

# **DNA Arrays for Analysis of Gene Expression**

Michael B. Eisen<sup>1</sup> and Patrick O. Brown<sup>2</sup>

Departments of Genetics<sup>1</sup> and Biochemistry<sup>2</sup> and the Howard Hughes  
Medical Institute<sup>2</sup>, Stanford University School of Medicine, Stanford, CA  
93405.

<b>Introduction</b>	<b>1</b>
<b>Genome-wide Expression Monitoring</b>	<b>3</b>
<b>Hardware</b>	<b>4</b>
<b>Preparation of Target DNA</b>	<b>5</b>
<b>Preparation of Slides</b>	<b>7</b>
<b>Printing of DNA Microarrays</b>	<b>9</b>
<b>Postprocessing of Slides</b>	<b>11</b>
<b>Preparation of Probe</b>	<b>13</b>
<b>Hybridization</b>	<b>16</b>
<b>Post-hybridization</b>	<b>16</b>
<b>Controls</b>	<b>16</b>
<b>Slide Imaging and Data Reduction</b>	<b>18</b>
<b>Array Problems</b>	<b>19</b>
<b>Conclusion</b>	<b>21</b>

## Introduction

DNA microarrays, microscopic arrays of large sets of DNA sequences immobilized on solid substrates, are valuable tools in areas of research that require the identification or quantitation of many specific DNA sequences in complex nucleic acid samples. DNA microarrays, which come in an ever increasing variety of flavors, have been used in genetic mapping studies, mutational analyses and in genome wide monitoring of gene expression<sup>1,2,3,4,5,6,7,8,9,10,11</sup>, and likely will become standard tools in research and clinical applications.

Microarrays are, in principle and practice, extensions of hybridization-based methods which have been used for decades to identify and quantitate nucleic acids in biological samples (e.g. Southern<sup>12</sup> and Northern blots<sup>13</sup>, colony hybridizations, dot blots<sup>14</sup>): samples of interest are labeled and allowed to hybridize to the array; after sufficient time for hybridization and following appropriate washing steps, an image of the array is acquired and the representation of individual nucleic acid species in the sample is reflected by the amount of hybridization to complementary DNAs immobilized in known positions on the array.

The idea of using ordered arrays of DNAs to perform parallel hybridization studies is not in itself new; arrays on porous membranes have been in use for years<sup>15,16,17</sup>. However, many parallel advances have

---

<sup>1</sup> J. DeRisi, L. Penland, P. O. Brown, M. L. Bittner, P. S. Meltzer, M. Ray, Y. Chen, Y. A. Su and J. M. Trent, *Nat Genet*, 14, 457-60 (1996).

<sup>2</sup> M. Chee, R. Yang, E. Hubbell, A. Berno, X. C. Huang, D. Stern, J. Winkler, D. J. Lockhart, M. S. Morris and S. P. Fodor, *Science*, 274, 610-4 (1996).

<sup>3</sup> J. L. DeRisi, V. R. Iyer and P. O. Brown, *Science*, 278, 680-6 (1997).

<sup>4</sup> F. Forozan, R. Karhu, J. Kononen, A. Kallioniemi and O. P. Kallioniemi, *Trends Genet*, 13, 405-9 (1997).

<sup>5</sup> R. A. Heller, M. Schena, A. Chai, D. Shalon, T. Bedilion, J. Gilmore, D. E. Woolley and R. W. Davis, *Proc Natl Acad Sci U S A*, 94, 2150-5 (1997).

<sup>6</sup> D. A. Lashkari, J. L. DeRisi, J. H. McCusker, A. F. Namath, C. Gentile, S. Y. Hwang, P. O. Brown and R. W. Davis, *Proc Natl Acad Sci U S A*, 94, 13057-62 (1997).

<sup>7</sup> L. Wodicka, H. Dong, M. Mittmann, M. H. Ho and D. J. Lockhart, *Nat Biotechnol*, 15, 1359-67 (1997).

<sup>8</sup> M. Schena, D. Shalon, R. Heller, A. Chai, P. O. Brown and R. W. Davis, *Proc Natl Acad Sci U S A*, 93, 10614-9 (1996).

<sup>9</sup> M. Schena, D. Shalon, R. W. Davis and P. O. Brown, *Science*, 270, 467-70 (1995).

<sup>10</sup> R. J. Lipshutz, D. Morris, M. Chee, E. Hubbell, M. J. Kozal, N. Shah, N. Shen, R. Yang and S. P. Fodor, *Biotechniques*, 19, 442-7 (1995).

<sup>11</sup> R. J. Sapolsky and R. J. Lipshutz, *Genomics*, 33, 445-56 (1996).

<sup>12</sup> E. M. Southern, *J. Mol. Biol.*, 98, 503-17 (1975).

<sup>13</sup> J. C. Alwine, D. J. Kemp and G. R. Stark, *Proc Natl Acad Sci U S A*, 74, 5350-4 (1977).

<sup>14</sup> F. C. Kafatos, C. w. Jones and A. Efstratiatis, *Nucl. Acids Res.*, 7, 1541-52 (1979).

<sup>15</sup> E. M. Southern, G. S. Case, J. K. Elder, M. Johnson, K. U. Mir, L. Wang and J. C. Williams, *Nucleic Acids Res*, 22, 1368-73 (1994).

<sup>16</sup> U. Maskos and E. M. Southern, *Nucleic Acids Res*, 21, 4663-9 (1993).

<sup>17</sup> E. M. Southern, U. Maskos and J. K. Elder, *Genomics*, 13, 1008-17 (1992).

occurred to transform these rather clumsy membranes into much more useful and efficient methods for performing parallel genetic analyses. First, large-scale sequencing projects have produced information and resources that make it possible to assemble collections of DNAs that correspond to all, or a large fraction of, the genes in many organisms from bacteria to humans. Second, technical advances have made it possible to generate arrays with very high densities of DNAs, allowing for tens of thousands of genes to be represented in areas smaller than standard glass microscope slides. Finally, advances in fluorescent labeling of nucleic acids and fluorescent detection have made the use of these arrays simpler, safer and more accurate.

The goal of this review is to describe one of the currently used microarray technologies, commonly called “spotting” or “printing” because DNAs, usually larger than oligonucleotides (100bp and up) are physically spotted on a solid substrate, in contradistinction to the other major current microarray technology, in which short oligonucleotides are synthesized directly on a solid support.<sup>18,19</sup> In standard spotting applications large collections of DNA samples (PCR products, plasmids, etc...) are assembled in 96 or 384 well plates. A robot positions a cluster of specially designed tips into adjacent sample wells, draws up approximately a microliter of DNA, and deposits a small spot of each onto coated glass microscope slides. By printing successive DNA samples in a staggered fashion, thousands of spots can be printed in an area corresponding to a cluster of 4 wells of a 96-well plate or 16 wells of a 384-well plate. Typically 100-200 copies of a given array can be printed simultaneously by successively touching the cluster of printing tips to each slide before washing and drying the tips and reloading with the next set of DNA samples. After a complete set of samples is spotted, the slides are treated to attach the DNA stably to the surface and to minimize non-specific binding of probes to the slide.

DNA microarrays have been used for a variety of purposes; essentially any property of a DNA sequence that can be made experimentally to result in differential recovery of that sequence can be assayed for thousands of sequences at once by DNA microarray hybridization. This review will focus on the application of DNA microarrays to gene expression studies and will discuss general principles of whole genome expression monitoring as well as detailing the specific process of making and using spotted DNA microarrays.

---

<sup>18</sup> A. C. Pease, D. Solas, E. J. Sullivan, M. T. Cronin, C. P. Holmes and S. P. Fodor, *Proc Natl Acad Sci U S A*, 91, 5022-6 (1994).

<sup>19</sup> S. P. Fodor, R. P. Rava, X. C. Huang, A. C. Pease, C. P. Holmes and C. L. Adams, *Nature*, 364, 555-6 (1993).

## Genome-wide Expression Monitoring

The ultimate technical goal of whole genome expression monitoring is to be able to determine to *absolute* representation of every RNA species in any cell or tissue sample of interest. One would like to be able to isolate RNA from a sample, make a labeled representation of the isolated RNA, hybridize this labeled probe to a microarray, and read out the amount of each RNA species in the original sample. However, there is a complex relationship between the amount of input RNA for a given gene and the intensity of the probe signal at a corresponding hybridization target. This relationship depends on a multitude of factors, including the labeling method, hybridization conditions, target features and the sequence of the gene. Therefore, microarray based methods are best used to assay the *relative* representation of RNA species in two or more samples. Fortunately, *differences* in gene expression between samples - i.e. where and when it is or is not expressed, and how it changes with perturbations, are what matters most about a gene's expression. Knowing the absolute abundance of the RNA provides only marginally further utility.

Two-color hybridizations: To maximize the reliability and precision with which we can quantitate differences in the abundance of each RNA species, we directly compare two samples by labeling them with spectrally distinct fluorescent dyes and mixing the two probes for simultaneous hybridization to one array. The relative representation of a gene in the two samples is assayed by measuring the ratio of the fluorescence intensities of the two dyes at the cognate target element. This ratio is relatively insensitive to the sources of spurious variation discussed above, which may affect the absolute amount of probe that hybridizes to a given target element, but not the relative amount of the two labeled species.

Experimental questions about gene expression to which DNA microarrays can be applied: The simplest kind of gene expression survey compares the relative abundance of mRNA corresponding to each arrayed gene, between two different cell or tissue samples. For example, an experiment might compare cells before and after an experimental perturbation, or at successive times during a temporally staged process, or between stages of differentiation, or one might compare the pattern of mRNA's expressed in a mutant cell with that of its wild-type counterpart. However, the flexibility of this experimental approach allows much more precise information about the regulation of each gene's expression to be collected. For example, translation rates can be measured for each gene by comparative hybridization of polysomal RNA and total mRNA to a

microarray. The distribution of each polyadenylated RNA between the nucleus and cytoplasm can be assessed by differential hybridization of nuclear and cytoplasmic polyadenylated RNA samples. A similar approach can allow other kinds of subcellular localization to be surveyed for the entire set of arrayed genes.

Type-I and Type-II experiments: Many interesting biological questions can be studied using the simplest form of two-color hybridization experiments, where two samples are directly compared on a single array (these are now designated Type-I experiments; c.f.<sup>20</sup>). However, such direct comparison is impractical for complex questions requiring the comparison of multiple samples. For these, we employ an alternative experimental design using a common reference sample in each of the hybridizations required for an experiment. This maintains the essential internal-control aspect of two-color hybridization, but allows for inferred comparisons to be made among large numbers of samples without requiring that every pairwise comparison be performed (such experiments are designated Type-II experiments). The most common use of this method has been in time-course experiments (c.f.<sup>21</sup>) where each timepoint is compared to an initial timepoint. Since the amount of a given gene at each timepoint is measured relative to a fixed reference (the initial timepoint), the behavior of each gene across the timecourse can be studied. The reference sample need not be related to the samples being examined. The most important attribute of the reference sample is that it should provide a hybridization signal, and thus a non-zero denominator for the hybridization ratio, at each target element on the array. For any large collection of samples to be compared, a convenient approximation of this ideal reference sample is an equal mixture of material from each of the samples. In such a mixture, any gene represented in one of the samples will be represented in the reference.

## Hardware

Making and using printed DNA microarrays requires two pieces of hardware: a robot to produce the microarrays (referred to as an **arrayer** [see Figure 1]), and a device to image the hybridized arrays (generally a scanning laser confocal microscope, or **scanner**). The excitement over microarray technology in large research labs and the biotech industry has been tempered elsewhere by concerns over the cost and expertise required to perform microarray experiments. To make the necessary hardware

---

<sup>20</sup> J. DeRisi, L. Penland, P. O. Brown, M. L. Bittner, P. S. Meltzer, M. Ray, Y. Chen, Y. A. Su and J. M. Trent, *Nat Genet*, 14, 457-60 (1996).

<sup>21</sup> J. L. DeRisi, V. R. Iyer and P. O. Brown, *Science*, 278, 680-6 (1997).

accessible and affordable, we have designed relatively inexpensive hardware that is easy for a molecular biologist without special engineering experience to assemble and operate, and are placing detailed parts lists, designs and construction and operating instructions on the web (at <http://rana.stanford.edu/hardware>). The total cost for both an arrayer and scanner is around \$60,000.

## Preparation of Target DNA

Choice of target: For large-scale gene expression studies, a specific hybridization target is required for every gene to be analyzed. As essentially any double-stranded DNA sample (and probably most single-stranded samples as well) can be printed onto a treated glass surface, the choice of target is largely dictated by the resources available for obtaining representations of the genes to be studied. For organisms whose genomes have been fully sequenced, the most straightforward approach is to amplify every known and predicted open reading frame in the genome (or any subset of this collection of particular interest to the experimenter) using PCR. The ever decreasing cost of synthetic oligonucleotides (now close to \$10 per primer pair), and the availability of 96-well thermocyclers in most departments, makes this a viable strategy for even relatively large genomes; our laboratory has used this strategy to print an array containing start-codon to stop-codon PCR products for nearly all of the approximately 6200 open reading frames from the fully sequenced yeast

*Saccharomyces cerevisiae* genome<sup>22</sup>. For unsequenced or partially sequenced genomes, or for genomes with large and numerous introns, this strategy is not practical. For many organisms, such as humans and mouse, extensive resources have been directed at determining partial sequences of clones from cDNA libraries. These expressed sequence tags (ESTs) can be used to identify distinct mRNA transcripts<sup>23</sup>, and individual cDNA clones corresponding to each of these transcripts can be used as the source of gene-specific targets in an array. Our laboratory has recently used this approach to make an array of approximately 10,000 human genes, where the DNA sequences that were spotted were the PCR-amplified inserts from cDNA clones in the I.M.A.G.E. library<sup>24</sup>. For unsequenced organisms without a genome sequencing project, or where no sequence is yet available, clones from cDNA libraries (preferably normalized libraries to minimize redundancy) can be used as targets. Expression data can be

---

<sup>22</sup>J. L. DeRisi, V. R. Iyer and P. O. Brown, *Science*, 278, 680-6 (1997).

<sup>23</sup>M. S. Boguski and G. D. Schuler, *Nat Genet*, 10, 369-71 (1995).

<sup>24</sup>G. Lennon, C. Auffray, M. Polymeropoulos and M. B. Soares, *Genomics*, 33, 151-2 (1996).

collected for each of these clones without any prior information about the sequence, map position or identity of the gene.

It is likely that some benefits could be found in using DNA fragments smaller than full open reading frames or 1000 basepair cDNA's to represent each gene. For example, with yeast genes having significant homology to each other, PCR products representing the most divergent parts of these genes could be used in place of the entire coding region to improve discrimination between the two genes. The optimal representation of each gene is a complicated problem currently under investigation.

PCR Amplification: In general, a 100  $\mu$ l PCR reaction provides sufficient DNA for printing up to 1000 arrays. Most of our reactions are done in this volume in 96-well thermocyclers (Perkin Elmer 9600s and 9700s). The actual parameters of the amplification will depend on the type of template and primers being used. It is important, however, to limit the components of the PCR to template, primers, nucleotides, magnesium, standard buffer and enzyme. Common additives such as glycerol or gelatin can interfere with the printing process by altering the surface tension of the drop or competing for surface attachment sites on the slide, and should not be used.

Preparing DNA for Printing: To clean up PCR products and prepare the DNA for printing, we generally isopropanol-precipitate the PCR products and resuspend the DNA at approximately 100 ng/ $\mu$ l in 3x SSC. Precipitation can be carried out either in the thermocycler plates used for PCR or in V-bottom 96-well tissue culture plates (Costar 3984). Start by adding, to each well, 10  $\mu$ l of 3M sodium acetate, pH 5.2, followed by 110  $\mu$ l isopropanol (we use isopropanol and not ethanol to reduce the total volume of the precipitation mixture to an amount that will fit in the thermocycler or tissue culture plates). Place the plates at -20C for one hour (if time is an issue, this step can be omitted). Precipitate the DNA by centrifugation at 4C (in a Sorvall RC-3B refrigerated centrifuge with an H6000A rotor fitted with P/N 11267 swinging bucket adapters) at 3500g (3500 rpm in this rotor) for 1 hour. Place two or three folded paper towels beneath the plate to prevent cracking. Stacks of up to four V-bottom plates can be spun in one adapter provided that a cushion of paper towels is placed between each plate and that the plates are carefully taped together to prevent slippage. After precipitation, carefully remove the liquid by aspiration or by inverting the plate. Wash the pellets with 100  $\mu$ l 70% ethanol (make from 95% ethanol as 100% ethanol may contain fluorescent impurities) and spin again for 30 minutes. Aspirate or decant the liquid again, and dry the pellets in a rotary evaporator fitted to accept 96-well plates (we use a Savant SpeedVac Plus SC210A; if such a device is not available, the pellets can be air-dried if care is taken to keep out dust).



Resuspend the pellets in 10-15  $\mu$ l 3x SSC (pH 7.0) at 4C for at least 12 hours; care should be taken to seal the plates so that little evaporation occurs, as excessive salt or DNA concentration can interfere with printing. After resuspension, transfer the solution to soft-bottom plates (Fisher 3911 MicroTest III Flexible Assay Plates) for printing. Seal the plates and store at 4C or -20C until use. Immediately prior to printing, the plates should be spun lightly to ensure that all material is at the bottom of the well.

Some users have had difficulty obtaining consistent results using this precipitation process. Many additional methods are available for recovering purified DNA from PCR reactions, and numerous kits for cleaning up PCR reactions can be purchased. We have not performed a comprehensive survey of these methods for their performance in microarray printing applications. The basic goal of this step is for the final DNA solution to be as free of contamination as possible; salts, detergents, glycerol, gelatin and particulate matter such as resins from columns can interfere at numerous later stages in the printing process. If these are used, extreme care must be taken to ensure that they are removed from the final DNA solution. Any alternative method for preparing DNA solutions for printing should be tested on a small scale, and taken through the entire printing and hybridization process to ensure that it is compatible. An alternate method that has been used successfully is described below.

Sephacryl purification: This method is based on using a simple gel filtration resin that retains salts and other small molecules while excluding PCR products. As the DNA elutes in water, it can simply be dried down, eliminating the need for precipitation. As mentioned above, great care must be taken to prevent the resin from contaminating the printing solutions. To purify 100  $\mu$ l PCR reactions, prepare Sephacryl S400 resin (Sigma S400 HR) by mixing well and diluting 2:1 in double-distilled H<sub>2</sub>O (ddH<sub>2</sub>O). Aliquot 100  $\mu$ l resin into 96-well filter bottom plates (Polyfiltronics UN800-PSC/mepp/D). Wash four times by adding 400  $\mu$ l ddH<sub>2</sub>O to plates and aspirating using a vacuum manifold (Polyfiltronics UNIVAC-S). Nest these plates in polypropylene 96-well plates and spin for five minutes at 4C at 2000 rpm (in a Beckman GS-6 centrifuge with a G.H. 3.8 rotor and Beckman Microplus carriers) to remove all remaining liquid. Add 100  $\mu$ l ddH<sub>2</sub>O and spin as above to test recovery. Add 100  $\mu$ l PCR product and spin into clean plates. Dry down in rotary evaporator and resuspend in 3x SSC as above.

### **Preparation of Slides**

The vast majority of our arrays have been printed on poly-L-lysine coated glass microscope slides. Many alternate coatings have been tried,

including numerous silanes, but none have as yet proven to be as simple and reliable as poly-lysine.

To prepare poly-lysine coated slides, place standard glass microscope slides (Gold Seal Products, catalog number 3010) in metal or glass racks (a wide variety of racks can be obtained from Wheaton) so that the slides are positioned vertically, and place the rack in an appropriately sized glass chamber. Make sure that you use slide racks that will fit in an available low-speed centrifuge as they must be spun dry at the end of the coating process, and it is best to avoid having to transfer the slides from one type of rack to another. Although most slides come in boxes labeled “pre-cleaned”, it is essential that they be stripped clean prior to coating. Prepare 500 ml alkaline wash solution by dissolving 50 g sodium hydroxide in 200 ml ddH<sub>2</sub>O and adding 300 ml 95% ethanol (again, never use 100% ethanol). Completely submerge the slides in this solution for at least two hours. After cleaning, extensively rinse the slides by five cycles of adding clean ddH<sub>2</sub>O, rocking for five minutes and rinsing. Keep the slides submerged in ddH<sub>2</sub>O; once the slides are washed, it is essential that they be exposed to air as little as possible, as dust particles will interfere both with the coating and printing processes. Prepare a coating solution containing (for 350 ml) 35 ml poly-L-lysine (0.1% w/v in H<sub>2</sub>O; Sigma P8920), 35 ml sterile-filtered PBS, and 280 ml ddH<sub>2</sub>O. Pour this solution into a clean glass chamber, and quickly transfer the slides from the final water wash into the coating solution. Rock the solution gently for one hour. Dunk the coated slides in a chamber containing clean ddH<sub>2</sub>O and plunge up and down five times. Transfer the slide rack to a low-speed centrifuge (place paper towels below the rack to absorb liquid) and spin the slides dry (5 minutes, room temperature, 1000 rpm in a Beckman GS-6 centrifuge with a G.H. 3.8 rotor and Beckman Microplus carriers). Wrap the dry slides in aluminum foil to keep out dust. Place the wrapped slide rack in a vacuum drying oven at 45C for 10 minutes with the vacuum on. Remove the slides from the foil and place in a clean plastic slide rack (wood and cork leave particles on the slides). Store at room temperature with the slide box tightly sealed. The slides are not immediately ready for printing, but must be “cured” for at least a few weeks to allow for the surface to become sufficiently hydrophobic (slides are usually best about one month after coating). The hydrophobicity of the coated slide surface is important in maintaining the small size of the printed DNA spots necessary for high-density arraying. A simple test for the readiness of a batch of slides is to take one slide and place 150  $\mu$ l H<sub>2</sub>O on the surface. Turn the slide 45 degrees to the horizontal and watch as the drop moves down the slide. If it leaves no noticeable trail, the slides are probably ready for use. However, the only reliable way to tell if a batch of slides is ready for printing is to print a test array on a few slides from all available batches (see spotting section below).

## Printing of DNA Microarrays

The arraying robot used in printing DNA microarrays is a variation of a standard “pick-and-place” robot. During standard arraying operations, a large number of microscope slides are placed on and secured to a platter. DNA samples in 3x SSC in 96 or 384 well microtiter plates are placed on a stand. The robot positions a cluster of specialized spring-loaded printing tips into adjacent wells of the DNA source plates filling the reservoir slot of each tip with approximately 1  $\mu$ l of DNA solution. The tips are then lightly tapped at identical positions on each slide leaving a small (<0.5 nl) drop of the DNA solution on the poly-lysine coated slide. After depositing DNA on every slide the tips are washed and dried and the process is repeated for the next set of DNA samples with the new spots offset a small distance relative to previous spots to produce a high-density grid. The spacing between spots is determined by the size of the deposited DNA droplet, which is a function of the sharpness and characteristics of the tip and the hydrophobicity of the poly-lysine surface. We routinely print arrays with spot-center to spot-center spacing of approximately 200  $\mu$ m and have achieved spacing approaching 100  $\mu$ m. At 200  $\mu$ m spacing the entire yeast genome can be printed in a 1.8 cm square area, and at 100  $\mu$ m spacing all of the approximately 75,000 human genes can be printed on a standard 1” by 3” microscope slide. With our current robot, each cycle (wash, dry, load, print on 110 slides) takes about two minutes, and we use a print-tip cluster with four tips spaced to fit in adjacent wells of standard 96 well microtiter plates. This results in a printing rate of 120 spots on 110 slides per hour. The cycle time can easily be cut in half with simple improvements in robotics, and we have a similarly sized print-tip cluster of 16 tips spaced for 384 well plates. This configuration results in a printing rate of approximately 1000 spots per 110 slides per hour, or six hours for the entire yeast genome. Expanding to 32 tips (and twice the area) will enable a set of all human genes to be spotted on each of 110 slides in about two days.

The printing tips: As the only part of the robot that touches the DNA and the slides, the printing tips are the most critical component of making DNA microarrays. The tips we use [see Figure 2] operate on the same principle as a quill pen; liquid is drawn up by capillary action and deposited when the tip makes contact with the slide surface. The tips are made by first cutting a thin slot in a cylinder of stainless steel. One end of the cylinder is sharpened to a cone with the slot passing through the point. The two sides of the slot are squeezed together until they are almost touching at the point, and the tips are further sharpened by hand. When the tips are immersed in the DNA solution, liquid is drawn up into the

reservoir. When the tips are tapped gently onto glass, DNA exits only at the point. While many other factors can influence the size of the spot produced during printing (hydrophobicity of the slides, humidity in the room), in general sharper tips produce smaller spots, so the tips must be kept as sharp as possible. Since we make all of our printing tips in a semi-manual manner, they do not behave identically. Furthermore, while good tips can survive hundreds of thousands of taps without noticeable loss of quality, they do eventually need maintenance. Thus, the first important step in printing good microarrays is to maintain high quality tips. Clean the tips by placing them in a bath sonicator (Branson 3200) with the tips floating, points submerged in clean water. Prime the tips by submerging the point in 3x SSC until it draws up liquid into the reservoir slot. Choose an initial set and place in the print head. Test the tips by doing a “test print”. A test print is a print run where a single slide is placed on the platter, the tips are submerged in a fresh test solution containing 100 ng/ $\mu$ l salmon sperm DNA in 3x SSC and are tapped a few hundred times on the single slide with the spot-center to spot-center spacing desired for the ultimate printing (the software for the arraying robot should have an option to do such a test print). After the test print, and examine the slide. The high salt content of the spotted DNA solution allows for easy visualization of the size and shape of each spot [see Figure 3]. Carefully examine the slide to ensure that: 1) each tip is depositing DNA on the slide for at least the same number of taps as there are slides to be printed, 2) that the spots are of the desired size and shape and stay roughly the same size and shape across the slide, and 3) that all the tips are producing spots of roughly the same size and shape. When a tip fails to print, the cause is either a failure of the tip to load DNA, or a failure to discharge DNA onto the slide. First, make sure that the tip is touching the glass when printing should occur. Second, make sure that the tip is loading DNA. This can be determined by removing the tip from the holder and slowly submerging it in the salmon sperm test solution described above while holding it up to the light. It should be possible to look through the slot to see if solution is being drawn into the slot. If the tip is not loading, make sure that the DNA solution is not too viscous (if prepared properly and made fresh this solution should be about as viscous as target DNA prepared as described above), and then clean out the slot either by continued sonication or using a thin (<100  $\mu$ m) stainless steel shim. If the tip is loading properly and is still not printing, it is likely that the tip is clogged. Remove the tip and examine it under a dissecting scope. The tip can be cleaned again either by sonication or by gently pulling a thin piece of metal through the end. Take care not to damage the tip's end. Make sure that the two ends of the tip are not touching, as this will prevent deposition of DNA.

If the spots are too large so that adjacent spots run together, the tips may not be sharp enough or they may be hitting the glass too hard. First

try to reduce the impact of the tap, and then try sharpening the tips with fine sandpaper or a sharpening stone under a dissecting scope. Large spots can also be caused by insufficiently hydrophobic glass or by surfactant contamination in the DNA solution. While it is aesthetically pleasing for the spots produced by each tip to be the same size, variation in the size and shape of spots is not a serious functional problem so long as the spots remain sufficiently large, and well-separated.

If everything looks good in this test print, the process should be repeated at least once, and then the printing should be initiated immediately. Once begun, the printing process involves only feeding plates to the robot and vigilantly monitoring the quality of the printing. As the printing tips behave best when in constant use, printing should be interrupted as infrequently as possible. As the plates are open while they are being printed and will evaporate rapidly (changing the viscosity of the solution and potentially adversely impacting spotting), it is not advisable to design a platter to feed multiple plates to the robot unless it is accompanied by a mechanism either to remove the plate cover immediately before use, or to cool the plates to minimize evaporation.

It is important that the printing process be closely monitored to make sure that spots are still being printed and that they are still an appropriate size and shape. [Note: often too much DNA is deposited on the first few slides; do not stop the print if the remaining slides look acceptable]. The most common problem is that a tip becomes clogged and stops printing. If this occurs, the print run should be interrupted immediately, and the problematic tip should be cleaned as described above. The test printing should be repeated until the tips are working well again. Another problem that occurs is that a dust particle or fiber gets stuck in the tip causing the spots to be irregularly shaped. This too should be cleared up by cleaning the tip. A different type of problem occurs if the tips become dulled during the print, and the spots will become larger, potentially causing adjacent spots to overlap. This can be solved either by replacing the tip or by sharpening it.

After printing is completed, the area containing spots should be delineated in some manner, as the spots will not be visible once the salt has been washed away. We generally do this by marking the boundaries of the area with a diamond tipped etching device. This pen should also be used to label the slides, as most other forms of labeling (ink pens and labels) will be removed by the solvents used in post-processing. After the slides are labeled they should be removed from the platter and stored in a clean slide box in a cool, dry place. The slides are stable at this stage for at least a month, and need not be post-processed immediately.

## **Postprocessing of Slides**

Four steps are required to prepare printed arrays for hybridization: rehydration and snap-drying, UV cross-linking, blocking, and denaturation.

Rehydration: The spotting process does not, in general, leave DNA evenly distributed throughout the spot. To more evenly distribute the DNA, the spots (which dry rapidly during arraying) are rehydrated and snap-dried. This is a delicate process, as insufficient rehydration can produce irregular spot sizes and can reduce overall hybridization intensity, while excessive rehydration can cause spots to fuse. Invert the metal block from a standard benchtop heating unit (VWR Scientific Products H2025-1) to produce a flat metal surface, and set the block to approximately 80C. Fill the reservoir of a plastic slide hydration chambers (Sigma H6644) with hot tap water (approximately 50-60C). Process each slide separately by placing it, array side down, over the reservoir, allowing the vapor to hydrate the elements until they all are glistening (this usually takes between 5 and 15 seconds). Immediately place the slide, array side up, on the heat-block until the spots have dried (usually about 5 seconds). This is a process that requires practice. More consistency can be achieved by carrying out the hydration over a temperature controlled water bath, although even then the amount of time required to hydrate different arrays, even within one batch, can vary considerably. Some groups prefer to hydrate at lower temperatures for longer periods of time. We have not tried this in our lab, but it is likely that the extreme temporal sensitivity of this process can be reduced at lower temperatures.

UV Crosslinking: Once the arrays have been rehydrated and snap-dried, the DNA is crosslinked to the slide by UV irradiation. While not essential, this step can increase the amount of hybridizable DNA stably attached at each spot, especially when the DNA concentration in the spotting solution is low. UV crosslinking is carried out by placing the slides, array side up, on plastic-wrap, in a Strategene UV 1800 Stratalinker. Switch from timed mode to total energy mode and expose the slides to 60 mJ of energy (as defined by the Stratalinker).

Blocking: The most critical step in post-processing is the blocking step, where the remaining free lysine groups are modified to minimize their ability to bind labeled probe DNA. If these groups are not blocked, labeled probe will bind indiscriminately and non-specifically to the surface and will produce excessively high background. We block by acylation with succinic anhydride (Aldrich 23,969-0). The charged amino group of the lysine carries out a nucleophilic attack on one of the carboxyl carbons, forming an amide bond and exposing a free carboxylate group on the other end of the chain. In addition to converting the amino groups of

the lysines to amides, this new moiety creates a negatively charged surface which further reduces non-specific binding of DNA.

As the blocking reagent has limited stability in solution, it is important that this process be carried out quickly. Find a glass chamber, and a slide rack with a handle that will fit easily in this chamber. Place slides in this rack, and place the rack in a fume hood right next to the chamber. The chambers we use hold 350 ml of solution, scale the following reagents to the volume you will need to completely cover the slides. Measure out 335 ml of water and pour into a 500 ml glass beaker. Mark the level of the solution with a pen on the outside of the beaker. Remove the water and rinse the beaker with 95% ethanol. Carefully dry the beaker with a paper towel, place a small magnetic stir bar in the beaker, and place on a stir-plate in a fume hood. Measure 15 ml 1M sodium borate pH 8.0 and place next to the beaker. Weigh out 6g of succinic anhydride and place in the beaker. Quickly fill the beaker to the 335 ml mark with 1,2-methyl pyrrolidinone (Sigma M6762) and begin stirring at high speed. (This solution should be clear at this point, although it will yellow slightly during the blocking reaction; if it is yellow or orange initially, the solvent has gone bad and should not be used. Immediately after of the succinic anhydride is dissolved, add the sodium borate and stir until the solutions mix (about 5 seconds). Pour the solution into the empty slide chamber, immediately dunk the slides into the solution and plunge the slides up and down 5 times. Cover the rack (with a glass cover or with aluminum foil) and shake gently for 15 minutes.

Denaturation: While (or before) the blocking reaction is taking place, prepare a large beaker or glass chamber (large enough so that the slide rack used in the blocking step can easily be submerged) filled with boiling ddH<sub>2</sub>O. The volume should be at least twice that of the blocking solution. Turn off the heat immediately shortly before the blocking reaction is complete, and when the water is no longer bubbling, quickly transfer the slides to the water, plunge the slides up and down 3 to 5 times and allow the slides to sit for 2 minutes. Immediately transfer the slides to a chamber containing 95% ethanol (again, never use diluted 100% ethanol as the benzene used to dehydrate this ethanol fluoresces), and again plunge the rack up and down 3 to 5 times. Transfer the slides to a centrifuge and spin dry. The slides are now ready for hybridization.

### **Preparation of Probe**

Isolation of RNA: We have had success with RNA prepared by a variety of methods including hot phenol extraction, guanidinium isothiocyanate stabilization and others; such methods have been reviewed elsewhere and will not be discussed here. We have had good success using

Invitrogen's FastTrak 2.0 mRNA isolation kits (Invitrogen K1593-02) for both yeast and human samples. The main criterion is the production of high quality RNA. For gene expression studies, the best results have been achieved using purified polyadenylated RNA as the template for reverse transcription, although, at least with yeast, total RNA has been used with good results. A good experiment requires at least 1  $\mu\text{g}$  (and preferably more) of poly-A RNA (or 100  $\mu\text{g}$  total RNA) per sample as template for synthesis of fluorescent cDNA probe. Methods for amplifying smaller amounts of RNA are being examined for their compatibility with the methods discussed here.

Labeling: A variety of methods have been used for preparation of labeled probe from RNA. In most cases we label cDNA directly by incorporating fluorescently labeled nucleotides during oligo-dT primed reverse transcription. Where appropriate, random-primed cDNA synthesis can be used for labeling, although it may not be appropriate for all array designs. For example, arrays composed of sequences from the 3' ends of transcripts are more suited for analysis of oligo-dT primed cDNA probes, while arrays whose elements consist of entire cDNA sequences or predicted open reading frames are well suited for probes prepared by either priming method. Samples can also be labeled during second-strand synthesis after first-strand synthesis of unlabelled cDNA. Protocols for all three methods are given below.

Fluorescent Dyes: A large number of fluorescently labeled deoxyribonucleotides are now commercially available. Excellent results have been obtained with Cy3- and Cy5-dUTP; they are spectrally well separated, can be incorporated with high specific activities with a variety of enzymes, and brightly fluoresce when dry which simplifies the image acquisition process. Other fluors we have used with good results include: R110 (Perkin Elmer), TAMRA (Perkin Elmer) and SpectrumOrange (Vysis).

oligo-dT and random primed labeling: All reagents should be sterilized and DEPC treated. Start with RNA (1-2  $\mu\text{g}$  poly-A RNA or 100-200  $\mu\text{g}$  total RNA) in 10  $\mu\text{l}$   $\text{H}_2\text{O}$ . This solution should include any control RNAs (see section below). Add 2.5  $\mu\text{l}$  (6  $\mu\text{g}$ ) primer (20-mer oligo-dT, anchored oligo-dT, random hexamer or random nonamer: 2.5  $\mu\text{g}/\mu\text{l}$  in  $\text{H}_2\text{O}$ ). Incubate in a 70C water bath for 10 minutes to denature RNA. Place on ice. Mix 2  $\mu\text{l}$  (400 activity units) MMLV reverse transcriptase (from GibcoBRL SuperScriptII Rnase H- Reverse Transcriptase kit; catalog number 18064-014), 6  $\mu\text{l}$  5x first-strand buffer (from SuperScript kit; 250 mM Tris-HCl, pH 8.3, 375 mM KCl, 15 mM  $\text{MgCl}_2$ ), 3  $\mu\text{l}$  0.1 M DTT



(also included in kit), 6  $\mu$ l 5x nucleotides (2.5 mM dATP, 2.5 mM dCTP, 2.5 mM dGTP, 1.0 mM dTTP), 3  $\mu$ l 1mM Cy3 or Cy5 linked dUTPs (Amersham). Add to RNA-primer solution to a final volume of 30  $\mu$ l. Incubate at 42C for one hour. Add 1  $\mu$ l SuperScriptII reverse transcriptase and incubate at 42C for an additional hour. Add 1.5  $\mu$ l 20 mM EDTA to stop the reaction and 15  $\mu$ l 0.1 M NaOH to degrade the template RNA. Incubate at 70C for 10 minutes. Immediately add 15  $\mu$ l 0.1 M HCl to neutralize the sample (the fluors, Cy5 in particular, are sensitive to high pH). Place the samples in a Microcon 30 microconcentrator (Amicon) and add 500  $\mu$ l TE (10 mM Tris, pH 8.0, 1 mM EDTA). Spin 5 to 10 minutes in a benchtop microcentrifuge at high speed to a volume of 10 to 20  $\mu$ l. Discard the flow-through. Add 500  $\mu$ l TE and repeat the spin. After the second spin, the probe retained by the Microcon should be significantly brighter than the flow-through. This is a very strong indicator of a successful labeling reaction. If no color is retained, it is very unlikely that the labeling will produce a good hybridization result. Collect the probes by inverting the filter and spinning for 1 minute. Mix the appropriate Cy3 and Cy5 labeled probes and bring up in 500  $\mu$ l TE. Spin to final volume of 8  $\mu$ l and collect probe (which should now be purple). Add 2  $\mu$ l of a blocking solution containing 10  $\mu$ g/ $\mu$ l oligo-dA 20-mers, 10  $\mu$ g/ $\mu$ l yeast tRNA and, for human probes, 10  $\mu$ g/ $\mu$ l human cot-1 DNA (GibcoBRL catalog number 15279-011). Add 2.1  $\mu$ l 20x SSC (to a final concentration of 3.5x) and 0.4  $\mu$ l 10% SDS (to 0.3%). Be careful that no SDS is present on the outside of the pipette tip when SDS is added to the sample, as extra SDS can adversely effect the hybridization. Denature the sample by placing in a 100C water bath for 1 minute. Allow the sample to cool on the benchtop. For human samples the sample should be left at room temperature for 30 minutes to allow for cot-1 hybridization to repetitive elements. For yeast, the probe is ready for hybridization as soon as it cools. 12  $\mu$ l is an appropriate volume for arrays around 2cm x 2cm. A larger final volume should be used for arrays covering a larger area.

Klenow labeling of cDNA: Prepare a cDNA template using the above procedure; except use a 5x nucleotide solution containing 2.5 mM of each dNTP (i.e. final concentration of 0.5 mM each dNTP in the labeling reaction) and do not include fluorescent nucleotides. Clean up the sample as above to a final volume of 20  $\mu$ l. Add 3  $\mu$ l of a random hexamer or nonamer (4  $\mu$ g/ $\mu$ l). Heat to 100C for one minute. Let cool on benchtop for 5 minutes. Add 5  $\mu$ l 10x buffer (100mM Tris-HCl, pH 7.4, 50 mM MgCl<sub>2</sub>, 75 mM DTT), 5  $\mu$ l 10x nucleotides (0.25 mM dATP, 0.25 mM dCTP, 0.25 mM cGTP, 0.09 mM dTTP), 1.5  $\mu$ l 1mM Cy3 or Cy5 dUTP (Amersham), 2  $\mu$ l (10 activity units) exonuclease free klenow (USB,

catalog number E700572), and 13.5  $\mu\text{l}$  ddH<sub>2</sub>O to a final volume of 50  $\mu\text{l}$ . Incubate at 37C for 4 hours. Clean up and prepare for hybridization as above.

### **Hybridization**

Spin the probe sample for one minute in a benchtop microcentrifuge at high speed to pellet any particulate matter. Pipette the probe onto the center of the array, being careful not to dislodge any pelleted material in the tube and to avoid forming bubbles. Place a cover slip (large enough to cover the entire array surface) over the probe, again being careful to avoid the formation of bubbles. Pipette three 5  $\mu\text{l}$  drops of 3x SSC on a separate part of the slide to provide humidity in the hybridization chamber and thus ensure that the probe mixture does not dehydrate during hybridization). Place the slide in a sealed chamber and submerge in a 65C water bath for 4 to 16 hours.

### **Post-hybridization**

Following hybridization, carefully dry the outside of the slide chamber and remove the slide. Place the slide in a slide rack submerged in the wash solution 1 (2x SSC, 0.1% SDS) with the array face of the slide tilted down so that, when the cover slip drops off, it does not scratch the array surface. After the cover slip comes off, plunge the slide up and down in this solution 2-3 times and transfer to a new slide rack submerged in wash solution 2 (1x SSC). Be careful to minimize the transfer of wash solution 1 to the second chamber as SDS can interfere with slide imaging. Gently rock the chamber for 2 minutes, transfer the slide to wash solution 3 (0.2x SSC) for 2 more minutes, then spin dry the slide. The wash solutions should all be at room temperature. Altering the temperature and/or time of these steps will change the stringency of the hybridization.

### **Controls**

Because of the highly parallel nature of microarray experiments, it is possible, and highly valuable, to include a large number of controls on every array and in every hybridization. Some useful controls are described below:

Doped RNAs: It is very useful to dope in known amounts of control RNAs at various stages of the probe preparation process to monitor and assess various aspects of each experiment. Clearly, doped RNAs should be chosen to have no cross hybridization with RNA from the organism being studied, but should be similar in their general

characteristics (GC content, length, poly-A tails). Many yeast genes can be used as controls on human gene expression arrays, as there is limited homology at the nucleotide level between yeast and humans. Hundreds of such genes can be spotted on human arrays with only a nominal increase in the time required for arraying. By cloning these genes into bacterial plasmids with phage RNA polymerase binding sites and artificial poly-A tails, large amounts of RNA can be made for each gene for incorporation at various experimental stages. Such controls can be used in a number of ways. For example, RNA from a number of such controls can be doped in at various ratios between the Cy3 and Cy5 samples and at varying total concentrations to assay the stringency of the hybridization, the calibration of output ratios to input ratios and the sensitivity of detection.

RNA quality controls: The amount of a given nucleic acid species that will bind to a complementary target element can be sensitive to changes in the size of the complementary region. Thus, the intensity of the hybridization signal for a given cDNA can vary with differences in RNA quality. This can confound microarray hybridization results as, for example, an apparent non-unity ratio can be observed when equal amounts of a given cDNA derived from RNA of differing quality are compared. Judicious choice of target elements can minimize this problem; for oligo-(dT) primed cDNA probe, target elements restricted to the 3' ends of genes will be minimally sensitive to RNA quality. RNA quality can be easily monitored with appropriate control elements. For example, target elements of DNA fragments of 100-200 basepairs tiled across an entire gene can be used to assay the relative RNA quality of that gene in the samples being compared (degraded RNA, labeled by oligo-dT primed cDNA synthesis, will give a proportionally weaker signal for the more 5' target elements. The experimental step that is the source of problems with RNA quality may be identified by making such tilings of genes used as doped controls, and then adding the corresponding RNA at various stages of the RNA and probe preparation process. For example, RNAs can be doped in when cells are harvested, when poly-A RNA is isolated, when labeled probe is made, and immediately before hybridization.

Blocking controls: As discussed above, most hybridization reactions require the addition of cold DNA to block hybridization due to non gene-specific sequence similarities between probe and target, especially those due to repetitive elements. Control target elements corresponding to such potentially confounding sequences can be used to monitor the success of the blocking. For example, on human gene expression arrays, a human cot-1 DNA spot should have little or no hybridization signal if repetitive elements have been successfully blocked.

Other useful negative controls include oligo-dA, oligo-dT, plasmid DNA and human genomic DNA.

Normalization controls: As it is difficult to equalize accurately the amount of RNA used to prepare labeled probe for multiple samples, it is useful to have control target elements whose hybridization signal is expected to be the same for both samples used in a hybridization. For example, for yeast gene expression experiments, target elements of total yeast genomic DNA can be used to normalize the signal between the two samples. Human gene expression experiments are more complicated in this regard; as only a small fraction of the human genome is expressed, the hybridization signal to total human genomic DNA should be very low.

### **Slide Imaging and Data Reduction**

After hybridization, a fluorescent image [see Figure 4] of the array must be acquired for both fluorescent dyes used. The device we use is a laser scanning confocal microscope (scanner). The scanner has a laser (or lasers) producing light with a wavelength appropriate for the excitation spectra of the dyes being used (for Cy3 this is green light around 540 nm; for Cy5 red light around 650 nm). The light passes through a standard microscope objective and illuminates a single point on the slide. The emitted light gathered by the objective passes through a series of filters (to remove the excitation beam), a collimating lens, and a pinhole (to minimize noise; this makes it a confocal device) and is quantified in a photomultiplier tube. The slide is rapidly scanned over the laser beam and a raster image of the array is acquired. It is possible to acquire a similar image using a fixed slide and a pair of galvanometers to position the beam at all possible points on the array surface, or using a wide-field illuminating source and a CCD camera. We have found that the confocal scanning device gives the best results as the signal to noise ratios are superior to those achieved with galvanometer or CCD based devices.

As discussed above, the experimental goal is to measure, for each spot on the array, the relative amount of fluorescence from each fluor used. There are many ways to accomplish this, the most obvious being segmenting the image into areas (boxes) corresponding to each spot and then determining the average fluorescence intensity in each box for each fluor. The relative amount of each element in the two samples is just the ratio of the average fluorescent intensities. This kind of calculation is confounded somewhat by issues involving the compensation for the non-zero fluorescent background observed on most arrays. We have developed software that implements a number of methods for accurately estimating this background, as well as more complex methods for directly

determining the relative fluorescent ratios. This software will be made public along with plans for an inexpensive laser scanning confocal microscope as discussed above.

### **Array Problems**

Experiments using DNA microarrays, can encounter technical problems at any step. The opportunity, provided by the microarray format, to include numerous internal controls, can greatly facilitate the recognition and correction of many kinds of problems. Indeed, even severely flawed experiments often provide plenty of useful new information. Still, it is wise, whenever possible, in designing an experiment, to plan so that any step of the experiment can be repeated without having to repeat the experiment in its entirety. For example, it is always wise to prepare more mRNA and more arrays than would be necessary if all went well.

Few microarray images are perfect. Many problems that we have encountered are difficult to trace, and do not manifest themselves in detectable ways until the array is scanned. It is advisable to perform multiple levels of quality control before hybridizing valuable experimental samples. Batches of poly-lysine coated glass should be tested by printing small arrays that are taken through the entire post-processing and hybridization process. Only batches of slides that have been used successfully for such small arrays should be used for “real” array printings. Once an array has been printed, post-processing should be performed on small batches of arrays and the success of the post-processing should be determined by performing a test hybridization with probe prepared from RNA that is known to produce good results. Finally, all arrays should be scanned prior to hybridization to eliminate slides with obvious blemishes (scratches or anything else that damages the poly-lysine coating can be detected as differences in the background fluorescence of unhybridized arrays) and to eliminate arrays with improperly high fluorescence prior to hybridization. Poly-lysine and some other coatings appear to fluoresce weakly at wavelengths used to detect Cy3; for unknown reasons this fluorescence varies from array to array and can occasionally reach intensities that will interfere with the hybridization signal. In addition, printed DNA also fluoresces weakly at these wavelengths and this fluorescence also varies from array to array. If either of these pre-hybridization sources of fluorescence are at levels near those of the signal on hybridized arrays, these arrays should be discarded. Some of the more common problems [see Figure 5] are discussed below.

Spot shape problems: One common problem is that some or all of the spots on the array have “comet tails”, where the fluorescence of a spot

tails off in a direction that is consistent across the array and where the color of a spot is reflected in the material tailing away from the spot. This phenomenon is likely caused by insufficiently rapid immersion of the slides in the succinic anhydride blocking solution, resulting in DNA binding to the slide as it is being washed off the spot. To the extent that spots (even with comet tails) can be resolved from each other arrays with comet tails are still usable. Another problem related to general spot appearance are “donut holes”. When DNA is deposited using tips that those described above, it is not left in a uniform circle; far more material accumulates near the outer edge of the spot which leaves a relatively empty area in the center. This phenomenon can usually be eliminated by longer rehydration. Even more so that comet-tails, donut holes do not make an array unusable and in fact only slightly compromise data quality (by reducing the fluorescent area being analyzed).

High Background: There is a nebulous class of array problems marked by abnormally high fluorescent backgrounds. High backgrounds are observed across the entire hybridized surface as well as in localized regions on the array. Many potential sources of high background can be described, but it is difficult to diagnose the source of a particular type of high background. In attempting to diagnose background problems, it is important to determine if the background is restricted to the area of the slide that was exposed to probe. Unusually high background that covers the whole slide can usually be avoided if arrays are systematically scanned prior to hybridization. High background restricted to the hybridized area indicates that labeled probe is adhering to the slide surface such that it is not being removed during the washing steps. Two possible causes of such an effect are insufficient blocking of the amino groups of the poly-lysine coating, or precipitation of the labeled probe during hybridization. High background from these sources is usually accompanied by a phenomenon known as “anti-probe”, where there the high fluorescent background occurs only in places where no DNA was spotted. This likely reflects the diminished capacity of the poly-lysine groups in areas containing spotted DNA to bind additional DNA. In some cases, the background is clearly worse towards one side of the array, indicating a relationship to how the slide was positioned in a liquid medium (most likely the blocking solution). Often, high background is observed near the edges of the coverslip, likely indicating that the probe dehydrated during the hybridization. This can be corrected by adding a larger volume of 3x SSC to the hybridization chamber and ensuring that the chambers are well sealed. It is also possible for uniformly high backgrounds to result from poor probe preparation, as small probe fragment and/or unincorporated fluorescent nucleotides can bind to the surface and contribute to background without contributing to the hybridization signal.

Locally weak signal: Most of the phenomena that produce high background are also accompanied by a diminution of specific signal, as much of the activity of the probe is being sequestered by the array surface. However, locally decreased signal that is not accompanied by high background is also observed. Our experience suggests that this is due to a local decrease in the thickness of the probe layer in that region. Air bubbles and inhomogeneities in the flatness of the slide and/or cover slip can cause such an effect. Care in choosing and positioning the cover slip can minimize this problem.

## **Conclusion**

DNA microarrays are tools for exploration and discovery on a genomic scale. Biologists have had the ability, for many years, to survey differential expression of a large number of genes under a small number of conditions, or a small number of genes under a larger set of conditions. DNA microarrays make it practical, for the first time, to survey the expression of thousands of genes under thousands of conditions. For organisms whose entire genome is known, (for example, *Saccharomyces cerevisiae*), this technology makes it possible to study the expression of all of the genes at once. For humans, the only limitation to applying the same approach is the number of known gene sequences; efforts to identify and sequence a complete set of human transcripts are likely to approach completion in the next few years. In the past, the limitations of our experimental tools have forced most investigations of gene expression patterns to be designed around specific hypotheses, framed on the basis of the very limited knowledge we already have. We can now readily design experiments to explore systematically the vast uncharted areas of biology. The use of DNA microarrays as tools for exploration and discovery relies on their low cost, flexibility and simplicity. Thus, in continuing to explore the many possible variations on this technology, it is useful to keep in mind that for many applications, where discovery is the goal, simplicity and low cost are critical. In our own efforts to develop this technology, a guiding principle has been to keep the economic, logistical and psychological barriers to each genomic-scale experiment as low as possible, so that a single laboratory can carry out hundreds to thousands of genome-wide surveys in a year.

Using a single microarrayer and scanner as described here, a single laboratory can, in one year, make millions of quantitative measurements of differential expression, or other properties, of specific genes under specific conditions. The opportunity to collect this unprecedented quantity of gene expression information is sufficiently new that appropriate mechanisms and standards for publishing, distributing, and archiving the results lag far

behind our ability to produce them. Furthermore, the analytical and computational tools for organizing, visualizing and interpreting this huge volume of information are still in the earliest stage of development. Far from being a deterrent to genome-wide exploration of gene expression and other functions, this new frontier presents a challenging new opportunity for pioneering biologists.



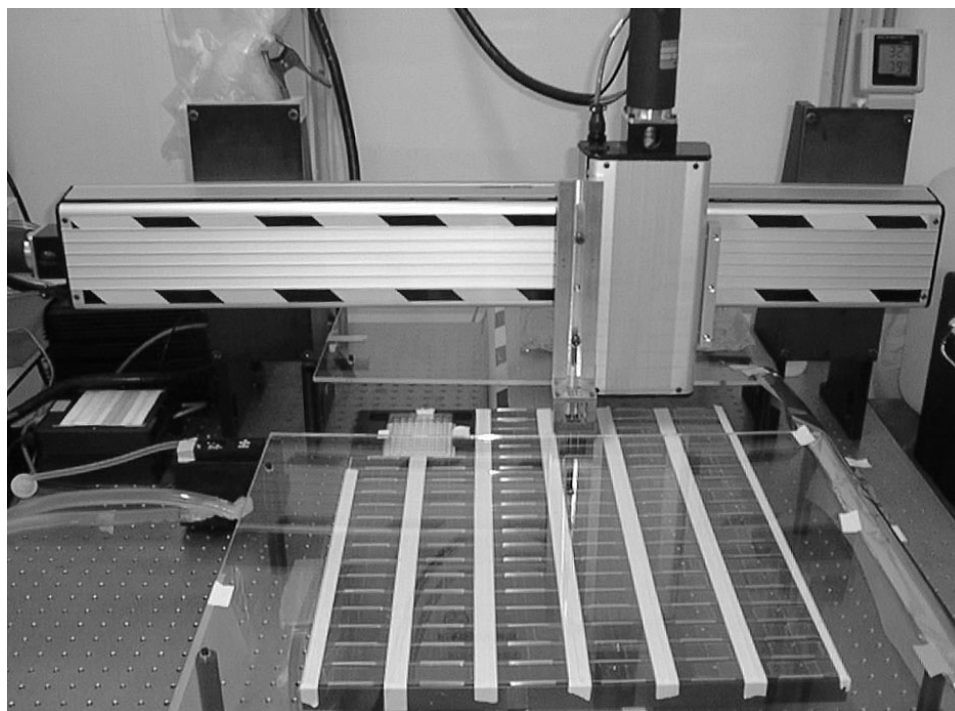


Figure 1. **The Arrayer.**

Photograph of Brown lab arraying robot primarily built by Dari Shalon. Movies of the robot printing an array can be viewed at <http://cmgm.stanford.edu/pbrown>.

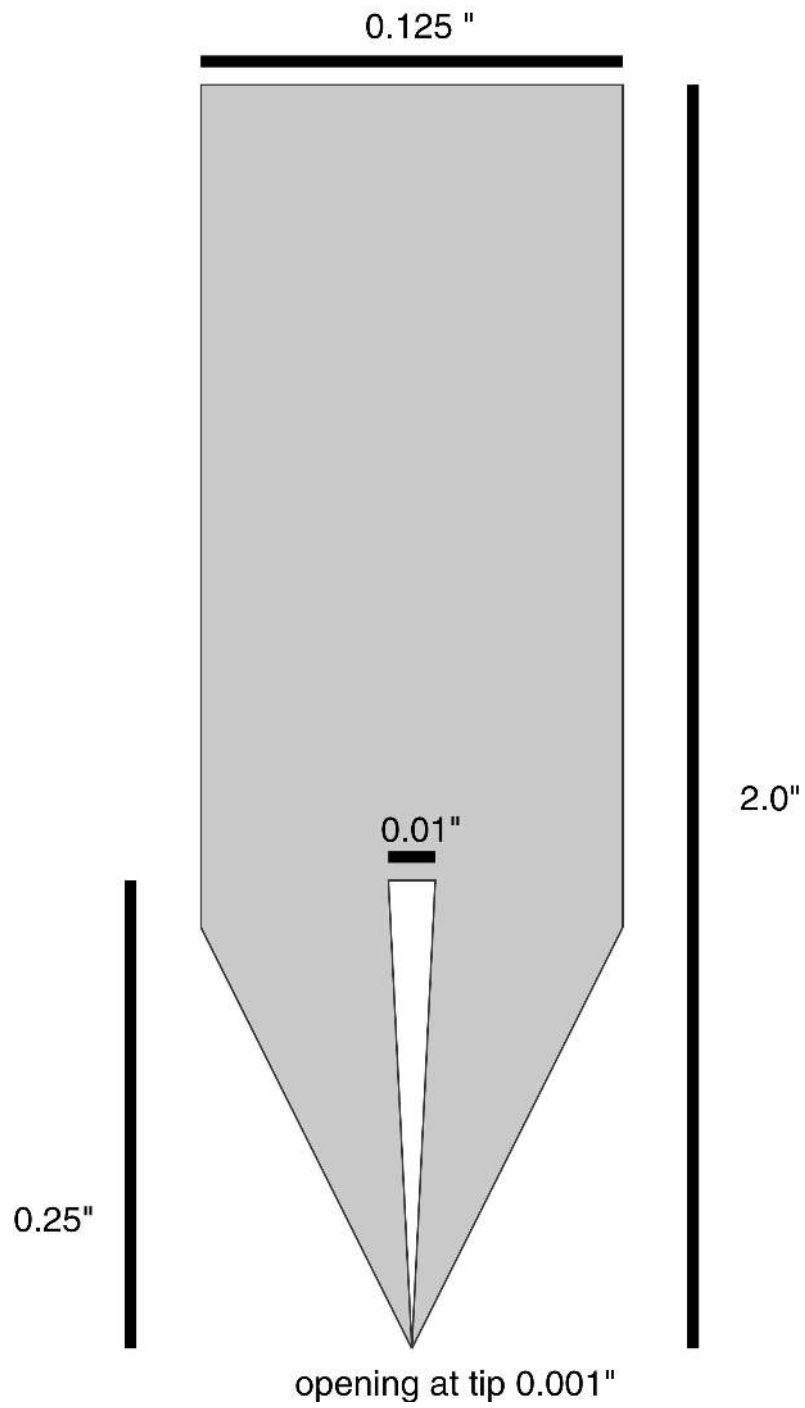


Figure 2. **Printing Tip.**

Diagram (not to scale) of printing tip.

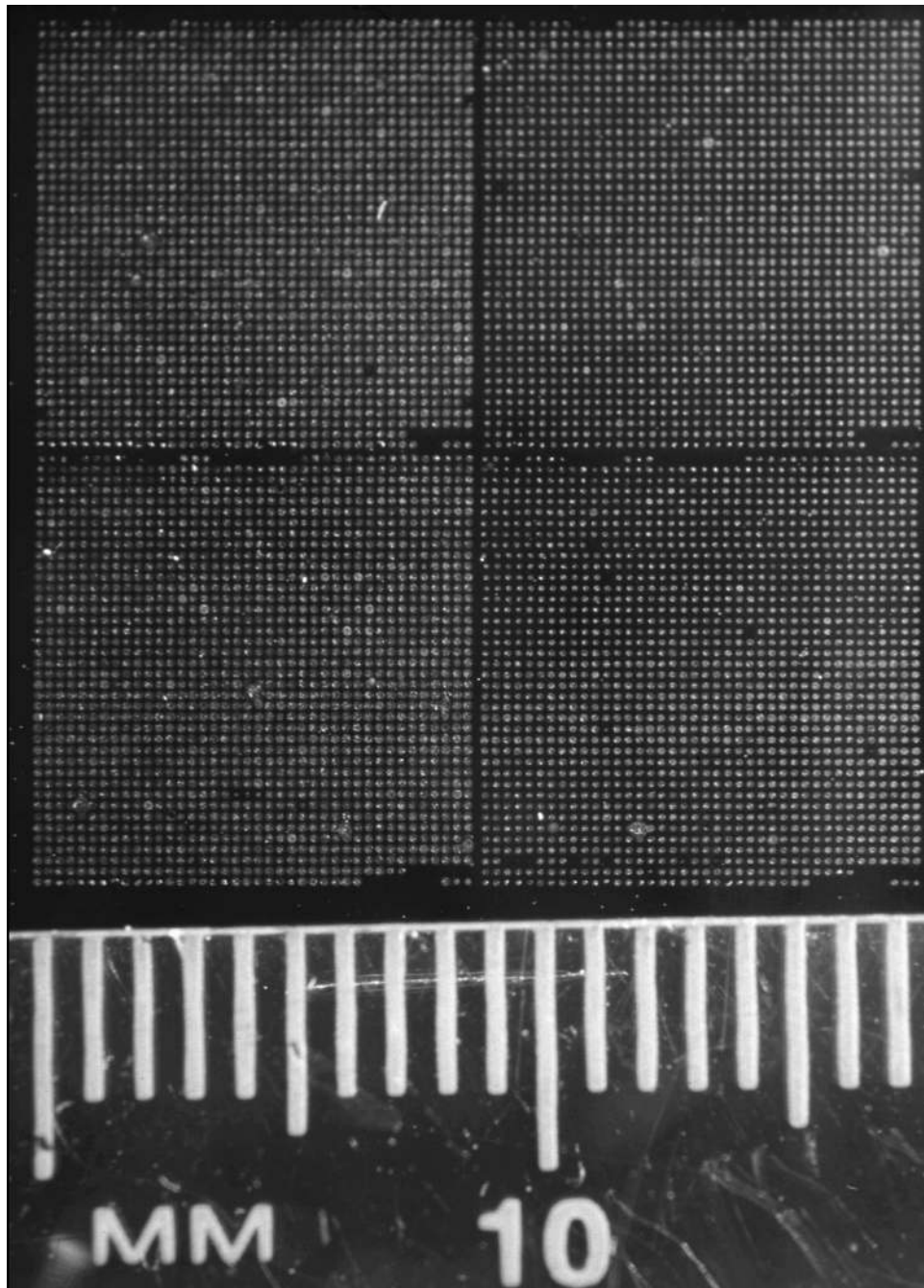


Figure 3. **A Microarray.**

Light micrograph of a DNA microarray with targets representing essentially all open reading frames in the yeast *Saccharomyces cerevisiae*. After printing, the spots are visible due their high salt content until they are post-processed. Each of the four quadrants of the image represents the spots deposited by one tip. Photograph courtesy of Joseph DeRisi.

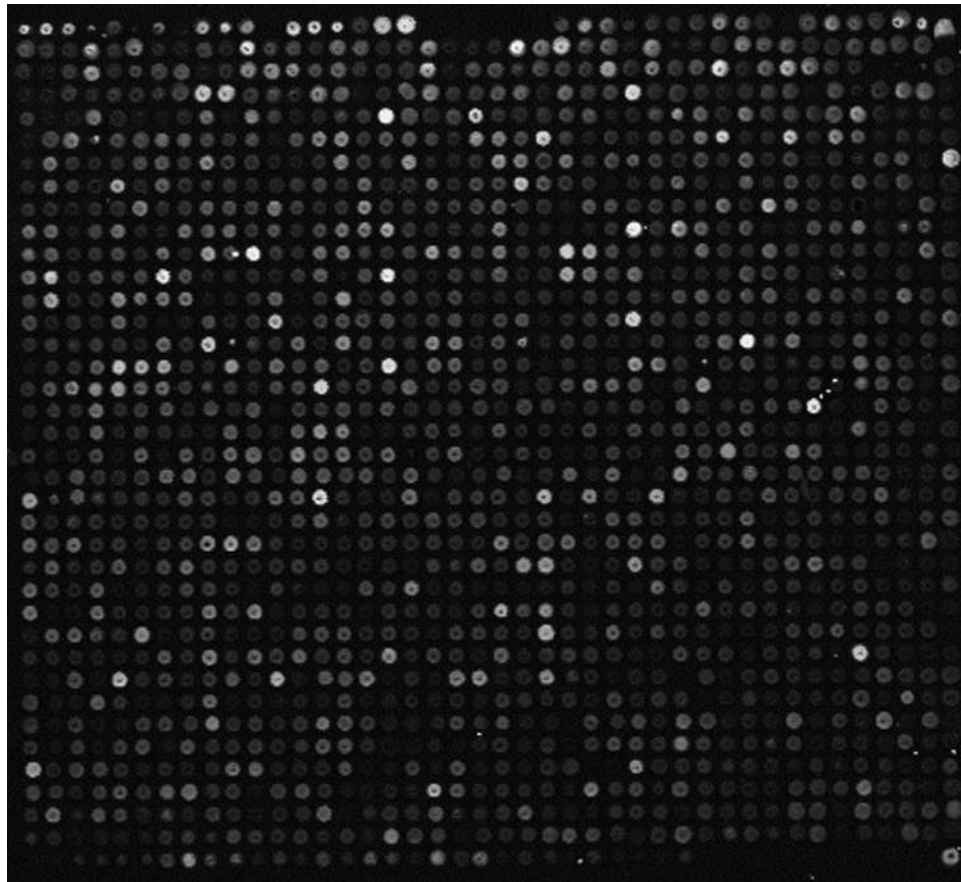


Figure 4. **Scanned Image.**

Image from laser scanning confocal microscope of slide pictured in Figure 3 (only one quadrant shown). A gray value is assigned to each pixel based on the observed fluorescence intensity (black represents no fluorescence).

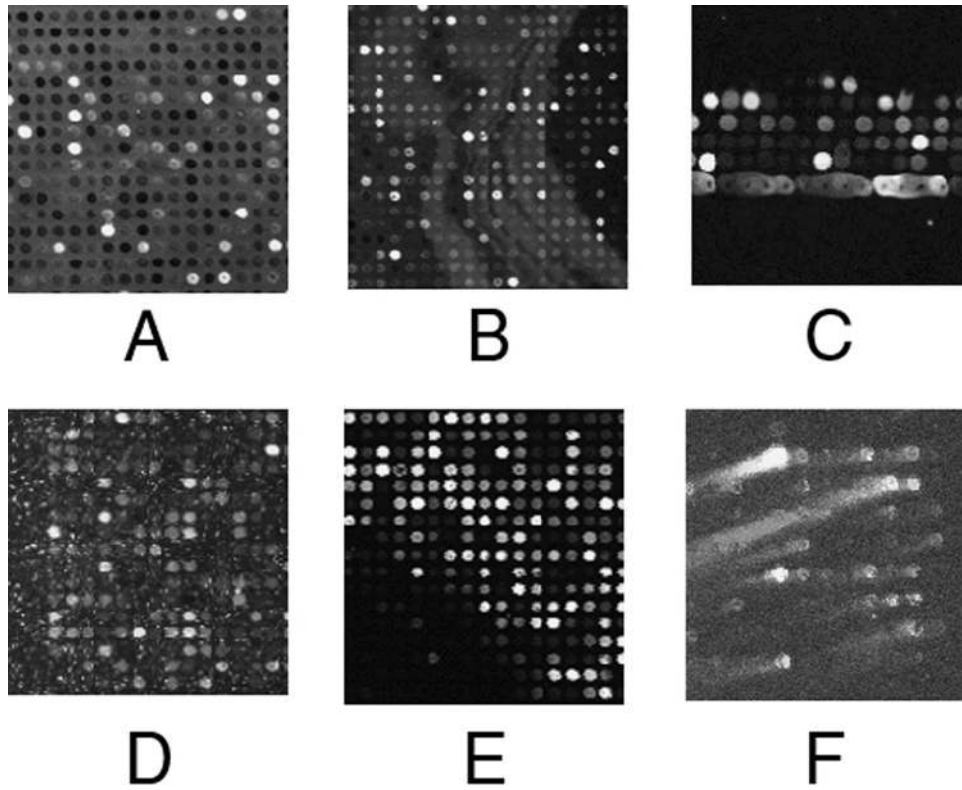


Figure 5. **Array Problems.**

Representative problems A) Anti-probe, B) Locally high background, C) Spot overlap, D) Precipitate, E) Locally low signal and F) Comet-tails.