

A Groundtruth Approach to Accurate Quantitation of Fluorescence Microarrays

*L. M. Kegelmeyer, L. Tomascik-Cheeseman, M. S.
Burnett, P van Hummelen, A. J. Wyrobek*

This article was submitted to
Society of Photo-Optical Instrumentation Engineers Photonics West
Conference, San Jose, CA., January 20-26, 2001

December 1, 2000

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

This work was performed under the auspices of the United States Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

This report has been reproduced directly from the best available copy.

Available electronically at <http://www.doc.gov/bridge>

Available for a processing fee to U.S. Department of Energy
And its contractors in paper from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-mail: reports@adonis.osti.gov

Available for the sale to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-mail: orders@ntis.fedworld.gov
Online ordering: <http://www.ntis.gov/ordering.htm>

OR

Lawrence Livermore National Laboratory
Technical Information Department's Digital Library
<http://www.llnl.gov/tid/Library.html>

A groundtruth approach to accurate quantitation of fluorescence microarrays

Laura Mascio Kegelmeyer^{a,*}, Lisa Tomascik-Cheeseman^a, Melinda S. Burnett^{a,†}, Paul van Hummelen^{a,‡}, Andrew J. Wyrobek^a

^aLawrence Livermore National Laboratory, Livermore, CA 94550

ABSTRACT

To more accurately measure fluorescent signals from microarrays, we calibrated our acquisition and analysis systems by using groundtruth samples comprised of known quantities of red and green gene-specific DNA probes hybridized to cDNA targets. We imaged the slides with a full-field, white light CCD imager and analyzed them with our custom analysis software. Here we compare, for multiple genes, results obtained with and without preprocessing (alignment, color crosstalk compensation, dark field subtraction, and integration time). We also evaluate the accuracy of various image processing and analysis techniques (background subtraction, segmentation, quantitation and normalization). This methodology calibrates and validates our system for accurate quantitative measurement of microarrays. Specifically, we show that preprocessing the images produces results significantly closer to the known groundtruth for these samples.

Keywords: groundtruth, cDNA, quantitative, expression, fluorescence, microarrays

1. INTRODUCTION

Expression microarrays provide a means for monitoring the operations of many genes at once; they can also be used for in-depth understanding of certain processes, such as DNA repair, cellular differentiation, and developmental biology. These areas are the focus of our efforts to accurately measure the expression levels demonstrated by microarrays.

Because our custom microarrays are designed to address pathway-specific topics, they require careful target selection. Targets are selected not only to represent the potentially interesting and relevant genes, but also to minimize sequence homology (and thus cross-hybridization) among spots. Subtle differences in gene expression measurements can have a large biological impact. Specificity and measurement accuracy are especially important under these conditions.

* Correspondence: Email: kegelmeyer1@llnl.gov; Telephone: 925 422-0924; FAX: 925 422-0924.

† Currently at University of California, San Francisco, CA 94143

‡ Currently at Flanders Institute of Biotechnology, Leuven, Belgium

There are many ways to perform the image processing and analysis steps needed to derive quantitative information from a microarray image. Examples can be found in the literature (Ref Brown , Chen 1997, Pie'tu 1996) and there are dozens of commercial packages with their own analysis methods. We evaluate a number of different methods for processing and analyzing images, and in addition, we show the effects of preprocessing. We define:

- Preprocessing is characterizing and accounting for the acquisition system parameters such as camera dark field, spectral crosstalk, image alignment, integration time and camera gain.
- Background subtraction removes the effects of autofluorescence and other effects that are not due to specific fluorescent hybridization.
- Segmentation delineates the extent of each spot, and thus distinguishes spots from surrounding background.
- Quantitation is the process of measuring intensities within the spot boundaries determined by segmentation.
- Normalization makes the measurements from the 2 color channels (e.g. red and green) commensurable so they may be sensibly compared or arithmetically manipulated relative to one another. It also enables slide to slide comparisons.

Each step can have optional techniques and implementations. Not only are there a number of different possible algorithms for each step, but the optimal combinations may differ from system to system. We use groundtruth to determine which techniques and which combination of the above options are optimal for correlating computed intensities with known probe amounts for our system. This methodology can be used to calibrate any acquisition and analysis system to improve measurement accuracy.

2. MATERIALS

2.1 The Arraying Robot for Microarray Preparation

A custom robotic high speed arrayer was used to grid cDNA onto glass microscope slides. The gridding system had a 3-axis DC servo driven gantry (GM2340R, Glentek, El Segundo, CA). The full travel of the Z-axis was 0.25m with 5 μ m resolution and of the X-and Y-axis (powered by Newport-Klinger MD4 servo motor driver), 2m and 1m, respectively, with 20 μ m resolution each. The system controller was a Newport-KlingerMM2000 card with 3DC modules which resides in an Intel 80486 PC. The spotting tool had 2 Beryllium-Copper plated pins with a spacing of 4.5mm. The grid density was 4.5 x 4.5 mm for each of 2 pins with a center to center spacing of 375 μ m. After spotting, the slides were humidified in a 37 $^{\circ}$ C incubator for 5 minutes, air dried and stored under vacuum at room temperature until hybridization.

2.2 Image Acquisition

Fluorescent microarray images were acquired with a full-field (15 mm square with a resolution of 0.015mm/pixel), white light imaging system purchased from Norgren Associates (Palo Alto, CA). Light from a 500 watt Xenon light source (400 nm - 600 nm) was scrambled through a fiber optic and passes through a bandpass excitation/emission filter pair before reaching a scientific grade CCD camera. The camera was controlled via a personal computer (PC) with a MATROX PULSAR digital acquisition and display board. Images were collected onto the PC and then transferred to a Unix workstation for analysis.

This system was modified by adding a second light source (a mercury arc lamp) for UV excitation. The UV wavelength allows the excitation of a third dye, called DAPI, which is described in Section 2.4.

2.3 Image analysis

Custom software was written within the SCIL-Image development platform (TNO/TPD, Delft, The Netherlands) for preprocessing, automated background subtraction, grid detection, spot detection and quantitation. Custom Perl scripts were used to ratio, normalize, summarize and plot data resulting from the image analysis.

2.4 Dyes & antifade

Fluorescent dyes conjugated to either dUTPs or dCTPs were incorporated as described in Section 2.7. Our system is equipped to image a wide variety of dyes with excitation wavelengths in the visible and ultraviolet range. We have used Cy3, Cy5, FITC, the Alexa dyes and DAPI (4',6-Diamidino-2-phenylindole). DAPI is a common nucleic acid stain used to mark the presence of DNA. We apply it as a third color to help locate all spots, including those that did not hybridize.

Vectashield mounting media (Vector Laboratories Inc., Burlingame, CA) was used to prevent photobleaching of the fluorescent dyes.

2.5 Generating amino modified clones

Expressed sequence tag (EST) clones were obtained from LLNL's I.M.A.G.E. consortium and used as the target cDNA for our microarrays. These clones were amplified using sequence-specific and vector-specific primers and verified on a 1-2% agarose gel. The ESTs had an average insert size ranging from 500 to 1500 base pairs. Prior to spotting, inserts from the selected clones were PCR amplified from plasmid preparations using 5' C6 amino-modified vector specific primers and purified using Qiagen PCR purification columns (Qiagen Inc., Chatsworth, CA). The amplicons were ethanol precipitated and resuspended for spotting in 0.1M Sodium Carbonate/Bicarbonate (pH10.2) to a final concentration of 2µg/µl.

2.6 Preparation of Slides

Slides were derivatized according to Guo et al. (Ref 1994). Briefly, glass slides were coated with 1% 3-aminopropyltrimethoxysilane [Sigma, St. Louis, MO] in 95%acetone/water for 2 min., washed 10 times in acetone and baked at 110°C for 45 min. Prior to spotting the silane was activated by incubating the slides for 2 hours in 0.2% 1,4 phenylene diisithiocyanate [Sigma, St. Louis, MO], 10% pyridine [Sigma, St. Louis, MO] and dimethyl formamide [Aldrich, Milwaukee, WI]. Slides were then washed in methanol (2x 10 min.) and acetone (2x 10 min.). Slides were air dried and used immediately for spotting.

2.7 Labeling Probes

To create groundtruth over a series of 10 slides, 10 probe mixtures were made using clones for the following genes: Globin, Skid, p53, Rad 50, Rad52, and Ku86. Each mixture contained a Cy3 and FITC labeled probe of each of the above genes. Globin and Skid probes had equal proportions across all slides to serve as controls and Globin was used for normalization. For the other 4 genes, the probe amount of one of the dyes was held constant while the same probe was varied in the other color, forming a dilution series as shown in Table 1. Each mixture was hybridized to one slide.

Globin	FITC	2.3 ng	→	2.3 ng	RAD52	FITC	18.2 ng	→	0.01 ng
	Cy3	2.3 ng		2.3 ng		Cy3	3.0 ng		3.0 ng
Skid	FITC	2.4 ng	→	2.4 ng	Ku86	FITC	3.2 ng	→	3.2 ng
	Cy3	2.4 ng		2.4 ng		Cy3	18.6 ng		0.01 ng
P53	FITC	20.4 ng	→	0.01 ng	RAD50	FITC	3.0 ng	→	3.0 ng
	Cy3	3.3 ng		3.3 ng		Cy3	17.256 ng		0.01 n

Table 1: 6 genes were used to build a groundtruth series on 10 slides. The ratio of the FITC and Cy3 probe amounts shown are expected to correlate with the subsequent ratio of intensity measurements. Globin and Skid have constant amounts in both colors for all slides, and should each yield a green/red ratio of 1, serving as positive controls for the whole series of slides. The arrows represent the serial change in the probe amount across the 10 slides.

The probes were generated by PCR amplification of the same I.M.A.G.E. EST clones that were spotted on the array using gene-specific primers. The probe labeling was performed by mixing serial dilutions for each gene together per color and afterwards replacing dCTP by Cy3-dCTP (Amersham Pharmacia Biotech, Piscataway, NJ) or dUTP by FITC-dUTP (dUTP (Roche, Indianapolis, IN) using a Nick Translation kit (Life Technologies, Rockville, MD). To incorporate labeled nucleotides during nick translation, the DNA is "nicked" by the enzyme DnaseI and DNA polymerase replaces the excised nucleotides. The Cy3 and FITC labeled probe mixtures were added together and purified with Qiaquick (Qiagen Inc., Chatsworth, CA).

2.8 Hybridizations

To block non-specific probe binding, slides were incubated for 10 min. in 1% NH₄OH (Sigma, St. Louis, MO) and washed 3 x 10 minutes in double distilled water prior to hybridization. Slides were then denatured for 6 minutes in 70% Formamide/2xSSC at 78°C, dehydrated through a 70- 85-100% ethanol series, and air dried at room temperature.

Labeled probes were concentrated using speed vacuum centrifugation and resuspended in 5-10µl hybridization mix containing 70% formamide, 2x SSC, 10% dextran sulfate and 1µl herring DNA. The hybridization mixtures were placed on the microarray under a 22x22 mm coverslip, sealed with rubber cement, and denatured at 72°C for 3 minutes. Hybridization was performed for 24 hours at 37°C in a moist chamber, followed by two 10 minute washes at 37°C in 2xSSC.

3. IMAGE PROCESSING & ANALYSIS METHODS

3.1 Preprocessing

To calibrate the acquisition system, we measure the inherent system response characteristics and correct or account for them in the analysis. For example, all CCD cameras have thermal noise which is present even when no light is incident on the CCD. This is called dark noise and it can be measured and then subtracted from images, with a method called dark field subtraction (Ref Mullikin). Similarly, the filters in the system do not have perfect bandpass cutoffs and light through the filters is not always parallel, reducing their effectiveness. This results in spectral crosstalk which, for example, can cause the green signal to leak into the red channel and vice versa. All systems

should measure and correct for this (Ref Castleman). Image alignment is another issue which can be corrected, via correlation methods, during preprocessing. Misalignment occurs when there is a physical shift of the microarray slide between the times that each dye is imaged. Lastly, we correct up-front for integration time differences between the two images. While it would be feasible to leave this correction until the end, during normalization, it is prudent to correct for known disparities in a straightforward manner. Then, the factors left to be corrected by normalization are those for which reliable characterization is not available.

Two of the system response characteristics, camera gain and non-uniform illumination, are difficult to correct, but can be ameliorated. The camera gain is non-linear for higher gain settings, so to avoid needing a correction for different acquisitions at different gain settings, we hold the gain constant during acquisition for all images in the same experiment. Correcting non-uniform illumination is difficult because of wavelength dependencies (chromatic aberration) and variation of the light source over time; therefore, the illumination pattern is optimized for uniformity as much as possible during acquisition and is not corrected computationally.

3.2 Image Analysis

3.2.1 Background subtraction

After preprocessing, one can begin to extract quantitative information from hybridization signals. The first step is usually background subtraction. Note that the camera dark field subtraction (discussed earlier in Section 3.1) is not the same as a background subtraction. The former is inherent to the camera and the latter to autofluorescence of the substrate, antifade or some of the blocking chemistries, as discussed below.

Before performing a background subtraction operation, it is important to know whether background subtraction is appropriate, at all, for the image at hand. This requires knowledge of the background and its physical composition. There are at least 3 sources of fluorescent background: autofluorescence inherent to the glass slide or other substrate, autofluorescence of the antifade, and the non-specific binding of fluorescent probe to material in the area surrounding the target spots of the array. Informal investigations have shown that different substrates have varying levels of autofluorescence. For some systems, high quality, polished quartz, such as Corning brand slides have shown lower contribution to background fluorescence than standard glass microscope slides. If background fluorescence can be attributed to autofluorescence of the substrate, it stands to reason that a slowly varying background intensity lies under the array and serves as a "DC offset" to the spot signals. (All spot intensities should be above this background.) For such a case, the background subtraction is a logical and necessary step for accurate quantitation of spot signals. This is the convention adopted and applied in quantitative analyses for fluorescence microscopy. (Reference Mascio, et al.)

If background fluorescence can be attributed to non-specific binding of probe, then the surrounding signal is not "under" the spots, but only around them, because the target DNA spotted on the slide prevents the non-specific binding at the spots. In this case, the background may have higher intensity than the spots! (See Figure 1.)

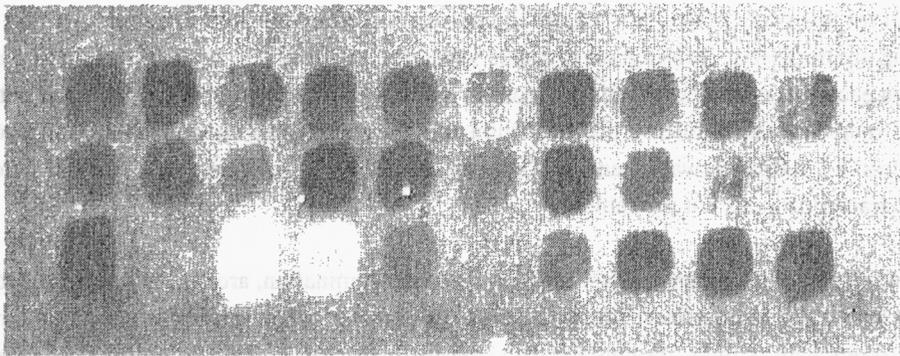


Figure 1: A contrast-enhanced example of a fluorescence microarray image with very high background. The dark appearance of many spots demonstrates that the high background is not additive to the spot intensities and thus should not be subtracted. Presumably, the target DNA that was robotically placed on the slide blocked those sites for non-specific hybridization.

In this case, background subtraction is unwarranted and will result in negatively valued spot intensities. A better approach for this type of image is to hybridize another array after applying an appropriate blocking procedure to prevent non-specific probe binding. Alternatively, spots can be compared to the negative controls (spots that had no hybridization probe) to get a relative sense of their brightness, but accurate ratios may be hard to obtain, since there may still be autofluorescence from the substrate which can't be characterized or subtracted in this case.

Once it has been determined that background subtraction is appropriate for a given image, there are a number of methods that can be used to perform the task: some involve sampling pixel intensities near the extremes of the image, or just outside of spots. The background subtraction method we employ is an adaptive, non-linear technique which continuously samples the background and creates a 2-dimensional, slowly changing "sheet" or blanket that follows the low frequency trend of intensity outside of the spots. This method is called the lower envelope subtraction and is described by Ref Verbeek, et al. It serves to lower the signal intensities relative to a base value near zero. This is presented visually in Figure 2.

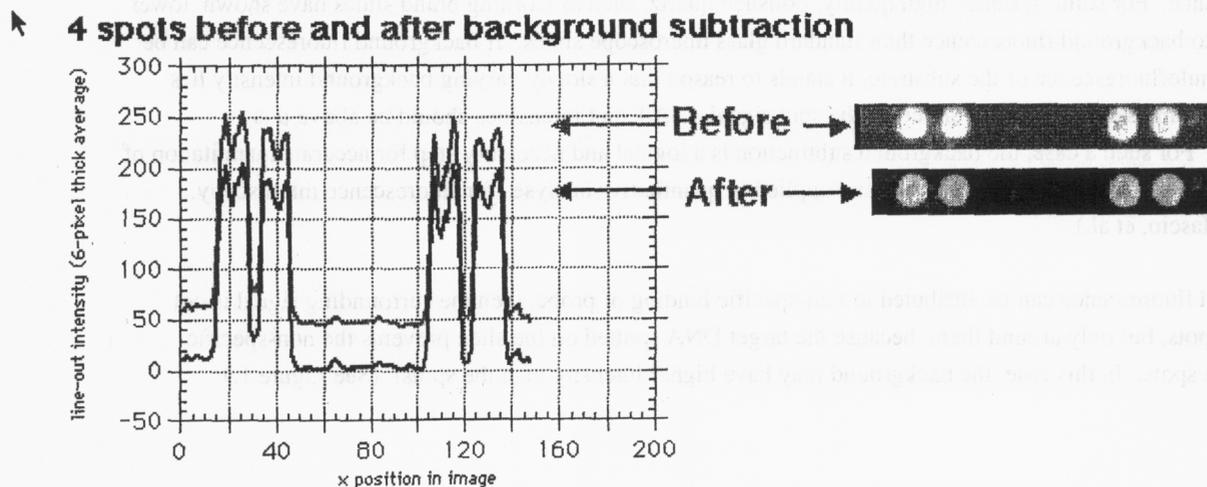


Figure 2. The grayscale images to the right represent a row of spots before and after background subtraction. The graph to the left plots intensity vs position along a line through the spots. The graph illustrates how background subtraction removes the effects of the underlying autofluorescence and lowers the spot intensity appropriately.

3.2.2 Semi-automated grid placement

Even though a precision robot spots the cDNA target, detecting spots and their extent is challenging because of shape and location anomalies. Our system uses a semi-automated grid placement algorithm which requires an operator to indicate the extent of the array and number of spots present. With this information, the system divides the area into self-adjusted, not necessarily uniformly sized grid squares, each containing one spot.

3.2.3 Three-color segmentation

Once a grid is overlaid on the array, one must utilize segmentation methods to determine which pixels inside a grid square belong to the spot under investigation and which to the background. Segmentation is more difficult for dim or non-expressing spots than for bright spots.

A good way to address this problem is to acquire a third image of the array after the first two images have captured data from the ratio probes. The third image should be captured after applying a DAPI DNA counterstain to the slide. This fluorescent stain is designed to mark the location of all DNA in a targeted location. The presence of DAPI stain removes any ambiguity about the presence of a valid target DNA in the case where no expression signal is observed. It also allows better determination of the size and shape of the DNA target for segmentation.

We found that DAPI is not always present in spots with a strong hybridization signal. (It may be that the hybridization of the labeled probes changes the shape of the DNA molecule and does not allow intercalation.) Therefore, after aligning them to each other, we combine the signals from all 3 dyes to get the best signal from each spot, and use this result for segmentation when DAPI is applied.

3.2.4 Methods for spot segmentation

Since microarray images can have very dim, "bagel-shaped" spots or other irregularities, especially when DAPI is not used, different segmentation methods will give different spot areas and will affect the total intensity measurement. We studied the following segmentation methods and show an example of each in Figure 3 below:

Circle Hough Transform: Find the best fit circle by using an edge detector to find edges and then determine their magnitude and direction. Use this information to vote for the center of the circle that best fits the edge information. (Ref. Ballard)

Trian Threshold: Take a histogram of the intensities in a grid square containing one spot. Find the low-intensity peak of the histogram (which represents background values) and fit a line to the falling slope on the right side, forming a triangle (hence the name trian) with the y axis. The clipping level (intensity threshold) is where this line crosses the x-axis of the histogram. (Ref Zack, et al)

Circles + Trian: Use a logical OR to combine the results from the two methods above.

Hulled Trian: Apply the morphological convex hull operation to the result of the Trian Threshold.

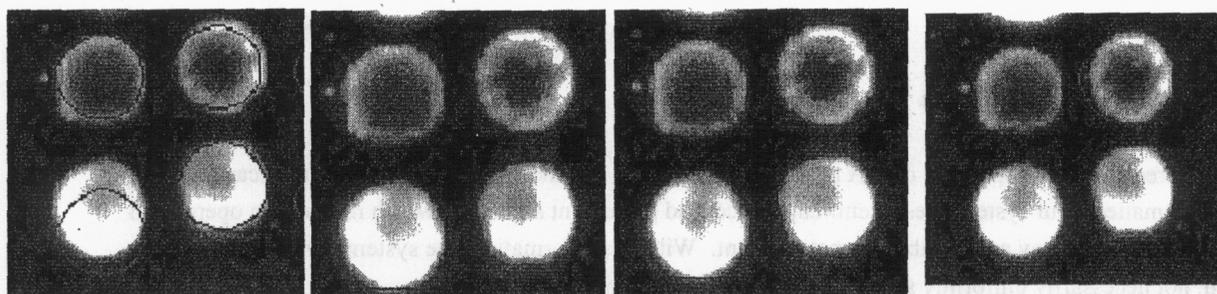


Figure 3. Countours of areas resulting from different segmentation methods. From left to right: Circle Hough transform, Trian threshold, Hulled Trian, and Circles + Trian.

3.2.5 Methods for quantitation of signal intensity

A number of methods can be used to compute the relative intensities of the red and green probe signal for a spot. The differences between the methods lie primarily in how well they accurately quantify the signal in the face of noise or other artifacts. We implemented and evaluated a number of methods:

- **Total Intensity.** Sum the intensities of all of the pixels assigned to a spot, for each color, and then divide them to get the ratio:

$$\frac{\sum_{i=1}^n G_i}{\sum_{i=1}^n R_i} \quad (1)$$

where G is the green signal, R is the red signal, i is a pixel in the spot and n is the total number of pixels in the segmented region.

- **Median.** Find the median pixel intensity in each color and take the ratio:

$$\frac{\text{(Median of red pixel values)}}{\text{(Median of green pixel values)}}$$

This method is superior to the total (or averaging) method if you have noisy outliers (i.e a few really bright or dim pixels). It works especially well if you have a similar number of low and high outliers.

- **Pixel-to-pixel mean and average ratios.** Divide the intensities of a spot on a pixel:pixel basis. Then, calculate the mean or the median of the pixel:pixel ratios for the entire spot. This method assumes that the two wavelength images are perfectly registered and that the spot shape is consistent in both colors.
- **Fit line.** Since the pixel:pixel ratios within a spot can vary a lot, especially for bagel-shaped spots, another approach is to plot the pixel intensities in one wavelength versus another for each pixel and then fit a line to the resulting scatterplot. Each point on the scatterplot represents a pixel, with the x location its intensity in one wavelength and the y location its intensity in another. The slope of the line that passes through the 0 y-intercept would represent the final ratio.
- **Log geometric mean.** The log geometric mean of the intensities in a spot is

$$10^L \text{ such that } L = \frac{1}{n} \sum_{i=1}^n \log Y_i / X_i \quad (2)$$

where Y represents the green signal, X represents the red signal, i represents one pixel in a spot and n is the number of pixels in the spot.

3.2.6 Normalization

Once an intensity ratio has been determined for a particular spot, it must be normalized to a standard if it is to be meaningfully compared to other spots on the same slide or spots on a different slide. Some color-to-color and slide-to-slide variations include differences in exposure time, amount of target, amount of probe, dye incorporation, rate of photobleaching, hybridization conditions, and imaging conditions. Normalization can be achieved by dividing all spot intensities on a slide by the average ratio of positive controls, which are spots that are designed to, ideally, have the same intensity and ratio on all slides. (We use equal amounts of Globin in both colors as the positive control for our groundtruth series.) When positive controls don't have the expected ratio, we assume the differences are due to the above variations and we use the factor by which they differ as a correction factor all other spots.

4. RESULTS

4.1 Preprocessing Results

In order to measure the effects of various analysis steps, we compared the computed intensity ratios against the known ratio of known probe quantities. In Figure 4, the two are plotted such that the distance between the curves represents the error between measured and known quantities, with and without preprocessing. Visually, one can see that the curves are much closer together over a greater range of probe quantities "with preprocessing". It is also possible to gauge the probe amounts at which the two curves diverge, and to note that the computed intensity ratio is more accurate down to lower probe amounts.

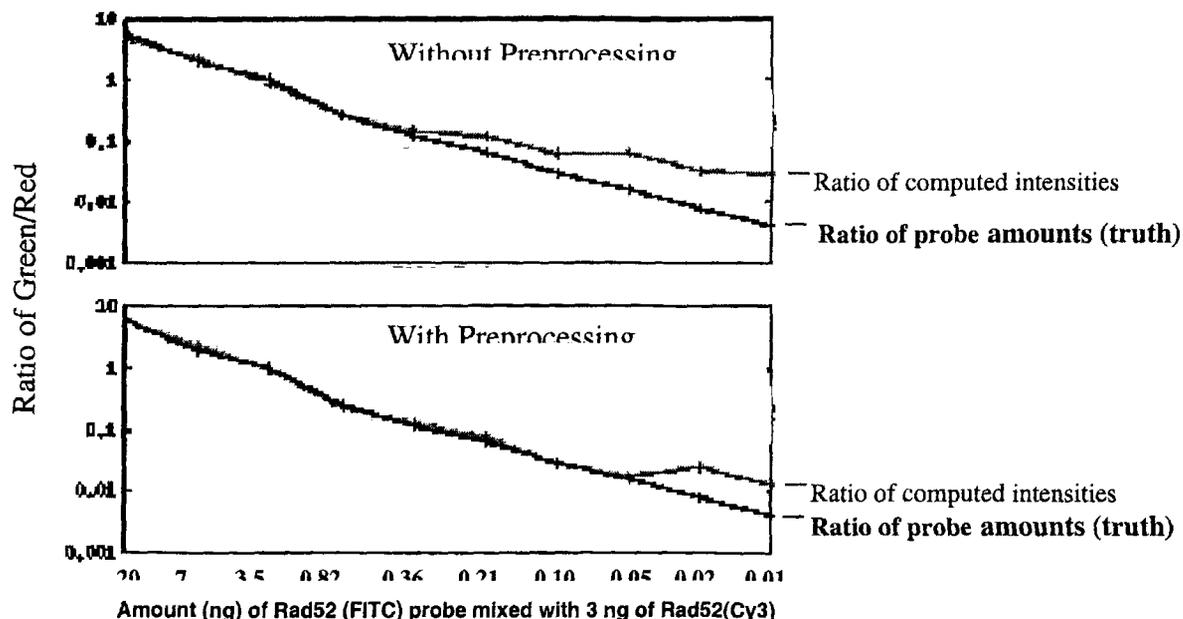


Figure 4: Groundtruth (ratio of known probe amounts) for the Rad52 gene is tracked among the 10 slides as shown by the (lower) dark black line. The ratio of measured intensities is shown as the (upper) grey line. The cumulative distance between these two lines gives an error measure (Section 4.2) so different methods can be applied and compared quantitatively.

4.2 Error evaluation for various analysis methods and combinations

In order to quantify the disparity among the curves, we calculate an error by summing the distance from measurement to the truth for all 10 data points:

$$E = \sqrt{\sum_{N=1}^{10} \{\log(\text{intensity ratio}) - \log(\text{quantity ratio})\}^2} \quad (3)$$

Numerically, the error for the graphs in Figure 4 is 1.3 for the top graph (without preprocessing) and 0.7 for the bottom graph (with preprocessing). These values are shown in the error chart in Table 2 with quantitation type "median".

	Full preprocessing			No preprocessing, except alignment			Log Gmean	Cumulative Error per gene
	Total Intensity	Median	Log Gmean	Total Intensity	Median	Log Gmean		
Globin:	0.1	0.1	0.1	4.3	0.1	0.1	4.7	
Skid:	0.3	0.4	0.4	0.3	0.3	0.3	1.9	
p53:	1.0	1.1	1.0	2.9	1.6	1.6	9.1	
Rad50:	1.3	1.3	1.3	1.9	1.7	2.1	9.6	
Rad52:	0.7	0.7	0.7	1.3	1.3	1.4	6.0	
Ku86:	1.3	1.4	1.4	1.9	1.8	2.1	10.0	
Cumulative error per method for the 4 genes of the dilution curve:	4.2	4.5	4.5	7.9	6.4	7.1		

Table 2: Values shown are the error, E (see Eq. 3), which represents the cumulative distance between truth and measured curves such as those in Figure 4. Similar curves were generated for the two control genes and the four genes of the dilution series for various combinations of analysis methods. Their error, E, is reported here. These serve as a metric for comparing the effects of different analysis techniques.

From this chart, we conclude that preprocessing is most valuable when signal strength is low and when there is an intensity disparity between the two probe amounts. Probes that were plentiful and equal in the two colors, such as Globin and Skid, did not benefit substantially from preprocessing. The other four genes, however, show significantly larger improvements with preprocessing. This is primarily due to the color correction which takes intensities from one color channel and replaces them in the other where they belong. For dim spots, this incremental change can be substantial, whereas bright spots are minimally affected.

Table 2 also shows the effect of different quantitation methods. As seen in the cumulative error, along the last row of the chart, the total intensity quantitation method can give the lowest error (4.6 with full preprocessing), but it can also give the highest error (no preprocessing, except alignment) when data is corrupted by debris or other noise. Therefore it is not the most robust quantitation method. Instead, the median and the log geometric mean error values were more consistent and overall lower than other methods tried (which are not shown on this chart, but are described in Section 3.2.5).

The chart also demonstrates the importance of normalizing in a valid way. Since all of the analyses used to generate the data for the chart were normalized via the Globin positive control spots, it stands to reason that the largest cumulative error occurs when the Globin control spot contained undetected debris and the quantitative method of "total" included those values in the intensity measure for Globin (Table 2). Since that measurement was flawed, and since that value was subsequently used to normalize all of the gene intensity ratios, all computed ratios using that method were flawed.

Just as these errors were generated for various combinations of preprocessing and quantitation parameters, we were able to generate similar values for the other analysis steps. For evaluation of background subtraction, we used our default settings for preprocessing, segmentation, quantitation and normalization and changed only whether

background was subtracted or not. For this data, the total error improved from 7.8 to 7.0 when background was subtracted. Further, almost all of the improvement was realized for the genes which had low probe amounts. As expected, the spots with strong intensity were least affected by background subtraction.

5. CONCLUSIONS

While there are numerous approaches to microarray analysis, accurate quantitation requires system calibration. The methodology outlined in this paper can be used to calibrate any acquisition and analysis system and can be tailored and optimized for specific data and specific experiments for improved measurement accuracy. In order to calibrate and validate our system, we generated a dilution series with known quantities for the red and green hybridization probes. We used that groundtruth sample to evaluate the results of a number of computational techniques for preprocessing and analysis. Computed ratios were compared to the known groundtruth and an error was calculated for each method. The combination of methods that gave the least overall error for our system were:

- preprocessing was superior to not preprocessing
- subtracting background was superior to not doing so
- spots were best segmented by combining a best circle fit with an intensity threshold
- quantitation was most accurate when using the median or the log geometric mean
- normalizing by a positive control was superior to not doing so.

Preprocessing had the greatest impact on bringing the computed ratios closest to the groundtruth. For the gene with the lowest cumulative error across all analysis methods, Rad52, preprocessing enabled accurate quantitation of DNA probe to 0.05 ng (improved from a limit of 0.36 ng without preprocessing). Future efforts will establish whether this translates to the detectable amount for samples using cDNA or RNA probes and will also evaluate metrics for signal brightness and uniformity.

ACKNOWLEDGEMENTS

We gratefully acknowledge David Nelson for suggesting statistical methods; Don Peters for hardware configuration and support; Damir Sudar for hardware and software consultation; and W. Philip Kegelmeyer for thoughtful reading and editing of the manuscript.

*This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48 with funding from NIH grant number ES09117-02, DOE grant number KP110202 and the University of California Campus Laboratory Collaboration with UCSF.

REFERENCES

MPS Brown, WN Grundy, D Lin, N Cristianini, CW Sugnet, TS Furey, M Ares, Jr. and D Haussler, "Knowledge-based analysis of microarray gene expression data by using support vector machines," *PNAS* **97**:1: pp. 262-267, 2000.

Kegelmeyer, BO 4266-6, page X/11 (Write this on the back of each page, in pencil)

Y Chen, ER Dougherty and ML Bittner, "Ratio-based decisions and the quantitative analysis of cDNA microarray images," *Journal of Biomedical Optics* 2(4) (1997), 364-374.

G Pie'tu, O Alibert, V Guichard, B Lamy, F Bois, E Leroy, R Mariage-Sampson, R Houlgatte, P Soularue and C Auffray, "Novel gene transcripts preferentially expressed in human muscles revealed by quantitative hybridization of a high density cDNA array," *Genome Research* 6: pp. 492-503, 1996.

Z Guo, RA Guilfoyle, AJ Thiel, R Wang, LM Smith. "Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports," *Nucleic Acids Res.* 22, pp. 5456-5465, 1994.

JC Mullikin, LJ van Vliet, H Netten, FR Boddeke, G van der Feltz, IT Young. "Methods for CCD camera characterization," *Image Acquisition and Scientific Imaging Systems*, Helen C. Titus and Amir Waks Eds. Vol 2173, pp 73-84, SPIE, San Jose, CA. 1994.

KR Castleman, TP Riopka, and Q Wu. "FISH image analysis." *IEEE Engineering in Medicine and Biology.* 15, No 1, pp. 67-75, 1996.

LN Mascio, PW Verbeek, D Sudar, W-L Kuo and JW Gray. "Semiautomated DNA probe mapping using digital imaging microscopy: I. System development," *Cytometry* 19: pp. 51-59, 1995.

PW Verbeek, HA Vrooman, LJ Van Vliet, "Low Level image processing by max-min filters. *Signal Process* 17:249-258, 1988.

DH Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111-122, 1981.

GW Zack, WE Rogers, and SA Latt, Automatic measurement of sister chromatid exchange frequency. 1977. 25(7): p. 741-753.