matlik, E. Nasiri, A. Zarfeshani

K. Deisseroth, B. Li

B. Stevens, D. Anderson

M. Platt, P.R. Montague

H. Hu, G. Gasque

Dorcas Cummings Lecture: Connecting Mind and Brain in a Computational Age
P.R. Montague, *Virginia Tech Carilion Research Institute, Roanoke*

Treatments
Chairperson: A. Zador, *Cold Spring Harbor Laboratory*

Anatomy, Connectivity, and Neurotechnology
Chairperson: H. Mayberg, *Icahn School of Medicine at Mount Sinai, New York*

Summary
A. Kepecs, *Cold Spring Harbor Laboratory*

MEETINGS

Double Helix Day: Forensics and DNA

February 20 80 Participants

ARRANGED BY **David Stewart, Bruce Stillman, and Jan Witkowski,** Cold Spring Harbor Laboratory

This special annual celebration ("Double Helix Day") is intended to coincide with the actual date that James Watson and Francis Crick discovered the double-helix structure of deoxyribonucleic acid (February 28, 1953) in Cambridge, England. Each year, a theme related to DNA science is explored through a series of review-style talks aimed at a broad scientific audience.

The 2018 gathering was a triple celebration: Dr. James Watson's 90th birthday, his 50th year at Cold Spring Harbor Laboratory, and the 65th anniversary of the discovery of the double helix.

The topic of forensics is a field that has been revolutionized by DNA-based analysis and has brought DNA to the attention of everyone. The program commenced in the early afternoon on Tuesday, February 20, and concluded with an early evening reception and celebratory double-helix feast in Blackford Hall.

PROGRAM

J. Butler, *National Institute of Standards and Technology, Gaithersburg, Maryland:* Introduction to current methods used in the lab, the FBI STR panel, and controversial issues.

G. Cooper, *Director of Science & Research, The Innocence Project, New York, New York:* Lessons from DNA exonerations: Shining a light on fallacies and flaws and the need for reforms.

E. Murphy, *New York University, New York:* Societal issues of DNA typing: Sample retention, familial identifications, deducing physical characteristics.

T. Parsons, *International Commission on Missing Persons:* Application of forensic genetics to large-scale identification of missing persons.

Systems Biology: Global Regulation of Gene Expression

March 20–23 230 Participants

ARRANGED BY **Barak Cohen,** Washington University School of Medicine in St. Louis, Missouri
Katherine Pollard, Gladstone Institutes/University of California, San Francisco
Bas Van Steensel, Netherlands Cancer Institute, Amsterdam, Netherlands
Julia Zeitlinger, Stowers Institute for Medical Research, Kansas City, Missouri

J. Zeitlinger, K. Pollard, B.V. Steensel, B. Cohen

Understanding how the cell orchestrates gene regulation across the genome remains a key goal of systems biology. This 12th conference highlighted the rapid evolution of this field, bringing together a large, diverse, and international group of scientists including students, postdocs, and PIs at all stages of their careers. Technology development—both experimental and computational—continues to play a large role in this field, and the conference featured many talks and posters that presented methodological improvements on old techniques or new techniques altogether. In many ways, the field of systems biology is technology-limited, or rather technology-enabled, as investigators must often invent new methods to address the questions they pose. Bringing together investigators with new technologies so that new methods can percolate through the community is one of the most important functions of this meeting. The 2018 conference, like its predecessors, was a success in this respect, highlighting advances in microscopy, single-cell genomics, computational models, and methods to probe chromatin structure.

Although technology was an important part of this meeting, several biological themes also emerged during the conference. One entire session was dedicated to the question of how paralogous transcription factors distinguish between closely related recognition motifs. Investigators presented evidence for several different answers to this question, including differences in core binding sites, differences in nucleotides flanking canonical sites, differential interactions with cofactors, and subtle structural differences in the protein motifs that contact DNA.

B. Lesch, G. Narlikar

M. Benton, A. Cuomo, A. Casella

M. Pazin, J. Vierstra

T. Baubec, S. Rao

The question of how a cell deals with nucleosome-occluded DNA continues to be an important focus in the field. One keynote speaker, Steve Hennikoff, presented studies detailing the configurations of nucleosomes on transcribed genes. Other speakers discussed the properties of transcription factors that can bind nucleosome-bound DNA, and several speakers discussed how patterns of DNA accessibility can guide the search for active regulatory elements. With new methods that profile nucleosome patterns on DNA at higher and higher resolution, this area of systems biology will continue to be an exciting area of inquiry.

New innovations in computational models of gene regulation played an important role at the conference. Investigators presented models that spanned a range of approaches, from powerful statistical and machine learning frameworks to detailed physical and biophysical models. Connecting the power of machine learning to detect patterns with the mechanistic insight of biophysical models will be an exciting challenge in the future meetings. Along these lines, several investigators presented high-throughput experimental methods to measure biophysical properties of DNA–protein interactions, such as free energies of interaction, as a way to populate and test mechanistic models of gene expression. The integration of new experimental methods with new computational models will continue to be the defining feature of this exciting meeting.

This meeting was supported by funds provided by the National Human Genome Research Institute.

H. Bussemaker, S. Henikoff

J. Kribelbauer, R. van der Lee

PROGRAM

Keynote Speaker
G. Narlikar, *University of California, San Francisco*

Emerging Technologies
Chairpersons: P. Fordyce, *Stanford University, California;* S. Henikoff, *Fred Hutchinson Cancer Research Center, Seattle, Washington*

Keynote Speaker
J.S. Weissman, *University of California, San Francisco*

Enhancers and Transcription I
Chairpersons: S. Sinha, *University of Illinois, Urbana-Champaign;* A. Stark, *Research Institute of Molecular Pathology (IMP), Vienna, Austria*

Enhancers and Transcription II
Chairpersons: S. Sinha, *University of Illinois, Urbana-Champaign;* A. Stark, *Research Institute of Molecular Pathology (IMP), Vienna, Austria*

Chromatin and Epigenetics
Chairpersons: D. Schuebeler, *Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland;* A. Pombo, *Max Delbrück Centre for Molecular Medicine, Berlin, Germany*

Development, Evolution and Variation
Chairpersons: A. Siepel, *Cold Spring Harbor Laboratory;* O. Troyanskaya, *Princeton University, New Jersey*

3D Architecture of the Nucleus
Chairpersons: W. Bickmore, *MRC Human Genetics Unit, Edinburgh, United Kingdom;* A. Boettinger, *Stanford University, California*

Networks, Global Analysis, and Computational Models
Chairpersons: H. El-Samad, *University of California, San Francisco;* R. Sandberg, *Karolinska Institutet, Stockholm, Sweden*

The PARP Family and ADP-ribosylation

April 3–6 144 Participants

ARRANGED BY **Michael O. Hottiger,** University of Zürich, Switzerland
W. Lee Kraus, University of Texas Southwestern Medical Center, Dallas
Susan Smith, The Skirball Institute/New York University, New York

W.L. Kraus, S. Smith, M.O. Hottiger

This third biennial meeting followed an initial meeting titled "The PARP Family and Friends: Gene Regulation and Beyond," which was held in 2014, and a second meeting titled "The PARP Family and ADP-ribosylation," which was held in 2016. These three meetings were built upon (1) previous international PARP meetings, which were organized without the support of a major conference organization; (2) growing interest in the biology of PARPs and ADP-ribosylation; (3) an expanded understanding of the PARP family; (4) an increased understanding of the role of NAD^+ metabolism in PARP function; and (5) the growing clinical promise of FDA-approved PARP inhibitors as therapeutic agents.

The recent PARP meeting featured 44 talks, as well as a poster session. The content presented in both the talks and posters showed that this field is advancing, evolving, and expanding. It included a greater focus on the PARP monoenzymes and their diverse biology, as well as new information about the nuclear PARP polyenzymes (e.g., PARP-1) and their role in DNA damage detection and repair. Other aspects of PARP biology that received greater attention than in the past included functional interactions with RNA, host–virus interactions, chromatin, and gene regulation. Methodological advances, including those related to proteomics (e.g., site-specific ADP-ribosylation) and genomics, were well represented in the presentations. Emerging controversies in the field, such as the repertoire of amino acids that are ADP-ribosylated by PARPs (e.g., Glu, Asp, Ser, Lys), were addressed.

As in the past, the faculty used this forum to meet and discuss issues important to the field, such as nomenclature (e.g., PARP vs. ARTD; monoPARP and polyPARP vs. PARP monoenzyme

C. Kenyon, S. Smith

T. Dawson, G. Poirier

A. Ladurner, S. Deind

Y. Fondufe-Mittendorf, K. Luger

and PARP polyenzyme), promoting and expanding the field, recruiting and retaining trainees in the field, and the relationship of this field to other NAD^+-related fields (e.g., NAD^+ metabolism, sirtuins). These interactions have led to a continuing initiative to standardize the assays, protocols, and reagents used by the PARP field. Dr. Peter Bai, University of Debrecen, Hungary, is organizing this initiative, with the goal of producing guidelines for the study of PARPs and poly(ADP-ribosyl)ation. These guidelines should help to standardize experiments and analyses, allowing for more facile comparison of results generated by different laboratories.

Overall, the sense among the attendees (determined from informal surveys) was that this meeting marked an important turning point for the field and should be continued.

The meeting was funded in part by the National Cancer Institute.

PROGRAM

ADP-Ribosylation and Cellular Signaling
Chairperson: J. Pascal, *University of Montréal, Canada*

C. Bonham, S. Krieg

DNA Repair, Genome Instability, Cancer
Chairperson: M. Hottiger, *University of Zürich, Switzerland*

PARPs in the Nucleus and Signaling
Chairperson: S. Smith, *Skirball Institute/New York University School of Medicine, New York*

Keynote Speaker
C. Kenyon, *Calico Life Sciences, South San Francisco, California*

Chemical Biology, Proteomics, and Substrates
Chairperson: K. Luger, *University of Colorado, Boulder*

NAD^+ Metabolism, Inhibitors, and Analogs
Chairperson: A. Ladurner, *Ludwig-Maximilians-University, Münich, Germany*

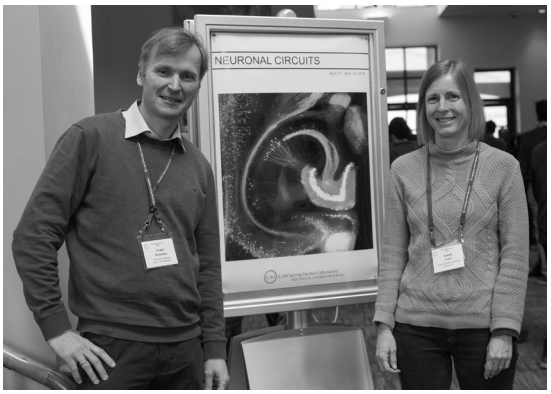
Structure and Mechanisms
Chairperson: C. Rabouille, *Hubrecht Institute, Utrecht, The Netherlands*

Neuronal Circuits

April 11–14 188 Participants

ARRANGED BY **Florian Engert,** Harvard University, Cambridge, Massachusetts
Fritjof Helmchen, Brain Research Institute, Zürich, Switzerland
Kristin Scott, University of California, Berkeley

The anatomical and functional connectivity of neurons underlies the simplest of behavioral decisions to the most complex cognitive tasks. Unprecedented advances in the ability to monitor the activity of large neural populations, to rapidly and precisely manipulate the activity of neurons with optogenetic approaches, and to trace connectivity with synaptic resolution are providing a new systems-level view of neural circuit processing. Coupling recent technological advances with computational modeling and quantitative behavioral approaches has the potential to reveal how the dynamic computations of the brain underlie thought and behavior. This meeting highlights recent progress in elucidating neural circuit function across nervous systems.



F. Helmchen, K. Scott

In keeping with the original goals of this meeting, researchers working in a variety of different organisms and systems were brought together to share their scientific and technological advances, focusing on recent advances in neuronal circuit processing. This meeting included six slide sessions covering a range of topics in sensory systems, motor control, circuit computations, internal states, technology, and plasticity—and a very interactive poster session.

For the seventh meeting of this kind, the response of the field was very enthusiastic. The meeting brought together 188 participants from all over the world, most of whom made either oral or poster presentations. Particularly impressive were the large numbers of students (34%) and postdoctoral fellows (33%) participating in the meeting. The invited presenters and their talks are listed below under Program. The meeting provided an important forum for the exchange of ideas


G. Jefferis, V. Ruta


M. Rubinov, A.A. Zarin

S. Paixao, S. Sayin

J.S. Li, L-F. Jiang-Xie

and approaches and helped scientists in the field get the most up-to-date information, as well as enabling them to meet, network, and establish collaborations. Based on the uniformly enthusiastic comments of the participants, the intensity of the oral and poster sessions, the productive question and answer sessions, and the large crowds that stayed up late to discuss science, the meeting was a great success.

A highlight of the meeting program is the Larry Katz Memorial Lecture, thanks to the generous donation of many colleagues in the field, which recognizes an outstanding research contribution by a graduate student. The overall quality of the seven nominees was truly spectacular, representing an international cohort of exceptional graduate students. The selection committee was composed of the three current organizers and three past organizers: Carl Petersen, École Polytechnique Federale de Lausanne; Ed Calloway, Salk Institute; and Rafa Yuste, Columbia University. The committee unanimously selected Justus Kebschull, Cold Spring Harbor Laboratory, as the 2018 Lecturer. He gave a wonderful presentation about his graduate research. A history of the Larry Katz prize, including background on Katz and a list of previous awardees, was included in the 2018 program booklet and should continue to be included in future programs.

This meeting was funded in part by the National Institute of Neurological Disorders and Stroke (NINDS), the National Eye Institute (NEI), and the National Institute of Biomedical Imaging and Bioengineering (NIBIB).

M. Zhen, M. Dickinson

A. LeMessurier, J. Roebber

PROGRAM

Sensory Transformations I
Chairperson: F. Helmchen, *Brain Research Institute, University of Zürich, Switzerland*

Circuit Computations
Chairperson: D. O'Connor, *Johns Hopkins University, Baltimore, Maryland*

Sensory Transformations II
Chairperson: V. Ruta, *The Rockefeller University, New York, New York*

Internal States
Chairperson: K. Scott, *University of California, Berkeley*

National Institutes of Health Lecture
J.W. Gnadt, *National Institutes of Health/Brain Initiative, Bethesda, Maryland*

Technology
Chairperson: F. Engert, *Harvard University, Cambridge, Massachusetts*

Motor Control
Chairperson: M. Zimmer, *Research Institute of Molecular Pathology, Vienna, Austria*

Learning
Chairperson: T. Komiyama, *University of California, San Diego*

Larry Katz Memorial Lecture
J.M. Kebschull, *Stanford University, California*

Protein Homeostasis in Health and Disease

April 17–21 263 Participants

ARRANGED BY **Judith Frydman,** Stanford University, California
F. Ulrich Hartl, Max Planck Institute for Biochemistry, Martinsreid, Germany
Harm Kampinga, University of Groningen, Netherlands
Richard Morimoto, Northwestern University, Evanston, Illinois

R. Morimoto, J. Frydman, H.H. Kampinga, F.U. Hartl

This meeting was the 26th in the series on this topic among Cold Spring Harbor Laboratory meeting. The topics and their presenters are listed below under Program. In addition to the traditional poster sessions, two sessions with rapid-fire presentations of 2-minute single-slide talks selected from the abstracts prior to the poster sessions were held. This provided an opportunity to highlight many of the poster presentations, encouraging people to attend the posters. In addition, there was an animated panel discussion with editors from *Nature Structure Molecular Biology*, the *Journal of Biological Chemistry*, and Cold Spring Harbor Laboratory Press on the topic of "New Mechanisms for Disseminating Scientific Information." Furthermore, several lunches for young scientists with speakers and organizers were organized to discuss career-related and science-related matters.

The high attendance, large number of posters, and lively audience participation in the sessions and in the informal activities that followed were again testament to the success of the meeting and to the successful balance of the sessions. The field of molecular chaperones, protein folding in vivo and in the cell, proteostasis networks, and diseases of protein misfolding is constantly evolving, and this meeting has been central in making this the cornerstone meeting for this rich and exciting field. It has reflected the many branches of biology, with newly emerging connections of the chaperone field with that of autophagy (highlighted by a talk of Nobel laureate Professor Ohsumi) and stress-induced phase separations in cell biology and its relevance to alternate protein states and human conformational diseases of aging.


L. Veenhoff, E. Sontag


R. Sawarkar, S. Carra

D. Ron, B. Bukau

A. Delury, student award winner

Among the cutting-edge themes explored in this meeting were presentations of the regulation of protein quality control at the ribosome and spatially controlled biogenesis of individual proteins of protein complexes. New insights were provided related to a long-standing debate on how activation of HSF-1, the master regulator of the stress response, is controlled. The importance of autophagy for maintaining proteostasis was also highlighted, and new data showed how chaperones are intimately linked to regulation of autophagic processes. Exciting new data revealed the contribution of phase transitions to protein homeostasis in cells and enlightened how failure in their dynamics can cause aggregation disease. Other data showed that a collapse of protein homeostasis is linked to many of the dysfunctions observed in aging, in neurodegeneration, and in cancer and other diseases related to genomic instability. Although many reports showed the importance of amyloid formation as cause for age-related disease, intriguing data were presented on the physiological function of regulated amyloid formation (e.g., in memory consolidation). Finally, several talks clearly demonstrated that (pharmacological) manipulations of protein quality control networks are at the horizon of becoming clinically applicable.

This meeting was funded, in part, by the National Institute on Aging, a branch of the National Institutes of Health, and by Proteostasis Therapeutics.

PROGRAM

Proteostasis at the Ribosome
Chairperson: F.U. Hartl, *Max Planck Institute of Biochemistry, Martinsried, Germany*

Chaperone Mechanisms I
Chairperson: J. Buchner, *Technical University Münich, Germany*

D. Pincus, J. Weissman

H. Itoh, Y. Ohsumi

Rapid Fire Presentations I
Chairperson: H.H. Kampinga, *University of Groningen, Netherlands*

Novel Mechanisms of Quality Control
Chairperson: U. Jakob, *University of Michigan, Ann Arbor*

Degradative Mechanisms
Chairperson: Y. Ohsumi, *Tokyo Institute of Technology, Yokohama, Japan*

Rapid Fire Presentations II
Chairperson: H.H. Kampinga, *University of Groningen, Netherlands*

Panel Discussion: New Mechanisms for Disseminating Scientific Information
Chairperson: L. Gierasch, *University of Massachusetts, Amherst*

Chaperone Mechanisms II
Chairperson: L. Gierasch, *University of Massachusetts, Amherst*

Pathogenic and Nonpathogenic Aggregates and Amyloids
Chairperson: J. Frydman, *Stanford University, California*

Organellar Proteostasis and Spatial Quality Control
Chairperson: I. Braakman, *Utrecht University, Netherlands*

Proteostasis Failure in Aging and Disease
Chairperson: R. Morimoto, *Northwestern University, Evanston, Illinois*

Gene Expression and Signaling in the Immune System

April 24–28 243 Participants

ARRANGED BY **Diane Mathis,** Harvard Medical School, Cambridge, Massachusetts
Stephen Nutt, The Walter & Eliza Hall Institute, Parkville, Victoria, Australia
Alexander Rudensky, Memorial Sloan Kettering Cancer Center, New York, New York
Art Weiss, HHMI/University of California, San Francisco

This ninth highly engaging and interactive 4-day conference hosted nearly 250 registrants, including many first-time as well as perennial attendees. The meeting is unique within the repertoire of immunology conferences in its focus on molecular and biochemical aspects of the development and function of the immune system, agnostic to cell type. It also attracted a broad range of scientists who exploit the immune system as a model to study basic principles of biological regulation. Talks were delivered by a mix of speakers, invited or selected from a group of more than 130 submitted abstracts, with 25 junior investigators giving talks. As is traditional for this meeting, most speakers focused almost exclusively on their unpublished work.

S. Nutt, A. Weiss

There were many exciting talks. For example, Leah Sibener (Stanford University Medical Center), one of the selected graduate student speakers, addressed the long-standing issue of why some high-affinity peptide:MHC (pMHC) ligands do not activate T cells. An impressive array of physical measurements revealed that agonist pMHC complexes form an additional "catch" bond with the TCR as it unbinds, extending bond half-life, whereas the nonagonist complexes do not form additional bonds, but instead they unbind, typical of a "slip" bond. Flavian Brown (Harvard Medical School) presented unpublished results from his recently completed Ph.D. studies investigating a role for fibroblastic reticular cells (FRCs), whose function was previously linked exclusively to negative regulation of the immune response. He discovered that these cells play an essential role in inducing protective $CD8^+$–T-cell-mediated immunity against viral infections via an IL-6-signaling-dependent mechanism. Marked chromatin

E. Rothenberg, I. Taniuchi

F. Lakkis, T. Golovkina

A. Rudensky, S. Dodgson

D. Mathis, A. Chawla

remodeling and metabolic changes are induced, enabling $CD8^+$–T-cell differentiation into potent effectors and long-lived tissue-resident memory cells. Richard Locksley (University of California, San Francisco) described an elegant circuit within the small intestine that is mobilized by parasitic colonization: Tuft cells make IL-25, which induces ILC-2s to secrete IL-13, which drives secretory epithelial cell generation (i.e., tuft and goblet cells) as well as bowel hypertrophy and enhanced motility. Miguel Soares (Gulbenkian Institute, Lisbon, Portugal) reported on a fascinating scenario of cross talk between iron and glucose metabolism that comes into play upon bloodstream infection—for example, by the malaria parasite, *Plasmodium*. This scenario turns out to be an important mechanism of disease tolerance, potentially useful for developing novel approaches to manipulation of pathogen–host interactions.

Such talks highlighted the relevance of the approaches discussed at this meeting not just to immunology but also to the broader biomedical scientific community. Oral presentations were supplemented by two afternoon poster sessions, which were extremely vibrant and provided a forum for all meeting participants (including many graduate students and postdocs) to share their most recent data. A new addition to the meeting was three sets of Chat-with-the-Speakers lunches, which proved stimulating to both the trainees and speakers who attended. Topics of discussion ranged across science, lab management, career development, and work/life balance. Participating trainees came from across the world, allowing interesting cross-cultural discourse.

This meeting was funded in part by the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health; and Amgen.

G. Natoli, C. Benoist

M. Boutet

PROGRAM

Control of Gene Expression
Chairperson: D. Mathis, *Harvard Medical School, Boston, Massachusetts*

Epigenetics and Chromatin
Chairperson: A. Rudensky, *Memorial Sloan Kettering Cancer Center, New York, New York*

Signaling at the Membrane
Chairperson: A. Weiss, *HHMI/University of California, San Francisco*

Intercellular Communication
Chairperson: R. Locksley, *University of California, San Francisco*

Cell Differentiation Pathways/Transcriptional Networks
Chairperson: S. Nutt, *The Walter and Eliza Hall Institute, Parkville, Victoria, Australia*

Molecular Basis of Host–Microbial Interaction
Chairperson: D. Littman, *HHMI/New York School of Medicine, New York*

Metabolism and the Immune System
Chairperson: R. Medzhitov, *HHMI/Yale University, New Haven, Connecticut*

Intracellular Signaling
Chairperson: H. Wu, *Boston Children's Hospital, Harvard Medical School, Massachusetts*

Nuclear Organization and Function

May 1–5 214 Participants

ARRANGED BY **Edith Heard,** Curie Institute, Paris, France
Martin Hetzer, Salk Institute for Biological Studies, San Diego, California
Tom Misteli, National Cancer Institute, Bethesda, Maryland
David L. Spector, Cold Spring Harbor Laboratory

M. Hetzer, E. Heard, D. Spector, T. Misteli

This 20th anniversary meeting highlighted the importance and impact this field has made over the last two decades. The richness of new science and the enthusiasm apparent at the meeting also impressively illustrated the foresight David Spector, the founder of this meeting, had when he initiated what is now one of the premiere events in this field. As in previous meetings, the presentations, poster, and interactive discussions highlighted the impressive recent progress made. A major direction in the field is the genome-wide mapping of higher-order chromatin interactions using biochemical methods in conjunction with imaging methods and with computational model approaches. These strategies are providing novel insights into higher-order genome architecture and nuclear organization. Using this approach, evidence now suggests that homotypic interactions between different types of chromatin (heterochromatin, euchromatin) are a major driver of large-scale 3D genome organization (Leonid Mirny). In addition, nuclear features such as the nuclear periphery and intranuclear bodies function as constraints in determining the arrangement of genes and chromosomes in 3D space (Karen Reddy, Mitchell Guttman). Many of the talks selected from abstracts provided exciting insights into the functional relationship between 3D organization, nuclear positioning, and gene expression, thanks to the use of powerful CRISPR-based approaches combined with imaging and genomics. Important discoveries about the mechanisms of action of specific architectural proteins such as CTCF, cohesin, and condensin in chromosome organization were also reported. The continued importance of the

P. Kaufman, D. Carone

E. Finn, S. Kundu

L. Ventimiglia, S. Bahmanyar

O. Gadal, E. Fabre

development of new methods was also evident. Although super-resolution microscopy is now routinely used, several high-end new imaging methods were highlighted, including ZOLA microscopy, which has solved the current limited spatial resolution along the optical axis (Christophe Zimmer), as well as the ability to combine multiplexed imaging of DNA, RNA, and protein both in cells and in tissues (Alistair Boettinger). In addition, global methods to probe the interaction of multiple genome regions have now been developed (Mitch Guttman, Ana Pombo), as well as methods that enable probing of interactions of chromatin with proteinaceous subnuclear structures (Andrew Belmont). Powerful tools for live-cell imaging of DNA, RNA, proteins, and their modifications were also presented (Kimura). The meeting also highlighted the importance of not just the structure of the genome in its function but also the contribution of architectural features of the nucleus, particularly the nuclear lamina and the nuclear pore complex. A breakthrough in this area has been the ability to resolve the structure of the lamina by cryo-electron microscopy (Ohad Medalia). Correlation of live-cell imaging with high-resolution electron tomography highlighted molecular differences in NPC assembly after mitosis and during interphase, the latter of which appears to occur from the inner nuclear membrane (Jan Ellenberg). Work in *Drosophila* and yeast suggests that nuclear pore proteins directly regulate the transcription machinery (Jason Brickner) and chromatin decondensation (Maya Capelson).

This meeting was funded, in part, by Avanti Polar Lipids, Inc.

E. Heard, A.D. Aguese

Nicholls Biondi Hall evening posters

PROGRAM

The Nuclear Membrane and Nuclear Lamina
Chairpersons: O. Medalia, *University of Zürich, Switzerland;* D. Starr, *University of California, Davis*

Chromatin Organization
Chairpersons: L. Mirny, *Massachusetts Institute of Technology, Cambridge;* K. Reddy, *Johns Hopkins University School of Medicine, Baltimore, Maryland*

Emerging Technologies to Explore Genome Organization
Chairperson: X. Darzacq, *University of California, Berkeley*

Gene Expression and Transcription
Chairperson: M. Dahan, *Institut Curie, Paris, France*

Nucleocytoplasmic Transport and Nuclear Pore Complexes
Chairpersons: J. Ellenberg, *European Molecular Biology Laboratory, Heidelberg, Germany*; M. Lusic, *Heidelberg University Hospital, Germany*

Nuclear RNAs and Nuclear Bodies
Chairperson: M. Guttman, *California Institute of Technology, Pasadena*

Nuclear Function in Differentiation and Development
Chairpersons: A. Fisher, *MRC London Institute of Medical Sciences, United Kingdom*; H. Kimura, *Tokyo Institute of Technology, Japan*

Epigenetics
Chairpersons: A. Pombo, *Max Delbrück Center for Molecular Medicine, Berlin, Germany;* J. Marko, *Northwestern University, Evanston, Illinois*

Biology and Genomics of Social Insects

May 5–8 97 Participants

ARRANGED BY **Sarah Kocher,** Princeton University, New Jersey
Seirian Sumner, University College London, United Kingdom
Amro Zayed, York University, Toronto, Ontario, Canada

S. Kocher, A. Zayed, S. Sumner

This meeting attracted attendees from the United States, Canada, Europe, and Australia. The program contained 43 talks organized into the 10 sessions that are listed below under Program. We also invited 30 poster submissions, which similarly covered a diversity of topics in social insect genomic research. Dr. Christina Grozinger delivered one of the keynote lectures, summarizing a large body of literature on the application of genomics to honeybee health. Dr. Jacobus Boomsma delivered the second keynote lecture, which provided a critical view on the importance of defining the diversity of behaviors and social organization found in insects for hypothesis testing in sociogenomics.

The conference attracted a good mix of senior and junior researchers; with graduate students and postdoctoral fellows representing ~34% and ~23.7% of all attendees. Participation of female scientists was markedly higher in 2018 (41.2%) relative to 2015 (34.5%).

This meeting was funded, in part, by the RCN-Research Coordination Network in the Genetics and Genomics of Social Behavior NSF Grant #OS-1256839.

PROGRAM

Keynote Speaker
C. Grozinger, *Pennsylvania State University, University Park*

Health and Immunity of Social Insects
Chairperson: S. Barribeau, *University of Liverpool, United Kingdom*

Emerging Genomic Tools for Non-Model Organisms
Chairperson: G. Zhang, *University of Copenhagen, Denmark*

Ecological Genomics of Social Insects
Chairperson: J. Purcell, *University of California, Riverside*

J. Snow, M. Hasselmann

H. Axen, E. Vargo

K. Boomsma, J. Evans

H. Helantera, W. Trible

Keynote Speaker
J. Boomsma, *University of Copenhagen, Denmark*

Integrating Social Insect Theory with Sociogenomics
Chairperson: H. Helantera, *University of Helsinki, Finland*

Symbionts and Microbiomes
Chairperson: N. Moran, *University of Texas. Austin*

Genomics of Social Plasticity
Chairperson: A. Toth, *Iowa State University, Ames*

Gene Regulation and Epigenetics in Social Organisms
Chairperson: R. Bonasio, *University of Pennsylvania, Philadelphia*

Genome Evolution and Social Evolution
Chairperson: K. Kapheim, *Utah State University, Logan*

Charting the Genotype–Phenotype Map for Social Insects
Chairperson: D. Kronauer, *The Rockefeller University, New York, New York*

Neurogenomics
Chairperson: A. Barron, *Macquarie University, Sydney, New South Wales, Australia*

C. Vernier, S. Karpe

A. Toth, K. Asahina

The Biology of Genomes

May 8–12 482 Participants

ARRANGED BY **Matthew Hurles,** Wellcome Trust Sanger Institute, Hinxton, United Kingdom
Elaine Ostrander, National Human Genome Research Institute, Bethesda, Maryland
Dana Pe'er, Memorial Sloan Kettering Institute, New York, New York
Jonathan Pritchard, Stanford University, California

J. Pritchard, E. Ostrander, D. Pe'er, M. Hurles

Participants at this meeting presented more than 300 abstracts describing a broad array of topics relating to the functional analysis, bioinformatics, comparative characterization, and interpretation of genomes from diverse organisms. The scope and applicability of genome science continues to grow. Even after almost three decades of Biology of Genomes, the talks were fresh with fierce competition for speaking slots. The meeting featured talks on a wide variety of topics that are listed below under Program. Session chairs tried to ensure a reasonable balance of genders, but this is something we can improve on next year. There was a strong focus on younger graduate students and postdoc presentations.

The specific talks spanned a wide range of topics within each area, including DNA sequence variation and its role in molecular evolution, population genetics and complex diseases, comparative genomics and behavior, large-scale studies of gene and protein expression, and genomic approaches to ecological systems. The impact of genomics in translation studies was exemplified through talks on decoding cancer genomes through hierarchical pathway architecture (Trey Ideker), studies of functional classification of *BRCA1* variants with saturation genome editing (Gregory Findlay), and discussion of rare disease studies (Pejman Mohammadi). Bioinformatics, technology, and applications, such as those associated with CRISPR-Cas9, were key to the success of the meeting. Particularly well attended talks were those that dealt with ancient DNA, such as discussions of sequence-based approaches that utilize complete modern and ancient genomes for investigating early human history (Mattias Jakobsson).

D. Garrido, B. Borsari

G. Sherlock, J. Granka

E. Dermitzakis, R. Hardison

R. Yadav, H. Wick

Three poster sessions allowed for comprehensive discussions of abstracts that did not make the talks and were well attended, with several participants commenting that many posters could easily have been talks.

All sessions were well attended, stretching the capacity of the CSHL facilities. The keynote presentations were from Wendy Bickmore, whose topic focused on special organization of the genome, and David Page, who discussed sex differences in health and disease. A brief presentation on the NHGRI strategic planning process was made.

The ELSI (Ethical, Legal, and Social Implications) panel was moderated by Nicole Lockhart, focusing on "Human Gene Editing: Traversing the Germline." Panelists included Ryan Fischer (Parent Project Muscular Dystrophy), Rosario Isasi (University of Miami), Debra Mathews (Johns Hopkins University), and Fyodor Urnov (Altius Institute for Biomedical Sciences).

This meeting was funded in part by the National Human Genome Research Institute, a branch of the National Institutes of Health; Illumina; and Oxford Nanopore Technologies.

PROGRAM

Genome Engineering, Editing
Chairpersons: J. Boeke, *New York University Langone Health, New York*; F. Zhang, *Broad Institute of MIT and Harvard, Cambridge, Massachusetts*

Population Genomics
Chairpersons: M. Jakobsson, *Uppsala University, Sweden;* S. Tishkoff, *University of Pennsylvania, Philadelphia*

INTRODUCTION TO THE NHGRI STRATEGIC PLANNING PROCESS

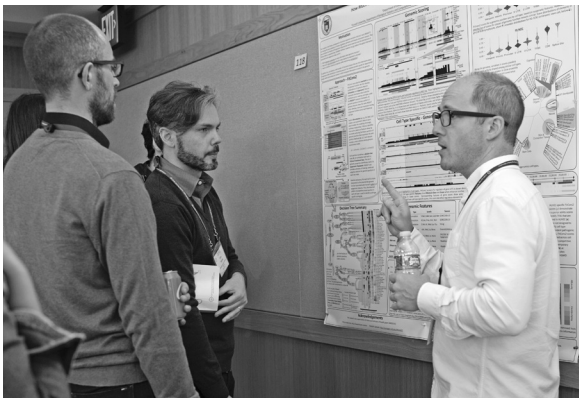
Functional Genomics and Epigenetics
Chairpersons: J. Dekker, *HHMI/University of Massachusetts Medical School, Worcester;* E. Farley, *University of California, San Diego*

Evolutionary and Non-Human Genomics
Chairpersons: M. Justice, *The Hospital for Sick Children, Toronto, Ontario, Canada*; G. Sherlock, *Stanford University, California*

ELSI Panel and Discussion: Human Gene Editing: Traversing the Germline
N.C. Lockhart, *National Institutes of Health/National Human Genome Research Institute, Bethesda, Maryland*

J. Cross, C. Hart

P. Carninci, D. Page



L. Pinello, A. Siepel

Cancer and Medical Genomics
Chairpersons: T. Ideker, *University of California, San Diego;* S. Plon, *Baylor College of Medicine, Houston, Texas*

Computational Genomics
Chairpersons: O. Stegle, *EMBL/EBI, Hinxton, United Kingdom;* C. Leslie, *Memorial Sloan Kettering Cancer Center, New York, New York*

Guest Speakers
W. Bickmore, *MRC/University of Edinburgh, United Kingdom*
D. Page, *HHMI/MIT Whitehead Institute, Cambridge, Massachusetts*

Complex Traits and Microbiome
Chairpersons: J. Barrett, *Wellcome Sanger Institute, Hinxton, United Kingdom;* A. Bhatt, *Stanford University, Palo Alto, California*

Regulatory and Noncoding RNAs

May 15–19 180 Participants

ARRANGED BY **Nicholas Proudfoot,** University of Oxford, United Kingdom
Erik Sontheimer, University of Massachusetts Medical School, Worcester

N. Proudfoot, E. Sontheimer

It is now widely recognized that RNA is much more than just a blueprint for proteins. A broad range of molecular functions have been defined for different noncoding RNA molecules, and likely many more roles remain to be discovered. In addition, the underlying mechanisms of noncoding (nc) RNA function often reveal new biological and biochemical principles. Characterization of the molecular machineries that enact this ncRNA biology have provided the basis for revolutionary technological advances such as RNA interference and CRISPR genome engineering. In this spring's meeting, a mix of established leaders, promising new investigators, and graduate and postdoctoral trainees gathered at Cold Spring Harbor to learn about the latest advances in this rapidly changing field. Oral presentations and poster sessions covered unpublished results from a wide variety of biological systems, including bacteria, single-celled eukaryotes, plants, insects, and mammals. The topics ranged from the basic to the applied, from the mechanistic to the ecological, and from individual genes to entire genomes and epigenomes. Throughout the conference, discussions were lively and vigorous. The meeting solidified the field's strongly communitarian ethic and reinforced the attendees' motivation to drive noncoding RNA research forward with maximum rigor. The meeting provided a superb opportunity to accelerate RNA-related discovery and to apply the advances to our understanding of basic biology and the improvement of human health.

Most of the currently identified ncRNA—both small RNAs, such as microRNA and piRNA, and long RNAs, such as lincRNAs and circular RNA—were well represented as well as several newer less well characterized ncRNAs. A strong feature of the meeting was that nearly the full set of invited speakers originally proposed by the organizers committed to attend and indeed stayed throughout

D. Cazalla, M. Simon

A. Jones, T. Cech

R. Ayana, S. Lopez-Gomollon

M. Suzawa, A.T. Matey

the meeting. Consequently, there was a great opportunity for more junior postdoctoral and graduate student researchers to meet up with and discuss their science with their peers in the field.

PROGRAM

Plenary Session
Chairpersons: N. Proudfoot, *University of Oxford, United Kingdom;* E. Sontheimer, *University of Massachusetts Medical School, Worcester*

microRNAs
Chairperson: N. Kim, *Seoul National University, Korea;* D. Bartel, *HHMI/MIT/Whitehead Institute, Cambridge, Massachusetts*

DNA Modification and Novel RNA Biology
Chairpersons: A. Aguilera, *University of Seville, Spain;* L. Landweber, *Columbia University, New York, New York*

Long Noncoding and Circular RNAs
Chairpersons: K. Adelman, *Harvard Medical School, Boston, Massachusetts;* N. Rajewsky, *Max Delbrück Center for Molecular Medicine, Berlin, Germany*

RNA Regulation in the Nucleus
Chairpersons: A. Akhtar, *Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany*; R. Martienssen, *Cold Spring Harbor Laboratory*

piRNAs and Transmissible RNAs
Chairpersons: J. Brennecke, *Institute of Molecular Biotechnology, Vienna, Austria*; E. Miska, *University of Cambridge, United Kingdom*

Microbial RNAs and Modifications
Chairpersons: A. Ke, *Cornell University, Ithaca, New York;* S. Jaffrey, *Weill Cornell Medical College, New York, New York*

D. Bartel, S. Gu

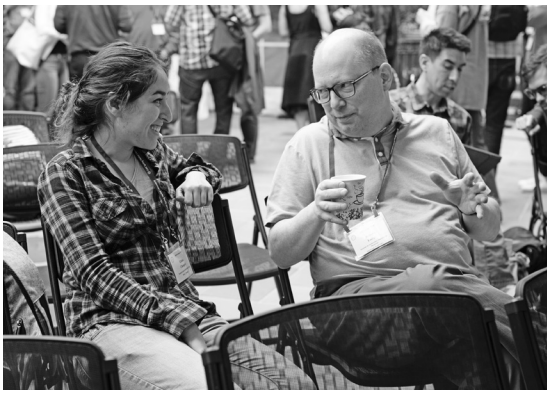
E. Lai, O. Rando

Retroviruses

May 21–26 377 Participants

ARRANGED BY **Theodora Hatziioannou,** The Rockefeller University, New York, New York
Walther Mothes, Yale University, New Haven, Connecticut

This meeting, considered the best on the basic biology of retroviruses including HIV, brings together scientists from around the world, fosters friendships and collaborations, and has frequently been the venue where major scientific breakthroughs have first been announced.



G. Chougui, R. Lodge

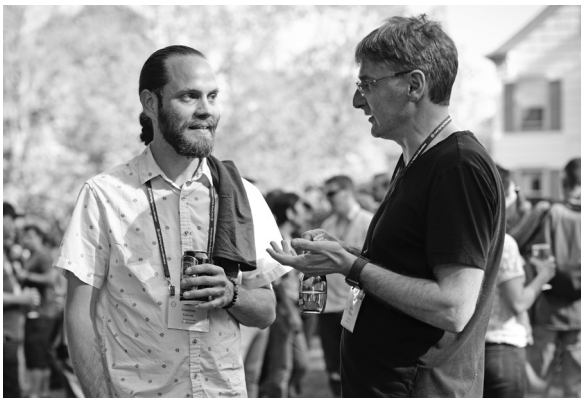
Groundbreaking presentations this year included novel methods for following viral dissemination in vivo, the role of IP6 in stabilizing the immature and mature HIV-1 capsid, the role of a CA-like protein in transmitting RNA between neurons, the evolution and mechanism of antiviral activity by ZAP, and multiple presentations on the role of innate immunity in controlling virus replication.

This year's keynote speakers were Drs. Heinrich Gottlinger and Pamela Bjorkman. Dr. Gottlinger presented an overview of his influential original discoveries, from the first description of the late domain and Vpu phenotypes in HIV assembly and budding and the characterization of the cellular machinery utilized for HIV assembly all the way to recent work on how the HIV Nef protein counteracts members of the SERINC family of proteins. Dr. Bjorkman gave an inspiring talk about her structural work that has revealed fundamental concepts of biology, from the structure of the MHC peptide complex and FC receptor to extensive work on the HIV envelope glycoproteins and how this knowledge influences vaccine studies.

The meeting maintained the overall arrangement of 13 sessions; however, presentations were grouped with an emphasis on the viral life cycle rather than specific proteins—resulting in more diverse and stimulating sessions that were interesting to larger fractions of the attendees. Additionally, two new features were introduced to the meeting: (1) a live recording of the podcast This Week in Virology (TWiV) by Drs. Racaniello and Despommiers, featuring veteran meeting

Z. Debyser, G. Xue

I. Rouzina, L. Cheluc



J. Grover, W. Mothes

C. Buffone, A. Selyutina

attendees Drs. Steven Hughes and John Coffin, as well as all the CSH Retrovirus prize winners, the Uta von Schwedler prize for Ph.D. students winner Matt Takata, and the Andy Kaplan prize for senior postdoc/young PI winner Ya-Chi Ho, which emphasized the historical importance of the meeting as well as the recent discoveries by the younger participants; and (2) a lunchtime panel discussion featuring Drs. Hughes, Ho, and Planelles and ViiV head of discovery Mark Cockett on HIV-1 reservoirs and latency, which allowed the participants of the meeting to discuss recent advances and ultimate goals in a developing field and provided the industry's point of view.

Note that a surprising number of scientists did not receive a visa to enter the United States to attend this year's meeting, and their presentations had to be canceled at the last minute. The historic success of basic science depends on a continued free exchange of ideas across the world that when impeded will not just weaken this meeting and the standing of the United States in the world, but also reduce our ability to defeat challenging human diseases such as the HIV/AIDS pandemic.

This meeting was funded, in part, by the National Institute of Allergy and Infectious Diseases, a branch of the National Institutes of Health. Corporate sponsorship provided by Viiv Healthcare.

PROGRAM

Virus Dissemination and Pathogenesis

Chairpersons: T. Hope, *Northwestern University, Chicago, Illinois;* G. Tachedjian, *Burnet Institute, Melbourne, Victoria, Australia*

Assembly

Chairpersons: E. Barklis, *Oregon Health & Science University, Portland;* O. Pornillos, *University of Virginia, Charlottesville*

T. Hatziioannou, M. Johnson

M. Naghavi, S. Goff

ENV and Restriction Factors
Chairpersons: A. Herschhorn, *University of Minnesota, Minneapolis;* D. Sauter, *Ulm University Medical Center, Germany*

Keynote Speaker
P. Bjorkman, *HHMI/California Institute of Technology, Pasadena*

Viral RNA Interactions
Chairpersons: N. Sherer, *University of Wisconsin, Madison;* S. Kutluay, *Washington University School of Medicine in St. Louis, Missouri*

Seventh Annual Uta von Schwedler Prize for Retrovirology
Presented by: T. Hope, *Northwestern University, Chicago, Illinois*
Awarded to: M.A. Takata, *The Rockefeller University, New York, New York*

Post-Entry
Chairpersons: M. Naghavi, *Northwestern University, Chicago, Illinois;* C. Aiken, *Vanderbilt University Medical Center, Nashville, Tennessee*

TWIV PODCAST (THIS WEEK IN VIROLOGY)

Nuclear Entry and Transcription
Chairpersons: K. Beemon, *Johns Hopkins University, Baltimore, Maryland;* V. KewalRamani, *National Cancer Institute, Frederick, Maryland*

Twelfth Annual Andy Kaplan Prize
Presented by: C. Jolly, *University College London, United Kingdom*
Awarded to: Y-C. Ho, *Yale University School of Medicine, New Haven, Connecticut*

Lunchtime Panel Discussion
HIV Reservoirs and Latency: What Are They and How Can We Tackle Them?

Keynote Speaker
H. Gottlinger, *University of Massachusetts Medical School, Worcester*

Discovering New Host Factors
Chairpersons: M. Johnson, *University of Missouri, Columbia;* M. Ohainle, *Fred Hutchinson Cancer Research Center, Seattle, Washington*

Innate Sensing
Chairpersons: R. Gummuluru, *Boston University School of Medicine, Massachusetts;* G. Towers, *University College London, United Kingdom*

Restriction Factors and Endogenous Viruses
Chairpersons: G. Arriagada, *Universidad Andres Bello, Vina del Mar, Chile;* J. Dudley, *University of Texas, Austin*

Eighth Annual Daniel Wolf Prize
Presented by: G. Arriagada, *Universidad Andres Bello, Santiago, Chile*
Awarded to: Best poster presentation

Integration and Transcription
Chairpersons: K. Bishop, *The Francis Crick Institute, London, United Kingdom;* P. Kumar, *Yale University School of Medicine, New Haven, Connecticut*

Glia in Health and Disease

July 19–23 296 Participants

ARRANGED BY **Nicola Allen,** Salk Institute, San Diego, California
Kelly Monk, Vollum Institute, Portland, Oregon

N. Allen, K. Monk

Glial cells are as abundant as neurons in the nervous system, but the diverse roles played by these cells under physiological and pathological conditions are not well understood. In this summer's eighth meeting, trainees and scientists from across the world gathered to discuss recent progress in this rapidly expanding field. Half of the session chairs were women, approximately half (46%) of the speakers were women, and most speakers were selected from the abstracts with a focus on trainees and junior faculty (~68% of speakers were trainees or junior faculty). Speakers presented exciting new data on a variety of important topics listed below under Program. The atmosphere was collegial and supportive, and there were lively discussions in oral sessions, poster sessions, and informal gatherings. Research involving vertebrate and invertebrate model systems was presented along with advances in state-of-the-art methodologies that are providing new insight into glial cell functions. A highlight of the meeting was the Ben Barres memorial session, which supplanted a more traditional keynote session. During this session, two talented trainees (one graduate student and one postdoctoral fellow) using diverse experimental approaches presented their recent work and were presented with an award. This session was convened in order to honor the long-standing contributions of the meeting's founder Dr. Ben Barres to the training of young scientists and his leadership in the field of glial cell biology; the current, previous, and currently designated future organizers desire for this session to take the place of the keynote session for all subsequent meetings. By the conclusion of the meeting, it was clear that information about properties and functions of glia is increasing exponentially, enabled by the many advances described at the meeting. These studies are yielding new insight into the roles

A. Klingseisen, M. Swire

S. Liddelow, M. Monje

E. Niederst, C. Smith

N. Huber, E. Cameron

of distinct groups of glia in nervous system development, neural circuits, disease initiation and progress, and nervous system repair.

This meeting was funded, in part by the National Institute of Neurological Diseases and Stroke, a branch of the National Institutes of Health.

PROGRAM

Glial Control of Circuits and Behaviors
Chairpersons: M. Freeman, *Oregon Health & Science University, Portland;* B. Khakh, *University of California, Los Angeles*

Glial Responses in Injury, Disease, and Repair
Chairpersons: M. Mokalled, *Washington University School of Medicine in St. Louis, Missouri;* S. Parrinello, *University College London, United Kingdom*

Glial Heterogeneity
Chairpersons: D. Rowitch, *University of Cambridge, United Kingdom;* S. Bilbo, *Harvard Medical School, Boston, Massachusetts*

Glial Development and Plasticity
Chairpersons: J. Trotter, *Johannes Gutenberg University of Mainz, Germany;* D. Lyons, *University of Edinburgh, United Kingdom*

Aberrant Properties of Glia in Disease
Chairpersons: B. Stevens, *Children's Hospital Boston/Harvard Medical School, Massachusetts;* M. Monje, *Stanford University, California*

Trainee Session: In Honor of Dr. Ben Barres
N. Allen, *The Salk Institute, La Jolla, California*

B. Zhang, S. Ding

D. Schafer, W.S. Chung

Wine and cheese on Grace Patio

Glia at the Synapse
Chairpersons: C. Eroglu, *Duke University Medical Center, Durham, North Carolina;* J. Dougherty, *Washington University School of Medicine in St. Louis, Missouri*

Glial–Vasculature Interactions
Chairpersons: D. Attwell, *University College London, United Kingdom*; S. Fancy, *University of California, San Francisco*

Glial–Glial Interactions
Chairpersons: X. Piao, *Boston Children's Hospital/Harvard Medical School, Massachusetts;* S. Liddelow, *New York University Langone Medical Center, New York*

Mechanisms and Models of Cancer

August 14–18 344 Participants

ARRANGED BY **Martin McMahon,** University of Utah/Huntsman Cancer Institute, Salt Lake City
Benjamin Neel, New York University School of Medicine, New York
Karen Vousden, Francis Crick Institute, London, United Kingdom
Eileen White, The Rutgers Cancer Institute of New Jersey, New Brunswick

B. Neel, E. White, M. McMahon, K. Vousden

Mechanisms that drive human cancers, along with cell- and animal-based models to test these mechanisms, remain central to our understanding of human malignancies. The current explosion in molecular cataloguing of mutations, gene expression, and epigenetic regulation that drives distinct human cancer types has carried with it an improved understanding of cancer physiology, offering new therapeutic opportunities. Our ability to understand cross talk between cancer cells and stromal cells continues to improve, leading to immunotherapies that have delivered cures for diseases that had been rapidly fatal.

This meeting convened an international group of investigators whose collective work focused on these themes. Oral and poster presentations provided beautiful examples of how new insights have been gained from application of interdisciplinary approaches that utilize genetics, developmental biology, genomics and proteomics, and model organisms (principally sophisticated mouse models, with significant advances in methods for rapid genomic editing of normal and cancer genomes since our 2012 meeting) to advance the development of rational therapeutics.

Keynote addresses were delivered by Drs. Joan Brugge and William Weiss. Dr. Brugge's topic was "Oxidative Stress Defenses in Cancer," and she detailed a recently uncovered pathway by which TRPA1 calcium channels unexpected help promote survival of anchorage-independent cancer cells by promoting Ca-evoked RAS activation. Dr. Weiss' address on "Mechanisms and Models of Glioblastoma" focused on the complex interactions between EGFRvIII and autocrine


E. Bernstein, A. Bagchi


C. Crews, C. Stahlhut

C-L. Chew, R. Petrie

C. Lutz, U. Herrmann

feedback pathways involving NF-κB, IL6, STAT3, and betacellulin. He also presented intriguing data on the role of macrophages in the GDM microenvironment. The Legacy Lecture was delivered by Dr. James Allison on "Immune Checkpoint Blockade in Cancer Therapy: New Insights, Opportunities, and Prospects for Cures." Allison, who subsequently shared the 2018 Nobel Prize in Medicine or Physiology for his discovery of the first checkpoint inhibitor, provided a beautiful overview of the history of the field, as well as recent provocative data from his own laboratory. Eight oral sessions and two poster sessions focused on the topics listed below under Program.

Generous support for this meeting was provided by Northwell Health; the meeting was funded in part by the National Cancer Institute, a branch of the National Institutes of Health. The meeting will be held again in the summer of 2020.

PROGRAM

Keynote Speaker
J. Brugge, *Harvard Medical School, Boston, Massachusetts*

Proliferation and Survival
Chairpersons: S. Courtneidge, *Oregon Health and Science University, Portland;* A. Saurin, *University of Dundee School of Medicine, United Kingdom*

Genetics/Heterogeneity and Evolution
Chairpersons: E. Mardis, *Nationwide Children's Hospital, Columbus, Ohio;* S. Nik-Zainal, *University of Cambridge, United Kingdom*

Experimental Therapeutics
Chairperson: C. Crews, *Yale University, New Haven, Connecticut*

F. Sesay, A. Ibrahim

J. Witkowski, M. McMahon, J. Allison, S. Courtneidge

Cancer Metabolism
Chairpersons: R. DeBerardinis, *University of Texas Southwestern Medical Center, Dallas;* J. Rutter, *HHMI/University of Utah School of Medicine, Salt Lake City*

Keynote Speaker
W. Weiss, *University of California, San Francisco*

Tumor–Host Interactions and Metastasis
Chairpersons: E. Sahai, *The Francis Crick Institute, London, United Kingdom;* V. Weaver, *University of California, San Francisco*

Gene Transcription and Epigenetics
Chairpersons: E. Bernstein, *Icahn School of Medicine at Mount Sinai, New York, New York;* C. Roberts, *St. Jude Children's Research Hospital, Memphis, Tennessee*

Keynote Lecture
J. Allison, *University of Texas MD Anderson Cancer Center, Houston*

Tumor Immunology
Chairperson: H. Levitsky, *Juno Therapeutics, New York*

Stem Cells and Cell Renewal
Chairpersons: S. Morrison, *HHMI/University of Texas Southwestern Medical Center, Dallas;* R. Levine, *Memorial Sloan Kettering Cancer Center, New York, New York*

Genome Engineering: The CRISPR-Cas Revolution

August 22–25 390 Participants

ARRANGED BY **Jennifer Doudna,** HHMI/University of California, Berkeley
Maria Jasin, Memorial Sloan Kettering Cancer Center, New York, New York
Stanley Lei Qi, Stanford University, California
Jonathan Weissman, HHMI/University of California, San Francisco

S.L. Qi, M. Jasin, J. Weissman, J. Doudna

This was the fourth consecutive conference in the series to be held at Cold Spring Harbor. Although the consideration last year was to move the meeting to alternating years, the robust attendance again this year led to the decision to hold the meeting again in 2019. The meeting is unusual in the number of corporate attendees (a record 116); there were also three workshops hosted by companies (Cell Microsystems, Agilent, Synthego). A mouse engineering workshop was also held during the meeting. A number of journalists were present representing *Nature, Nature Medicine, Nature Methods, Nature Communications, Science, Nature Cell Biology (China),* and *The CRISPR Journal.* In addition to the regular sessions, a panel discussion was held on human embryo engineering, bringing together a well known bioethicist (Art Caplan), a member of the NAS working group (Janet Rossant), and a scientist working with human embryos (Dieter Egli).

A goal of this meeting was to bring together researchers working in diverse fields to stimulate discussions and ideas to further exploit CRISPR-Cas9 and related technologies for biological discovery, organismal engineering, and medical applications. The presentations are listed below under Program. Twenty-four speakers were invited to cover these diverse topics. Another 27 speakers were chosen from submitted abstracts. Speakers represented institutions from the United States and abroad (United Kingdom, Switzerland, Australia, France, Canada, Sweden, and Japan) with a few representatives from industry. Speakers chosen from submitted abstracts included lab heads and staff scientists; postdoctoral fellows included

N. Rusk, M. Jinek

C. Fellmann, D. Savage

A. Russell, B. Wienert

K. Davies, S. Maragh

graduate students. Approximately 150 posters were presented in two sessions, complementing the oral presentations.

Talks throughout the meeting utilized molecular, cell, and computational biology in diverse model organisms, as well as economically important and some unconventional organisms. Much of the data presented was unpublished or only very recently published. Highlights related to human disease included the development of highly engineered CAR T cells and success in treating muscular dystrophy in both mice and dogs, suggesting promise for patient therapy. The CRISPR Biology session included vivid descriptions of CRISPR–anti-CRISPR wars between bacteria, phage, and prophage. Lineage analysis in the early mouse was also presented that made use of CRISPR tracking, and an ant knockout of the olfactory coreceptor was presented that led to defective social interactions, demonstrating how developmental systems can be interrogated. The DNA-repair session emphasized the complicated interplay of various double-strand-break repair pathways and highlighted approaches to alter pathway choice. CRISPR was also utilized for nonbreak applications, including to reprogram genome organization.

This meeting was funded in part by Addgene; Advanced Analytical Technologies, Inc.; Agilent Technologies; Bex Co., Ltd.; Cell Microsystems; Integrated DNA Technologies; Sigma Millipore; Synthego; and Twist Bioscience.

N. Jakimo, A. Davidson

J. McDonough, C. Cotta-Ramusino

PROGRAM

CRISPR Biology
Chairpersons: Alan Davidson, *University of Toronto, Ontario, Canada*; K. Makarova, *National Center for Biotechnology Information, National Institutes of Health, Bethesda, Maryland*

Screens and Technology
Chairpersons: D. Liu, *Harvard University, Cambridge, Massachusetts;* H. Fraser, *Stanford University, California*

Cell Engineering
Chairpersons: Z. Lippman, *Cold Spring Harbor Laboratory;* C. Desplan, *New York University, New York*

Repairing DNA Breaks
Chairpersons: J. Stark, *Beckman Research Institute of the City of Hope, Duarte, California;* L. Symington, *Columbia University Medical Center, New York, New York*

Panel Discussion: Human Embryo Editing
Chairpersons: A. Caplan, *New York University Medical Ethics, New York;* D. Pei, *Chinese Academy of Sciences, Beijing, China;* J. Rossant, *Hospital for Sick Children, Toronto, Ontario, Canada*

Immunotherapy
Chairperson: M. Sadelain, *Memorial Sloan Kettering Cancer Center, New York, New York*

Technology II
Chairpersons: E. Olson, *University of Texas Southwestern Medical Center, Dallas;* J.K. Joung, *Massachusetts General Hospital/Harvard Medical School, Boston*

Single Biomolecules

August 28–September 1 160 Participants

ARRANGED BY **Jennifer Lippincott-Schwarz,** HHMI/Janelia Research Campus, Ashburn, Virginia
Robert Singer, Albert Einstein College of Medicine, Bronx, New York
Robert Tjian, University of California, Berkeley

J. Lippincott-Schwarz, R. Tjian, R. Singer

This inaugural meeting hosted 160 participants of which 51% were new to Cold Spring Harbor Laboratory. Participants were from 15 countries, and 33% were female. The meeting was dedicated to the memory of Maxime Dahan, who died unexpectedly this summer. Maxime was one of the original contributors to the single-biomolecule field and head of physics at the Curie Institute.

Recent developments in this field have seen the advent of lattice light-sheet imaging, further advances in SIM (such as high-speed and multifocal modes), and a keen awareness of the power of microscopy as a tool for understanding biological mechanisms. Advances in dye technology (such as the JF dyes), together with in vivo labeling methods (e.g., HaloTag and SNAP-tag) have brought single-molecule imaging to bear on a wide and growing range of biological questions. The goal of this meeting was to explore the intersections between technology and biology that will lead us to mechanisms for various biological problems illustrated by the following: How does the study of single molecules provide an essential tool where the usual ensemble measurements are not sufficient to address important biological questions? The breadth of this question was designed to invite and attract all types of biological applications. This was borne out by the notable diversity evident in the program and posters. The 66 speakers all spoke for 15 minutes each and covered topics including molecular motors, the cytoskeleton, development, chromosome organization, membrane receptors, transcription, translation, replication, viral infection, brain function, energy metabolism, microscopy, image analysis, and stress response. We expect the next meeting to be held August 25–29, 2020 to demonstrate even more maturity of this evolving field.

D. Youmans, J. Mahadevan

D. Zenklusen, G. Neuert

D. Larson, X. Daracq

S. Upadhyayula, P. Bassereau

PROGRAM

Imaging Interfaces
Chairperson: I. Cisse, *Massachusetts Institute of Technology, Cambridge*

Imaging in the Cytoplasm
Chairperson: S. Chu, *Stanford University, California*

Imaging the Nucleus I
Chairperson: T. Ha, *Johns Hopkins University School of Medicine, Baltimore, Maryland*

Super-Resolution Imaging Approaches and Applications
Chairperson: E. Holzbaur, *University of Pennsylvania, Philadelphia*

Development
Chairperson: X. Zhuang, *HHMI/Harvard University, Cambridge, Massachusetts*

Cytoskeleton
Chairperson: P. Bassereau, *Institute Curie, Paris, France*

Imaging the Nucleus II
Chairperson: L. Waller, *University of California, Berkeley*

RNA Stress Granules and mRNA Trafficking
Chairperson: A. Dernburg, *University of California, Berkeley*

T. Harden, U. Boehm

Single Biomolecules conference

Translational Control

September 4–8 352 Participants

ARRANGED BY **Thomas Dever,** National Institutes of Health, Bethesda, Maryland
Rachel Green, Johns Hopkins University School of Medicine, Baltimore, Maryland
Davide Ruggero, University of California, San Francisco

This meeting attracted participants from around the world and included Antonio Giraldez, Lori Passmore, and Peter Walter as keynote speakers, eight platform sessions. and three poster sessions that covered 274 abstracts.

D. Ruggero, T. Dever, R. Green

Bringing together a wide variety of researchers with different areas of expertise, the meeting highlighted not only the molecular mechanism of protein synthesis and its quality control but also the impacts of misregulated protein synthesis on human disease, including cancer, and neurological disease. With an impressive mix of established and young investigators, the meeting maintained the central role of the Translational Control Meeting for researchers in the community.

This year, the program was organized into seven main topics. The program opened on Tuesday with the first session on Regulation. A highlight was the keynote address by Antonio Giraldez, who described genome-wide studies in zebrafish embryos that identified new translational regulatory elements that showed similarities to control elements in yeast that were also described in this session. Additional talks in this opening session identified distinct functional states of the ribosome during translation elongation, revealed control of translation elongation during stress, and described how short coding regions in the leaders of mRNAs can act to enhance or repress translation of the main open reading frame.

The following sessions of the meeting included two additional keynote addresses by Lori Passmore, who described her studies linking translation rates with shortening of the poly(A) tail on mRNAs, and by Peter Walter, who described his studies on the memory-enhancing compound ISRIB that targets a factor that functions in cellular protein synthesis. In addition, 67 short talks

A. Giraldez, M. Biggin

S. Vasudevan, M. Hatzoglou

W. Merrick, N. Sonenberg

T. Fujiwara, M. Hentze

were selected from among the submitted abstracts, and 202 abstracts were presented as posters divided among three sessions on Wednesday afternoon, Thursday night, and Friday morning. The poster sessions were well attended in both Bush Hall and Nicholls Biondi Hall, and they prompted many discussions and much exchange of ideas. A new addition to the meeting this year was a program entitled "Hour for Change," in which the attendees discussed ways to enhance the inclusiveness of the meeting and to help advance the careers of women and other underrepresented groups in the community.

Highlights of the meeting included new insights into the mechanism and physiological regulation of protein synthesis including how translation factors and different translational control pathways contribute to disease states. Ribosomal profiling studies examining the genome-wide distribution of ribosomes on mRNAs in a cell and single-molecule FRET studies revealed codon level resolution of translation. New biochemical and structural studies provided insights on ribosome scanning and start site recognition, on how ribosome dynamics regulate the accuracy and efficiency of translation, and how ribosomes adapt to stresses by forming hibernation assemblies. Additional highlights included the importance of codon usage in translation regulation and in disease, the cotranslational assembly of multiprotein complexes, and how translation is controlled by noncanonical GTPases, by nascent peptides that arrest translation elongation, and by proteins that prevent frameshifting. Finally, studies identifying a factor that recognizes collided ribosomes, the factors responsible for release of ubiquitinated nascent chains from stalled ribosomes, and the

C. Conn, R. Bithell

W. Gilbert, W. Faller

roles of nontemplated addition of amino acids to stalled polypeptides in nascent chain turnover provided new insights into ribosome quality control.

In summary, the 2018 meeting continues the key role of the Translational Control meeting as the major focal point for the ongoing excitement and growth of this field of research.

PROGRAM

Regulation I
Chairperson: W. Gilbert, *Yale University, New Haven, Connecticut*

Quality Control
Chairperson: S. Djuranovic, *Washington University in St. Louis, Missouri*

CNS and Development
Chairperson: M. Barna, *Stanford University, California*

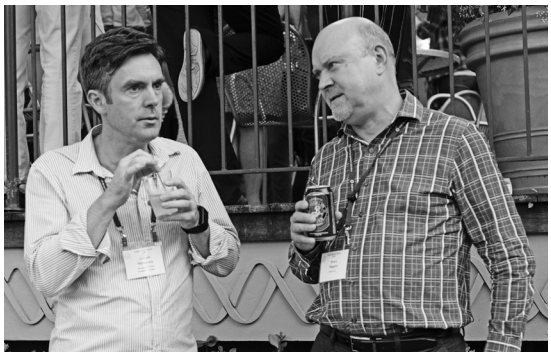
Elongation
Chairperson: E. Grayhack, *University of Rochester Medical School, New York*

Ribosome
Chairperson: A. Korostelev, *University of Massachusetts Medical School, Worcester*

Regulation II
Chairperson: S-B. Qian, *Cornell University, Ithaca, New York*

Disease
Chairperson: I. Topisirovic, *McGill University, Montréal, Canada*

Initiation
Chairperson: E. Jankowsky, *Case Western Reserve University, Cleveland, Ohio*

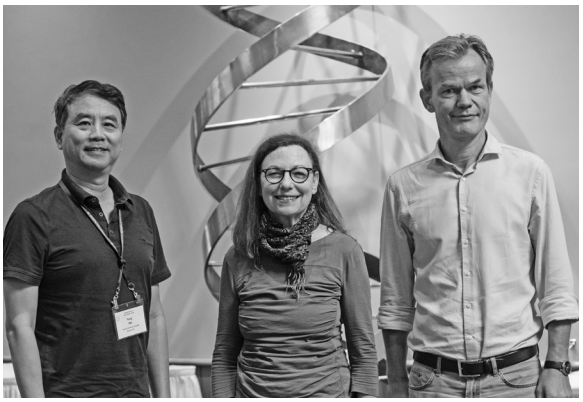


G. McInerney, B. Maguire

Epigenetics and Chromatin

September 11–15 358 Participants

ARRANGED BY **Shelley Berger,** University of Pennsylvania, Philadelphia
Kristian Helin, University of Copenhagen, Denmark
Yang Shi, Harvard Medical School, Boston, Massachusetts



Y. Shi, S. Berger, K. Helin

This fourth conference followed the successful previous meetings in 2012, 2014, and 2016. As in the earlier meetings, the conference brought together researchers working on different aspects of chromatin biology and epigenetics. The focus was on molecular mechanisms by which chromatin and chromatin-modifying protein complexes regulate gene transcription and the genome. The spectrum of oral and poster presentations demonstrated that the field has been making progress through studies that range from structure and function of chromatin proteins, to epigenome and genomic architecture, to functional analyses of regulatory networks at the organism level. New emerging areas of focus this year included metabolism and nuclear phase transition, as well as disease implications, which continues to be a burgeoning interest in the broad chromatin and epigenetics fields.

The high attendance, large numbers of posters (209), and the sustained involvement of the participants throughout the meeting demonstrated the success of the meeting. Scientists from 28 countries attended the meeting, and for many, this was their first visit to Cold Spring Harbor Laboratory. Several oral presentations selected from abstracts and given by junior PIs, postdocs, or students turned out to be scientific highlights, showing the success of this meeting format as a presentation platform for junior researchers. The emergence of novel mechanistic themes in the field, the progress in methodologies, and the presentations by junior scientists allowed for sustained interest by the large and diverse audience throughout the different sessions.

S. Xie, X. Zhuang, A. Rao, I. López-Moyado

K.J. Armache, H. Madhani

A. Schaefer, P. Casaccia

K. Monahan, M. Currie

There were eight plenary sessions. There were two sessions on chromatin processes controlling three-dimensional genome organization, and another on the role of enhancers in gene expression. The function and mechanisms of DNA and histone-modifying enzymes in gene activation and repression was covered in one session. Another session covered epigenetic mechanisms in germ cells, parental imprinting, and development at the organism level. Yet another session focused on inheritance of chromatin states. One session covered phase transition in the nucleus, and one covered the influence of metabolism and nuclear metabolic enzymes in the nucleus regulating chromatin states.

Among the highlights of the meeting were exciting presentations using new methods and approaches to query inheritance of nucleosomes and histone modifications through mitosis, which are issues at the heart of epigenetics as a means to preserve genomic memory. Talks about phase transition in the nucleus correlating with heterochromatin and euchromatin states led to spirited discussion among participants. Another area of current debate covered in the meeting was the function and regulatory pathways underlying 3D genome organization. Key questions and approaches were discussed about positioning of transcriptionally active genes within the nucleus and inheritance of these locations. Other talks covered new and accumulating evidence for metabolic enzymes in the nucleus to directly influence chromatin state and histone modifications. Rising issues and technologies for single-cell chromatin biology and 3D architecture were discussed, which are assuming increasing importance in the field. These numerous important and fascinating topics demonstrated the key role of chromatin in genome regulation and the crucial questions that still await exploration and discovery.

This meeting was funded, in part, by Abcam, Arima Genomics, and EpiCypher.

PROGRAM

Keynote Speaker
W. Bickmore, *MRC Human Genetics Unit, Edinburgh, United Kingdom*

Architecture
Chairperson: G. Narlikar, *University of California, San Francisco*

Gene Regulation and Phase Separation
Chairpersons: A. Rao, *La Jolla Institute for Allergy and Immunology, California;* C. Thompson, *Memorial Sloan Kettering Cancer Center, New York, New York*

Metabolic Epigenetics
Chairpersons: R. Kingston, *Massachusetts General Hospital/ Harvard Medical School, Boston;* M. Bartolomei, *University of Pennsylvania Perelman School of Medicine, Philadelphia*

Germ Cells and Development
Chairpersons: G. Crabtree, *Stanford University, California;* E. Furlong, *European Molecular Biology Laboratory, Heidelberg, Germany*

Keynote Speaker
G.R. Crabtree, *Stanford University, California*

S. Henikoff, G. Crabtree

A. Rojas, D. Reinberg

Chromatin Machines
Chairpersons: R. Young, *Whitehead Institute/MIT, Cambridge*; A. Groth, *BRIC, University of Copenhagen, Denmark*

Epigenetic Memory
Chairpersons: A. Shilatifard, *Northwestern University Feinberg School of Medicine, Chicago, Illinois;* J. Lee, *Massachusetts General Hospital/Harvard Medical School, Boston*

Enhancers, Chromatin and Disease
Chairpersons: X. Zhuang, *Harvard University, Cambridge, Massachusetts;* C. Vakoc, *Cold Spring Harbor Laboratory*

Architecture and Disease
Chairpersons: S. Berger, *University of Pennsylvania, Philadelphia;* Y. Shi, *Harvard Medical School, Boston, Massachusetts*

Molecular Mechanisms of Neuronal Connectivity

September 25–29 252 participants

ARRANGED BY **Linda Richards,** University of Queensland, Brisbane, Australia
Peter Scheiffele, University of Basel, Switzerland
Kang Shen, Stanford University, California
Yimin Zou, University of California, San Diego

Y. Zou, L. Richards, P. Scheiffele, K. Shen

As the field of neuroscience expands and evolves, the organizers worked together to redefine the identity of this meeting series by changing its name to "Molecular Mechanisms of Neuronal Connectivity." With this new name, research areas covered by the meeting include axon, dendrite, synapse, and glial biology under normal and disease conditions. We have put the word "molecular" up front to distinguish this meeting from other neural circuit meetings. We had a great turnout of experts in these related fields. The overall quality of science was high.

Speakers were chosen from among the participants submitting abstracts by session chairs who are leaders in the field. Forty-five percent of the participants were women, and this was reflected in the number of talks presented by women. Fifty-five abstracts in seven sessions were selected for talks, with the remaining abstracts being presented as posters. Advanced and starting assistant professors, postdoctoral fellows, and graduate students were well represented as speakers and participants. Session chairs at the meeting were well balanced between men and women, and the meeting had a clear international presence with participants from India, Israel, Germany, France, Japan, China, the United States, and Australia. Sixteen poster prizes were given out. The following are the highlights of the meeting.

All of the major areas of research in the field were covered, as were all of the major approaches (cellular, physiological, anatomical, molecular, dynamic imaging, biochemical, and genetic). In addition, there were two keynote addresses. The first was given by Professor Carol Mason, a leader in the field of axon guidance. She summarized the major contributions of her career in understanding axon guidance by her in-depth studies of mid-crossing of the optical nerves. She also pointed out that there are many questions remaining in axon guidance. The field is not "done." The second

R. Kellermeyer, T. Kidd

E. Ginger, L. Ma

S. Palumbos, D. Miller

C. Mason, K. Brose

keynote address by Noble laureate Eric Betizg dazzled us with his beautiful images and exciting new imaging modalities. One noticeable advance from several talks at the meeting was that the light sheet microscopy that Dr. Betzig helped develop has now dramatically improved the live imaging of axon development in vivo. He is also making progress on adaptive optics to allow deeper imaging overcoming tissue aberration. The third keynote was given by Marc Tessier-Lavigne, an internationally renowned scientist in the field of axon guidance. He detailed the discovery of netrin and recent progress in understanding the in vivo function of netrin in mouse. His talk summarized work both from his lab and from several other groups, which together enriched our understanding of how the nervous system is wired up, using netrin as an example. We have also included a special lecture from Dr. Katja Brose, who talked about her vision of philanthropy in neuroscience and how to organize large collaborative initiatives. She represents the Chan Zuckerberg Initiative.

This meeting also held several lunches targeted at professional development. Trainees were invited to interact with senior scientists to discuss topics as diverse as how to search for mentors (postdoc or faculty), prepare for job interviews (academic or industry), write grants, manage teaching and research, balance life and work, and handle paper reviews. The students and postdocs expressed high appreciation for the close interaction and valuable advice they gained through this event. The business meeting was conducted with open discussion.

Overall, this meeting provided an important forum for ideas and approaches in developmental neuroscience and regeneration and helped scientists in the field get the most up-to-date information, as well as enabling them to meet, network, and establish collaborations. Indeed, many

S. Emmons, T. Sengupta

S. Santiago, E. Stoeckli

expressed an appreciation for the smaller size of the meeting in promoting more interactions, as well as the social events including the wine and cheese, banquet, and informal lunches and dinners. The three poster sessions were extremely well attended. Based on the enthusiastic comments of the participants, the intensity of the oral and poster sessions, and the large crowds that stayed up late every night at the bar to discuss science, the meeting was truly a great success and highly appreciated by each of the attendees. Some people voiced their inclination to go back to the format in which only students and postdocs give talks.

This meeting was funded, in part, by the National Institute of Neurological Diseases and Stroke, a branch of the National Institutes of Health.

PROGRAM

Programs of Circuit Wiring
Chairpersons: S. Garel, *Institut de Biologie de l'ENS de L'École, Paris, France;* S. Millard, *University of Brisbane, St. Lucia, Queensland, Australia*

Axon Targeting
Chairpersons: R. Klein, *Max Planck Institute of Neurobiology, Martinsried, Germany;* R.J. Pasterkamp, *University Medical Center Utrecht, Netherlands*

Keynote Speaker
C.A. Mason, *Columbia University, New York, New York*

Keynote Speaker
E. Betzig, *HHMI/Janelia Research Campus, Ashburn, Virginia, and HHMI/University of California, Berkeley*

Axon Regeneration
Chairpersons: V. Cavalli, *Washington University in St. Louis, Missouri;* M. Hammarlund, *Yale University, New Haven, Connecticut*

Synapse Assembly
Chairpersons: S.X. Bamji, *University of British Columbia, Vancouver, Canada;* S.H. Soderling, *Duke University, Durham, North Carolina*

Synapses in Neural Circuits
Chairpersons: Y. Jin, *University of California, San Diego;* X. Yu, *Institute of Neuroscience, Shanghai, China*

Keynote Speaker
M. Tessier-Lavigne, *Stanford University, California*

Glial Control of Neuronal Circuitry
Chairperson: C. Eroglu, *Duke University Medical Center, Durham, North Carolina*

Neuronal Repair
Chairpersons: M. Rolls, *The Pennsylvania State University, University Park;* Z. He, *Children's Hospital, Boston, Massachusetts*

Special Lecture
K. Brose, *Chan Zuckerberg Initiative*

Synapses in Neural Circuits II
Chairpersons: Y. Goda, *Riken, Wako, Japan;* C. Gross, *EMBL, Monterotondo, Italy*

Neuronal Plasticity
Chairpersons: Y. Hayashi, *Kyoto University Graduate School of Medicine, Japan;* R. Yasuda, *Max Planck Florida Institute for Neuroscience, Jupiter*

Mechanisms of Aging

October 1–5 306 participants

ARRANGED BY **Vera Gorbunova,** University of Rochester, New York
Malene Hansen, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California
Scott Pletcher, University of Michigan, Ann Arbor

V. Gorbunova, S. Pletcher, M. Hansen

Aging is the most prominent risk factor for developing a number of serious diseases, ranging from neurodegeneration to cancer. The rate of aging and the appearance of disabilities that result from it are known to be influenced by genes and environment. Indeed, recent advances in the biology of aging have revealed several evolutionarily conserved molecular pathways and common environmental factors that modulate life span and susceptibility to age-related disorders across taxa, from yeast to humans. This conference provided a diverse and stimulating forum for presenting the latest research findings, for discussing new challenges in aging research at the molecular, cellular, and organismal levels, and for promoting educational opportunities for young scientists (e.g., via the new lunch event called "Meet the Speakers"). The conference opened with a session on longevity genes and pathways, which provided an update on the latest findings on specific genes linked to the aging process and their age-related effects. The second day opened with a morning session focused on exciting new developments in cellular senescence and DNA fidelity, with new emphasis on how transposable elements influence the aging process. The afternoon session introduced the new topic of intertissue communication, with research from a variety of model systems highlighting neural and environmental control of aging. A stimulating afternoon social gathering was followed by an evening poster session of 65 posters. The third day of the conference featured oral sessions on stem cell biology and homeostasis, in which a variety of model systems were used to highlight important areas such as inflammation, autophagy, and proteostasis. Again, the evening featured a poster session with 68 posters (133 posters in total). The fourth day of the

B. Rogina, D. Promislow

K. Kornfeld, Z. Kocsisova

D. Glass, D. Cai

N. Mazloum, J. Tyler

conference was marked by oral sessions that focused on epigenetics and metabolism. New discoveries linking modulation of DNA methylation with life span and characterizing epigenetic clocks were presented in the morning session, whereas the metabolic implications of nutrient signaling and gut microbiota were among the topics in the afternoon. The morning session on the fifth and final day focused on interventions, and it highlighted drugs and other noninvasive approaches to interfere with the aging process. A new initiative sponsored by the National Institute on Aging to study aging and intervention in a large cohort of dogs was introduced. The conference was attended by a large group of 306 researchers from many different fields who collectively created an exciting and interactive environment for established scientists and trainees to interact and who introduced new, thought-provoking ideas and research directions for understanding the molecular mechanisms of aging, so that better treatments can be developed to protect against its progression and the disease and disability with which it is associated.

This meeting was funded in part by the National Institute on Aging, a branch of the National Institutes of Health; *Aging Cell*; *Aging*; Calico; and The Glenn Foundation for Medical Research.

PROGRAM

Longevity Genes and Human Mutations
Chairpersons: K. Chua, *Stanford University School of Medicine, California;* C. Murphy, *Princeton University, New Jersey*

Senescence, DNA Repair and Transposable Elements
Chairpersons: J. Campisi, *Buck Institute for Research on Aging, Novato, California;* L. Harrington, *Université de Montréal,*

S-J. Le, A. Dillin

A. Meeker, S. Matsuyama

J. Plummer, S. Postnikoff

P. Spinelli, D. Palliyaguru

Quebec, Canada; V. Bohr, *National Institute on Aging, Baltimore, Maryland*

Inter-Tissue Communication
Chairpersons: A. Dillin, *HHMI/University of California, Berkeley;* D. Cai, *Albert Einstein College of Medicine, Bronx, New York*

Stem Cells
Chairpersons: T. Rando, *Stanford University, California;* G. Liu, *Institute of Biophysics, Chinese Academy of Sciences, Beijing;* L. Jones, *University of California, Los Angeles*

Homeostasis
Chairpersons: A. Antebi, *Max Planck Institute for Biology of Aging, Cologne, Germany;* R. Buffenstein, *Calico Life Sciences, LLC, South San Francisco, California*

Epigenetics
Chairpersons: S. Berger, *University of Pennsylvania, Philadelphia;* S. Helfand, *Brown University, Providence, Rhode Island;* S. Horvath, *University of California, Los Angeles*

Metabolism
Chairpersons: E. Verdin, *Buck Institute for Research on Aging, Novato, California;* D. Glass, *Novartis Institutes for Biomedical Research, Cambridge, Massachusetts;* D.R. Valenzano, *Max Planck Institute for Biology of Aging, Cologne, Germany*

Interventions
Chairpersons: R. Anderson, *University of Wisconsin, Madison;* R. Miller, *University of Michigan, Ann Arbor*

Germ Cells

October 9–13 196 Participants

ARRANGED BY **David Page,** HHMI/MIT/Whitehead Institute, Cambridge, Massachusetts
Yukiko Yamashita, HMMI/University of Michigan, Ann Arbor

D. Page, Y. Yamashita

This year's meeting marked the 20th anniversary (11th meeting) of the biennial meeting on Germ Cells first held at Cold Spring Harbor Laboratory in 1998. Traditionally, this meeting has emphasized the importance of studies from model organisms, integrating the knowledge from a broad spectrum of systems from lower eukaryotes to humans. Continuing this tradition, this year's meeting included talks on germ cells in diverse systems such as quail and fission yeast. The talks and posters were of high quality, and the poster sessions were well attended.

The keynote talk was presented by Dr. Cassandra Extavour, who spoke on the network analysis of signaling pathways that determine egg production in *Drosophila*. Her talk provided insights into how reproductive outputs are wired by modular interaction of signaling networks from the evolutionary perspective.

For this year's meeting, the co-organizers had a major goal of introducing a "new way of thinking" to the community of germ cell biology, in addition to maintaining the high quality of the meetings. To this end, the co-organizers introduced a new session ("Cheating"), which covered genetic conflicts that manipulate and modulate the process of gametogenesis. As gametogenesis is deeply and inevitably influenced by selfish genetic elements, their influence cannot be ignored in studying germ cell biology. This session was accepted by attendees extremely well. In addition, co-organizers intended to make session organization more inclusive, instead of dividing sessions by traditional themes. Each session included a broader range of topics than traditional session structures, tied by less-defined themes.

D. Katz, C. Mello

T. Oyewale, C. Eckmann, L. Jensen

D. Greenstein, S. Payne, G. Seydoux, E. Griffin

P. Rangan, E. Bach

The participants and presentations at this meeting were well balanced in multiple dimensions (gender, experimental model systems, career stage). Of a total of 54 talks, 28 presenters were male and 26 female; 20 talks used mouse as a model system, 20 fly, eight worm, three fish and other diverse organisms (parasitic worm [schistosome], quail, fission yeast); and 30 presenters were group leaders (at all ranks), 23 trainees, and one senior scientist. Additionally, of 197 participants, 53% were female and 52% were trainees (23% graduate students, 29% postdoctoral fellows). The two afternoon poster sessions were extremely well attended.

The meeting thus continues to be a great success judged by the criteria of exemplary science, the breadth and depth of its discussion, and attendance. The co-organizers have always been drawn from the broader germline community, each with appreciation of the strengths of this meeting. Thus, Ruth Lehman (New York University), Erez Raz (University of Munster, Germany), and Shosei Yoshida (NIBB, Japan), who eminently meet these criteria, were selected as the 2020 co-organizers. We are confident that they will organize a successful meeting that will continue to grow while retaining its participant, organism and system diversity.

This meeting was funded in part by the National Institute of Child Health and Human Development, a branch of the National Institutes of Health; and The Lalor Foundation.

M. de Cuevas, K. Lenhart

J. Escher, A. Spradling

PROGRAM

Keynote Session
D. Page, *HHMI/MIT Whitehead Institute, Cambridge, Massachusetts;* Y. Yamashita, *HHMI/University of Michigan, Ann Arbor*

Germ Cell Development
Chairperson: E. Raz, *University of Münster, Germany*

One Generation after Another
Chairperson: G. Seydoux, *HHMI/Johns Hopkins University School of Medicine, Baltimore, Maryland*

Affairs of the Nucleus
Chairperson: A. Dernburg, *HHMI/University of California, Berkeley*

Cheating
Chairperson: M. Lampson, *University of Pennsylvania, Philadelphia*

Something for Everybody
Chairperson: E. Bach, *New York University School of Medicine, New York*

I. Chambers, D. Laird

RNA Rules
Chairperson: M. Fuller, *Stanford University School of Medicine, California*

Genome Integrity
Chairperson: A. Aravin, *California Institute of Technology, Pasadena*

The Evolving Concept of Mitochondria: From Symbiotic Origins to Therapeutic Opportunities

October 18–21 159 Participants

ARRANGED BY **Mila Pollock,** Cold Spring Harbor Laboratory
Anu Suomalainen, University of Helsinki, Finland
John Walker, Mitochondrial Biology Unit, University of Cambridge, United Kingdom
Douglas Wallace, Children's Hospital of Philadelphia, Pennsylvania

In celebration of the 30th anniversary of the discovery that mitochondrial DNA (mtDNA) mutations can cause disease, Ludmila (Mila) Pollock, of Cold Spring Harbor Laboratory, proposed a Cold Spring Harbor Laboratory Symposium on mitochondrial biology and medicine. To assist in organizing this meeting, Mila Pollock recruited Drs. Suomalainen-Wartiovaara, Walker, and Wallace. This meeting was the seventh in a series of successful historically grounded symposia organized by Mila Pollock that emphasize outstanding research but are presented within an historical perspective.

A. Suomalainen, M. Pollock

The discovery that mtDNA mutations can contribute to diseases opened a broad new perspective on the etiology of previously mysterious rare and common complex diseases and promises an array of new therapeutic approaches. Moreover, the realization of the medical importance of mitochondrial function has revitalized investigations into the origins and biology of this critical human organelle. To enunciate the growing relevance of mitochondrial biology and evolution to human health and disease, the meeting was titled, "The Evolving Concept of Mitochondria: From Symbiotic Origins to Therapeutic Opportunities."

J. Walker, B. Spiegelman

H. Higgs, E. Shoubridge

The mitochondrion was first observed within cells at the end of the nineteenth century. Since then, the multifaceted mitochondrion has engendered fascinating research from a broad spectrum of physical, chemical, biological, and medical perspectives. As a result, the "Evolving Concept of Mitochondria" symposium encompassed a rich and varied history, a broad range of biomedical research approaches, and a plethora of powerful new concepts for understanding and treating common metabolic, inflammatory, and degenerative diseases, cancer, and aging. An illustration of the topical diversity is the range of concepts and experimental techniques encompassed by the speakers from philosophically oriented biochemical discussions on the origins of life and mitochondria, through mitochondrial cell biology using the most advanced molecular genetic and imaging techniques, to mitochondrial protein molecular dynamics as elucidated by X-ray crystallography and cryo-electron microscopy.

The main topics covered in the symposium are listed below under Program. These sessions encompassed 47 lectures by the very top investigators in the mitochondrial biology and medicine fields as well as 65 poster presentations on a striking range of novel mitochondrial topics. In attendance were approximately 200 people from a broad range of backgrounds including scientists, clinicians, historians, scholars, and science journalists.

In summary, this meeting was an exceptional success worthy of the high reputation of Cold Spring Harbor Laboratory Symposia. All talks are now available on the specially dedicated website at http://library.cshl.edu/Meetings/History-of-Science/.

Generous support of this meeting was received from Janssen Cardiovascular & Metabolism (Johnson & Johnson); Agilent; Biolog; Mitobridge (an Astellas company): and Neurovive.

J. Wagner, S. Iyer

D. Wallace, R. Youle

E. Schon, E. Shoubridge

N. Chandel, J. Hirst

PROGRAM

Mitochondrial Biology, Bioenergetics, Biogenesis, and Disease
Chairperson: A. Suomalainen, *University of Helsinki, Finland*

The Origins of the Organelle
Chairperson: V. Procaccio, *University of Angers, France*

Complex I: From Structure to Disease
Chairperson: P. Rich, *University College London, United Kingdom*

Mitochondrial DNA: Structure, Function, and Inheritance
Chairperson: A. Chomyn, *California Institute of Technology, Pasadena*

Mitochondrial Expression System and Proteome
Chairperson: M. Falkenberg, *University of Gothenburg, Sweden*

Mitochondrial Protein Import Systems and Metabolic Transporters
Chairperson: F. Palmieri, *University of Bari Aldo Moro, Italy*

Mitochondrial Diseases, Past and Present
Chairperson: M. Falk, *Children's Hospital of Philadelphia, Pennsylvania*

Mitochondrial Regulation of Metabolism
Chairperson: V. Carelli, *University of Bologna School of Medicine, Italy*

Banquet Speakers
A. Suomalainen, *University of Helsinki, Finland*
J. Walker, *University of Cambridge, United Kingdom*
D. Wallace, *Children's Hospital of Philadelphia, Pennsylvania*

Mitochondria, Calcium Signaling, and Cell Death
Chairperson: J. Walker, *University of Cambridge, United Kingdom*

Mitochondrial Dynamics and Quality Control
Chairperson: D. Wallace, *Children's Hospital of Philadelphia, Pennsylvania*

Novel Approaches to Mitochondrial Therapy
Chairperson: A. Suomalainen, *University of Helsinki, Finland*

Wrap-Up Discussion: Preserving the History of Mitochondrial Research
A. Suomalainen, *University of Helsinki, Finland*
J. Walker, *University of Cambridge, United Kingdom*
D. Wallace, *Children's Hospital of Philadelphia, Pennsylvania*
M. Pollock, *Cold Spring Harbor Laboratory*

Nutrient Signaling

October 25–28 117 Participants

ARRANGED BY **Anne Brunet,** Stanford University, California
Navdeep Chandel, Northwestern University Feinberg School of Medicine, Chicago, Illinois
David Sabatini, Whitehead Institute for Biomedical Research, Cambridge, Massachusetts
Lloyd Trotman, Cold Spring Harbor Laboratory

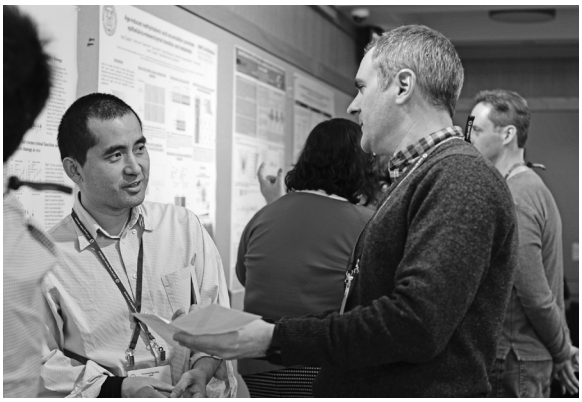
D. Sabatini, N. Chandel, A. Brunet, L. Trotman

This meeting hosted participants from 26 different countries, of which 49% were female. This meeting built and expanded on the prior successful meeting series on the PI3K-mTOR-PTEN network held biennially since 2006. Recent developments in this field have provided a new connection between nutrient signaling pathways and central metabolism, thereby providing a molecular understanding on how diet can influence cell and organismal responses. The nutrient signaling meeting brought cutting-edge technologies, including metabolomics, full genome screens, and single-cell approach, to bear on a wide range of biological processes, including aging and cancer. The goal of this meeting was to explore fundamental biological questions: How do nutrient sensing and signaling occur? Which organelles in the cells are implicated in nutrient sensing? How does nutrient signaling impact metabolism? Another goal of the meeting was to determine how these fundamental processes get dysregulated in pathological processes, with the following central questions: How are the connections between nutrient sensing and metabolism disrupted during aging and can this be leveraged to slow or revert aspects of aging and aging-related diseases? How do cancer cells hijack nutrient signaling, and its connection to metabolism, to develop, and can this be used to develop new therapeutic approaches? The breadth of these questions was designed to attract all types of applications. This was borne out by the notable diversity evident in the program and poster abstracts. The 45 speakers spoke for 20 or 12 minutes, as invited speakers or speakers selected from the abstracts, respectively. The speakers covered a broad range of topics including signaling pathways, autophagy, aging, mitochondria, metabolomics and cancer, and dietary restriction. We expect

T. Osawa, S. Sukumaram

B. Manning, M. Bland

C. Sing, Y. Rong



J. Du, R. DeBerardinis

the next meeting to be held in 2020 to reflect the growing interests in how changes in the diet mechanistically impact cellular metabolism to regulate cellular and organismal processes, such as aging and cancer.

Generous support was provided by Northwell Health.

PROGRAM

Metabolism I
Chairperson: N. Chandel, *Northwestern University Feinberg School of Medicine, Chicago, Illinois*

L. Abalde-Atristain, G. Hardie

Keynote Speaker
B.D. Manning, *Harvard T.H. Chan School of Public Health, Boston, Massachusetts*

Lysosomes/Endosomes in Nutrient Signaling
Chairperson: H. Christofk, *University of California, Los Angeles*

Cancer
Chairperson: R. Shaw, *The Salk Institute for Biological Studies, La Jolla, California*

Metabolism II
Chairperson: K. Wellen, *University of Pennsylvania, Philadelphia*

Metabolism and Physiology
Chairperson: A. Dillin, *University of California, Berkeley*

Aging
Chairperson: M. Hansen, *Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California*

Transposable Elements

November 1–4 140 Participants

ARRANGED BY **Rob Martienssen,** Cold Spring Harbor Laboratory
Phoebe Rice, University of Chicago, Hyde Park
Donald Rio, University of California, Berkeley

D. Rio, D. Schatz

This second meeting brought together leading experts, junior and senior investigators, postdoctoral fellows, and graduate students working on many different aspects and systems of transposons. The six oral sessions and one poster session included microbial, animal, and plant transposon presentations. Material was presented at the structural, biochemical, biological, medical, and applied levels and spanned a variety of methods, approaches, and experimental systems. The session chairs were not instructed to give introductory comments at the beginning of each session, but some did, and this was helpful because of the broad nature of the participants. There were 140 participants in attendance, with good numbers of female and younger, more junior, scientists participating. There was lively and extensive discussion following most of the talks and this was a highlight of the meeting, as was the poster session, which could have lasted two to three more hours. Given this was the second time this meeting was held, it was a success.

The meeting was funded in part by the Oliver Grace Fund.

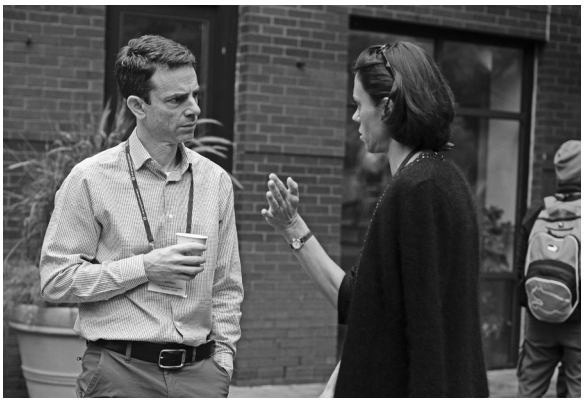
PROGRAM

Opening Keynote Session
C. Feschotte, *Cornell University, Ithaca, New York*

DNA Transposons
Chairperson: O. Barabas, *European Molecular Biology Laboratory, Heidelberg, Germany*

T. Kakutani, C. Pikaard

F. Karam Teixeira, V. Colot, A. Denli



M. Lorincz, M-E. Torres-Padilla

G. Schumann, C. DeLuca

Retrotransposons
Chairperson: K. Burns, *Johns Hopkins University School of Medicine, Baltimore, Maryland*

Transposon Control
Chairperson: H. Levin, *National Institute of Child Health and Development, National Institutes of Health, Bethesda, Maryland*

Host–Transposon Interactions
Chairperson: M. Gehring, *Whitehead Institute for Biomedical Research, Cambridge, Massachusetts*

Variety and Evolution of Transposable Elements
Chairperson: S. Wessler, *University of California, Riverside*

Probabilistic Modeling in Genomics

November 4–7 154 Participants

ARRANGED BY **Gerton Lunter,** University of Oxford, United Kingdom
Molly Przeworski, Columbia University, New York, New York
Adam Siepel, Cold Spring Harbor Laboratory

G. Lunter, A. Siepel

This fourth annual conference, and second one hosted at Cold Spring Harbor, grew out of two ad hoc meetings on a similar topic in 2013 at Janelia Farm and in 2014 at Merton College, Oxford. The main goal of the meeting is to provide a forum for presentation and exchange of ideas among researchers who are working in the general area of genomics but are particularly focused on the development of new probabilistic models, algorithms, and methods for inference. These researchers come from a variety of backgrounds, including computer science, statistics, applied mathematics, and physics. Previous workshops on this topic have strongly emphasized population genetics, but the program this time was designed to incorporate additional areas of interest such as functional genomics, systems biology, quantitative genetics, and cancer evolution.

The session topics are listed below under Program. There were six oral presentations per session for a total of 36 presentations. Two invited session chairs presented in each session, and four additional talks were selected from submitted abstracts. All talks were 20 minutes long plus five minutes for questions and answers. The quality of the presentations was very high overall, with considerable mathematical sophistication combined with a clear focus on biological relevance.

We were honored to recruit two distinguished senior scientists to attend the meeting and present keynote lectures: Katherine Pollard and Matthew Stephens. Dr. Pollard gave a broad survey of research on the genomics of the human microbiome, including her work and that of other groups. She emphasized an emerging line of research that treats the microbiome as an extension of the human genome itself and strives to fully characterize its genomic content, genetic variation, and

D. Enard, L. Cappello

R. Cartwright, A. Williams

S. Gravel, A. Clark

L. Hayward, R. Brown

association with disease. Dr. Stephens gave an overview of his work, which has made seminal impact on the field, focusing on three pieces of research in more detail and giving an interesting and thoughtful account of the statistical discoveries, as well as the nonlinear process of research itself.

In general, the meeting was characterized by a great deal of enthusiastic discussion, and a collaborative and collegial atmosphere. Several attendees remarked, often spontaneously, of their appreciation of this conference, which seems to be successful at filling the niche that was identified a few years ago, and indeed several attendees commented that this was their favorite conference. The next conference is planned in fall 2019 in France, organized by Laurent Duret and colleagues. Ian Holmes (University of California, Berkeley), who organized the meeting in Ohio that sowed the seed for the PROBGEN series, expressed an interest to be involved in organizing future meetings. Debra Marks (a session chair this time) also expressed an interest in playing an ongoing role in organization.

PROGRAM

Population Genetics I: Mutation, Recombination and Demography Inference
Chairpersons: K. Harris, *University of Washington, Seattle;* S. Sankararaman, *University of California, Los Angeles*

Transcriptomics and Epigenomics
Chairpersons: M. Hoffman, *University of Toronto, Ontario, Canada;* S. Krishnaswamy, *Yale University, New Haven, Connecticut*

M.E. Lauterberg, A.J. Laruson

J. Kelleher, A. Bloemendal

Keynote Speaker
K. Pollard, *Gladstone Institutes/University of California, San Francisco*

Cancer, The Microbiome, and Beyond
Chairpersons: E. Davenport, *Cornell University, Ithaca, New York;* Q. Morris, *University of Toronto, Ontario, Canada*

Quantitative Genetics and Association Mapping
Chairpersons: D. Balding, *University of Melbourne, Victoria, Australia;* B. Engelhardt, *Princeton University, New Jersey*

Systems and Structural Biology
Chairpersons: D. McCandlish, *Cold Spring Harbor Laboratory;* D. Marks, *Harvard Medical School, Boston, Massachusetts*

Keynote Speaker
M. Stephens, *University of Chicago, Hyde Park*

D. Bredikhin, L. Cappello

Population Genetics II: Natural Selection
Chairpersons: S. Ramachandran, *Brown University, Providence, Rhode Island;* K. Lohmueller, *University of California, Los Angeles*

Biological Data Science

November 7–10 280 Participants

ARRANGED BY **Christina Curtis,** Stanford University, California
Barbara Engelhardt, Princeton University, New Jersey
Jeff Leek, Johns Hopkins University, Baltimore, Maryland
Michael Schatz, Cold Spring Harbor Laboratory

J. Leek, E. Biederstedt

This conference was the third session in the new biennial series focusing on the infrastructure, software, and algorithms needed to analyze large data sets in biological research. It brought together researchers from diverse areas of biology and diverse computational domains, with interest in large-scale data analysis in common. The goal of the meeting was for attendees to step outside their usual domain and glean lessons from other areas of research. In this regard, the meeting was an overwhelming success with extensive discussion and exchange of ideas.

Two hundred and eighty researchers participated in the meeting, with 180 abstracts submitted from academic, government, and industry laboratories by researchers who regularly mine large data sets for genomics, imaging, translational, and clinical projects. There were 10 invited talks, two keynote lectures, and two panel discussions. The remaining 36 talks were selected from submitted abstracts along with 130 poster presentations. The meeting also featured 20 lightning talks to showcase a selection of posters presentations in the main auditorium. The discussion topics are listed below under Program. Immediately after the meeting, more than 25 attendees participated in a Biomedical Data Science Hackathon co-organized with NCBI in which four teams spent three days prototyping new computational tools for analyzing structural variations, for building and aligning to graph genomes, for studying cancer epitopes, and for ultrafast read mapping.

The keynote lectures were delivered by Anne Carpenter and Steven Salzberg. Dr. Carpenter, from the Broad Institute, delivered a lecture on "Making Microscopy Images as Computable as Genomes: Drug Discovery in the Age of Deep Learning." Dr. Salzberg gave a lecture on "Adventures and Misadventures in Genome Data Mining." The first panel was a forum on the Future of

L.C. Irber, Jr., F. Navarro

J. Webber, L. van't Veer

N. Matasci, V. Schneider

E. Biederstedt, F. Sedlazeck

Biological Data Sciences moderated by Adam Phillippy in which we discussed the definition of the field, future of funding, and key issues such as privacy. The second panel on the role of deep learning in biology was moderated by Michael Schatz and led to extensive discussions between the academic and industry researchers on the panel. The keynotes, discussion, and hackathon were unique strengths of the meeting with unanimous support to continue in future years.

Thanks to ePlus Technology Inc. for their support of this meeting.

PROGRAM

Algorithmics
Chairpersons: M. Gymrek, *University of California, San Diego;* A. Phillippy, *National Human Genome Research Institute, Bethesda, Maryland*

Machine Learning
Chairpersons: A. Goldenberg, *University of Toronto, Ontario, Canada;* M. Hoffman, *University of Toronto, Ontario, Canada*

PANEL: The Future of Biological Data Science
Chairperson: A. Phillippy, *National Human Genome Research Institute, Bethesda, Maryland*

Keynote Speaker
A.E. Carpenter, *The Broad Institute, Cambridge, Massachusetts*

Tools, Infrastructure, and Visualization
Chairpersons: H. Corrada Bravo, *University of Maryland, College Park;* S. Holmes, *Stanford University, California*

Single Cell
Chairpersons: A. Raj, *University of Pennsylvania, Philadelphia;* C. Vallejos Meneses, *University of Edinburgh, United Kingdom*

O. Botvinnik, J. Schreiber

M. Knudsen, S. Vang

PANEL: The Future of Deep Learning in Biology
Chairperson: M. Schatz, *Johns Hopkins University, Baltimore, Maryland*

Personalized Medicine and Biomarkers
Chairpersons: G. Parmigiana, *Dana-Farber Cancer Institute, Boston, Massachusetts;* L. van't Veer, *University of California, San Francisco*

Keynote Speaker
S.L. Salzberg, *Johns Hopkins University School of Medicine, Baltimore, Maryland*

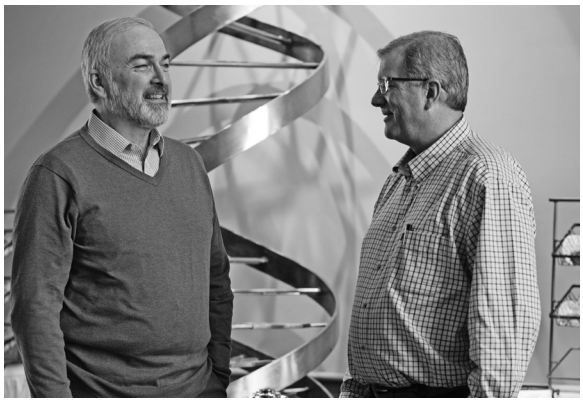
Imaging
Chairpersons: B. Andrews, *University of Toronto, Ontario, Canada*; J.H. Lee, *Cold Spring Harbor Laboratory*

T. Madden, C. Bolipata

Neurodegenerative Diseases: Biology and Therapeutics

November 28–December 1 202 participants

ARRANGED BY **Richard Ransohoff,** Third Rock Ventures, Boston, Massachusetts
Scott Small, Columbia University, New York, New York
John Trojanowski, University of Pennsylvania School of Medicine, Philadelphia
Li-Huei Tsai, Massachusetts Institute of Technology, Cambridge



A. Krainer, F. Bennett

Up to one-half of those aged 85 years or older will develop debilitating neurodegenerative diseases (NDDs) of the central nervous system. These diverse NDDs include Alzheimer's disease (AD), Parkinson's disease (PD), frontotemporal lobe degeneration (FTLD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), as well as combinations of these NDDs such as AD combined with PD and FTLD with ALS.

This biannual meeting series addressing NDD started in 2000 under the organizing leadership of Sam Gandy, Harry Levine, and Marcie McDonald. According to the plans for the launch of this new CSHL meeting, the explicit goals of the meetings were to focus on identifying disease pathways and to facilitate the translation of breakthrough science into effective medicines by bringing together established and younger scientists as well as trainees from academia and pharmaceutical/biotech companies. Because of the massive and unchanged unmet medical need, the meeting goals have not changed. This year's meeting was entitled Neurodegenerative Diseases: Biology and Therapeutics.

The topics in the six platform sessions and two poster sessions of this meeting are listed below under Program. These topics reflect the newest and most relevant research.

Although most forms of NDDS occur in the absence of obvious heritability or identifiable genetic mutations, it has been possible over the past 25 years to discover uncommon genetic mutations as well as risk-modifying DNA changes in some NDDs and predictable causative changes in others. From these findings, a genetic architecture of several NDDs has emerged, and this topic

M. Rohe, N. Kashikar

P. Taylor, J. Herz

J. Johansson, P. Nilsson

J-V. Haure-Mirande, F. Bard

was addressed in Alison Goate's keynote lecture. Additionally, transgenic technology has rapidly led to the development of mouse, fruit fly, and nematode model systems that partially recapitulate phenomena reminiscent of the human diseases as well as some of the hallmark molecular and morphological pathology of the conditions. Although not predictive for therapeutic outcomes, these varied models provide insight into the biological effects of pathogenic processes such as amyloid deposition.

As in past years, this year's meeting featured early-stage data from both academic labs and commercial drug discovery organizations aimed at identifying novel druggable pathways, as well as biomarkers for diagnosis and identification of patient populations of interest. In all, the 202 participants presented 16 invited talks, 20 presentations selected from abstracts, and one keynote lecture, as well as 95 posters. Particular highlights included a presentation from Eliezer Masliah, Director of the National Institute on Aging; Katja Brose's introduction to the CZI Neurodegeneration Challenge Network, and Frank Bennett's account of successful treatment for spinal muscular atrophy (SMA) using intrathecal antisense oligodeoxynucleotides that correct a splicing defect.

For this meeting, discussion of new, unpublished data was emphasized and adequate time was left for discussion of each presentation. Discussion sessions were robust and vigorous, with notable levels of input from graduate students and postdocs, who constituted 30% of those in attendance. Posters were displayed for an extended period during the meeting, and poster viewing was especially encouraged during the cocktail hours before the evening meals.

T. Cohen, L. Gan

V. Baekelandt, F. Ibrahim

A. Pensalfini, H. Sadlish

N. Ramalingam, D. Klenerman

At the close of the meeting, grateful thanks were offered to John Trojanowski, who is stepping down after a distinctly above-and-beyond involvement as a meeting co-organizer. Aaron Gitler was invited to join the co-organizer team and accepted that role for the next meeting, scheduled for December 2–5, 2020.

The 2018 meeting was funded in part by the National Institute on Aging, a branch of the National Institutes of Health; Amgen; and Biogen.

PROGRAM

Network Mechanisms of Neurodegenerative Disease
Chairperson: L-H. Tsai, *Massachusetts Institute of Technology, Cambridge*

Neuroinflammation and Glial Biology of Neurodegeneration
Chairperson: R. Ransohoff, *Third Rock Ventures, Boston, Massachusetts*

Lunchtime Panel Discussion
Opportunities and Challenges for Clinical Data in Neurodegenerative Disease R&D
Chairpersons: F. Bennett, *Ionis Pharmaceuticals*; K. Dave, *Michael J. Fox Foundation*; T. Plavina, *Biogen*; M. Raupp, *IQVIA*

Keynote Speaker
A. Goate, *Icahn School of Medicine at Mount Sinai, New York, New York*

Propagation of Neurodegenerative Disease
Chairperson: V. Lee, *University of Pennsylvania, Philadelphia*

The Endolysosome, Autophagy and Other Targets in Therapeutics
Chairperson: S. Small, *Columbia University, New York, New York*

ApoE and Lipid Metabolism
Chairperson: J. Herz, *University of Texas Southwestern Medical School, Dallas*

RNA Metabolism in Neurodegenerative Disease
Chairperson: A. Gitler, *Stanford University, California*

POSTGRADUATE COURSES

Cryo-Electron Microscopy

March 1–14

INSTRUCTORS
J. Kollman, University of Washington, Seattle
G. Lander, The Scripps Research Institute, La Jolla, California
M. Ohi, University of Michigan, Ann Arbor
D. Veesler, University of Washington, Seattle

CO-INSTRUCTOR
M. Vos, Thermofisher Scientific, Eindhoven, Netherlands

ASSISTANTS
M. Campbell, University of California, San Francisco
A. Hernandez, The Scripps Research Institute, La Jolla, California
M. Johnson, University of Washington, Seattle

Cryo-electron microscopy (cryo-EM) is an emerging technique in structural biology in which the biological sample of interest is prepared under cryogenic conditions. The utility of cryo-EM stems from the fact that it allows the observation of specimens that have not been stained or fixed in any way, showing them in their "native" environment. This is in contrast to X-ray crystallography, which requires crystallizing the specimens, which can be difficult, and placing them in non-physiological environments, which can occasionally lead to functionally irrelevant conformational

changes. The resolution of cryo-EM maps is improving steadily, and in recent years, near-atomic resolution of some structures had been obtained, including those of viruses, ribosomes, mitochondrial protein complexes, ion channels, and enzyme complexes. The spiraling number of cryo-EM publications suggests this to be true: In 2016 alone, the technique had been used to map the structures of more than 800 molecules. Researchers can use cryo-EM to calculate the structure of a biological macromolecule or complex that has been flash frozen in several conformations and so deduce the mechanism by which it works.

The course covered the theory, practice, and application of cryo-EM including single-particle analysis, tomography, and electron diffraction. Participants in the course had supervised access to CSHL's FEI Titan Krios and Falcon direct electron detector. This hands-on course included lectures by leading experts who discussed practical and conceptual approaches to structure determination using these techniques, as well as covered a wide range of state-of-the-art applications of cryo-EM in the biological sciences.

This course was supported with funds provided by the Helmsley Charitable Trust and the Howard Hughes Medical Institute.

PARTICIPANTS

Abraham, J., M.D./Ph.D., Harvard Medical School, Boston, Massachusetts
Barnes, C., Ph.D., California Institute of Technology, Pasadena
Joshua-Tor, L., Ph.D., Cold Spring Harbor Laboratory
Kober, D., Ph.D., University of Texas Southwestern Medical Center, Dallas
Mansoor, S., M.D./Ph.D., Oregon Health & Science University, Portland
Montemayor, E. , Ph.D., University of Wisconsin, Madison
Schubert, H., Ph.D., University of Utah, Salt Lake City
Sullivan, M., B.S., Yale University, New Haven, Connecticut
Worden, E., Ph.D., Johns Hopkins University School of Medicine, Baltimore, Maryland

SEMINARS

Adams, P., Lawrence Berkeley Laboratory, Berkeley, California: New tools for cryo-EM model building and refinement.
Carragher, B., New York Structural Biology Center, New York: A brief (personal) history of 30+ years of cryo-EM and some musings about the future.
Cianfrocco, M., University of Michigan, Ann Arbor: Software tools to deploy and manage cryo-EM jobs in the cloud.
DeRosier, D., Brandeis University, Boston, Massachusetts: Future directions of cryo-EM/CLEM.
Glaeser, R., Lawrence Berkeley Laboratory, Berkeley, California: Dangers of air–water interface. Advances in phase plates and other instrumentation.
Grant, T., Janelia Research Campus, Ashburn, Virginia: Image processing. Introduction to cryo-EM validation.
Johnson, G., Allen Institute for Cell Science, Seattle, Washington: Whole-cell multiscale atlassing.
Kollman, J., University of Washington, Seattle: A brief crash course in EM processing.
Lander, L., The Scripps Research Institute, La Jolla, California: Course overview: Structure determination of biological macromolecules using cryo-EM. Introduction to sample prep for cryo-EM. Using UCSF chimera with cryo-EM data.
Lyumkis, D., The Salk Institute for Biological Sciences, La Jolla, California: Anisotropy and preferred specimen orientation in single-particle cryo-EM.
Ohi, M., University of Michigan, Ann Arbor: Why use electrons? Introduction to cryo-EM workflow.
Russo, C., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom: Electron specimen interaction physics and radiation damage in biology. Basic principles of detecting high-energy electrons efficiently.
Veesler, D., University of Washington, Seattle: Introduction to atomic modeling in Coot. Model building with Rosetta.
Vos, M., Thermofisher Scientific, Eindhoven, Netherlands: Overview of the components in an electron microscopy facility. Electron microscope optics. Detectors: Applications and operation. Introduction to automated data acquisition with EPU.

Bioinformatics for Cancer Genomics

March 12–18

INSTRUCTORS **A. Meyer,** Ontario Institute for Cancer Research, Toronto, Canada
F. Ouellette, Genome Quebec, Montreal, Canada

TEACHING ASSISTANTS **H. Farooq,** Hospital for Sick Children, Toronto, Ontario, Canada
M. Johnston, University of Calgary, Alberta, Canada

With the introduction of next-generation sequencing platforms, it is now feasible to use high-throughput approaches to address many research questions. This is especially true in cancer research, which has rapidly embraced high-throughput technologies. Dedicated compute clouds, such as the Cancer Genome Collaboratory, also facilitate complex analyses on large cancer data sets from projects such as the International Cancer Genome Consortium (ICGC) and the Pan-Cancer Analysis of Whole Genomes (PCAWG). Now more than ever, it is crucial to know what bioinformatic tools and resources are available in cancer research, and it is necessary to develop informatic skills to analyze high-throughput data using those tools. The Canadian Bioinformatics Workshops (CBW), in collaboration with Cold Spring Harbor Laboratory, has developed a comprehensive 6-day course covering key bioinformatics concepts and tools required to analyze cancer genomic data sets. This course combined the material and concepts from three established CBW workshops. In this course, participants gained practical experience and skills to visualize genomic data; analyze cancer data sets for gene expression, genome rearrangement, somatic mutation, and copy-number variation; analyze and conduct pathway analysis on candidate gene lists;

integrate -omic and clinical data; launch, configure, customize, and scale virtual machines (VMs); navigate cloud repositories and work with data sets in them; and follow best practices in data and workflow management.

This course was supported with funds provided by the Helmsley Charitable Trust and the Howard Hughes Medical Institute and computational resources supported by Amazon Web Services.

PARTICIPANTS

Bais, P., Ph.D., the Jackson Lab, Farmington, Connecticut
Bhatia, S., Ph.D., Cold Spring Harbor Laboratory
Cheng, N., B.Sc., Ontario Institute for Cancer Research, Toronto, Canada
Cyrill, S., Ph.D., Cold Spring Harbor Laboratory
Demas, V., Ph.D., GRAIL, Menlo Park, California
Dunn, H., Ph.D., Clemson University, South Carolina
Eyre Sanchez, E., Ph.D., Karolinska Institutet, Stockholm, Sweden
Hardiman, T., B.Sc., King's College London, United Kingdom
Hoefges, A., M.S., University of Wisconsin, Madison
Jensen, S., B.S./M.S., Aarhus University, Denmark
Lazaris, C., M.S./Ph.D., University of Southern California, Los Angeles
Le, W., B.S., Tufts University, Medford, Massachusetts
Lihm, J., Ph.D., Cold Spring Harbor Laboratory
Lipsky, A., M.D., Weill Cornell Medical College, New York
Moore, A., B.Sc., Queen's University, Kingston, Ontario, Canada
Moss, M., Ph.D., Cold Spring Harbor Laboratory
Parisian, A., B.S., University of California, San Diego, La Jolla
Park, S., B.S., Icahn School of Medicine at Mount Sinai, New York, New York
Polyanskaya, S., M.S., Cold Spring Harbor Laboratory
Schwarzer, A., M.D./Ph.D., Hannover Medical School, Germany
Sun, L., Ph.D., New England Biolabs, Ipswich, Massachusetts
Trieu, T., Ph.D., Weill Cornell Medicine School, New York, New York
Tu, T., B.Sc., Icahn School of Medicine at Mount Sinai, New York, New York
Utama, R., Ph.D., Cold Spring Harbor Laboratory
Venkat, S., Ph.D., Roswell Park Cancer Institute, Buffalo, New York
Yoshimoto, J., B.Sc., Colorado State University, Fort Collins

SEMINARS

Erdman, L., Hospital for Sick Children, University of Toronto, Ontario, Canada: Module 14b: Clinical data integration.
Farooq, H., Hospital for Sick Children, Toronto, Ontario, Canada: Module 4: Genome alignment. Module 5: Genome assembly.
Haas, B., Broad Institute, Cambridge, Massachusetts: Module 9: Gene fusions and rearrangements.
Hammell, M., Cold Spring Harbor Laboratory: Guest lecture.
Haw, R., Ontario Institute for Cancer Research, Toronto, Canada: Module 14a: Variants to networks.
Meyer, A., Ontario Institute for Cancer Research, Toronto, Canada: Module 3b: Visualization tools.
Mihaiescu, G., Ontario Institute for Cancer Research, Toronto, Canada: Module 10: Sharing and scaling a VM.
Morrissy, S., University of Calgary, Canada: Module 6: Somatic copy-number changes. Module 7: Somatic mutations and annotations.
O'Connor, B., University of California, Santa Cruz: Module 11: Working reproducibly in the cloud.
Ouellette, F., Genome Quebec, Montreal, Canada: Module 3a: Cancer databases.
Phillips, M., McGill University, Montreal, Canada: Module 2: Ethics of data usage and security.
Pugh, T., University Health Network, Toronto, Canada: Module 1: Introduction to cancer genomics.
Reimand, J., Ontario Institute for Cancer Research, Toronto, Canada: Module 13: Genes to pathways.
Yousif, F., Ontario Institute for Cancer Research, Toronto, Canada: Module 8: Gene expression.
Yung, C., University of Chicago, Illinois: Module 12: Big data analytics in the cloud.

Workshop on Leadership in Bioscience

March 23–26

INSTRUCTORS **C. Cohen,** Science Management Associates, Newton, Massachusetts
S. Cohen, Science Management Associates, Newton, Massachusetts

This highly interactive 3.5-day workshop allowed students to develop the skills necessary to lead and interact effectively with others, in both one-on-one and group settings. The workshop focused on techniques, situations, and challenges that relate specifically to leading and managing in the scientific workplace. Participants gained a solid experience-based foundation in managing others, negotiating win/win outcomes, running effective meetings, selecting the best team members, and setting goals with mentees, direct reports, and teams. It emphasized learning by doing and involved role-playing, giving and receiving feedback, and group problem-solving. Much of the learning is peer-to-peer.

The workshop helped participants identify areas in which they need guidance and growth, as well as how to capitalize on areas of strength. Participants had the opportunity to share their experiences and challenges with others and to receive feedback and guidance from others with experience in leading scientists in a variety of settings. At the end of the course, participants linked through a unique online community in which they could continue learning from one another and from the course instructors.

Key focus areas of the workshop included recognizing and understanding leadership in a science setting; using negotiation as a tool in scientific discussions and problem solving; identifying and resolving conflicts in the lab; dealing with difficult people and situations; communicating

ideas and plans in a way that engages others; leading effective and productive meetings; becoming effective citizen scientists; and hiring and retaining a team.

The workshop was targeted to life scientists making, or recently having made, the transition to leadership or managerial positions. Many of the situations discussed were from the perspective of independent investigators running their own laboratories. As such, relatively new investigators (e.g., <3 years) were particularly encouraged to apply, as were senior postdoctoral scholars on the cusp of tenure-track research positions.

This course was supported with funds provided by the National Institute of General Medical Sciences.

PARTICIPANTS

Acar, H., Ph.D., University of Oklahoma, Norman
Babb, J., Ph.D., Boston Children's Hospital/Harvard Medical School, Boston, Massachusetts
Creixell, P., Ph.D., Massachusetts Institute of Technology, Cambridge
Elfenbein, J., Ph.D., North Carolina State University, Raleigh
Faltas, B., M.D., Weill Cornell Medicine, New York, New York
Han, A., M.D./Ph.D., Columbia University, New York, New York
Huang, C., Ph.D., New York University, New York
Huderson, B., Ph.D., University of the District of Columbia, Washington, D.C.
Hughey, J., Ph.D., Vanderbilt University, Nashville, Tennessee
Jadavji, N., Ph.D., Carleton University, Ottawa, Canada
Liu, C., Ph.D., University of Toronto, Ontario, Canada
Martin, J., Ph.D., Massachusetts Institute of Technology, Cambridge
Melas, P., Ph.D., Columbia University, New York, New York
Misra, J., Ph.D., Rutgers University, Piscataway, New Jersey
Mundt, F., Ph.D., Broad Institute, Cambridge, Massachusetts
Renwick, N., Ph.D., Queen's University, Kingston, Ontario, Canada
Rhodes, A., Ph.D., University of Vermont, Burlington
Schulz, A., M.D., Yale University School of Medicine, New Haven, Connecticut
Segarra, V., Ph.D., High Point University, North Carolina
Senatore, A., Ph.D., University of Toronto Mississauga, Ontario, Canada
Siebenthall, K., Ph.D., Altius Institute for Biomedical Sciences, Seattle, Washington
Thyme, S., Ph.D., Harvard University, Cambridge, Massachusetts
Titov, D., Ph.D., University of California, Berkeley
Wates, R., Ph.D., University of Kansas Medical Center, Kansas City
Weinhardt, V., Ph.D., University of California San Francisco

SEMINARS

Cohen, C., Science Management Associates, Newton, Massachusetts, and Cohen, S., Science Management Associates, Newton, Massachusetts: Session 1: Who we are. Session 2: Leadership challenges: Case study overview. Session 3: Difficult conversations and interactions: Fundamentals of negotiation. Session 4: Running productive scientific teams and project meetings. Session 5: Case study analysis: Part 1. Session 6: Hiring and retaining your science team: Interviewing, selecting, and orienting. Session 7: Projecting leadership. Session 8: Case study analysis: Part 2. Session 9: Managing your science team: Goal setting, feedback, motivation, and culture. Session 10: Concluding group discussion.

Cell and Developmental Biology of *Xenopus*

April 4–17

INSTRUCTORS **M. Khokha,** Yale University, New Haven, Connecticut
K. Liu, King's College London, United Kingdom

ASSISTANTS **P. Date,** Yale University, New Haven, Connecticut
R. Huebner, The University of Texas, Austin
M. Lane, Yale University, New Haven, Connecticut
E. Mis, Yale University, New Haven, Connecticut
H. Rankin Willsey, University of California, San Francisco
Z. Swider, University of Wisconsin, Madison

In vivo animal models are an important tool for the understanding of human development and disease. Studies using the frog *Xenopus* have made remarkable contributions to our understanding of fundamental processes such as cell cycle regulation, transcription, translation, and many other topics. *Xenopus* is remarkable for studying development and disease, including birth defects, cancer, and stem cell biology. Because *Xenopus* are easy to raise, producing many thousands of eggs per day, these frogs have emerged as a premiere model for understanding of human biology from the fundamental building blocks to the whole organism.

The recent development of CRISPR-Cas9 technology has made it easy to target genes of interest using *Xenopus.* This course was designed with that in mind. The goal of the course was for each student to design a set of experiments focusing on their gene or biological interest. Prior to starting the course, students were expected to choose gene(s) of interest, and the instructors generated sgRNAs targeting these genes. These were the students' own genes, or those chosen from a bank provided by the instructors.

During the course, the students analyzed any phenotypes generated from CRISPR-Cas9-based gene depletion and learned the diverse array of techniques available in *Xenopus*. In previous courses, students were guided in the ablation of a wide variety of genes and design-suitable assays for their biological interests. Most recently, students have targeted autism genes, thyroid genes, and immune modulators, several of which have already led to publications. Approaches covered included microinjection and molecular manipulations such as CRISPR-Cas9 knockouts, antisense morpholino-based depletions, transgenics, and mRNA overexpression. In addition, students combined these techniques with explant and transplant methods to simplify or test tissue-level interactions. Additional methods included mRNA in situ hybridization and protein immunohistochemistry as well as basic bioinformatic techniques for gene comparison and functional analysis. Biochemical approaches such as proteomics and mass spectrometry and biomechanical concepts were also discussed. Finally, to visualize subcellular and intercellular activities, a variety of imaging methods were introduced including time-lapse, fluorescent imaging, optical coherence tomography, and confocal microscopy. These were facilitated by state-of-the-art equipment from Nikon, Leica, Thorlabs, and Bruker.

This course was supported with funds provided by the National Institute of Child Health and Human Development.

PARTICIPANTS

Ahsan, A., B.S., Georgetown University, Washington, D.C.
Aztekin, C., M.Sc., University of Cambridge, United Kingdom
Baxi, A., M.S., George Washington University, Washington, D.C.
Blackburn, A., B.S., MD Anderson University of Texas Health Graduate School, Houston
Flowers, E., B.A., University of Maryland, Baltimore
Grand, K., M.Sc., University Medical Center Freiburg, Germany
Ivanova, A., Ph.D., Institute of Bioorganic Chemistry (IBCH RAS), Moscow, Russia
Landino, J., Ph.D., University of Michigan, Ann Arbor
Nguyen, T., B.S., Princeton University, New Jersey
Novikova, P., Ph.D., VIB, Ghent, Belgium
Shim, S., Ph.D., University of California, Davis
Tasca, A., M.S., University Clinic Freiburg, Germany
Ventura, G., M.S., University of Copenhagen, Denmark
Zhou, C., Ph.D., University of California, Berkeley
Zhou, J., B.S., University of California, Irvine

SEMINARS

Bement, B., University of Wisconsin, Madison: The cortical cytoskeleton: Imaging signaling and cell division in frogs for fun and nonprofit.
Blitz, I., University of California, Irvine: Intro to germ cell transplantation.
Cha, S-W., Cincinnati Children's Hospital Medical Center, Ohio: How to make a long gut tube.
Chang, C., University of Alabama, Birmingham: Embryonic induction and signaling: A walk of a century following Spemann.
Conlon, F., University of North Carolina, Chapel Hill: *Xenopus* proteomic approaches to development and disease.
Davidson, L., University of Pittsburgh, Pennsylvania: Leveraging *Xenopus* to explore the mechanics and mechanobiology of development.
Heald, R., University of California, Berkeley: Mechanisms of mitosis and size control in *Xenopus.*
Keller, R., University of Virginia, Charlottesville: Systems aspects of early *Xenopus* morphogenesis.
Khokha, M., Yale University, New Haven, Connecticut: Frogs can be shocking: Membrane voltage in embryonic development.
Mayor, R., University College London, United Kingdom: Neural crest.
Nemes, P., George Washington University, Washington, D.C.: Mass spec that single cell in the frog embryo.
Thomsen, G., Stony Brook University, New York: *Xenopus*: Show and tell demo.
Wallingford, J., University of Texas, Austin: *Xenopus:* In vivo imaging of cell behavior and protein dynamics.
Wills, A., University of Washington, Seattle: Remodeling tissue and gene expression during *Xenopus* regeneration.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.

Expression, Purification, and Analysis of Proteins and Protein Complexes

April 4–17

INSTRUCTORS

A. Courey, University of California, Los Angeles
S-H. Lin, MD Anderson Cancer Center/University of Texas, Houston
M. Marr, Brandeis University, Waltham, Massachusetts
S. Nechaev, University of North Dakota School of Medicine, Grand Forks

ASSISTANTS

N. Clark, Brandeis University, Waltham, Massachusetts
S. Gartland, Brandeis University, Waltham, Massachusetts
S. Ghosh Dastidar, University of North Dakota School of Medicine, Grand Forks
Y-C. Lee, MD Anderson Cancer Center, Houston, Texas
O. Molina, University of California, Los Angeles
M. Morales, University of California, Los Angeles
J. Rigal, Brandeis University, Waltham, Massachusetts
S. Rudraraju, University of North Dakota School of Medicine, Grand Forks
A. Sawyer, Brandeis University, Waltham, Massachusetts
T. Yau, University of California, Los Angeles

This course was for scientists, including graduate students, postdoctoral scholars, staff scientists, and principal investigators, who wanted a rigorous introduction to expression and purification of proteins as well as analysis of protein structure and function.

Through hands-on experience in the lab as well as extensive lecture and discussion, each student became familiar with key approaches in expression, purification, and analysis of soluble and membrane proteins and protein complexes from both natural sources and overexpression systems. The emphasis of the course was on the following:

1. Approaches in protein expression: Choosing the best bacterial or eukaryotic expression system tailored for the particular protein and experimental problem; determining how to optimize expression; understanding protein tagging—the advantages and pitfalls of various affinity and solubility tags.
2. Approaches in protein purification: Choosing the best strategy for a given protein including solubilization, bulk fractionation, and liquid chromatography and including conventional methods (ion exchange, size exclusion, and reverse phase) and affinity methods (e.g., MAC, DNA affinity, and immunoaffinity), as well as FPLC/HPLC.
3. Approaches in protein analysis: Introduction to common approaches for characterization of proteins, including binding assays, activity assays, and mass spectroscopy to identify protein interaction partners and posttranslational modifications.

In addition to purification, students also gained exposure to fundamental analytical approaches such as mass spectroscopy and protein structure determination (e.g., X-ray crystallography and cryo-EM).

Funding for this course was provided by NCI, Helmsley, and HHMI. We acknowledge the following companies that provided invaluable support: lab equipment from Bio-Rad Laboratories, GE Healthcare Corp., and Waters.

PARTICIPANTS

Arhar, T., B.S., University of California, San Francisco
Cabello Lobato, M.J., Ph.D., University of Manchester, United Kingdom
Carrasquillo, E., B.S., University of Puerto Rico, Rio Piedras, San Juan
Choi, C.H., B.S., The Rockefeller University, New York
Krska, D., M.Sc., Chalmers University of Technology, Göteborg, Sweden
Lyu, D., Ph.D., University of Arkansas for Medical Sciences, Little Rock
Makowski, M., B.S., Radboud University, Nijmegen, Netherlands
Nasrin, F., Ph.D., Cold Spring Harbor Laboratory
O'Donovan, K., B.S., St. Jude Children's Research Hospital, Memphis, Tennessee
Reyes, A., Ph.D., University of Buffalo, New York
Santos, V., M.S, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
Schnieder, T., Ph.D., Columbia University Medical Center, New York, New York
Soffers, J., M.A., Stowers Institute, Kansas City, Missouri
Spassibojko, O., B.S., Cold Spring Harbor Laboratory
Taeuber, S., M.S., Roche Diagnostics, Penzberg, Germany

SEMINARS

Courey, A., University of California, Los Angeles: System-wide analyses of Groucho and SUMO in *Drosophila*.
Jarvis, D., University of Wyoming, Laramie: The baculovirus expression system.
Lambert, C., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.
Lin, S-H., MD Anderson Cancer Center/University of Texas, Houston: Secreted factors in the tumor microenvironment confer therapy resistance.
Love, J., Expression Technologies, Newark, California: High-throughput eukaryotic protein production.
Marr, M., Brandeis University, Waltham, Massachusetts: Introduction to protein purification. Controlling gene expression in response to stress.
Nechaev, S., University of North Dakota School of Medicine, Grand Forks: Stable Pol II pausing is retained during gene activation to provide a platform for regulation.
Pappin, P., Cold Spring Harbor Laboratory: Introduction to mass spectrometry of proteins. Quantitative approaches to mass spectrometry of proteins.

Quantitative Imaging: From Acquisition to Analysis

April 4–17

INSTRUCTORS	**H. Elliott,** Harvard Medical School, Boston, Massachusetts **S. Manley,** Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland **J. Waters,** Harvard Medical School, Boston, Massachusetts
CO-INSTRUCTORS	**T. Lambert,** Harvard Medical School, Boston, Massachusetts
TEACHING ASSISTANTS	**M. Cicconet,** Harvard Medical School, Boston, Massachusetts **J. Hornick,** Northwestern University, Evanston, Illinois **A. Jost,** Harvard Medical School, Boston, Massachusetts **M. Weber,** Harvard Medical School, Boston, Massachusetts

Combining careful image acquisition with rigorous computational analysis allows extraction of quantitative data from light microscopy images that is far more informative and reproducible than what can be seen by eye. This course focused on advanced quantitative fluorescence microscopy techniques used for imaging a range of biological specimens, from tissues to cells to single molecules. The course was designed for quantitative cell and molecular biologists, biophysicists, and bioengineers.

The course provided a thorough treatment of the complete process of quantitative imaging, from the photons emitted from the sample to the extraction of biologically meaningful measurements from digital images. Material was covered in lectures, discussion groups, and hands-on quantitative exercises using commercial microscopes and open-source image analysis tools.

The concepts covered in this course included wide-field fluorescence microscopy; laser-scanning and -spinning disk confocal microscopy; CCD, EM-CCD, and sCMOS cameras; total internal fluorescence microscopy (TIRF); light-sheet microscopy; super-resolution microscopy (structured illumination, STED, and localization microscopy); imaging and analyzing ratiometric "biosensors" (including FRET); fluorescent proteins and live-sample imaging; image processing (filtering, denoising, corrections, and deconvolution; image segmentation; quantitative shape and intensity measurements; object detection and tracking; machine learning; and designing and troubleshooting quantitative imaging experiments. The course also included a series of seminars from guest speakers who apply the methods we discussed.

This course was supported with funds provided by the Helmsley Charitable Trust and Howard Hughes Medical Institute and major support by the National Cancer Institute.

PARTICIPANTS

Chan, E., M.Eng., Memorial Sloan Kettering Cancer Center, New York, New York
Chung, J., Ph.D., Harvard T.H. Chan School of Public Health, Boston, Massachusetts
Corbat, A., B.S., University of Buenos Aires, Conicet, Argentina
Eckenrode, K., B.S., City University of New York, Graduate Center, New York
Geldert, A., B.S., University of California, Berkeley
Lam, M., B.S., Baylor College of Medicine, Houston, Texas
Lamkin, E., B.S., Harvard University, Brookline, Massachusetts
Marquez, J., B.S., Yale University, New Haven, Connecticut
McBride, S., M.Sc., University of Ottawa, Ontario, Canada
Prunet, N., M.S./Ph.D., HHMI/California Institute of Technology Pasadena
Rao, T., Ph.D., University of Alabama, Birmingham
Rieder, L., Ph.D., Brown University, Providence, Rhode Island
Spindler, M-C., M.S., University of Würzburg, Biocenter, Germany
Taylor, J., Ph.D., Johns Hopkins University, Baltimore, Maryland
Young, A., Ph.D., Fred Hutchinson Cancer Research Center, Seattle, Washington
Zamponi, N., Ph.D., Weill Cornell Medicine, New York, New York

SEMINARS

Dernburg, A., HHMI/University of California, Berkeley: In vivo imaging reveals a novel mechanism that regulates meiotic recombination.
Dolman, N., Thermo Fisher: Fluorescent probes.
Elliott, H., Harvard Medical School, Boston, Massachusetts: 3D image analysis and deconvolution.
Elliott, H., Harvard Medical School, Boston, Massachusetts: Basics of image processing and digital microscopy: Resolution, SNR and diffraction-limited objects.
Elliott, H., Harvard Medical School, Boston, Massachusetts: Image processing 2: Image corrections and advanced filtering. Image segmentation and morphometry.
Elliott, H., Harvard Medical School, Boston, Massachusetts: Image correlation methods: Co-localization, registration, and stitching.
Elliott, H., Harvard Medical School, Boston, Massachusetts: Image time series analysis: Tracking, photo-bleach correction, and FRAP analysis. Machine learning in bioimage analysis.
Jug, F., Max Planck Institute CBG, Dresden, Germany: Content-aware image restoration and quantitative downstream analysis.
Keller, P., Howard Hughes Medical Institute, Ashburn, Virginia: Dissecting embryonic development by high-resolution whole-animal imaging.
Lambert, C., Cold Spring Harbor Laboratory: Ethics lecture.
Lambert, T., Harvard Medical School, Boston, Massachusetts: Confocal microscopy theory and hardware.
Lambert, T., Harvard Medical School, Boston, Massachusetts: Digital imaging.
Lambert, T., Harvard Medical School, Boston, Massachusetts: Light sheet microscopy: Part 2.
Lambert, T., Harvard Medical School, Boston, Massachusetts: Super-resolution microscopy II: Patterned illumination.
Manley, S., Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland: Super-resolution microscopy I: Localization.
Manley, S., Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland: Expanding horizons with large field-of-view, automated localization microscopy.
Shaner, N., The Scintillon Institute, San Diego, California: Fluorescent proteins.
Waters, J., Harvard Medical School, Boston, Massachusetts: Light sheet microscopy: Part 1.

Waters, J., Harvard Medical School, Boston, Massachusetts: Limitations on quantitative imaging of thick samples.
Waters, J., Harvard Medical School, Boston, Massachusetts: Live confocal microscopy and intensity measurements over time.
Waters, J., Harvard Medical School, Boston, Massachusetts: Live-cell imaging.
Waters, J., Harvard Medical School, Boston, Massachusetts: Multiphoton microscopy.
Waters, J., Harvard Medical School, Boston, Massachusetts: Quantitative microscopy basics. Objective lenses. Transmitted light microscopy. Fluorescence microscopy.
Waters, J., Harvard Medical School, Boston, Massachusetts: Quantifying fluorescence: image acquisition and controls.
Waters, J., Harvard Medical School, Boston, Massachusetts: Total internal reflection fluorescence (TIRF) microscopy.

Advanced Bacterial Genetics

June 5– 25

INSTRUCTORS **L. Bossi,** Institute of Integrative Biology of the Cell (I2BC), Paris, France
A. Camilli, Tufts University Medical School, Boston, Massachusetts
A. Grundling, Imperial College London, United Kingdom

ASSISTANTS **R. Balbontin Soria,** Instituto Gulbenkian de Ciência, Oeiras, Portugal
N. Figueroa-Bossi, Institute of Integrative Biology of the Cell (I2BC), Paris, France
M. Ramliden, Tufts University Medical School, Boston, Massachusetts
M. Zeden, Imperial College London, United Kingdom

This course presented logic and methods used in the genetic dissection of complex biological processes in diverse bacteria. Laboratory methods included classical and cutting-edge mutagenesis using transposons, allelic exchange, and TargeTron; recombineering with single- and double-stranded DNA; CRISPR-Cas genome editing; genome sequencing and assembly; mapping mutations using genetic and physical techniques; modern approaches to the generation and analysis of targeted gene disruptions and reporter gene fusions; and fluorescence microscopy. Key components of the course were the use of sophisticated genetic methods in the analysis of model bacteria (including *Escherichia coli, Salmonella, Staphylococcus aureus*, and *Vibrio cholerae*) and the use of the wealth of new genomic sequence information to motivate these methods. Invited lecturers presented various genetic approaches to study bacterial mechanisms of metabolism, development, and pathogenesis.

This course was supported with funds provided by the National Science Foundation.

PARTICIPANTS

Adomako, M., B.S., University of California, San Diego, La Jolla
Boeck, D., M.S., ETH Zürich, Switzerland
Cengher, L., B.S., Dartmouth College, Hanover, New Hampshire
Creasy, A., B.A., University of Florida, Gainesville
Han, N-C., B.S., Ohio State University, Columbus
Harper, C., B.S., Cornell University, Ithaca, New York
Lejars, M., M.S., The National Center for Scientific Research, UMR 8261/The Physical and Chemical Biology Institute, Paris, France
Merida-Floriano, A., M.S., University of Seville, Spain
Mortier, J., M.A., Katholieke Universiteit Leuven, Belgium
Patnode, M., Ph.D., Washington University in St. Louis, Missouri
Pike, C., B.S., University of Delaware, Newark
Pishchany, G., Ph.D., Harvard Medical School, Boston, Massachusetts
Povolo, V., M.S., Eawag ETH Zürich, Duebendorf, Switzerland
Suarez Cham, S., B.S., University of California, Irvine
Webster, S., B.A., Geisel School of Medicine at Dartmouth College, Hanover, New Hampshire
Yang, D., Ph.D., Georgia Institute of Technology, Atlanta

SEMINARS

Bernhardt, T., Harvard Medical School, Boston, Massachusetts: Genetic analysis of the cell wall assembly pathway.
Harshey, R., University of Texas, Austin: A dynamic *E. coli* genome uncovered by monitoring Mu transposition.
Kuzminov, A., University of Illinois, Urbana-Champaign: Thymine starvation in *E. coli*: The mystery of the resistance phase reveals a hidden treasure chest.
Moineau, S., Université Laval, Quebec, Canada: Phages and CRISPR-Cas systems: The ongoing battle.
Shen, A., Tufts University School of Medicine, Cambridge, Massachusetts: To spores and beyond: Explorations of *Clostridium difficile* spore formation and germination.
Tang, C., University of Oxford, United Kingdom: Going around in circles: What does plasmid maintenance tell us about *Shigella?*
Van Opijnen, T., Boston College, Chestnut Hill, Massachusetts: Tracking phenotypic, transcriptional, and metabolic responses on topological networks predicts survival success of pathogenic bacteria.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.

Ion Channels in Synaptic and Neural Circuit Physiology

June 5–25

INSTRUCTORS **T. Branco,** University College London, United Kingdom
C. Schmidt-Hieber, Pasteur Institute, Paris, France
A. Scimemi, University at Albany, New York
N. Wanaverbecq, Aix Marseille University, France

ASSISTANTS **K. Betsios,** Mantis 64, Athens, Greece
L. Brosse, Aix Marseille University, France
J. McCauley, University at Albany, New York
R.G. Ocadiz, Pasteur Institute, Paris, France
M. Petroccione, University at Albany, New York

Ion channels are the fundamental building blocks of excitability in the nervous system. The primary goal of this course was to demonstrate, through lectures and laboratory work, the different biophysical properties of ion channels that enable neurons to perform unique physiological functions in a variety of neural systems.

Areas of particular interest included (1) voltage- and ligand-gated ion channels at central and peripheral synapses, (2) synaptic integration and plasticity, (3) neural circuit function in vitro and in vivo, and (4) optogenetic strategies for circuit manipulation. A typical day consisted of morning

lectures followed by hands-on laboratory practical sessions in the afternoon and evening with guest lecturers available to give one-on-one practical advice.

The laboratory component of the course introduced students to state-of-the-art electrophysiological approaches for the study of ion channels in their native environments. The course provided students with hands-on experience in using patch-clamp electrophysiology to examine single-channel activity in cultured cells, ion channel biophysics in acutely dissociated neurons and synaptic integration, plasticity, and circuit dynamics in in vitro slice and in vivo preparations. Different recording configurations were used (e.g., cell-attached, whole-cell dendritic and somatic patch and voltage- and current-clamp configurations) and the advantages and limitations of each method were discussed in relation to specific scientific questions. The course also provided practical experience in cellular and circuit manipulation techniques (i.e., pharmacological, electrophysiological, and optogenetic) both in vitro and in vivo.

This course was supported with funds provided by the Helmsley Charitable Trust and Howard Hughes Medical Institute.

PARTICIPANTS

Beas Alvarez, B., Ph.D., National Institute of Mental Health, Bethesda, Maryland
Courtland, J., B.S., Duke University School of Medicine, Durham, North Carolina
Flores, J., B.S., University of California, Davis
Goff, K., B.S., University of Pennsylvania Perelman School of Medicine/The Children's Hospital of Philadelphia
Leitch, D., Ph.D., University of California, San Francisco
Tan, C., B.S., Duke University, Durham, North Carolina
Tasnim, A., B.A., Harvard Medical School, Boston, Massachusetts
Torwig, J., B.S., University of California, Berkeley
Warlow, S., Ph.D., University of Michigan, Ann Arbor
Wee, R., B.S., University College London, United Kingdom
Yalcinbas, B.S., University of California, San Diego
Yu, W., B.S., University of North Carolina, Chapel Hill

SEMINARS

Cohen, J., The Johns Hopkins University, Baltimore, Maryland: In vivo/in vitro extracellular recordings.
Diamond, J., National Institutes of Health, Bethesda, Maryland: Synaptic integration.
Dudek, S., National Institutes of Health, Research Triangle Park, North Carolina: Synaptic plasticity.
Duguid, I., University of Edinburgh, United Kingdom: In vivo intracellular recordings.
Gasparini, S., Louisiana State University, New Orleans: Ih and HCN channels.
Hull, C., Duke University, Durham, North Carolina: Synaptic inhibition.
Kammermier, P., University of Rochester, New York: Calcium channels.
Kozorovitskiy, Y., Northwestern University, Evanston, Illinois: Circuit interrogation/optogenetics.
Kullmann, D., University College London, United Kingdom: Plenary lecture: Ion channel biophysics and brain function.
Lampert, A., RWTH Aachen University, Germany: Na^+ channels.
Larsson, P., University of Miami, Florida: Electrical properties of excitable cells.
Nagel, G., Würzburg University, Germany: Channelrhodopsin biophysics/optogenetics.
Nimigean, C., Weill Cornell Medical College, New York, New York: K^+ channels.
Plested, A., FMP Berlin, Germany: Single-channel recording.
Spruston, N., Janelia Research Campus, Ashburn, Virginia: Fundamental concepts in neurophysiology.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.
Xu-Friedman, M., University of Buffalo, New York: Synaptic excitation.

Workshop on Schizophrenia and Related Disorders

June 6–13

INSTRUCTORS
A. Abi-Dargham, Stony Brook University, New York
J. Hall, Cardiff University School of Medicine, Wales, United Kingdom
A. Sawa, Johns Hopkins University, Baltimore, Maryland

ASSISTANTS
C. Aguirre-Chen, Cold Spring Harbor Laboratory
N. Clifton, Cardiff University, United Kingdom
K. Delevich, University of California, Berkeley

This workshop provided students with the most current understanding of the molecular, cellular. and neural systems underlying the disturbances in brain function in these devastating illnesses. During the 7-day workshop, students learned about the clinical aspects of schizophrenia, schizoaffective disorder, and bipolar disorder as well as explored in detail the genetic and neurobiological underpinnings of these complex psychiatric disorders. The workshop included sessions focused on the clinical syndrome, basic neurobiology, cognitive neuroscience, neuroimaging, genetics and genomics, endophenotypes, and gene expression and gene modulation. In addition to hearing about the most recent research in these areas, controversial topics and challenges to basic assumptions in the field were explored and discussed. A diverse faculty brought the most up-to-date results and theories to the students, making this workshop a valuable resource for young researchers starting out in this fast-moving and expansive field. Not only did it help them build the foundation for their future research, it also introduced them to many potential collaborators working to understand schizophrenia and similar disorders from different perspectives.

This course was supported with funds provided by Cold Spring Harbor Laboratory and the Stanley Foundation.

PARTICIPANTS

Ashok, A., M.B./B.S., King's College London, United Kingdom
de Filippis, R., M.D., University "Magna Graecia" of Catanzaro, Italy
Guillory, S., Ph.D., Icahn School of Medicine at Mount Sinai, New York
Gunasekaran, S., M.S., Rajiv Gandhi Centre for Biotechnology, Trivandrum, India
Jerotic, S., M.D., Faculty of Medicine, University of Belgrade, Serbia
Jonas, K., Ph.D., Stony Brook University Health Sciences Center, New York
Kepinska, A., B.Sc., Institute of Psychiatry, King's College London, United Kingdom
Kietzman, H., B.A., Emory University, Atlanta, Georgia
Lee, B., B.Sc., Johns Hopkins University School of Medicine, Baltimore, Maryland
Leskauskaite, I., M.S., Haukeland University Hospital, Bergen, Norway
Parker, J., Ph.D., Stanford University, Palo Alto, California
Reddaway, J., B.Sc., Cardiff University, United Kingdom
Rivero-Echeto, M.C., M.S., IFIBYNE, Ciudad Autonoma de Buenos Aires, Argentina
Rodriguez, V., M.Sc., King's College London, United Kingdom
Rutkowski, T., Ph.D., Emory University, Atlanta, Georgia
Schalbetter, S-M., M.S., University of Zürich–Vetsuisse, Switzerland
Schoonover, K., B.A., University of Alabama, Birmingham
Tripoli, G., M.Sc., King's College London, United Kingdom
Williams, J., B.E./ M.S., Stony Brook University School of Medicine, New York
Wright, A., Ph.D., University of Sussex, Brighton, United Kingdom
Yang, S., M.S., China Medical University, Taichung City, Taiwan
Zhu, X., B.S., University of Pittsburgh, Pennsylvania

SEMINARS

Abi-Dargham, A., Stony Brook University, New York: Pharmacology treatments and open discussion. PET imaging in schizophrenia.
Anderson, S., Children's Hospital of Philadelphia, Pennsylvania: Molecular studies of the GABA system.
Arnsten, A., Yale University, New Haven, Connecticut: Prefrontal cortex and cognition: Molecular vulnerabilities.
Barch, D., Washington University, St. Louis, Missouri: Measuring cognitive performance in schizophrenia.
Brandon, N., AstraZeneca, Waltham, Massachusetts: Developing medicines for psychiatric disorders through the lens of AstraZeneca.
Brennand, K., Mount Sinai School of Medicine, New York, New York: Modelling predisposition to schizophrenia using stem cells.
Correll, C., The Zucker Hillside Hospital, Glen Oaks, New York: Clinical trials in schizophrenia and The Prodrome.
Escola, G.S., Columbia University Medical Center, New York: Applying machine learning algorithms to the study of schizophrenia: Methods and basic concepts, applications.
Fletcher, P., University of Cambridge, United Kingdom: Neuropsychology.
Hall, J., Cardiff University, Wales, United Kingdom: Schizophrenia: A fearful madness.
Heimer, H., Schizophrenia Research Forum, Providence, Rhode Island: Schizophrenia Research Forum.
Horga, Columbia University Medical Center, New York: MRI imaging in schizophrenia: Brief review of methods and findings.
Jones, P., University of Cambridge, United Kingdom: The epidemiology of schizophrenia and related disorders.
Krystal, J., Yale University/VA Medical Center, New Haven, Connecticut: Glutamate, microcircuits, macrocircuits, and cortical dysfunction in schizophrenia: A computational and translational neuroscience perspective.
Law, A., University of Colorado, Denver: Genetics and neurobiology: Translational studies of the AKT pathway relevant to schizophrenia.
Lewis, D., University of Pittsburgh, Pennsylvania: GABA in schizophrenia: Basic and translational research.
Malhotra, A., The Zucker Hillside Hospital, Glen Oaks, New York: Biomarkers of antipsychotic drug response.
Meyer, U., University of Zürich, Switzerland: Immune involvement in schizophrenia: Neuroinflammation and beyond.
Miller, A., Emory University School of Medicine, Atlanta, Georgia: Impact of inflammation on neurotransmitters and neurocircuits: Relevance to psychiatric disorders.
Ming, G-I., University of Pennsylvania/Perelman School of Medicine, Philadelphia: 2D and 3D stem-cell biology and its application to study of mental illness.
Moghaddam, B., Oregon Health & Science University, Portland: Preclinical studies of the glutamate system.
Moore, H., Columbia University, New York, New York: Animal models in psychiatry.
Murray, R., King's College, London, United Kingdom: The clinical syndrome (unable to attend—presented by Jeremy and Peter Jones).
Sawa, A., Johns Hopkins University, Baltimore, Maryland: Translational approaches.
Sohal, V., University of California, San Francisco: Prefrontal circuits + cognitive deficits in schizophrenia.
Walters, J., Cardiff University, United Kingdom: Finding genes for schizophrenia.
Weinberger, D., Lieber Institute for Brain Development, Washington, D.C.: Genomic insights to the developmental origins of schizophrenia.

Mouse Development, Stem Cells, and Cancer

June 6–25

INSTRUCTORS **B. Allen,** University of Michigan Medical School, Ann Arbor
A. Ralston, Michigan State University, East Lansing

CO-INSTRUCTORS **T. Caspary,** Emory University School of Medicine, Atlanta, Georgia
C. Forsberg, University of California, Santa Cruz

ASSISTANTS **W. Cui,** University of Massachusetts, Amherst
M. Echevarria Andino, University of Michigan, Ann Arbor
N. Franks, University of Michigan Medical School, Ann Arbor
T. Frum, Michigan State University, East Lansing
E. Gigante, Emory University School of Medicine, Atlanta, Georgia
T. McCann, University of California, Santa Cruz
S. Smith-Berdan, University of California, San Francisco

This intensive lecture and laboratory course was designed for scientists interested in using mouse models to study mammalian development, stem cells, and cancer. The lecture portion of the course, taught by leaders in the field, provided the conceptual basis for contemporary research in embryogenesis, organogenesis in development and disease, embryonic, adult and induced pluripotent stem cells, and cancer biology.

The laboratory and workshop portions of the course provided hands-on introduction to engineering of mouse models, stem-cell technologies, and tissue analyses. Experimental techniques included genome editing by CRISPR-Cas9, ex vivo and in vivo embryo genome editing using adeno-associated viruses, pronuclear microinjection, isolation, and culture/manipulation of pre- and

postimplantation embryos, embryo transfer, embryo electroporation and roller bottle culture, chimera generation, generation and differentiation of mouse embryonic stem cells and fibroblasts, vibratome and cryosectioning, in situ RNA hybridization, immunostaining, FACS sorting and analysis of hematopoietic stem cells, skeletal preparation, organ explant culture, and fluorescent imaging, including live time-lapse microscopy.

This course was supported with funds provided by the National Cancer Institute.

PARTICIPANTS

Bayerl, J., M.Sc., Weizmann Institute of Science, Rehovot, Israel
Bentley, M., B.S., The University of Alabama, Birmingham
Burnet, G., M.S., University of Queensland, School of Biomedical Sciences, Brisbane, Australia
Erben, L., M.Sc., National Institutes of Health, National Institute of Child Health and Human Development, Bethesda, Maryland
Kondratova, A., Ph.D., Cleveland Clinic, Ohio
Migliori, V., Ph.D., University of Cambridge UK, Gurdon Institute, Cambridge, United Kingdom.
Mohad, V., B.S., University of Wisconsin, Madison
Narasimhan, A., Ph.D., Indiana University, Indianapolis
Reynolds, K., B.S., University of California. Davis, Sacramento
Rhee, H-S., Ph.D., University of Toronto, Mississauga, Canada
Richard, D., B.S., Harvard University, Cambridge, Massachusetts
Stapel, C., Ph.D., The Babraham Institute, Cambridge, United Kingdom
Watts, J., B.S., Michigan State University, East Lansing

SEMINARS

Allen, B., University of Michigan Medical School, Ann Arbor: Overview of mouse development. Hedgehog signaling in embryonic development, adult tissue homeostasis, and cancer. Ethics, rigor, and reproducibility.
Behringer, R., University of Texas MD Anderson Center, Houston: Mouse genetic tools.
Capel, B., Duke University Medical Center, Chapel Hill, North Carolina: The battle of the sexes: Setting the fate of the gonad.
Caspary, T., Emory University School of Medicine, Atlanta, Georgia: Neural patterning; listening to the mutants.
Cebrian, C., University of Michigan, Ann Arbor; Kidney culture and live imaging.
Charron, F., Montreal Clinical Research Institute, Canada: Hedgehog signaling in nervous system development and brain tumor formation.
Forsberg, C., University of California, Santa Cruz: Hematopoietic stem cells I. Hematopoietic stem cells II.
Hadjantonakis, K., Memorial Sloan Kettering Cancer Center, New York, New York: Gastrulation in vivo and in vitro: Quantitative approaches at single-cell resolution.
Hogan, B., Duke University Medical Center, Chapel Hill, North Carolina: History of the mouse model.
Katz, D., Emory University, Atlanta, Georgia: Novel epigenetic functions for the histone demethylase LSD1 in inherited disease and Alzheimer's disease.
Laird, D., University of San Francisco, California: Lineage and fate in the germ line.
Mager, J., University of Massachusetts, Amherst: Regulation of genomic imprinting during early development.
Mariani, F., Keck School of Medicine, Los Angeles, California: Skeletal stem cells: From bench to bedside.
Pourquié, O., Harvard University, Boston, Massachusetts: Segmentation and bilateral symmetry of the vertebrate embryo.
Ralston, A., Michigan State University, East Lansing: Overview of mouse development. Stem cell origins in the blastocyst.
Rivera-Perez, J., University of Massachusetts Medical School, Shrewsbury: Assessment of early lethal phenotypes, developmental constraints, and a novel method for generating gene edited mice.
Rock, J., Boston University School of Medicine, Massachusetts: Using mice to dissect cell lineage relationships in the lung.
Stover, P., Texas A&M Agrilife, College Station: In search of a common pathway for folic-acid-responsive neural tubule defects, cancers, and neurodegeneration.
Trainor, P., Stowers Institute for Medical Research, Kansas City, Missouri: Neural crest cells: Development, evolution, and disease.

Metabolomics

June 9–25

INSTRUCTORS **A. Caudy,** University of Toronto, Ontario, Canada
J. Cross, Memorial Sloan Kettering Cancer Center, New York, New York
A. Rosebrock, Stony Brook University, New York

ASSISTANTS **I. Abramovich,** Technion Institute of Technology, Haifa, Israel
M. Berisa, Memorial Sloan Kettering Institute, New York, New York
D. Sumpton, Cancer Research Beatson Institute, Glasgow, United Kingdom
S. Violante, Memorial Sloan Kettering Institute, New York, New York
O. Zaslaver, University of Toronto, Ontario, Canada

Metabolomics involves the systematic interrogation of the abundance of small chemical molecules (metabolites) within cells, tissues, organs, and organisms. In parallel with high-throughput technologies that facilitate genomic, transcriptomic, and proteomic analyses of cellular and organismal physiology, technologies for metabolite profiling represent an important source of information about the dynamic state of the cell or tissue that is relevant in both health and disease.

LC-MS metabolomics were the primary focus of the course and were applied for both targeted and untargeted analysis of endogenous metabolites and in vitro enzyme reactions. We used approaches for steady-state measurement of metabolite levels as well as assessment of metabolite flux. To complement these LC-MS analyses, there were experiments involving other methodologies. There were shorter-term activities with other methodologies, including GC-MS, polarimetric and Seahorse measurement of oxygen consumption, FRET sensors and/or MitoTracker measurements, enzymatic techniques for metabolite measurement, and uptake experiments.

The consistent and extended application of LC-MS reflects the emphasis of the course, and the exposure to other methodologies allowed students to appreciate the utility and complementarity of these methods. The objectives for students were to analyze quantitatively and qualitatively LC-MS data using currently available tools (vendor software, Rosebrock tools, XCMS online, Agilent Profinder/Genespring); to understand common interferences and limitations of LC-MS and GC-MS analysis; to recognize key issues in experimental design and sample preparation for metabolomics; to be aware of major biochemical pathways active in commonly used cell types; and to become familiar with methods for determining different types of oxygen consumption.

The laboratory exercises included:

- Full scan experiment on knockout/drug treatment. Students identified significantly changed metabolites and used MS/MS fragmentation and other methods for identification.
- Experimental treatments were selected in which the discovery of a phenotype was possible with only one of several analytic/separation methodologies.
- Identification of a significantly changed metabolite in blood/plasma/urine and development of a targeted method for its analysis by QQQ, including determination of LOD/LOQ/linearity and other appropriate method validation.
- Measurement of metabolite flux by pulse labeling (i.e., kinetic flux profiling).
- Enzyme assay to determine V_{max} and K_m (measure on MSD?).
- Enzymatic synthesis of a compound (e.g., sedoheptulose bisphosphate or ribose-1-phosphate) and purification by mass-based or HPLC fractionation.

The non-LC-MS methodologies included Roche kits for lactate and glucose measurement in culture supernatant; seahorse experiment, including all drugs for uncoupled, etc.; FRET experiment (by cytometry and/or microscopy) for NADH levels and proton gradient-sensitive Mitotracker staining; nutrient uptake by 14C (potentially amino acids?); and glycolytic flux in mammalian cells by 3H.

Students received hands-on training on Agilent QTOF and Thermo Orbitrap, Q-exactive, and Vantage triple quadrupole, and Waters SYNAPT G2-S mass spectrometers.

This course was supported with funds provided by the National Institute of General Medical Sciences, Helmsley Charitable Trust, and Howard Hughes Medical Institute.

PARTICIPANTS

Balbach, M., Ph.D., Weill Cornell Medical College, New York, New York
Chan, F., Ph.D., Brown University, Providence, Rhode Island
Heralde, F., Ph.D., University of the Philippines, Manila
Jinich, A., Ph.D., Weill Cornell Medicine, New York, New York
Konovalova, S., Ph.D., University of Helsinki, Finland
Kuhn, N., M.S., Memorial Sloan Kettering Cancer Center, New York, New York
Kulkarni, C., Ph.D., University of Rochester Medical Center, New York
Lewis, C., Ph.D., Whitehead Institute for Biomedical Research, Cambridge, Massachusetts
Mashabela, G., Ph.D., University of Cape Town, South Africa
McBrayer, S., Ph.D., Dana-Farber Cancer Institute, Boston, Massachusetts
Mitosch, K., Ph.D., European Molecular Biology Laboratory, Heidelberg, Germany
Mulder, S., B.A., University of Nebraska Medical Center, Omaha
Park, J.S., Ph.D., Harvard Medical School, Boston, Massachusetts
Pillai, S., Ph.D., Moffitt Cancer Center, Tampa, Florida
Sheldon, R., Ph.D., University of Iowa, Iowa City
Stalnecker, C., Ph.D., Lineberger Comprehensive Cancer Center, Chapel Hill, North Carolina
Xiong, J., Ph.D., National Institutes of Health, National Heart, Lung, and Blood Institute, Bethesda, Maryland

SEMINARS

Amador-Noguez, D., University of Wisconsin, Madison: Metabolomics in microbial biofuel production.

Clasquin, M., Pfizer, Boston, Massachusetts: Understanding disease through unlabeled and isotope tracer-based metabolomics and lipidomics.

Evans, A., Metabolon, Morrisville, North Carolina: Precision metabolomics: A single technology for understanding human health.

Fan, W-M.T., University of Kentucky, Lexington: NMR profiling of metabolites and their ^{13}C-labeling patterns in stable isotope-resolved metabolomics studies.

Gottlieb, E., Cancer Research UK Beatson Institute, Glasgow, United Kingdom: Metabolic advantages and vulnerabilities of cancer.

Gross, S., Weill Cornell Medical College, New York, New York: Untargeted metabolite profiling to discover mechanisms of drug actions, gene functions, and screen for inborn errors of metabolism.

Keshari, K., Memorial Sloan Kettering Cancer Center, New York, New York: Interrogating cancer metabolism using hyperpolarized magnetic resonance.

Mesaros, C., University of Pennsylvania, Philadelphia: Chiral lipidomics and metabolomics targeted on CoAs.

Vander Heiden, M., Massachusetts Institute of Technology, Cambridge: Environmental influence on cell metabolism to affect proliferation.

Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.

Zhu, J., Shanghai Institute of Organic Chemistry, Pudong, China: Metabolomics: A system biology approach to study disease pathogenesis.

Statistical Methods for Functional Genomics

June 29–July 12

INSTRUCTORS **H. Bussemaker,** Columbia University, New York, New York
S. Davis, National Institutes of Health, Columbia, Maryland
T. Lappalainen, New York Genome Center/Columbia University, New York
M. Love, University of North Carolina, Chapel Hill

CO-INSTRUCTORS **V. FitzPatrick,** Columbia University, New York, New York
H.T. Rube, Columbia University, New York, New York

ASSISTANTS **X. Li,** Columbia University, New York, New York
C. Rastogi, Columbia University, New York, New York

Over the past decade, high-throughput assays have become pervasive in biological research because of both rapid technological advances and decreases in overall cost. To properly analyze the large data sets generated by such assays and thus make meaningful biological inferences, both experimental and computational biologists must understand the fundamental statistical principles underlying analysis methods. This course was designed to build competence in statistical methods for analyzing high-throughput data in genomics and molecular biology.

The course curriculum included the R environment for statistical computing and graphics; introduction to Bioconductor; review of basic statistical theory and hypothesis testing; experimental design, quality control, and normalization; high-throughput sequencing technologies; expression profiling using RNA-Seq and microarrays; in vivo protein binding using ChIP-seq; high-resolution chromatin footprinting using DNase-seq; DNA methylation profiling analysis; integrative analysis of data from parallel assays; representations of DNA-binding specificity and motif discovery algorithms; predictive modeling of gene regulatory networks using machine learning; and analysis of posttranscriptional regulation, RNA-binding proteins, and microRNAs.

Detailed lectures and presentations by instructors and guest speakers were combined with hands-on computer tutorials. The methods covered in the lectures were applied to example high-throughput data sets.

This course was supported with funds provided by the National Institute of General Medical Sciences.

PARTICIPANTS

Chang, A., Ph.D., Harvard Medical School, Boston, Massachusetts
Clark, E., Ph.D., Brandeis University, Dorchester, Massachusetts
Deyssenroth, M., Dr.Ph., Icahn School of Medicine at Mount Sinai, New York, New York
Diedrich, J., Ph.D., St. Jude Children's Research Hospital, Memphis, Tennessee
Fang, X., Ph.D., University of California, San Diego
Filippi, C., Ph.D., National Institute for Agriculture Technology, Hurlingham, Argentina
Frueh, S., Ph.D., Cornell University, Ithaca, New York
Gholamalamdari, O., Ph.D., University of Illinois, Urbana-Champaign
Glasner, A., Ph.D., Memorial Sloan Kettering Cancer Center, New York, New York
Hui-Yuen, J., Ph.D., Feinstein Institute for Medical Research, CCMC, Lake Success, New York
Kim, Y., Ph.D., Sungkyunkwan University, Seongnam, South Korea
Levings, D., Ph.D., University of Minnesota Medical School, Duluth
Lorbeer, F., M.Sc., University of California, Berkeley
Lu, X., Ph.D., Cincinnati Children's Hospital Medical Center, Ohio
Mielko, Z., M.S., Duke University, Durham, North Carolina
Novikova, G., B.S., Icahn School of Medicine at Mount Sinai, New York, New York
Pippa, R., Ph.D., Thomas Jefferson University, Philadelphia, Pennsylvania
Reardon, P., B.A., National Institute of Mental Health, Bethesda, Maryland
Redman, L., B.S., Vanderbilt University, Nashville, Tennessee
Su, T., M.S., Ludwig-Maximilians-Universität München, Münich, Germany
Toral, M., B.A., Stanford University, Palo Alto, California
Vasudevan, D., Ph.D., New York University School of Medicine, New York
Wu, X., B.S., Cold Spring Harbor Laboratory
Zhang, H., Ph.D., Columbia University Medical Center, New York, New York
Zhang, Z., B.S., Gerstner Sloan Kettering Graduate School, New York, New York

SEMINARS

Adamson, B., University of California, San Francisco: CRISPR-based functional genomics performed with single-cell resolution: Parts 1 and 2.
All Instructors: Panel on Doing Comp Bio Research
Bussemaker, H., Columbia University, New York, New York: Normal distribution and multiple testing. Scoring differential expression. Gene ontology scoring. Lab: Gene ontology scoring. Basic discriminative motif discovery. Weight matrices. Regression-based motif discovery (REDUCE). Basics of ChIP-seq analysis. Linear models of regulatory networks. Lab: aQTL.
Davis, S., National Institutes of Health, Columbia, Maryland: Introduction to R: Parts 1 and 2. Bioconductor overview. Machine learning. Advanced bioc. Introduction to high-dimensional data. Principle component analysis, dimensionality reduction, and their applications. Cloud-based computation for genomics.
DePlancke, B., École Polytechnique Fédérale de Lausanne EPFL, Switzerland: A primer to single-cell transcriptomic analyses. Single cell–based dissection of somatic stem cells in adipose tissue and beyond. Dissecting the impact of noncoding variants on phenotypic diversity.
Lappalainen, T., New York Genome Center/Columbia University, New York: Introduction to NGS. Introduction to eQTLs and allele-specific expression. Lab: eQTLs and allele-specific expression.
Love, M., University of North Carolina, Chapel Hill: Basics of RNA-Seq analysis. Lab: RNA-Seq differential expression. RNA-Seq isoform-level analysis: Parts 1 and 2. Quantification and QC lecture/lab (Salmon, Tximport, MultiQC). [Combined Session]
Patro, R., Stony Brook University, New York: Quantifying transcript abundance. Quantification and QC lecture/lab (Salmon, Tximport, MultiQC). [Combined Session]
Risso, D., Weill Cornell Medicine, New York, New York: Statistical methods for single-cell RNA-Seq: Parts 1 and 2.
Rube, H.T, Columbia University, New York, New York: Student's T-distribution demystified. Basics of linear regression. Bayes' theorem.

Advanced Techniques in Molecular Neuroscience

June 29–July 14

INSTRUCTORS
C. Lai, Indiana University, Bloomington
J. LoTurco, University of Connecticut, Storrs
A. Schaefer, Icahn School of Medicine at Mount Sinai, New York, New York

ASSISTANTS
P. Ayata, Icahn School of Medicine at Mount Sinai, New York, New York
A. Battison, University of Connecticut, Storrs
A. Chan, Mount Sinai School of Medicine, New York, New York
M.K. Duff, Mount Sinai School of Medicine, New York, New York
G. Panahi, Indiana University, Bloomington
E. Perez, Indiana University, Bloomington
H. Strasburger, Icahn School of Medicine at Mount Sinai, New York, New York
J. Sullivan, Icahn School of Medicine at Mount Sinai, New York, New York
S. Veugelen, Icahn School of Medicine at Mount Sinai, New York, New York

This laboratory and lecture course was designed to provide neuroscientists at all levels with a conceptual and practical understanding of several of the most advanced techniques in molecular neuroscience. The course curriculum was divided into three sections: an extensive and up-to-date set of laboratory exercises, daily lectures covering the theoretical and practical aspects of the various methods used in the laboratory, and a series of evening research seminars. The informal and interactive evening lectures were given by leading molecular neuroscientists and served to illustrate the ways in which the various experimental approaches have been used to advance specific areas of neurobiology. In this year's course, the laboratory portion included topics such as an introduction to the design and use of animal virus vectors in neurobiology; the use of CRISPR genome editing and RNAi approaches for regulating the expression of specific genes in neurons; practical exercises

in gene delivery systems including mammalian cell infection and transfection and electroporation techniques for targeted gene transfer in vivo; an introduction to overall strategies, use, and design of BAC transgenic vectors; real-time RT-PCR analyses; assays of chromatin and chromatin structure in neurons; and mRNA isolation from specified neural subtypes by TRAP.

This course was supported with funds provided by the National Institute of Mental Health.

PARTICIPANTS

Anderson, L., Ph.D., Massachusetts General Hospital/Harvard Medical School, Charlestown
Cruz-Torres, I., B.S., University of Colorado Anschutz Medical Campus, Aurora
Ford, L., Ph.D., Columbia University, New York, New York
Fulenwider, H., B.A., University of Georgia, Athens
Gamache, J., B.A., University of Minnesota, Minneapolis
Guttmann, L., Ph.D., NASA, Mountain View, California
Hisey, E., Ph.D., Duke University, Durham, North Carolina
Kuan, C-Y., Ph.D., University of Virginia School of Medicine, Charlottesville
Maddox, S., Ph.D., McLean Hospital/Harvard Medical School, Belmont, Massachusetts
Maurinot, F., B.S., Sorbonne Université, Paris, France
Merritt, J., B.S., Emory University, Atlanta, Georgia
Myrum, C., Ph.D., National Institutes of Health, Baltimore, Maryland
Ravel-Godreuil, C., B.S., CIRB College de France, Paris, France
Tran Anh, K., M.D., Tokyo Medical and Dental University, Japan
Trotta, M., M.S., IFIBIO (UBA-CONICET), Buenos Aires, Argentina
Wang, X., Ph.D., Harvard University, Cambridge, Massachusetts

SEMINARS

Akbarian, S., Icahn School of Medicine at Mount Sinai, New York, New York: Chromosomal conformation and 3D genome mapping in brain.
Cai, D., University of Michigan, Ann Arbor: Neural circuit labeling and reconstruction in the mammalian brain.
Chan, K., Broad Institute of MIT and Harvard, Cambridge, Massachusetts: Enabling mutant Huntington repression through vector engineering and zinc finger proteins.
Cong, L., Broad Institute of MIT and Harvard, Cambridge, Massachusetts: In vivo genome engineering using CRISPR systems: Toolbox expansion and new development with neuro-applications.
Darnell, R., HHMI/The Rockefeller University/NY Genome Center, New York: RNA regulation in the brain.
Eberwine, E., University of Pennsylvania, Philadelphia: Transcriptome variability and theories of cellular phenotype: Multimodal subcellular genomics.
Haas, K., University of British Columbia, Vancouver, Canada: Single-cell electroporation for in vivo neuronal transfection.
Harwell, C., Harvard Medical School, Boston, Massachusetts: Generating neural diversity in the forebrain.
Hattar, S., National Institutes of Health, Bethesda, Maryland: Understanding behavior from molecules to circuits.
Jaffrey, S., Weill Cornell Medical College, New York, New York: Mapping RNA modifications and monitoring epitranscriptome dynamics.
Kanadia, R., University of Connecticut, Storrs: Understanding the role of minor intron splicing in disease and development.
Kenny, P., Icahn School of Medicine at Mount Sinai, New York, New York: microRNA regulation of drug craving.
Liddelow, S., New York University, New York: What do reactive astrocytes do?
Luikart, B., Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire: Impact of Pten dysfunction on neuronal physiology: A model for autism.
Maher, B., Lieber Institute for Brain Development/HHMI, Baltimore, Maryland: In utero electroporation for cellular transgenesis in the developing mammalian forebrain.
Schafer, D., University of Massachusetts Medical School, Worcester: Studying microglial function and dysfunction within neural circuits in health and disease.
Schmidt, E., The Rockefeller University, New York, New York: Molecular phenotyping of distinct cortical cell types using the translating ribosome affinity purification (TRAP) approach.
Silver, D., Duke University Medical Center, Durham, North Carolina: Dynamic posttranscriptional control of cortical development.
Tollkuhn, J., Cold Spring Harbor Laboratory: Sex differences in gene expression.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.

Single Cell Analysis

June 29–July 14

INSTRUCTORS

D. Chenoweth, University of Pennsylvania, Merion Station
M. McConnell, University of Virginia School of Medicine, Charlottesville
G. Yeo, University of California, San Diego

COURSE TEACHING ASSISTANTS

I. Burbulis, University of Virginia School of Medicine, Charlottesville
I. (Alex) Chaim, University of California, San Diego
W. Chronister, University of Virginia, Charlottesville
M. Haakenson, University of Virginia School of Medicine, Charlottesville
R. Marina, University of California, San Diego
S. Shaffer, University of Pennsylvania, Philadelphia
E. Wheeler, University of California, San Diego
M. Wolpert, University of Virginia, Charlottesville
D. Wu, University of Pennsylvania, Philadelphia
B. Yee, University of California, San Diego

The goal of this 2-week course was to familiarize students with cutting-edge technologies for characterization of single cells. Modules of the course were taught by scientists with expertise in distinct areas of single-cell analysis. Topics covered included quantitative single-cell analysis by RNA-seq, genomic DNA analysis, proteomics, and metabolomics. Multiple nucleic amplification methodologies including droplet-based RNA-seq, MALBAC, and MDA were used. In addition, students were instructed in basic bioinformatic analysis of next generation sequencing data. Other topics included single-cell genome, transcriptome, and proteome measurement; introductory next-generation sequencing data analysis; photoactivatable single-cell probes; and single-cell mass spectrometry/soft X-ray tomography.

This course was supported with funds provided by the National Institute of General Medical Sciences, Howard Hughes Medical Institute, and Helmsley Charitable Trust.

PARTICIPANTS

Bartlett, T., Ph.D., University College London, United Kingdom
Bonthuis, P., Ph.D., University of Utah, Salt Lake City
Gerdes, P., M.S., Mater Research Institute/University of Queensland, Woolloongabba, Australia
Glasauer, S., Ph.D., University of California, Santa Barbara
Jara, J.S., Burke Medical Research Institute, White Plains, New York
Kim, S., B.S., Harvard University/Boston Children's Hospital, Boston, Massachusetts
Krup, A., B.A., University of California/Gladstone Institutes, San Francisco
Lamanna, J., B.Sc., University of Toronto, Ontario, Canada
Lee, P.R., B.S., Seoul National University, Jongno-gu, South Korea
Maitra, M., B.A., McGill University, Montreal, Quebec, Canada
McCullough, K., Ph.D., Mclean Hospital, HMS, Belmont, Massachusetts
Mohammadi, S., Ph.D., Massachusetts Institute of Technology, Cambridge
Olingy, C., Ph.D., La Jolla Institute for Allergy and Immunology, San Diego, California
Petersen, B., B.S., University of North Carolina, Chapel Hill, Durham
Picciani, N., B.S., University of California, Santa Barbara
Quiñones-Al-Muhtaseb, S., B.S., Tufts University School of Medicine, Boston, Massachusetts
Settas, N., Ph.D., National Institutes of Health, Bethesda, Maryland
Tornini, V., Ph.D., Yale University, New Haven, Connecticut
Zhu, X., Ph.D., University of Hawaii at Manoa, Honolulu

SEMINARS

Allbritton, N., University of North Carolina, Chapel Hill: Microdevices for assaying and sorting single cells.
Chang, J., University of California, San Diego: Using single-cell approaches to elucidate the ontogeny of T-cell memory.
Eberwine, J., University of Pennsylvania, Philadelphia; Transcriptome variability and theories of cellular phenotype: Multimodal subcellular genomics.
Herr, A., University of California, Berkeley: Electrophoretic cytometry: Microfluidics advances single-cell protein analysis.
Larabell, C., University of California, San Francisco: Soft X-ray tomography: CT scans of single cells.
Raj, A., University of Pennsylvania, Philadelphia: Applications of single-cell analysis in cancer biology.
Romanova, E., University of Illinois, Urbana: Analysis of peptides by mass spectrometry.
Rubakhin, S., University of Illinois, Urbana: Introduction into single-cell mass spectrometry.
Sims, P., Columbia University, New York, New York: New tools for single-cell analysis in cancer.
Streets, A., University of California, Berkeley: Imaging and sequencing single cells.
Teichmann, S., Wellcome Trust Sanger Institute, Hinxton, United Kingdom: Immunogenomics one cell at a time.
Teruel, M., Stanford University, California: Controlling mammalian cell differentiation with feedback and noise.
Trotta, N., Cell Microsystems Inc., Research Triangle Park, North Carolina: CellRaft technology for single-cell analysis.
Voet, T., KU Leuven & Wellcome Trust Sanger Institute, Leuven, Belgium: Single-cell (multi-omics) sequencing to study the biology of cellular heterogeneity in health and disease.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.

Drosophila Neurobiology: Genes, Circuits, and Behavior

June 29–July 19

INSTRUCTORS **C.A. Frank,** University of Iowa, Iowa City
E. Heckscher, University of Chicago, Illinois
A. Keene, Florida Atlantic University, Jupiter

ASSISTANTS **K. Lembke,** University of Iowa, Iowa City
S. Ly, University of Pennsylvania, Philadelphia
J. Meng, University of Chicago, Illinois
M. Yurgel, Florida Atlantic University, Jupiter

This laboratory/lecture course was intended for researchers at all levels from beginning graduate students through established primary investigators who want to use *Drosophila* as an experimental system for nervous system investigation. This 3-week course was designed to introduce students to a wide variety of topics and techniques, including the latest approaches for studying nervous system development, activity, and connectivity, as well as complex behaviors and disease models. Daily research seminars presented comprehensive overviews of specific subfields of nervous system development or function or focused on state-of-the-art techniques and approaches in *Drosophila* neuroscience. Expert guest lecturers discussed their findings and approaches, and they brought along their own assays and techniques for students to learn in the laboratory part of the course. The hands-on portion of the course was centered on inquiry-based projects, utilizing the different morphological and physiological measurements and behavioral paradigms learned at the course. This included molecular-genetic analyses, immunocytochemistry, recording of activity using electrophysiology and genetically encoded calcium indicators, optogenetic and thermogenetic control of neural activity, and numerous quantitative behavioral measures. Collectively, the

course provided a comprehensive and practical introduction to modern experimental methods for studying the neural basis of behavior in *Drosophila*.

This course was supported with funds provided by the National Science Foundation, Helmsley Charitable Trust and Howard Hughes Medical Institute.

PARTICIPANTS

Botero, V., B.A., The Scripps Research Institute, Jupiter, Florida
Chin, M., B.A., Indiana University, Bloomington
Durkin, J., B.S., University of Michigan, Ann Arbor
Greenblatt, E., Ph.D., Carnegie Institute for Science, Baltimore, Maryland
Krätschmer, P., B.A., University College London, United Kingdom
Laws, K., Ph.D., University of Pennsylvania School of Medicine, Philadelphia
Ling, D., M.S., Washington University in St. Louis, Missouri
Palmateer, C., B.S., Florida State University, Tallahassee
Pardo-Garcia, T., B.S., University of Michigan, Ann Arbor
Prisco, L., M.Sc., Deutsches Zentrum für Neurodegenerative Erkrankung, Bonn, Germany
Schneider, J., Ph.D., Lehigh University, Bethlehem, Pennsylvania
Sizemore, T., B.S., West Virginia University, Morgantown

SEMINARS

Carrillo, R., University of Chicago, Illinois: Synaptic homeostasis at the *Drosophila* neuromuscular junction.
Crocker, A., Middlebury College, Vermont: Liberal arts.
Dubnau, J., Stony Brook School of Medicine, New York: Neurodegeneration.
Dus, M., University of Michigan, Ann Arbor: How fly brains regulate feeding behavior.
Frank, A., University of Iowa, Iowa City: Neurophysiology and synaptic neurotransmission. Fly neurofly bootcamp and fly base. Fly neuroanatomy and neurogenetic tools. Synaptic homeostasis at the *Drosophila* neuromuscular junction.
Gunay, C., Georgia Gwinnett College, Lawrenceville: Computational modeling in *Drosophila* motoneurons.
Hattar, S., National Institutes of Health, Bethesda, Maryland: Understanding behavior from molecules to circuits.
Heckscher, E., University of Chicago, Illinois: Development and function of neural circuits in the motor system. Single neuron labeling methods.
Jayaraman, V., Howard Hughes Medical Institute, Ashburn, Virginia: Abstract internal representations and attractor dynamics in the fly brain.
Kaun, K., Brown University, Providence, Rhode Island: Molecular and circuit mechanisms underlying addiction.
Kayser, M., University of Pennsylvania, Philadelphia: Building brains in our sleep: Using *Drosophila* to study sleep and neurodevelopment.
Keene, A., Florida Atlantic University, Jupiter: Scientific rigor and reproducibility in *Drosophila* neurobiology research. Designing behavior experiments and building behavior apparatus. Metabolic regulation of sleep.
Kohwi, M., Columbia University, New York, New York: Neuroblast development.
Liu, T., Columbia University, New York, New York: Sharing open neurophysiology data.
Louis, M., University of California, Santa Barbara: Orientation behavior in the *Drosophila* larva.
Masek, P., Binghamton University, New York: Neural basis of the interaction between sleep and feeding.
Mosca, T., Thomas Jefferson University, Philadelphia, Pennsylvania: Microscopy. The microscopy toolkit for a Drosophilist. Fly neurofly bootcamp and fly base. Fly neuroanatomy and neurogenetic tools.
Muraro, N., Biomedicine Research Institute of Buenos Aires, Argentina: Electrophysiology in central neurons of *Drosophila*/circadian rhythms and electrophysiology of clock neurons. Central nervous system electrophysiology.
Nagel, K., New York University School of Medicine, New York: The biophysics of sensory processing.
Prinz, A., Emory University, Atlanta, Georgia: Computational modeling.
Pulver, S., University of St. Andrews, United Kingdom: Neural control of locomotion in *Drosophila* larvae,
Reis, T., University of Colorado Medical School, Aurora: Neuronal control of energy balance.
Tomchik, S., The Scripps Research Institute, Jupiter, Florida: Central brain Ca-imaging.
Turkcan, M.K., Columbia University, New York, New York: Building the functional map of the fruit fly brain.
Wildenberg, G., University of Chicago, Illinois: Connectomics.
Wildonger, J., University of Wisconsin, Madison: Use of CRISPR-Cas9 for genome targeting and engineering.
Zhang, B., University of Missouri, Columbia: David Featherstone Memorial Lecture: Glia in *Drosophila*.

Frontiers and Techniques in Plant Science

June 29–July 19

INSTRUCTORS
S. Cutler, University of California, Riverside
J. Dinneny, Stanford University, California
U. Paszkowski, University of Cambridge, United Kingdom

ASSISTANTS
K. Brown, University of Kansas, Lawrence
H. Montero, University of Cambridge, United Kingdom
N.C. Sanden, Norwegian University of Science and Technology, Trondheim, Norway

This course provided an intensive overview of topics in genomics, genetics, physiology, biochemistry, development, and evolution and hands-on experiences in molecular imaging, computational, and high-throughput approaches to understanding plant biology. It emphasized recent results from model organisms including *Arabidopsis,* maize, and tomato, as well as a variety of other plants, and provided an introduction to current methods used in basic and applied plant biology, both theoretically and practically. The seminar series included plant morphology and anatomy, development, evolution, light and circadian biology, hormones, small RNAs and epigenetic inheritance, biotic and abiotic interactions, plant biochemistry, crop domestication, and applications addressing current agronomic problems. Speakers provided expert overviews of their fields, followed by in-depth discussions of their own work. The laboratory sessions provided exposure to cutting-edge experimental and computational techniques currently used in plant research. These included approaches for studying plant development, regulatory networks, transient gene expression, cell type–specific gene expression analysis, computational large-scale data analysis,

and applications of fluorescent proteins, including live imaging, genome editing, and chromatin immunoprecipitation.

Students gained hands-on experience in computational tools and environments for genome assembly; plant imaging and image analysis; design and use of fluorescent sensors; transcriptomics; identification of quantitative trait loci; mapping by sequencing; mathematical modeling of development and hormone action; purification of cell type–specific nuclei (INTACT); and high-throughput cloning.

This course was supported with funds provided by the National Science Foundation.

PARTICIPANTS

Chin, S., Ph.D., Friends' Central School, Wynnewood, Pennsylvania
Di, Z., B.S., University of Cambridge, United Kingdom
Erndwein, L., B.S., University of Delaware, Newark
Feehan, J., M.Sc., The Sainsbury Laboratory, Norwich, United Kingdom
Fischer, H., B.S., University of Arkansas, Fayetteville
Haque, T., B.S., University of Texas, Austin
Kapoor, K., McGill University, Quebec, Canada
Kolbeck, A., M.Sc., Université de Lausanne, Switzerland
Mijatovic, J., B.S., University of Georgia, Athens
Ogawa, S., Ph.D., Riken, Yokohama, Kanagawa, Japan
Okoro, M., M.Sc., Cold Spring Harbor Laboratory/New York University, New York
Ramachandran, P., M.S., Uppsala University, Sweden
Ricci, W., B.S., University of Georgia, Athens
Sgroi, M., B.Sc., University of Cambridge, United Kingdom
Sorkin, M., Ph.D., Washington University in St. Louis, Missouri
Yun, J., M.S., Massachusetts Institute of Technology, Cambridge

SEMINARS

Bailey-Serres, J., University of California, Riverside: INTACT and water extremes signaling.
Baulcombe, D., University of Cambridge, United Kingdom: Small RNAs.
Cutler, S., University of California, Riverside: Chemical and genetic dissection of plant hormone pathways.
Dinneny, J., Carnegie Institution for Science, Stanford, California: Stressed! How roots cope through dynamic behaviors.
Dolan, L., University of Oxford, United Kingdom: Development and evolution of land plant rooting structures.
Hibberd, J., Cambridge University, United Kingdom: Photosynthesis.
Jander, G., Boyce Thompson Institute for Plant Research, Ithaca, New York: Chemical ecology of plant insect interactions.
Johnson, M., Brown University, Providence, Rhode Island: Molecular dialogues between pollen and pistil.
Jones, A., Cambridge University, United Kingdom: Sensing plant hormones in action.
Koenig, D., University of California, Riverside: Computational genetics.
Law, J., The Salk Institute, La Jolla, California: Introduction to epigenetics and chromatin immunoprecipitation (ChIP). Epigenetics and DNA methylation.
Martienssen, R., Cold Spring Harbor Laboratory: Transposons.
Oldroyd, G., John Innes Centre, Norwich, United Kingdom: Nodulation and symbiosis.
Paszkowski, U., University of Cambridge, United Kingdom: Arbuscular mycorrhizal symbiosis.
Pedmale, U., Cold Spring Harbor Laboratory: Light perception and signaling in plants.
Provart, N., University of Toronto, Ontario, Canada: Hypothesis generation with big data.
Rogers, C., University of Cambridge, United Kingdom: An introduction to golden gate DNA assembly.
Shiu, S-H., Michigan State University, East Lansing: Genome evolution.
Sinha, N., University of California, Davis: An introduction to plant biology. Natural variation in leaf shape and complexity: Functional significance.
Surridge, C., SpringerNature, London, United Kingdom: Publishing without tears.
Voytas, D., University of Minnesota, Saint Paul: Precise genome engineering with sequence-specific nucleases.
Weigel, D., Max-Planck Institute for Developmental Biology, Tübingen, Germany: Genetic variation.
Wilkins, O., McGill University, Quebec, Canada: Regulatory networks.
Williams, J., Cold Spring Harbor Laboratory: Genomics data carpentry basic research skills in R.

Computational Neuroscience: Vision

July 9–22

INSTRUCTORS

G. Boynton, University of Washington, Seattle
M. Cohen, University of Pittsburgh, Pennsylvania
G. Horwitz, University of Washington, Seattle
J. Pillow, Princeton University, New Jersey

Computational approaches to neuroscience will produce important advances in our understanding of neural processing. Prominent success will come in areas in which strong inputs from neurobiological, behavioral, and computational investigation can interact. The theme of the course was that an understanding of the computational problems, the constraints on solutions to these problems, and the range of possible solutions can help guide research in neuroscience. Through a combination of lectures and hands-on experience with MATLAB-based computer tutorials and projects, this intensive course examined visual information processing from the retina to higher cortical areas, spatial pattern analysis, motion analysis, neuronal coding and decoding, attention, and decision-making.

This course was supported with funds provided by the Helmsley Charitable Trust and the Howard Hughes Medical Institute.

PARTICIPANTS

Beyeler, M., Ph.D., University of Washington, Seattle
Birman, D., B.A., Stanford University, California
Brezovec, L., B.S./B.A., Stanford University, California
Burlingham, C., B.H.A., New York University, New York
Bushnell, B., B.S., New York University, New York
Chang, K., B.A., University of Washington, Seattle

Dunn-Weiss, E., B.S., Johns Hopkins University, Baltimore, Maryland
Fernandez-Alonso, M., B.S./M.S., Newcastle University, Newcastle Upon Tyne, United Kingdom
French, R., B.S., University of Rochester, New York
Fu, J., B.A., Baylor College of Medicine, Houston, Texas
Gadot, M., M.Sc., Bar-Ilan University, Ramat Gan, Israel
Henderson, M., B.A., University of California, San Diego
Klindt, D., M.Sc., University of Tübingen, Germany
Kupers, E., B.S., New York University, New York
Lafer-Sousa, R., Ph.D., Massachusetts Institute of Technology, Cambridge
Mohan, K., B.S./M.S., The University of Chicago, Illinois
Ramos Traslosheros Lopez, L., M.Sc., European Neuroscience Institute, Göttingen, Germany
Roth, Z., Ph.D., National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland
Schütt, H., Ph.D., University of Tübingen, Germany
Stacy, A., B.A., Brandeis University, Waltham, Massachusetts
Wailes-Newson, K., B.Sc., University of York, United Kingdom
Waskom, M., Ph.D., New York University, New York
Whitmire, M., B.S., University of Texas, Austin
Xue, C., Ph.D., University of Pittsburgh, Pennsylvania

SEMINARS

Boynton, G., University of Washington, Seattle: fMRI, visual attention, visual prosthetics.
Cardin, J., Yale University, New Haven, Connecticut: Adaptation.
Chichilnisky, E. J., Stanford University, California: Retina II.
Churchland, A., Cold Spring Harbor Laboratory: Multimodal processing I.
Cohen, M., University of Pittsburgh, Pennsylvania: Attention and population coding.
Connor, E., Johns Hopkins University, Baltimore, Maryland: Object recognition.
DeAngelis, G., University of Rochester, New York: Motion and neurometric analyses.
Fine, I., University of Washington, Seattle: Multimodal processing II.
Glickfeld, L., Duke University, Durham, North Carolina: Rodent models of vision and attention.
Heeger, D., New York University, New York: Theory of cortical function.
Horwitz, G., University of Washington, Seattle: Color and related concepts.
Ma, W.J., New York University, New York: Bayesian perception.
Movshon, T., New York University, New York: What vision (and this course) is all about.
Osborne, L., University of Chicago, Illinois: Motion coding.
Palmer, S., University of Chicago, Illinois: Information theory and neural coding.
Pillow, J., Princeton University, New Jersey: Statistical models for neural coding, GLMs.
Read, J., Newcastle University, United Kingdom: Binocular vision.
Rieke, F., University of Washington, Seattle: Retina I.
Rust, N., University of Pennsylvania, Philadelphia: Visual memory.
Shlens, J., Google, San Francisco, California: Deep learning.
Simoncelli, E., New York University, New York: Natural image statistics.
Treue, S., German Primate Center, Göttingen, Germany: Physiology of attention.

Synthetic Biology

July 24–August 6

INSTRUCTORS
J. Chappell, Rice University, Houston, Texas
J. Dueber, University of California, Berkeley
E. Franco, University of California, Riverside
K. Haynes, Arizona State University, Tempe

ASSISTANTS
A. Ameruoso, Rice University, Houston, Texas
C. Cuba Samaniego, University of California, Riverside
S. Halperin, University of California, Berkeley
S. Tekel, Arizona State University, Tempe

Synthetic biology is a discipline in which living organisms are genetically programmed to carry out desired functions in a reliable manner. The field takes inspiration from our ever-expanding ability to measure and manipulate biological systems, as well as from the philosophical reflections of Schrödinger and Feynman, specifically that physical laws can be used to describe and rationally engineer biology to accomplish useful goals. Cells are the world's most sophisticated chemists, and their ability to adapt to changing environments offers enormous potential for solving modern engineering challenges. Nevertheless, biological systems are noisy, massively interconnected, and nonlinear, and they have not evolved to be easily engineered. The grand challenge of synthetic biology is to reconcile the desire for a predictable, formalized biological design process with the inherent "squishiness" of biology.

This course focused on how the complexity of biological systems can be combined with traditional engineering approaches to result in new design principles for synthetic biology. The

centerpiece of the course was an immersive laboratory experience in which students worked in teams to learn the practical and theoretical underpinnings of synthetic biology research. Broadly, the course explored how cellular regulation (transcriptional, translational, posttranslational, and epigenetic) can be used to engineer cells that accomplish well-defined goals. Laboratory modules covered the following areas: CRISPR technologies for genome editing and gene regulation; cell-free transcription and translation systems to characterize genetic circuits and RNA regulators; modeling gene expression using ordinary differential equations; and high-throughput DNA assembly techniques and genetic design principles.

Students first learned essential synthetic biology techniques in a 4-day "boot camp" at the beginning of the course. Following the boot camp, they rotated through research projects in select areas. Students also interacted closely with a panel of internationally recognized speakers who collectively provided a broad overview of synthetic biology applications, including renewable chemical production and therapeutics, state-of-the-art techniques, case studies in human practices, and socially responsible innovation.

This course was supported with funds provided by the National Institute of General Medical Sciences, Howard Hughes Medical Institute, Helmsley Charitable Trust, and National Science Foundation

PARTICIPANTS

Albin, D., B.S., University of Washington, Seattle
DuPrie, M., B.S., Ludwig Institute for Cancer Research/ University of California, San Diego
Farris, Y., B.S., Pacific Northwest National Laboratory, Richland, Washington
Heyde, S., M.S., Technical University of Denmark, Kgs. Lyngby, Denmark
Khakimzhan, A., B.S., University of Minnesota, Minneapolis
Kwon, U., B.A., Massachusetts Institute of Technology, Cambridge
Lakin, M., Ph.D., University of New Mexico, Albuquerque
Majumder, S., M.S., University of Michigan, Ann Arbor
Palmer, J., B.S., University of Massachusetts, North Dartmouth
Palomo, B., B.S., Texas Tech University, Lubbock
Papadopoulos, J., B.S., University of Wisconsin, Madison
Pereira e Calvo-Villamañán, A., M.S., Institut Pasteur, Paris, France
Pereirinha, J., M.S., CEDOC, Lisboa, Portugal
Moreno, J.S., Ph.D., University of Lausanne (UNIL), Switzerland
Vo, N., B.E., University of Michigan, Ann Arbor
Wimmer, F., M.S., HIRI, Würzburg, Germany

SEMINARS

Arnold, F., California Institute of Technology, Pasadena: Enzymes by evolution: Bringing new chemistry to life.
Avalos, J., Princeton University, New Jersey: Using mitochondria and light for spatial and temporal control of engineered metabolic pathways.
Bennett, M., Rice University, Houston, Texas: Spatiotemporal dynamics in synthetic microbial consortia.
Chen, Y., University of California, Los Angeles: Engineering next-generation T cells for cancer immunotherapy.
Deans, T., University of Utah, Salt Lake City: Synthetic biology approaches for studying hematopoiesis.
Gersbach, C., Duke University, Durham, North Carolina: Genome and epigenome editing for gene therapy and cell programming.
Kong, D., Massachusetts Institute of Technology, Cambridge: Community-driven biotechnology.
Moon, T.S., Washington University in St. Louis, Missouri: Systems and synthetic biology: Constructing programmable cells.
O'Malley, M., University of California, Santa Barbara: Engineering synthetic consortia inspired by the rumen microbiome.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor, and reproducibility.

Chromatin, Epigenetics, and Gene Expression

July 24–August 12

INSTRUCTORS **K. Adelman,** Harvard Medical School, Boston, Massachusetts
M. Mendillo, Northwestern University School of Medicine, Chicago, Illinois
G. Narlikar, University of California, San Francisco

ASSISTANTS **A. Field,** Harvard Medical School, Boston, Massachusetts
T. Henriques, Harvard Medical School, Boston, Massachusetts
M. Keenen, University of California, San Francisco
S. Takagishi, Northwestern University, Evanston, Illinois
E. Wong, University of California, San Francisco

This course was designed for students, postdocs, and principal investigators who have recently ventured into the exciting area of gene regulation. Emphasis was placed on exposing students to a broad array of methodologies to study gene regulation and chromatin structure and dynamics, including both state-of-the-art and well-developed methods.

Students performed widely used techniques such as chromatin immunoprecipitation (ChIP) coupled with sequencing (ChIP-seq), reporter assays of enhancer activity, and RNA expression analysis. They applied a basic pipeline to analyze sequencing results and discussed current informatics strategies.

Students isolated transcription factor complexes and assessed their activity in functional assays. Furthermore, they knocked out specific factors using CRISPR-Cas9 and evaluated the effects on gene expression.

This course provided the basic concepts behind different methods to analyze the chromatin architecture of the genome. Students were exposed to chromosome conformation capture (3C)

experiments, together with other approaches aimed to interrogate the 3D organization of genomes. Moreover, we discussed the computational methods required to analyze these data.

Students learned how to assemble recombinant chromatin and use biophysical methods such as FRET to assay the activity of chromatin remodeling enzymes. They also learned principles of enzyme kinetics and applied these to quantify chromatin remodeling reactions.

Given the broad biological roles for DNA-binding transcription factors and emerging roles of noncoding RNAs in transcription regulation, electrophoretic mobility shift assays (EMSAs) are again becoming widely used for assessing transcription factor binding to regulatory DNA or RNA elements. Students learned how to perform and interpret EMSA experiments, using quantitative gel-based methods.

Experience with basic recombinant DNA techniques was a prerequisite for admission to this course. Lectures by the instructors covered the current state of the gene expression field, theoretical aspects of the methodology, and broader issues regarding strategies for investigating the regulation of gene expression in eukaryotes. Emphasis was placed on advantages and limitations of specific techniques, and data interpretation. In the evenings, invited speakers who are experts in the field presented their work and interacted with students. The students were encouraged and expected to actively participate in these discussions and to take advantage of the many opportunities to network and receive input on their projects and future plans.

This course was supported with funds provided by the National Cancer Institute.

PARTICIPANTS

Bhalla, K., Ph.D., Cold Spring Harbor Laboratory
Dutta, A., B.S., University of Virginia Medical School, Charlottesville
Ferdous, S., B.S., Emory University, Atlanta, Georgia
Ivy, A., M.D/Ph.D., University of California, Irvine
Kelly, C., B.S., Duke University, Durham, North Carolina
Learman, L., B.A., Johns Hopkins University School of Medicine, Baltimore, Maryland
Lugo-Ramos, J., B.Sc., Max-Planck Institute for Evolutionary Biology, Plön, Germany
Matern, M., B.S., University of Maryland School of Medicine, Baltimore
Ochoa-Arenas, C., Ph.D., University of Texas Southwestern Medical Center, Dallas
Park, S., B.S., Johns Hopkins University, Baltimore, Maryland
Schmidt, S., M.S., Helmholtz Zentrum München, Neuherberg, Germany
Schmuecker, A., M.Sc., Gregor-Mendel-Institute, Vienna, Austria
Sorlien, E., B.S., Purdue University, West Lafayette, Indiana
Tu, L-C., Ph.D., University of Massachusetts Medical School, Worcester
Vizjak, P., M.S., Biomedical Center Münich, Planegg-Martinsried, Germany
Zhang, Y., B.S., Duke University, Durham, North Carolina

SEMINARS

Adelman, K., Harvard Medical School, Boston, Massachusetts: Introduction to transcription mechanisms.
Arndt, K., University of Pittsburgh, Pennsylvania: Using yeast to understand the role of chromatin in transcription elongation and termination.
Buratowski, S., Harvard Medical School, Boston, Massachusetts: Dynamics of RNA polymerase II transcription.
Di Croce, L., ICREA and Centre for Genomic Regulation, Barcelona, Spain: Gene regulation dynamics mediated by Polycomb and MLL complexes.
Kingston, R., Massachusetts General Hospital/Harvard Medical School, Boston: Nucleosome compaction, phase separation, and Polycomb-group repressive function during development.
Levine, M., Princeton University, New Jersey: Transcription networks in development and evolution.
Lis, J., Cornell University, Ithaca, New York: Dissecting mechanisms of transcription regulation.
Luger, K., University of Colorado, Boulder: Off to the races: Quantitating the recruitment of proteins to sites of DNA damage.

Madhani, H., University of California, San Francisco: Epigenetic memory over geological timescales.

Marti-Renom, M., CNAG-CRG, Barcelona, Spain: Structural determination of genomes and genomic domains by satisfaction of spatial restraints.

Neugebauer, K., Yale University, New Haven, Connecticut: Nascent RNA and the coordination of splicing with transcription.

Shilatifard, A., Northwestern University Feinberg School of Medicine, Chicago, Illinois: Principles of epigenetics and chromatin in development and human disease.

Stark, A., Research Institute of Molecular Pathology (IMP), Vienna, Austria: Decoding transcriptional regulation in *Drosophila*.

Yao, T., Colorado State University, Fort Collins: Regulation of chromatin dynamics by ubiquitin conjugation and deconjugation.

Witkowski, J., Cold Spring Harbor Laboratory: Ethics, rigor. and reproducibility.

Vierstra, J., and Wouter, M., Altius Institute for Biomedical Sciences, Seattle, Washington: A reference index of human regulatory DNA. Footprinting the human genome.

Imaging Structure and Function in the Nervous System

July 24–August 13

INSTRUCTORS
F. Albeanu, Cold Spring Harbor Laboratory
M. Orger, Champalimaud Foundation, Lisbon, Portugal
L. Palmer, University of Melbourne, Australia
P. Tsai, University of California, San Diego

ASSISTANTS
A. Bandyopadhyay, New York University Langone Medical Center, New York
H. Chae, Cold Spring Harbor Laboratory
J. Donovan, Max Planck Institute of Neurobiology, Martinsried, Germany
O. Gilday, The Hebrew University, Jerusalem, Holon, Israel
L. Knogler, Max Planck Institute of Neurobiology, Münich, Germany
A. La Chioma, Max Planck Institute of Neurobiology, Martinsried, Germany
D. LaTerra, University of Melbourne, Australia

Advances in light microscopy, digital image processing, and the development of a variety of powerful fluorescent probes present expanding opportunities for investigating the nervous system, from synaptic spines to networks in the brain. This intensive laboratory and lecture course provided participants with the theoretical and practical knowledge to use emerging imaging technologies. The primary focus of the course was on in vivo applications of light microscopy, particularly functional imaging with genetically encoded calcium indicators. Methods taught included multiphoton and light-sheet microscopy and the combination of imaging with optogenetics. Lectures by leading experts progressed through basic concepts to presentation of cutting-edge methods. Students learned the fundamentals of optics and microscopy, as well as the use of different types of cameras,

laser-scanning systems, in vivo preparations, and image processing and analysis software. A strong emphasis was placed on building exercises that allowed students to develop an understanding of basic principles while also introducing them to a variety of state-of-the-art commercial systems.

This course was supported with funds provided by the Helmsley Charitable Trust and Howard Hughes Medical Institute.

PARTICIPANTS

Chang, C-J., M.S.E, Massachusetts Institute of Technology, Cambridge
Dellazizzo Toth, T., M.Sc., University of British Columbia, Vancouver, Canada
El-Quessny, M., B.A., University of California, Berkeley
Goncalves, A., M.S., Champalimaud Foundation, Lisbon, Portugal
Huh, C.Y.L., Ph.D., University of California, Irvine
Ibrahim Marosh, L.A., Ph.D., Harvard Medical School, Boston, Massachusetts
Jenks, K., Ph.D., University of Utah, Salt Lake City
Licata, A., B.A., New York University, New York
Lottem, E., Ph.D., Champalimaud Centre for the Unknown, Lisbon, Portugal
McHenry, J., Ph.D., University of North Carolina, Chapel Hill
Mosberger, A., Ph.D., Columbia University, New York, New York
Printz, Y., M.S., Weizmann Institute of Science, Rehovot, Israel
Short, S., Ph.D., University of Utah, Salt Lake City
Weilinger, N., Ph.D., Centre for Brain Health/UBC, Vancouver, British Columbia, Canada
Zwaka, H., Ph.D., Harvard University, Cambridge, Massachusetts

SEMINARS

Aharoni, D., University of California, Los Angeles; and Cai, D. and Shuman, T., Mount Sinai/Icahn School of Medicine, New York, New York; and Golshani, P., University of California, School of Medicine, Los Angeles: Miniscope.
Albeanu, F., Cold Spring Harbor Laboratory: Intrinsic imaging/wide field. Introduction to building homebrew 2P microscope including laser safety.
Behnia, R., Columbia University, New York, New York: Imaging neural circuits in flies.
Denk, W., Max Planck Institute of Neurobiology, Martinsried, Germany: Block-face EM/connectomics.
Dieudonné, S., Ecole Normale Supérieure, Paris, France: AODs and random access multiphoton imaging.
Emiliani, V., CNRS and University Paris Decartes, France: SLMs and holographic photoactivation.
Hillman, E., Columbia University, New York, New York: High-speed volume imaging with SCAPE microscopy.
Ji, N., University of California, Berkeley: Super-resolution imaging, PALM, STED, deep imaging/adaptive optics.
Judkewitz, B., Charité and Humboldt University, Berlin, Germany: 2P microscopy; lasers. Applications of 2P microscopy.
Lichtman, J., Harvard University, Cambridge, Massachusetts: Part 1: Principles and practice of confocal microscopy. Part 2: Applications of confocal and other Imaging approaches to connectomics.
Lin, M., Stanford University, California: Genetically encoded voltage indicators.
Looger, L., HHMI/Janelia Research Campus, Ashburn, Virginia: Latest developments in molecular tools.
Narasimhan, A., Cold Spring Harbor Laboratory: Introduction to morphology.
Orger, M., Champalimaud Foundation, Lisbon, Portugal: Introduction to scanning/PMTs.
Pachitariu, M., University College, London, United Kingdom: Introduction to calcium imaging analysis using Suite2P.
Packer, University of Oxford, United Kingdom: Combining 2P optogenetics and imaging.
Peterka, D., Columbia University, New York: Incoherent emission point source, PSF/Airy pattern, aberrations. Phase contrast and DIC, Fourier/OTF, and introducing SLMs. Fluorescence, Jablonski diagrams, filters.
Smith, S., University of North Carolina, Chapel Hill: Designing and building custom optical systems.
Stringer, C., University College, London, United Kingdom: Introduction to calcium imaging analysis using Suite2P.
Tian, L., University of California, Davis: Genetically encoded tools for neuroscience.
Tsai, P., University of California, San Diego: Nature of light, coherence, wave description, polarization. Ray tracing, lenses. Introduction to Optics Bench Lab. Huygens principle, law of refraction, diffraction, Abbé resolution limit. Köhler illumination, aperture and field control. RECAP: Questions/discussion on scanning and confocal.
Waters, J., Allen Institute for Brain Science, Seattle, Washington: In vivo imaging in rodents/introduction to analyzing calcium imaging data.
Waters, J., Harvard Medical School, Boston, Massachusetts: Noise and detectors.

Yeast Genetics and Genomics

July 24–August 13

INSTRUCTORS **G. Brown,** University of Toronto, Ontario, Canada
G. Lang, Lehigh University, Bethlehem, Pennsylvania
E. Ünal, University of California, Berkeley

ASSISTANTS **J. Chen,** University of California, Berkeley
K. Fisher, Lehigh University, Bethlehem, Pennsylvania
B. Ho, University of Toronto, Canada

This modern and intensive laboratory course teaches students the full repertoire of genetic and genomic approaches needed to dissect complex problems using the yeast *Saccharomyces cerevisiae*. Both classical and modern approaches were emphasized, including the isolation and characterization of mutants, tetrad analysis, and complementation. Synthetic biology was explored through CRISPR-Cas9-directed engineering of heterologous biosynthetic pathways in yeast. Students learned genome-based methods of analysis facilitated by the *Saccharomyces* Genome Database, yeast genome sequences, the gene deletion collection, and other genomic resources available to the community. Molecular genetic techniques, including yeast transformation, gene replacement by PCR, construction and analysis of gene fusions, and generation of mutations, were also emphasized.

Students combined classical approaches with whole-genome sequencing to gain experience in identifying and interpreting genetic interactions, including suppression and synthetic lethality. They performed genome-scale screens using the synthetic genetic array (SGA) methodology. Students were immersed in yeast genomics and performed and interpreted experiments using colony arrays and whole-genome sequencing. Computational methods for data analysis were introduced.

Students gained first-hand experience in modern cytological approaches such as epitope tagging and imaging yeast cells using fluorescence microscopy with GFP–protein fusions and fluorescent indicators for different subcellular structures and organelles. Lectures on fundamental aspects of yeast genetics and genomics were presented along with seminars given by prominent experts in the field on topics of current interest.

This course was supported with funds provided by the National Science Foundation.

PARTICIPANTS

Abdullah, M., B.Sc., Concordia University, Montreal, Quebec, Canada
Arul Nambi Rajan, A.S., B.S., University of California, Davis
Delgado, J., M.S./Ph.D., Dow AgroSciences LLC, Indianapolis, Indiana
Enriquez Hesles, E., B.S., University of Virginia, Charlottesville
Leydon, A., Ph.D., University of Washington, Seattle
Martínez, A., B.S., Universidad Autónoma de México, Querétaro
Mojumdar, A., Ph.D., University of Calgary, Canada
Nunn, C., Ph.D., University of Toronto, Ontario, Canada
Österberg, L., M.S., Chalmers University of Technology, Gothenburg, Sweden
Pires, V., M.S., Johannes Guttenberg University, Mainz, Germany
Robinson, J., B.S., North Carolina Agricultural & Technical State University, Greensboro
Russell, J., Ph.D., Harvard University, Cambridge, Massachusetts
Sanford, E., B.S., Cornell University, Ithaca, New York
Schneider, K., M.S., Sahlgrenska Academy, University of Gothenburg, Göteborg, Sweden
Sterne-Marr, R., Ph.D., Siena College, Loudonville, New York
Zargar, A., Ph.D., University of California, Berkeley/Joint Bioenergy Institute, Emeryville

SEMINARS

Amon, A., Massachusetts Institute of Technology, Cambridge: Causes and consequences of aneuploidy.
Andrews, B., University of Toronto, Ontario, Canada: Mapping biological pathways and networks using yeast genomics.
Arndt, K., University of Pittsburgh, Pennsylvania: Using yeast to understand the role of chromatin in transcription elongation and termination.
Butler, G., University College Dublin, Ireland: Comparative genomics of *Candida* species.
Donaldson, A., University of Aberdeen, United Kingdom: Insights from yeast on eukaryotic chromosome replication.
Frydman, J., Stanford University, California: Understanding protein homeostasis: The complex cellular machinery that maintains a functional proteome.
Jarosz, J., Stanford University, California: Mapping biochemical drivers of phenotypic change.
Klavins, E., University of Washington, Seattle: Yeast synthetic biology.
Nash, R., Stanford University, Palo Alto, California: Navigating data at SGD with YeastMine.
Roth, F., University of Toronto, Ontario, Canada: Next-generation approaches to protein interaction mapping using yeast.
Sherlock, G., Stanford University, California: Fitness effects and trade-offs of beneficial mutation.
Symington, L., Columbia University College of Physicians and Surgeons, New York, New York: Role of the Mre11-Rad50-Xrs2 complex in maintaining genome integrity.
Tu, B., University of Texas Southwestern Medical Center, Dallas: From DNA to RNA to protein to metabolite: Using yeast and metabolomics methods to elucidate metabolic regulatory mechanisms.

Cellular Biology of Addiction (Cambridge, United Kingdom)

July 29–August 5

INSTRUCTORS **D. Belin,** University of Cambridge, United Kingdom
C. Evans, University of California, Los Angeles
B. Kieffer, Douglas Research Centre at McGill University, Montreal, Quebec, Canada

Drug addiction is the most costly neuropsychiatric disorder faced by our nation. Acute and repeated exposure to drugs produces neuroadaptation and long-term memory of the experience, but the cellular and molecular processes involved are only partially understood. The primary objective of this workshop was to provide an intense dialog of the fundamentals, state-of-the-art advances, and major gaps in the cell and molecular biology of drug addiction. Targeted to new or experienced investigators, the workshop combined formal presentations and informal discussions to convey the merits and excitement of cellular and molecular approaches to drug addiction research. With the advent of genomics and proteomics, an extraordinary opportunity now exists to develop comprehensive models of neuroadaptive processes fundamental to addiction, withdrawal, craving, and relapse to drug use and to brain function, in general. A range of disciplines and topics were represented, including noninvasive brain imaging to identify drug targets and adaptive processes; neuroadaptive processes at the molecular and cellular level; neural networks and their modulation; the relevance of genotype to susceptibility and drug response; tolerance and adaptation at the cellular level; and approaches to exploiting the daunting volume generated by neuroinformatics. This workshop provided an integrated view of current and novel research on neuroadaptive responses to addiction, fostered discussion on collaboration and integration, and provided critical information needed to construct a model of addiction as a disease and novel molecular targets for biological treatments. Beyond the plane of scientific endeavor, the information is vital for formulating public

policy and for enlightening the public on the neurobiological consequences of drug use and addiction. The workshop was designed to generate interest in this level of analysis, open conduits for collaborations, and present novel routes to investigating the neurobiology of addictive drugs.

This course was supported with funds provided by the U.S. National Institute of Drug Abuse.

PARTICIPANTS

Beldjoud, H., M.S., National Institute on Drug Abuse, Baltimore, Maryland
Berger, U., B.A., Yale University, New Haven, Connecticut
Campbell, R., B.S., University of California, Irvine
Ferragud, A., B.S., University of Cambridge, United Kingdom
Foster, S., B.S., Emory University, Atlanta, Georgia
Grigsby, K., B.S., University of Missouri, Columbia
Hart, E., B.A., University of California, Los Angeles
Jadhav, K., M.B.B.S., University of Lausanne, Prilly, Switzerland
Jones, J., B.Sc., University of Cambridge, United Kingdom
Moussawi, K., B.S., National Institute on Drug Abuse, Baltimore, Maryland
Sketriene, D., B.S., The University of Melbourne, New South Wales, Australia
Solis, E., B.S., National Institutes of Health/National Institute on Drug Abuse-Intramural Research Program, Baltimore, Maryland
Tzortzi, O., B.Sc., Karolinska Institutet, Stockholm, Sweden
Velazquez C., B.S., University of Cambridge, United Kingdom
Ward, L., B.A., University of Cambridge, United Kingdom
Wherry, J., B.A., State University of New York, Binghamton, New York
Zipperly, M., B.S., University of Alabama, Birmingham

SEMINARS

Belin, D., University of Cambridge, United Kingdom: Neural substrates of the interindividual vulnerability to develop compulsive drug-seeking habits.
Boutrel, B., Lausanne University Hospital, Switzerland: How homeostatic and hedonic regulations of feeding may concur to drive maladaptive pattern of food intake.
Bruchas, M., Washington University School of Medicine in St. Louis, Missouri: Dissecting neuromodulatory circuits and signaling in affective behavior.
Cheer, J., University of Maryland School of Medicine, Baltimore: Endogenous cannabinoids and the pursuit of reward.
Evans, C., University of California, Los Angeles: Opioids in pain and addiction.
Everitt, B., University of Cambridge, United Kingdom: Memory and addiction.
Kenny, P., Icahn School of Medicine at Mount Sinai, New York, New York: The habenula links nicotine addiction to tobacco-related diseases.
Kieffer, B., Douglas Research Centre at McGill University, Montreal, Quebec, Canada: Opioid receptors in addiction.
Kober, H., Yale University, New Haven, Connecticut: Human neuroscience of addiction.
Koob, G., National Institute on Alcohol Abuse and Alcoholism, Rockville, Maryland: Hedonic allostasis/stress.
Manglik, A., University of California, San Francisco: Structural dynamics of drug action.
Marisela Morales, F., National Institute on Drug Abuse, Baltimore, Maryland: Neuronal diversity and cotransmission.
Mason, B., The Scripps Research Institute, La Jolla, California: Workshop on bringing basic science to the clinic.
Otis, J., Medical University of South Carolina, Charleston: Innovative technologies for understanding the neural circuitry of addictive behaviors.
Picciotto, M., Yale University, Guilford, Connecticut: Molecular basis of nicotine addiction.
Robbins, T., University of Cambridge, United Kingdom: Impulsivity, compulsivity, and addiction: Neural basis.
Sommer, W., Central Institute of Mental Health (ZI)/ University of Heidelberg, Germany: Losing control: Alcohol, glutamate, and the prefrontal cortex.
von Zastrow, M., University of California, San Francisco: Drug actions viewed from a lonely neuron's perspective.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics, reproducibility, and rigor.

Genetics and Neurobiology of Language

July 30–August 5

INSTRUCTORS **S. Fisher,** Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands
D. Poeppel, New York University, New York
K. Watkins, Oxford University, United Kingdom

ASSISTANT **C. Wiltshire,** University of Oxford, United Kingdom

Why are children able to acquire highly sophisticated language abilities without needing to be taught? What are the neurobiological and neurophysiological processes that underpin human speech and language, and how do they go awry in developmental and acquired disorders? Which genetic factors contribute to this remarkable suite of human skills, and are there evolutionary precursors that we can study in other species? Can we trace connections between language skills and musicality? This unique CSHL course, in its third iteration, addressed these core questions about the bases and origins of speech and language, through talks, interactive sessions, keynotes, and debates, involving leading experts from a range of disciplines. It integrated the state of the art from complementary perspectives, including development, cognitive models, neural basis, gene identification, functional genomics, model systems, and comparative/evolutionary studies.

This course was supported with funds provided by the Nancy Lurie Marks Family Foundation.

PARTICIPANTS

Bajracharya, B., B.A., Washington University in St. Louis, Missouri
Blasi, D., Ph.D., University of Zürich/Max Planck Institute, Switzerland
Braden, R., M.A., Murdoch Children's Research Institute, Parkville, Victoria, Australia
Bradshaw, A., B.A., University of Oxford, United Kingdom
Bruce, L., M.S., Arizona State University, Tempe
Choi, D., M.S., The University of British Columbia, Vancouver, Canada
Cler, G., Ph.D., Boston University, Massachusetts
Eichert, N., B.Sc., University of Oxford, United Kingdom
Fazekas, J., B.A., University of Liverpool, United Kingdom
Hafri, A., B.A., University of Pennsylvania, Philadelphia
Haiduk, F., M.S., University of Vienna, Austria
Healey, M., Ph.D., University of Pennsylvania, Philadelphia
Ivanova, A., B.S., Massachusetts Institute of Technology, Cambridge
Jackson, V., Ph.D., Walter and Eliza Hall Institute, Parkville, Victoria, Australia
Kabdebon, C., Ph.D., Yale University, New Haven, Connecticut
Kepinska, O., Ph.D., University of California, San Francisco
Kim, J., Ph.D., Johns Hopkins University, Baltimore, Maryland
Mai, A., B.A., University of California, San Diego
Norton, P., Ph.D., Freie Universität Berlin, Germany
Patel, S., M.A., CUNY Graduate Center, New York
Tulling, M., M.A., New York University, New York
Whitten, A., B.A., Vanderbilt University, Nashville, Tennessee
Zhu, J., B.E., Macquarie University, Sydney, New South Wales, Australia

SEMINARS

Christophe, A., École Normale Supérieure, Paris, France: Synergies in early lexical and syntactic acquisition.
De Diego-Balaguer, R., University of Barcelona, Spain: Language in interaction: Recruiting working memory and attention for language learning.
Emmorey, K., San Diego State University, California: The neurobiology of language from the perspective of sign languages.
Fedorenko, E., Harvard Medical School/Massachusetts General Hospital/MIT, Belmont: The language network within the broader context of the human mind and brain.
Fisher, S., Max Planck Institute for Psycholinguistics, Netherlands: Translating the genome in human neuroscience.
Fitch, T., University of Vienna, Austria: The biology and evolution of language: A comparative and computational perspective.
Gordon, R., Vanderbilt University Medical Center, Nashville, Tennessee: The biology of musical rhythm and what it means for language development.
Jarvis, E., The Rockefeller University, New York, New York: Insight from nonhuman animals into the neurobiology of human language.
Lau, E., University of Maryland, College Park: Neural encoding of linguistic structure: Multimodal neuroimaging approaches.
Mac Sweeney, M., University College, London, United Kingdom: The impact of deafness on the neurobiology of spoken language.
Mahon, B., University of Rochester, New York: The neural representation of object concepts.
Morgan, A., Murdoch Children's Research Institute, Melbourne, Victoria, Australia: Genetics in the speech and language clinic: Phenotypes and genotypes.
Newbury, D., University of Oxford, United Kingdom: The role of common genetic variation in speech and language disorders.
Peelle, J., Washington University in St. Louis, Missouri: The role of cortical oscillations in auditory and audiovisual speech perception.
Poeppel, D., New York University, New York: How can we develop linking hypotheses between genetics, computational neuroscience, and language research?
Scharff, C., Freie Universität Berlin, Germany: Vocal learning: What is it, who has it, and why is it relevant to language?
Vernes, S., Max Planck Institute for Psycholinguistics, Netherlands: The building blocks of language: Genes, bats, and beyond.
Watkins, K., Oxford University, United Kingdom: Language (re)organization in the congenitally blind brain.

Brain Tumors

August 7–13

INSTRUCTORS **P. Brastianos,** Massachusetts General Hospital, Boston
M. Gilbert, National Institutes of Health, Bethesda, Maryland
S. Majumder, University of Texas MD Anderson Cancer Center, Houston

ASSISTANT **J. Zhang,** Columbia University Medical Center, New York, New York

This 1-week discussion course provided a clinical overview of brain tumors and emphasized molecular mechanisms involved in the growth and development of brain tumors with special emphasis on neural differentiation, signaling mechanisms, DNA replication, chromatin modulation, stem cells, mouse models, genomics, imaging techniques, genetically modified mouse techniques, nanotechnology, mechanism-based therapeutic strategies, and biobanks and ethical concerns. Attendees were able to interact with senior investigators on a one-to-one basis in an informal environment. All non-faculty students were considered for a generous stipend depending on stated need.

This course was supported with funds provided by the American Brain Tumor Association.

PARTICIPANTS

Alnahhas, I., M.D., Virginia Commonwealth University, Richmond
Ammoun, S., M.Sc., Plymouth University Peninsula Schools of Medicine, United Kingdom
Andersson, P., B.S., Harvard Medical School/Massachusetts General Hospital, Boston
Ayo, A., B.Sc., Research Programs Unit, Helsinki, Finland
Chatterjee, S., B.Sc., Massachusetts General Hospital/Harvard Medical School, Boston
Ellermann Jensen, K., B.Sc., Danish Cancer Society, Copenhagen, Denmark

Fan, J., B.Sc., University of Toronto, Ontario, Canada
Garcia Diaz, C., B.S., M.S., University College London, United Kingdom
Goenka, A., B.A., Northwell Health, Lake Success, New York
Hicks, M., B.A., Monmouth University, West Long Branch, New Jersey
Hwang, L., B.S., National Institutes of Health, Bethesda, Maryland
Kaushik, S., B.S., M.S., University of California, San Francisco
Krenciute, G., B.S., St. Jude Children's Research Hospital, Memphis, Tennessee
Liu, J., B.S., St. Jude Children's Research Hospital, Memphis, Tennessee
Merisaari, J., M.Sc., Turku Centre for Biotechnology, Finland
Pantazopoulou, V., B.S., Lund University, Sweden
Pianka, S., B.A., University of California, Los Angeles
Reynolds, P., B.S., Mayo Clinic Graduate School of Biomedical Sciences, Rochester, Minnesota
Shen, C., B.S., Northwestern University, Chicago, Illinois
Shuboni-Mulligan, D., B.A., National Institutes of Health, Bethesda, Maryland
Stine, C., B.S., Virginia Tech, Blacksburg
Zalenski, A., B.S., Ohio State University, Columbus

SEMINARS

Bachoo, R., University of Texas Southwestern Medical Center, Dallas: Adult brain tumors: Mechanisms, metabolism, imaging, viro-, and immunotherapy.
Bondy, M., Baylor College of Medicine, Houston, Texas: Clinical aspects of brain tumors.
Brastianos, P., Massachusetts General Hospital, Boston: Stem cells, chromatin, single-cell analysis, modeling, brain metastases.
Brennan, C., Memorial Sloan Kettering Cancer Center, New York, New York: Big data and clinical trials.
Cahill, D., Harvard/Massachusetts General Hospital, Boston: Pediatric brain tumors, IDH mutation.
Chiocca, E.T., Brigham and Women's Hospital, Boston, Massachusetts: Adult brain tumors: Mechanisms, metabolism, imaging, viro-, and immunotherapy.
Dirks, P., Hospital for Sick Children/University of Toronto, Ontario, Canada: Stem cells, chromatin, single-cell analysis, modeling, brain metastases.
Fuller, G., MD Anderson Cancer Center, Houston, Texas: Clinical aspects of brain tumors.
Furnari, F., University of California, San Diego: Adult brain tumors: Mechanisms, metabolism, imaging, viro-, and immunotherapy.
Gilbert, M., National Institutes of Health, Bethesda, Maryland: Big data and clinical trials.
Gladson, C., Cleveland Clinic, Ohio: Normal development and disease, angiogenesis and cell death, big data.
Gutmann, D., Washington University School of Medicine in St. Louis, Missouri: Normal development and disease, angiogenesis and cell death, big data.
Holland, E., Fred Hutchinson Cancer Research Center, Seattle, Washington: Normal development and disease, angiogenesis and cell death, big data.
Iavarone, A., Columbia University Medical Center, New York: Big data and clinical trials.
James, C.D., Feinberg School of Medicine, Northwestern University, Chicago, Illinois: Stem cells, chromatin, single-cell analysis, modeling, brain metastases.
Lasorella, A., Columbia University, New York, New York: Stem cells, chromatin, single-cell analysis, modeling, brain metastases.
Majumder, University of Texas MD Anderson Cancer Center, Houston: Normal development and disease, angiogenesis and cell death, big data.
Pieper, R., University of California, San Francisco: Adult brain tumors: Mechanisms, metabolism, imaging, viro-, and immunotherapy.
Reilly, K., National Cancer Institute, Bethesda, Maryland: Clinical aspects of brain tumors.
Sawaya, R., MD Anderson Cancer Center, Houston, Texas: Clinical aspects of brain tumors.
Snyder, E., Sanford Burnham, La Jolla, California: Pediatric brain tumors, IDH mutation.
Verhaak, R., MD Anderson Cancer Center, Houston, Texas: Big data and clinical trials.
Wechsler-Reya, R., Sanford Burnham Prebys Medical Discovery Institute, La Jolla, California: Pediatric brain tumors, IDH mutation.
Weiss, W., University of California, San Francisco: Pediatric brain tumors, IDH mutation.
Wen, P., Dana-Farber Cancer Institute, Boston, Massachusetts: Clinical aspects of brain tumors.

Proteomics

August 7–21

INSTRUCTORS	**G. Knudsen,** University of California, San Francisco **K. Medzihradszky,** University of California, San Francisco **D. Pappin,** Cold Spring Harbor Laboratory
PART-TIME INSTRUCTORS	**E. Soderblom,** Duke University, Durham, North Carolina **W. Thompson,** Duke Center for Genomic and Computational Biology, Durham, North Carolina
ASSISTANTS	**A. Greninger,** University of Washington, Seattle **E. Klement,** Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary **K. Rivera,** Cold Spring Harbor Laboratory

This intensive laboratory and lecture course focused on cutting-edge proteomic approaches and technologies. Students gained practical experience isolating, purifying, and identifying protein complexes: Sample preparation with in-solution digestion was performed; students were then trained using high-sensitivity nano-HPLC coupled with nanospray-ESI and tandem mass spectrometry analysis. Different search engines and bioinformatic approaches were introduced for data evaluation. Students were shown how to recognize unexpected posttranslational modifications. Diverse techniques for PTM peptide enrichment, including affinity chromatography for phosphopeptides and immunoenrichment of GlyGly-Lys, methyl-Arg, and phospho-Tyrpeptides, and the characterization of the resulting complex mixtures, including site assignments, were performed. For shotgun proteomic analysis sections, students used label-free and covalent isotopic-labeling quantitative approaches to profile changes in protein complexes and whole proteomes. In a section focused on targeted proteomics, students learned to analyze and process shotgun proteomic data for the development of

SRM/PRM assays that accurately identify and quantify targeted proteins. Students designed transitions for selected peptides and performed SRM/PRM analyses. They learned to process and interpret the acquired data to measure changing quantities of targeted peptides in a variety of biological samples and specifically spent significant time using Skyline for both MS1 and MS2 data analyses. For all sections of the course, a strong emphasis was placed on data analysis. There was opportunity to discuss and provide feedback on individual research projects, and students had the opportunity to learn to process their own data (acquired outside the course) in Skyline if so desired.

An industrial lecture series was delivered by drug discovery scientists. The students received in-depth knowledge about chemoproteomics techniques routinely used in industry and how they are used to profile compounds and potential protein targets. This was followed by lab work on in-lysate affinity enrichment techniques and a deep dive into data analysis. This broadened students' vision toward chemoproteomics applications in drug discovery programs.

A series of outside lecturers discussed various proteomics topics, including de novo sequence analysis, intact protein analysis, advanced mass spectrometry methods, glycosylation, and functional proteomics. The aim of the course was to provide each student with the fundamental knowledge and hands-on experience necessary for performing and analyzing proteomic experiments. The overall goal was to train students to identify new opportunities and applications for proteomic approaches in their biological research.

This course was supported with funds provided by the National Institute of Child Health and Human Development.

PARTICIPANTS

Aguilar, E., B.S., The Rockefeller University, New York, New York
Deol, K., B.S., University of Massachusetts, Amherst
Doddihal, V., B.S./M.S., Graduate School of the Stowers Institute, Kansas City, Missouri
Foschepoth, D., Ph.D., Technical University of Delft, Delft, Netherlands
Haq, M., B.S., East Tennessee State University, Johnson City
Khanal, O., Ph.D., University of Delaware, Newark
Lao, S., M.A., Amgen Inc., Cambridge, Massachusetts
Lorenz-Guertin, J., B.S., University of Pittsburgh, Pennsylvania
Ma, J., M.D., Virginia Commonwealth University, Richmond
Mattila, M., Ph.Lic., Medix Biochemica, Espoo, Finland
Mohler, K., Ph.D., Yale University, West Haven, Connecticut
Mokry, R., B.S./B.A., Medical College of Wisconsin, Milwaukee
Peterson, B., B.S., University of Michigan Medical School, Ann Arbor
Pham, H., B.S./M.S., Celgene Corporation, Cambridge, Massachusetts
To, M., B.A., University of California, Berkeley
Younis, S., Ph.D., Uppsala University, Sweden

SEMINARS

Anderson, L., NHMFL, Florida State University, Tallahassee: High-throughput intact protein characterization.
Clauser, K., Broad Institute of MIT and Harvard, Cambridge, Massachusetts: Manual de novo peptide MS/MS interpretation for evaluating database search results.
Farnsworth, C., Cell Signaling Technology, Danvers, Massachusetts: PTMScan® technology: An antibody-based proteomics discovery platform.
Jacob, R., Matrix Science Inc, Boston, Massachusetts: MS/MS-based protein identifications and the fine tuning of the search engines.
Kawatkar, A., Tomlinson, R., Hendricks, A., and Zhang, A., AstraZeneca, Waltham, Massachusetts: Introduction of in-lysate chemoproteomics and work flow. Industrial case study.
Klement, E., BRC of the Hungarian Academy of Sciences, Szeged, Hungary: Phosphopeptide-enrichment using metal-ion affinity chromatography.
Knudsen, G., University of California, San Francisco: On the isolation of protein complexes.
Medzihradszky, K., University of California, San Francisco: How to search for PTMs; how to interpret and evaluate MS/MS data of PTM-bearing peptides.
Pappin, D., Cold Spring Harbor Laboratory: Fundamentals of mass spectrometry for proteomics.
Thompson, W., Duke Center, Durham, North Carolina: Label-free quant, AUC vs. spectral counting, tools for data interpretation.
Trnka, M., University of California, San Francisco: Cross-linking mass spectrometry for exploring the structure and interactions of protein complexes.

Programming for Biology

October 15–30

INSTRUCTORS **S. Prochnik,** DOE/Joint Genome Institute, Walnut Creek, California
S. Robb, Stowers Institute for Medical Research, Kansas City, Missouri

ASSISTANTS **J. Bredeson,** University of California, Berkeley
K. Gotting, The University of Wisconsin, Madison
L. Kuderna, Institut de Biologia Evolutiva, Barcelona, Spain
J. Orkin, University of Calgary, Alberta, Canada
S. Webb, University of Calgary, Alberta, Canada

More often than not, today's biologist is studying data that is too complex or numerous to be analyzed without a computer and only boilerplate analysis can be performed with existing tools. Questions specific to the data set require novel analysis pipelines to be designed and written in computer code. Designed for lab biologists with little or no programming experience, this course gave students the bioinformatics and scripting skills necessary to exploit this abundance of biological data. The only prerequisite for the course was a strong commitment to learning basic UNIX and a scripting language.

This year, we offered the course in Python, an easy-to-learn scripting language with a growing code base and community of users. The course began with one week of introductory coding, continued with practical topics in bioinformatics—with plenty of coding examples, and ended with a group coding project. Formal instruction was provided on every topic by the instructors, teaching assistants, and invited experts. Students solved problem sets covering common scenarios in the

acquisition, validation, analysis, and visualization of biological data. They learned how to design, construct, and run powerful and extensible analysis pipelines in a straightforward manner. Final group projects were chosen from ideas proposed by students and were guided by faculty. Students were provided with a library of Python reference print and e-books that they were able to bring home with them.

Note that the primary focus of this course was to provide students with practical programming experience rather than to present a detailed description of the algorithms used in computational biology.

This course was supported with funds provided by the National Human Genome Research Institute. Access to cloud computational resources may be supported by an education grant from Amazon Web Services.

PARTICIPANTS

Alves, C., Ph.D., Cold Spring Harbor Laboratory
Bhargava, R., Ph.D., City of Hope, Duarte, California
Chaves, J., Ph.D., Oak Ridge National Laboratory, Tennessee
Chin, A., Ph.D., Institut de Recherches Cliniques de Montréal, Quebec, Canada
Feuerman, M., Ph.D., SUNY Downstate Medical Center, New York, New York
Forst J., Ph.D., University of California, Santa Cruz
Friedland, S., M.D./Ph.D., University of Rochester School of Medicine and Dentistry, New York
Gorsi, B., Ph.D., University of Utah, Salt Lake City
Hill, E., Ph.D., Stowers Institute for Medical Research, Kansas City, Missouri
Hodson, C., M.Sc., University of Edinburgh, United Kingdom
Jean, F., Ph.D., University of Alberta, Canada
Jibrin, M., Ph.D., University of Florida, Gainesville
Lam, S., Ph.D., Massachusetts General Hospital, Boston
Lewandowska, M., Ph.D., M.S., Hebrew University of Jerusalem, Israel
Miller, E., M.Sc., King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
Mouftah Ali. S., M.S., Faculty of Medicine and Health Science of the United Arab Emirates University
Moustafa, A., Ph.D., ARC, Children's Hospital of Philadelphia, Pennsylvania
Paffenholz, S., Ph.D., Memorial Sloan Kettering Cancer Center, New York, New York
Spaulding, C., Ph.D., Harvard School of Public Health, Boston, Massachusetts
Thies, A., B.S., Donald Danforth Plant Science Center, Creve Coeur, Missouri
Waas, M., Ph.D., Medical College of Wisconsin, Milwaukee
Wong, M., B.S., Cold Spring Harbor Laboratory

SEMINARS

Cain, S., Ontario Institute for Cancer Research, Toronto, Canada: Bioinformatics tools overview II: GMOD and JBrowse.
Campbell, M., Cold Spring Harbor Laboratory: Next-generation sequencing: File formats: fasta, fastq, bam, vcf.
Eilbeck, K., University of Utah, Salt Lake City: Ontologies: Controlled vocabulary.
Haas, B., Broad Institute, Northbridge, Massachusetts: RNA-Seq and transcript assembly.
Marques-Bonet T., and Kuderna, L., Institut Biologia Evolutiva, Barcelona, Spain: Structural variation.
Pearson, W., University of Virginia, Charlottesville: Sequence similarity I: Sequence similarity search I. Sequence similarity II: Sequence similarity search I: Alignment data problems.
Prochnik, S., Joint Genome Institute, Walnut Creek, California: Unix 1: Unix overview, the basics, advanced Unix, Unix cheat sheet. Python I: Python1 overview, running python, PyCharm, syntax, data types, and variables. Python III: Sequences, strings, lists, tuples. Python V: Iterables, I/O, and files. Python VII: Functions, scope, modules. Bioinformatics tools overview I: What's coming; define bioinformatics concepts; gene prediction, alignment, searching, and assembly. Python IX: Biopython.
Robb, S., Stowers Institute for Medical Research, Kansas City, Missouri: Unix 2: Text editors, git for beginners. Python II: Operators, truth, logic, numbers. Python IV: Loops, dictionaries, sets. Python VI: Regular expressions. Python VIII: Exceptions, data structures.
Triant, D., University of Virginia, Charlottesville: Genome assembly.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics and reproducibility in research.

X-Ray Methods in Structural Biology

October 15–30

INSTRUCTORS **P. Adams**, Lawrence Berkeley Laboratory, Berkeley, California
W. Furey, University of Pittsburgh, Pennsylvania
A. McPherson, University of California, Irvine
A. Perrakis, Netherlands Cancer Institute, Amsterdam
J. Pflugrath, Rigaku Americas, The Woodlands, Texas

ASSISTANT **M. Whitley,** University of Pittsburgh, Pennsylvania

X-ray crystallography has been the cornerstone of structural biology for half a century. This intense laboratory/computational course focused on the major techniques used to determine the three-dimensional structures of macromolecules. It was designed for scientists with a working knowledge of protein structure and function, but who were new to macromolecular crystallography. Topics that were covered included basic diffraction theory, crystallization (proteins, nucleic acids, complexes, and membrane proteins), synchrotron X-ray sources and optics, data collection and processing, structure solution by experimental phasing methods (SAD, MAD, MIR, and others) and molecular replacement, electron density maps improvement (solvent flattening, noncrystallographic averaging, etc.), model building and refinement, structure validation, coordinate deposition, and structure presentation. In addition, for the first time, the course extended to the theory and computation for small angle X-ray scattering (SAXS) and single-particle cryoelectron microscopy.

Participants learned through extensive hands-on experiments in fully equipped labs, crystallized multiple proteins, and determined their crystal structures by several methods while learning through extensive lectures on theory. Informal discussions behind the techniques were frequent, and students were responsible also for collecting questions to be answered in specific sessions.

This course was supported with funds provided by the National Institute of General Medical Sciences.

PARTICIPANTS

Candido Primi, M., B.S., The Scripps Institute, Jupiter, Florida
Fan, C., B.S., California Institute of Technology, Pasadena
Hatti, K., Ph.D., Cambridge Institute for Medical Research, United Kingdom
Isiorho, E., Ph.D., Auburn University, Alabama
Jao, C., Ph.D., Genentech, San Francisco, California
Kudryashov,D., Ph.D., Ohio State University, Columbus
Lyons, B., B.S., University of British Columbia, Vancouver, Canada
Otero, L., Ph.D., Leloir Institute Foundation, Buenos Aires, Argentina
Salustros, N., M.S., University of Copenhagen, Denmark
Shukla, S., M.Sc., University of Tennessee, Knoxville
Stevens, T., B.S., California Institute of Technology, Pasadena
Tamadonfar, K., B.S., Washington University in St. Louis, Missouri
Vo, J., B.S., La Trobe Institute for Molecular Sciences, Thomastown, Australia
Wei, K., Ph.D., University of California, Berkeley
Whitaker, A., Ph.D., University of Kansas Medical Center, Kansas City
Zhao, S., B.S., University of California, Irvine

SEMINARS

Adams P., Lawrence Berkeley Laboratory, Berkeley, California: Structure refinement: What is it? Low-resolution structure refinement.
Borek, D., University of Texas Southwestern Medical Center, Dallas: X-ray data processing, scaling, and merging. Anisotropy, order–disorder, and radiation damage in practice.
Caffrey, M., Trinity College, Dublin, Ireland: Membrane protein crystallization.
Clemons, B., California Institute of Technology, Pasadena: Routes toward crystallizing membrane proteins.
Conway, J., University of Pittsburgh, Pennsylvania: Introduction to cryo-EM.
Emsley, P., MRC Laboratory of Molecular Biology, Cambridge, United Kingdom: Map interpretation and model building. Advanced Coot.
Furey, W., University of Pittsburgh, Pennsylvania: Introduction to the phase problem, phasing methods, and Patterson space. Basic phasing theory: Isomorphous replacement and anomalous scattering. Phase improvement by solvent flattening and NCS averaging. Direct methods introduction.
Gilliland, G., Janssen Research & Development, Spring House, Pennsylvania: Maximizing crystallization success through seeding.
Hendrickson, W., Columbia University, New York, New York: MAD and SAD phasing.
Holton, J., University of California, San Francisco: Beamline basics/data collection strategies at the beamline. Radiation damage.
Kleywegt, G., European Bioinformatics Institute, Cambridge, United Kingdom: Structure validation.
McPherson, A., University of California, Irvine: Crystallization of macromolecules: Basic theory. Symmetry, periodicity, unit cells, space groups, Miller planes, and lattices. Waves, vectors, and complex numbers. Fundamental diffraction relationships and Bragg's Law, diffraction pattern formation, reciprocal space. Fourier transforms, electron density.
Newman, J., CSIRO, Parkville, Australia: Crystallization practice. Modern approaches toward crystallization.
Perrakis, A., Netherlands Cancer Institute, Amsterdam, the Netherlands: Overview of crystallography. Automated model building in ARP/wARP. Water molecules and side chains: Decision making. Automated data processing and structure solution in ISYS/CCP4i. Automated model rebuilding with PDB_REDO. The CSHL picture competition and software suggestions for pictures and movies.
Pflugrath, J., Rigaku Americas, The Woodlands, Texas: Cryo-crystallography. Data collection and processing.
Read, R., University of Cambridge, United Kingdom: Molecular replacement: How does it work? Molecular replacement: Making models.
Richardson, J., Duke University, Durham, North Carolina: Presentation of structures: History and perspectives. Validation with MolProbity. Cold Spring Harbor validation horror show. The CSHL picture competition and software suggestions for pictures and movies.
Smith, C., Stanford University, California: Serial femtosecond crystallography methods at XFELS and synchrotrons.
Sweet, R., Brookhaven National Laboratory, Shoreham, New York: I. Explaining the Ewald Sphere and its value. II. X-ray sources and detectors, and their properties.
Terwilliger, T., Los Alamos National Laboratory, New Mexico: Automated experimental phasing. Automated model building and other things.
Thorn, A., University of Hamburg, Germany: Twinning in crystallography. Assessing data quality. Experimental phasing with SHELXC/D/E.
Tronrud, D., The Light and Electron Works, Springfield, Oregon: Electron density maps/would you publish it?
Williams, C., Duke University, Durham, North Carolina: Validation with MolProbity. Cold Spring Harbor validation horror show.

Advanced Sequencing Technologies and Applications

November 6–18

INSTRUCTORS

M. Griffith, Washington University School of Medicine in St. Louis, Missouri
O. Griffith, Washington University School of Medicine in St. Louis, Missouri
E. Mardis, Nationwide Children's Hospital Research Institute, Columbus, Ohio
W. McCombie, Cold Spring Harbor Laboratory
A. Quinlan, University of Utah, Salt Lake City

ASSISTANTS

J. Belyeu, University of Utah, Salt Lake City
K. Cotto, Washington University in St. Louis, Missouri
A. Farrell, University of Utah, Salt Lake City
S. Goodwin, Cold Spring Harbor Laboratory
S. Kravitz, University of Utah, Salt Lake City
J. Preall, Cold Spring Harbor Laboratory
C. Regan, Cold Spring Harbor Laboratory
A. Wagner, Washington University in St. Louis, Missouri
J. Walker, Washington University School of Medicine in St. Louis, Missouri
R. Wappel, Cold Spring Harbor Laboratory
A. Ward, University of Utah, Salt Lake City

Over the last decade, massively parallel DNA sequencing has markedly impacted the practice of modern biology and is being utilized in the practice of medicine. The constant improvement of these platforms means that costs and data generation timelines have been reduced by orders of magnitude, enabling investigators to conceptualize and perform sequencing-based projects that heretofore were time-, cost-, and sample number–prohibitive. Furthermore, the application of these technologies to answer questions previously not experimentally approachable is broadening

their impact and application. However, data analysis remains a complex and often vexing challenge, especially as data volumes increase.

This intensive 2-week course explored the use and applications of massively parallel sequencing technologies, with a focus on data analysis and bioinformatics. Students were instructed in the detailed operation of several platforms, including library construction procedures, general data processing, and in-depth data analysis. A diverse range of the types of biological questions enabled by massively parallel sequencing technologies were explored, including DNA resequencing of known cancer genes, de novo DNA sequencing, and assembly of genomes, RNA sequencing, and others that were tailored to the student's research areas of interest.

Cloud-based computing was also explored. Guest lecturers highlighted unique applications of these disruptive technologies. We encouraged applicants from a diversity of scientific backgrounds, including molecular evolution, development, neuroscience, medicine, cancer, plant biology, and microbiology.

This course was supported with funds provided by the National Human Genome Research Institute. Access to cloud computational resources may be supported by an AWS in Education Grant award from Amazon.

PARTICIPANTS

Cadoo, K., M.D., Memorial Sloan Kettering Cancer Center, New York, New York
Carlo, M., B.A., Memorial Sloan Kettering Cancer Center, New York, New York
Donat, P., Stony Brook University, New York
Garcia, G., B.S., National Institutes of Health, Bethesda, Maryland
Gervasi, M., B.S., University of Massachusetts, Amherst
Graf Grachet, N., B.S., Oklahoma State University, Stillwater
Gupta, V., B.S., Emory University, Atlanta, Georgia
Kilfeather, P., BV.Sc., University of Oxford, United Kingdom
Kim, S., B.Sc., Western University, London, Ontario, Canada
Kumon, T., B.A., University of Pennsylvania, Philadelphia
Layman, D., B.S., Estée Lauder Companies Inc., New York, New York
Liu, C., B.S., University of California, San Diego
Malik, A., B.S., University of Michigan, Ann Arbor
Mouftah Ali, S., B.A., Faculty of Medicine and Health Science of the United Arab Emirates University
Rodriguez Terrones, D., B.S., Helmholtz Zentrum München, Germany
Sexton, A., B.S., Yale University, New Haven, Connecticut
Shi, Z., University of Notre Dame, Indiana
Simonetti, F., M.D., Johns Hopkins University, Baltimore, Maryland
Sun, Q., University of Illinois, Urbana-Champaign
Wu, X., B.S., Cold Spring Harbor Laboratory
Xu, T., B.S., Rosalind Franklin University, North Chicago, Illinois

SEMINARS

Chakravarti, A., New York University Langone, New York: Genetic regulatory control of cardiac diseases.
Chakravarty, D., Memorial Sloan Kettering Cancer Center, New York, New York: Identifying driver alterations and therapeutic options in cancer.
Dewar, K., McGill University/Genome QC Innovation Centre, Montreal, Quebec, Canada: Introduction to genome assembly of PacBio + Illumina data.
Farrell, A., University of Utah, Salt Lake City: Variant discovery lecture.
Griffith, M., Washington University School of Medicine in St. Louis, Missouri: Introduction to RNA-sequencing lecture.
Griffith, O., Washington University School of Medicine in St. Louis, Missouri: Alignment and visualization lecture.
Haas, B., Broad Institute, Northbridge, Massachusetts: Transcript assembly.
Khurana, E., Weill Cornell Medicine, New York, New York: Functional interpretation of noncoding sequence variants.
Kravitz, S., University of Utah, Salt Lake City: Monoallelic expression of human genes in diverse tissues.
Lappalainen, T., New York Genome Center & Columbia University, New York: Gene regulation, allelic expression, QC from GTeX.

Maher, C., Washington University School of Medicine in St. Louis, Missouri: Dissecting the RNA interactome with NGS.
Mardis, E., Nationwide Children's Hospital, Columbus, Ohio: Overview of next-generation sequencing technologies.
Morrissy, S., The Governors of the University of Calgary, Alberta, Canada: Introduction to NGS data analysis.
Pickrell, J., New York Genome Center, New York: Variant calling and disease risk prediction from low-pass sequencing data.
Preall, J., Cold Spring Harbor Laboratory: Single-cell sequencing technology and applications.
Quinlan, A., University of Utah, Salt Lake City: Disease variant discovery lecture.
Quinlan, A., University of Utah, Salt Lake City: Genome arithmetic with BEDTOOLS. Structural variation. Probability and statistics for genomics analysis.
Walker, J., Washington University School of Medicine in St. Louis, Missouri: RNA-Seq, expression and differential expression.
Witkowski, J., Cold Spring Harbor Laboratory: Ethics in genetics and genomics research.

Scientific Writing Retreat

November 14–18

INSTRUCTORS
C. Lambert, Cold Spring Harbor Laboratory
S. Matheson, *Cell Reports*, Cambridge, Massachusetts

WRITING COACHES
L. Connell, *Molecular Case Studies*, Senior Editor, Cold Spring Harbor Laboratory Press
S. Gary, Cold Spring Harbor Laboratory
E. Gaskell, Broad Institute of Harvard/MIT, Somerville, Massachusetts
J. Jansen, Cold Spring Harbor Laboratory
B. Plosky, Molecular Cell, Cambridge, Massachusetts
J. Rubin, Columbia University, New York, New York
J. Schaffhausen, Yale University, New Haven, Connecticut

The goal of this retreat was to have participants progress significantly on writing projects while improving their professional communication skills. The retreat included a mix of formal sessions and less structured writing time. The formal sessions covered publication writing for scientific journals from the perspectives of *Cell Press* and Cold Spring Harbor Laboratory Press; writing clearly and conversationally about your research in ways that engage diverse audiences, a skill particularly useful when developing lay summaries for NIH and NSF proposals; and style tips and considerations for clear professional writing in all forms.

The less-structured sessions of the retreat included small writing groups and dedicated individual writing time. For the small-group sessions, participants were preassigned to groups of three to four people for the purpose of soliciting peer feedback on writing samples they submitted ahead of time. For the individual writing sessions, coaches were on hand to work one-on-one with participants. As with all CSHL meetings and courses, participants were required to respect the confidentiality of any unpublished research they may have read during the retreat.

This course was supported with funds provided by the National Institutes of Health National Institute of General Medical Sciences.

PARTICIPANTS

Ames, K., Albert Einstein College of Medicine, Bronx, New York
Creixell, P., Massachusetts Institute of Technology, Cambridge
Dai, J., Fred Hutchinson Cancer Research Center, Seattle, Washington
Deutsch, D., Princeton University, New Jersey
Gathungu, G., Stony Brook University, New York
Ghanbarian, E., University of Lethbridge, Alberta, Canada
He, L., Weill Cornell Medical College, Flushing, New York
Hiratsuka, T., King's College London, United Kingdom
Jain, S., The Wenner-Gren Institute, Stockholm, Sweden
Marin-Valencia, I., Rockefeller University, New York, New York
Mazlo, J., University of North Carolina, Greensboro
Mendoza, A., Washington University in St. Louis, Missouri
Moiseeva, T., University of Pittsburgh, Pennsylvania
Osmanbeyoglu, H., Memorial Sloan Kettering Cancer Center, New York, New York
Perea-Resa, P., Massachusetts General Hospital/Harvard Medical School, Boston
Perez, O., Pontificia Universidad Católica del Ecuador, Quito
Ruan, H., SUNY Upstate Medical University, Syracuse, New York
Young, J., Massachusetts Institute of Technology, Cambridge
Zinoviev, A., SUNY Downstate Medical Center, Brooklyn, New York

SEMINARS

Lambert, C., Cold Spring Harbor Laboratory, and Matheson, S., *Cell Press,* Cambridge, Massachusetts: Session on top-10 tips. Lay summaries and writing for nonexpert audiences.
Matheson, S., *Cell Press,* Cambridge, Massachusetts, and Connell, L., Senior Editor, *Cold Spring Harbor Laboratory Press:* Publications and manuscripts.
Schaffhausen, J., Yale University, New Haven, Connecticut: All great writing starts with a story.

Computational and Comparative Genomics

November 28–December 5

INSTRUCTORS
D. Hawkins, University of Washington, Seattle
W. Pearson, University of Virginia, Charlottesville
J. Taylor, Johns Hopkins University, Baltimore, Maryland

ASSISTANTS
P. DeFord, Johns Hopkins University, Baltimore, Maryland
M. Heydarian, Johns Hopkins University, Baltimore, Maryland
L. Mills, University of Minnesota, Minneapolis
O. Sabik, University of Virginia, Charlottesville
T. Smith, University of Washington, Seattle
D. Triant, University of Virginia, Charlottesville

This course presented a comprehensive overview of the theory and practice of computational methods for the identification and characterization of functional elements from DNA sequence data. The course focused on approaches for extracting the maximum amount of information from protein and DNA sequence similarity through sequence database searches, statistical analysis, and multiple sequence alignment. Additional topics included alignment and analysis of "next-gen" sequencing data, with applications from metagenomic, RNA-Seq, and ChIP-seq experiments; the galaxy environment for high-throughput analysis; regulatory element and motif identification from conserved signals in aligned and unaligned sequences; integration of genetic and sequence information in biological databases; and genome browsers and genome features.

The course combined lectures with hands-on exercises. Students were encouraged to pose challenging sequence analysis problems using their own data. The course was designed for biologists seeking advanced training in biological sequence and genome analysis, for computational biology core resource directors and staff, and for individuals in other disciplines (e.g., computer science)

who wished to survey current research problems in biological sequence analysis. Advanced programming skills were not required.

The primary focus of this course was the theory and practice of algorithms in computational biology, with the goals of using current methods more effectively for biological discovery and developing new algorithms.

This course was supported with funds provided by the National Human Genome Research Institute.

PARTICIPANTS

Adewale-Fasoro, O., B.S., Redeemer's University, Osogbo, Nigeria
Cao, G., B.A., University of Chicago, Illinois
Chen, R., B.S., Cleveland Clinic, Ohio
Cheng, L., B.A., St. Jude Children's Research Hospital, Memphis, Tennessee
Cho, S., M.D., National Cancer Institute/National Institutes of Health, Frederick, Maryland
Davis, A., B.S., Moffitt Cancer Center, Tampa, Florida
Easton, F., B.Sc., University of Leicester, United Kingdom
Hynes, M., Stanford University, California
Jimenez-Barron, L., B.S., Max Planck Institute of Psychiatry, Münich, Germany
Kadow, Z., B.S./B.A., Baylor College of Medicine, Houston, Texas
Kim, H-D., B.S., University of Arizona College of Medicine, Phoenix
Malachowska, B., Medical University of Lodz, Poland
Martin del Campo, M.B., B.S., University of California, San Francisco
Niu, Z., Mayo Clinic, Rochester, Minnesota
Perez, O., Pontificia Universidad Católica del Ecuador, Quito
Phillips, H., B.S., University of California School of Medicine, Los Angeles
Riboldi, G.M., New York University School of Medicine, New York
Shunkwiler, L., B.A., Medical University of South Carolina, Charleston
Sun, M., B.S., University of California, Los Angeles
Yan, B., B.S., University of Florida, Gainesville

SEMINARS

Hawkins, D., University of Washington, Seattle: Chromatin states 1: Analysis of histone modifications. Chromatin states 2: Overlapping data sets.
Leek, J., Johns Hopkins University, Baltimore, Maryland: RNA sequencing technologies. Statistics for genomes I.
Mackey, A., HemoShear, LLC, Charlottesville, Virginia: Genome annotation (HMM basics). Gene lists to pathways. RStudio visualization.
Mahony S., Penn State University, University Park, Pennsylvania: Genomics of gene regulation 1: Analyzing protein–DNA-binding interactions. Genomics of gene regulation 2: Characterizing transcription-factor-binding dynamics.
Miller, D., Stowers Institute for Medical Research, Kansas City, Missouri: Genome assembly and analysis using Nanopore.
Pearson, W., University of Virginia, Charlottesville: Protein evolution and sequence similarity searching. Practical sequence similarity searching. PSSMs, HMMs, and phenotype prediction. Multiple sequence alignment.
Stubbs, L., University of Illinois, Urbana: Regulatory genomics.
Taylor, J., Johns Hopkins University, Baltimore, Maryland: Sequencing technologies: New genomics. Assembling genomes and transcriptomes. Galaxy for high-throughput analysis. Variation and SNP discovery. Probing higher-dimension chromatin structure.
Wilson Sayres, M., Arizona State University, Tempe: Sex bias in reference-based alignments.

The Genome Access Course

INSTRUCTORS **D. Fagegaltier,** New York Genome Center, New York
E. Hodges, Vanderbilt University School of Medicine, Nashville, Tennessee
B. King, University of Maine, Orono
S. Munger, The Jackson Laboratory, Farmington, Connecticut

INVITED SPEAKERS **C. dos Santos** (March), Cold Spring Harbor Laboratory
S. Maniatis (September), New York Genome Center, New York

The Genome Access Course (TGAC) is an intensive 2-day introduction to bioinformatics offered multiple times each year. The course is broken into modules that are each designed to give a broad overview of a given topic, with ample time for examples chosen by the instructors. Each module features a brief lecture describing the theory, methods, and tools, followed by a set of worked examples that students complete. Students are encouraged to engage instructors during the course with specific tasks or problems that pertain to their own research.

The core of TGAC is the analysis of sequence information framed in the context of completed genome sequences. Featured resources and examples primarily come from mammalian species, but concepts can be applied to any species. TGAC has been offered continuously since 2002, and it has evolved over that time to meet the needs of scientists venturing into the analysis of large sequencing data sets. In 2018, the topics covered by TGAC included genome browsers, the Galaxy Project, the R statistical computing environment, functional genomic elements and the ENCODE Project, high-throughput sequencing data analysis, RNA-Seq analysis, and pathway analysis.

Three iterations of TGAC took place in 2018—in March, September, and November. The November course was held at the CUNY Graduate Center in New York City, in partnership with the City University of New York's biomedical research programs.

March 26–28 (Cold Spring Harbor Laboratory)

21 PARTICIPANTS

Barber, A., Feinstein Institute for Medical Research, Glen Oaks, New York
Calof, A., University of California, Irvine
Chakraborty, P., New England Biolabs, Ipswich, Massachusetts
Hall Sedlak, R., Zoetis, Kalamazoo, Michigan
Holmes, D., Bristol-Myers Squibb, Princeton, New Jersey
Kizer, J., Albert Einstein College of Medicine, Bronx, New York
Liu, S., National Institutes of Health, Bethesda, Maryland
Luyten, Y., New England Biolabs, Ipswich, Massachusetts
Mbunwe, E., University of Pennsylvania, Philadelphia
McKimpson, W., Columbia University, New York, New York
Morrison, F., Boston University School of Medicine, Massachusetts
Parmalee, N., Albert Einstein College of Medicine, Bronx, New York
Rahman, M., Cold Spring Harbor Laboratory
Maria Riboldi, G., New York University School of Medicine, New York
Rodriguez-Martinez, J., University of Puerto Rico/Rio Piedras, San Juan
Schnieder, T., Columbia University, New York, New York
Shah, N., Nemours/AI Dupont Hospital for Children, Wilmington, Delaware
Tameire, F., University of Pennsylvania, Philadelphia
Varner, J., Duke University, Durham, North Carolina
Williams, A., University of California, Davis
Yuan, T., The Aaron Diamond AIDS Research Center, New York

September 23–25 (Cold Spring Harbor Laboratory)

33 PARTICIPANTS

Abdulla, A., University of Virginia, Charlottesville
Barnes, D., Genus PLC, Waunakee, Wisconsin
Basu, S., Wistar Institute, Philadelphia, Pennsylvania
Bavi, N., University of Chicago, Illinois
Beyaz, S., Cold Spring Harbor Laboratory
Chakravorty, S., Glaxo SmithKline, Sellersville, Pennsylvania
Chan, S-H., New England Biolabs, Ipswich, Massachusetts
Cheung, Y., Gilead Sciences, Foster City, California
Chin, M.S.S., Albert Einstein College of Medicine, Bronx, New York
Cohen, N., Northwell Health/Hofstra, Manhasset, New York
Conlon, D.M., University of Pennsylvania, Philadelphia

Dachon, A., National Research Council Canada, Montreal, Quebec, Canada
Floro, J., Boston University Medical School, Massachusetts
Gnanapradeepan, K., University of Pennsylvania, Philadelphia
Hackbarth, K., Dicerna Pharmaceuticals, Cambridge, Massachusetts
Hirschi, K., Yale University, New Haven, Connecticut
Johnson, W., Boston College, Chestnut Hill, Massachusetts
Kim, L-S., Johns Hopkins School of Medicine, Baltimore, Maryland
Mason-Osann, E., Boston University School of Medicine, Massachusetts
O'Malley, K., Jounce Therapeutics, Cambridge, Massachusetts
Penny, G., Washington University in St. Louis, Missouri
Pressl, C., The Rockefeller University, New York, New York
Qiu, Z., Stony Brook University, New York
Rymaszewski, A., Medical College of Wisconsin, Milwaukee
Salinas, Y., Yale School of Public Health, New Haven, Connecticut
Shi, J., Columbia University, New York, New York
Sookiasian, D., Homology Medicines Inc., Bedford, Massachusetts
Tramantano, M., Cold Spring Harbor Laboratory
van Soldt, B., Columbia University Medical Center, New York, New York
Walsh, K., University of Virginia, Charlottesville
Yang, G., Carnegie Mellon University, Pittsburgh, Pennsylvania
Zheng, B., Harvard Medical School/Massachusetts General Hospital, Charlestown
Zhu, Z., Glaxo SmithKline, Collegeville, Pennsylvania

November 29–30 (CUNY Graduate School, New York City)

35 PARTICIPANTS

Benchorin, G., Columbia University, New York, New York
Bennett, S., CUNY Graduate Center, New York, New York
Blitzblau, H., Novogy, Inc., Cambridge, Massachusetts
Carrocci, T., Yale University, New Haven, Connecticut
Dansu, D., CUNY Graduate Center, New York, New York
Das, M., Independent, Woodland, California
Davis, S., University of Massachusetts Medical School, Worcester
Desmeules, P., IUCPQ, Quebec City, Canada
Echeverria Gonzalez, A., Hospital for Special Surgery, New York, New York
Friday, A., University of Pittsburgh, Pennsylvania
Gerber, A., Albany Medical College, New York
Giakoumis, M., City University of New York, New York
Hoffman, J., City University of New York, New York
Hu, N., New York University School of Medicine, New York
Krivoshik, S.R., Baruch College/CUNY Graduate Center, New York, New York

Kurata, N., City University of New York/American Museum of Natural History, New York
LaMarre, E., CUNY Graduate Center, New York, New York
Lee, R., CUNY Graduate Center, New York, New York
Loetstedt, B., KTH, Stockholm, Sweden
Maguire, B., Independent, Chester, Connecticut
Michalski, M., Van Andel Research Institute, Grand Rapids, Michigan
Nayeem, N., CUNY Graduate Center, New York
Ndukwe, K., City University of New York, New York
Phizicky, D., Yale University, New Haven, Connecticut
Reimer, K., Yale University, New Haven, Connecticut
Seow, V.U., CUNY Graduate Center, New York, New York
Tieu, R., University of Pittsburgh, Pennsylvania
Trottier, M., New York University Langone, New York
Tsotakos, N., Penn State University, Harrisburg, Pennsylvania
Veerappa, A., New York University Langone Health, New York
Wu, S., Memorial Sloan Kettering Cancer Center, New York, New York
Yabut, O., University of California, San Francisco
Yang, J-S., The Graduate Center, New York, New York
Zuelke, D., CUNY Medical School, New York, New York
Zumajo, C., City University of New York, New York

SEMINARS

INVITED SPEAKERS PROGRAM ("CSHL SEMINAR SERIES")

Each year, Cold Spring Harbor Laboratory invites speakers from outside the institution to present their findings. These weekly seminars keep the CSHL staff current on the latest scientific developments and broaden their perspectives. Graduate students and postdoctoral fellows meet with the seminar speakers for lunch immediately after the seminar, providing an opportunity for the exchange of ideas in an informal setting.

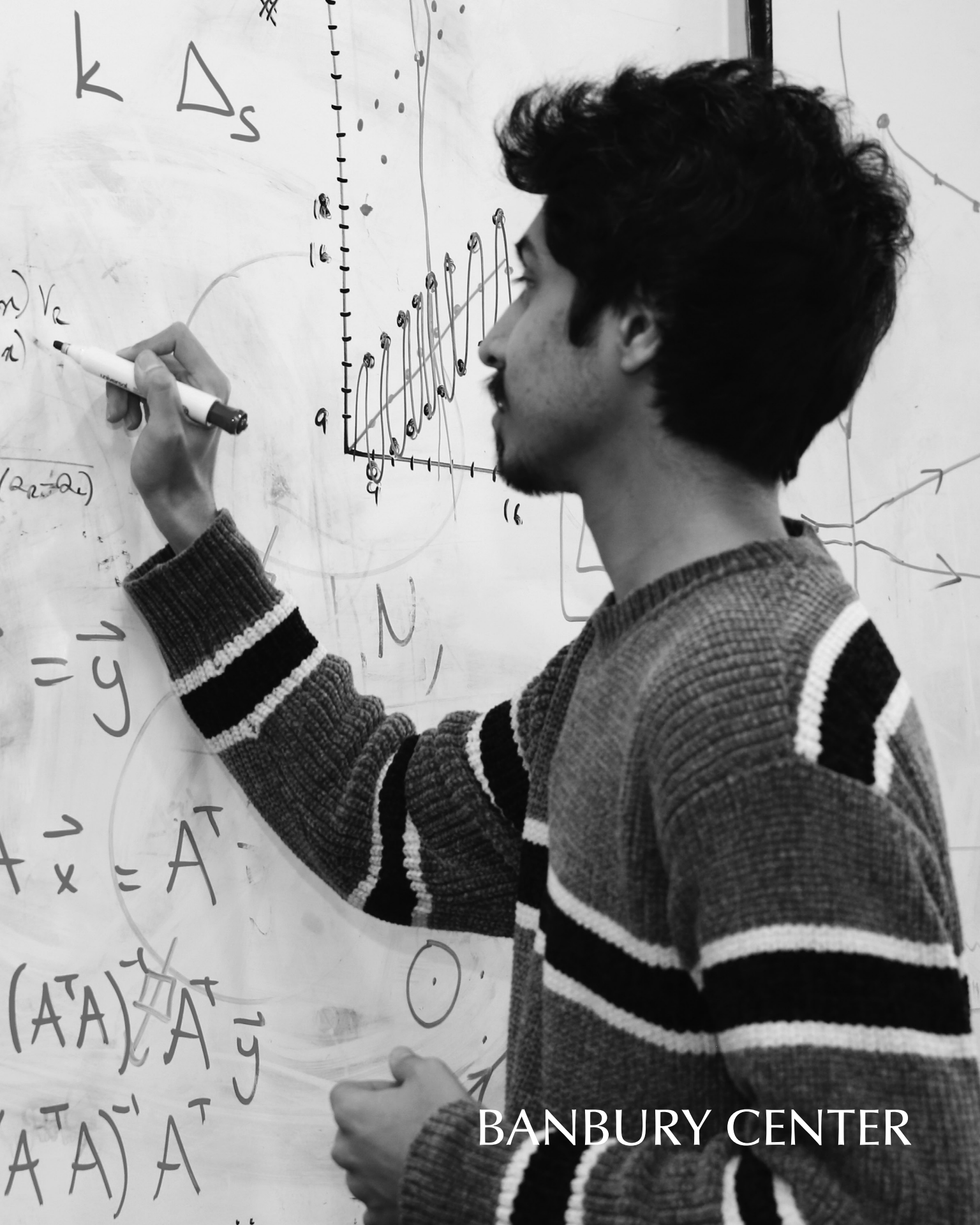
Speaker	Title	Host
January		
Kristi Anseth, Ph.D., Distinguished Professor of Chemical and Biological Engineering, Associate Faculty Director of the BioFrontiers Institute and HHMI Investigator, University of Colorado at Boulder	Dynamic hydrogel matrices: Cell biology in the fourth dimension	CSHL WiSE McClintock Lecture
Michale Fee, Ph.D., Department of Brain and Cognitive Sciences, McGovern Institute for Brain Research, Massachusetts Institute of Technology	Building a state space for song learning	Adam Kepecs
Patricia Janak, Ph.D., Bloomberg Distinguished Professor, Johns Hopkins University	Probing dopamine circuits during learning	Steve Shea
February		
David Pellman, M.D., Margaret M. Dyson Professor of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, HHMI	Job's dilemma for the genome: Why bad things happen to good chromosomes	Jason Sheltzer
Andrew Ewald, Ph.D., Associate Professor of Cell Biology, Oncology, and Biomedical Engineering, Johns Hopkins University	Cellular and molecular mechanisms of breast cancer metastasis	Mikala Egeblad
Michael Greenberg, Ph.D., Chair, Department of Neurobiology, Nathan Marsh Pusey Professor of Neurobiology, Harvard Medical School	How nature and nurture conspire to control brain development and function	Jessica Tollkuhn
Caroline Dean, Project Leader, John Innes Centre, Norwich, United Kingdom	Epigenetic switching and antisense transcription	CSHL WiSE
March		
Karel Svoboda, Janelia Senior Group Leader, HHMI's Janelia Research Campus, Ashburn, Virginia	Neural circuits underlying planning and movement	CSHL Postdocs
Samie Jaffrey, Ph.D., Greenberg-Starr Professor, Department of Pharmacology, Weill Medical College, Cornell University	The dynamic epitranscriptome: Encoding the fate and function of mRNA with reversible nucleotide modifications	Ullas Pedmale
Katherine S. Pollard, Ph.D., Gladstone Institutes, Chan-Zuckerberg Biohub, University of California, San Francisco	Transcription factors recognize DNA shape	Adam Siepel
April		
Joseph E. LeDoux, Ph.D., University Professor, New York University	Have we misunderstood fear?	Anthony Zador

Speaker	Title	Host
October		
René Hen, Ph.D., Professor, Departments of Neuroscience, Psychiatry and Pharmacology, Columbia University, College of Physicians and Surgeons; Director, Division of Systems Neuroscience, Department of Psychiatry; The New York State Psychiatric Institute and Research Foundation for Mental Hygiene, Inc.	The ventral hippocampus and mood	Bo Li
Mikhail G. Kolonin, Ph.D., Associate Professor, University of Texas Health Science Center at Houston	Targeting intercellular interactions with combinatorial peptides	Alea Mills
November		
Christine Mayr, M.D., Ph.D., Associate Professor, Memorial Sloan Kettering Cancer Center	Regulation of 3′-UTR-mediated protein–protein interactions by a membrane-less organelle	Chris Vakoc
Omar I. Abdel-Wahab, M.D., Associate Member, Memorial Sloan Kettering Cancer Center	Understanding and targeting spliceosomal gene mutations in cancer	Terri Grodzicker
Ali H. Brivanlou, Ph.D., Robert & Harriet Heilbrunn Professor, Head of Laboratory of Stem Cell Biology and Molecular Embryology, The Rockefeller University	Self-organization of synthetic human embryoids and organoids	Lloyd Trotman
Gina Poe		Graduate Students
December		
Botond Roska, Ph.D., Professor, University of Basel	Neuronal circuits and medicine	Anthony Zador William R. Miller Lecture
Elaine Fuchs, Ph.D., Rebecca C. Lancefield Professor, Mammalian Cell Biology and Development, The Rockefeller University	Stem cells: Coping with stress	CSHL WiSE
Gerd Blobel, M.D., Ph.D., Professor, Children's Hospital of Philadelphia; Perelman School of Medicine, University of Pennsylvania	Long-range chromatin contacts: Mechanisms and therapeutic applications	Chris Vakoc

IN-HOUSE SEMINAR PROGRAM

Cold Spring Harbor Laboratory's In-House Seminars were initiated to provide a semiformal avenue for communication among the various research groups at the Laboratory. The seminars also afford a necessary opportunity for graduate students and postgraduate staff to develop their skills in organizing, presenting, and defending their research.

Speaker	Title
January	
Chris Hammell	PQN-59/UBAP2/2L and their conserved roles in amplifying mRNA translation
Lloyd Trotman	Nutrient therapy of prostate cancer
February	
Bor-Shuen Wang	To die or not to die: Chandelier cells in the visual circuit
Xiaoyin Chen	Understanding neuronal circuits using DNA sequencing
March	
Siran Li	Droplet-based single-nucleus sequencing methods for cancer study
Jon Preall	Emerging applications of single-cell technology
Jay Lee	Nucleolar RNA architecture and epigenetics in cancer
Mandy Wong	Quantitative activity profile and context dependence of all human 5′ splice sites
Florin Albeanu	Parallel processing pathways in mammalian olfaction
April	
Pavel Osten	Of mice and twins: Partial penetrance phenotypes in a mouse model of 16p11.2 deletion
Jessica Tollkuhn	Leveraging defined neural circuits to link genes to behavior
October	
Chris Vakoc	Transcriptional addiction in diverse tumor lineages
Hervé Tiriac	Organoid profiling identifies common responders to chemotherapy in pancreatic cancer
Sebastian Soyk	Dissecting the effects of epistasis on plant stem cell control and crop productivity
November	
Doreen Ware	Biology enabled agriculture through the dissection of complex traits
Camila dos Santos	A role for pregnancy-induced epigenetic modifications on guiding development and oncogenesis of mammary epithelial cells
December	
Thomas Gingeras	Precision genomes and genomics in healthy individuals: Helpful or not


BANBURY CENTER

BANBURY CENTER EXECUTIVE DIRECTOR'S REPORT

Cold Spring Harbor Laboratory's Banbury Center marked another anniversary in 2018: 40 years since the first group of experts arrived at the Center for a meeting, "Assessing Chemical Mutagens: The Risk to Humans." This inaugural group set out to debate the impending health risks that could be tied to emerging and pervasive chemical technology. It is easy to find echoes of these first discussions of technology and health, as well as the conclusions that more data are needed, in the dialogues happening today.

The Numbers

The Center was humming in 2018, with more than 50 events utilizing the estate, including traditional Banbury meetings, Meetings & Courses Program (*Workshop on Schizophrenia and Related Disorders*, *Computational Neuroscience*, *Genetics and Neurobiology of Language*, *Brain Tumors*, *Workshop on Leadership in Bioscience*, *Scientific Writing Retreat*), Watson School of Biological Sciences courses (*Microbial Pathogenesis*, *Evolution*), and laboratory retreats.

The Center welcomed 465 individuals for Banbury meetings in 2018, with 71% marking their first visit. These participants were drawn from 26 countries and 32 states in the United States. The largest portion of Banbury attendees work in academic settings (72%); however, participants from industry (10%), not-for-profit organizations (7%), U.S. and foreign governments (7%), and publishing/writing (4%) brought diverse perspectives and new cross-sector relationships. Banbury continues to strive for gender equity: 41% of participants and 36% of meeting organizers in 2018 were women.

Funding continues to be a major hurdle in convening Banbury meetings. In 2018, Banbury secured financial support from not-for-profit organizations (46%), industry (22%), and government

Technology and Education Council: Opportunities for AI and Machine Learning for the Biotech Industry

B. Amulic, B. Barnes, M. Egeblad, M. Kaplan, and A. Zychlinsky at *Diverse Functions of Neutrophils in Cancer* meeting

(4%). The CSHL Corporate Sponsor Program remained a critical resource for cutting-edge meetings and represented 29% of funding.

One of the challenges in convening confidential, invitation-only meetings is ensuring relevant discussions and outputs reach target audiences beyond meeting participants. In 2018, Banbury made progress on this front with a new website and social media accounts. With the growing number of experts using social media to share ideas and engage in discussion of science and technology, the Banbury Center waded into these platforms as a way to share outputs from meetings, as well as historical context of the program and estate. The new Banbury Twitter and Instagram accounts (@CSHLbanbury) have been a productive tool to deliver information about recent and historical meetings, share outputs, showcase the beauty of the estate, and celebrate achievements of Banbury alumni. After only a few months on Twitter, and despite limits to what we can post (because of our confidentiality policy), we tweeted 46 times, added 173 followers, gained 69 mentions, and garnered nearly 60,000 impressions.

Meeting Themes

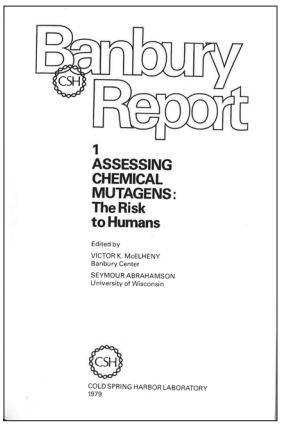
The year's meetings drew on two of Banbury's strengths: bridging interdisciplinary divides and hosting discussions at the frontiers of science and technology. Because of the highly diverse groups of experts, these meetings are often the only forum allowing meaningful engagement between groups. In the spring, *DNA for Digital Storage* saw synthetic biologists and computer scientists scrutinize opportunities and limitations for the use of DNA to store data. New questions emerged throughout the two days, and the group returns in 2019 to further these important discussions. Perhaps better suited for Halloween, bats were the subject of another spring meeting on *New Models for Aging Research*. The unique characteristics that allow this order to live far longer than its nonflying mammalian relatives were examined by experts spanning comparative biology, gerontology, immunology, genetics, and neuroscience; many new connections were made, which we expect will lead to new collaborations and progress in several fields. Finally, an especially diverse group of synthetic biologists, metabolic engineers, developmental biologists, and biochemists met at December's *Revolutionizing Agriculture with Synthetic Biology*, aiming to "think big" about using synthetic biology to improve crops and other plants. The year concluded with *Phase-Separated Assemblies in Cell Biology*, and lively debate between biologists and physicists over these so-called membraneless organelles.

R. Nkambule, P. Preko, H. Doyle, and A. Achrekar at *What Is Needed for a Comprehensive Community Response to HIV* meeting

In the context of neuroscience, four meetings tackled questions from basic research through training and policy. February's *The Evolving Phenomenon of Direct-to-Consumer Neuroscience* convened diverse stakeholders and experts to consider medical, ethical, and regulatory issues emerging with the availability of at-home devices and software to monitor and/or modulate brain function. The meeting's co-organizers used these discussions as the basis for a recent article in *Science*. The autumn brought issues of a more fundamental nature, including *Quantitative Approaches to Naturalistic Behaviors* in September, and *Why Does the Neocortex have Layers and Columns?* in October. The National Institute of Mental Health returned in 2018 for *Brain Camp IX,* an intensive scientific retreat for top psychiatry residents with an interest in a research career.

Cancer has always been a major target for Banbury discussions, and 2018 was not different, with four such meetings in the autumn. Three events gathered experts to discuss the ways in which cellular functions and pathways may play a role in the development or treatment of cancers: *Emerging Data on the Role of Wnt Biology in Cancer*, *Autophagy and Cancer*, and *Diverse Functions of Neutrophils in Cancer*. Further down the pipeline, *Towards a Cure for Advanced Stage Ovarian Carcinoma* reviewed current and prospective treatments for this disease in order to identify strategies best positioned to provide optimal outcomes for patients. Similarly working to support the best research and improving outcomes for patients, the Lustgarten Foundation returned to the Conference Room for their 2018 Scientific Meeting, providing an opportunity for the Scientific Advisory Board, as well as Foundation-supported investigators, to discuss research and strategy.

With an eye toward policy, two meetings targeted broader issues affecting the scientific community: trustworthiness and gender diversity. The former, *Signals of Trust in Science Communication*, was organized by Marcia McNutt (National Academy of Sciences), Richard Sever (CSHL Press and bioRxiv), and Kathleen Hall Jamieson (Annenberg Center). Participants debated emerging challenges to identifying whether reported research represents rigorous scientific standards, especially in the context of increasing interdisciplinarity in research, growth in the number of journals and other reporting outlets, and inconsistent peer review policies. In December, experts met at Banbury for *Increasing Gender Diversity in the Biosciences*, aiming to identify practical solutions to better recruit, promote, and support women. The meeting, led by Carol Greider (Johns Hopkins) and Jason Sheltzer (CSHL), resulted in a number of short- and long-term recommendations, including time banking, opt-out tenure clock extensions, and greater mentorship training, as well as a number of suggestions to improve policies surrounding sexual and gender harassment in academia.



A second set of policy-based meetings targeted public health: The Bill and Melinda Gates Foundation returned for May's *What Is Needed for a Comprehensive, Community Response to HIV?*, convening global leaders in HIV prevention, including organizers Rejoice Nkambule, Deputy Director at the Swaziland Ministry of Health, and Mark Dybul, former U.S. Global AIDS Coordinator. Whereas May's meeting focused on HIV prevention in Africa, September's *Non-Opioid Management of Chronic Pain* took aim at the opioid epidemic from a practical, policy approach: Reduce opioid prescriptions by improving effectiveness of a stepped care model. The meeting was co-organized by the NIH's Head of Pain Policy, Linda Porter, and the director of the Chronic Pain and Fatigue Research Center at the University of Michigan, Daniel Clauw. Among the participants were representatives from both healthcare providers and payers, including Centers for Medicare & Medicaid's Deputy Chief Medical Officer, and the Veterans Affairs National Program Director for Pain Management.

The Team. The Team. The Team.

The real engine of the Center are the professionals who ensure organization, finance, communication, and the estate are running at a high level. In 2018 we bade farewell to Pat Iannotti, who left the Banbury office after 6 years to head for the sunnier south, and to Hakon Heimer, who stepped away from consulting on mental disorders to take up a position with the University of Copenhagen. Michelle Corbeaux celebrated her 3-year anniversary at Banbury with a promotion to Finance and Development Coordinator, and we welcomed Jasmine Breeland as Communications and Special Projects Coordinator. Basia Polakowski continues to oversee our three residence buildings, ensuring our guests are comfortable, while the Culinary Services team keeps them well fed, and the Audiovisual staff ensures technology supports rather than distracts. Jose Peña-Corvera, John Shea, and Paulo Krizanovski look after 55 acres of impeccable grounds, and the entire Facilities team quite literally keeps us running.

Rebecca Leshan
Director

Publications Resulting from Banbury Meetings

Schutzer SE, Body BA, Boyle J, Branson BM, Dattwyler RJ, Fikrig E, Gerald NJ, Gomes-Solecki M, Kintrup M, Ledizet M, et al. 2018. Direct diagnostic tests for Lyme disease. *Clin Infect Dis* **68:** 1052–1057. doi:10.1093/cid/ciy614

Wexler A, Reiner PB. 2019. Oversight of direct-to-consumer neurotechnologies. *Science* **363:** 234–235. doi:10.1126/science.aav0223

BANBURY CENTER MEETINGS

Dates	*Title*	*Organizer(s)*
February 4–6	The Evolving Phenomenon of Direct-to-Consumer Neuroscience	P. Reiner, A. Wexler
March 4–6	DNA for Digital Storage	E. Birney, Y. Erlich, N. Goldman
March 11–14	Bats: New Models for Aging Research	S. Austad, E. Teeling
April 8–11	Signals of Trust in Science Communication	K. Hall Jamieson, M. McNutt, R. Sever
April 13–15	National Institute of Mental Health: Brain Camp IX	J. Chung, J. Gordon
May 13–16	What Is Needed for a Comprehensive, Community Response to HIV?	M. Dybul, R. Nkambule
June 28	Technology and Education Council: Opportunities for AI and Machine Learning for the Biotech Industry	J. Donaldson
September 16–19	Non-Opioid Management of Chronic Pain: Developing Value-Based Models for Diagnosis and Treatment	D. Clauw, L. Porter
September 23–26	Quantitative Approaches to Naturalistic Behaviors	W. Bialek, S. Palmer, S. Sober
October 7–10	Emerging Data on the Role of Wnt Biology in Cancer	J. Clevers, C. Mirabelli, R. Nusse, D. Tuveson, B. Williams
October 14–17	Autophagy and Cancer	R. Amaravadi, J. Debnath, A. Kimmelman
October 21–23	Towards a Cure for Advanced Stage Ovarian Carcinoma	J. Boyd, S. DeFeo, D. Levine, A. Moran
October 28–31	Why Does the Neocortex have Layers and Columns?	S. Ahmad, J. Gavornik, S. Mihalas
November 11–13	The Lustgarten Foundation Scientific Meeting	K. Kaplan, D. Tuveson, R. Vizza
November 27–30	Diverse Functions of Neutrophils in Cancer	K. de Visser, M. Egeblad, P. Kubes
December 2–5	Revolutionizing Agriculture with Synthetic Biology	A. Hanson, C. Vickers, E. Wurtzel
December 9–12	Increasing Gender Diversity in the Biosciences	C. Greider, J. Sheltzer
December 16–19	Phase-Separated Assemblies in Cell Biology	A. Chakraborty, G. Seydoux, P. Sharp, R. Young

BANBURY CENTER MEETINGS

The Evolving Phenomenon of Direct-to-Consumer Neuroscience

February 4–6

ARRANGED BY **P. Reiner,** University of British Columbia, Vancouver, Canada
A. Wexler, University of Pennsylvania, Philadelphia

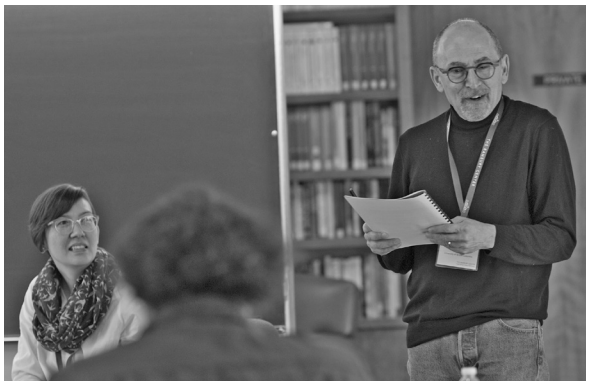
FUNDED BY **Cold Spring Harbor Laboratory Corporate Sponsor Program**

A growing array of devices, products, and software are being sold directly to consumers to monitor and modulate brain function. These products—noninvasive neurostimulation, EEG recording devices, brain-fitness software, and apps that diagnose mental health disorders—are allowing the public to gain access to technologies that were once held behind the closed doors of science and medicine. As there is currently little oversight over the effectiveness of these products and the claims made by their manufacturers, this phenomenon presents a host of novel regulatory and ethical questions. This meeting convened an interdisciplinary expert group of legal scholars, philosophers, bioethicists, sociologists, regulators, and industry representatives to discuss challenges posed by the evolving phenomenon of direct-to-consumer neuroscience and to develop solutions that foster best practices in the field. The meeting's organizers used discussions at this meeting as the foundation for a 2019 policy paper in *Science*.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: P. Reiner, University of British Columbia, Vancouver, Canada, and **A. Wexler**, University of Pennsylvania, Philadelphia

A. Seitz, A. Wexler



K. Rommelfanger, P. Reiner

SESSION 1: Overview of the Direct-to-Consumer Neuroscience Market

Chairperson: D. Dobbs, Science Journalist

A. Fernandez, SharpBrains, Washington, D.C.: What are consumers buying, and why, and how could we empower them to make better decisions in the brain/mental health space?

SESSION 2: Institutional Pushes and Pulls

Chairperson: A. Wexler, University of Pennsylvania, Philadelphia

C. Peña, U.S. Food and Drug Administration, Silver Spring, Maryland: A primer on FDA oversight for neurological devices.

A. Soberats, Federal Trade Commission, Washington, D.C.: FTC advertising law, brain training cases.

SESSION 3: Direct-to-Consumer Brain Training

Chairperson: P. Reiner, University of British Columbia, Vancouver, Canada

K. Rommelfanger, Emory University, Atlanta, Georgia: Internet-based brain training games, citizen scientists, and big data: Ethical issues in unprecedented virtual territories.

J. King, National Institute on Aging, NIH, Bethesda, Maryland: Integrating institutional assessment and communication of the effectiveness of cognitive training.

A. Seitz, University of California, Riverside: Experiences with carrot neurotechnology and FTC regulatory action.

SESSION 4: Direct-to-Consumer Recording

Chairperson: A. Fernandez, SharpBrains, Washington, D.C.

Y. Roy, University of Montreal/NeurotechX, Quebec, Canada: Overview of direct-to-consumer EEG products: Current state and the future.

R. Thibault, McGill University, Montreal, Quebec, Canada: Assessing neurofeedback claims made by consumer EEG companies.

B. Capestany, Duke University, Durham, North Carolina: Consumer concerns with the privacy of brain analytics.

SESSION 5: Direct-to-Consumer Electrical Stimulation

Chairperson: P. Zettler, Georgia State University, Atlanta

A. Wexler, University of Pennsylvania, Philadelphia: Direct-to-consumer brain stimulation: Ethical and regulatory issues.

J. King, B. Wingeier, K. Rommelfanger, C. Peña, D. Dobbs

C. O'Connor, H. Greely

B. Wingeier, Halo Neuroscience, San Francisco, California: Best practices for safe, effective, and credible consumer neurotech development.

SESSION 6: On the Horizon

Chairperson: H. Greely, Stanford University Law School, California

P. Reiner, University of British Columbia, Vancouver, Canada: Can technology be used to read our minds?

N. Farahany, Duke University, Durham, North Carolina; P. Reiner, University of British Columbia, Vancouver, Canada; and A. Wexler, University of Pennsylvania, Philadelphia: Introduction and aims for day 2.

SESSION 7: Perspectives from the Public

Chairperson: K. Rommelfanger, Emory University, Atlanta, Georgia

C. O'Connor, University College Dublin, Belfield, Ireland: Public engagement with brain optimization.

S. Lock, AARP, Washington, D.C.: Brain health.

J. Torous, Harvard University, Boston, Massachusetts: Informed decision-making around mental health apps: The American Psychiatric Association framework approach.

SESSION 8: Independent Third-Party Review of Products and Advertising

Chairperson: P. Reiner, University of British Columbia, Vancouver, Canada

B. Patten, Truth in Advertising, New York, New York: Mind games: The deceptive advertising of brain function products and audio services.

S. Schueller, Northwestern University, Chicago, Illinois: Identification and evaluation of consumer mental health apps.

N. Farahany, C. Peña, P. Reiner, R. Thibault (*back to camera*), B. Wingeier, B. Capestany (seated), K. Rommelfanger

T. Cooperman, ConsumerLabs.com, White Plains, New York: Experiences of a third-party evaluator of health and nutrition products.

SESSION 9: Lessons from DTC Health Products

Chairperson: H. Greely, Stanford University Law School, Stanford, California

P. Zettler, Georgia State University, Atlanta, Georgia: Reviewing the regulatory history of DTC Health Products.

N. Farahany, Duke University, Durham, North Carolina: Discussant.

SESSION 10: Summary Discussion and Wrap-Up

H. Greely, Stanford University Law School, California: Reflections and group discussion on regulation and oversight of direct-to-consumer neuroscience.

P. Reiner, University of British Columbia, Vancouver, Canada, and Anna Wexler, University of Pennsylvania, Philadelphia: Wrap-up and next steps.

DNA for Digital Storage

March 4–6

ARRANGED BY **E. Birney,** European Bioinformatics Institute, Hinxton, United Kingdom
Y. Erlich, Columbia University, New York, New York
N. Goldman, European Bioinformatics Institute, Hinxton, United Kingdom

FUNDED BY **Burroughs Wellcome Fund**; **Agilent Technologies, Inc.**; **Twist Bioscience Corporation**; with additional support from **Microsoft Corporation**

Progress in the use of DNA encoding for data storage has surged since the initial published descriptions of the technology 5 years ago. This Banbury meeting convened experts and thought leaders in order to (1) share current knowledge on the use of DNA for information storage, (2) examine limitations and potential opportunities, and (3) identify strategies to deploy the technology for more widespread research and commercial use. Discussions at the meeting were highly productive and underscored the need for more engagement and for input from groups not represented at the meeting. Ultimately, the group determined that a second meeting in 2019 was needed to continue the momentum.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

SESSION 1: Synthesis and Chemistry

Chairperson: E. Birney, European Bioinformatics Institute, Hinxton, United Kingdom

B. Peck, Twist Bioscience, San Francisco, California: Addressing the skeptics, how data storage will scale DNA synthesis.

J. Sampson and L. Whitman, Agilent Technologies, Santa Clara, California: Advanced oligonucleotide library synthesis.

J-F. Lutz, Institut Charles Sadron, Strasbourg, France: Digital polymers: Recent achievements and promises.

R. Grass, ETH Zürich, Zürich, Switzerland: The stability of DNA during storage.

O. Milenkovic, N. Goldman

R. Grass, J. Flatley, S. Kosuri, K. Strauss

O. Milenkovic, University of Illinois, Urbana: New directions in DNA-based data storage and computing.

SESSION 2: Systems

Chairperson: Y. Erlich, Columbia University, New York, New York

E. Zadok, Stony Brook University, New York: History and recent trends in data storage technologies.

S. Hickling, GCHQ, Cheltenham, United Kingdom: How GCHQ is planning to use DNA for data storage.

K. Strauss, Microsoft, USA, Redmond, WA; Digital data storage in synthetic DNA.

SESSION 3: Money

Chairperson: R. McKibbin, BBSRC, Swindon, United Kingdom

M. Biddle, Innovate UK, Swindon, United Kingdom: The UK approach to disruptive innovation such as DNA for digital storage.

B. Bramlett, Twist Bioscience, San Francisco, California: Economic Drivers for DNA Memory.

SESSION 4: New Science

Chairperson: N. Goldman, European Bioinformatics Institute, Hinxton, United Kingdom

Y. Erlich, Columbia University, New York, New York and Dina Zielinski, Institut Curie, Paris, France: DNA storage: A critical evaluation.

E. Birney, European Bioinformatics Institute, Hinxton, United Kingdom: Developments in digital DNA: New uses.

S. Kosuri, University of California, Los Angeles: DNA and storage: An optimist's view.

J. Flatley, Illumina, San Diego, California: Progress in DNA sequencing.

SESSION 5: Wrap-Up/Next Steps

Led by: Y. Erlich, Columbia University, New York, New York

M. Biddle, L. Whitman

E. Zadok, D. Rosenthal

Bats: New Models for Aging Research

March 11–14

ARRANGED BY **S. Austad,** University of Alabama at Birmingham
E. Teeling, University College Dublin, Ireland

FUNDED BY **The Cold Spring Harbor Laboratory Corporate Sponsor Program**

Aging, the nearly ubiquitous deterioration of physical and mental function that occurs with time, has the greatest impact on global health: People everywhere are experiencing longer life spans, but not necessarily longer "health spans." Thus, understanding the processes that underlie healthy aging remains a critical challenge. Although researchers have made substantial progress studying aging in short-lived mammals such as mice, there is little evidence that these methods will translate to more aging-resistant species such as humans. An alternative approach is to analyze species that are even more aging-resistant than humans: bats. This Banbury meeting convened a cross-disciplinary group of experts to explore the underlying molecular basis of extended health and longevity in bats and to identify strategies for integrating the discoveries from this model species into broader aging studies.

Welcoming Remarks: **R. Leshan,** Director, Banbury Center, Cold Spring Harbor Laboratory
S. Austad, University of Alabama at Birmingham: Introduction and meeting objectives; and Why bats are important for biological aging research.

SESSION 1: Bat Biology: An Aging Context

Chairperson: E. Teeling, University College Dublin, Ireland

N. Simmons, American Museum of Natural History, New York, New York: Bat diversity: What we know (and don't know).

G. Wilkinson, University of Maryland, College Park: Repeated evolution of longevity in bats.

D. Dechmann, Max Planck Institute for Ornithology, Radolfzell, Germany: The exception from the rule: Short life expectancy and unusual immune response of the Pallas free-tailed bat, *Molossus molossus*.

E. Teeling

N. Simmons, D. Dechmann

S. Puechmaille, University of Greifswald, Germany: Comparing what is ecologically comparable in interspecific ageing studies.

S. Vernes, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands: Bats as a model for ageing: What neurogenetic tools do we need?

SESSION 2: Comparative Approaches to Aging

Chairperson: S. Vernes, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

V. Gorbunova, University of Rochester, New York: Long-lived mammals as research models for healthy aging.

V. Gladyshev, Harvard Medical School, Boston, Massachusetts: Insights into lifespan control from the bat genome and comparative genomics of mammals.

R. Miller, University of Michigan, Ann Arbor: Multicladal cellular biogerontology.

W. Wright, UT Southwestern Medical Center, Dallas, Texas: The comparative biology of telomeres.

E. Teeling, University College Dublin, Ireland: Molecular basis of exceptional ageing in bats.

R. Anderson, University of Wisconsin, Madison: Nonhuman primate aging.

SESSION 3: Mechanisms to Measure Aging and its Intervention

Chairperson: R. Miller, University of Michigan, Ann Arbor, Michigan

D. Gems, University College London, United Kingdom: Evolutionary and proximate mechanisms of aging: New insights from *C. elegans*.

P. Fedichev, Gero, LLC and Moscow Institute of Physics and Technology, Russian Federation: Aging as dynamic instability of underlying regulatory network: The case for negligible senescence.

F. Sierra, National Institute on Aging, NIH, Bethesda, Maryland: Phylogenetic efforts at NIA.

S. Horvath, University of California, Los Angeles Epigenetic clock for mammals.

D. Promislow, University of Washington, Seattle: Systems biology approaches in aging research.

SESSION 4: Flight, Immunity, Hibernation and Longevity

Chairperson: C. Wright, Journalist

L. Dávalos, Stony Brook University, New York: Metabolism, immunity, and the emergent unified theory of survival and disease.

R. Miller, S. Austad

S. Horvath, P. Fedichev

V. Deep Dixit, Yale University, New Haven, Connecticut: Immune to aging.

K. Belov, The University of Sydney, Australia: Immunity and aging in marsupials.

K. Storey, Carleton University, Ottawa, Canada: Hibernation and aging.

E. Teeling, University College Dublin, Ireland: The bat immune system and its potential role in ageing.

SESSION 5: Breakout Groups

Chairperson: S. Austad, University of Alabama at Birmingham This interactive session used insights from meeting presentations to consider questions around the future of bats and aging research.

S. Austad, University of Alabama at Birmingham, Alabama: Overview of breakout group aims, and dividing into groups.

V. Gorbunova, S. Horvath

SESSION 6: Meeting Wrap-Up and Next Steps

Chairpersons: E. Teeling, University College Dublin, Ireland, and **S. Austad**, University of Alabama at Birmingham

Signals of Trust in Science Communication

April 8–11

ARRANGED BY **K. Hall Jamieson,** University of Pennsylvania, Philadelphia
M. McNutt, National Academy of Sciences, Washington, D.C.
R. Sever, Cold Spring Harbor Laboratory Press

FUNDED BY **The Alfred P. Sloan Foundation**

Although overall confidence in science remains relatively high, a number of factors are making it harder for scientists themselves to determine whether a source of scientific knowledge is trustworthy, a specific finding robust, and a scientific consensus confirmed. The expanding use of preprints in biomedical sciences in particular has the potential to confuse readers about the extent to which work has been vetted and/or is generally accepted by a field. At this Banbury Center meeting, leading scientists joined science communicators and information technology experts to explore how the scientific community can identify and institutionalize signals of trustworthiness on which the audiences can rely in assessing scientific information.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: M. McNutt, National Academy of Sciences, Washington, D.C.,
J. Greenberg, Alfred P. Sloan Foundation, New York, New York

SESSION 1: Social Context and Trust

Chairpersons: K. Hall Jamieson, University of Pennsylvania, Philadelphia, and **K. Prewitt**, Columbia University, New York, New York

M. Woolley, ResearchAmerica!, Arlington, Virginia: Signals of trust from the American public in the age of Twitter.

B. Spellman, University of Virginia School of Law, Charlottesville: Starting early by educating the consumers of scientific information.

C. Zimmer, B. Spellman

R. Schekman, C. Bargmann

G. Gray, University of Victoria, British Columbia, Canada: The social organization of trustworthiness in research.

A. Russell, Defense Advanced Research Projects Agency, Arlington, Virginia: Striking while the irony is hot: Lessons from trusting research for doing research on trust.

K. Prewitt, Columbia University, New York, New York: The sticky problem of self-inflicted wounds.

SESSION 2: Science in Practice

Chairperson: M. McNutt, National Academy of Sciences, Washington, D.C.

C. Bargmann, Chan Zuckerberg Initiative, Palo Alto, California: Aligning incentives in science.

K. Mitchell, Trinity College Dublin, Ireland: Differentiating signal from noise in the scientific literature.

R. Schekman, University of California at Berkeley: Challenge of reproducibility in biomedical science.

F. Lynch, Mayo Clinic, Rochester, Minnesota: Research storytelling at Mayo Clinic.

N. Thompson, Massachusetts Institute of Technology, Cambridge: Science is shaped by Wikipedia: Evidence from a randomized control trial.

SESSION 3: Academic Publication—Peer Review

Chairpersons: R. Sever, Cold Spring Harbor Laboratory Press, and **A. Casadevall**, Johns Hopkins School of Medicine, Baltimore, Maryland

R. Anderson, University of Utah, Salt Lake City: Peer review, What is it good for?: What peer review can do, what it can't, and why it seems (and may actually be) irreducible.

T. Bloom, The BMJ, London, United Kingdom: Peer review and trust: What is the role of journals in ensuring the trustworthiness of reported findings?

E. Phimister, *New England Journal of Medicine*, Boston, Massachusetts: How editorial requirements improve trust in communication.

J. Berg, Science Journals, Washington, D.C.: The role of publishers in communicating science to the scientific community and the public.

A. Casadevall, Johns Hopkins School of Medicine, Baltimore, Maryland: Creating a more robust rigorous research enterprise.

B. Nosek, Center for Open Science, Charlottesville, Virginia: Open science.

V. Kiermer, Public Library of Science, San Francisco, California: Screening content at scale before peer review—Respective roles of technology, publishers and community.

K. Jamieson, M. McNutt

A. Acharya, V. Kiermer, J. Sheehan, R. Sever, J. Greenberg

J. Dickerson, E. Phimister

U. Manber, A. Acharya

SESSION 4: Technology for Discovery and Assessment

Chairperson: U. Manber, University of California, San Francisco

J. Sheehan, U.S. National Library of Medicine, NIH, Bethesda, Maryland: Trusted health information: A tale of two systems.

U. Manber, University of California, San Francisco: The future of science is not what it used to be.

A. Acharya, Google, Inc., Mountain View, California: Leveraging aggregation to compute trustworthiness.

G. Bilder, Crossref, Oxford, United Kingdom: What color is your paratext?

J. Dickerson, Consumer Reports, Yonkers, New York: Consumer Reports: Smarter choices for a better world.

S. Hawke, World Wide Web Consortium, Waltham, Massachusetts: Toward cooperative verification: Standardizing credibility indicators on the web.

SESSION 5: Communicating with the Public

Chairperson: K. Hall Jamieson, University of Pennsylvania, Philadelphia

C. Zimmer, *The New York Times*, New York: How science journalists look for signals of good science.

S. Borenstein, The Associated Press, Washington, D.C.: We CAN handle the truth! Truth-telling tests and how science writers sort facts, fraud and fluff.

R. Harris, National Public Radio, Washington, D.C.: The role of science journalists in addressing the "reproducibility crisis."

L. Lindenfeld, Stony Brook University, New York: Linking research with practice to advance science communication.

SESSION 6: Wrap-Up and Next Steps

Facilitator: K. Hall Jamieson, University of Pennsylvania, Philadelphia

National Institute of Mental Health: Brain Camp XI

April 13–15

ARRANGED BY **J. Chung,** National Institute of Mental Health, Bethesda, Maryland
J. Gordon, National Institute of Mental Health, Bethesda, Maryland

FUNDED BY **National Institute of Mental Health, NIH**

Cold Spring Harbor Laboratory is renowned worldwide for its educational programs, from high school to the highest professional levels. One of the Banbury Center's contributions is to host the NIMH-sponsored "Brain Camp." The goal of the Brain Camp is to identify areas of neuroscience research with relevance to psychiatrists and to open discussions of these areas with a small group of outstanding psychiatry residents and research fellows. Some of the most distinguished and thoughtful neuroscientists in the country came as guest speakers to the meeting.

SESSION 1

B. Cuthbert, National Institute of Mental Health, Bethesda, Maryland: Welcome and Introductions.

H. Blair Simpson, Columbia University College of Physicians, New York, New York, and S. Ahmari, University of Pittsburgh Medical Center, Pennsylvania: ORAP: The next generation.

D. Ross, Yale School of Medicine, New Haven, Connecticut, and K. Dzirasa, Duke University, Durham, North Carolina: Brain camp ten years later.

SESSION 2

J. Javitch, Columbia University College of Physicians, New York, New York: Novel approaches to targeting G protein–coupled receptors.

S. Ahmari, University of Pittsburgh Medical Center, Pennsylvania: Using translational strategies to identify the neural substrates of OCD-like behaviors and treatment response.

K. Dzirasa, Duke University, Durham, North Carolina: Mapping emotions: Discovering structure in mesoscale electrical brain recordings.

M. Paulus, Laureate Institute for Brain Research, Tulsa, Oklahoma: The challenge of connecting units of analyses: Toward a multilevel description of psychiatric disorders.

SESSION 3

D. Hafeman, University of Pittsburgh, Pennsylvania: Steps toward prevention: Probing the bipolar prodrome in at-risk youth.

S. Vinogradov, University of Minnesota, Minneapolis: Neuroscience-informed cognitive training for neural system impairment in psychiatric illness.

What Is Needed for a Comprehensive, Community Response to HIV?

May 13–16

ARRANGED BY **M. Dybul,** Georgetown University, Washington, D.C.
R. Nkambule, Kingdom of Swaziland Ministry of Health, Swaziland

FUNDED BY **The Bill and Melinda Gates Foundation**

The HIV pandemic has created a long-term challenge for public health in low-income, severely affected countries and communities. In these countries, the lack of well-resourced and functioning health systems means that accelerating the declines in HIV infections will be a critical challenge requiring innovative thinking around community delivery and use of health promotion and healthcare. This Banbury meeting brought together experts in community responses to HIV, organizing frontline healthcare workers, designing and managing health systems, and political and management thinking in order to develop an agenda to promote sustainable, community adoption of effective HIV prevention, testing, and treatment.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

SESSION 1: Introductions and Framework for the Meeting

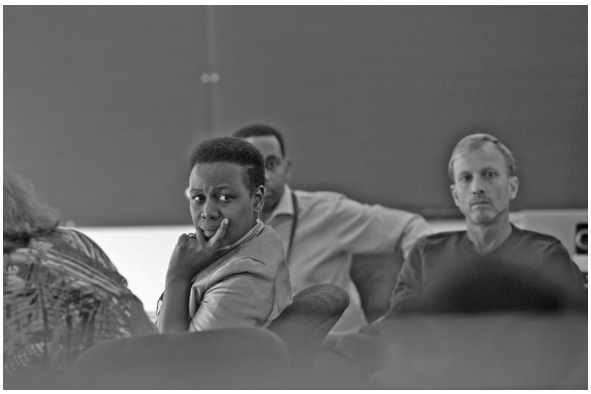
Introduction and Meeting Objectives: R. Nkambule, Ministry of Health, Swaziland, and **M. Dybul**, Georgetown University, Washington, D.C.

G. Garnett, Bill and Melinda Gates Foundation, Seattle, Washington: Gates Foundation strategy and perspectives on community-based service delivery for HIV.

SESSION 2: Core Meeting Themes: What Are Key Elements of Community-Based Care, and Multidisciplinary Perspectives for Person-Centered Care from Across a Health System

Chairperson: M. Dybul, Georgetown University, Washington, D.C.

C. Kinoti, National AIDS Control Council of Kenya, Nairobi: Engaging communities and county governments in Kenya in promoting a coordinated HIV response.



C. Kinoti, M. Osman, M. Dybul

C. Kinoti, A. Hardon, S. Fadiga-Branchi

S. Arbeiter, World Relief, Baltimore, Maryland: The impact of the faith community in addressing HIV.

S. Baptiste, International Treatment Preparedness, Gaborone, Botswana: What it really takes to work with and support communities.

C. Chikanda, Pulse Health, Midrand, South Africa: HIV epidemic control in South Africa—Is the private sector the missing link?

K. Sikkema, Duke University, Durham, North Carolina: Addressing mental health in HIV prevention and treatment.

L.-G. Bekker, Desmond Tutu HIV Centre, Cape Town, South Africa: Addressing reproductive health and beyond: Clinical interventions to engage adolescents and young people.

G. Mackie, University of California, San Diego, California: Learning from programs that organize change of moral, social, and legal norms.

A. Hardon, Amsterdam Institute Social Science, Amsterdam, Netherlands: New social forms in HIV prevention and care.

SESSION 3: Experiences in Community-Based Prevention for HIV

Chairperson: R. Nkambule, Ministry of Health, Swaziland

D. Birx, U.S. Department of State, Washington, D.C.: PEPFAR perspectives on prevention opportunities and challenges.

P. Bhattacharjee, University of Manitoba, Nairobi, Kenya: Scaling up a violence prevention and response program for key populations.

SESSION 4: Experiences in Community-Based Care and Treatment for HIV

Chairperson: C. Holmes, Georgetown University, Washington, D.C.

R. Barnabas, University of Washington, Seattle, Washington: Community-based HIV testing, ART initiation, monitoring, and retention in HIV care.

I. Sikazwe, Centre for Infectious Disease Research in Zambia, Lukasa, Zambia: Differentiated HIV service delivery in Zambia.

SESSION 5: Scaling Up Comprehensive Community-Based HIV Prevention, Care and Treatment for HIV

Chairperson: R. Nkambule, Ministry of Health, Swaziland

C. Holmes, P. Preko

R. Nkambule, S. Hamm Rush

S. Mukasa Monico, S. Baptiste

D. Birx

W. El-Sadr, ICAP at Columbia University, New York, New York: Paper presentation: Moving toward scaling up of differentiated service delivery models.

Panelist Responses

P. Preko, ICAP at Columbia University, New York, New York

S. Mukasa Monico, UNAIDS, Juba, South Sudan

SESSION 6: Achieving Sustainability in Community Engagement for HIV Prevention, Care, and Treatment for Greater Impact

Chairperson: S. Arbeiter, World Relief

H. Doyle, The Global Fund, Vernier, Switzerland: Innovations from the Global Fund.

S. Fadiga-Branchi, Ambassade de France en Côte d'Ivoire, Abidjan, Côte d'Ivoire: Empowering communities in the HIV response, what it means in our national, regional, global health architecture?

D. Rech, The Aurum Institute, Bedfordview, South Africa: Using digital tools for greater performance and impact in community programs.

M. Osman, Elton John AIDS Foundation, London, United Kingdom: Donor perspectives.

SESSION 7: Outstanding Topics and Breakout Groups

Chairperson: M. Dybul, Georgetown University and Rejoice Nkambule, Ministry of Health, Swaziland

SESSION 8: Meeting Wrap-Up, Next Steps

Chairpersons: M. Dybul, Georgetown University, Washington, D.C., and **R. Nkambule**, Ministry of Health, Swaziland

G. Garnett, H. Doyle

Technology and Education Council: Opportunities for AI and Machine Learning for the Biotech Industry

June 28

ARRANGED BY **J. Donaldson,** Cold Spring Harbor Laboratory

FUNDED BY **The Banbury Center, Cold Spring Harbor Laboratory's Meetings and Courses Program, and Cold Spring Harbor Laboratory's Technology and Education Council**

This one-day meeting brought together members of Cold Spring Harbor Laboratory's Technology and Education Council with experts and thought leaders for high-level, interdisciplinary engagement around the future of artificial intelligence, machine learning, and high-volume data for the biosciences. In addition to discussions of opportunities to better inform uptake and implementation strategies for these technologies, the meeting stimulated new cross-sector and cross-disciplinary relationships and insights.

Welcoming Remarks and Introduction: B. Stillman, Cold Spring Harbor Laboratory

SESSION 1

Chairperson: J. Crawford, Northwell Health, Lake Success, New York

K. Hudson, People-Centered Research Foundation, Washington, D.C.: How to do patient-centered research.

S. Khader, Northwell Health, New Hyde Park, New York: Therapeutic trajectories, repositioning candidates and companion diagnostics from electronic health records.

O. Elemento, Weill Cornell Medicine, New York, New York: An integrative AI framework that enables target identification, indication discovery, and drug safety predictions.

SESSION 2

Chairperson: P.J. Amini, Monsanto, St. Louis, Missouri

J. Wiens, University of Michigan, Ann Arbor: Increasing the utility of ML in clinical care: Leveraging big data and domain expertise.

J. Dutkowski, Data4Cure, San Diego, California: Combining systems biology and machine learning to continuously grow biomedical knowledge.

A. Heifets, Atomwise, San Francisco, California: Are we evaluating performance or just overfitting? How to assess the performance of Ligand-based algorithms on virtual screening benchmarks.

A. Vaughan, MapNeuro, Inc., Cold Spring Harbor, New York: ML for connectomics, and what to do with it.

M. Akerman, Envisagenics, New York, New York: Drug target discovery with splicing AI.

G. Yancopoulos, Regeneron Pharmaceuticals, Tarrytown, New York: General discussion and closing remarks.

Non-Opioid Management of Chronic Pain: Developing Value-Based Models for Diagnosis and Treatment

September 16–19

ARRANGED BY **D. Clauw,** University of Michigan, Ann Arbor
L. Porter, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland

FUNDED BY **The MAYDAY Fund**

The United States is in the midst of an opioid epidemic that is partly driven by the overprescribing of opioids for both acute and chronic pain. One of the challenges around removing opioids from the management of chronic pain patients is that nonopioid treatments, especially nonpharmacological therapies, are often not reimbursed by third-party payers and/or are difficult to deliver in the short office visits that are now the norm in primary care and other settings. Recent advances in our understanding of chronic pain provide an opportunity to introduce innovative care models that could increase the quality of care while reducing costs. This Banbury Center meeting convenes relevant stakeholders, thought leaders, and constituents to develop value-based models for diagnosis and treatment of chronic pain.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: **D. Clauw,** University of Michigan, Ann Arbor, and
L. Porter, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland

D. Clauw, T. Postma

M. Cheatle (*back to camera*), L. Porter

SESSION 1: Systems/Care Models I—Primary Care

Chairperson: D. Clauw, University of Michigan, Ann Arbor

B. McCarberg, University of California, San Diego, Poway, California: Chronic pain management in primary care.

K. Gebke, Indiana University School of Medicine, Indianapolis: Physician workflow and efficiency.

L. DeBar, Kaiser Permanente Washington Health, Seattle, Washington: Activating patients and promoting lifestyle approaches to pain management in frontline clinical care: Lessons from embedded research in integrated health plans and community health clinics.

K. Kroenke, Indiana University, Indianapolis: Telecare, stepped care, and collaborative care for chronic pain.

SESSION 2: Systems/Care Models II—Integrating into Specialty Care

Chairperson: L. Porter, NINDS, Bethesda, Maryland

H. Slater, Curtin University, Perth, Australia: Shifting the balance from low value to high value pain care: The role of models of care in driving system-wide reform.

F. Sandbrink, Veterans Affairs Medical Center, Washington, D.C.: Pain management and the opioid safety initiative in the Veterans Health Administration.

S. Stanos, Swedish Health Services, Seattle, Washington: Pain rehabilitation and interdisciplinary care: Our time has come.

A. Doorenbos, University of Washington, Seattle: Determinants of optimal dose and sequence of functional restoration and integrative therapies in service members with neuromusculoskeletal injury.

SESSION 3: Systems/Care Models III—Patient Perspective

Chairperson: L. Porter, NINDS, Bethesda, Maryland

C. Veasley, Chronic Pain Research Alliance, Brookfield, Wisconsin: The need to incorporate evidence and patient-centeredness into pain care models.

SESSION 4: Managing Pain Without Opioids

Chairperson: D. Clauw, University of Michigan, Ann Arbor

M.-A. Fitzcharles, McGill University, Montreal, Quebec, Canada: Can a multidisciplinary chronic pain program provide effective care without opioids?

J. D'Olimpio, Northwell Health, Lake Success, New York: Challenges in the nonopioid management of neuropathic pain.

R. Coakley, L. DeBar

A. Doorenbos, F. Sandbrink, M-A. Fitzcharles

M. Cheatle, University of Pennsylvania, Philadelphia: Managing pain while opioid sparing: Treating common pain comorbidities.

C. Buckenmaier, Uniformed Services University, Rockville, Maryland: Nonpharmacological approaches for pain.

SESSION 5: Psychological Interventions

Chairperson: L. Porter, NINDS, Bethesda, Maryland

J. Haythornthwaite, Johns Hopkins University, Baltimore, Maryland: Addressing psychosocial factors as a critical component of managing chronic pain.

C. Rini, Hackensack University Medical Center, Hackensack, New Jersey: Automated, web-based pain coping skill training: Potential to expand access to an evidence-based, nonpharmacologic pain treatment.

SESSION 6: Payers Perspective

Chairperson: D. Clauw, University of Michigan, Ann Arbor

S. Ling, Centers for Medicare & Medicaid Services, Baltimore, Maryland: Program and policy opportunities to build on the evidence: CMS perspective.

T. Postma, Centers for Medicare & Medicaid Services, Woodlawn, Maryland: Chronic pain management and CMS value-based models.

D. Knecht, Aetna, New York, New York: Aetna's data-driven approach to combating the opioid epidemic.

SESSION 7: Pain in Special Populations

Chairperson: L. Porter, NINDS, Bethesda, Maryland

R. Coakley, Boston Children's Hospital, Massachusetts: Psychological interventions for the treatment of chronic pediatric pain: Translating and scaling current science into widespread, accessible practice.

C. Chambers, Centre for Pediatric Pain Research, Halifax, Canada: Pediatric pain: Innovative care models that can or are being used.

C. Buckenmaier, M. Cheatle

C. Reid, Weill Cornell Medicine, New York, New York: Behavioral approaches to pain management in older adults: How helpful are they?

SESSION 8: Potential Future Directions Driven by New Research

Chairperson: L. DeBar, Kaiser Permanente Washington Health, Seattle

E. Bair, University of North Carolina, Chapel Hill, North Carolina: Using systems biology approaches to identify clusters of individuals with similar underlying mechanisms of chronic pain.

D. Clauw, University of Michigan, Ann Arbor: Innovative research approaches to chronic pain that can help improve care and reduce costs.

L. Porter, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland: The NIH pain research portfolio.

SESSION 9: Closing and Outputs

Chairpersons: L. Porter, NINDS, Bethesda, Maryland, and **D. Clauw**, University of Michigan, Ann Arbor

Quantitative Approaches to Naturalistic Behaviors

September 23–26

ARRANGED BY **B. Bialek,** Graduate Center at CUNY, New York, New York
S. Palmer, University of Chicago, Illinois
S. Sober, Emory University, Atlanta, Georgia

FUNDED BY **The Swartz Foundation**

The first thing that fascinates us about life is the macroscopic behavior of organisms. Recent years have seen an explosion of interest in quantitative approaches to study these real-world behaviors, taming their complexity through more powerful measurements and analyses. This Banbury workshop explored progress in this field, with examples drawn from many different systems ranging from worms to humans.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: **B. Bialek,** Graduate Center at CUNY, New York, New York;
S. Palmer, University of Chicago, Illinois; and
S. Sober, Emory University, Atlanta, Georgia

SESSION 1: Quantifying Behavior—Single Organisms

Chairperson: L. Osborne, Duke University, Durham, North Carolina

G. Berman, Emory University, Atlanta, Georgia: Measuring the hidden dynamics of animal behavior.

S. Reiter, Max Planck Institute for Brain Research, Frankfurt, Germany: Decomposing the control of skin patterning in cuttlefish.

S. Datta, Harvard Medical School, Boston, Massachusetts: Inferring internal from external state using motion sequencing.

E. Mackevicius, W. Bialek, G. Stephens

A. Kennedy, S. Datta

M. Mathis, Harvard University, Cambridge, Massachusetts: Markerless tracking of user-defined features with deep learning.

L. Osborne, Duke University, Durham, North Carolina: Sensation through the lens of behavior.

G. Stephens, VU Amsterdam and OIST, Amsterdam, Netherlands: Capturing the continuous complexity of behavior.

SESSION 2: Interacting Systems—Quantifying and Modeling Social Behavior

Chairperson: B. Bialek, Graduate Center at CUNY, New York, New York

A. Cavagna, Institute for Complex Systems, Rome, Italy: The relevance of scaling laws in natural groups.

P. Gonzalez-Bellido, University of Minnesota, St. Paul: How several predatory aerial insect groups intercept small, fast-moving targets, and why understanding reasons for the species-specific behavioral strategies matters.

N. Mhatre, University of Toronto, Ontario, Canada: Tree crickets optimize the acoustics of baffles to exaggerate. their mate-attraction signal.

G. Theraulaz, Université Paul Sabatier, Toulouse, France: Collective information processing in human phase separation.

A. Kennedy, California Institute of Technology, Pasadena: Quantifying social interactions in pairs of freely behaving mice.

SESSION 3A: The Neural Control of Behavior—Part 1

Chairperson: M. Carey, Champalimaud Center, Lisbon, Portugal

M. Carey, Champalimaud Center, Lisbon, Portugal: Cerebellar contributions to coordinated locomotion in mice.

R. Shadmehr, Johns Hopkins University, Baltimore, Maryland: Population coding in the cerebellum.

A. Churchland, Cold Spring Harbor Laboratory: A new view of decision-making neural activity from quantifying spontaneous movements.

SESSION 3B: The Neural Control of Behavior—Part 2

Chairperson: S. Sober, Emory University, Atlanta, Georgia

G. Theraulaz, G. Berman

S. Sober, S. Palmer, S. Reiter

M. Kaschube, Frankfurt Institute for Advanced Studies, Frankfurt, Germany: Sepia skin pattern control and development at chromatophore resolution.

E. Mackevicius, Columbia University, New York, New York: Unsupervised discovery of temporal sequences in high-dimensional datasets, with applications to neuroscience.

S. Sober, Emory University, Atlanta, Georgia: Spike timing codes for motor control and sensorimotor learning.

SESSION 4: Modeling at the Interface of Sensation and Action

Chairperson: I. Nemenman, Emory University, Atlanta, Georgia

A. Fairhall, University of Washington, Seattle: Sensory drivers of search behavior in mosquitoes.

C. Huang, University of Pittsburgh, Pennsylvania: Propagation and modulation of information in visual pathway.

I. Nemenman, Emory University, Atlanta, Georgia: Automated, predictive, and interpretable inference of *C. elegans* behavioral dynamics.

G. Tavoni, University of Pennsylvania, Philadelphia: Efficient strategies for predictive inference in dynamic environments.

SESSION 5: Challenges and Opportunities in the Physics of Behavior

Chairperson: S. Palmer, University of Chicago, Illinois

SESSION 6: Wrap-Up and Next Steps

Chairpersons: B. Bialek, Graduate Center at CUNY, New York, New York; **S. Palmer**, University of Chicago, Illinois; and **Sam Sober**, Emory University, Atlanta, Georgia

Emerging Data on the Role of Wnt Biology in Cancer

October 7–10

ARRANGED BY **H. Clevers,** Hubrecht Institute, Utrecht, Netherlands
C. Mirabelli, Leap Therapeutics, Cambridge, Massachusetts
R. Nusse, Stanford University, California
D. Tuveson, Cold Spring Harbor Laboratory
B. Williams, Van Andel Research Institute, Grand Rapids, Michigan

FUNDED BY **Leap Therapeutics**

Aberrations and mutations in Wnt-driven signaling pathways appear to play roles in a variety of human cancers. Far less clear are the specific molecular targets within canonical and noncanonical Wnt signaling pathways that drive cancer cell biology and the immune response to cancer. This Banbury Center meeting brought together experts in Wnt-related biology, immunology, pharmacology, and translational cancer medicine in order to assess the current state of the science, identify emerging themes, and prioritize potential therapeutic strategies.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: R. Nusse, Stanford University, California, and **B. Williams**, Van Andel Research Institute, Grand Rapids, Michigan

SESSION 1: Therapeutic Targeting in Oncology

Chairperson: L. Lum, Pfizer, San Diego, California

L. Lum, Pfizer, San Diego, California: Small molecule disruption of Wnt acylation in disease.

D. Virshup, Duke NUS Medical School, Singapore, Asia: Wnt signaling pathways in cancer revealed by PORCN inhibition.

M. Resh, Memorial Sloan Kettering Cancer Center, New York, New York: Biochemistry of Wnt acylation by porcupine: Insights into the mechanism of MBOAT acyltransferases.

M. Waterman, D. Virshup, R. Hannoush

R. Nusse, H. Varmus, X. He (*back to camera*)

SESSION 2: Regenerative Medicine Applications

Chairperson: B. Williams, Van Andel Research Institute, Grand Rapids, Michigan

L. Boulter, University of Edinburgh, Edinburgh, United Kingdom: Modulating noncanonical Wnt signaling in liver repair.

Y. Yang, Harvard School of Dental Medicine, Boston, Massachusetts: Mechanism of Wnt/planar cell polarity signaling in vertebrate embryonic morphogenesis.

B. Williams, Van Andel Research Institute, Grand Rapids, Michigan: Wnt signaling in the skeleton.

SESSION 3: Genetic Screens

Chairperson: S. Angers, University of Toronto, Ontario, Canada

S. Angers, University of Toronto, Ontario, Canada: Charting Wnt signaling networks in normal and cancer cells using CRISPR functional genomic screens.

R. Rohatgi, Stanford University School of Medicine, Stanford, California: Genetic analysis of WNT signaling using haploid human cells.

Y. Ahmed, Dartmouth Medical School, Hanover, New Hampshire: The guts of Wnt signaling in *Drosophila*.

SESSION 4: Targeting DKK1

Chairperson: C. Mirabelli, Leap Therapeutics, Cambridge, Massachusetts

R. Faccio, Washington University in St. Louis, Missouri: Immune suppressive effects of Dkk1 during tumor progression.

A. Bothwell, Yale University School of Medicine, New Haven, Connecticut: Complex immunoregulation by Dickkopf proteins.

D. Wise, New York University Langone Medical Center, New York: Circulating Dickkopf-1 (DKK1) is a marker of aggressive metastatic castration-resistant prostate adenocarcinoma with low PSA expression.

W. Newman, Leap Therapeutics, Cambridge, Massachusetts: Overview of preclinical data regarding the therapeutic targeting of DKK1 in cancer and ongoing translational medicine activities.

C. Sirard, Leap Therapeutics, Cambridge, Massachusetts: Clinical data with DKN-01.

M. Kagey, Leap Therapeutics, Cambridge, Massachusetts: Translational biomarkers for targeting DKK1 in oncology.

SESSION 5: Membrane Receptors and Signalosome Assembly

Chairperson: C. Janda, Princess Maxima Center for Pediatric Oncology, Utrecht, Netherlands

C. Niehrs, German Cancer Research Center (DKFZ), Heidelberg, Germany: Wnt signaling at the membrane.

E. Lee, Vanderbilt University, Nashville, Tennessee: Regulation of Wnt receptor activity.

C. Janda, Princess Maxima Center for Pediatric Oncology, Utrecht, Netherlands: Surrogate Wnt antagonists that phenocopy canonical Wnt ligands.

SESSION 6: Membrane Receptors: R-Spondin, Rnf43/Znrf3, and Frizzleds

Chairperson: M. Maurice, University Medical Center Utrecht, Netherlands

M. Maurice, University Medical Center Utrecht, Netherlands: Mechanisms of driver mutations in Wnt pathway tumor suppressors.

B. Reversade, A*STAR Institute of Medical Biology, Immunos, Singapore: R(e)SPONDIN' to WNT with or without LGR4/5/6.

F. Cong, Novartis Institute for Biomedical Research, Cambridge, Massachusetts: Regulation of the β-catenin destruction complex in colorectal cancer.

C. Janda, L. Lum

D. Tuveson, B. Williams

A. Gurney, OncoMed Pharmaceuticals, Redwood City, California: Therapeutic agents targeting the Wnt pathway.

SESSION 7: Wnt Signaling and Stem Cells

Chairperson: X. He, Boston Children's Hospital, Massachusetts

X. He, Boston Children's Hospital, Massachusetts: Wnt signaling in stem cells and cancer.

M. Waterman, University of California, Irvine: Modeling connections between WNT, stem cells, and the microenvironment in colorectal cancer.

R. Hannoush, Genentech, South San Francisco, California: Pharmacological targeting of Wnt-mediated stem cell function.

SESSION 8: Emerging Applications for Wnt Signaling in Cancer

Chairperson: H. Varmus, Weill Cornell Medicine, New York, New York, and David Tuveson, Cold Spring Harbor Laboratory

T. Tammela, Memorial Sloan Kettering Cancer Center, New York, New York: Wnt-producing niches in stem cell compartments and carcinomas.

J. Massagué, Memorial Sloan Kettering Cancer Center, New York, New York: Regenerative origin of L1CAM^{+}/LGR5^{+} metastatic stem cells.

S. Spranger, Koch Institute for Integrative Cancer Research, Cambridge, Massachusetts: Impact of tumor cell-intrinsic signaling on adaptive and innate immune responses.

SESSION 9: Wrap-Up: Conclusions and Next Steps

Chairpersons: R. Nusse, Stanford University, California, and **B. Williams**, Van Andel Research Institute, Grand Rapids, Michigan

R. Nusse, Stanford University, Stanford, California: Closing plenary.

B. Reversade, M. Maurice, T. Tammela, S. Spranger

Autophagy and Cancer

October 14–17

ARRANGED BY **R. Amaravadi,** University of Pennsylvania, Philadelphia
J. Debnath, University of California, San Francisco
A. Kimmelman, New York University, New York

FUNDED BY **Vescor Therapeutics; Deciphera Pharmaceuticals; Janssen Research** and **Development; and Sprint Bioscience, with additional funding from Genentech and the Cold Spring Harbor Laboratory Corporate Sponsor Program**

In March 2016, the first Banbury Center meeting on Autophagy and Cancer convened academic and industry leaders and resulted in a review article in *Genes & Development*. Since that time, numerous advances have emerged, including new understanding of autophagy's role in cytokine and metabolite secretion, cancer cell metabolism, metastases, and stem cells. However, controversy remains about the fundamental role of autophagy as a tumor suppressor or tumor promoter in cancer, as well as its role in tumor immunity. The goal of this meeting was to explore the mechanisms by which autophagy modulates cancer and to identify strategies to therapeutically target the autophagy pathway in order to best move the field forward.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: R. Amaravadi, University of Pennsylvania, Philadelphia;
J. Debnath, University of California, San Francisco; and
A. Kimmelman, New York University, New York

S. Tooze, A. Simon

R. Perera, J. Debnath

SESSION 1: Signaling Targets in the Autophagy Pathway

Chairperson: J. Debnath, University of California, San Francisco

S. Tooze, Francis Crick Institute, London, United Kingdom: Novel targets modulating autophagy in pancreatic cancer.

R. Zoncu, University of California, Berkeley: Regulation of lysosomal mTORC1 signaling by intracellular cholesterol transport.

J. Martinsson, Sprint Bioscience AB, Huddinge, Sweden: Vps34 inhibitors as immunomodulating agents.

R. Shaw, The Salk Institute for Biological Studies, La Jolla, California: AMPK and ULK1 control of metabolism and cancer.

D. Flynn, Deciphera Pharmaceuticals, Inc., Lawrence, Kansas: Probing the multiple mechanisms of ULK1/2 kinases in cancer cell autophagy, metabolism and survival.

C. Severin, Celgene, San Diego, California: Single cell sequencing analysis of autophagy pathway cross talk in pancreatic cancer.

SESSION 2: Selective Autophagy and Cancer

Chairperson: A. Kimmelman, NYU Langone Medical Center, New York, New York

A. Simonsen, University of Oslo, Norway: Lipid-binding proteins in selective autophagy.

W. Harper, Harvard Medical School, Boston, Massachusetts: Understanding selective autophagy.

J. Moscat, Sanford-Burnham Prebys Medical Discovery Institute, La Jolla, California: Autophagy adaptors in cancer metabolism and inflammation.

A. Thorburn, University of Colorado, Aurora [Presentation on behalf of Jean Mulcahy Levy, University of Colorado, Aurora, Colorado]: Autophagy inhibition for the pediatric brain tumor population. Autophagy, cell death and cancer treatment.

J. Debnath, University of California, San Francisco: NBR1, selective autophagy, and breast cancer metastasis.

SESSION 3: Insights on Autophagy and Cancer from Model Systems

Chairperson: A. Thorburn, University of Colorado, Aurora

T.E. Rusten, Oslo University Hospital, Norway: Autophagy and cancer—What flies tell us.

K. Ryan, Cancer Research UK Beatson Institute, Glasgow, United Kingdom: The connection between autophagy and pathways of tumor suppression.

B. Janji, J. Martinsson, W. Harper, A. Simon

R. Zoncu, A. Kimmelman

J. Guo, Rutgers Cancer Institute of New Jersey, New Brunswick: The role of autophagy in regulating lipid metabolism to support lung tumor growth.

J.-L. Guan, University of Cincinnati, Ohio: Regulation of different subtypes of breast cancer by autophagy genes.

A. Kimmelman, NYU Langone Medical Center, New York, New York: Autophagy and pancreatic cancer.

SESSION 4: Autophagy in Metabolism, Immunity and Inflammation

Chairperson: R. Amaravadi, University of Pennsylvania, Philadelphia

E. White, Rutgers University, New Brunswick, New Jersey: Autophagy-dependent metabolic and immune mechanisms to regulate cancer.

B. Janji, Luxembourg Institute of Health, Germany: Impact of targeting autophagy on the immune landscape of melanoma.

K. Simon, University of Oxford, United Kingdom: Autophagy in differentiating hematopoietic cells.

R. Perera, University of California, San Francisco: Lysosome mediated remodeling of the cellular proteome in pancreatic cancer.

D. Green, St. Jude Children's Research Hospital, Memphis, Tennessee: LC3-associated phagocytosis.

SESSION 5: Therapeutic Targeting of Autophagy in Cancer

Chairperson: A. Viale, MD Anderson Cancer Center, Houston, Texas

R. Amaravadi, University of Pennsylvania, Philadelphia: Targeting PPT1 in cancer.

A. Viale, T.E. Rusten

R. Amaravadi [on behalf of Peter O'Dwyer], University of Pennsylvania, Philadelphia: HCQ in colon/pancreatic cancer.

N. Bahary, McGowan Institute for Regenerative Medicine, Pittsburgh, Pennsylvania: Neoadjuvant chemotherapy combined with autophagy inhibition in pancreatic adenocarcinoma.

A. Viale, MD Anderson Cancer Center, Houston, Texas: Metabolic targeting of chemoresistance impacts clonal complexity in pancreatic tumors.

SESSION 6: General Discussion, Meeting Conclusions and Next Steps

R. Amaravadi, University of Pennsylvania, Philadelphia; **A. Kimmelman**, New York University, New York; and **J. Debnath**, University of California, San Francisco

Towards a Cure for Advanced Stage Ovarian Carcinoma

October 21–23

ARRANGED BY **J. Boyd,** Florida International University, Miami, Florida
S. DeFeo, Ovarian Cancer Research Alliance, New York, New York
D. Levine, New York University, New York
A. Moran, Ovarian Cancer Research Alliance, New York, New York

FUNDED BY **Ovarian Cancer Research Alliance, with additional funding provided by ImmunoGen and Clovis Oncology**

The large majority of epithelial ovarian carcinoma patients are diagnosed at an advanced stage (II-IV), and the great majority of these patients eventually succumb to their disease. However, clinical experience and epidemiologic evidence clearly indicate that a small fraction of these patients experience long-term survival (>12 years) and may effectively be considered as cured of disease. This Banbury Center meeting brought together a multidisciplinary group of thought leaders in the fields of ovarian cancer biology, genetics, epidemiology, surgery, and therapy in order to discuss existing data and to develop strategies that may provide optimal outcomes for the greatest proportion of ovarian cancer patients. In addition, strategies for exporting such a model(s) outside the context of major academic cancer centers were explored.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

SESSION 1: Overview/Intro

Chairperson: D. Levine, New York University, New York

Introduction and Meeting Objectives: **J. Boyd**, Florida International University, Miami; **D. Levine**, New York University, New York; and **A. Moran**, Ovarian Cancer Research Alliance, New York, New York

S. Narod, University of Toronto, Ontario, Canada: Strategy to avoid death.

C. Brown, A. Odunsi

D. Levine, S. Shah

M. Pike, Memorial Sloan Kettering Cancer Center, New York, New York: Risk factors related to long-term survivorship.
S. Shah, Memorial Sloan Kettering Cancer Center, New York, New York: Evolutionary dynamics of primary disease.

SESSION 2: Surgery

Chairperson: J. Boyd, Florida International University, Miami

D. Chi, Memorial Sloan Kettering Cancer Center, New York, New York: What are the limits of aggressive cytoreduction?
A. Fagotti, Catholic University of the Sacred Heart, Rome, Italy: Patient selection for primary cytoreduction—Biology and clinical?
C. Fotopoulou, Imperial College London, United Kingdom: Can you be cured after interval cytoreduction (complete path response).
C. Brown, Memorial Sloan Kettering Cancer Center, New York, New York: Optimal ovarian cancer care in underresourced populations.

SESSION 3: Adjuvant Therapy

Chairperson: S. Shah, Memorial Sloan Kettering Cancer Center, New York, New York

G. Mills, Oregon Health and Science University, Portland, Oregon: Systems biology perspectives for curing initial disease.
D. Armstrong, Johns Hopkins University, Baltimore, Maryland: Route of administration of cytotoxics.
A. Oza, Princess Margaret Hospital, Toronto, Ontario, Canada: Better initial treatments with targeted agents.

SESSION 4: Immunotherapy and Needs Assessment

Chairperson: D. Levine, New York University, New York

A. Odunsi, Roswell Park Cancer Institute, Buffalo, New York: Immunotherapy to cure in the primary setting.
A. Ellis, Ovarian Cancer Survivor and Research Advocate, White Plains, New York: Balancing hope versus reality in early survivorship?

SESSION 5: Meeting Summary and Next Steps

Chairpersons: J. Boyd, Florida International University, Miami, and **D. Levine**, New York University, New York

E. Baugh, S. Leighton

J. Boyd, S. Narod

Why Does the Neocortex have Layers and Columns?

October 28–31

ARRANGED BY **S. Ahmad,** Numenta, Redwood City, California
J. Gavornik, Boston University, Massachusetts
S. Mihalas, Allen Institute for Brain Science, Seattle, Washington

FUNDED BY **Numenta and Cold Spring Harbor Laboratory**

The neocortex is complex. Fortunately, most of this complex circuitry is remarkably preserved in all regions, suggesting that a canonical circuit consisting of columns and layers underlies much of what the neocortex does. Recent advances in recording technologies now enable detailed recording of activity in the microcircuitry of cortical columns, and new mapping technologies are rapidly increasing our knowledge of anatomical connections. However, despite these advances, the function of the laminar and columnar circuitry remains unclear and controversial. This Banbury Center meeting brought together experts from experimental, computational, and theoretical neurosciences to present their latest findings related to the anatomy, physiology, and function of cortical circuits. The goal was to develop a theoretical framework for understanding the function of stereotypical cortical circuits.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: S. Ahmad, Numenta, Redwood City, California;
J. Gavornik, Boston University, Massachusetts; and
S. Mihalas, Allen Institute for Brain Science, Seattle, Washington

SESSION 1: General Frameworks

Chairperson: S. Ahmad, Numenta, Redwood City, California

J. Hawkins, Numenta, Redwood City, California: Does the neocortex model objects in the same way that the entorhinal cortex models spaces?

M. Berry, Princeton University, New Jersey: Canonical computations in the neocortical microcircuit.

R. Rao, University of Washington, Seattle: Bayesian models of the neocortex: From predictive coding to POMDPs.

S. Mihalas, Allen Institute for Brain Science, Seattle, Washington: Hierarchical models with canonical local circuits.

SESSION 2: Anatomy/Thalamus I

Chairperson: S. Mihalas, Allen Institute for Brain Science, Seattle, Washington

K. Rockland, Boston University School of Medicine, Massachusetts: What if they're not stereotyped?

S. Aton, University of Michigan, Ann Arbor: State-dependent thalamocortical dynamics and visual system plasticity.

A. Zador, Cold Spring Harbor Laboratory: Statistics organization of long-range cortical projections.

M. Usrey, University of California, Davis: Cortical columns and layers facilitate feedforward and feedback network interactions between thalamus and cortex.

SESSION 3: Anatomy/Thalamus II

Chairperson: A. Pasupathy, University of Washington, Seattle

R. Bruno, Columbia University, New York, New York: The many input layers of the neocortex.

S. Brown, Johns Hopkins University, Baltimore, Maryland: Cortical layers and columns: Lessons from layer 6.

M. Sherman, University of Chicago, Illinois: Transthalamic corticocortical pathways.

M. Halassa, Massachusetts Institute of Technology, Cambridge, Massachusetts: Thalamic computations in cognitive control and flexibility.

SESSION 4: Sensory Systems

Chairperson: M. Geffen, University of Pennsylvania, Philadelphia

T. Engel, Cold Spring Harbor Laboratory: Cortical state and correlated variability across layers and columns.

R. Rao, A. Angelucci

A. Pasupathy, K. Nielsen

A. Pasupathy, University of Washington, Seattle: Cortical processing of occlusions: Role of feedback and inhibition.
A. Angelucci, University of Utah, Salt Lake City: Organization and function of feedback connections in early visual processing.
K. Nielsen, Johns Hopkins University, Baltimore, Maryland: Fine-scale organization of monkey visual cortex.
A. Hires, University of Southern California, Los Angeles: Circuit and behavioral mechanisms of feature learning in somatosensory cortex.

SESSION 5: Sequences, Prediction, and Cognition

Chairperson: A. Angelucci, University of Utah, Salt Lake City

M. Geffen, University of Pennsylvania, Philadelphia: Neuronal circuits for dynamic auditory processing.
J. Gavornik, Boston University, Massachusetts: Transient and durable temporal predictions in visual cortical circuits.
D. Schneider, New York University, New York: Learning, recalling, and ignoring self-generated sounds.
C. Constantinople, New York University, New York: Cortical computations during economic choice.

SESSION 6: Hippocampus/Navigation/Sensorimotor

Chairperson: J. Gavornik, Boston University, Massachusetts

M. Hasselmo, Boston University, Massachusetts: Coding in cortical circuits.

K. Rockland, G. Keller

C. Niell, University of Oregon, Eugene: Neural circuits for vision in action.
G. Shepherd, Northwestern University, Chicago, Illinois: Cortical circuit organization from a motor systems perspective.
S. Ahmad, Numenta, Redwood City, California: Interlaminar and intercolumnar models of sensorimotor prediction.
G. Keller, Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland: Internal models of the environment in the mouse cortex.

SESSION 7: Meeting Wrap-Up and Next Steps

S. Ahmad, Numenta, Redwood City, California; J. Gavornik, Boston University, Massachusetts; and S. Mihalas, Allen Institute for Brain Science, Seattle, Washington: Key conclusions and potential outputs.

The Lustgarten Foundation Scientific Advisory Board Meeting

November 11–13

ARRANGED BY **K. Kaplan,** Lustgarten Foundation, Woodbury, New York
D. Tuveson, Cold Spring Harbor Laboratory
R. Vizza, Lustgarten Foundation, Woodbury, New York

FUNDED BY **The Lustgarten Foundation**

Banbury was pleased to welcome back the Lustgarten Foundation for their 2018 Scientific Meeting, which provided an opportunity for the Scientific Advisory Board, as well as Foundation-supported investigators, to discuss research and strategy, evaluate performance, provide feedback for improvement, strengthen collaboration, and identify new ideas to bolster progress in the field.

R. Evans, Salk Institute for Biological Studies
D. Fearon, Cold Spring Harbor Laboratory
E. Fishman, Johns Hopkins Medicine
F. Froeling, Cold Spring Harbor Laboratory
C. Fuchs, Yale School of Medicine
L. Gruskiewicz, Lustgarten Foundation
T. Hunter, Salk Institute for Biological Studies
T. Jacks, Massachusetts Institute of Technology
E. Jaffee, Johns Hopkins School of Medicine
K. Kaplan, The Lustgarten Foundation for Pancreatic Cancer Research
D. Kelsen, Memorial Sloan Kettering Cancer Center
R. Mayer, Harvard University
D. Pellman, Dana-Farber Cancer Institute
B. Stillman, Cold Spring Harbor Laboratory
E. Stoeber, Lustgarten Foundation
H. Tiriac, Cold Spring Harbor Laboratory
D. Tuveson, Cold Spring Harbor Laboratory
F. Valsecchi, Lustgarten Foundation
R. Vizza, The Lustgarten Foundation for Pancreatic Cancer Research
B. Vogelstein, Howard Hughes Medical Institute and Johns Hopkins University
B. Wolpin, Harvard University Medical School
A. Yuille, Johns Hopkins University

Diverse Functions of Neutrophils in Cancer

November 27–30

ARRANGED BY **K. de Visser,** The Netherlands Cancer Institute, Amsterdam
M. Egeblad, Cold Spring Harbor Laboratory
P. Kubes, University of Calgary, Alberta, Canada

FUNDED BY **Northwell Health–Cold Spring Harbor Laboratory Affiliation**

Neutrophils are the most abundant leukocytes in blood, indispensable for combating microbial infections and facilitating wound healing. Recent studies have highlighted the diverse functions of neutrophils in cancer; however, it is still not clear when neutrophils are beneficial or detrimental to the host in the context of cancer. The goal of this Banbury meeting was to bring together cancer biologists with scientists and clinicians studying other aspects of neutrophils in order to facilitate discussions of recent findings on the functions of neutrophils, the classifications of neutrophils, and their potential as clinical biomarkers and therapeutic targets. A better understanding of the role of neutrophils is likely to provide opportunities for targeting of the antimetastatic effects of neutrophils, for immunomodulation acting via neutrophils, and, ultimately, for improving the treatment of cancer patients.

Welcoming Remarks: R. Leshan, Director, The Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: K. de Visser, The Netherlands Cancer Institute, Amsterdam;
M. Egeblad, Cold Spring Harbor Laboratory; and
P. Kubes, University of Calgary, Alberta, Canada

M. Egeblad, P. Kubes

P. Frenette, T. van den Berg

SESSION 1: Phenotyping Neutrophil Diversity and Targeting Neutrophils

Chairperson: P. Kubes, University of Calgary, Alberta, Canada

H. Goodridge, Cedars-Sinai Medical Center, Los Angeles, California: Myeloid cell heterogeneity, origins, and functional programming.

A. Hidalgo, Fundación CNIC, Madrid, Spain: Neutrophil diversity in homeostasis.

I. Udalova, University of Oxford, United Kingdom: Genomic control of neutrophil responses.

Z. Fridlender, Hadassah Medical Center, Jerusalem, Israel: Circulating neutrophils in human cancer—A functional and phenotypic (CyTOF) characterization.

E. Meylan, École Polytechnique Fédérale de Lausanne, Switzerland: Depletion strategies and targeting neutrophil metabolism in lung cancer.

SESSION 2: Sepsis, Stress, Infections, and Autoimmunity

Chairperson: B. Sherry, Feinstein Institute for Medical Research, Manhasset, New York

P. Kubes, University of Calgary, Alberta, Canada: Studying the neutrophil in health, injury, and repair.

P. Frenette, Albert Einstein College of Medicine, Bronx, New York: Influence of stress on neutrophil function.

P. Wang, Feinstein Institute for Medical Research, Manhasset, New York: CIRP Induces neutrophil reverse transendothelial migration.

M. Aziz, Feinstein Institute for Medical Research, Manhasset, New York: Neutrophils in sepsis: Role of cold-inducible RNA-binding protein.

M. Kaplan, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland: Neutrophil subsets and their role in systemic autoimmunity and organ damage.

K. de Visser, M. Pittet, Z. Fridlender

Z. Fridlender, A. Huttenlocher

P. Wang, M. Aziz

Z. Werb

B. Barnes, Feinstein Institute for Medical Research, Manhasset, New York: IRF5 genetic risk, spontaneous NETosis and autoimmunity.

SESSION 3: NETs and Imaging

Chairperson: P. Frenette, Albert Einstein College of Medicine, Bronx, New York

A. Zychlinsky, Max Planck Institute for Infection Biology, Berlin, Germany: NETs—The second function of chromatin.

B. Amulic, University of Bristol, United Kingdom: NETs in propagation of vascular inflammation.

D. Wagner, Boston's Children Hospital, Boston, Massachusetts: NETs in cancer.

M. Egeblad, Cold Spring Harbor Laboratory: Functions of neutrophil extracellular traps in metastasis.

A. Huttenlocher, University of Wisconsin, Madison: Live imaging of neutrophils in the tumor microenvironment.

SESSION 4: Roles of Neutrophils in Metastasis and Lung Cancer

Chairperson: M. Egeblad, Cold Spring Harbor Laboratory

M. Kaplan, B. Sherry, B. Barnes

V. Mittal, Weill Cornell Medicine, New York, New York: Mechanisms of neutrophil-mediated metastasis.

M. Pittet, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts: Neutrophils and lung cancer.

I. Malanchi, The Francis Crick Institute, London, United Kingdom: Cancer: The evil companion corrupting good behavior.

K. de Visser, The Netherlands Cancer Institute, Amsterdam: The genetic makeup of breast cancer dictates systemic neutrophilic inflammation.

D. Quail, McGill University, Montréal, Quebec, Canada: Obesity-associated inflammation and cancer metastasis.

SESSION 5: Antitumor and Proimmune Functions of Neutrophils

Chairperson: K. de Visser, The Netherlands Cancer Institute, Amsterdam

T. van den Berg, Sanquin Research and VU Medical Center, Amsterdam, Netherlands: Neutrophils kill antibody-opsonized cancer cells by trogoptosis.

T. Merghoub, Memorial Sloan Kettering Cancer Center, New York, New York: Contribution of innate immunity in T cell immunomodulatory antibody-based therapies.

E. Eruslanov, University of Pennsylvania, Philadelphia: Tumor-associated neutrophils with antigen-presenting cell features in human lung cancer.

Z. Granot, The Hebrew University of Jerusalem, Israel: Microenvironmental cues determine tumor cell susceptibility to neutrophil cytotoxicity.

Z. Werb, University of California, San Francisco: Regulation of neutrophils that are prometastasis or antimetastasis.

SESSION 6: Meeting Conclusions, Wrap-Up, and Next Steps

Chairpersons: M. Egeblad, Cold Spring Harbor Laboratory; **K. de Visser**, The Netherlands Cancer Institute, Amsterdam; and **P. Kubes**, University of Calgary, Alberta, Canada

Revolutionizing Agriculture with Synthetic Biology

December 2–5

ARRANGED BY **A. Hanson,** University of Florida, Gainesville
C. Vickers, CSIRO and The University of Queensland, Brisbane, Australia
E. Wurtzel, Lehman College, CUNY, Bronx, New York

FUNDED BY **Cold Spring Harbor Laboratory Corporate Sponsor Program**

Synthetic biology has the transformative potential to reconfigure metabolic pathways and other biological systems. Yet, thinking in the nascent plant synthetic biology sector tends still to be dominated by a "tinkering" mind-set and focuses on traditional mainline targets such as photosynthesis and producing plant pharmaceuticals in *Escherichia coli* or yeast. This Banbury meeting challenged an international group of experts to "think big" about using synthetic biology to install entirely new metabolic pathways and genetic circuitry in crops and other plants and to radically improve the efficiency of existing pathways and processes. The cross-sector discussions also touched on the necessity for continued advances in foundational knowledge, tools, and training as well as the real-world issues of government regulation and managing the public perception of plant synthetic biology.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: **E. Wurtzel,** Lehman College, The City University of New York, Bronx;
A. Hanson, University of Florida, Gainesville;
C. Vickers, CSIRO and The University of Queensland, Brisbane, Australia

R. Gershenzon, C. Vickers

J. Shanklin, J. Medford, J. Nemhauser, J. Matos

SESSION 1: Primary Metabolism

Chairperson: C. Vickers, CSIRO and The University of Queensland, Brisbane, Australia

T. Erb, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany: Fixing carbon fixation: Developing alternative solutions to the Calvin Benson Bassham cycle with synthetic biology.

S. Long, University of Illinois, Urbana: From math to field—Proof of concept in engineering photosynthesis for higher crop yield.

H. Millar, University of Western Australia, Crawley, Australia: Defining molecular targets for manipulation to improve efficiency of energy use processes to alter growth rate.

A. Hanson, University of Florida, Gainesville: "Maintenance respiration" as a next-gen target to improve crop productivity.

SESSION 2: Specialized Metabolism

Chairperson: J. Gershenzon, Max Planck Institute for Chemical Ecology, Jena, Germany

C. Vickers, CSIRO Synthetic Biology Future Science Platform, Brisbane, Australia: Carbon flux regulation for isoprenoid production.

E. Wurtzel, Lehman College, The City University of New York, Bronx: The complexity of engineering carotenoid biosynthesis.

T. Muranaka, Osaka University, Japan: Redesign of terpenoid biosynthetic pathway in plant by genome editing toward human health.

J. Shanklin, Brookhaven National Laboratory, Upton, New York: Controlling the diversity and yield of plant lipids.

B. Sattely, Stanford University, California: Discovery and engineering plant natural product biosynthesis.

SESSION 3: Eco-Interactions

Chairperson: A. Osbourn, John Innes Centre, Norwich, United Kingdom

P. Nikel, The Novo Nordisk Foundation Center for Biosustainability, Kongens Lyngby, Denmark: Engineering soil bacteria as biotechnological platforms.

J. Gershenzon, A. Hanson

D. Bhaya, P. Hines, E. Wurtzel

J. Nemhauser, N. Patron

A. Osbourn, H. Bouwmeester, N. Patron

G. Barbier, JoynBio, Boston, Massachusetts: Engineered microbes for agricultural use.

H. Bouwmeester, University of Amsterdam, Netherlands: Metabolic engineering to optimize the crop rhizosphere.

J. Gershenzon, Max Planck Institute for Chemical Ecology, Jena, Germany: Increasing the protective value of plant defense compounds: Plant-mediated deterrence of insect pest detoxification pathways.

SESSION 4: Radical Redesign

Chairperson: A. Hanson, University of Florida, Gainesville

B. Lindberg Møller, University of Copenhagen, Frederiksberg, Denmark: Bioengineering of structurally complex diterpenoids in yeast and photosynthetic cells.

J. Nemhauser, University of Washington, Seattle: Plant logic: Discovering and re-engineering design rules governing plant form.

J. Haseloff, University of Cambridge, United Kingdom: Marchantia as a simple prototype for bioengineering.

J. Medford, Colorado State University, Fort Collins: Plant synthetic biology: Following the path of electronics to produce genetic circuits with predictive functions and enabling synthetic biological desalination.

N. Patron, The Earlham Institute, Norfolk, United Kingdom: Toward predictable engineering of complex traits.

A. Kinney, Corteva Agriscience, Johnston, Iowa: Engineering oil seed composition.

SESSION 5: Parts-Prospecting and Tools

Chairperson: E. Wurtzel, Lehman College, The City University of New York, Bronx

A. Osbourn, John Innes Centre, Norwich, United Kingdom; Harnessing plant metabolism: From biosynthetic gene clusters to genomics and back.

M. Cooper, The University of Queensland, Brisbane, Australia: Genomic prediction and gene networks.

D. Orzáez, Universitat Politècnica de València, Spain: A toolbox of modular elements for orthogonal control of gene expression in plants.

R. Bock, Max Planck Institute of Molecular Plant Physiology, Potsdam, Germany: Taming plastids for synthetic biology.

N. Boyle, Colorado School of Mines, Golden: Next-generation metabolic models.

L. Nielsen, Novo Nordisk Foundation Center for Biosustainability, Kongens Lyngby, Denmark: A multitissue genome-scale metabolic modeling framework to guide plant metabolic engineering.

SESSION 6: Implementing SynBio

Chairperson: P. Hines, *Science*

D. Bhaya, National Science Foundation, Alexandria, Virginia: Plant synthetic biology: An NSF perspective.

S. Evans, Dow Agrosciences, Indianapolis, Indiana: Of plants and plants.

P. Rabinowicz, U.S. Department of Energy, Washington, D.C.: How BER supports synthetic biology: Research programs, user facilities, and enabling technologies.

SESSION 7: Wrap-Up

Chairperson: E. Wurtzel, Lehman College, The City University of New York, Bronx

Increasing Gender Diversity in the Biosciences

December 9–12

ARRANGED BY **C. Greider,** Johns Hopkins University, Baltimore, Maryland
J. Sheltzer, Cold Spring Harbor Laboratory

FUNDED BY **Cold Spring Harbor Laboratory**

The underrepresentation of women in bioscience careers is a fundamental problem because it represents a significant loss of talent and diversity. This meeting convened experts to identify practical institutional and extra-institutional approaches that can promote and support the advancement of women in science—connecting leaders from diverse fields to share lessons learned and to inspire innovative new ideas to achieve gender equity in biomedical research. The ultimate goal was to generate a list of general and adaptable "best practices" that institutions and communities can implement to stimulate and support the advancement of women in science.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor

Introduction and Meeting Objectives: **C. Greider,** Johns Hopkins University, Baltimore, Maryland, and **J. Sheltzer**, Cold Spring Harbor Laboratory

SESSION 1: Overview—Sexism and Academia

Chairperson: C. Greider, Johns Hopkins University, Baltimore, Maryland

J. Steitz, Yale University and Howard Hughes Medical Institute, New Haven, Connecticut: Reflections on (nearly) 50 years in academia.

C. Greider, L. Joshua-Tor

D. Ruebain, W. Copeland

S. Tilghman, Princeton University, New Jersey: Why leadership matters.

N. Hopkins, Massachusetts Institute of Technology, Cambridge: The leaky biology-to-biotech pipeline: Should universities step in?

SESSION 2: Implicit Biases and the Leaky Pipeline

Chairperson: D. Ruebain, Consultant, Equality, Diversity, and Inclusion in Higher Ed and Research Sectors, London, United Kingdom

J. Sheltzer, Cold Spring Harbor Laboratory: "Hidden" pipelines in biomedical research.

V. Valian, Hunter College, New York, New York: Remedying the (still too) slow advancement of women.

N. Dasgupta, University of Massachusetts at Amherst: Reducing bias and increasing diversity in STEM: What works and what doesn't.

L. Villa-Komaroff, Intersections SBD, Boston, Massachusetts: Why thinking fast makes changing slow: How cognitive processes impede gender equity.

SESSION 3: Academic Culture

Chairperson: J. Raymond, Stanford University, California

C. Greider, Johns Hopkins University, Baltimore, Maryland: Finally seeing the bubble: My experience as department chair.

L. Joshua-Tor, Cold Spring Harbor Laboratory: Let's get practical.

G. McDowell, Future of Research Organization, Abington, Massachusetts: Empowering the next generation of researchers to overcome adversity, bias, and sexual harassment.

J. Polka, ASAPbio, Cambridge, Massachusetts: Addressing gender bias in an evolving scientific publishing ecosystem.

SESSION 4: Government, Industry, and Funders

Chairperson: L. Villa-Komaroff, Intersections SBD, Boston, Massachusetts

H. Valantine, National Institutes of Health, Bethesda, Maryland: NIH's scientific approach to eliminating the gender leadership gap in biomedicine.

C. Greider, J. Sheltzer

J. Steitz, E. O'Shea

H. Valantine, N. Dasgupta

B. Stillman, N. Hopkins

A. Gammie, National Institute of General Medical Sciences, Bethesda, Maryland: Diversifying the biomedical research workforce.

E. O'Shea, Howard Hughes Medical Institute, Chevy Chase, Maryland: HHMI efforts to increase gender diversity in science.

W. Copeland, Celgene Corporation, Seattle, Washington: A corporate framework to quantify diversity and inclusion: Integrating metrics and business practice.

SESSION 5: Sexual Harassment and Legal Remedies

Chairperson: L. Joshua-Tor, Cold Spring Harbor Laboratory

M. Wadman, *Science*, Washington, D.C.: What I've learned while covering sexual harassment in science as a reporter.

B. McLaughlin, Vanderbilt University, Nashville, Tennessee: Disobedient ones.

A. Olivarius, McAllister Olivarius, Maidenhead, United Kingdom: War stories from a sexual harassment lawyer.

N. Chi Cantalupo, Barry University, Orlando, Florida: Taking a civil rights approach to gender-based violence in education.

C. Greider, Johns Hopkins University, Baltimore, Maryland [on behalf of Vicki Lundblad, Salk Institute of Biological Sciences, La Jolla, California]: Women scientists need to tell their stories.

SESSION 6: Organizational Approaches and Mentorship

Chairperson: N. Chi Cantalupo, Barry University, Orlando, Florida

J. Raymond, Stanford University, California: Diversity increases equity in peer review.

D. Ruebain, Consultant, EDI in Higher Education and Research Sectors, London, United Kingdom: The use of systemic change programmes to address chronic, long-standing underrepresentation and disadvantage.

J. Metcalf, Colorado State University, Fort Collins: 500 Women Scientists: A grassroots organization with a mission to serve society by making science open, inclusive, and accessible.

J. Wong, Boston University, Massachusetts: BU ARROWS: Creating organizational commitment and structured programs in academia to advance academic leaders in STEM.

SESSION 7: General Discussion, Meeting Conclusions, Outlining Next Steps

C. Greider, Johns Hopkins University, Baltimore, Maryland, and J. Sheltzer, Cold Spring Harbor Laboratory

J. Metcalf, J. Wong, G. McDowell, B. McLaughlin

Phase-Separated Assemblies in Cell Biology

December 16–19

ARRANGED BY **Arup K. Chakraborty,** Massachusetts Institute of Technology, Cambridge
G. Seydoux, Johns Hopkins University, Baltimore, Maryland
P. Sharp, Koch Institute for Integrative Cancer Research, Cambridge, Massachusetts
R. Young, Whitehead Institute for Biomedical Research, Cambridge, Massachusetts

FUNDED BY **Cold Spring Harbor Laboratory Corporate Sponsor Program**

An abundance of research has described intracellular condensation of proteins into liquids or hydrogels, and these studies have included work at the intersection of multiple fields, including molecular biology, chemistry, and physics. This Banbury meeting brought together an interdisciplinary group of experts to review functions and types of phase-separated assemblies in biology, develop a common conceptual framework and nomenclature, identify molecular code characteristics underlying assemblies, consider pathologies caused by aberrant phase-separated assemblies, and examine manipulation of phase-separated assemblies as a novel treatment target.

Welcoming Remarks: R. Leshan, Director, Banbury Center, Cold Spring Harbor Laboratory

Introduction and Meeting Objectives: **A. Chakraborty,** Massachusetts Institute of Technology, Cambridge;
G. Seydoux, Johns Hopkins University, Baltimore, Maryland;
R. Young, Whitehead Institute for Biomedical Research, Cambridge, Massachusetts

D. Lowe, J. Forman-Kay, G. Seydoux

P. Sharp, B. Stillman

SESSION 1: Physical Principles

Chairperson: R. Pappu, Washington University in St. Louis, Missouri

C. Brangwynne, Princeton University, New Jersey: Mechanics of phase separation.

A. Chakraborty, Massachusetts Institute of Technology, Cambridge: The role of phase separation in regulating transcription.

C. Fan Lee, Imperial College London, United Kingdom: Physics of passive and active emulsions.

C. Keating, Pennsylvania State University, University Park: Experimental model systems for compartmentalization based on phase separation.

T. Nott, University of Oxford, United Kingdom: Emergent properties of liquid-like membraneless organelles.

E. Siggia, Rockefeller University, New York, New York: Physical chemical properties of membrane bound organelles.

T. Mittag, St. Jude Children's Research Hospital, Memphis, Tennessee: How does the molecular grammar of low-complexity domains translate into phase transitions?

J. Forman-Kay, Hospital for Sick Children, Toronto, Ontario, Canada: Biophysical insights into neuronal granules and activity-dependent translation.

T. Hyman, Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany: A molecular grammar for phase separation of FUS family proteins.

R. Pappu, Washington University in St. Louis, Missouri: Connecting sequence to phase behavior using physical principles.

R. Tycko, National Institutes of Health, Bethesda, Maryland: Molecular structure of low-complexity protein assemblies: Information from magnetic resonance.

S. McKnight, University of Texas Southwestern Medical Center, Dallas, Texas: How do low-complexity domains achieve biological specificity?

SESSION 2: Molecular Determinants

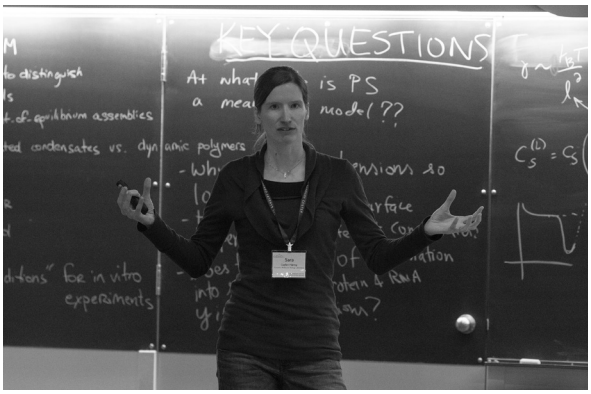
Chairperson: R. Parker, University of Colorado and Howard Hughes Medical Institute, Boulder, Colorado

SESSION 3: Imaging Dynamics

Chairperson: C. Fan Lee, Imperial College London, United Kingdom

C. Brangwynne, R. Pappu

R. Young



S. Cuylen-Häring

B. Tu, R. Tycko, S. McKnight, C. Keating

X. Darzacq, University of California, Berkeley: Imaging technologies provide new perspectives into phase separation.

I. Cissé, Massachusetts Institute of Technology, Cambridge: Superresolution imaging of transcription in live mammalian cells.

M. Botchan, University of California, Berkeley: *Drosophila* replication initiation factors that assemble on DNA through Cdk/cyclin regulated phase separation.

R. Parker, University of Colorado and Howard Hughes Medical Institute, Boulder, Colorado: Rnp granules.

R. Lehmann, New York University School of Medicine, New York: Nuclear and cytoplasmic germ granules in *Drosophila*: Connecting structure with function.

G. Seydoux, Johns Hopkins University, Baltimore, Maryland: RNA granule assembly in vivo and in vitro.

SESSION 4: Function

Chairperson: R. Lehmann, New York University School of Medicine, New York

R. Young, Whitehead Institute for Biomedical Research, Cambridge, Massachusetts: Transcriptional condensates.

G. Narlikar, University of California, San Francisco: Biophysical basis for phase separation processes in heterochromatin.

S. Cuylen-Häring, European Molecular Biology Laboratory, Heidelberg, Germany; Ki-67: From surfactant function to phase separation.

C. Mayr, Memorial Sloan Kettering Cancer Center, New York, New York: The interplay between the TIS granule and the ER creates a new subcellular compartment.

P. De Camilli, Yale University, New Haven, Connecticut: Phase separation as an organizing principle at neuronal synapses.

S. Petry, Princeton University, New Jersey: Phase separation enhances branching microtubule nucleation.

SESSION 5: Ensemble Properties

Chairperson: J. Forman-Kay, Hospital for Sick Children, Toronto, Ontario, Canada

B. Tu, University of Texas Southwestern Medical Center, Dallas, Texas: A metabolically regulated low-complexity domain.

P. Taylor, St. Jude Children's Research Hospital, Memphis, Tennessee: Dynamic RNA–protein assemblies and neurological disease.

D. Lowe, Novartis Institutes for BioMedical Research, Cambridge, Massachusetts: An industrial drug discovery perspective on intracellular phase condensates.

SESSION 6: Meeting Wrap-Up

Chairpersons: P. Sharp, Koch Institute for Integrative Cancer Research, Cambridge, Massachusetts, and **B. Stillman**, Cold Spring Harbor Laboratory

M. Botchan, G. Narlikar

DNA LEARNING CENTER

DNA LEARNING CENTER EXECUTIVE DIRECTOR'S REPORT

ADMINISTRATION

Lindsay Barone
Lauren Correri
Mary Lamont
Valerie Meszaros
David Micklos
Collette Riccardi

INSTRUCTION

Megan Capobianco
Elna Carrasco
Heather Cosel-Pieper
Alison Cucco
Cristina Fernandez-Marco
Melissa Lee
Christine Marizzi
Allison Mayle
Amanda McBrien
Pauline McGlone
Erin McKechnie
Bruce Nash
Michael Paul
Sharon Pepenella
Jeffry Petracca
Xiaoqun Catherine Zhang

BIOMEDIA

Cornel Ghiban
Susan Lauter
Jason Williams
Chun-hua Yang

Our education projects in Beijing and Suzhou, China, have matured to the point that they are now at the leading edge of biology lab instruction at the high school level.

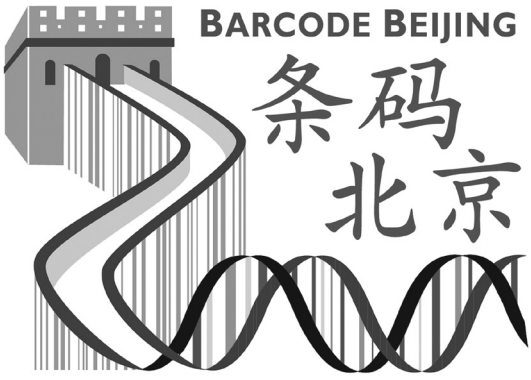
Secondary education in China is ruled by the omnipresent Gaokao, an extremely high-stakes exam taken in the 12th grade that determines a student's college placement. The rote memorization of facts required by this exam has made science labs a luxury throughout China and relegated what little practical work is done to after-school clubs. This was the case when, in 2011, we began working with Beijing 166 School, a large public school located within a mile of the Forbidden City and Tiananmen Square in the central district of Dongcheng. With training in biology, Principal Wang Lei was determined to live up to Beijing 166's new designation as the city's only "beacon" school in biology. She began by sending students for 3 weeks of summer workshops at the DNALC.

In 2014, we established a licensed DNALC at Beijing 166 with continuing grants from the Dongcheng Education Commission. The year-round licensed program now provides training for 275 students and 40 teachers per year: (1) 2-week winter workshops for students on DNA science and DNA barcoding, conducted at the DNALC; (2) 2-day spring workshops on human genetics and barcoding projects for students and teachers, at Beijing 166; (3) 3-week summer workshops on genetics, cloning, and forensic biology for students, at the DNALC; and (4) 2-day fall workshops on genome science for students and teachers at Beijing 166. In 2018, Beijing 166 also fulfilled its role as a beacon school by providing hands-on training for 63 teachers and webinar training for several thousand teachers from around China.

Dave Micklos (*left*) with Beijing 166 Principal Wang Lei (*right*)

We are now seeing encouraging results from this long-term collaboration, with more than 900 students currently taking required biology courses into which DNALC labs have been integrated. In addition, like the best high schools in the United States, Beijing 166 now offers a research course in DNA barcoding. With 20 students, this class meets two hours per week during the regular school day. This is to our knowledge the first elective lab science course in any high school in China.

The research course is complemented by *Barcode Beijing*, which is modeled after our *Urban Barcode Project* and is beginning to draw participants from around Beijing. Students have



Barcode Beijing logo

Visiting Beijing students during the *Forensic Detectives* summer camp

published 23 sequences in GenBank, the authoritative international database of DNA sequences. One Beijing 166 research team won first prize in the *Deng Feng Cup National High School Science and Technology Innovation Competition*, which is considered the most prestigious STEM award in China. Another Beijing 166 team won second place in the *Beijing Youth Science and Technology Innovation Competition*: Guo Li won the *Intel Young Talent* award, and Xiaoqiu An won the *Future Little Scientist* title.

Established in 2016 at the Suzhou Industrial Park (SIP), Cold Spring Harbor Asia (CSHA) DNALC is now delivering quality education programs. Under the leadership of U.S. scientist Jessica Talamas, 31 lab field trips, six summer camps, and two semester-long projects were completely translated for delivery in Chinese. In the past year, more than 8,500 people participated in science fairs and events at the DNALC; 2,683 students participated in lab field trips, and 221 students participated in weeklong summer and winter camps.

During the year, CSHA DNALC assumed an important role in reforming the Gaokao curriculum. Jiangsu Province—which includes Suzhou, Shanghai, and Nanjing—has the authority to administer its own Gaokao exam. This year the provincial government created a new biology major—counting 100 of 700 Gaokao points—to emphasize the importance of biotechnology in the Jiangsu economy. The DNALC collaborated with the SIP Bureau of Education and SIP Teacher Development Center to develop a set of molecular genetics labs to accompany the Gaokao. This is analogous to when the CSHL DNALC led in the incorporation of molecular genetics labs into Advanced Placement Biology, which catalyzed biotech lab teaching in the United States.

(*Left*) Education Director Jessica Talamas, Ph.D.; (*right*) teacher training at CSHA DNALC

During the fall semester, an experiment was conducted at Soochow University High School: 100 students took the traditional curriculum, and 100 students took the integrated lab curriculum, along with an intensive DNALC camp. Comparing pre- and post-scores on questions from previous Gaokao exams, students in experimental classes raised their scores 15% higher than students in the traditional classes. This is especially significant because most Chinese believe that time away from rote drilling will actually decrease performance.

The sorts of classroom experiments being done at Beijing 166 and Soochow University High School have been absent from Chinese secondary education to date. We are beginning to have some data that doing DNA labs can actually improve student Gaokao results, perhaps even beyond biology. The average Gaokao score of the Beijing 166 Life Science class, which receives intensive exposure to hands-on instruction, has increased in each of the last 3 years.

During the year, we also continued negotiations with officials in Guangzhou to develop a branded DNALC. Guangzhou, formerly known as Canton, is the capital of Guangdong Province, which has the fifth largest gross domestic product in the world. The Guangzhou center would be a hybrid of current Chinese licenses—located at a secondary school and serving local students, like Beijing 166, and providing field trips and summer camps for students throughout Guangdong Province, similar to CSHA DNALC.

DNA Barcoding and Metabarcoding

The DNALC administers three distinct programs for using DNA barcoding in high school research. *Barcode Long Island (BLI),* funded by the National Institutes of Health (NIH), involves students in "campaigns" to compare biodiversity across Long Island. The *Urban Barcode Project (UBP),* funded by the Thompson Family Foundation, and *Urban Barcode Research Program (UBRP),* funded by matching grants from the Pinkerton Foundation and Simons Foundation, involve students in independent research on biodiversity in New York City (NYC). *BLI* and *UBP* students are mentored by classroom science teachers, whereas *UBRP* students are mentored by scientists from NYC research institutions.

We continued to move students into the realm of "big data" by expanding our microbiome and environmental DNA (eDNA) efforts. Building on our NIH Big Data to Knowledge supplement efforts to adapt microbiome research for high school students, we started supporting fish eDNA experiments through a 1-year grant from the Lounsbery Foundation. In this pilot project, seven mentors worked with 11 teams to analyze fish diversity in the Hudson and East Rivers, eDNA degradation rates in response to increased levels of salinity, and effects of shoreline habitat, oyster beds, or eelgrass density on fish biodiversity. Like microbiome analyses, fish eDNA uses next-generation sequencing (NGS) to analyze thousands of DNA sequences representing the biodiversity in an environmental sample. DNA isolated from water is amplified by PCR to a variable region, and NGS reads identify the variety and abundance of microbial species from different locations.

For both eDNA and microbiome research, we are working to optimize and reduce the cost of DNA isolation, amplicon indexing, and sequencing library cleanup. An important component of these efforts is reducing the cost of indexing. Millions of metabarcoding sequences from hundreds of student samples can be sequenced together in one NGS run, theoretically making this research scalable and affordable. This kind of multiplex sequencing is made possible by attaching unique "index" sequences to the DNA in each sample, allowing all of the DNA sequences to be assigned to samples. We are developing an inexpensive quadruple index system that builds indexes into PCR primers in a two-step process, aiming to reduce the total cost of microbiome sequencing from about $50 to $10–$15 per sample.

In our previous microbiome workflow, students struggled to complete analyses. To make analysis more accessible, we developed a new Purple Line of *DNA Subway*, our online bioinformatics

Nineteen Long Island teachers attended a weeklong workshop in July to be trained to lead student teams conducting metabarcoding projects

platform developed through CyVerse. The Purple Line is a very approachable, browser-based graphical user interface for QIIME2, a microbiome analysis workflow. Data can be uploaded to CyVerse and analyzed using this web-based implementation. Importantly, the Purple Line has been modified to handle marine vertebrate eDNA, expanding the "traditional" use of QIIME2 and letting us support students who want to study microbes or fish.

During the summer, 19 *BLI* faculty mentors received training in metabarcoding research, covering project design and sample collection, biochemistry, and data analysis. As the year ends, 12 teams participating in *BLI* are working on metabarcoding projects, including microbiome (four teams, nine students) and eDNA (eight teams, 21 students) research—whereas two teams with six students are continuing work on eDNA research in New Jersey, extending their efforts as part of the Lounsbery grant, which ended in September.

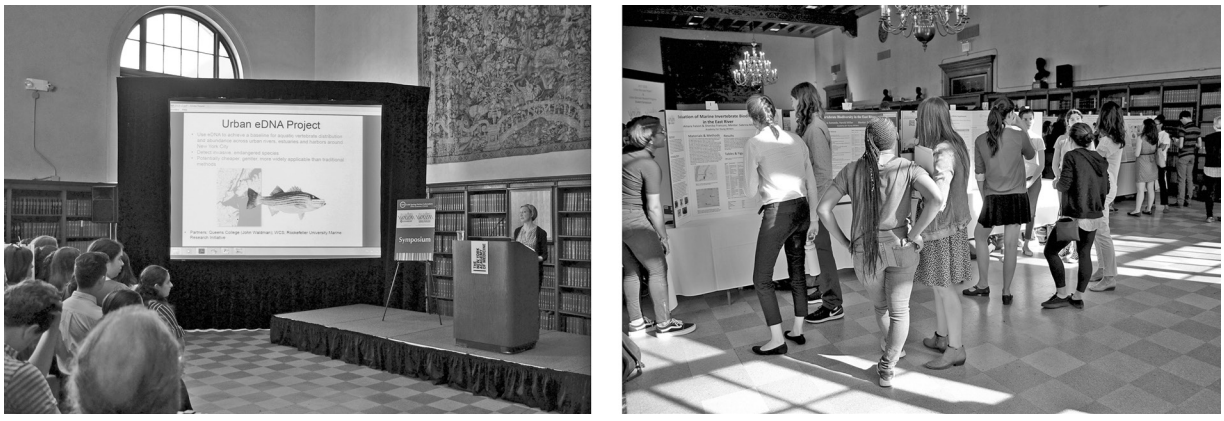
The 2018 *BLI* program included 382 students doing traditional DNA barcoding and microbiome projects. *BLI* teams represented 39 public and private high schools from Suffolk, Nassau, Kings, and Queens counties; 13% of participants were African American or Latino. To complete their research in time for the annual research symposium, 151 students from 55 teams attended 15 open lab sessions held at the Dolan DNALC, DNALC *West*, Stony Brook University (SBU), or Brookhaven National Laboratory (BNL), whereas 187 students from 64 teams borrowed equipment footlockers for use at school. Teams processed more than 1,200 samples, resulting in more than 1,900 sequencing reads. Students published 285 sequences in GenBank, including 30 novel barcode sequences and 52 with sequence polymorphisms.

One-hundred and nine DNA barcoding and microbiome projects were presented at the annual *BLI* research symposium on June 5, 2018, at CSHL, which included biodiversity studies of plants,

(*Left*) Dr. George Amato speaks to *BLI* Symposium attendees and (*right*) Dr. Bruce Stillman talks to *BLI* participants at their poster

invertebrates, fungi, algae, and lichens, plus microbiome studies of water, soil, invertebrates, plants, and disease vectors. Dr. George Amato of the Sackler Institute for Comparative Genomics at the American Museum of Natural History gave the keynote address on interesting applications of DNA barcoding. *BLI* students received a number of awards, including top honors at the Long Island Science and Engineering Fair (LISEF) and the Long Island Science Congress (LISC), in addition to awards at numerous other competitions. Two teams from William Floyd High School have accepted manuscripts in the *Journal of Emerging Investigators* (JEI) on their microbiome project data; a third team has a manuscript in preparation. Additionally, *BLI* mentors John Halloran from Connetquot High School and Daniel Williams from Shelter Island High School presented their students' DNA barcoding and microbiome results and experiences as mentors in the program at the NIH SEPA SciEd Conference; John attended on a scholarship awarded by the NIH SEPA program.

The 2018 *UBP* and *UBRP* programs included 200 students working in 74 teams and represented 35 public and eight private high schools from NYC. *UBP* students made ample use of DNALC resources: 142 students from 23 teams attended open lab sessions at *Harlem DNA Lab*, whereas 120 students from 40 teams borrowed equipment footlockers for use at school. Teams collected and processed more than 1,320 samples for DNA sequencing, resulting in more than 3,200 single sequences and 100 million NGS reads, and produced 22 new GenBank entries. The annual research symposium on May 24, 2018, at the New York Academy of Medicine showcased 74 projects and included a keynote speech by Dr. Elizabeth Alter, York University, on the most recent molecular approaches to surveying aquatic biodiversity in urban habitats. Two *UBRP*



(*Left*) Dr. Elizabeth Alter addresses *UBP* and *UBRP* student researchers who (*right*) later shared project results at a poster session

teams were recognized with an outstanding poster award at the event: one for the discovery of a novel RNA virus found in arthropods in the New York area that they named the "Jiminy Cricket" virus, and another project that used DNA barcoding to reclassify American sidewinders. The winners for the *UBP* examined the relationship between water quality at various locations of Van Cortlandt Park and the species of snail present. Results indicate that all locations investigated contained many more pollution-resistant lunged snails than pollution-sensitive gilled snails.

Several scientific, peer-reviewed publications have been coauthored by citizen scientists participating in DNALC's signature DNA barcoding programs in the last year. Twenty students were coauthors of a "bioblitz" study of the plants and animals of Marine Park, published in *PLOS One*. Three *UBRP* teams published studies about fraudulent herbal medicines and other marketplace replacements in *Open Life Sciences*, *The Finger Lakes Journal of Secondary Scien*ce, and MDPI's *Foods*. This demonstrates that the established DNA barcoding workflow can be used to teach molecular techniques and bioinformatics in a contextual situated learning design that is relevant personally as well as to the scientific community.

This year, 186 students across all three barcoding programs (*BLI, UBP,* and *UBRP*) took surveys as a part of our ongoing effort to monitor the impact of participation in science research. Participants were asked about their experiences in the programs, how much they had learned, and how they felt about science. The students were overwhelmingly proud of the research they had done (92.6%) and felt that the approach to problem-solving they learned through DNA barcoding research would be helpful in future science courses (92%) and careers (86.9%). They also reported that research participation had altered their desire to pursue science in the future, with 83% indicating they were more interested in continuing to study science and 87.1% indicating they were more interested in studying biology specifically. Overall, our results suggest that DNA barcoding effectively demystifies the process of science research and encourages students to continue pursuing science as a potential career path.

New NSF Faculty Training Program

In October we began a new DNA barcoding project with $2 million funding from the National Science Foundation program, Improving Undergraduate STEM Education (IUSE). The goal of this 5-year project is to move educators along a continuum of increasing expertise and broader student involvement in DNA barcoding course-based undergraduate research experiences (CUREs). Simultaneously, we will develop advanced CUREs in microbiome and eDNA analysis to introduce students to next-generation sequencing and data science. Project co-PIs have each implemented DNA barcoding in a range of undergraduate institutions: James Madison University (JMU), a public 4-year institution; City Tech, a dual 2- and 4-year urban university serving black and Hispanic students; Bowie State University (BSU), a historically black institution; and Austin Community College (ACC), a 2-year, Hispanic-serving institution. JMU has built a model DNA barcoding CURE serving 500 students per semester as a replacement for freshman biology. Building on this experience and strength in diversity, our IUSE project will disseminate and evaluate DNA barcoding as a flexible platform for scaling CUREs in a variety of contexts.

JMU students and faculty swabbed snakes to collect microbiomes as part of the DNA barcoding CURE at JMU

At the outset of the project, the partner institutions will test new technology to substantially simplify DNA barcoding and reduce the cost of metabarcoding. They will also develop mentoring strategies by helping one

another along the CURE continuum. A series of summer workshops will provide intensive training for 80 faculty from around the country. Then extensive mentoring, available DNA sequencing, and "extended collaborative support" (borrowed from the computational realm) will help ensure broad and lasting implementation. As faculty gain expertise, they will be invited to mentor other faculty with less experience, expanding the reach of the project.

This project will also increase our understanding of the impact of CUREs on student self-confidence and persistence in STEM. We will compare DNA barcoding CUREs with an established Howard Hughes Medical Institute program and the Freshman Research Initiative at UT-Austin to determine whether different types of CUREs achieve similar student outcomes. By comparing these programs, we also hope to determine which of the key research elements—engaging in scientific practices, collaborating, examining relevant problems, exploring questions with unknown answers, and reiterating experiments—most closely correlates with positive student outcomes. This could help provide a "prescription" for the most time- and cost-effective CUREs.

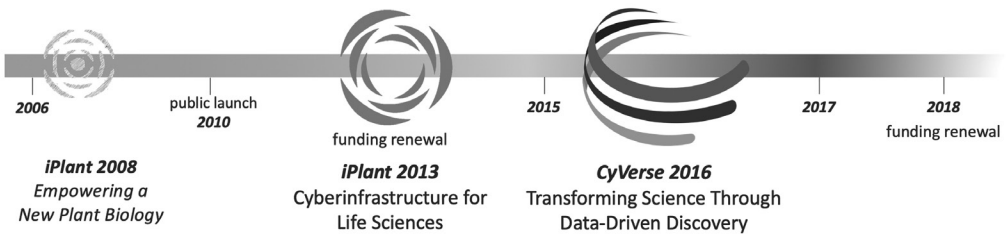
NSF CyVerse

The year 2018 marks 10 years of the CyVerse project, which was originally funded in 2008 as the *iPlant Collaborative* to bring plant biology into the era of genomics and big data. In our original proposal to the National Science Foundation, the DNALC was tasked with developing audio and video podcasts to communicate about the "Grand Challenges" of data-driven plant science and develop a nationwide workshop program to train 1,000 science teachers on the use of *iPlant*. Since that time, we have provided intensive 2-day training workshops on computational tools and data science for research and education to more than 3,446 participants. The DNALC shaped this project in significant ways—most prominently through the development of *DNA Subway*, which provides perhaps the most popular interface available to high-performance and cloud computing for biology education and course-based research.

As a rule, NSF does not fund projects for more than 10 years; however, the success of CyVerse was recognized as integral to the life science community. As a result, another 5-year award to the project was made in September. Although this continuing award was a reduction in scope, the decision was made to strengthen DNALC's involvement by formally making DNALC Executive Director Dave Micklos a co-principal investigator (although he previously served on the Executive Team for the project). The additional $800,000 award to the DNALC allows us to continue our education and outreach role with an increasing focus on data science education and course-based research, complementing many of our existing programs.

CyVerse Timeline 2008–2018

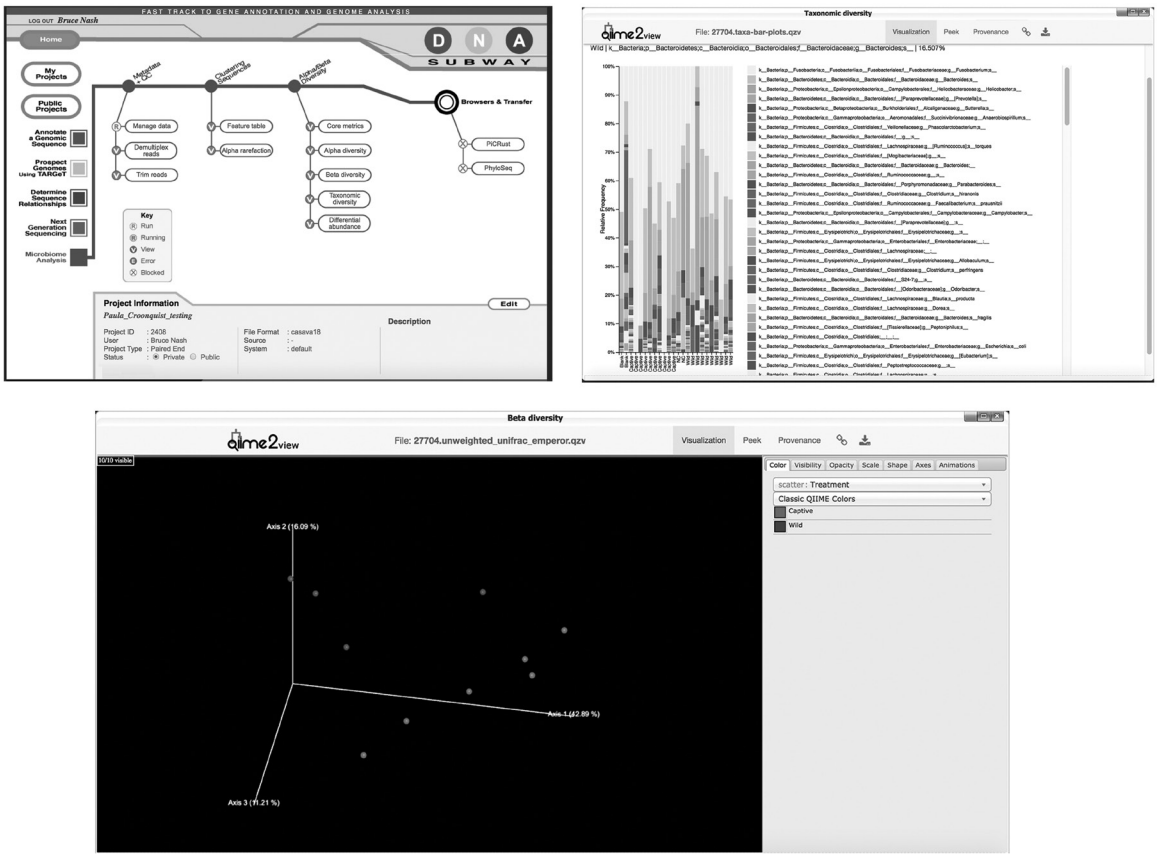
In 2018, *DNA Subway* had 30,563 registered users, 66,157 visits (~8% increase from 2017), and 1.25 million page views (~15% increase from 2017). Students created 39,283 projects (~7% increase from 2017) across the five *Subway* lines. A particular highlight for *DNA Subway* development was the completion of upgrades to the RNA-Seq Green Line. The Green Line now implements Kallisto,



one of the fastest and most accurate methods for quantifying transcript abundance, as well as Sleuth, which calculates differential abundance of transcripts and presents the results in a fully interactive R-Shiny app (including interactive tables and plots). Workflows that could take more than a week (mostly spent waiting on high-performance computing queues) now can be done in hours using cloud resources. We also implemented design upgrades to improve Green Line's user interface. These design improvements will also be carried over to the Purple Line eDNA/metabarcoding workflow, which was premiered this year with a workshop for 29 educators at the American Society for Microbiology's Conference on Undergraduate Education (ASMCUE) in Austin, Texas. In addition, we reached 336 researchers and educators at 12 2-day workshops across nine institutions and 1,550 attendees at 10 conferences and other training events.

NSF MaizeCODE

MaizeCODE is a collaboration between researchers at CSHL and New York University to create a comprehensive reference encyclopedia of DNA sequences that control transcription in maize and its wild ancestor, teosinte. MaizeCODE is developing data that will provide an important resource enabling breeders and plant scientists to improve crop traits, such as disease resistance, drought tolerance, and yield, by providing high-quality genome sequences paired with diverse molecular data. As outreach lead on this project, we continued our work to develop a simple system that allows students to participate in the annotation of the maize genome.



Purple Line tools (*top left*) show microbiome samples from wild and captive wolves (*top right*). The beta diversity shows that the two groups have different microbiomes, although both sets of wolves also vary within their group (*bottom*).

The sophistication of gene prediction programs and the abundance of RNA-based evidence for the maize genome would lead one to believe that manual curation of gene models is no longer necessary. However, we used a quality metrics algorithm to identify 17,225 of 130,330 (13%) protein-coding transcripts in the current maize reference genome that have discrepancies with available biological evidence. Also surprising, despite the fact that the maize genome is in its fourth version since its publication a decade ago, prior to our work only several hundred maize genes had been manually curated by a human being.

Over the year, we worked with eight graduate students to curate 86 transcript models flagged by quality metrics and a complementary method using the Gramene gene tree visualizer. All of the triaged models had significant errors—including missing or extra exons, noncanonical splice sites, and incorrect UTRs. We then used a graphical annotator, Apollo, to edit the gene models and display the corrected ones to Gramene. A correct transcript model existed for ~60% of genes (or transcripts) flagged by quality metrics; we attribute this to the convention of elevating the transcript with the longest coding sequence (CDS) to the canonical, or first, position. The remaining 40% of flagged genes resulted in novel annotations and represent a manual curation space of ~10% of the maize genome (approximately 4,000 protein-coding genes).

We presented the preliminary results of the MaizeCODE annotation project at the 60th Maize Genetics Conference in Saint-Malo, France (March) and the 60th American Society of Plant Biologists meeting in Montreal, Canada (July). We are now working to simplify the workflow so that our double triage system can be used to support the community curation of eukaryotic genomes by scientists, students, and potentially even citizen scientists.

Biotechnology in American High Schools: Then and Now

In 2017, we received NSF funding to administer a survey to 12,773 biology faculty—approximately one-quarter of active high school biology faculty in the country. The survey was based on earlier work done by the DNALC through an early grant from NSF's Advanced Technological Education (ATE) program to measure the integration of biotechnology into U.S. classrooms. By comparing responses, we were able to measure changes in teacher attitudes and training, teaching constraints, and student exposures to molecular biology techniques in the classroom, focusing our lab questions on six key techniques: transformation, restriction analysis, PCR, DNA sequencing, DNA recombination, and plasmid evolution.

Our preliminary analysis highlighted a number of important findings. Molecular biology labs are being integrated across the curriculum, most often in AP Biology (54% of the 2018 teachers), but increasingly in general biology courses as well (28% today vs. 21% in 1998). There has also been a major increase in schools offering biotechnology-based elective courses: 35% of today's teachers have some type of biotechnology elective in their schools (vs. 16% in 1998).

Teacher preparation and professionalization has changed, too. Today's teachers feel there are fewer opportunities to obtain biotechnology training at workshops and summer institutes than teachers in 1998, and there has been a large shift in the number of teachers who rely on formal graduate coursework to develop expertise (41% today vs. 24% in 1998). Teachers now are less likely to belong to any professional organizations and rarely attend professional meetings. Teachers are also less likely to engage with students in extracurricular science activities, including after-school research and science fairs and competitions.

Overall, the results of this project highlight a set of challenges for the future. An additional layer of analysis is set to be completed in 2019, when we enhance the data set with survey responses collected between 1987 and 1993 at our *DNA Science* Workshops. This third time point has a set of semantic differentials, which provide a unique way to measure attitudes. Teachers rated their reactions to "recombinant DNA, biotechnology" and "myself as a biology teacher" on 18 scales of polar

adjectives ("important–unimportant," "messy–neat," "dangerous–safe," etc.). This was repeated in the 2018 survey and will enable a longer perspective on shifts in teacher attitudes that may impact their ability or willingness to implement new and challenging materials in their classrooms.

Expanding Our Footprint in New York

During the year we continued on our long, and sometime tortuous, path to establish a *DNALC NYC* at City Tech in Brooklyn. After a year of back and forth between lawyers, the contract for a 30-year, no-cost lease of 17,500 square feet of space at City Tech was finished in September and sent to the Dormitory Authority of New York (DASNY) in Albany. In the meantime, Centerbrook Architects pushed toward finishing construction documents for a fantastic plan with six teaching labs, two bioinformatics labs, and an exhibit gallery.

When open sometime in 2020, *DNALC NYC* will run a full schedule of activities like those currently available at the Dolan DNALC—academic-year lab field trips, summer DNA camps, and weekend family activities. The no-cost lease means that all money raised for operating costs or endowment will go toward programs. Our business plan provides scholarships for at least 50% of students taking academic-year field trips.

City Tech President Russ Hotzler continued to encourage us and provided a temporary teaching lab in the Pearl Street Building, just two floors above our future permanent space. With City Tech faculty member Jeremy Seto as co-principal investigator of our new NSF program on DNA barcoding, we are establishing collaborations that will support a very large program to provide CUREs. With two labs dedicated to student research, we anticipate at least 500 CUNY students per year will make use of the DNALC's integrated systems for DNA barcoding, microbiomes, RNA sequence analysis, and genome annotation.

As we continued to face hurdles in Brooklyn, our project to develop a licensed DNALC at Regeneron Pharmaceuticals gained new momentum. Although the Regeneron Board formally approved the deal in June 2017, space designated for the center on its Tarrytown campus was taken over for mission-critical research. The project got back on track in October, when space was identified on their new administrative campus in a former IBM building in Sleepy Hollow. BAM architects then got busy on a lovely design for a 4,500 square foot facility, with two teaching labs and a large prep lab with space for assembling footlocker kits. Firmly on schedule to open in November 2019, the Regeneron DNALC center will serve Westchester, Rockland, Putnam, and Fairfield counties, which are demographically similar to Long Island.

International Partnerships

As we continued our quixotic quest for permanent space for a licensed DNALC in Mexico City, the symposium of the first *Mexico City Urban Barcode Project* was held in Mexico City in May 2018. The event was organized by Central DNA Company, Ministry of Agriculture (SAGARPA), National Institute of Genomic Medicine, and National Committee for the Knowledge and Use of Biodiversity (CONABIO). DNALC educator Cristina Fernandez Marco presented the keynote address. Eight schools fielded 23 teams with 68 students and 14 mentors. Student projects examined questions of ecology, biodiversity, physiology, medicine, and health—with three-quarters of barcodes done for plants or insects. The winning project was "Micro-inhabitants of the Soil of Mexico City," and the honorable mention was "Study of Rhizobacteria to Promote Plant Growth." Members of the winning teams and their teachers won free tuition to a *DNA Science* workshop at the DNALC in August:

Plantel Vallejo (UNAM), Escuela Nacional Colegio de Ciencias y Humanidades:
Students: Ximena Mendoza Tepach, Melanie Jimenez Lojero, Itzel Analí Sánchez Aguilar
Mentor: Sra. Issis Yolotzin Alvarado Sanchez

Michael Okoro (*left*) soon after arriving in New York in 2013 and (*right*) in May at his graduation from NYU

Universidad Nacional Autónoma de México:
Students: Raul Motte Nava, Agustin Celestino Lopez, Jonathan Perez Gonzalez, Marco Soriano Pimentel, Alejandro Gonzalez de la Luz
Mentor: Professor Monica Perez Ibarara

In February, Dave returned to Nigeria with our colleague George Ude to receive an honorary degree during commencement at Godfrey Okoye University (GOU). Several years ago, we worked with George to set up a modern biotech lab at this tiny Catholic university in the southeastern part of the country. While there, we taught a course in DNA barcoding and looked at a site for a DNALC on the university's new campus in Iguomo. Campus development is awaiting a tarmac access road, which will put it within 15 minutes of the international airport.

At year's end we bid farewell to Michael Okoro, who came first as a GOU student in 2013 to do his senior thesis on barcoding indigenous medicinal plants of Nigeria. After finishing his government service as a biology teacher in the unsettled north of the country, he returned in 2016 as our first DNALC Fellow—with the objective of preparing him for graduate school in the United States. Our plan worked, and in November of that year we learned that Michael had been accepted into the master's program at New York University! However, Michael had not realized the enormity of the tuition—he had misconstrued the per credit cost as the cost for an entire semester. At the same time, I realized that he had absolutely no resources to support this venture. The winter semester was to begin in six weeks, and we were both crestfallen.

Then someone had the brilliant idea to apply for a supplement to the MaizeCODE project. Program officer Diane Okamura suggested that Mike's college work would fit joint funding from the Gates Foundation's Bread for the World program. We put together a quick application, and before leaving for Christmas, we got word that the supplement to support Michael's fees and living expenses on the CSHL campus would be funded. It was truly a miracle. The rest, as they say, is history. Michael worked hard and finished his degree in molecular biotechnology in just 18 months, leaving time in the summer to take the CSHL course in Frontiers and Techniques in Plant Science and to work in the cornfield with Dave Jackson's group. During the upcoming year, he will continue his fellowship at GOU as he does the legwork in setting up a satellite DNALC there.

Dissemination at Professional Meetings

As in previous years, we disseminated our programs through numerous presentations at meetings. At the National Association of Biology Teachers (NABT), our new Ötzi the Iceman activity was

introduced along with presentations on DNA barcoding and personalized medicine. Our DNA barcoding approach was also presented at the Austrian Citizen Science Conference, Long Island Natural History Conference, Biocódigos de Barras Urbanos CDMX Symposium in Mexico City, and NIH SEPA SciEd Conference.

Our data science programs were presented at the International Plant and Animal Genome Conference, Bioinformatics Open Source Conference, and Biological Science Data Meeting, whereas our work on improving genome annotation with students, developed through MaizeCODE, was presented at the American Society of Plant Biology and Maize Genetics Conferences. Finally, our metabarcoding programs were presented at the American Fisheries Society Meeting, National Conference on Environmental Marine DNA, SciEd, and the Plant and Animal Genome (PAG) Conference.

Lab Instruction and Outreach

In 2018, 21,176 students attended lab field trips at our three facilities: Dolan DNA Learning Center, DNALC *West,* and *Harlem DNA Lab*. In-school instruction programs reached 7,905 students, and 1,347 students attended weeklong camps, including several from abroad, as part of collaborations in Mexico and Italy. Footlocker kits were used by 1,983 students, including 307 who conducted independent research through *UBP*, *UBRP*, or *BLI*.

Our instructional reach is increased by international licenses in China, discussed previously, as well as a license to the University of Notre Dame. In 2018, 1,275 students from 25 different schools participated in hands-on molecular biology labs supported by the DNALC at Notre Dame (DNALC-ND). Under the leadership of director Dr. Amy Stark, instructional programs included lab field trips to the DNALC-ND, in-school instruction, and engagement at regional and state-level science fairs. More than 100 students from around the world participated in weeklong residential and day camps and a new advanced workshop.

Grants from TEVA Pharmaceuticals and National Grid Foundation provided field trip and in-school instruction scholarships for more than 2,800 students from Long Island public school districts—including Amityville, Brentwood, Central Islip, Malverne, Roosevelt, Uniondale, and Valley Stream. The William Townsend Porter Foundation provided scholarships for a portion of the 2,451 students (75% of total students) from Title I schools who visited the *Harlem DNA Lab* for field trips. An additional 15 students from IS 59 in Queens received sequential lab instruction at DNALC *West* as part of an ongoing collaboration with Northwell Health.

Students from Partner Member St. David's School take advantage of field trips at the DNALC (*left*) as well as in school (*right*). (Images courtesy St. David's School)

(*Left*) We offered an agar art session in the fall *Saturday DNA!* series. (*Right*) Amanda McBrien leads WiSE young women at the *Fun with DNA* camp held on the CSHL campus

The *Partner Member Program* continued to provide custom science sequences and advanced electives for eight schools (primarily independent) in the tristate region.

- Our newest member, Grace Church School, offered a summer program that included using DNA barcoding to survey biodiversity of the East River.
- At Marymount School of New York, *Genome Science* experiments were incorporated as key parts of the biology curriculum, and students in molecular biology started a new project to analyze environmental DNA (eDNA) from NYC parks and the Hudson River.
- Research teams from Sacred Heart Greenwich used DNA barcoding to identify macro-invertebrates in the Gowanus canal and to assess the biodiversity of ants in both urban and suburban locations and used eDNA to identify fish species in the Mianus River.
- Lycée Français de New York continued to use DNA barcoding in research, developed a new forensics elective, and offered *Human Genomics* and *Green Genes* camps.
- The Chapin School implemented genetics programs at several grade levels, including the advanced Molecular Genetics elective.
- St. David's School integrated basic genetics and DNA barcoding programs with existing curricula in grades five and eight.

As part of ongoing partnerships, select students from Cold Spring Harbor High School and St. Dominic's High School received daily instruction from DNALC educators. Students enrolled in genomic and molecular biology electives visited the DNALC each afternoon for customized lab sequences in DNA and genome science, DNA barcoding, RNA interference, and gene expression. For the first time, all of the students in both classes presented posters at the annual *Barcode Long Island* Symposium.

We had 5,244 visitors to the *Ötzi the Iceman* exhibit, including the general public or as part of a field trip. This year we began to redesign the rest of our outdated exhibit space. The new exhibit—slated to open in 2019—will feature a look at the history of life on Earth and some of the key developments that have allowed life to flourish on our planet. The exhibit will explore the processes and outcomes of evolution, showcasing several interesting human evolution stories. A touch screen and interactive chromosome map will allow students to explore various parts of the human genome. Finally, a human variation wall will highlight how traits manifest themselves in different people—holding a literal mirror up to our visitors and allowing them to explore some of their own traits.

Nine *Saturday DNA!* sessions drew 238 participants. Through short hands-on labs, participants explored DNA isolation techniques, crime scene analysis, gel electrophoresis, the genetics of lactose intolerance, microscopy, genetic engineering, Mendelian inheritance, and industrial enzymology. Two microbial masterpieces created at our Agar Art session won the American Society for Microbiology (ASM) Agar Art contest in the Agar Art Kids and Agar Art Maker segments. In these workshops, participants learned how laboratory techniques that scientists commonly use to study the living world could also be used to create unique works of art. To further leverage this STEAM activity, ASM invited Christine Marizzi to coauthor a publication titled "The Many Dimensions of Microbial Art," which was published in the December issue of *SciArt Magazine*. In addition, DNALC staff presented student activities as part of Biodiversity Day at the Browning School in NYC and the SUBMERGE Science Festival at Hudson River Park.

As part of our ongoing partnership with CSHL Women in Science and Engineering (WiSE), we hosted the second WiSE *Fun with DNA* summer camp. Held on the main campus of CSHL in Delbrück Laboratory, 20 young female science enthusiasts enjoyed the standard *Fun with DNA* lab content, but also had the opportunity to meet and interact with enthusiastic female role models pursuing careers in the sciences. Each afternoon, the girls participated in WiSE activities on herd immunity, neuroscience, and astrophysics. They also took a "field trip" to Uplands Farm, toured the greenhouses, and learned about current CSHL plant research.

Our collaboration with the Watson School of Biological Sciences continued with the training of graduate students in the development of skills needed to communicate science to all audiences. As part of the required curriculum, first-year graduate students work with DNALC instructors to complete 12 half-day teaching sessions designed to prepare them to quickly assess an audience and subsequently customize a presentation accordingly. Graduate students interact equally with both middle and high school aged students during their required rotations, then choose three elective workshops to implement their new skills. This year, several students presented DNA extractions at the annual *Science Day* at Sagamore Hill.

BioMedia Visitation and Projects

In 2018, 5,617,903 visitors accessed our suite of multimedia resources. Google Analytics counted 4,044,799 visits to DNALC websites, our YouTube videos received 11,378 views, and the *3D Brain*, *Weed to Wonder*, and *Gene Screen* smartphone/tablet apps were downloaded 661,726 times. In-app purchases of 3D Brain HQ netted $7,060 for the year. As in the previous year, visitation to DNALC educational media decreased; in 2018 we started redesign of the DNALC.org site and began systematically updating the remainder of our online content. Over the next couple of years, the DNALC website will merge with the Cold Spring Harbor Laboratory (www.cshl.edu) site. We met with CSHL's team on that initiative several times during the year. The new dnalc.org under development will serve as an interim site until the final merger.

The *BioMedia* Group continues to support the initiatives of the DNALC through web and print design, photography, videography, exhibition development, and lab classroom layout planning for collaborators around the world. Highlights in 2018 included:

- Design work continued on the *DNA Subway* Purple Line and Green Line.
- At the annual *UBP*, *UBRP*, and *BLI* barcoding research symposia, student teams were photographed and video interviews were filmed of several groups of student participants.
- We initiated a video story on Michael Okoro.
- Design is under way on new displays for the rear gallery.

(*Left* to *right*) Colette Riccardi, Jeffry Petracca, Megan Capobianco, and Allison Mayle joined the DNALC staff in 2018.

Staff and Interns

DNALC staff was strengthened by the addition of Colette Riccardi to administration and Jeffry Petracca, Megan Capobianco, and Allison Mayle, Ph.D., to the instructional team.

Colette joined the administrative staff as our receptionist in March. A New Jersey native, she received her bachelor's degree in fine arts from Fairleigh Dickenson University, where she met her future husband, who happened to be a DNALC alumnus! Upon graduation, she worked in merchandising, marketing, and e-mail marketing. Colette relocated to Long Island, where she and her husband settled into a 100-year-old house in Huntington. Looking for a change of pace, Colette decided to switch from full- to part-time work. Not wanting just any job, she held out for something special, keeping CSHL at the top of her choices. When our receptionist position opened up, she knew it would be the right fit for her!

Jeffry's journey with the DNALC began in 2017, when he was called in as a taxonomy consultant for *BLI* to identify insects, spiders, and marine invertebrates. His talent was valued and he was offered an educator position; he joined the team part-time in March, spending the balance of his time as the Curator of Entomology at the Long Island Aquarium's Butterfly Exhibit and Insect Zoo in Riverhead. At 12 years old, Jeff began to fulfill his dreams of becoming an entomologist by volunteering at the Sweetbriar Nature Center's butterfly exhibit in Smithtown, where he spent 10 years teaching about butterflies, insects, and spiders, while using school breaks and weekends to volunteer for the butterfly exhibit at the American Museum of Natural History. He went on to graduate from Cornell University's School of Agriculture and Life Sciences in 2011 with a B.S. in Entomology, Plant Sciences, Neurobiology and Behavior, and Biological Statistics.

Inspired by the hit BBC TV series, *Planet Earth*, narrated by David Attenborough, Megan became fascinated by the environment, biology, and evolution. She graduated from SUNY Geneseo with a bachelor's in Childhood and Special Education with a minor in Environmental Science. She gained experience and a love for teaching as a middle school science teacher in Miami. While earning her master's in Environmental Policy and Management at Florida International University, she surveyed juvenile striped bass in the western bays of Long Island, including Cold Spring Harbor, with the NYS Department of Environmental Conservation. In addition to teaching, Megan will be taking on management of the *BLI* footlocker program and preparing equipment and reagents for off-site student use.

Allison joins our NYC team as we prepare to open our new location in Brooklyn. Passionate for scientific inquiry since childhood, she attended Michigan State University, where she could do research starting as a freshman through the Professorial Assistantship program. While earning her Ph.D. at Baylor College of Medicine in Houston, she learned techniques such as CRISPR and studied the processes that control stem cell self-renewal and how these processes go awry in blood cancers. During her time in Houston she volunteered for programs that introduced high school students to careers in science and medicine. At the Health Museum in the DeBakey Cell Lab, she guided

visitors through a variety of science experiments. These interactions fired up her passion for teaching. She moved on to a postdoc position at Memorial Sloan Kettering Cancer Center, where she continued to use CRISPR and study leukemia. She learned about CSHL from her postdoc advisor, saw there were openings for educators in NYC, and applied. She started in December of this year.

We said goodbye to three staff members in 2018: bioinformatics researcher Jorge Pérez de Acha Chavez, genetics educator Keil Thomas, and Nigerian exchange student Michael Okoro.

Jorge joined the DNALC's multimedia department as a Bioinformatics Researcher in 2017. During his tenure at the DNALC, Jorge made significant updates to *DNA Subway's* Green Line for RNA-Seq analysis and began development of a new Purple Line for metabarcode analysis. These updates became available in the fall. In September, he accepted a position as an associate computational biologist at the Broad Institute and Harvard.

Keil's first "real" science camp was *Fun with DNA* in fifth grade. It made such an impression that he returned in 2015 as a college intern. The lab experience he gained as an intern allowed him to teach scientific techniques to his college colleagues, making complex topics understandable for them. He was promoted to Middle School Educator after graduating with a bachelor's degree in Biomolecular Science from the NYU School of Engineering in 2016. During 3 years of teaching, he also managed the *BLI* Footlocker Program, equipment repair and supplies, purchasing, and all-around troubleshooting. In the fall, he moved on to the SUNY Stony Brook Biochemistry Department as the lab support technician.

Michael, native of Nigeria, was the DNALC's first exchange student, spending 3 weeks here in 2014 to complete a project on DNA barcoding of indigenous plants of Nigeria. We invited Michael back in 2016 as our first DNALC Fellow, with the objective of further developing his lab and teaching skills to prepare him for entry into a U.S. graduate program. Michael arrived in spring and by fall we had achieved our objective: Michael was accepted to the graduate program in molecular biotechnology at New York University (NYU), which he completed in 2018, achieving his Master's degree. He returned to Godfrey Okoye University in Nigeria this fall with plans to take the lead on starting up the first Africa-based DNA Learning Center.

Since the DNALC opened, we have relied on high school and college interns to support our day-to-day operations. An internship offers students the unique opportunity to gain real laboratory or design experience in an educational environment. The *BioMedia* Group and *Barcode Long Island* program also welcome interns for summer or longer-term roles. We gathered an amazing group of interns this year and said farewell as others left for college.

High School Interns

Gavin Calabretta, Cold Spring Harbor High School
Elijah Calle, Hempstead High School
Christopher Catalano, Garden City High School
Erika Mosso, Aquinas High School
Sibelle O'Donnell, Cold Spring Harbor High School
Jack O'Hara, St. Anthony's High School
Mina Samaras, Plainedge High School
Michael Stabile, Plainedge High School
Nicholas Stabile, Plainedge High School

High School Interns Departing for College

Duardo Akerle, New York University
Cassidy Alvarez, University of Delaware
Randy Diaz Arias, University of Rochester
Megan Erhardt, University of New Haven
Matthew Finkelberg, University of Massachusetts, Amherst
Brianna Hines, Tulane University
George Homenides, State University of New York at Albany
Brady Lyons, College of the Holy Cross
Jillian Maturo, Boston College
Jon Triscari, University of Rochester

College Interns

Gabrielle Blazich, Fordham University
Juliana Eastment, University of Richmond
Omotayo Ikuomenisan, Hunter College
William McBrien, Stony Brook University

Workshops and Visitors

January 6	NIH *Barcode Long Island* Open Lab, Brookhaven National Laboratory, Upton, New York
	NIH *Barcode Long Island* Special Open Lab, Stony Brook University, Stony Brook, New York
January 8–10	NSF INCLUDES Summit, "Broadening Participation with Bioinformatics, Big Data, and Data Science," Alexandria, Virginia
January 10	Lounsbery Fish eDNA Orientation Meeting, *Harlem DNA Lab*
January 11	NSF CyVerse Webinar Series, "Get Started with CyVerse," DNALC
January 12	Professional Development Workshop, "Bacterial Transformation," Michael J. Petrides School, Staten Island, New York
January 13	NIH *Barcode Long Island* Open Lab, DNA Learning Center *West*
	Ötzi the Iceman Tour, DNALC
	Saturday DNA! "Isolation Optimization," DNALC
	Urban Barcode Project Open Lab, *Harlem DNA Lab*
January 13–17	International Plant and Animal Genome XXVI Conference 2018, CyVerse Education Session: "Educational Workflows in Metagenomics: Microbiomes and Environmental (e)DNA," "Scaling Genomics and Data Science for the Biology Classroom," San Diego, California
January 17	Site visit by Xu Xun, Guangzhou Development District, Guangzhou, China
January 18	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
January 20	NIH *Barcode Long Island* Open Lab, DNALC
	NIH *Barcode Long Island—Microbiome Project* Open Lab, DNALC
January 23	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
Jan 29–Feb 9	*DNA Science, DNA Barcoding and Research* Workshops, Beijing 166 School, DNALC
February 1–3	Austrian Citizen Science Conference 2018, "Publishing with Citizen Scientists—Mission Impossible?" and "DNA Barcoding," University of Salzburg, Salzburg, Austria
February 5	Vienna Open Lab and Open Science at Vienna Biocenter, "DNA Barcoding," Austria
February 7	NSF CyVerse Webinar Series, "Get Started with CyVerse," DNALC
February 7–8	*DNA Barcoding* Workshop, Godfrey Okoye University, Enugu, Nigeria
February 8	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
February 10	NIH *Barcode Long Island—Microbiome Project* Open Lab, DNALC
	Ötzi the Iceman Tour, DNALC
	Saturday DNA! "Who Dunnit?" DNALC
	Urban Barcode Project Open Lab, *Harlem DNA Lab*
February 13	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
February 16	Microbiome eDNA Teacher Workshop, *Harlem DNA Lab*
February 19–23	*Urban Barcode Research Program, Conservation Genetics* Workshop, *Harlem DNA Lab*
February 20	*Ötzi the Iceman* Tour, DNALC
February 21	Site visit by Thomas Parsons, International Commission on Missing Persons, The Hague, The Netherlands
February 22	NIH *Barcode Long Island—Bioinformatics* Open Lab, DNALC
	Ötzi the Iceman Tour, DNALC
February 23	*Ötzi the Iceman* Tour, DNALC
February 26	CSH Asia DNA Learning Center Council Meeting, Suzhou Industrial Park, China
February 27	Site visit by Heather Savarese and Lauren Slaven, Cold Spring Harbor Educational Foundation, Cold Spring Harbor, New York
	CSHL *Gramene: A Resource for Comparative Plant Genomics* Webinar, DNALC
	Urban Barcode Project Open Lab, *Harlem DNA Lab*
February 28	Meeting with Guangzhou Development District, Guangzhou, China
March 2	*Urban Barcode Research Program* Update Event, The Irondale Center for Theater, Education, and Outreach, Brooklyn, New York
March 2–3	*DNA Barcode* Workshop, Science Center, Singapore
March 3	NIH *Barcode Long Island* Open Lab, Stony Brook University, Stony Brook, New York
March 5	Site visit by Mingdi Yang, Front Vision Magazine, New York, New York
March 6	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
March 7–8	NSF CyVerse *Genomics Data Carpentry* Workshop, North Carolina State University, Raleigh, North Carolina
March 9	NSF CyVerse *Tools and Services* Workshop, North Carolina State University, Raleigh, North Carolina
March 10	NIH *Barcode Long Island* Open Lab, DNALC
	Ötzi the Iceman Tour, DNALC
	Saturday DNA! "Agar Art," DNALC
March 14	NSF CyVerse Webinar Series, "Get Started with CyVerse," DNALC

March 14–18	NABT 2018 Professional Development Conference Presentations, "Sense in Molecules: Modeling Personalized Medicine," and "DNA Barcoding—Independent Research for All," Georgia World Conference Center, Juniper, Georgia
March 15	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
March 16	Professional Development Workshop, "Bacterial Transformation," Michael J. Petrides School, Staten Island, New York
March 20	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
March 22	NSF CyVerse Webinar Series, "Get Started with CyVerse," DNALC
March 22–25	Maize Genetics Conference, "Maize Tools & Resources," Saint-Malo, France
March 23	Long Island Natural History Conference, Brookhaven National Laboratory, Upton, New York Simons Foundation Science Outreach Working to Inspire the Next Generation (SOWING) Event, "Glowing Genes: Add Some DNA to Your Day," Caveat, New York, New York
March 24	NIH *Barcode Long Island—Microbiome Project* Bioinformatics Refresher, DNALC NIH *Barcode Long Island* Open Lab, Stony Brook University, Stony Brook, New York *Urban Barcode Research Program*, Mercy College, Dobbs Ferry, New York
March 24–26	American Society of Plant Biologists Southern Section Regional Meeting, "Data, Data Everywhere, nor Any a Drop to Drink," New Orleans, Louisiana
March 26	Site visit by Caren Gough Master Teacher Group, Science Education Group, Stony Brook University, Stony Brook, New York
March 26–27	Cold Spring Harbor Laboratory First Grade Science Fair Judging, CSHL
March 27	*Ötzi the Iceman* Tour, DNALC William Floyd Expo STEM Symposium, "DNA Extraction/DNA Barcoding," William Floyd Elementary School, Shirley, New York
March 29	NIH *Barcode Long Island—Bioinformatics* Open Lab, DNALC NIH *Barcode Long Island* Special Open Lab, Stony Brook University, Stony Brook, New York
April 2–6	*Urban Barcode Research Program, DNA Barcoding* Workshop, *Harlem DNA Lab*
April 2–3	eDNA Bioinformatics Training Workshop, City Tech, Brooklyn, New York
April 3	*Ötzi the Iceman* Tour, DNALC
April 5	*Ötzi the Iceman* Tour, DNALC
April 7	NIH *Barcode Long Island* Open Lab, DNA Learning Center *West*
April 11–12	NSF CyVerse *Genomics Data Carpentry* Workshop, Chan Zuckerberg BioHub, San Francisco, California
April 13	American Association of Physical Anthropologists 2018 Panel, "The Anthropologist's Academic Taboo II: Discussing Alternative Opportunities to Traditional R1 Anthropology Faculty Positions," Austin, Texas
April 14	NIH *Barcode Long Island* Open Lab, DNALC *Urban Barcode Project* Open Lab, *Harlem DNA Lab*
April 18	Site visit by Richard Baccari, Churchill & Banks, Providence, Rhode Island Webinar for Front Vision Magazine, New York, New York
April 18	*DNA Barcoding Research* Webinar for Parents and Students, Beijing, China NSF CyVerse Webinar Series, "Get Started with CyVerse," DNALC
April 19	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
April 21	NIH *Barcode Long Island* Open Lab, Brookhaven National Laboratory, Upton, New York *Saturday DNA!* "Enzyme Explosion," DNALC
April 24	*Urban Barcode Project* Open Lab, *Harlem DNA Lab*
April 25	Virtual Lab: *DNA Extraction from Wheat Germ*, DNALC
May 1–2	NOAA/Norwegian Institute of Marine Research eDNA Workshop, "Fish DNA and Student Research," Northeast Fisheries Science Center, Woods Hole, Massachusetts
May 3–4	Site visit to Centre for Interdisciplinary Science and Education, QuyNhon, Vietnam Beijing 166 Student Workshop, "Human Genome Science," Beijing 166 School, Beijing, China
May 5	NIH *Barcode Long Island* Open Lab, DNALC
May 6	Beijing 166 Teacher Workshop "Mentoring DNA Barcoding Research Projects," Beijing 166 School, Beijing, China
May 7–8	Beijing 166 Teacher Workshop, "Human Genome Science," Beijing 166 School, Beijing, China
May 8	NIH *Barcode Long Island—Bioinformatics* Open Lab, DNALC Beijing 166 Student Workshop, "DNA Barcoding Research Follow-Up," Beijing 166 School, Beijing, China
May 9	Site visit by Jyothi Nayar, Illumina, San Diego, California Site visit to Jade Spring School, Guangzhou, China
May 10–11	Site visit to CSH Asia DNA Learning Center, Suzhou Industrial Park, China
May 12	Biocodigos de Barras Urbanos CDMX Symposium, "DNA Barcodes: Genetic Challenges for Citizen Science," Mexico City, Mexico

	NIH *Barcode Long Island* Open Lab, Stony Brook University, Stony Brook, New York
	Saturday DNA! "Molecular Detectives," DNALC
	Ötzi the Iceman Tour, DNALC
May 16–17	NSF CyVerse *Genomics Data Carpentry* Workshop, University of New Orleans, New Orleans, Louisiana
May 18	NSF CyVerse *Tools and Services* Workshop, University of New Orleans, New Orleans, Louisiana
May 19	NIH *Barcode Long Island* Open Lab, Brookhaven National Laboratory, Upton, New York
May 24	*Urban Barcode Project*/Pinkerton *Urban Barcode Research Program* Symposium, New York Academy of Medicine, New York, New York
May 25	Amazon Research Conference, "Scaling Science by Scaling People, Purpose-Built Cyberinfrastructure for the Life Sciences," Amazon Research Development Center, Cambridge, United Kingdom
May 28–29	NSF CyVerse *Genomics Data Carpentry* Workshop, University College Dublin, Ireland
May 29–June 1	NIH SEPA SciEd Conference, "Barcode Long Island: Exploring Biodiversity in a Unique Urban Landscape," Washington, D.C.
June 2	*Saturday DNA!* "Got Lactase?," DNALC
	Ötzi the Iceman Tour, DNALC
June 4	ECSA Conference 2018, "Participatory Research/DNA Barcoding," Maison Communale de Plainpalais, Geneva, Switzerland
June 5	*Barcode Long Island* Student Symposium, CSHL
June 7	Oyster Bay Historical Society, "Barcode Long Island," Oyster Bay, New York
June 7–9	ECSITE Annual Conference 2018, "DIY Bio at Science Centers and Museums," Natural History Museum of Geneva, Geneva, Switzerland
June 11–15	*Human Genomics* Workshop, Lycée Français, New York, New York
	World of Enzymes Workshop, St. David's School, New York, New York
	Genome Science Workshop, Lycée Français, New York, New York
June 12	*Ötzi the Iceman* Tour, DNALC
June 14–29	*DNA Barcoding* and *Bioinformatics* Workshops, Grace Church High School, New York, New York
June 20	NSF CyVerse Webinar Series, "Get Started with CyVerse," DNALC
June 21	*Ötzi the Iceman* Tour, DNALC
	Eastport-South Manor Science Symposium, Eastport-South Manor High School, Manorville, New York
June 22	Enhancing the STEM Pipeline through Bioinformatics and Genomics Conference, Panels: "Introductory Bioinformatics/Genomics for High School Students, Community College and Undergraduate," "Bioinformatics/Genomics Curriculum and Research Opportunities," "The Formation of Partnerships for Bioinformatics and/or Genomics STEM Outreach," and Poster: "Barcoding and Metabarcoding for Independent Student Research Incorporating Bioinformatics," University at Buffalo, New York
June 22–24	NSF CyVerse *Genomics Data Carpentry* Workshop, Cold Spring Harbor Laboratory, New York
June 25–29	NIH *Barcode Long Island* Workshop, Brookhaven National Laboratory, Upton, New York
	DNA Science Workshop, DNALC (two sessions)
	Fun with DNA Workshop, DNALC
	Green Genes Workshop, DNALC
	Fun with DNA Workshop, DNA Learning Center *West*
	Genome Science Workshop, Grace Church High School, New York, New York
	Forensic Detectives Workshop, *DNALC Asia*, Suzhou, China
June 25–30	Bioinformatics Open Source Conference/Galaxy Community Conference, "Improving the Undergraduate Bioinformatics Curriculum," Reed College, Portland, Oregon
June 26	Site visit by Ana Jara Ettinger and Salvador Jara, Mexico City, Mexico
June 28	Bioinformatics Open Source/Galaxy Community Conference, "Improving the Undergraduate Bioinformatics Curriculum," Reed College, Portland, Oregon
July 2	NSF CyVerse *Genomics Data Carpentry* Course, CSHL Frontiers and Techniques in Plant Science, Cold Spring Harbor Laboratory, New York
July 2–3	*Genome Science* Workshop, Grace Church High School, New York, New York
July 2–6	*Forensic Detectives* Workshop, DNALC
	Fun with DNA Workshop, DNALC
	Genome Science Workshop, DNALC
	World of Enzymes Workshop, DNALC
	World of Enzymes Workshop, DNA Learning Center *West*
	Forensic Detectives Workshop, *DNALC Asia*, Suzhou, China
	Pinkerton *Urban Barcode Research Program, Conservation Genetics* Workshop, City Tech, Brooklyn, New York
July 9–12	NIH *Barcode Long Island* eDNA Workshop, DNALC
July 9–13	*DNA Science* Workshop, DNALC
	Green Genes Workshop, DNALC

	World of Enzymes Workshop, DNALC
	DNA Science Workshop, DNA Learning Center *West*
	DNA Science Workshop, *DNALC Asia*, Suzhou, China
	Pinkerton *Urban Barcode Research Program, DNA Barcoding and Bioinformatics* Workshop, DNALC at City Tech, Brooklyn, New York
July 14	American Society of Plant Biology 2018 Conference, Teacher Workshop "Finding and Fixing Genome Annotation Errors with MaizeCODE," Montreal, Canada
July 16–20	*Fun with DNA/World of Enzymes* Workshop, Beijing 166, DNALC (two sessions)
	DNA Barcoding Workshop, DNALC
	Fun with DNA Workshop, DNALC
	Green Genes Workshop, DNA Learning Center *West*
	DNA Science Workshop, *DNALC Asia*, Suzhou, China
	Pinkerton *Urban Barcode Research Program Conservation Genetics* Workshop, City Tech, Brooklyn, New York
July 19	NSF CyVerse *Genomics Data Carpentry* Workshop, "RNA Tutorial," Plant Biology 2018, University of Québec, Montreal, Canada
July 20	Agar Art activity and lecture, *DNALC Asia*, Suzhou, China
July 21	*Ötzi the Iceman* Tour with *DNA Extraction* Workshop, DNALC
July 23–27	*Forensic Detectives* Workshop, Beijing 166, DNALC
	Green Genes Workshop, Beijing 166, DNALC
	BioCoding Workshop, DNALC
	DNA Science Workshop, DNALC
	World of Enzymes Workshop, DNALC
	Fun with DNA Workshop, DNA Learning Center *West*
	Pinkerton *Urban Barcode Research Program, DNA Barcoding* Workshop, DNALC at City Tech, Brooklyn, New York
July 30–31	NSF CyVerse *Genomics in Education* Workshop, "DNA Subway and Microbiome Workshop," Austin Community College, Austin, Texas
July 30–August 3	*Forensic Detectives* Workshop, Beijing 166, DNALC
	Green Genes Workshop, Beijing 166, DNALC
	DNA Barcoding Workshop, DNALC
	Fun with DNA Workshop, DNALC
	DNA Science Workshop, DNA Learning Center *West*
	Pinkerton *Urban Barcode Research Program Conservation Genetics* Workshop, DNALC at City Tech, Brooklyn, New York
August 6–9	*DNA Barcoding and Bioinformatics* Workshop, Toms River High School East, Toms River, New Jersey
August 6–10	*BioCoding* Workshop, DNALC
	DNA Science Workshop, DNALC
	Fun with DNA Workshop, Peter Wang Chinese Group, DNALC
	Green Genes Workshop, DNALC
	World of Enzymes Workshop, DNALC
	Forensic Detectives Workshop, DNA Learning Center *West*
	Pinkerton *Urban Barcode Research Program* and *Science Sandbox DNA Barcoding and Bioinformatics* Workshop, DNALC at City Tech, Brooklyn, New York
August 7	Site visit by Kristen Wolslegel, Bay Area Bioscience Education Community, San Francisco, California
	Ötzi the Iceman Tour, DNALC
August 8–24	*Urban Barcode Project* Teacher Training Workshop, *DNA Barcoding*, DNALC at City Tech, Brooklyn, New York
August 13–15	NSF CyVerse *Data Carpentry* Workshop and *Tools and Services* Workshop, Colorado State University, Fort Collins, Colorado
August 13–17	NIH *Barcode Long Island* Teacher Workshop, Stony Brook University Institute for STEM Education, Stony Brook, New York
	Being Human Workshop, DNALC
	DNA Science Workshop, DNALC
	Fun with DNA Workshop, DNALC
	Green Genes Workshop, DNALC
	Green Genes Workshop, DNA Learning Center *West*
August 17	Site visit by John Tuke, Hotchkiss School, Salisbury, Connecticut
August 19–23	American Fisheries Society Meeting, "Independent Student Research Using eDNA Metabarcoding to Track Marine Fish," Atlantic City, New Jersey
August 20–24	*DNA Science* Workshop, DNALC
	Forensic Detectives Workshop, NALC
	Genome Science Workshop, DNALC

	World of Enzymes Workshop, DNALC
	Fun with DNA Workshop, DNA Learning Center *West*
	Urban Barcode Project DNA Barcoding Teachers Workshop, City Tech, Brooklyn, New York
August 22	Site visit by Alan Goldberg and family, Lindsay Goldberg LLC, New York, New York
August 23	Site visit by Janice Rolf and grandson, ALS Ride for Life, Stony Brook, New York
August 24	Site visit by Dr. Susan Bachle, Addgene, Watertown, Massachusetts
August 27–31	*DNA Science* Workshop, DNALC
	Fun with DNA Workshop, DNALC
	Green Genes Workshop, DNALC
	World of Enzymes Workshop, DNALC
	WiSE *Fun with DNA* Workshop, CSHL
	World of Enzymes Workshop, DNA Learning Center *West*
August 28	*Ötzi the Iceman* Tour, DNALC
September 19	DNA Barcoding Presentation, Leave No Trace Group, Brookhaven, New York, DNALC
September 22	Site visit by Guangzhou Delegation, Guangzhou, China
	Hudson River Park Trust SUBMERGE Science Festival, "Marine Diversity and Environmental DNA," Pier 84, New York, New York
September 24	*Human DNA Fingerprint: Genotyping a "Jumping Gene"* Teacher Workshop, Math for America, New York, New York
September 27	"Urban Barcode Research," Kickoff Event, The Irondale Center for Theater, Education, and Outreach, Brooklyn, New York
October 4	Science Outreach Conference 2018, "Models, Methods, and Measures," "BioArt as a Medium of Scientific Storytelling," Rockefeller University, New York, New York
October 8–19	*Genome Science* Student and Teacher Workshops, Beijing 166 School, Beijing, China
October 13	*Saturday DNA!* "The Magic of Microscopes," DNALC
	Ötzi the Iceman Tour, DNALC
	DNA Barcoding Teacher Workshop, Beijing 166 School, Beijing, China
October 15	*Human DNA Fingerprint: Genotyping a "Jumping Gene"* Teacher Workshop, Math for America, New York, New York
October 16	4th Regional NY/NJ eDNA Discussion Meeting, "Metabarcoding," Hudson River Foundation, New York, New York
October 24	Site visit by Dennis Liu, Wilson Biodiversity Foundation, Durham, North Carolina
October 24–26	2018 ATE Principal Investigators Conference, "Using a Supply Chain Model to Recruit and Educate Students" and "Models for Developing Faculty Leaders" poster, Omni Shoreham Hotel, Washington, D.C.
October 29	*Human DNA Fingerprint: Genotyping a "Jumping Gene,"* Teacher Workshop, Math for America, New York, New York
October 31	Site visit by Becky Gilmore, Wellcome Trust Genome Campus, Cambridge, England
Oct 31–Nov 5	Meetings with Armando Barriguete and Hugo Scherer, Mexico DNA Learning Center Development, Mexico City, Mexico
November 4–6	Midwest Big Data Hub Meeting, Case Western Reserve University, Cleveland, Ohio
November 6	NIH *Barcode Long Island* "Mentor Update and Bioinformatics Refresher Meeting," DNALC
November 7	Site visit by Regeneron, Tarrytown, New York, DNA Learning Center *West*
November 7–9	Biological Science Data Meeting Poster Presentation, "Improving the Bioinformatics Curriculum," CSHL
November 10–11	Genomics Education Alliance Meeting, St. Louis, Missouri
November 12	Site visit by Kiryn Hoffman, Margaret Honey, Elsbeth Pancrezi, and Sylvia Perez, The New York Hall of Science, New York, New York
Nov 29–30	EMBL-Australian Bioinformatics Resources All Hands & CyVerse Webinar, University of Melbourne, Australia
Nov 29–30	National Conference on Environmental Marine DNA, "Building and Implementing a Biochemical and Bioinformatic Workflow to Enable Authentic Student eDNA Research," Rockefeller University, New York, New York
December 4	NSF CyVerse BioInfo Summer Workshop, "Introduction to RNA-Seq with the Kallisto and Sleuth Workflows," "Improving the Bioinformatics Curriculum," University of Western Australia, Perth, Australia
	Erase Racism lecture, "Eugenics," Riverhead Senior Center, Riverhead, New York
December 8	Austrian Embassy Invitation to Austrian Research and Innovation Talks (ARIT), Poster Session, "BioArt as a Medium for Scientific Storytelling," Washington, D.C.
December 10	Erase Racism lecture, "Eugenics," Radisson Hotel, Hauppauge, New York
December 10–12	NSF CyVerse United Kingdom Workshop, "RNA Sequencing," Earlham Institute, Norwich, England
December 12	*Urban Barcode Research Program* Holiday Event, The Irondale Center for Theater, Education and Outreach, Brooklyn, New York
December 14	Site visit by William Harsh, New York, New York
December 19	DNALC Collaboration Visit to New York Hall of Science, New York, New York

Sites of Major Faculty Workshops

Program Key: *Middle School* High School **College**

State	*Institution*	*Year(s)*
ALABAMA	University of Alabama, Tuscaloosa	1987–1990
	Hudson Alpha Institute, Huntsville	**2014**
ALASKA	**University of Alaska, Anchorage**	**2012**
	University of Alaska, Fairbanks	1996
ARIZONA	Arizona State University, Tempe	2009
	Tuba City High School	1988
	University of Arizona, Tucson	**2011**
	United States Department of Agriculture, Maricopa	**2012**
ARKANSAS	Henderson State University, Arkadelphia	1992
	University of Arkansas, Fayetteville	**2017**
	University of Arkansas, Little Rock	**2012**
CALIFORNIA	**California State University, Dominguez Hills**	**2009**
	California State University, Fullerton	**2000**
	California State University, Long Beach	**2015**
	California Institute of Technology, Pasadena	**2007**
	Chan-Zuckerberg BioHub, San Francisco	**2018**
	Canada College, Redwood City	1997
	City College of San Francisco	2006
	City College of San Francisco	**2011, 2013**
	Contra Costa County Office of Education, Pleasant Hill	2002, 2009
	Foothill College, Los Altos Hills	**1997**
	Harbor-UCLA Research & Education Institute, Torrance	2003
	Los Angeles Biomedical Research Institute (LA Biomed), Torrance	**2006**
	Laney College, Oakland	1999
	Lutheran University, Thousand Oaks	1999
	Oxnard Community College, Oxnard	2009
	Pasadena City College	**2010**
	Pierce College, Los Angeles	**1998**
	Salk Institute for Biological Studies, La Jolla	2001, 2008
	San Francisco State University	1991
	San Diego State University	**2012**
	San Jose State University	**2005**
	Santa Clara University	2010
	Southwestern College, Chula Vista	**2014–2015**
	Stanford University, Palo Alto	**2012**
	University of California, Berkeley	**2010, 2012**
	University of California, Davis	1986
	University of California, Davis	**2012, 2014–2015**
	University of California, Long Beach	**2015**
	University of California, Northridge	**1993**
	University of California, Riverside	2011
	University of California, Riverside	**2012**
	University of California, San Francisco	**2015**
COLORADO	Aspen Science Center	2006
	Colorado College, Colorado Springs	1994, 2007
	Colorado State University, Fort Collins	**2013, 2018**
	Community College of Denver	**2014**
	United States Air Force Academy, Colorado Springs	**1995**
	University of Colorado, Denver	1998, 2009–2010
CONNECTICUT	Choate Rosemary Hall, Wallingford	1987
DELAWARE	**Jackson Laboratory, Farmington**	**2016**
	University of Delaware, Newark	**2016**
DISTRICT OF COLUMBIA	**Howard University, Washington**	**1992, 1996, 2009–2010**
FLORIDA	Armwood Senior High School, Tampa	1991
	Florida Agricultural & Mechanical University, Tallahassee	2007–2008

	Florida Agricultural & Mechanical University, Tallahassee	**2011**
	Florida SouthWestern State University, Fort Myers	**2015**
	North Miami Beach Senior High School	1991
	Seminole State College, Sanford	**2013–2014**
	University of Florida, Gainesville	1989
	University of Miami School of Medicine	2000
	University of Western Florida, Pensacola	1991
GEORGIA	Fernbank Science Center, Atlanta	1989, 2007
	Gwinnett Technical College, Lawrenceville	**2011–2012**
	Morehouse College	**1991, 1996**
	Morehouse College	1997
	Spelman College, Atlanta	**2010**
	University of Georgia, Athens	**2015**
HAWAII	Kamehameha Secondary School, Honolulu	1990
	University of Hawaii at Manoa	**2012**
IDAHO	University of Idaho, Moscow	1994
ILLINOIS	Argonne National Laboratory	1986–1987
	iBIO Institute/Harold Washington College, Chicago	2010
	Illinois Institute of Technology, Chicago	2009
	Kings College, Chicago	**2014**
	University of Chicago	**1992, 1997, 2010**
	University of Southern Illinois, Carbondale	**2016**
INDIANA	Butler University, Indianapolis	1987
	Purdue University, West Lafayette	**2012**
IOWA	Drake University, Des Moines	1987
KANSAS	University of Kansas, Lawrence	1995
KENTUCKY	**Bluegrass Community & Technical College, Lexington**	**2012–2014**
	Murray State University	1988
	University of Kentucky, Lexington	1992
	Western Kentucky University, Bowling Green	1992
LOUISIANA	Bossier Parish Community College	2009
	Jefferson Parish Public Schools, Harvey	1990
	John McDonogh High School, New Orleans	1993
	Southern University at New Orleans	**2012**
	University of New Orleans	**2018**
MAINE	Bates College, Lewiston	1995
	Southern Maine Community College	**2012–2013**
	Foundation for Blood Research, Scarborough	2002
MARYLAND	Annapolis Senior High School	1989
	Bowie State University	**2011, 2015**
	Frederick Cancer Research Center	1995
	McDonogh School, Baltimore	1988
	Montgomery County Public Schools	1990–1992
	National Center for Biotechnology Information, Bethesda	2002
	St. John's College, Annapolis	*1991*
	University of Maryland, School of Medicine, Baltimore	**1999**
MASSACHUSETTS	**Arnold Arboretum of Harvard University, Roslindale**	**2011**
	Beverly High School	1986
	Biogen Idec, Cambridge	2002, 2010
	Boston University	**1994, 1996**
	CityLab, Boston University School of Medicine	1997
	Dover-Sherborn High School, Dover	1989
	Randolph High School	1988
	The Winsor School, Boston	1987
	Whitehead Institute for Biomedical Research, Cambridge	2002
MICHIGAN	Athens High School, Troy	1989
	Schoolcraft College, Livonia	**2012**
MINNESOTA	**American Society of Plant Biologists, Minneapolis**	**2015**
	Minneapolis Community and Technical College, Madison	2009
	Minneapolis Community and Technical College, Madison	**2013**
	University of Minnesota, St. Paul	2005

	University of Minnesota, St. Paul	**2010**
MISSISSIPPI	Mississippi School for Math & Science, Columbus	1990–1991
	Rust College, Holly Springs	**2006–2008, 2010**
MISSOURI	**St. Louis Science Center**	**2008–2010**
	Stowers Institute for Medical Research, Kansas City	2002, 2008
	University of Missouri, Columbia	**2012**
	Washington University, St. Louis	1989
	Washington University, St. Louis	**1997, 2011**
MONTANA	**Montana State University, Bozeman**	**2012**
NEBRASKA	**University of Nebraska–Lincoln, Lincoln**	**2014**
NEVADA	University of Nevada, Reno	1992, **2014**
NEW HAMPSHIRE	Great Bay Community College, Portsmouth	2009
	New Hampshire Community Technical College, Portsmouth	**1999**
	St. Paul's School, Concord	1986–1987
NEW JERSEY	Coriell Institute for Medical Research, Camden	2003
	Raritan Valley Community College, Somerville	2009
NEW MEXICO	Biolink Southwest Regional Meeting, Albuquerque	2008
	Los Alamos National Lab	**2017**
	New Mexico State University, Las Cruces	**2017**
	Santa Fe Community College, Santa Fe	**2015**
NEW YORK	Albany High School	1987
	American Museum of Natural History, New York	2007, 2015
	Bronx High School of Science	1987
	Brookhaven National Laboratory, Upton	2015–2018
	Canisius College, Buffalo	2007
	Canisius College, Buffalo	**2011**
	City College of New York	2012
	Cold Spring Harbor High School	1985, 1987
	Cold Spring Harbor Laboratory	**2014–2015, 2018**
	Columbia University, New York	**1993**
	Cornell University, Ithaca	**2005**
	DeWitt Middle School, Ithaca	*1991, 1993*
	Dolan DNA Learning Center	1988–1995, 2001–2004, 2006–2009, 2015–2018
	Dolan DNA Learning Center	**1990, 1992, 1995, 2000–2011**
	Dolan DNA Learning Center	*1990–1992*
	DNA Learning Center *West*	2005
	Environmental Science Center, Bergen Beach, Brooklyn	2015–2016
	Fostertown School, Newburgh	*1991*
	Harlem DNA Lab, East Harlem	2008–2009, 2011–2013, 2016–2018
	***Harlem DNA Lab*, East Harlem**	**2015–2016**
	Huntington High School	1986
	Irvington High School	1986
	John Jay College of Criminal Justice	2009
	Junior High School 263, Brooklyn	*1991*
	Lindenhurst Junior High School	*1991*
	Math for America, New York	**2017–2018**
	Michel J. Petrides School, Staten Island	2018
	Mount Sinai School of Medicine, New York	1997
	Nassau Community College, Garden City	**2013**
	New York Botanical Garden, Bronx	**2013**
	New York City Department of Education	2007, 2012
	New York City Technical College (City Tech)	2018
	New York Institute of Technology, New York	2006
	New York Institute of Technology, New York	**2006**
	Orchard Park Junior High School	*1991*
	Plainview-Old Bethpage Middle School	*1991*
	School of Visual Arts, New York	**2017**
	State University of New York, Purchase	1989

	State University of New York, Stony Brook	1987–1990, 2015–2018
	State University of New York, Stony Brook	**2014, 2016**
	Stuyvesant High School, New York	1998–1999
	The Rockefeller University, New York	2003, 2015–2016
	The Rockefeller University, New York	**2010**
	Titusville Middle School, Poughkeepsie	*1991, 1993*
	Trudeau Institute, Saranac Lake	2001
	Union College, Schenectady	2004
	United States Military Academy, West Point	**1996**
	Wheatley School, Old Westbury	1985
NORTH CAROLINA	CIIT Center for Health Research, Triangle Park	2003
	North Carolina Agricultural & Technical State University, Greensboro	**2006–2007, 2009–2011**
	North Carolina School of Science, Durham	1987
	North Carolina State University, Raleigh	**2012, 2018**
NORTH DAKOTA	**North Dakota State University, Fargo**	**2012**
OHIO	Case Western Reserve University, Cleveland	1990
	Cleveland Clinic	1987
	Langston University, Langston	2008
	North Westerville High School	1990
	The Ohio State University, Wooster	**2016**
OKLAHOMA	**Oklahoma City Community College**	**2000**
	Oklahoma City Community College	2006–2007, 2010
	Oklahoma Medical Research Foundation, Oklahoma City	2001
	Oklahoma School of Science and Math, Oklahoma City	1994
	Tulsa Community College, Tulsa	2009
	Tulsa Community College, Tulsa	**2012–2014**
OREGON	Kaiser Permanente Center for Health Research, Portland	2003
	Linfield College, McMinnville	**2014**
PENNSYLVANIA	Duquesne University, Pittsburgh	1988
	Germantown Academy	1988
	Kimmel Cancer Center, Philadelphia	2008
RHODE ISLAND	Botanical Society of America, Providence	2010
SOUTH CAROLINA	**Clemson University**	**2004, 2015**
	Medical University of South Carolina, Charleston	1988
	University of South Carolina, Columbia	**1988**
SOUTH DAKOTA	**South Dakota State University, Brookings**	**2015**
TENNESSEE	NABT Professional Development Conference, Memphis	2008
TEXAS	Austin Community College—Rio Grande Campus	2000
	Austin Community College—Eastview Campus, —Roundrock Campus	**2007–2009, 2013**
	Austin Community College—Roundrock Campus	**2012**
	Austin Community College—Austin	**2018**
	Houston Community College Northwest	**2009–2010**
	J.J. Pearce High School, Richardson	1990
	Langham Creek High School, Houston	1991
	University of Lone Star College, Kingwood	**2011**
	Midland College	2008
	Southwest Foundation for Biomedical Research, San Antonio	2002
	Taft High School, San Antonio	1991
	Texas A&M University, College Station	**2013**
	Texas A&M University, Prairie View	**2013**
	Texas A & M, AG Research and Extension Center, Weslaco	2007
	Trinity University, San Antonio	**1994**
	University of Texas, Austin	**1999, 2004, 2010, 2012**
	University of Texas, Brownsville	**2010**
UTAH	**Brigham Young University, Provo**	**2012**
	University of Utah, Salt Lake City	1993
	University of Utah, Salt Lake City	**1998, 2000**
	Utah Valley State College, Orem	**2007**

VERMONT	University of Vermont, Burlington	1989
	Champlain Valley Union High School, Hinesburg	1989
VIRGINIA	Eastern Mennonite University, Harrisonburg	1996
	James Madison University, Harrisonburg	**2017**
	Jefferson School of Science, Alexandria	1987
	Mathematics and Science Center, Richmond	1990
	Mills Godwin Specialty Center, Richmond	1998
	Virginia Polytechnic Institute and State University, Blacksburg	**2005, 2008–2009**
WASHINGTON	Fred Hutchinson Cancer Research Center, Seattle	1999, 2001, 2008
	Shoreline Community College	2011, 2012
	University of Washington, Seattle	**1993, 1998, 2010**
WEST VIRGINIA	**Bethany College**	**1989**
WISCONSIN	Blood Center of Southeastern Wisconsin, Milwaukee	2003
	Madison Area Technical College/Madison Area College	**1999, 2009, 2011–2014**
	Marquette University, Milwaukee	1986–1987
	University of Wisconsin, Madison	1988–1989
	University of Wisconsin, Madison	**2004, 2012**
WYOMING	University of Wyoming, Laramie	1991
PUERTO RICO	**Universidad del Turabo, Gurabo**	**2011, 2012, 2014**
	University of Puerto Rico, Mayaguez	1992
	University of Puerto Rico, Mayaguez	**1992**
	University of Puerto Rico, Rio Piedras	**1993**
	University of Puerto Rico, Rio Piedras	1994

AUSTRALIA	Walter and Eliza Hall Institute and University of Melbourne	1996
	EMBL/Australian Bioinformatics Resource, University of Melbourne	**2016**
	University of Western Australia, Perth	**2018**
AUSTRIA	**Vienna Open Lab, Vienna**	**2007, 2012**
CANADA	Red River Community College, Winnipeg, Manitoba	1989
	University of Quebec, Montreal	**2018**
CHINA	Beijing No. 166 High School, Beijing	2013–2018
	Ho Yu College, Hong Kong	**2009**
DENMARK	**Faroe Genome Project, Torshavn, Faroe Islands**	**2013**
GERMANY	Urania Science Center, Berlin	2008
IRELAND	**European Conference on Computational Biology/Intelligent System for Molecular Biology Conference, Dublin**	**2015**
	University College Dublin	**2018**
ITALY	International Institute of Genetics and Biophysics, Naples	1996
	Porto Conte Research and Training Laboratories, Alghero	1993
MEXICO	ADN Mexico, Morelia	2016
	ASPB Plant Biology, Mérida	2008
	Langebio/Cinvestav, Irapuato	**2016**
NIGERIA	Godfrey Okoye University, Enugu	2013, 2018
PANAMA	**University of Panama, Panama City**	**1994**
PHILIPPINES	Eastern Visayas Campus, Philippine Science High School, Palo, Leyte	2017
RUSSIA	Shemyakin Institute of Bioorganic Chemistry, Moscow	1991
SINGAPORE	National Institute of Education	2001–2005
SOUTH AFRICA	Singapore Science Center	2013
	North-West University, Potchefstroom	**2016**
	South African Bioinformatics Society, Durban	**2016**
SWEDEN	Kristineberg Marine Research Station, Fiskebackgkil	1995
	Uppsala University	2004
THE NETHERLANDS	**International Chromosome Conference, Amsterdam**	**2007**
	Wageningen University and Research Center, Wageningen	**2014**
UNITED KINGDOM	**Earlham Institute, Norwich**	**2018**
	The Genome Analysis Center, Norwich	**2015**
	University of York	**2017**
	Wellcome Trust Conference Center, Hinxton	**2012–2013**
	University of Warwick, Coventry	**2013**

COLD SPRING HARBOR
LABORATORY PRESS

PRESS PUBLICATIONS

Serials

Genes & Development, Vol. 32 (www.genesdev.org)

Genome Research, Vol. 28 (www.genome.org)

Learning & Memory, Vol. 25 (www.learnmem.org)

RNA, Vol. 24 (www.rnajournal.org)

Cold Spring Harbor Symposia in Quantitative Biology, Vol. 82: *Chromosome Segregation and Structure*, edited by David Stewart and Bruce Stillman

Cold Spring Harbor Protocols (www.cshprotocols.org)

Cold Spring Harbor Perspectives in Biology (www.cshperspectives.org)

Cold Spring Harbor Perspectives in Medicine (www.perspectivesinmedicine.org)

Cold Spring Harbor Molecular Case Studies (www.molecularcasestudies.org)

Textbooks

RNA Worlds: New Tools for Deep Exploration, edited by Thomas R. Cech, Joan A. Steitz, and John F. Atkins

Monographs (Topic Collections from *Perspectives in Biology* and *Perspectives in Medicine*)

Immune Memory and Vaccines: Great Debates, edited by Shane Crotty and Rafi Ahmed

Circadian Rhythms, edited by Paolo Sassone-Corsi, Michael W. Young, and Akhilesh B. Reddy

Multiple Sclerosis, edited by Howard L. Weiner and Vijay K. Kuchroo

Bone: A Regulator of Physiology, edited by Gerard Karsenty and David T. Scadden

Cytokines: From Basic Mechanisms of Cellular Control to New Therapeutics, edited by Warren J. Leonard and Robert D. Schreiber

Enteric Hepatitis Viruses, edited by Stanley M. Lemon and Christopher M. Walker

RAS and Cancer in the 21st Century, edited by Linda Van Aelst, Julian Downward, and Frank McCormick

Translation Mechanisms and Control, edited by Michael B. Mathews, Nahum Sonenberg, and John W.B. Hershey

Prostate Cancer, edited by Michael M. Shen and Mark A. Rubin

Other

Faces of the Genome, portraits by Lewis Miller; edited by Ludmila Pollock, W. Richard McCombie, and Jan A. Witkowski

How Scientific Progress Occurs: Incrementalism and the Life Sciences, by Elof Axel Carlson

Cell Death: Apoptosis and Other Means to an End, Second Edition, by Douglas R. Green

Lab Dynamics: Management and Leadership Skills for Scientists, Third Edition, by Carl M. Cohen and Suzanne L. Cohen

CSHL Annual Report 2016, Yearbook Edition

E-books

Faces of the Genome, portraits by Lewis Miller; edited by Ludmila Pollock, W. Richard McCombie, and Jan A. Witkowski

How Scientific Progress Occurs: Incrementalism and the Life Sciences, by Elof Axel Carlson

Multiple Sclerosis, edited by Howard L. Weiner and Vijay K. Kuchroo

Bone: A Regulator of Physiology, edited by Gerard Karsenty and David T. Scadden

Cytokines: From Basic Mechanisms of Cellular Control to New Therapeutics, edited by Warren J. Leonard and Robert D. Schreiber

Cell Death: Apoptosis and Other Means to an End, Second Edition, by Douglas R. Green

Enteric Hepatitis Viruses, edited by Stanley M. Lemon and Christopher M. Walker

Cold Spring Harbor Symposia in Quantitative Biology, Vol. 82: *Chromosome Segregation and Structure*, edited by David Stewart and Bruce Stillman

RAS and Cancer in the 21st Century, edited by Linda Van Aelst, Julian Downward, and Frank McCormick

Lab Dynamics: Management and Leadership Skills for Scientists, Third Edition, by Carl M. Cohen and Suzanne L. Cohen

Translation Mechanisms and Control, edited by Michael B. Mathews, Nahum Sonenberg, and John W.B. Hershey

RNA Worlds: New Tools for Deep Exploration, edited by Thomas R. Cech, Joan A. Steitz, and John F. Atkins

Prostate Cancer, edited by Michael M. Shen and Mark A. Rubin

Next Generation Sequencing in Medicine, edited by W. Richard McCombie, Elaine R. Mardis, James A. Knowles, and John D. McPherson

Websites

Cold Spring Harbor Monographs Archive Online (www.cshmonographs.org)

Cold Spring Harbor Symposium on Quantitative Biology Archive (symposium.cshlp.org)

Services

BioSupplyNet, scientific supply directory (www.biosupplynet.com)

COLD SPRING HARBOR LABORATORY PRESS EXECUTIVE DIRECTOR'S REPORT

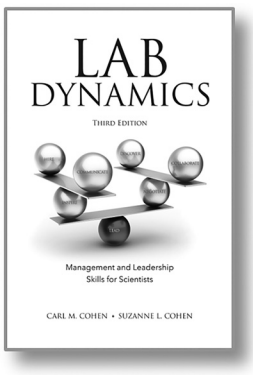
Progress in science has many dependencies, but among the most important is the timely, accurate, and ethical communication of ideas, experimental procedures, and results. Cold Spring Harbor Laboratory Press provides scientists worldwide with authoritative, affordable, and pertinent information to further their research and aid in their career development. This ambition continues the Laboratory's commitment to scientific communication that began with the publication of the proceedings of the first Annual Symposium in 1933.

The Press publishes nine journals and more than 200 books in print and electronic form. The long-established journals *Genome Research* and *Genes & Development* remain preeminent in their field, with in-house editorial teams adept in the assessment of new results and their interpretation. *RNA* and *Learning & Memory* serve specialized research communities in valuable ways. The review journals, *CSH Perspectives in Biology*, *CSH Perspectives in Medicine*, and *CSH Protocols*, continue to advance in stature and financial success. Each transforms content that in former years appeared only in print books, rendering it in a digital, readily discoverable, and reusable serial form. *CSH Molecular Case Studies* enables the open sharing of clinical insights that genomic and molecular analysis bring to the understanding and potential treatment of disease: in its second full year, the journal had a steady increase in submissions.

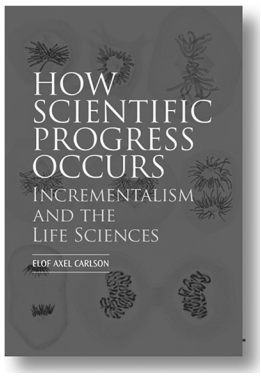
Overall, online usage of Press journals remained notable in 2018, with more than 17 million full-text article downloads worldwide, including more than 5.4 million via PubMed Central at the U.S. National Institutes of Health's National Library of Medicine.

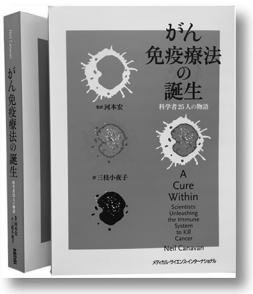
Launched in April 2018, *Life Science Alliance* (*LSA*) is a new open-access journal owned and published jointly by Cold Spring Harbor Laboratory, the European Molecular Biology Organization (EMBO), and The Rockefeller University. *LSA* provides an opportunity for the authors of papers that were not selected by the nine high-profile journals published by the partners. The "cascade" publishing model has become common in research communication in recent years and has benefitted publishers that gain revenue by retaining a greater proportion of submissions, and also authors whose papers find an appropriate vehicle for publication more efficiently. *LSA* is a unique cross-publisher cascade and, aside from its revenue potential for the Press, it helps provide a tactical counter to the growing cascade silos of large commercial publishers, which are an increasing factor in the competition for authors and journal content at all levels. The journal published 81 original research articles through January 2019. Usage of the journal's content during its first nine months was encouraging, with more than 114,000 full-text downloads and 1,300 sign-ups for e-mail content alerts.

Life Science Alliance



In the book-publishing program, 15 new print titles and 14 new eBooks were added in 2018. The handbook *Lab Dynamics* was published in a third edition, which has been updated to assist principal investigators in the challenging responsibilities of interviews with prospective team candidates and performance assessment. In December, *The RNA Worlds: New Tools for Deep Exploration* was published, a worthy successor to four previous well-received books on similar themes—editions that had been well-received.





The year's best-selling new books were *Symposium 82: Chromosome Segregation and Structure* and *How Scientific Progress Occurs: Incrementalism and the Life Sciences* by Elof Carlson.

A Japanese translation of *A Cure Within* by Neil Canavan was completed, with a Chinese translation in process. We have also contracted for Mongolian and Vietnamese translations of *The Annotated and Illustrated Double Helix*.

The strong performance of the book program in 2018 was driven by highly effective direct-to-customer marketing and e-commerce efforts. Direct sales through the Press website accounted for 21% of all book sales, exceeding any other wholesale or retail channel—including Amazon. This result is remarkable in today's retail environment and reflects the high level of affinity between the Press and its core audience of scientists worldwide. Providing direct customers with the added value of a bundled electronic edition greatly assisted this success. In 2018, >40% of purchases included an e-book as a companion to a print edition or as a stand-alone publication when available.

Staff

During the year, we welcomed Ahmet Denli—a member of the inaugural class of the Watson School of Biological Sciences—back to the Lab as Associate Editor of *Genome Research*, and Brandon Kelly joined us as a newly appointed Junior Accountant. We also said farewell to Miriam Fein, Reviews Editor, *Genes & Development*, and Jesse Ulloa, Staff Accountant.

The mission of the Press is to create publications and services that help scientists succeed while making a positive financial contribution of unrestricted funds to the Laboratory and supporting the institution's worldwide reputation for innovative research and first-class activities in scientific education and communication. As a publishing organization, the Press works with many of the world's most accomplished scientists. These relationships are successful because of the professionalism and dedication of the entire Press staff. I thank them all and recognize in particular those individuals who provide outstanding leadership in our diverse activities: Assistant Director Richard Sever, journal editors Terri Grodzicker and Hillary Sussman, and departmental directors Jan Argentine, Wayne Manos, Stephen Nussbaum, Marcie Siconolfi, and Linda Sussman. As always, I am extremely grateful for the extraordinarily deft, efficient, and generous way in which the Office of the Director is run by Mala Mazzullo.

John Inglis
Executive Director and Publisher

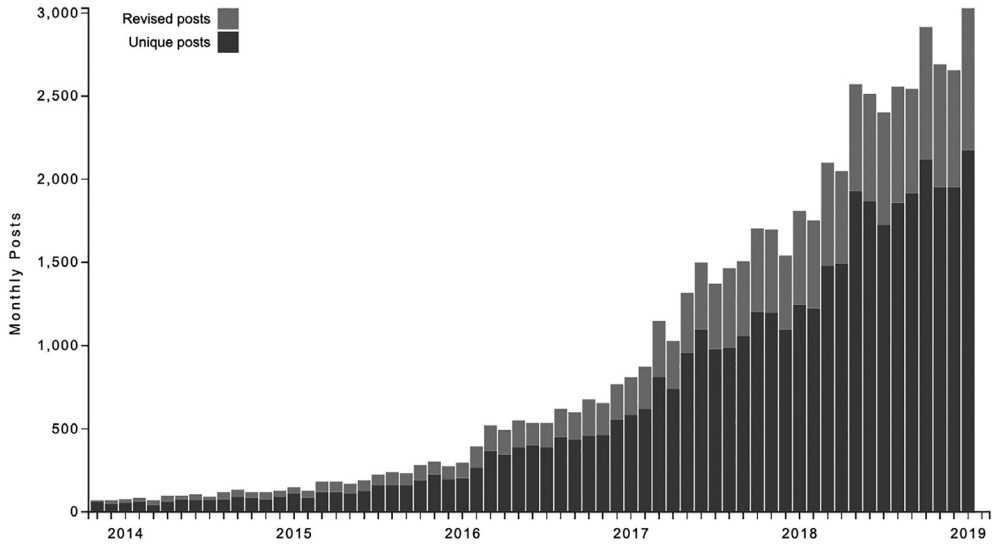
PREPRINT SERVER

bioRxiv

A preprint is a research manuscript distributed by its authors before certification by peer review and publication by a journal. The Laboratory's preprint server, bioRxiv, turned five in 2018 and, with major support from the Chan Zuckerberg Initiative (CZI), continued its rapid growth—doubling in size with 20,000 new submissions from more than 100 countries. Each month, papers on the server are read more than 4.5 million times and discussed in social networks and dedicated preprint assessment sites.

The server is clearly accelerating science: 70% of bioRxiv manuscripts are shared for community evaluation as long as two years (median 6 months) before journal publication. Thirty journals now enable authors to simultaneously submit a manuscript and post it on bioRxiv, and 130 journals accept automatic submission of preprints for editorial consideration. In 2018, an additional CZI grant made possible the conversion of all bioRxiv papers into XML format so they can be read in a web browser without downloading.

bioRxiv continues to transform the way biologists communicate their science.



bioRxiv content by month—all categories

FINANCE

FINANCIAL STATEMENTS

CONSOLIDATED BALANCE SHEET
December 31, 2018
(with comparative financial information as of December 31, 2017)

	2018	2017
Assets:		
Cash and cash equivalents	$ 121,078,957	64,179,016
Grants receivable	7,730,792	8,948,017
Contributions receivable, net	70,177,263	74,760,266
Investments	535,006,949	531,826,534
Investment in employee residences	6,811,348	6,348,606
Restricted use assets	1,975,292	4,070,570
Other assets	6,186,059	11,190,405
Land, buildings, and equipment, net	249,420,636	235,245,898
Total assets	$ 998,387,296	936,569,312
Liabilities and net assets:		
Liabilities:		
Accounts payable and accrued expenses	$ 12,948,173	15,888,844
Deferred revenue	81,386,316	5,601,953
Interest rate swap	25,281,037	31,345,495
Bonds payable	95,807,696	95,741,427
Total liabilities	215,423,222	148,577,719
Net assets:		
Without donor restrictions	426,827,607	403,766,549
With donor restrictions	356,136,467	384,225,044
Total net assets	782,964,074	787,991,593
Total liabilities and net assets	$ 998,387,296	936,569,312

CONSOLIDATED STATEMENT OF ACTIVITIES
Year ended December 31, 2018
(with summarized financial information for the year ended December 31, 2017)

	Without Donor Restrictions	With Donor Restrictions	2018 Total	2017 Total
Revenue and other support:				
Public support—contributions and nonfederal grant awards	$ 20,329,363	42,493,880	62,823,243	84,901,067
Federal grant awards	44,976,944	–	44,976,944	34,668,573
Indirect cost allowances	36,554,683	–	36,544,683	31,719,355
Investment return utilized	22,641,750	–	22,641,750	21,297,672
Royalty and license revenue	17,005,951	–	17,005,951	13,399,961
Program fees	7,990,826	–	7,990,826	8,741,475
Publications sales	9,564,563	–	9,564,563	9,411,102
Dining services	4,898,093	–	4,898,093	4,993,409
Rooms and apartments	3,739,034	–	3,739,034	3,922,034
Miscellaneous	751,769	–	751,769	663,437
Net assets released from restrictions	45,768,002	(45,768,002)	–	–
Total revenue and other support	214,220,978	(3,274,122)	210,946,856	213,718,085
Expenses:				
Research	106,880,122	–	106,880,122	96,658,762
Educational programs	18,950,499	–	18,950,499	18,392,576
Publications	8,768,554	–	8,768,554	9,223,370
Banbury Center conferences	2,301,320	–	2,301,320	2,407,363
DNA Learning Center programs	3,860,017	–	3,860,017	4,044,790
Watson School of Biological Sciences programs	3,199,759	–	3,199,759	3,111,700
General and administrative	30,841,520	–	30,841,520	22,149,504
Total expenses	174,801,791	–	174,801,791	155,988,065
Excess (deficiency) of revenue and other support over expenses	39,419,187	(3,274,122)	36,145,065	57,730,020
Other changes in net assets:				
Investment (loss) return (including) excluding amount utilized	(24,246,892)	(22,990,150)	(47,237,042)	44,890,565
Change in fair value of interest rate swap	6,064,458	–	6,064,458	1,368,278
Reclassification to implement ASU 2016-14	1,824,305	(1,824,305)	–	–
Increase (decrease) in net assets	23,061,058	(28,088,577)	(5,027,519)	103,988,863
Net assets at beginning of year	403,766,549	384,225,044	787,991,593	684,002,730
Net assets at end of year	$ 426,827,607	356,136,467	782,964,074	787,991,593

CONSOLIDATED STATEMENT OF CASH FLOWS
Year ended December 31, 2018
(with comparative financial information for the year ended December 31, 2017)

	2018	2017
Cash flows from operating activities:		
(Decrease) increase in net assets	$ (5,027,519)	103,988,863
Adjustments to reconcile (decrease) increase in net assets to net cash provided by operating activities:		
Change in fair value of interest rate swap	(6,064,458)	(1,368,278)
Depreciation and amortization	13,140,747	13,441,446
Amortization of deferred bond costs	66,269	66,270
Net depreciation (appreciation) in fair value of investments	29,986,414	(62,254,634)
Contributions restricted for long-term investment	(18,810,536)	(12,059,874)
Changes in assets and liabilities:		
Grants receivable	1,217,225	(881,047)
Contributions receivable, net	20,118,950	(23,650,821)
Restricted use assets	2,095,278	(548,515)
Other assets	5,004,346	(4,866,945)
Accounts payable and accrued expenses, net of financing activities	(6,994,843)	3,746,487
Deferred revenue	75,784,363	(3,523,528)
Net cash provided by operating activities	110,516,236	12,089,424
Cash flows from investment activities:		
Capital expenditures	(27,315,485)	(19,593,684)
Proceeds from sales and maturities of investments	108,718,791	100,074,927
Purchases of investments	(141,885,620)	(96,669,733)
Net change in investment in employee residences	(462,742)	(347,757)
Net cash used in investment activities	(60,945,056)	(16,536,247)
Cash flows from financing activities:		
Contributions restricted for long-term investment	900,022	4,065,789
Contributions restricted for investment in capital	17,910,514	7,994,085
Increase in contributions receivable	(15,535,947)	(3,755,827)
Decrease in accounts payable relating to capital expenditures	4,054,172	2,112,316
Net cash provided by financing activities	7,328,761	10,416,363
Net increase in cash and cash equivalents	56,899,941	5,969,540
Cash and cash equivalents at beginning of year	64,179,016	58,209,476
Cash and cash equivalents at end of year	$ 121,078,957	64,179,016
Supplemental disclosure:		
Interest paid	$ 3,762,449	3,886,138

FINANCIAL SUPPORT OF THE LABORATORY

Cold Spring Harbor Laboratory, Banbury Center, and the Dolan DNA Learning Center receive a substantial portion of funding through grants from the federal government and through grants, capital gifts, and annual contributions from New York state, private foundations, corporations, and individuals. The following section summarizes funding that occurred during 2018.

GRANTS January 1–December 31, 2018

COLD SPRING HARBOR LABORATORY

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
FEDERAL GRANTS					
NATIONAL INSTITUTES OF HEALTH					
Equipment Support	Dr. C. Hammell	03/01/16	12/31/20	$124,680	
Program Project and Center Support	Drs. Stillman/Egeblad/Krainer/McCombie/ Pappin/Spector/Vakoc	02/20/18	01/31/23	4,537,433	*
	Dr. Tuveson—Cancer Center Core	08/01/16	07/31/21	4,371,458	
	Drs. McCombie/Tuveson	09/25/15	08/31/19	154,589	
Cooperative Research Agreement Support[2]	Drs. Huang/Gillis/Mitra/Osten/Zador	09/21/17	05/31/22	4,526,788	
	Dr. Kepecs	09/30/15	08/31/19	105,622	
	Drs. Osten/Albeanu/Mitra	09/20/17	05/31/22	1,605,281	
	Dr. Tuveson	06/12/14	02/29/20	169,687	
	Dr. Tuveson	03/06/18	02/28/23	438,242	*
Research Support	Dr. Albeanu	02/09/16	01/31/21	408,000	
	Drs. Albeanu/Koulakov	09/11/18	08/31/23	993,218	*
	Dr. Churchland	05/01/18	04/30/22	480,000	*
	Drs. Dobin/Gingeras	08/18/17	05/31/22	480,000	
	Drs. Engel/Churchland	09/20/18	06/30/21	424,672	*
	Dr. Furukawa	03/01/15	02/29/20	487,140	
	Dr. Gillis	09/15/17	08/31/21	472,478	
	Drs. Gillis/Huang/Lee	07/13/17	05/31/22	472,477	
	Dr. C. Hammell	03/01/16	12/31/20	403,200	
	Dr. Huang	07/01/11	07/31/23	608,410	
	Drs. Huang/Gillis	08/01/16	07/31/21	797,983	
	Drs. Huang/Wigler	03/01/14	01/31/20	721,116	
	Dr. Joshua-Tor	06/10/16	03/31/20	345,600	
	Dr. Kepecs	09/15/17	07/31/22	420,000	
	Dr. Kepecs	04/01/14	03/31/19	480,000	
	Dr. Kepecs	08/15/15	05/31/20	420,000	
	Dr. Kepecs	09/15/15	07/31/20	432,772	
	Dr. Krainer	07/01/17	06/30/22	806,400	
	Dr. Lee	08/19/16	05/31/21	480,000	
	Dr. Li	03/01/14	02/28/19	480,000	
	Dr. Li	07/11/17	06/30/19	240,000	
	Dr. Li/Huang	09/28/15	06/30/20	660,707	

[1]Awarded, including direct and indirect costs
[2]Funding amounts include only CSHL's portion of the award
*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
	Dr. Martienssen	06/01/17	04/30/21	422,400	
	Dr. Mills	09/16/14	08/31/19	606,309	
	Dr. Mitra	09/30/16	06/30/19	423,985	
	Dr. Park	09/15/16	08/31/21	174,117	
	Dr. Pedmale	08/04/17	07/31/22	480,000	
	Dr. Shea	12/08/15	11/30/20	570,860	
	Dr. Sheltzer	09/18/15	08/31/20	480,000	
	Dr. Siepel	03/01/18	02/28/23	479,215	*
	Dr. D. Spector	04/01/15	03/31/20	766,080	
	Dr. Stenlund	05/01/15	04/30/20	504,989	
	Dr. Stillman	03/01/17	02/28/21	729,600	
	Drs. Tollkuhn/Gillis	03/01/18	12/31/22	480,000	*
	Dr. Tonks	05/14/15	04/30/20	776,676	
	Dr. Tonks	01/15/15	12/31/19	448,737	
	Dr. Trotman	07/30/14	06/30/19	398,400	
	Drs. Tuveson/M. Hammell/Pappin	12/07/16	11/30/21	658,780	
	Dr. Vakoc	04/01/13	11/30/23	552,186	
	Dr. Zador	07/01/14	04/30/19	420,000	
	Dr. Zador	09/01/17	08/31/20	907,874	
Research Subcontracts					
NIH/Case Western Reserve University Consortium Agreement	Dr. Tonks	09/08/17	08/31/19	48,325	
NIH/Envisagenics, Inc. Consortium Agreement	Dr. Krainer	04/05/18	03/31/20	49,801	*
NIH/Harvard Medical School Consortium Agreement	Dr. Osten	07/01/17	04/30/22	84,000	
NIH/Harvard Medical School Consortium Agreement	Dr. Osten	09/01/17	07/31/21	172,800	
NIH/Johns Hopkins University Consortium Agreement	Dr. Joshua-Tor	03/01/15	02/29/20	33,264	
NIH/New York Genome Center Consortium Agreement	Drs. Wigler/Iossifov/Levy/Siepel	01/14/16	11/30/19	422,972	
NIH/New York University Consortium Agreement	Dr. Koulakov	06/01/14	05/31/19	169,826	
NIH/Oregon Health & Science University	Dr. Li	12/15/17	11/30/22	856,610	
NIH/The Regents of the University of California, San Diego	Dr. Mitra	09/15/18	05/31/23	264,877	*
NIH/The Research Foundation for the State of New York - Stony Brook Consortium Agreement	Dr. Wigler	05/01/14	04/30/19	130,599	
NIH/The Research Foundation for the State of New York - Stony Brook Consortium Agreement	Dr. M. Hammell	09/15/17	06/30/22	89,885	
NIH/The Scripps Research Institute Consortium Agreement	Dr. Tuveson	08/01/16	07/31/19	79,680	
NIH/University of Minnesota Consortium Agreement	Dr. dos Santos	01/01/17	12/31/20	86,512	
NIH/University of Nebraska Consortium Agreement	Drs. Tuveson/Pappin	05/01/17	04/30/22	365,750	
NIH/University of Pittsburgh Consortium Agreement	Dr. Mitra	09/11/18	08/31/23	93,068	*
NIH/University of Texas at Austin Consortium Agreement	Dr. Osten	09/18/17	06/30/22	780,705	

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
Fellowship/Career Development Support	Dr. Bravo-Rivera	06/01/17	05/31/20	59,038	
	Dr. Crow	06/16/17	06/15/20	63,034	
	Dr. Fisher	03/01/18	02/28/21	58,654	*
	Dr. Jaremko	12/01/18	11/30/21	58,654	*
	Dr. Wang	05/02/16	05/01/19	65,158	
Institutional Training Program Support	Dr. Mills/Cancer Postdoctoral	09/16/16	08/31/21	146,168	
	Dr. Gann/Watson School of Biological Sciences	07/01/17	06/30/22	282,780	
Course Support	Advanced Sequencing Technologies and Applications	04/10/12	06/30/21	58,895	
	Advanced Techniques in Molecular Neuroscience	07/06/15	03/31/20	105,668	
	Cell and Development Biology of *Xenopus*	05/05/14	03/31/19	76,284	
	Cellular Biology of Addiction	08/01/16	07/31/21	40,260	
	Computational and Comparative Genomics	08/15/17	06/30/20	67,704	
	Empowering Nextgen: Advanced Biomedical Leadership	06/01/15	02/29/20	521,040	
	Eukaryotic Gene Expression	04/13/17	03/31/22	99,813	
	Molecular Embryology of the Mouse	04/13/17	03/31/22	130,582	
	Programming for Biology	09/01/17	06/30/20	83,641	
	Protein Purification and Characterization	04/13/17	03/31/22	86,387	
	Proteomics	08/03/18	04/30/23	123,254	*
	Quantitative Imaging: From Cells to Molecules	04/01/16	03/31/21	102,233	
	X-Ray Methods in Structural Biology	09/01/17	08/31/22	90,745	
Meeting Support	The Biology of Genomes	04/01/18	03/31/23	60,000	*
	Gene Expression and Signaling in the Immune System	03/14/18	02/28/19	8,000	*
	Germ Cells	07/01/14	06/30/19	6,000	
	Global Regulation of Gene Expression	02/22/16	01/31/19	29,951	
	Mechanisms and Models of Cancer	05/01/18	04/30/19	10,000	*
	Molecular Genetics of Aging	03/15/14	02/28/19	30,972	
	Molecular Mechanisms of Neuronal Connectivity	07/01/18	06/30/19	15,000	*
	Neuronal Circuits	09/25/17	09/29/18	3,000	
	Network Biology	09/18/18	08/31/19	31,268	*
	Neurodegenerative Diseases: Biology of Therapeutics	08/01/18	07/31/19	32,227	*
	Retroviruses	04/17/18	03/31/21	35,000	*
NATIONAL SCIENCE FOUNDATION					
Multiple Project Award Support	Drs. Gingeras/Jackson/Martienssen/McCombie/Ware	06/15/16	05/31/19	2,096,691	
	Drs. Jackson/Lippman	09/01/16	08/31/20	1,176,521	
	Drs. McCombie/Martienssen	09/01/18	08/31/22	1,307,159	*
Research Support	Dr. Albeanu	08/01/17	07/31/21	218,874	
	Dr. Jackson	08/01/18	07/31/21	224,055	*

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
	Drs. Jackson/Gillis	07/01/18	06/30/20	151,540	*
	Dr. Kepecs	09/01/17	08/31/19	402,752	
	Dr. Lippman	06/15/16	05/31/19	214,246	
	Dr. Lippman	07/01/17	06/30/21	1,230,008	
	Dr. Pedmale	09/01/18	08/31/21	327,434	*
	Dr. Siepel	03/01/16	02/28/19	19,377	
Research Subcontracts					
NSF/Cornell University Consortium Agreement	Dr. Siepel	09/01/18	08/31/22	9,965	*
NSF/Iowa State University Consortium Agreement	Dr. Jackson	03/01/13	02/28/19	60,000	
NSF/University of Georgia Consortium Agreement	Dr. Ware	01/15/18	12/31/19	216,376	*
Fellowship Support	B. Bibel	09/01/14	07/31/19	46,000	
	B. Berube	06/01/17	05/31/20	46,000	
	D. Johnson	06/01/17	05/31/20	46,000	
	K. O'Neill	06/01/18	05/31/21	46,000	*
	J. Werner	09/01/18	08/31/21	46,000	*
Institutional Training Program Support	Drs. Gillis/C. Hammell/Research Experiences for Undergraduates Program	05/01/16	04/30/19	156,467	
Course Support	Advanced Bacterial Genetics	06/15/17	05/31/22	90,000	
	Drosophila Neurobiology: Genes, Circuits, and Behavior	07/01/17	06/30/20	28,403	
	Frontiers and Techniques in Plant Science	08/01/18	07/31/21	106,434	*
	Synthetic Biology	09/01/18	08/31/21	62,305	*
	Yeast Genetics and Genomics	07/01/17	06/30/22	90,000	
UNITED STATES DEPARTMENT OF AGRICULTURE					
Research Support	Dr. Jackson	01/01/18	12/31/21	116,088	*
	Dr. Jackson	12/15/15	12/14/19	127,896	
	Dr. Lippman	11/01/15	10/31/19	110,989	
	Dr. Lippman	01/01/15	12/31/19	127,266	
	Dr. McCombie	09/15/17	09/14/19	1,161,472	
UNITED STATES DEPARTMENT OF THE ARMY					
Research Support	Dr. Egeblad	06/15/14	06/14/19	943,456	
Research Subcontracts					
DOA/Emory University Consortium Agreement	Dr. Trotman	09/30/16	09/29/19	54,348	
DOA/University of Southern California Consortium Agreement	Dr. Churchland	08/23/16	08/22/19	188,100	
Fellowship Support	Dr. Casanova Salas	08/15/17	08/14/19	120,000	
	Dr. Scaduto	08/15/18	08/14/21	184,564	*
UNITED STATES DEPARTMENT OF ENERGY					
Research Support	Dr. Martienssen	09/15/17	09/14/21	1,074,577	

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
Research Subcontracts					
DOE/Lawrence Berkeley National Laboratory Consortium Agreement	Dr. Ware	05/18/17	09/30/20	1,092,518	
DOE/New York University Consortium Agreement	Dr. McCombie	08/15/15	08/14/20	239,968	
MISCELLANEOUS SOURCES OF FUNDING					
Equipment Support					
The Claire Friedlander Family Foundation	Drs. Stillman/McCombie	07/24/18	07/23/19	50,000	*
CSHL Translational Cancer Support	Dr. Lyons	04/01/18	03/31/19	193,000	*
Dr. Robert Lourie	Dr. Stillman	12/01/18	11/30/19	275,000	*
Program Project Support					
The Simons Foundation/Autism	Dr. Wigler	01/01/17	12/31/20	1,728,834	
The Simons Foundation/Cancer	Dr. Wigler	01/01/17	12/31/19	2,535,542	
Research Support					
Rita Allen Foundation	Dr. dos Santos	09/01/16	08/31/21	100,000	
	Dr. M. Hammell	09/01/14	08/31/19	110,000	
American Association for Cancer Research	Dr. dos Santos	07/01/18	06/30/20	75,000	*
Anonymous	Dr. Shea	12/31/17	12/30/18	3,000	
Anonymous	Dr. Tuveson	11/01/18	10/31/21	1,350,000	*
Maret and Thaddeus Asaro	Dr. Tonks	12/01/18	11/30/19	50,000	*
Austins Purpose Corporation	Dr. Furukawa	12/29/16	12/28/19	10,000	
Babylon Breast Cancer Coalition Inc.	Dr. Egeblad	07/24/18	07/23/19	4,000	*
Beckman Research Institute of the City of Hope	Dr. Atwal	06/01/15	06/30/19	200,000	
Brain & Behavior Research Foundation	Dr. C. Hammell	09/15/17	09/14/19	50,000	
Breast Cancer Alliance	Dr. Sheltzer	02/01/17	06/30/19	62,500	
The Breast Cancer Research Foundation	Drs. Wigler/Levy	10/01/18	09/30/19	250,000	*
Calico Life Sciences LLC	Drs. Wigler/Levy/Mitra	09/01/16	03/31/19	174,453	
Case Western Reserve University/ Rett Syndrome Research Trust	Dr. Tonks	01/15/18	01/14/19	5,000	*
Cedar Hill Foundation	Dr. Fearon	11/16/18	11/15/19	60,000	*
CSHL Translational Cancer Support	Dr. Atwal	02/25/17	02/24/20	300,870	
	Dr. Chang	06/01/18	05/31/21	259,563	*
	Dr. dos Santos	05/26/17	05/25/20	288,000	
	Dr. Egeblad	06/15/16	06/14/20	453,021	
	Dr. Evans	11/01/18	10/31/19	282,881	*
	Dr. Gillis	03/01/18	02/28/22	133,981	*
	Dr. Gingeras	01/15/16	01/14/19	516,025	
	Dr. C. Hammell	03/01/18	02/28/21	356,678	*
	Dr. Janowitz	08/22/18	08/21/23	236,218	*
	Dr. Kepecs	07/01/17	06/30/20	145,369	
	Dr. Kinney	02/25/17	02/24/21	473,613	
	Dr. Krasnitz	03/01/18	02/28/22	194,179	*
	Dr. Levy	03/01/18	02/28/22	16,522	*
	Dr. Li	03/01/18	02/28/22	192,000	*

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
	Dr. Li	03/01/18	02/28/22	192,000	*
	Dr. Lyons	07/01/17	06/30/20	332,870	
	Dr. Lyons	04/01/18	03/31/21	482,611	*
	Dr. Mills	01/01/17	12/31/20	576,000	
	Dr. Mills	10/25/17	10/24/21	723,432	
	Dr. Mitra	07/01/17	06/30/20	169,710	
	Dr. Osten	07/21/17	07/20/20	250,802	
	Dr. Pappin	05/26/17	05/25/20	704,519	
	Dr. Preall	06/01/17	05/31/20	916,102	
	Dr. Sheltzer	03/01/18	02/29/20	638,064	*
	Dr. Sordella	04/27/17	04/26/19	507,846	
	Dr. D. Spector	12/01/18	11/30/21	534,588	*
	Dr. D. Spector	06/01/16	05/31/21	193,154	
	Dr. D. Spector	04/01/18	03/31/21	410,807	*
	Dr. Trotman	06/01/16	05/31/19	486,125	
	Dr. Tuveson	04/01/18	03/31/21	861,335	*
	Dr. Tuveson	07/14/18	07/13/19	629,760	*
	Dr. Tuveson	07/14/18	07/13/19	76,800	*
	Dr. Vakoc	01/15/16	07/14/19	1,205,807	
	Dr. Van Aelst	07/01/18	06/30/21	403,206	*
	Dr. Wigler	03/01/18	02/28/22	540,508	*
	Dr. Yeh	03/01/18	02/28/20	70,080	*
	Dr. Yeh	06/15/16	06/14/19	698,506	
	Dr. Yeh	06/01/16	05/31/21	418,445	
	Dr. Yeh	07/01/17	06/30/20	28,858	
	Dr. Zhang	01/13/17	01/12/21	347,608	
	Dr. Zhang	05/04/18	05/03/20	602,755	*
	Dr. Zheng	06/01/15	05/31/19	76,513	
Cystic Fibrosis Foundation	Dr. Krainer	02/01/18	01/31/20	106,119	*
Donaldson Charitable Trust	Drs. Tuveson/Beyaz/Froeling	12/21/16	12/20/19	462,564	
East-West International BV	Dr. Lippman	06/01/16	05/31/19	100,000	
Elstar Therapeutics Inc.	Dr. Beyaz	12/18/18	12/31/20	222,192	*
The Edward P. Evans Foundation	Dr. Zhang	07/01/18	06/30/19	125,000	*
Charitable Lead Annuity Trust under the Will of Louis Feil	Drs. Li/Osten/Zador	12/01/18	11/30/19	200,000	*
Douglas and Christine Fox	Dr. Furukawa	06/01/18	05/31/22	50,000	*
Bill & Melinda Gates Foundation	Drs. Sheltzer/McCombie	10/10/18	07/31/19	274,902	*
The Joni E. Gladowsky Breast Cancer Foundation	Dr. dos Santos	11/14/18	11/13/19	55,000	*
Glen Cove C.A.R.E.S.	Dr. dos Santos	01/27/18	01/26/19	5,000	*
The GoGo Foundation	Dr. Li	09/29/16	09/28/19	40,000	
Gyeongsang National University/ The Republic of Korea	Dr. Jackson	01/01/18	12/31/18	55,209	*
Irving Hansen Foundation	Dr. Tonks	08/01/18	07/31/19	40,000	*
Jo-Ellen and Ira Hazan	Dr. Tuveson	12/18/18	12/17/19	5,000	*
Heartfelt Wings Foundation Inc.	Dr. Furukawa	09/05/18	09/04/22	250,000	*
Robert and Cindy Higginson	Dr. Mills	12/17/18	12/16/19	1,000	*
The Hope Foundation	Dr. Yeh	03/01/18	02/28/19	53,400	*
Howard Hughes Medical Institute	Dr. Martienssen	08/10/18	08/09/19	38,451	*
Human Frontier Science Program	Dr. Li	09/01/16	08/31/19	100,000	
Inari Agriculture, Inc.	Dr. Lippman	02/14/18	02/13/19	10,000	*
Indian Institute of Technology Madras	Dr. Mitra	01/01/15	12/31/19	28,011	

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
Industry-Academic Cooperation Foundation of Wonkwang University in the Republic of Korea	Dr. Lippman	01/01/18	12/31/20	55,110	*
Ionis Pharmaceuticals, Inc.	Dr. D. Spector	07/01/15	12/31/20	945,794	
M. Leidner	Dr. D. Spector			100	*
The Leukemia & Lymphoma Society	Dr. Vakoc	07/01/15	06/30/20	110,000	
Long Island Bioscience Hub	Dr. Zhang	05/16/18	05/15/19	120,000	*
The Lustgarten Foundation	Dr. Fearon	07/01/14	12/31/19	1,000,000	
	Dr. Tuveson	09/01/17	06/30/19	333,334	
	Dr. Tuveson	06/15/12	06/30/19	70,000	
John and Patti Maroney	Dr. Wigler	12/18/18	12/17/19	5,000	*
The G. Harold and Leila Y. Mathers Charitable Foundation	Dr. Mitra	01/01/17	12/31/19	330,000	
Breast Cancer Awareness Day in Memory of Elizabeth McFarland	Dr. Wigler	09/01/00	12/31/19	127,917	
Memorial Sloan Kettering Cancer Center/Henry and Marilyn Taub Foundation	Dr. Krainer	12/15/16	12/14/19	66,666	
Louis Morin Charitable Trust	Dr. Gingeras	12/12/18	12/11/19	115,000	*
New York State Department of Health	Dr. Egeblad	01/01/17	06/30/19	170,400	
Northwell Health, Inc.	Drs. Tuveson/Froeling	03/14/18	10/16/18	83,209	*
Ono Pharmaceutical Co., Ltd.	Drs. Tuveson/Chang	04/01/18	09/30/19	1,027,005	*
The Pershing Square Foundation	Dr. dos Santos	07/01/18	06/30/21	200,000	*
	Dr. Egeblad	07/01/17	06/30/20	400,000	
	Dr. Vakoc	09/09/16	06/30/20	180,000	
The Pew Charitable Trusts	Dr. Hannon	01/01/11	09/30/19	134,015	
Christina Renna Foundation Inc.	Dr. Vakoc	09/16/14	09/15/19	30,000	
W.J. Riley Memorial Foundation	Dr. Vakoc	09/16/14	09/15/19	10,000	
Charles and Marie Robertson Foundation	Dr. dos Santos	02/16/18	02/15/19	40,000	*
Eleanor Schwartz Charitable Foundation	Dr. Churchland	07/26/18	07/25/19	100,000	*
Edith and Alan Seligson	Drs. Tuveson/Koike	05/23/18	05/22/19	100,000	*
Seven Bridges Genomics Inc.	Dr. Lyon	09/01/16	08/30/20	9,000	
The Simons Foundation	Dr. Iossifov	05/01/17	04/30/19	177,842	
	Dr. Zador	07/01/17	06/30/22	140,400	
	Dr. Churchland	07/01/17	06/30/22	280,800	
	Dr. Tollkuhn	10/01/18	09/30/19	80,000	*
	Dr. Zador	01/01/17	06/30/18	75,000	
The Simons Foundation/CSHL Innovative Cancer Research	Dr. Stillman	07/01/17	06/30/19		
	Drs. dos Santos/Egeblad/Van Aelst	07/01/17	06/30/19	660,000	
	Dr. Lyons	07/01/17	06/30/20	12,601	
	Drs. Lyons/Trotman	07/01/17	06/30/20	225,318	
	Dr. Tonks	07/01/17	06/30/20	666,959	
	Drs. Tuveson/Vakoc	07/01/17	06/30/20	250,000	
Stand up for Suzanne	Dr. dos Santos	12/19/18	12/18/19	25,000	*
Starr Cancer Consortium	Dr. Joshua-Tor	01/01/17	12/31/19	142,200	
Swim Across America	Dr. Sordella	12/26/18	12/25/19	60,000	*
Anne D. Thomas	Dr. dos Santos	08/08/18	08/07/19	10,000	*
The Thompson Family Foundation, Inc.	Drs. Tuveson/Chang/Egeblad/Fearon/Lyons/Vakoc/Yeh	03/06/17	03/05/21	538,440	
Three Strohm Sisters Family Foundation	Dr. Egeblad	12/13/18	12/12/19	5,000	*

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
Friends of TJ Foundation Inc.	Dr. Vakoc	09/16/14	09/15/19	50,000	
U.S.–Israel Binational Agricultural Research and Development Foundation	Dr. Lippman	12/01/18	11/30/21	44,000	*
	Dr. Shea	09/01/16	08/31/20	20,834	
University College London/Prostate Cancer UK	Dr. Heavey	01/22/18	01/21/19	4,890	*
The V Foundation	Dr. Tuveson	11/01/16	11/01/19	200,000	
Kathleen and Paul Van Valkenburg	Dr. Li	12/31/18	12/30/19	50,000	*
The Wasily Family Foundation	Dr. Lyons	06/27/18	06/26/19	50,000	*
Joan & Sanford I. Weill Medical College	Dr. Fearon	07/01/14	06/30/19	260,334	
Women's Partnership in Science	Dr. Stillman	01/01/18	12/31/19	194,392	*
The Bradley Zankel Foundation, Inc.	Dr. Mills	11/15/18	11/14/19	20,000	*
	Dr. Zheng	01/21/18	01/20/19	15,000	*
Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation	Dr. M. Hammell	12/01/18	11/30/23	500,000	*
Fellowship Support					
Agency for Science, Technology and Research	Watson School of Biological Sciences	09/01/17	08/31/19	89,258	
Rita Allen Foundation	Dr. Fearon	01/01/17	12/31/18	6,000	
Autism Speaks, Inc.	D. Rupert	10/01/18	09/30/20	40,000	*
Brain & Behavior Foundation	Dr. Crow	01/15/18	01/14/20	35,000	*
	Dr. Gschwend	01/15/18	01/14/20	35,000	*
	Dr. Huilgol	01/15/18	01/14/20	35,000	*
	Dr. Sturgill	01/15/18	01/14/20	35,000	*
	Dr. Wang	01/15/17	01/14/19	35,000	
CSHL Association Fellowship	Dr. Stillman	01/01/18	12/31/19	280,000	*
German National Academy of Sciences Leopoldina	Dr. Schmack	05/01/18	04/30/20	45,071	*
German Research Foundation (DFG)	Dr. Starosta	09/01/18	08/31/20	18,200	*
Lola A. Goldring	Dr. Stillman	10/01/18	09/30/19	100,000	*
Google Inc.	Women in Science and Engineering Initiative	04/04/18	04/03/19	650	*
Howard Hughes Medical Institute	Dr. Tseng	01/02/18	01/01/20	10,000	*
Human Frontier Science Program	Dr. Carnevale	01/01/16	01/12/18	1,635	
	Dr. Xu	04/01/16	03/31/19	55,620	
Annette Kade Charitable Trust	Watson School of Biological Sciences	12/27/18	12/26/19	31,250	*
Betty and Jim Karam	Watson School of Biological Sciences	02/21/18	02/20/19	10,000	*
The Susan G. Komen Breast Cancer Foundation, Inc.	Dr. Albrengues	11/29/16	11/28/19	60,000	
The Meier & Linnartz Family Foundation	Dr. Li	04/05/18	04/04/19	25,000	*
The Pew Charitable Trusts	Dr. Rodriguez-Leal	08/01/16	07/31/20	70,000	
John and Amy Phelan Foundation	Watson School of Biological Sciences	01/01/18	08/31/22	100,000	*
The Research Foundation for State University of New York, Stony Brook	D. Cheng	01/16/17	01/14/21	4,200	
	N. Gallo	11/16/18	11/15/19	4,200	*
	Y. Kim	08/01/18	07/31/19	4,200	*
	R. Raudales	02/16/18	02/15/19	4,200	*
	Y. Allen	07/16/17	07/15/19	4,200	

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]
Charles H. Revson Foundation, Inc.	Dr. Balasooriya	08/01/18	07/31/20	104,013 *
Charles and Marie Robertson Foundation	Dr. Yu	02/16/18	02/15/19	10,000 *
Jason Sheltzer and Joan Smith	Women in Science and Engineering Initiative	12/16/17	12/14/19	1,000
Lauri Strauss Leukemia Foundation, Inc.	Dr. Tarumoto	01/12/18	01/11/19	10,000 *
The Swartz Foundation	Drs. Albeanu/Gupta	01/01/18	12/31/18	50,000 *
	Drs. Engel/Genkin	01/01/18	12/31/18	50,000 *
	Dr. Koulakov/B. Baserdem	01/01/18	12/31/18	25,000 *
Swiss National Science Foundation	Dr. Musall	08/01/17	01/31/19	69,600
University of Southern California	N. Anaparthy	12/27/16	04/30/19	42,665
Training Support				
CUNY Research Foundation	Undergraduate Research Program	08/16/17	08/15/18	15,480
New York State Department of Economic Development	Dr. Mills/CSHL Cancer Gene Discovery & Post Doctoral Research Training Program	02/01/16	01/31/21	70,000
William Townsend Porter Foundation	Undergraduate Research Program	01/03/18	01/02/19	11,800 *
Trinity College	Undergraduate Research Program	08/16/17	08/15/18	247
University of Notre Dame	Undergraduate Research Program	04/01/16	03/31/21	20,000
Course Support				
The Grass Foundation	Neuroscience Scholarship Program in honor of Dr. Ben Barres	11/01/18	10/31/19	11,000 *
The Leona M. & Harry B. Helmsley Charitable Trust	Course Program	02/15/18	02/14/21	1,210,000 *
Howard Hughes Medical Institute	Course Program	08/01/15	07/31/19	600,000
Estée Lauder Inc.	Course Scholarship Program	06/26/18	06/25/21	20,000 *
The Nancy Lurie Marks Family Foundation	Genetics and Neurobiology of Language	06/14/18	06/13/19	38,000 *
Society for Neuroscience/ International Brain Research Organization	Brain Tumor Course	07/01/18	06/30/19	4,255 *
Meeting Support				
Abcam Plc	Epigenetics and Chromatin	06/27/18	06/26/19	2,500 *
Amgen Inc.	Gene Expression and Signaling in the Immune System	04/13/18	04/12/19	5,000 *
	Neurodegenerative Diseases: Biology and Therapeutics	12/20/17	12/19/18	3,000
Art Guild, Inc.	History of Mitochondrial Research	09/14/16	09/13/19	6,000
Avanti Polar Lipids, Inc.	1st International Ferroptosis	08/01/18	07/31/19	3,500
	Nuclear Organization and Function	01/24/18	01/23/19	3,000 *
Biolog, Inc.	History of Mitochondrial Research	09/14/16	09/13/19	4,000
Burroughs Wellcome	1st International Ferroptosis	08/01/18	07/31/19	10,000 *
Calico Life Sciences LLC	Mechanisms of Aging	12/30/15	12/29/18	2,500
Tianqiao & Chrissy Chen Institute	SYMPOSIUM: Brains and Behavior: Order and Disorder in the Nervous System	01/19/18	01/18/19	150,000 *
Children and Screens: Institute of Digital Media and Child Development, Inc.	2nd Digital Media and Developing Minds	09/01/18	08/31/19	158,635 *
CSHL Translational Cancer Support	Diverse Functions of Neutrophils in Cancer	06/01/18	05/31/19	112,437 *
	Mechanisms and Models of Cancer	06/01/18	05/31/19	115,200 *
	Nutrient Signaling	06/01/18	05/31/19	96,000 *
ePlus Technologies Inc.	Biological Data Science	10/01/18	09/30/19	20,000 *

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

Grantor	*Program/Principal Investigator*	*Duration of Grant*		*2018 Funding*[1]	
Genentech, Inc.	Blood Brain Barrier	11/01/18	10/31/19	5,000	*
Georgia State University	Biology and Genomics of Social Insects	03/13/18	03/12/19	10,000	*
Glenn Foundation for Medical Research	Mechanisms of Aging	12/30/15	12/29/18	20,000	
Marnie Halpern & Charles Ippolito	The Zebrafish Neural Circuits and Behavior Meeting	03/13/18	03/12/19	1,713	*
Impact Journals	Mechanisms of Aging	12/30/15	12/29/18	7,000	
Johnson & Johnson Services, Inc.	History of Mitochondrial Research	06/25/18	06/24/19	20,000	*
The Lalor Foundation, Inc.	Germ Cell	07/12/18	07/11/19	8,000	*
Mitobridge, Inc.	History of Mitochondrial Research	09/14/16	09/13/19	3,000	
Modis Therapeutics	History of Mitochondrial Research	09/14/16	09/13/19	3,000	
Neurovive Pharmaceutical AB	History of Mitochondrial Research	09/14/16	09/13/19	3,000	
Proteostasis Therapeutics, Inc.	Protein Homeostasis in Health and Disease	04/17/18	04/16/19	1,000	*
Ribon Therapeutics, Inc.	The PARP Family and ADP-Ribosylation	02/27/18	02/26/19	1,000	*
The Simons Foundation	Molecular Mechanisms of Neuronal Connectivity	09/13/18	11/29/18	3,000	*
Uniqure Biopharma B.V.	RNA & Oligonucleotide Therapeutics	12/21/18	12/20/19	2,500	*
ViiV Healthcare Company	Retrovirus	02/26/18	02/25/19	20,000	*
Wellcome Trust Ltd.	Molecular Mechanisms of Neuronal Connectivity	09/18/18	09/17/19	9,529	*
Library Support					
The Ellen Brenner Memorial Fund		12/15/17	12/16/19	2,000	
The New York State Education Department		07/01/18	06/30/19	4,225	*
The Ensemble Studio Theatre, Inc.		06/28/18	06/27/19	5,000	*
Francis Goelet Charitable Trust		12/15/18	12/14/19	150,000	*
Dr. and Mrs. Philip Goelet		12/21/16	12/20/18	28,800	*
Preprint Server for Biology					
Anonymous	Dr. Inglis	05/01/17	04/30/22	1,228,568	
Anonymous	Dr. Inglis	06/01/18	05/31/20	1,127,572	*

[1]Awarded, including direct and indirect costs

*New or competing renewals or supplements awarded in 2018

DNA LEARNING CENTER GRANTS

Grantor	*Program*	*Duration of Grant*	*2018 Funding*†
FEDERAL GRANTS			
National Institutes of Health	*Barcode Long Island*	7/14–3/20	$252,346
National Science Foundation	*Biotechnology in American High Schools: Then and Now*	9/17–8/18	33,733
National Science Foundation	*Biotechnology in American High Schools: Continuing Research*	9/18–1/19	21,884
National Science Foundation	*Implementing DNA Barcoding for Course-Based Undergraduate Research Experiences*	10/18–9/23	37,744
National Science Foundation	*MaizeCODE: An Initial Analysis of Functional Elements in the Maize Genome*	6/16–5/19	215,716
National Science Foundation	*CyVerse: Cyberinfrastructure for the Life Sciences*	8/18–7/23	49,744
National Science Foundation	*The iPlant Collaborative: Cyberinfrastructure for the Life Sciences*	9/13–8/19	467,517
National Science Foundation	*RCN-UBE: Establishing a Genomics Education Alliance: Steps Towards Sustainability*	9/18–8/20	12,359
NONFEDERAL GRANTS			
Alfred P. Sloan Foundation	*DNA Center NYC* Start-up	12/13–6/19	31,208
Beijing No. 166 High School	Chinese Collaboration Agreement	5/14–6/19	23,902
Breakthrough Prize Foundation	Laboratory Design and Teacher Training for Breakthrough Junior Challenge Prize Winners	12/15–12/18	43,384
Ashley and Frank O'Keefe	Support for Eastwood School and Greenvale School	12/16–12/18	5,000
Health Park	Health Park Agreement	12/15–12/20	17,315
National Grid Foundation	Genetics Education Program	9/16–1/19	12,000
Pinkerton Foundation	*Urban Barcode Research Program*	1/13–5/19	95,863
Richard Lounsbery Foundation	Developing Independent Student Marine Biodiversity Research Using eDNA	10/17–10/19	90,513
Teva Pharmaceuticals	The DNALC Stem Access Fund to Support Usage of the DNALC by Under-Represented Minorities and Disadvantaged Students	10/16–10/19	10,000
The Simons Foundation	*Urban Barcode Research Program*	12/17–8/20	112,841
William Townsend Porter Foundation	*Harlem DNA Lab for Underprivileged Students*	4/16–1/19	13,500

The following schools and school districts participated in the *Curriculum Study* program:

Bellmore–Merrick Central High School District	$2,100	Long Beach Union Free School District	$3,150
East Meadow Union Free School District	$2,100	Massapequa Union Free School District	$3,150
East Williston Union Free School District	$3,150	North Shore Central School District	$2,100
Elwood Union Free School District	$2,100	Oceanside Union Free School District	$2,100
Fordham Preparatory School	$2,100	Oyster Bay–East Norwich Central School District	$2,100
Half Hollow Schools Central School District	$2,100	Plainedge Union Free School District	$2,100
Harborfields Central School District	$2,100	Plainview–Old Bethpage Central School District	$2,100
Herricks Union Free School District	$2,100	Port Washington Union Free School District	$2,100
Island Trees Union Free School District	$2,100	Portledge School	$3,150
Jericho Union Free School District	$2,100	Roslyn Union Free School District	$3,150
Levittown Union Free School District	$2,100	Syosset Central School District	$2,100
Locust Valley Central School District	$2,100	Yeshiva University High School for Girls	$2,100

The following schools and school districts participated in the *Genetics as a Model for Whole Learning* program:

Bayshore Union Free School District	$2,695	Great Neck Union Free School District	$12,700
Berkeley Carroll School, Brooklyn	$4,280	Green Vale School, Old Brookville	$1,662
Cold Spring Harbor Central School District	$15,260	Greenwich Country Day School, Connecticut	$5,280
Commack Union Free School District	$2,530	Half Hollow Hills Central School District	$11,070
East Williston Union Free School District	$1,347	Hicksville Union Free School District	$1,540
Elwood Union Free School District	$9,537	Hofstra University Science and Technology Entry Program	$2,200
Floral Park–Bellerose Union Free School District	$8,250	Holy Child Academy, Old Westbury	$1,800
Garden City Union Free School District	$12,600	Huntington Union Free School District	$2,880

†Includes direct and indirect costs.

Island Park Union Free School District	$2,160
Kings Park Central School District	$3,300
Locust Valley Central School District	$8,817
Massapequa Union Free School District	$1,000
North Bellmore Union Free School District	$3,850
Northport–East Northport Union Free School District	$1,100
Oceanside Union Free School District	$1,800
Our Lady of the Hamptons Regional Catholic School, Southampton	$1,440
Oyster Bay–East Norwich Central School District	$1,320
Port Washington Union Free School District	$6,820
PS/IS 178, Queens	$1,600
Rockville Centre Union Free School District	$11,880
Roslyn Union Free School District	$5,775
Scarsdale Union Free School District	$6,924
School of the Holy Child, Rye	$1,555
Smithtown Union Free School District	$8,800
South Huntington Union Free School District	$7,315
St. Patrick's School, Huntington	$2,160
Syosset Union Free School District	$46,255
Three Village Central School District	$3,190
Wantagh Union Free School District	$3,140

BANBURY CENTER GRANTS

Grantor	*Program*	*2018 Funding*
FEDERAL SUPPORT		
National Institute for Mental Health	NIMH Brain Camp IX	$34,460
NONFEDERAL SUPPORT		
Agilent Technologies, Inc	DNA for Digital Storage	15,000
Alfred P. Sloan Foundation	Signals of Trust in Science Communication	67,100
Bill & Melinda Gates Foundation	What Is Needed for a Comprehensive, Community Response to HIV?	85,959
Burroughs Wellcome Fund	DNA for Digital Storage	15,000
Cold Spring Harbor Laboratory	Increasing Gender Diversity in the Biosciences	50,081
Cold Spring Harbor Laboratory Corporate Sponsor Program	Phase Separated Assemblies in Cell Biology	59,269
Cold Spring Harbor Laboratory Corporate Sponsor Program	Revolutionizing Agriculture with Synthetic Biology	72,990
Cold Spring Harbor Laboratory Corporate Sponsor Program	The Evolving Phenomenon of Direct-to-Consumer Neuroscience	34,949
Cold Spring Harbor Laboratory Corporate Sponsor Program	Bats: New Models for Aging Research	45,820
Cold Spring Harbor Laboratory Corporate Sponsor Program	Autophagy and Cancer	2,344
Northwell Health–Cold Spring Harbor Laboratory Affiliation	Diverse Functions of Neutrophils in Cancer	58,565
Deciphera Pharmaceuticals, Inc.	Autophagy and Cancer	10,000
Genentech, Inc.	Autophagy and Cancer	3,500
Janssen Research & Development	Autophagy and Cancer	10,000
Leap Therapeutics	Emerging Data on the Role of Wnt Biology in Cancer	63,547
Lustgarten Foundation	The Lustgarten Foundation Scientific Meeting	23,813
Mayday Fund	Non-Opioid Management of Chronic Pain: Developing Value-Based Models for Diagnosis and Treatment	55,000
Microsoft Corporation	DNA for Digital Storage	5,000
Numenta, Inc.	Why Does the Neocortex Have Layers and Columns?	57,000
Ovarian Cancer Research Alliance	Towards a Cure for Advanced Stage Ovarian Carcinoma	30,402
Sprint Bioscience	Autophagy and Cancer	10,000
The Swartz Foundation	Quantitative Approaches to Naturalistic Behaviors	40,000
Twist Bioscience Corporation	DNA for Digital Storage	10,000
Vescor Therapeutics	Autophagy and Cancer	20,000

CORPORATE SPONSOR PROGRAM FOR MEETINGS SUPPORT

Contributions from the following companies provide core support for the Cold Spring Harbor meetings program: **Corporate Benefactors:** Regeneron, Posco Research Institute, Estée Lauder Companies; **Corporate Sponsors**: Agilent Technologies, Bristol-Myers Squibb Company, Calico Labs, Celgene, Genentech, Inc., Merck, Bayer, New England BioLabs, Pfizer, Thermo Fisher Scientific; and **Corporate Partners**: Alexandria Real Estate, Gilead Sciences, Lundbeck, Novartis Institutes for BioMedical Research, Sanofi.

The Laboratory acknowledges the generosity of the following companies who loaned equipment and reagents to the various courses: 3i; 10× Genomics; 89North; AB SCIEX LLC; AD Instruments, Inc.; Advanced Analytical; Advanced Illumination; Agilent Technologies, Inc.; ALA Scientific Instruments; A-M Systems; Andor Technologies, Inc.; Applied Scientific Instrumentation, Inc.; Axiom Optics; BD Biosciences; Berthold Technologies USA, LLC; BioLegend; Bioline USA; Bio-Rad Laboratories; BioTek; Biotool; Biovision; Bitplane; BMG Labtech; Bruker Daltonics, Inc.; Cairn Research Ltd.; Campden Instruments; Cell Microsystems, Inc.; Charles River Laboratories, Inc.; Chroma Technology Corporation; CNRS; Coherent; ConOptics; Covaris, Inc.; CrystaLaser; CrystalGen; Diagenode; Drummond Scientific Company; eBioscience; EMD Millipore Corp.; Electron Microscopy Sciences; Epicentre Technologies Corp.; Eppendorf North America; Excelitas Technologies Corp.; Exiqon, Inc.; Fotodyne, Inc.; GE Healthcare; Hamamatsu Photonic Systems; Hamilton Drayage, Inc.; Hampton Research; Harlan; Harvard Apparatus, Inc.; Holoeye; Homebrew; Illumina, Inc.; Intan Technologies; Intelligent Imaging, Inc.; Invitrogen; Life Technologies; Kinetic Systems, Inc.; Labcyte; Leica Biosystems; Leica Microsystems, Inc.; LI-COR; Life Technologies; Lumen Dynamics; Lumencor; Luxendo; MDS Analytical Technologies; MiTeGen; Molecular Devices Corp.; Morrell Instrument Co.; Nanodrop Technologies; Nano-Temper Technologies; Narishige International USA, Inc.; Nasco; New England Biolabs, Inc.; Newport Corporation; Nikon, Inc.; OKO-Lab; PerkinElmer Life and Analytical Sciences; PCO; Photometrics; Physik Instruments; Precisionary Instruments; Prior Scientific; Promega Corporation; ProteinSimple; Qiagen, Inc.; QSonica, LLC.; RC Testing Services; Sage Science, Inc.; Scientifica Ltd.; Singer Instruments; SONY Biotechnology, Inc.; Spectra-Physics; Sutter Instruments; Taconic Farms, Inc.; Takara Bio, Inc.; Tecan Group US, Inc.; the Jackson Laboratories; Thermo Fisher Scientific; Thor Labs; Tokai Hit; Vidrio Technologies; Waters Corporation; White Labs; World Precision Instruments; Xenopus1; Zenith Biotech.

DEVELOPMENT

With help from our loyal supporters, 2018 was a record-breaking year for Cold Spring Harbor Laboratory's annual fund with $7.3 million raised, including $4.5 million for the Double Helix Medals. These vitally important funds provide flexibility to Laboratory President Bruce Stillman to invest in the most innovative research projects that might not yet be eligible for federal funding.

In 2018, the Laboratory continued to receive transformative support from New York State when Governor Andrew Cuomo announced a $30 million gift toward the Neuroscience Research Complex (NRC). The NRC will comprise three state-of-the-art research buildings to further CSHL's leadership in the areas of neuroscience and computer science for the betterment of human health. Scientists at the NRC will focus on cognition, including mapping the wiring diagram of the human brain—a game-changing CSHL technology that will have a profound impact on both developmental neurological disorders such as autism and neurodegenerative diseases including Alzheimer's and Parkinson's diseases.

Most excitingly, longtime Development Chair Marilyn Simons assumed the role of chairwoman of the Laboratory's Board. Her vast experience and focus on philanthropy bring many opportunities for the Laboratory to continue the growth it experienced under the leadership of Jamie Nicholls.

This is truly an important time for science and especially for Cold Spring Harbor Laboratory. Thank you for partnering with us as we continue to transform human health.

Charles V. Prizzi
Vice President for Development and Community Relations

Chairwoman Marilyn Simons, President Bruce Stillman, and former Chairwoman Jamie Nicholls at the 2018 Women's Partnership for Science lunch

Cold Spring Harbor Laboratory Corporate Advisory Board

The Corporate Advisory Board (CAB) comprises prominent business leaders from the tristate community and is a vital source of funding and outreach for Cold Spring Harbor Laboratory. Board members are the driving force behind the Laboratory's annual golf outing at Piping Rock Club that raises critical unrestricted funding for research and education programs. CAB president Eddie Chernoff chaired the 25th annual CSHL outing, which honored CAB member Victoria Meagher. The CAB members also participate in other events and fund-raisers for the Lab and are instrumental "ambassadors" to the community.

Corporate Advisory Board 2018

Edward A. Chernoff, MARS, *Chair*

Michael Aboff, Aboff's Inc.
David Altman, Brown & Altman, LLP
Paul Amoruso, Oxford & Simpson Realty
Todd Andrews, Centerbrook Architects and Planners
Rocco S. Barrese, Dilworth & Barrese
Edward Blaskey, Sterling National Bank
Thomas J. Calabrese, Daniel Gale Sotheby's International Realty
Christopher Callaghan
John D. Catalano, Catalano Enterprises, LLC
Richard A. Catalano, KPMG, LLP
Jonathan Connors, Wells Fargo
Marian Conway, Ph.D., New York Community Bank Foundation
Philip D'Avanzo, Cushman & Wakefield
J. Daniel Decker, Eppendorf North America
Gregory DeRosa, Roanoke Holdings
Robert Dickstein, Ph.D.
Brian Fox, McKinsey & Company
Lawrence Goodman, Curtis, Mallet-Prevost, Colt & Mosle LLP
Mary Jane Helenek
Richard W. Humann, H2M architects + engineers
Robert Isaksen, Bank of America
Alan L. Jakimo, Sidley Austin LLP
Patricia Janco-Tupper, Capital Group
John C. Kean III, Kean Development Company
Michael Keenan, Wells Fargo Bank, N.A.
Norman Kelker, Enzo Life Sciences
Errol Kitt, GEI Consultants
Amy Koreen, M.D.
Andrew Kurita, Kettle Hill Capital Management, LLC
Laurie J. Landeau, V.M.D.
Brian D. Lee, Newmark Grubb Knight Frank
David Lessing, Lessing's
Kyle Markland, Bethpage Federal Credit Union
Jeffrey L. Martin, M.D., Sight MD
Michael Maturo, RXR Realty LLC
Victoria Sagona Meagher
Stephen F. Melore, Farrell Fritz, P.C.
Frank T. O'Keefe, Infinitree LLC
Robert Palatnick, DTTC
John G. Passarelli, M.D.
David Peikon
Patricia Petersen, Daniel Gale Sotheby's International Realty
Joseph Roberto, BankUnited, N.A.
Gareth Roberts, HSBC
Stephen Ross, Nikon
Don Saladino, Drive 495
Raju Sarwal, M.D.
John Topolovec, TD Bank
Craig A. Weiss, T. Weiss Realty Corp.
Joseph M. Willen, Advantage Title

Advisory Board 2018 Honoree Victoria Meagher and her family at the Golf Tournament

Cold Spring Harbor Laboratory Association

Under the leadership of CSHL Association President Michele Celestino, the CSHLA community raised $7.3 million dollars in vital, unrestricted funding to support the globally recognized research and education programs at CSHL. The CSHLA Board of Directors continues its efforts as ambassadors to the community at large. The Directors continue to host dinner parties in their homes for the Dorcas Cummings Symposium and participate in events. The 25th Annual Golf Tournament was once again chaired by Eddie Chernoff and the honoree was Tori Meagher. The Women's Partnership for Science Luncheon was September 16 and featured CSHL's Amanda McBrien, Assistant Director of the CSHL DNA Learning Center, who took us on a fascinating forensic journey through time with the story of Anastasia Romanov. The Double Helix Medals Dinner was November 7 at the American Museum of Natural History, where we honored Larry Norton, Priscilla Chan, and Mark Zuckerberg.

Executive Committee Officers

Michele Munn Celestino, *President*
Kristin Olson Smith, *Vice President*
James A. Woods, *Treasurer*
Virginia Knott, *Secretary*

Directors

Elizabeth Ainslie
Mary Auersperg
Hans E.R. Bosch
Vince Breitenbach
Lisa Broadbent
Barbara W. Callaghan
Eric Carlstrom
Edward A. Chernoff
Frank DellaFera
Nelson DeMille
Anna Saladino Gevinski
Gina Gherlone
David Goldring
Ronald A. Gottlieb
Mark Hamer
Nancy Lippman Israeli, M.D.
Carissa Jordan
Ashley Jostrom
Joel Kassimir, M.D.
Terri Keogh
Peter J. Klein
Michael Maturo
Marcia Kramer Mayer
Michèle Bahnik Mercier
Mickie J. Nagel
Eileen Otto
Whitney F. Posillico
Alicia Zarou Scanlon
Lonnie Shoff
Hope Geier Smith
Debbie Parmet Sondock
Heather Spehr
David M. Stark
Mary Striano
Nancy M. Szigethy

Honorary Directors

Mary D. Lindsay
Anne R. Meier
Cathy Cyphers Soref
Pearl F. Staller
Cynthia R. Stebbins

CSHLA Directors

Honor Roll of Donors

Lifetime Contributions of $5 million +

Estate of Donald Everett Axinn
Dr. and Mrs. Lalit Rai Bahl
Arnold and Mabel Beckman Foundation
Jamie Nicholls and Fran Biondi
David Boies, Chairman of Boies, Schiller and Flexner LLP
The Breast Cancer Research Foundation
Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation
Lavinia Clay
The Dana Foundation
Mrs. Norris W. Darrell
Estate of Kathryn W. Davis
DeMatteis Family Foundation
Mr. and Mrs. Charles F. Dolan, Dolan Family Foundation
The William Stamps Farish Fund
Charitable Lead Annuity Trust under the will of Louis Feil
Jacob Goldfield
Kate Medina Guthart and Leo A. Guthart
The Leona M. and Harry B. Helmsley Charitable Trust
Howard Hughes Medical Institute
W.M. Keck Foundation
Mr. and Mrs. David H. Koch
Laurie J. Landeau Foundation
Mr. and Mrs. Robert D. Lindsay and Family
Dr. Robert W. Lourie
The Lustgarten Foundation
Nancy Abeles Marks
The G. Harold and Leila Y. Mathers Charitable Foundation
Don Monti Memorial Research Foundation
New York State Empire State Development Corporation
Pershing Square Foundation
Cynthia Hazen Polsky and Leon Polsky, Hazen Polsky Foundation
The Quick Family
Charles and Marie Robertson Family
Drs. Marilyn and James Simons
The Simons Foundation
Alfred P. Sloan Foundation
St. Giles Foundation
The Stanley Medical Research Institute
The Starr Foundation
The Swartz Foundation
Thompson Family Foundation
Dr. and Mrs. James D. Watson
Mr. and Mrs. Henry Wendt III
Roy J. Zuckerberg Family Foundation

Contributions of $100,000 +

Mr. and Mrs. Lee S. Ainslie III
The Bahnik Foundation, Inc.
Tianqiao & Chrissy Chen Institute
Mr. and Mrs. Edward A. Chernoff, MARS
Helene Comfort and Liete Urry
Oliver S. and Jennie R. Donaldson Charitable Trust
Edward P. Evans Foundation
Douglas and Christine Fox
Miriam and Alan E. Goldberg Foundation
Lola Goldring
Michael J. Griffin
Heartfelt Wings Foundation Inc.
Mr. and Mrs. Jeffrey E. Kelter
Betsy and Bryan H. Lawrence
Louis Morin Charitable Trust
John Ma, Cheerland Investment Group
Manhasset Women's Coalition Against Breast Cancer
Mr. and Mrs. Yuri Milner
Dr. and Mrs. Howard L. Morgan
Estate of Florence and Harold and Ethel McNeill
William C. and Joyce C. O'Neil Charitable Trust
Amy and John Phelan
Bruce Ratner Family Fund
Mr. and Mrs. Thomas A. Saunders III
Elaine and Eduardo Saverin
Eleanor Schwartz Charitable Foundation
Alison Holtzschue and Douglas Schloss
Edith and Alan Seligson
Dr. and Mrs. James M. Stone
Mr. and Mrs. Paul J. Taubman

Contributions of $30,000 +

Maret and Ted Asaro
Geoffrey Beene Foundation
Mr. and Mrs. Hans E.R. Bosch
Rita and Frank Castagna Family Foundation
Cedar Hill Foundation
Ellen and Casey Cogut
Mr. and Mrs. Roderick H. Cushman
Michel David-Weill
Janet and Frank Della Fera, Fera Pharmaceuticals, LLC
Drs. Pamela Hurst-Della Pietra and Stephen Della Pietra
Laura and John Desmarais
Dr. Lester Dubnick
Ike, Molly and Steven Elias Foundation
Genentech
Joni Gladowsky Breast Cancer Foundation
The Go Go Foundation
Lynn Merrill Gray, The Merrilyn Foundation
The Marc Haas Foundation
The Irving A. Hansen Memorial Foundation
Elizabeth McCaul and Frank Ingrassia
Annette Kade Charitable Trust
Peter J. Klein, Claire Friedlander Family Foundation
Knott Family Foundation
Vesna and Tomislav Kundic
Judy and Leonard Lauder
Rita and Arthur Levinson
Mary D. Lindsay
Breast Cancer Awareness Day in memory of Elizabeth McFarland

2018 Women's Partnership for Science lunch

CAB Chairman Eddie Chernoff and CSHL President Bruce Stillman at the Annual Meeting

Gillian and Eduardo Mestre
Katie Oppo Research Fund
Michelle Paternoster Foundation
Pfizer, Inc.
POSCO Research Institute
Regeneron Pharmaceuticals
Christina Renna Foundation
Robertson Family Fund
Charles and Marie Robertson Foundation
Jonathan Rothberg
Susan Wasserstein and George Sard
Mr. and Mrs. Herbert J. Siegel
Mr. and Mrs. Theodore Smith III, The Geier Foundation
Kristin and Dustin Smith
Mr. and Mrs. Erwin P. Staller
Swim Across America Nassau/Suffolk
Friends of TJ Foundation
Mr. and Mrs. Edward Travaglianti
Mr. and Mrs. Howard J. Tytel
Kathleen and Paul Van Valkenburg
The Wasily Family Foundation
Dr. and Mrs. Michael Wigler
Elisabeth R. Woods Foundation
The Bradley Zankel Foundation

Contributions of $10,000 +

Austin's Purpose
Drs. Cori Bargmann and Richard Axel
Mr. and Mrs. W. Dillaway Ayres, Jr.
Mr. and Mrs. Robert A. Bailenson
Ed Blaskey, Sterling National Bank
Mr. and Mrs. Vincent A. Breitenbach
Broad Hollow Bioscience Park, Inc.
Lisa and Tim Broadbent
David P. Brown
Mr. and Mrs. Thomas J. Calabrese, Jr.
Mr. and Mrs. Eric Carlstrom
Bernadette Castro
Michele and Jim Celestino
Dr. and Mrs. Lan Bo Chen and Lin Huey Chen
Ann Thompson and Jim Childress
Dr. and Mrs. Bayard Clarkson
Dr. and Mrs. William T. Comer
The Conway Family Charitable Fund
Ann Dau Conway, Dau Family Foundation
Barbara J. Amonson and Vincent J. Della Pietra
Marlo Thomas and Phil Donahue
Drs. Miki and David Donoho
Kathleen P. Eden
Eppendorf North America, Inc.
Fortunato Sons Contracting, Inc.
Samuel Freeman Charitable Trust
Mr. and Mrs. Robert A. Gay
Carolyn and Leonard Gero
Celia and Wally Gilbert
Dr. and Mrs. Philip Goelet, Francis Goelet Charitable Trust
Drs. Joan and Alfred Goldberg
Mr. and Mrs. Ronald A. Gottlieb
Mr. and Mrs. John S. Grace
The Randall and Mary Hack Foundation
Robin Hadley
Mr. and Mrs. Mark W. Hamer
Yaoqing & Shouying Huang Family Foundation
Joan Morthland Hutchins
Nancy and Ron Israeli
Carissa and James Jordan
Dr. and Mrs. James D. Karam
Dr. and Mrs. Joel J. Kassimir
Terri and Peter Keogh
Andrea B. and Peter D. Klein
Gordon Lamb, Jefferson Family Charitable Foundation
Mr. and Mrs. Stephen M. Lessing
LUI Che Woo Prize on behalf of Sir James D. Wolfensohn
Mr. and Mrs. John Macaskill, Macaskill Foundation
Rachel and Tom Maniatis
Edward E. Matthews, The Starr Foundation
Mr. and Mrs. Michael Maturo, RXR Realty
Dr. Marcia Kramer Mayer
Diahn and Tom McGrath
Mr. and Mrs. Edward C. Meagher, Carter, DeLuca, Farrell and Schmidt LLP
The Meier and Linnartz Family Foundation
William R. Miller
The Tom and Judy Moore Foundation
Mr. and Mrs. Douglas P. Morris
National Grid Foundation
NBTY Helping Hands Charity
Northwell Health
University of Notre Dame
OBX, Inc.
Gerry Ohrstrom
Susanne Olin
Mr. and Mrs. Jonathan Otto
Drs. Nancy Wexler and Herbert Pardes
Dr. and Mrs. John G. Passarelli, Sight MD
David P. Pearson
Patricia J. Petersen, Daniel Gale Sotheby's International Realty
William Townsend Porter Foundation
Mrs. Richardson Pratt, Jr. and family, The Spionkop Charitable Trust
Mr. and Mrs. Thomas Purcell
Rauch Foundation
Mr. and Mrs. John R. Reese
William J. Rily Memorial Fund
Mr. and Mrs. Douglas E. Rogers
Amy C. Falls and Hartley R. Rogers
Alicia Zarou Scanlon and Richard T. Scanlon
Dr. and Mrs. Phillip A. Sharp
Lonnie Shoff

CSHL President Bruce Stillman, Leonard Lauder, Laura Slatkin, and 2018 Honoree Larry Norton at the Double Helix Medals

CSHLA Directors Eileen Otto and Joel Kassimir

Mr. and Mrs. Harry Slatkin
John Sobolewski
Deborah Solbert
Mr. and Mrs. Douglas S. Soref
Lucy H. Caldwell-Stair and Thomas Osborne Stair
Stand Up For Suzanne
David Stark, Teva Pharmaceutical Industries
Mr. and Mrs. James F. Stebbins
Dr. and Mrs. Bruce W. Stillman
The Lauri Strauss Leukemia Foundation
Nancy Szigethy
Anne D. Thomas
Dori and Peter Tilles
Joan P. Tilney
Mr. and Mrs. Paul A. Vermylen, Jr.
Estate of Marjorie M. von Stade
Anne Wojcicki Foundation
Ann Eden Woodward Foundation

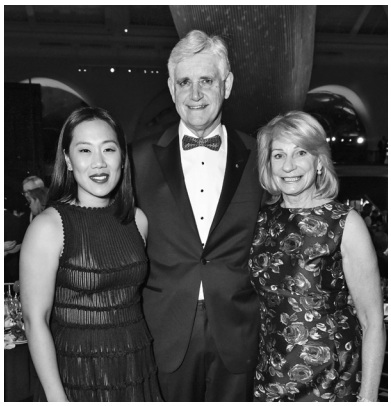
Contributions of $5,000 +

Nancy and Michael Aboff, Aboff's, Inc.
Paul Amoruso, Oxford & Simpson Realty Services, Inc.
Mr. and Mrs. Paul Auersperg
Mr. and Mrs. Fred H. Baron
Deanna Bahel
Deborah C. Benjamin
Judy Bernhang
Kathy and Gene Bernstein
Dr. Steven Brenner
Mr. and Mrs. Christopher C. Callaghan
Jim Carey, Clickit Inc.
Mr. and Mrs. George W. Carmany III
Centerbrook Architects and Planners, LLC
Rita M. Cleary
Mr. and Mrs. Jonathan Connors
Dr. Marian Conway. New York Community Bank Foundation
Mr. and Mrs. James C. Cook
Denise R. Coyle
Maria and Geoff de Lesseps, Alpha Omega Charitable Foundation
Nelson R. DeMille
Dilworth & Barrese, LLP
Jane Duggan, TD Bank
Lee MacCormick Edwards Charitable Foundation
Michael L. Focazio, Carissa Maringo Fund
Putney W. Cloos and Brian C. Fox
Gina and Al Gherlone
The Deane A. and John D. Gilliam Foundation
Glen Cove C.A.R.E.S.
Carl Goldfischer
H2M architects + engineers
The Hastings Foundation, Inc.
Jo-Ellen and Ira Hazan
Hearts Who Care
Robin and Jim Herrnstein
Nancy Hopkins
Alan L. Jakimo
Harvey L. Karp, The Karp Foundation, Inc.
Errol Kitt, GEI Consulting Engineers & Scientists
Charles and Jane Klein Family Fund
The Klinsky Family Charitable Fund
Drs. Amy and Roger Koreen
Andrew Y. Kurita
David Lessing, Lessing's Food Service Management
Dr. and Mrs. Arnold J. Levine
Robert T. Lindsay
Mr. and Mrs. Thomas H. Lister
The David & Sondra Mack Foundation
Mrs. David J. Mahoney
Robert L. Marcus
Kyle Markland - Bethpage Federal Credit Union
Patti and John Maroney
Masthead Cove Yacht Club
Stephen F. Melore, Farrell Fritz, P.C.
Mickie and Jeff Nagel
Mr. and Mrs. Wright Ohrstrom
Wendy O'Neill
Tracy Dellomo and Clifford Packingham
Mr. and Mrs. Robert Palatnick
Mr. and Mrs. John Pufahl
Kathleen Rawlings
Joe Roberto, BankUnited, N.A.
Gareth Roberts, HSBC Bank USA
Stephen Ross, Nikon Instruments, Inc.
Lari and Frank Russo
UBS Financial Services, Inc.
Susan and Charles Sawyers
Debbie Parmet Sondock
Mr. and Mrs. Richard A. Spehr
Mr. and Mrs. James F. Stebbins
Lee Steinberg
Carl J. Studer and Michele Miroff
Three Strohm Sisters Family Foundation
Trinity College
Dr. George F. Vandewoude, Dr. Susan Vandewoude
John B. Vermylen
Dr. and Mrs. Stuart Weisbrod
Alanna and Craig Weiss
West Islip Breast Cancer Coalition for L.I.
Georgene and Steven Winick

Contributions of $1,000 +

Dr. and Mrs. Bruce M. Alberts
Debra and Scott Arenare
Tricia and Hal Avidano
Babylon Breast Cancer Coalition
Mrs. Robert H. Baldwin
Baranay Family Foundation
Mr. and Mrs. Dominic P. Bencivenga
Elizabeth Bertani
Jan and Don Bettencourt
Dr. William Braden, Porter Braden Fund at the Rhode Island Foundation
Katharine Bradford Foundation
Barbara Taylor Bradford and Robert Bradford
Mr. and Mrs. Matthew J. Bruderman
Josephine Marrali and Richard Bryant
C.M.S. Electrical Services Inc.
Canon Solutions America
Carol and Stephen Canter
CardWorks Servicing, LLC
Louise M. Parent and John Casaly
Mary and Richard Catalano
Susan and Glenn Cohen
Maggie and John Cooley
Cosel-Pieper Family Foundation
Kristin Holby Darnell and Dr. James E. Darnell
Marilyn H. DeLalio
Dime Community Bank
David Drillick
Frederick W. Engel
Diane and Bob Fagiola
Floral Park Police Benevolent Association
Jane Furse and John H. Friedman
Bess and David Fuchs
Mary Jane Helenek and Norman Futterman
E. Maxwell Geddes, Jr.
Dr. and Mrs. Walter L. Goldschmidts
Joseph L. Goldstein
Mr. and Mrs. Pierre Gonthier
Lorna and Edwin Goodman
Lawrence Goodman, Curtis, Mallet-Prevost, Colt & Mosle LLP
Katy and Marc Goodrich
Mr. and Mrs. William H. Grover
Qing Gu
Hachette Book Group

CSHLA Directors Michele Celestino, Heather Whitman, and Mickie Nagel



Honoree Priscilla Chan, President Bruce Stillman, and Chairwoman-elect Marilyn Simons at the 2018 Double Helix Medals

Patricia and Doug Hammond
Reverend Linda Peyton Hancock
Margaret M. Hargraves
Dr. and Mrs. William Healy III
Drs. Nouria Hernandez and Winship Herr
William A. Herzog
Cindy and Bob Higginson
Mary Jo and Mark Hoffman
Regina Imperio
Marjorie and Fredd Isaksen
Walter B. James Fund No. 2
Dr. Leemor Joshua-Tor
Mr. and Mrs. Gabriel Jostrom
John C. Kean, Kean Development Company, Inc.
Dawn and Alan Kisner
Daniel F. Klessig
John and Patricia Klingenstein Fund
Heather and Daniel Knopf
Laureen and Ragnar M. Knutsen
Dr. Adrian Krainer
Mr. and Mrs. James M. Large, Jr.
Isabel B. Leib
Mr. and Mrs. Michael L. Leidner
Maureen and Tony Leness
Fred Leonardo, Electronix Systems
Nancy Metz and George N. Lindsay, Jr.
The Litwin Foundation
Mr. and Mrs. Carl Lombardi
Long Island 2 Day Walk to Fight Breast Cancer
Ruth S. McElheny and Victor K. McElheny
Mr. and Mrs. Neil T. McGoldrick
William McHenry
Thomas McNally
Donald G. McNeil, Jr.
Medical Arts Radiological Group, P.C.
Gary Melius
John Miller
Walter Mondschein
Mr. and Mrs. Stephen V. Murphy
Holly Myers and Kirk Neely
Norwood Foundation, Inc.
Jared Nussbaum
Patricia and Hugh O'Kane, Jr.
Mr. and Mrs. Frank T. O'Keefe
Karen H. Orzel
Lisa and Gil Ott
Mr. and Mrs. Richard S. Overton
Ann B. Goodman and Arthur B. Pardee
Rhoda Parmet
Mr. and Mrs. Nicholas B. Paumgarten
Terry and Dave Peikon
Peto Maintenance and Development, LLC
Dr. and Mrs. Anthony F. Piazza
Pilkington Family Fund
Platsky Company, Inc.
Mrs. Joseph D. Posillico, Jr.
Nicole and Charlie Prizzi
Lucy Gordon and Mark Ptashne, Joe Lewis Jefferson Foundation
R2D Consulting LLC
Bonnie Rechler
Mr. and Mrs. Lawrence Remmel
Dr. Jeffrey Roberts
Linda K. Rodgers
Dr. and Mrs. Gerald M. Rubin
Mr. and Mrs. Keith Ryan
Dr. and Mrs. Raju Sarwal
The Schiff Foundation
Robin and Enrique Senior
Joan Smith and Jason Sheltzer
Mr. and Mrs. H. Brooks Smith
Bonnie Sollog
Mr. and Mrs. Howard Solomon
Silda Wall Spitzer
Frank Spokane
Mr. and Mrs. Vincent T. Striano
Syosset Limousine Service, Inc.
Judith and Irwin Tantleff
Mrs. David S. Taylor
Jeremy W. Thorner
Pamela M. Thye
Mr. and Mrs. Bart T. Tiernan
Mrs. Carl W. Timpson, Jr.
Susumu Tonegawa, Mayumi Tonegawa
Dr. and Mrs. Alexander T. Traykovski
Mr. and Mrs. Stanley S. Trotman, Jr.
Sarah Vermylen Trust
Barbara Ulman
Uni-Data & Communications, Inc.
Universe Kogaku (America), Inc.
Thomas Vladeck
Richard John Walk
Mr. and Mrs. Thomas J. Walsh
Mr. Douglas A. Warner III
Gerald I. White
Mr. and Mrs. Erby Wilkinson
Renee and Richard Worrall

In-Kind Gifts

Americana Manhasset
Sophie and Dill Ayres
Banfi Vintners
Besito
Bottles and Cases
Lisa Broadbent
Brooks Brothers
Butera's Restaurant of Woodbury
Barbara Callaghan
John D. Catalano
Edward Chernoff
Claytime
Cold Spring Harbor Wine Shoppe
Country Club Studios
Curtis, Mallet-Prevost, Colt and Mosle LLP
Daniel Gale Sotheby's International Realty Inc.
Nelson DeMille
Drive 495
Eileen Kathryn Boyd Interiors
Eleanor Stone NYC
John W. Engeman Theater
Estée Lauder
Fiorello Dolce
Esther Fortunoff
Frassanito Jewelers
Candido Fuentes-Felix, M.D.

Kate Medina Guthart, CSHL Trustee Leo Guthart, Leslie Stahl, Dill Ayres, and Dave Tuveson at the 2018 Double Helix Medals

Members of Manhasset Women's Coalition Against Breast Cancer pictured with Drs. David Spector and Sonam Bhatia. They donated $100,000 to support Dr. Bhatia's breast cancer research.

GarageTek
Gina and Al Gherlone
Girl Obsessed
Dagnia and Walter Goldschmidts
Harbor Mist Restaurant
Harlan Estate
Huntington Country Club
Huntington Indoor Tennis
Ron Israeli, M.D.
J. McLaughlin
Jan Tozzo, Glass.Pieces.Studio
Jewel
Jonathon's Ristorante
Kimberly Gorman Muto Photography
Living
MARS
Mission Nutrition
Mitchells Huntington
Molton Brown
Nest
Nikon Instruments, Inc.
Oheka Castle Hotel and Estate
Heidi and John Passarelli
David Peikon
Poll Restaurants
Prime Restaurant
Red Restaurant
Annabel Romero-Hernandez
Safavieh
Sandbar
Raju Sarwal, M.D.
Scarsella's
Tom Schaudel
SGS Original Art, Susan Graham Switala
The Shed
Sports Illustrated Golf Group
Priya Sridevi
Sterling National Bank
Grace and Bruce Stillman
TD Bank
Title Boxing Club of Huntington
Cathy Freudenberg Traykovski
Tres Jolie Salon
Marjorie van de Stouwe, M.D.
Walters Faith
W.B. Mason
Wempe
Windham House

Contributions in honor of

Jamie Nicholls and Fran Biondi
Kristina Perkin Davison
Lester Dubnick
Arash and Mahsan Farhadieh
Michael Gynizio
Robert Lindsay
Mardi Matheson
Victoria S. Meagher
Alea Mills
Kate O'Phelan
Mr. and Mrs. Mitchel Rossman
Susan Schultz
Fay Siegel
Bruce W. Stillman
Mr. and Mrs. Peter R. Tauberer
Dr. James D. Watson
Michael Wigler

Contributions in memory of

Sandy Aranoff
TJ Arcati
Henry Austin
Priscilla Averbuck
Stu Bailin
Anne E. Beaudreault
Dr. Joseph Bennett
Arthur Bernhang
Giuseppe Bertani
Vito Bianco
Anthony Boateng-Mensah
William M. Brady
Ellen Brenner
Carol A. Buonaiuto
Vincent Carey
Sabi Castro
John P. Cleary
Daniel W. Coakley
Lila Dannenberg
Katya H. Davey
Sandy DeMille
Phillip A. Dissinger
Claire Martin Fairman
Christine Fazio
Jean Fitzpatrick
Jerry Frankel
Stephen Franz
Marilyn Friedland
Shirley Friedman
Nancy S. Gay
Joni Gladowsky
Christopher G. Goff
Norman Goldfinger
Jerry Goldman
Melvin Gorelick
Teresa Haire
Charles E. Harris II
Murray Hazan
Barbara Hoover
Joseph E. Houle
Erik P. Jacobsen
Bernice Kahn
Joseph J. Leto
Allan Lisse
David L. Luke III
Laura Malchick
Carol Marcincuk
Larry Margott
Carissa Maringo
Marjorie A. Matheson
Marie Matthews
Frank Morelli
Barbara Novak
Francis O'Connor
Kevin O'Connor
Robert H. Olney
Barry Osborn
Michelle Paternoster
Rhoda Phillips
Christina Renna
Maria Replak
William J. Riley
Diane Emdin Sachs
Daniel Schecter
Andrew Seligson
Sheila Silverstein
Adelaide Simpson
Michelle Smith
Jason Stewart
Morris Sunshine
Sylvia Taubman
Shirley and Melvin Thorner
Helene P. Victor
Marjorie von Stade
Steven M. Walkowski
William A. and Ellastine V. Weekes
Elizabeth Brunson Wilson
Mark Zoller

LABORATORY MANAGEMENT

Bruce W. Stillman, Ph.D., President & CEO
John P. Tuke, Chief Operating Officer

RESEARCH

David L. Spector, Ph.D.
Director of Research

Walter L. Goldschmidts, Ph.D.
Vice President, Executive Director of Sponsored Programs

Sydney Gary, Ph.D.
Director, Research Operations

Professors

Douglas Fearon, M.D.
Hiroyasu Furukawa, Ph.D.
Thomas Gingeras, Ph.D.
Professor and Head Functional Genomics
Z. Josh Huang, Ph.D.
David Jackson, Ph.D.
Leemor Joshua-Tor*, Ph.D.
Adam Kepecs, Ph.D.
Program Chair, Neuroscience
Alexei Koulakov, Ph.D.
Adrian Krainer, Ph.D.
Program Chair, Cancer and Molecular Biology
Bo Li, Ph.D.
Zachary Lippman*, Ph.D.
Robert Maki, M.D., Ph.D.
Robert Martienssen*, Ph.D.
Program Chair, Genomics and Plant Biology
W. Richard McCombie, Ph.D.
Alea Mills, Ph.D.
Partha Mitra, Ph.D.
Adam Siepel, Ph.D.
Program Chair, Simons Center Quantitative Biology
Nicholas Tonks, Ph.D.
Deputy Director, CSHL Cancer Center
Lloyd Trotman, Ph.D.
David Tuveson, M.D., Ph.D.
Director, CSHL Cancer Center
Linda Van Aelst, Ph.D.
Michael Wigler, Ph.D.
Anthony Zador, M.D., Ph.D.

Associate Professors

Dinu Albeanu, Ph.D.
Gurinder Atwal, Ph.D.
Anne Churchland, Ph.D.
Mikala Egeblad, Ph.D.
Jesse Gillis, Ph.D.
Christopher Hammell, Ph.D.
Molly Hammell, Ph.D.
Ivan Iossifov, Ph.D.
Alexander Krasnitz, Ph.D.
Dan Levy, Ph.D.
Pavel Osten, M.D., Ph.D.
Stephen Shea, Ph.D.
Raffaella Sordella, Ph.D.
Arne Stenlund, Ph.D.
Christopher Vakoc, M.D., Ph.D.
Doreen Ware**, Ph.D.

Assistant Professors

Alexander Dobin, Ph.D.
Camila dos Santos, Ph.D.
Tatiana Engel, Ph.D.
Tobias Janowitz, M.D., Ph.D.
Justin Kinney, Ph.D.
Je Hyuk Lee, M.D., Ph.D.
David McCandlish, Ph.D.
Vincent Pedmale, Ph.D.
Jessica Tollkuhn, Ph.D.

Research Professor

Darryl Pappin, Ph.D.

Research Assistant Professors

Kenneth Chang, Ph.D.
Scott Lyons, Ph.D.
Jonathan Preall, Ph.D.
Michael Ronemus, Ph.D.
Johannes Yeh, D.D.S., Ph.D.
Seungtai Yoon, Ph.D.

Senior Computer Scientists

Peter Andrews, Ph.D.
Boris Yamrom, Ph.D.

Computer Scientists

Osama El Demerdash, Ph.D.
Vivek Kumar, Ph.D.

Senior Research Investigator

Zihua Wang, Ph.D.

Research Investigators

Joan Alexander, M.D.
Lindsey Baker, Ph.D.
Clary Evans, M.B.Ch.B.
Christopher Faehnle, Ph.D.
Qing Gao, M.D., Ph.D.
Carla Gauss, Ph.D.
Sara Goodwin, Ph.D.
Gilbert Henry, Jr., Ph.D.
Manzar Hossain, Ph.D.
Yoon-Ha Lee, Ph.D.
Youngkyu Park, Ph.D.
Michael Regulski, Ph.D.
Katharina Schmack, M.D., Ph.D.
Yi-Jun Sheu, Ph.D.
Damianos Skopelitis, Ph.D.
Priya Sridevi, Ph.D.
Asya Stepansky, Ph.D.
Kai Yu, Ph.D.
Huiqing Zhan, Ph.D.

Manager, Clinical and Translational Research

Jessica Gotterer

Professor Emeritus

James D. Watson, Ph.D.
Oliver R. Grace Professor Emeritus/Chancellor Emeritus

Adjunct Professors

Gregory Hannon, Ph.D.
Scott Lowe, Ph.D.
Larry Norton, M.D.
Scott Powers, Ph.D.

Adjunct Associate Professor

Michael Schatz, Ph.D.

Visiting Scientists/ Collaborative Scientists

Ebru Demir, Ph.D.
Hyun Jae Pi, Ph.D.
Michael Wulf

Visiting Professors

Matthew Ridley, Ph.D.
Stuart Weisbrod, Ph.D.

QB Visiting Scholar

Bhubaneswar Mishra, Ph.D.

Director, Gene Targeting Shared Resources

Leyi Li, Ph.D.

Clinical Fellows

Rintaro Koike, M.D.
Kenneth Yu, M.D.

CSH Fellows

Semir Beyaz, Ph.D.
Jason Sheltzer, Ph.D.
Lingbo Zhang, Ph.D.

Translational Fellow

Fieke Froeling, M.D., Ph.D.

Postdoctoral Fellows

Jean Albrengues, Ph.D.
Xu An, Ph.D.
Balasooriya Balasooriya, Ph.D.
Sara Ballouz, Ph.D.
Leah Banks, Ph.D.
Prasamit Baruah, Ph.D.
Walter Bast, Ph.D.
Pascal Belleau, Ph.D.
Yael Berstein, Ph.D.
Kuhulika Bhalla, Ph.D.
Sonam Bhatia, Ph.D.

*Employee of the Howard Hughes Medical Institute
**Employee of the United States Department of Agriculture

Sonali Bhattacharjee, Ph.D.
Giulia Biffi, Ph.D.
Amit Blumberg, Ph.D.
Christopher Bonham, Ph.D.
Marion Bosc, Ph.D.
Christian Bravo-Rivera, Ph.D.
Robert Bronstein, Ph.D.
Irene Casanova Salas, Ph.D.
Hong Goo Chae, Ph.D.
Chen Chen, Ph.D.
Wei Chia Chen, Ph.D.
Xiaoyin Chen, Ph.D.
Yushu Chen, Ph.D.
Quentin Chevy, Ph.D.
Tsung Han Chou, Ph.D.
Tae Mun Chung, Ph.D.
Hannes Claeys, Ph.D.
Megan Crow, Ph.D.
Samantha Cyrill, Ph.D.
Juliane Dassler Plenker, Ph.D.
Edgar Demesa Arevalo, Ph.D.
Hanfei Deng, Ph.D.
Cristiane de Santis Alves Rosa, Ph.D.
Jakub Dolata*, Ph.D.
Noah Dukler, Ph.D.
Roman Dvorkin, Ph.D.
Elad Elkayam*, Ph.D.
Ela Elyada, Ph.D.
Mary Feigman, Ph.D.
Miriam Fein, Ph.D.
Stephan Fischer, Ph.D.
Matthew Fisher, Ph.D.
Daniel Fuerth, Ph.D.
Akihiro Funamizu, Ph.D.
Alessandro Furlan, Ph.D.
Yuan Gao, Ph.D.
Mikhail Genkin, Ph.D.
Fabian Gerth, Ph.D.
Debarati Ghosh, Ph.D.
Nicholas Gladman, Ph.D.
Olivier Gschwend, Ph.D.
Wuqiang Guan, Ph.D.
Priyanka Gupta, Ph.D.
Jaynee Hart, Ph.D.
Rasmani Hazra, Ph.D.
Jianping He, Ph.D.
Xueyan He, Ph.D.
Anat Hendelman, Ph.D.
Diego Hernandez Trejo, Ph.D.
Longwen Huang, Ph.D.
Yifei Huang, Ph.D.
Yuhan Huang, Ph.D.
Dhananjay Huilgol, Ph.D.
Stellar Hur, Ph.D.
Jonathan Ipsaro, Ph.D.
Anbalagan Jaganathan, Ph.D.
Matt Jaremko, Ph.D.
Yinping Jiao, Ph.D.
James Jung*, Ph.D.
Gukhan Kim, Ph.D.
Brian Kinney, Ph.D.
Munenori Kitagawa, Ph.D.
Olaf Klingbeil, Ph.D.
Hideki Kondo, Ph.D.
Choon-Tak Kwon, Ph.D.
Seung Cho Lee*, Ph.D.
Bo Li, Ph.D.
Jiayun Li, Ph.D.
Qian Li, Ph.D.
Shujing Li, Ph.D.
Siran Li, Ph.D.
Yi Li, Ph.D.
Jayon Lihm, Ph.D.
Kuan-Ting Lin, Ph.D.
Bodu Liu, Ph.D.
Lei Liu, Ph.D.
Bin Lu, Ph.D.
Wai Kit Ma, Ph.D.
Grinu Mathew, Ph.D.
Katherine Matho, Ph.D.
Olivia Mendivil Ramos, Ph.D.
Kevin Michalski, Ph.D.
Koji Miyabayashi, M.D., Ph.D.
Andrea Moffitt, Ph.D.
Hemanth Mohan, Ph.D.
Rodrigo Munoz Castaneda, Ph.D.
Adriana Munoz Jimenez, Ph.D.
Simon Musall, Ph.D.
Farzaneh Najafi, Ph.D.
Arun Narasimhan, Ph.D.
Farhana Nasrin, Ph.D.
Gal Nechooshtan, Ph.D.
Ellen Louise Noren Lindback, Ph.D.
Kin On*, Ph.D.
Torben Ott, Ph.D.
Huiwu Ouyang, Ph.D.
Meng Ouyang, M.D., Ph.D.
Chand Parvez Danka Mohammed, Ph.D.
Sapan Patel, Ph.D.
Dennis Plenker, Ph.D.
Anna Posfai, Ph.D.
Yongjun Qian, Ph.D.
Mohammad Rahman, Ph.D.
Nissim Ranade, Ph.D.
Ashlan Reid, Ph.D.
Benjamin Roche, Ph.D.
Nikolay Rozhkov, Ph.D.
Subramanian Sankaranarayanan, Ph.D.
Thiago Santos Gouvea, Ph.D.
Christine Scaduto, Ph.D.
Andrea Schorn, Ph.D.
Harrison Seidner, Ph.D.
Yanliang Shi, Ph.D.
Atsushi Shimada*, Ph.D.
Timothy Somerville, Ph.D.
Sebastian Soyk, Ph.D.
Sarah Starosta, Ph.D.
James Sturgill, Ph.D.
Lijuan Sun, Ph.D.
Xueqin Sun, Ph.D.
Johanna Syrjanen, Ph.D.
Yilin Tai, Ph.D.
Julian Taranda, Ph.D.
Yusuke Tarumoto, Ph.D.
Claudia Tonelli, Ph.D.
Michael Tramantano, Ph.D.
Roger Tseng, Ph.D.
Anne Urai, Ph.D.
Raditya Utama, Ph.D.
Anand Vasudevan, Ph.D.
Prabhadevi Venkataramani, Ph.D.
Ledong Wan, Ph.D.
Bo Wang, Ph.D.
Bor-Shuen Wang, Ph.D.
Minghui Wang, Ph.D.
Xingang Wang, Ph.D.
Xinyi Wang, Ph.D.
Zhikai Wang, Ph.D.
Yiliang Wei, Ph.D.
Sze Wong, Ph.D.
Qingyu Wu, Ph.D.
Xiong Xiao, Ph.D.
Fang Xu, Ph.D.
Xiaosa Xu, Ph.D.
Alexander Xue, Ph.D.
Tao Yang, Ph.D.
Zhaolin Yang, Ph.D.
Min Yao, Ph.D.
Li Yuan, Ph.D.
Dinar Yunusov, Ph.D.
Xian Zhang, Ph.D.
Yixin Zhao, Ph.D.
Juannan Zhou, Ph.D.
Imanol Zubiete Franco, Ph.D.

Graduate Students

Disha Aggarwal
Lyndsey Aguirre
Naishitha Anaparthy
Seamus Balinth
Batuhan Baserdem
Giorgia Battistoni
Benjamin Berube
Brianna Bibel
Anna Lien Bouhuis
Sara Boyle
Emilis Bruzas
Jason Carter
Yuxin Cen
Kung-Chi Chang
Derek Cheng
King Hei Cheng
Hsiang-Chen Chou
Danielle Ciren
Alberto Corona
Marie Dussauze
Ilgin Ergin
Connor Fitzpatrick
Talitha Forcier, Ph.D.
William Galbavy
Nicholas Gallo
Jordan Gegenhuber
Sanchari Ghosh
Hamza Giaffar
Jacqueline Giovanniello
Kristina Grigaityte
Michael Gutbrod
Amritha Varshin Hanasoge Somasundara
Benjamin Harris
Kelly Hills-Muckey
Yixin Hu
Yuzhao Hu
Elizabeth Hutton
Shruti Iyer
David Johnson
Young Jin Kim
Navyateja Korimerla
Christopher Krasniak
Asad Aziz Lakhani
Matthew Lee
Cristian Lopez-Cleary
Shaina Lu
Devon Lukow
Lital Lynn, Ph.D.
Dennis Maharjan
Diogo Maia e Silva, M.D.
Laura Maiorino
Paul Masset, Ph.D.
Katarina Meze
Judith Mizrachi
Ziyi Mo
Manojit Mosur Swamynathan
Alexandra Nowlan
Kathryn O'Neill
Tobiloba Oni
Alexa Pagliaro
Tzvia Pinkhasov
Sashank Pisupati
Sofya Polyanskaya
Zhe Qian
Ricardo Raudales
Suelynn Ren

*Employee of the Howard Hughes Medical Institute

Jenelys Ruiz
Deborah Rupert
Danilo Segovia
Luqun Shen
Margaret Shevik
Padmina Shrestha
Sergey Shuvaev
Dongyan Song
Olya Spassibojko
Martyna Sroka
Joshua Steinberg
Colin Stoneking
Jennifer Thalappillil
Ngoc Tran
Gaurang Trivedi
Jue Xiang Wang
Jonathan Werner
Xiaoli Wu
Cole Wunderlich
Wenbo Xu
Ran Yan
Georgi Yordanov
Allen Yu
Zhezhen Yu
Chengxiang Yuan
Anqi Zhang
Qian Zhang
Siwei Zhang, Ph.D.

Computational Science

Samik Banerjee
Daniel Blibaum
Mihail Bota, Ph.D.
Kapeel Chougule
Astrid Deschenes
Evan Ernst*
Daniel Ferrante, Ph.D.
Elena Ghiban
Yuan Hao, Ph.D.
Hussein Hijazi, Ph.D.
Ying Jin, Ph.D.
Jude Kendall
Melissa Kramer
Sunita Kumari, Ph.D.
Xu Li
Zhenyuan Lu
Steven Marks
Ivar Meijs
Matthew Moss
Andrew Olson
Jason Palmer
Xiaoli Qi, Ph.D.
Ritika Ramani
Manthan Shah
Joshua Stein, Ph.D.
Oliver Tam, Ph.D.
Ammar Tareen, Ph.D.
Marcela Tello-Ruiz, Ph.D.
Kannan Umadevi Venkataraju
Liya Wang, Ph.D.
Xiaofei Wang, Ph.D.
Xuehong Wei
Matthew Wroten, Ph.D.
Fengyuanshan Xu
Miu ki Yip

Research Associates

Habeeb Al Sudani
Carmelita Bautista
Barry Burbach
Hilary Cox
Jorg Drenkow
Eric Gamache
Hyun Soo Kim, Ph.D.
Lisa Manche
Joseph Merrill
Ramesh Palaniswamy
Anirban Paul, Ph.D.
Suzanne Russo
Caizhi Wu, M.D.
Melody Wu, Ph.D.
Scilla Wu
Lifang Zhang, Ph.D.

Scientific Project Managers

Katherine Brenner
Kathleen Kelly, Ph.D.
Alexander Vaughan, Ph.D.

Laboratory Facility Managers

Sabrina Boettcher
Amy Brady
Patricia Brady
Julie Cheong
Robert Eifert
Nour El-Amine
Stephanie Goldsmith*
Calvin Lee
Pamela Moody
Timothy Mulligan
Keith Rivera
Joseph Simorowski*
Dennis Thomas, Ph.D.
Ming Wang

Scientific IT Support

John Pane
David Trimboli
Peter Van Buren

Scientific Administrators

Lorraine Baldwin
AnnMarie Ballew
Christine Bedell
Patricia Bird
Jessica Bisceglio*
Barbara Cascone
Melissa Daley
Becky Dong
Susan Fredericks
Irene Gill
Denise Hallinan
Lisa Kimoto
Lynn Leonard
Idee Mallardi
Lindsay Monaco
Sandra Moran
Kylie Parker
Catherine Scalise
Dorothy Tsang Hu
Madeline Wisnewski

Research Technicians

Valeria Aguillon Rodriguez
Priya Alagesan
Andalus Ayaz
Lukas Benjamin
Anastasiia Bludova
Lily Brine
Denise Cahn
Yon Chang
Devon Cowan
Jenesis Curtis
Cassidy Danyko
Martin Davis
Michael Dworjan
Amanda Epstein*
Audrey Fahey
Margaret Falkowski
Christy Felice
Eugene Fong
Jared Freedman
Ya Gao
Libia Garcia
Steven Georgiou
Elizabeth Gibson
Steven Gluf
Inessa Hakker
Kaarina Hanington
Joshua Hatfield
Kendall Horst
Jason Hsu
Fangle Hu
Qingting Hu, M.D.
Cassandra Kanzler
Clancy Kelahan
Anup Khanal
Mirna Kheir Gouda
Jennifer Kim
Maja Kornaj
Yan Li
Ying Hsiu Liu
Landon Lodato
Kathleen Lucere
Matthew Lucito
Beicong Ma
Arthur Makarenko
Krisann Marquart
Senem Mavruk Eskipehlivan
Joseph Milazzo
Kristin Milicich
Stephanie Muller
David Ng
Aigoul Nourjanova
Eleni Noutsos, Ph.D.
Joseph O'Rourke
Jie Ouyang
Ann Palladino
Sarah Park
Hardik Patel
Umamaheswari Ramu
Claire Regan
Michael Riggs
Gina Robitaille
Julie Rosenbaum
Taylor Russo
Polona Safaric Tepes
Stephen Savoia
John Scheuring, Jr.
Jennifer Shapp
Regina Shaw
Aubrey Siebels
Noriko Simorowski
Tara Skopelitis
Junho Song
Benjamin Spielman
Logan Stecher
Yu-Chi Sun
Alexander Tolpygo
Claire Uhlman
Wiktor Wadolowski
Jing Wang
Robert Wappel
Kaitlin Watrud
Simone Weinmann
Shih Ting Yang
Melissa Yao

Laboratory Facility Staff

Rachel Rubino, D.V.M.
Director of Lab Animal Resources/ Attending Veterinarian
Alexander Aguilera

*Employee of the Howard Hughes Medical Institute

Carlos Ballon
Damir Begonja
Lisa Bianco
Alexis Blanco
Erin Brosnan
Oral Bryan
Amy Burrows
Manuel Canas
Jodi Coblentz
Steven Collins
Kristin Cox
Charlene De Poto
Maria Demott
Jared Downing
Rhonda Drewes
Eileen Earl
Stephanie Flanagan
Wilson Garciaguirre
Susan Guterman
Jill Habel
Thomas Hicks
Tristan Jarrett
Andrew Labriola
Alyssa Molloy
Maria Mosquera
Gustavo Munoz
Jason Pellegrini
Bryan Rose
Ann Smith
Laura Stonebridge
Walter Vigil Garcia
Peter Wimett, Jr.

Media Makers/Farm Assistants

Nancy Bolanos
Terrance Chisum
Adriana Hincapie
Kelly Ann Holfester
Michele Murrell
Shouyun Qiao
Amy Qiu

EDUCATION

Banbury Center

Rebecca Leshan, Ph.D.
Director
Jasmine Breeland
Michelle Corbeaux
Barbara Polakowski

CSHL Press

John R. Inglis, Ph.D.
Executive Director
Barbara Acosta
Janet Argentine
Kevin John Black
Cynthia Blaut
Carol Brown
Kathleen Bubbeo
Kathleen Cirone
Laureen Connell, Ph.D.
Marie Cotter
Jennifer DeLeon, Ph.D.
Ahmet Denli, Ph.D.
Corrisa Farmer
Bibiane Garite
Michael Henigman
Samantha Hindle, Ph.D.
Nancy Hodson
Deborah Jarski
Brandon Kelly
Catherine Kelly
Tara Kulesa
Dana Macciola
Wayne Manos
Mala Mazzullo
Joanne McFadden
Mary Mulligan
Stephanie Novara
Stephen Nussbaum
Dorothy Oddo
Robert Redmond
Theodore Roeder
Mary Sabala
Eric Sawey, Ph.D.
Diane Schubach
Richard Sever, Ph.D.
Inez Sialiano
Marcie Siconolfi
Maria Smit, Ph.D.
Hillary Sussman, Ph.D.
Linda Sussman
Denise Weiss

DNA Learning Center

David A. Micklos
Executive Director
Lindsay Barone, Ph.D.
Megan Capobianco
Elna Carrasco
Lauren Correri
Heather Cosel-Pieper
Alison Cucco
Jenna Dorey
Cristina Fernandez Marco, Ph.D.
Cornel Ghiban
Mary Lamont
Susan Lauter
Melissa Lee
Christine Marizzi, Ph.D.
Allison Mayle, Ph.D.
Amanda McBrien
Pauline McGlone
Colleen McKechnie
Valerie Meszaros
Eric Nash, Ph.D.
Lydia Paradiso
Michael Paul
Sharon Pepenella, Ph.D.
Jeffry Petracca
Colette Riccardi
Jason Williams
Chun-hua Yang
Xiaoqun Zhang

Library and Archives

Ludmila T. Pollock
Executive Director
Paula Abisognio
Thomas Adams
Tara Bonet-Black
Clare Clark
Rosemary DeRidder
Matthew Dunn, Ph.D.
Adrian Gomez, Ph.D.
Elizabeth Pessala, J.D.
Stephanie Satalino
Gail Sherman

Meetings & Courses Program

David J. Stewart, Ph.D.
Executive Director
Stephen Alexander
Alfred Baman
Margaret Bernhardt
Gerard Burlandi
Edward Campodonico
Cathleen Carr
Joseph Carrieri
Meredith Cassuto
Betty Cavallaro
Reginald Cayo
Katiria Ceballos
Maoyen Chi, Ph.D.
Wendy Crowley
Matthew Denninger
William Dickerson
Jane Donaldson, Ph.D.
Alicia Franco
Bradley Frey
Constance Hallaran
Patricia Klieger
Charla Lambert, Ph.D.
Rachel Lopez
Dennis McGuire
Andrew Mendelsohn
Maureen Morrow
Andrea Newell
Kenneth Orff
Valerie Pakaluk
Jonathan Parsons
Ira Russo
Stephen Silver
Mary Smith
Erica Stang
Vivian Vuong
James Whitaker
Barbara Zane

Office of the President

Terri I. Grodzicker, Ph.D.
Dean of Academic Affairs/ Editor, Genes & Development
Jan A. Witkowski, Ph.D.
WSBS Professor

Watson School of Biological Sciences

Alexander Gann, Ph.D.
Lita Annenberg Hazen *Dean*
Kimberly Creteur
Kimberley Geer
Alyson Kass-Eisler, Ph.D.
Monn Monn Myat, Ph.D.

ADMINISTRATIVE DEPARTMENTS

Finance and Accounting

Lari C. Russo
Chief Financial Officer
Sharon Bense
Elise Best
Karine Boyce
Damian Desiderio
Maryanne Detweiler
Karen Filasky
Mary Ellen Goldstein
Evelyn Imbriani
Mary Mastropaolo
Alison McDermott
Denise Miller
Kristine Murphy
Patricia Penner
Tara Ruocco
Nicole Solomon
Patricia Urena
Gisela Webb
Ann Williams

Central Administration

Maureen Berejka
Carol Gundersen
Carol Link
Elizabeth Panagot
Denise Roberts, Ph.D.

Culinary Services

James Hope
Director
Doris Abrego
Inolka Abreu
Ana Adames
Alan Aguirre
Jeannette Amato
Eliseo Amaya
Ingrid Amaya
Yamic Ambra
Julie Barkan
Francis Bowdren
Maryam Brinkwart
Elmer Canales
William Carmona
Dilicia Castro Najera
Francis Cayamcela
James Charvis
Marvin Chavez
Rene Chavez Contreras
Christina DeDora
Erica Douglas
Suzanne Esposito
Bartola Garcia
Luis Garcia
Andres Gonzalez
Silvia Gonzalez
Jennifer Gordon
Bartholomew Isaac
Kamal Johnson
Elizabeth Kirsch
Deborah Koslosky
Pamela Lancellotti
Jennifer LoSordo
Juan Magana
Jose Maradiaga Gonzalez
Thomas McIlvaine
Jason Moss
Frederik Munter
Joel Pena Corvera
Jose Reyes
Jesus Rivera
Susana Roman
Marlenis Romero
Carol Rupp
Robert Scherer
Karen Scola
Shannon Silva
Lisa Spero
Jose Suarez
Jose Velasquez
Hector Vigil Garcia
Maria Villatoro
Maria Vilorio
Tracy Woolfson
Briana Yim
Jennifer Ziegler
Lidia Zuleta

Development

Charles V. Prizzi
Vice President, Development, Government and Community Relations
Diane Fagiola
Stephanie Garrett
Bridgette Johnson
Karen Orzel
Jill Stone

Facilities

Arthur Brings
Vice President, Chief Facilities Officer
Robert Albinson
Leslie Allen
Peter Anderson
Lucas Arevalo
Jerry Armstrong
William Bishop
Ralph Byers
Caitlin Carroll
Dessie Carter
Daniel Chapman
Anthony Ciccarella
Christopher Clark
Robert Collins
Drew Comer
Timothy Cotone
Peter Dale
John Damico
Nicholas Dassau
Robert Dickerson
Glen DiMaria
Kevin Donohue
Sarah Drayzen
Wesley Dreusike
Paul Edwards
Elias Eifert
Rafael Estrella
Matthew Ferraro
Cody Garing
Vito Gassi
Robert Gensel
Erick Greene
Leo Greene
Stephen Gregorovic
Philip Grella, Jr.
Wilman Gutierrez
Wayne Hamilton
William Harding
Gerard Holler
Jason Holzman
Adam Hopper
Joseph Houser
Matthew Hoyer
Christopher Hubert
Louis Hunter
Michael Hutchinson
Randal Jones
Teri Jurkschat
Daniel Jusino
Jeffrey Klaverweiden
Robert Kopetic, Jr.
Paulo Krizanovski
Joseph Leonard
Walter Leute
Jian Lin
Charles Mandell
Gregory Marotta
Patricia McAdams
Tahid McCoy
Christopher McEvoy
Kevin McHugh
Thomas McNally
Daniel Miller
Susan Montalvo
Oscar Moreira
Guido Mosquera
Mary Muno
Michael Murphy
Kenneth Nastri
Herbert Neazer
Michael Nolasco
Wayne Pav
Christopher Pearson II
Jose Pena Corvera
Saul Pena Corvera
John Pisciotta
Manuel Ramones
Wilson Ramones
Steven Randall
Elisha Rendal Sestoso
Jaqueline Reynolds
Alfredo Rivera, Jr.
Ronald Romani
Frank Russo
Richard Rutherford
Jeffrey Sanchez
Eric Sappell
Steven Sardone
John F. Schindler
John L. Schindler
Claudia Schmid
Lauren Schmidt
Peter Schuman
Bethany Scola
Cesar Sisalima
Daniel Stahl
Raymond Stewart
John Sundman
Steven Tang
Hock Yew Teo
Leigh Tyrrel
James Van Wie
Fredy Vasquez
Marden Vasquez
Trinidad Velasquez
Wilfredo Velasquez
Benjamin Veneable III
Kyle Welton
Leslie Wenzel
Tracey Anne White
Graham Wildt
Daniel Zimmer

Human Resources

Katherine G. Raftery
Vice President
Andres Alarcon
Wendy Alexander
Lynn Carmen
Camille Cava-Belluscio
Patricia Crawford
Jessica Heckel
Lesley Inglis
Laura Magri
Laura Moran
Elsy Padro
Victoria Panebianco
Victoria Roberts
Patricia Schoeffler
Samantha Walbridge
Lauren Warmbrand

Information Technology

Hans Erik Aronson
Vice President, Chief Information Officer
Alex Barton
David Jonathan Castillo Torres
Daniel Drayzen
Frank Falco
Todd Heywood, Ph.D.
Thomas Karow
Leila Kelly
Sean Kelly
Caryn Koza
Michael Malave
Louis Malfi
Paul Marcley
Bhargav Patel
Barbara Peters
Vincent Reale
Salvatore Serafino
Michael Skuthan
Steven Zielinski

Procurement

Michael Marchesiello
Vice President

Deborah Aufiero
Kimberly Bronson
Susan De Angelo
Jeffrey DuPree
Elizabeth Janow
Anne Knoch
Christopher Oravitz
Barbara Purcell
Jorge Ramirez
Krystyna Rzonca
Felicia Wolf

Public Affairs

Dagnia Zeidlickis
Vice President, Communications
Margot Bennett
Christopher Fergo
Jessica Giordano
Philip Renna
Sara Roncero-Menendez
Susan Runkowski
Brian Stallard

Research Operations

Diane Esposito, Ph.D.
Director of Research Compliance /Research Investigator
Robert Gerdes
Jaclyn Jansen, Ph.D.
Laura Lynn
Joan O'Connor
Julie Sutherland

Sponsored Programs

Walter L. Goldschmidts, Ph.D.
Vice President, Executive Director of Sponsored Programs
Carol Amella
Irina Armstrong
Carol DuPree
David Garner
Donna Gavin
Jill Hemish, Ph.D.
Joanne Janelli
Philip King
Ruta Ledins
Josephine Marrali, J.D.
Cynthia McCormack
Dave Neale
Catherine Perdikoylis
Jaclyn Tenaglia

Technology Transfer

Andrew Whiteley
Vice President, Business Development/Technology Transfer
Vladimir Drozdoff, J.D., Ph.D.
Michael Gurtowski
Elizabeth Hand
Cynthia Miller
Radhakrishnan Narayanan, Ph.D.

Legal

Debra Arenare, J.D.
Vice President, General Counsel
Alexander Sosnovski

RESEARCH STAFF DEPARTURES DURING 2018

Professor

Marja Timmermans, Ph.D.

Assistant Professors

Gholson Lyon, M.D., Ph.D.
Hongwu Zheng, Ph.D.

Research Investigators

Gayatri Arun, Ph.D.
Mona Spector, Ph.D.
Herve Tiriac, Ph.D.

Adjunct Professors

Grigori Enikolopov, Ph.D.
James Hicks, Ph.D.

Adjunct Associate Professor

Glenn Turner, Ph.D.

Visiting Scientists/ Collaborative Scientists

Sandra Correia, Ph.D.
Ahmed Elewa, Ph.D.
Ruth Goldschmidt, Ph.D.
Barbara Grunwald, Ph.D.
Susan Heavey, Ph.D.
Giyong Jang, Ph.D.
Whitney Johnson, Ph.D.
Rotem Karni, Ph.D.
Risa Kawaguchi, Ph.D.
Jieun Lee, Ph.D.
Young Koung Lee, Ph.D.***
Dawid Nowak, Ph.D.***
Xiaoli Qi
Nikolina Radulovich, Ph.D.
Daniel Rodriguez-Leal, Ph.D.***
Alan Schoenfeld, Ph.D.
Payton Stevens, Ph.D.
Laura Tamblyn, Ph.D.

Senior Fellow

Dannielle Engle, Ph.D.

Visiting Clinical

Adam Kressel, M.D.

Postdoctoral Fellows

Sandra Ahrens, Ph.D.
Federico Carnevale, Ph.D.
Cristina Chen, Ph.D.
Filipe de Sousa Borges*, Ph.D.
Luana Guerriero, Ph.D.
Change-Il Hwang, Ph.D.
Ashley Juavinett, Ph.D.
Matthew Kaufman, Ph.D.
Virginie Kergourlay, Ph.D.
Young Koung Lee, Ph.D.***
Zachary Lemmon, Ph.D.
Jiangteng Lu, Ph.D.
Luis Meija, Jr., Ph.D.
Dawid Nowak, Ph.D.***
Jean-Sebastien Parent, Ph.D.
Michael Regan, Ph.D.
Jie Ren, Ph.D.
Daniel Rodriguez-Leal, Ph.D.***
Jae Seok Roe, Ph.D.
Elena Rozhkova, Ph.D.
Juergen Scharner, Ph.D.
Mario Shields, Ph.D.
Nami Tajima, Ph.D.
John Wilson, Ph.D.

Computational Science

Arjun Dass
Aashna Jhaveri
Demitri Muna, Ph.D.
Toufiq Parag, Ph.D.
James Thomason III

Senior Research Associate

Saikat Nandi, Ph.D.

*Employee of the Howard Hughes Medical Institute
***These individuals are both Postdoctoral Fellows and Visiting Scientists/Collaborative Scientists.