Optimized Normalization for Antibody Microarrays and Application to Serum-Protein Profiling*[S]

Darren Hamelinck‡§, Heping Zhou‡§¶, Lin Li‖, Cornelius Verweij**, Deborah Dillon‡‡, Ziding Feng‖, Jose Costa§§, and Brian B. Haab‡¶¶

The measurements of coordinated patterns of protein abundance using antibody microarrays could be used to gain insight into disease biology and to probe the use of combinations of proteins for disease classification. The correct use and interpretation of antibody microarray data requires proper normalization of the data, which has not yet been systematically studied. Therefore we undertook a study to determine the optimal normalization of data from antibody microarray profiling of proteins in human serum specimens. Forty-three serum samples collected from patients with pancreatic cancer and from control subjects were probed in triplicate on microarrays containing 48 different antibodies, using a direct labeling, two-color comparative fluorescence detection format. Seven different normalization methods representing major classes of normalization for antibody microarray data were compared by their effects on reproducibility, accuracy, and trends in the data set. Normalization with ELISA-determined concentrations of IgM resulted in the most accurate, reproducible, and reliable data. The other normalization methods were deficient in at least one of the criteria. Multiparametric classification of the samples based on the combined measurement of seven of the proteins demonstrated the potential for increased classification accuracy compared with the use of individual measurements. This study establishes reliable normalization for antibody microarray data, criteria for assessing normalization performance, and the capability of antibody microarrays for serum-protein profiling and multiparametric sample classification. ***Molecular & Cellular Proteomics 4:773–784, 2005.***

Antibody microarrays may be very effective for the discovery and application of molecular diagnostics tools. A key feature of microarrays is multiplexing, the ability to measure multiple proteins in low volumes, consuming small amounts of both precious clinical samples and expensive antibodies. Because many antibodies can be tested in parallel, many candidate molecular markers may be efficiently screened. In addition, the relationships between multiple analytes may be observed so that the use of combinations of proteins in disease diagnostics may be assessed. The use of combinations of proteins for disease diagnostics may produce fewer false positive and false negative results relative to tests based on single proteins. Antibody microarrays can be run efficiently in parallel, enabling studies on the large populations of samples that are necessary for marker discovery and validation. Additional advantages of the platform are good reproducibility, high sensitivity, and quantitative accuracy over large concentration ranges (1). We and others have demonstrated the experimental feasibility of antibody and protein microarrays in applications such as protein profiling of cancer tissue (2, 3), autoimmune diagnostics (4), protein interaction screening (5–8), and antibody-based detection of multiple antigens (1, 9–12).

The routine application of antibody microarrays to biological and marker-based research requires establishing optimized experimental and analysis methods. Experimental optimization can help to improve the accuracy and reproducibility of measurements, but the analysis methods must be properly developed and applied to ensure the proper interpretation of the data. A common data processing procedure applied to microarray data is normalization, which adjusts the data from each microarray to account for possible systematic experimental variation in factors such as sample labeling efficiency, scanner readout efficiency, and microarray quality (13, 14). Several normalization procedures have been developed for DNA microarrays. A method developed early on for DNA microarrays is global normalization, which normalizes each array by the median or mean of the intensity log ratios on the array. Other normalization methods purport to correct for systematic errors that may affect arrays non-globally when not all of the spots on an array have the same bias. Intensity-dependent normalization adjusts two-color microarray ratios to account for intensity-based bias in the ratios (14, 15) either linearly or non-linearly. Print tip normalization has been used

From the ‡The Van Andel Research Institute, Grand Rapids, Michigan 49503, ‖Public Health Science Division, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, **Dept. of Molecular and Cell Biology, The University of Amsterdam, 1081 BT Amsterdam, The Netherlands, ‡‡Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, and §§Dept. of Pathology, Yale University School of Medicine, New Haven, Connecticut 06510

Received November 17, 2004, and in revised form, February 7, 2005

Published, MCP Papers in Press, March 25, 2005, DOI 10.1074/mcp.M400180-MCP200

to account for bias caused by variation associated with differences in the tips used for printing (14). Scale normalization makes the assumption that the spread of the distribution of log ratios should be the same for all print tip groups (14). Statistical regression models of microarray data also have been developed for normalization (16).

The two-color comparative fluorescence detection method that we and others have used for antibody microarray experiments is similar to the two-color labeling strategies used for cDNA microarray experiments, so the normalization methods for cDNA arrays may be useful for two-color antibody microarray experiments. However, the differences in antibody microarray experiments, such as a smaller and more selected set of targets and a different labeling method, may mean that the optimal normalization methods may be different. A systematic, detailed comparison and evaluation of the various normalization options for antibody microarrays have not yet been performed. Given the importance of this procedure for subsequent data analysis and interpretation, an in-depth analysis of normalization methods for antibody microarray data is necessary before performing large scale biomarker studies. Therefore, we conducted studies to evaluate various normalization methods for antibody microarray data. Three replicate sets of antibody microarray measurements from serum samples of patients with pancreatic cancer and of control subjects were acquired, and we evaluated seven different normalization methods. The methods represented a variety of major classes of normalization types. Modifications of these types exist, but by evaluating representative methods from a range of classes we could broadly survey the effects of normalization on the data. Newly developed methods with computations requiring special software were not tested.

Each normalization method makes use of assumptions of how "correct" data should behave, and the comparison and evaluation of normalization methods must be independent of those assumptions. Previous comparisons of DNA normalization methods have used the criteria of reproducibility between replicate data sets (13, 17, 18), the linearity of signals from spiked-in standards (17), and the levels of biases in simulated data (13). In this study, we examined several different parameters to get a broad picture of the affects of normalization. The criteria for evaluating and comparing the methods were reproducibility between replicate data sets, accuracy in comparison with known values, and the integrity of overall trends in the data sets. In addition, using optimally normalized data, we investigated the potential benefit of using combinations of measurements for the classification of the samples.

MATERIALS AND METHODS

Serum Samples

Serum samples were collected from 43 patients. Sixteen samples from patients with pancreatic adenocarcinoma and 11 samples from patients with benign diseases, such as ampullary adenoma, pancreatitis, cystadenoma, pseudocyst, or diverticulosis, were collected at the Yale University School of Medicine. Sixteen control samples from healthy persons were collected at the VU University Medical Center in Amsterdam, The Netherlands. All samples were stored frozen in aliquots at −80 °C, and each aliquot had been thawed no more than twice before use. All samples were collected under protocols approved by local Institutional Review Boards for human subjects research.

Antibodies, ELISA, and Protein Concentration Measurements

Antibodies were purchased from various sources (see Supplemental Table I for the complete list of antibodies and sources). Antibodies that were supplied in ascites fluid or antisera were purified using protein A beads (Affi-gel Protein A MAPS kit, Bio-Rad) according to the manufacturer's protocol. The antibodies were prepared at concentrations of 100–1000 μg/ml in 1× phosphate-buffered saline. Two antibodies targeting HGF and one antibody targeting MUC-1 were kindly contributed to the project (Drs. Brian Cao and Ilan Tsarfaty, respectively). ELISAs were performed using commercially available kits from Bethyl Corporation (Montgomery, TX) for the detection of hemoglobin, IgM, IgG, IgA, transferrin, and albumin, and from Cedarlane Laboratories (Hornby, ON, Canada) for the detection of von Willebrand factor. Total serum protein concentrations were measured in duplicate using the BCA assay kit (Pierce).

Fabrication of Antibody Microarrays

Antibodies were deposited eight times each onto slides coated with a polyacrylamide hydrogel (HydroGel, PerkinElmer Life Sciences) using a high-throughput, custom built contact arrayer. Before printing, the hydrogel-coated slides were hydrated for 10 min each in three changes of purified water, dried by centrifugation, and incubated at 40 °C for 20 min. Each printed microarray was circumscribed using a hydrophobic marker. The slides were incubated overnight at room temperature in a humidified chamber to induce binding of the antibodies to the hydrogel matrix. They were washed for 30 s, 3 min, and 30 min in 1× phosphate-buffered saline/0.5% Tween 20, blocked for 1 h at room temperature in 1% BSA/phosphate-buffered saline/0.5% Tween 20, and washed briefly two times in phosphate-buffered saline/0.5% Tween 20 before use.

Sample Labeling

An aliquot from each of the 43 serum samples was labeled with *N*-hydroxysuccinimide-Cy3 (Amersham Biosciences), and another aliquot was labeled with *N*-hydroxysuccinimide-Cy5 (Amersham Biosciences). Each serum aliquot was diluted 1:20 into a 200 mM carbonate buffer at pH 8.3, and a twentieth volume of 6.7 mM *N*-hydroxysuccinimide-Cy3 or -Cy5 in Me_2SO was added. This labeling mix gave approximately a 5–10-fold molar excess of dye relative to the serum proteins (assuming an average serum protein molecular mass of 70 kDa). The concentrations, time, and pH of the labeling reaction were designed to label each serum protein thoroughly but not to completion in case overlabeling of certain proteins might interfere with antibody-antigen interactions. The carbonate buffer contained 1.5 μg/ml BSA labeled with 2,4-dinitrophenol (DNP),[1] as a normalization spike-in. After the reactions proceeded for 2 h on ice, a twentieth volume of 1 M Tris-HCl, pH 8.0, was added to each tube to quench the reactions, and the solutions were allowed to sit for another 20 min. The unreacted dye was removed by passing each solution through a size exclusion chromatography spin column (Bio-Spin P6, Bio-Rad) under centrifugation at 1000 × *g* for 2 min. The

[1] The abbreviations used are: DNP, 2,4-dinitrophenol; CV, coefficients of variation.

Cy5-labeled samples were pooled, and equal amounts of the pool were transferred to each of the Cy3-labeled samples. Each dye-labeled protein solution was supplemented with nonfat milk to a final concentration of 3%, Tween 20 to a final concentration of 0.1%, and 1× phosphate-buffered saline to yield a final serum dilution of 1:100.

Processing of Antibody Microarrays

100 μl of each labeled serum sample mix was incubated on a microarray with gentle rocking at room temperature for 2 h. After incubation, the slides were rinsed briefly in 1× phosphate-buffered saline with 0.1% Tween 20 to remove the unbound sample and then subsequently washed three times for 10 min each in 1× phosphate-buffered saline with 0.1% Tween 20. The slides were spun dry before scanning for fluorescence at 543 and 633 nm using a microarray scanner (ScanArray Express HT, PerkinElmer Life Sciences).

Data Analysis

The software program GenePix Pro 5.0 (Axon Instruments, Foster City, CA) was used to quantify the image data. An intensity threshold for each antibody spot was calculated by the formula $3 \times B \times CV_b$, where B is the median local background of each spot, and CV_b is the average coefficient of variation (S.D. divided by the average) of all the local backgrounds on the array. Spots that either did not surpass the intensity threshold in both color channels, had a regression coefficient (calculated between the pixels of the two color channels) of less than 0.3, or had more than 50% of the pixels saturated in either color channel were excluded from analysis. The ratio of background-subtracted, median sample-specific fluorescence to background-subtracted, median reference-specific fluorescence was calculated, and the ratios from replicate antibody measurements within the same array were averaged using the geometric mean.

Hierarchical clustering and visualization were performed using the programs Cluster and Treeview (see rana.lbl.gov). Ratios were log transformed (base 2) and median centered by genes. Antibodies that did not have measurements in at least 80% of the samples were removed from the clusters.

Normalization Methods

Multiple normalization methods were applied to the microarray data. The details of each are given below.

DNP—The averaged ratios were multiplied by a normalization factor N for each array that was calculated by $N = 1/R_{DNP}$, where R_{DNP} is the average ratio of the replicate anti-DNP antibody spots on the array.

IgM-ELISA—The averaged ratios were multiplied by a normalization factor N for each array that was calculated by $N = (S_{IgM}/\mu_{IgM})/R_{IgM}$, where R_{IgM} is the average ratio of the replicate anti-IgM antibody spots on the array, S_{IgM} is the ELISA-measured IgM concentration of the serum sample on that array, and μ_{IgM} is the mean ELISA-measured IgM concentration of all of the samples.

IgM Set to 1—The averaged ratios were multiplied by a normalization factor N for each array that was calculated by $N = 1/R_{IgM}$, where R_{IgM} is the average ratio of the replicate anti-IgM antibody spots on the array.

Mean Centering—The averaged ratios were multiplied by a normalization factor N for each array that was calculated by $N = 1/\mu$, where μ is the mean ratio of all of the antibody spots on the array.

Loess—For each array, the log-transformed ratios of the antibody measurements were plotted with respect to the average intensities of the spots (averaged over both the 543 and 633 channels), and a regression line was fit as implemented by the marray package for the R environment (19). The ratios of the individual spots were adjusted so that the regression line centered around zero (14).

Loess/IgM-ELISA—The averaged ratios were first processed using the Loess method described above. The resulting array values were normalized by the IgM-ELISA method as described above.

RESULTS

Reproducible, Accurate Serum Protein Profiling Using Antibody Microarrays—Forty-three different serum samples (16 from patients with pancreatic cancer, 11 from patients with other types of gastrointestinal diseases, and 16 from healthy persons) were incubated on a microarray containing 48 different antibodies targeting known serum proteins and putative cancer markers. The amount of protein binding to each antibody from each serum sample was measured relative to the protein binding from a common reference pool using two-color comparative fluorescence (1, 20). The set of 43 samples was performed in triplicate, each set performed on a different day, using batches of microarrays that had been printed on different days. The antibodies on the microarrays and a summary of the performance of each are presented in Supplemental Table I.

The variation across the samples in relative protein binding to the antibodies can be visualized using hierarchical clustering (21). The non-normalized ratios of sample-specific to reference-specific fluorescence from the three replicate experiment sets were grouped and clustered (Fig. 1). Each of the columns contains the data for a given serum sample over all the antibodies and replicate experiment sets, and each row contains the measurements from a given antibody in a given experiment set. The cluster includes only antibodies that gave reproducible measurements between at least two of the three data sets, defined by a 99% confidence threshold for correlation (22). Of the 48 antibodies, 29 surpassed the threshold using non-normalized data. The clustering algorithm ordered the rows and columns by similarity, placing similar rows or columns close to each other. Many replicate antibody measurements, such as those from anti-complement C4 and C3, anti-alkaline phosphatase, and anti-von Willebrand factor cluster immediately adjacent to each other, showing the high reproducibility and distinct profiles of those measurements. Other replicate antibody measurements, such as those from anti-α1-antitrypsin and anti-vascular endothelial growth factor, are more scattered, showing lower reproducibility or less distinction from the other profiles.

Independently collected ELISA measurements from seven of the proteins were included in the cluster and allowed to cluster by similarity to the microarray measurements. Each ELISA measurement set clustered immediately adjacent to its corresponding microarray measurements. This agreement between the ELISA and microarray measurements provided validation of the specificity and accuracy of the microarray measurements for those proteins.

Evaluation of Normalization Methods—We evaluated the effects of seven different normalization procedures on measurement reproducibility, measurement accuracy, and trends

FIG. 1. **Two-way hierarchical clustering of non-normalized microarray data.** Microarray data from 43 serum samples (*horizontal axis*) and the combined antibodies from all three experiment sets (*vertical axis*) were clustered. Each *colored square* represents one antibody measurement from one array, and the color and intensity of each square represents the relative expression level: *red*, high; *green*, low; *black*, medium; *gray*, no data. The color of the column labels indicates the clinical category of the patient from which the serum was taken: *red*, pancreatic adenocarcinoma; *blue*, benign and other gastrointestinal diseases; *green*, healthy. Independently collected ELISA measurements are included for the proteins IgM, IgG, IgA, hemoglobin, von Willebrand factor, albumin, and transferrin (*row labels colored red*). The total protein concentration of each serum sample was also included in the cluster.

in the data sets. "DNP" normalizes each array by setting values from a spiked-in standard (DNP-labeled BSA in this case) to a fixed value; "IgM-ELISA" normalizes each array by setting an internal standard (IgM in this case) to the standard's known values (from ELISA); "IgM set to 1" normalizes each array by setting an internal standard (IgM in this case) to a fixed value; "Mean centering" sets the mean of the ratios in each array to a fixed value; "Loess" uses intensity-based correction to account for biases in the data that may arise from non-linearity in the ratios at certain intensities (14); and "Loess/IgM-ELISA" uses intensity-based correction followed by normalization to the known values of an internal standard (using the IgM-ELISA method in this case). Each of the methods except for the Loess methods corrects for factors that affect the arrays globally, such as labeling or scanner effects. Print tip-based methods were not tested because each antibody was printed by all the tips, and the replicate spots were averaged.

Properly normalized data should reduce variability caused by systematic noise between experiments. The effect of normalization on the reproducibility between replicate data sets was evaluated by examining both the coefficients of variation (CV) and the correlations between the replicate experiments. The CV of each antibody (S.D. divided by average) between the triplicate measurements from each serum sample was calculated for each normalization method. The average CVs for each antibody were compared between the non-normalized data and each set of normalized data using a two-tailed, paired *t* test (Table I). The average CVs ranged from 0.16

TABLE I
CV analysis

The standard deviation was divided by the average of the three replicate experiments for each antibody and each serum sample after normalization by each method. Column 2 gives the average CVs for each method. Column 3 gives the *p* value of the comparison between each method and the non-normalized data. Columns 4 and 5 give the number of antibodies with statistically ($p < 0.05$) lower or higher CVs, respectively, than the non-normalized data.

Normalization method	Average CV	*p* Value	No. of antibodies with	
			Lower CV	Higher CV
Non-normalized	0.18			
IgM-ELISA-normalized	0.17	0.34	10	6
DNP-normalized	0.20	0.01	1	7
IgM Set-to-1-normalized	0.18	0.98	11	3
Mean-centered	0.16	<0.01	16	0
Loess-normalized	0.18	0.55	7	4
Loess/IgM-ELISA-normalized	0.22	<0.01	4	17

TABLE II
Correlation analysis

The Pearson correlation was calculated between measurements from the replicate sets of 43 samples for each antibody and after each normalization method. A pairwise comparison among all three of the experiment sets (1 *versus* 2, 1 *versus* 3, 2 *versus* 3) was performed. Only the antibodies that had correlations over the 99% confidence threshold between at least two of the three data sets were used in the analysis (column 2). The correlations from the normalized data were compared to the correlations from the non-normalized data. The antibodies that had either a higher or lower correlation in comparison with the non-normalized data were counted (columns 3 and 4), and those with a correlation difference greater than 0.1 were also counted (columns 5 and 6).

	Total no. of Ab passed corr.	No. with higher corr.	No. with lower corr.	No. with higher corr. difference >0.1	No. with lower corr. difference >0.1
Set 1 *vs.* Set 2					
Non-normalized					
DNP-normalized	22	19	1	11	0
IgM-ELISA-normalized	22	14	3	7	1
IgM Set-to-1-normalized	22	18	3	11	0
Mean-centered	22	11	8	3	4
Loess-normalized	22	4	16	2	9
Loess/IgM-ELISA-normalized	22	4	15	3	10
Set 1 *vs.* Set 3					
Non-normalized					
DNP-normalized	23	10	3	6	0
IgM-ELISA-normalized	23	15	3	5	1
IgM Set-to-1-normalized	23	19	1	12	0
Mean-centered	23	13	6	2	3
Loess-normalized	23	5	15	1	11
Loess/IgM-ELISA-normalized	23	4	12	2	4
Set 2 *vs.* Set 3					
Non-normalized					
DNP-normalized	27	11	8	4	3
IgM-ELISA-normalized	27	14	3	3	0
IgM Set-to-1-normalized	27	25	0	12	0
Mean-centered	27	7	9	4	5
Loess-normalized	27	2	21	1	18
Loess/IgM-ELISA-normalized	27	5	19	2	14

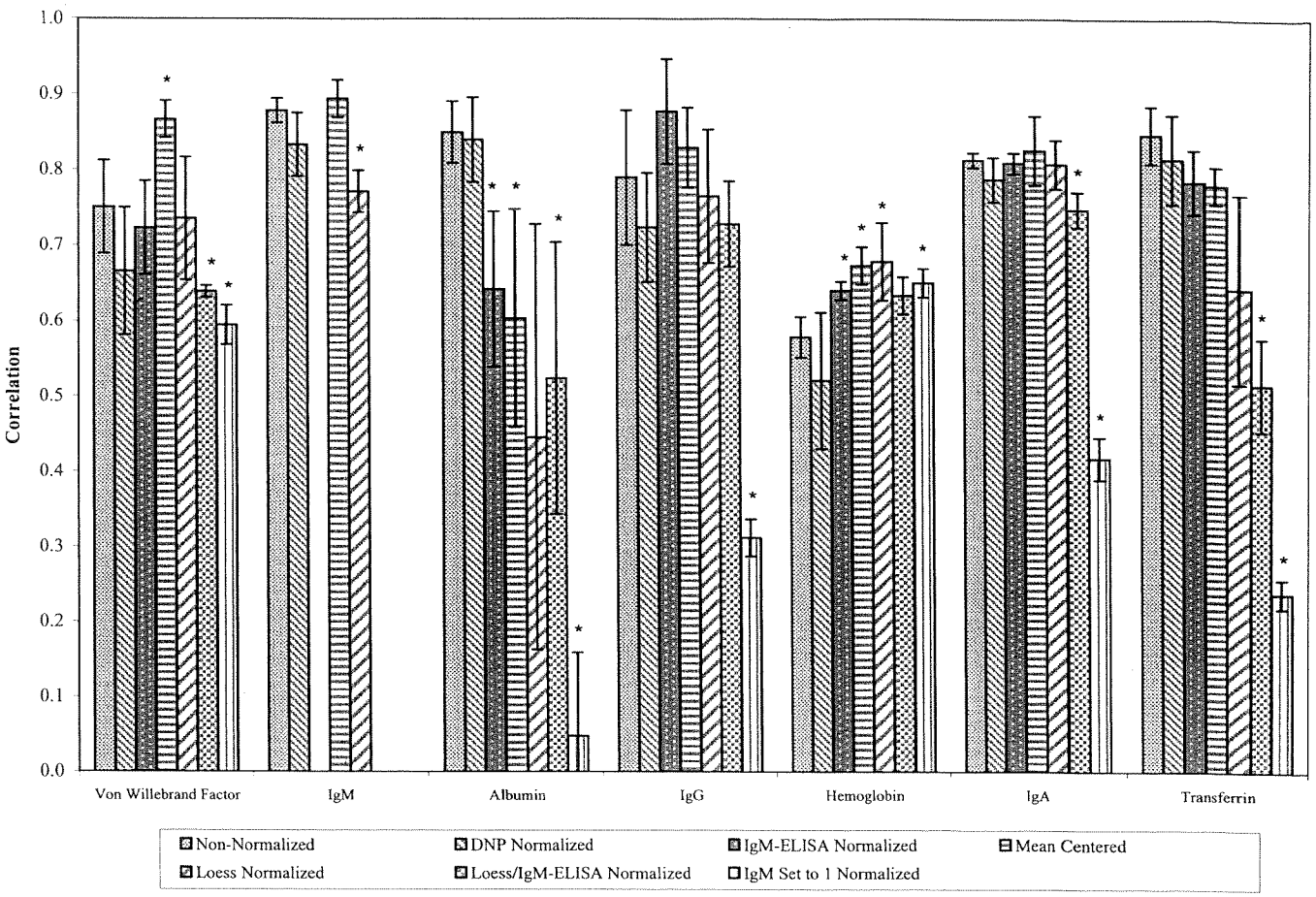

FIG. 2. **Correlation of microarray data with ELISA measurements.** The average correlations (over all three experiment sets) between the microarray data and the ELISA measurements (averaged over three replicates) for each of the seven proteins are shown for each normalization method. The *error bars* represent the S.D. between the correlations of the three replicate experiments. The correlation values of the IgM protein for both normalization methods involving IgM (IgM-ELISA and IgM set to 1) have been excluded because both methods have a correlation of 1 because of the nature of the method used. The *asterisks* refer to a statistically significant difference in correlation between normalized and non-normalized data for a given protein.

(mean centered) to 0.22 (Loess/IgM-ELISA). Normalization by mean centering was the only method that had a significantly lower ($p < 0.05$) average CV in comparison with the non-normalized data (Table I, column 3). Normalization by Loess/IgM-ELISA and DNP resulted in an average CV that was significantly higher than non-normalized. We also counted the number of individual antibodies that had significantly higher or lower CVs in the normalized data compared with the non-normalized data. A two-tailed, paired *t* test was used to compare the CVs between the non-normalized data and each set of normalized data for each antibody. Normalizing by mean centering or by IgM set to 1 produced an abundance of antibodies with a lower CV than in the non-normalized data (Table I, column 4). IgM-ELISA and Loess normalizing had similar numbers of antibodies with higher and lower CVs relative to the non-normalized data. Normalizing by DNP or by Loess/IgM-ELISA yielded a high number of antibodies with a higher CV and a low number with a lower CV than the non-normalized data.

A complementary approach for evaluating reproducibility is to calculate a correlation between duplicate experiment sets. Pearson correlations were calculated between the replicate sets of 43 arrays after normalization by each method for each antibody. The correlations from each of the normalized data sets were compared with the correlations from the non-normalized data (Table II). In the comparison of experiment set 1 with set 2, normalization by DNP, IgM-ELISA, and IgM set to 1 produced many antibodies (19, 14, and 18, respectively) that had higher inter-set correlations and few antibodies (1, 3, and 3, respectively) that had lower correlations than the non-normalized data. Many of the antibodies (11, 7, and 11, respectively) had correlation coefficients that increased by 0.1 or more over the non-normalized data. In contrast, normalization by Loess or Loess/IgM-ELISA resulted in only a few antibodies with higher correlations and many with lower correlations than the non-normalized data. Normalization by mean centering did not seem to significantly increase the antibody correlations in comparison with the non-normalized correlations. The pairwise comparisons between all three of

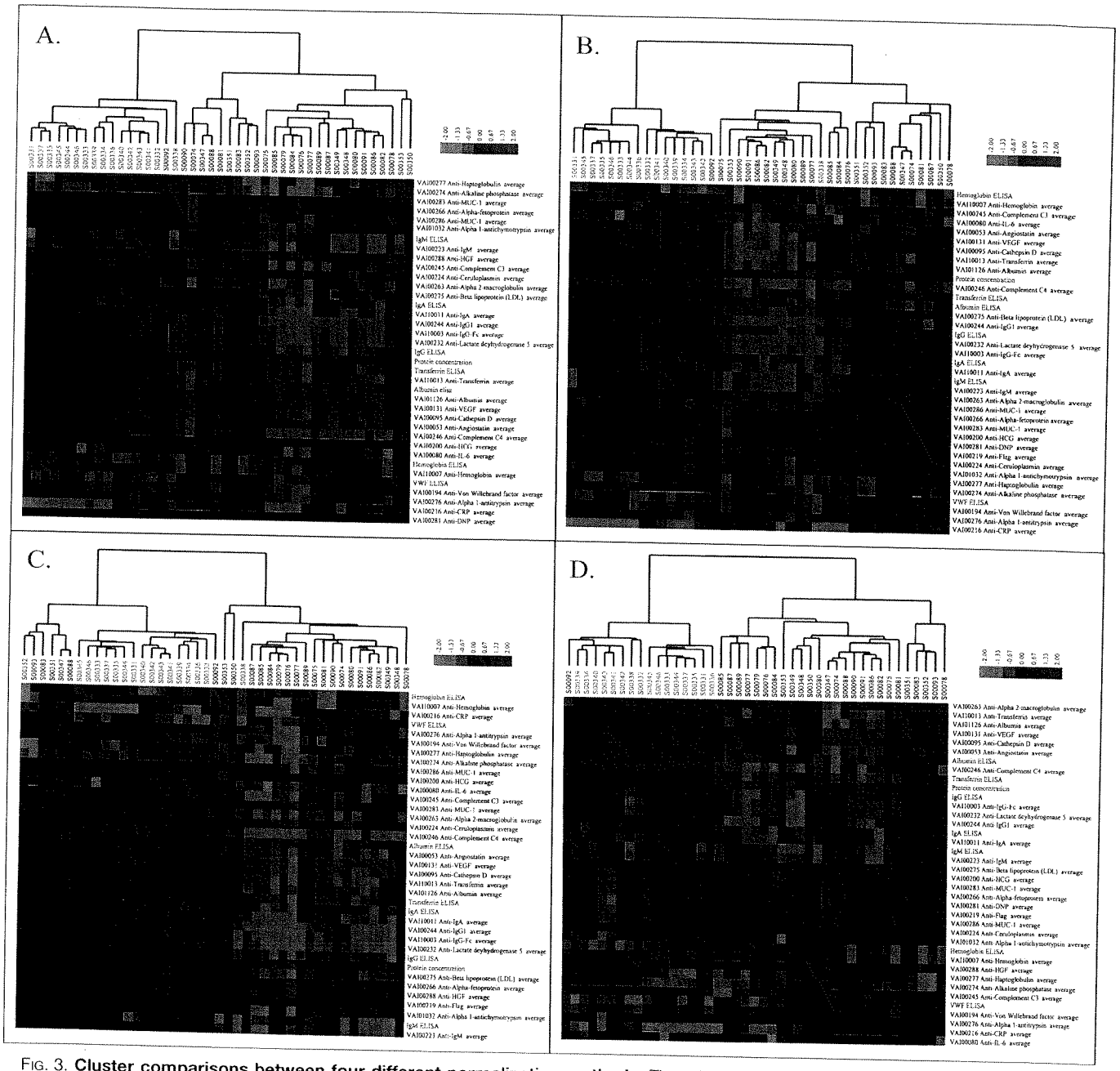

FIG. 3. **Cluster comparisons between four different normalization methods.** The microarray data were normalized using the indicated method, averaged over the three replicate experiments, and clustered. *A*, non-normalized data; *B*, IgM-ELISA-normalized data; *C*, DNP-normalized data; *D*, mean centered normalized data. Only the antibodies that produced reproducible measurements (defined as having correlations over the 99% confidence threshold between at least two of the three data sets) were averaged over the three replicate sets. The *column labels* are *colored* according to the descriptions in Fig. 1.

the experiment sets (1 *versus* 2, 1 *versus* 3, 2 *versus* 3) produced similar results. Taking the two analyses together, the reproducibility of the replicate data seems slightly improved compared with the non-normalized data after normalization by mean centering, IgM set to 1, or IgM-ELISA.

To evaluate the effects of normalization on measurement accuracy, ELISA measurements for seven of the proteins were used as a standard against which the microarray measurements were compared. The correlations between the microarray measurements (averaged over the three replicate sets) and the ELISA measurements over the 43 samples were calculated with each normalization method. The agreement between the microarray and ELISA data for all seven of the normalization methods is depicted in a bar graph (Fig. 2). Normalization by five of the seven methods produced statistically similar correlations, generally between 0.75 and 0.9.

Normalization based on fixing the IgM values to a constant produced consistently lower correlations with the ELISA values for five of the six comparisons. Microarray data adjusted by Loess/IgM-ELISA normalization had a statistically significant reduction in correlation with ELISA values for four of the six antibodies in comparison with the non-normalized data. Based on these comparisons with ELISA measurements, normalization by fixing an internal standard to a constant seems to negatively affect measurement accuracy, and measurement accuracy may be lowered after normalization by Loess.

Finally, we examined the effect of normalization on trends in the data sets to determine whether certain normalization methods altered overall trends more than others. These effects were evaluated by comparing clusters and comparing the proteins that distinguished the sample groups after normalization by each method. Clusters of the averaged data (over the three replicates) from four of the normalization methods (the methods with the best reproducibility and accuracy, as determined in the previous analyses) are shown in Fig. 3, and the other three are available in the supplemental data. The clusters may be inspected to identify the effects of normalization on overall trends in the data sets. In all four clusters, the control samples (*green labels*) cluster together, and the other two classes are intermixed. In all four, the ELISA values (*row labels highlighted red*) also generally cluster adjacent to their respective microarray values. Trends that define the patient groups are fairly consistent between the non-normalized, the IgM-ELISA normalized, and the DNP-normalized data (Fig. 3, *A–C*, respectively), although ordered differently. The non-control samples seem to divide into two or three groups in Fig. 3, *A–C*; one group has high levels of most of the proteins (*e.g.* samples 351, 88, and 93), another group is low in most of the proteins (*e.g.* samples 84, 76, and 77), and another is low in some proteins and high in C-reactive protein, von Willebrand factor, and α1-antichymotrypsin (*e.g.* samples 86, 91, and 82). The above patterns are somewhat altered in the mean centered data because the groups of patients that are high in most proteins or low in most proteins are not seen as they are in the other clusters.

Because a major use for antibody microarray experiments is to identify differences between sample classes, we examined the effect of normalization on the statistical differences between the sample classes. The antibody measurements that distinguished between the sample classes ($p < 0.05$, using data averaged between the three replicate sets) were identified by a two-tailed *t* test, and the number of antibodies that were found in common between any two methods was tabulated (Table III). Using the non-normalized data, 17 antibody measurements differentiated the serum samples from patients with pancreatic cancer and the samples from healthy control subjects. Of those 17 antibodies, 16 were common to the analysis from DNP-normalized data and 15 were common from IgM-ELISA-normalized data. The other four normalization methods had between 3 and 12 common discriminators with the non-normalized data. Therefore, the trends observed after normalization by DNP and IgM-ELISA are similar to each other and similar to the non-normalized data, but normalization by the other methods resulted in a divergent set of antibodies that distinguished the patient groups. Because of the high accuracy and reproducibility of normalization by IgM-ELISA and because of its minimal alteration to the trends in the data sets, this normalization method was used for the analysis below.

TABLE III

Common discriminators between normalization methods

A two-tailed *t* test was used to identify antibodies with statistically different ($p < 0.05$) measurements between the pancreatic cancer and healthy sample sets. The column and row labels indicate the type of normalization method used: 1) non-normalized, 2) DNP, 3) IgM-ELISA, 4) Loess, 5) Loess/IgM-ELISA, 6) mean centered, 7) IgM set-to-1. The values within the table indicate the number of antibodies that were significantly different among the patient classes and that were common among the normalization methods intersecting in the rows and columns.

	1	2	3	4	5	6	7
1	17						
2	16	19					
3	15	14	16				
4	10	10	11	14			
5	12	11	12	10	12		
6	12	12	13	14	10	16	
7	3	4	4	7	3	7	8

Evaluation of Multiparametric Classification—A major benefit of the ability to measure multiple proteins is the evaluation of using combinations of proteins for improved sample classification relative to the use of single proteins. We investigated this application of the data using the IgM-ELISA-normalized data. The data were averaged over the three replicate sets using only the antibodies that produced reproducible measurements (defined as having correlations over the 99% confidence threshold between at least two of the three data sets). Twenty-nine of the 48 antibodies passed the reproducibility threshold. Sixteen antibodies produced measurements with statistically different means ($p < 0.05$) between the samples from patients with pancreatic cancer and those from healthy control subjects. The distributions of the measurements from each of these antibodies in the cancer and control samples are shown in Fig. 4. Three antibodies, anti-von Willebrand factor, anti-α1-antitrypsin, and anti-C-reactive protein, showed higher binding in the pancreatic cancer samples, and the rest showed lower binding in the pancreatic cancer samples; anti-albumin, anti-transferrin, and anti-complement C3 showed the greatest significance. Fig. 4 shows that all of the distributions overlapped significantly between the cancer and control groups.

The potential benefit from using combined antibodies to classify the samples was examined using a generic version of the Real Boosting algorithm (23, 24). Boosting is an appealing

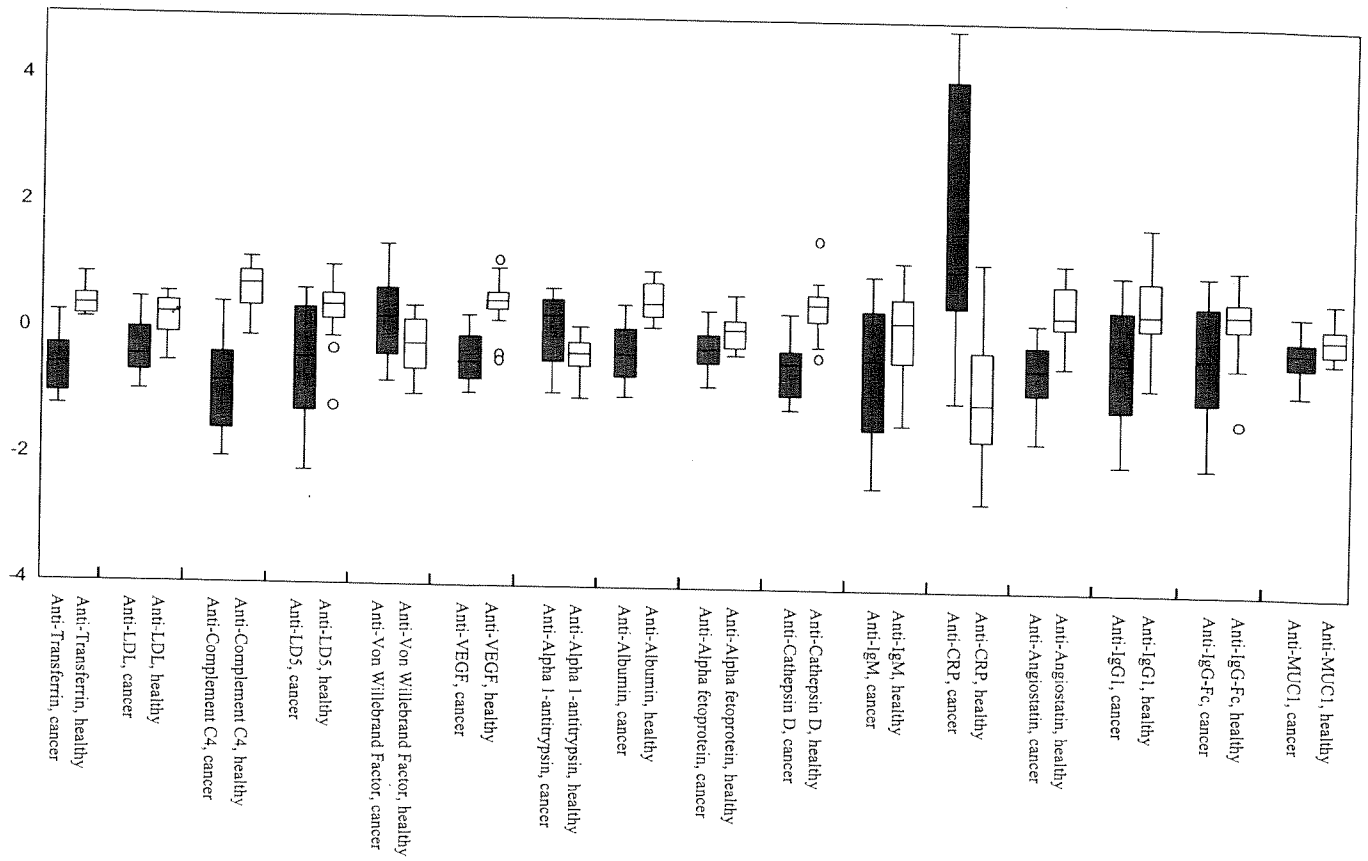

FIG. 4. **Distributions of measurements for the discriminating antibodies.** The distributions of measurements that were statistically different ($p < 0.05$) between the samples from pancreatic cancer patients (*dark boxes*) and healthy controls (*light boxes*) are shown. The *boxes* give the upper and lower quartiles of the measurements with respect to the median value (*horizontal line in each box*). The *lines* give the ranges of the measurements, excluding outliers, which are represented by the *circles*.

method for combining "weak" individual classifiers into a "powerful" combined (or summary) classifier, and it was shown empirically to be resistant to overfitting (23), a problem often encountered by other widely used classification algorithms. A 10-fold cross-validation determined the optimal number of antibodies in a final combined classifier. 90% of the samples were used as a training set to define a model for classification, whereas 10% of the samples were reserved as a testing set to determine the error rate of the model. This process was repeated 10 times, each time using a different group of 90% for classification. The classifier was considered final when further addition of an antibody into the model increased the cross-validation error. The cross-validation process simulates the uncertainty in classification of future samples and estimates the prediction error of the selected combined classifier from the current data set. This validation gives extra protection against the chance of overfitting.

Real boosting was used to classify each of the cancer and control samples as belonging to one of the two classes. The antibodies used in the multiparametric classification and the performance of the classifier is presented in Table IV. The errors, sensitivities, and specificities (columns 2–4) were averaged over the 10 cross-validation iterations. Each row gives the cumulative result after using the antibodies at and above that row. When using all seven antibody measurements (one antibody was used twice), the cumulative results are 0.00 error, 1.00 sensitivity, and 1.00 specificity, or perfect classification, which was not possible using any single antibody.

DISCUSSION

The major goal of this work was to determine the optimal normalization procedure for antibody microarray data. This work was necessary because no systematic evaluation of normalization methods for antibody microarray experiments had yet been performed. Our thorough analysis revealed some differences between the normalization methods in the resulting data quality, and a comparison of the methods provided valuable information about which normalization method is optimal. Our study also was valuable to develop an objective approach for assessing and comparing normalization methods.

Some of the normalization methods performed well in one category but not very well in another, showing the value of using multiple criteria for the evaluation. For example, normal-

TABLE IV
Sample classification using multiple proteins

A 'real boosting' algorithm was used to identify a linear combination of protein classifiers that could distinguish pancreatic cancer from control samples using the IgM-ELISA-normalized data averaged over the three replicates. The error rate, sensitivity (the percentage of true positives over all positives), and specificity (the percentage of true negatives over all negatives) were calculated from cross-validation analysis as described in the text. Antibodies with measurements that were higher in pancreatic cancer relative to control samples are indicated by bold type.

Antibody	Error	Sensitivity	Specificity
Anti-complement C4	0.08	0.90	0.95
Anti-C-reactive protein	0.11	0.85	0.95
Anti-transferrin	0.06	0.95	0.95
Anti-complement C3	0.08	0.90	0.95
Anti-complement C4	0.05	0.95	0.95
Anti-α1-antichymotrypsin	0.02	1.00	0.95
Anti-cathepsin D	0.02	1.00	0.95
Anti-α1-antitrypsin	0.00	1.00	1.00

izing by IgM set to 1 yielded more reproducible data than the non-normalized data. However, the accuracy of the measurements was greatly compromised as determined by comparison with the ELISA values. Because IgM concentrations vary from sample to sample, setting this value to a constant is not valid and reduces the accuracy of the measurements. It will probably be impossible to find a "housekeeping" protein in the serum that could be used as a constant reference because it seems that all serum proteins are subject to significant change between individuals as supported by our current analyses. Normalizing by IgM set to 1 also greatly altered the trends in the data sets in comparison with the other normalization methods (Supplemental Fig. 1).

Other methods performed well in some categories but not in others. Loess normalization produced reasonably accurate data, but the reproducibility was lower than the non-normalized data as assessed by the correlation analysis. Applying IgM-ELISA normalization to the Loess-adjusted data did not improve reproducibility. The list of proteins that discriminated cancer from healthy and the general structure of the cluster (Supplemental Fig. 2) also was altered after Loess normalization. The Loess method was developed for DNA microarray data and relies on having a large number of data points to produce an accurate picture of intensity-based biases in the ratios. With fewer data points, as with our data, such adjustments may be erroneous and may actually add noise to the data. Other normalization methods that use regression calculations of trends in data therefore also may not perform well on smaller, more selected arrays. Likewise, scaling methods, which adjust the variances (the spread) in groups of ratios, also might not perform well on these arrays because the variances in small numbers of proteins could legitimately change.

Both DNP and IgM-ELISA normalization had good accuracy, and neither altered the trends in the data sets relative to the non-normalized data, but the reproducibility after IgM-ELISA normalization was slightly higher. Normalization by mean centering also performed well in reproducibility and accuracy, although the normalization seemed to alter trends in the data more than normalization by DNP or IgM-ELISA. Normalization by mean centering is accurate if the average concentration of the measured proteins is constant between samples. Because the average concentration may change, especially if measuring a small number of proteins, normalization by mean centering may occasionally produce results that inaccurately reflect the trends in the data sets.

Therefore, taking all of the information together, the IgM-ELISA normalization method, of the methods evaluated, seems to have performed the best. Normalizing by the known values of an internal standard such as IgM is attractive because these values are inherent to the sample. A spiked-in standard like DNP-labeled BSA is not inherent to the sample, so the standard would not correct for sources of bias that occurred before the standard was introduced. The accuracy of a known standard is independent of the size or selection of the rest of the array, and it makes no assumptions about the behaviors of particular housekeeping proteins. Drawbacks of normalization by an internal standard are that highly accurate ELISA values for that protein must be obtained for every sample and that one relies on the quality of the microarray measurements for that protein.

Further improvements in the normalization method are still possible and necessary for antibody microarrays. The high reproducibility of the mean centered data showed that normalizing by many proteins may be valuable. We are currently investigating variants on the mean centering method. In addition, different spike-in proteins, such as plant or peanut proteins that have no homology to human proteins, may perform better than DNP-labeled BSA. A panel of three or four highly specific spike-in proteins may produce less variable results than the use of a single protein. Other antibody or protein array techniques may have other optimal normalization methods; the methods presented here provide a strategy for determining which is optimal. In addition, other sample types, such as those from tissue or cell culture sources, may behave differently than serum, and the normalization would need to be independently optimized.

In addition to being useful for evaluating normalization methods, these data served the additional purpose of exploring the value of combined measurements for sample classification. Multiple markers may be grouped together to improve diagnostic performance if the markers contribute complementary, non-overlapping discriminatory information. The improvement of the sample classification when using the multiparametric method, compared with the use of single proteins, showed the potential value of antibody microarray data for more accurate diagnostics. This particular classifier is not likely to be specific for pancreatic cancer because most of the proteins used in the classifier had similar distributions between the cancer and other gastrointestinal disease sam-

ples. The development of a specific classifier for pancreatic cancer will require measurements from additional proteins that are more specifically associated with pancreatic cancer. A more sensitive detection method, such as the two-color rolling circle amplification method demonstrated previously (22), would allow the measurement of lower abundance proteins that may contribute to a specific signature for pancreatic cancer. Studies using that approach are ongoing.

No firm conclusions on the nature of specific serum protein alterations in pancreatic cancer can be made from these data because of the small sample size and potential bias between the case and control samples, but the observed differences between the cancer and control samples were consistent with the high levels of inflammation usually associated with pancreatic cancer. The higher levels of C-reactive protein and von Willebrand factor in the disease samples probably reflect a positive acute phase response (25, 26), and a reduction in the levels of albumin and transferrin as observed here is also commonly observed in an acute phase response (27). Decreased levels of serum IgG and IgM have been observed in cancer (28, 29), and higher α1-antitrypsin has also been associated with pancreatic cancer (30).

In summary, this work established reliable methods for normalizing antibody microarray data and established objective criteria for assessing normalization methods. Furthermore, we showed that many different proteins in serum samples can be reliably measured using antibody microarrays and that this capability is useful for multiparametric sample classification. These developments lay the foundation for larger-scale studies that could lead to improved diagnostics for pancreatic cancer and other cancers.

Acknowledgments—We thank Dr. Kyle Furge (Van Andel Research Institute) for analytical support and Drs. Brian Cao (Van Andel Research Institute) and Ilan Tsarfaty (Tel Aviv University) for the contribution of antibodies.

* This research was funded in part by the Early Detection Research Network of the National Cancer Institute and the Michigan Proteome Consortium of the Michigan Life Sciences Corridor, and by the Van Andel Research Institute. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains Supplemental Tables I and II and Supplemental Figures 1 and 2.

§ Both authors contributed equally to this work.

¶ Current address: Seton Hall University, Dept. of Biology, South Orange, NJ 07079.

¶¶ To whom correspondence should be addressed: 333 Bostwick, NE, Grand Rapids, MI 49503. Tel.: 616-234-5268; Fax: 616-234-5269; E-mail: brian.haab@vai.org.

REFERENCES

1. Haab, B. B., Dunham, M. J., and Brown, P. O. (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biol.* **2,** 1–13
2. Knezevic, V., Leethanakul, C., Bichsel, V. E., Worth, J. M., Prabhu, V. V., Gutkind, J. S., Liotta, L. A., Munson, P. J., Petricoin, E. F. I., and Krizman, D. B. (2001) Proteomic profiling of the cancer microenvironment by antibody arrays. *Proteomics* **1,** 1271–1278
3. Sreekumar, A., Nyati, M. K., Varambally, S., Barrette, T. R., Ghosh, D., Lawrence, T. S., and Chinnaiyan, A. M. (2001) Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. *Cancer Res.* **61,** 7585–7593
4. Joos, T. O., Schrenk, M., Hopfl, P., Kroger, K., Chowdhury, U., Stoll, D., Schorner, D., Durr, M., Herick, K., Rupp, S., Sohn, K., and Hammerle, H. (2000) A microarray enzyme-linked immunosorbent assay for autoimmune diagnostics. *Electrophoresis* **21,** 2641–2650
5. MacBeath, G., and Schreiber, S. L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science* **289,** 1760–1763
6. Ge, H. (2000) Upa, a universal protein array system for quantitative detection of protein-protein, protein-DNA, protein-rna and protein-ligand interactions. *Nucleic Acids Res.* **28,** e3
7. Lueking, A., Horn, M., Eickhoff, H., Buessow, K., Lehrach, H., and Walter, G. (1999) Protein microarrays for gene expression and antibody screening. *Anal. Biochem.* **270,** 103–111
8. 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., and Snyder, M. (2001) Global analysis of protein activities using proteome chips. *Science* **293,** 2101–2105
9. deWildt, R. M. T., Mundy, C. R., Gorick, B. D., and Tomlinson, I. M. (2000) Antibody arrays for high-throughput screening of antibody-antigen interactions. *Nat. Biotechnol.* **18,** 989–994
10. Arenkov, P., Kukhtin, A., Gemmell, A., Voloshchuk, S., Chupeeva, V., and Mirzabekov, A. (2000) Protein microchips: use for immunoassay and enzymatic reactions. *Anal. Biochem.* **278,** 123–131
11. Huang, R.-P., Huang, R., Fan, Y., and Lin, Y. (2001) Simultaneous detection of multiple cytokines from conditioned media and patient's sera by an antibody-based protein array system. *Anal. Biochem.* **294,** 55–62
12. Mendoza, L. G., McQuary, P., Mongan, A., Gangadharan, R., Brignac, S., and Eggers, M. (1999) High-throughput microarray-based enzyme-linked immunosorbent assay (ELISA). *BioTechniques* **27,** 778–788
13. Park, T., Yi, S. G., Kang, S. H., Lee, S., Lee, Y. S., and Simon, R. (2003) Evaluation of normalization methods for microarray data. *BMC Bioinformatics* **4,** 33
14. Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30,** e15
15. Workman, C., Jensen, L. J., Jarmer, H., Berka, R., Gautier, L., Nielser, H. B., Saxild, H. H., Nielsen, C., Brunak, S., and Knudsen, S. (2002) A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biol.* **3,** research0048
16. Kepler, T. B., Crosby, L., and Morgan, K. T. (2002) Normalization and analysis of DNA microarray data by self-consistency and local regression. *Genome Biol.* **3,** research0037
17. Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19,** 185–193
18. Yoon, D., Yi, S. G., Kim, J. H., and Park, T. (2004) Two-stage normalization using background intensities in cdna microarray data. *BMC Bioinformatics* **5,** 97
19. Ihaka, R., and Gentleman, R. (1996) R: A language for data analysis and graphics. *J. Comp. Graph. Stat.* **5,** 299–314
20. Miller, J. C., Zhou, H., Kwekel, J., Cavallo, R., Burke, J., Butler, E. B., Teh, B. S., and Haab, B. B. (2003) Antibody microarray profiling of human prostate cancer sera: Antibody screening and identification of potential biomarkers. *Proteomics* **3,** 56–63
21. Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 14863–14868
22. Zhou, H., Bouwman, K., Schotanus, M., Verweij, C., Marrero, J. A., Dillon, D., Costa, J., Lizardi, P. M., and Haab, B. B. (2004) Two-color, rolling-circle amplification on antibody microarrays for sensitive, multiplexed serum-protein measurements. *Genome Biol.* **5,** R28
23. Friedman, J. H., Hastie, T., and Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting. *Ann. Stat.* **28,** 337–407

24. Yasui, Y., Pepe, M., Thompson, M. L., Adam, B. L., Wright, G. L., Jr., Qu, Y., Potter, J. D., Winget, M., Thornquist, M., and Feng, Z. (2003) A data-analytic strategy for protein biomarker discovery: profiling of high-dimensional proteomic data for cancer detection. *Biostatistics* **4,** 449–463
25. Barber, M. D., Ross, J. A., and Fearon, K. C. (1999) Changes in nutritional, functional, and inflammatory markers in advanced pancreatic cancer. *Nutr. Cancer* **35,** 106–110
26. Fearon, K. C., Barber, M. D., Falconer, J. S., McMillan, D. C., Ross, J. A., and Preston, T. (1999) Pancreatic cancer as a model: Inflammatory mediators, acute-phase response, and cancer cachexia. *World J. Surg.* **23,** 584–588
27. Barber, M. D., Ross, J. A., Preston, T., Shenkin, A., and Fearon, K. C. (1999) Fish oil-enriched nutritional supplement attenuates progression of the acute-phase response in weight-losing patients with advanced pancreatic cancer. *J. Nutr.* **129,** 1120–1125
28. Deture, F. A., Deardourff, S. L., Kaufman, H. E., and Centifanto, Y. M. (1978) A comparison of serum immunoglobulins from patients with non-neoplastic prostates and prostatic carcinoma. *J. Urol.* **120,** 435–437
29. Gahankari, D. R., and Goldhar, K. B. (1993) An evaluation of serum and tissue bound immunoglobulins in prostatic diseases. *J. Postgrad. Med.* **39,** 63–67
30. Trachte, A. L., Suthers, S. E., Lerner, M. R., Hanas, J. S., Jupe, E. R., Sienko, A. E., Adesina, A. M., Lightfoot, S. A., Brackett, D. J., and Postier, R. G. (2002) Increased expression of alpha-1-antitrypsin, glutathione S-transferase pi and vascular endothelial growth factor in human pancreatic adenocarcinoma. *Am. J. Surg.* **184,** 642–647, discussion 647–648

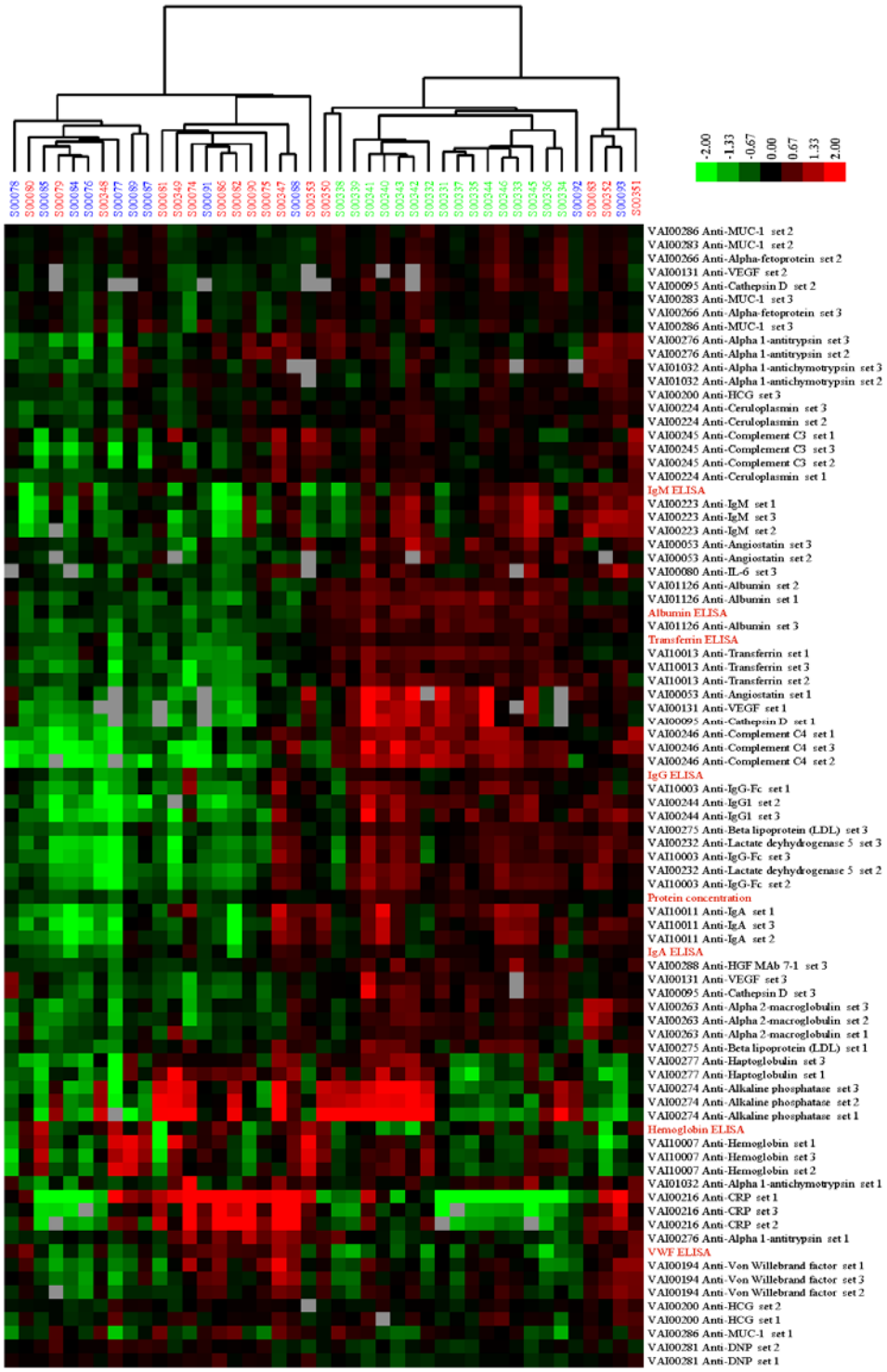
-2.00
-1.33
-0.67
0.00
0.67
1.33
2.00
VAI00286 Anti-MUC-1 set 2
VAI00283 Anti-MUC-1 set 2
VAI00266 Anti-Alpha-fetoprotein set 2
VAI00131 Anti-VEGF set 2
VAI00095 Anti-Cathepsin D set 2
VAI00283 Anti-MUC-1 set 3
VAI00266 Anti-Alpha-fetoprotein set 3
VAI00286 Anti-MUC-1 set 3
VAI00276 Anti-Alpha 1-antitrypsin set 3
VAI00276 Anti-Alpha 1-antitrypsin set 2
VAI01032 Anti-Alpha 1-antichymotrypsin set 3
VAI01032 Anti-Alpha 1-antichymotrypsin set 2
VAI00200 Anti-HCG set 3
VAI00224 Anti-Ceruloplasmin set 3
VAI00224 Anti-Ceruloplasmin set 2
VAI00245 Anti-Complement C3 set 1
VAI00245 Anti-Complement C3 set 3
VAI00245 Anti-Complement C3 set 2
VAI00224 Anti-Ceruloplasmin set 1
IgM ELISA
VAI00223 Anti-IgM set 1
VAI00223 Anti-IgM set 3
VAI00223 Anti-IgM set 2
VAI00053 Anti-Angiostatin set 3
VAI00053 Anti-Angiostatin set 2
VAI00080 Anti-IL-6 set 3
VAI01126 Anti-Albumin set 2
VAI01126 Anti-Albumin set 1
Albumin ELISA
VAI01126 Anti-Albumin set 3
Transferrin ELISA
VAI10013 Anti-Transferrin set 1
VAI10013 Anti-Transferrin set 3
VAI10013 Anti-Transferrin set 2
VAI00053 Anti-Angiostatin set 1
VAI00131 Anti-VEGF set 1
VAI00095 Anti-Cathepsin D set 1
VAI00246 Anti-Complement C4 set 1
VAI00246 Anti-Complement C4 set 3
VAI00246 Anti-Complement C4 set 2
IgG ELISA
VAI10003 Anti-IgG-Fc set 1
VAI00244 Anti-IgG1 set 2
VAI00244 Anti-IgG1 set 3
VAI00275 Anti-Beta lipoprotein (LDL) set 3
VAI00232 Anti-Lactate deyhydrogenase 5 set 3
VAI10003 Anti-IgG-Fc set 3
VAI00232 Anti-Lactate deyhydrogenase 5 set 2
VAI10003 Anti-IgG-Fc set 2
Protein concentration
VAI10011 Anti-IgA set 1
VAI10011 Anti-IgA set 3
VAI10011 Anti-IgA set 2
IgA ELISA
VAI00288 Anti-HGF MAb 7-1 set 3
VAI00131 Anti-VEGF set 3
VAI00095 Anti-Cathepsin D set 3
VAI00263 Anti-Alpha 2-macroglobulin set 3
VAI00263 Anti-Alpha 2-macroglobulin set 2
VAI00263 Anti-Alpha 2-macroglobulin set 1
VAI00275 Anti-Beta lipoprotein (LDL) set 1
VAI00277 Anti-Haptoglobulin set 3
VAI00277 Anti-Haptoglobulin set 1
VAI00274 Anti-Alkaline phosphatase set 3
VAI00274 Anti-Alkaline phosphatase set 2
VAI00274 Anti-Alkaline phosphatase set 1
Hemoglobin ELISA
VAI10007 Anti-Hemoglobin set 1
VAI10007 Anti-Hemoglobin set 3
VAI10007 Anti-Hemoglobin set 2
VAI01032 Anti-Alpha 1-antichymotrypsin set 1
VAI00216 Anti-CRP set 1
VAI00216 Anti-CRP set 3
VAI00216 Anti-CRP set 2
VAI00276 Anti-Alpha 1-antitrypsin set 1
VWF ELISA
VAI00194 Anti-Von Willebrand factor set 1
VAI00194 Anti-Von Willebrand factor set 3
VAI00194 Anti-Von Willebrand factor set 2
VAI00200 Anti-HCG set 2
VAI00200 Anti-HCG set 1
VAI00286 Anti-MUC-1 set 1
VAI00281 Anti-DNP set 2
VAI00281 Anti-DNP set 1

Figure 1

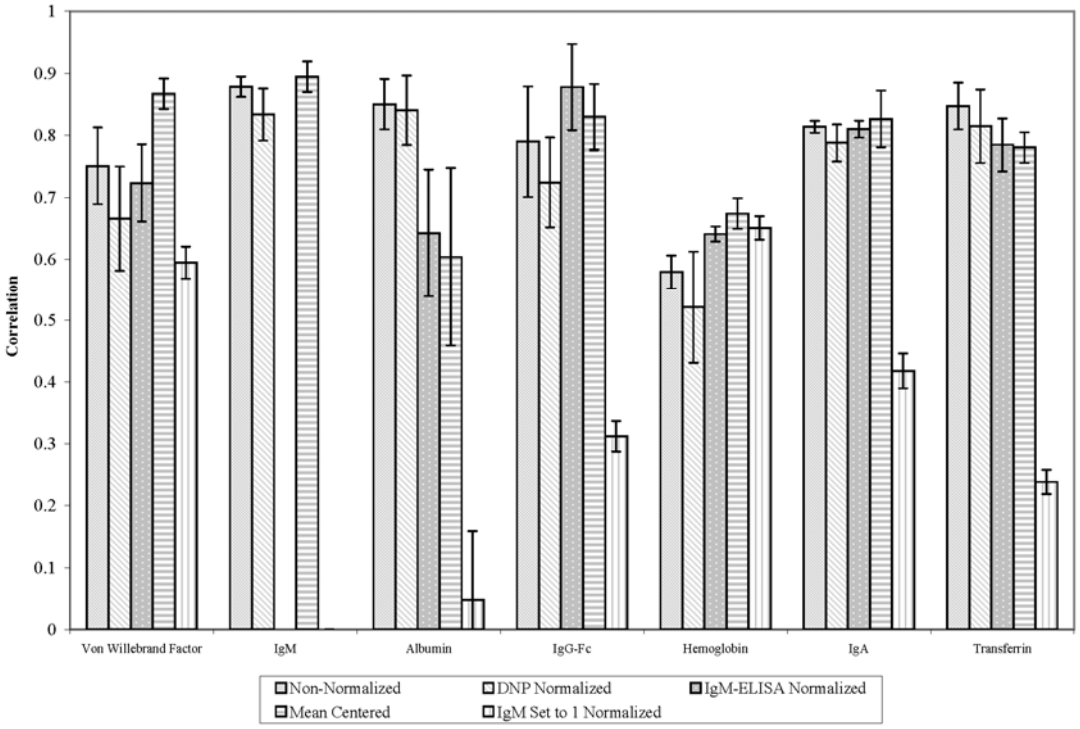
Correlation
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Von Willebrand Factor
IgM
Albumin
IgG-Fc
Hemoglobin
IgA
Transferrin
Non-Normalized
DNP Normalized
IgM-ELISA Normalized
Mean Centered
IgM Set to 1 Normalized

Figure 2

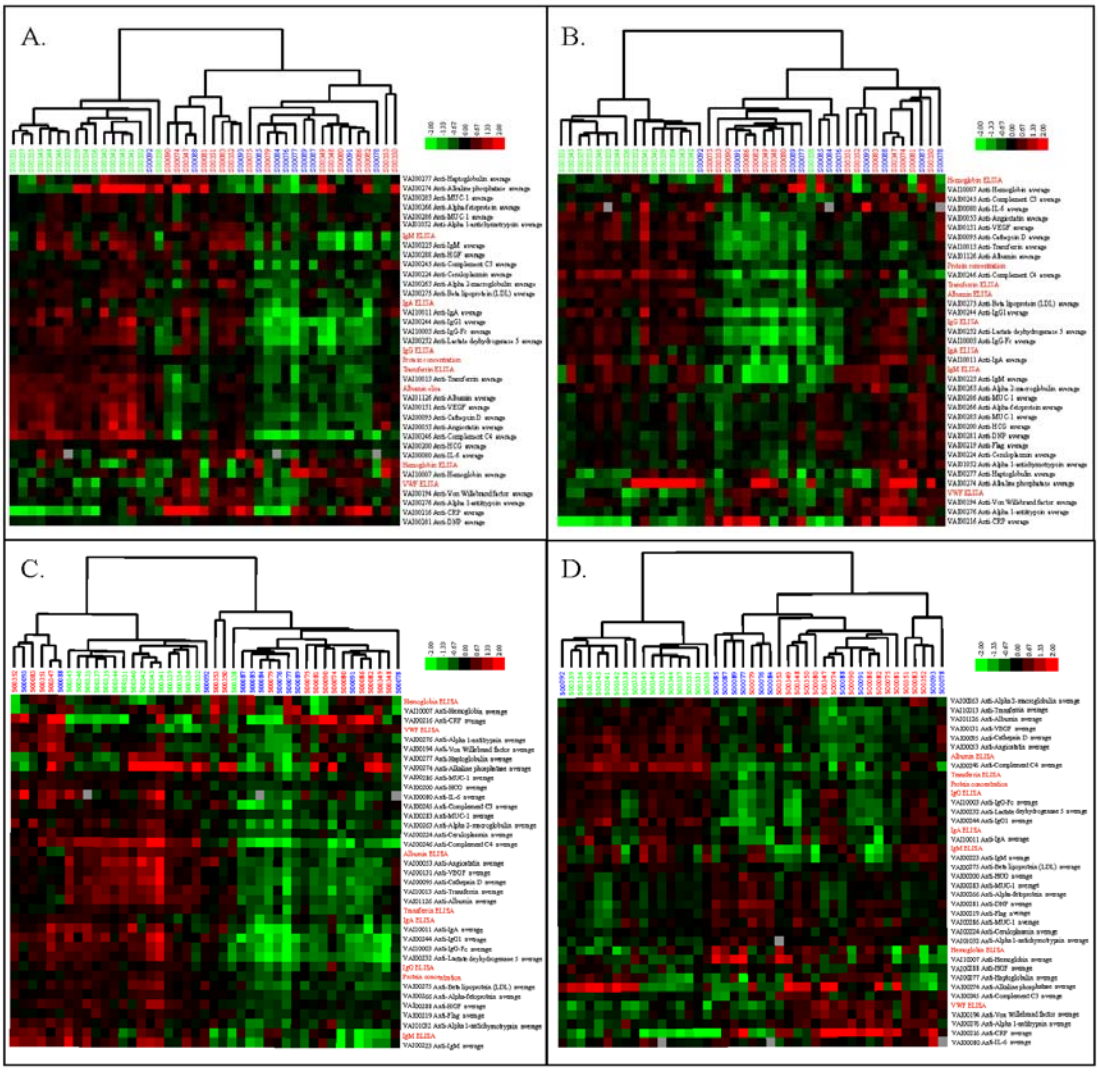

Figure 3

A.

Normalization Method	Avg CV	p-Value	# Ab with Lower CV	# Ab with Higher CV
Non-Normalized	0.18			
IgM-ELISA Normalized	0.17	0.34	10	6
DNP Normalized	0.20	0.01	1	7
IgM Set to 1 Normalized	0.18	0.98	11	3
Mean Centered	0.16	< 0.01	16	0
Loess Normalized	0.18	0.55	7	4
Loess & IgM-ELISA Normalized	0.22	< 0.01	4	17

B.

		Total # of ab Passed Corr.	# with Higher Corr.	# with Lower Corr.	# of ab having Higher Corr. with Difference >0.1	# of ab having Lower Corr. with Difference >0.1	% Passed Corr.
Set 1 vs Set 2	Non-Normalized						44
	DNP Normalized	22	19	1	11	0	61
	IgM-ELISA Normalized	22	14	3	7	1	54
	Loess Normalized	22	4	16	2	9	44
	Loess & IgM-ELISA Normalized	22	4	15	3	10	46
Set 1 vs Set 3	Non-Normalized						58
	DNP Normalized	23	10	3	6	0	58
	IgM-ELISA Normalized	23	15	3	5	1	63
	Loess Normalized	23	5	15	1	11	55
	Loess & IgM-ELISA Normalized	23	4	12	2	4	63
Set 2 vs Set 3	Non-Normalized						57
	DNP Normalized	27	11	8	4	3	63
	IgM-ELISA Normalized	27	14	3	3	0	61
	Loess Normalized	27	2	21	1	18	54
	Loess & IgM-ELISA Normalized	27	5	19	2	14	57

Table 1

	1	2	3	4	5	6	7
1	**17**	16	15	10	12	12	3
2	16	**19**	14	10	11	12	4
3	15	14	**16**	11	12	13	4
4	10	10	11	**14**	10	14	7
5	12	11	12	10	**12**	10	3
6	12	12	13	14	10	**16**	7
7	3	4	4	7	3	7	**8**

Table 2

Antibody	Error	Sensitivity	Specificity
Anti-Complement C4	0.08	0.90	0.95
Anti-CRP	0.11	0.85	0.95
Anti-Transferrin	0.06	0.95	0.95
Anti-Complement C3	0.08	0.90	0.95
Anti-Complement C4	0.05	0.95	0.95
Anti-Alpha 1-antichymotrypsin	0.02	1.00	0.95
Anti-Cathepsin D	0.02	1.00	0.95
Anti-Alpha 1-antitrypsin	0.00	1.00	1.00

Table 3