

MAMMALIAN CELL GENETICS

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

Our overall research objectives are to define the genetic alterations that cause human disease, especially cancer, and to determine their biochemical consequences. This effort is focused on three areas: signal transduction, in which we study the pathways of communica-

tion that become deranged in cancer cells; genomics, in which we apply and develop tools for the genetic analysis of human disease; and combinatorial chemistry, in which we are trying to develop a general methodology that will enable us to manipulate the bio-

chemical pathways involved in cancer. The ras pathway has figured prominently over the years in our studies of signal transduction, and RDA (representational difference analysis) has figured prominently in our genomic studies. RDA is a tool for genomic difference analysis developed in previous years in collaboration with Nikolai Lisitsyn (Lisitsyn et al., *Science* 259: 946 [1993]).

Studies of MAP Kinase Cascades

H. Tu, D. Dong

The MAP protein kinase cascade is a conserved motif in the signal transduction repertoire of probably all eukaryotic organisms. Our first encounter with it was in the ras-controlled sexual differentiation pathway in the yeast *Schizosaccharomyces pombe*. In that organism, ras1 acts directly on byr2, the kinase that activates byr1, itself a kinase that activates spk1, the kinase that is the homolog of the mammalian MAP/ERK family of kinases. In the control of this cascade, ras1 is joined by ste4, a leucine zipper protein that we have studied in previous years (Barr et al., *Mol. Cell. Biol.* 16: 5597 [1996]). We have demonstrated that ras1 and ste4 act independently and jointly upon the regulatory domain of byr2. We have now demonstrated the direct involvement of yet a further protein kinase, shk1, which appears to act by releasing the catalytic domain of byr2 from inhibition by its regulatory domain, thereby "opening" its conformation (Tu et al. 1997). We assess these changes in conformation by the two-hybrid system of Fields and Song (*Nature* 340: 245 [1989]). Our results suggest that ras1 exerts its influence upon byr2 through two routes: one by directly acting upon byr2 (Van Aelst et al., *Proc. Natl. Acad. Sci.* 90: 6213 [1993]), and another through a pathway involving scd1, cdc42, and shk1 (Chang et al., *Cell* 79: 131 [1994]).

byr2 is a member of the MEKK family of protein kinases, and shk1, which acts upon it, has a similar structure and belongs to another family of protein kinases, 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 a member of the ras family of GTPases. We have tested whether the tools for analyzing the regulatory components of byr2 could be used for the study of the PAK family of kinases. Preliminary studies encourage us in

this. The regulatory domain of shk1 binds its catalytic domain, and mutations in the regulatory region that abolish its binding to the catalytic region "open" the conformation of shk1 and genetically "activate" the kinase. Using the opening of the wild-type shk1 as a genetic screen, we have identified a novel open reading frame that may encode an upstream activator of this kinase.

PTEN, a Tumor Suppressor Encoding a Phosphoprotein Phosphatase

J. Strolarov, K. Chang, C. Yen

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. 1997). In collaboration with Michael Myers and Nick Tonks at CSHL, PTEN was shown to encode a protein tyrosine phosphatase with a somewhat atypical substrate preference (Myers et al. 1997). In addition to its relation to phosphatases, the PTEN protein contains motifs also found in cytoskeletal proteins. The discovery by Parsons and collaborators that the PTEN gene is mutated in a hereditary disease predisposing the afflicted with neoplastic growths completes the demonstration that PTEN is a tumor suppressor gene (Liaw et al., *Nat. Genet.* 16: 64 [1997]; Marsh et al., *Nat. Genet.* 16: 333 [1997]).

Three lines of investigation into the biology of PTEN are being pursued. First, we have expressed the PTEN protein in cultured tumor cells that lack a functional PTEN gene. As yet we have not observed a clear phenotypic response in cell culture. Second, PTEN has a homolog in the yeast *Saccharomyces cerevisiae*, and experiments looking at the effects of deleting or overexpressing the gene in such yeast have begun. Again, as yet no phenotype is apparent. Third, we have searched for binding partners of PTEN using the two-hybrid system of Fields and Song and have found a number of binding partners, including proteins containing PDZ domains. Many PDZ domain proteins act as scaffolding proteins for proteins involved in signal transduction (Ponting et al., *BioEssay* 17: 5876 [1997]). We are attempting to explore the physiological relevance of our findings.

In a collaboration with Richard McCombie's lab at CSHL, we initiated the sequencing of the PTEN locus. Nearly 180 kbp of genomic sequence have been completed by his group.

Loci Amplified in Breast Cancer

M. Nakamura, R. Lucito, C. Yen

From the analysis of breast cancer biopsies, we have derived many probes derived from amplified loci. These loci are suspected to contain oncogenes that drive the growth and malignancy of breast cancers. Along with loci that were known previously, such as the ErbB2, c-myc, and cyclin D loci, several new loci have been observed. We are in the midst of trying to identify the genes encoded in these regions and have been concentrating on a region that lies near to but is distinct from the c-myc region.

As a first step in characterizing amplified loci, we need to perform "epicenter" mapping, a process that delineates the minimum region which is commonly amplified. 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 of the samples for present and future analyses (Lucito et al. 1998). The technique is based on a high-complexity "representation" of tumor DNA, which entails its cleavage, ligation to adaptors, and subsequent PCR amplification. We have shown that such representations are useful for measuring gene copy number and, in collaboration with Joe Grey's laboratory (University of California, San Francisco), can be used for comparative genomic hybridization, a technique useful for the global assessment of genomic changes (Kallioniemi et al., *Genes, Chromosomes & Cancer* 10: 231 [1994]). "Quantitative" PCR can also be used to measure gene copy number in high-complexity representations (Gibson et al., *Genome Res.* 6: 995 [1996]).

Loci Deleted in Breast Cancer

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

Several loci have now been identified by RDA that are homozygously deleted in breast cancer. Among these

are the PTEN locus and the locus encoding p16, the inhibitor of the cdc2 and cdc4 cyclin-dependent protein kinases. Ten loci remain in which we are searching for genes. This process also entails epicenter mapping, which as described above, narrows the region in which we need to search for genes. Unfortunately, the regions of homozygous deletion are quite large, often greater than a megabase upon initial characterization, and hence the process of gene searching has become the rate-limiting step in discovery. Several steps have been taken to accelerate this phase of our work. The first is the use of high-complexity representations, which has already been described. The second is the development of a solution hybridization technique, called RICH, for the cloning of cDNAs that hybridize to a given large-insert chromosomal vector such as a BAC or YAC (Hamaguchi et al. 1998).

RICH (rapid isolation of cDNAs by hybridization) is meant to complement the searching of genes by exon trapping and sequence analysis. The ends of fragments of restriction-endonuclease-cleaved cDNA libraries are modified so that when they hybridize to fragments from the insert of large cloning vectors, they can be selected and amplified and cloned. The application of this technique has been successfully tried on the c-myc locus (an example of amplification) and the PTEN locus. In a "mock" comparison with exon trapping and sequence analysis, the method fares well.

Rearrangements in Cancer and Spontaneous Genetic Disease

L. Serina, E. Hatchwell, C. Yen

We are beginning to adapt RDA to detect rearrangements in cancer. Several reasons for attempting this are, first, the present use of RDA for the discovery of tumor suppressors is biased to the discovery of large regions of homozygous deletion, and these may be "gene-poor" regions of the genome. Second, every deletion will create a rearrangement, and small deletions or large deletions do so equally. Small deletions are more likely to involve directly the rearrangement of a target gene, and hence the bridging region is more likely to be directly linked to a target gene. Although we have encountered rearranged regions with RDA as we now practice it, the method was not optimized for that task. We are now testing a variety of improved methods for the "co-representation" of tumor and nor-

mal DNAs that are designed to find differences in the mobility of restriction endonuclease fragments in the respective genomes. The same kinds of methods may find application to the discovery of the spontaneous genetic lesions that can occur in the germ line of a parent and be transmitted as de novo disease to offspring.

Combinatorial Chemistry

D. Dong

In collaboration with Peter Nestler's lab at CSHL, we have been investigating the binding capacity of a class of branched peptides, or "forcep" molecules. These have been generated as members of encoded combinatorial libraries, a method developed in collaboration with C. Still's lab 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 to be exposed as a peptide. Several members of this forceps library have been shown to bind specifically to the carboxy-terminal sequences of H-ras and to recognize these sequences when they are fused to other proteins. A subset of these forceps also block the farnesylation of H-ras. These results, although very preliminary at the moment, suggest that branched pep-

tidic molecules will have the general capacity to recognize peptides and may indicate that this will be a fruitful class of molecules that can modulate biological systems.

PUBLICATIONS

- Li, J., C. Yen, D. Liaw, K. Podsypanina, S. Bose, S.I. Wang, J. Puc, C. Miliareis, L. Rodgers, R. McCombie, S.H. Bigner, B.C. Giovanella, M. Ittmann, B. Tycko, H. Hibshoosh, M.H. Wigler, and R. Parsons. 1997. PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast and prostate cancer. *Science* **275**: 1943-1947.
- Myers, M.P., J.P. Stolarov, C. Eng, J. Li, S.I. Wang, M.H. Wigler, R. Parsons, and N.K. Tonks. 1997. P-TEN, the tumor suppressor from human chromosome 10q23, is a dual-specificity phosphatase. *Proc. Natl. Acad. Sci.* **94**: 9052-9057.
- Tu, H., M. Barr, D.L. Dong, and M. Wigler. 1997. Multiple regulatory domains on the Byr2 protein kinase. *Mol. Cell Biol.* **17**: 5876-5887.

In Press

- Hamaguchi, M., E.A. O'Connor, T. Chen, L. Parnell, R.W. McCombie, and M.H. Wigler. 1998. Rapid isolation of cDNA by hybridization. *Proc. Natl. Acad. Sci.* **95**: 3764-3769.
- Lucito, R., M. Nakamura, J.A. West, Y. Han, K. Chin, K. Jensen, R. McCombie, J.W. Gray, and M. Wigler. 1998. Genetic analysis using genomic representations. *Proc. Natl. Acad. Sci.* **95**: (in press).