

Multiple Comparisons Model-based Clustering and Ternary Pattern Tree Numerical Display of Gene Response to Treatment: Procedure and Application to the Preclinical Evaluation of Chemopreventive Agents¹

Thomas R. Sutter,² Xiao-Rui He, Peter Dimitrov, Lijing Xu, Giri Narasimhan, E. Olusegun George, Carrie Hayes Sutter, Clinton Grubbs, Richard Savory, Markus Stephan-Gueldner, Dirk Kreder, Michael J. Taylor, Ronald Lubet, Tricia A. Patterson, and Thomas W. Kensler

W. Harry Feinstone Center for Genomic Research [T. R. S., X-R. H., P. D., L. X., G. N., E. O. G., C. H. S.], Department of Mathematical Sciences [X-R. H., P. D., G. N., E. O. G.], University of Memphis, Memphis, Tennessee 38152; Chemoprevention Center, University of Alabama at Birmingham, Birmingham, Alabama 35294 [C. G.]; Roche, Welwyn, United Kingdom [R. S.]; F. Hoffman-La Roche AG, Basel, Switzerland [M. S-G.]; Roche Bioscience, Palo Alto, California 94304 [D. K., M. J. T.]; Division of Cancer Prevention, National Cancer Institute, Rockville, Maryland 20852 [R. L.]; and Department of Environmental Health Sciences, Johns Hopkins School of Hygiene and Public Health, Baltimore, Maryland 21205 [T. A. P., T. W. K.]

Abstract

Microarray technology has greatly aided the identification of genes that are expressed differentially. Statistical analysis of such data by multiple comparisons procedures has been slow to develop, in part, because methods to cluster the results of such comparisons in biologically meaningful ways have not been available. We isolated and analyzed, by Northern blot and GeneChip, replicate liver RNA samples ($n = 4/\text{group}$) from rats fed with control diet or diet containing one of three chemopreventive compounds, selected because their pharmacological activities, including RNA expression response, are relatively well understood. We report on a classification tree, based on the results of nonparametric multiple comparisons, which results in the bipolar hierarchical clustering of genes in relation to their response to treatment. In addition to identifying treatment-responsive genes, application of this procedure to our test study identified the known pharmacological relationships among the treatment groups without supervision. Also, small treatment-specific subsets of genes were identified that may be indicative of additional pharmacophores present in the test compounds.

Received 5/20/02; revised 10/10/02; accepted 10/22/02.

¹ Supported in part by National Cancer Institute contract N01-CN-95114-MAO, R01 CA39416, R01 ES08148, U01 AA13515, the W. Harry Feinstone Center for Genomic Research, and Roche Bioscience.

² To whom requests for reprints should be addressed, at W. Harry Feinstone Center for Genomic Research, 201 Life Science Building, University of Memphis, Memphis, TN 38152. Phone: (901) 678-8391; Fax: (901) 678-4457; E-mail: tsutter@memphis.edu.

Introduction

High density gene microarrays permit the exploration of the relative expression of thousands of genes through hybridization-based analyses of RNA samples that have been isolated under differing conditions of interest such as time, treatment, or disease state. On the basis of the rapid advances in identifying expressed sequence tags and open reading frames present in several genomes, such methods will soon allow for genome-wide comparisons of gene expression for any type of experimental variable (1, 2).

Large scale gene expression data are most often tabulated as an $m \times n$ matrix of intensity values measured for m genes (rows) in n samples/experimental groups (columns). Various unsupervised methods that cluster genes based on their expression profiles have been developed for the analysis of such gene expression matrices including hierarchical clustering (3), k-means clustering (4), self-organizing map (5), and others. Whereas the clusters of genes determined by these methods reflect the internal structure of the data in different ways, these algorithmic techniques share a common feature: they treat the expression values for each gene in n samples as a vector in n -dimensional space and then calculate the distance or similarity between each of the points in n -dimensional space using a measure such as Euclidean distance or Pearson correlation coefficient, producing a single value from two multidimensional vectors. While effective in capturing the "shape" of the gene expression data, such methods are not able to discriminate the contributions of individual dimensions to the measure of distance or similarity. This characteristic diminishes the power of these procedures to distinguish differences in the magnitude of responses between sample groups, for example, treatment effects.

Recently, increased attention has been given to the importance of replication in microarray gene expression studies. A recent study was designed to investigate a single and perhaps the simplest potential source of variation, that of the location of the cDNA spots on the slide. In this study, it was shown that any single microarray output is subject to considerable variability and that rates of misclassification can be greatly reduced by replication of the cDNA spots at other locations on the same slide (6). This experiment demonstrates the importance of considering the variability in microarray data, which can come from many sources including the actual printing and location of the probes, the preparation of different arrays on different slides, the preparation of different target samples, as well as the more commonly considered sources of variation associated with experiments such as interanimal variation or variation in agent dose de-

livery. As better datasets are becoming available, attention is now shifting toward the development of statistical procedures for microarray analysis (7, 8). Whereas such procedures are well established in the scientific literature, their application to microarray analysis is not trivial, and often requires careful consideration of data transformation to achieve normal errors with homogenous variance in the treatment groups.

Gene expression intensity values of different genes usually range over 3 to 4 orders of magnitude. To make the data comparable between genes, a feature common to analysis procedures is the log ratio transformation of the intensity values (3, 6), where the treatment group intensity value is divided by the corresponding control group intensity value, and the resulting value is then transformed by taking its natural log. In fact, this procedure may be inadequate as a normalizing transformation, and the failure of this transformation to achieve normality partially explains the use of bootstrap resampling procedures to compute adjusted P s (8). Another concern with the use of this transformation procedure is that it often requires arbitrary censoring of the expression data, allowing only positive intensity values and producing larger variation in measurements of which the intensity values are small. As a result, data that may contain valuable information is often discarded. Finally, dividing treated gene expression values by a common control expression value introduces additional correlation into the data, requiring additional adjustments when calculating P s (6).

We report on a controlled experiment that is based on a test case of three chemopreventive compounds having relatively well-understood pharmacological activities, including knowledge of the RNA expression responses of a few genes that were used to validate the test case. Interest in these compounds stems from their chemopreventive actions, that is, their ability to block or diminish early stages of carcinogenesis. Oltipraz has been shown to exhibit 20–50% of the efficacy of D3T for inhibition of aflatoxin-induced hepatic foci (9). Nonetheless, because Oltipraz is an approved drug, it is currently being evaluated in human Phase II intervention trials of biomarkers of aflatoxin-related hepatocellular carcinoma (9). More recently, BNF³ was shown to be an effective chemopreventive agent in the rat mammary carcinogen model, inhibiting 7,12-dimethylbenz(a)anthracene DNA adduct formation in liver and mammary cells by 96 and 83%, respectively (10).

Of particular importance to this study is the fact that the mechanisms of action of these compounds are relatively well understood and known to include RNA expression re-

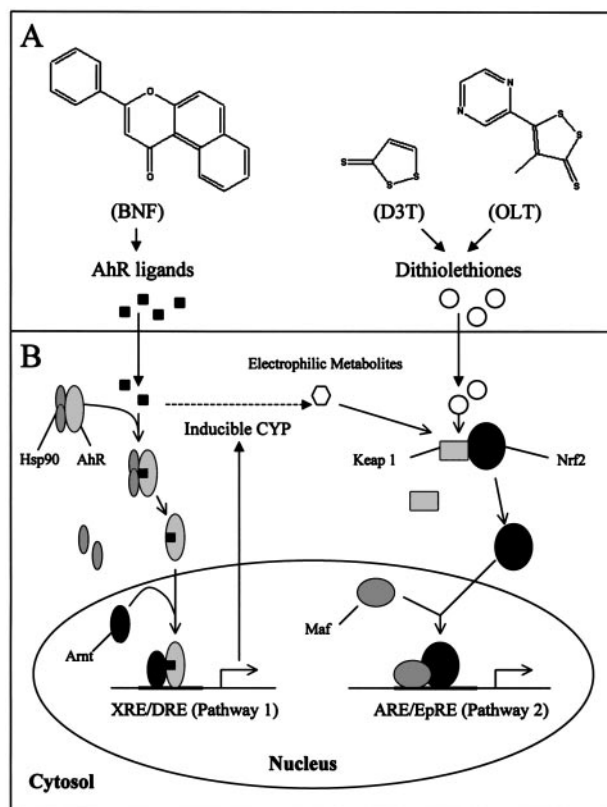


Fig. 1. Compounds and mechanisms of action. A, chemical structures. B, mechanisms of action. Two families of Phase II, or conjugating, enzyme inducers exist, based on their capacity to coordinately elevate Phase I, or functionalizing, enzymatic profiles (9, 14). Prochaska *et al.* and Talalay (14, 15), using mutant hepatoma cell lines defective in the regulation of certain Phase I enzymes, were able to experimentally define bifunctional and monofunctional inducers representing these two families. BNF, a bifunctional inducer, is an AhR ligand. In its latent form the cytosolic AhR is associated with Hsp90. Upon ligand binding, the AhR is activated and translocates to the nucleus where it forms a heterodimer complex with a second basic-helix-loop-helix transcription factor, Arnt. This complex binds to specific DNA sequences termed xenobiotic or dioxin response elements (XRE or DRE) that are found upstream of several known target genes, for example CYP1A1, and acts to enhance the rates of gene transcription (Pathway 1). Also, BNF can be metabolized to electrophilic intermediates capable of activating Pathway 2. Dithiolethiones such as D3T and OLT are monofunctional inducers and activate only Pathway 2. These compounds affect the dissociation of a Cap-N-Collar type leucine zipper transcription factor, Nrf2, from its cytosolic protein partner Keap1. Upon dissociation, Nrf2 translocates to the nucleus where it complexes with small Maf transcription factors, affecting gene transcription through DNA sequences termed electrophilic or antioxidant response elements (EpRE or ARE; Pathway 2).

sponses (11–13). As depicted in Fig. 1, BNF is known to act through pathways 1 and 2, whereas D3T and OLT are known to act through pathway 2 (14–18). Using these validated samples we measured the expression levels of several hundred genes by Affymetrix oligonucleotide microarray (GeneChip). We show that the data from such an experiment can be analyzed using nonparametric statistical procedures without censoring, transformation of the data, or division by the control value. Most importantly, we describe a novel clustering procedure that provides information about treatment effects as well as gene responses.

³ The abbreviations used are: BNF, 5,6-benzoflavone; D3T, 3H-1,2-dithiole-3-thione; OLT, 4-methyl-5-pyrazinyl-3H-1,2-dithiole-3-thione (oltipraz); CYP, cytochrome; GSTY_a, glutathione S-transferase Ya subunit; GSTY_p, glutathione S-transferase Yp subunit; GSTY_{c2}, glutathione S-transferase Yc2 subunit; DIG, dithiolethione-inducible gene; QR, NADPH quinone reductase; AFAR, aflatoxin aldehyde reductase; FL, ferritin light subunit; UGT1, UDP-glucuronosyltransferase 1; hsp, heat shock protein; gadd45, growth arrest and DNA damage-inducible gene 45; PCA, principal components analysis; AhR, aryl hydrocarbon receptor.

Materials and Methods

Chemicals and Materials. BNF, D3T, and OLT were provided by the Chemoprevention Branch of the National Cancer Institute. Rat cDNA probes for RNA blot hybridizations: CYP1A1, CYP1A2, and CYP1B1 were as described previously (11); GSTY α , and GSTY β , microsomal epoxide hydrolase, DIG1 and DIG2, QR, AFAR, FL, and albumin were measured as described previously (12). UGT1*6, and GSTYc2 were measured using oligonucleotide probes that were described previously (19). The Roche TOXCHIP containing 383 preselected rat genes was manufactured by Affymetrix (Santa Clara, CA), and included genes involved in metabolism, transport, signal transduction pathways, and housekeeping. Details of the gene responses identified in this work and the complete list of genes can be obtained.⁴

Animals and Treatments. Female Sprague Dawley rats, 4 weeks old, were obtained from Harlan Sprague-Dawley, Inc. (Indianapolis, IN) and housed under controlled conditions of temperature, humidity, and lighting at the University of Alabama at Birmingham. Food and water were available *ad libitum*, and animals were allowed to acclimate for 14 days before beginning treatments. Animals were randomly assigned to treatment groups ($n = 4/\text{group}$) and fed Harlan Teklad Rodent Diet, 4% mash diet (control), or the same formulation containing BNF (1650 mg/kg diet), D3T (600 mg/kg diet), or OLT (500 mg/kg diet) for a period of 12 days. The diets were prepared using a Patterson-Kelley liquid-solid blender and were stored at 4°C until fed to the animals. The rats were sacrificed with carbon dioxide. The livers were harvested within 3 min, rapidly frozen in liquid nitrogen, and stored at -80°C until they were processed.

RNA Isolation and Analyses. Total RNA isolation and RNA blot analysis followed procedures described previously (12), using the guanidinium thiocyanate-phenol-chloroform extraction procedure (20). Total RNA was reacted with glyoxal, electrophoretically separated, and transferred to NytranPlus membrane according to the manufacturer's instructions. Hybridizations were performed according to standard procedures using 100 ng cDNA probes that were random primed labeled with [α -³²P]dCTP to a specific activity of 1×10^8 cpm/ μg . Blots were washed twice with $1 \times \text{SSC}$ (0.15 M NaCl and 0.025 M sodium citrate) and 1.0% SDS at 50°C. Signals were visualized by autoradiography using Kodak X-Omat film with a Cronex LightningPlus intensifying screen (DuPont, Wilmington, DE). Levels of RNA were quantitated using an Epson-1200C scanner and Fuji BAS1000 software with correction for local background and normalized for RNA loading by stripping and reprobing the blots with a cDNA probe for albumin. For GeneChip expression analysis, total RNA samples were treated with RNase-free DNase, precipitated with ethanol, and resuspended in RNase-free water. Five μg of total RNA and the T7-polyT primer were used to synthesize cDNA (Ambion Megascript System) by *in vitro* transcription in the presence of biotinylated UTP and CTP (Enzo Diagnostics). Forty μg of transcripts, fragmented to 50–150 nt size, were placed in a solution of 1.0 M NaCl, 10

mM Tris-HCl (pH 7.6), and 0.005% Triton X-100, containing herring sperm DNA (0.1 mg/ml) and 5 nM control oligonucleotides. Targets were hybridized for 16 h at 45°C, and the arrays were washed at 50°C with $6 \times \text{SSET}$ [0.9 M NaCl, 60 mM NaH₂PO₄, 6 mM EDTA, and 0.005% Triton X-100 (pH 7.6)] and then at 40°C with $0.5 \times \text{SSET}$. The arrays were developed with streptavidin-phycoerythrin (Molecular Probes), and the fluorescence intensities were measured with a laser confocal scanner (Agilent/Affymetrix). Intensity for each feature of the array was used to calculate a single raw expression value for each gene using the trimmed mean algorithm of the GeneChip Software (Affymetrix). Each gene was represented by 20 oligonucleotide probe pairs, each containing one oligonucleotide perfectly matched to the gene sequence (PM and a second oligonucleotide containing a single base mismatch, MM). Because gene expression levels were computed as the average of the sum of the PM – MM, the software sometimes reported the expression levels as negative numbers. Each liver RNA sample was analyzed individually by both Northern blot and GeneChip analysis. Samples were not pooled, and each sample was analyzed on a single GeneChip.

Computer Software and Statistics. Data files were stored and managed in EXCEL (Microsoft Corp., Seattle, WA). The normal probability plot and the Shapiro Wilks tests were implemented in SAS (Statistical Analysis System, Cary, NC). The Kruskal-Wallis Statistic and the nonparametric variant of the Student-Newman-Keuls multiple comparisons test (21) were implemented using SigmaStat (SPSS Inc., Chicago, IL). Additional information about classification trees can be found in Ref. 22. PCA followed the procedure described by Raychaudhuri *et al.* (23) as implemented in the Stanford University Software package. The following algorithm was used: Given n gene expression measurements and m experimental conditions, find $k \leq m$ orthogonal vectors that can be best used to represent the data. Using the data of replicated gene expression measurements, principal components are obtained as the eigenvectors of the sample variance-covariance matrix for each gene. Each calculated eigenvector defines a principal component, and the variance accounted for by each component is defined by the size of its associated eigenvalue, *i.e.*, the eigenvectors associated with large eigenvalues contain most of the information.

Results

Study Design and Test Case. Equal numbers of animals were randomly assigned to one of four groups: control diet, BNF, D3T, or OLT. The concentrations of these compounds in the diet were selected on the basis of several ongoing studies of chemoprevention (for examples see Refs. 9, 10). In the current study, animals were administered the indicated doses of compounds for periods up to 48 days. There were no observed differences in average body weight gains, or in the color and sizes of the livers, indicating that these concentrations were below the maximum tolerated doses of these agents (data not shown). There were not significant differences noted in food consumption over the treatment period. One can estimate the daily dose of each compound based on a study of the effect of body-weight gain on the

⁴ E-mail address: genomics@memphis.edu.

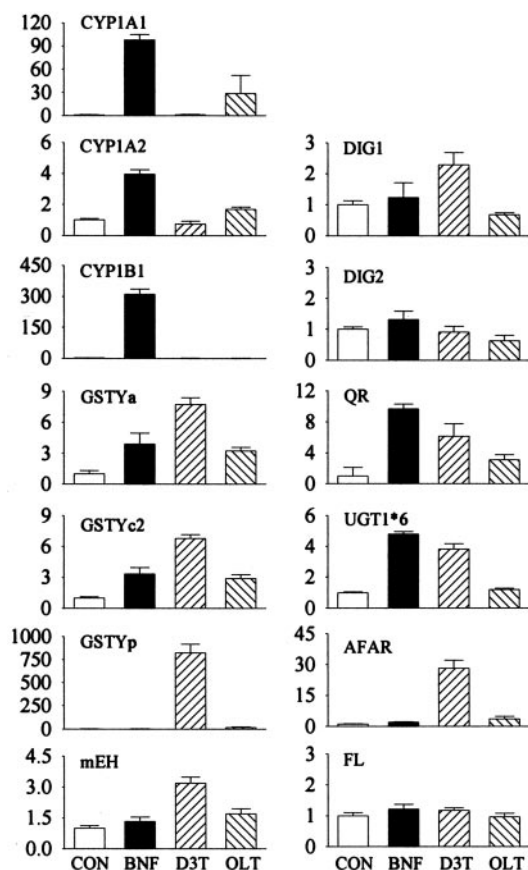


Fig. 2. RNA blot validation of gene response by Northern blot. Shown is the relative fold induction of hepatic gene expression levels (Y axis) for each treatment group (X axis), where the control value is assigned a value of 1.0. Rats were fed the chemopreventive agents for 12 days as described in "Materials and Methods;" bars, \pm SD.

evaluation of chemopreventive agents (24) where the average food intake of female Sprague Dawley rats on the same Harlan Teklad rodent diet used in our study was \sim 17g diet/day/rat. On the basis of this estimate, the calculated daily doses of BNF, D3T, and OLT are 28.1 mg, 10.2 mg, and 8.5 mg, respectively.

In parallel pilot studies, using Northern blot analysis, 12 days of treatment was shown to result in near maximal RNA responses for all three of the agents (data not shown). On this basis, we selected this time of treatment for additional analysis by GeneChip.

RNA Blot Validation of Gene Response. To ensure the effectiveness of the treatments and the quality of the RNA samples, we determined the relative expression of 13 genes known to respond to one or more of the test compounds. Shown in Fig. 2 is the relative fold induction of each gene in response to each of the chemical treatments. Such a descriptive model (relative fold induction or change factor) is often used to identify treatment-responsive genes, to determine whether a compound is active, and to assess whether two or more compounds act similarly. However, the expression of response as a ratio of treated:control introduces

correlation into the data, thereby introducing more complexity into the analysis.

Statistical Analysis and Classification of GeneChip Data. A primary way to infer the role of a factor in a biological process is by comparing the relative level of that factor under multiple experimental conditions. Initial analysis of the data indicated that there was unequal variance in each of the four treatment groups and, based on normal probability plots, that the data were not normally distributed even after log transformation (data not shown). To check the statistical validity of this graphical analysis, we performed the Shapiro-Wilks test, a formal test of normality. Even after log transformation of the data, the P s of the Shapiro-Wilks tests were <0.0001 . Setting the level of significance of this test to 0.01, we reject the hypothesis that the observed gene expression values come from a normal distribution.

We selected the Kruskal-Wallis rank procedure to test the differences between the expression levels of each gene in samples of liver from four groups of rats. Advantages of this nonparametric procedure are that it is robust to outliers and requires no assumption of distribution for the expression levels. A P of <0.05 was considered statistically significant. If the Kruskal-Wallis statistic rejects the null hypothesis that a treatment had no effect, then all of the pairwise comparisons were conducted using the nonparametric variant of the Student-Newman-Keuls test, selected because there were an equal number of observations in each treatment group. Unlike most cluster algorithms that only compare expression profiles between genes, this procedure provides specific information about the treatment effects for each gene. The latter forms the basis for gene clustering and offers pharmacological meaning to each cluster.

To cluster genes exhibiting common patterns of treatment effects, we developed a strategy that can be expressed as a ternary classification tree that is based on the outcomes of the test as follows: Represent the null hypothesis H_0 by $H_0: a = b$, and the two-sided alternative hypothesis by $H_A: a \neq b$. The classification tree assigns a value of 1 to comparisons for which the H_0 is accepted, case where a is not different from b , and assigns either a value of 2 to the case where $a < b$, or 0 to the case where $a > b$. For genes where the Kruskal-Wallis test accepts the null hypothesis H_0 , the values of 1 are assigned to each of the possible group comparisons for that gene.

The number (N) of pairwise comparisons for n treatments is $N = {}_n C_2 = n! / (2(n-2)!) = n(n-1) / 2$. For example, for three treatment groups there are three pairwise comparisons: (a) a versus b ; (b) a versus c ; and (c) b versus c . Because there are three possible outcomes for each comparison ($a < b$, $a = b$, $a > b$) the possible number of permutations is 3^3 or 27 for the example of three treatment groups. However, not all of these permutations are mathematically meaningful. For example, if $b > a$ and $c < a$, then $c > b$ is not allowed. The class of mathematically meaningful comparisons $F(a, b, c)$ satisfies the condition $F(a, b, c) = |(a \text{ versus } b) + (b \text{ versus } c) - (a \text{ versus } c) - 1| \leq 1$. Furthermore, the class of mathematically meaningful comparisons is not dependent on the order of the comparisons. The pattern tree describing the classification of the pairwise comparisons in an experiment with a

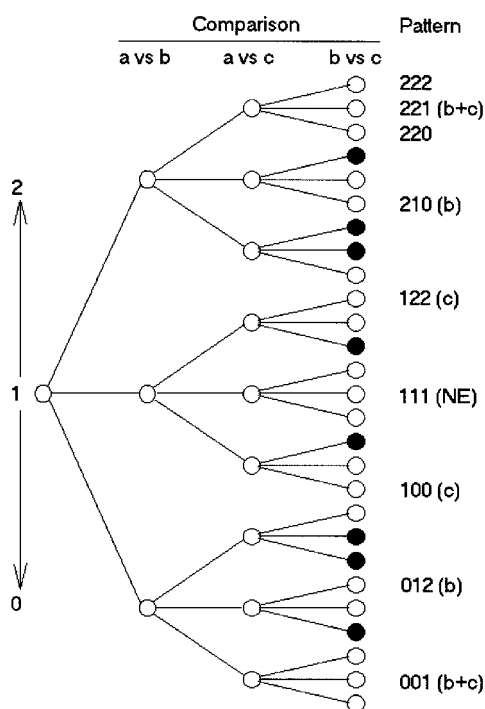


Fig. 3. The ternary pattern tree describing the permutations of possible outcomes of the pairwise comparisons in an experiment with three treatments. The height of the tree equals the number of possible pairwise comparisons. The filled nodes (●) identify permutations that are not allowed. Every nonleaf node has three children, the three branches correspond to comparisons with results 0, 1, or 2, respectively. The result of type 0 is placed below its parent node. The result of type 2 is placed above its parent node. The result of type 1 is placed in the middle. See "Results" for additional discussion.

control and two additional treatment groups is shown in Fig. 3. This ternary classification tree results in the bipolar hierarchical clustering of genes into the allowed pattern permutations with the central node pattern of 111, identifying those genes for which there was no significant treatment effect. In Fig. 3, treatment "a" represents a control group, whereas treatments "b" and "c" represents two different treatment groups. In this figure, the patterns above and below the 111 pattern are mirror images and reflect either relative increases (assigned value of 2) or decreases (assigned value of 0) in gene expression. Furthermore, in this example of three groups, when "a" is the control group, the first two comparisons indicate the properties of the gene response to treatment; the last comparison indicates the relative strength of the response to each treatment. Perhaps of greater importance, each cluster identifies genes of which the responses are indicative of specific and discernible pharmacological activity. In Fig. 3, the pattern 111 identifies genes for which there was no statistically significant treatment effect. Patterns 122 and 100 identify genes responding only to treatment "c," whereas patterns 210 and 012 identify genes responding only to treatment "b." Finally, groups of related clusters, for example, patterns 222, 221, and 220, identify genes that share similar responses to treatment as compared with control, but differ in their relative strengths of response.

Application to the Test Case of Chemopreventive Compounds.

We applied this procedure to our test case of 383 genes measured for four groups ($n = 4/\text{group}$): control, BNF, D3T, and OLT. An illustration of the data analysis performed is shown for one gene, CYP1A1, in Table 1, and the results of this analysis of all 383 genes are summarized in Table 2. For four treatment groups there are six possible pairwise comparisons for each gene. Only 219 of the possible $3^6 = 729$ decisions about the pairwise comparisons are meaningful. Furthermore, only 29 of the 219 allowed patterns are observed in this analysis (Table 2). As expected, the largest cluster, number 22, containing 194 genes, corresponds to the 111111 pattern indicating no statistically significant response to any of the treatments. The next largest cluster, number 16, contains 108 genes. This pattern, 211001, identifies genes of which the expression levels increased significantly in response to BNF only. The significance of observing a cluster of 108 genes may be appraised as follows: if the expression levels of the 383 genes are completely random then each gene expression profile has a chance of $1/219$ of being classified into one of the 219 possible patterns. In Table 2, we report the classification of genes into seven patterns (Numbers 1, 5, 8, 12, 16, 22, and 25) containing 6, 5, 11, 6, 108, 194, and 18 genes, respectively. Using a Poisson model, the probability of classifying 5 or more genes into a pattern, at random, is ~ 0.01 . Thus, using 0.05 as critical level for P_s , it is highly unlikely that any one of these gene clusterings could have been observed by chance; moreover, the probability of jointly classifying 6, 5, 11, 6, 108, 194, and 18 into the seven patterns randomly is essentially zero. Therefore, we can conclude that these patterns represent statistically significant clusters of genes. These patterns include the following responses: (a) no effect; (b) BNF only; (c) BNF and D3T; (d) BNF and OLT; and (e) BNF, D3T, and OLT. Furthermore, in our example of four experimental groups (control and three treatments) having six possible pairwise comparisons, if we ignore the last three pattern numbers (identified as xxx below), indicating relative strength of the treatment responses, and focus on the first three numbers, indicating gene response to treatment relative to control, we observe the following: 132 genes respond to BNF only (patterns: 211xxx and 011xxx); 4 genes respond to either D3T only or OLT only (patterns: 121xxx, 112xxx, and 110xxx); and the remaining 53 genes respond to BNF and D3T, BNF and OLT, or BNF and D3T and OLT.

What is evident from these results is that the clustering of genes into patterns of treatment effects is consistent with the known mechanisms of action of the test compounds. As BNF acts through both pathway 1 and pathway 2 (Fig. 1), BNF-responsive gene clusters constitute the largest set, and include the subset of D3T- and OLT-responsive gene, regulated through only pathway 2. This is additionally supported by the fact that patterns 122xxx, 100xxx, 120xxx, and 102xxx are absent in the observed patterns.

To additionally explore the treatment effects of D3T and OLT, two structurally related dithiolethiones that differ in their efficacy, we removed the BNF group data and clustered according to the pattern tree for three treatment groups (Table 3). Of the 19 allowed permutations, 13 clusters were

Table 1 Data and statistical analyses of one gene: cytochrome P4501A1

A. Gene data and Kruskal-Wallis statistic: $H = 12.904$ with 3 degrees of freedom ($P = 0.005 < 0.05$)									
gID	Probe Set	Fluorescence Intensity				Rank			
		CON	BNF	D3T	OLT	CON	BNF	D3T	OLT
1007	rCYP1A1.at	56	10815	79	564	2	16	5	9
		84	8522	36	1016	6	14	1	10
		59	9616	187	3147	3	15	8	12
		76	7103	90	1610	4	13	7	11

B. SNK analysis						
Comparison	C vs B	C vs D	C vs O	B vs D	B vs O	D vs O
SNK: q^a	4.516	1.225	3.744	5.131	3.266	4.287
$P < 0.05^b$	Yes	No	Yes	Yes	Yes	Yes
Class ^c	2	1	2	0	0	2

^a q is the Student-Neuman-Keuls test statistic for multiple comparisons of the average sum of ranks of each treatment group.

^b This is the family-wise type I error, i.e. corrected for multiple testing within a gene.

^c Assigned as described in Fig. 3.

Table 2 The expression patterns consisting of 6 pairwise comparisons found in 383 genes

Number	Pattern						Cluster size
	CB ^a	CD ^b	CO ^c	BD ^d	BO ^e	DO ^f	
1	2	2	2	1	1	1	6
2	2	2	2	1	0	1	1
3	2	2	2	1	0	0	1
4	2	2	2	0	1	2	1
5	2	2	2	0	0	1	5
6	2	2	1	2	0	0	2
7	2	2	1	1	0	1	1
8	2	2	1	1	0	0	11
9	2	2	1	0	0	1	1
10	2	2	1	0	0	0	4
11	2	2	0	0	0	0	1
12	2	1	2	0	1	2	6
13	2	1	2	0	0	2	3
14	2	1	1	1	1	1	4
15	2	1	1	1	0	1	1
16	2	1	1	0	0	1	108
17	2	1	0	0	0	1	1
18	2	0	1	0	0	2	1
19	2	0	0	0	0	1	4
20	1	2	1	2	1	0	1
21	1	1	2	1	2	2	2
22	1	1	1	1	1	1	194
23	1	1	0	1	0	0	1
24	0	1	2	2	2	1	1
25	0	1	1	2	2	1	18
26	0	1	1	1	1	1	1
27	0	1	0	2	1	0	1
28	0	1	0	1	1	1	1
29	0	0	1	1	2	2	1

^a CB, control vs. BNF.

^b CD, control vs. D3T.

^c CO, control vs. OLT.

^d BD, BNF vs. D3T.

^e BO, BNF vs. OLT.

^f DO, D3T vs. OLT.

observed. As expected, the largest increase in gene number occurred in cluster 9, pattern 111, corresponding to the no-effect group. This cluster contains genes that were identified previously either as no effect or as responding only to BNF.

Table 3 Thirteen clusters determined by 3 pairwise comparison^a

Number	Pattern						Cluster size
	CD ^b	CO ^c	DO ^d	BD ^e	BO ^f	CB ^g	
1	2	2	2	—	—	—	1
2	2	2	1	—	—	—	12
3	2	2	0	—	—	—	1
4	2	1	1	—	—	—	2
5	2	1	0	—	—	—	18
6	2	0	0	—	—	—	1
7	1	2	2	—	—	—	11
8	1	2	1	—	—	—	1
9	1	1	1	—	—	—	326
10	1	0	1	—	—	—	2
11	1	0	0	—	—	—	2
12	0	1	2	—	—	—	2
13	0	0	1	—	—	—	4

^a The data corresponding to the BNF treatment group were removed, and the remaining data were analyzed as in Table 2.

^b CD, control vs. D3T.

^c CO, control vs. OLT.

^d DO, D3T vs. OLT.

^e BD, BNF vs. D3T.

^f BO, BNF vs. OLT.

^g CB, control vs. BNF.

Consistent with the known efficacy of these compounds there is a large cluster of 18 genes responding only to D3T and several smaller clusters containing genes responding to both D3T and OLT. Of interest, clusters 1, 7, and 11 (Table 3) contain genes where the effect of OLT is greater than the effect of D3T or where the genes responded only to OLT. These clusters identify a deviation from the established structure activity relationship of D3T and OLT. Of the 14 genes present in these clusters, four are cytochromes P450 (*CYP3A4*, *CYP17A*, *CYP1A1*, and *CYP2B1*). Similar results were observed in the Northern blot analysis of CYP1A1 and CYP1A2 (Fig. 2). One plausible explanation for this observed deviation in dithiolethione structure activity lies with the 5-pyrazinyl substituent on the dithiolethione ring of OLT. OLT has been shown to be a much more potent inhibitor of hepatic cytochrome P450 than D3T (25), and such enzyme inhibition has been associated with subsequent increases in

Table 4 Comparison of RNA blot and microarray methods as tools to identify and describe responsive genes

Gene	r^a	$P < 0.05^b$	Pattern ^b	Information ^{b,c}
GSTYp	0.92	0.005	221200	B + D, B < D
GSTYa	0.80	0.008	221100	B + D, B = D
UGT1*6	0.80	0.009	221100	B + D, B = D
FL	0.55	0.006	221000	B + D, B > D
QR	0.90	0.004	221000	B + D, B > D
CYP1A1	0.96	0.005	212002	B + O, B > O
CYP1A2	0.94	0.001	211001	B only
CYP1B1	0.89	0.027	211001	B only

^a This is the correlation between the Northern blot and microarray datasets for a subset of genes identified previously as responsive to one or more of the treatments.

^b These parameters correspond only to the microarray analysis as described in "Results" and Table 2.

^c The parameter labeled "information" refers to the pharmacological meaning of the assigned pattern. For example, the pattern 221200 describes the pattern of the genes responding with an increase relative to control to either treatment with BNF (B) or D3T (D), but no response to OLT. Furthermore, in this case, the response to D3T was greater than the response to BNF.

cytochrome P450 gene expression (26). Although the specific underlying mechanism of these effects on cytochrome P450 expression are not known, these results demonstrate the bipolar hierarchical clustering of multiple comparisons of gene response to treatment groups genes that share common patterns of treatment effects. These clusters provide insight into the pharmacological activities inherent in the test compounds, including the potential presence of unique pharmacophores. Thus, the output of such an analysis includes specific new hypotheses about the activities of the test compounds, and can aid in both drug discovery and development.

The most common use of clustering algorithms, to date, has been as a tool to identify responsive genes. To test this function in our procedure, we compared the RNA blot (Fig. 2) and microarray (Table 2) analyses for the eight genes that were common to both methods. The results of this comparison are presented in Table 4. For seven of the eight genes there is a strong correlation ($r \geq 0.80$) between the raw values of the RNA blot and microarray intensity values, indicating that these independent methods give similar estimates of gene expression levels. The gene showing the poorest correlation between these two methods was *FL* ($r = 0.55$); this gene also showed the lowest response to treatment of any of the genes that were measured by both methods (Fig. 2).

When comparing the information provided by each of these models, it can be seen that the pattern generated from the classification tree provides the same kind of information as the descriptive model of relative fold induction. However, although the measurements made by Northern blot and GeneChip are correlated, the description of gene expression provided by relative fold induction (Fig. 2) and multiple comparisons model-based clustering (Table 4) are not always the same. For example, if one compares the pattern of response of CYP1A1 in Fig. 2 and Table 4, there is good agreement in the description of relative gene response. In contrast the descriptions of the relative gene response for GSTYp appears different in these two models, reflecting fundamental

Table 5 Genes selected by statistical significance and any relative fold-induction value ≥ 10 -fold

Gene	Ratio ^a		
	B:C	D:C	O:C
CYP1A1	131.1 ^b	1.4	23.0 ^b
CYP1B1	10.5 ^b	1.1	2.4
rCYP4B1	16.9 ^b	2.9	2.9
GSTYp	5.1 ^b	70.2 ^b	1.3
QR	9.5 ^b	4.3 ^b	1.6
UGT1*6	29.3 ^b	10.1 ^b	0.9
UGT1*7	19.8 ^b	7.8 ^b	2.8
hsp27	17.9 ^b	3.8	11.7 ^b
hsp70	14.4 ^b	1.0	0.5
gadd45	331.8 ^b	10.3	48.5

^a The ratio (relative fold induction) is the average gene expression value for the indicated treatment group divided by the average gene expression value for the control group. B indicates BNF; D indicates D3T, and O indicates OLT treatment, respectively. C indicates the control group.

^b Significantly different than control as determined in Table 2 ($P \leq 0.05$).

differences in these models. The major factors influencing such differences are the effect of the magnitude of scale in the relative fold induction (Fig. 2), which can mask small, yet significant effects within one group relative to another, and the fact that the pattern assigned in bipolar hierarchical clustering is based on a level of statistical significance. Nonetheless, the relative fold induction is easily calculated from the data behind the pattern assigned for each gene in Table 2. As an example, a subset of the 189 genes identified in Table 2 as having a statistically significant response to one or more of the treatments is presented in Table 5.

PCA, a multivariate statistical technique that can be used to simplify and explore complex data sets, has been applied to the analysis of microarray data (23). Because PCA can be used to describe the key components of variation in multi-dimensional data sets, we chose this analysis as an exemplary comparison of our procedure to other clustering procedures. Given our goal of understanding treatment effects, we selected the experimental conditions as the variable and applied PCA as described in "Materials and Methods." Shown in Fig. 4 are the results of this analysis. In Fig. 4A, one sees that PCA analysis identifies two distinct clusters that are separated primarily in component 1. The eigenvalues of the larger group center ~ -0.08 , whereas the eigenvalues of the smaller group center ~ 0.09 . In Fig. 4B, we have color-coded the 7 significant clusters identified by our procedures to reveal several interesting observations. The first observation is that the genes associated with cluster 111111, the no-effect cluster, are distributed throughout the PCA analysis. This indicates either that these values have been misclassified as no effect or that they are variable and would distribute to different locations with repeated experiments. The second observation is that the two major groups of genes identified by PCA are represented by clusters from our analysis that contain genes responding in opposite directions, *i.e.*, increased expression or decreased expression relative to control. For example, clusters 012221 (pink) and 211001 (red) represent genes showing decreased or increased expression in response only to the BNF treatment. The third observation is that individual clusters identified by our procedure tend to

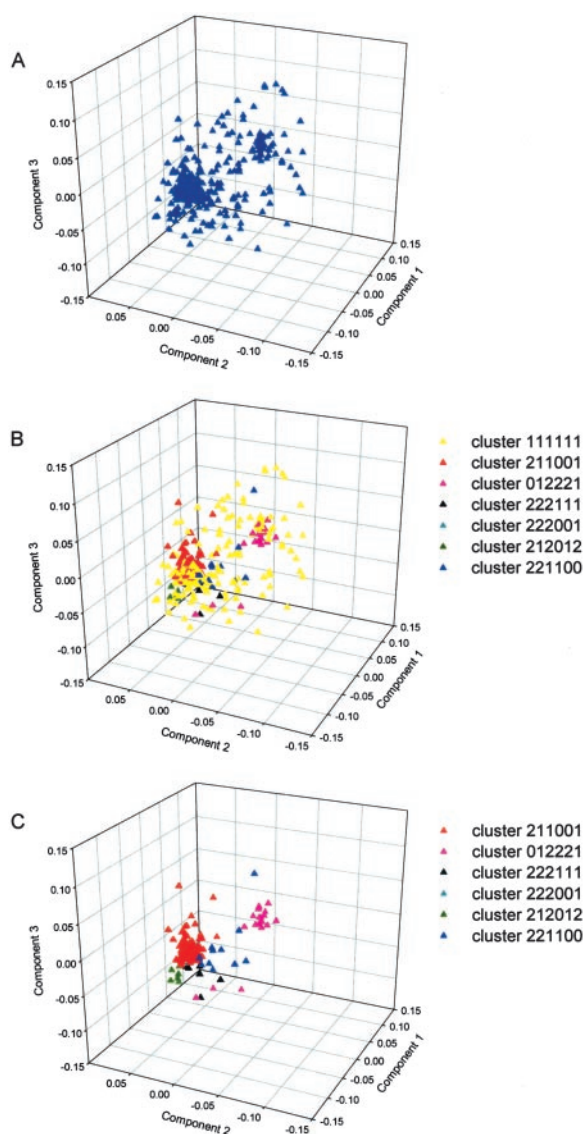


Fig. 4. PCA of the GeneChip data. A, we selected the experimental condition (CON, BNF, D3T, and OLT) as the variable and performed PCA as described in "Materials and Methods." It was observed that the first two components contain >70% of the variability, and the first three components contain ~90% of the total variability. We chose 90% of the total variance in the data as our cutoff and, therefore, plotted the data in three dimensions. The eigenvalues corresponding to components 1, 2, and 3 are shown on the X, Y, and Z axis, respectively. B, PCA color-coded to identify the genes contained in the 7 significant clusters identified by the multiple comparisons-model based clustering (Table 2; "Results"). C, PCA color-coded as in B with removal of the no-effect group (cluster 111111).

group toward distinct regions of the PCA clusters (Fig. 4C). We interpret this observation to indicate a general concordance of the results of the two methods. However, Fig. 4C helps to illustrate the finer level of clustering that is achieved by our procedure, as well as the understanding of the cluster information that is conveyed by the ternary pattern tree numerical display. Other methods such as self organizing map procedures, whereas different in their approach to PCA, share the apparent difficulty seen in this example, namely the

problem of assigning biological meaning to the clusters. Therefore, our procedure may aid in the interpretation of other clustering methods such as PCA.

Discussion

As shown in Table 5, the greatest increases in expression ratios, >10-fold, were most commonly observed in several metabolizing enzymes, including several cytochromes P450, GSTYp, UGT1*6, and UGT1*7, known to respond to these chemicals (11–13, 19). Of interest, a 16.9-fold increase in the expression of CYP4B1 was observed in response to BNF. This enzyme has been shown to catalyze the mutagenic activation of several urinary bladder carcinogens (27), and its induction by polycyclic aromatic hydrocarbons present in complex mixtures such as cigarette smoke could be an important determinant of tissue sensitivity to other mutagens such as 4-aminobiphenyl. In addition to its effects on metabolizing enzymes, treatment with BNF resulted in marked increases in the expression levels of several stress response genes (28), including *hsp27* and *hsp70*, as well as the growth-arrest and damage-inducible gene, *gadd45*. A possible explanation for the elevated expression of these genes is that they are indicative of cellular oxidative stress and DNA damage, a response associated previously with the strong induction of cytochrome P4501A1 in response to Ah receptor agonists in a hepatoma cell line (29). However, *hsp70* has been found recently to be associated with another BNF-inducible protein, QR (30), which may indicate additional or alternate roles of these chaperone proteins in maintaining the stability and function of proteins belonging to inducible oxido-reductase pathways. The remarkable number of 162 BNF-responsive genes observed in this study (Table 2) both confirms and expands the observations of Puga *et al.* (31) about the complexity of the cellular response to Ah receptor agonists. In their study (31), they identified >300 known genes of which the levels of expression were altered by at least a factor of 2.1 in human hepatoma HepG2 cells treated with 10 nM 2,3,7,8-tetrachlorodibenzo-*p*-dioxin for 8 h. Whereas it is likely that such complexity in gene response underlies the pleiotropic effects of dioxin-like compounds, the sheer number of genes indicates a major challenge to developing testable hypotheses about the relationships between changes in gene expression and toxicity.

Whereas DNA microarray technology offers investigators a powerful tool to simultaneously monitor the expression of a many genes, it also presents enormous analytical challenges. A fundamental step in extracting nontrivial information from a large collection of data is to organize the data into meaningful structures. Although numerous algorithms have been described for clustering genes that share similar response characteristics, the use of statistical methods of analysis has received less attention, in part, because methods to cluster the results of such tests in meaningful ways have not been available.

Here we describe a novel clustering approach that classifies genes on the basis of their response to treatment. We developed a ternary classification tree that assigns a value of 1 to comparisons for which the null hypothesis $H_0: a = b$ is accepted (case where a is not different from b), and assigns

either a value of 2 (case where $a < b$) or 0 (case where $a > b$) to comparisons for which the H_0 is rejected. This classification tree results in the bipolar hierarchical clustering of genes in relation to their response to treatment. A unique aspect of this microarray analysis procedure is that the treatment effect for each gene is first tested using accepted statistical methods, and then each gene is clustered into one of a finite set of bins. Each bin is identified by a specific string of numbers (pattern) that conveys explicit pharmacological and/or biological meaning.

In this study we chose a nonparametric statistical method because of the resistance of such a procedure to outliers and because the method allowed us to proceed with inference without specifying a probability model for our data, either in its original form or after log transformation. This Kruskal-Wallis statistic is a robust procedure that retains much of the information about the relative sizes of the responses without making any assumptions about how the population from which the samples were drawn is distributed. Nonetheless, it is noteworthy that this clustering procedure can be similarly applied to ANOVA methods when the assumptions of a normally distributed population are justified.

When a study includes several experimental groups, multiple comparison procedures can indicate whether a gene responds similarly to the different treatments, as well as the relative strengths of the observed responses. In the present study, a control and three treatment groups were compared. The multiple comparisons of four groups resulted in a pattern of six numbers representing the outcome of each of the six possible pairwise comparisons (Table 1). In this example, the first three numbers of the pattern identify the gene response to each of the three treatments, whereas the last three numbers of the pattern identify the relative strength of the treatment responses.

In our analysis of four treatment groups (Table 2), only 29 of the allowed 219 clusters were observed, indicating that the distribution of genes into these clusters was not occurring at random. Using a Poisson probability model we showed that 7 of the 29 observed clusters, containing 348 of the 383 analyzed genes, were significant. Whereas additional study to improve the methods to quantify the sources of error associated with such clustering is warranted, both the probability model and the independent knowledge of the mechanisms of action of the three test compounds (Fig. 1) support the utility and validity of this algorithm.

Inherent to this statistical analysis, the level of significance is predefined, clearly identifying the chance of indicating an effect of treatment when, in fact, there is none. When performing multiple comparisons, we correct for these multiple tests by computing a family-wise type 1 error. Nonetheless, as noted by Callow *et al.* (32), the possibility of observing extreme test statistic values by chance increases with the number of hypotheses being tested, *i.e.*, increases with the number of RNAs that are being measured. Currently, our model does not address this issue. However, following the example of Callow *et al.* (32), one approach would be to visually analyze the normal quantile-quantile plot of the test statistics to reveal whether the observed test statistics are likely to represent real differences between the comparison

groups. The use of resampling-based P adjustment methods should improve the classification of genes into expression patterns. When these statistical procedures are coupled with experimental designs that include reference compounds having known gene responses, as well as selective retesting of subsets of genes with alternative methods such as real-time PCR, then the predictive classification of compounds based on these procedures should improve. Additional studies will be required to determine the cost:benefit ratios of these additional analyses.

In addition to the above concerns, another limitation of the method is that type 2 error is not specifically accounted for, which can lead to misclassification of the pattern type 1 (no effect), by concluding that the treatment had no effect when, in fact, it did. However, this type of error can be lessened by increasing the power of the study, most often by increasing the sample number in each group.

There are several advantages of multiple comparison model-based clustering and ternary pattern tree numerical display including: (a) the ability to understand and compare treatment effects; (b) the observed symmetries in the response patterns; (c) the simple ability to change the priorities of the comparisons; and (d) the specific pharmacological meaning of the clusters. In addition to identifying treatment-responsive genes, application of this procedure to our test study identified the known pharmacological relationships among the treatment groups without supervision. Also, small treatment-specific subsets of genes were identified that may be indicative of additional pharmacophores present in the test compounds. In comparison to PCA, we observed a general concordance between the clusters that were identified by these two models. However, the clustering technique described here appears to produce a more refined pattern tree for the data analyzed in this study.

In summary, we have shown that multiple comparison model-based clustering and ternary pattern tree numerical display can be used to identify active compounds, to compare the activities of compounds in a structure series, and to identify unique pharmacophores present in such structure series. Because this procedure requires neither censoring nor transformation of the raw data before analysis, it will permit the comparisons of different studies, even when these studies are based on different microarray platforms.

Acknowledgments

We thank John Schuetz (St. Jude Children's Research Hospital, Memphis, TN) and Alvaro Muñoz (Johns Hopkins University, Baltimore, MD) for review of this manuscript before submission.

References

- Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (Wash. DC)*, 270: 467–470, 1995.
- Lockhart, D. J., and Winzler, D. A. Genomics, gene expression and DNA arrays. *Nature (Lond.)*, 405: 827–836, 2000.
- Eisen, M. B., Spellman, P. T., Brown, P. O., and Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA*, 95: 14863–14868, 1998.
- Herwig, R., Poustka, A. J., Muller, C., Bull, C., Lehrach, H., and O'Brien, T. Large-scale clustering of cDNA-fingerprinting data. *Genome Res.*, 9: 1093–1105, 1999.

5. Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitaveewan, S., Dmitrovsky, E., Lander, E. S., and Golub, T. R. Interpreting patterns of gene expression with self-organizing methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA*, **96**: 2907–2912, 1999.
6. Lee, M. L., Kuo, F. C., Whitmore, G. A., and Sklar, J. Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc. Natl. Acad. Sci. USA*, **97**: 9834–9839, 2000.
7. Tusher, V. G., Tibshirani, R., and Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA*, **98**: 5116–5121, 2001.
8. Kerr, M. K., and Churchill, G. A. Bootstrapping cluster analysis: assessing the reliability of conclusions from microarray experiments. *Proc. Natl. Acad. Sci. USA*, **98**: 8961–8965, 2001.
9. Kensler, T. W., Groopman, J. D., Sutter, T. R., Curphey, T. J., and Roebuck, B. D. Development of cancer chemopreventive agents: oltipraz as a paradigm. *Chem. Res. Toxicol.*, **12**: 113–126, 1999.
10. Izzoti, A., Camoirano, A., Cartiglia, C., Grubbs, C. J., Lubet, R. A., Kelloff, G. J., and De Flora, S. Patterns of DNA adduct formation in liver and mammary epithelial cells of rats treated with 7, 12-dimethylbenz(a)anthracene, and selective effects of chemopreventive agents. *Cancer Res.*, **59**: 4285–4290, 1999.
11. Walker, N. J., Gastel, J. A., Costa, L. T., Clark, G. C., Lucier, G. W., and Sutter, T. R. Rat CYP1B1: an adrenal cytochrome P450 that exhibits sex-dependent expression in livers and kidneys of TCDD-treated animals. *Carcinogenesis (Lond.)*, **16**: 1319–1327, 1995.
12. Primiano, T., Gastel, J. A., Kensler, T. W., and Sutter, T. R. Isolation of cDNAs representing dithiolethione-responsive genes. *Carcinogenesis (Lond.)*, **17**: 2297–2303, 1996.
13. Sutter, T. R., Guzman, K., Dold, K. M., and Greenlee, W. F. Targets for Dioxin: Genes for plasminogen activator inhibitor-2 and interleukin-1 β . *Science (Wash. DC)*, **254**: 415–419, 1991.
14. Prochaska, H. J., Santamaria, A. B., and Talalay, P. Rapid detection of inducers of enzymes that protect against carcinogens. *Proc. Natl. Acad. Sci. USA*, **89**: 2394–2398, 1992.
15. Talalay, P. Mechanisms of induction of enzymes that protect against chemical carcinogens. *Adv. Enzyme Regul.*, **28**: 237–250, 1998.
16. Hayes, J. D., Ellis, E. M., Neal, G. E., Harrison, D. J., and Manson, M. M. Cellular response to cancer chemopreventive agents: contribution of the antioxidant response element to the adaptive response to oxidative and chemical stress. *Biochem. Soc. Symp.*, **64**: 141–168, 1999.
17. Itoh, K., Wakabayashi, N., and Yamamoto, M. Regulatory mechanisms of cellular response to oxidative stress. *Free Radic. Res.*, **31**: 319–324, 1999.
18. Jaiswal, A. K. Regulation of genes encoding NAD(P)H:quinone oxidoreductases. *Free Radic. Biol. Med.*, **29**: 254–262, 2000.
19. Buetler, T. M., Gallagher, E. P., Wang, C., Stahl, D. L., Hayes, J. D., and Eaton, D. L. Induction of phase I and phase II drug-metabolizing enzyme mRNA, protein, and activity by BHA, ethoxyquin, and oltipraz. *Toxicol. Appl. Pharmacol.*, **135**: 45–57, 1995.
20. Chomczynski, P., and Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*, **162**: 156–169, 1987.
21. Glantz, S. A. *In: Primer of Biostatistics*, pp. 322–355. New York: McGraw-Hill, 1981.
22. Cormen, T. H., Leiserson, C. E., and Rivest, R. L. *In: Introduction to Algorithms*. Cambridge: MIT Press, 1990.
23. Raychaudhuri, S., Stuart, J. M., and Altman, R. B. Principal components analysis to summarize microarray experiments: application to sporulation time series. *Pac. Symp. Biocomput.*, **2000**: 455–466, 2000.
24. Rodriguez-Burford, C., Lubet, R. A., Eto, I., Juliana, M. M., Kelloff, G. J., Grubbs, C. J., and Steele, V. E. Effects of reduced body weight gain on the evaluation of chemopreventive agents in the methylnitrosourea-induced mammary cancer model. *Carcinogenesis*, **20**: 71–76, 1999.
25. Langouet, S., Furge, L. L., Kerriguy, N., Nakamura, K., Guillouzo, A., and Guengerich, F. P. Inhibition of human cytochrome P450 enzymes by 1, 2-dithiole-3-thione, oltipraz and its derivatives, and sulforaphane. *Chem. Res. Toxicol.*, **13**: 245–252, 2000.
26. Langouet, S., Makeo, K., Berthou, F., Morel, F., Lagadic-Gossman, D., Glaize, D., Colles, B., Ketterer, B., and Guillouzo, A. Effects of administration of the chemoprotective agent oltipraz on CYP1A and CYP2B in rat liver and rat hepatocytes in culture. *Carcinogenesis (Lond.)*, **18**: 1343–1349, 1997.
27. Imaoka, S., Yoneda, Y., Matsuda, T., Degawa, M., Fukushima, S., and Funae, Y. Mutagenic activation of urinary bladder carcinogens by CYP4B1 and the presence of CYP4B1 in bladder mucosa. *Biochem. Pharmacol.*, **54**: 677–683, 1997.
28. Fornace, A. J., Jr., Amundson, S. A., Bittner, M., Myers, T. G., Meltzer, P., Weinstein, J. N., and Trent, J. The complexity of radiation stress responses: analysis by informatics and functional genomics approaches. *Gene Expr.*, **7**: 387–400, 1999.
29. Park, J.-Y., K., Shigenaga, M. K., and Ames, B. N. Induction of cytochrome P4501A1 by 2, 3, 7, 8-tetrachlorodibenzo-*p*-dioxin or indolo(3, 2-*b*)carbazole is associated with oxidative DNA damage. *Proc. Natl. Acad. Sci. USA*, **93**: 2322–2327, 1996.
30. Anwar, A., Siegel, D., Kepa, J. K., and Ross, D. (2002) Interaction of the molecular chaperone Hsp70 with human NAD(P)H:quinone reductase 1. *J. Biol. Chem.*, **277**: 14060–14067, 2002.
31. Puga, A., Maier, A., and Medvedovic, M. The transcriptional signature of dioxin in human hepatoma HepG2 cells. *Biochem. Pharmacol.*, **60**: 1129–1142, 2000.
32. Callow, M. J., Dudoit, S., Gong, E. L., Speed, T. P., and Rubin, E. M. Microarray expression profiling identifies genes with altered expression in HDL-deficient mice. *Genome Res.*, **10**: 2022–2029, 2000.