

Prediction of Compound Signature Using High Density Gene Expression Profiling

Hisham K. Hamadeh,* Pierre R. Bushel,* Supriya Jayadev,† Olympia DiSorbo,† Lee Bennett,* Leping Li,* Raymond Tennant,* Raymond Stoll,† J. Carl Barrett,* Richard S. Paules,* Kerry Blanchard,† and Cynthia A. Afshari*¹

*National Institute of Environmental Health Sciences, P.O. Box 12233, MD2-04, Research Triangle Park, North Carolina 27709; and †Boehringer-Ingelheim Pharmaceuticals, Inc., Ridgefield, Connecticut 06877

Received November 26, 2001; accepted January 8, 2002

DNA microarrays, used to measure the gene expression of thousands of genes simultaneously, hold promise for future application in efficient screening of therapeutic drugs. This will be aided by the development and population of a database with gene expression profiles corresponding to biological responses to exposures to known compounds whose toxicological and pathological endpoints are well characterized. Such databases could then be interrogated, using profiles corresponding to biological responses to drugs after developmental or environmental exposures. A positive correlation with an archived profile could lead to some knowledge regarding the potential effects of the tested compound or exposure. We have previously shown that cDNA microarrays can be used to generate chemical-specific gene expression profiles that can be distinguished across and within compound classes, using clustering, simple correlation, or principal component analyses. In this report, we test the hypothesis that knowledge can be gained regarding the nature of blinded samples, using an initial training set comprised of gene expression profiles derived from rat liver exposed to clofibrate, Wyeth 14,643, gemfibrozil, or phenobarbital for 24 h or 2 weeks of exposure. Highly discriminant genes were derived from our database training set using approaches including linear discriminant analysis (LDA) and genetic algorithm/K-nearest neighbors (GA/KNN). Using these genes in the analysis of coded liver RNA samples derived from 24-h, 3-day, or 2-week exposures to phenytoin, diethylhexylphthalate, or hexobarbital led to successful prediction of whether these samples were derived from livers of rats exposed to enzyme inducers or to peroxisome proliferators. This validates our initial hypothesis and lends credibility to the concept that the further development of a gene expression database for chemical effects will greatly enhance the hazard identification processes.

Key Words: toxicogenomics; gene expression database; discriminant genes; prediction; algorithms; DNA arrays.

Development of novel approaches for high-throughput screening for potential adverse effects of chemicals is a major goal in the drug development process and is also now a part of environmental health research programs. In the past, structural data, mutagenicity assays, and a host of other endpoints have

been used as measures for prediction of potential adverse effects of chemical exposure, but with limited success (Ashby, 1994; Ashby and Paton, 1993; Cunningham *et al.*, 1998; Enslin *et al.*, 1994; King and Srinivasan, 1996; Klopman and Rosenkranz 1994). The need for advancing prediction processes has made technologies exploiting advances in genomics, proteomics, and metabonomics promising approaches to achieve this goal. One strength of these genomic based approaches is that gene and protein expression analyses, or analyses of metabolites offer multivariate data sets that theoretically increase the chance of generating unique profiles associated with chemical exposure and effects, which should in turn, increase the potential power on which to make predictions of unknowns. This is in marked contrast to current approaches to predictive analysis that use structural information dealing with physical attributes of compounds; these attributes are frequently relatively low in their numbers of variables corresponding to a chemical. In this study, we wanted to test whether gene expression profiling could be used to classify RNA samples derived from livers of rats exposed to coded compounds.

The genomics approach for predictive toxicology mandates the successful interrogation of databases populated with gene expression profiles corresponding to biological responses to well characterized, known compounds and comparing those with expression profiles from biological responses to exposures corresponding to unknown chemicals (Lovett, 2000; Nuwaysir *et al.*, 1999; Hamadeh *et al.*, 2002a). The hypothesis that underlies this approach is that similarities among profiles will indicate shared mechanisms of action and/or toxicological responses among the chemicals being compared. It has been demonstrated that compounds with similar pharmacological or toxicological effects produced similar gene expression profiles following either *in vitro* (Waring *et al.*, 2001a) or *in vivo* (Waring *et al.*, 2001b) exposure conditions. We have previously demonstrated that gene expression profiles corresponding to livers of rats exposed to either peroxisome proliferators or an enzyme inducer, clustered based on the mechanism of toxicant action (Hamadeh *et al.*, 2002b). Gene expression measurements corresponding to the *in vitro* response of rat hepatocytes to 15 known compounds revealed that profiles of

¹ To whom correspondence should be addressed. Fax: (919) 316-4535. E-mail: afshari@niehs.nih.gov.

chemicals with similar toxic mechanisms clustered together (Waring *et al.*, 2001a). Another study demonstrated a strong correlation between the histopathology, clinical chemistry, and gene expression profiles corresponding to livers derived from chemically exposed rats (Waring *et al.*, 2001b).

The use of gene expression profiles for classification and predictive purposes has been demonstrated in the field of oncology (Alaiya *et al.*, 2000; Alizadeh *et al.*, 2000; Golub *et al.*, 1999; Perou *et al.*, 1999; Perou *et al.*, 2000). Tumor samples from human patients were classified in a blinded fashion, based on learning data sets that provided knowledge on the tumor categorization and allowed for objective classification of the unknown samples. This approach, however, has not been robustly applied toward the determination of the identity of biological samples derived from *in vivo* chemical exposure models. A challenging question facing the validity of the use of transcript profiling to reveal chemically induced responses in treated animals is whether profiles can be used to predict the classification of coded samples generated from exposures to compounds that have not been profiled before.

To test our hypothesis, we investigated, in a blinded study, gene expression profiles from liver samples of chemically treated Sprague-Dawley rats. Specific compound identities, mechanistic classes, or doses of the compounds were coded to the team members involved in the gene expression profiling and data interpretation throughout the analysis and prediction process. In addition, no grouping of samples was provided in cases in which multiple samples were derived from animals treated with the same agent. The only information provided was that the duration of exposure to the agents varied from 24 h, to 3 days, to 2 weeks. The study included 23 coded samples. The knowledge derived from previous studies about key discriminator genes that correlated highly with their mechanism of action (Bushel *et al.*, unpublished data) was used to interpret the gene expression profiles of the blinded samples in order to generate predictions about the identity of these samples. Using these discriminative genes, we were able to predict that 13 of the samples were similar to either the class of enzyme inducers (phenobarbital-like) or to peroxisome proliferators. The remaining 10 compounds were classified as being not similar to profiles in our database. Upon completion of the prediction, the sample identifiers were decoded, and we found that "correct" statements were made regarding 22 of the 23 samples. These results provide strong evidence that the classification of unknown compounds, based on *in vivo* gene expression profiles by comparison to a limited known data set, is possible, and provides validation of the strategy that underlies a toxico- or pharmacogenomic approach to classification of agent action.

MATERIALS AND METHODS

Animal treatment and sample collection. Male Sprague-Dawley VAF⁺ albino rats (CRL:CD(SD) BR; Charles River, Kingston, NY), approximately

5–7 weeks old, were treated with phenytoin (5,5-diphenylhydantoin, CAS # 57–41–0 [612, 616, 618: 300 mg/kg body weight/day for 24 h; 672, 674, 676, 678, 300 mg/kg/day for 2 weeks; 3462, 3464, 3468: 150 mg/kg/day for 2 weeks]), hexobarbital (CAS # 56–29–1 [630, 632, 634: 200 mg/kg/day for 24 h; 688, 690, 692, 694: 200 mg/kg/day for 2 weeks]), or DEHP (di-(2-ethylhexyl)phthalate, CAS # 117–81–7 [270, 272, 274, 276: 1200 mg/kg/day for 3 days; 4216, 4218: 1200 mg/kg/day for 2 weeks]). In-life study protocols, including animal housing, dosage, sacrifice, and tissue harvesting, were identical to the methods described earlier (Hamadeh *et al.*, 2002b). Experiments were performed according to guidelines established in the NIH *Guide for the Care and Use of Laboratory Animals*. On necropsy days, liver portions were collected in RNase-free tubes and snap frozen in liquid nitrogen. Frozen tissues were stored at –70°C until processed for RNA extraction. A control sample was generated by pooling livers of 9 vehicle-treated rats.

RNA isolation and DNA microarray hybridization and analysis. RNA isolation protocols are identical to those reported earlier in (Hamadeh *et al.*, 2002b). The cDNA Rat Chip software, v1.0, developed in-house at NIEHS, was used for gene expression profiling experiments. A complete listing of the genes on this chip is available at the following website: <http://dir.niehs.nih.gov/microarray/chips.htm>. cDNA microarray chips were prepared as previously described. (DeRisi *et al.*, 1996) and are also described in (Hamadeh *et al.*, 2002b). Each RNA pair, from coded control and treated livers, was hybridized to at least 2 arrays yielding at least 4 measurements on each gene. The raw pixel intensity images were analyzed using the ArraySuite, v1.3, extensions of the IPLab image processing software package (Scanalytics, Fairfax, VA) (Chen *et al.*, 1997). The ratio intensity data from all of the 1700 spots printed on the NIEHS Rat Chip, v1.0, was used to fit a probability distribution to the ratio intensity values and estimate the normalization constants that this distribution provides. Genes having normalized ratio intensity values outside of the 95% confidence interval were considered significantly differentially expressed and deposited into the NIEHS MAPS database (Bushel *et al.*, 2001). For each exposure condition, a query of the database yielded a list of genes that were differentially expressed in at least 3 of the 4 replicate measurements. A calculation using the binomial probability distribution indicated that the probability of a single gene appearing on this list when there was no real differential expression is approximately 0.0025.

Training set. The training set used in this study comprised of RNA samples derived from livers of Sprague-Dawley rats exposed to one of 3 peroxisome proliferators (clofibrate, Wyeth 14,643, gemfibrozil), or an enzyme inducer (phenobarbital) for 24 h or 2 weeks. A detailed description of this set is provided in Hamadeh *et al.* (2002b).

Genetic algorithm/K-nearest neighbor (GA/KNN). The GA/KNN method combines a genetic algorithm (GA) as a searching tool and the K-nearest neighbor (KNN) approach for nonparametric pattern recognition. The method not only selects a subset of informative genes that jointly discriminate among different classes of specimens but also assesses the relative predictive importance of all the genes for specimen classification. The methodology of GA/KNN is briefly described below; see Li *et al.*, 2001 and the website <http://dir.niehs.nih.gov/microarray/datamining/> for details. Let $G_m = (g_{1m}, g_{2m}, \dots, g_{im}, \dots, g_{qm})$, where g_{im} is the *log* expression ratio of the *i*th gene in the *m*th specimen; $m = 1, \dots, M$ (M = number of samples in the training set = 27; 9-clofibrate, 9-Wyeth, 9-gemfibrozil, and 9-phenobarbital). In the KNN method, one computes the Euclidean distance between each specimen, represented by its vector G_m , and each of the other specimens. Each specimen is classified according to the class membership of its *k*-nearest neighbors. In this study, we set $q = 30$ and $k = 3$ and required all of the 3 nearest neighbors to agree. If the 3 nearest neighbors was not of the same chemical class, the specimen was considered unclassified. A set of q ($q = 30$) of genes was considered discriminative when at least 25 of 27 specimens were correctly classified. A total of 10,000 such subsets of genes were obtained. Genes were then rank-ordered according to how many times they were selected into these subsets. The top 100 genes were subsequently used for prediction purposes.

Linear discriminant analysis (LDA). Standard ANOVA models (Kerr and Churchill 2001), were used to identify genes that have significantly different

mean expression values between classes of compounds in the training set of peroxisome proliferators and enzyme inducers (Hamadeh *et al.*, 2002b). Any genes that were identified by ANOVA but had a global standard deviation of 0.3 (\log_2 units) or higher were excluded. Linear discriminant analysis (LDA) was then used to test all pairs of genes to identify those that can jointly discriminate between the classes, again using a minimum variability criterion to reduce the number of pairs selected. Additional genes that had high similarity ($r > 0.95$) in their expression profile across known samples were determined using GeneSpring software (Silicon Genetics, Wood, CA) and added to this list of discriminatory genes.

Prediction criteria. Genes, found to be highly discriminatory between peroxisome proliferators and the enzyme inducer in the training set, using LDA and the 100 top-ranked class discriminatory genes selected from the GA/KNN, were compared. The intersection of these gene lists was generated and resulted in a list of 22 genes. For each, the calibrated ratios (log-transformed) were then averaged across the replicate hybridizations in the training set. Next, a pairwise Pearson correlation coefficient was calculated for each of the training samples and each of the coded samples, according to the expression ratios of all 22 genes using JMP software (SAS, Cary NC). Samples were determined to be similar if $r \geq 0.8$.

RESULTS

Sprague-Dawley rats were treated with a series of compounds, and RNA was isolated, coded, and submitted for global gene expression analysis in a single-blinded fashion. No prior knowledge of the pharmacological/toxicological class of the blinded samples or the relationship of compounds and sample groupings was provided to the array analysis team. RNA from blinded samples, corresponding to *in vivo*, chemically treated Sprague-Dawley livers, were analyzed according to the same protocols used for previous studies on the effects of peroxisome proliferators and phenobarbital (Hamadeh *et al.*, 2002b).

In order to make a prediction on properties of the blinded samples, we used the gene expression profile data set (Hamadeh *et al.*, 2002b) corresponding to livers from rats exposed to 4 known compounds (Wyeth 14,643, clofibrate, gemfibrozil, phenobarbital) as a training set. Multiple approaches were used to find highly discriminatory/informative genes whose expression pattern could distinguish RNA samples derived from livers exposed to different chemicals. LDA and GA/KNN were useful in revealing single genes or groups of genes that could separate known samples based on the class of chemical involved in the exposure. Table 1 lists 22 highly informative genes that clearly exhibited different patterns of expression between the 2 pharmacological/toxicological classes of compounds, peroxisome proliferators and enzyme inducers.

We found that visualization of the profiles of these discriminative genes was useful for interpretation. For example, the *tripeptidylpeptidase II* gene (UniGene accession # AII11901) was identified as a highly discriminating gene between peroxisome proliferators and enzyme inducers, based on LDA, GA/KNN. We found 5 clones [*Mitochondrial 3,2 transenoyl isomerase* (AA965078, AA997009); *p55cdc* (AA957359); *3-oxoacyl-CoA thiolase* (AA964573); *Mitochondrial long chain 3-ketoacyl CoA thiolase* (AI070082)] on our Rat Chip

TABLE 1
Genes Determined by LDA and GA/KNN to Discriminate between Peroxisome Proliferators and Enzyme Inducers

GenBank accession no.	Description
AA818043	Rat cytochrome P450 2C7
AA818339	Rat liver glutathione S-transferase Ya subunit
AA818412	Rat cytochrome P450 2B2, phenobarbital-inducible
AA818637	Rat transthyretin
AA819345	Rat parvalbumin alpha, calcium binding protein
AA819465	Rat apolipoprotein C-III
AA819595	Rat corticosteroid 11-beta-dehydrogenase
AA858966	Rat cytochrome P450 2C6
AA859478	<i>Rattus norvegicus</i> , Sprague-Dawley Ah receptor
AA899180	Rat NADPH:menadione oxidoreductase
AA924883	<i>Rattus norvegicus</i> CTP:phosphoethanolamine cytidyltransferase
AA955106	<i>Rattus norvegicus</i> mitochondrial aldehyde dehydrogenase 1
AA957359	<i>Rattus norvegicus</i> p55cdc, probable cell division protein
AA964573	Rat mitochondrial ketoacyl CoA thiolase
AA965078	Rat mitochondrial 3-2trans-enoyl-CoA isomerase
AA997009	ESTs
AA998734	Rat glutathione S-transferase Yb2
AI043655	SP24 rat secreted phosphoprotein 24
AI044782	Rat complement C3 precursor
AI070028	Rat UDP-glucuronosyltransferase
AI070082	Rat long-chain 3-ketoacyl-CoA thiolase beta-subunit
AII11901	<i>Rattus norvegicus</i> tripeptidylpeptidase II

that had a minimum of 95% correlation with the expression pattern of *tripeptidylpeptidase II* across known samples. The expression pattern corresponding to this set of genes across known and blinded samples was plotted (Fig. 1). Visually, the plot indicated a similarity in the pattern of expression of the *tripeptidylpeptidase II*-like genes among known peroxisome proliferator samples and blinded samples 270, 272, 274, 276, 4216, and 4218 and provided evidence that none of the other blinded samples were likely to be peroxisome proliferators.

We performed set correlation analysis, which compares 2 sets of multiple variables by pairing each blinded sample with every known sample using JMP software (SAS, Cary, NC). The procedure is previously described in detail (Johnson, 1998; Neter, 1996). Tables 2, 3, and 4 show the correlation coefficient (r) values for the comparisons. Relatively higher values of r indicate strong correlation and potential similarity between the compared samples. Since times of exposure of blinded samples were furnished to us, r values corresponding to comparisons between samples from same times of exposures were considered for predictive purposes.

Because there were no previously published reports on the application of microarray analyses to prediction of identity of coded samples, we had to decide initially on a cutoff to determine where we would accept that a correlation was truly strong enough to accurately predict similarity. Based on our

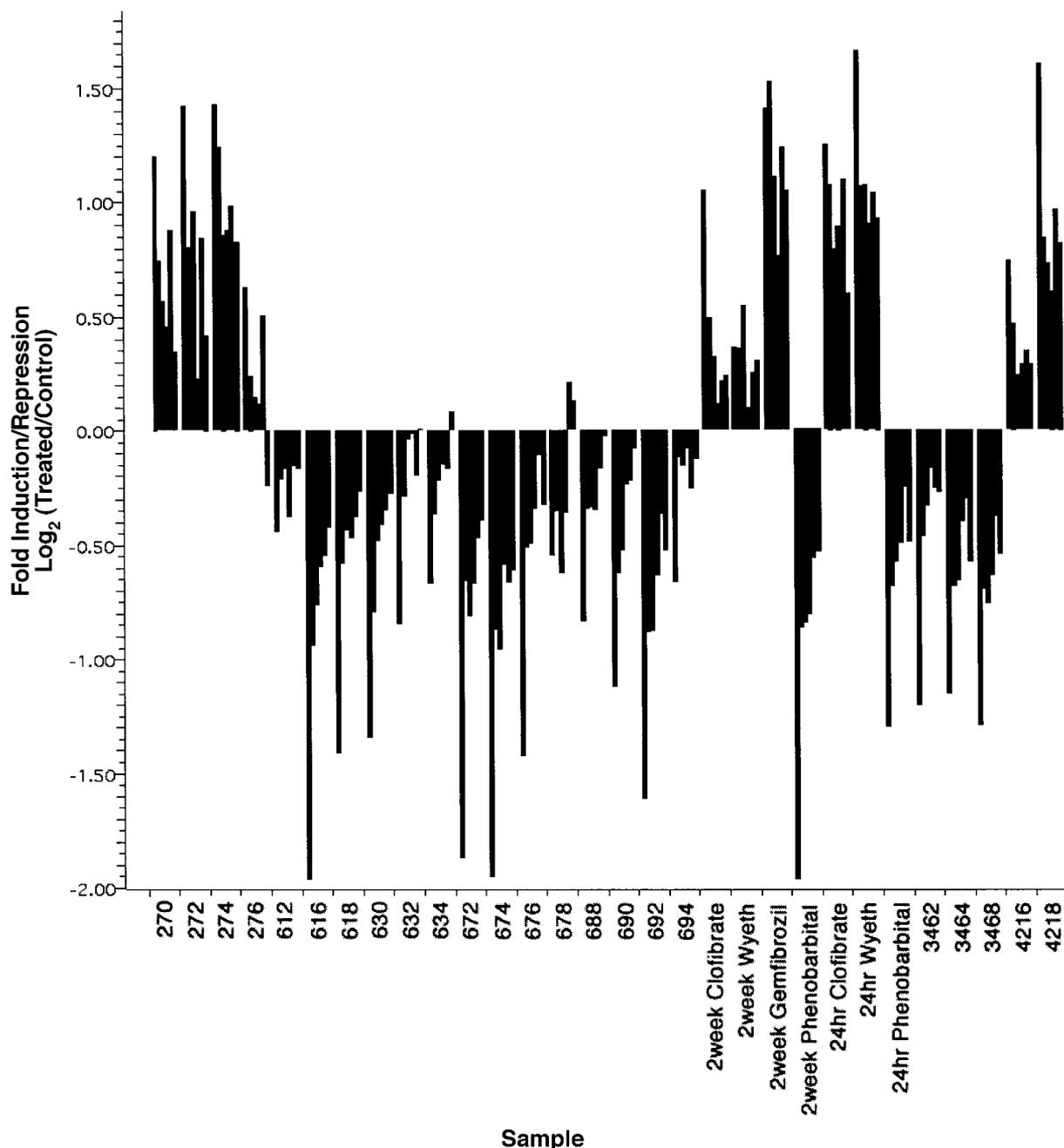


FIG. 1. Statistical and computational tools such as linear discriminant analysis, genetic algorithm/K-nearest neighbor, and single gene ANOVA enabled the identification of genes/clones (*tripeptidylpeptidase II* [AI111901]; *Mitochondrial 3,2 transenoyl isomerase* [AA965078], EST [AA997009]; *p55cdc* [AA957359]; *3-oxoacyl-CoA thiolase* [AA964573]; *Mitochondrial long chain ketoacyl CoA thiolase* [AI070082]) on our chip that had relatively high discriminative properties between the 2 classes of compounds, namely peroxisome proliferators and enzyme inducers. The plot shows that those genes were upregulated by peroxisome proliferators only and indicates a visual similarity in the pattern of expression of those genes among known peroxisome proliferator samples and blinded samples predicted to have similar properties to those samples.

previous experience, where we had studied the correlation of animals treated with similar and different compounds (Hama-deh *et al.*, 2002b), we decided to use $r \geq 0.8$ as our cutoff. For example, blinded samples 4216 and 4218, which were derived from 2-week exposures, displayed very high correlation (Table

2, 0.8809 and 0.9595, respectively) with the 2-week Wyeth 14,643 gene expression profile, while having no correlation with the 2-week phenobarbital sample (Table 2, 0.0510, and -0.0315 , respectively) and thus were classified as peroxisome proliferators (Table 5). Blinded samples 616, 618, 672, 674,

TABLE 2
Pairwise Correlation Matrix between Known and Blinded Samples (2-Week) Based on the Expression of Highly Discriminant Genes

Unknown samples	Known samples			
	Clofibrate	Wyeth 14,643	Gemfibrozil	Phenobarbital
672	0.3644	-0.0743	0.0970	<i>0.8886</i>
674	0.3859	-0.0867	0.1417	<i>0.9489</i>
676	0.4045	0.1476	0.2803	<i>0.8670</i>
678	0.4172	0.2009	0.3156	<i>0.8502</i>
688	0.4650	0.1967	0.3632	<i>0.8590</i>
690	0.2333	-0.4689	-0.2955	0.5221
692	0.4098	-0.1080	0.0212	0.7120
694	0.3171	0.4843	0.5496	0.6138
3462	0.2314	-0.2938	-0.0650	0.7294
3464	0.1059	-0.5528	-0.2662	0.5155
3468	0.3786	-0.5044	-0.2708	0.7524
4216	0.2093	<i>0.8949</i>	0.7259	0.0263
4218	0.1250	<i>0.9666</i>	0.7871	-0.0690

Note. Values in italics met the criteria of $r \geq 0.8$, indicating similarity between known and coded samples.

676, 678, and 688 had high correlation (Tables 2 and 3, $r > 0.8$) when compared to samples derived from phenobarbital-treated animals at the respective times of exposure. In addition, these samples had either negative or low correlation with samples derived from known peroxisome proliferator-treated animals, and therefore classified as similar to the enzyme inducer class of compounds (Table 5). Blinded samples 270, 272, 274, and 276 that were generated from 3-day exposures were compared to both 24-h and 2-week known samples, because the time of exposure was not common to any of the samples in the learning set. The highest correlation was found between the 3-day coded samples and the known sample derived from the 2-week Wyeth 14,643 treatment (Table 4, $r > 0.83$). The correlation of the 3-day coded samples with samples derived from phenobarbital treatments (Table 4, $r < 0.2$) was low. Therefore, these samples were also classified as being similar to peroxisome proliferators (Table 5).

We were unable to make a "positive" call with high confidence on several samples that were not similar to any of the training samples based on correlation (Table 3, $r < 0.8$).

Blinded samples 612, 630, 632, and 634 were derived from rats chemically exposed for 24 h and were compared to 24-h chemically exposed known samples. While these samples were negatively correlated with 24-h known peroxisome proliferator samples, correlation coefficients with the 24-h phenobarbital sample ranged from 0.33 to 0.73 (Table 3). Similar observations were noted for the 2-week blinded samples 690, 692, 694, 3462, 3464, and 3468 (Table 2). Because of the negative correlation between these samples and the known peroxisome proliferator-derived samples in the learning set, it was concluded that these samples were most likely not similar to peroxisome proliferators. There was a positive correlation between these samples and phenobarbital gene expression profiles; however, the similarity was not sufficient to make a certain determination. Therefore we predicted that it was highly unlikely these coded samples were derived from exposure to peroxisome proliferators, and we documented the prediction of the identity of these unknowns as being not similar to peroxisome proliferators (Table 5).

In summary, we were successful in correctly making a

TABLE 3
Pairwise Correlation Matrix between Known and Blinded Samples (24-h) Based on the Expression of Highly Discriminant Genes

Unknown samples	Known samples			
	Clofibrate	Wyeth 14,643	Gemfibrozil	Phenobarbital
612	0.21	0.09	0.05	0.66
616	-0.06	-0.12	-0.16	<i>0.85</i>
618	0.02	-0.15	0.06	<i>0.93</i>
630	-0.29	-0.41	-0.30	0.67
632	-0.05	0.01	-0.17	0.31
634	0.04	-0.11	0.12	0.75

Note. Values in italics met the criteria of $r \geq 0.8$, indicating similarity between known and coded samples.

TABLE 4
Pairwise Correlation Matrix between Known and Blinded Samples (3-day) Based on the Expression of Highly Discriminant Genes

Unknown samples	2-week known samples				24-h known samples			
	Clofibrate	Wyeth 14,643	Gemfibrozil	Phenobarbital	Clofibrate	Wyeth 14,643	Gemfibrozil	Phenobarbital
270	0.20	<i>0.87</i>	0.74	0.06	<i>0.81</i>	0.79	0.53	0.20
272	0.13	<i>0.87</i>	0.72	-0.02	<i>0.81</i>	<i>0.83</i>	0.53	0.14
274	0.08	<i>0.89</i>	0.74	-0.05	<i>0.83</i>	<i>0.84</i>	0.62	0.07
276	0.29	<i>0.88</i>	0.77	0.18	<i>0.83</i>	0.74	0.61	0.33

Note. Values in italics met the criteria of $r \geq 0.8$, indicating similarity between known and coded samples.

positive prediction regarding the classes of 12 out of 13 of the blinded samples. We were also successful at noting that 10 other blinded samples did not belong to the class of peroxisome proliferators, as evidenced by the lack of similarity in pattern to compounds in that class. A summary of classifications is listed in Table 5, which shows the list of blinded samples, our corresponding prediction, and their actual identity. The results show that, using the approach described, we had a 92.3% accuracy rate in class prediction. The animals that had a gene expression profile similar to phenobarbital were treated with high-dose phenytoin, a drug in the same class as phenobarbital. Similarly, the coded samples that were identified as being similar to samples from clofibrate- and Wyeth 14,643-treated

rats were from animals exposed to diethylhexylphthalate (DEHP), a known peroxisome proliferator. Finally, the samples that were noted to be definitely unrelated to peroxisome proliferators but weakly similar to phenobarbital were from animals treated with low doses of either phenytoin or hexobarbital. The one sample that appeared to be classified incorrectly was that of a hexobarbital-exposed animal as phenobarbital-exposed. Further investigation of interanimal variation is being conducted that might help to explain this inaccuracy.

After the decoding of the samples, we were interested in visualizing the gene expression profiles of all of the coded samples in the context of the discriminator genes. Cluster analysis was performed to visualize the grouping of known

TABLE 5
Classification of Blinded Samples Derived from Chemically Treated Rat Livers Based on Predictive Analyses of Gene Expression Profiles Generated Using cDNA Microarrays

Sample	Actual	Exposure	Prediction	Validity of call
612	Phenytoin 300 mg/kg/day	24-h	Not peroxisome proliferator	✓
616	Phenytoin, 300 mg/kg/day	24-h	Phenobarbital-like	✓
618	Phenytoin, 300 mg/kg/day	24-h	Phenobarbital-like	✓
3462	Phenytoin, 150 mg/kg/day	2-week	Not peroxisome proliferator	✓
3464	Phenytoin, 150 mg/kg/day	2-week	Not peroxisome proliferator	✓
3468	Phenytoin, 150 mg/kg/day	2-week	Not peroxisome proliferator	✓
672	Phenytoin, 300 mg/kg/day	2-week	Phenobarbital-like	✓
674	Phenytoin, 300 mg/kg/day	2-week	Phenobarbital-like	✓
676	Phenytoin, 300 mg/kg/day	2-week	Phenobarbital-like	✓
678	Phenytoin, 300 mg/kg/day	2-week	Phenobarbital-like	✓
270	DEHP, 1200 mg/kg/day	3-day	Clofibrate/Wyeth-like	✓
272	DEHP, 1200 mg/kg/day	3-day	Clofibrate/Wyeth-like	✓
274	DEHP, 1200 mg/kg/day	3-day	Clofibrate/Wyeth-like	✓
276	DEHP, 1200 mg/kg/day	3-day	Clofibrate/Wyeth-like	✓
4216	DEHP, 1200 mg/kg/day	2-week	Wyeth-like	✓
4218	DEHP, 1200 mg/kg/day	2-week	Wyeth-like	✓
630	Hexobarbital, 200 mg/kg/day	24-h	Not peroxisome proliferator	✓
632	Hexobarbital, 200 mg/kg/day	24-h	Not peroxisome proliferator	✓
634	Hexobarbital, 200 mg/kg/day	24-h	Not peroxisome proliferator	✓
688	Hexobarbital, 200 mg/kg/day	2-week	Phenobarbital-like	X
690	Hexobarbital, 200 mg/kg/day	2-week	Not peroxisome proliferator	✓
692	Hexobarbital, 200 mg/kg/day	2 week	Not peroxisome proliferator	✓
694	Hexobarbital, 200 mg/kg/day	2-week	Not peroxisome proliferator	✓

Note. ✓, valid; X, not valid.

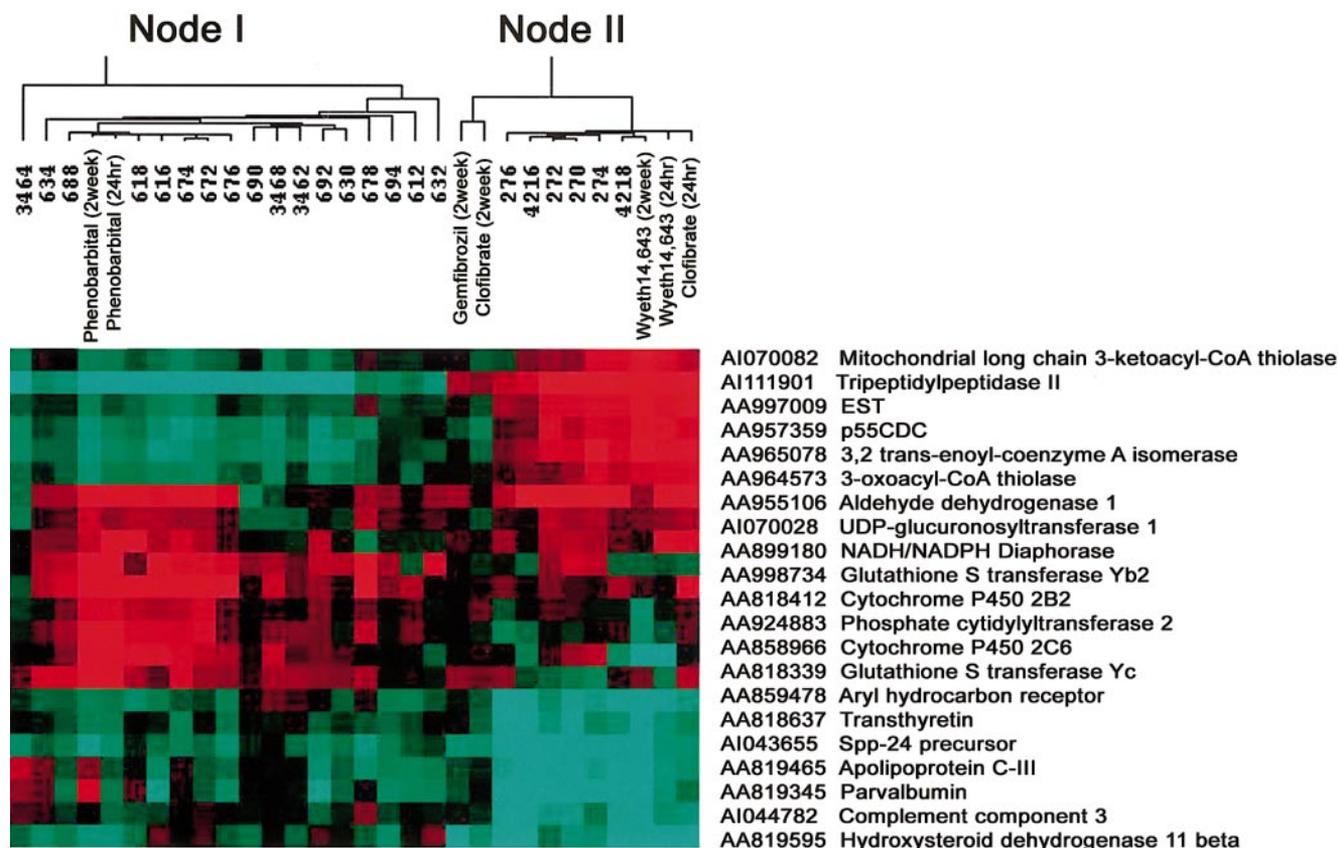


FIG. 2. Clustering diagram of samples in the study. The algorithm (Eisen *et al.*, 1998) was used to cluster the gene expression profiles for known and coded compounds for the set of derived discriminatory genes used in this study. The diagram illustrates that all of the compounds separate into 2 distinct nodes. Node I represents samples that are related to phenobarbital and Node II represents samples classified as peroxisome proliferators. Red indicates genes that are induced by treatment and green indicates repression of expression by treatment.

samples within unknowns according to the expression levels of the highly discriminant genes. Figure 2 shows the hierarchical dendrogram that clearly separated known and blinded samples into 2 major nodes according to their major mechanistic classes. One node (Fig. 2, node I) contained all of the animals exposed to the barbiturates phenobarbital, phenytoin, and hexobarbital. Upon further inspection of this node, one can see that all of the high-dose phenytoin animals (616, 618, 674, 672, 676) were tightly clustered with the phenobarbital, and that animals treated with low-dose or short-time phenytoin or hexobarbital were clustered slightly apart. The second major node (Fig. 2, node II) contained all of the animals that had been exposed to peroxisome proliferators.

DISCUSSION

This study illustrates the successful classification of coded RNA samples derived from the livers of chemically exposed animals. A unique feature of this study is that the classification was done in a blinded fashion. This provided a challenge, because we had to determine cutoffs for assignment of identity

($r > 0.8$) in the absence of being able to learn from the model we were developing. However, the success of our predictions demonstrates that if one can reduce the dimension of their data set with various statistical models, then it is possible to perform predictions for unknown compounds that are similar to compounds present in a database. We are now using the data set to develop more formal and rigorous models for prediction (i.e., neural networks) to make this task less arbitrary in future studies.

It is of interest to note that samples classified as phenobarbital-like (616, 618, 672, 674, 676, 678) were actually derived from rats exposed to high doses of phenytoin (5,5-diphenylhydantoin). Phenytoin and phenobarbital belong to the same pharmacological class of compounds that act as anticonvulsants with enzyme-inductive properties (Brodie 1992; Fitzsimmons *et al.*, 1990; Liu and Delgado 1995; Patsalos and Duncan 1993; Pichard *et al.*, 1990; Riva *et al.*, 1996), which attributes validity to the prediction made on the identity of these samples. Our findings corroborate numerous previous studies that have reported on the proximity of responses to phenytoin and phenobarbital in different biological models (Brodie 1992;

Fitzsimmons *et al.*, 1990). Likewise, coded samples that were correctly classified as similar to clofibrate or Wyeth 14,643 corresponded to rat livers exposed to DEHP (di-(2-ethylhexyl)phthalate). DEHP belongs to the peroxisome proliferator class of compounds (Lake *et al.*, 1986; Mitchell *et al.*, 1985) and produces a multitude of effects that are shared by other peroxisome proliferators such as clofibrate and Wyeth 14,643 (Crane *et al.*, 1990; Lake *et al.*, 1984).

We provided positive classification rather than the exact identity of the coded samples since no gene expression profiles corresponding to DEHP or phenytoin were present in our learning set/database. However, we were successful in classifying coded samples according to their pharmacological/toxicological effects or modes of action. A prediction on the classification of samples 630, 632, 634, 690, 692, 694, 3462, 3464, and 3468 was made by the definite assessment that none of these samples were similar to peroxisome proliferator compounds. The first 6 of these samples were derived from animals treated with hexobarbital, a compound that is structurally related to phenobarbital but is not carcinogenic in rodents and elicits a less potent enzyme-inducing response (Nims *et al.*, 1987). The last 3 samples corresponded to a 2-week exposure to a low dose of phenytoin and were negatively correlated with known peroxisome proliferator samples, but did not meet our stringent criteria to be classified as potential phenobarbital samples. If our database had contained expression profiles of rat livers exposed to low-dose phenobarbital, positive identification of those samples may have been possible. This highlights the importance of building multiple doses and time points into studies designated to populate a database developed for the purpose of screening and prediction.

Another challenge for this study was the lack of information regarding biological replicates among the blinded set. In our previous studies (Hamadeh *et al.*, 2002b) we utilized multiple animals for each dose and time group to determine what gene expression changes occurred robustly in all exposed animals vs. those that reflected variation between animals. In this prediction study, we did not know which samples represented biological replicates. This lack of replicate knowledge contributed to the difficulty for making our predictions. For example, we classified one of the hexobarbital-treated blinded samples (688) as being weakly similar to phenobarbital and did not positively classify sample 612 (phenytoin exposure) as similar to phenobarbital (Table 5). These calls might be due to inter-animal variation where a relatively higher- or lower-amplitude response in gene expression was evident in those particular samples, respectively, since their biological replicates were classified correctly.

Inherent in this data set are a variety of time- and dose-dependent, as well as independent, changes that might be exploited in further studies. By studying samples treated with phenytoin at the low and high dose, we found numerous genes that appeared to respond in a dose-dependent manner. *GST Yb2* (AA998732), *carboxylesterase precursor* (AI070587), *cyto-*

chrome p450 2C6 (AA858966), *palmitoyl-protein thioesterase* (AA818995) and *cytochrome p450 2B2* (AA818412) were among genes that were apparently induced in a dose-dependent fashion by phenytoin. Likewise, the expression levels of *diazepam-binding inhibitor* (AA925794), *parvalbumin* (AA819345), *growth hormone receptor* (AA819745), *p450 1A2* (AA924594), and *cytochrome p450 2C7* (AA818043) were repressed in a dose-dependent manner as a result of phenytoin exposure. If, in additional studies, these genes continue to appear to have dose- and time-independent responses, they may provide valuable identifiers for classifying compounds at unique dose or time points.

There were several challenges we encountered in this study. Ideally, the information housed in a database should be large so that one can interrogate this large sum of information with a relatively small query, to derive more knowledge from the database. However, we were challenged to interrogate a database with a large data set that outweighed the relevant database set that we could query against. Our query of 24 gene expression profiles, potentially belonging to 24 different chemicals, exceeded the database set of 4 chemicals that was used for the purpose of this study. Our learning set contained gene expression profiles corresponding to 24 h or 2 weeks of chemical exposure, however, 4 of the blinded samples were generated in rats exposed for 3 days. This required us to reduce the dimension of the data set (Bushel *et al.*, unpublished data) and to find time-independent, highly discriminative genes that helped to separate different compounds, so that we could query 24-h and 2-week data with 3-day profiles (Table 1). This investigation demonstrates the first example of a successful query of a database with gene expression profiles to predict classification of unknown compounds of this number.

In summary, this work illustrates the successful prediction of properties of blinded samples using gene expression profiling. It demonstrates that large gene expression profile databases will be able to be successfully queried to help classify unknown compounds from exposed tissues. It also highlights the importance for beginning, even at this early stage, to develop analysis models for these types of data. Our analyses informed us on important considerations for our experimental study designs in the future. It is now foreseeable that in the future, gene expression profiling or other high density genomics analyses will prove valuable in the screening of compounds for mechanistic classification in a high-throughput fashion. In turn, this will ultimately better our understanding in selecting chemicals for advanced stages of target testing in commercial settings, and allow the advancement of predictive information on uncharacterized human health hazards.

REFERENCES

- Alaiya, A. A., Franzen, B., Hagman, A., Silfversward, C., Moberger, B., Linder, S., and Auer, G. (2000). Classification of human ovarian tumors using multivariate data analysis of polypeptide expression patterns. *Int. J. Cancer* **86**, 731–6.

- Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., *et al.* (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511.
- Ashby, J. (1994). International Commission for Protection Against Environmental Mutagens and Carcinogens. Two million rodent carcinogens? The role of SAR and QSAR in their detection. *Mutat. Res.* **305**, 3–12.
- Ashby, J., and Paton, D. (1993). The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures. *Mutat. Res.* **286**, 3–74.
- Brodie, M. J. (1992). Drug interactions in epilepsy. *Epilepsia* **33**(Suppl. 1), S13–22.
- Bushel, P. R., Hamadeh, H., Bennett, L., Sieber, S., Martin, K., Nuwaysir, E. F., Johnson, K., Reynolds, K., Paules, R. S., and Afshari, C. A. (2001). MAPS: A microarray project system for gene expression experiment information and data validation. *Bioinformatics* **17**, 564–565.
- Chen, Y., Dougherty, E. R., and Bittner, M. L. (1997). Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Optics* **2**, 364–374.
- Crane, D. I., Zamattia, J., and Masters, C. J. (1990). Alterations in the integrity of peroxisomal membranes in livers of mice treated with peroxisome proliferators. *Mol. Cell. Biochem.* **96**, 153–161.
- Cunningham, A. R., Klopman, G., and Rosenkranz, H. S. (1998). Identification of structural features and associated mechanisms of action for carcinogens in rats. *Mutat. Res.* **405**, 9–27.
- DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A., and Trent, J. M. (1996). Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* **14**, 457–460.
- 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.
- Enslin, K., Gombar, V. K., and Blake, B. W. (1994). International Commission for Protection Against Environmental Mutagens and Carcinogens. Use of SAR in computer-assisted prediction of carcinogenicity and mutagenicity of chemicals by the TOPKAT program. *Mutat. Res.* **305**, 47–61.
- Fitzsimmons, W. E., Ghalib, R., and Kaizer, H. (1990). The effect of hepatic enzyme inducers on busulfan neurotoxicity and myelotoxicity. *Cancer Chemother. Pharmacol.* **27**, 226–228.
- Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., Bloomfield, C. D., and Lander, E. S. (1999). Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537.
- Hamadeh, H. K., Amin, R. P., Paules, R. S., and Afshari, C. A. (2002a) An overview of toxicogenomics. *Curr. Issues Mol. Biol.*, **4**, 45–56.
- Hamadeh, H. K., Bushel, P. B., Paules, R., and Afshari, C. A. (2001). Discovery in toxicology: Mediation by gene expression array technology. *J. Biochem. Mol. Toxicol.* **15**, 231–242.
- Hamadeh, H. K., Bushel, P. R., Jayadev, S., Martin, K., DiSorbo, O., Sieber, S., Bennett, L., Tennant, R., Stoll, R., Barrett, J. C., Blanchard, K., Paules, R. S., and Afshari, C. A. (2002b). Gene expression analysis reveals chemical-specific profiles. *Toxicol. Sci.* **67**, 219–231.
- Johnson, R. A., and Wichern, D. W. (1998). *Applied Multivariate Statistical Analysis*, 4th ed. Prentice-Hall, Upper Saddle River, NJ.
- Kerr, M. K., and Churchill, G. A. (2001). Statistical design and the analysis of gene expression microarray data. *Genet. Res.* **77**, 123–128.
- King, R. D., and Srinivasan, A. (1996). Prediction of rodent carcinogenicity bioassays from molecular structure using inductive logic programming. *Environ. Health Perspect.* **104**(Suppl. 5), 1031–1040.
- Klopman, G., and Rosenkranz, H. S. (1994). International Commission for Protection Against Environmental Mutagens and Carcinogens. Approaches to SAR in carcinogenesis and mutagenesis. Prediction of carcinogenicity/mutagenicity using MULTI-CASE. *Mutat. Res.* **305**, 33–46.
- Lake, B. G., Gray, T. J., and Gangolli, S. D. (1986). Hepatic effects of phthalate esters and related compounds—*in vivo* and *in vitro* correlations. *Environ. Health Perspect.* **67**, 283–290.
- Lake, B. G., Tredger, J. M., Gray, T. J., Stubberfield, C. R., Hodder, K. D., Gangolli, S. D., and Williams, R. (1984). The effect of peroxisome proliferators on the metabolism and spectral interaction of endogenous substrates of cytochrome P450 in rat hepatic microsomes. *Life Sci.* **35**, 2621–2626.
- Li, L., Darden, T. A., Weinberg, C. R., and Pedersen, L. G. (2001). Gene assessment and sample classification for gene expression data using a genetic algorithm/k-nearest neighbor method. *Combinator. Chem. High Throughput Screening* **4**, 727–739.
- Liu, H., and Delgado, M. R. (1995). Interactions of phenobarbital and phenytoin with carbamazepine and its metabolites' concentrations, concentration ratios, and level/dose ratios in epileptic children. *Epilepsia* **36**, 249–254.
- Lovett, R. A. (2000). Toxicogenomics. Toxicologists brace for genomics revolution. *Science* **289**, 536–537.
- Mitchell, A. M., Lhuguenot, J. C., Bridges, J. W., and Elcombe, C. R. (1985). Identification of the proximate peroxisome proliferator(s) derived from di(2-ethylhexyl) phthalate. *Toxicol. Appl. Pharmacol.* **80**, 23–32.
- Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996). *Applied Linear Statistical Models*, 4th ed. Irwin, Chicago.
- Nims, R. W., Devor, D. E., Henneman, J. R., and Lubet, R. A. (1987). Induction of alkoxyresorufin O-dealkylases, epoxide hydrolase, and liver weight gain: Correlation with liver tumor-promoting potential in a series of barbiturates. *Carcinogenesis* **8**, 67–71.
- Nuwaysir, E. F., Bittner, M., Trent, J., Barrett, J. C., and Afshari, C. A. (1999). Microarrays and toxicology: The advent of toxicogenomics. *Mol. Carcinog.* **24**, 153–159.
- Patsalos, P. N., and Duncan, J. S. (1993). Antiepileptic drugs. A review of clinically significant drug interactions. *Drug Safety* **9**, 156–184.
- Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T., Pergamenschikov, A., Williams, C. F., Zhu, S. X., Lee, J. C., Lashkari, D., Shalon, D., Brown, P. O., and Botstein, D. (1999). Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 9212–9217.
- Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L., *et al.* (2000). Molecular portraits of human breast tumours. *Nature* **406**, 747–752.
- Pichard, L., Fabre, I., Fabre, G., Domergue, J., Saint Aubert, B., Mourad, G., and Maurel, P. (1990). Cyclosporin-A drug interactions. Screening for inducers and inhibitors of cytochrome P450 (cyclosporin-A oxidase) in primary cultures of human hepatocytes and in liver microsomes. *Drug Metab. Dispos.* **18**, 595–606.
- Riva, R., Albani, F., Contin, M., and Baruzzi, A. (1996). Pharmacokinetic interactions between antiepileptic drugs. Clinical considerations. *Clin. Pharmacokinet.* **31**, 470–493.
- Waring, J. F., Ciurlionis, R., Jolly, R. A., Heindel, M., and Ulrich, R. G. (2001a). Microarray analysis of hepatotoxins *in vitro* reveals a correlation between gene expression profiles and mechanisms of toxicity. *Toxicol. Lett.* **120**, 359–368.
- Waring, J. F., Jolly, R. A., Ciurlionis, R., Lum, P. Y., Praestgaard, J. T., Morfitt, D. C., Buratto, B., Roberts, C., Schadt, E., and Ulrich, R. G. (2001b). Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicol. Appl. Pharmacol.* **175**, 28–42.