Chemistry & Biology, Vol. 11, 161–171, February, 2004, ©2004 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2004.02.003

Toxicogenomics in Predictive Toxicology in Drug Development

Review

Laura Suter,[1,*] Lee E. Babiss,[2] and Eric B. Wheeldon[2]
[1]Department of Non-Clinical Drug Safety
F. Hoffmann-La Roche, Ltd.
4070 Basel
Switzerland
[2]Department of Preclinical R&D
F. Hoffmann-La Roche, Inc.
Nutley, New Jersey 07110

Summary

The goal of toxicology is the assessment of possible risk to man. An emerging technology with the potential to have a major impact on risk assessment is toxicogenomics. In this review, we provide an overview of the many possibilities for toxicogenomics including technology platforms, data interpretation, and regulatory perspective and we give examples of toxicogenomics investigations. Toxicogenomics is a powerful tool for compound classification, for mechanistic studies, and for the detection of toxicity markers. Thus, toxicogenomics helps in the extrapolation of findings across species and increases predictability. Biomarkers are valuable in the evaluation of compounds at earlier development phases, improving clinical candidate selection. Caution regarding the interpretation of the results is still necessary. Nevertheless, toxicogenomics will accelerate preclinical safety assessments and improve the prediction of toxic liabilities, as well as of potential risk accumulation for drug-drug or drug-disease interactions.

Traditional Approaches for Assessing Toxicity

Toxicology, the study of adverse effects of chemicals on living organisms, has traditionally been evaluated by the dosing of animals to define well-established cytologic, physiologic, metabolic, and morphologic endpoints. The evaluation of the risk to humans can not be performed in human individuals initially and thus must be derived from studies performed in other species. Typically, rodents are used to identify toxic substances such as carcinogens, reproductive toxins, and neurotoxins. Follow-up studies in nonrodent species can then be used to further define the effects of low doses as well as species extrapolation and mechanism of action.

Although it is well recognized that intact animals are needed to reflect physiologic changes and mirror the effects of chronic dosing, such studies have disadvantages. Animals may not be fully predictive of the response in humans due to species variation in physiology, anatomy, and metabolism. Also, toxicology studies require large numbers of animals to allow statistically significant conclusions to be drawn. Nevertheless, these numbers are still very small compared to the human population potentially at risk. In order to compensate for this relatively small sample size in these animal studies, the future risk to humans at therapeutic dosages is inferred by giving large doses of compound to these groups of animals. Finally, depending on the anticipated duration of exposure in the population, studies of up to two years are currently mandated to determine the carcinogenic potential. Thus, the current approach to toxicologic testing is costly, in terms of time, labor, compound synthesis and, not least, the large numbers of animals. Any approach that offers savings in (any of) these areas would represent a significant advance in the development of new drugs.

Problems of Attrition Facing the Pharmaceutical Industry Today

Developing new drugs is becoming more expensive and more difficult. There has been a noticeable shift in the fortunes of the pharmaceutical industry since the 1990s due to two main factors. First, there is the increasing cost of developing new drugs. Several estimates put the cost of bringing a new drug to market at around $800 million—up from around $230 million in 1987. Investment in drug development in the US has tripled to over $30 billion over the last 10 years but with fewer drugs coming onto the market. The number of New Drug Applications (NDA) made to the FDA peaked at 131 in 1996 but had fallen to 78 by 2000. Second, there is the continuing problem of adverse drug reactions (ADR) seen in the postmarketing period, resulting in withdrawal or restricted use of the drug. Thus, 16 of 548 marketed drugs were withdrawn in the period 1975–1999, while another 45 of these 548 drugs had warning labels added, restricting the therapeutic indications and so limiting the potential market. Any advance in technology that can shorten development times and/or more accurately reflect the liabilities of a drug once it enters the general population would be of tremendous benefit.

It should be noted that the productivity seen today is a reflection of the strategies that the pharmaceutical companies put in place 10–15 year ago. As such, the industry has begun to reinvent the process of drug discovery by factoring in new technological advances in parallel with general process improvements and new management models. As a consequence, we are beginning to see trends in improvement, given the quantity of new clinical candidates, first-time studies in humans, and the indication of improved success rates.

Success of development depends in part on the rigorous attrition of unsuitable compounds in the early stages. In this early phase, successes in screening and selection are relatively quick and cheap to obtain. Beyond this phase, experimental designs (including clinical trials) become increasingly lengthy and expensive. Traditionally, these expensive later phases have been considered to be the most reliable for the selection of the best compounds. Genomics offers the possibility of moving the selection process upstream, bringing a new

*Correspondence: laura.suter-dick@roche.com

Table 1. Survey of Reasons for Failure of Compounds in Development

Industry Median	Preclinical	Phase I	Phase II	Phase III	Registration
Clinical Safety	0.5%	27.9%	13.4%	9.8%	30.0%
Efficacy	5.6%	17.5%	52.0%	72.5%	20.0%
Formulation	5.1%	5.8%	1.6%	0.0%	0.0%
Market potential	6.2%	3.9%	7.9%	3.9%	30.0%
PK/bioavailability	11.8%	14.9%	2.4%	0.0%	0.0%
Strategic	14.4%	12.3%	13.4%	5.9%	20.0%
Resources	1.5%	1.3%	0.8%	3.9%	0.0%
Toxicology	44.1%	10.4%	2.4%	3.9%	0.0%
Cost of goods	1.5%	1.3%	0.0%	0.0%	0.0%
Unknown	7.2%	1.3%	4.7%	0.0%	0.0%
Other	2.1%	3.2%	1.6%	0.0%	0.0%
Number of projects	195	154	127	51	10

and rigorous filtering process into the very early low-cost phase. Table 1 illustrates the importance of toxicology in the preclinical phase of drug development. During this stage of development, toxicity is the major cause of failure, causing over 40% of failures. Later, efficacy in clinical trials becomes the leading cause of attrition—approximately 75% by Phase 3. Thus, recognizing toxicity earlier is critically important if the early pipeline is to be filled with more of the most promising candidates so that these can be developed most efficiently with no extra cost in resource.

Concept of Toxicogenomics

Toxicogenomics applies genomics concepts and technologies to study adverse effects of chemicals. These studies use global gene expression analyses to detect expression changes that influence, predict, or help define drug toxicity. Technological advances have enabled scientists to simultaneously analyze thousands of genes of several species, including humans and rodents, quickly and in a reproducible manner. By evaluating and characterizing differential gene expression after exposure to drugs, it is possible to use complex expression patterns to predict toxicologic outcomes and to identify mechanisms involved with or related to the toxic event. In short, the technology now exists to potentially revolutionize toxicity testing.

Toxicogenomics thus combines conventional toxicology with the emerging technologies of genomics and bioinformatics. Gene and protein expression respond specifically to external stimuli such as pathological conditions or exposure to drugs. The corresponding genomic and proteomic technologies thus provide a new way of understanding biological systems and their response to toxic insult. This leads to a better understanding of the mechanisms of toxicity, by the identification of toxicity-related gene expression signatures (fingerprints) and the prediction of the toxic potential of unknown compounds by comparing their gene-expression profiles to the fingerprints of known, similar compounds. In addition, the identification of "toxicity-related" genes, together with the rapidly growing understanding of the human genome, is providing a basis for identifying and characterizing sequence variations in genes that might affect responses to chemicals. This is already having a great impact in pharmacology and toxicology, since it allows the prediction/differentiation of species-specific responses and also the identification of populations of responders and nonresponders [1]. The most optimistic estimates predict that the replacement of traditional methods of toxicology by toxicogenomics could eventually shorten the safety assessment of a new chemical entity from years to days and reduce costs by an estimated factor of 4 to 6 times [2]. A more realistic picture with the data currently available suggests that toxicogenomics will reduce failure rates by helping select the right compounds for development early on and by accelerating toxicology testing and identifying suitable biomarkers amenable to screening using the generated data [3]. Currently, there seems to be general agreement that this technology will greatly accelerate the detection of toxic liabilities by replacing long-term exposure experiments for the selection of clinical candidates. However, the technology is unlikely ever to be able to predict idiosyncratic ADR.

Aim of Predictive Toxicology: Lead Selection and Liability Identification

ADR is a major problem from a public health perspective as well as for the development of new medicines. In a meta-analysis of literature, Lazarou et al. showed that ADR is between the fourth and sixth cause of death in the United States, accounting for more than 100,000 deaths in 1994 [4]. Failure of compounds in late preclinical development or, even worse, in the clinic, represents a very important economic burden for the pharmaceutical industry. The current pharmaceutical paradigm concentrating efforts on enhancing efficacy is inherently inefficient, since increased potency does not imply reduced toxicity. Indeed, more potent compounds might lead to more severe nonspecific effects, increasing the likelihood of failure due to toxicity. Thus, to increase productivity and minimize failure, efficacy and safety would ideally be determined simultaneously at very early stages [5]. In order to achieve this, a change in the traditional drug discovery process must take place, enabling scientists to integrate toxicology in the earlier discovery phases. Thus, toxicity would ideally be detected earlier at the time of lead selection or shortly thereafter.

Overview of the Approaches: Technologies and Platforms

The concept of using gene and protein expression analysis for mechanistic and predictive toxicology is not completely new. However, it is the amount of information that can be gathered using genomics tools that has

transformed molecular toxicology. It was less than 50 years ago that the structure and function of DNA was deduced, and by the end of 2002 the genomes of 800 organisms, including the human, had been nearly fully sequenced [6]. Alongside the major genome sequencing projects, the technological progress led to the development of microarrays allowing the simultaneous monitoring of the expression of thousands of genes and leading to a rapid progress in functional genomics research. It is now expected that gene expression analysis will not only reveal the pharmacological action of a compound, but also provide insights into possible safety issues. Two approaches have been followed to better understand and avoid ADR in the future: (1) the identification of genes or gene variations (SNPs) that put specific populations at risk of ADR to a given drug, and (2) the understanding of the molecular mechanisms (gene and protein expression) underlying toxicity. The first approach has provided evidence supporting a role for polymorphisms in ADR [7–9]. Using the second approach (toxicogenomics), a substantial amount of data has been generated in animal models with known toxicants, mainly hepatotoxins, proving that gene expression analysis can provide information to allow classification of compounds according to their mechanism of toxicity as well as identifying cellular pathways related to the toxic event [10–16]. More recently, this global gene expression analysis has been applied to the evaluation of nephrotoxicity [17, 18], genotoxicity [19], and testicular toxicity [20].

Several technologies to characterize gene expression are currently available and additional approaches are still being developed. It is outside the scope of this review to give a full account of microarray technologies and providers, but it might be useful to give the reader an idea of the most commonly used platforms. Currently, widely used toxicogenomics platforms are DNA microarrays able to monitor the expression level of a large number of known genes simultaneously. There are several providers of microarrays, but the basic principle of DNA arrays is to attach a probe to a solid support such as glass, nylon, or gel matrix [21–23]. Probes are either oligomeres (16- to 70-mers, depending on the platform) or 500–1500 bp cDNA sequences selected to be specifically complementary to the sequence that needs to be interrogated. These probes can either be directly synthesized onto the support (Affymetrix), spotted by fluid microdispensers (Clontech, Agilent, ABI, and several in-house array designs from research institutes and pharmaceutical companies), or targeted electronically to the appropriate spot coordinates (Nanogen, Matrixarray). The microarrays are incubated with the samples at conditions appropriate to allow hybridization between the target sample and the immobilized probe. The detection of the hybridization is usually accomplished through fluorescent staining, since the samples are labeled before hybridization. Depending on the platform, the expression level for a gene is given as an absolute intensity value or as a relative ratio with respect to a baseline sample. The number of genes on a single array varies according to the chip design. Low-density arrays comprise a few hundred genes, while high-density chips contain probes specific for thousands of target genes. In addition to microarrays, kinetic RT-PCR (Taqman), differential display, amplified fragment length polymorphism (AFLP), serial analysis of gene expression (SAGE), and Northern blots are also frequently employed methods for the analysis of gene expression [24].

Databases and Data Analysis

Predictive toxicogenomic studies usually compare the gene expression patterns elicited by chemicals with unknown toxic potential to the profiles of model compounds with known toxicity. Thus, for the use of toxicogenomics as a predictive tool, the prior knowledge of gene expression patterns related to toxicity is absolutely necessary. Consequently, this approach depends on the availability of a reference gene expression database (DB). Differential expressions of gene signatures are thus derived by analyzing expression levels of the compound under scrutiny and the compounds in the DB. A high-quality DB and robust software with appropriate algorithms for the comparison of complex gene expression fingerprints are vital for the interpretation and utilization of toxicogenomics data [25, 26].

Several toxicogenomics DBs are currently being built. The main focus is on the liver, since due to its physiological functions this organ is highly exposed to xenobiotics and is thus an important target organ for toxicity. There are several commercially available DBs such as those provided by GeneLogic, CuraGen, Iconix, and Phase I. A review of the currently available companies has recently been published [2]. These vendors offer several modes of subscription for pharmaceutical companies and usually include varying amounts of microarray data on organs of interest (liver, kidney, etc.) and bioinformatic tools incorporating predictive algorithms. In addition to these commercially available DBs, most of the major pharmaceutical companies now have internal toxicogenomics initiatives. Results obtained from these initiatives with model compounds reach the public domain through presentations at meetings and publications [16, 18, 23, 27–29], while results on development compounds remain largely unpublished. In addition to the commercially available DBs and those generated by pharmaceutical companies, there are some institutional efforts to create publicly available DBs. In 2000, the National Institute of Environmental Health Sciences (NIEHS) established the National Center for Toxicogenomics (http://www.niehs.nih.gov/nct/home.htm). Also, the International Life Science Institute (ILSI) created in 1999 a "Technical Committee on Application of Genomics to Mechanism-Based Risk Assessment" (http://www.ilsi.org). This committee is currently working together with the European Bioinformatics Institute (EBI, http://www.ebi.ac.uk/) to create a toxicogenomics database within the more general gene expression DB from EBI called ArrayExpress (from the Microarray Gene Expression Data Society [MGED]). These DBs are intended to provide the framework to make microarray data, including toxicogenomics data, publicly available. In addition, EBI has made a major effort to standardize the description and annotation of microarray data and has come up with the Minimal Information about Microarray Experiments (MIAME) standard to enable scientist to compare results obtained at different centers [30, 31].

As regards actual data analysis, the massive amount of genomics data that has been generated has given

biostatisticians a challenge that has yet to be resolved. On the one hand and due to the relatively high costs, most scientists minimize the number of biological replicates and sometimes even resort to pooling samples to minimize the number of microarrays. On the other hand, a typical microarray experiment generates thousands of data points. Additional confounding factors are time and dose responses. Moreover, groups of genes might be coregulated by a stimulus while redundancy in microarray design leads to multiple probes examining the same mRNA sequence. Also, gene expression for certain genes might be switched on or off, rather than following a normal distribution pattern. Thus, the data is very complex and highly multivariate. Several statistical methods to analyze microarray data are available, but there is no single paradigm that fits all situations [6, 26, 32–36]. Most commercial providers of array readers supply software that allows basic analysis of the arrays to be performed and exported to calculation tables such as Microsoft Excel. More sophisticated analysis is usually performed by additional software such as Eisen Clustering Tool (Stanford University), GeneSpring (Silicon Genetics), Spotfire, Mineset, SIMCA-P (Umetrics), Rosetta Resolver (Merk), GeneData Expressionist, which has been tailored to deal with microarray data from a variety of platforms. Several pharmaceutical companies have recognized the need for their own in-house microarray analysis tools. Finally, toxicogenomics DB providers such as GeneLogic, CuraGen, and Iconix supply software tools as part of their subscription packages. Each of these software programs offer one or several analysis methods and tend to leave the biologist with the difficult choice of which one to employ. Rigorous attention to data quality and transformation steps is crucial and will have a major impact on any analysis. Among the statistical methods most frequently used for microarray analysis are conventional parametric or nonparametric statistics, used to analyze the results gene by gene. Unsupervised analysis tools such as clustering algorithms, principal component analysis, and self-organizing maps are used to determine if gene expression patterns allow the discrimination of natural subpopulations such as treated/untreated or healthy/diseased. In order to use the knowledge stored in the toxicogenomics DBs, several supervised methods such as discriminant analysis, neural networks, nearest neighbors, and support vector machines are based on algorithms that learn from the training data set in the DB and use previously acquired knowledge to classify unknown compounds.

Current Approaches

Regulatory Perspective

In spite of the hopes placed on gene expression analysis for the eventual prediction of toxicity, there is considerable apprehension within the pharmaceutical industry regarding the use that regulatory authorities could make of this type of data. The technologies used are reproducible and robust, but are also known to provide a certain amount of false positive and negative results. Also, the ability to investigate the whole transcriptome does not necessarily imply that we understand all the answers. In truth, our understanding of the biological significance of gene expression changes is still very incomplete. An additional source of concern is the sensitivity of gene expression profiling. If toxicogenomics results in higher sensitivity and earlier detection of possible ADR, we might well face the need for a new definition of NOEL and NOAEL, while not understanding the relevance of these more sensitive findings.

Such issues have made the pharmaceutical industry reluctant to use this new technology on compounds in (advanced) development phases, since they would be obliged to submit difficult-to-interpret data to the authorities, which might cause delay in the development program. In order to try to assuage this concern, the FDA introduced the concept of "safe harbor" at a joint meeting with the Pharmaceutical Industry held in May 2002 [37]. The Safe Harbor concept was designed to nonvalidated, exploratory, hypotheses-generating data from toxicogenomics, pharmacogenomics, and pharmacogenetic testing of animals or human subjects. This principle was clearly intended to encourage companies to present these kinds of data to the FDA, thus fostering dialog by allowing data sharing and improving collaboration between the FDA and the pharmaceutical industry. However, some of the points originally included in the Safe Harbor draft document were not fully clear regarding the wider legal and regulatory framework. Among the major points raised by industry are the definition of Safe Harbor data and the sharing of such data (legal problems regarding informed consents and confidentiality of such data). Also, there is an ethical issue of the FDA ignoring data that have been generated on compounds that may be administered to humans. An additional point of concern is the handling of the transition from exploratory to confirmed data: how and when does Safe Harbor finish and what are the possible implications? A recent public meeting of the Science Advisory Board of the FDA on April 9, 2003, provided additional clarification regarding the submission and interpretation of toxicogenomics data. The CDER (Center for Drug Evaluation and Research) is currently gaining insight and understanding of microarray data by performing experimental work and by collaborating with several commercial providers and users. Among these efforts, "mock" submissions of toxicogenomics data have been submitted and then used as a basis for discussion. Additional clarification on the current policy of the FDA refers to three categories of pharmacogenomics data: (1) not required to be submitted (e.g., early drug development data); (2) data required to be submitted but without regulatory impact; and (3) data required to be submitted and with regulatory impact (e.g., metabolizer genotype used for dose selection).

The position of the FDA/CDER regarding submission of genomics data has been further clarified in a subsequent meeting of the Advisory Committee for Pharmaceutical Science (Pharmacology Toxicology Subcommittee) held on June 10, 2003, but discussions on the "voluntary genomic data submission" without regulatory impact are still ongoing. Similarly to the FDA, EMEA (European Agency for the Evaluation of Medicinal Products) has established an ad hoc pharmacogenetics working group. A concept paper was released in January

2003 (CPMP/4445/03), stating that EMEA/CPMP has created briefing sessions as an informal forum for discussion between sponsors and regulators at EU level.

In spite of these ongoing discussions, the separation of the data categories remains dependent on interpretation and will probably be readjusted with growing experience. Also, pharmacogenomics seems to include quite different kinds of data, from the genotyping of patients to gene expression data obtained in preclinical animal studies. The FDA is planning to clarify these points further and to establish an Interdisciplinary Pharmacogenomics Review Group (IPGRG) to perform periodically public reevaluation of this decision making process. To date, an additional discussion point is the standardization of protocols to ensure data comparability. It is expected that the authorities will release a guidance document soon. Despite this very active discussion, pharmaceutical companies are already starting to generate toxicogenomics data on compounds in development. In this context, it is worth mentioning that the process of submission of exploratory data without regulatory consequences is currently considered as an addition to the conventionally used toxicity testing. It is going to take a considerable amount of time until validation of these new methods has been completed. It is too early to address the question whether or not toxicogenomics will be able to replace some of the conventional toxicology tests needed for a NDA, but it is hoped that long-term studies might be considerably shortened by an increase in sensitivity. For now the main advantage of a toxicogenomics submission from a regulatory point of view is the insight it provides into mechanisms and the identification and validation of suitable biomarkers.

ILSI Initiative

The International Life Sciences Institute's (ILSI) Health and Environmental Sciences Institute (HESI), a nonprofit research and educational organization, has created a Technical Committee on Application of Genomics to Mechanism-Based Risk Assessment (http://www.ilsi.org). The mission of this Genomics Committee is to "advance the scientific basis for the development and application of genomic methodologies to mechanism-based risk assessment; to address scientific issues relating to the use of these new technologies as a means for understanding toxic response and mechanisms; and to provide a scientific forum for a consensus-based approach to interpreting and applying these data." This committee was established in 1999 as a collaborative program between industrial, government, and academic institutions including nearly 40 laboratories in the US, Europe, and Japan [38]. This international collaborative effort is evaluating relevant experimental protocols and technologies with the main foci of interest being hepatotoxicity, nephrotoxicity, and genotoxicity. For each of these topics, working groups are conducting experimental studies with known toxicants and evaluating them using microarray technologies, as well as conventional toxicology endpoints. In addition, the Committee has formed a database working group that is working together with EBI (European Bioinformatics Institute) to make the obtained data publicly available.

Some of the data obtained by the working groups has already been published by subgroups of participants [39], and a special issue of Environmental Health Perspectives with several articles has been published [40–52]. In addition, a subset of the data produced and analyzed by the hepatotoxicity working group is being submitted to the FDA/CDER in the form of a "mock" submission. This exercise has shown the critical need for rigorous attention to experimental design including methodology, data evaluation, and the need to reduce sources of experimental and intersite variability. Also, the use of a variety of platforms for the analysis of gene expression has shown that there are unique aspects to each platform. Thus, the findings of different platforms do not fully overlap, but the combination of the outcomes enriches the data set [38].

Successes

So far, toxicogenomic data on proprietary compounds in development are scarce. At the First FDA-Pharmacogenomics Workshop held in May 2002, data supporting investigative and mechanistic toxicogenomics applications were presented. A typical set of data showed the identification of gene expression markers indicative of acute phase response in isolated mesenteric arteries from rats with fenoldopam-induced vasculitis [37]. Another example identified patterns of gene expression showing that administration of a 5-lipooxygenase inhibitor repressed synthesis of cholesterol and that lens proteins were targets of drug-induced cataractogenesis [37]. In a recent publication by our group, gene expression analysis was used to distinguish two compounds with comparable pharmacology but with distinct toxicological profiles. The results provided possible markers for compound-induced steatosis, which were amenable to testing using higher throughput methods such as PCR [53]. These early examples provide evidence that toxicogenomics can give insights into toxicological mechanisms and affected pathways.

Cancer is one of the major disease areas in which genomics investigations have made considerable advances, probably due to the close relationship between the disease and genetic factors. An early report described the development and progression of malignant melanoma using microarrays [54]. More recent publications present results differentiating neoplastic and normal tissue as well as differentiating types of cancers and generating predictive markers using gene expression profiles [1]. Golub et al. demonstrated gene expression profiles characteristic of two types of leukemia (ALL and AML) using the technique of class prediction. This approach provided an improvement in early diagnostic techniques that led in some cases to an adjustment of the therapy with a direct benefit for the patient [55]. Also, studies on diffuse large cell lymphoma (DLBCL) using microarrays showed that variability in the disease progression and survival outcome could be correlated with gene expression data [56, 57]. The group of John Weinstein at the NIH has published results classifying tumor cell lines derived from a variety of tumors according to their gene expression patterns and to their response to anticancer drugs, providing an important set of data on gene-drug relationships [58–60]. Microarrays are also being employed for investigations aiming to identify diagnostic biomarkers for different types of cancer. As an example, gene expression analysis of

adrenocortical carcinomas has recently provided a set of genes that are likely to be specific of malignant lesions and are therefore potential diagnostic biomarkers [61]. Hence, in the field of cancer research, the analysis of differential gene expression has helped to increase the diagnostic power and the prediction of the clinical outcome as well as to adapt the therapy, providing direct benefits to patients.

In addition to the use of gene expression analysis in the field of cancer research and diagnosis, increasing interest is emerging in the use of microarrays for the detection of specific genotypes. Currently, DNA testing is moving rapidly into widespread use to allow better and more individualized choices of treatments. Variations in cytochrome P450 are known to play a major role in drug response and have been examined in relation to safety and efficacy of drugs for years. A microarray-based genotyping assay allows the simultaneous detection of over two dozen allelic variants affecting CYP450 enzyme activity, including those caused by SNPs, frame shifts, multiple base repeats, and even complete gene deletion or duplication. The new AmpliChip, developed by Roche and Affymetrix, tests the most common variations in two genes, CYP2D6 and CYP2C19, which play roles in the metabolism of about 45% of the prescription drugs on the market. The metabolic analyses performed by the chip will eventually offer practitioners a tool to categorize patients according to their metabolic type, thereby aiding them in prescribing more effective dosages of medication and avoiding adverse side effects. In spite of these promising developments, several obstacles need to be overcome before achieving the full potential of this kind of test. Technical challenges such as automation and reduction of sample processing time need to be improved. Additionally, education of the general practitioners regarding the technology and its benefits to patients is necessary. However, it is expected that the technology will develop into a fully automated, certified diagnostic tool within the next 3–5 years.

Application of Toxicogenomics at Roche

Compound Classification Using Gene Expression Profiles

As outlined above, it has been widely recognized that a DB with known reference toxicants is necessary for accurate compound classification. The toxicogenomics group at Roche has generated a database containing hepatic gene expression data using the Affymetrix microarray. Model compounds, administered to male rats, were categorized into three main groups: direct acting (necrosis/apoptosis), steatotic (lipidosis), and cholestatic (bile acid transport impairment). Preliminary validation of the data was performed by analyzing similarities among fingerprints obtained from several compounds within the same category. We determined the effect of five hepatotoxic compounds with a similar mechanism of toxicity ("direct acting") on hepatic gene expression profiles and determined a direct acting fingerprint by including genes regulated by at least 4 of these 5 compounds. The specificity of these fingerprints for the common mechanism of toxicity was assessed by comparing the effects of steatotic and cholestatic and nontoxic compounds on the same genes. Figure 1 depicts a heat map representing the induction or repression of the selected genes by a variety of compounds. The results proved that gene expression profiles can be used to distinguish direct acting compounds from other types of compounds. Furthermore, the assessment of the modulation of gene expression in two experiments performed independently with the same compound showed good reproducibility of the gene expression profiles. Hence, gene expression profiles in liver after exposure to hepatotoxins are reproducible and characteristic of classes of toxins.

In addition to the in vivo DB, in vitro approaches are also being conducted at Roche. Results from these in vitro models, including primary cultures and cell lines, show significant discrepancies when compared to in vivo results [62]. Briefly, it has been shown that the tested hepatic cell lines do not express many of the crucial metabolic enzymes and are, therefore, not comparable to in vivo systems or to the primary in vitro models. Primary cell culture models, on the other hand, are basically unstable, since hepatocytes undergo a dynamic process including an initial phase of isolation stress followed by a dedifferentiating phase ending in cell death. These time-related changes increase variability and so might mask some toxicologically relevant effects. Nevertheless, results obtained in our laboratory using Affymetrix GeneChips and RT-PCR demonstrated that some key genes that are regulated in vivo are also changed in the cell culture system. In this example, represented in Table 2, the induction of CYP2B, CYP3A1, and UDP2B in vitro were in good agreement with the effects on the same genes in the livers of the animals treated with the same test compound.

Transcriptional effects in vitro on a subset of appropriately selected marker genes will provide information regarding the potential liability in vivo. In addition, the in vitro hepatotoxicity DB provides evidence showing that compound classification according to gene expression profiles is possible, as long as the experimental conditions (in particular cell culture model) are standardized. The enormous advantage of in vitro systems is that they allow the number of test animals to be reduced, increase the assay throughput, and make possible experimentation on human tissues, as well as on animal tissues.

Mechanistic Explanations

In addition to the classification of compounds based on gene expression fingerprints obtained from tissue samples after exposure to toxicants, much has been learned about the underlying mechanisms of toxicity. Indeed, the identification of genes and/or pathways that are modulated by certain toxicants provides insight into possible mechanisms of toxicity. As typical examples, Clofibrate and WY 14643 are known to cause peroxisome proliferation and induce cytochrome P450 and other specific genes in rodents. These compounds were chosen as proof of concept for a toxicogenomics investigation [63], where the induction of some known PPAR α target genes by both compounds could be confirmed [64, 65].

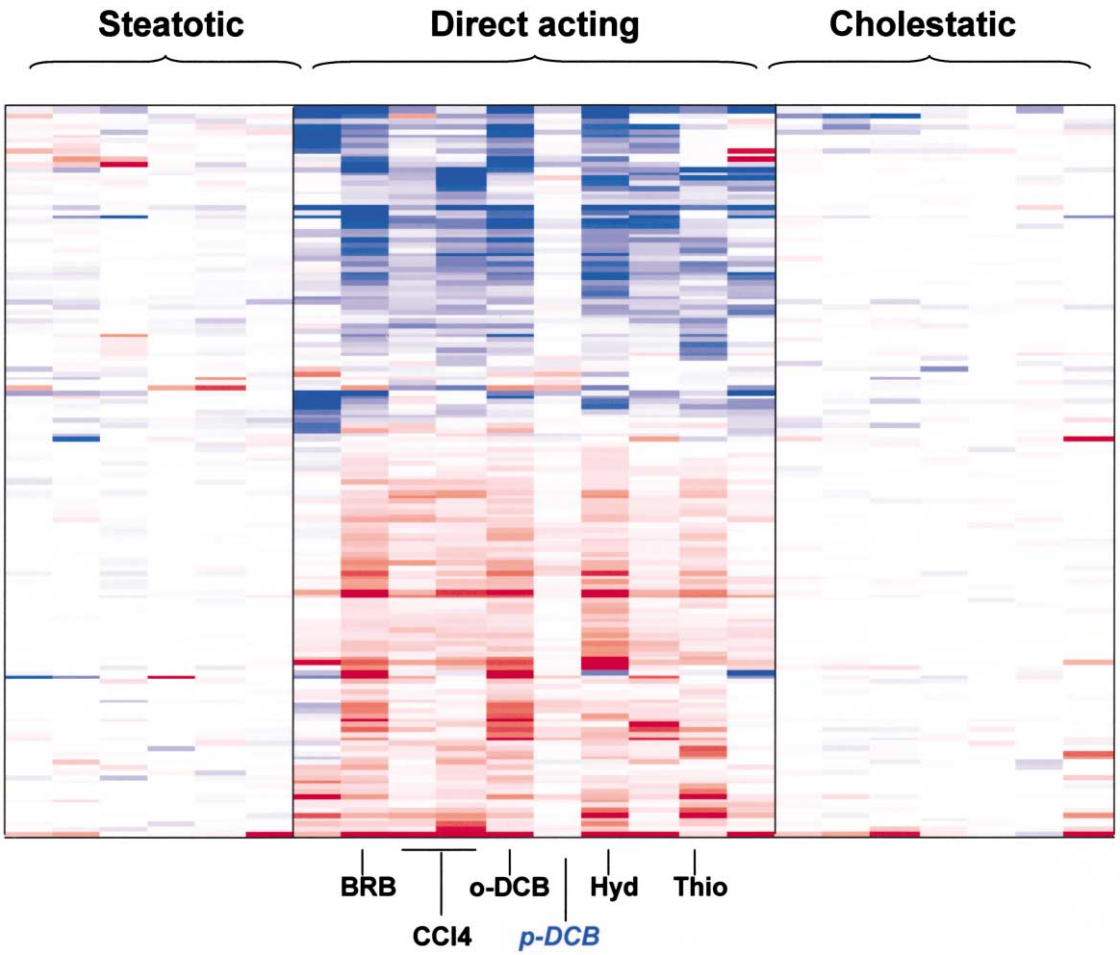

Figure 1. Regulation of Gene Expression by Different Types of Hepatotoxicants

Heat map representing gene regulation for a selection of genes caused by different types of compounds (steatotic, direct acting, and cholestatic). Blue boxes represent downregulated genes, and red boxes represent upregulated genes. Labels under the lanes indicate the five compounds used for the selection of genes (BRB, Bromobenzene; CCl4, carbon tetrachloride; o-DCB, 1,2-dichlorobenzene; Hyd, Hydrazine; and Thio, Thioacetamide) as well as the nontoxic analog p-DCB (1,4-dichlorobenzene).

Follow-Up Experiments for the Confirmation of Hypothesis

Confirmation of results obtained by microarray analysis using adequate follow-up assays with lower cost and/or higher throughput and increased accuracy is an important task in toxicogenomics. Nevertheless, one needs to distinguish between the technical validation of microarray data to confirm the alteration of a given transcript (using RT-PCR or Northern blots) and functional tests that confirm that increase of transcript is translated into an increase of protein content and/or activity. Typical examples from our laboratory include the confirmation of induced messenger levels are the induction of Gadd-45 and other stress-inducible genes

Table 2. Gene Expression of Example Genes In Vivo and In Vitro Assessed Using RT-PCR with Sybr Green as the Reporting Fluorophore

	In Vivo, Single Administration of 400 mg/kg Dose		In Vitro, 24 hr after Single Exposure	
Gene Name	6 hr	24 hr	20 μM	100 μM
Cytochrome P450 CYP2B2	31	72	260	2
Cytochrome P450 CYP2B, exon 9	91	95	240	10
Cytochrome P450 CYP3A1	1	2	8	2
UDP-glucuronosyltransferase 2b1	4	5	101	5

Values are expressed as fold changes with respect to the control. Note that in vitro the strongest induction occurs at a concentration of 20 μM, while the cytotoxic concentration of 100 μM has a weaker effect. Also noteworthy is the fact that the direct exposures of the hepatocyte monolayers elicit a stronger response than the one observed in vivo.

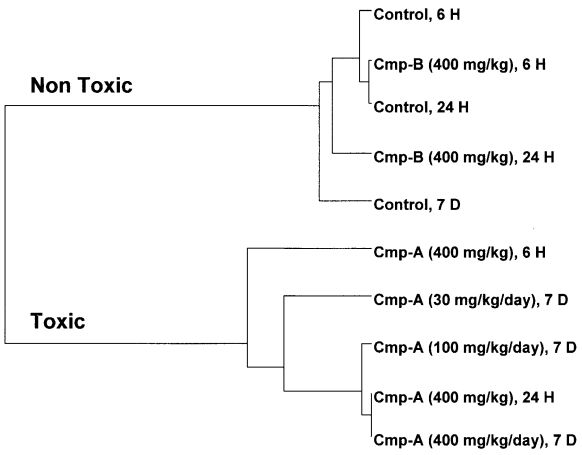

Figure 2. Cluster of Different Treatment Groups According to the Expression Levels of Selected Genes

The nonsteatotic Compound B (Ro-CmpB) clusters together with the vehicle-treated animals, while the steatotic Compound A (Ro-CmpA) forms a separate branch.

by compounds causing macromolecule damage and classified as direct acting compounds. As regards functional confirmation, the induction of CYP2B by 5-HT6 receptor antagonist was confirmed by Western blot analysis [53], while the induction of acyl-Coa oxidase by WY 14643 was confirmed by measuring its enzymatic activity in liver extracts (unpublished observations).

Use of Toxicogenomics for Pharmaceuticals: A Retrospective Analysis

Two compounds with similar pharmacological profiles but differing hepatotoxic potential were analyzed using a toxicogenomics approach [53]. In order to clarify whether toxicogenomics could recognize the hepatotoxic potential of Compound A and to gain insight into the possible molecular mechanisms underlying the hepatic findings (steatosis), gene expression profiles of the livers of rats treated with this compound and a nonhepatotoxic 5-HT6 receptor antagonist (Compound B) were obtained. Analysis using hierarchical clustering clearly showed that the two compounds could be distinguished based on their effect on gene expression patterns in the liver. As depicted in Figure 2, gene expression levels in the liver provide sufficient information to distinguish animals treated with a nonhepatotoxic compound or vehicle, from animals exposed to a steatotic compound. Moreover, side effects can be detected earlier with gene expression analysis than with conventional toxicology measurements. This finding is vital in order to further assess the usefulness of toxicogenomics approaches to improve the process of detecting the toxic potential of compounds quicker and more accurately than with conventional toxicity studies. Also, the obtained gene expression patterns provide possible marker genes amenable to being validated as biomarkers.

Future Perspectives

Advantages

Toxicogenomics represents an exciting new approach to toxicology and has a great potential to influence the predictability and speed of preclinical safety assessments. Published results so far show that genome-wide gene expression analysis is a powerful tool for compound classification and for the detection of new, specific, and sensitive markers for given mechanisms of toxicity.

In addition, preliminary results support the theory that gene expression might be more sensitive than conventional toxicology endpoints. Therefore, compound classification could be performed during early, short-term (i.e., single-dose) animal studies. Hence, time, cost, and number of animals needed to identify the toxic potential of a compound would be greatly minimized. It is also expected that in the very near future, regulatory authorities will evaluate toxicogenomics data as supplementary information supporting NDAs. Thus, the growing acceptance by the scientific community and the regulatory authorities, together with a variety of publications validating the technology, is providing the foundation for a successful implementation of toxicogenomic approaches in safety assessment.

The potential identification and validation of possible marker genes is also gaining momentum. Such markers could be employed in automated, high throughput assay systems that will provide indications regarding toxicity potential fast and accurately, without incurring the high costs commonly associated with microarray analysis. Appropriately chosen markers are amenable to being tested in cell-based assays that will allow scientists to evaluate compounds much earlier in the development process, improving clinical candidate selection.

The understanding of the molecular mechanisms underlying toxicity obtained through gene expression analysis after exposure of model systems (animals or cell cultures) to test compounds will also provide more insight into species-specific response to drugs regarding efficacy and toxicity. Hence, it is expected that extrapolation across species will become more accurate by enhancing the interpretation of preclinical observations and their meaning for the human situation. This should immensely increase the predictability of toxic liabilities and of potential risk accumulation for drug combinations or drug-disease interactions.

Challenges

In spite of the enormous opportunities offered by the available technologies, caution regarding the interpretation of the obtained results is still necessary. Gene expression data in the absence of excellent technical and analytical procedures or correct interpretation of the toxicological significance of the results can be misleading. Bearing this in mind, and that toxicogenomics is a young and fast developing field, there are two major issues that need to be overcome.

First, technical validation. This refers to the comparability of the data and standardization of the methods. It includes issues such as quality control, gene annotations, hybridization procedures, and statistical data analysis. Can results generated at different sites be compared, and is a reference DB generated at one site a valid reference for data generated at other sites? These issues have been recently addressed by ILSI, and discussions regarding the need for "standard procedures" are ongoing between the FDA and the pharmaceutical

industry. Lastly, it is paramount to stay abreast of the technological developments in this rapidly evolving field.

Second, biological and toxicological relevance of the findings. Currently, there is little understanding of the relationship between gene expression and dose-dependent induction of toxicity. The use of gene expression data in hazard identification in the absence of a correct interpretation of the toxicological significance of the results could lead to the generation of erroneous hypotheses. Suitable biological systems should be put in place to further support the mechanistic conclusions drawn from alterations in transcription. Basic questions of relevance need to be addressed. How many and which genes should be measured to characterize a toxic response and distinguish it from pharmacologic or physiologically adaptive responses that are not linked to toxicity? Do indicators of toxic responses always indicate that there is toxicity? How do we use gene expression studies in conjunction with commonly used in vitro and in vivo toxicity studies? And finally, how do we incorporate gene expression data and conventional tests for more accurate risk assessment?

Other "-omics"

Besides the efforts focusing on gene expression analysis, other "-omic" approaches, mainly proteomics and metabonomics, also need to be taken into consideration. The effects of compounds on protein patterns and on the excreted metabolite profiles provide additional information that is complimentary to the transcriptome analysis. This adds another dimension to the understanding of the molecular events underlying toxicity. Combined, gene and protein expression together with metabolite profile analysis are becoming a very powerful tool for mechanistic and predictive toxicity studies.

Acknowledgments

The authors would like to acknowledge the Toxicogenomics Team from Hoffmann-La Roche Ltd., Basle, Switzerland, for their excellent work. In addition, we are indebted to the Bioinformatics Group for their constant support and to the members of the Roche Center for Medical Genomics (RCMG) for many fruitful scientific discussions. In particular, we would like to mention our colleagues U. Certa, F. Boess, R. Gasser, and S. Albertini.

References

1. Watters, J.W., and McLeod, H.L. (2003). Cancer pharmacogenomics: current and future applications. Biochim. Biophys. Acta *1603*, 99–111.
2. Boggs, W.M. (2003). Toxicogenomics: Genome-Based Assessment of Drug Toxicity (Waltham, MA: Decision Resources Inc.).
3. Goodsaid, F.M. (2003). Genomic biomarkers of toxicity. Curr. Opin. Drug Discov. Dev. *6*, 41–49.
4. Lazarou, J., Pomeranz, B.H., and Corey, P.N. (1998). Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA *279*, 1200–1205.
5. Ulrich, R., and Friend, S.H. (2002). Toxicogenomics and drug discovery: will new technologies help us produce better drugs? Nat. Rev. Drug Discov. *1*, 84–88.
6. Butte, A. (2002). The use and analysis of microarray data. Nat. Rev. Drug Discov. *1*, 951–960.
7. Phillips, K.A., Veenstra, D.L., Oren, E., Lee, J.K., and Sadee, W. (2001). Potential role of pharmacogenomics in reducing adverse drug reactions: a systematic review. JAMA *286*, 2270–2279.
8. Acuna, G., Foernzler, D., Leong, D., Rabbia, M., Smit, R., Dorflinger, E., Gasser, R., Hoh, J., Ott, J., Borroni, E., et al. (2002). Pharmacogenetic analysis of adverse drug effect reveals genetic variant for susceptibility to liver toxicity. Pharmacogenomics J. *2*, 327–334.
9. Alvarado, I., Wong, M.L., and Licinio, J. (2002). Advances in the pharmacogenomics of adverse drug reactions. Pharmacogenomics J. *2*, 273.
10. Amin, R.P., Hamadeh, H.K., Bushel, P.R., Bennett, L., Afshari, C.A., and Paules, R.S. (2002). Genomic interrogation of mechanism(s) underlying cellular responses to toxicants. Toxicology *181–182*, 555–563.
11. Bartosiewicz, M.J., Jenkins, D., Penn, S., Emery, J., and Buckpitt, A. (2001). Unique gene expression patterns in liver and kidney associated with exposure to chemical toxicants. J. Pharmacol. Exp. Ther. *297*, 895–905.
12. Bulera, S.J., Eddy, S.M., Ferguson, E., Jatkoe, T.A., Reindel, J.F., Bleavins, M.R., and De La Iglesia, F.A. (2001). RNA expression in the early characterization of hepatotoxicants in Wistar rats by high-density DNA microarrays. Hepatology *33*, 1239–1258.
13. Hamadeh, H.K., Bushel, P.R., Jayadev, S., DiSorbo, O., Bennett, L., Li, L., Tennant, R., Stoll, R., Barrett, J.C., Paules, R.S., et al. (2002). Prediction of compound signature using high density gene expression profiling. Toxicol. Sci. *67*, 232–240.
14. Morgan, K.T. (2002). Gene expression analysis reveals chemical-specific profiles. Toxicol. Sci. *67*, 155–156.
15. Hamadeh, H.K., Bushel, P.R., Jayadev, S., Martin, K., DiSorbo, O., Sieber, S., Bennett, L., Tennant, R., Stoll, R., Barrett, J.C., et al. (2002). Gene expression analysis reveals chemical-specific profiles. Toxicol. Sci. *67*, 219–231.
16. Waring, J.F., Ciurlionis, R., Jolly, R.A., Heindel, M., and Ulrich, R.G. (2001). Microarray analysis of hepatotoxins in vitro reveals a correlation between gene expression profiles and mechanisms of toxicity. Toxicol. Lett. *120*, 359–368.
17. Huang, Q., Dunn, R.T., 2nd, Jayadev, S., DiSorbo, O., Pack, F.D., Farr, S.B., Stoll, R.E., and Blanchard, K.T. (2001). Assessment of cisplatin-induced nephrotoxicity by microarray technology. Toxicol. Sci. *63*, 196–207.
18. Luhe, A., Hildebrand, H., Bach, U., Dingermann, T., and Ahr, H.J. (2003). A new approach to studying ochratoxin A (OTA)-induced nephrotoxicity: expression profiling in vivo and in vitro employing cDNA microarrays. Toxicol. Sci. *73*, 315–328.
19. Aardema, M.J., and MacGregor, J.T. (2002). Toxicology and genetic toxicology in the new era of "toxicogenomics": impact of "-omics" technologies. Mutat. Res. *499*, 13–25.
20. Cheng, R.Y., Alvord, W.G., Powell, D., Kasprzak, K.S., and Anderson, L.M. (2002). Microarray analysis of altered gene expression in the TM4 Sertoli-like cell line exposed to chromium(III) chloride. Reprod. Toxicol. *16*, 223–236.
21. Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467–470.
22. Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., et al. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat. Biotechnol. *14*, 1675–1680.
23. Young, M.B., DiSilvestro, M.R., Sendera, T.J., Freund, J., Kriete, A., and Magnuson, S.R. (2003). Analysis of gene expression in carbon tetrachloride-treated rat livers using a novel bioarray technology. Pharmacogenomics J. *3*, 41–52.
24. Steiner, S., and Anderson, N.L. (2000). Expression profiling in toxicology—potentials and limitations. Toxicol. Lett. *112–113*, 467–471.
25. Tseng, G.C., Oh, M.K., Rohlin, L., Liao, J.C., and Wong, W.H. (2001). Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. Nucleic Acids Res. *29*, 2549–2557.
26. Salter, A.H., and Nilsson, K.C. (2003). Informatics and multivariate analysis of toxicogenomics data. Curr. Opin. Drug Discov. Dev. *6*, 117–122.
27. Ruepp, S.U., Tonge, R.P., Shaw, J., Wallis, N., and Pognan, F. (2002). Genomics and proteomics analysis of acetaminophen toxicity in mouse liver. Toxicol. Sci. *65*, 135–150.
28. Iida, M., Anna, C.H., Hartis, J., Bruno, M., Wetmore, B., Dubin,

J.R., Sieber, S., Bennett, L., Cunningham, M.L., Paules, R.S., et al. (2003). Changes in global gene and protein expression during early mouse liver carcinogenesis induced by non-genotoxic model carcinogens oxazepam and Wyeth-14,643. Carcinogenesis *24*, 757–770.

29. Fountoulakis, M., de Vera, M., Crameri, F., Boess, F., Gasser, R., Albertini, S., and Suter, L. (2002). Modulation of gene and protein expression by carbon tetrachloride in the rat liver. Toxicol. Appl. Pharmacol. *183*, 71–80.
30. Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A., Causton, H.C., et al. (2001). Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat. Genet. *29*, 365–371.
31. Cunningham, M.L., Bogdanffy, M.S., Zacharewski, T.R., and Hines, R.N. (2003). Workshop overview: use of genomic data in risk assessment. Toxicol. Sci. *73*, 209–215.
32. Raychaudhuri, S., Stuart, J.M., and Altman, R.B. (2000). Principal components analysis to summarize microarray experiments: application to sporulation time series. Pac. Symp. Biocomput. 455–466.
33. Ambroise, C., and McLachlan, G.J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. Proc. Natl. Acad. Sci. USA *99*, 6562–6566.
34. Liu, A., Zhang, Y., Gehan, E., and Clarke, R. (2002). Block principal component analysis with application to gene microarray data classification. Stat. Med. *21*, 3465–3474.
35. 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. USA *95*, 14863–14868.
36. Schoelkopf, B. (1999). Advances in Kernel Methods—Support Vector Learning (Cambridge, MA: MIT Press).
37. Lesko, L.J., Salerno, R.A., Spear, B.B., Anderson, D.C., Anderson, T., Brazell, C., Collins, J., Dorner, A., Essayan, D., Gomez-Manzilla, B., et al. (2003). Pharmacogenetics and pharmacogenomics in drug development and regulatory decision making: report of the first FDA-PWG-PhRMA-DruSafe workshop. J. Clin. Pharmacol. *43*, 342–358.
38. Robinson, D., Pettit, S., and Morgan, D.G. (2003). Use of genomics in mechanism based risk assessment. In Toxicogenomics, T. Inoue, W.D. Pennie, eds. (Tokyo: Springer-Verlag), pp. 194–203.
39. Hamadeh, H.K., Knight, B.L., Haugen, A.C., Sieber, S., Amin, R.P., Bushel, P.R., Stoll, R., Blanchard, K., Jayadev, S., Tennant, R.W., et al. (2002). Methapyrilene toxicity: anchorage of pathologic observations to gene expression alterations. Toxicol. Pathol. *30*, 470–482.
40. Chu, T., Deng, S., Wolfinger, R., Paules, R., and Hamadeh, H. (2004). Cross-site comparison of gene expression data reveals high similarity. Environ. Health Perspect. Toxicogenomic *112*, in press.
41. Newton, R., Aardema, M., and Aubrecht, J. (2004). The utility of DNA microarrays for characterizing genotoxicity. Environ. Health Perspect. Toxicogenomic *112*, in press.
42. Mattes, W. (2004). Annotation and cross-indexing of array elements on multiple platforms. Environ. Health Perspect. Toxicogenomic *112*, in press.
43. Mattes, W., Pettit, S., Sansone, S., Bushel, P., and Waters, M. (2004). Database development in toxicogenomics: issues and efforts. Environ. Health Perspect. Toxicogenomic *112*, in press.
44. Goodsaid, F., Smith, R., and Rosenblum, I. (2004). Quantitative PCR deconstruction of discrepancies between results reported by different hybridization platforms. Environ. Health Perspect. Toxicogenomic *112*, in press.
45. Rosenzweig, B., Pine, P., Domon, O., Morris, S., Chen, J., and Sistare, F. (2004). Dye bias correction in dual-labeled cDNA microarray gene expression measurements. Environ. Health Perspect. Toxicogenomic *112*, in press.
46. Amin, R., Vickers, A., Sistare, F., Thompson, K., Roman, R., Lawton, M., et al. (2004). Identification of putative gene based markers of renal toxicity. Environ. Health Perspect. Toxicogenomic *112*, in press.
47. Baker, V., Harries, H., Waring, J., Duggan, C., Ni, H., Jolly, R., Yoon, L., Souza, A.D., Schmid, J., Brown, R., et al. (2004). Clofibrate-induced gene expression changes in rat liver: a cross-laboratory analysis using membrane cDNA arrays. Environ. Health Perspect. Toxicogenomic *112*, in press.
48. Thompson, K., Afshari, C., Amin, R., Bertram, T., Car, B., Cunningham, M., Kind, C., Kramer, J., Lawton, M., Mirsky, M., et al. (2004). Identification of platform-independent gene expression markers of cisplatin nephrotoxicity. Environ. Health Perspect. Toxicogenomic *112*, in press.
49. Ulrich, R., Rockett, J., Gibson, G., and Pettit, S. (2004). Overview of an inter-laboratory collaboration on evaluating the effects model hepatotoxicants on hepatic gene expression. Environ. Health Perspect. Toxicogenomic *112*, in press.
50. Pennie, W., Pettit, S., and Lord, P. (2004). Toxicogenomics in risk assessment: an overview of an ILSI HESI collaborative research program. Environ. Health Perspect. Toxicogenomic *112*, in press.
51. Kramer, J., Pettit, S., Amin, R., Bertram, T., Car, B., Cunningham, M., Curtiss, S., Davis, J., Kind, C., Lawton, M., et al. (2004). Overview on the application of transcription profiling using selected nephrotoxicants for toxicology assessment. Environ. Health Perspect. Toxicogenomic *112*, in press.
52. Waring, J., Ulrich, R., Flint, N., Morfitt, D., Kalkuhl, A., Staedtler, F., Lawton, M., Beekman, J., and Suter, L. (2004). Interlaboratory evaluation of rat hepatic gene expression changes induced by methapyrilene. Environ. Health Perspect. Toxicogenomic *112*, in press.
53. Suter, L., Haiker, M., de Vera, M.C., and Albertini, S. (2003). Effect of two 5-HT6 receptor antagonists on the rat liver: a molecular approach. Pharmacogenomics J. *3*, 320–334.
54. 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.
55. 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., et al. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science *286*, 531–537.
56. 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.
57. Rosenwald, A., Wright, G., Chan, W.C., Connors, J.M., Campo, E., Fisher, R.I., Gascoyne, R.D., Muller-Hermelink, H.K., Smeland, E.B., Giltnane, J.M., et al. (2002). The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N. Engl. J. Med. *346*, 1937–1947.
58. Ross, D.T., Scherf, U., Eisen, M.B., Perou, C.M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S.S., Van de Rijn, M., Waltham, M., et al. (2000). Systematic variation in gene expression patterns in human cancer cell lines. Nat. Genet. *24*, 227–235.
59. Scherf, U., Ross, D.T., Waltham, M., Smith, L.H., Lee, J.K., Tanabe, L., Kohn, K.W., Reinhold, W.C., Myers, T.G., Andrews, D.T., et al. (2000). A gene expression database for the molecular pharmacology of cancer. Nat. Genet. *24*, 236–244.
60. Weinstein, J.N., Myers, T.G., O'Connor, P.M., Friend, S.H., Fornace, A.J., Jr., Kohn, K.W., Fojo, T., Bates, S.E., Rubinstein, L.V., Anderson, N.L., et al. (1997). An information-intensive approach to the molecular pharmacology of cancer. Science *275*, 343–349.
61. Giordano, T.J., Thomas, D.G., Kuick, R., Lizyness, M., Misek, D.E., Smith, A.L., Sanders, D., Aljundi, R.T., Gauger, P.G., Thompson, N.W., et al. (2003). Distinct transcriptional profiles of adrenocortical tumors uncovered by DNA microarray analysis. Am. J. Pathol. *162*, 521–531.
62. Boess, F., Kamber, M., Romer, S., Gasser, R., Muller, D., Albertini, S., and Suter, L. (2003). Gene expression in two hepatic cell lines, cultured primary hepatocytes, and liver slices compared to the in vivo liver gene expression in rats: possible implications for toxicogenomics use of in vitro systems. Toxicol. Sci. *73*, 386–402.

63. Esbenshade, T.A., Kamanna, V.S., Newman, H.A., Tortorella, V., Witiak, D.T., and Feller, D.R. (1990). In vivo and in vitro peroxisome proliferation properties of selected clofibrate analogues in the rat. Structure-activity relationships. Biochem. Pharmacol. *40*, 1263–1274.
64. Frederiksen, K.S., Wulf, E.M., Wassermann, K., Sauerberg, P., and Fleckner, J. (2003). Identification of hepatic transcriptional changes in insulin-resistant rats treated with peroxisome proliferator activated receptor-alpha agonists. J. Mol. Endocrinol. *30*, 317–329.
65. Kramer, J.A., Blomme, E.A., Bunch, R.T., Davila, J.C., Jackson, C.J., Jones, P.F., Kolaja, K.L., and Curtiss, S.W. (2003). Transcription profiling distinguishes dose-dependent effects in the livers of rats treated with clofibrate. Toxicol. Pathol. *31*, 417–431.