Journal of Toxicology Research, Vol. 1, Issue 1, 2010, PP-01-11

Functional Toxicogenomics: Technological Aspects

Gomase V.S.[1], Kapoor R.A.[2], Ladak S.S.[2]
[1]School of Technology, S.R.T.M. University, Sub-Centre, Latur, 413512, India
[2]Padmashree Dr. D.Y. Patil University, Navi Mumbai, 400614, India

Abstract- Toxicogenomics is a field of science that deals with the collection, interpretation, and storage of information about gene and protein activity within particular cell or tissue of an organism in response to toxic substances. Toxicogenomics combines toxicology with genomics or other high throughput molecular profiling technologies such as transcriptomics, proteomics and metabolomics. Toxicogenomics endeavors to elucidate molecular mechanisms evolved in the expression of toxicity, and to derive molecular expression patterns i.e., molecular biomarkers that predict toxicity or the genetic susceptibility to it.

Introduction

Toxicogenomics is a rapidly developing field that promises to aid scientists in understanding the molecular and cellular effects of chemicals in biological systems. This field encompasses global assessment of biological effects using technologies such as DNA microarrays or high throughput NMR and protein expression analysis, genomic, proteomic, metabonomic that may extend our understanding of toxicology and highlights the toxicity studies.

This broad definition is supported by the United States Environmental Protection Agency stating that "the term "genomics" encompasses a broader scope of scientific inquiry and associated technologies than when genomics was initially considered. A genome is the sum total of all an individual organism's genes. Thus, genomics is the study of all the genes of a cell, or tissue, at the DNA (genotype), RNA (transcriptome), or protein (proteome) levels. Genomics methodologies are expected to provide valuable insights for evaluating how environmental stressors affect cellular/tissue function and how changes in gene expression may relate to adverse effects. However, the relationships between changes in gene expression and adverse effects are unclear at this time and may likely be difficult to elucidate.

In pharmaceutical research toxicogenomics is more narrowly defined as the study of the structure and function of the genome as it responds to adverse xenobiotic exposure. It is the toxicological subdiscipline of pharmacogenomics, which is broadly defined as the study of inter-individual variations in whole-genome or candidate gene single-nucleotide polymorphism maps, haplotype markers, and alterations in gene expression that might correlate with drug responses (Lesko and Woodcock 2004, Lesko et al. 2003). Though the term toxicogenomics first appeared in the literature in 1999 (Nuwaysir et al.) it was already in common use within the pharmaceutical industry as its origin was driven by marketing strategies from vendor companies. The term is still not universal accepted, and others have offered alternative terms such as chemogenomics to describe essentially the same area (Fielden et al., 2005).

The nature and complexity of the data (in volume and variability) demands highly developed processes for of automated handling and storage. The analysis usually involves a wide array of bioinformatics and statistics., regularly involving classification approaches.

In pharmaceutical Drug discovery and development toxicogenomics is used to study adverse, i.e. toxic, effects, of pharmaceutical drugs in defined model systems in order to draw conclusions on the toxic risk to patients or the environment. Both the EPA and the U.S. Food and Drug Administration currently preclude basing regulatory decision making on genomics data alone. However, they do encourage the voluntary submission of well-documented, quality genomics data. Both agencies are considering the use of submitted data on a case-by-case basis for assessment purposes (e.g., to help elucidate mechanism of action or contribute to a weight-of-evidence approach) or for populating relevant comparative databases by

encouraging parallel submissions of genomics data and traditional toxicologic test results.

History and Background

The field of toxicology is defined as the study of stressors and their adverse effects. One sub-discipline deals with hazard identification, mechanistic toxicology, and risk assessment. Increased understanding of the mechanism
of action of chemicals being assayed will improve the efficiency of these tasks. However, the derivation of mechanistic knowledge traditionally evolves from studying a few genes at a time in order to implicate their function in mediation of toxicant effects. Undoubtedly, this process has to be accelerated to monitor and discern the effects of the thousands of new compounds developed by the chemical and pharmaceutical industries. There is a need for a screening method that can offer some insight into the potential adverse outcome(s) of new drugs allowing the intelligent advancement of compounds into late stages of safety evaluation.

Mechanism of Toxicity

The study and use of toxicogenomics approaches aids our understanding of the mechanisms of toxicity. Bulera and coworkers identified several groups of genes reflective of mechanisms of toxicity and related to a hepatotoxic outcome following treatment. An example of the advantage of using a toxicogenomics approach to understand mechanisms of chemical toxicity was the observation of liver tumor promoters microcystin-LR and phenobarbital, induced a parallel set of genes. Based on this information the authors speculated that liver tumor promotion by both compounds may occur by similar mechanisms. These observations derived through the application of microarrays to toxicology to understand mechanisms and our ability to identify compounds with similar mechanisms of toxicity. However, it would have been advantageous to utilize gene expression data to map relevant pathways depicting mechanism(s) associated with the hepatotoxicity of each compound. Collectively, in the future, researchers may attempt to build “transcriptome” or “effector maps” to aid to visualization of pathway activation.Huang and coworkers utilized cDNA microarrays to investigate gene expression patterns of cisplatin-induced nephrotoxicity. In these studies, rats were treated daily for 1 to 7 days with cisplatin at a dose that resulted in necrosis of the renal proximal tubular epithelial cells but no hepatotoxicity at day 7. Gene expression patterns for transplatin, an inactive isomer, was examined and revealed little gene expression change in the kidney, consistent with the lack of nephrotoxicity of the compound. Cisplatin-induced gene expression alterations were reflective of the histopathological changes in the kidney i.e. gene related to cellular remodeling, apoptosis, and alteration of calcium homeostasis.

Technological Aspects

Protein Expression

Gene expression alone is insufficient to understand the toxicant and the disease they cause. Abnormalities in protein production or function are expected in response to toxicant exposure and the onset of disease states. To understand the complete mechanism of toxicant action, it is necessary to identify the protein alterations associated with that exposure and to understand how these changes affect protein/cellular function. Unlike classical genomic approaches that discover genes related to toxicant induced disease, proteomics can aid to characterize the disease process directly by capturing proteins that participate in the disease. The lack of a direct functional correlation between gene transcripts and their corresponding proteins necessitates the use of proteomics as a tool in toxicology. Proteomics is the systematic analysis of expressed proteins in tissues, by isolation, separation, identification and functional characterization of proteins in a cell, tissue, or organism. Proteomics, under the umbrella of toxicogenomics, involves the comprehensive functional annotation and validation of proteins in response to toxicant exposure. Understanding the functional characteristics of proteins and their activity requires a determination of cellular localization and quantitation, tissue distribution, post-translational modification state, domain modules and their effect on

protein interactions, protein complexes, ligand binding sites and structural representation. Currently, the most commonly used technologies for proteomics research are2-dimensional (2-D) gel electrophoresis for protein separation followed by mass spectrometry analysis of proteins of interest. Analytical protein characterization with multidimensional liquid chromatography/mass spectrometry improves the throughput and reliability of peptide identity. Matrix Assisted Laser Desorption Mass Spectrometry (MALDI-MS) has become a widely used method for determination of various biological molecules including peptides. Other technologies such as Surface-Enhanced Laser Desorption/Ionization (SELDI) and antibody arrays are also proving to be useful. Various projects are conducted investigate biochemical changes and identify biomarkers associated with acute renal injury following a single dose of puromycin aminonucleoside to Sprague Dawley rats using a combination of 2-D PAGE, reverse phase HPLC, mass spectrometry, amino acid analysis and 1H-NMR spectroscopy of urine as well as routine plasma clinical chemistry and tissue histopathology. The 2-D PAGE of urine showed patterns of protein change according to the limited profiles for glomerular toxicity derived by use of other techniques and allowed a more detailed understanding of the nature and progression of the proteinuria associated with glomerular toxicity. The 2-D PAGE approach taken by the various investigators, coupled with computational analysis of the accompanying data gradually on the collected samples, lead to the detection of proteinuria at a considerably earlier time point than has typically been reported following puromycin aminonucleoside exposure, thus potentially defining relatively early biomarkers which are superior to the traditional gross urinary protein determination procedure. A serious limitation of proteomic analysis using 2-D gel electrophoresis is the sensitivity of detection. Analysis of low abundance proteins by 2-D electrophoresis is challenging due to the presence of high abundant proteins such as albumin, immunoglobulin heavy and light chains, transferrin, and haptoglobin in the sera or actin, tubulin, and other structural proteins when analyzing tissue. Selective removal of these proteins from protein samples via column-based immunoaffinity procedures allows for more sample to be loaded on gels thereby facilitating visualization of low abundant proteins that would otherwise be obscured by more abundant ones.

Metabolite Analysis by NMR

Genomic and proteomic methods do not offer the information needed to gain understanding of the resulting output function in a living system. Neither approach addresses the dynamic metabolic status of the whole animal. The metabonomic approach is based on the premise that toxicant-induced pathological or physiological alterations result in changes in relative concentrations of endogenous biochemicals. Metabolites in body fluids such
as urine, blood, or cerebrospinal fluid (CSF), are in dynamic equilibrium with those inside cells and tissues, thus toxicant induced cellular abnormalities in tissues should be reflected in altered biofluid compositions. An advantage of measuring changes in body fluids is that these samples are much more readily available from human subjects. High resolution NMR spectroscopy (1H NMR) has been used in a high-throughput fashion to simultaneously detect various biological chemicals in urine, bile, blood plasma, milk, saliva, sweat, gastric juice, seminal, amniotic, synovial and cerebrospinal fluids. In addition, intact tissue and cellular suspensions have also been successfully analyzed for metabolite content using magic-angle-spinning 1H NMR spectroscopy.

Toxicogenomics and Cancer

DNA encodes for RNA, RNA encodes for proteins, and these proteins regulate all the processes. Most proteins are enzymes which have a catalytic function in cell, whereby substances are converted metabolised into another substance. This may also imply that another protein is modified; this modification may lead to activation or inactivation of protein. Controlling the activity of genes, and with that of proteins such as enzymes and therefore also the metabolites which reside in cell, is very complex and takes place at several steps: 1) By the cell itself, for example apoptosis, 2) By other cells in the

body, for example in the immune system, where one immune cell may influence the functioning of another and as a result of that, becomes a regulator of transcription and raises the activity of a many genes drastically, thereby eventually leading to the development of cancer. In the last decade, methods have become available, which enable to simultaneously analyse the activity of all genes, or the quantities of all proteins or of all cellular metabolites. For measuring the expression of all genes, DNA microarrays have been developed, which eventually results in genome-wide gene expression profiles. Changes in these gene expression profiles reflect modifications in the activities of all genes. For proteins and metabolites other methods have been developed, but these will not be discussed here. Since exposure of humans and animals to chemical substances leads to changed gene activities in the body, analysing these modulations may therefore have very interesting applications for toxicology. Toxicology namely investigates whether chemical substances have hazardous impact on humans or animals, and importantly also, to retrieve the mechanisms of action. Establishing these so-called genomic risk profiles has a central role in toxicogenomics immune cell, and 3) Influences by outside the body, like as a result of exposure to a toxic substance. For example of the latter is dioxin, a carcinogenic substance which binds to a sensor in the cell .

Applications of Toxicogenomics

There are two main applications for a toxicogenomic approach, comparative and functional.

Comparative Toxicogenomics (Predictive Toxicogenomics)

Comparative genomic, proteomic, or metabonomic studies measure the quantity and types of genes, proteins, and metabolites respectively that are present in normal and toxicant-exposed cells, tissues, or biofluids. This approach is useful in defining the composition of the assayed samples in terms of genetic, roteomic or metabolic variables. Thus a biological sample derived from toxicant, or sham treated animals can be regarded as an n-dimensional vector in gene expression space with genes as variables along each dimension. The same analogy can be applied for protein expression or NMR analysis data thereby providing ndimensional fingerprints or profiles of the biological sample under investigation. Thus, this aspect of toxicogenomics deals with automated pattern recognition analysis aimed at studying trends in data sets rather than probing the individual genes for mechanistic information. The need for pattern recognition tools is mandated by the volume and complexity of data generated by genomic, proteomic and metabonomic tools, and human intervention, in required repetitive computation, is kept to a minimum. Automatic toxicity classification methods are very desirable and prediction models are well suited for this task. The data profiles reflect the pharmacological or toxicological effects, such as disease outcome, of the toxicant being utilized. The underlying goal is that a sample from an animal exposed to unknown chemical, or inducing a certain pathological effects, can then be compared to a database of profiles corresponding to exposure conditions with well-characterized chemicals, or to well defined pathological endpoints, to predict some properties regarding the studied sample. These predictions, as we view them, fall into two major categories, namely, classification of samples based on the class of compound to which animals were exposed, or classification of samples based on the histopathology and clinical chemistry that the treated animals displayed. Such data will allow insight into the gene, protein, or metabolite perturbations associated with pharmacologic effects of the toxic endpoints that ensue. If array data can be “phenotypically anchored” to conventional indices of toxicity, it will be possible to search for evidence of injury prior to its clinical or pathological manifestation. This approach could lead to the discovery of potential early biomarkers of toxic injury. “Supervised” predictive models have been used for many years in the financial sectors for evaluating future economic prospects of companies, and in geological institutes for predicting adverse weather outcomes using past or historical knowledge. They have also been utilized for predictions, using clinical and radiographic data, for the diagnosis of active pulmonary tuberculosis at the time of

presentation at a health-care facility that can be superior to physicians' opinion. Predictive modeling will undoubtedly revolutionize the field of toxicology by recognizing patterns and trends in high-density data, and forecasting gene-, protein-, or metabolic interactions relying on historical data from well studied compounds and their corresponding profiles. During the development of a predictive model, a number of issues must be considered. These include the representativeness of the variables to the entity being modeled and the quality of databases consulted. The National Center for Toxicogenomics (NCT), at the National Institute of Environmental Health Sciences, is building a database to store many data and observations related with the process of compound evaluation studies. Recording these parameters will greatly enhance the process of parameter selection in subsequent efforts such as predictive modeling or mechanism of action interpretation. Predictive modeling can be fragmented into a multistage process. The primary stage predictive modeling includes hypothesis development, organization and data collection. Secondary stage modeling includes initial model development and testing. Tertiary stage modeling includes continued application of the model, ongoing refinement, and validation. Ideally, tertiary stage modeling is a perpetual process whereby lessons learned from previous model applications are incorporated into new and future applications maintaining or increasing the predictive robustness of the model.

1) Data Collection

The development of the primary stage of a predictive model involves activities such as data collection strategies based on proposed hypothesis. Data can be generated from in vivo or in vitro experiments, depending on the suitability of the biological system for studying effects of the targeted compound. In the case of in vivo studies, hypothesis must be generated regarding the compounds and endpoint effects so that other measures, such as pathology, serum markers, and carcinogenicity potential, are made and can contribute to the ensuing model development. Data on animal weight fluctuations, serum markers, pathological alterations, and mortality rates corresponding to a chemical exposure study should be documented and be the primary source of such information for the constructed predictive model. Pertinent data and analytically useful variables gathered from other sources can be evaluated and incorporated into the model. These data are important in developing a theoretical framework in which to interpret the results of the predictive model as well as to provide a guide for the data to be collected.

2) Model Development

The next step in the predictive model construction involves a deductive phase that incorporates collected data into the second stage of the model. The degree of correlation between gene-, protein- or metabolite-related profiles of different compounds or different toxicological/pathological outcomes and the accompanying variables can be measured and ranked. Computational and statistical approaches would be applied to the data set to glean relationships and dissimilarities among the variables studied. Neural networks, which have been used in models predicting health status of HIV/AIDS patients, can be trained with a set of available profiles from previously studied compounds or pathophysiological states. This allows the automation of all the actions aimed at searching the interrelationships and producing predictions regarding unknown or new profiles. Every effect is characterized by many parameters describing its gene expression pattern. Thus, a pattern may be represented by a vector in space whose components could represent various parameters that drive the decision of classification. Dimensionality of this space is the number of vector components or parameter involved and is based on the analysis of multiple parameters that can correlate similar expression profiles. A multitude of available algorithms satisfactorily cluster objects in 3- or dimensional space based on computational approaches (ex. PCA). We can then construct similarity zones around various preset chemical or adverse endpoint nodes. Such similarity zones would allow the classification, with a defined level of confidence, of the identity of unknown

samples which neighbor samples in the training data set. Thus possessing the map and information about the analyzed compounds, we can reliably judge the compounds with which we are less familiar. The initial predictive model can be tested using the data collected in the primary stage. Based upon the outcome, variables such as toxicant induced lesion severity or organ weight fluctuations can be introduced or removed from the process, or the weighting of the variables can be adjusted until the model is able to predict the highest percentage of chemicals possible. This highlights the need for the consulted database to contain enough parameters such as histopathological observations or clinical chemistry data that accompany an experimental design to facilitate this dynamic model optimization process. Developed models should ideally allow the distinction of gene expression profiles associated with outcome of pathology depending on the querying preferences of the user and the question being asked. Once this has been achieved, tertiary stage modeling may begin.

3) Utility

The use of genomic resources such as DNA microarrays in safety evaluation will facilitates "in silico" testing. In silico experimentation can define this relationship through rigorous computation and mining of high-density gene expression data. Developments in computer modeling and expert systems for the prediction of biological activity and toxicity will revolutionize the process of drug discovery and development, by reducing the need to use animals for the pre-screening of almost limitless numbers of potential drug candidates. It is not foreseeable that in the near future predictive models will take the place of actual testing. However, in the context of toxicogenomics, and with the increasing number of chemicals to be tested, better prioritization can be used to select the compounds for animal testing. The most promising efficacious compounds with the least probability of an adverse outcome would be selected for further development.

Functional Toxicogenomics

Functional toxicogenomics is the study of genes' and their products' activities on an organism. Gene and protein expression profiles are analyzed for information that might provide insight into specific mechanistic pathways. Mechanistic inference is complex when the sequence of events following toxicant exposure is viewed in both dose and time space. Gene and protein expression patterns can indeed be highly dependent on the toxicant concentrations furnished at the assessed tissue and the time of exposure to the agent. Expression patterns are only a snapshot in time and dose space. Thus, a comprehensive understanding of potential mechanisms of action of a compound requires establishing patterns at various combinations of time and dose. This will minimize the misinterpretation of temporary responses and allow the discernment of delayed alterations that could be related to adaptation events or may be representative of potential biomarkers of pathophysiological endpoints. Studies that target temporal expression of specific genes and protein in response to toxicant exposure will lead to a better understanding of the sequence of events in complex regulatory networks. Algorithms, such as selforganizing maps can categorize genes or proteins based on their expression pattern across a continuous time points. These analyses might suggest relationships in the expression of some genes or proteins depending on the concerted modulation of these variables. An area of study which is of great interest to toxicologists is the mechanistic understanding of toxicant induced pathological endpoints. The premise that uneasiness in gene, protein, or metabolite levels are reflective of adverse phenotypic effects of toxicants offers an opportunity to phenotypically anchor these uneasiness. This is quite challenging due to the fact that phenotypic effects often vary in the time-dose space of the studied agent and may have regional variations in the tissue. Furthermore, very few compounds exist that result in only one phenotypic alteration at a given coordinate in dose and time. Thus, objective assignment of measured variables to multiple phenotypic events is not possible under these circumstances. However study of multiple structurally and

pharmacologically unrelated agents that have similar pathological endpoints of interest, one could tease out gene, protein, or metabolite modulations that are in common between the studied compounds. Laser capture microdissection may also be used to capture regional variations such as zonal patterns of hepatotoxicity. This concept will allow the objective assignment of measurable variables to phenotypic observations that will supplement traditional pathology.

The stand-alone, gene and protein expression, or metabolite fluctuation analyses are not expected to produce decisive inferences on the role of genes or proteins in certain pathways or regulatory networks. However, these tools constitute powerful means to generate viable and testable hypotheses that can direct future endeavors on proving or disproving the involvement of genes, proteins, and metabolites in cellular processes. Ultimately, hypothesized mechanistic inferences have to be validated by the use of traditional molecular biology techniques that include the use of specific enzyme inhibitors and the examination of the effects of over expression or deletion of specific genes or proteins on the studied toxic endpoint or mechanism of compound action.

Future of Predictive Toxicology

From the rapid screening perspective, it is neither cost effective, nor is it practical to survey the abundance of all genes, proteins, or metabolites in a sample of interest. It would be prudent to conduct cheaper, more highthroughput measurements on variables that are of most interest in the toxicological evaluation process. Thus, this reductionist strategy mandates the selection of subsets of genes, proteins or metabolites that will yield useful information in regards to classification purposes such as hazard identification or risk assessment. The challenge is finding out what these minimal variables are and what data we need to achieve this knowledge. Election of these subsets by surveying existing toxicology literature is inefficient because the role of most genes or proteins in toxicological responses is poorly defined. Moreover, there exists a multitude of undiscovered or unknown genes (ESTs) that might ultimately be key players in toxicological processes.

The use of genes, proteins, or metabolites that are found to be most different between stressor induced-specific profiles, for efficient screening purposes

Discriminative potential of genes, proteins, or metabolites is inferred when comparing differences in the levels of these parameters across toxicant exposure scenarios. In the case of samples derived from animals treated with one of few chemicals; the levels of one gene, protein, or metabolite might be sufficient to distinguish samples based on the few classes of compounds used for the exposures. However, multiple parameters are needed to separate samples derived from exposures to a larger variety of chemical classes. Finding these discriminatory parameters requires the use of computational and mining algorithms that extract this knowledge from a database of chemical effects.

Commonality across animals revealed by cluster analysis of gene, protein, or metabolite levels would indicate a potential association between the altered parameters and the shared histopathological endpoint. Linear discriminant analysis (LDA) and single gene ANOVA can be used to test single parameters (ex. genes) for their ability to separate profiles corresponding to samples derived from different exposure conditions (ex. chemical identity, biological endpoint). Higher order analyses such as genetic algorithm/K-nearest neighbor (GA/KNN) are able to find a user defined number of parameters that would, as a set, highlight the most difference between biological samples based on the levels of genes, proteins, or metabolites. Once the profile of a parameter, or a set of parameters, is found to distinguish between samples in a data set, it can be used to interrogate the identity of unknown samples for screening purposes in a highthroughput fashion.

Conclusion

Toxicogenomic tools will inevitably improve the way data is extracted from classical toxicology studies. Ultimately, through the use of computational tools encompassed within the comparative branch of toxicogenomics, environmental hazards may be identified in a high-throughput and efficient fashion. These achievements will be facilitated through the development of gene,

protein, or metabolite markers whose levels can be monitored in samples derived from exposed populations. Compound analysis will also improve our understanding of toxicant induced adverse effects in biological systems by providing information about the basis of molecular pathways that are involved in response to expose compound. This knowledge will lead to a more informed and precise classification of compounds for their safety evaluation.

References

[1] Afshari, C.A., Nuwaysir, E.F., and Barrett, J.C. 1999. Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation. Cancer Res. 59: 4759-4760.

[2] Anderson, N.L., Matheson, A.D., and Steiner, S. 2000. Proteomics: applications in basic and applied biology. Curr. Opin. Biotechnol. 11: 408-412.

[3] Bonner, R.F., Emmert-Buck, M., Cole, K., Pohida, T.,Chuaqui, R., Goldstein, S., and Liotta, L.A. 1997. Laser capture microdissection: molecular analysis of tissue. Science. 278: 1481-1483.

[4] Borrebaeck, C.A., Ekstrom, S., Hager, A.C., Nilsson, J., Laurell, T., and Marko-Varga, G. 2001. Protein chips based on recombinant antibody fragments: a highly sensitive approach as detected by mass spectrometry. Biotechniques. 30: 1126-1130, 1132.

[5] Bulera, S.J., Eddy, S.M., Ferguson, E., Jatkoe, T.A., Reindel, J.F., Bleavins, M.R., and De La Iglesia, F.A. 2001. RNA expression in the early characterization of hepatotoxicants in wister rats by high-density DNA microarrays. Hepatology. 33: 1239-1258.

[6] Bundy, J.G., Osborn, D., Weeks, J.M., Lindon, J.C., and Nicholson, J.K. 2001. An NMR-based metabonomic approach to the investigation of coelomic fluid biochemistry in earthworms under toxic stress. FEBS Lett. 500: 31-35.

[7] Carroll, K., Ray, K., Helm, B., and Carey, E. 2000. Twodimensional electrophoresis reveals differential protein expression in high- and low-secreting variants of the rat basophilic leukaemia cell line. Electrophoresis. 21: 2476-2486.

[8] Corcoran, O., Spraul, M., Hofmann, M., Ismail, I.M., Lindon, J.C., and Nicholson, J.K. 1997. 750 MHz HPLC-NMR spectroscopic identification of rat microsomal metabolites of phenoxypyridines. J. Pharm. Biomed. Anal. 16: 481- 489.

[9] De Beer, R., Van den Boogaart, A., Cady, E., Graveron- Demilly, D., Knijn, A., Langenberger, K.W., Lindon, J.C., Ohlhoff, A., Serrai, H., and Wylezinska-Arridge, M. 1998. Absolute metabolite quantification by in vivo NMR spectroscopy: V. Multicentre quantitative data analysis trial on the overlapping background problem. Magn. Reson. Imaging. 16: 1127-1137.

[10] DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A., and Trent, J.M. 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer [see comments]. Nat. Genet. 14: 457-460.

[11] Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P., and Trent, J.M. 1999. Expression profiling using cDNA microarrays. Nat. Genet. 21: 10-14.

[12] Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. 1998. Cluster analysis and display of genome-wide expression patterns. Proc. Natl. Acad. Sci. USA. 95:14863-14868.

[13] El-Solh, A.A., Hsiao, C.B., Goodnough, S., Serghani, J., and Grant, B.J. 1999. Predicting active pulmonary tuberculosis

using an artificial neural network. Chest. 116: 968-973.

[14] Emmert-Buck, M.R., Bonner, R.F., Smith, P.D., Chuaqui, R.F., Zhuang, Z., Goldstein, S.R., Weiss, R.A., and Liotta, L.A. 1996. Laser capture microdissection. Science. 274: 998-1001.

[15] Farr, S.R.n.D. 1999. Concise review: gene expression applied to toxicology. Toxicol. Sci. 50: 1-9. Fend, F., Kremer, M., and Quintanilla-Martinez, L. 2000. Laser capture microdissection: methodical aspects and applications with emphasis on immuno-laser capture microdissection. Pathobiolo. 68: 209-214.

[16] Fountoulakis, M., Berndt, P., Boelsterli, U.A., Crameri, F., Winter, M., Albertini, S., and Suter, L. 2000. Two dimensional database of mouse liver proteins: changesin hepatic protein levels following treatment with acetaminophen or its nontoxic regioisomer 3-acetamidophenol. Electrophoresis. 21: 2148-2161.

[17] Foxall, P.J., Mellotte, G.J., Bending, M.R., Lindon, J.C., and Nicholson, J.K. 1993. NMR spectroscopy as a novel approach to the monitoring of renal transplant function. Kidney Int. 43: 234-245.

[18] Garrod, S., Humpfer, E., Spraul, M., Connor, S.C., Polley, S., Connelly, J., Lindon, J.C., Nicholson, J.K., and Holmes, E. 1999. High-resolution magic angle spinning 1H NMR spectroscopic studies on intact rat renal cortex and medulla. Magn. Reson. Med. 41: 1108-1118.

[19] Giacomini, M., Ruggiero, C., Maillard, M., Lillo, F.B., and Varnier, O.E. 1997. Application of neural networks to the follow-up of AIDS patients. Stud. Health Technol. Inform. 43: 386-390.

[20] Griffin, J.L., Walker, L.A., Shore, R.F., and Nicholson, J.K. 2001. Metabolic profiling of chronic cadmium exposure in the rat. Chem. Res. Toxicol. 14: 1428-1434.

[21] Haab, B.B., Dunham, M.J., and Brown, P.O. 2001. Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. Genome Biol. 2.

[22] Hamadeh, H., and Afshari, C.A. 2000. Gene chips and functional genomics. American Scientist. 88: 508-515.

[23] Hamadeh, H.K., Bushel, P., Paules, R.S., and Afshari, C.A. 2001. Discovery in toxicology: Mediation by gene expression array technology. J. Biochem. Mol. Toxicol. 15: 231-242.

[24] Hamadeh, H.K., Bushel, P., Tucker, C.J., Martin, K., Paules, R., and Afshari, C.A. 2002a. Detection of diluted gene expression alterations using cDNA microarrays. Biotechniques. In Press.

[25] Hamadeh, H.K., Bushel, P., Jayadev, S., DiSorbo, O., Bennett, L., Li, L., Tennant, R., Stoll, R., Barrett, J.C., Blanchard, K., Paules, R.S., and Afshari, C.A. 2002b. Prediction of compound signature using high density gene expression profiling. Toxicol. Sci. In press.

[26] Hamadeh, H.K., Bushel, P., Jayadev, S., Martin, K., DiSorbo, O., Sieber, S., Bennett, L., Tennant, R., Stoll, R., Barrett, J.C., Blanchard, K., Paules, R.S., and Afshari, C.A. 2002c. Gene expression analysis reveals chemicalspecific profiles. Toxicol. Sci. In press.

[27] Hamadeh, H.K., Nuwaysir, E.F., Paules, R.S., Barrett, J.C., and Afshari, C.A. 2002d. cDNA microarray technology: Merging toxicology and genomics. Comments Toxicol. In Press.

[28] Henry, C.M. 1999. DNA microarrays in toxicology. Anal. Chem. 71: 462A-464

[29] A. Holmes, E., Caddick, S., Lindon, J.C., Wilson, I.D.,Kryvawych, S., and Nicholson, J.K. 1995. 1H and 2H NMR spectroscopic studies on the metabolism and biochemical effects of 2-bromoethanamine in the rat. Biochem. Pharmacol. 49: 1349-1359.

[30] Holmes, E., Nicholson, J.K., and Tranter, G. 2001. Metabonomic characterization of genetic variations in toxicological and metabolic responses using probabilistic neural networks. Chem. Res. Toxicol. 14: 182-191.

[31] Hooker, T.M. 2001. The Brave New World of Toxicogenomics. Environmental Compliance Litigation. 16: 5.

[32] Huang, Q., Dunn, R.T., 2nd, Jayadev, S., DiSorbo, O., Pack, F.D., Farr, S.B., Stoll, R.E., and Blanchard, K.T. 2001. Assessment of cisplatin-induced nephrotoxicity by microarray technology. Toxicol. Sci. 63: 196-207.

[33] Huang, R. 2001. Detection of multiple proteins in an antibody-based protein microarray system. J. Immunol. Meth. 255: 1-13. Iannaccone, P.M. 2001. Toxicogenomics: “The call of the wild chip”. Environ. Health Perspectives. 109: A8-A11.

[34] Ioannidis, J.P., McQueen, P.G., Goedert, J.J., and Kaslow, R.A. 1998. Use of neural networks to model complex immunogenetic associations of disease: human leukocyte antigen impact on the progression of human immunodeficiency virus infection. Am. J. Epidemiol. 147:464-471.

[35] Johnson, R.A. and Wichern, D.W. 1998. Applied Multivariate Statistical Analysis. 4th ed. Prentice-Hall, Upper Saddle River,

[36] NJ. Jonic, S.,ankovic, T., Gajic, V., and Popovic, D. 1999. Three machine learning techniques for automatic determination of rules to control locomotion. IEEE Trans.Biomed. Eng, 46: 300-310.

[37] Kaji, H., Tsuji, T., Mawuenyega, K.G., Wakamiya, A., Taoka, M., and Isobe, T. 2000. Profiling of Caenorhabditis elegans proteins using two-dimensional gel electrophoresis and matrix assisted laser desorption/ionization-time of flightmass spectrometry. Electrophoresis. 21: 1755-1765.

[38] Kohonen, T. 1999. Comparison of SOM point densities based on different criteria. Neural Comput. 11: 2081-2095.

[39] Kreuzer, K.A., Lass, U., Bohn, A., Landt, O., and Schmidt, C.A. 1999. LightCycler technology for the quantitation of bcr/abl fusion transcripts. Cancer Res. 59: 3171-3174.

[40] Kuwata, H., Yip, T.T., Yip, C.L., Tomita, M., and Hutchens, T.W. 1998. Direct detection and quantitative determination of bovine lactoferricin and lactoferrin fragments in human gastric contents by affinity mass spectrometry. Adv. Exp. Med. Biol. 443: 23-32.

[41] Kwak, N.K., and Lee, C. 1997. A neural network application to classification of health status of HIV/AIDS patients. J. Med. Syst. 21: 87-97. Li, L., Darden, T.A., Weinberg, C.R., and Pedersen, L.G. Gene assessment and sample classification for gene expression data using a genetic algorithm/k-nearest neighbor method. Combinatorial Chemistry and High Throughput Screening, In press.

[42] Li, X., Mohan, S., Gu, W., Miyakoshi, N., and Baylink, D.J. 2000. Differential protein profile in the ear-punched tissue of regeneration and non-regeneration strains of mice: a novel approach to explore the

candidate genes for softtissue regeneration. Biochim. Biophys. Acta. 1524: 102- 109.

[43] Liang, X., Zheng, K., Qian, M.G., and Lubman, D.M. 1996. Determination of bacterial protein profiles by matrixassisted laser desorption/ionization mass spectrometry with high-performance liquid chromatography. RapidCommun. Mass Spectrom. 10: 1219-1226.

[44] Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H., and Brown, E.L. 1996. Expression monitoring by hybridization to high-density oligonucleotide An Overview of Toxicogenomics 55 arrays. Nat. Biotechnol. 14: 1675-1680.

[45] Lueking, A., Horn, M., Eickhoff, H., Bussow, K., Lehrach,H., and Walter, G. 1999. Protein microarrays for gene expression and antibody screening. Anal Biochem, 270: 103-111.

[46] Merchant, M., and Weinberger, S.R. 2000. Recent advancements in surface-enhanced laser desorption/ionization- time of flight-mass spectrometry. Electrophoresis. 21: 1164-1177.

[47] Murakami, H., Liotta, L., and Star, R.A. 2000. IF-LCM: laser capture microdissection of immunofluorescently defined cells for mRNA analysis rapid communication. Kidney Int. 58: 1346-1353.

[48] Neter, J., M.H. Kutner, M.H., Nachtsheim, C.J., and Wasserman, W. 1996. Applied Linear Statistical Models. 4th ed. Irwin Press, Chicago.

[49] Nicholls, A.W., Holmes, E., Lindon, J.C., Shockcor, J.P., Farrant, R.D., Haselden, J.N., Damment, S.J., Waterfield, C.J., and Nicholson, J.K. 2001. Metabonomic investigations into hydrazine toxicity in the rat. Chem. Res. Toxicol. 14: 975-987.

[50] Nuwaysir, E.F., Bittner, M., Trent, J., Barrett, J.C., and Afshari, C.A. 1999. Microarrays and toxicology: the advent of toxicogenomics. Mol. Carcinog. 24: 153-159.

[51] Smith, L.L. 2001. Key challenges for toxicologists in the21st century. Trends Pharmacol. Sciences. 22: 281-285.

[52] Solberg, H.E. 1978. Discriminant analysis. CRC Crit. Rev.Clin. Lab. Sci. 9: 209-242.

[53] Sreekumar, A., Nyati, M.K., Varambally, S., Barrette, T.R., Ghosh, D., Lawrence, T.S., and Chinnaiyan, A.M. 2001.

[54] Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. Cancer Res. 61: 7585-7593.

[55] Steiner, S., and Anderson, N.L. 2000. Pharmaceutical proteomics. Ann. N.Y. Acad. Sci. 919: 48-51. Stults, J.T. 1995. Matrix-assisted laser desorption/ionization mass spectrometry (MALDI- MS). Curr. Opin. Struct. Biol. 5: 691-698.

[56] Tafeit, E., and Reibnegger, G. 1999. Artificial neural networks in laboratory medicine and medical outcome prediction. Clin. Chem. Lab. Med. 37: 845-853. Tennant. R.W. The National Center for Toxicogenomics: Using new technologies to inform mechanistic toxicology. Environ. Health Perspectives. In Press.

[57] Tokunaga, E., Maehara, Y., Oki, E., Koga, T., Kakeji, Y., and Sugimachi, K. 2000. Application of quantitative RTPCR using “TaqMan” technology to evaluate the expression of CK 18 mRNA in various cell lines. J. Exp. Clin. Cancer Res. 19: 375-381.