Baldocchi, Russ [15]

Hybridization to oligonucleotide arrays for the assessment of gene copy number

Russ Baldocchi[1], Richard Glynne[2], Dave Kowbel[1], Ed Tom[2], Colin Collins[1], David Mack[2] & Joe Gray[1]

[1]*University of California San Francisco Cancer Center, San Francisco, California, USA*
[2]*Eos Biotechnology, Inc., South San Francisco, California, USA*

Array-based comparative genomic hybridization (CGH) provides a higher-resolution and more quantitative alternative to chromosome CGH for the assessment of abnormalities in genomic copy number. In array-based CGH analyses published to date, array elements have been composed of cloned DNA sequences (such as P1s, BACs or complementary DNAs), and thus copy number abnormalities can be mapped from data on the genomic mapping of such clones. These advanced methods are limited by difficulties in the production of cloned DNAs and distortions in the data arising from elements in the arrays that are present at multiple copies in the normal genome. To overcome some of these limitations, we have produced a low-complexity probe based on the polymerase chain reaction (PCR). In oligonucleotide array CGH, probes are prepared by multiplex PCR amplification using 10–50 primer pairs. Two 50-base oligonucleotides that map between the primer pairs serve as array elements. Because the linearity of this method apparently depends on primer complexity during PCR, we substituted adapter sequences (which make up the 5′ end of our PCR primers). We compared the performances of T7 and R24 adapters in a 24-plex PCR with one another and with previously obtained copy number data on several chromosome 20 loci in MCF7 cells obtained from fluorescence *in situ* hybridization. We found that use of the T7 adapters improved the amplitude of test-to-reference ratios 2.3-fold, suggesting less-dampened ratios and increased linearity in our assessment of copy number by this method. We will demonstrate the large-scale application of oligonucleotide array CGH to ovarian and breast cancer samples.

Basilion, J.P. [16]

Magnetic resonance imaging of HSV-based *in vivo* transgene expression

Dagmar Hoegemann[1], Tomo Ichikawa[2], E.A. Chiocca[2], Ralph Weissleder[1] & James P. Basilion[1,2]

[1]*Center for Molecular Imaging Research and* [2]*Neurosurgical Service, Molecular Neuro-Oncology Laboratory, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA*

Real-time imaging of gene expression *in vivo* at high spatial resolution would significantly aid our ability to study and understand transgene expression in live animals and potentially in a clinical gene therapy setting. Magnetic resonance (MR) imaging has recently achieved spectacular image resolution (50 μm^3 voxel resolution *in vivo*), making possible imaging at very high resolutions in small animals, during development and in clinical practice. Recently we have developed gene-targeted MR imaging contrast agents that allow us to use clinical MR imaging systems to monitor gene expression *in vivo*. We have demonstrated that expression of an altered form of the transferrin receptor (ETR) can be imaged *in vivo* using MR imaging and ETR-targeted imaging probes[1]. ETR expression could be used as a surrogate marker to monitor viral delivery and expression of other transgenes in real time, *in vivo* and potentially in a clinical setting. To demonstrate the feasibility of this approach we have constructed amplicons, based on the herpes simplex virus, that co-express the ETR and different therapeutic or marker genes under the control of different constitutive promoters. *In vitro* we have shown that these amplicons drive the expression of multiple inserted genes and that the expression of the ETR is correlated with the expression of other genes contained within the same amplicon. Using these amplicons in a mouse xenograft brain tumor model we have demonstrated non-invasive MR imaging of *in vivo* transferred gene products. The results demonstrate the feasibility and power of this new approach; identification of targets overexpressed in different disease states will expand the potential application of this technology as a noninvasive diagnostic tool.

1. Hoegemann, D. *et al. Nature Med.* **6**, 351–355 (2000).

Barrett, Michael [17]

Transcriptional analysis of tetraploid epithelial cells in Barrett's esophagus

Michael Barrett[1], David Pritchard[2], Judy Anderson[2], Brian Reid[1] & Peter Rabinovitch[1,2]

[1]*Fred Hutchinson Cancer Research Center, Seattle, Washington, USA*
[2]*University of Washington, Seattle, Washington, USA*

Increased 4N (G2/tetraploid) fractions in patients with premalignant Barrett's esophagus are associated with loss of p53 function and are highly predictive of progression to aneuploidy and esophageal adenocarcinoma. Studies in Barrett's esophagus and model systems have suggested that tetraploid cells are unstable intermediates in the development of aneuploidy and cancer. We developed primary epithelial cultures from biopsies of patients with early stages of Barrett's esophagus. Each of four cultures had p53 lesions, p16 lesions or both that were also present *in vivo*, and each culture contained 12–25% 4N cells, of which approximately 50% were G1 tetraploid cells. To study the molecular phenotypes of these tetraploid cells, we sorted 2N and 4N cells from each culture viably using Hoechst 33342 labeling. We stored diploid cells while sorted 4N fractions were recultured for two weeks, then resorted to purify tetraploid populations. Total RNA from each sorted sample was used to interrogate Affymetrix FL6800 arrays. No genes were consistently underexpressed in 4N compared with 2N cells, whereas 32 genes (0.9%) showed consistently twofold increased expression in 4N cells; 24 of 32 were named genes and all had functions associated with G2/M and related checkpoints. These included genes associated with centrosome regulation (*nek-2, plK-1*), the G2/M transition (*CDC2, cyclin B, Wee1, CDC25*) and mitosis (genes for topoisomerase II, mitosin, CENP-A, kinesin-related protein). In contrast there were no changes in the expression of a series of G1/S-specific genes, including *RRM1, CDK4, PRAD1, CAD* and *MAPKKK5*. Tetraploid BE G1 cells in Barrett's esophagus have dysregulated expression of G2/M genes.

Baumbach-Reardon, Lisa [18]

Completed *BRCA1/BRCA2* mutation analysis reveals a low rate of germline mutation in at-risk African American families

Lisa Baumbach-Reardon[1], Luis Gayol[1], Tom Scholl[2], Hugo Basterrechea[1], Selina Smith[1] & J. Fernando Arena[1]

[1]*University of Miami School of Medicine, Miami, Florida, USA*
[2]*Myriad Genetics, Inc., Salt Lake City, Utah, USA*

The incidence of *BRCA1* germline mutations in at-risk individuals is controversial. In Caucasians the rate of detection of *BRCA1* mutations varies from 5% to 40%. The incidence of *BRCA1* mutations and genetic variants in at-risk African Americans has

been reported to range from very low to levels equal to those in Caucasians. We report the results of completed *BRCA1* and *BRCA2* analyses in 20 African American families at risk for breast or ovarian cancer. Families were selected on the basis of a history of breast cancer or breast and ovarian cancer and further subdivided into the following categories: high-risk (three affected first-degree relatives; ten families), moderate-risk (two affected first-degree relatives; seven families) and undetermined-risk (single affected first-degree relative, with medical information being updated). *BRCA1* and *BRCA2* germline alterations were first detected using a series of exon-specific polymerase chain reaction primers for single-strand conformation polymorphism analysis; this was followed by DNA sequencing of polymorphism variants. A limited number of *BRCA1* polymorphic intronic variants detected as a result of these studies were also analyzed for their effect on *BRCA1* mRNA splicing using an assay developed by Myriad Genetics. In this cohort we detected one protein-truncating mutation in either *BRCA1* or *BRCA2* (1/20; 5%). However, we detected splice mutations, missense mutations and several polymorphic variants in both *BRCA1* and *BRCA2,* with a much higher frequency in *BRCA2.* Many of these variants were both new and specific to African American patients; they also occurred in the absence of another disease-causing mutation. Moreover, a new *BRCA1* missense mutation in one of the high-risk families, exon 19(W1718C), seems to co-segregate with breast cancer. We will report the relative frequencies of these *BRCA1* and *BRCA2* variants in patient and control populations. These results agree with previous observations that deleterious mutations in *BRCA1* or *BRCA2* are uncommon in at-risk African American patients. They indicate that more benign variants should be further evaluated for their potential role in the disease process in these patients and that additional, as yet unidentified, genetic factors may contribute to breast cancer risk in African American families.

Baxevanis, Andreas D. [19]

An integrated approach to the extraction, storage, processing and analysis of microarray gene expression data

Andreas D. Baxevanis[1], Izabela Makalowska[1], Kenneth Trout[1], Qien Zhou[2], Zheng-Zheng Zhou[2], Jaime Stein[2], Edward R. Dougherty[3], Paul S. Meltzer[4], Yidong Chen[4], Michael L. Bittner[4] & Jeffrey M. Trent[4]

[1]*Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*
[2]*NuTec Services, Stafford, Texas, USA*
[3]*M.D. Anderson Cancer Center, University of Texas, Houston, Texas*
[4]*Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland*

A single microarray experiment provides thousands of individual pieces of data; it is therefore essential to focus on effective informatics methods in order to draw new biological conclusions efficiently and confidently. We have developed a unified software package that meets several important needs in the area of microarray data analysis. The software achieves several goals: (1) Integrating algorithms developed at the National Human Genome Research Institute into NuTec's GLEAMS software system, creating a single, integrated core platform that is capable of handling the computational aspects of complementary DNA and oligonucleotide array analysis, including imaging, signal processing, data extraction, database management and higher-order data analysis. (2) Developing new bioinformatics strategies to deduce the sequence- and structure-based relationships that are currently hidden within cDNA and oligonucleotide array data, including clustering methods and connections to gene ontologies. (3) Linking microarray data to data being amassed in clinical laboratories, to combine pedigree, phenotypic and genotypic data with gene expression data. (4) Using parallel computing to facilitate the analysis of microarray data. An overriding goal of this project is that the software be as user-friendly as possible, keeping in mind the level of computer training of most of the biologists who will be using it as well as the hardware that is typically available in most laboratories.

Beaulieu, Martin [20]

High-throughput genotyping using candidate region mismatch scanning (PCR-CRMS)

Martin Beaulieu[1], Ching Ouyang[1], Garry Larson[1], Louis Geller[1], Steven Flanagan[2] & Theodore Krontiris[1]

[1]*Division of Molecular Medicine, City of Hope National Medical Center, Duarte, California, USA*
[2]*Division of Neurosciences, City of Hope National Medical Center, Duarte, California, USA*

To identify loci affecting disease susceptibility, linkage and association analyses are performed by scoring previously characterized sequence variation, such as microsatellites and single-nucleotide polymorphisms. However, such analyses can be expensive and time-consuming. Moreover, examination of particular candidate gene regions for potential new linkages and associations may be limited by the infrequent occurrence of known markers. The typing by currently available methods of a sufficiently large number of markers to achieve a robust analysis is often beyond the reach of the typical research laboratory. Various techniques that score either unknown or known sequence variation have been developed, yet none of these strategies is equally well suited for mutation detection and the scoring of genotypes. To perform both tasks, a technique should be both sensitive enough to detect all mutation types and sufficiently quantitative that the translation of the data into allele-sharing status is feasible. We investigated the possibility of replacing conventional genotyping with a mutation-detection approach. Genome mismatch scanning is a hybridization-based mutation-detection technique employing the *Escherichia coli* mismatch-detection enzymes MutHLS. It was originally developed to enrich for identical-by-descent regions between two whole genomic DNA samples. To accommodate polymerase chain reaction products of candidate regions, we have conducted a comprehensive biochemical optimization of the technique for target sizes ranging from 260 to 1,250 base pairs. Our modifications, which we now collectively designate polymerase chain reaction–candidate region mismatch scanning, have simplified the assay, rendered it quantitative and demonstrated its potential for cost-effective, high-throughput genotyping. Strategies for exploiting the method in the study of candidate regions or in genome-wide studies will be discussed.

Belbin, Thomas [21]

Molecular classification of head and neck squamous cell carcinoma using cDNA microarrays

Thomas Belbin[1,2], Bhuvanesh Singh[3], Ilana Barber[2], Bruce Wenig[2], Richard Smith[2], Mike Prystowsky[2] & Geoffrey Childs[1]

[1]*Department of Molecular Genetics, Albert Einstein College of Medicine, New York, New York, USA*
[2]*Department of Pathology, Albert Einstein College of Medicine, New York, New York, USA*
[3]*Head and Neck Service, Memorial Sloan-Kettering Cancer Center, New York, New York, USA*

Head and neck squamous cell carcinoma (HNSCC) is the fifth most common cancer worldwide and accounts for 95% of the head and neck cancer cases in the Western world. Over the past three decades advancements in management have