Making and reading microarrays

Vivian G. Cheung[1], Michael Morley[1], Francisco Aguilar[1], Aldo Massimi[2], Raju Kucherlapati[2] & Geoffrey Childs[2]

[1]*Department of Pediatrics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.* [2]*Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA (e-mail: vcheung@mail.med.upenn.edu & childs@aecom.yu.edu)*

There are a variety of options for making microarrays and obtaining microarray data. Here, we describe the building and use of two microarray facilities in academic settings. In addition to specifying technical detail, we comment on the advantages and disadvantages of components and approaches, and provide a protocol for hybridization. The fact that we are now making and using microarrays to answer biological questions demonstrates that the technology can be implemented in a university environment.

One high throughput method by which to gain information about gene function is the gridded cDNA microarray[1–3], in which microscope slides containing hundreds to thousands of immobilized DNA samples are hybridized in a manner very similar to the northern and Southern blot (see page 5 of this issue (ref. 4)). On realizing the power of this approach, we decided to build high-speed, high-precision arrayers in our respective laboratories—at the University of Pennsylvania (Penn) and the Albert Einstein College of Medicine (AECOM). The robot built by Pat Brown at Stanford Medical School (http://cmgm.stanford.edu/pbrown/index.html) first demonstrated the feasibility of this approach. Our goals are (i) to make it possible to eventually measure the expression of every gene in a mammalian cell using one or a few arrays and at a reasonable cost, (ii) to develop an array-based mapping method[5], and (iii) to optimize sensitivity with regard to both hardware and protocols.

The advantages of glass slides

An ideal support allows effective immobilization of probe onto its surface, and robust hybridization of target with the probe. Glass has many of the same advantages as nylon, the other standard support used for making microarrays (see pages 7 (ref. 4) and 11 (ref. 5) of this issue). It also has unique advantages. First, DNA samples can be covalently attached onto a treated glass surface. Second, glass is a durable material that sustains high temperatures and washes of high ionic strength. Third, it is non-porous so the hybridization volume can be kept to a minimum, thus enhancing the kinetics of annealing probes to targets. Fourth, as a consequence of its low fluorescence, it does not significantly contribute to background 'noise'. Finally, two different probes can be labelled with different fluors, and simultaneously incubated with a microarray in a single reaction; nylon arrays are restricted to serial or parallel hybridizations.

Robots (arrayers) are required to place (or array) a large number of probes onto slides. Here, we describe the AECOM arrayer (http://sequence.aecom.yu.edu/bioinf/funcgenomic.html) and scanner, and discuss technical considerations of their design and operation. For information about the arrayer at Penn (Fig. 1*a*), please refer to the web site http://w95vcl.neuro.chop.edu/vcheung, which houses a video of the arrayer in action.

Robotic features

The AECOM arrayer, Albert (Fig.1*b,c*), generates high-density, gridded arrays of cDNA, genomic DNA or similar biological material on glass surfaces. Its principal components are a computer-controlled three-axis robot and a unique pen tip assembly.

Design features. The robot is designed to automatically collect samples from either 96- or 384-well microtitre plates, with up to 12 pens simultaneously. Each pen collects from between 250 to 500 nl of solution per pen and deposits 0.25–1 nl on each slide, creating spots that range from 100–150 μm in diameter. The robot is programmed so that successive spots are spaced so that each avoids contact with adjacent spots, with approximately 200 to 250 μm separating the centres of each spot (Fig. 2). The precision of this measurement is about 10 μm. The robot rests on an optical table (Newport Corporation) which allows many configurations of microscope slides and microtitre plates. An example of a full table is one supporting 230 slides and 5 microtitre plates, along with up to 3 washing stations and a drying station.

The wash stations are stationary basins containing distilled water that is replaced after every two microtitre plates. When the pen tips are immersed, the robot shakes the pen assembly back and forth (at about 5 Hz) to enhance cleaning. A computer-controlled water bath sonicator and/or flowing water bath could be substituted, although we have not found this to be necessary. The dryer is essentially a computer-controlled wet/dry vacuum cleaner (Sears Company, USA) and an adapter fitted with restricting inlet holes into which the pen tips are inserted. Drying is accomplished by the rapid air flow around the tips and the partial vacuum this creates.

An important goal when designing a robot is to obtain both high speed and precision with minimum vibrations. We achieved this by using a well-damped optical table, mechanical slides with precision screw drives, robust servos with high-resolution encoders, and an outrigger supporting rail along the X-axis that avoids the cantilever design seen in some systems. Use of the second X-axis slide increases system stiffness, which in turn yields more rapid position settling and uniform accuracy across the table surface. These features permit rapid motion at rated precision, allowing the robot to service two microscope slides per second.

A critical component is the pen holder assembly with pen tips (see pages 11 (ref. 6) and 31 (ref. 7) of this issue). Our design incorporates linear motion bearings for each pen that allows precise vertical shaft motion with minimal resistance while preventing displacement in the other axes. Another unique feature of our design is an adjustable threaded end piece that allows alignment

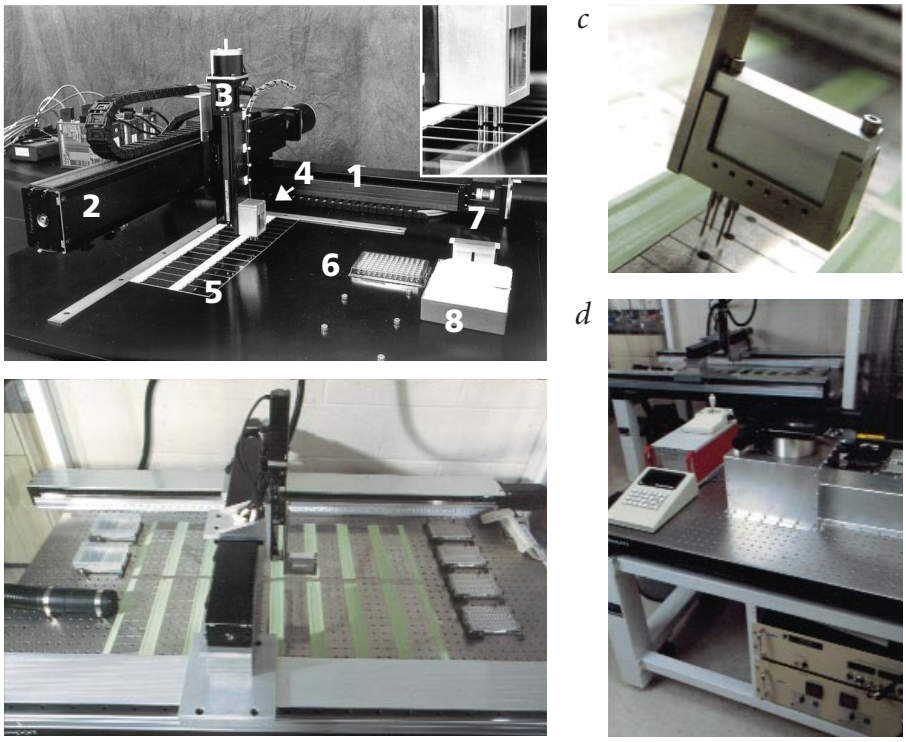

Fig. 1 ***a***, Penn microarray robot. The X-, Y-, Z- axes are labesled 1, 2, and 3, respectively. The key component of the arrayer is the print-head, containing pens (4). Microscope glass slides are placed on the slide station (5). Samples are prepared and arrayed from 96-well sample plates (6). The pins are cleaned between sample acquisitions at the washing (7) and drying (8) stations. ***b***, AECOM microarray robot. The table configuration shown contains 160 slides with four microtitre plates, two wash stations and the dryer. The print-head (***c***) shows four of the possible twelve pen tips in use. ***d***, AECOM laser scanner. Visible are the optical table, power supplies for lasers and PMT cooling, the Ludl stage, and lasers. The 20× microscope objective is inside the ludl stage while lenses, mirrors and other optics are enclosed in the metal casing. PMTs are to the right and outside the photo.

of each pen shaft within 10 µm to ensure that all 12 pen tips touch the microscope slide at the same time. Other designs without this feature require matched sets of pens with precise lengths for multi-pen printing. Each pen shaft is backed with a low-tension spring to ensure return to the extended position when not contacting a surface. The pen tips are made from approximately 1.6 mm diameter stainless steel tapered at the point to approximately 100 µm diameter, with a vertical notch cut along the center so as to create two walls that, at the tip, are 25 µm apart.

The system is controlled by a Visual Basic program running in a Microsoft Windows NT environment. The software provides a user-friendly front end to: configure the printing sequence; perform system alignments and calibrations; display real time position, velocity, following errors, servo system status as well as other functional parameters; and dynamically display key parameters of the printing process. A Servo control card in the computer makes it possible to dynamically control the high-speed, complex robot dynamics and is programmed using its own motion control program language. The visual basic and the servo card motion control programs operate interactively and exchange parameters, status and commands as required. The identity of the microtitre plate is determined by scanning it with a barcode reader as it is placed on the table. Due to initial problems with pen tip clogging by dust and fibers, the printer is now enclosed in a softwall chamber that allows easy access from three sides and incorporates a HEPA filter with blower for recirculating humidity-controlled air (a measure that has solved the problem effectively).

Operation. The printer is first prepared by aligning the microscope slides in a uniform pattern on the table using the 1" spaced holes on the laser table as guides and then taping them down. The modular microtitre plate holders snap into position anywhere on the table and use the same holes to align themselves. The same microtitre plate holders are used to hold the wash stations, which can also be placed in any convenient location. The user either selects a saved configuration or enters parameters into a configuration table. A slow walk-through mode is used to verify that key coordinates have been entered correctly. Finally, the user is prompted to add the microtitre plates and the robot proceeds with the automated spotting operation. The pens collect samples from the microtitre plate and deposit them, sequentially, on each of the slides (Figs 1*c*,2). This is followed by a wash/dry operation and the cycle is repeated with a new set of samples. When all the samples in the microtitre plates have been processed, the user is again prompted to add a new batch of plates and the cycle is repeated until all the samples in the run have been spotted. During the spotting operation, the program automatically saves to disk the identity of the source microtitre plate, the well number for each spot, and its X–Y destination on the slide. This file is later merged with a gene description file to produce a composite description of each spot printed on that slide.

Observations. The precision of the spot size is critically dependent on the specification of the pen tips. The fine notch at the tip requires special micro-machining tools such as Wire EDM. At the moment, we are using tips from TeleChem International, adapted to our pen shafts. Their performance is acceptable, although they are very fragile; we hope to obtain improved, more durable versions.

Scanner features

We designed and built a laser scanner, IRIS (Fig.1*d*), a derivative of instruments built at Stanford and the National Institutes of Health (NIH), to maximize sensitivity and dynamic range. We also sought to incorporate a degree of operative flexibility into its design, so that it can be modified in the future, and will allow the evaluation of more efficient fluorescent dyes (recently introduced to automated DNA sequencing instruments). Following hybridization with two dye-tagged targets, the slide is scanned to generate two 16-bit TIF images. The pixel intensity of each spot is proportional to the number of dye molecules and hence the number of probes hybridized with the spotted PCR product.

Design features. There are several key components of the laser scanner (Table 1). The software developed with the HPVEE graphical programming language synchronizes the program to the stage motion, manages the A/D converter data capture,

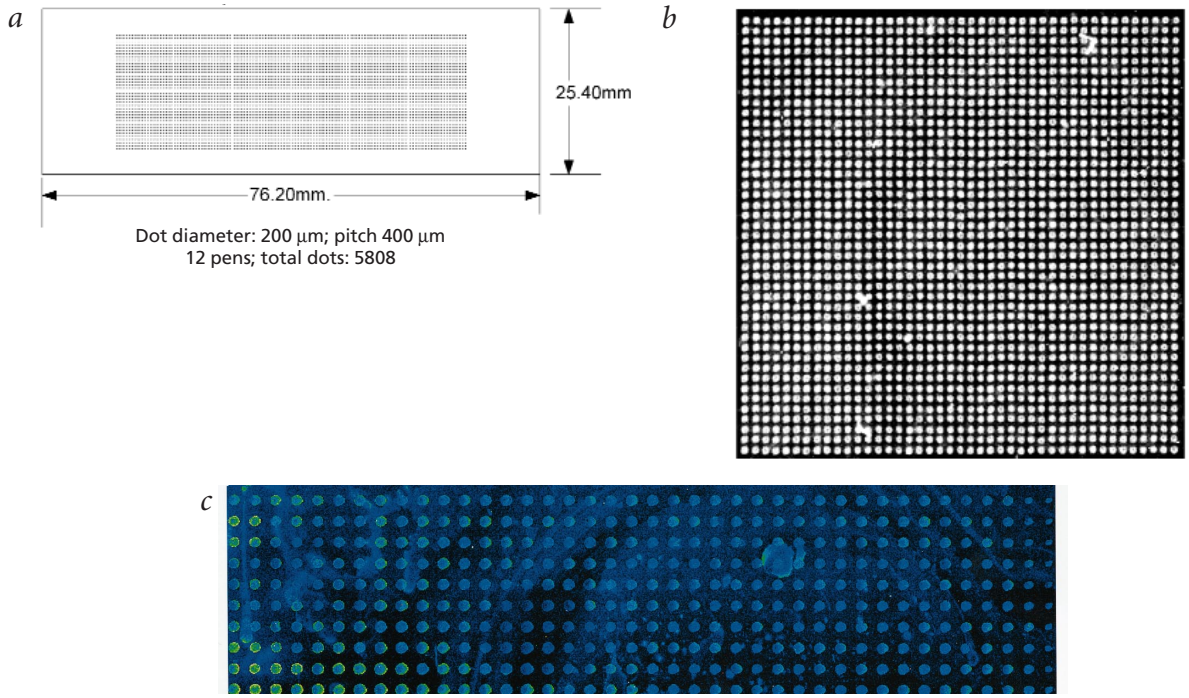

Fig. 2 ***a***, Schematic representation of a 12 pin array. This printing configuration makes maximal use of the area of a 1" X 3" microscope slide. ***b***, Salt deposits from printing with a single pen tip with spacing between spots of 200 μm. This 43 X 43 grid contains 1849 spots. This array was produced before the installation of a filtered enclosure—dust contamination is evident. ***c***, A genomic microarray, printed by the Penn robot. Cye3-labelled genomic DNA was hybridized onto a microarray containing clones DNA fragments. All spots were deposited from one 'loading' of the pen; a pen can hold enough samples to array 350 spots (500 μm in diameter).

processes signals for the two channels, displays real time parameters on a per scan basis, generates the TIF file header and saves the results as TIF images. Eight samples of incoming data are averaged per pixel and converted to binary integers, order-reversed in alternate scans to compensate for the change in direction of the stage and saved. The user is presented with large screen oscilloscope waveforms as well as average, minimum and maxiumum statistics for each scan.

Operation. The motorized stage executes a programmed comb scan pattern that sequentially traverses the microscope slide in the X direction, and then steps a pixel width in the Y direction, producing a bi-directional raster pattern. The X-axis encoder signals are processed by a specially designed trigger circuit that blanks out the servo oscillations at the ends of each scan and generates clean triggers for the A/D converter in both directions. The circuit assures sub-micron spatial resolution and image linearity. The two laser beams are made co-linear and reflected by the dual beam-splitter filter into the objective to form a narrowly focused beam that stimulates the dye molecules on the microscope slide to fluoresce. Part of that fluorescence is captured by the objective and routed through the dual beam-splitter filter, separated into red and green signals in the filter cube, bandpass filtered and routed to their respective photomultiplier tubes (PMTs) where they are converted to electrical signals. The output of each PMT is amplified, filtered and sampled by the A/D converter. The converter performs 8× oversampling and the software averages 8 samples per pixel, yielding a true 16-bit resolution image.

Observations. We initially obtained an inadequate signal-to-noise ratio, and therefore had a customized dual element filter built to our specifications (by Chroma), which reduced the DC component of the noise level by a factor of ten. Elimination of the spatial filter components improved the sensitivity still further. Our initial design included a spatial filter consisting of two matched focusing lenses, a pinhole and various mounting components, which, when put together, make a confocal microscope. We found, however, that confocal optics do not enhance detection. In fact, the lenses doubled the noise level—due to auto-fluorescence—while the entire assembly caused a sizable attenuation of the desired signal, resulting in an intolerable reduction in overall signal-to-noise ratio. We therefore conclude that spatial filtering in the Z-axis does not serve any useful purpose in this application. We found that clean-up filters at the output of the lasers are essential for noise reduction. Finally, with careful alignment of the optical components and the precise determination of optimal focus, using a calibration slide and software, we have achieved the dual objective of high sensitivity and a wide dynamic range.

A problem sometimes encountered is the presence of specular noise, consisting of very bright signals that are smaller (<1 μm–25 μm) than the spots we are attempting to image. We think that this was due to the presence of dust and unincorporated dye molecules. Operating in a clean environment and rigorous adherence to hybridization protocol is essential to minimize these effects.

Table 1 • IRIS components

- Green (532 nm) and Red (633 nm) lasers for simultaneous two colour dye excitation
- Ludl motorized X,Y,Z microscope stage and electronics with sub-micron resolution
- Dual PMTs and housings (with cooling capability) for two colour detection and preamplifiers
- Nikon 20×infinity focus objective, filter cube and Z axis slide
- Dual beam splitter filter, laser clean-up filters, beam combiner, and bandpass filters
- A PC based A/D converter and anti alias filters
- Digital trigger circuit for conversion of stage encoder signals to A/D converter triggers
- 300 MHz Pentium PC running software developed in HPVEE graphical language
- Laser table, various fabricated metal parts, optical mounts and mirrors

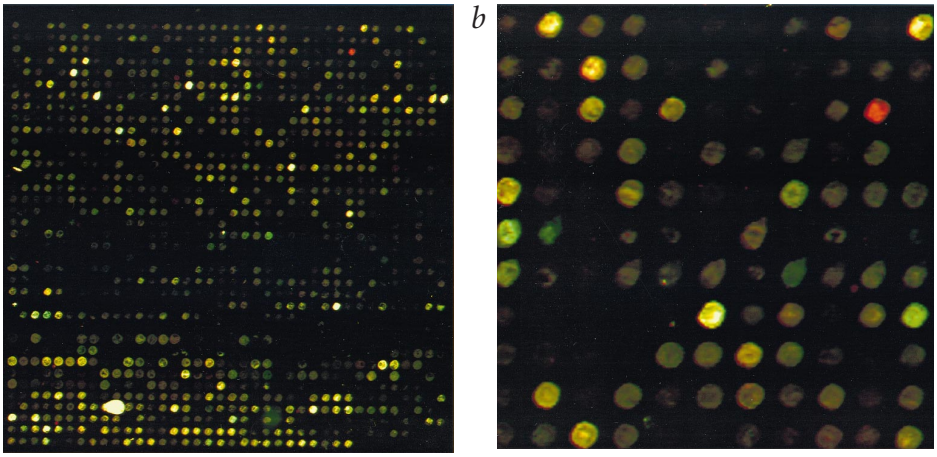

Fig. 3 ***a***, An hybridized microarray printed by the AECOM robot. A 5550-gene mouse cDNA microarray was printed and hybridized to Cye3-dUTP and Cye5-dUTP probes from wild-type and mutant mouse cell-lines and imaged using the AECOM laser scanner. Shown is one out of four of the pen tip printing areas. ***b***, An enlargement of an 11×11 spot region of the array.

The waveforms on the screen display are highly repeatable. When compared to a repeated scan of the same line, we see excellent reproducibility. From this we deduce that the scanner electronics and optics do not introduce any significant signal distortion. This display capability also allows us to accurately measure effects on signal-to-noise ratio when adjusting various control parameters. PMT cooling was included in the design to help keep electronic noise to a minimum. So far, we have not found any significant improvement in sensitivity when the PMT is cooled down to –18 °C. The system noise level has not yet approached the electronic noise level and its dominant component is still optical noise. It is likely that the hybridization process leaves some residue on the slide surface that auto-fluoresces. We are exploring methods to further reduce the noise level.

Laser power of 10 mW appears to be more than sufficient, given the current system sensitivity; 5 mW yields satisfactory results and indicates that the system gain is linear in this region. PMT voltage and laser power can be traded interchangeably without any apparent degradation in signal-to-noise ratio.

Performance. At the moment, there is no commonly accepted standard with which to compare scanner performance. To gauge the performance of our scanner, we measured sensitivity by using several slides containing calibrated amounts of Cye3 dye of various concentrations. Results indicate that the scanner reliably detects concentrations of less than 10^{-18} mole of dyes on a 100-μm spot. We intend to carry out additional tests to determine the efficiencies of dye incorporation and hybridization, but performance suggests that we will be able to detect low-abundance mRNA species using current probe preparation protocols. Preliminary results show that our scanner has a sensitivity of approximately fourfold that of a commercial scanner, while simultaneously handling signals about three times larger (before saturation). Our scanner has a dynamic range of over 1000 fold, which obviously compares well with high-density filters whose dynamic range is no greater than tenfold. We can do two-colour simultaneous imaging; however, this has its limitations because of optical crosstalk between the two channels. This can be minimized by scanning one colour at a time, although cross excitation can still produce some crosstalk when more than one dye is used per slide. This imposes an upper and lower limit on the measurable green/red ratio when using two colours on one slide. A typical array generated by our system, scanned using a typical pixel size of 12.8 μm, creates an image of 2048 by 1550 pixels with 16 bit resolution (Fig. 3). Each spot covers an area of about 100 pixels and scan time for the image is about 40 minutes.

Table 2 • DNA quantity and spot size

Radius of spot (μm)	Volume of spot (nl)	Amount of DNA in spot (ng)
250	32.71	16.35
200	16.75	8.37
150	7.07	3.53
100	2.09	1.05
50	0.26	0.13

The DNA content of each spot is calculated assuming a sample concentration of 500 ng/μl

Hybridization considerations

Quantity of DNA. The amount of DNA in each spot of the microarray can be estimated. Assuming that each spot is deposited as a hemisphere, its volume and content can be calculated:

$$\text{The volume of a spot} = 1/2 \times (4/3\ \pi r^3)$$

$$\text{The amount of DNA per spot} = \text{sample concentration} \times \text{volume of the spot}$$

The small volume of the spot means that the amount of probe for hybridization is also small, even if the sample concentration is high (Table 2). Every effort must thus be made to make good this limitation. Several factors should be taken into consideration. In addition to the amount of probe DNA, the proportion of probe DNA that is complementary to the target, the length and specific activity of the target as well as the sensitivity of the method used to detect the signal all affect signal intensity[11].

The strength of the hybridization signal is proportional to the specific activity of the target and inversely proportional to its length, so it is important to use targets with high specific activity. Hybridization time should also be titrated for each experiment.

The mechanics of deposition. The narrow slit of the spotting pen allows samples to be drawn by capillary action from microtitre plates. The force generated from the downward motion of the pen and the surface tension of the slide 'pulls' the sample from the slit onto the slide. Spot size depends on the acceleration of the pen towards and away from the slide, and the surface tension of the slide. The rate of acceleration of the pen towards the slide is proportional to the size of the spot; the arrayer can therefore be adjusted to achieve the desired spot size. As the pen withdraws from the slide, a column of fluid is formed between the sample in the slit and the sample that has been deposited on the slide. If the rate of withdraw is fast, that column can be interrupted at several points, creating a larger spot that is also less likely to be a perfect hemisphere. However, if the rate of withdraw is sufficiently slow, that column is likely to 'pinch off' at only one point, leaving behind a small spot (Howard Hu, pers. comm.).

DNA samples. Samples are prepared in 96-well plates, ethanol precipitated and washed in 70% ethanol, then reconsti-

Table 3 • Hybridization solutions

Solution A (used at 42 °C)	50% formamide 6×SSC 0.5% SDS 5×Denhardt's reagent (0.5 g Ficoll, 0.5 g polyvinylpyrrolidone, 0.5 g bovine serum albumin)
Solution B (used at 65 °C)	6×SSC 0.5% SDS 5×Denhardt's reagent
Solution C (used at 65 °C)	10% SDS 7% PEG-8000

Table 4 • Protocol for hybridization

- moisten array prior to UV-cross linking
- cross-link at 0.25 J/cm^2 to 0.35 J/cm^2
- moisten array again and heat snap for 3 s
- wash in 0.1% SDS for 30 s
- rinse in water for 1 min
- denature by placing array in 95 °C water for 3 min
- immediately transfer to ice-cold ethanol
- air dry
- prepare targets and reconstitute in 15 µl hybridization solution
- prehybridize array with hybridization solution containing blocking agent for 30 min
- incubate probes with the array for 8–24 h
- wash the array in 0.1% SDS, 0.2×SSC for 5 min at RT, then in 0.2×SSC for 5 mins at RT
- dry by centrifugation at 2000 rpm for 2 min

tuted in 2×saline sodium citrate (SSC). The concentration to which the samples are reconstituted depends on the size of the desired spot and on the viscosity of the sample. If samples are to be arrayed in very small spots, then their concentration should be high. A limiting factor, however, is the pen-tip design, which makes very viscous samples difficult to array; those with a concentration greater than 2 µg/µl are usually too viscous to array. Whenever possible, we aim for a final concentration that allows us to array about 15 ng of sample per spot (Table 2). Although we reconstitute the samples in 2×SSC, the ionic strength of the solvent can vary from 1×SSC to 5×SSC without affecting the hybridization, but samples dissolved in solvents with an ionic strength higher than 5×SSC attach poorly onto the slides.

Immobilization of DNA onto glass slides. After DNA samples are arrayed onto slides, they are air-dried. The samples are immobilized by ultraviolet (UV)-irradiation to form covalent bonds between the thymidine residues in the DNA and the positively charged amine groups on the silane slides. A similar protocol is used for attaching DNA samples onto nylon membranes[8,9]. To achieve maximal hybridization, the arrays are slightly dampened before ultraviolet (UV)-crosslinking by exposing the 'arrayed' side to boiling water (Table 4). This is followed by UV-crosslinking at 254 nm by exposing the slide to 0.27 J/cm^2. We have found it advantageous to titrate the amount of irradiation, so as to determine an optimal level of irradition for producing the best signal. Under- and over-irradiation cause excessive loss of DNA by insufficient binding and over-nicking of the DNA samples, respectively. After crosslinking, excess DNA molecules are removed by washing the arrays in 0.1% SDS at room temperature and arrayed samples are denatured in water at 95 °C before hybridization.

Honing in on hybridization. There are many methods for hybridizing targets and probes (for more information on target preparation, see page 11 of this issue (ref. 6)). We found that three solutions work well for hybridization of fluorescent probes onto immobilized DNA on glass (Table 3); they differ with respect to the solvents and temperatures used. In general, formamide-based hybridization at 42 °C works better than aqueous solutions at 65 °C, as it favours a higher signal-to-noise ratio. The kinetics of hybridization in formamide, however, are slower than in aqueous solution[10,11], inspiring us to use an aqueous solution with dextran sulfate or polyethylene glycol when using a target of low copy number.

Familiar blocking agents are used to minimize 'noise': Denhardt's reagents, sodium dodecyl sulfate (SDS), sheared salmon sperm DNA, tRNA and Cot_1DNA. When using cDNA probes, we also include Poly(A)$^+$ RNA or poly dA to bind T-rich sequences.

Conclusions

Our collective efforts show how microarray technology can be implemented in an academic environment. At the AECOM, the microarrayer and scanner were built by an in-house engineer; one advantage of having an in-house engineer is that s/he is on hand to upgrade and repair the instruments. At the University of Pennsylvania, the microarrayer was built by personnel in the lab with the help of the university machine-shop; a commercial scanner (General Scanning, Inc.) was purchased.

Improved technologies for construction of microarrays must be matched by an increase in the availability of mapped genomic clones, well-characterized cDNA clones and user-friendly analysis tools. Happily, this would appear to be the case[12–14]. The integration and accessibility of these tools will be critical to the speed of with which developments in genomics inform genetic studies.

Acknowledgements

The AECOM group is indebted to S. Kneitz, R. Yang, S. Chen and T. Harris for their strenuous efforts towards building our facility. We thank H. Hu, J. Gregg, M. Bittner and P. Brown for stimulating discussions and R. Spielman and A. Bruzel for comments on the manuscript. The work at AECOM was supported by the Human Genetics Program and the Comprehensive Cancer Center at AECOM (NIH CA 13330). Work at the University of Pennsylvania was supported by grants (to V.G.C.) from Merck Genome Research Institute and the NIH (DC00154).

1. Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467–470 (1995).
2. DeRisi, J. *et al*. Use of a cDNA microarray to analyze gene expression patterns in human cancer. *Nature Genet.* **14**, 457–460 (1997).
3. DeRisi, J.L., Iyer, V.R. & Brown, P.O. Exploring gene expression on a genomic scale. *Science* **278**, 680–686 (1997).
4. Southern, E. Mir, K. & Shchepinov, M. Molecular interactions on microarrays. *Nature Genet*. **21**, 5–9 (1999).
5. Cheung, V.G. *et al*. Linkage disequilibrium mapping without genotyping. *Nature Genet.* **18**, 225–230 (1998).
6. Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P. & Trent, J. Expression profiling using cDNA microarrays. *Nature Genet.* **21**, 10–14 (1999).
7. Bowtell, D.L. Options available—from start to finish—for obtaining expression data by microarray. *Nature Genet.* **21**, 25–32 (1999).
8. Church, G.M. & Gilbert, W. Genomic sequencing. *Proc. Natl Acad. Sci. USA* **81**, 1991–1995 (1983).
9. Vollrath, D. & Davis, R.W. Resolution of DNA molecules greater than 5 megabases by contour-clamped homogeneous electric fields. *Nucl. Acid Res.* **15**, 7865–7876 (1987).
10. Sambrook, J., Fritsch, F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Press, Cold Spring Harbor, New York, 1989).
11. Casey, J. & Davidson, N. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. *Nucl. Acid Res.* **4**, 1539–1545 (1977).
12. Boguski, M.S. & Schuler, G.D. ESTablishing a human transcript map. *Nature Genet*. **10**, 369–371 (1995).
13. Schuler, G.D. *et al*. A gene map of the human genome. *Science* **274**, 540–546 (1996).
14. Ermoleava, O. *et al*. Data management and analysis for gene expression arrays. *Nature Genet*. **20**, 19–23 (1998).