MICROARRAY TECHNOLOGIES:

BENCH TO BEDSIDE

DNA microarray technology has matured to the point where some applications are deemed reliable enough for use in patient care. At the same time, microarrays are evolving to help expand the understanding of transcriptome complexity: single nucleotide polymorphisms, copy number variation, CpG methylations, microRNAs—so many genetic and epigenetic variations and a slew of microarrays to investigate each one. Furthermore, newer DNA sequencing technologies now threaten to do to microarray technology what automobiles did to the horse and buggy. For now, however, the two approaches appear to coexist happily, and microarrays remain a perfectly reasonable way to get around the genome.

By Emma Hitt

The first report on microarray technology appeared in *Science* barely more than a decade ago, and since its inception, microarray technology has been burdened with issues of reproducibility and standardization. Recently, however, the technology has received a vote of confidence in the form of US Food and Drug Administration approval, laying the groundwork for its broader application in a clinical setting. The newly approved test, the MammaPrint assay, was developed by **Agendia** in the Netherlands and measures the expression of 70 genes that predict the likelihood of breast cancer metastasis. Several such tests have already received the "CE mark" for clinical use in Europe, and a handful of tests that analyze DNA polymorphisms (as opposed to gene expression) have been available clinically in the United States since 2004, when **Roche's** AmpliChip CYP450 assay became the first to receive approval.

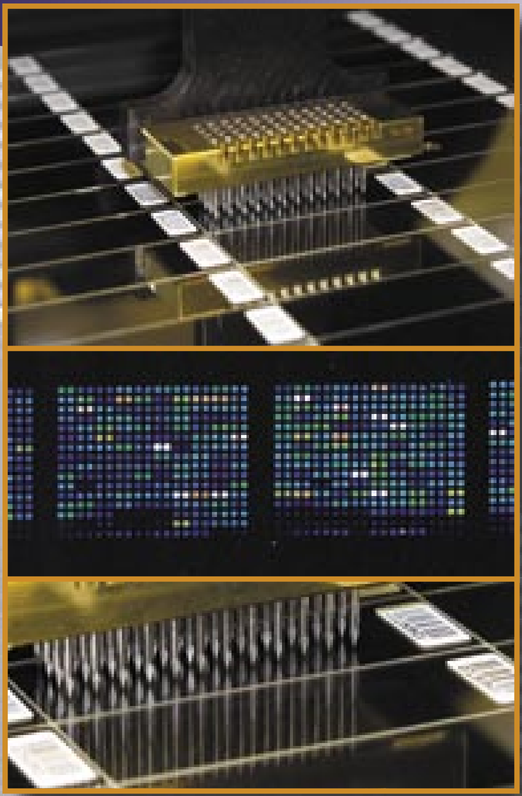
As is the case for most biological paradigms, the entire transcriptome within a given cell has turned out to be much more complicated than ever imagined. As a result, tools to study epigenetic variations and other regulatory mechanisms within the cell have proliferated. Microarrays can now measure changes in gene copy number (copy number variation, CNV), changes in the expression levels of microRNAs (miRNAs), contributions from gene splice variants, and changes in methylation patterns. "There are really a plethora of new assay types, and each of these is complementary to one another and informs from a different perspective," notes Kevin Meldrum, director of genomics marketing for **Agilent**.

DNA Sequencing Technologies

Perhaps one of the most exciting new developments in the microarray field is one that—paradoxically—may render the technology obsolete. DNA sequencing—the remarkably robust kind that has allowed a pioneer like James Watson to obtain his personal genome sequence on a DVD in about two months—is becoming the five-lane superhighway of genomic analysis. A 454 Life Sciences (now a part of Roche Diagnostics) sequencer was used to complete Watson's analysis, but several other companies now have their own sequencing approaches including **Illumina's** Solexa technology, **Applied Biosystems'** SOLiD System, **Helicos BioSciences'** True Single Molecule Sequencing (tSMS), and others.

"An interesting paradigm shift is taking place," says Steve Lombardi, chief operating officer and executive vice president with Helicos BioSciences. "You can do virtually any genetic analysis with sequencing technology—ranging from quantification of DNA, RNA, or cDNA, to measuring copy number variation and methylation," he says. "Microarrays, which require probes with known sequences, have limitations that simply do not exist with sequencing technology."

continued ›



> "There are really a plethora of new assay types, and each of these is complementary to one another and informs from a different perspective."

Look for these **Upcoming** Articles

Genomics: SNPs — October 19
RNAi: Therapeutic Applications — November 2
Protein-Protein Interaction — November 30

> “What this massively parallel approach to sequencing allows us to do is to generate about a hundred-fold more data for about 1 percent of the cost of current technology.”

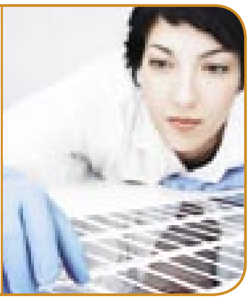


IMAGE: ILLUMINA, INC.

According to Lombardi, True Single Molecule Sequencing has the advantage of a simple sample preparation process that does not require amplification, and the image processing software and informatics will be open source allowing users to innovate.

A popular application of sequencing technology is to replace classic microarray chromatin immunoprecipitation (ChIP)-chip analyses. Microarrays to measure protein binding to DNA surfaced about five years ago and involve the placement of immunoprecipitated DNA fragments as probes on a DNA microarray. There is a wide range of products from Agilent, **Affymetrix,** Illumina, **Aviva Systems Biology, Nimblegen Systems,** and others for this purpose. With sequencing technology, the whole process becomes scaled up and digitized. First, binding proteins are cross-linked in place to DNA, the chromatin is isolated and fragmented, a protein-specific antibody is added, and the protein-DNA complexes are purified. The binding proteins are then chemically removed, and the remaining DNA is isolated. The DNA strands are then size selected, a process that isolates short sequences of, say, 25 bases—millions and millions of them. Finally, binding sequences are identified by virtue of the fact that their sequences are enriched several-fold relative to other sequences.

“What this massively parallel approach to sequencing allows us to do is to generate about a hundred-fold more data for about 1 percent of the cost of current technology,” notes Omead Ostadan vice president of marketing with Illumina. Some proponents suggest that the sequencing approach will ultimately replace microarrays. “Microarrays are most popular for performing gene expression and genotyping applications, and in some cases, for sequencing applications such as ChIP-chip analysis. Over time some of these applications may migrate over to a sequencing platform because of the added power and quality you get from sequencing. However, we see these two approaches as complementary, offering researchers the broadest solutions for studies,” says Todd Dickinson, director of product marketing at Illumina.

It may not be prudent to throw out the array chips just yet, however. “Sequencing technology is young, and issues regarding cost and ease of use still need to be worked out,” notes Steven Bodovitz, an industry analyst with **BioPerspectives**. “Nonetheless, sequencing is certainly more comprehensive than microarray technology, and getting a sequence without knowing what you are looking for ahead of time is a big advantage,” he says. “It will be interesting to see how the two approaches play out—the technologies are likely to both complement and compete with one another.”

Still, microarrays remain ideally suited for experiments that require analysis of large numbers of samples against known sequences—and will likely remain the predominant approach in diagnostic applications. “Not all companies are jumping on board with the sequencing approach,” says Bodovitz.

Optimizing Arrays

Of late, developments in gene expression arrays have focused on optimizing coverage of the entire human genome, which includes well over 30,000 genes, processed into hundreds of thousands of different variations. Newer arrays are aiming to capture that variability. Affymetrix’s GeneChip Human Gene 1.0 ST Array released this year measures the overall expression of all transcripts derived from a gene, in contrast to traditional arrays that measure only the 3´ end of a gene. “More than 60 percent of human genes undergo alternative splicing, resulting in multiple transcript variants with potentially distinct functions,” says Yan Zhang, associate director of product marketing with Affymetrix. Each of 28,869 genes is represented on the array by approximately 26 probes per gene directed at the full length of the gene. “We estimate that the accuracy is greater compared with the previous generation’s 3´ approach, and the format is smaller, which reduces cost without reducing content,” Zhang says.

Illumina’s whole-genome expression arrays employ the BeadArray technology, which uses 50-mer oligos attached to 3 µm beads, allowing six or eight samples to be profiled per BeadChip simultaneously. This multisample approach provides increased throughput and decreased cost per sample. Also allowing for higher throughput and lower costs, Nimblegen’s Human Whole Genome Expression Microarrays include 60-mer oligo probes for all human genes, eight per gene, including 47,633 different targets, and Agilent has a whole human genome chip containing 60-mer oligonucleotide probes arranged as microarrays in the format of 1 x 244K, 2 x 105K, 4 x 44K, or 8 x 15K on individual glass slides.

Microarray Manufacture

Arrays can be loosely classified into ready-made versions and custom “home-brew” forms that contain DNA spotted onto slides. Custom microarrays are fabricated largely by pin-based direct deposition printers.

UK-based company **Arrayjet** recently launched its benchtop Sprint Inkjet Microarrayer. The new machine can handle 20 microarray slides and allows printing from two microtiter plates (96- or 384-well) simultaneously.

“Due to their contact with the surface, pin printers can have problems with speed, quality, and reproducibility,” notes Howard Manning, Arrayjet founder and technical director. “With the ink jet approach, the Jetspyder enables the inkjet print head to draw in multiple samples before arraying spots are printed on the slide without the print head stopping and without making contact,” says Manning. Another advantage of inkjet printing is that spotting pins may damage some of the newer substrates used for arrays, such as membrane-coated substrates and hydrogels.

The processes utilized to manufacture arrays are also being enhanced. “Updates to the Agilent SurePrint technology were implemented last year,” Meldrum said, “enabling us to manufacture arrays with up to 244,000 unique features. We have also improved the efficiency of the nucleic acid synthesis process, signficantly reducing depurination side reactions.” **Nanogen** NanoChip arrays use a proprietary technology consisting of individually controllable microelectrodes that, when positively charged, can attract negatively charged DNA and RNA molecules. “The use of electronics to drive the concentration reaction mitigates the variability arising from spotting small volumes onto a surface,” says Nanogen’s Elaine Weidenhammer. “Consequently, the Nanogen system is subject to less variability than some other formats,” she claims.

SNPs and CNV

A recent publication from The ENCyclopedia Of DNA Elements (ENCODE), an international research consortium organized by the National Human Genome Research Institute (NHGRI) indicates that the DNA-makes-RNA-makes-protein paradigm is too simplistic and that 99 percent of the genome is not, in fact, junk.

continued ›

Gene transcripts appear to overlap, and genes, regulatory elements, and other types of DNA sequences interact as a network in a nonlinear fashion. New technologies are now evolving to define this complex variability.

This year, Affymetrix introduced the Genome-Wide Human SNP Array 6.0, which includes about 1.8 million markers that can be used to detect genetic variation in whole-genome association studies. Specifically, the newest SNP Array 6.0 contains 906,600 single nucleotide polymorphisms (SNPs) and about 946,000 nonpolymorphic probes that "will aim to detect CNVs in both known and unknown regions," says Jessica Tonani, Genotyping Specialist with Affymetrix. "The array tiles known CNVs very heavily, with an average of about 61 probes per known CNV region. For the remainder of the genome we have tiled probes at a resolution of about a probe every 700 bases."

"Our overall hope is to identify the part of the genome that has previously been considered a 'genomic wasteland,' or junk DNA, but may in fact be involved in regulating gene expression," she says, "although we are reaching a point where there is more to be gained by studying more samples as opposed to putting more SNPs on an array."

Newest to Illumina's Infinium portfolio is the Human 1M BeadChip, which provides expanded genomic coverage for whole-genome association studies. This chip adds several hundred thousand individually selected SNPs in genes and functional areas as well as CNV content not available on any other platform, according to Illumina. Tag SNPs represent a group of SNPs that tend to be inherited together. "By targeting tag SNPs, we can vastly increase the power of the study by looking at the genotype of one SNP, which effectively represents the genotype of many," Dickinson says. "The power of this array is increased dramatically because of this coverage and has reduced costs for researchers conducting genome-wide association studies," he adds. "For example, in 2000, genotyping a single SNP cost $1 or more. Today, the cost is less than $0.001 per SNP."

Featured Participants

Affymetrix
www.affymetrix.com

Agendia
www.agendia.com

Agilent
www.agilent.com

Almac Diagnostics
www.almacgroup.com/diagnostics

Applied Biosystems
www.appliedbiosystems.com

Arrayjet Ltd.
www.arrayjet.co.uk

Aviva Systems Biology
www.avivasysbio.com

BioMerieux
www.biomerieux.com

BioPerspectives
www.bioperspectives.com

Helicos BioSciences Corporation
www.helicosbio.com

Illumina
www.illumina.com

Kreatech Biotechnology
www.kreatech.com

MicroArray Quality Control (MAQC)
www.fda.gov/nctr/science/centers/toxicoinformatics/maqc

Minimal Information about Microarray Experiment (MIAME)
www.mged.org/Workgroups/MIAME/miame.html

Nanogen
www.nanogen.com

Nimblegen Systems
www.nimblegen.com

Pathwork Diagnostics
www.pathworkdx.com

Roche
www.roche.com

Veridex
www.veridex.com

Regulatory miRNA

MiRNAs have recently been shown to play an important role in regulating gene expression. Over 500 human miRNAs have already been identified and validated, and are thought to regulate about one-third of all human genes. Several companies now make kits for labeling and detection of miRNAs utilizing array platforms. Agilent recently introduced its miRNA assay which uses a direct labeling protocol on as little as 100 ng of total RNA. "Because miRNAs are typically 25–30 nucleotides in length and have high sequence homology, utilization of conventional probe design methods results in significant cross-hybridization of RNAs within the sample," notes Agilent's Meldrum, "but we have developed novel probe design algorithms that enable us to achieve extremely high specificity for each unique sequence."

Kreatech Biotechnology in the Netherlands recently introduced its miRacULS II Kit labeling technology, which uses a nonenzymatic labeling reaction independent of fragment size. The kit enables isolation and labeling of miRNAs from samples of less than 105 cells or 5 mg tissue as well as samples with a low miRNA expression level.

Ready for Prime Time?

Efforts are ongoing to standardize data input and reporting of microarray experimental data. **Minimal Information about Microarray Experiment (MAIME)** standards for microarray experiments have been put in place to define the minimal amount of information needed to interpret and reproduce a microarray experiment.

The first phase of the **MicroArray Quality Control** (MAQC) project, with participants from academia, government, and industry, recently concluded that results mostly were reproducible among labs; the second phase of the MAQC is under way and will try to establish the applicability of microarray data to clinical settings, with results expected in about a year. "When moving something from the bench to medical practice, there is an absolute requirement that the consistency of the performance of the test system can be guaranteed in different people's hands," says Uwe Scherf, with the FDA's Center for Devices and Radiological Health. "The MAQC studies were one of the first indications that microarrays could be suitable for medical evaluation."

The External RNA Controls Consortium (ERCC), established in 2003, is developing a set of external RNA spike-in controls for use in all microarray platforms and quantitative polymerase chain reaction (PCR) assays. If successful, this will allow a reliable comparison of expression profiling among more than 10 microarray and PCR platforms.

According to Affymetrix, at least 20 diagnostic tests are being developed for use on its platform—mostly expression-based tests evaluating tumors—and they are partnering with several companies, including Roche, **Almac Diagnostics**, **Veridex**, **BioMerieux**, and **Pathwork Diagnostics**, for the development of molecular diagnostic tests that use its platform. VeraCode, a new technology introduced by Illumina earlier this year, is available for researchers conducting low- and midplex biomarker validation studies. Using digital holographic codes embedded in cylindrical glass beads, VeraCode allows for efficient and cost-effective screening of hundreds to thousands of samples against panels of one to 384 biomarkers.

"It will be important to develop tools that can be utilized to analyze data sets for different applications and start to merge that information together," Agilent's Meldrum says. "Within the next few years the field is going to go from one-dimensional [e.g., gene expression profiling] to a multidimensional view, offering a richer picture of what's really happening in the cell."

Emma Hitt is a freelance medical and science writer living in Marietta, Georgia.

New Products

Multiplexed Assay System

The BeadXpress System is a high throughput, dual color laser detection system that enables scanning of a broad range of multiplexed assays developed using the VeraCode digital microbead technology. VeraCode technology features uniquely inscribed digital microbeads that provide high-quality data, broad multiplexing capability, and assay flexibility. The instrument scans the VeraCode microbeads for their code and fluorescent signals, generating data quickly and efficiently. Using the system, researchers can assay tens to hundreds of analytes in a single sample at one time. Application areas include biomarker research and validation, pharmaceutical development, industrial testing, agriculture, clinical research, and the development of molecular diagnostic assays. For example, BeadXpress technology can be used to develop screening assays that profile inherited and complex diseases. Drug discovery researchers can use it to validate RNA and protein biomarkers. Genetic researchers can use it to replicate single nucleotide polymorphism studies from large disease association studies to their own populations, for example, quickly retesting the top 50 or 100 polymorphisms in very large sample sizes. The system can scan assays ranging from a single-plex to 384-plex in a single well.

Illumina
For information 312-997-2436
www.illumina.com

DNA Methylation Arrays

NimbleGen DNA methylation analysis arrays allow researchers to rapidly identify methylated regions in a high throughput manner. These arrays can be used to correlate promoter and genic methylation with gene expression and phenotype and detect differential DNA methylation between normal tissues and tumor samples as potential diagnostic and prognostic markers. Whole-genome DNA methylation arrays are available for human, mouse, rat, dog, and chicken genomes, tiling at an average probe spacing of 100 bp to provide an unbiased approach to DNA methylation analysis. These arrays survey not only regions surrounding known genes but also intergenic regions, allowing distal regulatory elements to be surveyed and providing the maximum opportunity to discover relevant biomarkers. For studies focusing on promoter regions, one- and two-array sets are available for surveying promoter regions of all known genes in human, mouse, rat, and *Arabidopsis*.

NimbleGen Systems
For information 608-218-7623
www.nimblegen.com

Network Analysis Tools

The integration of Oncomine Professional and Ingenuity Pathways Analysis (IPA) will allow users to navigate directly from cancer gene signatures discovered in Oncomine Professional to the curated pathway information available from IPA. Oncomine Professional is a rapidly growing compendium of transcriptome profiles coupled with analysis functions and a web application for data mining and visualization. Comprised of thousands of profiles representing every major type of cancer, Oncomine Professional was developed to exploit these data for therapeutic target discovery, validation, and prioritization. IPA is software that enables biologists and bioinformaticians to identify the biological mechanisms, pathways, and functions most relevant to their experimental datasets or genes of interest. The integrated workflow based on IPA's capabilities for rapid and reliable biological interpretations of Oncomine's high-quality cancer transcriptome profiles is expected to provide a powerful tool for oncology researchers focused on understanding mechanisms of disease, and on the discovery and validation of cancer biomarkers.

Compendia Bioscience and Ingenuity Systems
For information 650-381-5150
www.compendiabio.com and www.ingenuity.com

Human Genomewide SpliceArray

The SpliceArray provides a unique approach to detecting all known and novel expressed transcripts. The Human GW SpliceArray monitors more than 23,000 human genes covering more than 130,000 expressed transcripts and requires more than 6 million probes to help elucidate the enormous diversity of the human proteome. This new design, manufactured on the Affymetrix platform, provides researchers the opportunity to obtain gene expression results at a resolution previously unattainable on a single microarray by interrogating both exons and splice junctions.

ExonHit Therapeutics
For information +33 1 58 05 47 00
www.exonhit.com

Dynamic Light Scattering Instrument

The DynaPro Plate Reader is a temperature controlled, high throughput dynamic light scattering (DLS) plate reader. It allows researchers to perform classical thermal DLS studies automatically on a large number of therapeutic macromolecules over a wide range of solution conditions. Conventional DLS studies are laborious and time consuming, with each solution condition studied manually, one at a time, over several hours. In contrast, the DynaPro controls the temperature for hundreds of samples simultaneously, capturing DLS data quickly, easily, and directly from standard microplates (96-well, 384-well, and 1,536-well formats) without perturbing precious samples. As little as 5 µl of sample per well is required in the 1,536-well format. The instrument includes an array of state-of-the-art technologies, including software controlled, Peltier-based temperature systems; compact digital attenuators; precision mechanical stages; proprietary multimode optical fibers; and novel software control and processing algorithms.

Wyatt Technology
For information 805-681-9009
www.wyatt.com

Microarray Data Analysis

TMA Foresight is a statistical tool for tissue microarray data analysis. Users can explore the relatedness of prognostic marker expression and clinicopathological variates with the outcome. Cox's Proportional Hazard analysis can be used for prognostic marker identification. Hierarchical clustering can be used to group patients on the basis of clinicopathological parameters and biomarkers. Kaplan-Meier plots and log-rank tests are used to identify prognostically significant clusters.

Premier Biosoft International
For information 650-856-2703
www.premierbiosoft.com

Correction

Life Science Technologies Feature:
"Cell Signaling: Phosporylation, Kinases, and Kinase Inhibitors" (6 April 2007, p. 125). The trademark for the homogeneous time-resolved fluorescence (HTRF) technology was incorrectly attributed to Millipore Corporation. The HTRF trademark is held by Cisbio International of Gif-sur-Yvette, France.