571
573
574
576
576

Microarrays: branching out from expression analysis

The main application of microarrays is gene expression analysis, but other uses for these arrays continue to grow in popularity. Laura Bonetta reports on the use of this technology to analyze other entities, from carbohydrates to tissue samples.

The term "microarray analysis" is almost synonymous with the study of gene expression. Since scientists at Stanford University first reported the use of DNA microarrays to simultaneously monitor the expression of 45 *Arabidopsis thaliana* genes¹, the technology has made its way into every area of biomedical research. Today, slides containing the entire complement of human and animal genes are hybridized to labeled targets representing transcripts from different cells and tissues to monitor global geneexpression profiles.

Such studies represent the main market for DNA microarrays, but other applications of the technology are available to researchers. They range from identifying differences in RNA and DNA sequences in order to type microorganisms, to finding gene regulatory sequences, to monitoring the expression of translation modifiers, such as microRNAs (miRNAs; **Box 1**). Additionally, many of the basic techniques developed for analyzing nucleic acids have been adapted to other molecules, such as proteins, carbohydrates, tissue samples and even living cells.

This article highlights several of the newer applications of both DNA and other types of microarrays, and the tools and services commercially available to researchers. "The way we think about all these different applications is to take the 1995 *Science* publication, and all that we have learned about expression analysis since then, as a starting point," says Mark Schena, who was first author on the seminal paper and current visiting scholar of ArrayIt Life Sciences Division at TeleChem International. "We then allow the scientific community to expand the applications in different areas."

Pathogen detectives

Traditionally, the detection and identification of microbes has relied on phenotypic traits, such as assaying the ability of microbes to grow on certain carbon sources. Today, these methods have been largely replaced by genotypic ones, such as DNA sequencing and PCR analysis, which provide greater discrimination and precision. Microarrays have taken the process a step further, providing greater ease and speed of detection. CombiMatrix recently launched the Influenza A Research Microarray to detect and type flu strains. The array contains DNA segments corresponding to different versions of the viral neuraminidase and hemagluttin genes, to discriminate between different strains and subtypes based on their gene sequences. The microarray also provides information on new strains of flu produced by rapid mutation or recombination between existing strains. The influenza array, which is available from

BOX 1 IT IS A SMALL RNA WORLD

miRNAs are a class of short endogenous RNAs that are important in cancer development and in tissue differentiation. They are found in the mature form as 19-25 nucleotide RNAs in the cytoplasm of cells, where they act as posttranscriptional regulators of gene expression by base-pairing with their target mRNAs. Exigon has developed an array system that uses nucleotide analogs, called locked nucleic acids (LNAs), for the large-scale investigation of miRNA expression. The incorporation of LNAs into an oligonucleotide probe increases the specificity of that probe for its complementary DNA or RNA target. For every LNA incorporated into an oligonucleotide, the melting temperature of the oligonucleotide bound to its target increases by 2-8 °C. Exiqon's miRCURY Arrays are designed so that the capture probes on the slides contain optimal amounts of LNAs to produce an equal hybridization affinity for all their miRNA targets. Thus, these arrays provide sensitive miRNA profiles, requiring smaller amounts of sample and less handling than would be needed for most DNA-based miRNA arrays. "Most other platforms require that you do an miRNA enrichment step before the hybridization, but with our arrays you can take the total RNA fraction from a sample, label it and then hybridize directly to our arrays," says Peter Roberts, brand manager at Exigon. On May 19, 2006 Exigon announced the launch of its miRNA profiling service, promising fast turnaround times for results of 2-3 weeks.

The capture oligonucleotides on Exiqon arrays are based on the miRBase sequence database (http://microrna.sanger.ac.uk). "As miRBASE updates, we update our chips," says Roberts. Ambion linked with Rosetta Genomics to access Rosetta's proprietary miRNA sequence database to add to its current range of microarrays targeting human, mouse and rat miRNAs. According to the company, Ambion's mirVana miRNA probe sets have a polynucleotide spacer that ensures maximal hybridization efficiency to labeled mRNAs. Other companies that provide miRNA microarray products or services include Agilent Technologies, which launched their line of products this year, LC Sciences and CombiMatrix.

NATURE METHODS | VOL.3 NO.7 | JULY 2006 | 571



The ElectraSense Reader from CombiMatrix. (Courtesy of CombiMatrix.)

the company both as a tool and a service, is one of several pathogen-detection arrays produced using CombiMatrix's *in situ* oligonucleotide synthesis method. The company sells the CustomArray 4X2K, which contains four arrays on one slide, with each array having 2,240 sites for probe synthesis; the CustomArrays 12K and 90K have 12,000 and 90,000 such sites, respectively. "These arrays allow for very rapid analysis of the type of disease for any RNA virus," says Michael Tognotti, vice president for sales and marketing. "Customers can order the actual chips from us, or order the entire synthesis system so that they can make their own arrays." The technology for making the arrays uses "a porous surface that the DNA is welded on," says Tognotti, allowing multiple uses of the same array. "If you try to reuse a glass slide, the intensity will drop off. With our system, you can wash bound targets off the array very aggressively and use the array again, up to four times."

For several years, researchers have been using Affymetrix resequencing arrays, which allow the detections of single-base differences in the sequence of specific genes, to identify and type pathogens. In a recent example, David Stenger's group at the Naval Research Laboratory, and other members of the Epidemic Outbreak Surveillance Consortium, used a custom-designed Respiratory Pathogen Microarray (RPM v.1) to identify a wide range of pathogens². The array contains probes to detect human



BOX 2 GOING WITH THE FLOW

DNA arrays are traditionally printed on glass or glass-like slides, hybridized to a fluorescently labeled target, and the slide is read by a scanner that can detect multiple fluorescent signals. An alternative method for making arrays is the Continuous Flow Microspotter developed by scientists at the University of Utah and sold by Wasatch Microfluidics. It uses a parallel network of microchannels to cycle small volumes of fluid samples over a surface, binding the samples to the surface in an array of spots. The spots are sealed at the surface so they are physically protected from one another. "You can start with a very dilute solution of the material you want to spot on the array and run it back and forth on the surface until you have the concentration you want," says Bruce K. Gale of the University of Utah. "Variation between spots is less than 5%," he adds. One of the company's arrays, which became available last month, can hold 48 spots. "The technology works especially well for proteins that don't bind well to surfaces and denature when they dry out. The device keeps them wet until they are ready to use," says Gale. The platform has been used for spotting proteins, DNA, cells and lipids by changing the coating on the slides.

MetriGenix has also developed a microarray device based on a three-dimensional matrix of microchannels. Similarly to Biacore's platform, MetriGenix Flow-Though Chip allows molecular interactions between probes and targets to occur in flow. Microchannels connect the upper and lower faces of the chip. As fluids containing different samples pass through these channels, binding reagents attach to the walls of channels, and target molecules are then captured in a fully automated process. Multiple spots are deposited in a regular grid to permit parallel analysis of up to 400 spots per sample. Because the binding reagents cover a wider surface area than they would on more traditional two-dimensional microarrays, binding occurs faster, resulting in shorter assay times. The technology has been used for protein arrays and pathogen detection. "We sell array systems, which include the flowthrough chip and all the necessary reagents for hybridization," says Michael L. Cohen, CEO for MetriGenix. "The researchers just have to add their sample."

At the detection end, CombiMatrix has developed an electrochemical detection method as an alternative to traditional fluorescence-based platforms. "We have developed a smaller less expensive detection system that can measure 12,000 features on a microarray using electrochemical signals," says Tognotti. "The whole machine is the size of a toaster and can be taken out into the field." According to Tognotti the assay is slightly more sensitive than fluorescence detection and does not produce any images. "The data comes straight out of the instrument as electronic signals. You can get very quick answers in 40 seconds," he explains.

adenoviruses, several subtypes of influenza A virus, 12 other common respiratory pathogens, and 6 US Centers for Disease Control and Prevention category A bioterrorism pathogens known to cause flulike symptoms. Two years ago, Affymetrix started working on developing a microarray capable of detecting hundreds of different bacteria and viruses from the US National Institute of Allergy and Infectious Diseases high-priority pathogen list.

Where transcription factors go

Gene expression analysis provides only a partial glimpse into how transcription is regulated. Additional information can be gained by determining the precise locations the binding sites for regulatory proteins on genomic DNA. Chromatin immunoprecipitation (ChIP), an assay used to detect interactions between proteins and DNA, has been combined with DNA microarray technology in a method dubbed ChIP-on-chip or location analysis. In this technique, a protein is cross-linked to DNA upon binding. The protein is then used as a tag to pull out the bound DNA by means of an antibody. The DNA is then eluted, labeled and hybridized to a genomic tiling array, which consists of DNA fragments spaced at regular intervals across the entire genome. The DNA fragment that 'lights up' identifies the genomic position of the protein binding site. This type of array provides key information not only about transcriptional start sites, but also DNA replication, modification and repair.

NimbleGen was the first company to offer a microarray service for custom ChIP-onchip assays. Additionally, NimbleGen offers "both *de novo* tiling and targeted arrays for specific regions of the genome," says Emile F. Nuwaysir, vice president for business development. "Our platform is the only high-density long-oligonucleotide arrays for covering up to 40 megabases of genomic DNA per array. It gives a maximum number of data points."

One of the most recent validations of ChIP-on-chip was performed by Richard Young's group at the Whitehead Institute for Biotechnology. The study identified DNA segments bound by a master regulator of human embryonic stem cells differentiation³. In January 2005 Agilent Technologies purchased Computational Biology Corp.,



The Biacore T100 system. (Courtesy of Biacore.)

BOX 3 A CONSORTIUM FOR GLYCOMICS

The Consortium for Functional Genomics is an initiative funded by the US National Institute of General Medical Sciences at the National Institutes of Health to understand the functions of carbohydrate-protein interactions. First funded in 2000, the effort is halfway through its planned 10-year lifespan. Two of the consortium's cores are charged with building a glycan library and printing the microarrays. Another core conducts the analyses, explains Richard Alvarez, assistant professor in the department of biochemistry at the University of Oklahoma Health Sciences Center in Oklahoma City, who directs the protein-carbohydrate interaction core.

The teams' first-generation systems consisted of microtiter plates containing bound carbohydrates. When a protein in a test sample bound to one of the carbohydrates, one of the wells would give off a fluorescent signal. The microtiter plates "worked remarkably well" but the system was difficult to scale up, explains Alvarez. For this reason, about two and a half years ago, they adopted a slide-based system, similar in principle to traditional DNA microarrays. Currently, each slide accommodates 260 unique glycan structures and 5 glycoproteins printed in replicates of six and at two concentrations. The system uses all of the same supporting technologies as gene arrays.

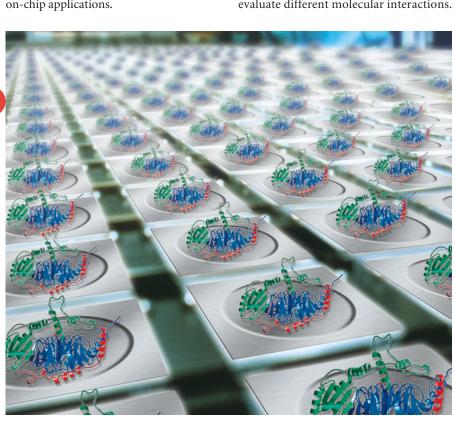
One of the biggest bottlenecks in making carbohydrate microarrays is synthesizing the actual carbohydrates, a laborious and expensive process, out of the reach of most academic labs. Additionally, one of the limitations of the technique is that there is no way of knowing how much sample is bound at each spot. "The carbohydrates are printed at the same concentrations but different sugars may have different coupling efficiencies," says Alvarez.

Researchers interested in screening a target for glycanbinding can use the system at no charge, provided that the data is made public in the consortium's database. Last year alone the consortium received 100 such requests from researchers. "Anything that anyone believes has putative carbohydrate binding sites is open to screening," says Alvarez, who has conducted studies using proteins, bacteria and viruses.

the start-up company founded by Young and colleagues to develop the ChIP-on-chip technology. Today, Agilent offers several 60mer oligonucleotide microarrays, as well as reagents and microarray hardware and software, all designed specifically for ChIPon-chip applications. As an alternative to the ChIP-on-chip assay, microarrays can be used to monitor the direct binding of proteins to DNA fragments, bypassing the immunoprecipitation step. Since 1990, Biacore has been supplying instruments and software to detect and to evaluate different molecular interactions.

50

© 2006 Nature Publishing Group http://www.nature.com/naturemethods



Qproteome Glycoarray kit for assaying glycosylation. (Courtesy of Qiagen.)

In their systems, one type of molecule is immobilized in flow cells onto a sensor surface equipped with the appropriate 'chemical handle'. Interacting partners are delivered to the flow cells in solution through a microfluidics system (**Box 2**). The sensor detects interactions in real time and without the use of labels.

Biacore sells a variety of systems optimized for different applications, including the Biacore T100 system, which allows the recovery and characterization of binding partners. "Our systems provide a range of information about molecular interactions: 'Does a partner bind or not bind? How fast and how stable is the binding? How quickly does it dissociate?'" says Helena Nilshans, marketing director at Biacore. "Two array– based systems, the Biacore A100 and the Flexchip, can be used to screen for interactions involving proteins, nucleic acids, small molecules, carbohydrates as well as cells and viruses, for a wide range of applications."

Hitting the carbs

Besides DNA, other types of molecules can be spotted onto slides. Among them, carbohydrates are the new kids on the block. Carbohydrates are a large group of compounds consisting of sugar units, and include monosaccharides, disaccharides, oligosaccharides and polysaccharides (or glycans). The way they interact with proteins allows carbohydrates to have important roles in a variety of biological events.

BOX 4 TISSUE-ARRAY SOFTWARE

The scientist who first developed tissue microarrays as a postdoc at the US National Institutes of Health, has now turned his attention to software development. "We produce intelligent pattern recognition software to recognize different cellular features and quantify the staining," says Juha Kononen, chief scientific officer at Beecher Instruments. Normally tissue-microarray analysis requires manual scoring of stained slides by microscopy. Beecher's software offers modules that recognize different cellular compartments and can determine, for example, whether the staining is in the nucleus or cytoplasm. Other modules can measure the fraction of positively stained cells in a sample or the number of tumor versus normal cells. "For a typical experiment the output file will look at about 20 different parameters for each assay," says Kononen.

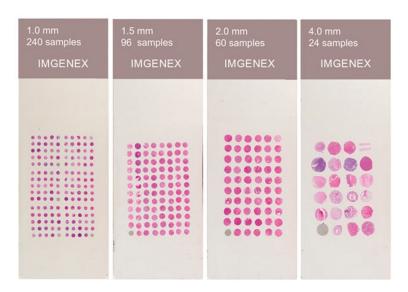
The software, says Kononen, does not replace what a pathologist would do, but "takes out the more repetitive aspects of manually looking at the spots." So far the focus for the company has been to provide software tools for the analysis of cancer tissues, but it is moving toward other approaches. "We are going to cover noncancer diseases and normal tissue structures," says Kononen. Another area of development is data integration. "We want to develop software that explores and links observations, and can combine data from tissue microarrays with data from other microarray platforms," he adds.

© 2006 Nature Publishing Group http://www.nature.com/naturemethods

For example, the binding of proteins on the membrane of a human or bacterial cell to a specific carbohydrate on the surface of a target tissue can initiate an immune response or allows bacteria or viruses to infect the tissue. Thus, there is growing interest among researchers and pharmaceutical companies to delineate proteincarbohydrate interactions.

To aid such studies, Glycominds produces the GlycoChip, the first commercially available microarray of simple and com-

plex carbohydrates. To detect interacting proteins, carbohydrates are constructed throughout the chip surface and incubated with a mixture of labeled proteins or cells. The company's Glycochip Master Analyzer software facilitates the interpretation of the resulting data. According to Rom Altstock, vice president for research and development, the unique feature about the company's platform is the manner in which the carbohydrates are bound to the surface of the slide. "The sugars are presented in a way that they can be appropriately accessed by the target protein or cell," he explains. "We spent a lot of time and effort on determining how best to present both short and long



A variety of tissue microarrays from Imgenex. (Courtesy of Imgenex.)

complex structures." Some of the smaller sugar molecules, the mono- and disaccharides, don't have large three-dimentional structures. "Thus, they do not protrude well above the surface of the slide. Trying to fashion them in a way that large molecules will bind to them is not trivial," explains Altstock. Glycan arrays are also available from several consortia (**Box 3**).

QIAGEN recently launched a different product, the Qproteome Glycoarray Kit, for characterizing proteins that are modified by the addition of carbohydrate moieties (glycosylation), which is an important mechanism for changing protein conformation, clearance rate and activity. "Some proteins are only biologically functional if glycosylated, so you can investigate the functional status of proteins by determining their glycosylation status," says Kerstin Steinert, director for research and development at Qiagen. "By using an array of spotted lectins and denconvoluting their binding profile, we can simultaneously obtain qualitative and semiquantitative analysis of glycosylation." Normally, identifying and characterizing glycosylation on proteins involves a multistep procedure, which includes enzymatic modification, liquid chromatography and mass spectrometry. This process is greatly simplified by using the glycoarray. Each array is spotted with more than 20 proteins, called lectins, which specifically bind to different monosaccharide motifs. Because these lectins have overlapping specificities, the array and associated software provide a detailed profile of glycosylation of the proteins in a sample.

Of cells and tissues

Gene expression provides information about which genes are expressed, but another type of microarray provides information about the cells, and even the cellular compartments, expressing these genes. "If you take a small specimen from a tissue and produce a homogenate for gene-expression analysis, you can get completely different results from assay to assay due to sampling variation," says Joseph P. Brown, CEO of LifeSpan Biosciences, a company that offers tissue-microarray services.

Since their first description⁴ in 1998, the use and availability of tissue microarrays has skyrocketed. The arrays are slides containing numerous small tissue sections of various types. The diameter of each spot section can range from 0.6 to 2 mm, and the number of spots from tens to several

hundreds. These spots provide every feature of whole tissue sections, allowing the detection of protein, RNA and DNA, by immunohistochemistry and fluorescence *in situ* hybridization, and can be analyzed manually or with the help of software (**Box 4**). "The advantage of the arrays is that you can look at multiple specimens on the same slide under the same conditions," says Brown.

Although several companies sell tissue microarrays, not all of them are suited for every application. "One of the critical issues is the quality of the tissues used," says Brown. "Some applications require having specimens with associated clinical information or, in some cases, having sufficiently large pieces of tissue." For some tissues it may be necessary to have larger spots (and, consequently, fewer of them per slide), to have a good representation of all the cells in that tissue in every spot. With solid tumors, spots can typically be smaller, with a single slide fitting up to 60 spots.

"What drives all microarray technologies

is parallelism. If you wanted to compare 100 patients, you had to compare 100 different slides processed differently," says ArrayIt's Mark Schena. "Now you can look at them all in one or two slides" In addition to a variety of protein microarrays, ArrayIt sells the AccuMax microarray slides, containing 10–300 elements of various tissue types. The company also provides a wide range of services to its customers. "Some customers send us plates of samples; we print those out without even knowing what they are," he says.

Over 50 different human or animal tissue array slides are also available from Imgenex, with tissue-sample densities ranging from 24 to 146 spots per slide. Each slide includes different cancer or adjacent normal tissues from different organs, as well as patient data on all tissue samples, including age, sex, diagnosis and any applicable staging codes. "The availability of normal adjacent tissues enables researchers to study the tumor microenvironment, which is increasingly recognized as playing a key role in tumorogenesis and chemotherapeutic responses," says Dan Jones, manager for sales and marketing.

Images of ordered arrays of colorful red, green and yellow dots have become symbolic of biomedical science in the postgenomics era. Although microarrays continue to be most widely used to study gene expression, a variety of new applications are making their way into the market place. As these continue to mature, and as researchers increasingly turn their attention from nucleic acids to proteins and cells, gene expression arrays may eventually lose their front-row seats.

Laura Bonetta is a freelance writer based in the Washington, DC area (lbonetta@nasw.org).

- 1. Schena, M., Shalon, D., Davis, R.W. & Brown,
- P.O. Science **270**, 467–470 (1995).
- 2. Lin, B. et al. Genome Res. 16, 527–535 (2006).
- 3. Lee, T.I. *et al. Cell* **125**, 301–313 (2006).
- 4. Kononen, J. et al. Nat. Med. 4, 844-847 (1998).

SUPPLIERS GUIDE: VENDORS OF MICROARRAY PRODUCTS AND SERVICES

Company	Web address	Company	Web address
Abbott Molecular	http://www.vysis.com	Imstar	http://www.imstar.fr
Affymetrix	http://www.affymetrix.com	Invitrogen	http://www.invitrogen.com
Agilent	http://www.agilent.com	Illumina	http://www.illumina.com
Alpha Innotech	http://www.alphainnotech.com	JPT Peptide Technologies	http://www.jpt.com
Alphelys	http://www.alphelys.com	LC Sciences	http://www.lcsciences.com
Ambion	http://www.ambion.com	LifeSpan Biosciences	http://www.lsbio.com
Applied Biosystems	http://www.appliedbiosystems.com	Luminex Corporation	http://www.luminexcorp.com
Applied Precision	http://www.api.com	MetriGenix	http://www.metrigenix.com
ArrayJet	http://www.arrayjet.co.uk	MiraiBio	http://www.miraibio.com
Asterand	http://www.asterand.com	Molecular Devices	http://www.moleculardevices.com
BD Biosciences	http://www.bdbiosciences.com	MOGene	http://www.mogene.com
Beecher Instruments	http://www.beecherinstruments.com	MTR Scientific	http://www.mtrscientific.com
Biacore	http://www.biacore.com	Nanogen	http://www.nanogen.com
BioAware	http://www.bio-aware.com	New England Peptide	http://www.newenglandpeptide.com
BioCat	http://www.biocat.de	NimbleGen Systems	http://www.nimblegen.com
BioChain	http://www.biochain.com	Oxford Gene Technology	http://www.ogt.co.uk
BioDiscovery	http://www.biodiscovery.com	Panomics	http://www.panomics.com
Bioforce Nanosciences	http://www.bioforcenano.com	PerkinElmer Life Sciences	http://las.perkinelmer.com
BioMicro	http://www.biomicro.com	Perlegen Sciences	http://www.perlegen.com
Bio-Rad	http://www.bio-rad.com	Premier Biosoft International	http://www.premierbiosoft.com
Clondiag	http://www.clondiag.com	Qiagen	http://www1.qiagen.com
CombiMatrix	http://www.combimatrix.com	Rosetta Genomics	http://www.rosettagenomics.com
DNAStar	http://www.dnastar.com	Scanalytics	http://www.scanalytics.com
Eppendorf Array Technologies	http://www.eppendorf.com	Schleicher & Schuell	http://www.schleicher-schuell.com
Eurogentec	http://www.eurogentec.be	Sigma-Aldrich	http://www.sigmaaldrich.com
Exiqon	http://www.exiqon.com	Strand Life Sciences	http://www.strandgenomics.com
Folio Bioscience	http://www.foliobio.com	Stratagene	http://www.stratagene.com
GE Healthcare	http://www.amershambiosciences.com	SuperBioChips Laboratories	http://www.tissue-array.com
Genaco	http://gene.genaco.com	SuperArray Bioscience	http://www.superarray.com
Genetix	http://www.genetix.com	TeleChem International	http://www.arrayit.com
Genomic Solutions	http://www.genomicsolutions.com	TriStar Technology Group	http://www.tristargroup.us
Gentaur	http://www.gentaur.com	US Biomax	http://www.biomax.us
Glycominds	http://www.glycominds.com	Wasatch Microfluidics	http://www.microfl.com
Icoria	http://www.icoria.com	Zeptosens	http://www.zeptosens.com
Imgenex	http://www.imgenex.com		

npg