African Journal of Pharmacy and Pharmacology Vol. 4(11). pp. 763-774, November 2010
Available online http://www.academicjournals.org/ajpp
ISSN 1996-0816 ©2010 Academic Journals

Full Length Research Paper

Toxicogenomics revolution in the optimisation of pharmaceutical drug development and drug safety evaluations

Fokunang C. N.[1]*, Tembe-Fokunang E. A.[1], Ngameni B.[1], Barkwan S. S.[2], Tomkins P. T.[2], Asongalem E. A.[1], Ngadjui B. T.[1], Ngogang J. Y.[1], Abena O. M. T.[1] and Asonganyi T.[1]

[1]Department of Pharmacy and African Pharmacotherapeutics, Faculty of Medicine and Biomedical Sciences, University of Yaoundé 1, P.O. Box 33032, Yaoundé, Centre Region, Cameroon.
[2]Athlone Institute of Technology, Centre for Biopolymer and Bio-molecular Research, County Westmeath, Republic of Ireland.

Accepted 09 November, 2010

Modern high throughput screening methods in toxicology need to be developed. These new approaches are necessary to provide more insight into potential human toxicity than the old traditional methods. Toxicogenomics investigate the changes in gene expression profile following exposure to a toxicant. It offers the potential to identify a human toxicant earlier in drug development and to detect human-specific toxicants that can cause no adverse reaction in rats. New transcript profiling technologies enable simultaneous measurement of the changes in expression of many hundreds or many thousands of genes. The availability of these methods has brought about significant revolutionary changes in many areas of investigative biology, where analyses of patterns of gene expression, rather than of individual genes, are being used. Toxicogenomics offers new opportunities for both mechanistic toxicity research and predictive toxicology. This paper gives an overview of the impact of toxicogenomics on the revolution of pharmaceutical drug development and drug safety evaluation. The basic approach within the frame work of high throughput screening is also a debate for discussion. The challenging issues with this novel technology in the selection screening of biomarkers for different therapeutic end points are also examined in this review.

Key words: Toxicogenomics, genomics, transcript profiling, biomarkers, microarray, drug development, drug safety.

INTRODUCTION

Toxicology, the study of adverse effects of chemicals on living organisms, has traditionally been evaluated by the dosing of animals to define well-established cytological, physiological, metabolical, and morphological endpoints aspects. The evaluation of the risk to humans cannot be performed in human individuals initially and thus must be derived from studies performed in other species (Ulrich and Friend, 2002; Collings and Vaidya, 2008). Typically, rodents are used to identify toxic substances such as carcinogens, reproductive toxins, and neurotoxins. Follow-up studies in non-rodent species (other animal models) can then be used to further define the effects of low doses as well as species extrapolation and mechanism of action (Smith, 2001; Reilly et al., 2005).

Although it is well recognized that intact animals are needed to reflect physiological 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 (Scherf et al., 2000; Steven et al., 2006). Thus, the current approach

*Corresponding author. charlesfokunang@yahoo.co.uk. Tel: +237 9421 8670.

to toxicological al testing is costly, in terms of time, labour, 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 (Holmes and Nicholson, 2001; Gibson et al., 2008).

TOXICOGENOMICS

The biggest improvement needed in the drug development process is in the field of toxicology, which is the point where most developmental bottlenecks occur (Fielden and Zacharewski, 2001; Wills, 2007). One promising area of advancement is the new field of toxicogenomics (Foster et al., 2007; Collings and Vaidya, 2008). Detection of changes at the molecular level provides insight into a toxicant's mechanism-of-action and its potential to cause human toxicity (Harries et al., 2001; Petricoin et al., 2002; Burczynski, 2009). Toxicogenomics has grown quickly with the number of articles published approximately tripling in 2009, (estimated) over that of 2005. However, the field is still in its infancy, as shown by the majority of review articles describing the promise of this technology. As toxicogenomics data grows, a developing challenge is the analysis of large datasets and the building of predictive toxicogenomic databases (Van Dijck et al., 2003; Gibson et al., 2008). Here, this challenge is addressed together with how toxicogenomics can predict toxicity faster than classical measurements, such as serum chemistry perturbations (Altman and Raychaudhuri, 2001; Shah, 2006).

In vitro hepatocyte models represent very useful systems in both fundamental research and various application areas. Primary hepatocytes appear as the closest model for the liver *in vivo*. However, they are phenotypically unstable, have a limited life span and in addition, exhibit large interdonor variability when of human origin. Hepatoma cell lines appear as an alternative but only the HepaRG cell line exhibits various functions, including major cytochrome P450 activities, at levels close to those found in primary hepatocytes (Aardema and MacGregor, 2002). *In vitro* hepatocyte models have brought a substantial contribution to the understanding of the cell biology of the normal and diseased liver and in various application domains such as xenobiotic metabolism and toxicity and more generally cell therapies (Cunningham, 2000). In the future, new well-differentiated hepatocyte cell lines derived from tumors or from either embryonic or adult stem cells might be expected and although hepatocytes will continue to be used in various fields, these *in vitro* liver models should allow marked advances, especially in cell-based therapies and predictive and mechanistic hepatotoxicity of new drugs and other chemicals. All models will benefit from new developments in throughput screening based on cell chips coupled with high-content imaging and in toxicogenomics technologies (Guguen-Guillouzo and Guillouzo, 2010).

Gene Logic (Gaithersburg, MD, USA) has built a reference database called the ToxExpress™ Module (Wen and Fitch, 2009). Examples of its use to predict toxicity of drugs and chemicals before standard signs of toxicity in rats have been highly studied (Sherlock, 2000; Foster et al., 2007). Gene expression profiling lends itself readily to two highly topical areas of drug development: "predictive" toxicology and mechanism-based risk assessment (Lennon, 2000; Ayensu et al., 2004; Blomme et al., 2009). The acceleration of efforts in predictive toxicology in particular has been largely due to the technological and scientific advances made in the last decade in genomics research. Hence, the advent of arrayed gene platforms for gene expression analysis has led to much investment by drug companies, government agencies and technology providers in applying genomics-based approaches in drug development (Harill and Rusyn, 2008).

Turning the promise of genomics into practice in drug development is not without its obstacles (Cunningham et al., 2003). Cellular response to xenobiotics and consequent pathogenesis represent a dynamic process, gene transcriptional responses being just one component part. A key challenge is to measure gene expression at time points (and dose levels) at which changes are meaningful for the response to a drug, for the adaptation to the response, and for the down-stream consequence of adverse drug reaction (Smith, 2000). The high expectations for toxicogenomics coupled with the complexity of the task of putting genomics into practice in drug development has understandably raised a number of concerns within the drug industry and within drug regulatory agencies (Ulrich and Friend, 2002). For the genomic data to gain basic acceptance there needs to be a confidence in the technology (Cunningham, 2000). As the field advances there is a need to built confidence in the meaning of modulation of gene expression as it applies to assessment of toxicity (Crosby et al., 2000; Ganter et al., 2008). Techniques of molecular biology have improved diagnostic sensitivity, accuracy and validity in forensic medicine very much, especially in the field of identification (paternity testing, stain analysis). Since more than 10 years these techniques - meanwhile well established in clinical disciplines - are used also for other applications in forensic medicine: determination of cause and manner of death, tissue identification by mRNA and microRNA, examination of gene expression levels (survival time, time since death, cause of death), toxicogenetics (Madea et al., 2010) Development of genomics and bioinformatics enable us to analyze the global gene expression profiles of cells by DNA microarray. Changes in gene expression patterns indicate changes in its physiological conditions. Following the exposure of an organism or cell to toxic chemicals or

other environmental stresses, the global genetic responses can be expeditiously and easily analyzed. Baker's yeast, *Saccharomyces cerevisiae*, is one of the most studied and useful model eukaryotes. The biggest advantage of yeast genomics is the available functional information for each gene and a considerable number of data are accumulating in the field of toxicity assessment using yeast DNA microarray (Yasokawa et al., 2010).

BIOMARKER GENES

Selection of marker genes often entails ranking with respect to differences between experimental groups. Therefore, the definition of difference becomes important. Ideally, this should be based on statistical inference where genes are ranked by the probability that their expression values are different in a reproducible sense and not the result of random noise (Alvarado et al., 2002; Szyf, 2007). However, the ability to infer significance depends on the sample size and population assumptions. Not all genes are expected to have the same dynamic range or variability; therefore, they do not belong to the same population (Amin et al., 2002; Guillouzo and Guguen-Guillouzo, 2008). A large database enables direct determination of biological variability and dynamic range for every gene. For this reason, the company generally collects data from at least thousand biological samples per organ to directly determine dynamic range and variability (Morgan, 2000). Information on each gene can thus stand alone without the assumption that all genes behave in a similar way and therefore belong to the same population. Both parametric and non-parametric methods are available to help rank genes based on inference or discrimination ability (Hamadeh et al., 2002).

The rapid decline in the cost of dense genotyping is paving the way for new DNA sequence-based laboratory tests to move quickly into clinical practice, and to ultimately help realize the promise of 'personalized' therapies (Rusyn et al., 2010). These advances are based on the growing appreciation of genetics as an important dimension in science and the practice of investigative pharmacology and toxicology. On the clinical side, both the regulators and the pharmaceutical industry hope that the early identification of individuals prone to adverse drug effects will keep advantageous medicines on the market for the benefit of the vast majority of prospective patients (Smith, 2001; Wen and Fitch, 2009). On the environmental health protection side, there is a clear need for better science to define the range and causes of susceptibility to adverse effects of chemicals in the population, so that the appropriate regulatory limits are established. In both cases, most of the research effort is focused on genome-wide association studies in humans where de novo genotyping of each subject is required. At the same time, the power of population-based preclinical safety testing in rodent models (e.g., mouse) remains to be fully exploited (Ryan et al., 2008).

Some attempts have been made to improve gene expression variability estimates. Variability between different genes has been measured and reapplied to each individual gene (Bulera et al., 2001) A probability for a fold change or relative difference is then calculated based on this pooled variability estimate. Genes that are most variable between selected conditions or successive rounds of clustering have also been used to select target genes. However, all these methods assume that all genes belong to a population with similar biological variability and dynamic range (Bartosiewicz et al., 2001; Ryan et al., 2008). With a thousand or more observations for every gene in our database, direct examination of gene expression distributions are possible, as are the selection of targets for determining both the presence of toxicity and its probable mechanism (Waring et al., 2001; Lockhart et al., 2007).

SAMPLE CLASSIFICATIONS

Once target genes have been selected, information from multiple genes is used to classify a drug as toxic and to determine its potential toxicity types (Aardema and MacGregor, 2002). Many methods have been used to classify samples and genes in microarray experiments. Sample classification starts with known samples called a training set. Combinations of known sample genes that change reproducibly are identified, and then used to determine the treatment of unknown samples (Cheng et al., 2002; Wills, 2007). Classification success depends on the reliability that changes in known samples accurately reflect what will be seen in the population being tested. An additional challenge is proving that these combinations of gene expression changes are relatively unique to the classification being attempted. Classification as normal or pathological must consider all possible normal states and abnormal pathologies (Young et al., 2003; Ryan et al., 2008). Therefore, a large number of toxin-treated and control samples must be used to approximate the uniqueness of the markers to measuring the toxic events (Steiner et al., 2000; Gibson et al., 2008).

GENES CLASSIFICATION

Genes can be classified according to how close their expression profiles are related across many samples. Clustering is a method that is widely used to analyze gene expression. However, a general problem with clustering is that the same measure of similarity is used with every gene (Tseng et al., 2001; Cunningham et al., 2003). The differences in dynamic range and variability

between genes are often not known and assumed to be similar so no accurate level of confidence can be assigned to distinguish an expression pattern from biological noise (Fountoulakis et al., 2002). Most clustering methods also assume that a gene can belong to only one cluster; a condition that is inconsistent with the intricate interconnections of pathways known to occur in biology (Ambroise and McLachlan, 2002). In addition, hierarchical clustering assumes that a parent–child type relationship exists in the regulation of all genes. Although some of these assumptions might not be valid for gene expression, clustering can identify new relationships between genes that can then be validated by experimental manipulation (Ruepp et al., 2002).

DATABASES AND DATA ANALYSIS

Predictive toxicogenomic studies attempts usually to compare the gene expression patterns elicited by chemicals with unknown toxic potential to the profiles of model compounds with known toxicity (Salter and Nelson, 2003; Shah, 2006; Wen and Pitch, 2009). 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) (Steiner and Anderson, 2000; Ruepp et al., 2002). 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 (Liu et al., 2002; Lockhart et al., 2007).

There are many toxicogenomics DBs either fully constructed or 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 serving as an important target organ for toxicity. 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 (Waring et al., 2001; Morgan, 2002). On the other hand, a typical microarray experiment generates thousands of data points. Additional confounding factors are time and dose responses. Groups of genes might be co-regulated by a stimulus while redundancy in microarray design leads to multiple probes examining the same mRNA sequence. However, 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 (Alvarado et al., 2002; Collins and Vaidya, 2008). 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.

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 (Cunningham, 2000; Petricoin et al., 2002; Shah, 2006). 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 (Waring et al., 2001; Wills, 2007). 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 (Liu et al., 2002; Salter et al., 2003; Blomme et al., 2009). Rigorous attention to data quality and transformation steps is crucial and will have a major impact on any analysis. In order to use the knowledge stored in the toxicogenomics DBs, several supervised multivariate methods of sample classification has been described in Table 1. This include the discriminate analysis, neural networks, nearest neighbours and support vector machines which are based on algorithms that learn from the training data set in the DB and use previously acquired knowledge to classify unknown compounds (Smith, 2001; Young et al, 2003).

GENOMIC PLATFORM

Currently, technological innovation resulting in the sequencing of the human genome (Bullera et al., 2001) has revolutionalized toxicogenomic research. Progress in genomic research has allowed for the study of gene expression of thousands of gene simultaneously (Aardema and MacGregor, 2002; Wen and Fitch, 2009). Toxicology research can now be performed with DNA microarrays to redefine the pharmaceutical testing platform (Bullera et al., 2001; Ambroise and MacLachlan, 2000). The process of Microarray is well outlined in Figure 1, and the data that is generated for statistical interpretation shown in Figure 2. The use of toxicogenomics in parallel with other technological advancements, such as metabonomics (Ruepp et al., 2002) and proteomics (Wen and Fitch, 2009), has enhanced investigative research by providing an understanding of the complex mechanistic pathways of toxicity and identifying molecular biomarkers capable of predicting toxicity early in drug development testing process as well as providing biomarkers of toxicity,

Table 1. Multivariate methods of sample classification.

Methods	How it works	Reference
Logistic regression	Weight assigned to each gene based on distribution of known samples in the training sets. Requires enough samples to adequately estimate distribution. Estimate a linear log likelihood function for distribution values. Robust against departure from normal distribution	Fielding and Zacharewski, 2001
Linear discriminant analysis	Weight assigned to each gene based on distribution of in the training sets. And requires enough samples to adequately estimate distribution. Uses probability functions based on normal distribution and requires the assumption of normality	Shena et al., 1995
Neutral networks	Use functional units called nodes that calculate inputs, process relationships between input nodes and calculate output. Weights assigned are learned from a training set. Requires many samples to accurately estimate weights. No assumption on underlying distributions is needed. Can learn from new information	Liu et al., 2002
Clustering approaches	Uses guilt by association method. Samples are classified as to which they most closely resemble. Can be useful method when the number of samples is small. Assumes the same measure of closeness can be applied equally to all genes. The measure of closeness (metric) can greatly influence results.	Salter and Nelson, 2003
Support vector machine	Draws an optimal complex boundary through multi-dimensional space that separates sample groups in the training set. Depends on each sample being good representatives of groups Highly sensitive to measurement errors or outliers.	Waring et al., 2000

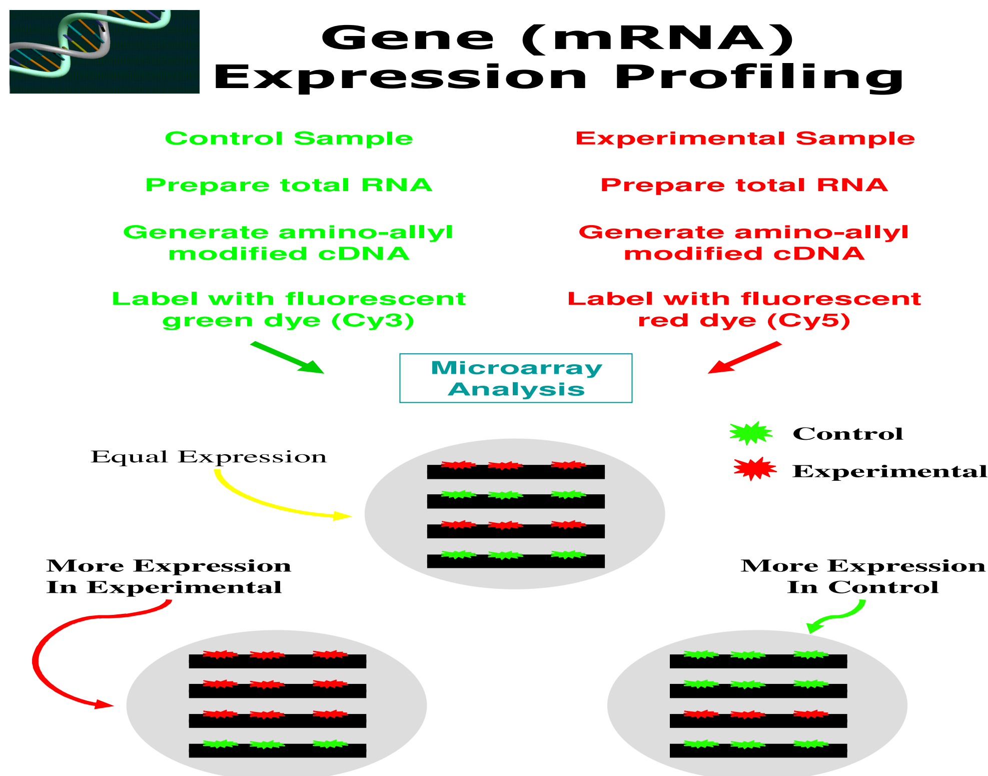

Figure 1. Microarray technology in gene expression transcript profiling (Bulera et al., 2001).

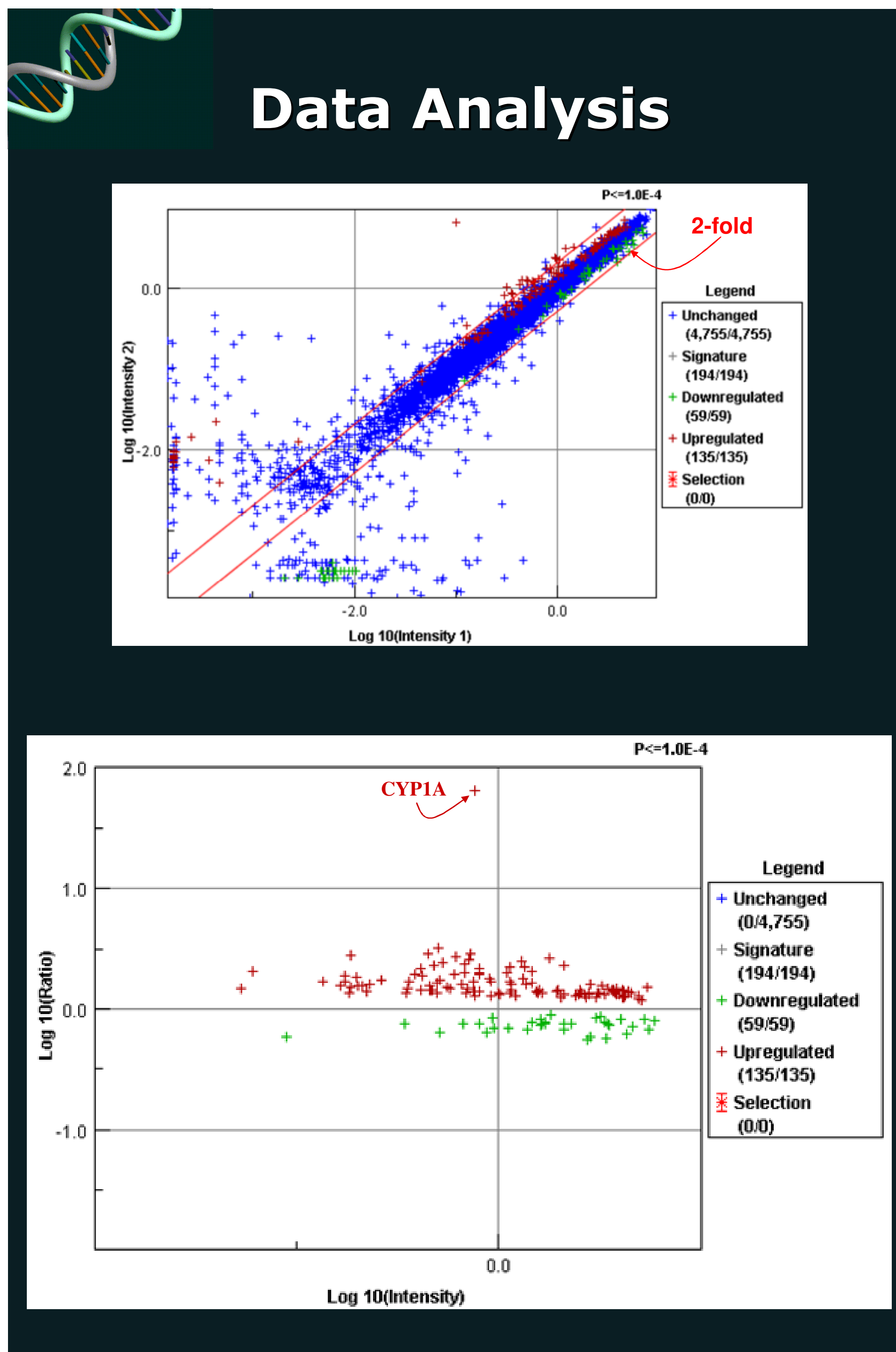

Figure 2. Statistical data analysis of Microarray results (Kerr and Churchill, 2001).

efficacy and exposure in preclinical and clinical trials (Will, 2007).

Gene expression can be monitored one gene at a time or by the thousands. Quantitative real-time PCR (Q-RT-PCR) which enables the quantitation of one gene at a time and is robust; useful traits when the genes of interest are known (Ayensu et al., 2004). However, to identify genes of interest, a broad survey approach is more useful. Microarrays (closed systems) and open platforms, such as READS™ (Gene Logic) (Van Dijck et al., 2003; Blomme et al., 2009) differential display technology, enable the simultaneous evaluation of

thousands of genes (Harill and Rusyn, 2008). Microarray platforms offer greater speed and more quantitative information than the differential display technology but this platform provides advantages such as: (1) detecting novel as well as known genes and (2) enabling the profiling of any species, even plants (Fielden and Zacharewski 2001). Gene expression profile in the microarray test battery is shown in Figure 1. Microarray platforms use either cDNA clones, or oligonucleotides, which can discriminate small sequence variations among highly homologous genes (e.g. the cytochrome P450 family), DNA mutations and polymorphic variations between individuals (Van Dijck et al., 2003; Foster et al., 2007).

Numerous commercial microarray platforms offer genome-wide coverage for model systems such as rat, mouse, *Caenorhabditis elegans*, and humans. Commercial microarrays are also available for genes that are highly expressed in specific tissues (e.g., liver, breast) and during specific biological processes such as metabolism (e.g., P450 enzymes) (Foster et al., 2007). Both genome wide and dedicated arrays can be used with RNA samples from *in vivo* and *in vitro* (tissue and cell culture) systems, enabling parallel studies to be conducted with a single microarray platform. This is important because the results of microarray experiments can vary depending on the array design and the selection and performance of gene probes on the array (Kerr and Churchill, 2001).

Reference RNAs can be derived from tissue extracts, cell lines, or both and serve a variety of purposes. Workshops sponsored by governments and industry have focused on defining the specifications for reference RNAs for clinical and regulatory applications (Joseph, 2004). The consensus is that multiple RNA standards are needed to measure the accuracy, dynamic range, sensitivity, and specificity of varied technology platforms under varied conditions. Important questions are whether regulatory agencies will define preferred sources of RNA standards, and, if so, who will generate and maintain baseline information about these standards. Although the selection of a given RNA standard depends primarily on the purpose and application, all RNA standards should be tested for a clearly defined number of copies of a given sequence within an RNA preparation over some linear range (Cronin et al., 2004). Some initiatives are raising awareness of the effects of variables that might hamper data comparability and are working toward developing best practice guidelines for microarray-based measurements (Hopkins et al., 2004). For example, recommenddations for best practice in array normalization, together with performance characteristics in terms of sensitivity, accuracy, and comparability of different array platforms (cDNA and oligo, spotted and in situ synthesis), are beginning to emerge together with proposals for transparency and availability through publicly accessible databases (Morgan, 2000). Other initiatives are considering the use of quality metrics for standardizing and validating array-based toxicogenomics measurements. The extent to which such efforts will be pursued and the impact they will have upon the standardization issues that are a necessary prerequisite to the validation exercises remain to be seen.

PREDICTIVE TOXICOLOGY

The last few years have seen a lot of progress being made in linking the profiles of gene expression induced by drugs with their toxicities. By developing databases of expression profiles for a wide variety of toxic compounds and toxic models it has been possible to create statistical and computational methods which provide an indication of the toxic potential of a drug from the pattern of gene expression changes it elicits *in vitro* or *in vivo* systems (Shena et al., 1995; Cunningham, 2003). Using an array representing around 1000 rat genes, it was possible to show patterns of rat liver gene expression that distinguished between 15 different hepatotoxins, directly from the data (Scherf et al., 2000; Reilly et al., 2005). They tried several statistical/informatics tools to analyse the data to reveal how the gene expression profiles fell into clusters. Each method highlighted similarities between compounds which were expected as well as a few that were not expected (Olson et al., 2000; Shah, 2006). The specific genes whose expression was modulated by the treatments were found to be indicative of several known toxic mechanisms. This initial study of a small set of compounds therefore showed utility for comparing and discriminating between compounds as well as for investigating underlying mechanism (Schena, 1996; Gibson, 2008).

MECHANISTIC TOXICOLOGY

In recent times, an understanding of the mechanism of toxicity of a new drug has become a major part of its risk assessment. Because gene expression is central to many responses to xenobiotics, genomic approaches are highly applicable to mechanistic toxicology studies (Waring et al., 2000; Colling and Vaidya, 2008). A transcriptional response can give a preliminary indication of the biochemical or biological mechanism being affected by a xenobiotic. By examining changes in gene expression in cells and tissues in response to drugs it is possible to generate hypotheses as to the underlying mechanism (Cunningham, 2003). Used in this way, gene expression data should be viewed as starting points rather than as end-points in a toxicological examination (Salter and Nilson, 2003; Gibson et al., 2008). If the mechanism is unknown then the genomic data can help to identify more definitive end-points which may be proteomic or enzymatic in nature (Thomas et al., 2000; Holmes and

Table 2. Experimental design parameters (Morgan 2002).

Primary research objective	Academia	Industry
	Explain underlying mechanisms of toxicity	**Determine expression profiles for specific compounds or families**
Most frequently used model system	Mouse	Mouse
Most frequently used sample type	Liver tissue	Liver tissue
Doses per compound tested	3	3
Replicates per compound tested	3	3
Time points per compound tested	Either 3 or more than 5	2

Nicholson, 2001; Stevens et al., 2006).

TOXICOGENOMICS OF HUMAN-SPECIFIC TOXICANT

Toxicogenomic studies are conducted based on different experimental designed parameters as described in Table 2. The design within the academia is different from an industrial approach. Although the replicate per compound, model system and doses of compound are uniform across the board, the primary research objective and time points per compound tested vary from the academia and the industry (Farr and Dunn, 1999; Reilly et al., 2005; Blomme et al., 2009). It is known that preclinical studies in animals do not always detect drugs that prove to be toxic in humans (DeRisi et al., 1996; Harill and Rusyn 2008). To determine if toxicogenomics can detect some of these harmful drugs before they enter the clinical studies were conducted on, tacrine (Cognex®; Sigma, St Louis, MO, USA). Controls included donepezil (Aricept®; Pfizer, New York, NY, USA) and physostigmine (Antilirium®; Sigma), which have a mechanism-of-action similar to tacrine but do not induce hepatotoxicity. Tacrine and donepezil are used clinically to treat patients with Alzheimer's disease. Physostigmine was used in a clinical trial but has not been approved by the FDA for this indication (Scherf et al., 2000; Fountoulakis et al., 2002; Shah, 2006). These drugs are reversible inhibitors of acetylcholinesterase although selectivity for cholinesterases varies between the drugs. Using models built with gene expression data, toxicity can be predicted at early phases before classical toxicological changes occur (e.g. alterations in serum parameters or pathological evidence of toxicity). (Amin et al., 2002; Liu et al., 2002; Lockhart et al., 2007).

SUCCESSES

Currently, 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 (Steven et al., 2006, Harill and Rusyn, 2008). 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 (Van Dijck et al., 2003; Ryan et al., 2008). 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 (Fielden and Zacharewski, 2001; Ayensu et al., 2004).

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 (Farr and Dunn, 1999; Liu et al., 2002; Young et al., 2003). An early report described the development and progression of malignant melanoma using microarrays (Schena, 1996; Burczynski, 2009). 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 (Sherlock, 2000; Salter, 2003; Harill and Rusyn, 2008). Information on the carcinogenic potential of chemicals is primarily available for High Production Volume (HPV) products. This is due to the limited knowledge gain from routine cancer bioassays and the fact that HPV chemicals are tested only where there is the need for more cost effective and informative testing strategies. Advanced genomics application has been developed for a cellular transformation assay to identify toxicity pathways and gene signatures predictive for carcinogenicity (Rohrbeck et al., 2010). Specifically, genome wide gene expression analysis and qRT-PCR has been applied to untransformed and transformed mouse fibroblast Balb/c 3T3 cells that were exposed to either 2, 4-diaminotoluene, benzo (a) pyrene, 2-acetylaminoflourene or 3-methycholanthrene at IC20 conditions for 24 h and 120 h, respectively. Then, bioinformatics was applied to define toxicity pathways and a gene signature predictive of the carcinogenic risk of these chemicals. While bioinformatics revealed distinct differences for individual chemicals at the gene level pathway analysis identified

common perturbation that resulted in an identification of 14 genes whose regulation in cancer tissue had already been established (Rohrbeck et al., 2010).

Microarrays are also being employed for investigations aiming to identify diagnostic biomarkers for different types of cancer (Tseng et al., 2001). 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 (Aardema et al., 2002). 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 (Suter et al., 2003; Rusyn et al., 2010). A typical set of data microarray data can show the identification of gene expression markers indicative of acute phase response in isolated mesenteric arteries from rats with fenoldopam-induced vasculitis (Lesko et al., 2003). 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 (Lesko et al., 2003).

Microarray-based classifiers and associated signature genes generated from various platforms are abundantly reported in the literature; however, the utility of the classifiers and signature genes in cross-platform prediction applications remains largely uncertain (Fan et al., 2010). As part of the MicroArray Quality Control Phase II (MAQC-II) project, the study has shown 80-90% cross-platform prediction consistency using a large toxicogenomics data set by illustrating that: (1) the signature genes of a classifier generated from one platform can be directly applied to another platform to develop a predictive classifier; (2) a classifier developed using data generated from one platform can accurately predict samples that were profiled using a different platform. The results suggest the potential utility of using published signature genes in cross-platform applications and the possible adoption of the published classifiers for a variety of applications. The study reveals an opportunity for possible translation of biomarkers identified using microarrays to clinically validated non-array gene expression assays. Gene expression analysis has been 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 (Suter et al., 2003). These early examples provide evidence that toxicogenomics can give insights into toxicological mechanisms and affected pathways. 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 single nucleotides polymorphisms (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 (Rusyn et al., 2010). The metabolic analyses performed by the chip offers 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 (Scherf et al., 2000). 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.

IMPACT ON THE PHARMACEUTICAL INDUSTRY

The potential benefits of a successful toxicogenomics program have been described briefly. Treatment with animals or cells in culture with new chemical entities (NCE) and examination of the resulting gene expression profiles can influence several areas of drug development (Shena et al., 1995; Collings and Vaidya, 2008). First, it can impact the quality of drug development pipelines by improving the science of toxicology, providing more specific information as to the mechanisms of drug pathologies and providing it earlier in the discovery–development continuum (Altman and Raychaudhuri, 2001). Second, it can improve the efficiency of the process because toxicogenomics information complements genomic target identification and characterization methods used in discovery and leads to reduced attrition during drug development for unfavourable compounds (Tseng et al., 2001; Reilly et al., 2005). Toxicogenomics can be applied at any stage in the drug development process, but appears to have greatest potential use when used in one or more of the following settings:

The risk to a pharmaceutical company for misunderstanding incurred with toxicogenomics approaches will depend on multiple factors, such as the technology and the type of studies employed (that is, *in vivo* and *in vitro*, examining reactions in animals or humans) (Stevens et al., 2006; Rusyn et al., 2010). In general, *in vivo* studies using global gene expression profiling platforms with compounds that have already advanced into clinical trials are considered to have the highest risk of uncovering some unexplained or uninterpretable toxicogenomics data (Olson et al., 2000;

Wen and Fitch, 2009). The least amount of risk to drug development would be seen with *in vitro* studies using only model compounds from the literature and clinical trial failures or less potent analogs from the discovery program of interest (Thomas, 2000; Salter and Nilsson, 2003). Alternatively, *in vivo* or *in vitro* systems that use a targeted approach in which only a few genes of known function are measured should be of low risk.

CONCERN ABOUT THE USE OF TOXICOGENOMICS

There are still a number of concerns around the use of gene expression data in drug risk assessment. There are technical concerns about the sensitivity and reliability of the methods (Ambroise and McLachlan, 2002; Steven et al., 2006). There are also concerns about the interpretation of the data, especially if genomic data are taken out of context. For example, genes such as c-myc, c-fos and c-Ha-ras which are associated with carcinogenesis may be found to have increased expression (Altman and Raychaudhuri, 2001; Ganter et al., 2008). These genes are not oncogenic by nature but are found to be mutated or highly over-expressed in tumours (Crosby et al., 2000; Foster et al., 2007). The increased expression in response to drug treatment may simply reflect an acute, and probably benign, stress response. They are, after all, genes for normal cellular functions in cell growth and viability (Smith, 2001; Guillouzo and Guguen-Guillouzo, 2008). The availability of practically the whole genome for expression analysis also brings difficulties in interpretation. There just is not enough information in the literature to interpret the modulation of expression of every single gene (Colling and Vaidya, 2008).

Until the knowledge base is complete, it must be accepted that toxicogenomic data will provide a starting point for further investigations and not necessarily give definitive answers. To address these concerns (with particular attention to using genomic data in the regulatory environment) a consortium of academic, governmental and industrial representatives formed a committee on the use of genomics in mechanism based risk assessment coordinated by the International Life Sciences Institute (ILSI) Health and Environmental Sciences Institute (HESI) (Derisi et al., 1996; Olson et al., 2000; Reilly et al., 2005; Wills, 2007). The committee's findings have shed much light on the technical issues and have shown the relevance of the data in understanding several mechanisms of toxicity.

CONCLUSION

Toxicogenomics can no longer be regarded as an emerging technology in drug development. The investments made in applying the technology growing within the pharmaceutical drug development process and through open debate in the toxicological community, highlight the fact that there is a stronger initiative to consolidate this technology into drug confidence in safety, either in selecting new chemical drug candidates in the evaluation of drug safety. It is now well established that gene expression profiles can discriminate between classes of compound and some toxicities, hence showing value in predicting toxicity. The development and applications of toxicogenomics will be facilitated by collaborations on technical issues, such as platform comparisons, data analysis/bioinformatics, and broad species coverage. Appropriate genetically defined mouse models may be combined with the limited data from human studies to not only discover the genetic determinants of susceptibility, but to also understand the molecular underpinnings of toxicity.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Medical and Pharmacotherapeutic Research Group Faculty of Medicine and Biomedical Science, University of Yaoundé 1, for their excellent work and many fruitful scientific discussions on this topic. In addition, we are indebted to the Athlone Institute of Technology for their technical, and financial support, and their North-South research collaboration efforts.

REFERENCES

Aardema MJ, MacGregor JT (2002). Toxicology and genetic toxicology in the new era of "toxicogenomics": impact of "-omics" technologies. Mutat. Res., 499: 13–25.
Altman RB, Raychaudhuri S (2001). Whole-genome expression analysis: challenges beyond clustering. Curr. Opin. Struct. Biol., 11: 340–347.
Alvarado ML, Wong I, Licinio J (2002). Advances in the pharmacogenomics of adverse drug reactions. Pharmacogenom, J. 2: 273-278.
Ambroise C, McLachlan GL (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. Proc. Natl. Acad. Sci. USA., 99: 6562–6566.
Amin RP, Hamadeh hK, Bushel PR, Bennett L, Afshar CA, Paules RS (2002). Genomic interrogation of mechanism(s) underlying cellular responses to toxicants. Toxicol., 181: 555–563.
Ayensu WK, Tchounwou PB, McMurray RW (2004). Molecular and cellular mechanisms associated with autoimmune diseases. Int. J. Environ. Res. Public Health., 1: 39-73.
Bartosiewicz MJ, Jenkins D, Penn S, Emery J, Juckpitt R (2001). Unique gene expression patterns in liver and kidney associated with exposure to chemical toxicants. J. Pharmacol. Exp. Ther., 297: 895–905.
Blomme EA, Yang Y, Waring JF (2009). Use of toxicogenomics to understand mechanisms of drug-induced hepatotoxicity during drug discovery and development.Toxicol Lett., 186: 22-31.
Bulera SJ, Eddy SM, Ferguson E, Jatkoe TA, Reindel BH, Bleavins JF, De La Iglesia FA (2001). RNA expression in the early characterization of hepatotoxicants in Wistar rats by high-density DNA microarrays. Hepatol., 33: 1239–1258.
Burczynski ME (2009). Pharmacogenomic approaches in clinical studies to identify biomarkers of safety and efficacy.Toxicol.

Lett., 186: 18-21.
Cheng RY, Alvord WG, Powell D, Kasprzak KS, Anderson LM (2002). Microarray analysis of altered gene expression in the TM4 Sertoli-like cell line exposed to chromium (III) chloride. Reprod. Toxicol., 16: 223–236.
Collings FB, Vaidya VS (2008). Novel technologies for the discovery and quantitation of biomarkers of toxicity. Toxicol., 245: 167-174.
Cronin M, Ghosh K, Sistare F, Quackenbush J, Vilker V, O'Donnell C (2004). Universal standards for gene expression. Clin. Chem., 50(8): 1464–1471.
Crosby LM, Hyder KS, DeAngelo AB, Kepler TP, Gaskill B, Benavides GR, Yoon R, Morgan KT (2000). Morphologic analysis correlates with gene expression changes in cultured F344 rat mesothelial cells. Toxicol. Appl. Pharmacol., 169: 205–221.
Cunningham MJ (2000). Genomics and proteomics. The new millennium of drug discovery and development. J. Pharmacol. Toxicol. Methods., 44: 291–300.
Cunningham ML, Bogdanffy MS, Zacharewski TR, Hines RN (2003). Workshop overview: use of genomic data in risk assessment. Toxicol. Sci., 73: 209–215.
DeRisi JL, Penland PO, Brown ML, Bittner PS, Meltzer M, Ray Y, Chen YA, Trent JM (1996). Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat. Genet., 14: 457–460.
Farr S, Dunn RT (1999). Concise review: gene expression applied to toxicology. Toxicol. Sci., 50: 1-9.
Fielden M, Zacharewski TR (2001). Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology.Toxicol Sci., 60: 6-10.
Foster WR, Chen JS, He A, Truong A, Bhaskaran V, Nelson DM, Dambach LD, Lehman-McKeeman P, Car BD (2007). A retrospective analysis of toxicogenomics in the safety assessment of drug candidates.Toxicol. Pathol., 35: 621-35.
Fountoulakis MM, de Vera F, Crameri F, Boess R, Gasser S, Albertini L, Suter E (2002). Modulation of gene and protein expression by carbon tetrachloride in the rat liver. Toxicol. Appl. Pharmacol., 183: 71–80.
Ganter B, Zidek N, Hewitt PR, Müller D, Vladimirova A (2008). Pathway analysis tools and toxicogenomics reference databases for risk assessment. Pharmacogenom, 9: 35-54.
Gibson RJ, Bowen JM, Keefe DM (2008). Technological advances in mucositis research: new insights and new issues. Cancer Treat. Rev. 34: 476-82.
Guguen-Guillouzo C, Guillouzo A (2010). General review on *in vitro* hepatocyte models and their applications. J. Biosci. Bioeng., 10: 5-11.
Guillouzo A, Guguen-Guillouzo C (2008). Evolving concepts in liver tissue modeling and implications for *in vitro* toxicology. Expert Opin Drug Metab.Toxicol., 4: 1279-1294.
Hamadeh HK, Bushel PR, Jayadev S, DiSorbo O, Bennett L, Li S, Tennant R, Stoll JC, Barret P, Paules RS (2002). Prediction of compound signature using high density gene expression profiling. Toxicol. Sci., 67: 232–240.
Harries HM, Fletcher ST, Duggan CM, Baker VA (2001). The use of genomics technology to investigate gene expression changes in cultured human liver cells. Toxicol. *In Vitro*, 15: 399–405.
Harill AH, Rusyn I (2008). Systems biology and functional genomics approaches for the identification of cellular responses to drug toxicity.Expert Opin Drug Metab Toxicol., 4: 1379-89.
Holmes E, Nicholson JK (2001). Characterization of genetic variations in toxicological and metabolic responses using probability neural networks. Chem. Res. Toxicol., 14: 182-191.
Joseph LJ (2004). RNA reference materials for gene expression studies. RNA metrology: forecast calls for partial clearing. Clin Chem., 50(8): 1290–1292.
Kerr MK, Churchill GA (2001). Statistical design and the analysis of gene expression Microarray data. Genet. Res., 77: 123–128.
Lennon GG (2000). High-throughput gene expression analysis for drug discovery. Drug Discov. Today, 5: 59–66.
Lesko LJ, Salerno RA, Spear BB, Anderson DC, Brazell C, Collins J, Dorner A, Essayan D, Gomez-Manzilla B (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.
Liu A, Zhang, Gehan E, Clarke R (2002). Block principal component analysis with application to gene microarray data classification. Stat. Med., 21: 3465–3474.
Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H (2007). Expression monitoring by hybridization to high-density oligonucleotide arrays. Nat. Biotechnol., 14: 1675–1680.
Madea B, Saukko P, Oliva A, Musshoff F (2010). Molecular pathology in forensic medicine-Introduction. Forensic Sci. Int. 2010 Aug 19 Pharmacogenomics. 11(8): 1127-36.
Morgan KT (2002). Gene expression analysis reveals chemical-specific profiles. Toxicol. Sci., 67: 155–156.
Olson H, Betton G, Robinson D (2000). Concordance of the toxicity of pharmaceuticals in humans and animals. Regul. Toxicol. Pharmacol., 32: 56-67.
Petricoin EF, Hackett JL, Lesko LJ, Puri RK, Gutman SI, Chumakov PG, Woodcock JJ, Feigal DW, Zoon KC, Sistare FD (2002). Medical applications of microarray technologies: a regulatory science perspective. Nature Genet. Suppl., 32: 474–479.
Reilly TP, Bourdi M, Brady JN (2005). Expression profiling of acetominophen liver toxicity in mice using microarray technology. Biochem. Biophys. Res. Common., 27: 321-328.
Rohrbeck A, Salinas G, Maaser K, Linge J, Salovaara S, Corvi R, Borlak J 2010.Toxicogenomics applied to *in vitro* carcinogenicity testing with Balb/c 3T3 cells revealed a gene signature predictive of chemical carcinogens. Toxicol. Sci., 16: 201-212.
Ruepp SU, Tonge RP, Shaw J, Wallis N, Pognan F (2002). Genomics and proteomics analysis of acetaminophen toxicity in mouse liver. Toxicol. Sci., 65: 135–150.
Rusyn I, Gatti DM, Wilshire T, Kleeberger SR, Threadgill DW (2010). Toxicogenetics: population-based testing of drug and chemical safety in mouse models. Pharmacogenomics J. Aug., 10(4): 247-257.
Ryan TP, Stevens JL, Thomas CE (2008). Strategic applications of toxicogenomics in early drug discovery.Curr. Opin. Pharmacol., 8(5): 654-660.
Salter AH, Nilsson KC (2003). Informatics and multivariate analysis of toxicogenomics data. Curr. Opin. Drug Discov. Dev., 6: 117–122.
Schena M (1996). Genome analysis with gene expression Microarray. Bioessays., 18: 427-431.
Scherf U, Ross DT, Waltham M, Smith LH, Lee JK, Tanabe L,. Kohn KW, Reinhold. WC, Myers TG, Andrews DT (2000). A gene expression database for the molecular pharmacology of cancer. Nat. Genet., 24: 236–244.
Shah RR (2006). Can pharmacogenetics help rescue drugs withdrawn from the market? Pharmacogenom, 7: 889-908.
Sherlock G (2000). Analysis of large-scale gene expression data. Curr. Opin. Immunol., 12: 201–205.
Smith LL (2001). Key challenges for toxicologists in the 21st century. Trends Pharmacol. Sci., 22: 281–285.
Steiner S, Anderson NL (2000). Expression profiling in toxicology—potentials and limitations. Toxicol. Lett., 112: 467–471.
Stevens JL, Liu H, Halleck M, Bower RC, Chen QM, van De WB (2006). Linking gene expression to mechanisms of toxicity. Toxicol. Lett., 112-113: 479-486.
Suter L, Haiker M, de Vera MC, Albertini S (2003). Effect of two 5-HT6 receptor antagonists on the rat liver: a molecular approach. Pharmacogenomics J. 3: 320–334.
Szyf M (2007). The dynamic epigenome and its implications in toxicology.Toxicol. Sci., 100: 7-23.
Thomas RS, Rank DR, Penn SG (2000). Identification of toxicologically predictive gene sets using cDNA microarrays. Mol. Pharmacol., 60: 1189-1194.
Tseng GC, Rohlin L, Liao JS, Wong WH (2001). Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. Nucleic Acids Res., 29: 2549–2557.
Ulrich R, Friend SH (2002). Toxicogenomics and drug discovery: will new technologies help us produce better drugs? Nat. Rev. Drug Discov., 1: 84–88.
Van Dijck PW, Selten GS, Hempenius RA (2003). On the safety of a new generation of DSM Aspergillus niger enzyme production strains. Regul. Toxicol. Pharmacol., 38:27-35.

Waring JF, Ciurlionis R, Jolly RA, Heindel M, Ulrich.RG (2001). Microarray analysis of hepatotoxins *in vitro* reveals a correlation between gene expression profiles and mechanisms of toxicity. Toxicol. Lett., 120: 359–368.

Wen B, Fitch WL (2009). Analytical strategies for the screening and evaluation of chemically reactive drug metabolites.Expert Opin Drug Metab. Toxicol., 5: 39-55.

Wills Q (2007). SimuGen Ltd: reliable, early prediction of drug toxicity with toxicogenomics, human cell culture and computational models. Pharmacogenom, 8: 1081-1084.

Yasokawa D, Iwahashi H (2010). Toxicogenomics using yeast DNA microarrays. Toxicol. Appl. Pharmacol., 2010 Jul 16. [Epub ahead of print].

Young MB, DiSilvestro MR, Sendera TJ, Freund J, Kriete A, Magnuson SR (2003). Analysis of gene expression in carbon tetrachloride-treated rat livers using a novel bioarray technology. Pharmacogenom. J., 3: 41–52.