Toxicogenomics in Drug Development

Nelson Guerreiro, Frank Staedtler, Olivier Grenet, Jeanne Kehren and Salah-Dine Chibout

Toxicol Pathol 2003 31: 471
DOI: 10.1080/01926230390224656

Toxicologic Pathology, 31:471–479, 2003

ISSN: 0192-6233 print / 1533-1601 online
DOI: 10.1080/01926230390224656

Toxicogenomics in Drug Development

NELSON GUERREIRO, FRANK STAEDTLER, OLIVIER GRENET, JEANNE KEHREN, AND SALAH-DINE CHIBOUT

Novartis Pharma AG, Pharmacogenomics, Preclinical Safety, Basel, Switzerland

ABSTRACT

Toxicogenomics represents the merging of toxicology with technologies that have been developed, together with bioinformatics, to identify and quantify global gene expression changes. It represents a new paradigm in drug development and risk assessment, which promises to generate a wealth of information towards an increased understanding of the molecular mechanisms that lead to drug toxicity and efficacy, and of DNA polymorphisms responsible for individual susceptibility to toxicity. Gene expression profiling, through the use of DNA microarray and proteomic technologies will aid in establishing links between expression profiles, mode of action and traditional toxic endpoints. Such patterns of gene expression, or 'molecular fingerprints' could be used as diagnostic or predictive markers of exposure, that is characteristic of a specific mechanism of induction of that toxic or efficacious effect. It is anticipated that toxicogenomics will be increasingly integrated into all phases of the drug development process particularly in mechanistic and predictive toxicology, and biomarker discovery. This review provides an overview of the expression profiling technologies applied in toxicogenomics, and discusses the promises as well as the future challenges of applying this discipline to the drug development process.

Keywords. DNA microarrays; PCR; bioinformatics; gene expression profiling; genomics; proteomics; biomarkers; toxicology.

INTRODUCTION

The evolution of new innovative technologies in parallel with recent dramatic increases in genomic knowledge is anticipated to revolutionize toxicological studies by providing significant advances in the understanding and prediction of the toxicity and efficacy of new drugs (38, 39, 52). In classic toxicology, potential adverse effects resulting from drug exposure are evaluated using endpoints such as body and organ weight changes, biochemical and histopathological observations. Such observations, however, do not provide information about a drug's mode of action. To better evaluate the adverse effects associated with drug exposure, one needs to understand the drug's specific mode of action. Drugs are expected to induce a multitude of complex molecular perturbations in a wide variety of pathways, involving differential gene expression at the transcript and functional protein level, leading to efficacious and/or pathological outcomes. These changes in gene expression are often more sensitive and characteristic of the toxic response or process than currently employed endpoints of pathology, and have the potential to indicate toxicity already at lower doses or at earlier time points.

Technological advances derived from genomic research have made it possible to follow transcriptional and translational events of genes and even the entire genome. The application of genome-wide expression profiling technologies to toxicology has created a new subdiscipline coined 'toxicogenomics,' which has the potential to provide a more comprehensive understanding of the mechanisms of pharmacology and toxicity than has been possible in classical toxicology approaches. The shotgun approach is exemplified by DNA microarray technologies that have received much attention recently by the toxicology community (13). These genomic technologies, along with the development of specialized bioinformatics tools, rely on the use of platforms that allow the high-throughput quantitative comparison of the transcriptional activity of potentially thousands of individual genes from untreated and treated groups in a single experiment, thereby providing a holistic interrogation of cellular responses to drug exposure. Transcripts that are modulated in response to drug treatment, represent a characteristic gene expression profile or 'molecular signature' of the drug effect, and may potentially serve as a diagnostic or predictive signature of drug efficacy or for certain forms of toxicity. It is anticipated that by comparing the expression profiles of drug candidates to those from reference compounds with well-characterized pharmacological and toxicological endpoints more informed decision can be made on the prioritization of new drugs early in the development phase (34, 79).

Other shotgun approaches for genome-wide analysis of cellular constituents will also have a major impact on the field of toxicology. Proteomics is a field that deals with the global separation, quantitation and functional characterisation of expressed proteins. Improvements in classical two-dimensional gel electrophoresis (2DGE), continued development of mass spectrometry techniques for rapid identification and characterization of proteins, introduction of multidimensional liquid chromatography, and protein chip systems, have allowed proteomics to provide complementary information to genomics data (53).

Preclinical studies are especially suited for application of these technologies as the repeated administration of a low dose of the compound is generally in the pharmacological range and the high dose clearly toxic in order to detect toxicological changes in target organs. Should toxicological findings occur, mechanistic studies employing toxicogenomics might be undertaken to generate information on the pathogenesis. The messenger RNA (mRNA) or protein expression profiles are then linked to data from measurements of toxicological or pharmacological endpoints obtained by

This paper evolved from a presentation at the STP/IFSTP Annual Symposium "Toxicologic Pathology in the New Millenium" held in Orlando, Florida, June 24–28, 2001.

Address correspondence to: Nelson Guerreiro, Novartis Pharma AG, WKL-136.1.18, 4002 Basel, Switzerland; e-mail: nelson.guerreiro@pharma.novartis.com

conventional methods such as histopathology or clinical biochemistry to generate hypotheses about the drugs mode of action. Understanding the mode of action may provide information on whether the change observed is species specific or predictive of human risk. Since the initial publications (47, 65, 69, 83), DNA microarrays have found widespread interest in drug research. The tremendous potential for the application of DNA microarrays in drug development and in particular in toxicology is just going to be realized (7, 33, 40, 51, 64). This review will provide an overview of mRNA and protein expression array technologies used in the field of toxicogenomics, and their application to facilitate discovery of mechanistic pathways associated with drug related effects, outlining both the strengths and challenges of this approach.

High Density DNA Microarrays Technology

For genome wide transcript profiling several established methods exist such as differential display of mRNA (43–45), subtractive hybridization (72, 77), serial analysis of gene expression (76) or amplifications fragment length polymorphism (48). Examples for the application of these approaches in toxicology have recently been reviewed (80). In the recent years the availability of sequence information for whole genomes from humans and other species has allowed the construction of DNA microarrays, that enable quantitative comparisons of the transcriptional activity of potentially tens of thousands of individual genes between different biological samples (65).

Microarray platforms have been developed by commercial vendors, pharmaceutical companies and academic institutions. Although there are a number of different technical platforms, the basic principles are similar. A high-density microarray is a small, solid support, onto which hundreds or thousands of oligonucleotides or cDNAs, each representing an individual gene are covalently attached at defined positions. From cells or tissues of interest, mRNA or total RNA is extracted, reverse transcribed into cDNA, fluorescently labeled and then hybridized to a DNA microarray. Complementary molecules bind to their counterparts via Watson-Crick base pairing on the microarray. Following laser excitation, the microarray is scanned, and signal intensities using specialised software are normalised to adjust for labeling and detection efficiencies for different fluorescent labels and differences in the quantity of starting mRNA between samples. Subsequently, a comparison of control with test samples permits quantitative assessment of changes in gene expression associated with treatment or disease (16, 49, 81). The data of experiments with DNA microarrays are represented as a matrix of fluorescent intensities, each value corresponding to a spot on the microarray. Multicondition experiments are represented by a matrix of gene expression values, with genes in rows and conditions (or cDNA samples) in columns. Any biological sample from which high quality mRNA can be isolated may be used for microarray analysis.

Two major types of high density DNA microarrays are used in gene expression analysis: cDNA and oligonucleotide arrays. Oligonucleotide-based microarrays are fabricated using photolithographic processes to generate the sequence order in the oligonucleotide synthesis on-chip, resulting in the generation of high density arrays of short oligonucleotide (~20 bases) probes that are synthesized on the glass surface or substrate in predefined positions (55, 71). Affymetrix GeneChips (Affymetrix, Inc, Santa Clara, CA) are a popular type of oligonucleotide array (84). cDNA microarrays are generated by robotic deposition of PCR fragments (0.5–2 kb in length) amplified from cDNA clones, onto a glass substrate or chip in predefined positions (65, 66). Up to 25,000 spots of cDNAs may be arrayed on a single chip. Microarrays have rapidly become commercially available from a number of sources for mice, rats, and humans. In the near future microarrays will also become available for other species. Another interesting development has been in the manufacture of custom-designed chips to profile genes involved in a variety of toxic responses or pathways such as apoptosis. TOXCHIP v1.0 is one such chip that has been developed by the National Institute of Environmental Health Sciences in the USA. Similarly, other chips have been developed using genes selected among potential toxic markers involved in basic cellular processes and drug metabolism (17, 56, 57).

Rapid and Quantitative Methods for Validating Microarray Gene Expression Data

After a list of genes has been compiled and annotated, it is usual to select a subset of these genes to independently validate changes in their expression. Real-time, quantitative RT-PCR (qRT-PCR) is especially useful in this capacity. For highly quantitative results the TaqMan probe method (36) is the preferred approach. Alternatively, to reduce the cost of this expensive technique, SyBr Green can also be used to label PCR products (67). The advantages of the PCR gene expression analysis lies in the speed of the technique as well as the quantitative data obtained. The dose and time-dependencies are very important for the understanding and the evaluation of the biological relevance of a particular gene marker. The quantitative results obtained by PCR allow quantitative correlation with the extent of the pharmacological effect of a compound or with the grading of a pathological finding. This correlation is mandatory to prove the relevance of a marker for a pharmacological or toxicological effect of a compound. In addition, once the relevant markers are identified, in situ-hybridization (ISH) allows the localization of the expression of the gene marker at the mRNA level in a specific target tissue. Tissue microarrays are now commonly used for high throughput ISH screening for tissue-specific gene expression patterns on a microarray of paraffin-embedded tissues (22).

Two-Dimensional Gel Electrophoresis Based Proteomics Technology

The interpretation of expression data from microarrays requires caution as the use of mRNA levels by themselves is insufficient for understanding protein abundance and function. Posttranscriptional mechanisms, including protein translation, co- and posttranslational modification, alternative mRNA splicing, protein turnover, and posttranscriptional regulation of gene expression are essential cellular processes that make it difficult to extrapolate from mRNA to protein profiles and cellular function (5, 32). For example, a recent study compared the relationship between mRNA and protein expression for a cohort of genes in the same lung adenocarcinomas using oligonucleotide microarray and 2DGE. Among the 69 genes for which only a single 2D gel protein spot was known, only 9 genes were observed to have a statistically significant

relationship between protein and mRNA abundance (14). In addition, discrepancies between mRNA and protein levels were also noted among proteins with multiple isoforms, potentially indicating isoform specific mechanisms for the regulation of protein abundance. The author concluded that posttranslational mechanisms play a major role controlling the complex patterns of protein abundance and posttranslational modifications. These conclusions, however, were based on the simultaneous measurement of mRNA and protein at just a single time point. Lian et al (42) recently investigated the global relationship between mRNA and protein over multiple time points during myeloid differentiation. In this study a moderately good correlation was found for the 51 proteins analyzed, which was sufficiently strong enough to indicate that the regulation of transcript levels is likely a major determinant of changes in protein levels during myeloid differentiation (42). Nonetheless, both these studies represent the importance of proteomics as a complementary technology to DNA microarrays, and in extending and validating genomics data (37).

To date, the most widely used proteomics platform for protein profiling has been 2DGE. This technique involves the separation of proteins in the first dimension according to their charge, by isoelectric focusing, and then in a second dimension according to their molecular mass, by sodium dodecyl sulfate polyacrylamide gel electrophoresis. These 2 protein parameters make possible the separation and quantitation of thousands of proteins, from complex protein mixtures, as a constellation pattern of spots on a single gel. By comparing the profiles of proteins present in drug-treated samples with those present in untreated tissue, it is possible to identify changes in expression of proteins that can yield information on response to drug treatment. Specific spots whose intensities differ between two gel groups can then be targeted for identification. Proteomics has only recently surfaced as a viable entity in drug development because of advances in high-throughput analytical tools for protein identification (25). Using mass spectrometry (MS) and robotics it is now possible to identify proteins with high speed and sensitivity. Peptide mass fingerprinting using MALDI-TOF (matrix-assisted laser desorption/ionization-time-of-flight) MS has emerged as a protein identification tool particularly suited to automated, high-throughput screening of proteins separated on 2D gels (24). Once proteins are separated by 2DGE, they are visualized and the protein spot profiles are analyzed using dedicated image analysis tools. Robotic systems excise the proteins of interest and carry out digestion with a protease, usually trypsin, followed by transfer in microtiter format to a MALDI-TOF plate for subsequent analysis of the tryptic peptide masses by MS. The peptide mass values are used as unique fingerprints (peptide mass fingerprints) and compared with databases of theoretical tryptic digests of known or hypothetical proteins (15, 27). In situations where no match is obtained with the peptide mass fingerprint approach, more sophisticated MS analysis, typically tandem mass spectrometry (MS/MS), can be performed to generate sequence information. Recently developed MS approaches such as quadrupole time-of-flight mass spectrometer fitted with a MALDI ion source (MALDI qTOF) (70), and MALDI-TOF/TOF instruments (85) have remarkable potential for high-throughput de novo sequencing of tryptic peptides and characterization of posttranslational modifications. Used in combination with peptide mass analysis, MS/MS data can greatly reduce the risk of ambiguity arising when MS data match more than one protein sequence.

Bioinformatics

The amount of information generated when using the array technologies is not amenable to manual analysis. Sophisticated data management and bioinformatics tools are essential to efficiently process and analyze large amounts of data, and to facilitate pattern recognition across several time points or dose levels. Thus, bioinformatics has emerged as a major catalyst in the development of genomic and proteomic technologies. Different academic and commercial software packages have been developed for this application. Packages typically include spot quantitation, data storage and retrieval, and higher level analysis.

For DNA microarray analysis, algorithms comparing the data between rows (genes) and columns (samples), in order to sorting out the "noise" (nonspecific variations) from the "signal" are used. These include GeneSpring (Silicon Genetics, Redwood City, CA, USA), GeneSight (Biodiscovery, Marina Del Rey, CA, USA). For an experiment usually involving the study of time series or dosage series, once a measure of similarity (or distance) between individual gene profiles has been assigned, these may be divided into groups or clusters to detect which genes display highly correlated expression patterns. The grouping is usually performed using clustering methods. Brazma et al (10) have provided a comprehensive review of clustering algorithms in the analysis of expression data. GeneSpring, like other common software packages, provides statistical clustering tools to assist with the interpretation and visualisation of complex multivariate gene expression data, and to facilitate the sorting of the most significant variations.

Clustering was first described by DeRisi et al (19), who discovered that genes with similar expression profiles during metabolic shift in yeast were functionally related. Viewing the clustered expression data in an interactive graphical interface helps to understand the global trends in the data set and can be useful for generating hypotheses. The graphic display of the data can draw attention to an interesting profile during a time course experiment. Once the profile is identified the software can find all the genes following the same variation pattern. The analysis is based on matrix calculation: each gene is defined by its coordinate in a n-dimensional space, n being the number of measures for the genes. Clustering algorithms can then group genes according to the distance from one to another and define groups of genes with a similar behavior over the different experiments. This analysis will group genes exhibiting no significant changes together and sort those genes exhibiting a monotone decrease or increase; genes decreasing at one point and increasing at the following point. In GeneSpring, different parameters can be integrated in order to link the expression data with other data such as histopathological grading. The data are then analyzed as a function of the new parameters and the same sorting of profiles can be performed. Another significant development, especially well suited to industry, is in enterprise-wide solutions to array data storage, retrieval and high level queries. Rossetta Inpharmatics, Inc. (Kirkland,

Washington, USA) and others such as GeneData AG (Basel, Switzerland), have developed integrated computational tools for queries and interrogation of large array data sets in a GenBank-like DNA microarray database. This assimilation of large amount of data collected in a controlled fashion could be employed to identify highly defined expression patterns with diagnostic or predictive potential.

For 2-D based proteomics, there are two levels where bioinformatics plays an integral part both in the speed and value of the analysis. The first of these levels is in the mapping, quantification, and analysis of the gel images using specialized software programs capable of analyzing the pattern of protein spots on the gel and identifying differences between gels. However, considerable improvements are still needed to overcome the labour intensive and often troublesome process of image analysis. Better spot recognition, alignment and matching algorithms that facilitate comparison of 1,000s of proteins across multiple gel images in a comprehensive, reproducible, and automated fashion are required. Some companies, such as Scimagix (Redwood Shores, CA, USA), Definians (Munich, Germany) and NonLinear Dynamics (Newcastle, UK) have engaged in development of new algorithms and automated packages that look promising. Matching data can be directly imported into sophisticated statistical packages, such as Impressionist (GeneData), for the identification of statistically significant changes in protein expression. The rate of detecting false-positives can also be reduced by performing replicate analysis on individual samples. In the next level, sophisticated algorithms are employed for MS data annotation, and for the identification and characterisation of proteins using MS derived data. MS data are used by software programs to search both publicly available and internal protein sequence databases, comparing what would be expected if the protein entry in the database was subjected to the same pattern of fragmentation.

Toxicogenomics Applications

Toxicogenomics can greatly influence the drug development process by: increasing our knowledge of molecular mechanisms of toxicity and efficacy; providing sensitive biomarkers for better monitoring of compound effects in clinical trials; providing new indications for drug candidates; providing more informed decisions regarding safety as well as efficacy of compounds; enhancing the ability to extrapolate accurately between experimental animals and humans in the context of risk assessment; and providing a better understanding of the influence of genetic variation on toxicological outcomes.

In Novartis, toxicogenomics is being increasingly used as a tool for mechanistic or exploratory studies in preclinical and clinical studies, with the general purpose of generating hypothesis to aid in decision making. Target organs and/or body fluids are selectively sampled, and toxicogenomics data is integrated with data from conventional endpoint measurements (clinical chemistry, biochemistry, and histopathology). In addition to traditional parameters, knowledge of the molecular mechanisms of drug response is essential to a drug's development, as many regulatory agencies are now placing increased value on mechanistic information for improving the risk assessment process. A better understanding of the molecular mechanisms of drug response is key to establishing the relevance of animal data to humans, and also in identifying species specific responses. An improved understanding of the toxicity manifested in experimental animal models will improve the predictive accuracy of extrapolating from animal models to humans, and therefore, greatly reduce this source of uncertainty in the risk assessment process. For example, in our labs proteomics and quantitative RT-PCR played a key role in the mechanistic understanding of cyclosporine A (CsA) induced nephrotoxicity (28, 74), which is a common side effect observed for immunosuppressant drugs. In kidneys of CsA-treated rats, a profound downregulation of the calcium binding protein calbindin-D28kDa was found to correlate with the accumulation of calcium in the tubules (Figure 1).

The role of calbindin-D28kDa in the kidney is to bind and transport the calcium in renal tubules, and its downregulation was found to be the cause of the observed intratubular corticomedullary calcification in the kidney following CsA treatment. Indeed, ISH hybridization showed that the expression of calbindin-D28kDa is restricted to distal tubules in the kidney (Figure 2). Subsequent studies involving various CsA derivatives and other immunosuppressive compounds such as FK506 and rapamycin also revealed a downregulation of calbindin-D28kDa (3). In addition, marked decrease in renal calbindin-D28kDa protein level was found in most of the kidney biopsy sections from CsA-treated human kidney-transplant recipients with renal vascular or tubular toxicity, but not in dogs and monkeys, which are generally devoid of CsA-mediated nephrotoxicity (4). Prior to these toxicogenomic studies, the relationship between CsA induced kidney toxicity and calbindin-D28kDa downregulation was not known, and emphasizes the role toxicogenomics can have in providing essential information in mechanistic toxicology and facilitating the identification of novel, clinically relevant biomarkers of toxicity or potential adverse effects.

The possibility that a specific group or class of compounds (grouped by toxic endpoint, mechanism, target organ, structure, etc) may induce 'molecular fingerprints' of gene expression changes is the basis for the application of toxicogenomics to predictive toxicology and to prioritize new drug candidates in early development (34, 35, 75). Having defined 'molecular fingerprints' that are diagnostic for certain forms of toxicity, gene expression profiles induced by candidate drugs in the same model system can then be compared with the established and validated signatures. Several laboratories have recently shown that expression profiles do indeed align with known toxic mechanisms for tissues from treated animals (8, 12, 58). Waring et al (78, 79) demonstrated the ability to cluster known hepatotoxins using the gene expression profiles from the livers of rats treated with these agents. Using computer algorithms and statistical approaches gene expression profiles induced by the agents were found to form clusters. These clusters strongly correlated with the histopathology findings and clinical chemistry values induced by the hapatotoxins. These results show that toxicogenomics has promise as a diagnostic or predictive tool for toxicology.

Companies such as GeneLogic (Gaithersburg, USA) is a leading provider in gene expression-based biological information. Extensive gene expression databases of reference compounds whose toxicological and pathological endpoints

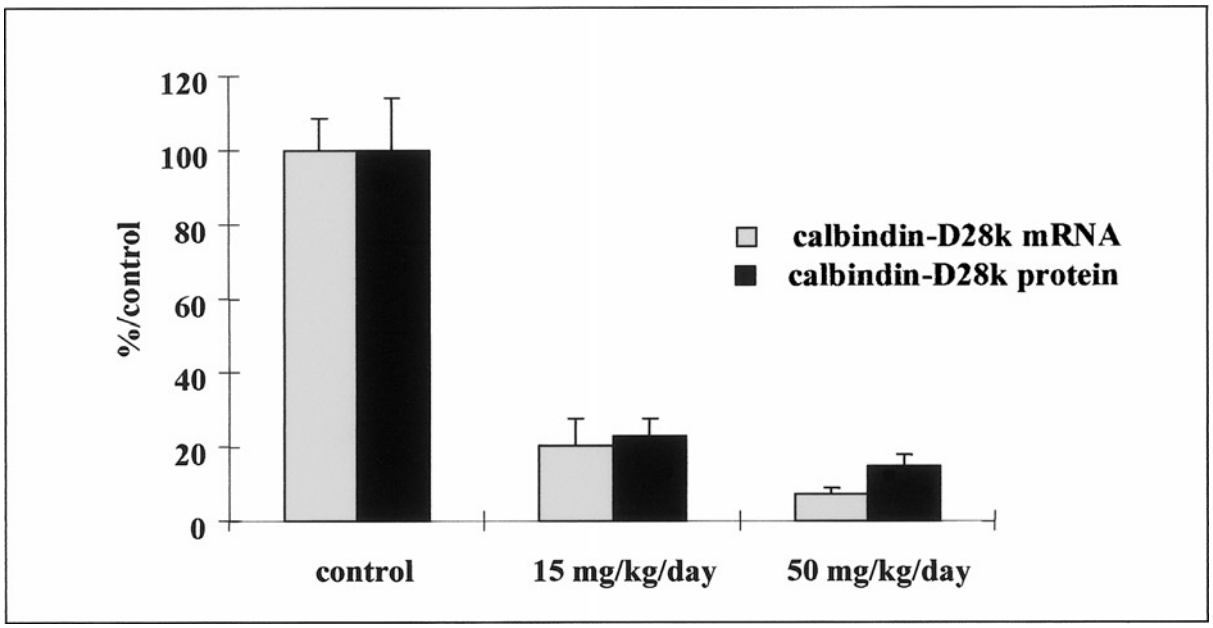

FIGURE 1.—Calbindin-D28kDa mRNA and protein levels in kidneys of CsA-treated rats. The ELISA and the RT-PCR data from treated rats are presented as percent of the control rats. The bars represent the means ± SD of 8–10 rats. From Grenet et al (28).

are well characterized are now available for interrogation with expression profiles of candidate drugs. BioExpress (GeneLogic) is a repository of gene expression profiles from normal and diseased human tissue and experimental animal tissues treated with reference compounds. A positive correlation with an archived profile could lead to some knowledge regarding the potential adverse effects of new drugs and also lead to new indications. In our labs we have recently filed a patent pertaining to the use of a diagnostic gene expression pattern for monitoring renal disorders (Patent no: 4-32567P1). Using clustering algorithms (GeneSpring), 5 genes were uniquely clustered in response to test compounds and Neoral, and their expression levels strongly correlated to renal pathology status. One practical application of this is to help rank a series of compounds based upon the expression of a set of genes indicative of renal toxicity. This approach holds promise for aiding decision making on the likelihood of compound success early in the discovery process before initiation of costly development studies, for monitoring treatment-related renal effects in a clinical setting, and allowing for a rapid reevaluation of dose regimes for individual patients.

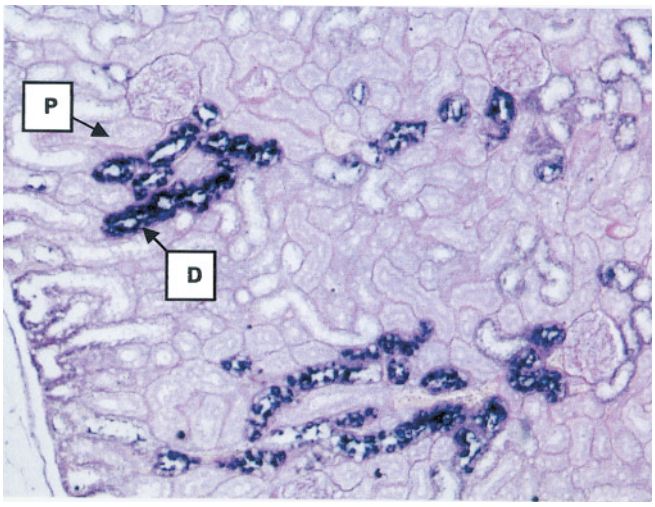

FIGURE 2.—Detection of the calbindin-D28kDa mRNA in the rat kidney. ISH hybridization shows that the expression of this gene is restricted to distal tubules in the kidney. P = proximal tubule. D and black staining = distal tubules. Magnification ×25.

One of the driving forces of toxicogenomics in drug development and risk assessment is the discovery and utilization of biomarkers in body fluids to advance clinical diagnostics and therapeutic monitoring. Proteomics technologies are particularly amenable to the analysis of body fluids, which is of particular importance for 2 main reasons: body fluids such as plasma, urine, cerebrospinal fluid, or synovial fluid are easily obtained and represent a rich source for identifying biomarkers of pathological states as the protein composition in body fluids is derived from many tissues and processes. Body fluids such as plasma, however, is exceptionally difficult to analyze due to the large proportion of albumin (55%) and a dynamic range of more than 10 orders of magnitude in concentration that separate the highest and lowest abundant proteins (6). Considering the combinatorial effect of posttranslation modifications, plasma is estimated to contain many thousands perhaps millions of peptides. To address this, approaches such as those that utilize liquid chromatography techniques coupled online to MS instruments have been developed to analyze peptides and proteins present in body fluids (68, 73).

Recently, Adkins et al (1) used ion exchange and reverse-phase chromatography coupled with MS to identify a total of 490 different proteins in an immunoglobulin-depleted tryptic digest of human serum. Among the proteins identified included the low abundance serum proteins interleukin-12

and human growth hormone demonstrating the sensitivity and capacity of this approach to comprehensively analyze proteins in body fluids. An extension of this approach would be to employ protein or peptide labeling strategies for accurate quantitation and expression profiling by MS (29–31, 61). SELDI-TOF MS is another approach to profile proteins in body fluids, and has recently been used to detect tumor markers in human plasma that have predictive value in ovarian cancer (59). Novartis and GeneProt (Geneva, Switzerland) recently announced a partnership to analyze the proteome in the plasma of patients with coronary artery disease. Novartis intends to use proteomics information generated by GeneProt to obtain potential therapeutic proteins, or drug targets to be used in Novartis' lead discovery, or as diagnostic biomarkers. Having identified the presence of potential protein or peptide markers in body fluids, samples can be acquired by noninvasive techniques and converted into standard immunoassays for routine screening in a clinical setting. Alternatively, the screening of validated protein signatures in body fluids that discriminate pathological states is envisioned in the future to play an important role in diagnostics and therapeutics.

Due to the novelty of toxicogenomics as a tool in drug development, to date, there are few published studies showing the applicability of gene expression profiling in clinical trials. Recently, however, gene expression profiling was incorporated into a proof of concept phase I/II study in which the efficacy, safety and tolerability of orally administered pimecrolimus was evaluated in psoriasis patients (60). Gene expression profiling of blood cells of individuals during therapy identified a common genomic profile with a downregulation of genes associated with the known target pathway of pimecrolimus, inflammation, proliferation, chemotaxis, and migration of leukocytes, but no changes in gene expression that might be linked to treatment-related immunosuppression and toxicity. This expression profile was highly consistent with the observed clinical efficacy and tolerability of pimecrolimus, allowing the authors to conclude that pimecrilimus taken orally is highly effective on patients with psoriasis and well tolerated (60). In addition, the gene expression findings from this study suggest that pimecrolimus has unique anti-inflammatory qualities that may be effective for other inflammatory skin diseases.

Clearly, the integration of toxicogenomics into this proof of concept phase I/II study proved to be a powerful approach in obtaining a broad and detailed description of patient response to systemic treatment with pimecrolimus. Furthermore, an important by-product from the application of toxicogenomics in preclinical and clinical studies is that comprehensive drug response profiling can potentially lead to the serendipitous discovery of novel disease indications. In another study, DePrimo et al (18) described expression profiling of peripheral blood cells as an approach for clinical biomarker discovery. Expression profiling using DNA microarrays was applied to blood mononuclear cells obtained from patients with advanced colorectal cancer participating in Phase III clinical trials. Data mining of expression profiles from samples before and at the end of one treatment cycle with a VEGF receptor tyrosine kinase inhibitor, SU5416, revealed changes in 4 transcripts that consistently correlated with SU5416 administration. The analysis of additional clinical samples indicated that the expression profile of these transcripts could be used to predict the treatment regime to which patients belonged (18).

An important impact toxicogenomics will have on drug development is on the identification and mapping of variant genetic sequences (polymorphisms) responsible for individual susceptibility to toxicity from particular drugs (11, 23, 63). Methodologies such as pharmacogenetics are being applied to define polymorphic responses (62). Gene chip arrays designed to examine single nucleotide polymorphisms (SNPs) in individuals have been developed. For example, the GeneChip CYP450 array (Affymetrix, Inc, Santa Clara, CA) contains genotype information for the cytochrome P450 drug metabolizing enzyme known to bear variations that affect the metabolism of 90% of commercially available drugs. Other arrays such as GeneChip HuSNP (Affymetrix, Inc, Santa Clara, CA) enables whole-genome typing by simultaneously tracking nearly 1,500 genetic variations.

Challenges

As with all novel technologies, there are technical issues associated with the use of genomics and proteomics technologies that can complicate the interpretation of data. A recent review addressed the common technical problems associated with DNA microarrays (41). However, compared to proteomic technologies, these technical issues are much less substantive mainly because of the chemical homogeneity of RNA. Although considerable improvements in 2DGE technologies have occurred in recent years, limitations remain (21, 26, 54). Unlike the ordered DNA microarrays, 2DGE does not produce a comprehensive display of all expressed proteins. Highly hydrophobic proteins, such as integral membrane proteins, remain poorly soluble even in specialized detergent cocktails and fail to resolve during the isoelectric focusing step. Other classes of proteins such as those with extremes in isoelectric point and molecular mass, and those present in low copy number are underrepresented on a 2D gel. These limitations arise from the wide protein dynamic range and tremendous chemical heterogeneity. While some of these problems can be at least partially mitigated by preenrichment or prefractionation steps, using multiple narrow pH range gels, and applying different solubilisation mixtures for hydrophobic proteins, any future developments in proteomics must consider these key issues. For these reasons, gel-free proteomics technologies such as isotope-coded affinity tagging (ICAT) and multidimensional liquid chromatography online coupled to MS/MS techniques for quantitative analysis of complex protein mixtures (2, 31, 46), and microarray-based approaches to protein detection (20, 50, 82) are currently emerging as alternative and complementary techniques to 2DGE. These analytical tools, which are amenable to automization and allow low sample consumption and accurate protein identification, hold great promise for proteomic research in drug development, especially for biomarker profiling in body fluids.

The application of genome-wide expression technologies in toxicology is based on the underlying assumption that there are no toxicologically relevant outcomes in vitro or in vivo, with the possible exception of rapid necrosis, that do not require differential gene expression. In toxicology, the 'gene targets' to describe the adverse effects may constitute a large number of complex pharmacological, physiological and

biochemical processes, most of them interacting with each other, and related to a multitude of toxicological endpoints. Numerous processes not necessarily involved in the toxic process, but are secondary or indirect consequences are also involved. It is critical that the right conclusions are drawn from the multitude of observed relevant and irrelevant changes. The challenge encountered when such approaches are applied to toxicological investigations is, therefore, to separate those alterations associated with toxicity from changes that are adaptive, beneficial, and/or unrelated to the mechanism of toxicity and development of lesions. Moreover, how to know when and how much of a change on the mRNA or protein level does have any biological relevance. Clearly, the possibility of toxicogenomics giving an early alert or indication to potential adverse effects at low levels of exposure, while powerful, raises many issues in the context of interpreting differential gene expression data with respect to risk assessment. The difficulties in assessing the relevance of this data must, therefore, not be underestimated. Consequently, expression data for a given toxic response needs to be put into perspective with other (biochemical, physiological, clinical, histopathology) data and integrated expert evaluation is essential (9). Adequate study designs together with good sample quality can significantly aid in the interpretation of differential gene expression results, such as in deciphering which genes are adaptive, beneficial or toxicologically relevant. The use of different doses and time points and organ selection can generate comprehensive data to a drug response that can aid the interpreter in drawing correct conclusions.

Another consideration important to interpretation of gene expression data is to include in the studies reference compounds. This is especially important when new chemical entities are being examined. The number of animals required for each time point or dose is also an important consideration for a given study. In particular, the degree of interanimal variation is a major parameter to consider. From our experience 4 to 6 animals per group is usually adequate to derive statistically and biologically significant data that overcomes interanimal variability. Additional study designs can be envisioned in which genomic tools are first used to screen a multitude of organs, after which, interesting findings are followed up with proteomics on the selected organ(s). Focused technologies, including RT-PCR, in situ hybridization, immunohistochemistry or immunoassays, are then critical to confirm the importance of differentially expressed mRNA or proteins and to make the link with histopathology. Such an approach can verify the relevance and importance of the expression data. Furthermore, the potential of genome-wide expression technologies to generate enormous amounts of data, may engender additional problems in the analysis of data. Further advancements in bioinformatics are necessary to data mine the vast amount of genomics and proteomics data, and convert it into useful comprehensible information.

Much debate has taken place concerning the suitability of expression profiling technologies to quantitative risk assessment studies in toxicology. Studies in toxicology are traditionally conducted at pharmacological and toxic doses that are clearly toxic in order to minimise the chance of missing a toxic effect. Genome-wide mRNA and protein expression analysis has the potential to provide critical information regarding responses at dose ranges below those required for the induction of a toxic endpoint to doses that induce a toxic endpoint. Gene expression data obtained across a range of doses can provide a quantitative measurement of drug response that can greatly facilitate data interpretation and extrapolation. Theoretically, this can aid in the identification of sensitive biomarkers that indicate that a toxicity is being approached, before it actually becomes manifest. Currently, however, such expression data needs to be supported by validated and accepted regulatory assays before it can be used for quantitative risk assessment.

Toxicogenomics is an emerging discipline with data that is not validated. Preclinical development studies are highly regulated by government guidelines, and as toxicogenomics is being integrated into drug development by most, if not all, pharmaceutical research companies, regulatory issues are being raised by both industry and regulatory agency concerning the compatibility of regulatory evaluations with the level of validation of the technology. The FDA continues to encourage the industry to submit their toxicogenomics data in their applications for drug investigations and approval. However, one concern is that information from this evolving technology would hinder drug development and put drugs at risk. Regulatory agencies, however, are willing to provide 'safe harbour' until more is known about the validation and interpretations of what toxicogenomics can and cannot predict in terms of adverse events, toxicity, and efficacy. This concept would safeguard toxicogenomics data from regulatory decision-making, and concurrently facilitate further development of the field toward the goal of increasing the utility of toxicogenomics data for risk assessment. Industry and agency, however, still need to define how data submitted under 'safe harbour' concept would impact clinical study designs that are for registration. In the future, with greater understanding of the intricate nature of molecular events and their physiological consequences, together with the development of extensive knowledge data bases it can be envisaged that gene expression profiling can be realized as a validated tool in drug development.

Conclusion

Toxicogenomics is a new scientific discipline describing the combination of a systematic and comprehensive study of gene expression in response to a drug treatment in a biological system. High expectations are set on this new discipline to fundamentally change the process of drug development, especially in toxicity assessment. The use of expression profiling technologies to mechanistic and predictive toxicology, and biomarker discovery, will enable us to ask detailed questions and generate hypotheses. Ideally, such toxicology research must integrate itself into the discovery phase rather than following it to improve the quality of drug candidates and reduce the overall costs due to attrition during development. However, sound interpretation is required in this new area of data generation to ensure that toxicologically relevant changes are distinguished from those that are not. The regulatory aspects of toxicogenomics must also be seriously considered. Due to the exploratory and non-validated nature of current gene expression profiling studies, it is not currently

an intrinsic part of today's regulatory toxicology. Currently, the value of toxicogenomic data is in mechanistic studies, hypotheses generation, and as a source for providing candidates for validation as biomarkers of toxicity.

REFERENCES

1. Adkins JN, Varnum SM, Auberry KJ, Moore RJ, Angell NH, Smith RD, Springer DL, Pounds JG (2002). Toward a human blood serum proteome: Analysis by multidimensional separation coupled with mass spectrometry. *Mol Cell Proteomics* 1: 947–955.
2. Aebersold R, Mann M (2003). Mass spectrometry-based proteomics. *Nature* 422: 198–207.
3. Aicher L, Meier G, Norcross AJ, Jakubowski J, Varela MC, Cordier A, Steiner S (1997). Decrease in kidney calbindin-D 28kDa as a possible mechanism mediating cyclospori. *Biochem Pharmacol* 53: 723–731.
4. Aicher L, Wahl D, Arce A, Grenet O, Steiner S (1998). New insights into cyclosporine A nephrotoxicity by proteome analysis. *Electrophoresis* 19: 1998–2003.
5. Anderson L, Seilhamer J (1997). A comparison of selected mRNA and protein abundances in human liver. *Electrophoresis* 18: 533–537.
6. Anderson NL, Anderson NG (2002). The human plasma proteome: History, character, and diagnostic prospects. *Mol Cell Proteomics* 1: 845–867.
7. Bandara LR, Kennedy S (2002). Toxicoproteomics—a new preclinical tool. *Drug Discov Today* 7: 411–418.
8. Bartosiewicz MJ, Jenkins D, Penn S, Emery J, Buckpitt A (2001). Unique gene expression patterns in liver and kidney associated with exposure to chemical toxicants. *J Pharmacol Exp Ther* 297: 895–905.
9. Boorman GA, Anderson SP, Casey WM, Brown RH, Crosby LM, Gottschalk K, Easton M, Ni H, Morgan KT (2002). Toxicogenomics, drug discovery, and the pathologist. *Toxicol Pathol* 30: 15–27.
10. Brazma A, Vilo J (2000). Gene expression data analysis. *FEBS Lett* 480: 17–24.
11. Brookes AJ (1999). The essence of SNPs. *Gene* 234: 177–186.
12. Bulera SJ, Eddy SM, Ferguson E, Jatkoe TA, Reindel JF, Bleavins MR, De La Iglesia FA (2001). RNA expression in the early characterization of hepatotoxicants in Wistar rats by high-density DNA microarrays. *Hepatology* 33: 1239–1258.
13. Castle AL, Carver MP, Mendrick DL (2002). Toxicogenomics: A new revolution in drug safety. *Drug Discov Today* 7: 728–736.
14. Chen G, Gharib TG, Huang CC, Taylor JM, Misek DE, Kardia SL, Giordano TJ, Iannettoni MD, Orringer MB, Hanash SM, Beer DG (2002). Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* 1: 304–313.
15. Cottrell JS (1994). Protein identification by peptide mass fingerprinting. *Pept Res* 7: 115–124.
16. Davis RE, Staudt LM (2002). Molecular diagnosis of lymphoid malignancies by gene expression profiling. *Curr Opin Hematol* 9: 333–338.
17. de Longueville F, Surry D, Meneses-Lorente G, Bertholet V, Talbot V, Evrard S, Chandelier N, Pike A, Worboys P, Rasson JP, Le Bourdelles B, Remacle J (2002). Gene expression profiling of drug metabolism and toxicology markers using a low-density DNA microarray. *Biochem Pharmacol* 64: 137–149.
18. DePrimo SE, Wong LM, Khatry DB, Nicholas SL, Manning WC, Smolich BD, O'Farrell AM, Cherrington JM (2003). Expression profiling of blood samples from an SU5416 Phase III metastatic colorectal cancer clinical trial: A novel strategy for biomarker identification. *BMC Cancer* 3: 1–12.
19. DeRisi JL, Iyer VR, Brown PO (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278: 680–686.
20. Eickhoff H, Konthur Z, Lueking A, Lehrach H, Walter G, Nordhoff E, Nyarsik L, Bussow K (2002). Protein array technology: The tool to bridge genomics and proteomics. *Adv Biochem Eng Biotechnol* 77: 103–112.
21. Fey SJ, Larsen PM (2001). 2D or not 2D. Two-dimensional gel electrophoresis. *Curr Opin Chem Biol* 5: 26–33.
22. Frantz GD, Pham TQ, Peale FV, Jr, Hillan KJ (2001). Detection of novel gene expression in paraffin-embedded tissues by isotopic in situ hybridization in tissue microarrays. *J Pathol* 195: 87–96.
23. Ganguly NK, Bano R, Seth SD (2001). Human genome project: Pharmacogenomics and drug development. *Indian J Exp Biol* 39: 955–961.
24. Gevaert K, Vandekerckhove J (2000). Protein identification methods in proteomics. *Electrophoresis* 21: 1145–1154.
25. Godovac-Zimmermann J, Brown LR (2001). Perspectives for mass spectrometry and functional proteomics. *Mass Spectrom Rev* 20: 1–57.
26. Gorg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21: 1037–1053.
27. Gras R, Muller M, Gasteiger E, Gay S, Binz PA, Bienvenut W, Hoogland C, Sanchez JC, Bairoch A, Hochstrasser DF, Appel RD (1999). Improving protein identification from peptide mass fingerprinting through a parameterized multi-level scoring algorithm and an optimized peak detection. *Electrophoresis* 20: 3535–3550.
28. Grenet O, Varela MC, Staedtler F, Steiner S (1998). The cyclosporine A-induced decrease in rat renal calbindin-D28kDa protein as a consequence of a decrease in its mRNA. *Biochem Pharmacol* 55: 1131–1133.
29. Griffin TJ, Gygi SP, Rist B, Aebersold R, Loboda A, Jilkine A, Ens W, Standing KG (2001). Quantitative proteomic analysis using a MALDI quadrupole time-of-flight mass spectrometer. *Anal Chem* 73: 978–986.
30. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17: 994–999.
31. Gygi SP, Rist B, Griffin TJ, Eng J, Aebersold R (2002). Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags. *J Proteome Res* 1: 47–54.
32. Gygi SP, Rochon Y, Franza BR, Aebersold R (1999). Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19: 1720–1730.
33. Hamadeh HK, Bushel P, Paules R, Afshari CA (2001). Discovery in toxicology: Mediation by gene expression array technology. *J Biochem Mol Toxicol* 15: 231–242.
34. Hamadeh HK, Bushel PR, Jayadev S, DiSorbo O, Bennett L, Li L, Tennant R, Stoll R, Barrett JC, Paules RS, Blanchard K, Afshari CA (2002). Prediction of compound signature using high density gene expression profiling. *Toxicol Sci* 67: 232–240.
35. Hamadeh HK, Bushel PR, Jayadev S, Martin K, DiSorbo O, Sieber S, Bennett L, Tennant R, Stoll R, Barrett JC, Blanchard K, Paules RS, Afshari CA (2002). Gene expression analysis reveals chemical-specific profiles. *Toxicol Sci* 67: 219–231.
36. Heid CA, Stevens J, Livak KJ, Williams PM (1996). Real time quantitative PCR. *Genome Res* 6: 986–994.
37. Heijne WH, Stierum RH, Slijper M, van Bladeren PJ, van Ommen B (2003). Toxicogenomics of bromobenzene hepatotoxicity: A combined transcriptomics and proteomics approach. *Biochem Pharmacol* 65: 857–875.
38. Kennedy GC (2000). The impact of genomics on therapeutic drug development. *EXS* 89: 1–10.
39. Kennedy S (2002). The role of proteomics in toxicology: Identification of biomarkers of toxicity by protein expression analysis. *Biomarkers* 7: 269–290.
40. Lakkis MM, DeCristofaro MF, Ahr HJ, Mansfield TA (2002). Application of toxicogenomics to drug development. *Expert Rev Mol Diagn* 2: 337–345.
41. Li X, Gu W, Mohan S, Baylink DJ (2002). DNA microarrays: Their use and misuse. *Microcirculation* 9: 13–22.
42. Lian Z, Kluger Y, Greenbaum DS, Tuck D, Gerstein M, Berliner N, Weissman SM, Newburger PE (2002). Genomic and proteomic analysis of the myeloid differentiation program: Global analysis of gene expression during induced differentiation in the MPRO cell line. *Blood* 100: 3209–3220.
43. Liang P, Averboukh L, Pardee AB (1993). Distribution and cloning of eukaryotic mRNAs by means of differential display: Refinements and optimization. *Nucleic Acids Res* 21: 3269–3275.
44. Liang P, Pardee AB (1992). Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257: 967–971.
45. Liang P, Pardee AB (1998). Differential display. A general protocol. *Mol Biotechnol* 10: 261–267.

46. Liu H, Lin D, Yates JR, III (2002). Multidimensional separations for protein/peptide analysis in the post-genomic era. *Biotechniques* 32: 898, 900, 898–902.
47. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H, Brown EL (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol* 14: 1675–1680.
48. Money T, Reader S, Qu LJ, Dunford RP, Moore G (1996). AFLP-based mRNA fingerprinting. *Nucleic Acids Res* 24: 2616–2617.
49. Morikawa S, Takabe W, Mataki C, Kanke T, Itoh T, Wada Y, Izumi A, Saito Y, Hamakubo T, Kodama T (2002). The effect of statins on mRNA levels of genes related to inflammation, coagulation, and vascular constriction in HUVEC. Human umbilical vein endothelial cells. *J Atheroscler Thromb* 9: 178–183.
50. Mouradian S (2002). Lab-on-a-chip: Applications in proteomics. *Curr Opin Chem Biol* 6: 51–56.
51. Mukherjee D, Topol EJ (2002). Pharmacogenomics in cardiovascular diseases. *Prog Cardiovasc Dis* 44: 479–498.
52. Nuwaysir EF, Bittner M, Trent J, Barrett JC, Afshari CA (1999). Microarrays and toxicology: The advent of toxicogenomics. *Mol Carcinog* 24: 153–159.
53. Patterson SD, Aebersold RH (2003). Proteomics: The first decade and beyond. *Nat Genet* 33(Suppl): 311–323.
54. Patton WF (2002). Detection technologies in proteome analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* 771: 3–31.
55. Pease AC, Solas D, Sullivan EJ, Cronin MT, Holmes CP, Fodor SP (1994). Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc Natl Acad Sci USA* 91: 5022–5026.
56. Pennie WD (2000). Use of cDNA microarrays to probe and understand the toxicological consequences of altered gene expression. *Toxicol Lett* 112–113: 473–477.
57. Pennie WD (2002). Custom cDNA microarrays; Technologies and applications. *Toxicology* 181–182: 551–554.
58. Pennie WD, Woodyatt NJ, Aldridge TC, Orphanides G (2001). Application of genomics to the definition of the molecular basis for toxicity. *Toxicol Lett* 120: 353–358.
59. Rai AJ, Zhang Z, Rosenzweig J, Shih I, Pham T, Fung ET, Sokoll LJ, Chan DW (2002). Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med* 126: 1518–1526.
60. Rappersberger K, Komar M, Ebelin ME, Scott G, Burtin P, Greig G, Kehren J, Chibout SD, Cordier A, Holter W, Richter L, Oberbauer R, Stuetz A, Wolff K (2002). Pimecrolimus identifies a common genomic anti-inflammatory profile, is clinically highly effective in psoriasis and is well tolerated. *J Invest Dermatol* 119: 876–887.
61. Reynolds KJ, Yao X, Fenselau C (2002). Proteolytic 18O labeling for comparative proteomics: Evaluation of endoprotease Glu-C as the catalytic agent. *J Proteome Res* 1: 27–33.
62. Ring HZ, Kroetz DL (2002). Candidate gene approach for pharmacogenetic studies. *Pharmacogenomics* 3: 47–56.
63. Rockett JC (2001). Chipping away at the mystery of drug responses. *Pharmacogenomics J* 1: 161–163.
64. 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.
65. Schena M, Shalon D, Davis RW, Brown PO (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270: 467–470.
66. Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW (1996). Parallel human genome analysis: Microarray-based expression monitoring of 1,000 genes. *Proc Natl Acad Sci USA* 93: 10614–10619.
67. Schneeberger C, Speiser P, Kury F, Zeillinger R (1995). Quantitative detection of reverse transcriptase-PCR products by means of a novel and sensitive DNA stain. *PCR Methods Appl* 4: 234–238.
68. Schrader M, Schulz-Knappe P (2001). Peptidomics technologies for human body fluids. *Trends Biotechnol* 19: S55–S60.
69. Shalon D, Smith SJ, Brown PO (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res* 6: 639–645.
70. Shevchenko A, Loboda A, Shevchenko A, Ens W, Standing KG (2000). MALDI quadrupole time-of-flight mass spectrometry: A powerful tool for proteomic research. *Anal Chem* 72: 2132–2141.
71. Singh-Gasson S, Green RD, Yue Y, Nelson C, Blattner F, Sussman MR, Cerrina F (1999). Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat Biotechnol* 17: 974–978.
72. Sive HL, St John T (1988). A simple subtractive hybridization technique employing photoactivatable biotin and phenol extraction. *Nucleic Acids Res* 16: 10937.
73. Spahr CS, Davis MT, McGinley MD, Robinson JH, Bures EJ, Beierle J, Mort J, Courchesne PL, Chen K, Wahl RC, Yu W, Luethy R, Patterson SD (2001). Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. I. Profiling an unfractionated tryptic digest. *Proteomics* 1: 93–107.
74. Steiner S, Aicher L, Raymackers J, Meheus L, Esquer-Blasco R, Anderson NL, Cordier A (1996). Cyclosporine A decreases the protein level of the calcium-binding protein calbindin-D 28kDa in rat kidney. *Biochem Pharmacol* 51: 253–258.
75. Uetrecht JP (2000). Is it possible to more accurately predict which drug candidates will cause idiosyncratic drug reactions? *Curr Drug Metab* 1: 133–141.
76. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995). Serial analysis of gene expression. *Science* 270: 484–487.
77. Wang Z, Brown DD (1991). A gene expression screen. *Proc Natl Acad Sci USA* 88: 11505–11509.
78. 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.
79. Waring JF, Jolly RA, Ciurlionis R, Lum PY, Praestgaard JT, Morfitt DC, Buratto B, Roberts C, Schadt E, Ulrich RG (2001). Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicol Appl Pharmacol* 175: 28–42.
80. Waring JF, Ulrich RG (2000). The impact of genomics-based technologies on drug safety evaluation. *Annu Rev Pharmacol Toxicol* 40: 335–352.
81. Watanabe H, Suzuki A, Mizutani T, Khono S, Lubahn DB, Handa H, Iguchi T (2002). Genome-wide analysis of changes in early gene expression induced by oestrogen. *Genes Cells* 7: 497–507.
82. Wilson DS, Nock S (2002). Functional protein microarrays. *Curr Opin Chem Biol* 6: 81–85.
83. Wodicka L, Dong H, Mittmann M, Ho MH, Lockhart DJ (1997). Genome-wide expression monitoring in Saccharomyces cerevisiae. *Nat Biotechnol* 15: 1359–1367.
84. Yap G (2002). Affymetrix, Inc. *Pharmacogenomics* 3: 709–711.
85. Yergey AL, Coorssen JR, Backlund PS, Jr, Blank PS, Humphrey GA, Zimmerberg J, Campbell JM, Vestal ML (2002). De novo sequencing of peptides using MALDI/TOF-TOF. *J Am Soc Mass Spectrom* 13: 784–791.