June 13, 2007

2007. U.S. Environmental Protection Agency's activities to prepare for regulatory and risk assessment applications of genomics information.

William H. Benson
Kathryn Gallagher
J. Thomas McClintock

Environmental and Molecular Mutagenesis 48:359–362 (2007)

Perspective

U.S. Environmental Protection Agency's Activities to Prepare for Regulatory and Risk Assessment Applications of Genomics Information

William H. Benson,[1] Kathryn Gallagher,[1]* and J. Thomas McClintock[2]

[1]*Office of the Science Advisor, U.S. Environmental Protection Agency, Washington, D.C.*

[2]*Office of Science Coordination and Policy, Office of Prevention, Pesticides, and Toxic Substances, U.S. Environmental Protection Agency, Washington, D.C.*

Genomics is expected to have significant implications for risk assessment and regulatory decision making. Since 2002, the U.S. Environmental Protection Agency (EPA) has undertaken a number of cross-agency activities to further prepare itself to receive, interpret, and apply genomics information for risk assessment and regulatory purposes. These activities include: (1) the issuance of an Interim Genomics Policy on the use of genomics information in risk assessments and decision making, (2) the release of the 2004 Genomics White Paper, which outlines potential applications and implications of genomics for EPA, and (3) the recent release of the external review draft of the Interim Guidance on Microarray-Based Assays, which outlines data submission, quality, analysis, management, and training considerations for such data. This manuscript discusses these activities and more recent follow-up activities with the aim of further communicating these efforts to the broader scientific and stakeholder community. Environ. Mol. Mutagen. 48:359–362, 2007.

Key words: EPA; genomics; risk assessment; regulatory applications

INTRODUCTION

Genomics will provide important information to be considered in risk assessment and regulatory decision making. In 2002, the U.S. Environmental Protection Agency (EPA) issued its Interim Policy on Genomics identifying the Agency's initial approach to the use of genomics data and information in risk assessment and decision making [US EPA, 2002]. The Interim Policy describes genomics as the study of all the genes of a cell or tissue, at the DNA, mRNA, or protein level. The policy notes that genomics data may be considered in the EPA decision making process, as part of the weight of the evidence on a case-by-case basis, but that these data alone are currently considered insufficient as a basis for decisions.

Following the release of the Interim Policy, EPA's Science Policy Council (SPC), which is comprised of senior Agency managers, charged an intra-Agency Genomics Task Force with examining the implications genomics is likely to have on EPA programs and policies, and with identifying examples of instances where genomics data might be applied in risk assessment and decision making. The resulting Genomics White Paper entitled "Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA" identified several areas likely to be influenced by the generation of genomics information within EPA and the submission of such information to EPA [US EPA, 2004]. These areas are: (1) regulatory applications including: (a) prioritization of contaminants and contaminated sites, (b) monitoring, and (c) reporting provisions; and (2) risk assessment applications.

Regulatory Applications

Prioritization of Contaminants and Contaminated Sites

There is a large number of chemical and biological stressors that EPA must prioritize for further evaluation.

*Correspondence to: Kathryn Gallagher, US EPA, Mail Code 8105R, 1200 Pennsylvania Ave, NW, Washington, DC 20460, USA.
E-mail: gallagher.kathryn@epa.gov

Received 30 January 2007; provisionally accepted 2 March 2007; and in final form 5 March 2007

DOI 10.1002/em.20302

Published online in Wiley InterScience (www.interscience.wiley.com).

For example, there are over 80,000 chemicals currently listed in the Toxic Substances Control Act (TSCA) Inventory. There is sufficient information to allow a thorough evaluation of risk for only a fraction of these chemicals since most of them have not undergone extensive toxicological testing. Nevertheless, EPA program and regional offices need to make a variety of decisions about these chemicals. These decisions may include prioritization of the chemical(s) for further evaluation or a decision that no further research is needed. A number of approaches have been developed to aid in prioritization decisions. Currently, chemical prioritization may be determined by factors including production volume, exposure information, persistence, chemical class, analysis of structural analogues, and consideration of more formal structure-activity relationships (SARs). However, all these attributes have limitations, and an enhanced, knowledge-based approach would be useful.

There are also approximately 60 contaminants and contaminant classes, including 10 microbial contaminants, on EPA's Office of Water 1998 final Contaminant Candidate List. There currently is no rapid, comprehensive method for prioritizing which chemicals or microbes should be tested based on the potential for toxicity, and it is recognized that it is not possible to test all stressors. Genomics technologies may be useful in providing more mechanistic, molecular-based data for risk-based prioritization of such stressors. Further, these technologies are likely in the future to offer more efficient, potentially high throughput, and low cost alternatives to tests EPA currently relies on for prioritization.

However, there is currently little scientific consensus concerning which tests would be most appropriate for the Agency's different prioritization needs. The selection of appropriate model system(s) and endpoints would be important considerations for using genomics information in prioritization.

Monitoring

EPA obtains, requests, and receives many types of environmental data for both assessment and compliance purposes, including but not limited to the following: chemical and physical analyses of air, water, soil, and sediment; toxicity testing of various environmental media or chemicals; plant, animal, and human tissue residues of various chemicals or their breakdown products; community structure analyses, microbial community and pathogenic microorganism analyses of air, water, soil, and sediment. The cost and time required in collecting and analyzing the large number of conventional environmental samples needed to make sound regulatory decisions and to evaluate environmental status is enormous. Genomics technologies may ultimately yield rapid, efficient, and cost-effective methods for environmental monitoring.

Reporting Provisions

Reporting of certain adverse effects/risks for industrial chemicals and pesticides already on the market is mandated under specific provisions of both the TSCA and the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). There is a need to interpret how these TSCA and FIFRA reporting provisions apply to genomics data. There are already certain types of conventional tests whose data are not considered to present indication of substantial risk to health or the environment and are not required by the Agency as stand-alone submissions. As the predictability and validity of genomics methods increase, EPA may need to re-evaluate its stance on these reporting provisions. Because these provisions address the reporting of adverse effects, the issue of what genomic changes mean in terms of adversity must be addressed before reporting for genomic responses may be required.

The Toxics Release Inventory (TRI) database was established under the Emergency Planning and Community Right-to-Know (EPCRA) Act of 1986. EPCRA requires certain industrial facilities to annually report information on toxic chemical releases and other waste management activities to EPA and the states to inform communities of chemical hazards in their area. The statutory chemical listing/delisting criteria for TRI are primarily based on hazard, not risk. The emphasis of EPA's hazard assessment is on a chemical's inherent toxicity rather than the potential risks from exposure to the chemical. If there is a linkage to adverse effects in humans or on the environment, genomics data may be considered in the hazard assessment when determining whether or not a chemical meets the TRI chemical listing/delisting criteria. The EPA Interim Policy on Genomics allows such information to be used in the overall assessment on a case-by-case basis, but genomics information alone currently cannot be used to determine hazard at this time.

Risk Assessment Applications

The risk assessment applications listed in the 2004 EPA Genomics White Paper included understanding chemical modes of action, evaluating susceptible populations, analyzing the effects of mixtures, and improving exposure assessment. These applications are relevant to both human health and ecological risk assessment.

Genomics data may allow the development of gene, protein, or metabolite profiles that can advance the screening of individual chemicals and allow faster and more accurate categorization into defined classes according to their mode of action (MOA). Understanding the MOA of environmental agents that induce toxic effects would inform the assessment of the relevance of these findings in protecting human health and safeguarding the environment. MOA information derived from genomic studies

may be useful in improving hazard identification, dose-response assessment, and understanding and informing the use of extrapolations within and between species.

Genomics and related technologies also offer an opportunity to define and identify differences in susceptibility to many environmental contaminants in both humans and animals. Genomics technologies are also expected to be useful in the identification of unique patterns of gene expression in ecological species and human cell models induced by exposure to multiple environmental stressors, including mixtures of chemicals. Furthermore, genomics technologies are likely to lead to the development of useful biomarkers of exposure that can be used in exposure assessments, particularly in the evaluation of potential occupational exposures for human health assessments and for environmental exposures for both human health and ecological risk assessments.

Ecological Risk Assessment

Perhaps one area less discussed in the open scientific literature regarding the use of genomics is the application of these data to ecological risk assessments. While all of the applications identified in the Genomics White Paper are relevant to both human health and ecological risk assessment, it is likely that some of the earliest and most fruitful applications of genomics to ecological risk assessment, as it is currently practiced in the regulatory arena, would be in the areas of extrapolation and exposure assessment.

In ecological risk assessment, it is necessary to extrapolate results from a very limited set of test species to a very wide range and number of species present in the environment. For example, for the registration of pesticides within EPA's Office of Pesticide Programs, in evaluating the acute effects of pesticides on the range of freshwater fish throughout the U.S., two representative species are generally tested [US Code of Federal Regulations]. While the test species have been selected to represent sensitive species, and to include differences in habitat (coldwater vs. warmwater species), the uncertainties inherent in extrapolating these results to all freshwater fish of varying life stages and susceptibilities, throughout a wide geographic and climatic range, are great. When evaluations of the variability in response to chemicals have been made across species, the differences in acute mortality for pesticides across freshwater fish species may range over several orders of magnitude, even under controlled laboratory conditions (e.g., same temperature, feeding regime, etc.) [Mayer and Ellersieck, 1986]. In evaluating the acute effects of pesticides on invertebrates, only one species is generally tested. These are simple examples of the typical paucity of data available to risk assessors who are required to make broad assessments of risks for the large number of fish and invertebrate species that might be adversely affected by acute chemical exposures. There is a similar lack of data for understanding potential acute effects of chemicals on salt water fish, birds, and other wildlife, as well as potential chronic effects on species. The development of reliable methods for extrapolating toxicity information from test species to those that are of concern but cannot be directly tested is necessary. As in human health assessments, an important issue is determining whether the MOA in the test species is feasible for other species under consideration (i.e., whether the target genes are conserved and operative across species). Genomics data that show little or no similarity in key genes or patterns of gene expression between species would indicate interspecies differences and support a possible conclusion of nonrelevance to species of concern in the environment. Conversely, data showing good agreement in key genes or expression patterns between species would provide higher confidence in the relevance of the findings for species of concern. Similarly, interspecies comparison of pharmacodynamic responses enhanced by the use of genomics data could be used to define toxicological pathways in a quantitative sense. This information could be compared across species by the choice of appropriate molecular markers. Gene expression profiling is one approach that may be useful in evaluating responses to a specific environmental chemical or mixture in test species relative to species of concern in the environment. Use of genomics tools for the development of quantifiable molecular markers should significantly enhance species-species extrapolations and may reduce the current reliance on the application of uncertainty factors.

Regarding exposure assessment of ecological species, genomics technologies are likely to lead to the development of sensitive and informative biomarkers of exposure that can be used in exposure assessments, particularly in the evaluation of chemical exposures. Current methods rely on residue analyses or modeled scenarios and a few well-documented biomarkers of exposure (e.g., CYP1A, metallothionein). Molecular techniques, such as the use of microarrays, are likely to provide tools for documenting actual exposures to ecological species of concern using identified biomarkers for which there is a good understanding of the relationship of the level of biomarker to the level of exposure. Genomics may also aid in the identification of new biomarkers that can identify exposure to more stressors or more MOAs, and potentially enhance their quantification. When these biomarkers become sufficiently well-documented and understood to be able to link exposure endpoints to whole organism adverse effects, risk assessment predictions will become significantly more accurate and defensible.

Challenges Implementing Genomics Information in Risk Assessments and Regulatory Applications

The Genomics White Paper also identified several challenges regarding areas where EPA would need to further

prepare itself to receive and utilize genomics information: research, capacity (personnel and training), and technical development. For research, the challenges are to link genomics information to adverse outcomes; and interpret genomics information for risk and hazard assessment. It is important to note that significant research will be necessary to fully understand and apply genomics technologies to human health and ecological risk assessment. Two challenges were identified with respect to capacity, including human capital: recruiting individuals who possess genomics expertise, and training EPA risk assessors and managers to interpret and understand genomics data in the context of a risk assessment. The challenge of technical development, regarding the need to establish an EPA framework for analysis and acceptance of genomics information for scientific and regulatory purposes, was focused on as one of the first key tasks to be undertaken by EPA. Thus, in 2004, the EPA Genomics Technical Framework Workgroup was formed. The Genomics Technical Workgroup considered all of the "omics" technologies and applications and decided that an interim guidance document on the use of data generated by DNA microarray technology would be the most useful to the Agency and regulated community at this time, since this is currently perhaps the most established and commonly used technology in the area of genomics. The Workgroup developed a draft guidance document entitled the "Interim Guidance for Microarray-Based Assays: Data Submission, Quality, Analysis, Management, and Training Considerations" [US EPA, 2007]. The external review draft of this document is available at http://www.epa.gov/osa/spc/genomicsguidance.htm. The purpose of this draft guidance document is to provide information to the regulated community and other interested parties regarding submitting microarray data to the Agency and to provide guidance for reviewers in evaluating and utilizing such data and/or information. Microarray technologies are expected to change, but the need to ensure consistency and quality in generating, analyzing and using the data will remain a constant. As the state of the science develops, EPA plans to revisit this guidance as necessary. Included in the recommendations of the guidance document is a proposed plan for a series of genomics training modules, the work of a cross-EPA Genomics Training Workgroup. The training plan is intended to begin to address the need to build genomics capacity in the EPA through providing genomics training to risk assessors and managers.

As a result of the development of the Draft Interim Guidance for Microarray-Based Assays, EPA is undertaking a number of follow-up activities. These include (1) the implementation of the Genomics Training Plan to prepare EPA staff and managers to understand and utilize genomics information, (2) further evaluation of the options for development of an EPA-wide genomics data management solution, and (3) an effort to develop practical experience in the application of genomics data to risk assessments and regulatory applications through the development of internal case studies. It is expected that these activities will be ongoing for several years, and that future activities will build upon these.

ACKNOWLEDGMENTS

The authors thank the members of the Genomics Task Force and Genomics Technical Framework and Training Workgroups for their work in moving forward EPA genomics activities, and in the generation of the documents described herein.

REFERENCES

Mayer FL, Ellersieck, MR 1986. Manual of acute toxicity: interpretation and data base for 410 chemicals and 66 species of freshwater animals, United States Department of the Interior, U.S. Fish and Wildlife Service, Resource Publication 160, 1986. Available at http://www.cerc.usgs.gov/data/acute/acute.html

U.S. Environmental Protection Agency (US EPA), Science Policy Council. 2007. External review draft of Interim guidance for microarray-based assays: Data submission, quality, analysis, management, and training considerations. EPA 100/B-06/003. Draft currently available at: www.epa.gov/osa/genomicsguidance.html

U.S. Environmental Protection Agency (US EPA), Science Policy Council. 2004. Potential implications of genomics for regulatory and risk assessment applications at EPA. EPA 100/B-04/002. Available at: www.epa.gov/osa/genomics.html

U.S. Environmental Protection Agency (US EPA), Science Policy Council. 2002. Interim policy on genomics. Available at http://www.epa.gov/osa/spc/genomics.html

U.S. Code of Federal Regulations (40 CFR Part 158.490, 2003). Protection of environment, environmental protection agency, data requirements for registration, wildlife and aquatic organism data requirements. Available at http://www.gpoaccess.gov/cfr/index.html