Toxicogenomics in Risk Assessment: An Overview of an HESI Collaborative Research Program

William Pennie,[1] Syril D. Pettit,[2] and Peter G. Lord[3]

[1]Drug Safety Evaluation Department, Pfizer Inc, Groton, Connecticut, USA; [2]ILSI Health and Environmental Sciences Institute, Washington, DC, USA; [3]Johnson & Johnson Pharmaceuticals, Raritan, New Jersey, USA

The value of genomic approaches in hypothesis generation is being realized as a tool for understanding toxicity and consequently contributing to an assessment of drug and chemical safety. In 1999 the membership of the International Life Sciences Institute Health and Environmental Sciences Institute formed a committee to develop a collaborative scientific program to address issues, challenges, and opportunities afforded by the emerging field of toxicogenomics. Experts and advisors from academia and government laboratories participate on the committee, along with approximately 30 corporate member organizations from the pharmaceutical, agrochemical, chemical, and consumer products industries. The committee has designed, conducted, and analyzed numerous toxicogenomic experiments within the broad fields of hepatotoxicity, nephrotoxicity, and genotoxicity. The considerable body of data generated by these programs has been instrumental in increasing understanding of sources of biological and technical variability in the alignment of toxicant-induced transcription changes with the accepted mechanism of action of these agents and the challenges in the consistent analysis and sharing of the voluminous data sets generated by these approaches. Recognizing the importance of standardized microarray data formats and public repository databases as the mechanism by which microarray data can be compared and interpreted by the scientific community, the committee has partnered with the European Bioinformatics Institute to develop a database to house the data generated by its collaborative research. *Key words:* genomics, HESI, microarrays, risk assessment, toxicogenomics. *Environ Health Perspect* 112:417–419 (2004). doi:10.1289/txg.6674 available via *http://dx.doi.org/* [Online 15 January 2004]

Gene transcription lies at the beginning of a response of a cell to a xenobiotic. Thus, a transcriptional response can give a preliminary indication of the biochemical or biological mechanism being affected by a xenobiotic, and gene expression data can provide starting points in a toxicological examination. The use of genomics technologies, particularly gene arrays, as tools for identifying profiles of gene expression associated with particular compounds and/or toxicities has shown increasing merit. If a good correlation exists between gene expression and a toxic mechanism, then the genomic data may provide supportive evidence for that mechanism (Chevalier and Roberts 2001; Harries et al. 2001; Lord et al. 2001). Even when the mechanism is unknown, genomic data can help identify components (i.e., proteins or enzymes) of pathways that may be involved in the biological process under study (Crosby et al. 2000). Developing databases of expression profiles for a wide variety of toxic compounds and toxic models makes it possible to create statistical and computational methods that can indicate the toxic potential of a drug or chemical from the pattern of gene expression changes it elicits in *in vitro* (Brooks and Pennie 2001; Burczynski et al. 2000; Waring et al. 2001) or *in vivo* systems (Hamadeh et al. 2002).

Over the past several years there has been considerable investment by chemical and pharmaceutical companies, government agencies, and technology providers in the application of gene array–based approaches in chemical and drug development. The value of genomic approaches to generate hypotheses is being realized for understanding toxicity and consequently contributing to an evaluation of drug and chemical safety both in predictive toxicology and in mechanism-based risk assessment.

As the field of toxicogenomics emerged in the late 1990s, it became clear there was a need to establish a body of available knowledge to serve as a foundation for applying the data generated by gene array methodologies to risk assessment. To this end, in 1999 the membership of the nonprofit scientific research organization, the International Life Sciences Institute (ILSI) Health and Environmental Sciences Institute (HESI) initiated a collaborative scientific program—the HESI Committee on the Application of Genomics to Mechanism-Based Risk Assessment—to address issues, challenges, and opportunities afforded by toxicogenomics (Robinson et al. 2003). The articles in this mini-monograph include overviews of the design and objectives of this experimental program (Kramer et al. 2004; Mattes et al. 2004; Newton et al. 2004; Ulrich et al. 2004) as well as a series of technical articles detailing data generated and analyzed as part of the HESI genomics research programs (Amin et al. 2004; Baker et al. 2004; Goodsaid et al. 2004; Mattes 2004; Rosenzweig et al. 2004; Thompson et al. 2004; Waring et al. 2004). Additional manuscripts detailing the experimental findings of the HESI Genotoxicity Working Group will be published in the journal *Mutation Research* in March 2004. The complete data set is currently being submitted to ArrayExpress (European Bioinformatics Institute, Hinxton, UK; http://www.ebi.ac.uk/arrayexpress) and will be available for public download by second quarter 2004. Accession numbers referencing this data set will be available on the HESI website (http://hesi.ilsi.org/index.cfm?pubentityid=120).

Toxicogenomics has progressed considerably during the course of the HESI Genomics Committee's research program (1999–2003). Publications on the subject evolved from illustrating the theoretical promise of the technologies (Burchiel et al. 2001; Fielden and Zacharewski 2001; Nuwaysir et al. 1999; Simmons and Portier 2002; Smith 2001; Storck et al. 2002; Tennant 2002; Ulrich and Friend 2002; Waring and Halbert 2002) to illustrating the practical use of the technologies in toxicology (Brooks and Pennie 2001; Hamadeh et al. 2002; Lord et al. 2001; Waring et al. 2001). Many of the initial concerns about the practical use of toxicogenomics (e.g., oversensitivity, lack of comparability, availability of analytical tools) have been addressed,

This article is part of the mini-monograph "Application of Genomics to Mechanism-Based Risk Assessment."

Address correspondence to S. Pettit, ILSI Health and Environmental Sciences Institute, 9th Floor, One Thomas Circle NW, Washington, DC 20005 USA. Telephone: (202) 659-3306. Fax: (202) 659-3617. E-mail: spettit@ilsi.org

We thank the participants of the HESI Genomics Committee for their contributions of time, expertise, and experimental research. Special thanks go to our public-sector steering committee advisors for their assistance in the review of this publication and their leadership on the committee and to G. Morgan and D. Robinson for their important roles in initiating this research program.

The authors declare they have no competing financial interests.

Received 14 August 2003; accepted 15 December 2003.

although several issues still merit resolution before such data are used fully in the risk assessment process.

The HESI Collaborative Research Program

The HESI genomics research program was executed via a multinational team of scientists from academic and government laboratories, along with scientists from more than 30 corporate member organizations from the pharmaceutical, agrochemical, chemical, and consumer products industries (Table 1). Participation in the activities of the committee has afforded the collaborators an unprecedented opportunity to share experiences, best-operating practices, and actual data from a wider cross-section of commercially available and proprietary platforms, protocols, instrumentation, and analysis methods than they would have had access to individually.

Table 1. Organizations participating on the HESI Committee on the Application of Genomics to Mechanism-Based Risk Assessment.

- Private-sector participants
  - Abbott Laboratories
  - Amgen Inc.
  - AstraZeneca
  - Aventis Pharmaceuticals Inc.
  - Bayer AG
  - Bayer CropScience AG
  - Berlex Laboratories
  - Biogen, Inc.
  - Boehringer-Ingelheim Pharmaceuticals, Inc.
  - Bristol-Myers Squibb Co.
  - E.I. duPont de Nemours & Co.
  - Eisai Co., Ltd.
  - Eli Lilly and Co.
  - GlaxoSmithKline
  - Hoffmann-La Roche Inc.
  - Johnson & Johnson Pharmaceutical Research and Development, LLC
  - Meiji Seika Kaisha, Ltd.
  - Merck & Co., Inc.
  - Mitsubishi Pharma Corp.
  - Novartis Pharmaceuticals Corp.
  - Pfizer Inc
  - The Procter & Gamble Co.
  - Sankyo Co., Ltd.
  - Sanofi-Synthélabo Research
  - Schering AG
  - Schering-Plough Research Institute
  - Sumitomo Chemical Co., Ltd.
  - Syngenta Central Toxicology Laboratory
  - Tanabe Seiyaku Co., Ltd.
  - Wyeth Research
- Public-sector participants
  - Government
    - European Agency for the Evaluation of Medicinal Products
    - National Institute of Public Health and the Environment (RIVM)—Netherlands
    - U.S. Air Force
    - U.S. Army Center for Environmental Health Research
    - U.S. Environmental Protection Agency
    - U.S. Food and Drug Administration
    - U.S. National Cancer Institute
    - U.S. National Center for Toxicological Research
    - U.S. National Center for Toxicogenomics
  - Academic
    - McArdle Laboratory for Cancer Research
    - Medical College of Wisconsin
    - Michigan State University
    - The Institute for Genomic Research
    - University of Surrey

Since its inception in 1999, the committee has conducted and analyzed toxicogenomics experiments within the broad fields of hepatotoxicity, nephrotoxicity, and genotoxicity. These experiments were designed *a*) to evaluate responses to prototypical toxicants to determine if known mechanisms of toxicity could be associated with characteristic gene expression profiles, *b*) to identify technological and biological sources of variability associated with toxicogenomic experimental protocols, and *c*) to ultimately evaluate the utility of this technology for risk assessment applications.

The experience highlighted the challenges inherent in the analysis and sharing of the data sets generated by these approaches. As reported in this mini-monograph, the experimental program has characterized the biological and technical sources of variability such as the isolation and labeling of mRNA samples, the detection hardware settings, the analysis software threshold settings, the microarray lot number, and the differences in gene coverage and probe annotation across different technical platforms. Nevertheless, in support for toxicogenomics as a valuable tool in assessing toxicity, the experimental programs have shown that *a*) patterns of gene expression relating to biological pathways are robust enough to allow insight into mechanisms of toxicity; *b*) gene expression data can provide meaningful information on the physical location of the toxicity; *c*) dose-dependent changes can be observed; and *d*) concerns about oversensitivity of the technology may be unfounded (particularly when compared with existing *in vitro* assays for direct- and indirect-acting genotoxicants).

Toxicogenomic Data Storage and Exchange

As part of its mission, the HESI Genomics Committee also recognized the importance of standardized microarray data formats and public repository databases for better comparison and interpretation by the broader scientific community. Because of this, the committee has partnered with the European Bioinformatics Institute, to develop a database (Tox-MIAMExpress; www.ebi.ac.uk/tox-miamexpress/) based on ArrayExpress database structure (Brazma et al. 2003) and MIAME (Minimum Information About a Microarray Experiment) data format standards (Brazma et al. 2001). The database will be consistently annotated and integrated with other relevant information (e.g., histopathology, clinical chemistry, gross observations), employ standard controlled vocabulary, and be supported by a query and data analysis interface. This database will house all the data generated by the HESI genomics consortium as part of the project described herein and will be made available to the public in early 2004. See Mattes et al. (2004) in this issue for more information.

Toxicogenomics and Risk Assessment

As the investments made in the application of genomic technologies mature, there is a determined effort to bring the full force of the technology into risk assessment. The committee's practical experimental experience has provided a valuable substrate for discussions on the technical and logistical challenges associated with the use of such data in the regulatory environment. A workshop with invited academic and regulatory participation (representing North America, Japan, and Europe) was held 5–6 June 2003 in Fairfax, Virginia, to discuss the findings of the committee's working groups (Pettit 2003). The participants were able to exploit these findings to form opinions on the implementation and interpretation of genomic microarray data in the risk assessment process.

Scientists from government agencies are encouraging greater input from both the academic and private sectors in the development of data-quality standards and commonly accepted analysis methods. The committee's work can help answer the many questions related to the use of toxicogenomic data in regulatory submissions or other risk evaluations. For example, what depth of analysis (and in what format) of gene expression changes is required for scientifically meaningful risk assessment? Genomic microarrays used in the pharmaceutical industry typically are limited to early-stage predictive assays and are not used for advanced mechanistic analysis of compounds in later stages of assessment. However, it is becoming clear that microarray data results should be placed in an appropriate biological context (i.e., with other biological end points) for researchers to understand mechanisms underlying toxicity. The relevance of single gene expression changes in the absence of pathway-level gene expression data cannot be assured.

Raw data from publicly referenced experiments needs to be available via public toxicogenomic databases [for example, Tox-MIAMExpress or the Chemical Effects in Biological Systems (CEBS) being developed by the National Center for

Toxicogenomics at the National Institute of Environmental Health Sciences, Research Triangle Park, NC] to create a knowledge base that could be used to support genomics applications in hazard identification. Furthermore, if successfully implemented with the appropriate depth of data content, such databases could serve as robust resources for advanced queries (e.g., genetic patterns of toxicity within/across compound classes, extrapolation of toxicity across species, ages, or durations of exposure). Challenges inherent in populating and using these resources include limitations in the availability of sequence information about probes on microarray platforms, use of nonstandard ontologies for toxicology end points, and relatively few public submissions of genomic data. However, as reported in this mini-monograph, several public consortia and organizations (including HESI and the NCT) are actively engaging the toxicological community in addressing these challenges.

Conclusion

Genomics, and more specifically toxicogenomics, can no longer be regarded as a new technology. The technologies are maturing to the extent that we now have considerable experience in their use. As evidenced in the research conducted by the HESI Genomics Committee and presented in this mini-monograph, it is clear that transcriptional profiles can discriminate between classes of compounds and some toxicities. These very preliminary data can provide users with valuable information for mechanistic evaluation and may even facilitate commercial decisions concerning the compounds to be developed and how to develop them. This committee's collective experience presented in this mini-monograph also provides awareness of the limits of sensitivity and reproducibility of the methods and an understanding of how transcriptomic data can be interpreted in the context of the pathology and other biological data from a toxicology study. As discussed in this mini-monograph, the experimental work and ongoing interactions of the participants on the HESI Genomics Committee represent a unique opportunity for the integration and distillation of this collective experience for the benefit of the regulators and regulated industries as well as for the toxicology community as a whole.

References

Amin RP, Vickers AE, Sistare F, Thompson KL, Roman RJ, Lawton M, et al. 2004. Identification of putative gene-based markers of renal toxicity. Environ Health Perspect 112:465–479.

Baker VA, Harries HM, Waring JF, Duggan CM, Ni HA, Jolly RA, et al. 2004. Clofibrate-induced gene expression changes in rat liver: a cross-laboratory analysis using membrane cDNA arrays. Environ Health Perspect 112:428–438.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, et al. 2001. Minimum Information About a Microarray Experiment (MIAME)—toward standards for microarray data. Nat Genet 29:365–371.

Brazma A, Parkinson H, Sarkans U, Shojatalab M, Vilo J, Abeygunawardena N, et al. 2003. ArrayExpress—a public repository for microarray gene expression data at the EBI. Nucleic Acids Res 31:68–71.

Brooks AN, Pennie WD. 2001. Transcript profiling of the response to environmental hormone mimics. Comments Toxicol 7:303–315.

Burchiel SW, Knall CM, Davis JW, Paules RS, Boggs SE, Afshari CA. 2001. Analysis of genetic and epigenetic mechanisms of toxicity: potential roles of toxicogenomics and proteomics in toxicology. Toxicol Sci 59:193–195.

Burczynski ME, McMillian M, Ciervo J, Li L, Parker JB, Dunn RT, Hicken S, et al. 2000. Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells. Toxicol Sci 58:399–415.

Chevalier S, Roberts RA. 2001. Altered mRNA and protein expression in non-genotoxic hepatocarcinogenesis. Comments Toxicol 7:317–331.

Crosby LM, Hyder KS, DeAngelo AB, Kepler TB, Gaskill B, Benavides GR, et al. 2000. Morphologic analysis correlates with gene expression changes in cultured F344 rat mesothelial cells. Toxicol Appl Pharmacol 169:205–222.

Fielden MR, Zacharewski TR. 2001. Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology. Toxicol Sci 60:6–10.

Goodsaid FM, Smith RJ, Zairov H, Rosenblum IY. 2004. Quantitative PCR deconstruction of discrepancies between results reported by different hybridization platforms. Environ Health Perspect 112:456–459.

Hamadeh HK, Bushel PR, Jayadev S, Martin K, DiSorbo O, Sieber S, et al. 2002. Gene expression analysis reveals chemical-specific profiles. Toxicol Sci 67:219–231.

Harries HM, Fletcher ST, Duggan CM, Baker VA. 2001. The use of genomics technology to investigate gene expression changes in cultured human liver cells. Toxicol In Vitro 15:399–405.

Kramer JA, Pettit SD, Amin RP, Bertram TA, Car B, Cunningham M, et al. 2004. Overview of the application of transcription profiling using selected nephrotoxicants for toxicology assessment. Environ Health Perspect 112:460–464.

Lord PG, Barne KA, Kramer K, Bacon EJ, Mooney J, O'Brien S, et al. 2001. cDNA microarrays in investigative toxicology: a study of differential gene expression in compound induced cardiac hypertrophy. Comments Toxicol 7:381–392.

Mattes WB. 2004. Annotation and cross-indexing of array elements on multiple platforms. Environ Health Perspect 112:506–510.

Mattes WB, Pettit SD, Sansone S-A, Bushel PR, Waters MD. 2004. Database development in toxicogenomics: issues and efforts. Environ Health Perspect 112:495–505.

Newton RK, Aardema M, Aubrecht J. 2004. The utility of DNA microarrays for characterizing genotoxicity. Environ Health Perspect 112:420–422.

Nuwaysir EF, Bittner M, Trent J, Barrett JC, Afshari CA. 1999. Microarrays and toxicology: the advent of toxicogenomics. Mol Carcinog 24:153–159.

Pettit S. 2003. Toxicogenomics in risk assessment: assessing the utility, challenges, and next steps. Workshop sponsored by the ILSI HESI Committee on the Application of Genomics in Mechanism-Based Risk Assessment, 5–6 June 2003, Fairfax, Virginia. Washington, DC:ILSI Press. Available: http://hesi.ilsi.org/index.cfm?pubentityid=120 [accessed 13 December 2003].

Robinson DE., Pettit SD, Morgan DG. 2003. Use of genomics in mechanism based risk assessment. In: Toxicogenomics (Inoue T, Pennie WD, eds). Tokyo:Springer-Verlag, 194–203.

Rosenzweig BA, Pine PS, Domon OE, Morris SM, Chen JJ, Sistare FD. 2004. Dye-bias correction in dual-labeled cDNA microarray gene expression measurements. Environ Health Perspect 112:480–487.

Simmons PT, Portier CJ. 2002. Toxicogenomics: the new frontier in risk analysis. Carcinogenesis 23:903–905.

Smith LL. 2001. Key challenges for toxicologists in the 21st century. Trends Pharmacol Sci 22:281–285.

Storck T, von Brevern MC, Behrens CK, Scheel J, Bach A. 2002. Transcriptomics in predictive toxicology. Curr Opin Drug Discov Devel 5:90–97.

Tennant RW. 2002. The National Center for Toxicogenomics: using new technologies to inform mechanistic toxicology. Environ Health Perspect 110:A8–A10.

Thompson KL, Afshari CA, Amin RP, Bertram TA, Car B, Cunningham M, et al. 2004. Identification of platform-independent gene expression markers of cisplatin nephrotoxicity. Environ Health Perspect 112:488–494.

Ulrich R, Friend SH. 2002. Toxicogenomics and drug discovery: will new technologies help us produce better drugs? Nat Rev Drug Discov 1:84–88.

Ulrich RG, Rockett JC, Gibson GG, Pettit SD. 2004. Overview of an interlaboratory collaboration on evaluating the effects of model hepatotoxicants on hepatic gene expression. Environ Health Perspect 112:423–427.

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.

Waring JF, Halbert DN. 2002. The promise of toxicogenomics. Curr Opin Mol Ther 4:229–235.

Waring JF, Ulrich RG, Flint N, Morfit D, Kalkuhl A, Staedtler F, et al. 2004. Interlaboratory evaluation of rat hepatic gene expression changes induced by methapyrilene. Environ Health Perspect 112:439–448.