

TOXICOLOGICAL HIGHLIGHT

Gene Expression Analysis Reveals Chemical-Specific Profiles

Kevin T. Morgan¹

Toxicogenomics Mechanisms, GlaxoSmithKline, Inc., Research Triangle Park, North Carolina 27709

The articles highlighted in this issue are “Gene Expression Analysis Reveals Chemical-Specific Profiles” by Hisham K. Hamadeh, Pierre R. Bushel, Supriya Jayadev, Karla Martin, Olimpia DiSorbo, Stella Sieber, Lee Bennett, Raymond Tennant, Raymond Stoll, J. Carl Barrett, Kerry Blanchard, Richard S. Paules, and Cynthia A. Afshari (pp. 219–231) and “Prediction of Compound Signature Using High Density Gene Expression Profiling” by Hisham K. Hamadeh, Pierre R. Bushel, Supriya Jayadev, Olimpia DiSorbo, Leping Li, Raymond Tennant, Raymond Stoll, J. Carl Barrett, Richard S. Paules, Kerry Blanchard, and Cynthia A. Afshari (pp. 232–240).

For it is the greatest truth of our age: information is not knowledge.
—C. Carr (2000)

The scientific and popular press has been replete with optimistic promises concerning the technological products of the “-omics revolution.” We are still, however, on a very steep learning curve with respect to the application of these powerful and sometimes overwhelming technologies, such as large-scale differential gene expression (LSDGE), that generate many thousands of data points in a single small experiment. Interesting examples of the application of LSDGE to toxicology are appearing with increasing frequency. Examples include initial attempts to address basic issues, such as data quality, artifacts and normalization (Crosby *et al.*, 2000), temporal changes in an important *in vitro* system, primary rat hepatocytes (Baker *et al.*, 2001), and expression of mRNA levels for rat hepatic drug metabolizing enzymes (Gerhold *et al.*, 2001).

The highlighted articles describe LSDGE studies of chemically induced hepatotoxicity in rats, and address data interpretation, an important aspect of such research. There are currently two major philosophies being applied to interpretation of these large data sets.

Functional analysis. Transcriptional changes are mechanistically linked to probable physiological and pathophysiological responses. For this approach, only genes with established downstream (generally protein) functions are considered.

Pattern recognition. Statistical correlations are made between groups of gene changes that have a discernable motif (with respect to which ones are regulated up or down). This

work can be undertaken without consideration of any mechanistic link between gene and function, and all messenger RNA transcripts, including those for which the function is unknown, can be used.

These approaches require very different experimental procedures and skills. Functional analysis requires extensive knowledge of the relevant cell biology of the system under investigation. Pattern recognition is more dependent upon computer science, statistics and mathematics. These approaches will eventually merge. It is to be hoped that researchers using the new technologies in this “infocentric” time will be wise enough to take heed of lessons learned from “old” techniques, and also to distinguish clearly between information and knowledge (Brown and Duguid, 2000). Both approaches are being applied successfully today, and Hamedeh *et al.* have incorporated an element of each into their publications.

With respect to functional interpretation, readers are encouraged to look closely at Figure 6, in the first article. The difficulty associated with creating such diagrams, and their potential value, should not be underestimated. The approach taken here was to (a) thoroughly review the published literature on the compound to be studied, (b) display the data on a map of toxicant effector pathways, and (c) portray the data in such a way as to highlight new contributions provided by their study. Such a map or diagram places the data in context and provides a guide to follow-up research needed to confirm potential functional changes associated with the reported transcriptional responses. From personal experience, it is clear that a gene list, which contains the same information, turns people off. A good diagram draws people into the knowledge-rich realm of the transcriptome. Other diagramming procedures have been developed (Crosby *et al.*, 2000; Kohn 1999), and some excellent examples are appearing on the Internet (<http://www.its.caltech.edu/~mirsky/endomeso.htm>). The use of a standardized approach to diagram preparation could markedly accelerate interdisciplinary communication in toxicogenomics.

Many research groups are exploring the pattern recognition approach, as it holds hope for quick answers. For instance, the majority of research groups engaged in LSDGE use clustering algorithms to group genes with respect to similarity of “behavior.” We do so as the first stage of our expression interpretation process. A number of commercial programs are available, and

¹ For correspondence via fax: (919) 315-8391. E-mail: ktm42807@gsk.com.

an excellent clustering tool provided by Eisen and colleagues (1998) is available via the Internet (<http://rana.lbl.gov/EisenSoftware.htm>). If you are considering embarking in the field of toxicogenomics, I recommend that you first try this tool on the time course data of Iyer *et al.* (1999), which is available at <http://www.sciencemag.org/feature/data/984559.shl>.

Hamadeh *et al.*, by applying a pattern recognition approach to their data, demonstrated their ability to distinguish “chemical-specific gene expression profiles.” Using an initial training set comprised of gene expression profiles derived from livers of rats exposed to four compounds for 24 hours or 2 weeks, they tested the hypothesis that “knowledge can be gained regarding the nature of blinded samples.” It was concluded that this study successfully predicted whether samples were derived from rats treated with enzyme inducers or peroxisome proliferators. In spite of the fact that no expression responses detected by the arrays in these studies were confirmed by other techniques (real time PCR or Northern Blot), and that this work is limited to a very small set of compounds, I concur with the authors’ encouraging conclusions.

An important feature of LSDGE is a need to include more than one time point. This I consider essential while toxicogenomics is still in an early research phase. Waring *et al.* (2001, p. XX) stated that they “believe that if gene expression cluster analysis is to be used to predict mechanisms of toxicity for thousands of compounds, it needs to be robust enough to cluster based on a single time point.” Given the dynamic nature of the transcriptome (Klevecz and Bowse, 2000), single time point or (for that matter) single dose (exposure concentration) studies should be avoided where resources permit more meaningful experiments. The use of only two time points in the study of Hamadeh *et al.* contributed significantly to the success of their venture. I do concur with Waring *et al.*, who suggested that “these assays may prove to be a highly sensitive technique for safety screening of drug candidates and for the classifica-

tion of environmental toxins.” The work reported by Hamedeh *et al.* provides further support for this proposal. These authors are to be congratulated on the completion of such a thorough and wide-ranging piece of research in the still technically challenging and “embryonic” discipline of LSDGE-based toxicogenomics.

REFERENCES

- Baker, T. K., Carfagna, M. A., Gao, H., Dow, E. R., Li, Q., Searfoss, G. H., and Ryan, T. P. (2001). Temporal gene expression analysis of monolayer cultured rat hepatocytes. *Chem. Res. Toxicol.* **14**, 1218–1231.
- Brown, J. S., and Duguid, P. (2000). *The Social Life of Information*. Harvard Business School Press, Cambridge, MA.
- Carr, C. (2000). *Killing Time, A Novel of the Future*. Warner Books, New York.
- Crosby, L. M., Hyder, K., DeAngelo, A. B., Kepler, T. B., Gaskill, B., Benavides, G., Yoon, L., and Morgan, K. T. (2000). Morphologic analysis correlates with gene expression changes in cultured F344 rat mesothelial cells. *Toxicol. Appl. Pharmacol.* **169**, 205–221.
- 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. U.S.A.* **95**, 14863–14868.
- Gerhold, D., Lu, M., Xu, J., Austin, C., Caskey, C. T., and Rushmore, T. (2001). Monitoring expression of genes involved in drug metabolism and toxicology using DNA microarrays. *Physiol. Genomics* **5**, 161–170.
- Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C. F., Trent, J. M., Staudt, L. M., Hudson, J. J., Boguski, M. S., Lashkari, D., Shalon, D., Botstein, D., and Brown, P. O. (1999). The transcriptional program in the response of human fibroblasts to serum. *Science* **283**, 83–87.
- Klevecz, R. R., and Dowse, H. B. (2000). Tuning in the transcriptome: Basins of attraction in the yeast cell cycle. *Cell Prolif.* **33**, 209–218.
- Kohn, K. W. (1999). Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol. Biol. Cell* **10**, 2703–2734.
- Waring, J. F., Jolly, R. A., Ciurlionis, R., Lum, P. Y., Praestgaard, J. T., Morfitt, D. C., Buratto, B., Roberts, C., Schadt, E., and Ulrich, R. G. (2001). Clustering of hepatotoxins based on mechanism of toxicity using gene expression profiles. *Toxicol. Appl. Pharmacol.* **175**, 28–42.