MAMMALIAN CELL GENETICS

M. Wigler C. Yen, M. Hamaguchi, M. Nakamura, R. Lucito, L. Serina, D. Esposito, E. Hatchwell, D. Dong, J. Stolarov, H. Tu, J. West, L. Rodgers, M. Riggs, J. Troge, B. O'Conner, V. Tu, A. Buglino, M. Katari, D. Almanzar, J. Brodsky, J. Douglas

Our overall research program is only modestly modified from last year's. Our general aims remain the same: to define the mutations that result in cancer and other human diseases; to understand the mechanisms by which these mutations cause disease; and to develop strategies to modify the course of disease. Our effort is specifically focused in three areas: *combinatorial chemistry*, in which we are trying to develop a general methodology that will enable us to manipulate the biochemical pathways involved in cancer; *signal transduction*, in which we study the pathways of intracellular communication that become deranged in cancer cells; and *genomics*, in which we apply and develop tools for the genetic analysis of human disease.

The combinatorial program (Dennis Dong), conducted in collaboration with Peter Nestler and colleagues here at the Laboratory, has demonstrated the feasibility of finding small branched peptidic molecules that recognize the peptide epitopes of proteins and alter their biochemical interactions.

Our signal transduction studies are focused in two areas: cellular pathways that relate to the oncogenic Ras proteins (Hua Tu and Dennis Dong) and the analysis of the PTEN tumor suppressor (Javor Stolarov, Ken Chang, and Dennis Dong, in collaboration with Mike Myers and Nick Tonks at CSHL).

Our genomic studies are now divided into three parts. First, we are searching for cancer genes at the many loci that we have discovered by the application of representational difference analysis, or RDA. We are seeking tumor suppressor genes at six of the loci that are commonly deleted in cancer (Clifford Yen, Masaaki Yamaguchi, Robert Lucito, and Diane Esposito, in collaboration with Dick McCombie and Larry Parnell at CSHL). The PTEN tumor suppressor was found this way. In collaboration with Scott Powers and colleagues at Tularik, we are also seeking oncogenes in the regions that are commonly amplified in cancer (Clifford Yen and Mariko Nakamura). In the second part, we are expanding the use of genomic representations and microarrays to develop "genomic chips" (Robert Lucito and Joe West, in collaboration with Andy Reiner at CSHL). We expect to develop thereby a vastly more powerful engine for the discovery of genes mutated in cancer. Additionally, we hope that genomic chips will provide a method for rapidly genotyping clinical human cancers, leading to improvements in the diagnosis and the selection of therapies. The cancer genomics studies are conducted in collaboration with Larry Norton of the Sloan Kettering Memorial Cancer Institute. In the third part, we are adapting the techniques of representational analysis to find spontaneous mutations that cause germ-line transmission of sporadic hereditary diseases (Eli Hatchwell and Lidia Serina).

Combinatorial Chemistry

D. Dong

In collaboration with Peter Nestler's lab here at the Laboratory, we are investigating the binding capacity of a class of branched peptidic molecules, or molecular "forceps." These "forceps" have been generated as members of encoded combinatorial libraries, a method developed in collaboration with C. Still's laboratory at Columbia University (Ohlmeyer et al., *Proc. Natl. Acad. Sci. 90:* 10922 [1993]). As targets, we have chosen the carboxyl end of H-RAS, because it is the substrate of farnesyl transferase, the first step in the essential processing of RAS proteins (Gibbs, *Cell 65:* 1 [1991]), and because this end is predicted by crystallographic studies to be an exposed peptide. Several forceps libraries were screened against both H-RAS

protein and H-RAS terminal peptides using "on-bead" assays. Several members of this forceps library thus found were shown to bind specifically to the carboxy-terminal sequences of H-RAS and to recognize these sequences when they are fused to other proteins. At high concentrations, these forceps block the enzymatic farnesylation of H-RAS, in a manner strongly suggesting competition with enzyme for substrate binding. These results (Dong et al. 1999) suggest that branched peptidic molecules have the general capacity to recognize peptides and could, in principle, be used to modulate biological systems. However, improvements in the chemical backbone and binding affinities are clearly necessary before this end can be achieved. Our current efforts are now focused on screening more selective libraries and validating the strategy of specifically blocking H-RAS maturation as a way to modulate tumor growth.

Studies of MAP Kinase Cascades

H. Tu, D. Dong

The MAP protein kinase cascade, consisting of components from the MEKK, MEK, and MAPK protein kinase families, is a conserved triad found in all eukaryotic organisms and is under the control of Ras proteins in multiple species from diverse phyla. Our first encounter with the triad was in the sexual differentiation pathway in the yeast *Schizosaccharomyces pombe*. In that organism, Ras1 acts directly on Byr2, an MEKK, by binding to its regulatory domain. We have conducted many studies of the integration and control of this pathway during the past years, and in previous years demonstrated that Shk1, a conserved protein kinase, activates Byr2 by "opening" its conformation. Shk1 is a member of the PAK family of protein kinases. Like the MEKK kinases, these kinases also have a carboxy-terminal catalytic domain and an amino-terminal regulatory domain that contains the binding site for certain RHO proteins, members of the Ras superfamily of GTPases. We have demonstrated that the tools for analyzing the regulatory components of Byr2 can be used for the study of the PAK family of kinases (Tu and Wigler 1999). In particular, the RHO proteins bind the regulatory domain of Shk1, and this domain of Shjk1 binds the catalytic domain. Moreover, mutations in the Shk1 regulatory region that abolish binding to the catalytic region open the conformation of Shk1 and activate it.

The PTEN Tumor Suppressor

J. Stolarov, K. Chang, C. Yen, D. Dong

The use of RDA led to the identification of a region on chromosome 10 deleted in a variety of human cancers, including breast, prostate, and brain cancers. We therefore suspected that this region harbored a tumor suppressor gene. In a collaboration with Ramon Parsons of Columbia University, a gene from this region, PTEN, was identified and found to be mutated in many cancers (Li et al., *Science 275:* 1943 [1997]). Parsons and collaborators found that the PTEN gene is mutated in the germ line of patients with Cowdens syndrome, an inherited predisposition to multiple neoplasms (Liaw et al., *Nat. Genet. 16:* 64 [1997]; Marsh et al., *Nat. Genet. 16:* 333 [1997]). More recently, other investigators have shown that transgenic mice hemizygous for PTEN are cancer prone, completing the evidence that PTEN is a tumor suppressor.

In a collaboration with Mike Myers and Nick Tonks here at the Laboratory, we showed that PTEN is a protein tyrosine phosphatase with a somewhat atypical substrate preference (Myers et al., *Proc. Natl. Acad. Sci. 90:* 10922 [1997]). More recently, we have shown that PTEN has a lipid phosphatase activity, specific for phosphatidylinositol-3,4,5 phosphate, PIP_3, (Myers et al. 1998). In fact, one of the mutations in Cowdens syndrome destroys lipid phosphatase activity, leaving the protein phosphatase activity apparently untouched, suggesting that the lipid phosphatase activity is essential for PTEN function. This hypothesis is made more plausible because PIP_3 is a second messenger that has been previously implicated in the growth and survival of cancer cells.

We are pursuing two lines of investigation into the biology of PTEN. First, we have expressed the PTEN protein in cultured human tumor cells that lack a functional PTEN gene. Constitutive expression of PTEN appears to slow the growth and flatten these cells, and significantly suppresses tumorigenicity of these cells in experimental animals. Our current efforts are directed at constructing an inducible expression system for

PTEN in cultured cells, so that these effects can be examined more accurately. Second, we are studying the PTEN homologs in the yeasts *Saccharomyces cerevisiae* and *S. pombe*. As yet, no phenotype is apparent in *S. cerevisiae* when the PTEN homolog, Tep1, is deleted or overexpressed. We are just initiating similar experiments in *S. pombe*.

Loci Deleted or Amplified in Breast Cancer

C. Yen, M. Hamaguchi, R. Lucito, Y. Han,
D. Esposito, M. Nakamura

Many loci that are homozygously deleted or amplified in breast cancer have now been identified by RDA. These loci are presumed to carry tumor suppressor genes or oncogenes, respectively. Among the loci we have found deleted are those encoding the PTEN gene and the p16/ARF (alternate reading frame) tumor suppressors. Virtually all of the loci known to be amplified in breast cancer have also been identified by RDA, including ErbB2, c-Myc, and the cyclin D loci. Six deletion loci and an equal number of amplified loci remain that are uncharacterized and in which we are searching for genes.

Unfortunately, the regions of homozygous deletion and amplification are quite large, often greater than a megabase upon initial characterization. Therefore, the first step in characterizing these loci entails "epicenter" mapping, a process that delineates the minimum region that is commonly deleted in cancers. To achieve this, it is necessary to search through large archives of stored breast cancer biopsies. For this purpose, we have developed a technique for "immortalizing" and amplifying the DNA from small tumor samples (Lucito et al. 1998). The immortalized DNAs can then be analyzed for gene copy gains or loss, which we currently do by quantitative polymerase chain reaction (Q-PCR).

Once the epicenter has been determined, a variety of techniques will be employed for searching for genes, including exon trapping, mining the existing databases for expressed sequences, and a new method for hybridization screening that we have developed called rapid isolation of cDNAs by hybridization (RICH). In the latter procedure, cDNAs from libraries are identified by their ability to hybridize to a given large-insert chromosomal vector such as a BAC (bacterial artificial chromosome) or YAC (yeast artificial chromosome) (Hamaguchi et al. 1998).

Another method for gene searching entails DNA sequencing and sequence analysis. In a collaboration with Richard McCombie's lab at CSHL we have finished the sequencing of the PTEN locus. PTEN appears to arise from a gene-poor region of the genome. Nevertheless, we are using these studies to model methods for gene finding from large contiguous tracts of genomic sequence.

Genomic Microarrays

R. Lucito, J. West

Although RDA was an advance over existing methods for finding lesions in cancer cells, we need to accelerate the pace of cancer gene discovery. Several opportunities have moved our genomic studies into the arena of microarrays. DNA microarrays are spatially organized collections of probes on a surface ("chip") that are used for measuring DNA concentrations of complementary sequences by hybridization (DeRisi et al., *Nat. Genet. 14:* 457 [1996]; Chee et al., *Science 274:* 610 [1996]). We are beginning to use this technology to perform RDA on a chip. Eventually, we expect that this approach will allow us to scan the genome for amplifications and deletions more thoroughly, vastly more rapidly, and in a more quantitative manner than ever before. This application of microarrays to genomics is made feasible by the successful implementation of representations, which allows us to sample the genome reproducibly, yielding DNA populations with reduced nucleotide complexity and hence improved hybridization kinetics. With the help of Raju Kucherlapati, Geoffery Childs, and Aldo Massimi, from the Albert Einstein School of Medicine, we have constructed and hybridized our first pilot chips. Andy Reiner of CSHL's Bioinformatics division is directing the software development needed to interpret and analyze the data. We expect that the genomic chips we develop will also have application to cancer and genetic clinical diagnosis.

Applications of RDA to Spontaneous Genetic Disease

L. Serina, E. Hatchwell

In principle, RDA can be used to detect genetic rearrangements. Such events occur in cancer, but they also sometimes occur as spontaneous genetic lesions in the germ line of a parent. When these germ-line mutations are transmitted to offspring, "sporadic" genetic disease, such as some forms of neurofibromatosis, Duchenne's muscular dystrophy, William's Syndrome, and a host of other apparent anomalies such as autisms may result. We have begun to apply RDA to families of children with suspected sporadic genetic disease. Success could open a new chapter in the study of human genetic disorders.

Although we have encountered rearranged regions with RDA in cancers, the method was not optimized for that task. We are now testing a variety of improved methods for the "corepresentation" of progeny and parental DNAs that are designed to find restriction endonuclease fragments in the progeny that are not present in either parent. In principle, such differences should arise predominantly by spontaneous mutation. At present, we have been tackling technical obstacles that are formidable but can be overcome.

PUBLICATIONS

Hamaguchi M., O'Connor E.A., Chen T., Parnell L., McCombie R.W., and Wigler M.H. 1998. Rapid isolation of cDNA by hybridization. *Proc. Natl. Acad. Sci.* **95:** 3764–3769.

Lucito R., Nakimura M., West J.A., Han Y., Chin K., Jensen K., McCombie R., Gray J.W., and Wigler M. 1998. Genetic analysis using genomic representations. *Proc. Natl. Acad. Sci.* **95:** 4487–4492.

Myers M.P., Pass I., Batty I.H., Van Der Kaay J., Stolarov J.P., Hemmings B.A., Wigler M.H., Downes C.P., and Tonks N.K. 1998. The lipid phosphatase activity of PTEN is critical for its tumor suppressor function. *Proc. Natl. Acad. Sci.* **95:** 13513–13518.

Webb C.P., Van Aelst L., Wigler M.H., and Vande Woude G.F. 1998. Signaling pathways in Ras-mediated tumorigenicity and metastasis. *Proc. Natl. Acad. Sci.* **95:** 8773–8778.

In Press

Dong D.I., Liu R., Sherlock R., Wigler M.H., and Nestler H.P. 1999. Molecular forceps from combinatorial libraries prevent the farnesylation of Ras by binding to its carboxyl terminus. *Chem. Biol.* **6:** 133–141.

Tu H. and Wigler M. 1999. Genetic evidence for Pak1 autoinhibition and its release by Cdc42. *Mol. Cell. Biol.* **19:** 602–611.