research articles Journal of proteome research

Robust-Linear-Model Normalization To Reduce Technical Variability in Functional Protein Microarrays

Andrea Sboner,[†] Alexander Karpikov,[†] Gengxin Chen,[‡] Michael Smith,[‡] Mattoon Dawn,[⊥] Lisa Freeman-Cook,[‡] Barry Schweitzer,[¶] and Mark B. Gerstein*[,†,§,||]

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, Invitrogen, part of Life Technologies, 5791 Van Allen Way, Carlsbad, California 92008, Invitrogen, part of Life Technologies, 29851 Willow Creek Road, Eugene, Oregon 97402, Invitrogen, part of Life Technologies, 5781 Van Allen Way, Carlsbad, California 92008, Program in Computational Biology and Biochemistry, Yale University, New Haven, Connecticut 06520, and Department of Computer Science, Yale University, New Haven, Connecticut 06520

Received May 8, 2009

Protein microarrays are similar to DNA microarrays; both enabling the parallel interrogation of thousands of probes immobilized on a surface. Consequently, they have benefited from technologies previously developed for DNA microarrays. However, assumptions for the analysis of DNA microarrays do not always translate to protein arrays, especially in the case of normalization. Hence, we have developed an experimental and computational framework to assess normalization procedures for protein microarrays. Specifically, we profiled two sera with markedly different autoantibody compositions. To analyze intra- and interarray variability, we compared a set of control proteins across subarrays and the corresponding spots across multiple arrays, respectively. To estimate the degree to which the normalization could help reveal true biological separability, we tested the difference in the signal between the sera relative to the variability within replicates. Next, by mixing the sera in different proportions (titrations), we correlated the reactivity of proteins with serum concentration. Finally, we analyzed the effect of normalization procedures on the list of reactive proteins. We compared global and quantile normalization, techniques that have traditionally been employed for DNA microarrays, with a novel normalization approach based on a robust linear model (RLM) making explicit use of control proteins. We show that RLM normalization is able to reduce both intra- and interarray technical variability while maintaining biological differences. Moreover, in titration experiments, RLM normalization enhances the correlation of protein signals with serum concentration. Conversely, while quantile and global normalization can reduce interarray technical variability, neither is as effective as RLM normalization in maintaining biological differences. Most importantly, both introduce artifacts that distort the signals and affect the correct identification of reactive proteins, impairing their use for biomarker discovery. Hence, we show RLM normalization is better suited to protein arrays than approaches used for DNA microarrays.

Keywords: protein arrays • global normalization • quantile normalization • robust linear model normalization • intra-array variability • interarray variability

Introduction

In the past decade, there has been a research paradigm shift from a traditional hypothesis-driven approach to a discovery-driven approach employing high-throughput (HT) methodologies. These new HT methods enable the collection of large data sets, greatly accelerating the progress of many investigations. Among the different HT technologies, microarrays are one of the first and most widely used methods. Microarrays allow the simultaneous query of thousands or millions of probes (nucleic acids, proteins, etc.), which are spatially arranged in a grid-like pattern. DNA microarrays are the most common type of microarray and typically consist of a series of DNA fragments representing genes, deposited in a grid on a solid-state surface. A fluorescently labeled sample, for example, the transcriptome of a cell, can be used to assess the expression profile of the cell. By measuring the fluorescence intensity of each probe bound to its cognate gene probe on the array, an estimate of the expression for each gene is obtained, and multiple samples

* Corresponding Author: Mark B. Gerstein, Yale University, Bass 432, 266 Whitney Avenue, New Haven, CT 06520 USA. Tel: +1 203 432 8189. Fax: +1 203 432 6946. E-mail: asbmg@gersteinlab.org.
[†] Department of Molecular Biophysics and Biochemistry, Yale University.
[‡] 5791 Van Allen Way, Carlsbad, CA 92008.
[⊥] 29851 Willow Creek Road, Eugene, OR 97402.
[¶] 5781 Van Allen Way, Carlsbad, CA 92008.
[§] Program in Computational Biology and Biochemistry, Yale University.
[||] Department of Computer Science, Yale University.

10.1021/pr900412k CCC: $40.75

Published on Web 10/09/2009

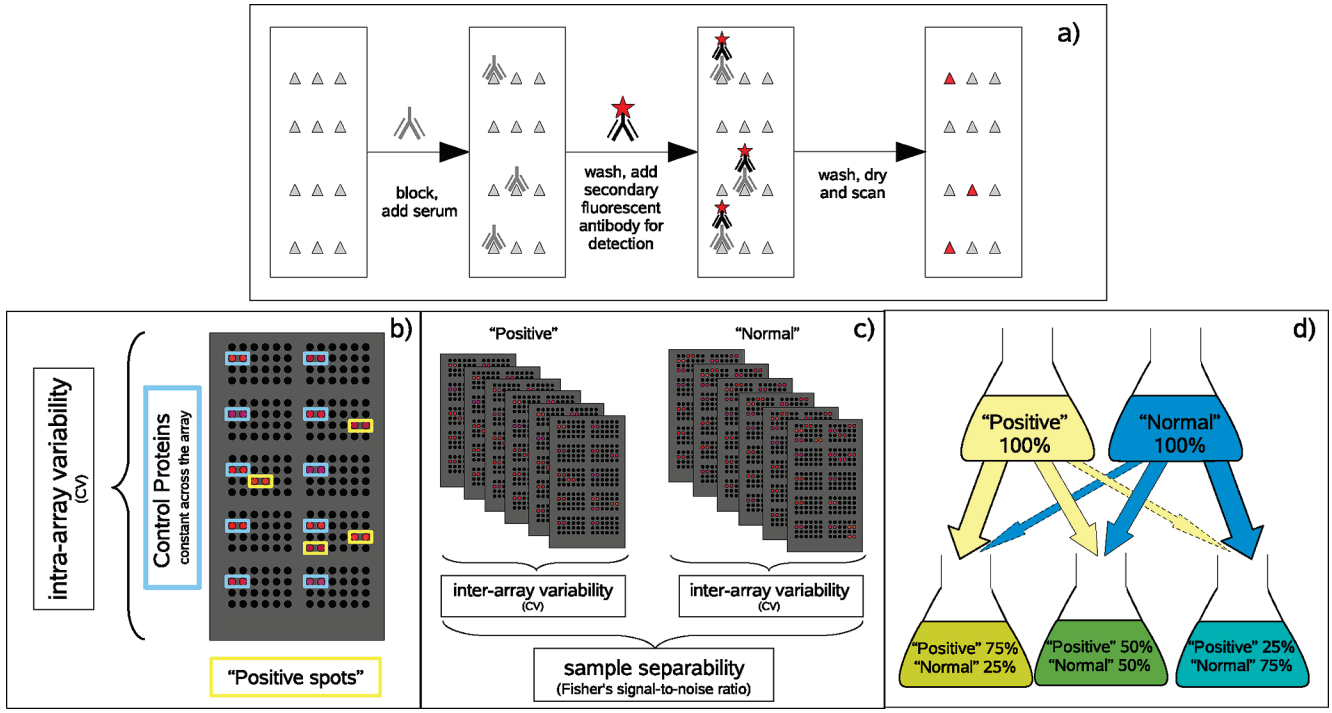


Figure 1. Serum profiling assay and experimental design. (a) Schematic of the assay. The array is probed with the serum sample. A second fluorescently labeled antibody is added after washing the array to remove unbound antibody. The amount of bound serum antibody is measured with the fluorescent signal intensity. (b) The array contains a series of subarrays. Control proteins are spotted in each subarray, enabling the estimation of the intra-array variability. They can also be used for the normalization procedure. The "positive spots" are proteins which reacted with antibodies in the serum or the secondary antibody. (c) Two different sera were utilized, a "positive" serum—Positive Control Sera (PCS)—and a "normal" serum, known to have distinct proteomic profiles. The "positive" serum is a mixture of sera qualified for reactivity to different proteins, whereas the "normal" serum is a sample with no known pathology. Six technical replicate arrays were probed for each serum sample. Interarray variability can be estimated for each serum. In addition, the two sets of arrays enable the estimation of sample separability, i.e., how well differentiated the proteomic profile of the two sera is. (d) The two sera were mixed in 5 different proportions, including pure samples, and each was assayed 4 times.

can be profiled to assess which genes are differentially expressed under different conditions.

Following a HT methodology, other types of microarrays have been developed, including Single Nucleotide Polymorphism (SNP) arrays,[1] tiling arrays for Comparative Genomic Hybridization (array-CGH),[2] and for Chromatin Immunoprecipitation (ChIP-on-chip).[3,4] Although targeted to different applications, they require the same basic components: probes placed on a grid, labeled samples, and a system to detect the fluorescent signal.

Recently, protein arrays have been introduced to further functional genomics studies.[5−8] Protein arrays are composed of hundreds or even thousands of proteins immobilized on a solid surface. Protein arrays can be used for a wide variety of applications, including detection of antibodies in blood, protein−protein interactions, protein−small molecule interactions, and enzyme activities.[9−11] Two types of protein arrays can be defined: functional protein arrays and protein profiling arrays.[12,13] The latter usually comprise spotted antibodies and are typically used to measure protein abundance and/or alterations. Functional protein arrays can include any type of protein and therefore have broader applicability. Here, we focus on the analysis of data from biomarker discovery efforts utilizing functional protein arrays. The premise of these experiments is that autoantibodies present in serum are correlated with diseases such as autoimmune disorders, infectious disease, and cancer. This is well documented for autoimmune diseases where high levels of antibodies are produced by the immune system against autoantigens, and it has also been shown that anomalously high levels of cancer-related proteins lead to the production of the corresponding autoantibodies.[14−16]

The immune response biomarker profiling assay is performed by probing protein arrays with serum samples in order to detect antibodies present in the serum that bind to their corresponding target protein on the array (see Figure 1a).

Bioinformatics is essential for analysis of the large data sets arising from these experiments. Several software tools have been developed for the detection and analysis of the hybridization signal on microarrays. The most common software tools for signal detection are GenePix Pro,[17] ImaGene,[18,19] ArrayVision,[20] ScanAlyze,[21] and SpotFinder.[22] Many more computational tools have been developed for other aspects of data analysis, including preprocessing, quality controls, normalization procedures, and actual statistical analysis of differential expression (e.g., SAM, Bioconductor, cyberT).[23−25] The majority of the hardware and software for the analysis of protein arrays has been adapted from DNA microarray tools. Although this has been advantageous in the early stage of protein arrays, more tailored tools are needed to tackle the specific issues of protein arrays.

We reasoned that many sources of variability can be introduced in the printing, probing, and scanning of protein arrays, resulting in differences in the measured signals that are not caused by differences in biological samples. This technical variability can be greater than the actual biological variability and thus hide the real differences between samples (false

negatives) or can introduce spurious effects leading to the identification of false biomarkers (false positives).

An effective normalization procedure should reduce the variability in the signal caused by systematic artifacts without losing useful biological information. One way to measure systematic variation is to use control proteins spotted on the arrays, which we assume to react independently from the serum and therefore return constant signals across and within arrays. For this reason, control proteins can be employed both for normalization and for evaluation of array variability.

We used a set of specific replicated experiments (Figure 1b,c) and titration experiments (Figure 1d) to examine the technical variability of protein arrays and demonstrated that traditional normalization procedures employed in DNA microarray analysis may not be suitable for protein arrays. Briefly, we employed two sera known to have different reactivity against specific proteins on the array. A set of replicated experiments was carried out to assess intra- and interarray variability by measuring the coefficient of variation of the control proteins within one array or the spots at corresponding positions across multiple arrays (see Materials and Methods for details). Moreover, the distinct reactivity of the two serum samples allowed us to evaluate sample separability by means of Fisher's signal-to-noise ratio and to analyze the correlation of the signal with the serum concentration within the context of the titration experiments. Indeed, signals of proteins reacting to a particular serum sample should show a positive correlation with increasing concentrations of the corresponding reactive serum. Three normalization procedures were employed. Global and quantile normalization are commonly used with DNA microarrays. The third method is a novel normalization approach based on the robust linear model specifically developed for functional protein arrays.

Study Rationale

Sources of Variability. Sources of variability in protein microarrays can be either systematic or random. For example, systematic spatial artifacts can be introduced by the printing process of the array, and the scanning process can introduce systematic noise if the scanner is not properly aligned. Heterogeneity of the array surface can introduce random variability. In addition, differences in the total quantity of serum or of the secondary antibody can affect the comparison between different arrays. Other types of artifacts can be introduced by the experimental procedure itself. Indeed, a common procedure employed by researchers is the adjustment of the Photo-Multiplier Tube (PMT) gain which is responsible for the acquisition and amplification of the fluorescent signal.[26] This procedure is typically used to avoid signal saturation or to enhance lower signals, but it can be a significant source of variability if arrays acquired at different PMT settings are compared.[27,28]

Experimental and Computational Approach. To assess those diverse sources of noise affecting protein binding signals, a set of replicated controlled experiments was devised. Protein arrays with selected control proteins spotted in each subarray were probed with each of two known different sera to examine intra- and interarray variability as well as the separability of the two sera. A set of titration experiments were also conducted to analyze the correlation of protein signals with serum concentration. Different normalization techniques were then evaluated on these data.

Array variability is estimated by means of coefficient of variations (CVs, see Intra- and Interarray Variability) computed considering either the replicated experiments (for interarray variability) or duplicated control protein spots (for intra-array variability). This measure provides an estimate of the level of variability in the signal. The lower the CV, the lower signal variation and, typically, the better the quality of the measurements. Controls proteins are expected to have the lowest CVs since they should be independent from the sera. Hence, they are ideal to estimate intra- and interarray variability due to technical reasons and assess the performance of different normalization procedures.

Although advisable, a low CV does not always translate to good quality array. Indeed, one could develop a normalization technique which sets each spot on the arrays to a fixed value. In this rather extreme example, CV would be zero, but obviously, all relevant information is completely lost. Hence, in addition to CVs, we also assessed sample separability. Since we have replicated experiments for each serum, which are quite distinct, we can quantify how well the two groups of serum are separated by means of Fisher's Signal-to-Noise ratio S (see Sample Separability Score). Ideally, the normalization procedure should lower the variability due to technical artifacts (lower CVs), while enhancing the separability of the samples (higher S).

Finally, with the set of titration experiments, the two sera were mixed in different concentrations and repeatedly tested. The two sera are sufficiently distinct that they react to different proteins. The intensity of the signals is expected to correlate with the serum concentration. Therefore, since the relationship of signals between titrated samples is known, we can evaluate the normalization procedures. For example, the signal of proteins known to react with the "positive" serum should yield a positive correlation with increasing concentration of the "positive" serum.

The aim of high-density protein microarrays is the identification of candidate proteins that are differentially reactive between two conditions, such as normal and disease sera. Therefore, an optimal normalization procedure should result in a reduction of both false negatives (undetected proteins) and false positives (proteins incorrectly identified as differentially reactive) with respect to the original unnormalized data. To assess this point, we first identified differentially reactive proteins, that is, created a hit list, by means of M-statistic (see Hit list section below) and then compared the hit lists of the three normalization procedures with the one obtained from the original data looking for the less disruptive normalization procedure.

Materials and Methods

Human Protein Collection. Human clones were obtained from Invitrogen's Ultimate ORF (open reading frame) collection or from a Gateway collection of kinase clones developed by Protometrix. The nucleotide sequence of each clone was verified by full length sequencing. All clones were transferred into a system for expressing recombinant proteins in insect cells via baculovirus infection. Using a proprietary high-throughput insect cell expression system, thousands of recombinant human proteins were produced in parallel. Each protein is tagged with Glutathione-S-Transferase (GST), which enables high-throughput affinity purification under conditions that retain activity. After purification, a sample of every purified protein is checked to ensure that the protein is present at the predicted molecular

weight. In addition, all proteins are printed onto arrays and the concentration of each protein is determined. The printed concentration of proteins on the array ranges from 5 ng/μL up to over 1000 ng/μL with an average of approximately 50 ng/μL.

Human Protein Microarray Manufacture. The output of the protein purification process described above is thousands of purified proteins that are ready to be printed on arrays. A contact-type printer equipped with 48 matched quill-type pins is used to deposit each of these proteins along with a set of control proteins in duplicate spots on 1 in. $\times$ 3 in. glass slides that have been coated with a thin layer of nitrocellulose. APiX slides are manufactured by GenTel BioSciences, Inc. ApiX technology is covered by US Patent #6,861,251. APiX and GenTel are registered trademarks of GenTel BioSciences, Inc. The printing of these arrays is carried out in a cold room under dust-free conditions in order to preserve the integrity of both protein samples and printed microarrays. The arrays are designed to accommodate more than 19 200 spots. Protein samples are printed in 130 μm spots arrayed in 48 subarrays (4000-μm^2 each) and are equally spaced in vertical and horizontal directions with 22 columns and 22 rows per subarray (Figure 1). Spots are printed with a 200 μm spot-to-spot spacing. Control proteins including human-IgG, anti-human-IgG and V5-CMK1 series (printed in an 8 step gradient: 0.125, 0.25, 0.5, 1, 2, 4, 8, and 16 μg/mL) are replicated on each subarray at defined positions and are used as reference controls in our normalization method.

Immune Response Biomarker Profiling Protocol. Microarray slides were blocked in 5 mL of PBS/1% BSA/0.1% Tween 20 in quadriPERM 4-well trays for 1 h at 4 °C with gentle agitation. After blocking, the blocking solution was removed from the arrays by vacuum aspiration and 5 mL of each serum sample diluted 1:500 in freshly prepared Serum Probing Buffer (1$\times$ PBS, 5 mM $MgCl_2$, 0.5 mM DTT, 0.05% Triton X-100, 5% glycerol, 1% BSA) or buffer alone (negative control) was applied to the arrays. Samples were allowed to incubate for 90 min at 4 °C with gentle agitation. After incubation, the samples were removed by vacuum aspiration and arrays were washed five times (5 min per wash) in 5 mL of Probing Buffer. Alexa Fluor647-conjugated anti-human IgG at 1.0 μg/mL diluted in a total volume of 5 mL of Probing Buffer was then added to each array and allowed to incubate with gentle shaking at 4 °C for 90 min. Alexa Fluor647- and Alexa Fluor555-conjugated anti-V5 antibody (0.26 μg/mL each) were added with the anti-human IgG as a serum-independent control for use in the robust linear model normalization algorithm (see Normalization Methods). After incubation, the secondary antibody was removed, and arrays were washed as described above. Arrays were dried by spinning in a table top centrifuge equipped with a plate rotor at 200*g* for 1 min.

Design of Experiments. All experiments were performed with two serum samples:

- "Normal" sample: a serum sample acquired by Bioreclamation (Hicksville, NY) with no known pathology.
- "Positive" sample: a Positive Control Sera (PCS, Immuno-Vision, Cat # HAP-0600); an equal mixture of six independent human sera, each qualified for a different antigen, Ro/SS-A, La/SS-B, Smith Antigen, RNP Complex, DNA Topoisomerase I (Scl-70), and Jo-1.

Two sets of experiments were performed to evaluate the variations in the data and the normalization effectiveness:

1. Two-sample replicates: six arrays were probed with the "positive" serum, and six arrays were probed with the "normal" serum. This represents a two-class scenario to assess the relative signals between samples and variations among replicates.
2. Sample titrations: a total of 20 arrays were used. "Positive" and "normal" sera were mixed at 0:100%, 25:75%, 50:50%, 75:25%, 100:0% proportions. Four arrays were probed with each mix of sera. This represents a more continuous situation with a known relationship in the signals between titrated samples.[29]

Data Acquisition and Analysis. After probing the sample, each array was scanned using Axon GenePix 4000B fluorescent microarray scanner which yields a Tagged Image File Format (TIFF) image. The arrays were scanned at two different wavelengths: 635 nm (red) and 532 nm (green) since two fluorescently labeled antibodies were used. Each image was then processed by GenePix version 6.1,[17] which quantitatively assesses the intensity of each spot on the array. The output of GenePix is a GenePix Pro Results File Format (GPR) text file containing protein reactivity measurements. It also provides flags and quantities reflecting the spot quality, such as SNR (signal-to-noise ratio) and percent saturation, which were used in data preprocessing to filter out protein features that were saturated or without significant signal. We did not filter control proteins used for training the robust linear model algorithm.

Each array was scanned twice at two different PMT gains, High (700/800 V), and Low (550/700 V) for 635 nm/532 nm channels, respectively, to simulate PMT drifting that may occur in practice and to evaluate the performance of normalization procedures in that situation.

Image Processing. Image processing is another example where tools developed for DNA microarray analysis cannot be applied "as-is" to protein arrays, but require careful considerations. Features on DNA arrays usually consist of a square containing a well-defined circle, where the actual foreground signal is measured. Background signal is then computed from the remaining region that does not include the foreground region. The process of identifying objects in an image, such as the foreground and the background region, is called segmentation. In protein arrays, features may not contain circular shapes because of smears, if the signal is quite strong, or small speckles. This situation may easily affect the proper segmentation of foreground and background regions. We compared two different segmentation procedures provided by GenePix: circular and irregular. The former assumes that the foreground signal in the spot is circular and identifies the best diameter. The latter does not assume any specific shape and thus segments the foreground signal from the background with any irregular shape. We determined that irregular segmentation had better performance, and we used the irregular segmentation option for all the analysis reported in this paper (see Supporting Information for more details).

Intra- and Interarray Variability. We assess the extent of intra- and interarray variability by measuring the Coefficient of Variation (CV), defined as

$$\mathrm{CV} = \frac{\sigma}{\mu}$$

where σ is the standard deviation and μ is the mean of the signal at original scale. Regarding intra-array variability, the computation of CV is based on the measurements of control

proteins spotted in each subarray within an array. Concerning interarray variability, CV is computed for the protein spot at the same location across multiple arrays.

Sample Separability Score. To examine the separability of the two serum samples, we apply Fisher's signal-to-noise ratio to $\log_2$-transformed signals, which is defined as

$$S = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$$

where μ_1 and μ_2 are the mean values of a protein feature for the two classes, and σ_1 and σ_2 are the corresponding standard deviations. Along with CV, we use it to evaluate and compare different normalization procedures. Ideally, normalization procedures should preserve the real biological difference while reducing the variation between replicates, yielding higher values of S with respect to the original data.

Normalization Methods. We compared three different normalization procedures:

1. *Global*: The signal levels of each array are scaled with a factor such that the signal medians of all the arrays are the same as the overall median. In the DNA microarray context, where global normalization is commonly used, this strategy assumes that the overall amount of transcription in each sample is the same.[30]
2. *Quantile*: Quantile normalization has been developed mainly in the context of DNA microarrays, where it is one of the most common normalization procedures.[31] This method assumes that the distribution of actual signals is the same in all samples and adjusts the observed data accordingly. Algorithmically, the largest signal for each array is replaced by a median value of the largest signals; the second largest signal is replaced by a median value of the second largest signals, and so forth. Many software packages for DNA microarray analysis include quantile normalization as a standard preprocessing step. We used the implementation provided by the Bioconductor package "limma".[32]
3. *Robust Linear Model* (*RLM*):[33] We applied a robust linear model to normalize the data by exploiting the control proteins. The fitting to a linear model is performed through a robust regression by means of an iteratively reweighted least-squares procedure with a robust estimator, like the median. Generally, RLMs are less sensitive to outliers in the data because outliers are weighted down in the model fitting procedure and the median is less affected than the mean by potential outliers. Preliminary results suggest that the systematic variations across subarrays or slides appear to be a multiplicative factor to the protein binding signals. Hence, we propose a linear model on log-transformed signals to estimate and correct the variations as the following formula:

$$y_{ijkr} = \alpha_i + \beta_j + \tau_k + \varepsilon_{ijkr}$$

where

- y_{ijkr} is the observed signal in logarithmic scale for spot r of protein feature k located in subarray j on the i-th array. Thus, r ranges from 1 to 2, the spot replicate; k from 1 to n_f, where n_f is number of protein features; j from 1 to 48 subarrays; and i from 1 to n_s, where n_s is the number of arrays;
- α_i is the slide effect of slide i, accounting for the overall differences among arrays possibly due experimental and/or scanning conditions;
- β_j represents the subarray j effect, accounting for spatial and printing pin effects;
- τ_k is the effect of the protein feature k, accounting for the spotted protein amount and binding affinity of different protein features;
- ε_{ijkr} is the random error. We assume that $\varepsilon_{ijkr} \sim N(0,\sigma^2)$; that is, that the random error has a normal distribution centered around zero with the same variance σ across all spots.

This model can be applied simultaneously to both intraslide and interslide normalization. Once the best parameters of the model are estimated, the logarithmic intensity of each spot is corrected according to

$$y'_{ijkr} = y_{ijkr} - \alpha_i - \beta_j$$

For intraslide normalization of a single slide, term α_i is dropped. In this study, we evaluated normalization using different sets of control proteins, including human-IgG series and anti-human-IgG series (together denoted as IgG), which respond to serum and secondary antibody, or V5-CMK1 series (denoted as V5), which responds to spike-in anti-V5 antibody, or the combination of all three control series (denoted as IgG+V5). Human-IgG and anti-human-IgG spots can be used as normalization references because they are chemically saturated in our experiment condition and their observed signals are expected to reflect only technical variations of the arrays or assays. Normalization was typically carried out using the control signals acquired at the same wavelength (same-channel normalization). In addition, since the arrays were scanned at two wavelengths, we were able to evaluate cross-channel normalization. In this case, the model was fitted using the control proteins acquired at one wavelength and then signals acquired at the second wavelength were normalized with the coefficients derived from this model.

Weighted or iteratively reweighted least-squares algorithms have also been used in robust local regression methods (i.e., LOWESS) for normalization of cDNA microarrays.[34,35] However, in those settings, the procedure has to be applied to a sufficiently large number of genes with the assumption that only a small portion of those genes may be differentially expressed across conditions. In a different context, a robust linear model fitting procedure was applied to estimate gene expression level or difference on Affymetrix microarrays.[36] To the best of our knowledge, a robust linear model has never been implemented for protein array normalization. We used the "rlm" function in the R MASS package to implement our normalization procedure in this study.[37]

Hit List via M-Statistic. The final goal of serum profiling on functional protein arrays is to identify a set of candidate autoantibodies that are differentially reactive between two different populations, such as normal and disease. To this goal, Pearson's t test (or its robust nonparametric counterpart: Mann−Whitney test) could be applied. By identifying a proper cutoff on p-values, a hit list of differentially reactive proteins can be constructed. The t test and Mann−Whitney specifically take into account the entire set of samples in groups. However, due to biological heterogeneity of the immune response, it is possible that an antibody is reactive to a certain protein in only

Table 1. List of Abbreviations

abbreviation	description
CV	Coefficient of variation, i.e., standard deviation divided by the mean
S	Fisher's Signal-to-Noise ratio
orig	Original data, i.e., raw data without any normalization
global	Global normalization
quantile	Quantile normalization
rlm.IgG.r	Robust Linear model normalization with Human IgG or Anti-human IgG control proteins, using the red channel, i.e., same-channel normalization
rlm.V5.r	Robust Linear model normalization with V5 control proteins, using the red channel, i.e., same-channel normalization
rlm.IgG+V5.r	Robust Linear model normalization with Human IgG or Anti-human IgG and V5 control proteins, using the red channel, i.e., same-channel normalization
rlm.V5.g	Robust Linear model normalization with V5 control proteins, using the green channel, i.e., cross-channel normalization

a subset of samples in one group. Hence, to identify protein features that show different reactivity between the two groups, we use a statistical method called M-statistic that can identify proteins reactive in only a subset of a group. Briefly, all of the signal values for a given protein in one group (e.g., group 1, control samples) are iteratively compared to the values in the comparison group (e.g., group 2, diseased samples). The number of values in group 2 which are above the highest value in group 1 are counted (M^*) and a *p*-value is determined by computing the probability of having M greater than M^* via hypergeometric distribution. Then, the number of values in group 2 which are above the second highest value in group 1 are counted and a *p*-value is determined and so on until all of the values in group 1 are considered. A cutoff value (or threshold) is defined as the comparison which returns the lowest *p*-value. This lowest *p*-value is the significance of the protein in consideration. Hits identified via M-statistic by ProtoArray Immune Response Biomarker Profiling assays have been validated using a variety of approaches.[38,39]

Data analysis was performed with R version 2.5.[40] Table 1 lists the set of definitions and abbreviations used throughout the paper.

Results

Intra-Array Variability. Intra-array variability was estimated by computing the CV using only the control features replicated in each subarray that had median signal intensities significantly above the background and were not used in the training of the robust linear model (RLM). In the original data for arrays scanned at the High PMT, the median CV was 0.24 (Figure 2), whereas RLM normalization reduced the median CV to between 0.19 and 0.20 when the reference control signals in the same channel were used to fit the model. Cross-channel RLM normalization, that is, using the control signals acquired at a different wavelength than the signal to be normalized, is less effective than same-channel normalization, resulting in a median CV of 0.23. On the other hand, global and quantile normalization are typically applied on several arrays, and thus, they are not designed to address intra-array variability. In fact, when global normalization is performed on a single array, a single scaling factor is applied to all features, resulting in no effect at all to intra-array CV. Theoretically, quantile normalization may affect the intra-array CV depending on the data distribution. But in this experiment, there was minimal effect.

Signal Distribution. We compared the distribution of the signals across the arrays in each group before and after applying the normalization procedures. As an illustration, we also compared it with a DNA microarray profiling mRNA (Figure

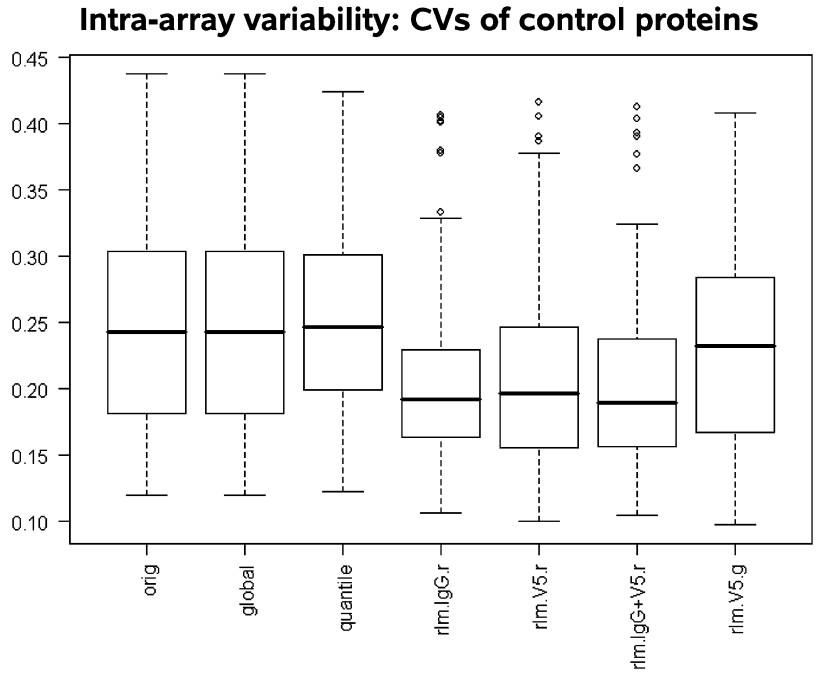
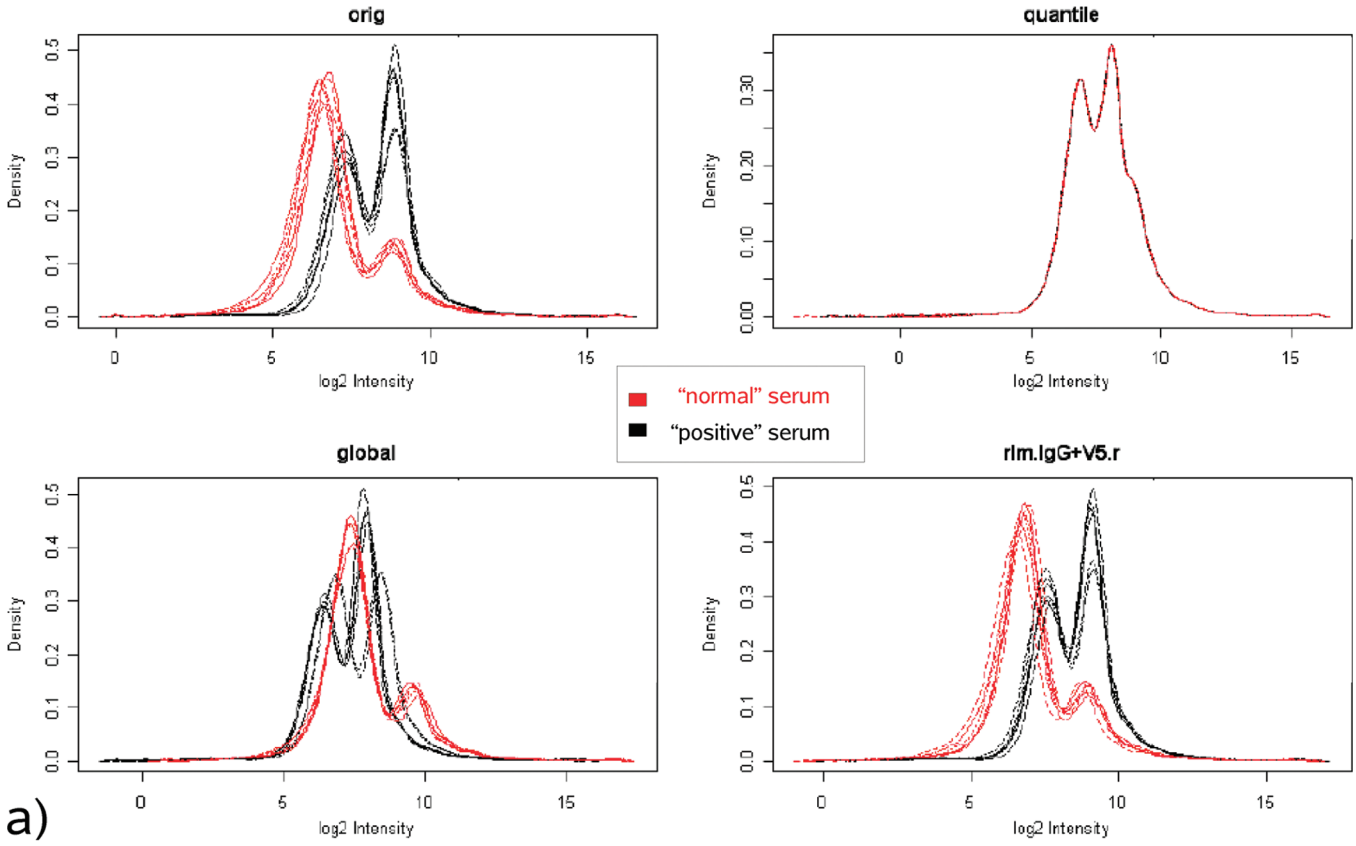


Figure 2. Effect of normalization on intra-array CV. The computation of CV was performed by looking at the control proteins spotted on each subarray that were not used for normalization. The box plots show the results for all 12 slides (6 replicates of the "positive" serum and 6 replicates of the "normal" serum), see Table 1 for definitions.



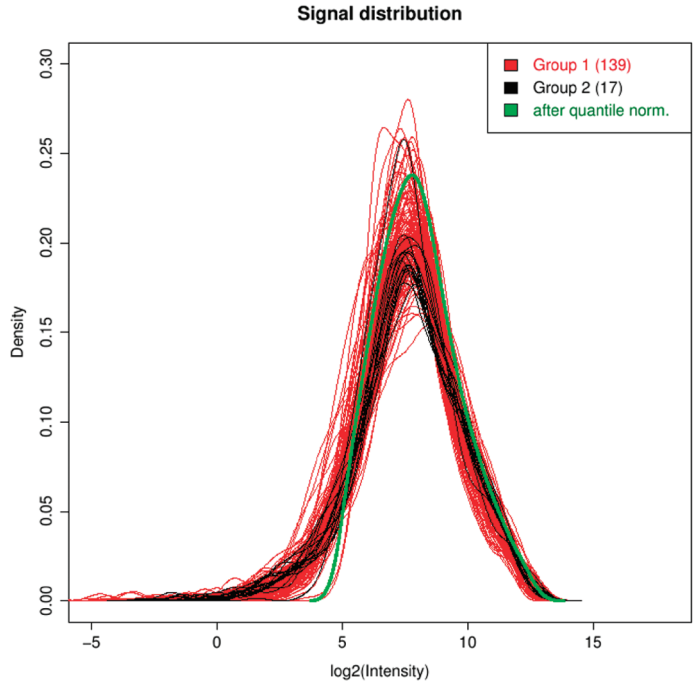


Figure 3. Signal distribution. (a) Signal distribution for noncontrol proteins before and after normalization. Red and black lines represent the signals of "normal" and "positive" samples in each replicated experiment, respectively (see Table 1 for definitions). The result of RLM with control spots is reported in the Supporting Information. (b) Signal distribution of a DNA microarray including 156 samples divided in two classes: group 1 has 139 samples, whereas group 2 contains 17 samples. The green line represent signal distribution after quantile normalization.

3).[41] It is worth noting that, while we are comparing technical replicates within the "positive" and "normal" samples for the protein array, the DNA microarray data set includes tissues from different subjects, thus, adding biological variability to the technical variability.

Quantile normalization assumes that the signal distribution is the same across all samples. In fact, the distribution of signals on DNA arrays is similar between multiple samples, and thus, quantile normalization is appropriate in this context (Figure 3b). However, protein arrays show rather different signal

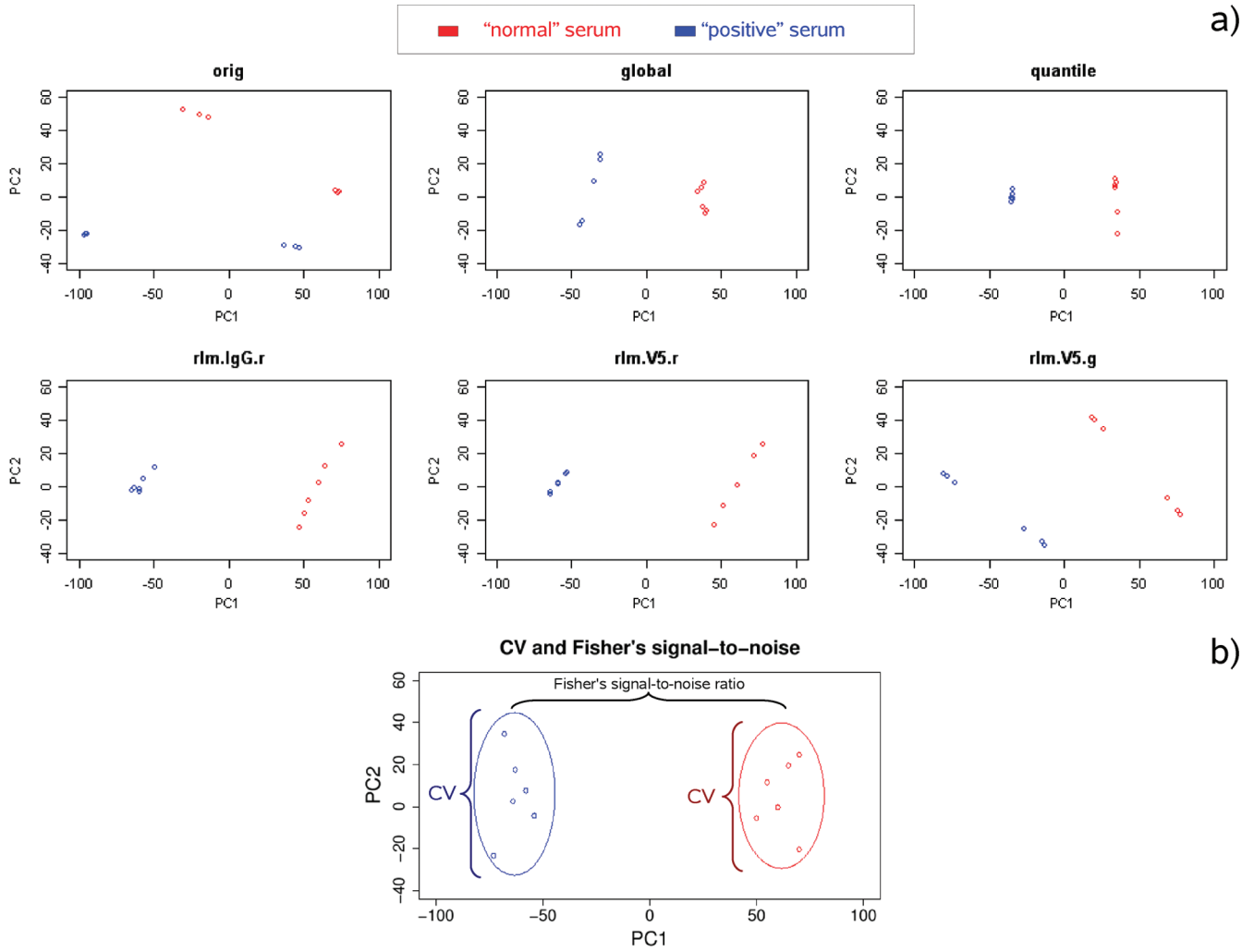


Figure 4. Normalization comparison using the principal component analysis (PCA) for samples acquired at different PMTs. Log_2-transformed signals were used in PCA (see Table 1 for definitions). (a) The plots show the results for the actual data using different normalization methods. (b) Schematic of plot interpretation: Fisher's signal-to-noise ratio is equivalent to the distance between the two clusters, whereas CV can be represented by the size of the cluster. The actual details of the computation are reported in the text.

distribution for the two samples (Figure 3a). Hence, the assumptions of quantile normalization do not hold. As expected, quantile normalization completely removes any variation in the signal distribution by translating the signal into a unique distribution which does not resemble the original distribution of either of the two sera. Global normalization simply shifts the signal distributions. Conversely, RLM is able to reduce variability while preserving the original distribution of the signals.

Normalization of Interarray Variation due to PMT Difference. One of the common experimental procedures to avoid saturated (or too low) signals in the scanned image is to change PMT gains. This may have a particularly detrimental effect on comparative analysis of different arrays. Hence, we examined the effect of normalization on PMT variation. We randomly selected 3 arrays acquired with High PMT and 3 arrays with Low PMT for both "positive" and "normal" samples.

We calculated Fisher's Signal-to-Noise ratio *S* and performed Principal Component Analysis (PCA) to examine the separability of the samples both before and after normalization. PCA provides a visual illustration of sample separability, whereas Fisher's Signal-to-Noise ratio yields a more quantitative description.

Figure 4 shows the results of PCA. Each plot shows the projection of the replicated samples along the first two components of PCA. PCA components are linear combination of the features, that is, proteins. The first two components capture most of the variability of the data and can be used to draw a 2D representation of the data, as in this case.

Clearly, in the original data, the replicated arrays of the same sample acquired at the same PMT cluster closely together, and arrays scanned at High PMT are separated from those scanned at Low PMT. Although the distinction of the two classes, that is, "positive" and "normal", is still present, this might not be the case for more subtle differences as in normal versus cancer patients. All normalization procedures seem to be able to correct the PMT effect; however, both global and quantile normalization also reduce the distance between the two classes.

We took a closer look at these data by comparing Fisher's signal-to-noise ratio and interarray CV before and after normalization (see Figure 5). Although global and quantile normalization improve the signal-to-noise ratio with respect to the original data, RLM with either IgG and/or V5 controls in the same channel achieves the best improvement. This result is particularly relevant when differences between samples may be subtler than in this study.

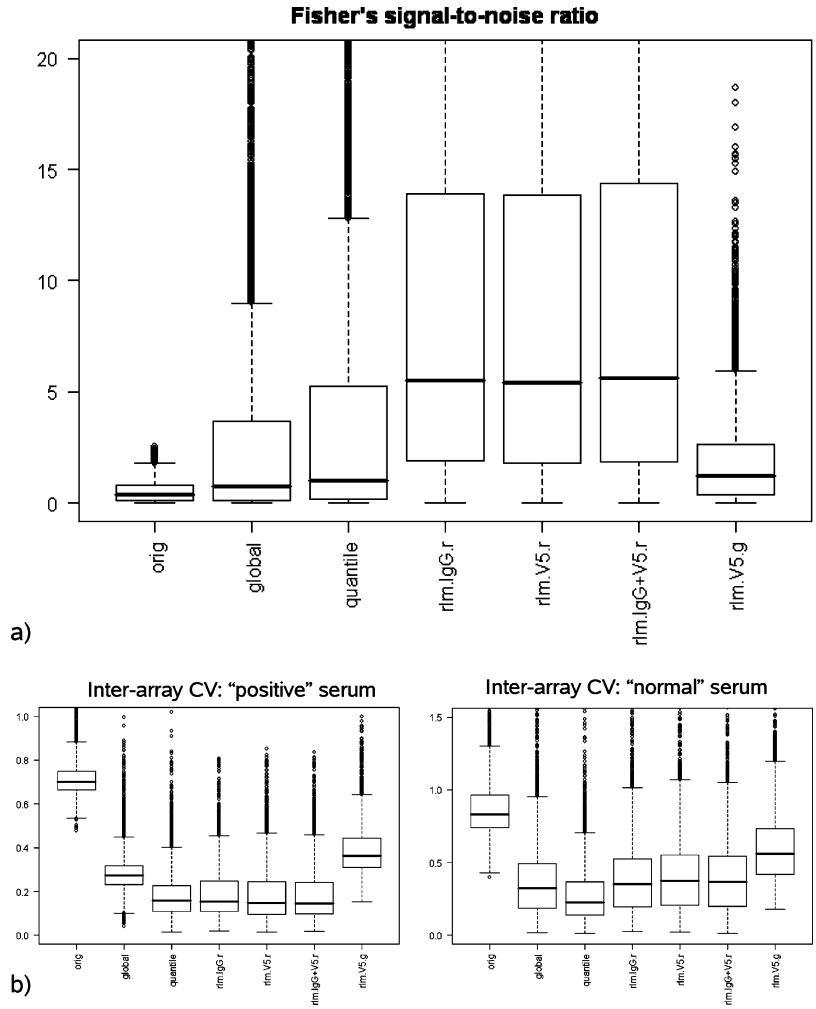


Figure 5. Interarray variability and sample separability. (a) Fisher's signal-to-noise-ratio before and after normalization. The *y*-axis is trimmed at 20. (b) Comparison of normalization methods by means of interarray CV (see Table 1 for definitions).

Furthermore, all normalization procedures significantly reduce the rather high interarray variability seen in this experiment (Figure 5b). We again observe that cross-channel normalization is less effective than same-channel normalization.

Sample Titration Experiments. To evaluate the effectiveness of normalization in more detail, we performed a titration experiment in which the two serum samples were mixed at defined ratios (see Figure 1d). We preprocessed the data to identify noncontrol protein features that were reactive with at least one of the sera used (SNR $\geq$ 2, not saturated). As expected, most feature reactivity is linearly proportional to the titration even with the raw data as shown in the top-left graph of Figure 7. Therefore, we calculated the correlation coefficients between the signal and the input (titration proportion) as an indicator of the signal accuracy. Given an increasing concentration of the "positive" sample, proteins that are more reactive to the "positive" serum should show positive correlation and those that are more reactive to the "normal" serum should show negative correlation. Proteins with similar reactivity to "positive" and "normal" sera should have correlation coefficients close to zero, that is, show no correlation. Given the nature of the serum samples used, we would expect more proteins showing a positive correlation with "positive" serum concentration.

Our results show that RLM normalization improved the correlation as the distribution peaks shift toward 1 or -1 in Figure 6 relative to the original data, suggesting an improvement of data accuracy. This is likely achieved by reducing the variation between replicates while retaining the signal range, as also indicated by the results shown in Figure 7a,b, bottom right.

On the other hand, quantile normalization reduced, eliminated or even reversed the correlation between the signal and the input as reflected by the flat distribution in Figure 6. Figure 7 shows the results for two proteins on the array. In Figure 7a, top-right, although quantile normalization effectively reduced the variation between replicates, it did so at the cost of reducing the signal range. In Figure 7b, top-right, quantile normalization generated an artificial negative correlation in a protein feature with nearly constant reactivity across the titrated samples.

For global normalization, the impact is even more dramatic: the majority of the proteins reversed their correlation (Figures 6 and 7a,b, bottom-left). These results demonstrate that global normalization drastically distorted the data in this experiment.

Impact of Normalization on Hit Calls. Ultimately, the function of protein arrays is to distinguish samples based on their reactivity. In this experiment, we probed arrays with two serum samples (in 6 replicates) and determined the significant protein features (hits) using M-statistic (see Materials and Methods). We determined how many spotted proteins had increased reactivity with one serum sample or the other using

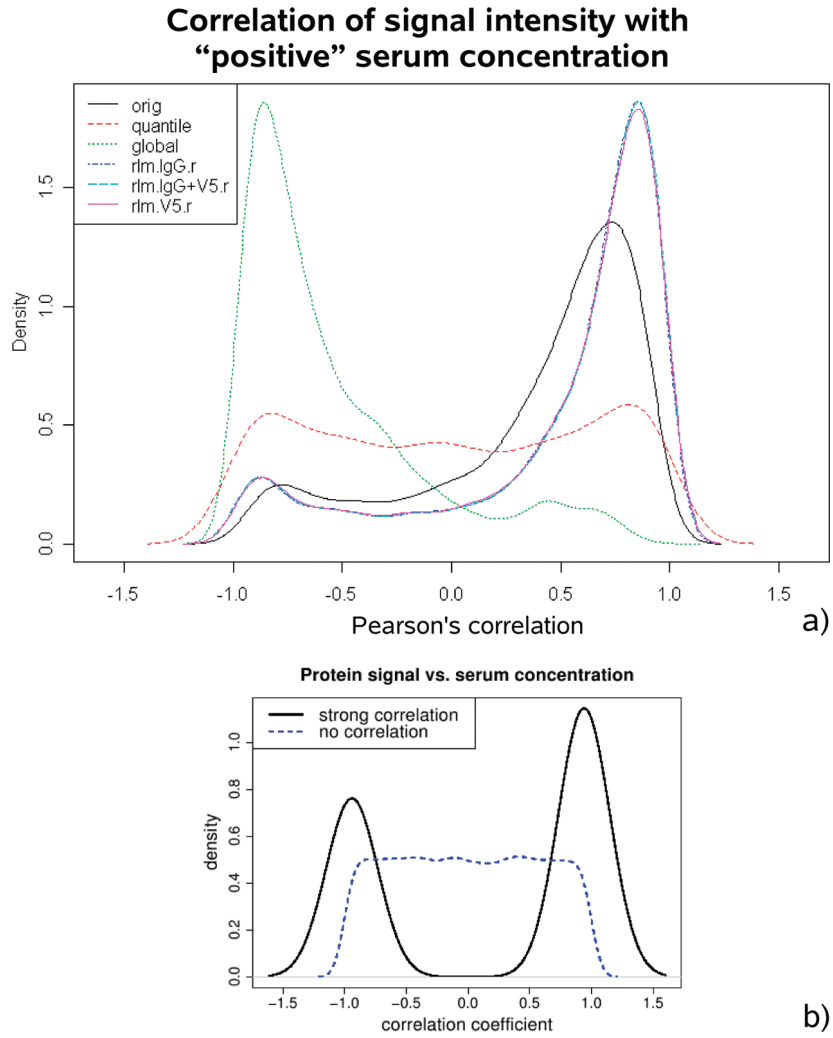


Figure 6. Density distribution of Pearson's correlation between signal and titration proportion. (a) The plot includes 1390 proteins satisfying the following criteria across all 20 replicates: no saturated spots, less than 20% flagged spots; more than 80% spots with a Signal-to-noise ratio ≥2 (see Table 1 for definitions about the legend). (b) The two extreme situations are depicted in a simulated example. If there is no correlation between the entire set of proteins and serum concentration, we will observe an almost flat line ranging from −1 to +1 (dashed blue line). On the other hand, if there is a correlation between reacting proteins and serum concentration, we will see two peaks around −1 and +1 (for normal and positive correlated proteins, respectively). Since we have more proteins reactive to the "positive" serum than to "normal" one, in this example, the peak centered around +1 is higher than the one at −1, as expected.

either the raw data acquired at the same PMT or signals normalized with quantile, global or RLM. The results are summarized in Figure 8. We expect that the "positive" serum should show a higher number of reactive proteins compared to the "normal" serum. When comparing the raw data, 2574 proteins have elevated reactivity with the "positive" serum as compared to the "normal" serum. Surprisingly, global normalization reduced that number to 289 and quantile reduced the "positive" hit-list to 1048, therefore, increasing the number of potential false negatives. RLM, however, extended the hit-list to 4110 proteins. Conversely, 510 proteins were found to be more reactive to the "normal" serum in the raw data. Global normalization raised that list to 2503 proteins, and quantile to 1210 proteins. RLM reduced the "normal" hit-list to 439 proteins. Thus, RLM maintains a similar proportion of hits between the two samples as seen in the original data and as expected from the experimental design.

Since the hit-lists changed dramatically after normalization (especially after quantile and global normalization), we investigated if also the classification of the hits defined in the raw data changed after normalization. We found that of the 2574 "positive"-serum reactive proteins, global normalization classifies 1396 of them as being more reactive with the "normal" serum. This redefinition of reactive proteins occurred with quantile normalization too, although to a lesser extent. Indeed, quantile normalization redefined 464 of the "positive"-serum reactive proteins as "normal"-serum reactive ones. Conversely, RLM changed the classification of only 3 proteins (from "normal" to "positive" serum reactive). By this measure, RLM again better maintains the trends seen in the original data and expected by experimental design, whereas quantile and global normalization appear to introduce artifacts.

Overall, the results from this set of experiments are very consistent with those in the previous sections. Reduction of technical variability without losing biological signals with RLM normalization improved the sensitivity of M-statistic calls, resulting in more significant hits. Quantile and global normalization dramatically changed the hit lists due to signal distor-

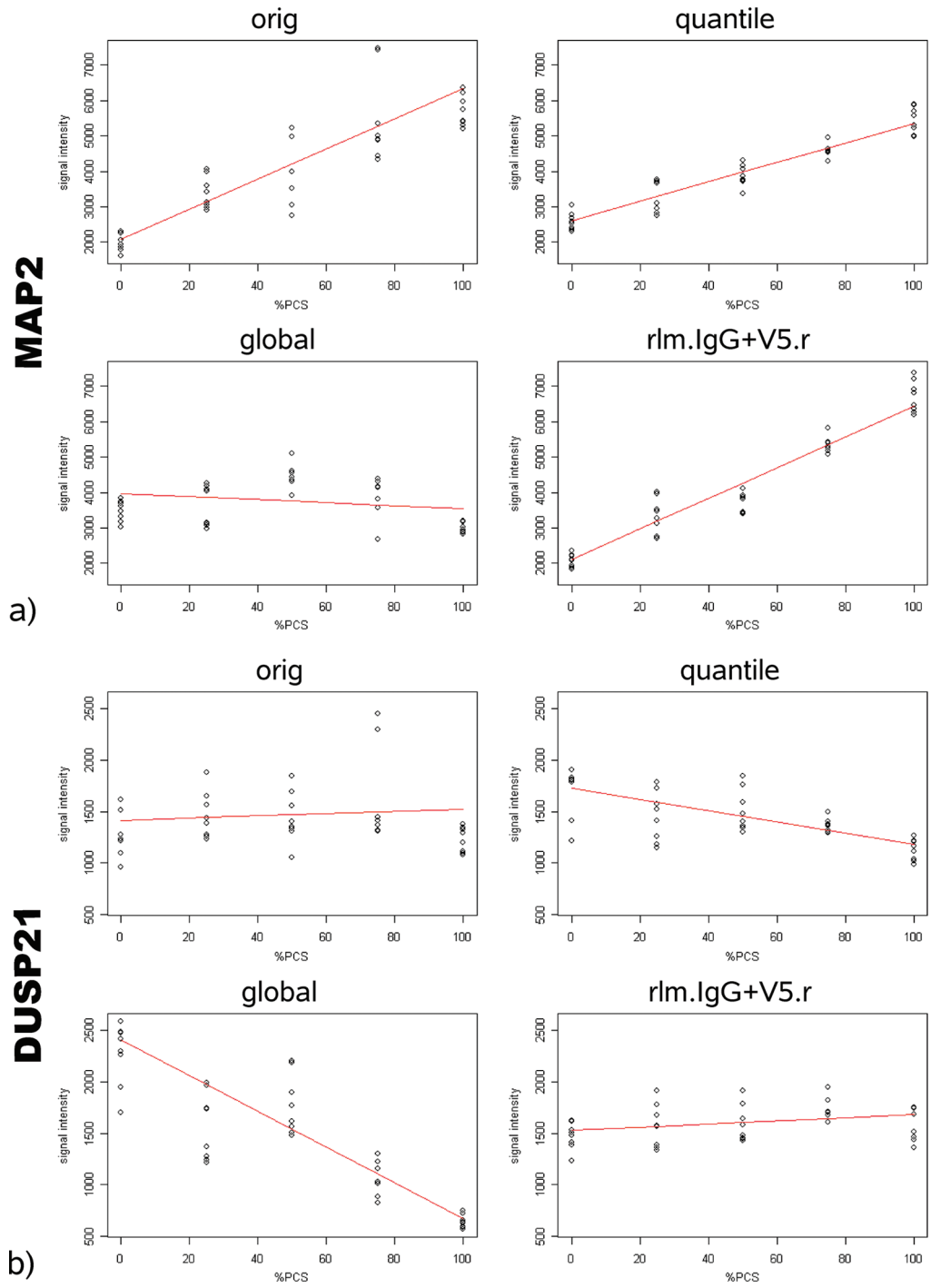


Figure 7. Correlation of signal with titration proportion for protein MAP2 (a) and protein DUSP21 (b) as measured by the protein arrays. Each panel shows the original data, quantile normalization, global normalization and RLM normalization with IgG and V5 (see Table 1 for definitions). Red lines represent fitted regression lines.

tion, including strong artifacts that reversed the polarity of the hits, as shown in Figure 6.

Discussion and Conclusions

Protein microarrays have become important tools for large-scale analysis of protein function and interactions, often taking advantage of the well-developed DNA microarray technology. Here, by means of a comprehensive comparative analysis, we show that some assumptions made by studying DNA microarrays may not be true for functional protein microarrays, although they share the common framework of immobilizing probes on a surface. Some artifacts affecting protein arrays may play a different role with respect to DNA microarrays. For example, local spatial effects are a substantial source of noise for protein arrays due to differences in the printing process or uneven probing conditions, whereas this aspect is not as critical for DNA microarrays. Other artifacts are related to the nature of binding between antibodies and proteins, which is markedly different from the hybridization process on DNA microarrays. For example, highly reactive proteins can yield strong signals

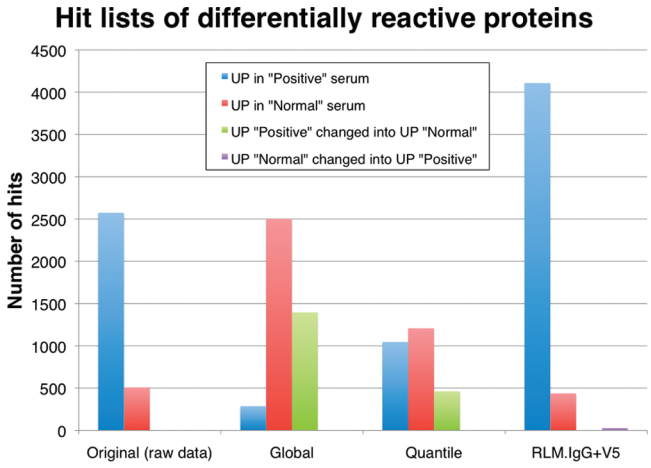


Figure 8. Effect of normalization on hit-list. The first two bars in each category show the number of reactive proteins, i.e., hits, detected by M-statistic. The normalization methods also report the number of reactive proteins that changed classification, i.e., up in "positive" or "negative" serum, with respect to the original data.

that exceed the size of the spot. This technical variability and the differences in binding characteristics contribute to one of the most distinct characteristics between the two array types: signal distribution. In DNA microarrays, the overall amount of mRNA is assumed to be nearly the same across samples, therefore enabling global or quantile normalization approaches. The assumption of the global normalization method is that arrays have the same median expression value, whereas the quantile normalization assumes the same signal distribution across arrays. These assumptions hold true for DNA microarrays, since large numbers of "neutral" genes in the samples show levels of expression that cover the entire dynamic range of the microarray. However, this cannot be considered a reasonable hypothesis for protein arrays. When serum is profiled on functional protein microarrays, typically only a small portion of protein features on the array shows reactivity to the probed serum samples. Furthermore, the number of reactive proteins and their reactivity can vary significantly from sample to sample, leading to markedly different signal distributions. These two observations imply a violation of the assumptions underlying quantile and global normalization. Indeed, we demonstrated that global and quantile normalization methods are not always suitable for protein microarrays and blindly applying those normalization methods may result in loss of real signal differences between samples or production of artifacts.

The experimental design of this study resulted in more proteins on the array that were reactive to the "positive" than to the "normal" serum, although the reactive proteins were still only a small fraction of the number of expressed genes typically seen on DNA microarrays. The selection of these two samples may have exaggerated the problems identified with global and quantile normalization to some extent, but similar artifacts were introduced by quantile normalization when comparing autoimmune disease patients with normal controls (unpublished results). Global and quantile normalization procedures distorted the signal such that they impaired the detection of reactive proteins. Conversely, RLM was able to improve the detection of those proteins.

Among the main advantages of RLM are its flexibility to capture many sources of errors, such as local spatial variations (subarrays), overall differences in the brightness of the arrays, and its ability to exploit control proteins, either internal or spike-in controls. Moreover, the robust model fitting algorithm used in RLM makes this method less prone to artifacts introduced by outliers.

Some other normalization strategies for protein microarrays have been proposed or evaluated in recent studies. Zhu et al. proposed ProCAT: a normalization procedure for functional protein microarrays in the context of radiometric kinase assays.[42] ProCAT specifically focuses on a single array by reducing the effects of noise in the background signal of a spot via a window of neighboring spots. The use of this window allows the elimination of potential artifacts in the background signal due to smears and smudges, which are prevalent in radiometric assays. This procedure enhances the signal-to-noise ratio of each spot. Moreover, a neighboring window is used also to normalize the foreground signal of a spot by using robust statistics less prone to outliers, helping to reduce spatial effects. For kinase assays, ProCAT is particularly useful because the main purpose is the analysis of catalytic activities of a protein. In our setting, however, signal scattering is not a concern in the fluorescent detection system. The main problem with fluorescent-based detection strategies is interarray variability since we need to identify reactive proteins across multiple arrays where differences between protein reactivities may be subtle.

Hamelinck et al. reported that global normalization with ELISA-determined IgM concentration performed the best among the six normalization methods evaluated on their antibody microarray platform.[43] Although this strategy can be easily extended to subarray level as long as a sufficient number of reference protein spots are present in each subarray, the requirement to perform ELISA on every sample is a limiting factor. In addition, it may be challenging to identify a good reference protein to calibrate with ELISA in the immune response profiling on functional protein microarrays. Sundaresh and colleagues applied DNA microarray analysis techniques to their pathogen protein microarrays, including variance stabilizing normalization (VSN), which performs data transform and normalization at the same time.[44−46] The parameters of the transform/normalization were estimated with either all content or only control proteins using a robust algorithm. However, in this case too, subarray effects are not taken into account in the method. Given that VSN employs a nonlinear function (arsinh), it would be difficult to extend VSN to fit a multifactor model such as ours, although it may be an interesting future direction to explore.

Marina et al. addressed the protein concentration variation in protein microarrays, an orthogonal source of variation to the ones we have considered in this study.[39] The core of their concentration-dependent analysis (CDA) method is a modified *Z*-score transform that is localized to a neighborhood of protein concentration and robust against outliers. The variation of protein concentrations may be more important for intra-array analysis, that is, when comparing different proteins on the same array to determine significant hits, than when the same protein is compared across arrays, that is, interarray comparison. As long as each protein is printed consistently across arrays, the absolute concentration of a protein will only affect the relative signal-to-noise ratio and thus power or sensitivity of detection. CDA for comparison between arrays would be similar to applying global normalization in a concentration neighborhood, assuming that a sufficiently large proportion of proteins

in any given concentration neighborhood have a similar distribution across arrays or conditions. Therefore, in cases where such an assumption is violated, comparison of the signals across two RLM normalized arrays may provide more accurate results than the comparison of *Z*-scores.

In this manuscript, we show that some assumptions of DNA microarray studies are not applicable to protein arrays. To this end, we have proposed a new normalization protocol able to properly address the variability seen in protein array experiments. RLM reduced technical variability, but unlike other normalization strategies, RLM was able to preserve biological difference. Although the results we present in the paper refer to a particular application and platform, such as functional protein microarrays, we believe that the problems we highlight might apply in general to other protein array applications or platforms. However, further analysis should be carried out to examine to what extent the issues reported here are common across different platforms.

Data Sets and Source Code Availability. RLM is included in the Prospector v5.2, the software tool by Invitrogen for analyzing protein chips. The software is available for free download at: http://www.invitrogen.com/site/us/en/home/Products-and-Services/Applications/Protein-Expression-and-Analysis/Biomarker-Discovery/ProtoArray/Online-Tools.html. The data sets used in this study can be downloaded from http://archive.gersteinlab.org/proj/proteinchipRLM/.

Acknowledgment. Conflict of Interests: Mark B. Gerstein held an appointment as consultant at Invitrogen, Inc. for parts of 2008.

Supporting Information Available: Additional information on image processing. Figure S1, comparison of two segmentation procedures to identify foreground and background signals. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Fan, J.; Chee, M. S.; Gunderson, K. L. Highly parallel genomic assays. *Nat. Rev. Genet.* **2006**, *7* (8), 632–644.
(2) Urban, A. E.; Korbel, J. O.; Selzer, R.; Richmond, T.; Hacker, A.; Popescu, G. V.; Cubells, J. F.; Green, R.; Emanuel, B. S.; Gerstein, M. B.; Weissman, S. M.; Snyder, M. High-resolution mapping of DNA copy alterations in human chromosome 22 using high-density tiling oligonucleotide arrays. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (12), 4534–4539.
(3) Lieb, J. D.; Liu, X.; Botstein, D.; Brown, P. O. Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat. Genet.* **2001**, *28* (4), 327–334.
(4) Ren, B.; Robert, F.; Wyrick, J. J.; Aparicio, O.; Jennings, E. G.; Simon, I.; Zeitlinger, J.; Schreiber, J.; Hannett, N.; Kanin, E.; Volkert, T. L.; Wilson, C. J.; Bell, S. P.; Young, R. A. Genome-Wide Location and Function of DNA Binding Proteins. *Science* **2000**, *290* (5500), 2306–2309.
(5) Hall, D. A.; Ptacek, J.; Snyder, M. Protein microarray technology. *Mech. Ageing Dev.* **2007**, *128* (1), 161–167.
(6) Zhu, H.; Snyder, M. Protein chip technology. *Curr. Opin. Chem. Biol.* **2003**, *7* (1), 55–63.
(7) LaBaer, J.; Ramachandran, N. Protein microarrays as tools for functional proteomics. *Curr. Opin. Chem. Biol.* **2005**, *9* (1), 14–19.
(8) Bertone, P.; Snyder, M. Advances in functional protein microarray technology. *FEBS J.* **2005**, *272* (21), 5400–5411.
(9) Hueber, W.; Utz, P. J.; Steinman, L.; Robinson, W. H. Autoantibody profiling for the study and treatment of autoimmune disease. *Arthritis Res.* **2002**, *4* (5), 290–295.
(10) Zhu, H.; Bilgin, M.; Bangham, R.; Hall, D.; Casamayor, A.; Bertone, P.; Lan, N.; Jansen, R.; Bidlingmaier, S.; Houfek, T.; Mitchell, T.; Miller, P.; Dean, R. A.; Gerstein, M.; Snyder, M. Global analysis of protein activities using proteome chips. *Science* **2001**, *293* (5537), 2101–2105.
(11) Smith, M. G.; Jona, G.; Ptacek, J.; Devgan, G.; Zhu, H.; Zhu, X.; Snyder, M. Global analysis of protein function using protein microarrays. *Mech. Ageing Dev.* **2005**, *126* (1), 171–175.
(12) Schweitzer, B.; Kingsmore, S. F. Measuring proteins on microarrays. *Curr. Opin. Biotechnol.* **2002**, *13* (1), 14–19.
(13) Schweitzer, B.; Predki, P.; Snyder, M. Microarrays to characterize protein interactions on a whole-proteome scale. *Proteomics* **2003**, *3* (11), 2190–2199.
(14) Hong, S.; Misek, D. E.; Wang, H.; Puravs, E.; Giordano, T. J.; Greenson, J. K.; Brenner, D. E.; Simeone, D. M.; Logsdon, C. D.; Hanash, S. M. An autoantibody-mediated immune response to calreticulin isoforms in pancreatic. *Cancer Cancer Res.* **2004**, *64* (15), 5504–5510.
(15) Nam, M. J.; Kee, M. K.; Kuick, R.; Hanash, S. M. Identification of Defensin {alpha}6 as a Potential Biomarker in Colon Adenocarcinoma. *J. Biol. Chem.* **2005**, *280* (9), 8260–8265.
(16) van Rhee, F.; Szmania, S. M.; Zhan, F.; Gupta, S. K.; Pomtree, M.; Lin, P.; Batchu, R. B.; Moreno, A.; Spagnoli, G.; Shaughnessy, J.; Tricot, G. J. NY-ESO-1 is highly expressed in poor prognosis multiple myeloma and induces spontaneous humoral and cellular immune responses. *Blood* **2005**, *105* (10), 3939–3944.
(17) Molecular Devices. GenePix Pro 6.1 [Internet] [accessed 2009 May 7] Available from: http://www.moleculardevices.com/pages/software/gn_genepix_pro.html.
(18) Imagene. BioDiscovery Microarray Analysis Software and Services [Internet] [accessed 2009 May 7] Available from: http://www.biodiscovery.com/index/imagene.
(19) Medigue, C.; Rechenmann, F.; Danchin, A.; Viari, A. Imagene: an integrated computer environment for sequence annotation and analysis. *Bioinformatics* **1999**, *15* (1), 2–15.
(20) Imaging Research Inc. GE Healthcare Life Sciences - ArrayVision" [Internet] [accessed 2009 May 7] Available from: http://www4.gelifesciences.com/APTRIX/upp01077.nsf/content/microarrays_analysis_arrayvision?OpenDocument.
(21) Eisen, M. EisenLab [Internet] [accessed 2009 May 7] Available from: http://rana.lbl.gov/EisenSoftware.htm.
(22) TM4. Spotfinder [Internet] [accessed 2009 May 7] Available from: http://www.tm4.org/spotfinder.html.
(23) Tusher, V. G.; Tibshirani, R.; Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (9), 5116–5121.
(24) Gentleman, R.; Carey, V.; Bates, D.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y.; Gentry, J.; Hornik, K.; Hothorn, T.; Huber, W.; Iacus, S.; Irizarry, R.; Leisch, F.; Li, C.; Maechler, M.; Rossini, A.; Sawitzki, G.; Smith, C.; Smyth, G.; Tierney, L.; Yang, J.; Zhang, J. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **2004**, *5* (10), R80.
(25) Baldi, P.; Long, A. D. A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes. *Bioinformatics* **2001**, *17* (6), 509–519.
(26) Schermer, M. Confocal scanning microscopy in microarray detection. In *DNA Microarrays—a Practical Approach*; Oxford University Press: Oxford, U.K., 1999; pp 17−42.
(27) Lyng, H.; Badiee, A.; Svendsrud, D.; Hovig, E.; Myklebost, O.; Stokke, T. Profound influence of microarray scanner characteristics on gene expression ratios: analysis and procedure for correction. *BMC Genomics* **2004**, *5* (1), 10.
(28) Bengtsson, H.; Jonsson, G.; Vallon-Christersson, J. Calibration and assessment of channel-specific biases in microarray data with extended dynamical range. *BMC Bioinf.* **2004**, *5* (1), 177.
(29) Shippy, R.; Fulmer-Smentek, S.; Jensen, R. V.; Jones, W. D.; Wolber, P. K.; Johnson, C. D.; Pine, P. S.; Boysen, C.; Guo, X.; Chudin, E.; Sun, Y. A.; Willey, J. C.; Thierry-Mieg, J.; Thierry-Mieg, D.; Setterquist, R. A.; Wilson, M.; Lucas, A. B.; Novoradovskaya, N.; Papallo, A.; Turpaz, Y.; Baker, S. C.; Warrington, J. A.; Shi, L.; Herman, D. Using RNA sample titrations to assess microarray platform performance and normalization techniques. *Nat. Biotechnol.* **2006**, *24* (9), 1123–1131.
(30) Quackenbush, J. Microarray data normalization and transformation. *Nat. Genet.* **2002**, *32*, 496–501.
(31) Bolstad, B.; Irizarry, R.; Astrand, M.; Speed, T. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **2003**, *19* (2), 185–193.
(32) Smyth, G. K. Limma: linear models for microarray data. In: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*; Spinger: New York, 2005; pp 397−420.
(33) Hampel, F. R.; Ronchetti, E. M.; Rousseeuw, P. J.; Stahel, W. A. *Robust Statistics: The Approach Based on Influence Functions*; John Wiley and Sons: New York, NY, 1986.

(34) Yang, Y. H.; Dudoit, S.; Luu, P.; Lin, D. M.; Peng, V.; Ngai, J.; Speed, T. P. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **2002**, *30* (4), e15.
(35) Kepler, T.; Crosby, L.; Morgan, K. Normalization and analysis of DNA microarray data by self-consistency and local regression. *Genome Biol.* **2002**, *3* (7), 0037.1−0037.12.
(36) Bolstad, B. M. Low Level Analysis of High-density Oligonucleotide Array Data: Background, Normalization and Summarization. Ph.D. Thesis, University of California, Berkeley, CA, 2004.
(37) Venables, W. N.; Ripley, B. D. *Modern Applied Statistics with S*; Springer: New York, 2002.
(38) Hudson, M. E.; Pozdnyakova, I.; Haines, K.; Mor, G.; Snyder, M. Identification of differentially expressed proteins in ovarian cancer using high-density protein microarrays. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104* (44), 17494–17499.
(39) Marina, O.; Biernacki, M. A.; Brusic, V.; Wu, C. J. A concentration-dependent analysis method for high density protein microarrays. *J. Proteome Res.* **2008**, *7* (5), 2059–2068.
(40) R Development Core Team. R: A Language and Environment for Statistical Computing [Internet] **2007** Vienna, Austria: R Foundation for Statistical Computing; [accessed 2009 May 7] Available from: http://www.R-project.org.
(41) Bhattacharjee, A.; Richards, W. G.; Staunton, J.; Li, C.; Monti, S.; Vasa, P.; Ladd, C.; Beheshti, J.; Bueno, R.; Gillette, M.; Loda, M.; Weber, G.; Mark, E. J.; Lander, E. S.; Wong, W.; Johnson, B. E.; Golub, T. R.; Sugarbaker, D. J.; Meyerson, M. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (24), 13790–13795.
(42) Zhu, X.; Gerstein, M.; Snyder, M. ProCAT: a data analysis approach for protein microarrays. *Genome Biol.* **2006**, *7* (11), R110.
(43) Hamelinck, D.; Zhou, H.; Li, L.; Verweij, C.; Dillon, D.; Feng, Z.; Costa, J.; Haab, B. B. Optimized normalization for antibody microarrays and application to serum-protein profiling. *Mol. Cell. Proteomics* **2005**, *4* (6), 773–784.
(44) Sundaresh, S.; Doolan, D. L.; Hirst, S.; Mu, Y.; Unal, B.; Davies, D. H.; Felgner, P. L.; Baldi, P. Identification of humoral immune responses in protein microarrays using DNA microarray data analysis techniques. *Bioinformatics* **2006**, *22* (14), 1760–1766.
(45) Sundaresh, S.; Randall, A.; Unal, B.; Petersen, J. M.; Belisle, J. T.; Hartley, M. G.; Duffield, M.; Titball, R. W.; Davies, D. H.; Felgner, P. L.; Baldi, P. From protein microarrays to diagnostic antigen discovery: a study of the pathogen Francisella tularensis. *Bioinformatics* **2007**, *23* (13), i508–518.
(46) Huber, W.; von Heydebreck, A.; Sültmann, H.; Poustka, A.; Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **2002**, *18 Suppl 1*, S96–104.

PR900412K