Glass slides to DNA microarrays

by Samuel D. Conzone* and Carlo G. Pantano†

A tremendous interest in deoxyribonucleic acid (DNA) characterization tools was spurred by the mapping and sequencing of the human genome. New tools were needed, beginning in the early 1990s, to cope with the unprecedented amount of genomic information that was being discovered. Such needs led to the development of DNA microarrays; tiny gene-based sensors traditionally prepared on coated glass microscope slides. The following review is intended to provide historical insight into the advent of the DNA microarray, followed by a description of the technology from both the application and fabrication points of view. Finally, the unmet challenges and needs associated with DNA microarrays will be described to define areas of potential future developments for the materials researcher.

*Director R&D-Schott Nexterion AG,
Ma/086, Otto-Schott-Str. 2,
D-55127 Mainz, Germany
E-mail: sam.conzone@schott.com

†Distinguished Professor of Materials Science and Engineering and Director, Materials Research Institute,
The Pennsylvania State University,
199 Materials Research Institute Building,
University Park, PA 16802-6809 USA
E-mail: cgp1@psu.edu

Most individuals, outside of academic circles focused on genomics, became aware of the potential commercial, technical, and social importance of the human genome project during the late 1990s. The human genome project was formally initiated in 1990[1] and was expected to last 15 years. It had the major goals of identifying all the genes in human DNA, determining the sequences of those genes, and storing the information in public databases. However, the project moved quickly from the onset and, by 1998, the Department of Energy (DOE) and the National Institutes of Health (NIH) predicted that the human genome project would be completed by 2003.

The big buzz about biotech

The tremendous success in rapidly mapping and sequencing the human genome (a working draft sequence of the human genome was completed in 2000), has led many commentators to predict that similar achievements would follow on the applications side, giving rise to unprecedented discoveries related to human health[2,3]. Gaudy promises of high-tech clinics with the ability to prescribe drugs based on the genetic make-up of the patient were well ahead of their time. This normal lag from discovery (the sequenced human genome) to true applications (genetically engineered drugs) is partially attributable to the lack of tools, which could enable researchers to utilize effectively the tremendous amount of information that was generated during the human genome project.

ISSN:1369 7021 © Elsevier Ltd 2004

The broad goal of the human genome project was to create a library of knowledge (mapped, sequenced genes) that would allow researchers to take an active, discovery-focused approach to understanding human physiology at the molecular level. This new, more scientific approach should ideally replace and/or provide synergies to the trial and error methods that are presently used for drug discovery in high-throughput laboratories[4]. The 30 000+ genes that comprise the human genome are the starting point for scientifically understanding physiology and, hence, disease proliferation.

The example in Fig. 1 shows the important role that genes play in human physiology. Some external stimulus (disease, air pollution, overexposure to sunlight, medication, ingestion of a toxic substance, etc.) induces a genetic response known as gene transcription. During transcription, those genes (sections of human DNA) affected by the stimulus create messengers (mRNA), which produce proteins through a process called translation. These translated proteins are responsible for carrying out the physiological response to the external stimulus. As DNA transcription is the first process in the sequence of events that comprises human physiology, it is important for drug development researchers to understand three key issues. Firstly, what genes are present in the human genome? Secondly, how does gene activity fluctuate? Thirdly, what is the function of those genes with fluctuating activity?

The human genome project has led to the public dissemination of sequenced human DNA, answering the first question. The library of data has been built, and the medical discoveries of the future will now depend on how this information is used to characterize gene activity and function. Proper use of information starts when scientists are given the equipment needed to succeed. A simple analogy can be drawn for the materials researcher, using X-ray diffraction (XRD) as an example. The technology for probing the crystal structure of materials by XRD was invented after the turn of the century and used to create a database of knowledge about the structure of those materials. Although much of this knowledge was available in the 1940s, it was not until the late 1960s that XRD tools became commonplace in research laboratories for powder and related diffraction methods. Thanks to this technique, materials science took a new form, leading to discoveries such as semiconductors, ferroelectrics, glass ceramics, high-purity single crystals, high-temperature superconductors, liquid crystals, superalloys, and intermetallics.

Likewise, the world of biotechnology is now poised to take advantage of the great depth of knowledge provided by the human genome project using the DNA microarray as an integral tool. Thousands of researchers, hundreds of biotechnology companies, and all major pharmaceutical companies are using DNA microarrays to characterize the genetic state (data from tens of thousands of genes) of a biological specimen (human, rat, mouse, etc.) with a single miniaturized test, in an attempt to link gene activity and function to human health.

Microarrays: what are they and how are they applied?

Prior to the human genome project, the library of known, sequenced genes was small and the characterization of gene

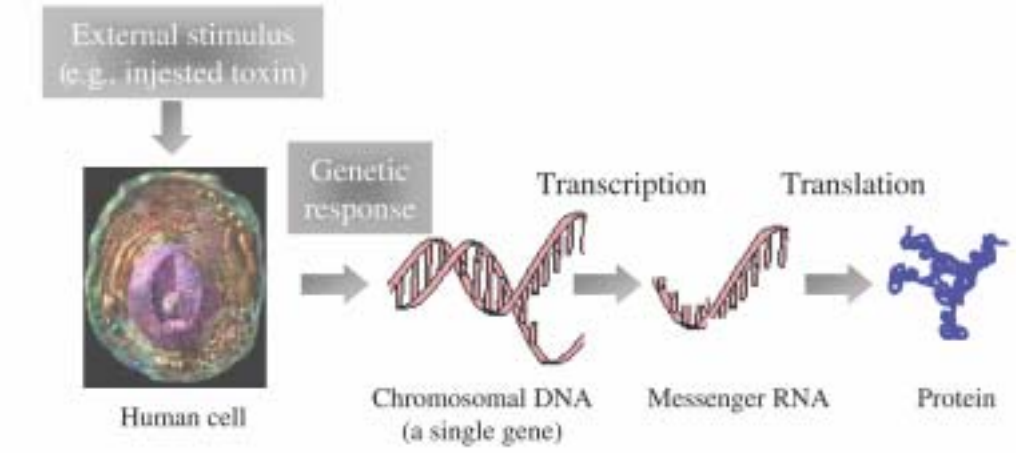

Fig. 1 Pictorial representation of the important role that genes play in human physiology. Gene transcription leads to the formation of messenger RNAs, which then translate to form functional proteins.

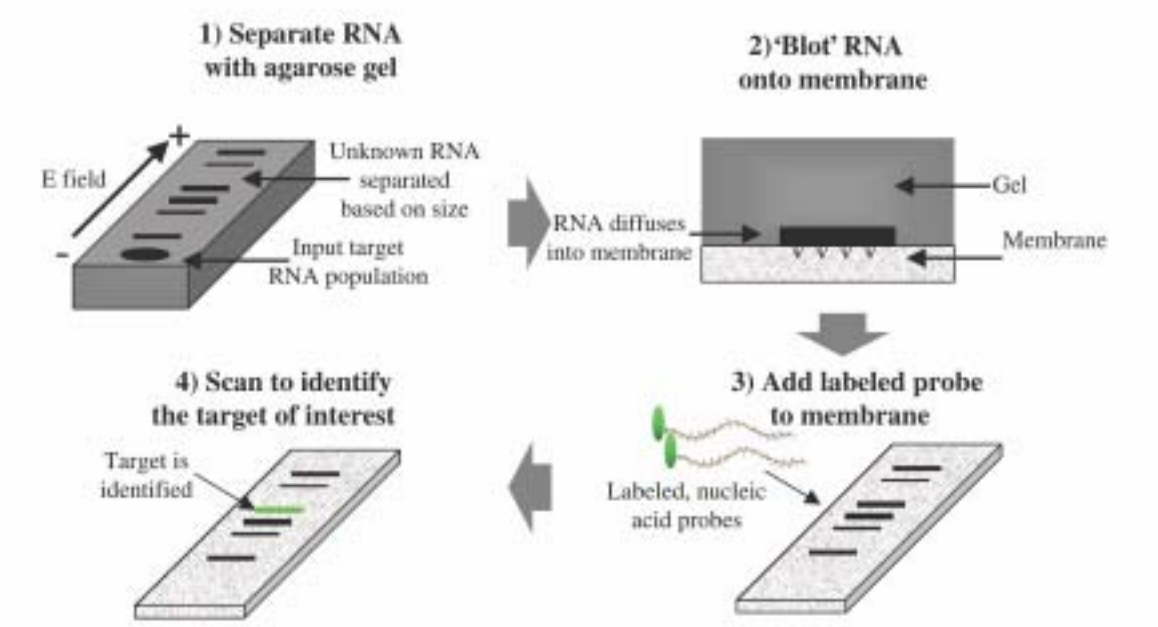

Fig. 2 Pictorial representation of the Northern blotting process, which enables single RNA target characterization.

activity was typically conducted one gene at a time. Northern blotting is a technique developed in the 1970s[5,6], which enables researchers to characterize gene activity based on RNA expression (Fig. 1) in a complex genomic mixture, such as a cell extract. Northern blotting consists of four primary steps (Fig. 2), whereby a complex mixture of RNA (targets) is placed on an agarose gel and separated (in terms of size) by electrophoresis. The separated RNA target population is then transferred and immobilized on a nylon or nitrocellulose membrane. The unknown RNA targets on the membrane then interact with a labeled probe (fluorescent or radioactive labeling), which is comprised of a known genetic sequence that only interacts with one single target of interest within the complex RNA mixture. After hybridization, the 'stained' membrane is scanned to determine (i) whether the target is present in the complex mixture and (ii) the relative abundance of that target. This method provides researchers with a tool for characterizing gene activity, albeit very slowly, since only one gene is typically probed per experiment, and human DNA is comprised of over 30 000 genes.

A higher throughput tool is needed to assess global gene activity (expression) simultaneously and it was this need that led to the development of DNA microarrays in the early 1990s[7-9]. A DNA microarray consists of a flat, solid substrate (typically glass) with an organic coating, typically an organo-functional alkoxysilane. The coated glass is then grafted (by printing or *in situ* synthesis) with various known DNA probes at predefined locations, as shown in Fig. 3. The DNA microarray can be thought of as a glass-based, biological sensor that can contain over 30 000 distinct, known probes at prespecified locations. This powerful, glass-based, biological sensor provides researchers with the ability to characterize the human genomic state fully with a single, miniature experiment.

A typical functional genomic experiment may be based on the need to understand the difference in gene activity between healthy and cancerous tissues, as shown in Fig. 4. During such an experiment, complex RNA populations are first extracted and separated from healthy and cancerous tissues using laser microdissection and conventional cell

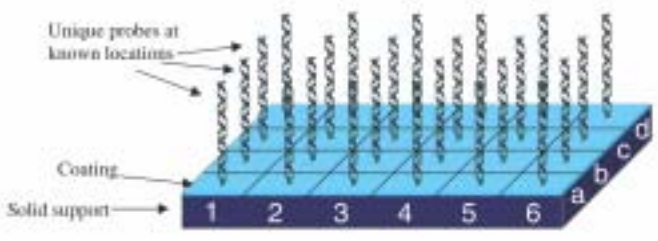

Fig. 3 Simplified pictorial representation of a spotted, DNA microarray. Unique nucleic acid probes are immobilized at distinct locations (i.e. position 1-a) on a coated support. Note that the image is an oversimplified example, as a DNA microarray can have a much higher density (>500 probes/cm^2) and many copies of a unique probe are found at each location.

lysing/RNA isolation processes[10]. The two separate RNA populations (target populations) are then labeled with dissimilar fluorescent dyes (i.e. the healthy targets can be labeled with a green dye, while the cancerous targets are labeled with red)[11]. After labeling, the healthy and cancerous target populations are combined and applied to the surface of a DNA microarray. This step of introducing unknown, labeled targets onto a DNA microarray initiates the hybridization process, whereby the labeled targets begin to form strong hybrids with complementary probes that are located on the array surface[8]. Hybridization occurs when the sequences of the target and probe are complementary, resulting in strong hydrogen bonding between the target and probe. Noncomplementary targets and probes do not form such strong hybrids and are subsequently removed from the array surface using stringent washing.

A scanning process is conducted after hybridization and washing to image the red and green dyes simultaneously, which represent labeled targets that have hybridized with their probe complements on the microarray surface (Fig. 4). After scanning, the image can be analyzed quantitatively to characterize the differences in gene activity corresponding to the healthy and cancerous tissues. Spots fluorescing strongly in the green or red channels are indicative of strong gene activity in the healthy or cancerous tissues, respectively, while spots exhibiting a yellow appearance indicate minimal difference in gene activity between the healthy and cancerous tissues (scanned image in Fig. 4).

This example shows the true utility of a DNA microarray for the characterization of global gene activity with a single experiment. This powerful technology is now allowing researchers to link gene activity to a disease state at a pace never before possible. Furthermore, scientists are now combining the principle of microarrays with novel protein applications and traditional histologic examinations (which is beyond the scope of this review to detail) to correlate gene activity with function. Thus, microarrays are providing researchers with one of the first key tools needed to realize the promise initially provided by the sequencing of the human genome. In the next section, the microarray will be

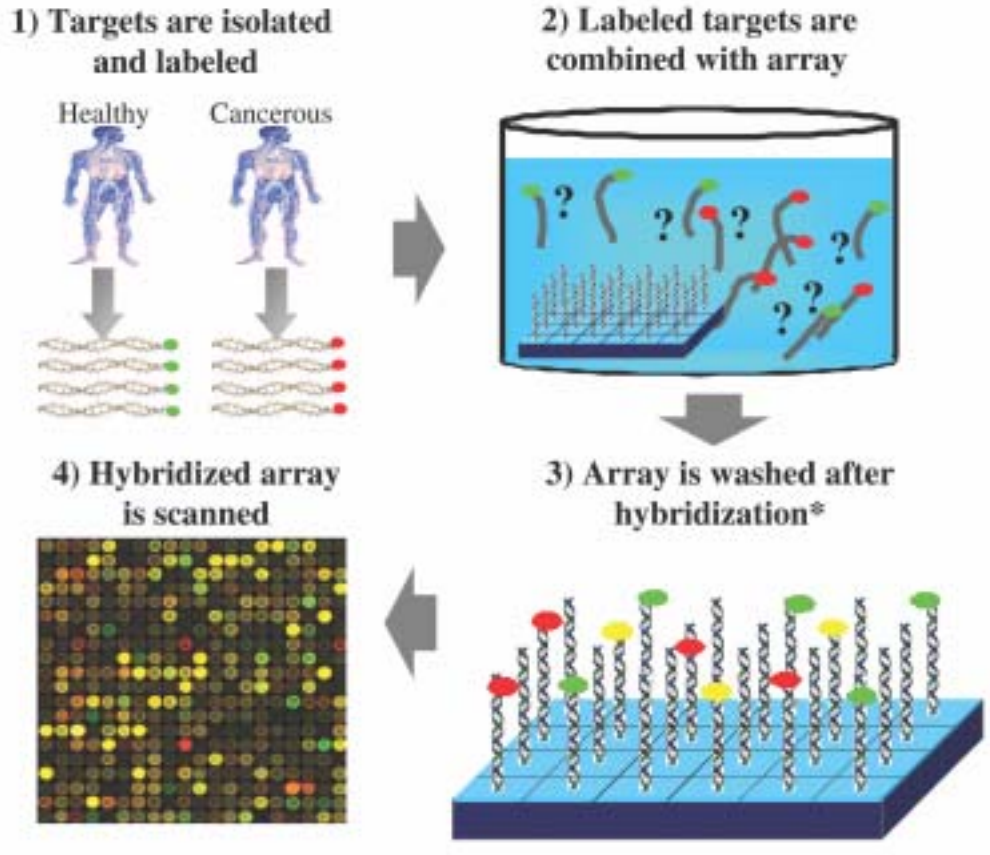

*Fig. 4 Simplified example of a two-color DNA microarray hybridization experiment. *Yellow dyes correspond to probes that have have hybridized with green- and red-labeled targets.*

broken down into its separate physical components to describe the properties and characteristics that are of interest to the materials developer.

The importance of glass and coatings

The DNA microarray can essentially be broken down into three key components:

- A solid, inert, impermeable support;
- A coating that links the inorganic glass to the organic biological molecules; and
- An array of biomolecular probes suited for hybridization.

Standard 25 mm x 75 mm glass microscope slides were the first supports commonly used for microarray assays, mainly because of their availability in most molecular biology laboratories where DNA microarray technologies were invented. However, the widespread adoption of glass as the substrate of choice for microarray applications is also attributable to its intrinsic properties, including low fluorescence, excellent flatness, chemical inertness, and low cost. Low intrinsic fluorescence is an especially important property for microarray supports, since most assays are characterized using fluorescent imaging equipment with excitation wavelengths in the visible spectrum (most commonly 523 nm and 635 nm), and high background would result in unacceptable microarray sensitivity. Fortunately, both borosilicate and soda lime silicate technical flat glasses produced by major producers, such as Schott Glass, Corning, and Erie Scientific, exhibit low fluorescence when excited in the visible spectrum, especially when compared to other low cost substrates such as polymers.

The excellent flatness achieved by floating, down-drawing, and overflow fusion of glass is also important for microarray applications. A total flatness of <50 μm (peak-to-valley over a 25 mm x 75 mm area) is required during a microarray assay to ensure precise biomolecule immobilization (by spotting or photosynthesis) and accurate scanning, since some laser scanners have a depth of focus of only ±30 μm.

Another important property of glass for microarraying applications is its excellent chemical durability[12]. The glass must provide an inert support for biomolecule immobilization and hybridization and should not dissolve or leach alkali ions, which could change the ionic strength of the buffer solutions used during the microarray printing and hybridization processes. The alkali leaching resistance of borosilicate glass is about 15 times better than soda lime silicate under acidic conditions and about five times better in hot water. For this reason, borosilicate glasses are highly suited for microarraying applications. In addition, borosilicate technical flat glasses are low cost materials because of the economies of scale resulting from their use in a number of consumer applications, ranging from microelectronics and displays to fire protection barriers.

The next important component of a microarray is the coating, which provides a uniform surface for biomolecule immobilization. Organo-functional alkoxysilane coatings are widely used and accepted within the microarray industry, as they are well suited to the coating of glass surfaces and can provide a variety of possible chemical functionalities. The general structure of an organo-functional alkoxysilane is shown in Fig. 5. During silanization, which is typically achieved by dip coating or thermal chemical vapor deposition, the silanol group reacts covalently with the glass surface via hydrolysis and condensation reactions. Once attached, the silane molecule can extend from the glass surface, providing organic functionality to the glass surface (see Figs. 5 and 6). Such silanization of glass has been studied since the early 1970s and a great deal of data are available, which show that these coatings can be highly uniform, robust, and reliable. However, the substrate chemistry, surface preparation, and organo-silane deposition technique must all be carefully controlled[13].

Although the coating structures shown in Figs. 5 and 6 are highly idealized, such uniform coatings are the goal. In order to approach this ideally smooth and organo-functionalized surface, it is important to eliminate the possible formation and deposition of polymeric globules of condensed silane. Further, it is important to minimize the interaction and

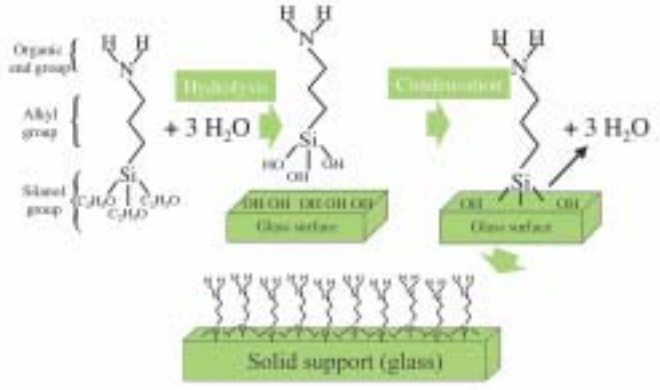

Fig. 5 Hydrolysis and condensation reactions associated with an amino alkyl silane group (top), which lead to a silane-coated surface (bottom).

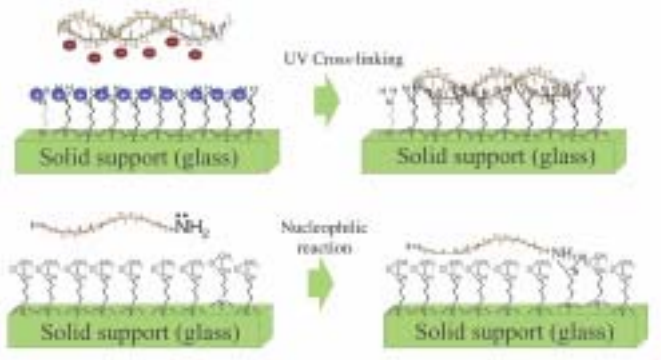

Fig. 6 Ionic attraction of a negatively charged, double-stranded DNA probe with a positively charged (protonated) aminosilane coating (top), where covalent attachment is achieved by UV cross-linking. Reaction of a single-stranded, amino-modified DNA probe with an epoxysilane coating (bottom), where covalent attachment is achieved through a nucleophilic reaction.

bonding of the organic end group of the silane with the glass surface, as the organic functionality should ideally protrude from the coated surface[14]. Such proper orientation of silane molecules can be especially challenging to achieve in the case of aminosilanes because of the ionic interactions between the negatively charged glass surface and positively charged amino groups. Finally, while the close packing of the silane molecules is desirable to stabilize the coating and prevent undercutting by water, it is difficult to achieve except in the case of self-assembled monolayers[15]. In current commercial practice, monolayer silane coatings are generally disordered and display a distribution of bonding configurations (including self-condensation/cross-linking), although skewed such that the majority of the organic end groups extend from the coating and are free to interact with the DNA.

Although silane molecules with many different organic end groups are commercially available, amino and epoxy chemistries are the most commonly used for printed DNA microarray applications. Aminosilane coatings were originally used for DNA microarray applications because the high concentration of primary amino groups available at the coated surface become protonated and, thus, positively charged when placed in contact with a near-neutral, aqueous solution (Fig. 6). This positively charged aminosilane surface is well suited to the ionic attraction of negatively charged DNA probes during the probe spotting and immobilization stages of DNA microarray fabrication. After DNA probes are immobilized on an aminosilane-coated surface by ionic attraction, ultraviolet (UV) cross-linking is generally used to achieve covalent linkages. Epoxysilane-coated substrates do not function on the premise of ionic attraction, but are instead based on the irreversible, covalent, ring-opening interaction between epoxide end groups in the coating and nucleophilic groups, such as amino groups, on the DNA probe (Fig. 6). Although amino- and epoxysilane-coated substrates differ mechanistically, they both result in the covalent immobilization of DNA probes onto a coated, solid support.

The final component of a DNA microarray is the biomolecule probe. There are two primary methods of grafting these probes onto the surface of a coated substrate: robotic printing and *in situ* synthesis. Robotic printing is commonly used by institutes, universities, and small biotechnology companies who fabricate their own microarrays. Printing is a 'pick and place' process, whereby robotic spotting equipment is used to 'pick up' nanoliter quantities of a DNA-probe-containing solution that is located in the wells of a 96- or 384-well microtiter plate. This DNA-probe-containing solution is then deposited onto a coated substrate to form spots with a diameter of 80-200 μm via a contact (pin) or noncontact (piezoelectric) print head. This 'pick and place' process is repeated tens, hundreds, or even thousands of times using single or multiple pin print heads to achieve microarrays with densities ranging from less than 100 to over 10 000 spots per array.

In contrast to 'pick and place' printing, *in situ* microarray fabrication processes are used by the two most commercially successful DNA microarray companies, Affymetrix and Agilent. During *in situ* synthesis, techniques such as combinatorial photochemistry (Affymetrix) or phosphoramidite coating supported by ink jet printing (Agilent) are used to build up unique DNA probe sequences at specific locations on a coated glass substrate surface[16-18]. Such *in situ* processes can lead to more probe versatility, as the probes are synthesized directly on the coated surface, thus eliminating the need to work with large probe libraries, which have to be stored in microtiter plates. Further, the use of photochemistry for *in situ* synthesis can result in reduced spot sizes (<20 μm for the Affymetrix system) and increased microarray densities.

Not perfect, but promising

The previous text may lead one to think that DNA microarray technologies are robust and poised to provide major contributions in the areas of drug discovery and molecular diagnostics immediately. However, major problems still exist with DNA microarrays, mainly related to the lack of

standardization of this young technology and adequate computer-based tools (bioinformatics) for quick and efficient analysis of the tremendous data sets created with each experiment. Such lack of standardization and 'turn-key' bioinformatic solutions are often manifested in the form of microarrays with insufficient quality, sensitivity, and reproducibility to be fully exploited for critical applications, such as diagnostics.

Materials researchers working in the area of DNA microarrays are struggling to come up with advancements to overcome some of the problems related to inadequate quality, sensitivity, and reproducibility. Methods to characterize the quality of microarrays fully are still lacking, yet are absolutely necessary to ensure that microarray products are fully functional prior to use and will provide both the precision and accuracy to answer a biological question adequately. Most silane coatings used for DNA microarrays are <1 nm thick, amorphous, transparent, and push the detection limit of most surface spectroscopic tools such as ellipsometry, infrared spectroscopy, X-ray photoelectron spectroscopy, and other thin film characterization methods. Further unmet challenges still exist with regards to the nondestructive quality assessment of the 10-200 μm spots that are printed onto these coated substrates, sometimes with densities of more than 500 spots/cm^2. Thus, the quality assessment of most microarray products is based on destructive tests that are conducted on some statistical, batch-wise analysis of products coming from multi-step production processes, which are often not well characterized or understood.

Advancements in sensitivity are also needed so that microarray users can characterize the activity of low expression genes, which may play a major role in drug discovery and/or molecular diagnostics without amplification. Although most microarray scanners offer four to five decades of dynamic range, microarrays often only exhibit useful data over two to four decades of signal intensity. Materials researchers are currently developing nanostructured surfaces and/or coatings capable of higher probe loading capacities or that exhibit reflective properties, which can lead to increased microarray sensitivity and (useful) dynamic range.

Increased reproducibility in microarray results is also necessary before such tools can become commonplace in clinical laboratories. Array to array coefficients of variation (CVs) of 10-20% still plague the industry, yet the true combined sources, which cause poor reproducibility in the multi-step microarray hybridization processes, are not well understood. The hybridization process still involves a great deal of manual handling steps, which must be eliminated eventually to remove human error. Materials researchers are developing microfluidic and lab-on-a-chip technologies that will eventually be combined with microarrays to reduce manual handling and could lead to easy-to-use, highly reproducible microarrays.

Developments by materials researchers in the areas of quality, sensitivity, and reproducibility could lead to breakthroughs that will eventually be adopted and standardized within the DNA microarray community. Such developments, along with subsequent standardization and bioinformatic improvements, are needed to realize fully the promise of microarrays for drug discovery and molecular diagnostic applications. The proverbial stage has been set, and most of the information and tools are available for making major advancements in human health where it relates to medication and disease diagnosis. However, such promise will be dependent upon the further standardization and utilization of microarray tools, along with better education on how these tools can best be put to work. MT

REFERENCES

1. Moore, S. K., *IEEE Spectrum* (November 2000), 33
2. Carey, J., *Business Week* (June 2000), 147
3. Service, R. F., *Science* (1998) **282**, 396
4. Patterson, S. D., *BioTechniques* (2003) **35**, 440
5. Alwine, J. C., *et al.*, *Proc. Natl. Acad. Sci. USA* (1977) **74**, 5350
6. Sambrook, J., *et al.*, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1989)
7. Schena, M., *et al.*, *Science* (1995) **270**, 467
8. Southern, E., *et al.*, *Nat. Genet.* (1999) **21** (supplement), 5
9. Duggan, D. J, *et al.*, *Nat. Genet.* (1999) **21** (supplement), 10
10. Chomczynski, P., and Sacchi, N., *Anal. Biochem.* (1987) **162**, 156
11. Manduchi, E., *et al.*, *Physiol. Genomics* (2002) **10** (3), 169
12. Pantano, C. G., *et al.*, *Corrosion of Glass*, Books for Industry, New York (1979)
13. Pantano, C. G., What do we know about glass surfaces? In *Proceedings of 61st Conference on Glass Problems*, Drummond, C. H., (ed.), American Chemical Society (2001), 137
14. Shallenberger, J. R., *et al.*, *Surf. Interface Anal.* (2003) **35**, 667
15. Ostuni, E., *et al.*, *Langmuir* (2001) **17** (18), 5605
16. Fodor, S. P. A., *et al.*, *Nature* (1993) **364** (6437), 555
17. Richter, A., *Semiconductor Manufacturing* (August 2003), 29
18. Masi, C. G., *R&D Magazine* (April 2000), 22