

## FORUM

# Challenges and Limitations of Gene Expression Profiling in Mechanistic and Predictive Toxicology

Mark R. Fielden and Tim R. Zacharewski<sup>1</sup>

*Department of Biochemistry and Molecular Biology, National Food Safety and Toxicology Center, Institute of Environmental Toxicology,  
Michigan State University, 223 Biochemistry Building, Wilson Road, East Lansing, Michigan 48824*

Received August 16, 2000; accepted October 23, 2000

---

RNA and protein expression profiling technologies have revolutionized how toxicologists can study the molecular basis of adverse effects of chemicals and drugs. It is expected that these new technologies will afford efficient and high-throughput means to delineate mechanisms of action and predict toxicity of unknown agents. To reach these goals, a more thorough understanding of the constraints of the methodology is needed to design genome-scale studies and to interpret the vast amount of data collected. This paper addresses some of the limitations and uncertainties of gene expression profiling in mechanistic and predictive toxicology with respect to the expectations of toxicogenomics. The challenges associated with interpreting information from large-scale gene expression experiments *in vivo* is also discussed.

**Key Words:** gene expression profiling; mechanistic and predictive toxicology; toxicogenomics; genome analysis.

---

cDNA and oligonucleotide arrays and high-throughput 2-D electrophoresis systems have quickly emerged as the premier tools to enable genomewide analysis of gene expression at the RNA and protein level. These new technologies are heavily influencing drug discovery and preclinical safety in the biotechnology and pharmaceutical industry (Freeman, 2000). Toxicologists are also promoting genomic expression technologies as a superior alternative to traditional rodent bioassays to identify and assess the safety of chemicals and drug candidates for human safety (Afshari *et al.*, 1999; Nuwaysir *et al.*, 1999; Pennie *et al.*, 2000). It is expected that gene expression profiling will identify mechanisms of action that underlie the potential toxicity of chemicals and drug candidates. Other touted applications include the identification of biomarkers of toxicity to predict potential hazardous substances and therapeutics. Ultimately, toxicogenomics (the integration of genomics, bioinformatics, and toxicology) is expected to accelerate

drug development and aid risk assessment. Recent experiments applied to cancer genetics have demonstrated the potential of gene expression profiling to accurately classify disease phenotypes (Alizadeh *et al.*, 2000; Bittner *et al.*, 2000), thus lending hope that expression profiling may classify and thus predict phenotypes of toxicity. Despite these expectations, it is still uncertain how gene expression profiling experiments will ultimately contribute to our understanding of toxicity and allow us to realize the full potential of this new technology. Although there has been much review and hyperbole surrounding the potential applications of toxicogenomics, these novel and unverified approaches to toxicological problems require an awareness of the constraints of the methodology in order to design and interpret gene expression profiling data. Pennie *et al.* (2000) have also discussed the possibilities and caveats of gene expression profiling in the context of mechanistic and predictive toxicology and have addressed the certainty, biological relevance, and need for validation of microarray data. The purpose of this paper is to illustrate the current constraints of gene expression profiling in mechanistic and predictive toxicology and to stress how current experimental designs may confound accurate interpretation of genome-scale data. The limitations described are not intended to discourage the application of gene expression profiling technologies to mechanistic or predictive toxicology, but rather guide experiments that will produce more interpretable and useful data.

### Gene Expression Profiling in Mechanistic Toxicology— A Hypothesis-generating Tool

There is a certain degree of faith that gene expression profiling will reveal the mechanisms of action of chemicals and drugs despite the inherent limitation of genomic and proteomic experiments, which measure single end points (i.e., RNA or protein levels), albeit for thousands of genes at a time. Consider the many experiments and end points that have been employed to explain the mechanism of action of some previ-

<sup>1</sup> To whom correspondence should be addressed. Fax: (517) 353-9334. E-mail: tzachare@pilot.msu.edu.

ously characterized chemicals and drugs. It is understood that mechanisms of action are far more complex and affect more than simply the levels of cellular macromolecules. Many toxicants affect enzyme activity, DNA integrity, redox status, membrane integrity, and other processes that are not amenable as yet to genomewide measurements. Although alterations in the above processes are likely to indirectly affect the expression of genes and proteins, the question remains how we extrapolate a mechanism of action from the one end point. Similarly, predictive toxicology attempts to infer the potential mechanism(s) of action of an unknown agent on the basis of correlation to large databases of activity or expression profiles (Hughes *et al.*, 2000; Paull *et al.*, 1989; Scherf *et al.*, 2000).

Can mechanisms of action be determined or predicted from gene expression profiling? To answer this question, we need to first define what is meant by a mechanism of action, as the term is often used with many connotations. The mechanism of action of a chemical or drug is described by the series of molecular events following interaction of a chemical with its cellular target(s) and the subsequent alteration(s) in target function that precedes a cascade of cellular events that ultimately leads to the observed effect. The challenge of trying to determine the mechanism of action from measuring steady-state mRNA or protein levels is that many toxicants and drugs initiate toxicity by binding to proteins and/or altering macromolecules (although with exceptions, as noted below), and not by directly inducing gene expression or altering gene product stability or turnover. For example, the mechanism of action of acetaminophen (APAP)-induced hepatocellular necrosis is due to cytochrome P450-catalyzed activation of APAP to the electrophilic NAPQI intermediate, leading to arylation and thiol oxidation of cellular proteins. These events in turn lead to nonspecific and/or undefined alterations in protein function and subsequent changes in nuclear and organelle structure and function leading to irreversible cell injury and oncotic necrosis (Cohen and Khairallah, 1997). The mechanism of action of APAP has been delineated through many detailed chemical and biochemical experiments that could not have been revealed through observation of gene expression changes alone.

This is a limited view of the complete spectrum of toxic effects initiated by APAP, as it has been observed to cause chromosomal aberrations, apoptotic DNA fragmentation, unscheduled DNA synthesis, oxidative stress, altered calcium homeostasis, and inhibition of cell proliferation (Boulares *et al.*, 1999, and references therein). The fact that multiple cellular signaling pathways may converge to alter the expression of the same gene products also makes it difficult to identify the affected pathway from observing gene expression changes. The above arguments illustrate the point that most chemicals and drugs will act through multiple mechanisms of action that will depend on dose, timing and duration of exposure, and cell phenotype. Each individual mechanism represents an initiating event which by itself is inadequate to drive progression of toxicity, but these mechanisms together act in concert to cause

cell injury and/or death. Although gene expression would be expected to be altered as a result of APAP exposure, the changes in gene expression will reflect secondary outcomes due to primary upstream events starting with the interaction of APAP with its target protein(s). Therefore, our ability to define the mechanism of action of a compound using gene expression profiling technologies will be highly limited in resolution. In the best-case scenario, gene expression changes in cellular perturbation experiments will lead to many new testable hypotheses that will require subsequent molecular and biochemical experiments to reveal and confirm precise mechanisms of action.

Ultimately, being able to define the mechanisms of organismal toxicity will depend on our understanding of cellular and tissue level effects and how they are related to the molecular changes in target cells. Because changes in gene expression do not necessarily imply toxicity, gene expression profiling experiments need to be integrated into larger studies that examine multiple end points at the molecular, cellular, tissue, and physiological levels in the context of the whole organism. As noted (Pennie *et al.*, 2000), this creates a further challenge in trying to integrate knowledge at all levels of biological organization, and highlights the need for an interdisciplinary approach in mechanistic toxicology.

In some instances of toxicity, a direct and primary response affecting gene expression and subsequent initiation of toxicity, due likely to a receptor-mediated pathway, may be used to explain the mechanism of action of chemicals, including non-genotoxic carcinogens or endocrine disruptors. This will be particularly true for therapeutics and drug candidates, as it has been estimated that close to 50% of marketed drugs act through receptors (Drews, 2000). These observations will be complicated, of course, by parallel mechanisms of toxicity that may or may not be receptor mediated, yet may augment the receptor-mediated events. The challenge will be to distinguish the therapeutic effects from the pathological changes. This will require establishing time-dependent relationships between dose and toxicity, which may or may not be linear. Where alterations in gene expression precede or coincide with toxicity, our ability to understand the mechanism of action will be limited to our understanding of the pathways that regulate transcription of the affected genes and their kinetics of expression. This is currently a major limitation in understanding why a particular gene or cluster of genes is observed to be up- or down-regulated, since only a small fraction of the estimated 100,000+ human genes have been studied at the level of transcriptional regulation. Combining the identification of gene regulatory elements with expression profiles in microarray experiments (Barzma *et al.*, 1998; Zhang, 1999) represents an industrious approach to begin to understand what transcription factors and upstream signaling molecules are governing the observed response in gene expression following chemical or drug exposure.

Gene expression profiling has possibly a greater potential to

reveal modes of action through the analysis of secondary responses and/or the series of contingent regulatory events induced by chemical or drug exposure. The mode of action of a chemical or drug can be described, in part, by a fundamental obligatory step directing toxicity, or adverse cell fate, be it reversible cell injury, apoptotic or oncotic necrosis, or malignant transformation. Farr and Dunn (1999) have noted that organismal manifestations of toxicity can be explained by combinations of a limited number of cellular outcomes from a limited number of cell and/or tissue types. Furthermore, multiple mechanisms of action may converge at common points to trigger the same molecular response. If this is true, then the number of possible modes of action will be limited to the number of molecular responses that can drive the obligatory step toward a discrete cellular outcome. It follows, then, that gene expression profiles cannot be used as an explanation or predictor of toxicity unless correlated with an adverse effect. Again, this underscores the need to integrate genomic experiments with experiments examining effects at higher levels of biological organization that are intended to assess toxicity in the context of the whole organism. By understanding the gene expression changes that direct a unique cellular outcome (i.e., the mode of action), we can begin to use gene expression profiles to explain and potentially predict toxicity.

#### Predictive Toxicology—Fact or Fiction

It has been proposed that each chemical that acts through a particular mechanism of action will induce a unique and diagnostic gene expression profile under a given set of conditions (Nuwaysir *et al.*, 1999). Indeed, proof-of-principle experiments in *S. cerevisiae* have revealed that the response to inhibitory compounds mimics the loss of function of its target or pathway for at least six compounds (Hughes *et al.*, 2000, and references within; Marton *et al.*, 1998). For example, genetic disruption of calcineurin in *S. cerevisiae* resulted in a gene expression profile highly correlated with the expression profile of wild-type cells treated with FK506 or cyclosporin, antagonists of the calcineurin-signaling pathway. To estimate the significance of the relationship, the FK506 treatment profile was compared to more than 40 randomly selected deletion strains or drug-treated cells and found to be uncorrelated (Marton *et al.*, 1998). Whether predictive patterns in gene expression can be observed in mammalian systems remains to be shown, although preliminary studies suggest they can (Blanchard *et al.*, 2000; McMillian *et al.*, 2000). Therefore, there is significant potential for chemicals and drugs to be classified based on the similarity of their induced gene expression profile by comparison with expression profiles induced by chemicals or drugs with known mechanism of action using multivariate statistical methods and correlation metrics. In some cases, however, their classification may be limited to the affected signaling or metabolic pathway rather than by target protein in the pathway. By extension of this observation, gene expression profiles are anticipated to

produce knowledge of a subset of commonly regulated genes that can be used as biomarkers to predict modes of action.

While it has been pointed out that the number of possible patterns of differential gene expression, even when expressed as binary variables, is enormous (Farr and Dunn, 1999), subtle differences in the number and magnitude of gene expression changes have proven to be sufficient to classify expression profiles into distinct clusters when applied to *S. cerevisiae* (Hughes *et al.*, 2000). The utility of this approach, however, may be lost when outside the context of a large database, or compendium, of expression profiles, as subtle changes in relative expression level (i.e., less than 2-fold) are usually considered unreliable in isolation (Hughes *et al.*, 2000). Based on gene expression profiles of yeast mutants, it has been estimated that there exist 300 to 700 distinct full genome transcriptional patterns from a full set of 5000 yeast deletion mutants profiled under a single condition (Hughes *et al.*, 2000). Although this was a crude prediction, an extrapolation to mammalian systems may predict substantially more distinct transcriptional patterns under a single condition. Classifying transcriptional responses into distinct diagnostic clusters may prove more problematic if responses under different conditions do not extrapolate under different conditions. For example, transcriptional responses may differ between one target cell and another, from cell culture to *in vivo* conditions, or from rodent models to humans. Thus, the predictive power of gene expression profiling may be limited to the model system employed and the prototypical compound with known mechanisms used to generate the diagnostic expression profile. As yet there is no published data to support that predictive expression profiles will extrapolate to other tissues or *in vivo* settings.

#### The Challenge of Interpreting Gene Expression Data

Currently, there is a significant knowledge gap in our understanding of the molecular events that govern toxicologically relevant outcomes. In any event, the changes in gene expression directing cell fate will reflect, in part, an active physiological response that is nontoxic. These responses may include, but are not limited to, host-defense responses (e.g., acute phase proteins, cytokines, DNA repair enzymes), adaptive responses (e.g., hyperplasia, metaplasia, hypertrophy, atrophy), and regenerative or protective responses (proliferation, differentiation). In addition, there will be secondary responses following toxicity that will reflect pathology as a result of disturbances in cell function. These responses are likely to be idiosyncratic and diverse across cell types due to the interaction of pathological responses with the physiological mechanisms of detoxification and repair that are cell specific. Again, the challenge then lies in differentiating the physiological responses from the diagnostic pathological changes in light of confounding experimental artifacts inherent in the model system and the experimental design.

Consider, for example, an experiment designed to measure

time- or dose-dependent changes in gene expression following an  $EC_{50}$  dose of a cytotoxic chemical in cultured cells. When administering a dose that kills half the cell population, the measured response (i.e., mRNA or protein level) in the affected culture will be a combination of multiple factors, including the gene expression changes in dying cells due to treatment, adaptive changes in surviving cells due to treatment, and normal responses in living cells due to adjacent necrotic cells. This would be particularly relevant *in vivo*, as necrosis can induce a regenerative or inflammatory response in some populations of unaffected or resistant cell types. The heterogeneous responses are likely to be highly dependent on the tissue or cell type affected, again highlighting the limitation of extrapolating one model system to another. Measuring gene expression changes following sublethal exposure concentrations may be more likely to reveal treatment-induced changes that initiate toxicity before heterotypic cellular responses obscure interpretation. This will require a complete characterization of the full dose- and time-response relationship, including a qualitative description of cellular changes as correlates.

Artificial complications may also apply to other classes of chemicals, particularly chemicals that act through receptor-mediated pathways, as receptor expression is usually restricted to discrete cell types. Subsequent changes in paracrine signaling may have dramatic effects that could lead to misinterpretation of gene expression profiles in cultured cells. This would also be particularly relevant *in vivo*, where cellular complexity plays a dominant role in adaptation and defense, or when target tissues are affected secondary to primary targeting of a proximal endocrine gland such as the pituitary or thyroid. Furthermore, when analyzing gene expression profiles from whole tissues as part of a whole-animal toxicology study, the relevant gene expression changes in the specific cell types targeted by chemical or drug may be masked or diluted by the benign changes in surrounding cell types. For example, consider the cell type-specific toxicity of alloxan or streptozotocin on the  $\beta$  cells of the pancreas and the fact that the  $\beta$  cells represent less than 2 % of the pancreatic cell population. The ability to detect changes in gene expression within 2 % of an RNA sample derived from whole pancreas is likely below the limits of sensitivity of current genomic profiling platforms. Compensatory changes in other, more abundant cell types may also negate any changes in the targeted cell and could even result in the opposite conclusion regarding message or protein abundance. Being able to measure gene expression profiles in individual targeted cells or cell types, by using laser capture microdissection, for example, would be more desirable in these instances (Emmert-Buck *et al.*, 2000; Luo *et al.*, 1999). However, prior knowledge of the target tissues and/or cell types from pathology studies are typically required for this level of investigation. This would preclude its utility in higher throughput predictive assays that are currently desired, but would prove useful for mechanistic studies. Reducing the number of observations (i.e., gene expression profiles) and correlating

them with a binary response (i.e., apoptosis, DNA damage) may allow for the identification of a more robust set of predictive markers with utility in higher throughput systems. Realization of such a scenario will be heavily dependent on the standards of known modes of action that are available, the reproducibility of the model system, statistical robustness of the data, and the application of multivariate methods of analysis to reduce the data set into a comprehensible and manageable number of components for purposes of classification.

Other limitations to consider arise when adverse cellular or tissue functions are observed in the absence of discrete cellular outcomes such as cell injury or death. In these instances, subtle changes in cell function, such as reduced responsiveness to endocrine signals or altered secretion or production of signaling molecules, will be more difficult to observe, as relevant changes may be transient, posttranslational, and/or in non-target organs. Many endocrine disruptors will likely fall into this category. Perhaps the greatest source of complexity and variability in gene expression profiling experiments *in vivo* stems from non-treatment-related phenomena, or intrinsic variability, which is difficult, if not impossible, to control and reproduce. Normal fluctuations in gene expression will occur as a result of differences in age, gender, temperature, light, diet, and hormonal status. Although age, gender, and the external environment can be tightly controlled within experiments, comparisons between laboratories using similar treatment protocols may be more challenging when environmental factors are not strictly adhered to. Differences in nutritional or hydration status, time of last meal, hormonal fluctuations during estrus, and seasonal and light-induced changes in hormone levels are more difficult to control within experiments. Such intrinsic variation is likely to interact with timing, duration, and frequency of treatments to alter the observed response in gene expression. As with any experiment designed to test a hypothesis, there must be sufficient replication to assure certainty in the experimental results.

The expectation that toxicogenomics will enable us to define mechanisms of action and predict toxicity of unknown agents is supported by recent studies in lower eukaryotes. However, our current ability to define a mechanism of action or accurately predict toxicity in mammalian systems is still in its infancy. Incorporating genomic experiments into larger studies designed to assess effects at higher levels of biological organization is a must if one is to begin to understand and predict organismal outcomes and possibly incorporate gene expression data into mechanism-based risk assessment. The progression of expression profiling into whole-animal studies also presents a higher level of complexity that challenges our understanding of biological systems and the interpretation of what changes in gene expression are relevant. It is expected that experience and interdisciplinary collaborations will continue to advance the utility of gene expression profiling in mechanistic and predictive toxicology. However, continued discussion, debate, and

the sharing of knowledge and data are vital for toxicogenomics to move ahead rapidly.

## REFERENCES

- 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.
- Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., Grever, M., Byrd, C., Botstein, D., Brown, P., and Staudt, L. M. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiles. *Nature* **403**, 503–511.
- Barzma, A., Jonassen, I., Vilo, J., and Ukkonen, E. (1998). Predicting gene regulatory elements in silico on a genomic scale. *Genome Res.* **8**, 1202–1215.
- Blanchard, K., DiSorbo, O., Burris, R., Dunn, R., Farr, S., and Stoll, R. (2000). Toxicogenomics: Understanding the use of microarrays for toxicology studies *in vivo*. *Toxicologist* **54**, 195.
- Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, J., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D., and Sondak, V. (2000). Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* **406**, 536–540.
- Boulares, H. A., Giardina, C., Navarro, C. L., Khairallah, E. A., and Cohen, S. D. (1999). Modulation of serum growth factor signal transduction in Hepa 1–6 cells by acetaminophen: An inhibition of *c-myc* expression, NF-kappaB activation, and Raf-1 kinase activity. *Toxicol. Sci.* **48**, 264–274.
- Cohen, S. D., and Khairallah, E. A. (1997). Selective protein arylation and acetaminophen-induced hepatotoxicity. *Drug Metab. Rev.* **29**, 59–77.
- Drews, J. (2000). Drug discovery: A historical perspective. *Science* **287**, 1960–1964.
- Emmert-Buck, M. R., Gillespie, J. W., Paweletz, C. P., Ornstein, D. K., Basrur, V., Appella, E., Wang, Q. H., Huang, J., Hu, N., Taylor, P., and Petricoin, E. F., III (2000). An approach to proteomic analysis of human tumors. *Mol. Carcinog.* **27**, 158–165.
- Farr, S., and Dunn, R. T., II (1999). Concise review: Gene expression applied to toxicology. *Toxicol. Sci.* **50**, 1–9.
- Freeman, T. (2000). High throughput gene expression screening: its emerging role in drug discovery. *Med. Res. Rev.* **20**, 197–202.
- Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D., Kidd, M. J., King, A. M., Meyer, M. R., Slade, D., Lum, P. Y., Stepaniants, S. B., Shoemaker, D. D., Gachotte, D., Chakraburty, K., Simon, J., Bard, M., and Friend, S. H. (2000). Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126.
- Luo, L., Salunga, R. C., Guo, H., Bittner, A., Joy, K. C., Galindo, J. E., Xiao, H., Rogers, K. E., Wan, J. S., Jackson, M. R., and Erlander, M. G. (1999). Gene expression profiles of laser-captured adjacent neuronal subtypes. *Nat. Med.* **5**, 117–122.
- Marton, M. J., DeRisi, J. L., Bennett, H. A., Iyer, V. R., Meyer, M. R., Roberts, C. J., Stoughton, R., Burchard, J., Slade, D., Dai, H., Bassett, D. E., Jr., Hartwell, L. H., Brown, P. O., and Friend, S. H. (1998). Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat. Med.* **4**, 1293–1301.
- McMillian, M., Ciervo, J., Li, L., Dunn, R., Fairfield, E., Farr, S., and Johnson, M. (2000). Comparison of toxicant-induced gene expression patterns in HepG2 human hepatoma cells. *Toxicologist* **54**, 384.
- 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.
- Pennie, W. D., Tugwood, J. D., Oliver, G. J., and Kimber, I. (2000). The principles and practice of toxicogenomics: Applications and opportunities. *Toxicol. Sci.* **54**, 277–283.
- Paull, K. D., Shoemaker, R. H., Hodes, L., Monks, A., Scudiero, D. A., Rubinstein, L., Plowman, J., and Boyd, M. R. (1989). Display and analysis of patterns of differential activity of drugs against human tumor cell lines: Development of mean graph and COMPARE algorithm. *J. Natl. Cancer Inst.* **81**, 1088–1092.
- Scherf, U., Ross, D. T., Waltham, M., Smith, L. H., Lee, J. K., Tanabe, L., Kohn, K. W., Reinhold, W. C., Myers, T. G., Andrews, D. T., Scudiero, D. A., Eisen, M. B., Sausville, E. A., Pommier, Y., Botstein, D., Brown, P. O., and Weinstein, J. N. (2000). A gene expression database for the molecular pharmacology of cancer. *Nat. Genet.* **24**, 236–244.
- Zhang, M. Q. (1999). Large-scale gene expression data analysis: A new challenge to computational biologists. *Genome Res.* **9**, 681–688.