

Gene expression

Analysis of dose–response effects on gene expression data with comparison of two microarray platformsJianhua Hu¹, Mini Kapoor², Wei Zhang³, Stanley R. Hamilton³ and Kevin R. Coombes^{1,*}¹Department of Biostatistics and Applied Mathematics, ²Department of Cancer Genetics and ³Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

Received on February 7, 2005; revised on May 22, 2005; accepted on July 18, 2005

Advance Access publication August 4, 2005

ABSTRACT

Motivation: The problems of analyzing dose effects on gene expression are gaining attention in biomedical research. A specific challenge is to detect genes with expression levels that change according to dose levels in a non-random manner, but nonetheless may be considered as potential biomarkers.

Method: We are among the first to formally apply a tool that uses an isotonic (monotonic) regression approach to this area of study. We introduce a test statistic to select genes with significant dose–response expression in a monotonic fashion based on a permutation procedure. We then compare the results with those achieved from the application of a likelihood ratio-based test.

Results: We apply the isotonic regression approach to a study of gene expression in the RKO colon carcinoma cell line in response to varying dosage levels of the chemotherapeutic agent 5-fluorouracil. A feature of both Affymetrix and printed 75mer oligomer cDNA arrays produced from the same samples provides an opportunity to compare the two microarray platforms.

Availability: Statistical software S-plus Code to implement the method is available from the authors.

Contact: kcoombes@mdanderson.org

INTRODUCTION

The exploration of dose–response associations is an important objective in dose efficacy experiments. In recent years, biologists have integrated dose–response studies with microarray technologies, and more researchers are beginning to appreciate the usefulness of this application in cancer risk assessment and other biomedical research (Haber *et al.*, 2001; Preston, 2002). A specific challenge is to detect genes with expression levels that change according to dose levels in a non-random manner, but nonetheless may be considered as potential biomarkers. Several papers on this topic have been published. Methods commonly used in this type of analysis include hierarchical clustering and partitioning clustering (Cheng *et al.*, 2003; Zhao *et al.*, 2004), pairwise *t*-statistics (Braam *et al.*, 2004), Pearson correlation coefficient (Huffman *et al.*, 2004; Nichols and Sanders-Bush, 2004), analysis of variance (ANOVA) (Kato *et al.*, 2004) and linear regression approaches (Boverhof *et al.*, 2004). The problem posed is how

best a monotonic dose–response relationship between a response and an ordered exposure can be investigated; i.e. a relationship in which the dependent variable increases (or decreases) with each increment of exposure. Limitations of the common methods include the inability to differentiate changing patterns with ANOVA and the capability to capture only the linear associations between the predictors and the dependent variable with linear regression-like approaches. The use of pairwise comparisons may confront the multiple testing issue, particularly for a large number of discrete dose levels. Naciff *et al.* (2003) have implemented a trend analysis via the Jonckheere–Terpstra test, which is a non-parametric test for ordered differences among classes in a contingency table analysis.

Seeking improved methods for this research problem, we applied the isotonic regression method to the analysis of gene expression dose–response data. This method (also known as the monotonic regression method) is a powerful tool with a well-established theoretical foundation (Barlow *et al.*, 1972; Robertson *et al.*, 1988). It does not require any specific mean structure assumptions and can capture all possible shapes that represent potential dose responses in a single statistical test.

The isotonic regression method has been applied to many other areas in biomedical research, as discussed in Breslow and Day (1987), Wu *et al.* (1989), and Maclure and Greenland (1992). This type of model assumes that the expected value of the dependent variable is constant within disjointed regions of the independent variables, and is monotonic in increasing independent variables.

In this paper we illustrate an application of the isotonic regression approach using a study of gene expression in the RKO colon carcinoma cell line in response to varying dosage levels of the chemotherapeutic agent 5-fluorouracil (5-FU). The *in vitro* study was designed to determine which genes in colon carcinoma cells exhibit a dose-dependent response to this fluoropyrimidine drug that is used commonly to treat patients who have colorectal carcinoma (Zhang *et al.*, 2003). We introduce an algorithm to obtain estimated dose–response curves. We also propose a new test statistic and use real data to compare our results with those achieved using a statistic based on the traditional likelihood ratio. A feature of this dataset is that data from Affymetrix arrays and printed long oligomer cDNA arrays were produced from the same samples in the study, which allows for direct comparisons between results obtained with the two array platforms.

*To whom correspondence should be addressed.

METHODS

Model fitting

A simple regression model can be built for each gene of the form

$$y_{ij} = \beta_0 + \beta_1 d_i + \epsilon_{ij},$$

where d_i for $i = 0, \dots, I$ is the i -th dose level administered in increasing order, y_{ij} is the observed log gene expression level at dose i for the replicate j , $j \leq n_i$ for dose i , where n_i is the number of observations at dose i and $\sum_{i=1}^I n_i = N \cdot \epsilon_{ij} \sim N(0, \sigma^2)$ is the residual error term. A gene has an interesting dose response according to this model if the parameter β_1 is significantly non-zero.

The assumption of a linear relationship, however, is so restrictive that it might not be able to capture all the possible patterns of interest. In fact, researchers are interested in genes whose expression varies with the dose level in a monotonic fashion rather than in a strictly linear fashion. To explore this problem, we use an isotonic regression model. The model can be written as

$$y_{ij} = f(d_i) + \epsilon_{ij},$$

where f is an unknown monotone function (non-decreasing or non-increasing), and the other terms are defined as stated above. Using an isotonic regression model, we can capture all possible shapes that represent potential dose responses in a single statistical test.

To estimate a monotonic dose–response curve, we estimate a monotonic sequence $\hat{f}(d_0) \leq \hat{f}(d_1) \leq \dots \leq \hat{f}(d_I)$. The estimated expression level at dose i is,

$$\hat{y}_i = \hat{f}(d_i).$$

The ‘pool-adjacent-violators’ algorithm (Barlow *et al.*, 1972) can be applied to estimate the isotonic regression in a linear interpolating manner. The algorithm starts with the average response value, f^* , of each dose-level group or ‘solution block’; then the blocks are joined together step-by-step until the final non-decreasing (non-increasing) order is reached. If $f^*(d_0) \leq f^*(d_1) \leq \dots \leq f^*(d_I)$, then we also have the final partition $\hat{f}(d_i) = f^*(d_i)$. Otherwise, we select any of the pairs of violators of the ordering, i.e. we select an i such that $f^*(d_i) > f^*(d_{i+1})$, and then we pool these two values of f^* and replace them by the weighted average of the two. The procedure is repeated until the final estimates are in the required order.

Test statistics

In this section, we describe two test statistics. The hypothesis test of interest is $H_0 : f(d_0) = f(d_1) = \dots = f(d_I)$ versus $H_1 : f(d_0) \leq f(d_1) \leq \dots \leq f(d_I)$, with at least one strict inequality.

Assume y_{ij} to be normally distributed with unknown variance σ . Barlow *et al.* (1972) proposed a one-sided test based on the likelihood ratio test, defined by

$$\lambda = \frac{\sum_{i=1}^I n_i (\hat{f}(d_i) - \bar{y})^2}{\sum_{i=1}^I \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2},$$

where \bar{y} is the overall sample mean of the expression level for an individual gene. Note that λ is the ratio of the between-groups sum of squares, based on an amalgamation, to the total sum of the squares. H_0 is rejected for a certain large value of λ . Barlow *et al.* (1972) also derived the asymptotic distribution of λ if H_0 is true:

$$\Pr(\lambda \geq C) = \sum_{l=2}^I P(l, I) \Pr(B_{(l-1)/2, (N-l)/2} \geq C), \quad C > 0$$

$$\Pr(\lambda = 0) = P(1, I),$$

where $P(l, I)$ is the probability that \hat{f} takes exactly l distinct values [see Barlow *et al.* (1972) for the formula], and $B_{a,b}$ denotes a random variable having the beta distribution with parameters a and b . The asymptotic property holds only for a reasonably large sample size.

For this study, we may be more interested in genes that undergo large-fold changes of expression level when comparing the baseline level (no treatment)

to the highest dose level, conditional on the monotonic trend. Accordingly, we propose a new test statistic M . On the log-expression level,

$$M = \frac{\hat{f}(d_I) - \hat{f}(d_0)}{v},$$

where $v = \sqrt{\sum_{i=1}^I \sum_{j=1}^{n_i} (y_{ij} - \hat{f}(d_i))^2 / (N - I)}$. We treat v as an estimate of the standard error, which measures the variability by collecting information from all doses. The null distribution of the M -statistic, as well as an empirical threshold for rejecting genes, can be straightforwardly estimated through permutations. We also compare the performance of the M -statistic with λ .

RESULTS

We examined both the Affymetrix and the printed long oligo array data, and we compared the results between the two microarray platforms.

Laboratory study for analysis

The dose–response study was conducted in the Department of Pathology at The University of Texas M. D. Anderson Cancer Center in Houston to examine the effects of 5-FU on the RKO colon carcinoma cell line, as reported in detail previously (Zhang *et al.*, 2003). The RKO cells were treated with 0, 20, 40 or 80 μM 5-FU. Treated and control cells were harvested at 48 h, and RNA was isolated.

Both Affymetrix and printed long oligomer cDNA microarrays were used for transcriptome analysis of gene expression. Eight Affymetrix Hu133A GeneChipsTM (Affymetrix, Sunnyvale, CA) were hybridized, with a duplicate chip for each of the RNAs from the untreated control cells and cells treated with the three dosages. There were 22 283 probes on each array. In addition, the same four RNA samples were hybridized with cDNA arrays printed on glass slides prepared in the Genomics Core Facility at the M. D. Anderson Cancer Center. RNA from 5-FU-treated RKO cells was labeled with Cy5 red fluorescent dye, and RNA from untreated control RKO cells was labeled with Cy3 green fluorescent dye. The Cy5-labeled RNA from one of the three dose levels and the Cy3-labeled RNA from the untreated control were hybridized twice in equal amounts with each of the three printed arrays, resulting in six measurements of each gene for analysis. Each 75mer oligo was printed in duplicate on the array, resulting in 3840 spots that represented 1824 genes and 192 quality control spots. Details of the CG11 printed array are available on the website www.mdanderson.org/genomics.

Affymetrix array

Method and data quality examination Gene-expression levels were estimated using the PM-only model in dChip version 3.1 (<http://biosun1.harvard.edu/complab/dchip>). The dChip software automatically performs quality control checks, looking for individual probe outliers and entire array outliers. No arrays were flagged as outliers using dChip.

We performed hierarchical clustering of the samples using average linkage and the Pearson correlation coefficient between the log expression levels for all genes that were called ‘present’ by dChip in all six samples. We found that the untreated control sample was most different from the treated samples; and that the replicate pairs of experiments clustered as nearest neighbors. Based on both the dChip array outlier analysis and the unsupervised clustering analysis, we found the quality of this set of experiments to be high.

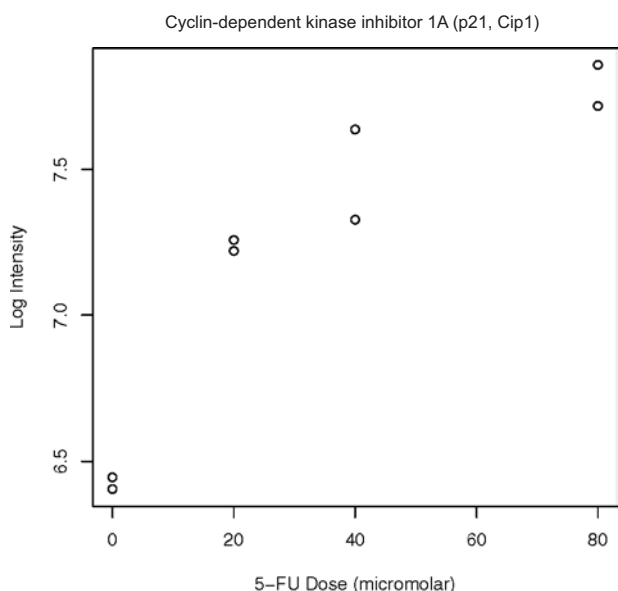


Fig. 1. Expression of *p21* from Affymetrix GeneChip in RKO cells in response to increasing doses of 5-FU.

Linear regression We fitted the simple linear regression model to all 22 283 genes, omitting those absent in all 6 samples. We identified 50 genes to control for the false-discovery rate (FDR) at 0.3 by applying the beta-uniform mixture method (Pounds and Morris, 2003); more stringent FDR criterion could detect no interesting genes. An observation was that the linear model concentrated heavily on genes expressed at low levels with small fold changes. Moreover, there was a definite bias toward downregulation: among the 50 genes, 47 were downregulated and 3 were upregulated. In addition, the analysis did not detect two important genes *Mdm2* and *p21*, which were expected to be affected by the treatment. In reviewing the data, we realized that the two genes did not increase linearly over the range of doses tested in the experiment; instead, they directly reached their maximum values in response to the smallest dose, retaining the same magnitude of expression for increasing doses. Figure 1 shows the expression pattern of *p21* along the dose levels.

Isotonic regression Since the isotonic regression model can only account for one direction at a time (increasing or decreasing), we implemented the modeling procedure twice, based on the original and sign-reversed Y . For each gene i , the corresponding test statistic in each direction was defined as M_{i+} and M_{i-} . We then chose the test statistic M_i between M_{i+} and M_{i-} that produced a closer fit to the data, or a smaller sum of squared errors. This procedure has the effect of biasing M away from zero (Fig. 2). Since the specific distribution form of M_i was not clear, it was convenient to use permutations to estimate the null distribution of M based on all the genes that had at least one expression value at each dose level (15 397): there was a total of $384(2^4 \times 4!)$ permutations, sustaining the array replicates in the same dose-level group but in a random order. The distributions of the observed and null test statistics are shown in Figure 2. By taking the signs into account, we used the threshold of the M -statistics ± 10.67 to control for the type-I error rate α at 0.01 for the two-sided test. We detected 148 genes based on

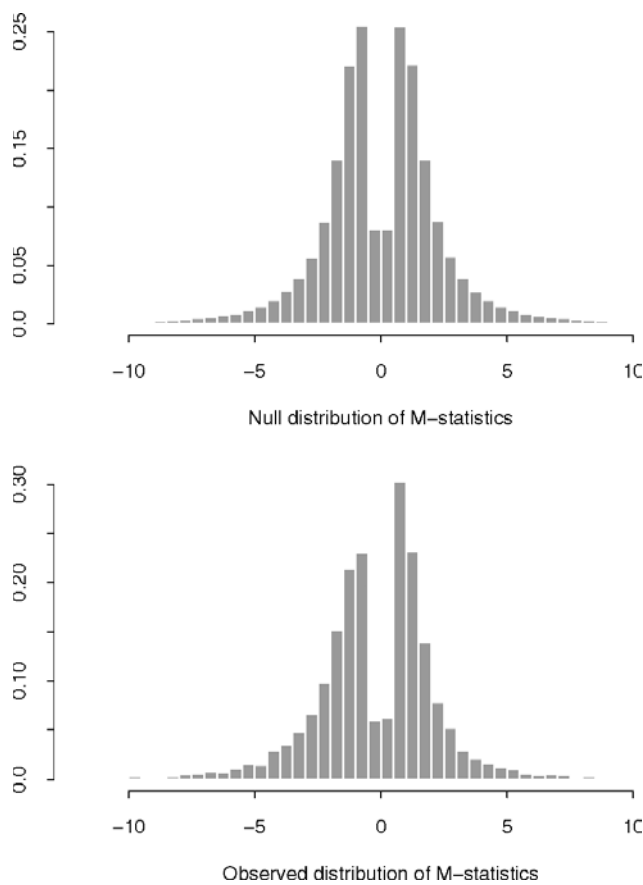


Fig. 2. Distribution of M -statistics using Affymetrix arrays.

the cutoff, forming the candidate gene list of greatest interest to biologists. The four highest-ranking probe sets are shown in Figure 3a. We noted that *Mdm2*, which was represented multiple times on the array, consistently showed up with a high value of the M -statistic.

We also computed the likelihood ratio test statistic λ . Since M and λ were in different ranges of magnitude ($\lambda \leq 1$), it was intuitive to compare the ranks of the two test statistics with respect to gene detections. The rank correlation between them was very high (0.991). The rank plots of the M -statistic versus λ are shown in Figure 4. We examined the list of 148 genes identified by M ; 134 genes also appeared in the top list of 148 genes in terms of λ ; the other 14 genes that could not be detected by λ were at the bottom of the list of the M -statistic. Hence, we found the agreement between M and λ to be high. However, we preferred the M -statistic because the genes of larger relative difference or fold change in expression between dose levels of 0 and 80 μM were of greater interest to our study. We note that the asymptotic distribution property of λ may not hold with such a small sample size. However, we were still able to compute the empirical threshold via the permutation procedure, which led to 112 genes detected at $\alpha = 0.01$. All of the 112 genes are in the list of genes detected by M -statistic.

Printed 75mer long oligo array

Method and data quality examination We preprocessed the glass microarray data: we first subtracted the background estimates at each

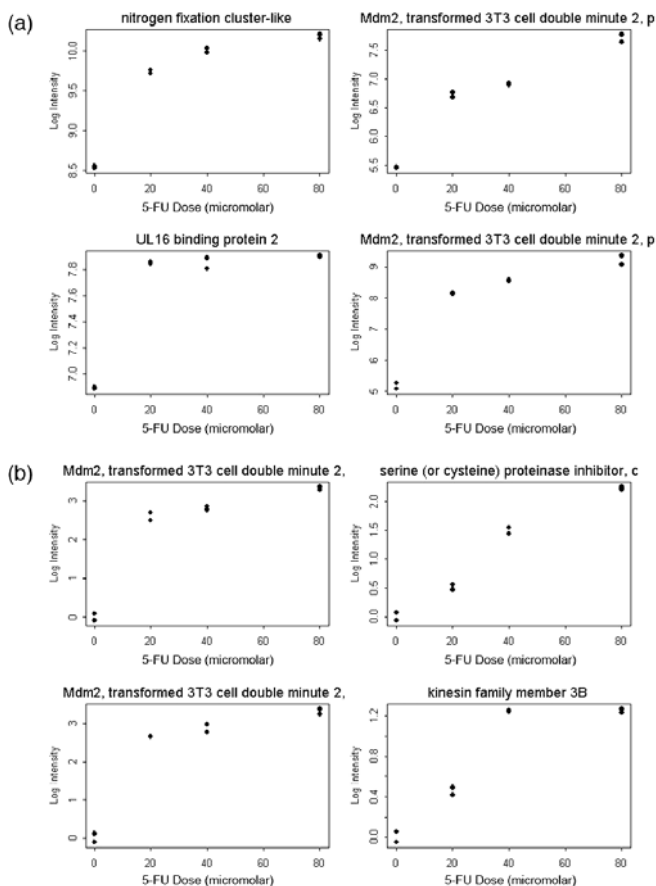


Fig. 3. Four genes with the largest M -statistics using Affymetrix GeneChip (a) and printed arrays (b).

spot from the signal; we then normalized individual channels and truncated the data by setting any normalized intensity value below a threshold of 25 equal to the threshold value; we transformed the data by computing the logarithm (base two).

Each array was examined using our standard quality control metrics. There were no obvious problems in quality; the distributions of signal intensities across channels looked consistent, with fairly uniform background noise. Standard M -versus- A plots showed no deviations from linearity. An important feature of the design of glass microarrays is that the hybridization of two samples allows us to compute ratios (or, preferably, log ratios), which eliminates some of the variation in hybridization efficiency coming from spatial variation across the array. Accordingly, we computed the log ratios of the Cy5 (red) channel, which contained the treated sample, to the Cy3 (green) channel, which contained the untreated control. The duplicate spots on each array provided two measurements of the log ratios at each of the treatments (20, 40 or 80 μM), normalized to the expression levels of the untreated control.

Isotonic regression Data from the printed glass slide arrays were similar to the Affymetrix data, except that the latter had two intensity levels at each of the four dose levels (0, 20, 40, 80 μM), without normalizing to the untreated controls. In fact, we found that the printed array data could be converted to be even more similar to the

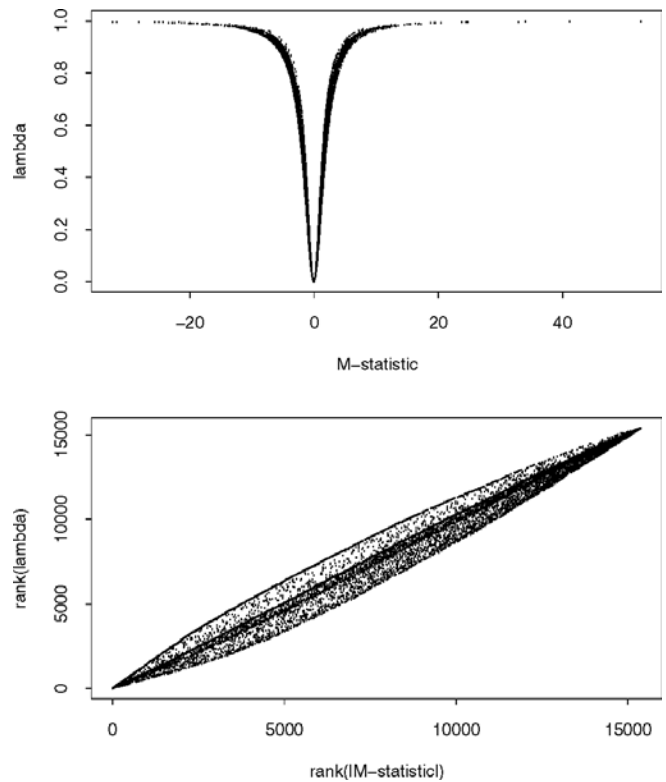


Fig. 4. M -statistics versus λ using Affymetrix arrays.

Affymetrix data, which allowed us to use identical analytical tools. We noticed that the log ratio for each gene in the untreated control should be 0. The simplest approach was to add two extra columns of all zeros to the six log ratio measurements. However, this ignored some of the inherent variability in measuring the baseline expression level. Because we had measured the actual intensity of each gene in the untreated control six times (twice in the green channel of each of three arrays), we computed the standard error of these six measurements, defined as S_0 . Thus, we were able to create two different columns of data, $S_0/2$ and $-S_0/2$, for the zero-level log-transformed ratios. In this manner, the average of these two columns is zero; the estimated variability in the measurements is identical to the observed variance of log intensities. Through this transformation, we produced a dataset that was identical in structure to the Affymetrix data.

We fitted isotonic regression models for both increasing and decreasing functions on the log ratio scale. We chose the direction with the smallest standard error of the model fitting as the final fit. We then computed the M -statistic using the same formula as before. To determine the list of potentially interesting genes, we implemented the permutation procedure in a manner similar to that used in treating the Affymetrix arrays. We selected 7.42 as the threshold of controlling α at 0.01. Using this threshold, we detected 57 genes on the printed array that showed a statistically significant dose-dependent response to the treatment of RKO cells with 5-FU. The four spots on the printed array with the highest M -statistic are shown in Figure 3b.

Platform comparison

In the comparison of the two platforms, our first observation was that two instances of the gene *Mdm2* were among the top four scoring

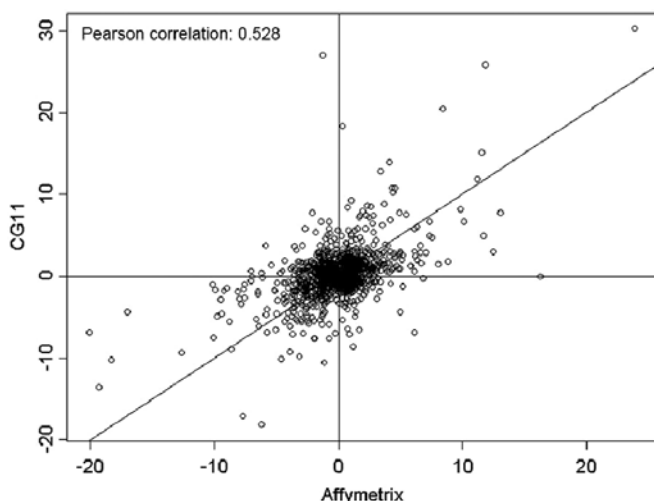


Fig. 5. Comparisons of *M*-statistics between Affymetrix GeneChip and printed arrays.

genes on both the Affymetrix and the printed platforms. We concluded fairly convincingly that treatment of RKO cells with 5-FU at any dose level above 20 μ M leads to a substantial increase in the expression level of *Mdm2*.

In order to make a more thorough comparison, we determined how many genes were common to the two platforms. For this purpose, we identified a ‘gene’ with a UniGene (Boguski and Schuler, 1995) cluster number from build 160 of UniGene (<http://www.ncbi.nlm.nih.gov/UniGene>). The 1824 spots on the printed microarray represented 1282 distinct UniGene clusters, whereas the 22 283 probe sets on the Affymetrix U133A GeneChip represented 13 794 distinct UniGene clusters. There were 1236 UniGene clusters represented on both microarrays, so that >96% of the genes on the printed microarray were also present on the Affymetrix microarray.

The fact that many UniGene clusters were represented by multiple probe sets on the individual platforms made it difficult to perform a one-to-one comparison of the measurements on the two platforms. We computed a ‘database join’ that paired every representative of a UniGene cluster on one array with every representative of the same cluster on the other array. This comparison generated 3005 such pairings. Visual inspection of the values coming from such pairings suggested that multiple probe sets on the same platform yielded fairly consistent estimates of *M*. Based on this observation, we averaged *M* over multiple probe sets within each platform in order to obtain a single number that could be used to compare data across the platforms. Figure 5 shows the comparison of the *M* values of the 1236 unique genes common to both platforms. The Pearson correlation between the *M* values was 0.528. This correlation was better than that resulting from measurements of the ‘absolute’ intensity of gene expression compared across the platforms [average correlation <0.3, described in Kuo et al. (2002)].

Another approach to assessing the agreement between platforms is to compare the number of genes found on both platforms that are detected as ‘significantly responding to the treatment’ in the separate platforms. Table 1 shows the number of genes with *M* exceeding some threshold on each platform, along with the number

Table 1. Comparisons of two platforms: number of genes detected by thresholds of *M*-statistics

Threshold	Affymetrix	Printed array	Common in two platforms	By chance
5	78	93	26	5
5.5	66	78	23	4
6	60	62	19	3
6.5	47	57	16	2
7	41	46	13	1
7.5	36	37	13	1
8	29	29	10	0
8.5	26	26	8	0
9	22	22	7	0
9.5	19	19	6	0
10	16	18	6	0

Table 2. Comparisons of two platforms: percentage of *M*-statistics having the same sign

Threshold	Printed + if Affy +	Printed – if Affy –	Affy + if printed +	Affy – if printed –
5	86	90	83	87
6	89	94	89	81
7	94	96	93	89
8	92	100	94	91
9	90	100	92	100
10	89	100	92	100

found in common among the platforms. We observed that the number of genes detected on the individual platforms were fairly similar, whereas the number of genes detected in common was only about one-third of the total number of genes selected by the Affymetrix arrays at any given cutoff level. The inherent variability in the biology and the technology means that any arbitrary cutoff will always produce some disagreement in the lists. The level of agreement we found, however, was significantly higher than one would expect by chance.

We also examined whether *M* tended to ‘point in the same direction’ on one platform whenever the gene was found to be ‘significant’ on the other platform. For instance, we looked at genes with significantly positive *M* on the Affymetrix platform, indicating that their expression levels increased in response to treatment with 5-FU. For those genes, we observed how often the sign of *M* was positive on the printed platform. If the platforms were measuring nothing in common, then we expected the signs of *M* to be 50% positive and 50% negative. If they were measuring similar features, then we expected to observe significantly more agreement. The results at different cutoffs of significance for the four kinds of comparisons are shown in Table 2. We observed that the percentage of *M* with the ‘correct’ sign, given that the gene was declared significant on one platform, was typically in the upper 80–90% range. This implied a good level of agreement between the two platforms in terms of this criterion.

CONCLUSION AND DISCUSSION

In the area of microarray screening, we are the first to formally apply the isotonic regression approach. We applied the isotonic regression method to both Affymetrix and printed 75mer long oligo array data. Moreover, we proposed the M -statistic to detect genes of large fold changes in expression between the untreated samples and those treated with the highest dose level. It served well the primary interest of our study, which was also the advantage over the likelihood ratio-based test that can only test if gene-expression intensities and dose levels are monotonically associated. We observed that the correlation between the M -statistic and the likelihood ratio-based test was high. We implemented all the possible permutations exhaustively to estimate the null distribution of the M -statistic, and then selected genes based on a certain empirical threshold determined by the permutations.

This problem can also be pursued using the non-negative least squares estimation approach (Lawson and Hanson, 1974). In practice, it can be implemented directly in S-PLUS (Insightful Company, Seattle) using the function 'npls.fit.' Interestingly, we found that the pool-adjacent-violators algorithm and the non-negative least squares estimation approach produced identical estimated dose–response curves on this dataset.

The availability of Affymetrix and printed long oligo array data obtained from the same samples allowed for comparisons between the two different platforms. Among the genes measured on both platforms, we detected similar numbers of genes differentially regulated in a dose-dependent manner in response to treatment with 5-FU. The number of genes detected in common on the two platforms was greater than that which could be attributed to chance. The correlation between the M -statistics was reasonably good. The signs of the M -statistics for genes observed to be significant were highly correlated. Moreover, from both platforms, we were able to detect the same gene, *Mdm2*, as the most important gene. Overall, both platforms appear to be able to detect some genes in common. It is not clear if we would find better agreement between experiments on the same platform by simply counting the number of genes that show up on both lists. A better test of agreement between the two lists is worth exploring. Biological confirmation of the genes detected on one platform but not the other platform would complement the study nicely, confirming that each platform can accurately detect some differences that might be more difficult to detect with the other platform.

ACKNOWLEDGEMENTS

This study was supported by P30-21115 from the National Cancer Institute, National Institutes of Health, Department of Health and

Human Services; by the Favrot Fund; and by the Wilson Fund. S.R.H is the recipient of the Frederick F. Becker Distinguished University Chair in Cancer Research from The University of Texas.

Conflict of Interest: none declared.

REFERENCES

- Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972) *Statistical Inference Under Order Restrictions*. Wiley, New York.
- Boguski, M.S. and Schuler, G.D. (1995) Establishing a human transcript map. *Nat. Genetics*, **10**, 369–371.
- Boverhof, D.R. *et al.* (2004) Temporal- and dose-dependent hepatic gene expression changes in immature ovariectomized mice following exposure to ethynyl estradiol. *Carcinogenesis*, **25**, 1277–1291.
- Braam, B. *et al.* (2004) Nitric oxide donor induces temporal and dose-dependent reduction of gene expression in human endothelial cells. *Am. J. Physiol. Heart Circ. Physiol.*, **287**, H1977–H1986.
- Breslow, N.E. and Day, N.E. (1987) *Statistical Methods in Cancer Research. Vol. II. The Design and Analysis of Cohort Studies*. IARC Scientific Publications, Lyon.
- Cheng, R.Y. *et al.* (2003) Gene expression dose-response changes in microarrays after exposure of human peripheral lung epithelial cells to nickel(II). *Toxicol. Appl. Pharmacol.*, **191**, 22–39.
- Haber, L.T. *et al.* (2001) Applications of mechanistic data in risk assessment: the past, present, and future. *Toxicol. Sci.*, **61**, 32–39.
- Huffman, D.L. *et al.* (2004) Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. *Proc. Natl Acad. Sci. USA*, **101**, 10995–11000.
- Kato, N. *et al.* (2004) Gene expression profile in the livers of rats orally administered ethynylestradiol for 28 days using a microarray technique. *Toxicology*, **200**, 179–192.
- Kuo, W.P. *et al.* (2002) Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics*, **18**, 405–412.
- Lawson, C.L. and Hanson, R.J. (1974) *Solving Least-Squares Problems*. Prentice-Hall, New York.
- Maclure, M. and Greenland, S. (1992) Tests for trend and dose-response: misinterpretations and alternative. *Am. J. Epidemiol.*, **135**, 96–104.
- Naciff, J.M. *et al.* (2003) Gene expression profile induced by 17 α -ethynyl estradiol in the prepubertal female reproductive system of the rat. *Toxicol. Sci.*, **72**, 314–330.
- Nichols, C.D. and Sanders-Bush, E.J. (2004) Molecular genetic responses to lysergic acid diethylamide include transcriptional activation of map kinase phosphatase-1, c/ebp-beta and ilad-1, a novel gene with homology to arrestins. *Neurochemistry*, **90**, 576–584.
- Pounds, S. and Morris, S.W. (2003) Estimating the occurrence of false positives and false negatives in microarray studies by approximating and partitioning the empirical distribution of P -values. *Bioinformatics*, **19**, 1236–1242.
- Preston, R.J. (2002) Quantitation of molecular endpoints for the dose-response component of cancer risk assessment. *Toxicol. Pathol.*, **30**, 112–116.
- Robertson, T., Wright, F.T. and Dykstra, R. (1988) *Order Restricted Statistical Inference*. Wiley, New York.
- Wu, M.M. *et al.* (1989) Dose–response relation between arsenic concentration in well water and mortality from cancers and vascular diseases. *Am. J. Epidemiol.*, **130**, 1123–1132.
- Zhang, W. *et al.* (2003) Apoptotic response to 5-fluorouracil treatment is mediated by reduced polyamines, non-autocrine fas ligand, and induced tumor necrosis factor receptor 2. *Cancer Biol. Ther.*, **2**, 572–578.
- Zhao, H. *et al.* (2004) Diverse effects of methylseleninic acid on the transcriptional program of human prostate cancer cells. *Mol. Biol. Cell*, **15**, 506–519.