

## **A Concise Guide to cDNA Microarray Analysis – II**

Priti Hegde, Rong Qi, Kristie Abernathy, Cheryl Gay, Sonia Dharap, Renee Gaspard, Julie Earle-Hughes, Erik Snesrud, Norman Lee, and John Quackenbush<sup>†</sup>

Additional protocols developed by Benjamin Schroder and Ivana Yang

The Institute for Genomic Research, Rockville, MD 20850

Microarray expression analysis has become one of the most widely used functional genomics tools. Efficient application of this technique requires the development of robust and reproducible protocols. We have optimized all aspects of the process, including PCR amplification of target cDNA clones, microarray printing, probe labeling, and hybridization, and we have developed strategies for data normalization and analysis.

<sup>†</sup>Address correspondence to:  
John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Drive  
Rockville, MD 20850  
<johnq@tigr.org>

## **Introduction**

Recently, a variety of techniques including SAGE (12), differential display (5), oligonucleotide arrays (6), and cDNA microarrays (8), have been developed that allow mRNA expression to be assessed on a global scale, allowing the parallel assessment of gene expression for hundreds or thousands of genes in a single experiment. The most common use of these is for the determination of patterns of differential gene expression, comparing differences in mRNA expression levels between identical cells subjected to different stimuli or between different cellular phenotypes or developmental stages. Microarray expression analysis (8) has a number of features that have made it the most widely used method for profiling mRNA expression. DNA segments representing the collection of genes to be assayed are amplified by PCR and mechanically spotted at high density on glass microscope slides using relatively simple *x-y-z* stage robotic systems, creating a microarray containing thousands of elements. Microarrays containing the entire set of genes from a microbial genome or tens of thousands of eukaryotic cDNA clones can be easily constructed. The microarrays are queried in a co-hybridization assay using two or more fluorescently labeled probes prepared from messenger RNA from the cellular phenotypes of interest (10). The kinetics of hybridization allows relative expression levels to be determined based on the ratio with which each probe hybridizes to an individual array element. Hybridization is assayed using a confocal laser scanner to measure fluorescence intensities, allowing simultaneous determination of the relative expression levels all the genes represented in the array.

Efficient expression analysis using microarrays requires the development and successful implementation of a variety of laboratory protocols and strategies for fluorescence intensity normalization. The process of expression analysis can be broadly divided into three stages:

- I. Array Fabrication
- II. Probe Preparation and Hybridization
- III. Data Collection, Normalization and Analysis

Below we present protocols that we have standardized and that have been used regularly in our laboratory for microarray analysis. The procedures described in this article have been tested and refined over the past year and have been optimized using hybridization of RNA derived from cell lines to give reproducible and consistent results. It should be noted that a number of alternative protocols have been published (4) or are available via the World Wide Web (see for example, Table 1), but the system that we describe here has a number of advantages over these. In particular, the combination of printing, labeling, and hybridization conditions that we have derived have allowed a significant reduction in the quantity of starting total RNA required for analysis.

### **I. Array Fabrication**

Microarrays are constructed by arraying PCR amplified cDNA clones or genes at high density on derivatized glass microscope slides. For the analysis of expression in most eukaryotes, expressed sequence tag (EST) data represent the most extensive data for gene identification. ESTs are single-pass, partial sequences of cDNA clones, and they have been used extensively for gene discovery and mapping in humans and other organisms. The EST approach has been widely adopted; more than 71% of all GenBank entries and 40% of the individual nucleotides in the database are derived EST sequences (9).

Generally, cDNA clones are selected to represent as many unique transcripts as possible. There are a number of analyses of these data that attempt to identify unique human transcripts within the EST data,

the two most widely used are UniGene (1) (<<http://www.ncbi.nlm.nih.gov/UniGene/>>) at the National Center for Biotechnology Information (NCBI) and the TIGR Human Gene Index (7) (HGI; <<http://www.tigr.org/tbd/hgi/hgi.html>>) at The Institute for Genomic Research (TIGR). While both UniGene and HGI are based on EST clustering, the TIGR protocol assembles the ESTs within the clusters, producing Tentative Human Consensus (THC) sequences.

We selected cDNA clones for array construction using the TIGR HGI as part of a program to assemble a 30,000 gene clone set. THCs were chosen for representation in the clone set with preference given to those containing known genes or those with mapped positions; additional THCs were selected to represent as yet uncharacterized transcripts. For each target THC, a single cDNA clone was identified based on the EST content of the THC assembly.

### **I.A. PCR Amplification and Clone Preparation**

cDNA clone inserts can be amplified by PCR from plasmid miniprep DNA or directly from clones in culture. In high-throughput applications, amplification of clones from culture has the advantage of being both more cost efficient and less labor intensive with lower cross-contamination rates than amplification from plasmid DNA. Our amplification success rate from culture is equivalent to that we have achieved using plasmid templates. For more than 30,000 clones, our success rate for single-band amplification is approximately 87.5%; 6.3% of reactions yield multiple or weak bands and 6.2% fail to amplify. Results from a typical amplification using the protocol described below are shown in Figure 1.

#### *PCR amplification*

The cDNA clones that are widely available through the IMAGE consortium distributors – The American Type Culture Collection (ATCC), Research Genetics, and Genome Systems – have been cloned into a variety of vectors. While the majority have both M13(-21) and M13REV priming sites. However, many have point mutations in either of these two “universal” priming sites. We have designed alternative M13 primers that avoid these point mutations and that have allowed amplification of clone inserts from all of the vectors we have encountered to date. These new universal amplification primers are:

M13 FWD: 5' GTT TTC CCA GTC ACG ACG TTG 3'  
M13 REV: 5' TGA GCG GAT AAC AAT TTC ACA CAG 3'

Clone inserts are amplified using the following protocol:

1. Selected clones are inoculated into 96 well deep-well blocks (Qiagen; Cat # 19573) containing 1.2 ml of LB/Ampicillin (50 µg/ml) and incubated for 16 hours at 37°C and 200 rpm in a shaking incubator. A 100 µl aliquot of each is archived for future use in microtiter plates containing 10% glycerol at -80°C.
2. Following overnight growth, 5µl of culture suspension are transferred into a 96 well Falcon U-bottom plate (BD Biosciences; Cat # 353077) containing 95µl of MilliQ water.
3. Microtiter plates containing diluted culture are heated to 95°C for 10 minutes in a laboratory oven to lyse the cells and release the plasmid clones.
4. Prior to PCR, cellular debris is removed by centrifugation at 1200×g for 3 minutes in a centrifuge equipped with microtiter plate carriers.

- Clone inserts are amplified in 50µl PCR reactions in 96 well reaction plates (Perkin-Elmer Applied Biosystems Cat # N801-0560). A reaction master mix is prepared for each reaction plate:

MilliQ water	3.2ml	
10× PCR buffer	500µl	
MgCl <sub>2</sub> (25mM)	500µl	
M13 Forward primer (10µM)	100µl	
M13 Reverse primer (10µM)	100µl	
dNTP mix (2mM per dNTP)*	500µl	(Life Technologies Cat# 10297-018)
(* 10µl of each dNTP as 100 mM stock in 460µl of MilliQ water)		
Platinum Taq (5U/µl)	20µl	(Life Technologies; Cat# 10966-018)
Total:		4.92 ml

- For each clone, add 48µl of master mix to 2µl of culture supernatant in 96 well PCR plate.
- Reactions are amplified in an thermocycler (MJ Research; PTC-225 Tetrad) using the following cycling protocol:

95°C × 2min	Initial Denaturation
95°C × 30sec	Denaturation
52°C × 30sec	Annealing
72°C × 2 min	Extension
} × 30 cycles	
4°C forever	

#### *Reaction clean-up*

For efficient binding of the amplified clone inserts to the slides, it is essential to remove unincorporated nucleotides and primers from the reaction products. While there are a variety of techniques that can be used, we have found filtration using 96 well multiscreen filter plates (Millipore; Cat # MANU 03050) to give excellent DNA product recovery without any significant contamination at relatively low cost.

PCR products are cleaned using the following filtration protocol:

- Transfer PCR product (50µl) to the Millipore filter plate.
- Place the filter plate on a vacuum manifold filtration system (Qiagen, Cat # 19504 or Millipore Cat # MAVM0960R) and filter at a pressure of 15in (380 mm) Hg for 10 minutes or until the plate is dry.
- Add 30µl of MilliQ water to each well and filter at 15mm (380 mm) Hg for 5-10 minutes or until the plate is dry.
- Repeat step 3.
- Remove plate from the manifold filtration system. Add 60µl of MilliQ water to each well and place on a shaker. Shake vigorously for 10 minutes to resuspend the DNA.
- Manually pipet the purified product to a new 96 well plate.
- Plates containing the purified PCR products are then sealed using a cap mat (VWR; Cat # 40002-002) and stored at 4°C for future arraying.

## **I.B. Array Printing**

Microarrays are prepared by printing PCR amplicons suspended in either a high salt or other denaturing buffer onto poly-L-lysine or aminosilane coated glass microscope slides using a high-speed robotic system. This process was originally described by Patrick Brown and collaborators (8) at Stanford University and they provide plans so that others can replicate their arraying robot (<http://cmgm.stanford.edu/pbrown/mguide/index.html>). However, there are a number of companies that are selling robotic systems for microarraying and these are listed in Table 2. We use a microarray robot built by Intelligent Automation Systems (IAS) of Cambridge, Massachusetts. Based on a high precision, four-axis Seiko robotic arm, the IAS arrayer uses a 12-tip print head to array DNA samples from either 96- or 384-well microtiter plates onto as many as 100 silanized glass microscope slides. With an average spot size of 130  $\mu\text{m}$  and the capability to adjust the spot-to-spot spacing, the IAS arrayer can spot 19,200 elements (the contents of 200 microtiter plates) or more onto a single slide.

Both the slide surface and the spotting buffer are critical components for reproducible, high-fidelity microarray analysis. Most published reports have used high salt buffers such as 3 $\times$ SSC to print DNA on poly-L-lysine coated slides (10). Our analysis suggests, however, that aminosilane offers a more consistent surface with lower background fluorescence. There are a number of commercial vendors for aminosilane coated slides (see Table 2), but Telechem SuperAmine<sup>TM</sup> aminosilane coated glass microscope slides have been the most consistent. Using our protocol in number of side-by-side comparisons, the SuperAmine<sup>TM</sup> slides produce approximately half the background fluorescence of poly-L-lysine slides while yielding signal intensities that are consistently higher (data not shown). In addition, the spot morphology on SuperAmine<sup>TM</sup> slides is much more uniform, with fewer “doughnuts” than on any of the alternatives we have investigated.

We also investigated a number of different spotting chemistries to determine which provides the best results in subsequent hybridization assays. Figure 2 shows the results of a comparison between 50% dimethyl-sulfoxide (DMSO) and 3 $\times$ SSC as a spotting buffer. We have found that PCR products printed onto SuperAmine<sup>TM</sup> aminosilane coated glass microscope slides (Telechem, Cat# SMM-100) using 50% DMSO as a printing buffer provides the best substrate for hybridization, giving the greatest hybridization intensities. Using 50% DMSO as a printing solution has a number of additional advantages. DMSO denatures the DNA allowing better binding to the slide and providing more single-stranded targets for hybridization. Further, DMSO is hygroscopic and has a low vapor pressure, allowing DNA prepared for arraying to be stored for long periods of time without significant evaporation.

The print head on our arrayer and most others use “quill” pens that use capillary action to draw fluid into the spotting pens and surface tension interactions to dispense solution onto the slide. The Arrayit ChipMaker3<sup>TM</sup> microspotting pins (TeleChem International Inc.) are very durable and can reproducibly generate high-quality spots with good precision; all array images shown were printed with the same set of ChipMaker<sup>TM</sup> 3 over more than six months. A variety of parameters such as the robot arm acceleration, temperature, and humidity control both spot morphology and size. We have found printing to be optimal at approximately 45% relative humidity and a constant temperature of 72°F (22°C). Changes in humidity and temperature have a significant impact on the size and morphology of spots, as well as the efficiency of DNA binding to the slides and these must be carefully controlled to provide the

consistent spotting. Figure 3 shows the effects of varying humidity and temperature on spot morphology and DNA retention. DNA samples were spotted onto the slides as described above while temperature and humidity levels were recorded on a chart recorder. During the printing, temperature and humidity levels were allowed to vary continuously from 72°F (22.2°C) and 45-50% to a low of 62°F (16.7°C) and 40-45% and a high of 80°F (26.7°C) and 80-85% respectively. Following hybridization with a vector specific probe, we were able to reconstruct the optimal printing conditions by using the chart recorder data to assign temperature and humidity values to the spots.

### *Arraying*

1. Add equal volumes of purified PCR product to DMSO in a 96 well V-bottom plate (Corning; Cat # 3897). Typically, 5µl of each are used to prepare spotting plates that can be used to print 100 or more slides.
2. Slides to be printed are marked with a diamond-tipped pen, dust is removed by blowing the slides with high-pressure nitrogen gas, and the slides are placed in the arrayer. Care must be taken not to touch the surface of the slides as oils adversely affect the ability of the slide surface to bind DNA.
3. Microtiter spotting plates are loaded into the arrayer and PCR products are spotted onto the slides at 72°F and 45% relative humidity.
4. Following printing, the slides are allowed to dry and spotted DNA is bound to slide by UV-crosslinking at 90-350 mJ using a Stratalinker™ (Stratagene, Cat# 400071).
5. Printed slides are stored in a light-tight box in a bench-top dessicator at room temperature until they are to be used for hybridization.

## **II. Probe Preparation and Hybridization**

Microarrays assay differential gene expression by co-hybridization of fluorescently labeled probes prepared from different RNA sources. As with many other RNA-based assays, the purity and quality of the starting RNA has a significant effect on the results of the assay. Further, the products of the labeling reactions must be cleaned to remove unincorporated labeled nucleotides that can produce a significant background on the slides following hybridization. Finally, hybridization conditions and wash must be optimized to provide high specificity to minimize cross-hybridization. We have developed probe preparation and hybridization protocols using RNA derived from human carcinoma cell lines as a model system; variations of this protocol have been applied to the study of expression in rat and other systems.

### **II.A. RNA Extraction**

Impurities in RNA preparations can have an adverse effect both on labeling efficiency and the stability of the fluorescent labels that are used for microarray expression analysis. We have found that Trizol™ (Life Technologies; Cat# 15596-014) gives consistently high quality RNA from cell culture and many tissue samples, although additional steps must be taken to remove polysaccharides when extracting RNA from some tissues. Trizol extraction is quick and produces a high yield of total RNA.

#### *RNA Extraction*

1. Aspirate media from the cells and wash once with Phosphate Buffered Saline (PBS).
2. Add 5ml PBS and scrape cells from the plate.

3. Transfer cell suspension to a 50ml polypropylene conical-bottom tube (Falcon Cat# 352070).
4. Wash the plate with an additional 1ml PBS and add the suspension to the tube.
5. Pellet the cells by centrifugation at 2300 rpm (900×g) for 3 minutes at 4°C and discard the supernatant.
6. Add 2ml Trizol™ per  $\sim 2 \times 10^6$  cells (approximately one 150mm plate of fibroblasts) to the pellet and pass the suspension through an 18 gauge syringe several times to disrupt the pellet.
7. Incubate the sample at room temperature for 5 minutes.
8. Add 0.4ml of chloroform (0.2ml/1ml Trizol) and shake vigorously for 1 minute
9. Incubate at room temperature for 2 minutes 30 seconds.
10. Remove cellular debris by centrifugation at 4000rpm (2700×g) for 15min at 4°C.
11. Transfer the supernatant to 1.2 ml microfuge tubes (0.5ml/tube) and an equal volume of isopropanol to precipitate the RNA.
12. Incubate at room temperature for 15 minutes.
13. Centrifuge at 15,000rpm (21,000×g) for 15 minutes to pellet the RNA.
14. Discard the supernatant and resuspend the pellet in 70% ethanol. The RNA can be stored in 70% ethanol at -20°C until use.
15. Prior to use, centrifuge at 15,000 rpm (21,000×g) for 15 minutes at 4°C and discard supernatant.
16. Resuspend the pellet in diethylpyrocarbonate (DEPC) treated water or RNase-free TE buffer for labeling.

## **II.B. RNA Labeling**

The ability to label small quantities of starting material is an important consideration for the study of expression in rare patient samples and consequently, we have focused on decreasing the quantity of starting material required. Probes for microarray analysis are prepared from RNA templates by incorporation of fluorescently labeled deoxyribonucleotides during first strand cDNA synthesis. Either total or poly(A<sup>+</sup>) RNA can be used in the reverse transcription reaction. Oligo(dT) labeling of total RNA provides consistently high-quality probes from smaller quantities of starting RNA and without the expense of poly(A<sup>+</sup>) purification. Figure 4 shows the results of microarray hybridizations using labeled total or poly(A<sup>+</sup>) RNA prepared from the same cell lines. An analysis of the fluorescence intensities for the elements in arrays hybridized with probes prepared from 1.5μg of poly(A<sup>+</sup>) RNA (the equivalent 50-100μg of starting total RNA) and 4μg of total RNA indicate that total RNA labeling provides comparable probe activity without any increase in background fluorescence.

In the past, we typically prepared direct-labeled probes using cyanine3- and cyanine5-dUTP (NEN Cat#s NEL578, NEL579), although cyanine-labeled dCTP (NEN Cat#s NEL576, NEL577) can be used with an appropriate change in the concentrations of unlabeled dNTPs in the reaction. While this method has been used for generating the bulk of the microarray data published to date, it is known to suffer from differential incorporation of Cy-3 and Cy-5 labeled nucleotides in the first-strand cDNA synthesis and from the generation of some dye-specific hybridization artifacts. An alternative, and one that we have recently adopted, is first strand synthesis of cDNA containing amino-allyl-labeled nucleotides for both samples, followed by a covalent coupling to the NHS-ester of the appropriate Cyanine fluor. This method has proven highly efficient for both eukaryotic and prokaryotic expression analysis and is what we now generally use. It has the disadvantage of being more labor-intensive than direct incorporation, but it is also less expensive.

Regardless of the incorporation method used, the performance of the reverse transcriptase is crucial for successful probe synthesis. We have investigated a number of reverse transcriptases, including AMV and MMLV and have found that Superscript™ II RT (LifeTechnologies; Cat# 18064-014) generates probes with significantly greater activity (data not shown).

It should be noted that both Cy3 and Cy5 are photosensitive and care should be taken to minimize exposure to light during the labeling, hybridization, washing, and scanning processes. Upon receipt, Cy-labeled nucleotides, or the NHS-esters, should be aliquotted into single-use light- proof tubes and stored at –20°C until needed. All reactions should be carried out in foil-wrapped tubes and all hybridizations and washes in foil-wrapped containers.

#### *Direct Incorporation Probe Labeling and Purification*

1. Prepare a labeling reaction master mix containing 500μM dCTP, 500μM dATP, 500μM dGTP, 100μM dTTP, Cy 3-dUTP/Cy 5-dUTP, 400U Superscript™ II RT, 1mM dithiothreitol (DTT) and 1× RT buffer. We typically prepare sufficient quantities for 20 labeling reactions and store the unused solution at -20°C:

5× RT buffer	120μl
DTT (5mM)	60μl
dATP (100mM)	3μl
dCTP (100mM)	3μl
dGTP (100mM)	3μl
dTTP (100mM)	0.6μl
DEPC'd	110.4μl
WATER	
Total:	300μl

2. To 10μg of total RNA (or 2μg poly(A<sup>+</sup>)) in a microfuge tube, add 2μg of oligo(dT) (18-20mer; Life Technologies Cat# Y012120) and DEPC-treated WATER to a total volume of 10μl.
3. Incubate the reaction mixture at 70°C for 10 minutes, and chill on ice for one minute.
4. To the RNA, add:

RT labeling mix	15μl
Cy3- or Cy5-dUTP (1 mM)	3μl
Superscript™ II RT (200U/μl)	2μl

5. Mix thoroughly and incubate at 42°C for 2 hours.
6. Briefly centrifuge the reaction and add 1.5μl of 20mM EDTA to stop the reaction.
7. Add 1.5μl of 500mM NaOH and heat at 70°C for 10 minutes to degrade the RNA
8. Neutralize the reaction by adding 1.5μl of 500mM HCl
9. Unincorporated fluorescent nucleotides are removed by glass fiber filtration using GFX columns (Pharmacia Cat# 27-9602-01) and the instructions provided by the manufacturer.
10. Elute the purified products using 50μl of TE pH 8.0 and dry the probe to completion in a speedvac.



### *Aminoallyl/Secondary Coupling Probe Labeling and Purification*

#### **Aminoallyl labeling of RNA**

This protocol is based on that developed by Chris Seidel for the initial protocol (<<http://www.pangloss.com/seidel/Protocols/amino-allylRT.html>>) and modified by Benjamin Schroeder, Erik Snesrud and Ivana Yang of TIGR

#### *A. Labeling*

1. Prepare a 100mM aminoallyl dUTP (Sigma Cat # A0410) by dissolving 1 mg in 19.1  $\mu$ l 0.1 M KPO<sub>4</sub> buffer, pH 7.5. The concentration of this stock solution can be measured by diluting an aliquot 1:5000 in the same buffer and measuring OD<sub>289</sub>. Stock concentration in mM = OD<sub>289</sub>×704.
2. Prepare 50× labeling mix containing a 2:3 ratio of aminoallyl-dUTP to unlabeled dTTP containing with a final concentration of 25 mM dATP, 25 mM dCTP, 25 mM dGTP, 15 mM dTTP, 10 mM aa-dUTP:

dATP (100mM)	5 $\mu$ l
dCTP (100mM)	5 $\mu$ l
dGTP (100mM)	5 $\mu$ l
dTTP (100mM)	3 $\mu$ l
aa-dUTP (100mM)	2 $\mu$ l
<hr/>	
Total:	20 $\mu$ l

Unused solution can be stored at -20°C.

3. To 10  $\mu$ g of total RNA (or 2  $\mu$ g poly(A<sup>+</sup>) eukaryotic RNA or 2.0  $\mu$ g Bacterial RNA, which has been DNaseI-treated, Qiagen RNeasy purified), add 2  $\mu$ l Random Hexamer primers (3mg/ml; LifeTech #48190-011) and RNase-free water to 18.5  $\mu$ l.
4. Mix well and incubate at 70°C for 10 minutes.
5. Snap-freeze in dry ice/ethanol bath for 30 seconds, microfuge for 1 minute and continue at room temperature.
6. To this, add

5× Superscript II buffer	6 $\mu$ l
0.1 M DTT	3 $\mu$ l
50× aminoallyl-dNTP mix	0.6 $\mu$ l
SuperScript II RT (200U/ $\mu$ l)	2 $\mu$ l

7. Mix and incubate at 42°C for 3 hours to overnight.
8. To hydrolyze the RNA, add:

1 M NaOH	10 $\mu$ l
0.5 M EDTA	10 $\mu$ l

Incubate at 65°C for 15 minutes.

9. Add 10  $\mu$ l 1 M HCl to neutralize. (Alternatively, one can add 25  $\mu$ l 1 M HEPES pH 7.0 or 25  $\mu$ l 1 M Tris pH 7.4.)

### *B. Cleanup*

The aa-cDNA must be purified away from unincorporated aa-dUTP and free amines (such as HEPES and Tris) prior to conjugation with NHS-ester Cy dyes to prevent coupling to free amine groups in solution. This can be done with Microcon 30 filtration or with a modified Qiagen PCR spin column protocol. We use the Qiagen protocol, although we provide the Microcon alternative for completion.

#### Modified Qiagen QIAquick PCR purification

1. Prepare a Tris-free Phosphate wash buffer consisting of 5 mM KPO<sub>4</sub>, pH 8.0 and 80% EtOH. *The use of this buffer is necessary because the Qiagen-supplied PE elution buffer contains Tris.* A protocol for the preparation of this Phosphate buffer, as well as the phosphate elution buffer, is provided at the end of this section.
2. Mix cDNA reaction with 300 µl (5× reaction volume) buffer PB, transfer to QIAquick column.
3. Place the column in a collection tube and centrifuge for 1 minute at 14,000 rpm in a microfuge.
4. Wash with 750 µl Phosphate wash buffer.
5. Spin at 14,000 rpm for 1 minute in a microfuge.
6. Repeat the wash and centrifugation steps in 4 and 5 above, emptying the collection tube before centrifugation.
7. Transfer column to fresh tube, add 30 µl 4 mM KPO<sub>4</sub>, pH 8.5.
8. Incubate for 1 minute.
9. Elute by centrifuging 14,000 rpm for 1 minute in a microfuge.
10. Add an additional 30 µl 4 mM KPO<sub>4</sub>, pH 8.5, incubate for 1 minute, and centrifuge to elute with to a total volume of 60 µl.
11. Dry the sample in a Speed vac.

#### Microcon 30 cleanup

1. Add 375 µl water to cDNA reaction and place in Microcon 30.
2. Centrifuge at 12,000 rpm for 12 minutes in a microfuge.
3. Wash twice by adding 450 µl water and centrifuging at 12,000 rpm for 12 minutes in a microfuge tube, emptying the collection tube each time.
4. Invert Microcon column into new collection microfuge tube and centrifuge at 12,000 rpm for 1 minute in a microfuge tube.
5. Dry the sample in a Speed vac.

### *C. Coupling the labeled cDNA to the Cy Dye Ester*

1. Prepare the Cy-dye esters for labeling. The Cy3-ester (Cat# PA23001) and Cy5-ester (Cat# PA25001) are each provided by AmershamPharmacia as dried samples in 5 tubes. As each tube of dye about to be used, resuspend the dye ester in 73 µl DMSO. The dye esters can spontaneously hydrolyze, so this must either be used immediately or stored at -80°C as 4.5 µl single-use aliquots.
2. Prepare a 1 M Na<sub>2</sub>CO<sub>3</sub>, pH 9.0 carbonate buffer for the coupling reaction. Dissolve 10.8 g Na<sub>2</sub>CO<sub>3</sub> in 80 ml water and bring the pH to 9.0 with 12 N HCl; bring the final volume to 100 ml with water. Dilute 1:10 with water for to make the 0.1M solution used in the coupling reaction. Carbonate buffer changes composition over time so make sure you make it fresh every couple of weeks to a month.
3. Resuspend the aminoallyl-labeled cDNA in 4.5 µl 0.1 M carbonate buffer, pH 9.0.

4. Add 4.5 µl of the appropriate NHS-Cy suspended in DMSO.
5. Incubate the reaction for 1 hour in the dark at room temperature.

#### *D. Reaction Cleanup II – Removal of uncoupled dye*

Following the covalent coupling of the dye-ester to the aminoallyl-labeled cDNA, uncoupled dye must be removed prior to hybridization. We have found the QIAquick PCR purification kit to provide a simple and effective purification.

1. To the reaction, add 35 µl 100 mM NaOAc pH 5.2.
2. Add 250 µl PB buffer.
3. Follow the PCR protocol provided by Qiagen, eluting twice in 30 µl EB for a total of 60 µl.
4. Following analysis of incorporation below, dry the products to completion in a speedvac.

#### *E. Analysis of Labeling Reaction*

It is extremely useful to measure the efficiency of these reactions prior to microarray hybridization; if the reactions did not work well, then a hybridization with these labeled probes is not likely to provide useful data. We use a 50 µl Beckman quartz MicroCuvette to analyze the entire undiluted sample in a spectrophotometer.

1. For each sample, measure absorbance at 260 nm and either 550 nm for Cy3 or 650 nm for Cy5, as appropriate.
2. Calculate the total cDNA synthesis for each sample using:  

$$\# \text{pmol nucleotides} = [\text{OD}_{260} * \text{volume (in } \mu\text{l}) * 37 * 1000] / (324.5)$$

$$(1\text{OD}_{260} = 37 \text{ ng}/\mu\text{l for cDNA; } 324.5 \text{ pg/pmol average molecular weight of a dNTP; } 1000 \text{ to convert ng to pg})$$
3. Calculate total dye incorporation in pmol as appropriate, using  

$$\# \text{pmol Cy3} = \text{OD}_{550} * \text{volume (in } \mu\text{l}) / 0.15$$

$$(\text{The extinction coefficient for Cy3 is } 150,000 \text{ M}^{-1}\text{cm}^{-1}.)$$

$$\# \text{pmol Cy5} = \text{OD}_{650} * \text{volume (in } \mu\text{l}) / 0.25$$

$$(\text{The extinction coefficient for Cy5 is } 250,000 \text{ M}^{-1}\text{cm}^{-1}.)$$

We have found that approximately 150 pmol of dye incorporation per sample and a ratio of less than 50 nucleotides/dye molecules is near optimal for hybridizations.

#### *Preparation of Phosphate buffers*

1. Prepare 2 solutions: 1 M KH<sub>2</sub>PO<sub>4</sub> and 1M K<sub>2</sub>HPO<sub>4</sub>.
2. Combine

1M K <sub>2</sub> HPO <sub>4</sub>	9.5 ml
1 M KH <sub>2</sub> PO <sub>4</sub>	0.5 ml

to make 1 M KPO<sub>4</sub> (Phosphate buffer) The pH of this solution should be 8.5-8.7.

3. For 100 ml Phosphate wash buffer, mix:
 

1M KPO <sub>4</sub> pH 8.5	0.5 ml
MilliQ water	15.25 ml
95% ethanol	84.25 ml

Note: this solution will be slightly cloudy.

4. Phosphate elution buffer is made by dilution of the 1 M KPO<sub>4</sub> pH 8.5 to 4 mM.

## **II.C. Hybridization**

The goal in any hybridization is to obtain high specificity while minimizing background. We have developed protocols that give reproducible, high-quality hybridization results while maximizing the measured fluorescence from the array.

Aminosilane coated slides bind DNA with high efficiency. Prior to hybridization, the free amine groups on the slide must be blocked or inactivated, otherwise nonspecific binding of labeled cDNA to the slide can deplete the probe and produce high background. Although the slides can be blocked chemically, we have found a simple prehybridization in a solution containing 1% bovine serum albumin to be extremely effective in eliminating nonspecific binding of the probe to the slide.

Prehybridization has the additional advantage of washing unbound DNA from the slide prior to the addition of the probe. Any DNA that washes from the surface during hybridization competes with DNA bound to the slide. As the kinetics of solution hybridization is much more favorable than surface hybridization, this can dramatically decrease the measured fluorescence signal from the microarray. All prehybridization and hybridization washes are carried out in microscope slide staining trays (VWR Cat# 25461-003).

### *Prehybridization*

1. Prepare prehybridization buffer containing 5×SSC, 0.1% SDS and 1% bovine serum albumin (BSA; Sigma Cat# A-9418).
2. Prepare 1× hybridization buffer containing 50% formamide, 5×SSC, and 0.1% SDS.
3. Place slides to be analyzed into a Coplin jar (VWR Cat# 25457-200), fill with prehybridization buffer, and incubate at 42°C for 45 minutes.
4. Wash the slides by dipping five times in room temperature MilliQ water.
5. Dip the slides in room temperature isopropanol and air dry.

Slides should be used immediately following prehybridization. We have found that hybridization efficiency decreases rapidly if the slides are allowed to dry for more than one hour.

### *Hybridization*

1. Resuspend each labeled probe in 12μl of 1× hyb buffer (50% formamide, 5×SSC, and 0.1% SDS).
2. Combine 12μl of each of purified Cy3- and Cy5-labeled probes, mix well and add  
COT1-DNA (20μg/μl)      1μl      (LifeTechnologies; Cat# 25279-011).  
Poly(A)-DNA (20μg/μl)      1μl      (Pharmacia; Cat# 27-7836-01).  
to block nonspecific hybridization.
3. Heat the probe mixture at 95°C for 3 minutes to denature.
4. Centrifuge the probe in a microfuge set at maximum angular velocity for 1 minute.
5. Apply the labeled probe to a prehybridized microarray slide and cover with a 22mm×60mm glass coverslip (Fisher Cat# 12-545J).

6. Place the slide in a sealed hybridization chamber (Corning Costar Cat #2551), add 20 $\mu$ l of water to the chamber at the end of the slide.
7. Place the sealed chamber in a 42°C water bath and incubate for 16-20 hours.
8. Remove the array from the hybridization chamber, taking care not to disturb the coverslip.
9. Place the slide in a staining dish containing low-stringency wash buffer containing 1 $\times$ SSC and 0.2% SDS at 42°C.
10. Gently remove the coverslip while the slide is in solution and agitate for 4 minutes.
11. Wash the slide at high-stringency in a staining dish containing 0.1 $\times$ SSC and 0.2% SDS at room temperature, agitating for 4 minutes.
12. Wash the slide in 0.1 $\times$ SSC, agitating for 4 minutes.
13. Allow the slides to air dry.

### **III. Data Collection, Normalization, and Analysis**

Differential gene expression is assessed by scanning the hybridized arrays using a confocal laser scanner capable of interrogating both the Cy3- and Cy5-labeled probes and producing separate TIFF images for each. As is the case with arraying robots, there are a number of manufacturers that produce scanners capable of detecting Cy3 and Cy5 (see Table 4) and most are planning to release instruments capable of detecting additional dyes.

#### *Slide Scanning*

We are currently using the ScanArray 3000 produced by GSI Lumonics. This scanner uses red and green Helium-Neon lasers operating at 633nm and 543nm to excite Cy5 and Cy3, respectively. Hybridized slides are scanned first in the Cy5 channel, and then the Cy3 channel, as Cy5 is more susceptible to photodegradation than Cy3. Data from each fluorescence channel is collected and stored as a separate 16-bit TIFF image. These images are analyzed to calculate the relative expression levels of each gene and to identify differentially expressed genes. The analysis process can be divided into two steps – image processing and data analysis. Figure 5 shows a typical hybridization image produced when things work well. The contrast in this image has been adjusted to allow faint spots to be easily visualized. Important aspects of the hybridization to note are the low level, uniform background and the good signal-to-noise

#### *Image Processing*

Image processing involves three stages. First, the spots representing the arrayed genes must be identified and distinguished from spurious signals that can arise due to precipitated probe or other hybridization artifacts or contaminants such as dust on the surface of the slide. This task is simplified to a certain extent because the robotic arraying systems used to construct the arrays produce a regular arrangement of the spotted DNA fragments. However, variable intensities and uneven slide backgrounds as well as some irregularities in the gridded arrays complicate the problem slightly. Generally, problem of grid spot location is coupled with estimation of the fluorescence background. For microarrays, it is important the background be calculated locally for each spot, rather than globally for the entire image as uneven background can often arise during the hybridization process. The second step in analysis of the array images is the estimation of background.

Following spot identification and local background determination, the background-subtracted hybridization intensities for each spot must be calculated. There are currently two schools of thought regarding the calculation of intensities – the use of the median or the mean intensity for each spot. As array analysis generally uses ratios of measured Cy3 to Cy5 intensities to identify differentially expressed genes, the mean and the integrated intensities are operationally equivalent. In comparisons of intensities measured for normalization controls spiked into the labeling reactions, we have found mean intensities to give more consistent results and consequently we use these in subsequent calculations (V. Sharov and J. Quackenbush, in preparation).

A number of image processing software packages are available and are listed in Table 5. We have developed a software package called TIGR\_Spotfinder for image processing (<<http://www.tigr.org/softlab/>>). TIGR\_Spotfinder uses a thresholding algorithm that separates spots from the background, allowing a grid to be laid across the spots. Having found a grid, spots are found within each grid element, local background is calculated, and background-subtracted, integrated intensities are calculated in both the Cy3 and Cy5 channels. Measured intensities are entered into the Molecular Analysis of Gene Expression (MAGE) database, a Sybase relational database specifically designed to capture gene expression data.

#### *Data Normalization and Analysis*

Following image processing, the data generated for the arrayed genes must be further analyzed before differentially expressed genes can be identified. The first step in this process is the normalization of the relative fluorescence intensities in each of the two scanned channels. Normalization is necessary to adjust for differences in labeling and detection efficiencies for the fluorescent labels and for differences in the quantity of starting RNA from the two samples examined in the assay. These problems can cause a shift in the average ratio of Cy5 to Cy3 and the intensities must be rescaled before an experiment can be properly analyzed.

The normalization strategies that can be used are based on some underlying assumptions regarding the data and the strategies used for each experiment should be adjusted to reflect both the system under study and the experimental design. The primary assumption is that for either the entire collection of arrayed genes or some subset of the genes such as housekeeping genes, or for some added set of controls, the ratio of measured expression averaged over the set should be one.

Depending on the experimental design, there are three useful approaches for calculating normalization factors. The first simply uses total measured fluorescence intensity. The assumption underlying this approach is that the total mass of RNA labeled with either Cy3 or Cy5 is equal. While the intensity for any one spot may be higher in one channel than the other, when averaged over thousands of spots in the array, these fluctuations should average out. Consequently, the total integrated intensity across all the spots in the array should be equal for both channels. Alternatively, one could add a number of controls in increasing but equimolar concentrations to both the labeling reactions and the sum of the intensities for these spots should be equal. A second approach uses linear regression analysis. For closely related samples, one would expect many of the genes to be expressed at nearly constant levels. Consequently, a scatterplot of the measured Cy5 versus Cy3 intensities should have a slope of one. Measured intensities for added equimolar controls should behave similarly. Under this assumption, one can use regression

analysis techniques to calculate the slope. This is then used to rescale the data and adjust the slope to one. A third approach has been described by Chen *et al.* (2). They assume that some subset of housekeeping genes exists and that for these, the distribution of transcription levels should have some mean value  $\mu$  and standard deviation  $\sigma$  independent of the sample. In this case, the ratio of measured Cy5 to Cy3 ratios for these genes can be modeled and the mean of the ratio adjusted to 1. Chen and collaborators describe an iterative procedure to achieve this normalization and we have implemented their algorithm and a variation of it that uses the entire data set, as well the total intensity and linear regression normalization, into a data visualization and analysis tool called TIGR ArrayViewer. TIGR ArrayViewer is freely available and can be obtained through <http://www.tigr.org/softlab/>. In any normalization approach, care must be taken in handling genes expressed at low levels. Statistical fluctuations in the measured levels can cause a significant variation in the ratios that are calculated and inefficiencies in labeling for either of the two dyes can cause these low intensity genes to disappear from the arrays. Typically, we only use spots in the final analysis where the intensities in both channels are two standard deviations above background.

Following normalization, data are typically analyzed to identify genes that are differentially expressed. Most published studies have used a post-normalization cutoff of two-fold up- or down-regulation to define differential expression; the approach defined by Chen *et al.* (2) provides confidence intervals that can be used to identify differentially expressed genes. In order to separate genes that are truly differentially expressed from stochastic changes, we typically conduct three independent microarray assays starting from independent mRNA isolations and define differential expression based on their consensus.

## Conclusion

The examination of gene expression using microarrays holds tremendous promise for the identification of candidate genes involved in a variety of processes. Indeed, the experiments that have been described to date have confirmed known patterns of expression and provided information on genes of unknown function. However, most applications have to date only allowed the identification of genes differentially expressed at significant levels. The true challenge, and the promise of this technique, will be to use it to identify genes that are consistently up- or down-regulated by 10 or 20% yet play significant roles in the development and progression of disease. This will require the analysis of data from multiple experiments and the correlation of patterns of gene expression with additional experimental and clinical information. Recently a variety of techniques including hierarchical clustering (3) and self-organizing maps (11) have been applied to the analysis of microarray expression data across multiple experiments. However, each of these depends on having reliable and reproducible data from each microarray assay. The laboratory techniques outlined here have allowed reproducible hybridization results such as those shown in Figure 5. Although these protocols will likely continue to evolve, we believe that they represent a reliable starting point for those beginning microarray experimentation.

## Acknowledgements

This work was supported with funding from the National Cancer Institute's Cancer Genome Anatomy Project (R01 CA77049-01; PI: J. Quackenbush). The authors wish to thank V. Sharov, A. Saeed, R.T. Cline, and S. Peterson for valuable comments and contributions.

## References

1. Boguski, M.S., and Schuler, G.D. 1995. ESTablishing a human transcript map. *Nature Genet.* **10**:369-371.
2. Chen, Y., Dougherty, E.R., and Bittner, M.L. 1997. Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Optics* **24**:364-374
3. Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**:14863-14868.
4. Eisen, M.B. and Brown, P.O. 1999. DNA arrays for analysis of gene expression. *Methods Enzymol.* **303**:179-205.
5. Liang, P., and Pardee, A.B. 1992. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* **257**:967-971.
6. Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., Brown, E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nature Biotechnol.* **14**:1675-1680.
7. Quackenbush, J., Liang, F., Holt, I. Perte, G., and Upton, J. 2000. The TIGR Gene Indices: Reconstruction and Representation of Expressed Gene Sequences. *Nucleic Acids Res.* **28**:141-145.
8. Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. 1995. Quantitative monitoring of gene expression patterns with complementary DNA microarray. *Science* **270**:467-470.
9. Schuler, G.D. 1997. Pieces of the puzzle: expressed sequence tags and the catalog of human genes. *J. Mol. Med.* **75**:694-698.
10. Shalon, D., Smith, S.J. and Brown, P.O. 1996. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* **6**:639-645.
11. Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E. Lander, E.S., and Golub, T.R. 1999. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA* **96**:2907-2912.
12. Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. 1995. Serial analysis of gene expression. *Science* **270**:484-487.



National Human Genome Research Institute	<a href="http://www.nhgri.nih.gov/DIR/LCG/15K/HTML/protocol.html">&lt;http://www.nhgri.nih.gov/DIR/LCG/15K/HTML/protocol.html&gt;</a>
Stanford University	<a href="http://cmgm.stanford.edu/pbrown/protocols/index.html">&lt;http://cmgm.stanford.edu/pbrown/protocols/index.html&gt;</a>
Telechem	<a href="http://arrayit.com/DNA-Microarray-Protocols/">&lt;http://arrayit.com/DNA-Microarray-Protocols/&gt;</a>
University of Pennsylvania	<a href="http://genomics.med.upenn.edu/vcheung/protocols.htm">&lt;http://genomics.med.upenn.edu/vcheung/protocols.htm&gt;</a>
The Institute for Genomic Research	<a href="http://www.tigr.org/tdb/microarray">&lt;http://www.tigr.org/tdb/microarray&gt;</a>

**Table 1. Protocols for microarray analysis.**

Beecher Instruments	< <a href="http://www.beecherinstruments.com">http://www.beecherinstruments.com</a> >
BioRobotics	< <a href="http://www.BioRobotics.com/">http://www.BioRobotics.com/</a> >
Cartesian Technologies	< <a href="http://www.cartesiantech.com/">http://www.cartesiantech.com/</a> >
Engineering Services	< <a href="http://www.ESIT.com/">http://www.ESIT.com/</a> >
Genetic Microsystems	< <a href="http://www.geneticmicro.com">http://www.geneticmicro.com</a> >
Genetix	< <a href="http://www.genetix.co.uk/">http://www.genetix.co.uk/</a> >
Gene Machines	< <a href="http://www.genemachines.com">http://www.genemachines.com</a> >
Genomic Solutions	< <a href="http://www.genomicsolutions.com/">http://www.genomicsolutions.com/</a> >
Intelligent Automation Systems	< <a href="http://www.ias.com">http://www.ias.com</a> >
Packard	< <a href="http://www.packardinst.com/">http://www.packardinst.com/</a> >

**Table 2. Manufacturers of microarray spotting robots.**

Amersham Pharmacia Biotech	< <a href="http://www.apbiotech.com">http://www.apbiotech.com</a> >
Corning Costar	< <a href="http://www.cmt.corning.com">http://www.cmt.corning.com</a> >
Telechem	< <a href="http://www.arrayit.com/">http://www.arrayit.com/</a> >

**Table 3. Aminosilane coated slide suppliers.**

Axon	< <a href="http://www.axon.com">http://www.axon.com</a> >
Beecher Instruments	< <a href="http://www.beecherinstruments.com">http://www.beecherinstruments.com</a> >
GSI Lumonics	< <a href="http://www.gsilumonics.com">http://www.gsilumonics.com</a> >
Genetic Microsystems	< <a href="http://www.geneticmicro.com">http://www.geneticmicro.com</a> >
Genomic Solutions	< <a href="http://www.genomicsolutions.com/">http://www.genomicsolutions.com/</a> >
Molecular Dynamics	< <a href="http://www.mdyn.com">http://www.mdyn.com</a> >
Virtek	< <a href="http://www.virtek.ca/">http://www.virtek.ca/</a> >

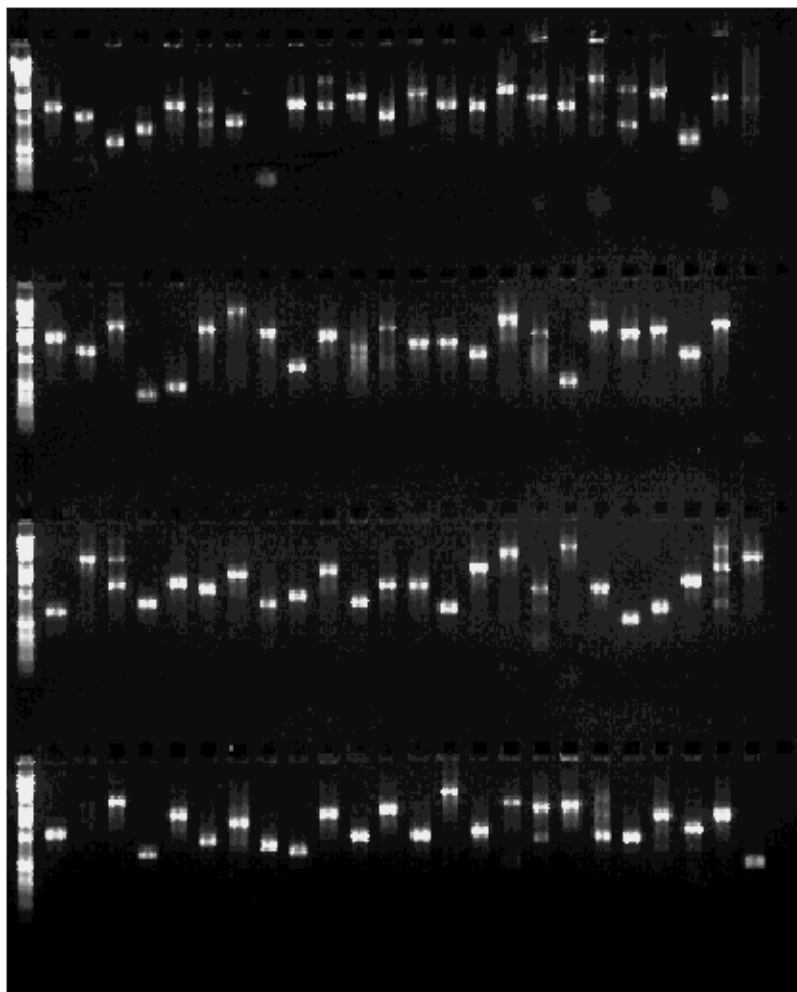
**Table 4. Manufacturers of microarray scanners.**

BioDiscovery	< <a href="http://www.biodiscovery.com/">http://www.biodiscovery.com/</a> >
Imaging Research	< <a href="http://imaging.brocku.ca/Arrayvision.html">http://imaging.brocku.ca/Arrayvision.html</a> >
National Human Genome Research Institute	< <a href="http://www.nhgri.nih.gov/DIR/LCG/15K/HTML/img_analysis.html">http://www.nhgri.nih.gov/DIR/LCG/15K/HTML/img_analysis.html</a> >
Stanford University	< <a href="http://rana.Stanford.EDU/software/">http://rana.Stanford.EDU/software/</a> >
The Institute for Genomic Research	< <a href="http://www.tigr.org/softlab/">http://www.tigr.org/softlab/</a> >

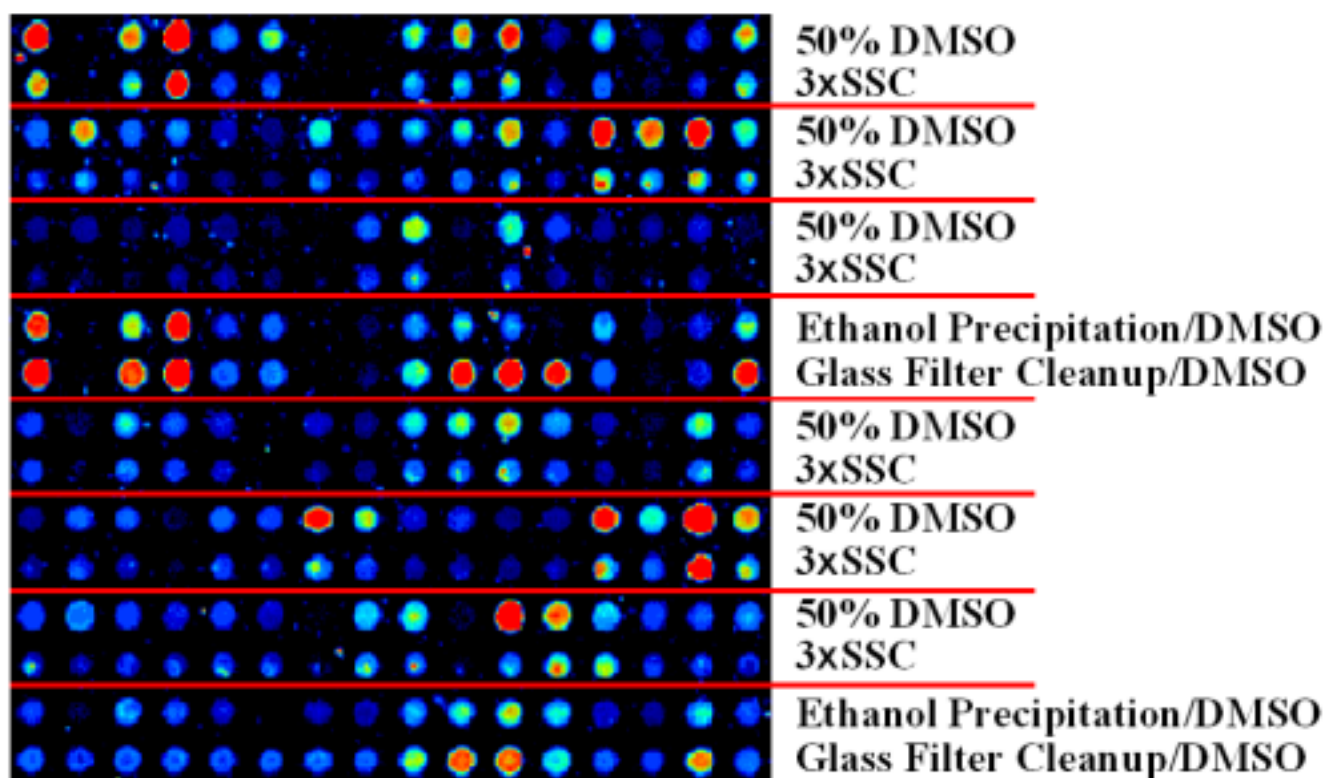
**Table 5. Microarray image processing software sources.**

BioDiscovery	< <a href="http://www.biodiscovery.com/">http://www.biodiscovery.com/</a> >
Silicon Genetics	< <a href="http://www.sigenetics.com/">http://www.sigenetics.com/</a> >
Spotfire	< <a href="http://www.spotfire.com/">http://www.spotfire.com/</a> >
Stanford University	< <a href="http://rana.Stanford.EDU/software/">http://rana.Stanford.EDU/software/</a> >
TIGR	< <a href="http://www.tigr.org/softlab/">http://www.tigr.org/softlab/</a> >

**Table 6. Microarray data analysis sources.**

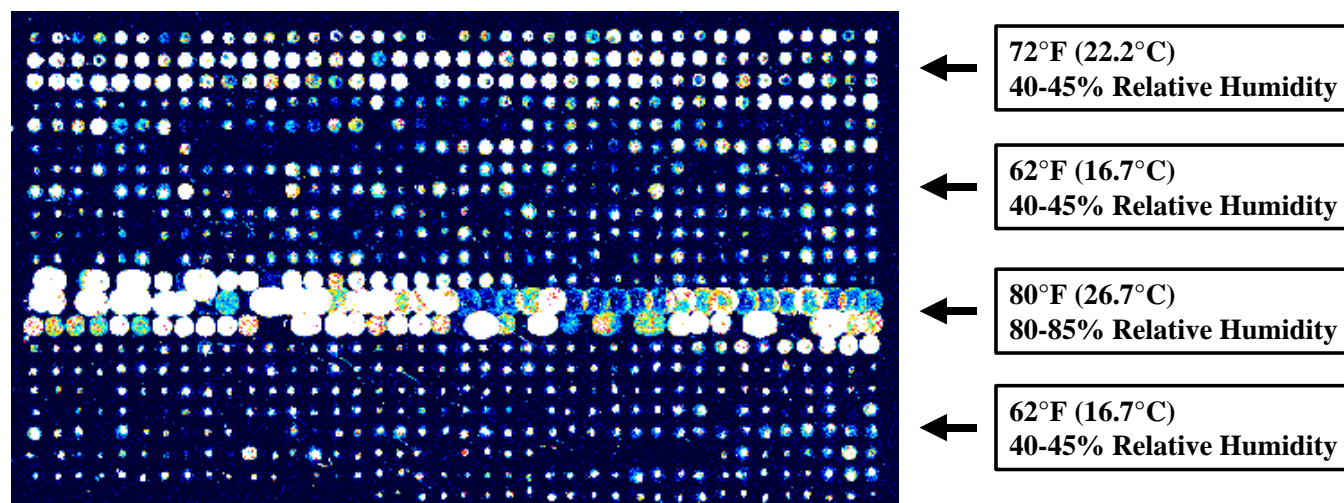


**Figure 1. Results of PCR amplification of 96 cDNA clones directly from bacterial culture.** The leftmost lane in each tier of the gel contains a 1 kb ladder; each tier contains 24 distinct samples.



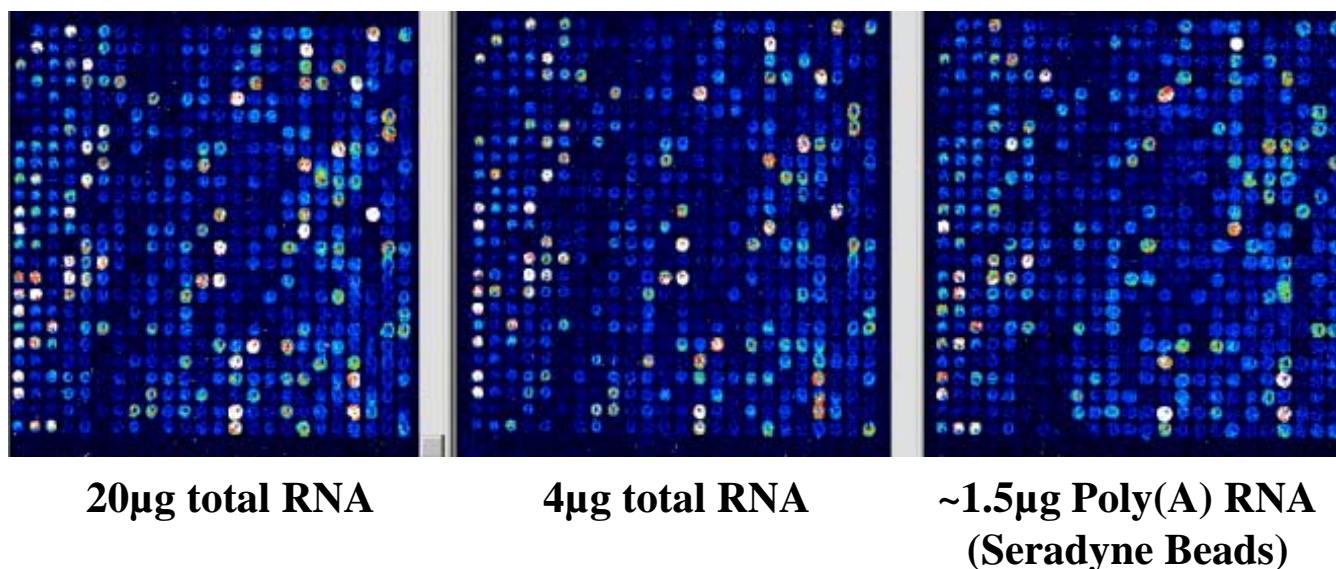
**Figure 2. Effects of various spotting buffers and DNA clean-up protocols on DNA binding and hybridization using Corning CMT-GAPS™ aminosilane coated slides.** This false color image was generated by spotting identical samples in adjacent rows and hybridizing with a labeled mRNA probe; red lines separate paired rows. Paired rows 1-3 and 5-7 contain samples spotted using either 50% DMSO or 3×SSC as a spotting buffer. Comparing spots vertically adjacent to each other, it is clear that spotting with DMSO allows hybridization with significantly higher affinity than does spotting with SSC. In our evaluation, DMSO consistently gives 1.5-fold or greater hybridization intensities with 10% fewer “drop out” spots. Paired rows 4 and 8 show the effects of different clean-up protocols on DNA binding and hybridization. The glass-filter method described in the text gives visibly better results than does simple ethanol precipitation.





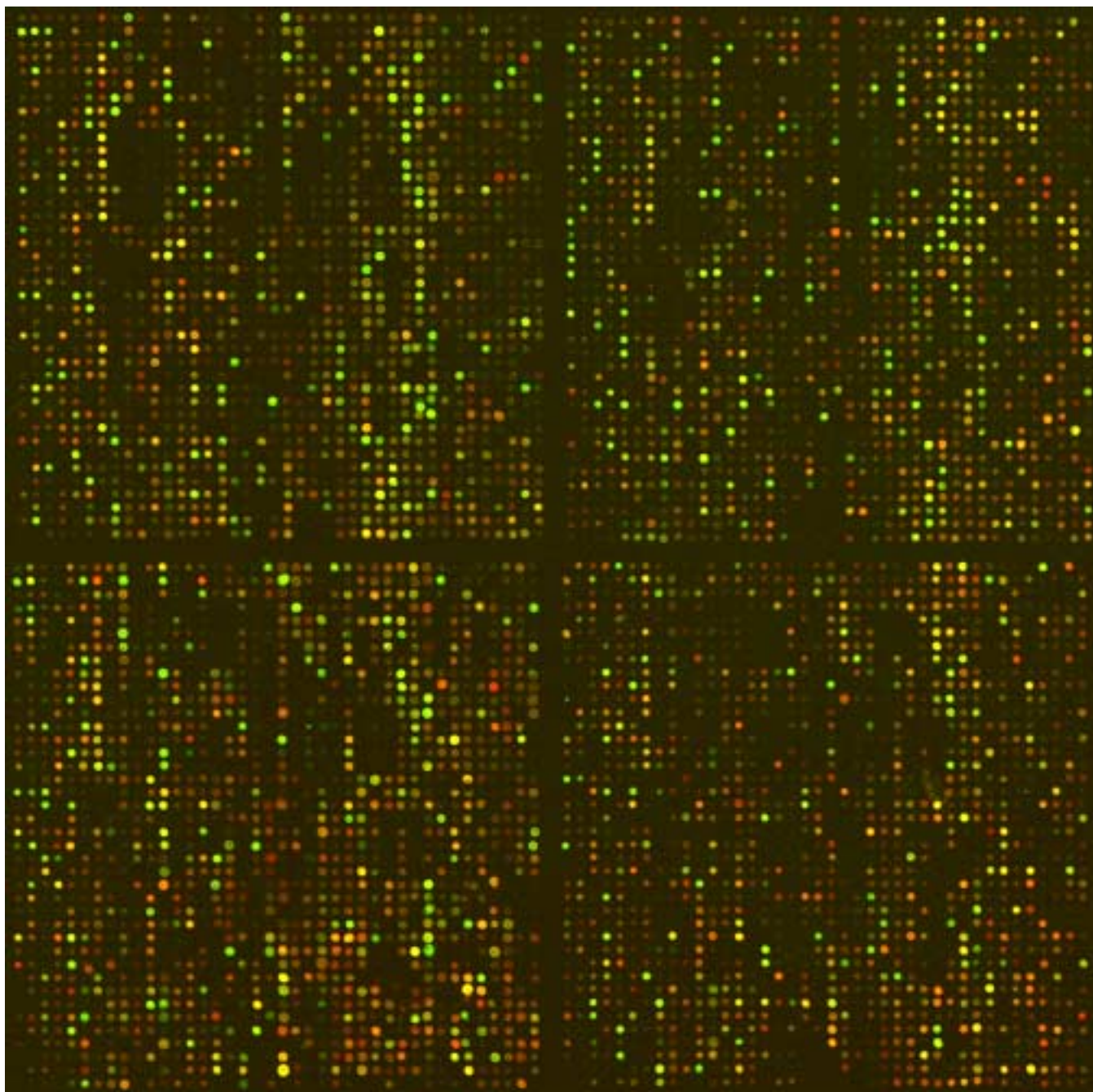
**Figure 3. Effect of temperature and humidity on slide morphology**

Printing on this slide began at 72°F (22.2°C) and 45-50% relative humidity, which we had determined to be optimal. Under these conditions, the spots have a uniform appearance. As printing progressed, the temperature was reduced to approximately 62°F (16.7°C), resulting in smaller, less distinct spots. Temperature and humidity were then increased. The rows of large spots in the center of the slide (rows 13-15) were printed at 80°F (26.7°C) and 80-85% relative humidity. As the temperature and humidity were decreased once more, optimal conditions were again achieved. By correlating data from a chart recorder with the spot number, we were able to determine the conditions that subsequently gave the best hybridization performance. Representative temperature and humidity levels are shown. The cDNA clones in this array were hybridized with a Cy3 labeled vector-specific probe.



**Figure 4. Comparison of hybridization probes made using total and poly(A<sup>+</sup>) RNA.**

Corresponding areas of cDNA microarrays containing 7,200 elements, each hybridized with labeled probe prepared from 20µg of total RNA, 4µg of total RNA, and 1.5µg of poly(A<sup>+</sup>) selected RNA. Note that the relative hybridization intensities are similar for each of the total RNA samples and that both are slightly greater than for the poly(A<sup>+</sup>) hybridized sample. In our experience, 4-10µg of total RNA gives consistently high quality hybridization results. All images were obtained using the same laser and PMT settings during scanning and are displayed using the same parameters.



**Figure 5. Hybridization of mRNA to a portion of a 19,200 element array.** Using the protocols described in this manuscript, 10 $\mu$ g of RNA extracted from related human colon carcinoma test and reference cell-lines (KM12L4A and KM12C respectively) was reverse transcribed and labeled with Cy-5 dUTP and Cy-3 dUTP, respectively. These were then hybridized to a microarray containing 19,200 distinct human cDNA clones. The contrast on the image has been adjusted to allow the majority of the spots in the array to be easily visualized. The protocols outlined in this manuscript consistently provide hybridization results similar to this.