Genetic Complexity and Neuropsychiatric Disorders

Gholson J. Lyon, M.D. Ph.D.

NURTURING GENETICS: REFLECTIONS ON A CENTURY OF SCIENTIFIC AND SOCIAL CHANGE

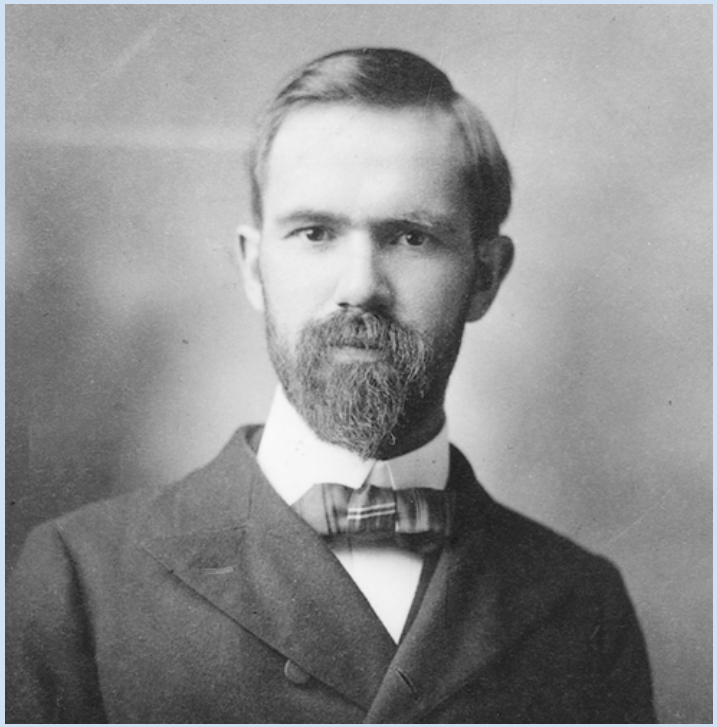
An international and interdisciplinary symposium

University of Leeds, 29 June-2 July 2014

Cold Spring Harbor Laboratory

Stanley Institute for Cognitive Genomics


CSH
Cold Spring Harbor Laboratory
Stanley Institute for Cognitive Genomics

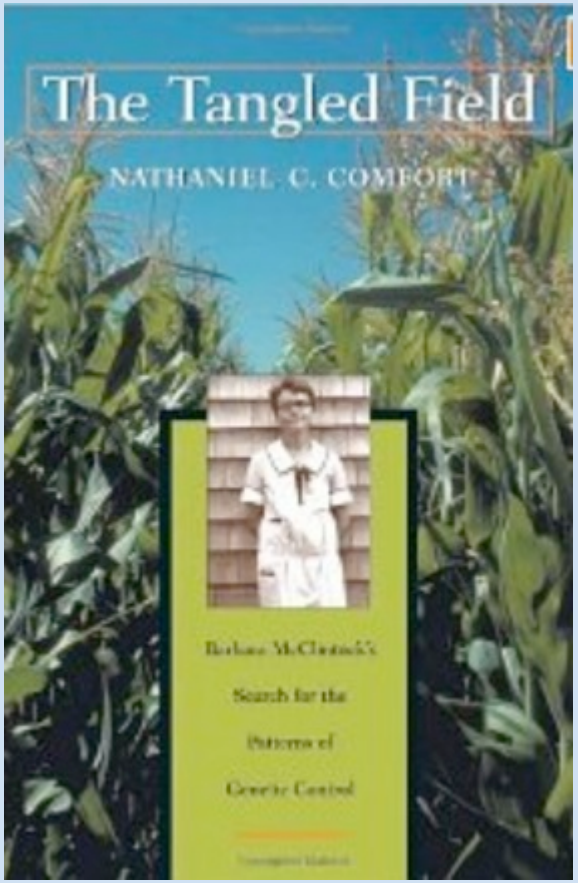


Director, 1898-1934
Charles Davenport

Station for Experimental Evolution – 1904
(Carnegie Institution of Washington)


CSH
Cold Spring Harbor Laboratory
Stanley Institute for Cognitive Genomics


CSH
Cold Spring Harbor Laboratory
Stanley Institute for Cognitive Genomics


The Tangled Field
NATHANIEL C. COMFORT

From Base Pair to Body Plan:
Celebrating 60 years of DNA

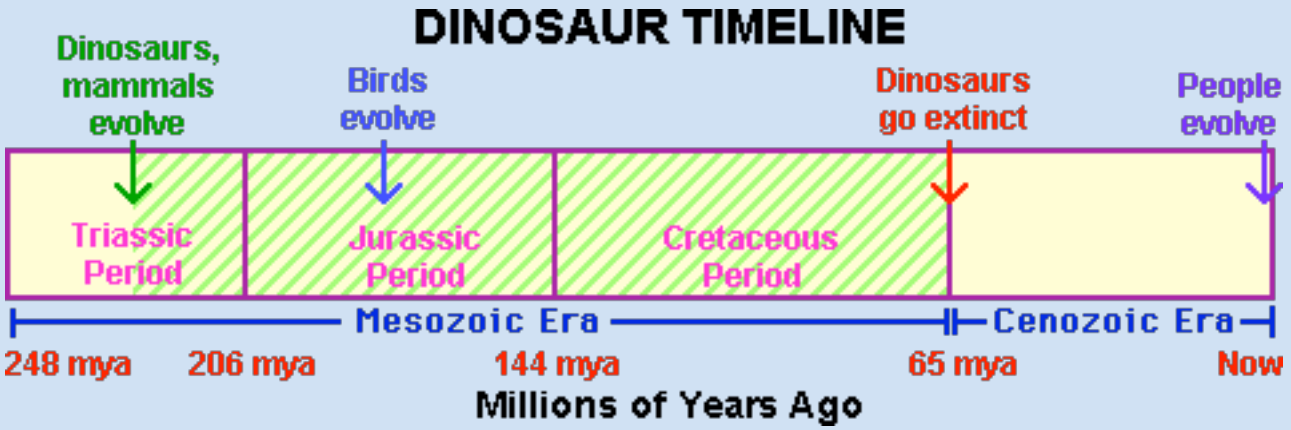
Organizers:

Alex Gann, Cold Spring Harbor Laboratory

Robert Martienssen, Cold Spring Harbor Laboratory/HHMI


The Earth
is 4.5 Billion
Years Old


Gulf of Mexico
Campeche Escarpment
Outer rim
of buried
impact crater
Yucatán
Peninsula


DINOSAUR TIMELINE
Dinosaurs, mammals evolve
Birds evolve
Dinosaurs go extinct
People evolve
Triassic Period
Jurassic Period
Cretaceous Period
Mesozoic Era
Cenozoic Era
248 mya
206 mya
144 mya
65 mya
Now
Millions of Years Ago

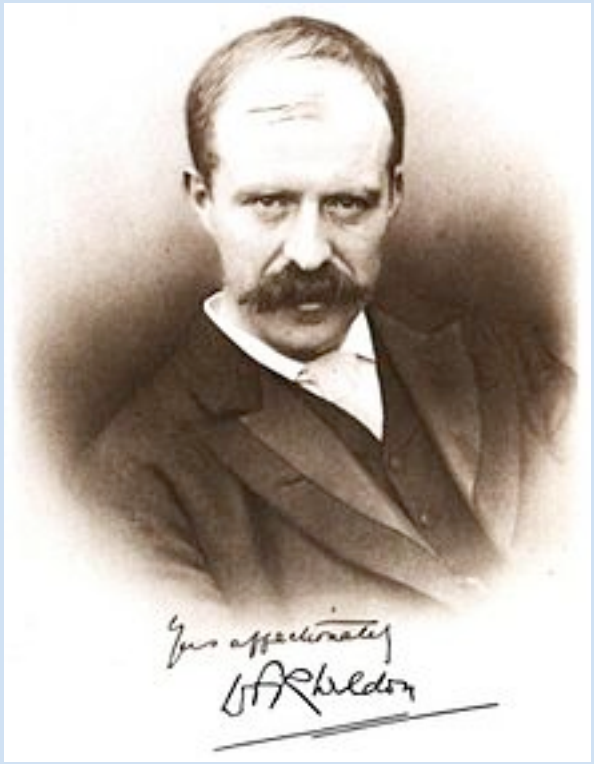
Trend in Beliefs About Human Evolution, by Party Affiliation

% of U.S. adults in each group

2013	Humans have evolved over time NET	Due to natural processes	Supreme being guided	Unclear/ don't know	Humans have existed in present form	Don't know	N
All adults	60	32	24	4	33	7	1983
Republican	43	21	20	2	48	9	455
Democrat	67	37	27	4	27	6	705
Independent	65	35	25	5	28	7	674
Other (vol.)/none (vol.)/don't know	50	28	18	5	38	12	149
2009							
All adults	61	32	22	7	31	8	2001
Republican	54	23	26	5	39	7	504
Democrat	64	36	22	6	30	5	747
Independent	67	38	20	9	27	6	579
Other (vol.)/none (vol.)/don't know	46	19	12	14	30	24	171

Source: Pew Research Center survey March 21-April 8, 2013 and April 28-May 12, 2009. Q54/Q55. Figures may not sum to NET due to rounding and net figures may not sum to 100% due to rounding.

PEW RESEARCH CENTER



Vs.

Walter Frank Raphael Weldon

William Bateson

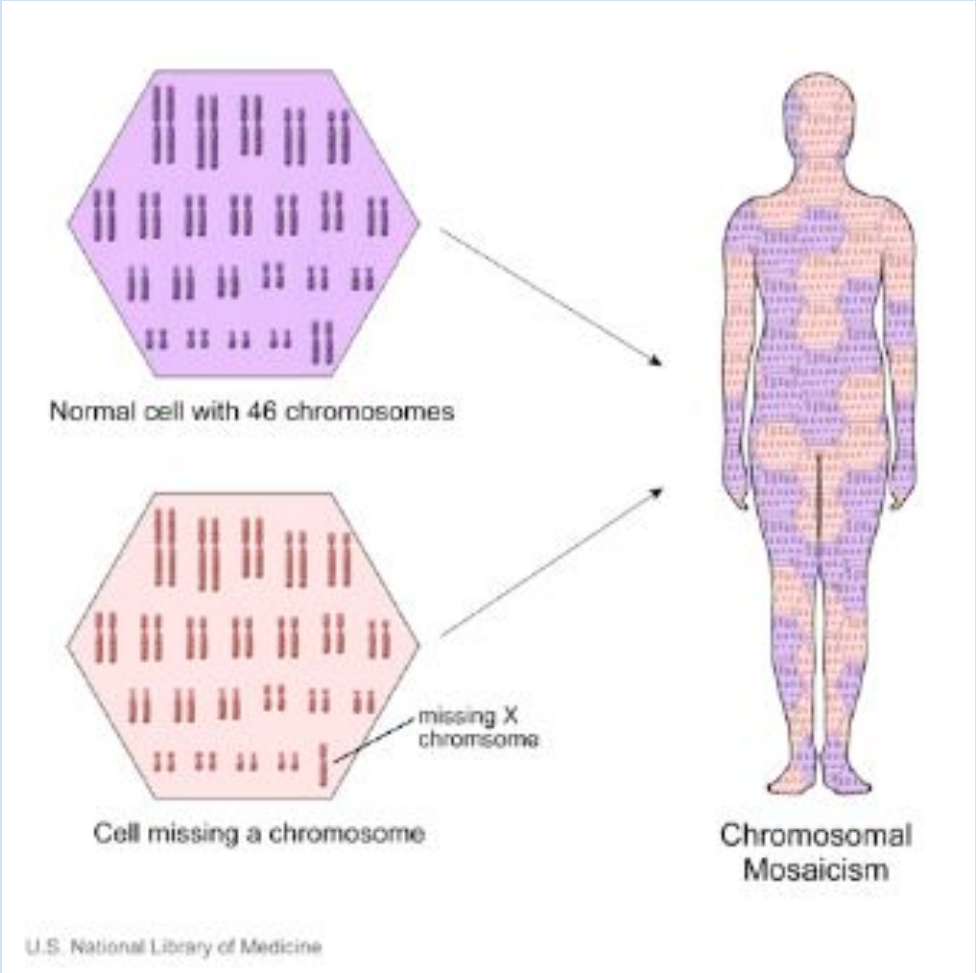
Forthcoming by Greg Radick. Scholarly edition of W. F. R. Weldon's Theory of Inheritance (1904-1905), coedited with Anne Jamieson.

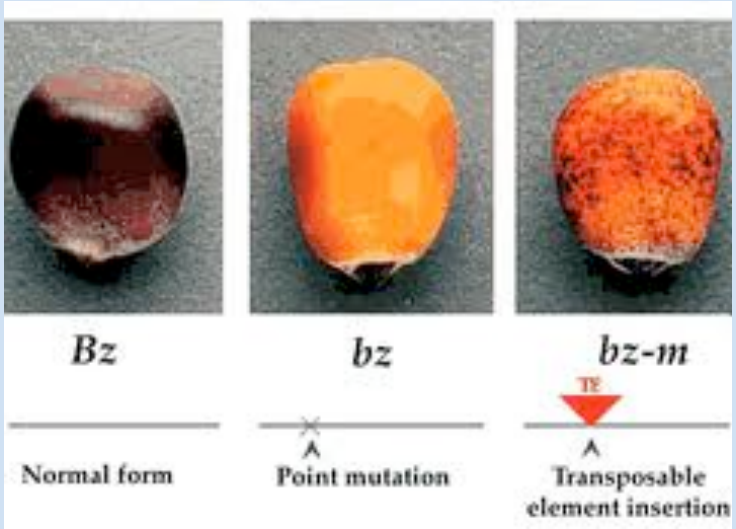
Plate I.

Weldon, W. F. R. 1902. Mendel's laws of alternative inheritance in peas. *Biometrika*, 1:228-254.

Complexity

- There are ~25-100 TRILLION cells in each human body, with ~6 billion nucleotides per cell.
- There is extensive modification of DNA, RNA and proteins both spatially and temporally.
- There are higher level mechanisms of somatic mosaicism, heterosis, and likely ancestral inheritance.





Source: http://www.thenakedscientists.com/HTML/features/article/jamilcolumn1.htm/

OPEN ACCESS Freely available online

PLoS one

Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types

Julia Salzman[1☯], Charles Gawad[1,3☯], Peter Lincoln Wang[1], Norman Lacayo[3], Patrick O. Brown[1,2*]

1 Department of Biochemistry, Stanford University School of Medicine, Stanford, California, United States of America, **2** Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California, United States of America, **3** Department of Pediatric Hematology/Oncology, Stanford University School of Medicine, Stanford, California, United States of America

Abstract

Most human pre-mRNAs are spliced into linear molecules that retain the exon order defined by the genomic sequence. By deep sequencing of RNA from a variety of normal and malignant human cells, we found RNA transcripts from many human genes in which the exons were arranged in a non-canonical order. Statistical estimates and biochemical assays provided strong evidence that a substantial fraction of the spliced transcripts from hundreds of genes are circular RNAs. Our results suggest that a non-canonical mode of RNA splicing, resulting in a circular RNA isoform, is a general feature of the gene expression program in human cells.

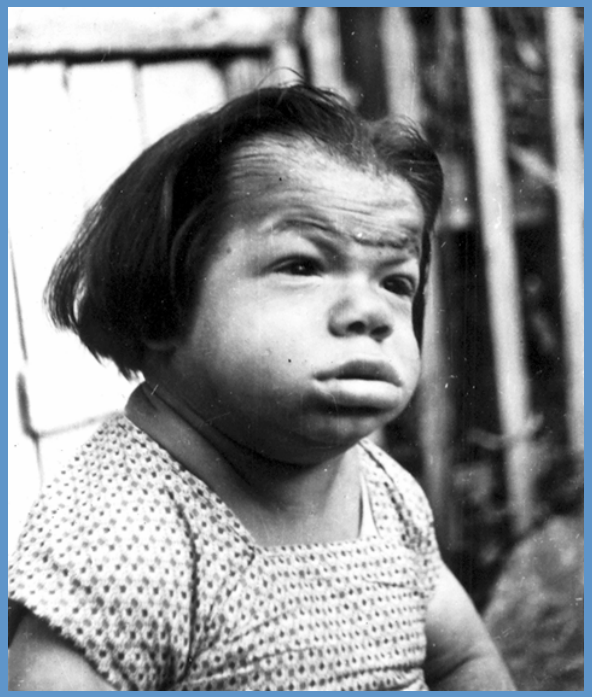
Citation: Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO (2012) Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. PLoS ONE 7(2): e30733. doi:10.1371/journal.pone.0030733

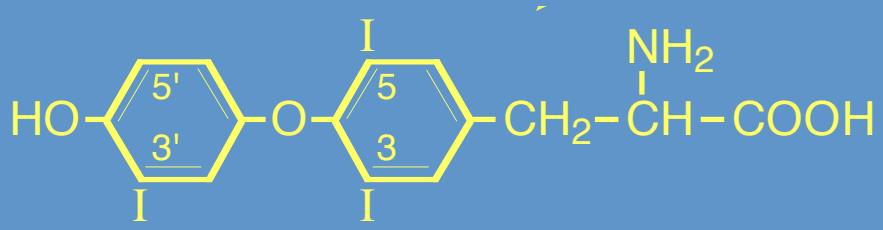
Editor: Thomas Preiss, The John Curtin School of Medical Research, Australia

Received November 7, 2011; **Accepted** December 28, 2011; **Published** February 1, 2012

The Story began for me around 1993, at ~age 19….

when I joined the lab of Don St.Germain to study the role of thyroid hormone in cretinism, which is caused by lack of iodine during maternal pregnancy, so this is an environmentally triggered disease.



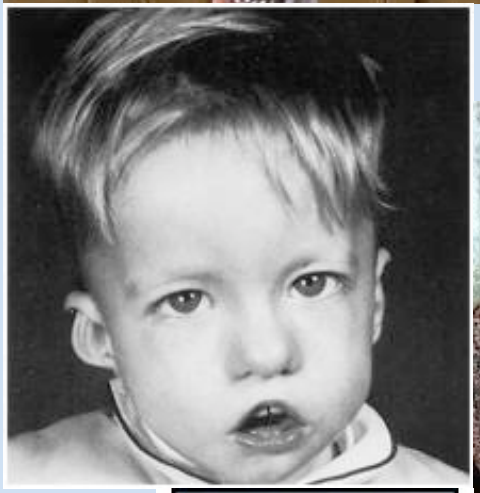


Thyroid Hormone

Down Syndrome

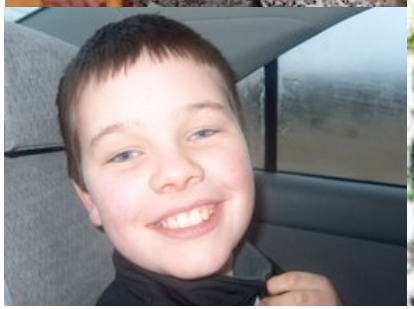
Velocardiofacial (22q11.2) Syndrome









16p11.2 deletion, not in mother or father, only in child.

5 years old, but developmental age of 2 year old.
Speaks a few words, almost unintelligible.
Very hyperactive.
Can be withdrawn and has at times been diagnosed with “autism”.

Current Diagnoses under Evaluation (DSM IV-TR)		
AXIS I	299.00	Autism Disorder
	314.01	Attention-Deficit-Hyperactivity Disorder, Combined Type
AXIS II	V71.09	No Diagnosis
AXIS III	16p11.2	Microdeletion
AXIS IV		Psychosocial Stressors: Moderate(Adaptive/Behavioral and Educational/Learning Problems)
AXIS V		Current GAF: 60

Discovering a new syndrome and its genetic basis.

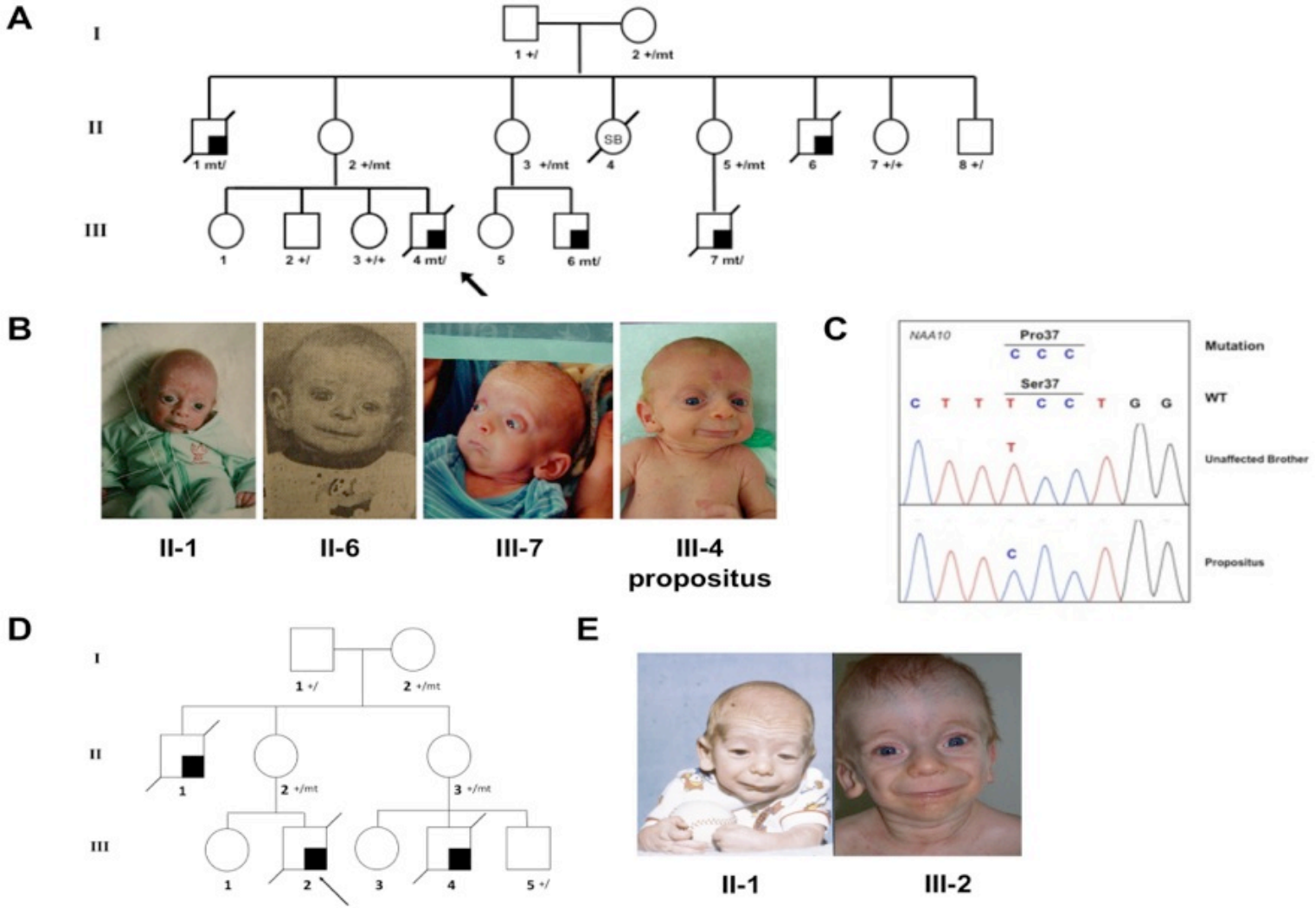
ARTICLE

Using VAAST to Identify an X-Linked Disorder Resulting in Lethality in Male Infants Due to N-Terminal Acetyltransferase Deficiency

Alan F. Rope,[1] Kai Wang,[2,19] Rune Evjenth,[3] Jinchuan Xing,[4] Jennifer J. Johnston,[5] Jeffrey J. Swensen,[6,7] W. Evan Johnson,[8] Barry Moore,[4] Chad D. Huff,[4] Lynne M. Bird,[9] John C. Carey,[1] John M. Opitz,[1,4,6,10,11] Cathy A. Stevens,[12] Tao Jiang,[13,14] Christa Schank,[8] Heidi Deborah Fain,[15] Reid Robison,[15] Brian Dalley,[16] Steven Chin,[6] Sarah T. South,[1,7] Theodore J. Pysher,[6] Lynn B. Jorde,[4] Hakon Hakonarson,[2] Johan R. Lillehaug,[3] Leslie G. Biesecker,[5] Mark Yandell,[4] Thomas Arnesen,[3,17] and Gholson J. Lyon[15,18,20,*]

The American Journal of Human Genetics *89*, 1–16, July 15, 2011

Ogden Syndrome



We found the SAME mutation in two unrelated families, with a very similar phenotype in both families, helping prove that this genotype contributes to the phenotype observed.

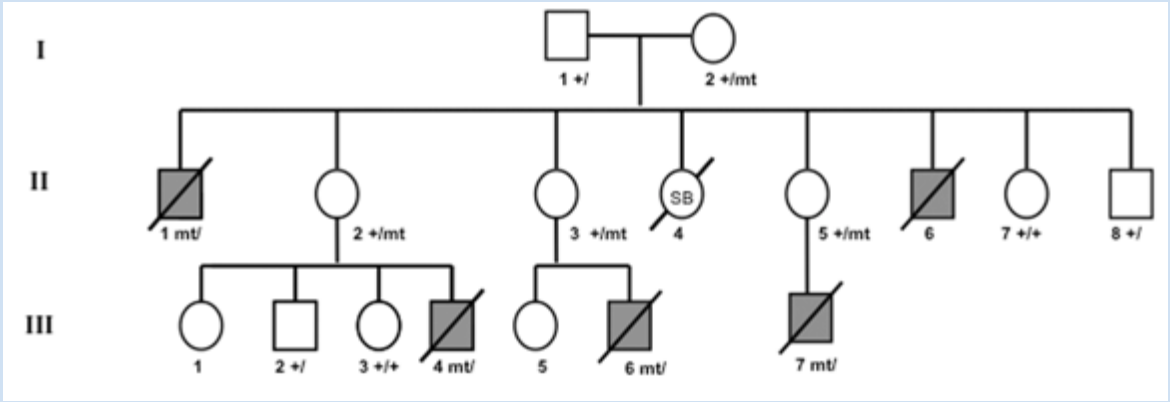
This is the first boy in the late 1970's.

First boy. Called "a little old man" by the family. Died around ~1 year of age, from cardiac arrhythmias.

This is the "Proband" photograph presented at Case Conference.

prominence of eyes, down-sloping palpebral fissures, thickened eyelids, large ears, beaking of nose, flared nares, hypoplastic nasal alae, short columella, protruding upper lip, micro-retrognathia

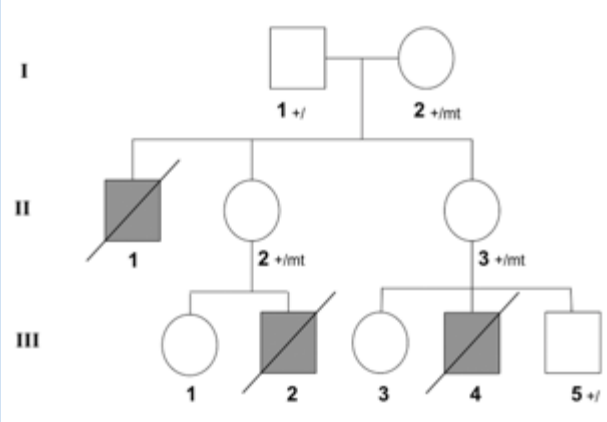
A



B

II-1 II-6 III-4 III-6 III-7

C

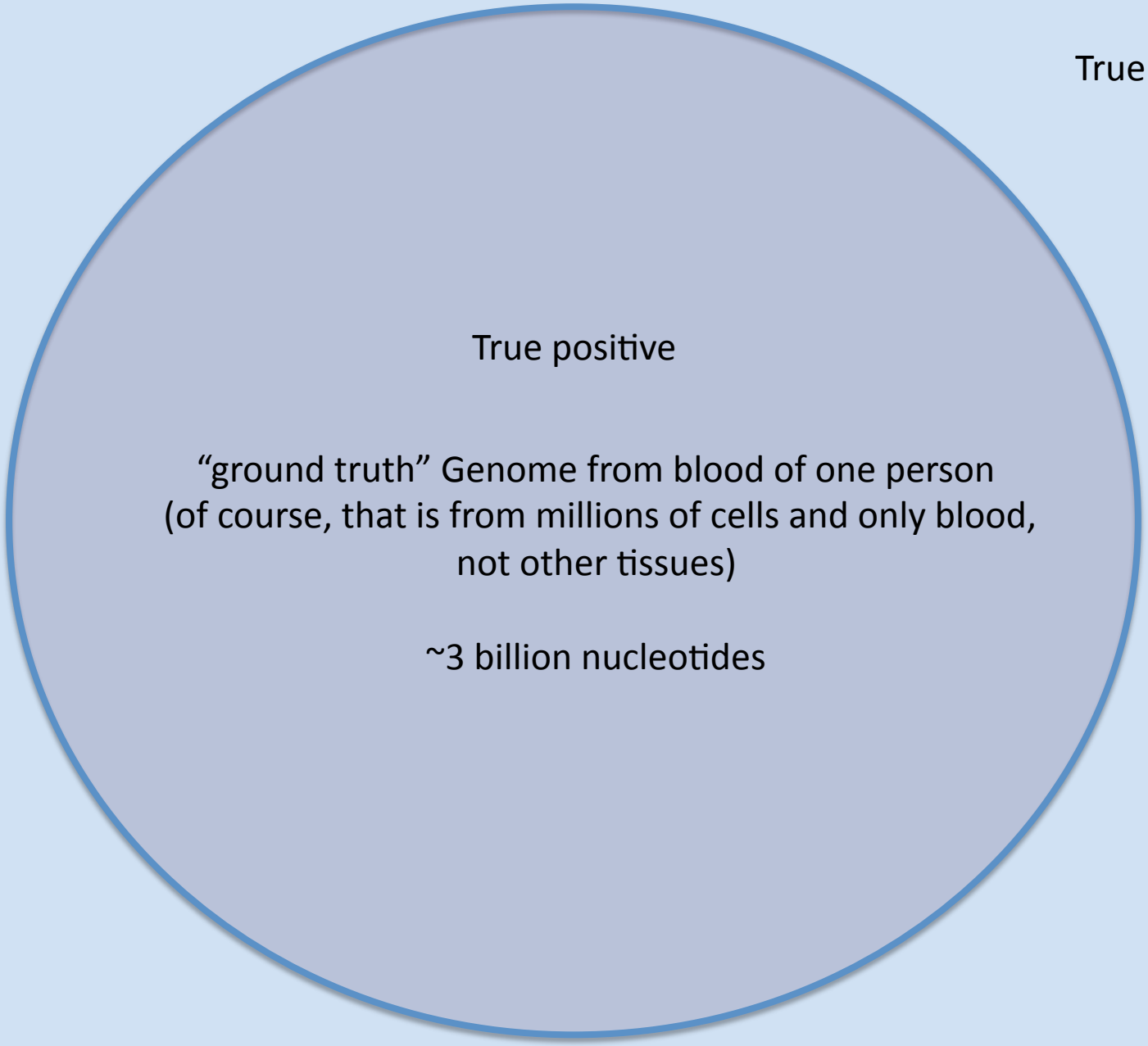


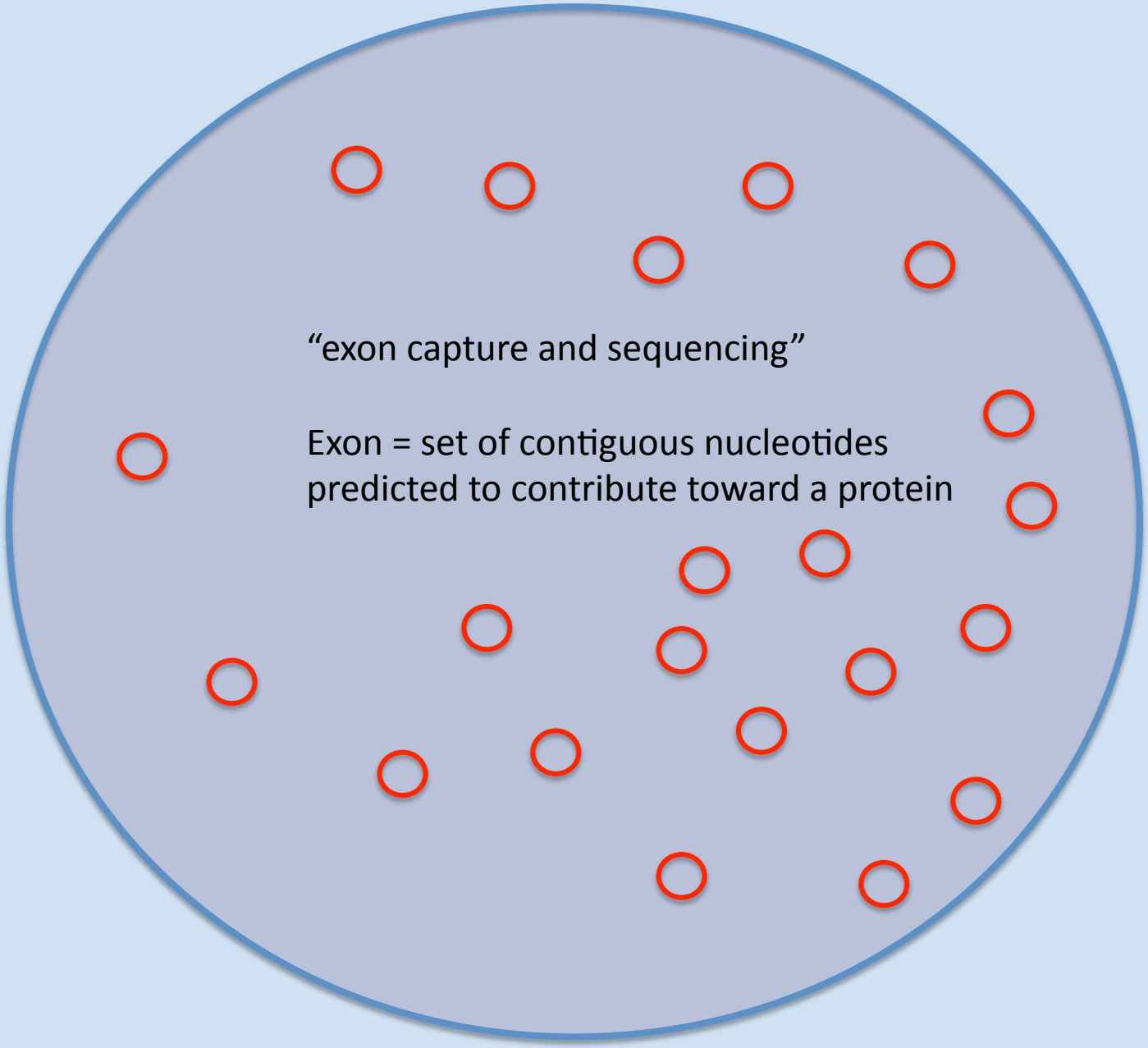
D

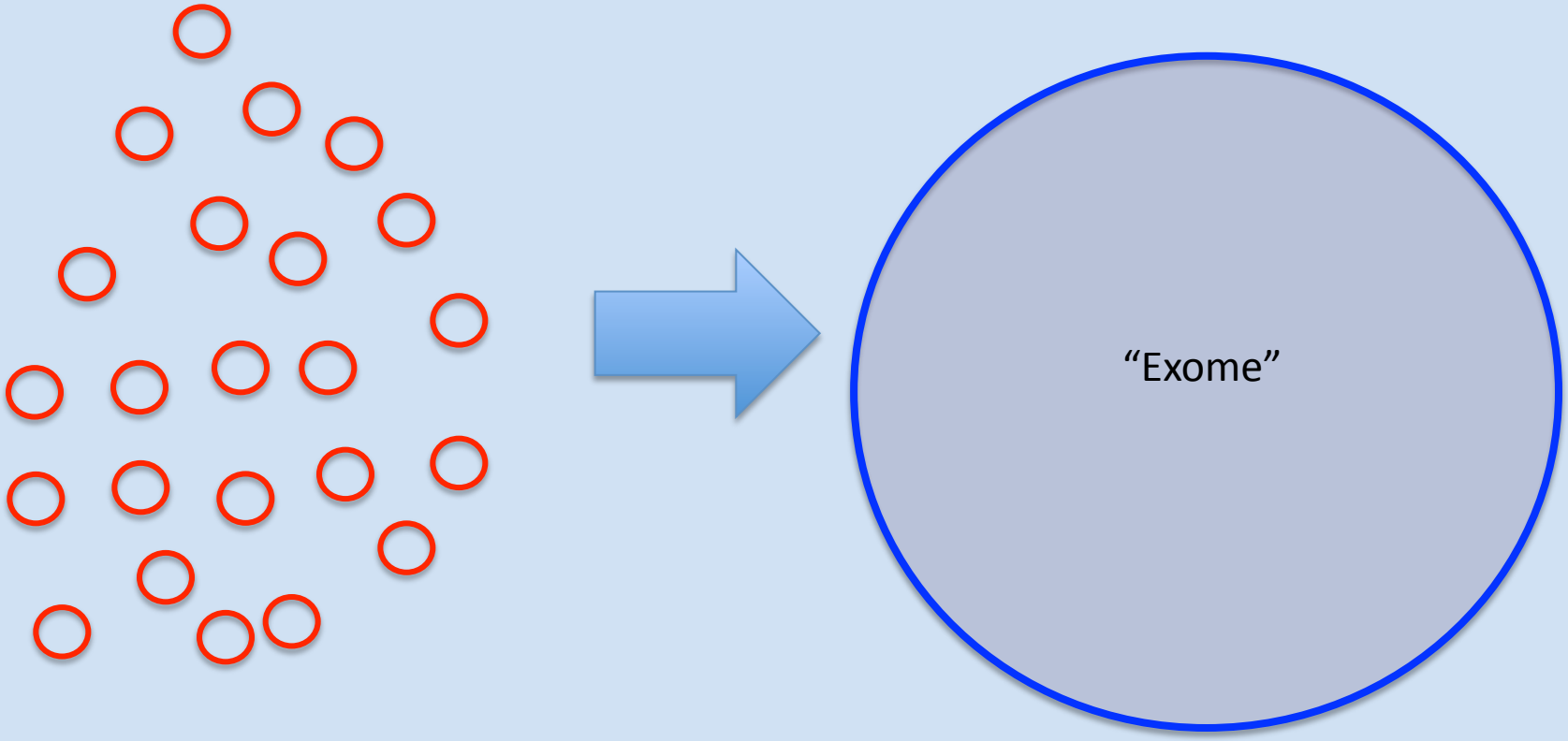
II-1 III-2

These are the Major Features of the Syndrome.

Table 1. Features of the syndrome	
Growth	post-natal growth failure
Development	global, severe delays
Facial	prominence of eyes, down-sloping palpebral fissures, thickened lids large ears beaking of nose, flared nares, hypoplastic alae, short columella protruding upper lip micro-retrognathia
Skeletal	delayed closure of fontanels broad great toes
Integument	redundancy / laxity of skin minimal subcutaneous fat cutaneous capillary malformations
Cardiac	structural anomalies (ventricular septal defect, atrial level defect, pulmonary artery stenoses) arrhythmias (Torsade de points, PVCs, PACs, SVtach, Vtach) death usually associated with cardiogenic shock preceded by arrythmia.
Genital	inguinal hernia hypo- or cryptorchidism
Neurologic	hypotonia progressing to hypertonia cerebral atrophy neurogenic scoliosis
Shaded regions include features of the syndrome demonstrating variability. Though variable findings of the cardiac, genital and neurologic systems were observed, all affected individuals manifested some pathologic finding of each.	


True negative
True positive
“ground truth” Genome from blood of one person
(of course, that is from millions of cells and only blood,
not other tissues)
~3 billion nucleotides


“exon capture and sequencing”
Exon = set of contiguous nucleotides
predicted to contribute toward a protein


“Exome”

- **We performed X-chromosome exon capture with Agilent, followed by Next Gen Sequencing with Illumina.**

- **We analyzed the data with ANNOVAR and VAAST (Variant Annotation, Analysis and Search Tool). New computational tools for identifying disease-causing mutations by individual genome sequencing.**

Yandell, M. *et al.* 2011. "A probabilistic disease-gene finder for personal genomes." *Genome Res*. 21 (2011). doi:10.1101/gr.123158.111.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38, e164.

VAAST integrates AAS & Variant frequencies in a single probabilistic framework

- non-coding variants scored using allele frequency differences

- n_i : frequency of variant type among all variants observed in Background and Target genomes

- a_i: frequency of variant type among disease causing mutations in OMIM

- This approach means that *every* variant can be scored, non-synonymous, synonymous, coding, and non-coding. Phylogenetic conservation not required.

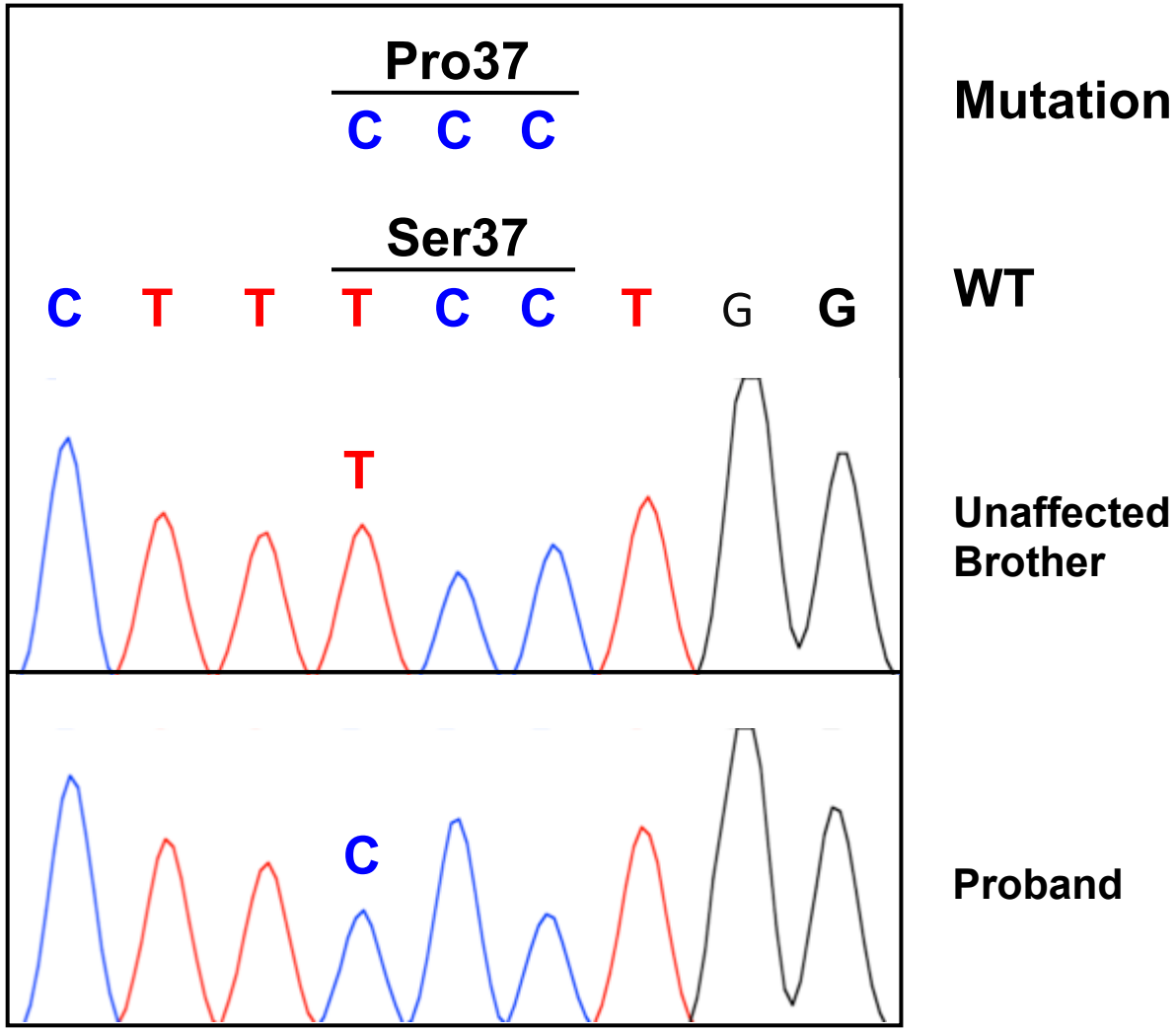
Omicia™ Unlocking Individualized Medicine

THE UNIVERSITY OF UTAH

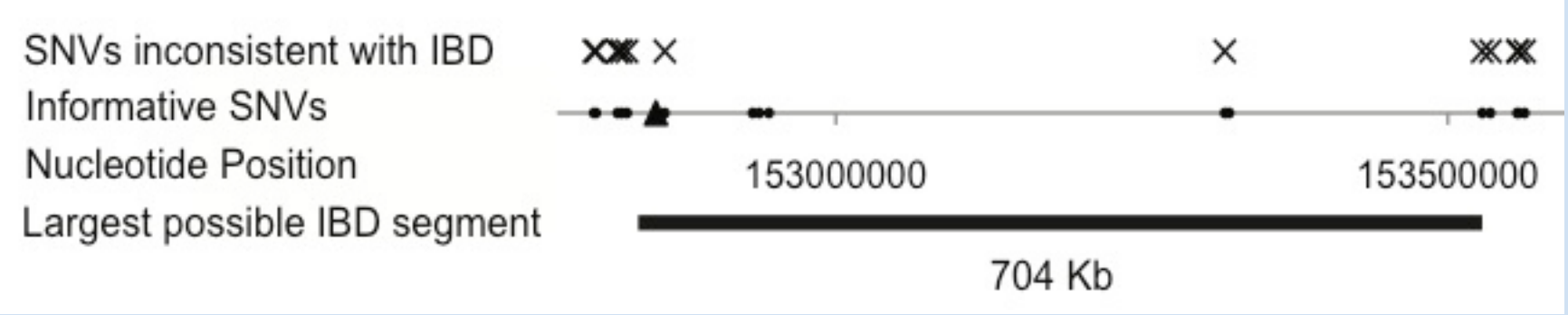
Analysis with VAAST readily identified a few likely candidates.

Table 3. Summary of the filtering procedure and candidate genes using VAAST			
SNV calling pipeline	GATK	Samtools	GNUMAP
III-4 (total SNVs)	1546	1499	2168
III-4 (nsSNVs)	146	114	155
VAAST candidate genes (NAA10 ranking)	4 (3)	3 (2)	5 (2)
Present in III-4 and mother II-2 (nsSNVs)	122	107	116
VAAST candidate genes (NAA10 ranking)	3 (2)	2 (1)	2 (2)
Present in III-4, mother II-2, and grandmother I-2 (nsSNVs)	115	95	104
VAAST candidate genes (NAA10 ranking)	2 (1)	2 (1)	1 (1)
Present in III-4, II-2, and I-2, absent in brother III-2 and uncle II-8 (nsSNVs)	8	6	8
VAAST candidate genes (NAA10 ranking)	1 (1)	1 (1)	2 (1)

This is the mutation we found… one nucleotide change out of 6 billion nucleotides in a diploid genome…



Identity by Descent Analysis shows that the mutation must have arisen independently in two different families.



Courtesy of Chad Huff and Lynn Jorde

Published online 23 June 2011 | Nature | doi:10.1038/news.2011.382

News

Software pinpoints cause of mystery genetic disorder

Genome analysis tools speedily track down previously unknown mutation.

Brendan Maher

Halena Black's first son, Kenny Rae, was born in November 1979. He struggled to put on weight, and had thin, wrinkled skin, big eyes and a broad mouth. In October the following year, he died from heart problems.

After Kenny Rae died, Black — a Mormon living in Ogden, Utah — had three healthy daughters before giving birth to another son in 1987. He had the same problem, and a similarly short lifespan. Her third son is healthy.

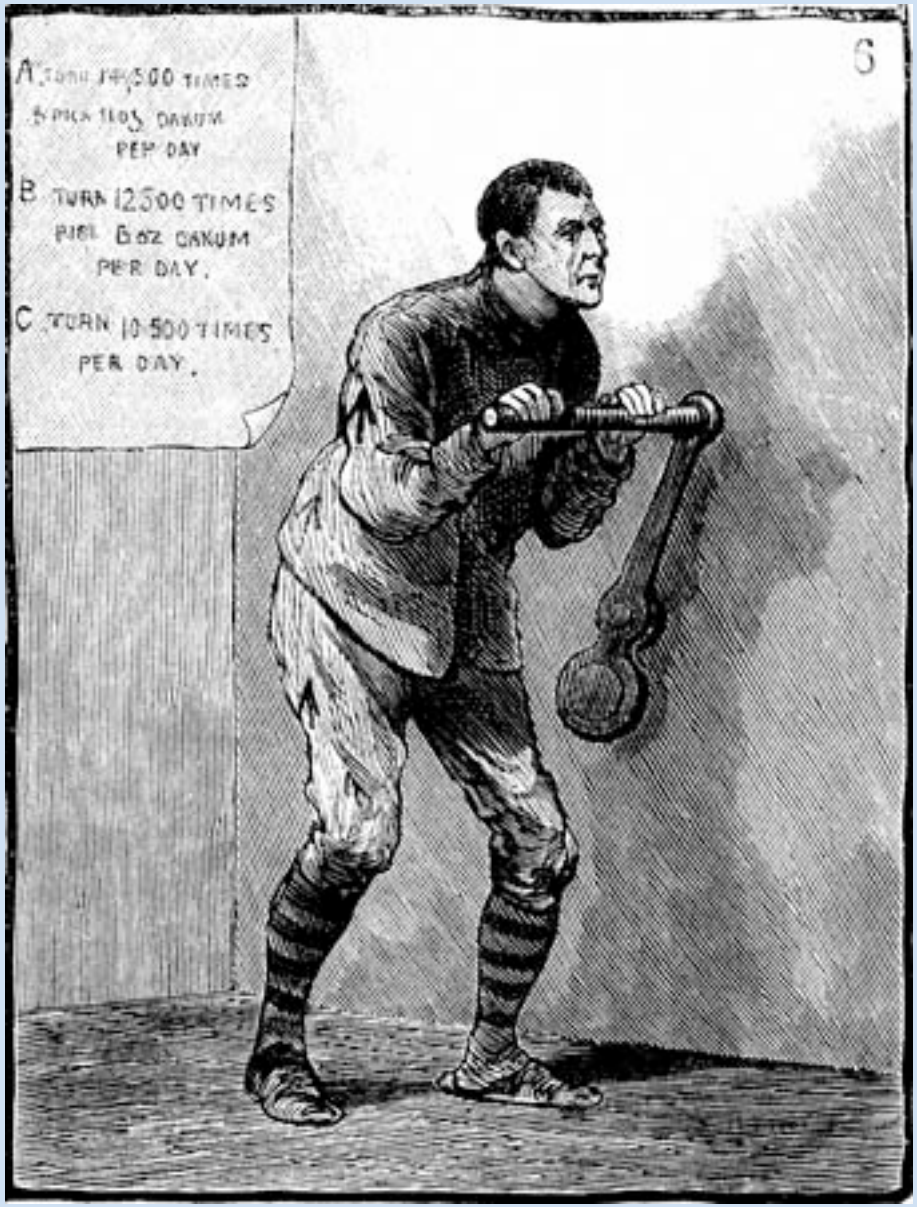
Four boys from the same family born wirh 'Ogden Syndrome'. Sufferers rarely survive for more than a year.

A. Rope et al./Am. J. Hum. Genet.

Black says she didn't dwell much on why her sons died until one of her daughters gave birth to a boy who looked just like Kenny Rae. "We didn't think that it passed on to the next generation. We didn't think that this would be a problem for them," says Black. All three daughters have since given birth to what the family calls 'little old men', one of whom died just last Sunday.

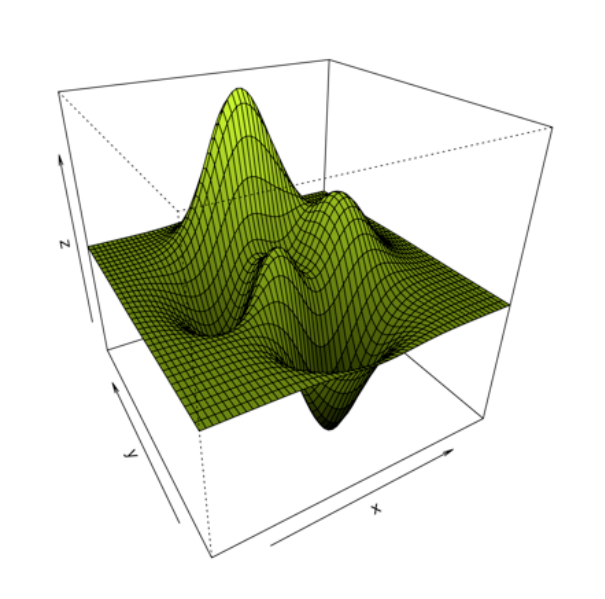
"This exemplifies an exceptionally rare disease, but the same type of strategy is now going to be applied to more common diseases to get the root cause," says Eric Topol, a medical geneticist at the Scripps Research Institute in La Jolla, California.

"This is one of the most exciting things in medicine," says Topol. "We're going to take the term 'idiopathic' which, basically means 'we don't' know,' and eliminate it."

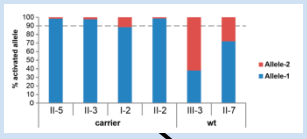


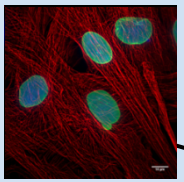
A prisoner at Dartmoor is forced to turn a crank handle repeatedly as a form of punishment, as depicted in an illustration dated 1884.

Once one finds a validated high-effect size mutation, functional analysis is appropriate.



patient cells

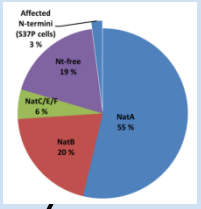




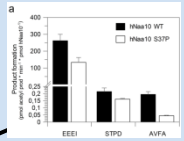
X-chromosome skewing in carriers

normal nuclear envelope

molecular level

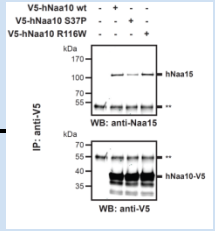


affected *in vivo* substrates



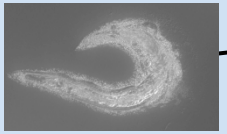
reduced enzymatic activity

Naa10 S37P



disrupted NatA complex

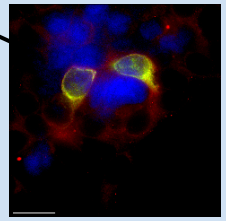
C. elegans



initial experiments

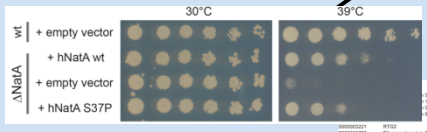
HEK293 cells

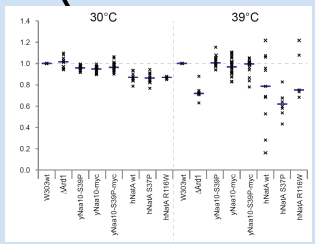
normal localization



growth defect at 39°C
ribosome defect?
no difference cyclo/doxo?

S39 mutation?
growth defect of "humanized" strains





S. cerevisiae

NAA10 Alignments

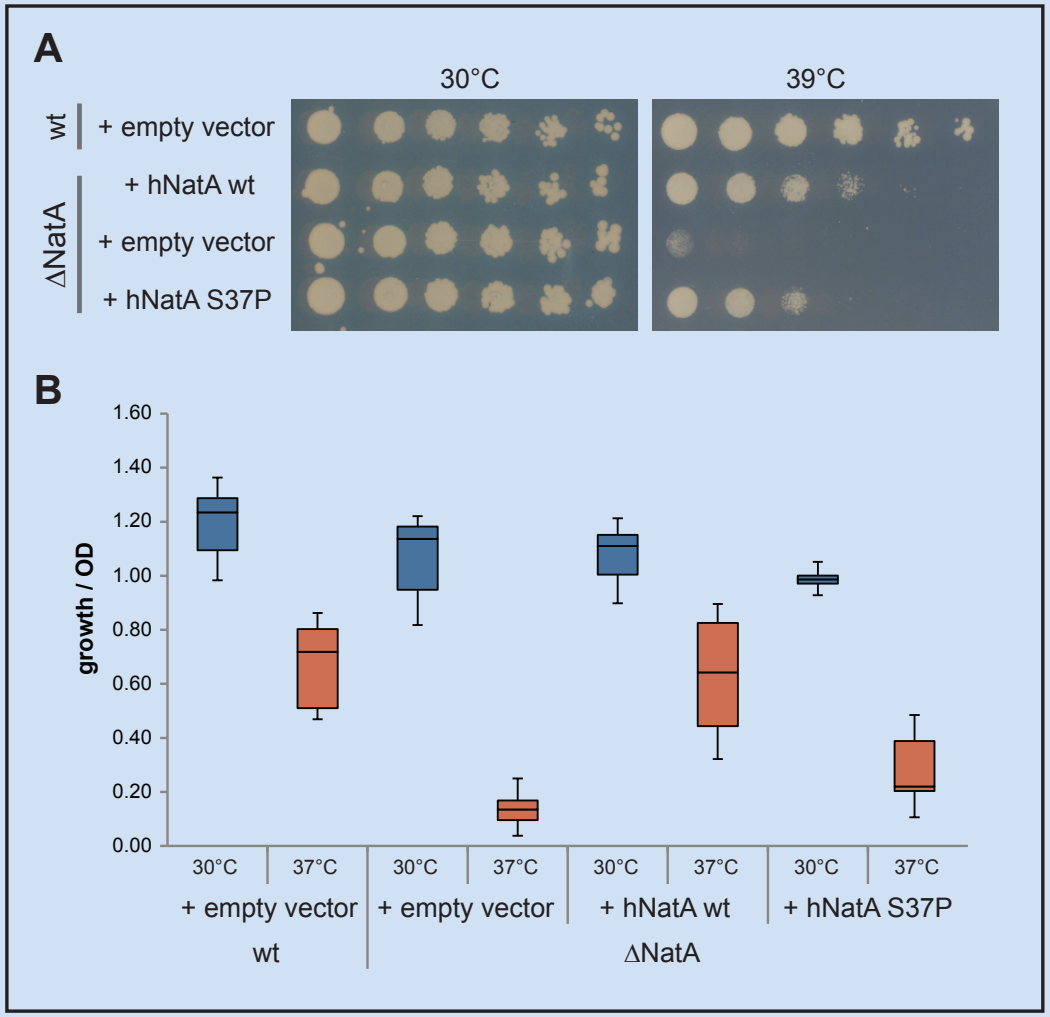
Gene	Length		Gene	Length	Identity Score %
human	235	vs	c_elegans	181	68
human	235	vs	s_pombe	177	57
human	235	vs	s_cerevisiae	238	38
c_elegans	181	vs	s_pombe	177	57
c_elegans	181	vs	s_cerevisiae	238	48
s_pombe	177	vs	s_cerevisiae	238	52

Sequences were aligned using Clustal multiple sequence alignment tool. The Scores Table shows the pairwise scores calculated for every pair of sequences that is to be aligned. Pairwise scores are simply the number of identities between the two sequences, divided by the length of the alignment, and represented as a percentage. This alignment is only a precursor to the full multiple alignment and might not be preserved.

```
CLUSTAL 2.1 multiple sequence alignment

h_sapiens_V1      --MNIRNARPEDLMNMQHCNLLCLPENYQMKYYFYHGLSWPQLSYIAE------------ 46
c_elegans         --MNIRCARVDDLMSMQNANLMCLPENYQMKYYFYHALSWPQLSYIAE------------ 46
s_pombe           --MDIRPARISDLTGMQNCNLHNLPENYQLKYYLYHAISWPMLSYVAT------------ 46
s_cerevisiae      MPINIRRATINDIICMQNANLHNLPENYMMKYYMYHILSWPEASFVATTTTLDCEDSDEQ 60
                    ::** *  .*:  **:.**  ***** :***:** :***  *::*

h_sapiens_V1      DENG--------------------------KIVGYVLAKMEEDPDD---VPHGHITSLAV 77
c_elegans         DHKG--------------------------NVVGYVLAKMEEDPGE---EPHGHITSLAV 77
s_pombe           DPKG--------------------------RVVGYVLAKMEEEPKDG--IPHGHITSVSV 78
s_cerevisiae      DENDKLELTLDGTNDGRTIKLDPTYLAPGEKLVGYVLVKMNDDPDQQNEPPNGHITSLSV 120
                  * :.                          .:*****.**:::* :    *:*****::*

h_sapiens_V1      KRSHRRLGLAQKLMDQASRAMIENFNAKYVSLHVRKSNRAALHLYSNTLNFQISEVEPKY 137
c_elegans         KRSYRRLGLANKMMDQTARAMVETYNAKYVSLHVRVSNRAALN-YKNTLKFEIVDTEPKY 136
s_pombe           MRSYRHLGLAKRLMVQSQRAMVEVYGAKYMSLHVRKSNRAAIHLYRDTLQFDVQGIESKY 138
s_cerevisiae      MRTYRRMGIAENLMRQALFALREVHQAEYVSLHVRQSNRAALHLYRDTLAFEVLSIEKSY 180
                   *::*::* :*:.:* *:  *: * . *:*:***** *****:: * :** *::   * .*

h_sapiens_V1      YADGEDAYAMKR--DLTQMADELRRHLELKEKGRHVVLGAIENKVESKGNSPPSSGEACR 195
c_elegans         YADGEDAYAMRR--DLAKWAEE--RNIEPADR---------------------------- 164
s_pombe           YADGEDAYAMHK--DFSTLKFD---TPETN------------------------------ 163
s_cerevisiae      YQDGEDAYAMKKVLKLEELQISNFTHRRLKENE---------------------------- 213
                  * ********::  .:     .     .

h_sapiens_V1      EEKGLAAEDSGGDSKDLSEVSETTESTDVKDSSEASDSAS-- 235
c_elegans         ------------------EAYTTAKTTDDKKKNRS------- 181
s_pombe           -----------------DELAKTVQSLALNN----------- 177
s_cerevisiae      -----------------EKLEDDLESDLLEDIIKQGVNDIIV 238
                                  :     ::   :.
```

yeast growth

This is with plasmid over-expressing the various proteins. A 5 ml overnight culture was grown in SD^{-URA} at 30°C. Cells were diluted to an OD_{600} of 0.1 and either spotted in 1:5 serial dilutions on plates for 48 h (upper panel) or grown in 2 ml cultures at 30°C or 39°C under constant agitation for 24 h (lower panel). Optical density was plotted.
n=11

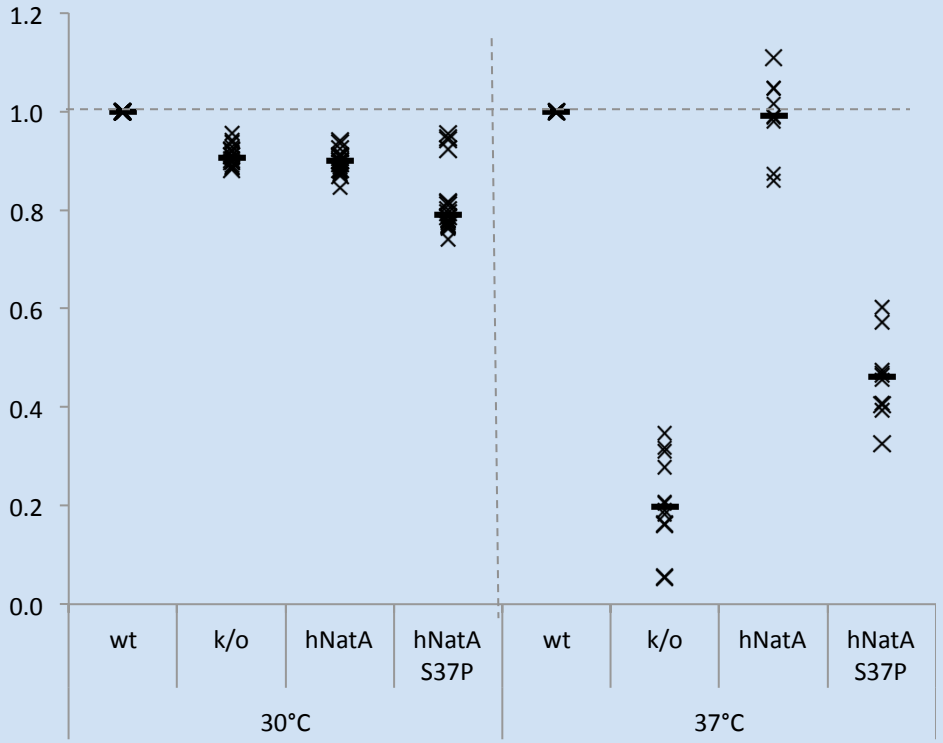


Same data (B), just presented differently.

yeast growth

A 5 ml overnight culture was grown in SD^{-URA} at 30°C. Cells were diluted to an OD_{600} of 0.1 and either spotted in 1:5 serial dilutions on plates for 48 h (upper panel) or grown in 2 ml cultures at 30°C or 39°C under constant agitation for 24 h (lower panel). Optical density was plotted.

n=11

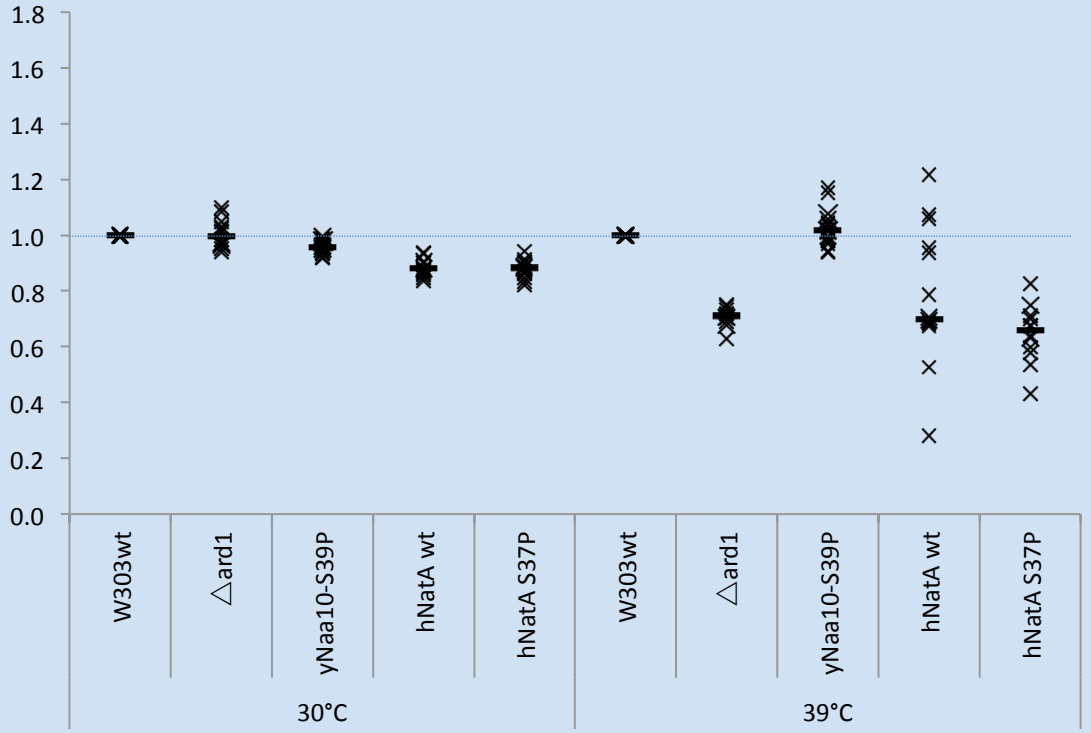


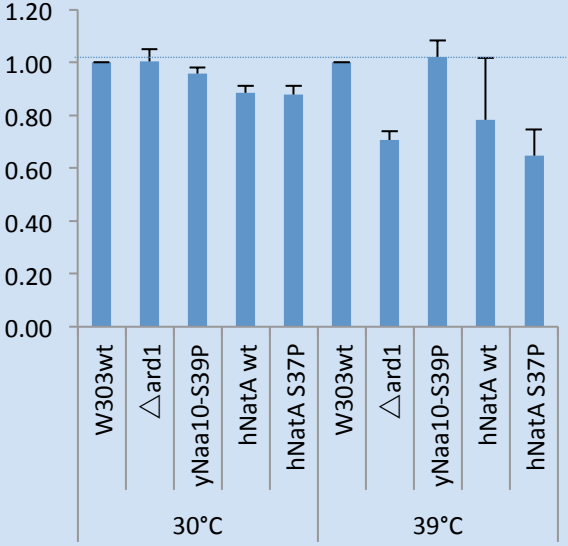
Endogenous, single-copy genes in yeast.

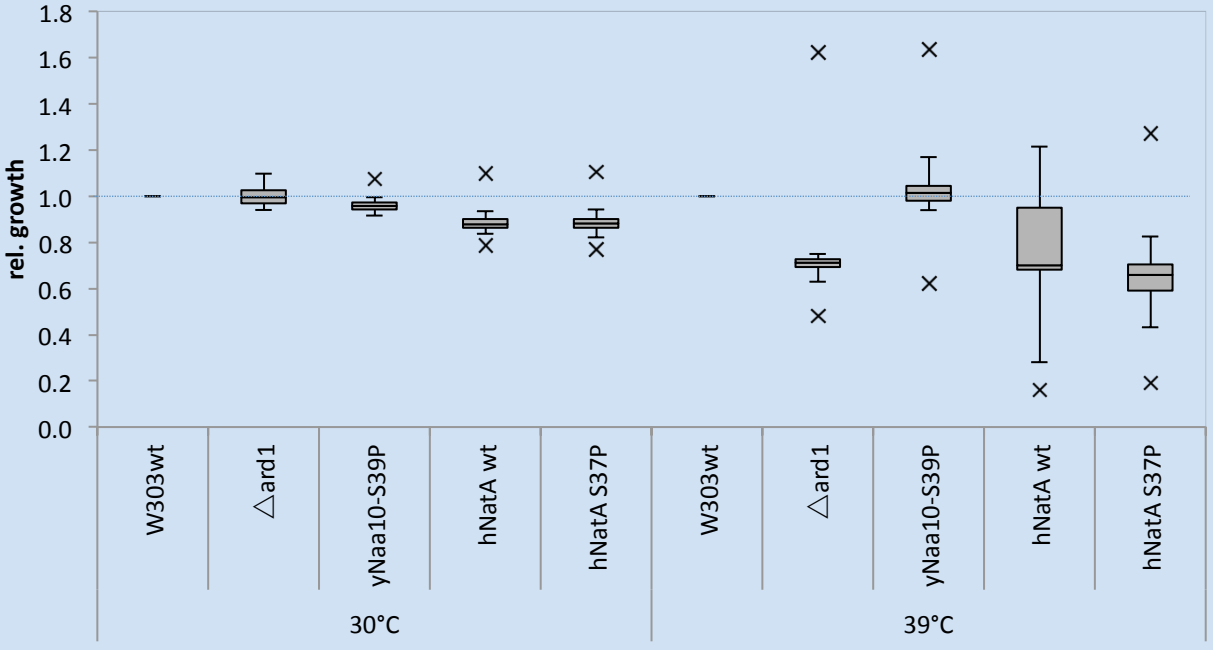
yeast growth

A 5 ml overnight culture was grown in SD^{-URA} at 30°C. Cells were diluted to an OD_{600} of 0.1 and either spotted in 1:5 serial dilutions on plates for 48 h (upper panel) or grown in 2 ml cultures at 30°C or 39°C under constant agitation for 24 h (lower panel). Optical density was normalized to the W303 wt strain for every independent experiment and plotted (X). The median of all experiments is shown as a short line

n=22 for S39P
n=11 for all other strains
n=11







Remarkably:

- We do not really know the expression of pretty much ALL mutations in **humans**, as we have not systematically sequenced or karyotyped any genetic alteration in **Thousands to Millions** of **randomly** selected people.

A Genotype-First Approach to Defining the Subtypes of a Complex Disease

Holly A. Stessman,[1] Raphael Bernier,[2] and Evan E. Eichler[1,3,*]
[1]Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA
[2]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA 98195, USA
[3]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA
*Correspondence: eee@gs.washington.edu
http://dx.doi.org/10.1016/j.cell.2014.02.002

Medical genetics typically entails the detailed characterization of a patient's phenotypes followed by genotyping to discover the responsible gene or mutation. Here, we propose that the systematic discovery of genetic variants associated with complex diseases such as autism are progressing to a point where a reverse strategy may be fruitful in assigning the pathogenic effects of many different genes and in determining whether particular genotypes manifest as clinically recognizable phenotypes. This "genotype-first" approach for complex disease necessitates the development of large, highly integrated networks of researchers, clinicians, and patient families, with the promise of improved therapies for subsets of patients.

Prioritization of neurodevelopmental disease genes by discovery of new mutations

Alexander Hoischen[1], Niklas Krumm[2] & Evan E Eichler[2,3]

Advances in genome sequencing technologies have begun to revolutionize neurogenetics, allowing the full spectrum of genetic variation to be better understood in relation to disease. Exome sequencing of hundreds to thousands of samples from patients with autism spectrum disorder, intellectual disability, epilepsy and schizophrenia provides strong evidence of the importance of *de novo* and gene-disruptive events. There are now several hundred new candidate genes and targeted resequencing technologies that allow screening of dozens of genes in tens of thousands of individuals with high specificity and sensitivity. The decision of which genes to pursue depends on many factors, including recurrence, previous evidence of overlap with pathogenic copy number variants, the position of the mutation in the protein, the mutational burden among healthy individuals and membership of the candidate gene in disease-implicated protein networks. We discuss these emerging criteria for gene prioritization and the potential impact on the field of neuroscience.

Table 4 Recurrent identical *de novo* mutations in 6 genes identified in 11 exome studies with different neurodevelopmental phenotypes

Gene	Coding effect	Mutation (genomic DNA level)	Mutation (cDNA level)	Mutation (protein level)	Study	Disorder
ALG13	Missense	ChrX(GRCh37):g.110928268A>G	NM_001099922.2:c.320A>G	p.Asn107Ser	de Ligt *et al.*[1]	ID
ALG13	Missense	ChrX(GRCh37):g.110928268A>G	NM_001099922.2:c.320A>G	p.Asn107Ser	Allen *et al.*[11]	EE
ALG13	Missense	ChrX(GRCh37):g.110928268A>G	NM_001099922.2:c.320A>G	p.Asn107Ser	Allen *et al.*[11]	EE
KCNQ3	Missense	Chr8(GRCh37):g.133192493G>A	NM_001204824.1:c.328C>T	p.Arg110Cys	Rauch *et al.*[2]	ID
KCNQ3	Missense	Chr8(GRCh37):g.133192493G>A	NM_001204824.1:c.328C>T	p.Arg110Cys	Allen *et al.*[11]	EE
SCN1A	Splice donor	LRG_8:g.24003G>A	NM_006920.4:c.602+1G>A	p.?	Allen *et al.*[11]	EE
SCN1A	Splice donor	LRG_8:g.24003G>A	NM_006920.4:c.602+1G>A	p.?	Allen *et al.*[11]	EE
CUX2	Missense	Chr12(GRCh37):g.111748354G>A	NM_015267.3:c.1768G>A	p.Glu590Lys	Rauch *et al.*[2]	ID
CUX2	Missense	Chr12(GRCh37):g.111748354G>A	NM_015267.3:c.1768G>A	p.Glu590Lys	Allen *et al.*[11]	EE
SCN2A	Missense	Chr2(GRCh37):g.166198975G>A	NM_021007.2:c.2558G>A	p.Arg853Gln	Allen *et al.*[11]	EE
SCN2A	Missense	Chr2(GRCh37):g.166198975G>A	NM_021007.2:c.2558G>A	p.Arg853Gln	Allen *et al.*[11]	EE
DUSP15	Missense	Chr20(GRCh37):g.30450489G>A	NM_080611.2:c.320C>T	p.Thr107Met	Neale *et al.*[7]	ASD
DUSP15	Missense	Chr20(GRCh37):g.30450489G>A	NM_080611.2:c.320C>T	p.Thr107Met	Fromer *et al.*[10]	SCZ

EE, epileptic encephalopathies; ASD, autism spectrum disorder; ID, intellectual disability; SCZ, schizophrenia.

Mutations as “Difference Makers”

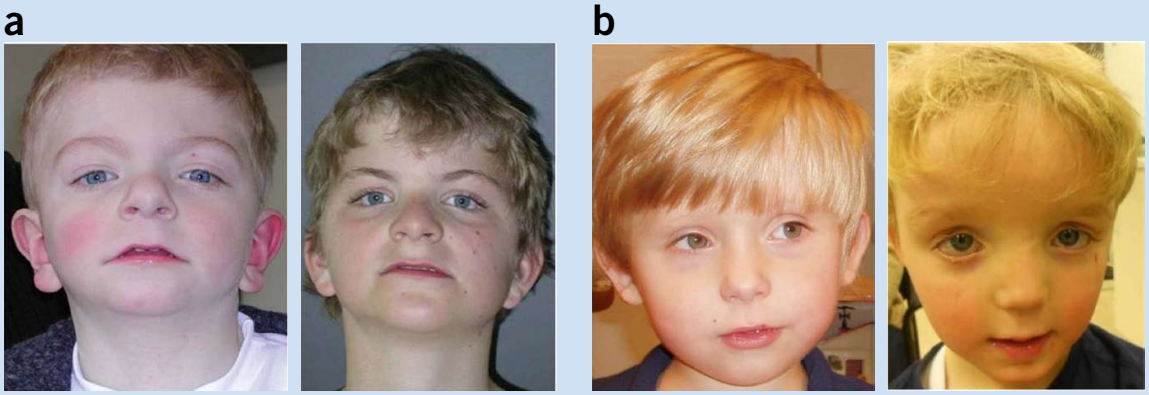


Figure 3 Phenotypic similarity of two patients with identical *PACS1 de novo* mutations and two patients with similar *ADNP* mutations. (**a**) These two unrelated patients show identical *de novo* point mutations (c.607C>T; p.Arg203Trp) in *PACS1* (RefSeq NM_018026.3)[53]. The striking similarity in phenotype includes low anterior hairline, highly arched eyebrows, synophrys, hypertelorism with downslanted palpebral fissures, long eyelashes, a bulbous nasal tip, a flat philtrum with a thin upper lip, downturned corners of the mouth and low-set ears. Reprinted from ref. 53, Copyright (2012), with permission from The American Society of Human Genetics. (**b**) These two unrelated patients both show LoF mutations in *ADNP* (c.2496_2499delTAAA; p.Asp832Lysfs*80 and c.2157C>G; p.Tyr719*)[44] resulting in a new SWI-SNF–related autism syndrome. Patients present with clinical similarities, including a prominent forehead, a thin upper lip and a broad nasal bridge. Reprinted from ref. 44.

Molecular Genetics & Genomic Medicine

Open Access

ORIGINAL ARTICLE

Disease variants in genomes of 44 centenarians

Yun Freudenberg-Hua[1,2], Jan Freudenberg[3], Vladimir Vacic[4], Avinash Abhyankar[4], Anne-Katrin Emde[4], Danny Ben-Avraham[5], Nir Barzilai[5], Dayna Oschwald[4], Erika Christen[1], Jeremy Koppel[1,2], Blaine Greenwald[2], Robert B. Darnell[4,6], Soren Germer[4], Gil Atzmon[5] & Peter Davies[1]

[1]The Litwin-Zucker Research Center for the Study of Alzheimer's Disease and Memory Disorders, The Feinstein Institute for Medical Research, North Shore-LIJ, Manhasset, New York 11030

[2]Division of Geriatric Psychiatry, Zucker Hillside Hospital, North Shore-LIJ, Glen Oaks, New York 11040

[3]Robert S. Boas Center for Genomics and Human Genetics, The Feinstein Institute for Medical Research, North Shore-LIJ, Manhasset, New York 11030

[4]New York Genome Center, 101 Avenue of the Americas, New York, New York 10013

[5]Institute for Aging Research Departments of Medicine and Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

[6]Department of Molecular Neuro-Oncology, Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, New York 10065

“Number of genes causing autism”

- Exome sequencing on 3000 quad families, i.e. mother, father, two children.
- Looking for newly arising mutation in child with autism, not found in parents or unaffected sibling.
- Estimating ~500 “genes” involved.

Researchers develop genetic profile of the Netherlands

- “The Genome of the Netherlands can greatly accelerate research into genes that play a key role in the development of chronic and age-related diseases. We can now focus specifically on the disease-causing genes”.
- “A noticeable result is that every participant in this research on average turned out to have twenty mutations that were thought to cause rare diseases, although the participants were perfectly healthy”.

http://medicalxpress.com/news/2014-06-genetic-profile-netherlands.html

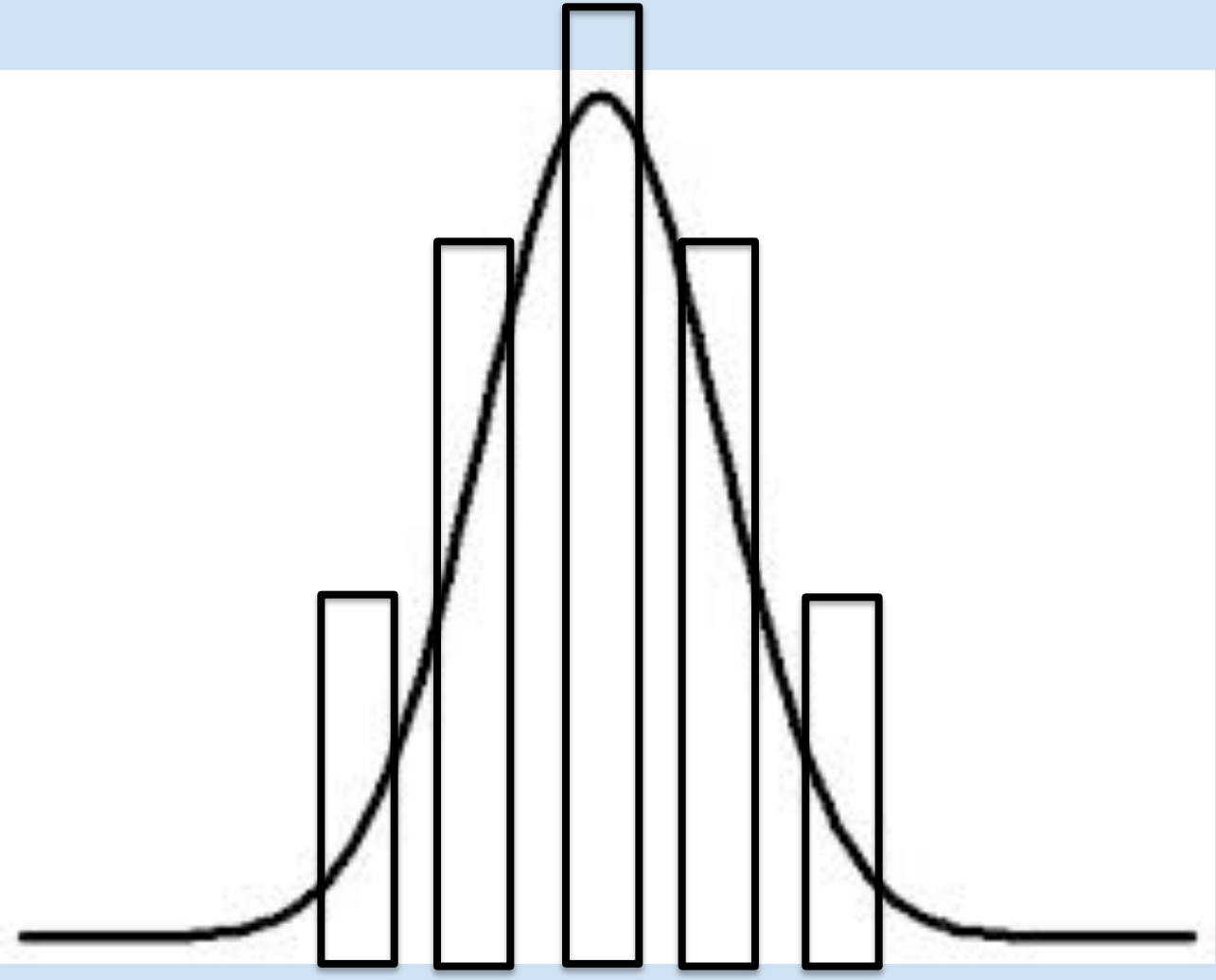
Clinical genetics of neurodevelopmental disorders

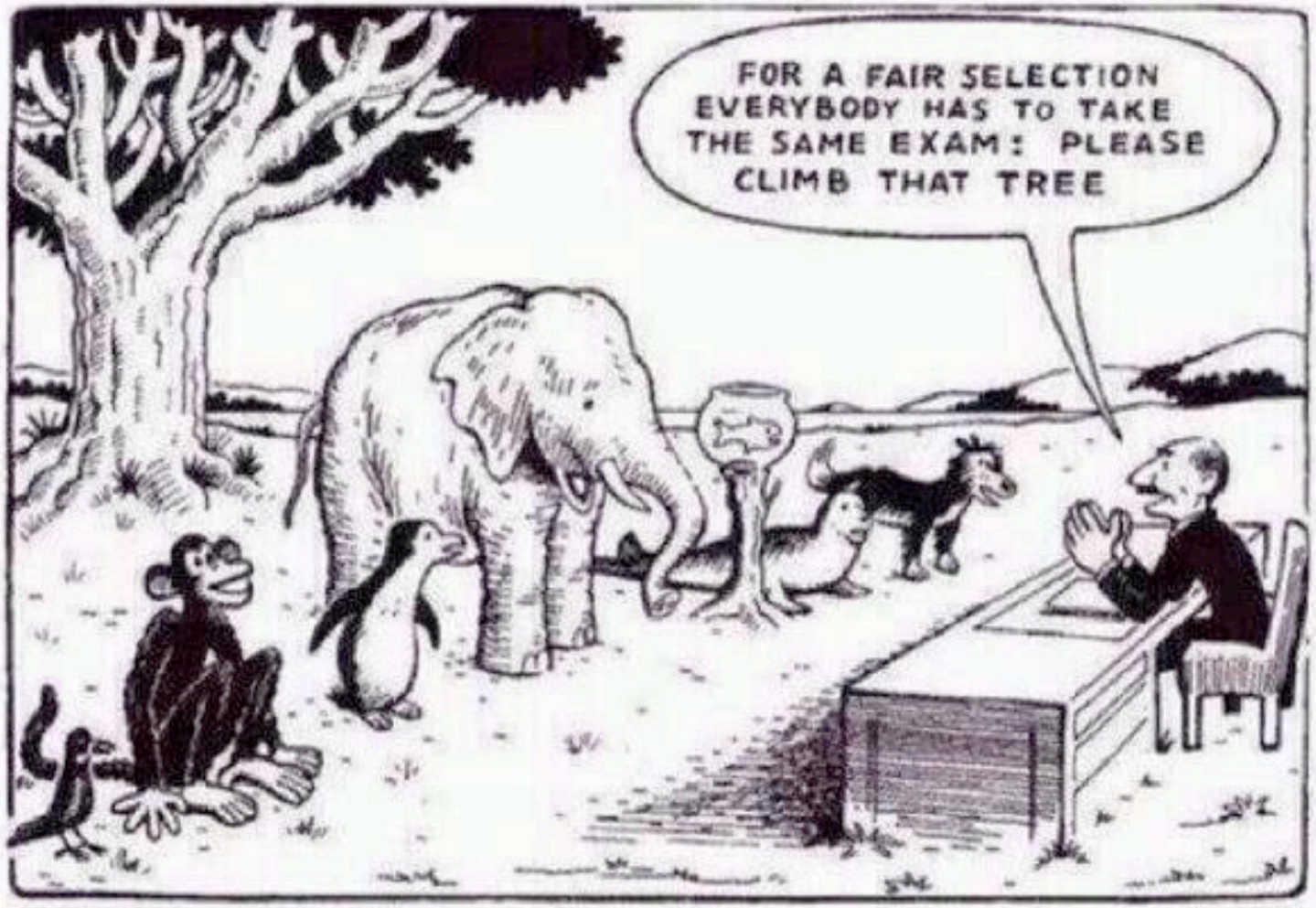
Gholson J Lyon and Jason O'Rawe

bioRxiv posted online November 18, 2013
Access the most recent version at doi:10.1101/000687

"There are ~12 billion nucleotides in every cell of the human body, and there are ~25-100 trillion cells in each human body. Given somatic mosaicism, epigenetic changes and environmental differences, no two human beings are the same, particularly as there are only ~7 billion people on the planet".

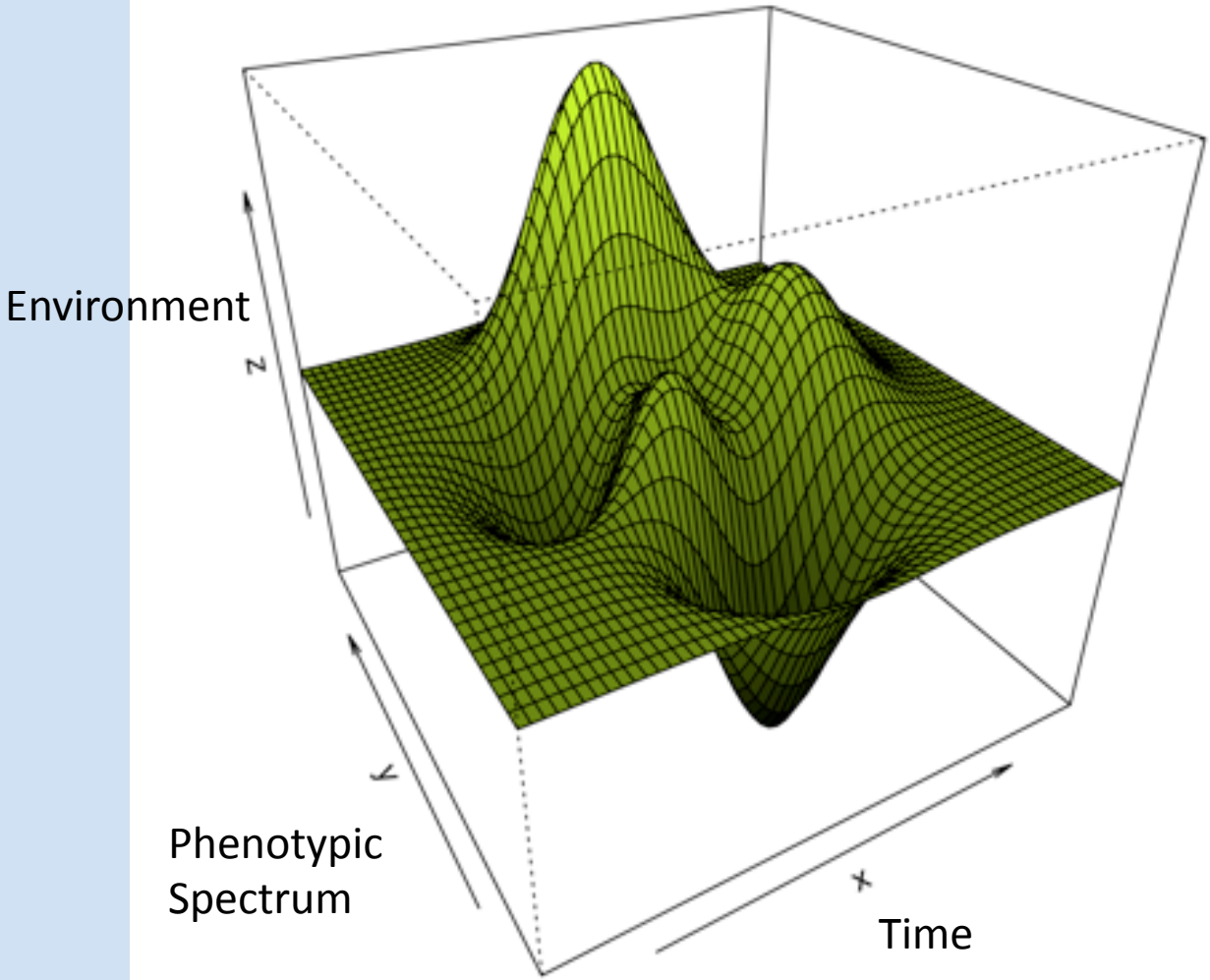
Categorical Thinking Misses Complexity



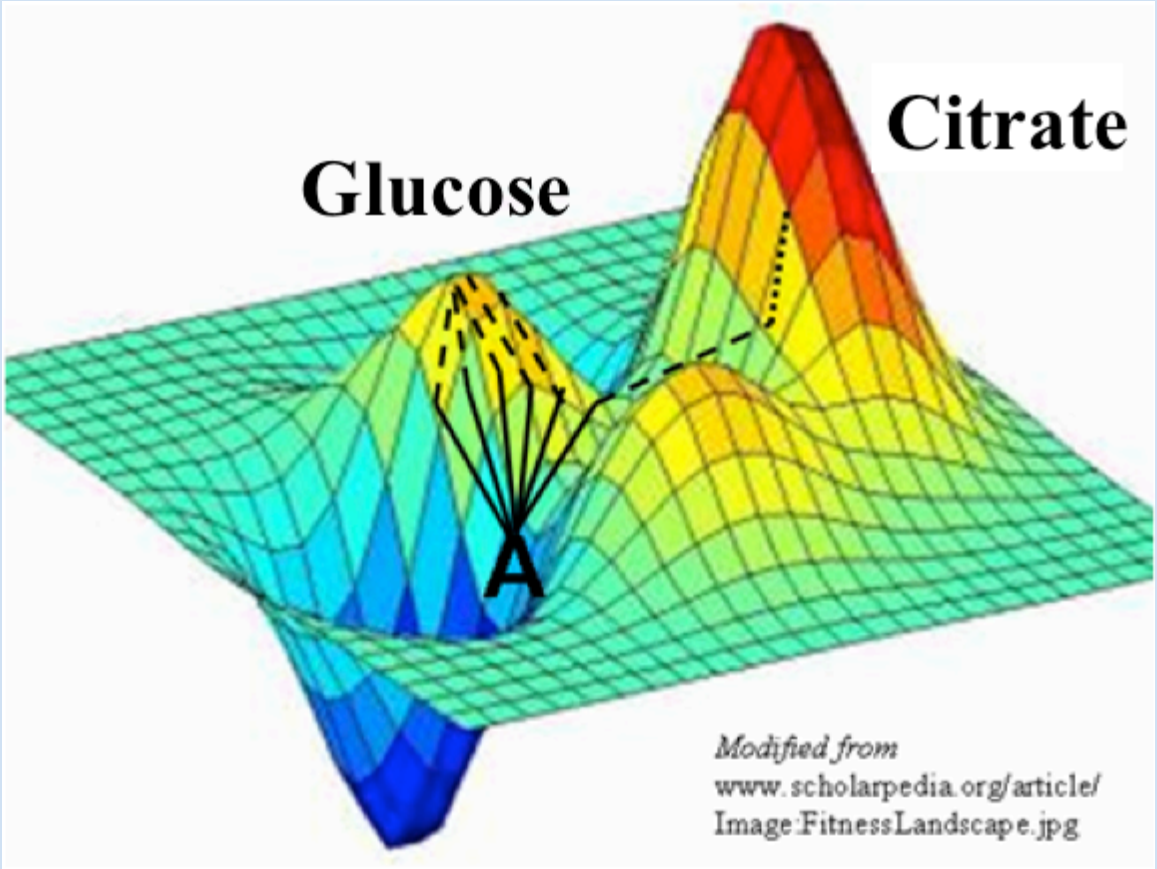


Our Education System

Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid.



A conceptual model of genotype-phenotype correlations. The *y* plane represents a phenotypic spectrum, the *x* plane represents the canalized progression of development through time, and the *z* plane represents environmental fluctuations.



E. coli adapting to low glucose conditions, in the context of media containing citrate. – Richard Lemski experiment

Limits of our current technology & knowledge

Analytic Validity

- Sequencing “clinical-grade genomes”

- Bioinformatics analysis

Clinical Validity

- **Genetic architecture of illness**

False Positives in the Literature

REPORT

XLID-Causing Mutations and Associated Genes Challenged in Light of Data From Large-Scale Human Exome Sequencing

Amélie Piton,[1,2,4,*] Claire Redin,[1,2,4] and Jean-Louis Mandel[1,2,3,*]

Because of the unbalanced sex ratio (1.3–1.4 to 1) observed in intellectual disability (ID) and the identification of large ID-affected families showing X-linked segregation, much attention has been focused on the genetics of X-linked ID (XLID). Mutations causing monogenic XLID have now been reported in over 100 genes, most of which are commonly included in XLID diagnostic gene panels. Nonetheless, the boundary between true mutations and rare non-disease-causing variants often remains elusive. The sequencing of a large number of control X chromosomes, required for avoiding false-positive results, was not systematically possible in the past. Such information is now available thanks to large-scale sequencing projects such as the National Heart, Lung, and Blood (NHLBI) Exome Sequencing Project, which provides variation information on 10,563 X chromosomes from the general population. We used this NHLBI cohort to systematically reassess the implication of 106 genes proposed to be involved in monogenic forms of XLID. We particularly question the implication in XLID of ten of them (*AGTR2*, *MAGT1, ZNF674, SRPX2, ATP6AP2, ARHGEF6, NXF5, ZCCHC12*, *ZNF41*, and *ZNF81*), in which truncating variants or previously published mutations are observed at a relatively high frequency within this cohort. We also highlight 15 other genes (*CCDC22*, *CLIC2*, *CNKSR2*, *FRMPD4*, *HCFC1*, *IGBP1*, *KIAA2022*, *KLF8*, *MAOA, NAA10*, *NLGN3*, *RPL10, SHROOM4*, *ZDHHC15*, and *ZNF261*) for which replication studies are warranted. We propose that similar reassessment of reported mutations (and genes) with the use of data from large-scale human exome sequencing would be relevant for a wide range of other genetic diseases.

Bring clinical standards to human–genetics research

Study protocols need to be rigorous, because more than science is at stake. Sometimes participants' lives depend on the results, writes **Gholson J. Lyon**.

Lyon and Wang *Genome Medicine* 2012, 4:58
http://genomemedicine.com/content/4/7/58

REVIEW

Identifying disease mutations in genomic medicine settings: current challenges and how to accelerate progress

Gholson J Lyon[*1,2] and Kai Wang[*2,3]

"It is perhaps naive to expect that these obstacles can be overcome within the next 20 years, and it may very well be the case that there might be a 50-year time horizon on the secure implementation of clinical genomics and individualized medicine. We certainly hope that every newborn will have the vast majority of their genome sequenced and digitally available by the year 2062".

Summary

- Ancestry, i.e. genetic background, matters.
- Collectively, we need to improve the accuracy of “whole” genomes, and also enable the sharing of genotype and phenotype data broadly, among researchers, the research participants and others.
- We need to sequence accurate whole genomes of large pedigrees, and then construct super-family structures.

The End– extra slides to follow

MICROBIAL EVOLUTION

Global epistasis makes adaptation predictable despite sequence-level stochasticity

Sergey Kryazhimskiy,[1,3]*† Daniel P. Rice,[1,3]* Elizabeth R. Jerison,[2,3] Michael M. Desai[1,2,3]†

Epistatic interactions between mutations can make evolutionary trajectories contingent on the chance occurrence of initial mutations. We used experimental evolution in *Saccharomyces cerevisiae* to quantify this contingency, finding differences in adaptability among 64 closely related genotypes. Despite these differences, sequencing of 104 evolved clones showed that initial genotype did not constrain future mutational trajectories. Instead, reconstructed combinations of mutations revealed a pattern of diminishing-returns epistasis: Beneficial mutations have consistently smaller effects in fitter backgrounds. Taken together, these results show that beneficial mutations affecting a variety of biological processes are globally coupled; they interact strongly, but only through their combined effect on fitness. As a consequence, fitness evolution follows a predictable trajectory even though sequence-level adaptation is stochastic.

SCIENCE sciencemag.org

27 JUNE 2014 • VOL 344 ISSUE 6191

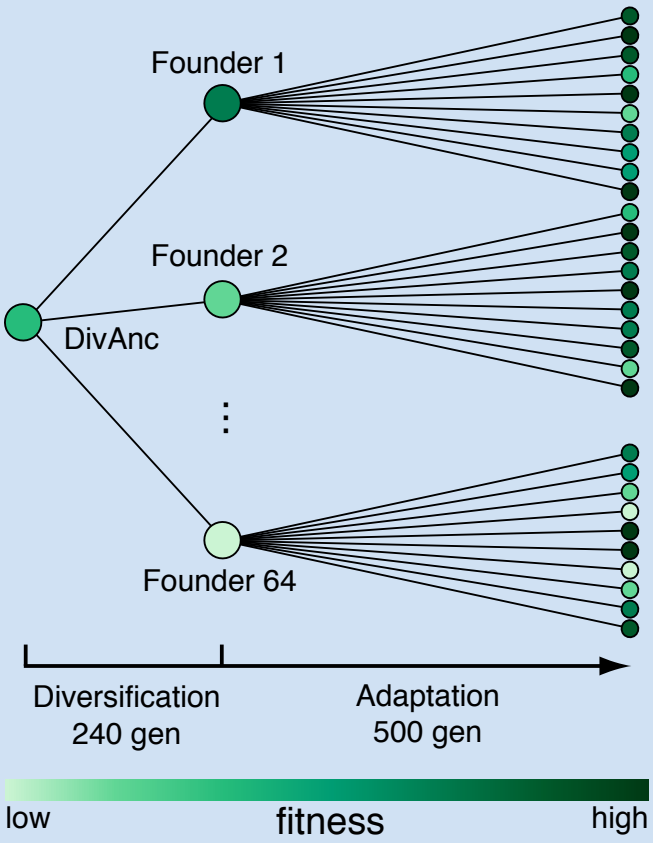


Figure S1. Experimental design. We created many independent lines from a single clone (DivAnc) which came from a previous evolution experiment in the same environment (15) and evolved each of them for 240 generations (Diversification). We then selected a single "Founder" clone from 64 of these lines (chosen to span a range of fitness) and evolved 10 independent replicate populations descended from each Founder for 500 generations (Adaptation).

- “Yet despite their lack of apparent functional relationship, these mutations are globally coupled by diminishing-returns epistasis; their effects are strongly mediated by background fitness but are otherwise essentially independent of the specific identity of mutations present in the background. The biological basis of this global coupling remains unknown”.