

Methodology article

Open Access

Utilization of a labeled tracking oligonucleotide for visualization and quality control of spotted 70-mer arrays

Martin J Hessner*^{1,2}, Vineet K Singh^{†3}, Xujing Wang^{1,2}, Shehnaz Khan², Michael R Tschannen² and Thomas C Zahrt³

Address: ¹The Max McGee National Research Center for Juvenile Diabetes, Department of Pediatrics, The Medical College of Wisconsin and Children's Hospital of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, USA, ²The Human and Molecular Genetics Center, The Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, USA and ³Department of Microbiology and Molecular Genetics, The Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, USA

Email: Martin J Hessner* - mhessner@mcw.edu; Vineet K Singh - vsingh@mcw.edu; Xujing Wang - xujing@mcw.edu; Shehnaz Khan - skhan@mcw.edu; Michael R Tschannen - mtschann@mcw.edu; Thomas C Zahrt - tzahrt@mcw.edu

* Corresponding author †Equal contributors

Published: 09 February 2004

Received: 05 December 2003

BMC Genomics 2004, 5:12

Accepted: 09 February 2004

This article is available from: <http://www.biomedcentral.com/1471-2164/5/12>

© 2004 Hessner et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: Spotted 70-mer oligonucleotide arrays offer potentially greater specificity and an alternative to expensive cDNA library maintenance and amplification. Since microarray fabrication is a considerable source of data variance, we previously directly tagged cDNA probes with a third fluorophore for prehybridization quality control. Fluorescently modifying oligonucleotide sets is cost prohibitive, therefore, a co-spotted *Staphylococcus aureus*-specific fluorescein-labeled "tracking" oligonucleotide is described to monitor fabrication variables of a *Mycobacterium tuberculosis* oligonucleotide microarray.

Results: Significantly ($p < 0.01$) improved DNA retention was achieved printing in 15% DMSO/1.5 M betaine compared to the vendor recommended buffers. Introduction of tracking oligonucleotide did not effect hybridization efficiency or introduce ratio measurement bias in hybridizations between *M. tuberculosis* H37Rv and *M. tuberculosis* mprA. Linearity between the mean log Cy3/Cy5 ratios of genes differentially expressed from arrays either possessing or lacking the tracking oligonucleotide was observed ($R^2 = 0.90$, $p < 0.05$) and there were no significant differences in Pearson's correlation coefficients of ratio data between replicates possessing (0.72 ± 0.07), replicates lacking (0.74 ± 0.10), or replicates with and without (0.70 ± 0.04) the tracking oligonucleotide. ANOVA analysis confirmed the tracking oligonucleotide introduced no bias. Titrating target-specific oligonucleotide ($40 \mu\text{M}$ to $0.78 \mu\text{M}$) in the presence of $0.5 \mu\text{M}$ tracking oligonucleotide, revealed a fluorescein fluorescence inversely related to target-specific oligonucleotide molarity, making tracking oligonucleotide signal useful for quality control measurements and differentiating false negatives (synthesis failures and mechanical misses) from true negatives (no gene expression).

Conclusions: This novel approach enables prehybridization array visualization for spotted oligonucleotide arrays and sets the stage for more sophisticated slide qualification and data filtering applications.

Background

Historically, the investigation of genetic alterations has focused on the study of single genes. In recent years, development of DNA microarrays has allowed researchers to study the complete genome of an organism and profile transcriptional expression patterns of up to tens of thousands of genes in a single experiment [1-3]. Consequently, this technology is already providing important insights into the biological properties of various organisms, making it a mainstream component of biomedical research [3-10].

DNA microarrays are currently based on one of two general fabrication formats: (i) light-directed *in vitro* synthesized oligonucleotide arrays (Affymetrix) [11,12], or (ii) spotted cDNA or oligonucleotide arrays [3]. While both platforms allow researchers to compare gene expression profiles, each format offers unique advantages and disadvantages. *In situ* synthesized oligonucleotide arrays offered by Affymetrix are part of a turn-key system that includes both the DNA microarray platform and the corresponding bioinformatic tools. Consequently, investigators can rapidly begin gene expression studies using "off the shelf" reagents and optimized protocols. However, the system and reagents are expensive, thereby limiting the extent of necessary technical and biological replication [13]. In addition, Affymetrix-based arrays cannot be fabricated in the academic laboratory, and offer less flexibility in their design and content compared with spotted cDNA- or oligonucleotide-based arrays. In contrast, spotted cDNA- or oligonucleotide-based arrays utilize high-speed robotics to mechanically or piezoelectrically [14] deposit small volumes of desired DNA probes onto solid support surfaces. Although substantial optimization of the fabrication process by the user laboratory is required, mechanically spotted arrays can be fabricated "in house", are highly flexible in their design and content, and can be manufactured in reasonably large numbers providing a marked cost savings over commercial array platforms.

Spotted oligonucleotide arrays exhibit a number of advantages over cDNA arrays. For example, oligonucleotides can be synthesized such that homologous sequences between genes can be excluded, thereby enhancing specificity. In addition, a given gene can be represented by a set of different oligonucleotides targeting different regions or exons, allowing for the detection of splice variants, or discrimination of closely related genes, strains, or species. Finally, oligonucleotides can be purchased "ready-to-print", by-passing the labor-intensive probe preparation steps required for utilization of cDNA-based microarrays.

A substantial limitation in the utilization of spotted cDNA- or oligonucleotide-based arrays are their susceptibility to quality control issues, resulting primarily from

variable DNA probe deposition and retention on the solid support surfaces. To minimize variations using this fabrication platform, a number of approaches have been described that allow direct visualization of array integrity following printing and blocking procedures. Commonly used methods include the staining of microarrays with DNA-binding fluorescent dyes, or the hybridization of "universal" targets (i.e. random 9-mers) to the spotted DNA elements [15,16]. While these techniques provide useful information regarding the physical characteristics of the array, its integrity may be compromised during subsequent de-staining or stripping procedures required prior to hybridization of labeled targets [16]. Consequently, investigators typically only examine one or a few representative slides to access the quality of a printed batch.

Previously, we have reported the development and use of a novel three-color cDNA array platform that allows immobilized probes to be directly visualized [17-19]. Utilizing this format, oligonucleotide primers used to amplify cDNA targets are labeled at their 5' end with fluorescein, a dye compatible with commonly used cyanine labeling dyes and confocal laser scanners possessing narrow bandwidths [18,20]. Element/array morphology, surface DNA deposition/retention, and surface background can be monitored on each slide. Thus, in our laboratory, all cDNA arrays are imaged for quality control prior to hybridization, maximizing the use of quality arrays for subsequent experimental procedures. It is likely that many or all of the benefits to using a directly-coupled fluorophore are also applicable to oligonucleotide-based microarrays; however, synthesis costs make this approach unfeasible. In this report, we describe the use and evaluation of a *Staphylococcus aureus*-specific fluorescein-labeled 70-mer "tracking" oligonucleotide as a third-color quality control measure of a *Mycobacterium tuberculosis*-specific oligonucleotide-based microarray.

Results and Discussion

Rationale and design for tracking oligonucleotide

Utilization of oligonucleotide-based DNA microarrays is limited by the inability to monitor quality control for every array prior to hybridization. If costs were no consideration, it would be ideal to label each oligonucleotide in a given microarray set with a third fluorescent dye so that array fabrication variables could be directly monitored, analogous to the direct labeling methodology we have utilized with cDNA-based microarrays [18,19]. However, since this practice is currently cost-prohibitive, we reasoned that introduction of a third-color "tracking" oligonucleotide into the printing buffer at low molarity would allow printing fidelity of a target-specific oligonucleotide set to be indirectly monitored. A *M. tuberculosis* oligonucleotide probe set consisting of 4,269 70-mer oligonucleotides (Operon) was chosen for these studies as our

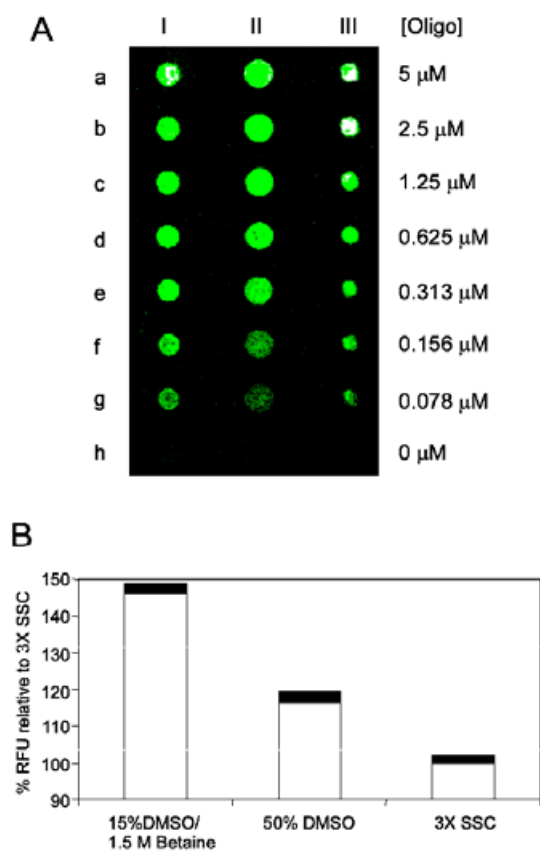


Figure 1
 Evaluation of spotting buffers for 70-mer oligonucleotides. A: Serial dilutions of 5'-fluorescein labeled *S. aureus* tracking oligonucleotide were printed in 15% DMSO/1.5 M betaine (I), 50% DMSO (II), or 3X SSC (III) print buffers at the indicated concentrations and spotted onto poly L-lysine coated glass slides. Slides were processed, blocked, and scanned for fluorescein fluorescence. B: Percent retention of 5'-fluorescein-labeled tracking oligonucleotide on poly L-lysine coated glass slides were determined for each print buffer using 25 replicate arrays. Retention of oligonucleotide in each print buffer was calculated by determining relative fluorescence units (RFU's) at 0.625 μ M (row d), 0.313 μ M (row e), and 0.156 μ M (row f) from replicate arrays. Values are expressed as percent RFU (+/- standard deviation) after normalization to the 3X SSC print buffer. Highest retention was obtained with the 15% DMSO/1.5 M betaine print buffer.

laboratory is currently conducting expression profiling studies on *M. tuberculosis* strains (Singh and Zahrt, unpublished data).

Probe retention studies

Given the cost of probe material and dependence of data quality/reproducibility on the amount of support-bound

probe available for hybridization [16,19], maximizing DNA retention is a critical first step for laboratories fabricating either cDNA- or oligonucleotide-based microarrays. An advantage to directly labeling the material being spotted is the ability to easily monitor probe deposition efficiency, as there is a direct correlation between the amount of DNA probe immobilized on the slide surface and the corresponding fluorescence intensity [18,19]. Here, the utility of a fluoresceinated "tracking" oligonucleotide was investigated to indirectly monitor probe deposition efficiency of a target-specific oligonucleotide set. To first determine the optimal printing buffer for the 70-mer oligonucleotide probe set on poly-L-lysine coated glass slides, a series of 2-fold dilutions of a 5'-fluorescein-labeled *S. aureus*-specific tracking oligonucleotide ranging from 5 μ M to 0.078 μ M were prepared in two commonly used printing buffers 50% DMSO and 3X SSC, as well as 15% DMSO/1.5 M betaine. Arrays ($n = 25$) were evaluated for DNA retention and spot morphology (size/shape) by measuring fluorescein fluorescence after post-processing procedures. Oligonucleotides printed between 0.625 μ M and 0.156 μ M were used for direct comparison in these experiments, as fluorescence from these dilutions were unsaturated in all three printing buffers [Figure 1; panel A, rows d-f]. As previously observed with cDNAs [18], significantly higher retention of the 70-mer tracking oligonucleotide was achieved utilizing the 15% DMSO/1.5 M betaine (5158 +/- 2686 RFU/element; $p < 0.01$) [Figure 1A and 1B] compared to the vendor recommended 50% DMSO (4100 +/- 2792 RFU/element) or 3X SSC (3578 +/- 2241 RFU/element) print buffers. The standard deviations are large since averages are derived from three different concentrations of the dilution series. Because microarray image analysis is dependent on the spot size and spot shape, array element morphology obtained with the various print buffers was also evaluated. Probes printed in 15% DMSO/1.5 M betaine print buffer consistently produced round spots of the desired size (150.6 +/- 19.6 total pixels; ~120 microns in diameter) [Figure 1A]. In contrast, probes deposited with 50% DMSO produced larger spots (188.9 +/- 23.1 total pixels) [Figure 1A], which frequently resulted in merging of adjacent array elements (data not shown). Probes deposited with 3X SSC produced the smallest spots of the three buffers evaluated (120.8 +/- 17.8 total pixels) and exhibited more variable spot morphology [Figure 1A].

Observation of competitive binding between the tracking and target-specific 70-mers

In determining a printing buffer that provided high probe retention and good spot morphology, we observed that spotting 0.5 μ M tracking oligonucleotide resulted in elements that were easy to visualize using our standard imaging protocol [Figure 1A]. Therefore we used this concentration to co-deposit tracking oligonucleotide with

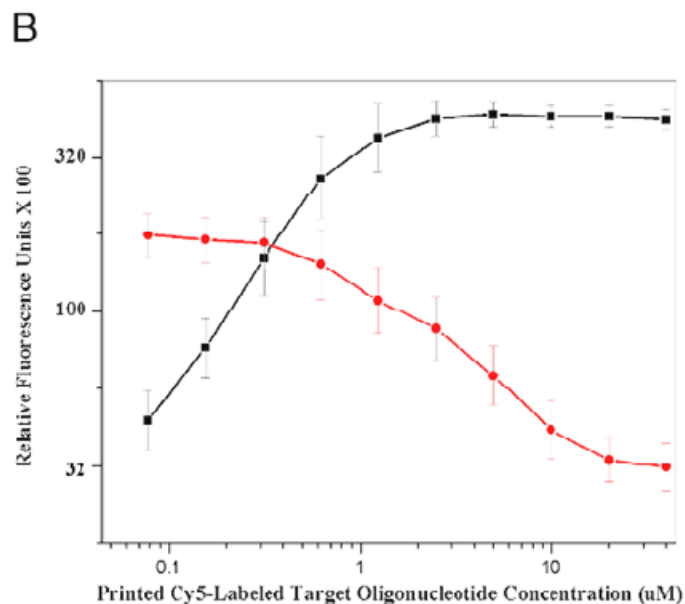
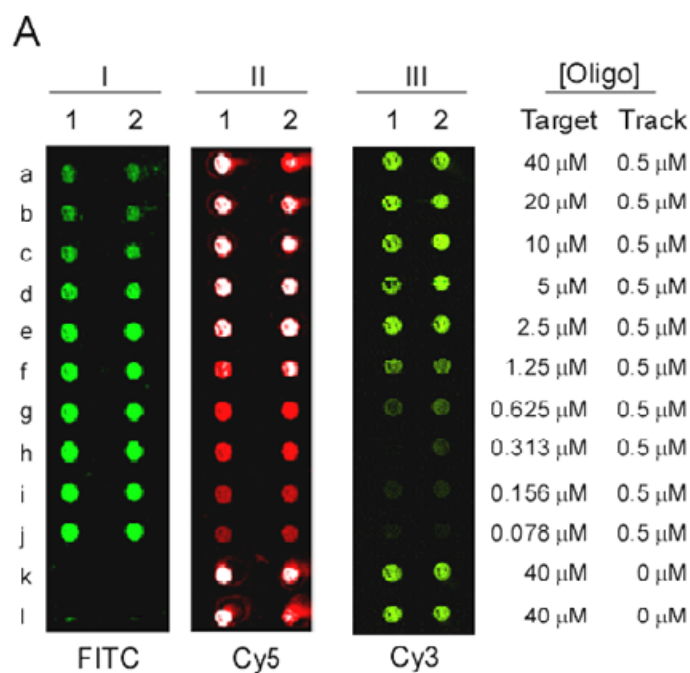


Figure 2

Determination of optimal *M. tuberculosis* target-specific oligonucleotide concentration. A: Serial dilutions of 5'-Cy5-labeled 70-mer oligonucleotides to *M. tuberculosis* *sodA* (column I) or *aceA* (column 2) were generated in 15% DMSO/1.5 M betaine in the presence or absence of 0.5 μM 5'-fluorescein-labeled tracking oligonucleotide and spotted onto poly L-lysine coated glass slides. Slides were processed, blocked, and scanned for fluorescein fluorescence (panel I) or Cy5 fluorescence (panel II). A subset of arrays was also hybridized with Cy3-labeled *sodA*- and *aceA*-specific PCR products and scanned for Cy3 fluorescence (panel III). No loss of hybridization signal is observed following addition of tracking oligonucleotide. B: Fluorescein (red) or Cy5 (black) mean relative fluorescence intensities from printed oligonucleotide were plotted as a function of Cy5-labeled target-specific oligonucleotide concentration. Note that the fluorescein intensity derived from the tracking oligonucleotide (red line) is informative over approximately two orders of magnitude of *M. tuberculosis*-specific 70-mer concentration. Bars represent S.D. Data were generated from 100 dilution series over 25 arrays.

40 μM *M. tuberculosis* target-specific oligonucleotides. To ensure that a 0.5 μM tracking oligonucleotide concentration would not interfere with detection of hybridization, even if the target specific oligonucleotide was at a concentration less than 40 μM , a set of arrays were printed in which the concentration of two 5'-Cy5-labeled *M. tuberculosis* target-specific 70-mers (*aceA* and *sodA*) were serially diluted from 40 μM to 0.078 μM while maintaining the tracking oligonucleotide concentration at 0.5 μM . When visualizing these mixed dilutions immediately after printing, we observed equal fluorescein spot intensity among all spots in the dilution series, which is consistent with the manufacturers specifications that the TeleChem SMP3 printing pin deposits approximately 6 nl of probe volume per contact with the glass surface (image not shown). However after post-processing, the fluorescein signal intensity derived from the tracking oligonucleotide was observed to be inversely related to the intensity of the Cy5 signal derived from the labeled target-specific oligonucleotide (Figure 2A, panels I and II), indicating that as molarity of the Cy-5 labeled target-specific oligonucleotide decreased from the recommended 40 μM (Figure 2A, panel II, rows a-j), the tracking oligonucleotide was able to more effectively compete for open sites on the glass surface, resulting in increased fluorescein fluorescence (Figure 2A, panel I, rows a-j). The observation of this relationship for 100 dilution series printed over 25 arrays is plotted in Figure 2B, where the fluorescein intensity derived from the tracking oligonucleotide (red line) is shown to be informative over approximately two orders of magnitude of *M. tuberculosis*-specific 70-mer concentration. Cy3-labeled PCR targets complementary to the immobilized Cy5-labeled *aceA* and *sodA* probes were hybridized to these arrays and no significant difference in Cy3 hybridization signal intensity between elements possessing no tracking oligonucleotide versus those possessing 0.5 μM tracking oligonucleotide at each of the 40 μM , 20 μM , or 10 μM target-specific oligonucleotide levels were detected [Figure 2A, panel III, compare rows a, b and c to k and l]. Thus, utilization of the 5'-fluorescein-labeled *S. aureus*-specific tracking oligonucleotide at 0.5 μM does not affect hybridization characteristics of *M. tuberculosis*-specific target oligonucleotides printed at the manufacturers recommended concentrations of 40 μM or 20 μM .

Introduction of tracking oligonucleotide does not impact Cy3/Cy5 ratio measurement

To demonstrate equivalency between data obtained from individual arrays possessing the tracking oligonucleotide and arrays lacking the tracking oligonucleotide, experiments were performed on replicate *M. tuberculosis* array sets in which tracking oligonucleotide was either excluded or included in the print buffer. A total of 100 arrays each containing 5376 total spots (4269 *M. tuberculosis* probe set plus 1107 additional positive and negative controls)

were initially printed at 40 μM in 15% DMSO/1.5 M betaine print buffer lacking tracking oligonucleotide. Immediately after printing, the oligonucleotide set was dried down, reconstituted with print buffer possessing 0.5 μM fluoresceinated tracking oligonucleotide, and reprinted on an additional 100 poly-L-lysine coated glass slides at a final concentration of 40 μM . A pre-hybridization fluorescein image of one sub array from a representative slide containing the tracking oligonucleotide is shown (Figure 3A; panel I). As can be observed, incorporation of tracking oligonucleotide into the print buffer allowed spotted probes to be efficiently visualized and utilized for quality control purposes. Array elements containing negative controls consisting of fluoresceinated tracking oligonucleotide alone (Figure 3A; panel I, black arrowheads) were visibly more fluorescent than those elements containing both tracking oligonucleotide and *M. tuberculosis* target-specific oligonucleotides (Figure 3A; panel I, remaining rows).

We investigated the impact of the tracking oligonucleotide on Cy3/Cy5 ratio measurements ($n = 12$ direct replicates, 6 arrays with and 6 arrays without the tracking oligonucleotide) using heterotypic hybridizations between wild-type *M. tuberculosis* H37Rv and the *mprA* mutant strain. The generation of sufficient labeled target for this experiment required multiple labeling reactions, and pooling all the labeled targets prior to distribution over the arrays normalized potential variation between these reactions. Visual examination of Figure 3A panels I and II indicate that the image data derived from arrays printed with or without the tracking oligonucleotide are equivalent and no cross-hybridization of target to the tracking oligonucleotide is observed. The resulting hybridized image data was analyzed with the Marray software, which employs a unique quality control algorithm that defines a quality score (q_{com}) for each spot [20,21]. Because it was our objective to investigate data variability inherent to the slide itself versus variability introduced through hybridization, we filtered, on average 15% of spots on each array that possessed low hybridized q_{com} . These spots possessed low scores mainly due to intensity levels lower than background. After filtering these spots (many of which were negative controls), 4852 spots common to all arrays were used to determine ratio correlation values. Linearity between the mean log Cy3/Cy5 ratio from replicate arrays possessing the tracking oligonucleotide ($n = 6$) versus those lacking the tracking oligonucleotide ($n = 6$) was observed (Figure 3B; $R^2 = 0.67$, $p < 0.001$). When the analysis was restricted to only those genes differentially expressed, the coefficient of determination became more significant, reaching $R^2 = 0.90$. Furthermore, no significant difference in Pearson's correlation coefficients of mean log Cy3/Cy5 ratio between replicates possessing the tracking

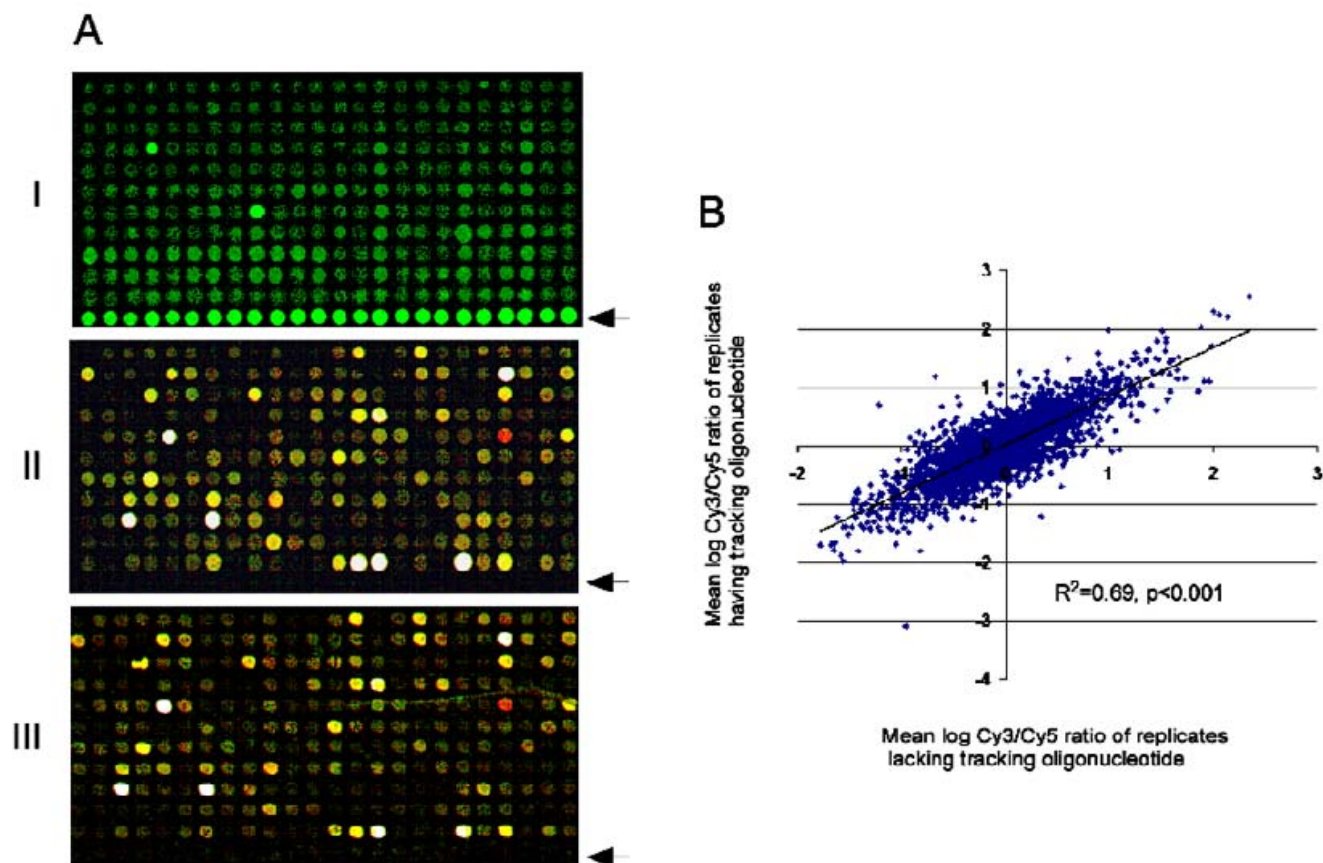


Figure 3

Comparison of *M. tuberculosis* 70-mer oligonucleotide array set printed in the presence or absence of 5'-fluorescein-labeled tracking oligonucleotide. A: Single sub array images of the *M. tuberculosis* 70-mer oligonucleotide array set printed in the presence (panels I, fluorescein; and II, Cy3/Cy5) or absence (panel III, Cy3/Cy5) of 5'-fluorescein-labeled tracking oligonucleotide. Black arrowheads, denote rows containing tracking oligonucleotide only, note high fluorescein intensity (Panel I) and lack of cross-hybridization with labeled target (Panel 2). Images shown in panel I and panel II are the same array before and after hybridization. B: Scatter plot of mean Cy3/Cy5 log intensity ratios of fluorescence from hybridized arrays printed in the absence (x-axis) or presence (y-axis) of 5'-fluorescein-labeled tracking oligonucleotide.

oligonucleotide (0.72 ± 0.07), between replicates lacking the tracking oligonucleotide (0.74 ± 0.10), and between replicates with and without the tracking oligonucleotide (0.70 ± 0.04) were observed. Finally, we identified a total of 131 genes that were significantly regulated between wild-type *M. tuberculosis* H37Rv and the isogenic *mprA* mutant at $p = 0.05$ and performed ANOVA to determine if individual genes, arrays, or inclusion of the tracking oligonucleotide have a significant effect on the gene expression measurements [22,23]. The results [Table 1] indicate that no significant impact on the ratio measurement is

introduced by the tracking oligonucleotide since no significant difference is observed between hybridizations including or excluding the tracking oligonucleotide (Table 1; row C-, $p = 0.2091$). Therefore, introduction of the third-color tracking oligonucleotide into the *M. tuberculosis* oligonucleotide set does not affect hybridization efficiency or introduce ratio measurement bias through cross-hybridization, fluorescein signal contamination of the cyanine dye channels, or quenching interaction with the cyanine dyes.

Table 1: Analysis of Variance (ANOVA) introduced through the addition of tracking oligonucleotide [22].

Row	Variance Source	Sum ²	Degrees of Freedom	Mean ²	F	Prob > F
A	Gene	89.745	130	0.69035	1.58	0.0002
B	Slide	5.189	5	1.03781	3.38	0.0372
C	Tracking Oligo	0.689	1	0.6886	1.58	0.2091
D	Gene*Slide	282.851	650	0.43516	1	0.5053
E	Gene*Tracking Oligo	57.254	130	0.44041	1.01	0.4558
F	Slide*Tracking Oligo	8.275	5	1.65504	3.8	0.0021
G	Error	283.146	650	0.43561		
H	Total	727.148	1571			

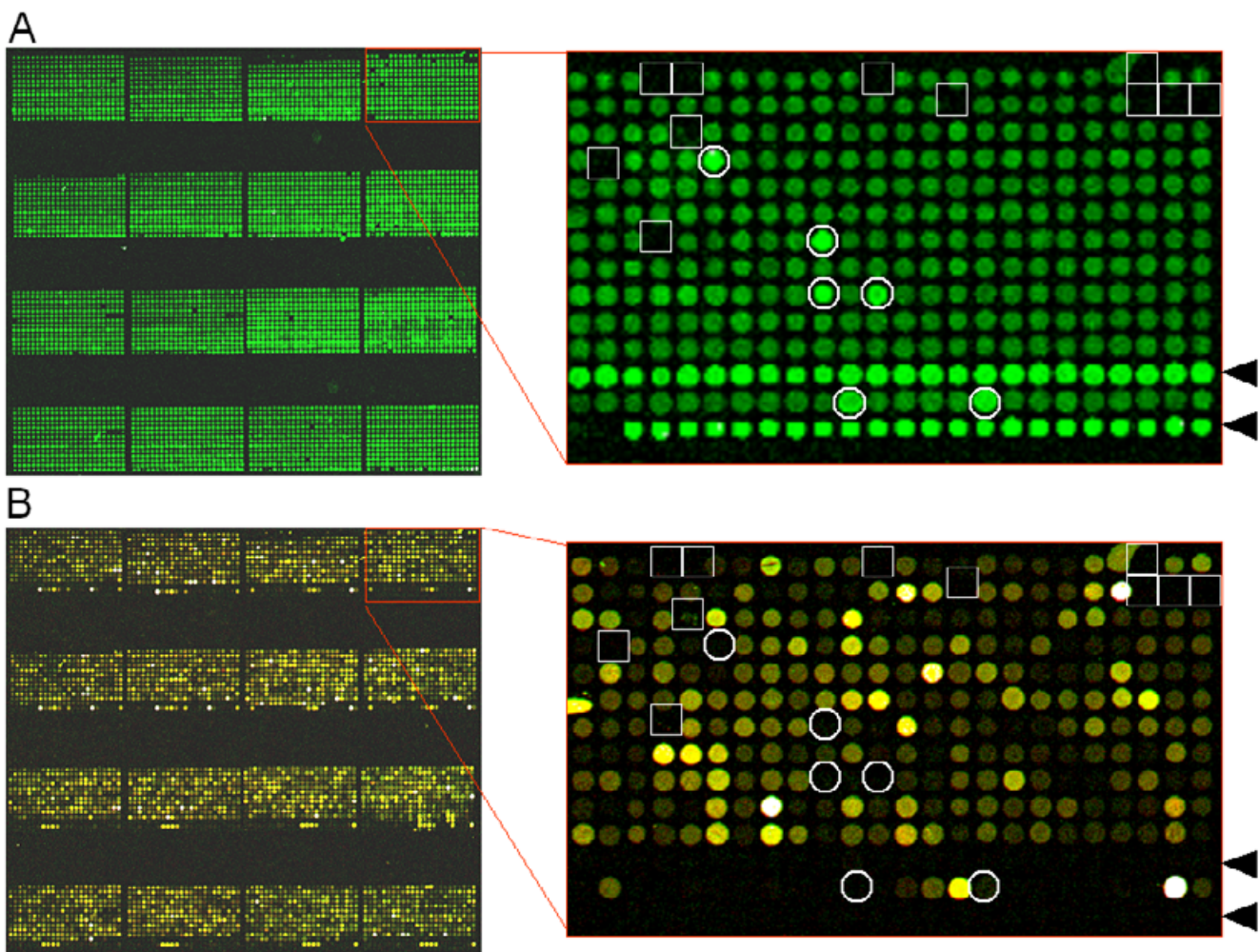


Figure 4
 Use of third-color tracking oligonucleotide for quality control and data filtering. Pre-hybridization fluorescein image (A) and post-hybridization Cy3/Cy5 image (B) of *M. tuberculosis* array possessing the tracking oligonucleotide. Open squares; denote elements where oligonucleotides were not deposited on slide surface. Open circles; denote elements where target-specific *M. tuberculosis* oligonucleotides were present in low concentration. Black arrowheads, negative controls containing tracking oligonucleotide only.

Quality control and data filtering using the tracking oligonucleotide

A major advantage to using a third-color tracking oligonucleotide is the ability to efficiently monitor printing fidelity on individual microarray slides in a given batch. To monitor printing fidelity in our *M. tuberculosis* oligonucleotide set containing the 5'-fluoresceinated tracking oligonucleotide, individual slides were analyzed for fluorescein fluorescence before, or Cy5 and Cy3 fluorescence after heterotypic hybridizations between wild-type *M. tuberculosis* H37Rv and the isogenic *mprA* mutant. Images from a representative microarray are depicted in Figure 4. As expected, rows containing negative controls consisting of tracking oligonucleotide alone were visibly more fluorescent on the pre-hybridization fluorescein image relative to spots containing tracking oligonucleotide co-spotted with target specific oligonucleotide [Figure 4; panel A, black arrowheads]. Interestingly, similar levels of fluorescein fluorescence were also observed from a subset of printed spots that contained both tracking- and target-specific oligonucleotides [Figure 4; panel A, compare open circles with black arrowheads]. This indicates that the concentration of target-specific oligonucleotide present in these spots were at a lower concentration than expected, perhaps due to synthesis failure. Consistent with this hypothesis, little Cy5 or Cy3 fluorescence could be observed from these spots following hybridization [Figure 4; panel B, open circles]. Furthermore, fluorescein, Cy3, or Cy5 fluorescence were not observed on pre-hybridization or post-hybridization images for some spotted oligonucleotides [Figure 4; panels A and B, open squares], suggesting that probe DNA for these genes was not efficiently deposited or immobilized on the glass surface. Thus, utilization of a third-color tracking oligonucleotide is useful for quality control measurements and differentiating false negatives (synthesis failures and mechanical misses) from true negatives (no gene expression).

This sets the stage for more sophisticated data filtering, however, the signal intensity derived from the tracking oligonucleotide will vary as a function of the binding capacity of the slide, which is influenced by many variables including slide coating chemistry, slide coating lot, printing conditions (temperature and humidity), as well as post-processing variables. Therefore, the binding capacity of the array must first be established using the tracking oligonucleotide alone, to identify the maximum fluorescein intensity for the slide. Quality slides with high retention capacities will have high intensities, whereas low quality/retention slides will not. Likewise, it is also necessary to establish for each slide the fluorescein intensity obtained when co-spotting 0.5 μM tracking oligonucleotide with target-specific oligonucleotide at its proper 40 μM concentration, which is accomplished with a dilution series of positive controls, ie quantified, high synthe-

sis quality oligonucleotides, in known molar ratio with tracking oligonucleotide printed on each array, such as those in Figure 2. The thousands target-specific elements of the array, which include a range of high to low synthesis yield as well as mechanical misses, can now be gauged against these positive and the negative controls.

We have previously shown for cDNA arrays that hybridized data quality is highly dependent on the amount of immobilized probe [19], and this variable can be controlled using directly labeled probe material and filtering those spots that do not meet defined intensity thresholds. Our current efforts include defining the lowest amount of support-bound target oligonucleotide necessary to obtain reliable ratio measurements, however there are a number of issues to consider. First, cDNA probes typically average about 1 kb in length, which is more than 10-fold the size of a 70-mer oligonucleotide. Therefore it is reasonable to assume that a given area of glass support will bind >10-fold the number of copies of a 70-mer oligonucleotide versus a 1 kb double-stranded PCR product and if ratio data quality is dictated by the number of support-bound copies of a probe sequence, it will be easier to achieve this threshold with spotted oligonucleotides. The second consideration is the signal dynamic range of the tracking oligonucleotide and whether or not data impacting differences in bound target-specific oligonucleotide (aside from synthesis failures) can be measured with this approach. We are currently investigating these issues with the objective of defining the threshold of support-bound probe necessary for reliable gene expression measurements so that these compromising elements can be filtered from data sets.

Variability of microarray data can arise from both biological and technical sources. In this report, use of a fluoresceinated tracking oligonucleotide for quality control purposes has been described. This three-color microarray methodology can be utilized for pre-hybridization quality assessment [18], to facilitate automated post-hybridization image analysis [24], as well as post-hybridization data filtering [19]. The ability to perform QC filtering on printed slides allows identification of potentially problematic slides containing low amounts of probe available for hybridization or high variation in spot morphology/deposition across the array.

Methods

Mycobacterium tuberculosis 70-mer oligonucleotide set

The Operon (Alameda, CA) 4,269 *Mycobacterium tuberculosis* 70-mer oligonucleotide probe set was printed on in-house poly-L-lysine coated slides prepared as previously described [25]. Oligonucleotide probes present in this collection were designed from the genomic sequence of *M. tuberculosis* strain H37Rv and *M. tuberculosis* strain

CDC1551, and include an additional 25 negative and 12 positive oligonucleotide controls. Lyophilized probes were shipped in 384-well micro titer plates. Upon arrival, probes were resuspended to a final concentration of 80 μ M in deionized water and resuspended by incubation overnight at 4 °C with gentle shaking. Probes were diluted 2-fold into one of three 2 \times printing buffers: 6x SSC (saline sodium citrate), 100% DMSO (dimethyl sulfoxide), or 30% DMSO containing 3.0 M betaine, thereby generating oligonucleotide sets at a final concentration of 40 μ M. When required, a *Staphylococcus aureus*-specific 70-mer "tracking" oligonucleotide synthesized with a 5'-fluorescein tag was added to the printing buffer at a final molar ratio of 1:80 relative to the *M. tuberculosis*-specific oligonucleotides.

Array Fabrication

Printing was performed with a GeneMachines Omni Grid printer (San Carlos, CA) using Telechem International SMP3 pins (Sunnyvale, CA) at 40% humidity and 22 °C. To control pin contact force and duration, the instrument was set with the following Z motion parameters, velocity: 7 cm/sec, acceleration: 100 cm/sec², deceleration: 100 cm/sec². Following printing, oligonucleotides were UV cross-linked to slides and processed using the previously described nonaqueous blocking protocol [26]. Image files on all arrays were collected before hybridization using excitation and emission spectra specific for fluorescein using a ScanArray 5000 (GSI Lumonics, Billerica, MA).

Generation of third-color *Staphylococcus aureus*-specific 70-mer oligonucleotide and 5'-Cy5-labeled *M. tuberculosis*-specific 70-mer oligonucleotides

A 5'-fluoresceinated 70-mer oligonucleotide [5'-(FITC)-ATGAAGAACTATATACATCTTATGGCACTTATGGATTTTACATCAGATAAAAATCAATAAC-CCGACCC-3'] was generated to the *S. aureus* response regulator *trap* and used as a third-color tracking oligonucleotide in the *M. tuberculosis* 70-mer array. This gene has been shown to be specific to *S. aureus* [27], and exhibits no homology to any gene in the *M. tuberculosis* genome (data not shown). For some experiments, two 5'-Cy5 labeled 70-mer oligonucleotides (Operon) were generated to the *aceA* (*Rv0467*) [5'-Cy5-GGATCAACAACGCACTGCAGCGCGCCGACCA-GATCGCCAAGATCGAGGGCGATACTTCGGTGGA-GAACTG-3'] and *sodA* (*Rv3846*) [5'-Cy5-CAAGCTGCTGATATTCCAGGTTTACGACCACCAGACGAACCTC-CCGCTAGGCATTGTTCCGCTGCTGCG-3'] genes of *M. tuberculosis* H37Rv. These genes are expressed during growth of *M. tuberculosis* in laboratory medium *in vitro* [28,29].

Growth of *Mycobacterium tuberculosis* H37Rv and preparation of DNA or RNA

M. tuberculosis H37Rv and the isogenic *mprA* mutant [30] were grown at 37 °C to an $A_{600} \sim 1.2$ in Middlebrook 7H9 broth (Difco, Sparks, MD) supplemented with 10% (vol/vol) albumin-dextrose-catalase (Difco, Sparks, MD) and 0.05% (vol/vol) Tween 80 (Sigma, St. Louis, MO). DNA was prepared from *M. tuberculosis* as previously described [31]. For preparation of RNA, bacterial cells from a 50.0 ml culture were transferred to a conical tube and concentrated by centrifugation. Cells were resuspended in Tel-Test RNA-Bee (Friendswood, TX) and mechanically lysed using a bead beater (BioSpec Products, Bartlesville, OK) as described [32]. Total RNA was chloroform extracted and precipitated by the addition of isopropanol. The resulting RNA pellet was resuspended in DEPC-treated water and treated with DNase I (Ambion Inc., Austin, TX) to remove contaminating DNA, and re-purified by affinity chromatography using RNeasy purification columns (Qiagen, Valencia, CA). To reduce ribosomal RNA content, 10 μ g of DNase I-treated total RNA was enriched for mRNA using the MICROExpress™ bacterial mRNA purification kit (Ambion).

Labeling and hybridization of *M. tuberculosis* RNA or DNA

Enriched mRNA or PCR-amplified DNA was used as templates for labeling reactions microarray hybridization experiments. Test hybridizations utilized *aceA* (*Rv0467*) and *sodA* (*Rv3846*) Cy 3-labeled PCR products amplified from genomic *M. tuberculosis* H37Rv DNA using primers [*aceAF* (5'-GATCCAGCAGGAATGGGACAC-3') and *aceAR* (5'-CAGACTAGTGGAACTGGCCCTCT-3')] and primers [*sodAF* (5'-GCCAGACTGGACTGGGACTA-3') and *sodAR* (5'-CCGAATATCAACCCCTTGGTCT-3')]. For heterotypic hybridizations, labeled cDNAs were synthesized using SuperScript™ III Reverse Transcriptase (Invitrogen, Carlsbad, CA), a cocktail of synthetically generated decamers [33], and Cy3- or Cy5-dCTP (Amersham, Piscataway, NJ). Cy3- and Cy5-labeled templates were purified using a Qiagen MinElute kit, and concentrated using Amicon YM 30 columns. The concentrated mixed probes were added to a solution containing 50% formamide, 0.1% SDS, and 5X SSC. Hybridization of *M. tuberculosis* array slides were performed in sealed humid hybridization cassettes for 16–20 h at 42 °C. All slides were pre-hybridized with 0.1% BSA solution in 5X SSC and 0.1% SDS for 30 min at 42 °C. Following hybridization, array slides were washed at room temperature for 1 minute each in 2X SSC/0.1% SDS, 1X SSC, 0.2X SSC, and 0.05X SSC and then dried by centrifugation. Following hybridization reactions, arrays were scanned using excitation and emission spectra specific for Cy3 and Cy5. Scanning was performed using a ScanArray 5000 (GSI Lumonics, Billerica, MA).

Data analyses and statistics

Microarray images were processed with Matarray, which has been extended to process three images simultaneously, in particular utilizing the advantages of the pre-hybridization FITC image to assist signal segmentation and quantification [20,24]. Data filtering and normalization were performed utilizing the quality score q_{com} defined in Matarray [21]. Statistical analysis was performed in Matlab using its statistical toolbox and modules developed in house.

Authors' Contributions

MJH, VKS, and TCZ conceived of the study, and participated in its design and coordination, and drafted the manuscript. VKS, SK, MRS carried out the array fabrication and gene expression studies. XW performed the statistical analysis. All authors read and approved the final manuscript.

Acknowledgments

M.J.H is supported by a National Institute of Biomedical Imaging and Bioengineering Grant (EB001421) and by a special fund from the Children's Hospital of Wisconsin Foundation. T.C.Z. is supported by the National Institutes of Health (AI51669). We thank the Clement J. Zablocki VA Medical Center (Milwaukee, WI) for utilization of their BSL-3 bio-containment laboratory and Shuang Jia, M.S., for data preparation.

References

- Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL: **Expression monitoring by hybridization to high-density oligonucleotide arrays.** *Nat Biotechnol* 1996, **14**:1675-1680.
- Chee M, Yang R, Hubbell E, Berno A, Huang XC, Stern D, Winkler J, Lockhart DJ, Morris MS, Fodor SP: **Accessing genetic information with high-density DNA arrays.** *Science* 1996, **274**:610-614.
- Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with complementary DNA microarray.** *Science* 1995, **270**:467-470.
- Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurchi K, Pienta KJ, Rubin MA, Chinnaiyan AM: **Delineation of prognostic biomarkers in prostate cancer.** *Nature* 2001, **412**:822-826.
- Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, van de Rijn M, Rosen GD, Perou CM, Whyte RI, Altman RB, Brown PO, Botstein D, Petersen I: **Diversity of gene expression in adenocarcinoma of the lung.** *Proc Natl Acad Sci U S A* 2001, **98**:13784-13789.
- Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilfond B, Borg A, Trent J: **Gene-expression profiles in hereditary breast cancer.** *N Engl J Med* 2001, **344**:539-548.
- Trent JM, Stanbridge EJ, McBride HL, Meese EU, Casey G, Araujo DE, Witkowski CM, Nagle RB: **Tumorigenicity in human melanoma cell lines controlled by introduction of human chromosome 6.** *Science* 1990, **247**:568-571.
- Southern EM, Maskos U, Elder JK: **Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models.** *Genomics* 1992, **13**:1008-1017.
- Sorlie T, et. al: **Gene expression patterns of breast carcinomas distinguish tumor subclass with clinical implications.** *Proc Natl Acad Sci* 2001, **98**:10869-10874.
- Hegde P, Qi R, Gaspard R, Abernathy K, Dharap S, Earle-Hughes J, Gay C, Nwokekeh NU, Chen T, Saeed AI, Sharov V, Lee NH, Yeatman TJ, Quackenbush J: **Identification of tumor markers in models of human colorectal cancer using a 19,200-element complementary DNA microarray.** *Cancer Res* 2001, **61**:7792-7797.
- Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D: **Light-directed, spatially addressable parallel chemical synthesis.** *Science* 1991, **251**:767-773.
- Fodor SP, Rava RP, Huang XC, Pease AC, Holmes CP, Adams CL: **Multiplexed biochemical assays with biological chips.** *Nature* 1993, **364**:555-556.
- Pritchard CC, Hsu L, Delrow J, Nelson PS: **Project normal: defining normal variance in mouse gene expression.** *Proc Natl Acad Sci U S A* 2001, **98**:13266-13271.
- Ramakrishnan R, Dorris D, Lublinsky A, Nguyen A, Domanus M, Prokhorova A, Gieser L, Touma E, Lockner R, Tata M, Zhu X, Patterson M, Shippy R, Sendera TJ, Mazumder A: **An assessment of Motorola CodeLink microarray performance for gene expression profiling applications.** *Nucleic Acids Res* 2002, **30**:e30.
- Battaglia C, Salani G, Consolandi C, Bernardi LR, De Bellis G: **Analysis of DNA microarrays by non-destructive fluorescent staining using SYBR green II.** *Biotechniques* 2000, **29**:78-81.
- Yue H, Eastman PS, Wang BB, Minor J, Doctolero MH, Nuttall RL, Stack R, Becker JW, Montgomery JR, Vainer M, Johnston R: **An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression.** *Nucleic Acids Res* 2001, **29**:E41-1.
- Hessner M.J., K. Dowling, O. Kokanovic, L. Meyer, S.H. Nye, X. Wang, J. Waukau, S. Ghosh: **Use of fluorescein-labeled probes as a quality control tool for cDNA microarrays.** *The American Journal of Human Genetics* 2001, **Supplement**, **69**:468.
- Hessner MJ, Wang X, Hulse K, Meyer L, Wu Y, Nye S, Guo SW, Ghosh S: **Three color cDNA microarrays: quantitative assessment through the use of fluorescein-labeled probes.** *Nucleic Acids Res* 2003, **31**:e14.
- Hessner MJ, Wang X, Khan S, Meyer L, Schlicht M, Tackes J, Datta MW, Jacob HJ, Ghosh S: **Use of a three-color cDNA microarray platform to measure and control support-bound probe for improved data quality and reproducibility.** *Nucleic Acids Res* 2003, **31**:e60.
- Wang X, Hessner MJ, Wu Y, Pati N, Ghosh S: **Quantitative quality control in microarray experiments and the application in data filtering, normalization and false positive rate prediction.** *Bioinformatics* 2003, **19**:1341-1347.
- Wang X, Ghosh S, Guo S-W: **Quantitative quality control in microarray image processing and data acquisition.** *Nucleic Acids Research* 2001, **29**:E75-82.
- Zar JH: **Biostatistical Analysis.** 4thth edition. Upper Saddle River, New Jersey, Prentice-Hall; 1999.
- Pati N., S. Ghosh, M.J. Hessner, H. Khoo, X. Wang: **Difference in gene expression profiles between human CD4+CD25+ and CD4+CD25- T cells.** *Annals of the New York Academy of Sciences* 2003, **1005**:279-283.
- Wang X., N. Jiang, X. Feng, Y. Xie, P. Tonellato, S. Ghosh, M.J. Hessner: **A novel approach for high quality microarray processing using third-dye array visualization technology.** *IEEE Transactions on Nanoscience* 2003 in press.
- Eisen MB, Brown PO: **DNA arrays for analysis of gene expression.** *Methods in Enzymology* 1999, **303**:179-205.
- Diehl F, Grahlmann S, Beier M, Hoheisel JD: **Manufacturing DNA microarrays of high spot homogeneity and reduced background signal.** *Nucleic Acids Research* 2001, **29**:e38.
- Balaban N, Goldkorn T, Gov Y, Hirschberg M, Koyfman N, Matthews HR, Nhan RT, Singh B, Uziel O: **Regulation of Staphylococcus aureus pathogenesis via target of RNAIII-activating Protein (TRAP).** *J Biol Chem* 2001, **276**:2658-2667.
- Graham JE, Clark-Curtiss JE: **Identification of Mycobacterium tuberculosis RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS).** *Proc Natl Acad Sci U S A* 1999, **96**:11554-11559.
- Tullius MV, Harth G, Horwitz MA: **High extracellular levels of Mycobacterium tuberculosis glutamine synthetase and superoxide dismutase in actively growing cultures are due to high expression and extracellular stability rather than to a protein-specific export mechanism.** *Infect Immun* 2001, **69**:6348-6363.

30. Zahrt TC, Deretic V: **Mycobacterium tuberculosis signal transduction system required for persistent infections.** *Proc Natl Acad Sci U S A* 2001, **98**:12706-12711.
31. Zahrt TC, Deretic V: **An essential two-component signal transduction system in Mycobacterium tuberculosis.** *J Bacteriol* 2000, **182**:3832-3838.
32. Talaat AM, Howard ST, Hale W. th, Lyons R, Garner H, Johnston SA: **Genomic DNA standards for gene expression profiling in Mycobacterium tuberculosis.** *Nucleic Acids Res* 2002, **30**:e104.
33. Fisher MA, Plikaytis BB, Shinnick TM: **Microarray analysis of the Mycobacterium tuberculosis transcriptional response to the acidic conditions found in phagosomes.** *J Bacteriol* 2002, **184**:4025-4032.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

