EPA/600/R-04/039F
March 2006

Summary of the NCEA Colloquium on Current Use and Future Needs of Genomics in Ecological and Human Health Risk Assessment

Prepared by

Rebecca Klaper
AAAS Fellow 2002-2004
and Susan Euling
NCEA, EPA

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

NOTICE

This document is a general record of discussions during the workshop. The document captures the main points and highlights of the discussions and may include brief summaries of work group sessions. It is not a complete record of all details discussed, nor does it interpret or elaborate upon matters that were incomplete or unclear. Statements represent the individual views of the workshop participants; except as specifically noted, none of the statements represent analyses by or positions of the EPA.

Preferred Citation:
U.S. Environmental Protection Agency (EPA). (2006) Summary of the NCEA colloquium on current use and future needs of genomics in ecological and human health risk assessment. National Center for Environmental Assessment, Washington, DC; EPA/600/R-04/039F. Available online at http://www.epa.gov/ncea.

CONTENTS

1. EXECUTIVE SUMMARY

Many federal agencies recognize the great potential of genomics technologies to change the way in which human health and environmental exposure and effects are measured. It is anticipated that the use of genomics technologies may improve risk assessment by providing more sensitive measures of toxic agent-induced changes in the physiology of various organisms, including humans. It is also anticipated that data from genomic studies may identify genes that cause susceptibility and biomarkers of exposure and/or effect and provide information to extrapolate across species. Although some immediate applications of genomics have been defined (e.g., as biomarkers of disease), more work is needed to determine how this information will be used in risk assessment.

In 2002, the U.S. Environmental Protection Agency's (EPA's, or the Agency's) Science Policy Council (SPC) developed the Interim Policy on Genomics, allowing genomics to be used on a case-by-case basis in a weight-of-evidence approach for risk assessments (U.S. EPA, 2002). Genomics technologies currently need further refinement (e.g., reduced experimental variability), development, and validation before data from these experiments can be used in risk assessment. Even when these issues have been addressed, it is unclear how genomics data will affect individual risk assessments and whether the current risk assessment process will accommodate the integration of genomics data.

Because one role of the National Center of Environmental Assessment (NCEA) is to develop and improve EPA risk assessment methods, NCEA held a colloquium to provide a forum to assess current thinking about the use of genomics in risk assessment (e.g., how this technology will be applicable to risk assessors) and to determine the needs of the EPA program offices and regions that NCEA serves. NCEA defined two goals for the colloquium: (1) to identify how genomics data may improve risk assessment, and (2) to identify the current and future needs (e.g., tools, data, case studies) of the EPA program offices and regions in the area of genomics and risk assessment.

Colloquium participants consisted of scientists, risk assessors, and managers from various EPA offices and laboratories, including NCEA; the Office of Water; the Office of Prevention, Pesticides and Toxic Substances; EPA regional offices 2 and 4; the National Health and Environmental Effects Laboratory; the National Exposure Research Laboratory; the Office of the Administrator; and the Office of the Chief Financial Officer.

In preparing for and developing the colloquium, information was gathered from several EPA offices to gauge the understanding of genomics technologies and how these types of data could be used–currently and in the future–in risk assessments. The responses to questions indicated that many offices recognized the future importance of genomics in risk assessment in general.

The colloquium was designed to initiate discussions among risk assessors, managers, and scientists developing and using genomic technologies. Participants recognized the need for future, extended interaction between laboratory scientists and risk assessors, so that experiments are designed to produce data in a form that will be useful for risk assessment purposes.

An overall conclusion of the colloquium discussions was that genomics data will most likely play a role in several aspects of the risk assessment process, including hazard identification, defining mode(s) and mechanism(s) of toxicity, identification of genetic susceptibilities, and prioritization for screening and testing of environmental agents. However, participants considered it unlikely that gene expression data will be used as the sole indicator of an adverse effect. Rather, such data will be used in conjunction with in vivo endpoints. For "omics" (defined as genomics, proteomics, and metabonomics) data to be useful in risk assessment, several issues will need to be addressed, including (1) validation of methodology, including data analysis; (2) development of interpretation tools for risk assessors; (3) development of criteria for cross-species extrapolation from model organisms to humans; (4) linkage of traditional in vivo endpoints to genomics data; (5) development of a method for communicating this information both within and outside the Agency; and (6) development of criteria for the inclusion of genomics data in risk assessment.

2. INTRODUCTION

2.1. COLLOQUIUM PURPOSE

A colloquium entitled "Current Use and Future Needs of Genomics in Ecological and Human Risk Assessment" was held on May 8, 2003, in Alexandria, VA, to discuss how data from current genomics technologies and their future refinements (e.g., reduced experimental variability) could be used in risk assessment. The overall goal of the colloquium was to provide Agency scientists, researchers who are using genomics, and risk assessors an opportunity to

share perspectives, to discuss how genomics may improve risk assessment, and to identify current needs.

Because the colloquium represented a scoping step in identifying future needs in the area of genomics data in risk assessment, this summary report of the colloquium was not peer-reviewed.

2.2. INITIAL INVESTIGATION OF KNOWLEDGE AND USE OF GENOMICS IN RISK ASSESSMENT

In preparing for and developing this colloquium, Rebecca Klaper, an American Association for the Advancement of Science (AAAS) fellow at NCEA, requested responses to seven questions sent via e-mail to several EPA program and regional office contacts. The questions were designed to gauge an office staff's level of understanding of genomics and to determine whether they had received or were currently using genomics data and whether they had discussed how genomics data might be used in their health or risk assessments.

Overall, the responses to the questions indicated that many offices recognize the future importance of genomics for their needs and for risk assessment in general. However, only a few staff members in each office had a reasonable comprehension of the technologies and types of data that will be generated and the current limitations of the technology. Those who responded indicated that they believe genomics will eventually contribute to risk assessment through a better understanding of the mechanisms and/or modes of chemical toxicity, the shape of dose-response curves for many pollutants, the basis for extrapolations from model organisms to species of interest, identification of susceptible populations, and estimates of uncertainty factors. At the time of the colloquium, none of the program offices had received microarray data to support a risk assessment or decision; however, genomics data, in the form of single gene expression changes, as well as protein and genetic biomarker data, had been submitted in the past. All respondents expressed an interest in having genomics training provided to their office. The questions and the peer consultant's answers and comments are included in Appendix A.

The EPA Science Policy Council's definition of genomics was used at the colloquium and is used in this document: Genomics is the study of all the genes of a cell, or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) level (U.S. EPA, 2002). By extension, toxicogenomics is defined as the study of gene expression (mRNA and/or protein products) changes after exposure to a toxic agent.

2.3. COLLOQUIUM PARTICIPANTS

The colloquium was an internal EPA meeting designed to bring together risk assessors, scientists, and managers at the EPA program offices and regions and laboratory scientists within EPA's Office of Research and Development who are currently developing or using genomics technologies. Forty-two participants attended the colloquium, including scientists and risk assessors from many different EPA offices. Their areas of expertise spanned a broad spectrum of knowledge and agency understanding. The colloquium participants are listed in Appendix B.

2.4. COLLOQUIUM FORMAT AND SCOPE

Presentations in the morning session included discussions on how genomics data might be used in risk assessment and EPA's policy on their use (see Appendices C and D). Each presentation was followed by a question-and-answer period.

For the afternoon session, participants were divided into four breakout groups to discuss the implications of genomics technologies in four specific risk assessment areas: (1) ecological risk assessment, (2) human health risk assessment, (3) identification of sensitive or susceptible subpopulations, and (4) screening and prioritization of chemicals and microbes. The breakout group discussions focused on charge questions designed by the conference organizers and session co-chairs. Summaries of each breakout group responses to the charge questions are presented in Chapter 3.

Topics discussed included the potential use of genomics data in risk assessments and how to make these data useful from the risk assessor's perspective. The following questions were addressed: (1) If we assume that issues such as standardization of genomics techniques have been solved, then how will this information be used in risk assessment in the future? (2) What genomics data set format will be most useful to risk assessors? (3) How can experiments be designed that will provide the most useful information for risk assessment?

3. BREAKOUT GROUP DISCUSSIONS: RESPONSES TO CHARGE QUESTIONS

3.1. ECOLOGICAL RISK ASSESSMENT BREAKOUT GROUP

Moderators: Bob Frederick, Sig Degitz

Participants: Rebecca Klaper, Tala Henry, Greg Toth, Ann Miracle, Thomas Baugh, Michael Brody, Greg Susanke

Question 1. Given that it is critical to establish a link between gene expression data and an endpoint of concern, what specific information or links would be needed to use toxicogenomics data in risk assessment?

Response:

Ecological risk assessment focuses on the population rather than the individual, so a primary goal for ecological risk assessment will be to extend genomics data on exposure and effects from the individual level to the population level. Ecological risk assessment, by its nature, involves determining exposure and effects for many different organisms in an environment rather than for just one, as in human health risk assessment. Because it is impossible to determine the exposure and effects for every species in an ecosystem, species extrapolation is a key issue. Therefore, it is necessary to determine the genomic (i.e., the global gene expression profile) homologies and similarities in biochemical mechanisms and metabolism among species to be able to extend genomic technologies developed for one species to another. For a given chemical, it would be helpful to understand the degree of cross-species conservation among genes whose expression pattern has changed after chemical exposure.

How can genomics be used in assessing population effects in general? Genetic diversity is critical to sustaining populations. Thus, genomics will likely provide insight into the role that genetic diversity plays in sustainability. Currently, it is difficult to gain information for a wildlife species using genomics tools unless the species of interest is genetically well-defined, which is rare. Researchers are defining stressor-response relationships for markers known to be relevant to population viability (e.g., survival, development, fitness). When these relationships have been defined, then links between genomics changes (e.g., gene expression patterns of response) and the response can be assessed. For species within an ecosystem, genomics data

from a well-defined species may be used to predict physiologic responses in a less well-defined but related species. Eventually, genomics will provide a sensitive means to directly compare diverse species within an ecosystem.

Patterns of gene expression will likely provide new and more specific indicators of exposure or effects. It is necessary to define what genomic changes (e.g., in what genes and at what level of change) are relevant to adverse outcomes. The amount and type of information needed to link gene expression to an endpoint of interest will need to be determined for each scenario.

Case studies demonstrating the linkage (i.e., proof of concept) between gene expression and adverse outcomes would be valuable. For example, it would be useful to begin by linking a well-defined stressor to a genomic response to a known adverse outcome. As one example, there are data linking estrogen exposure, vitellogenin gene expression, and male feminization effects. In addition, it will be important to link genomic changes to toxicity pathways and use this information to inform mechanism or mode of action (MOA). Genomic endpoint information will not be used in isolation but may be used to inform other, higher-level effects.

Question 2. What are the current limitations of the technology for use in ecological risk assessment? Will these be overcome in the near term (less than 5 years) or in the long term (5 or more years)?

Response:

Currently, information connecting genomics data and effects data is lacking at the individual and the population levels. In order for genomics to become a viable tool for ecological risk assessment, genomics data must be developed for species that represent organisms of ecological interest (for chemical testing and field work), and data will need to be extrapolated to population- and community-level effects.

For ecological risk assessment, there is a need for genomic information for ecologically relevant species and resources to support this need. At present, there are genomics data for a few species relevant to ecological risk assessment, including the fathead minnow (*Pimephales promelas*), the African clawed frog (*Xenopus laevis*), daphnia (*Daphnia pulex*), zebrafish (*Danio rerio*), and Japanese medaka (*Oryzias latipes*). However, genomics information is lacking for

other ecologically relevant species used in chemical testing by many of the program offices, including the Office of Water and the Office of Prevention, Pesticides and Toxic Substances.

Technical issues for genomics technologies include an inadequate reproducibility within and across laboratories, expression level variability, and the ability of genomics data to be quantitative. For this technology to be useful in the near future, further validation of the techniques is needed. Reproducibility of data will improve as the understanding of the techniques and experimental variables improves, and this is currently being addressed by several studies.

Another limitation is the lack of interaction between scientists using genomics and risk assessment scientists, contributing to roadblocks in use and acceptance of the technology in risk assessment. In the short term, the Agency is unlikely to use genomics data for quantitative aspects of risk assessment. However, the data may be used to inform qualitative aspects of risk assessment, including mode or mechanism of action or exposure.

3.2. HUMAN HEALTH RISK ASSESSMENT BREAKOUT GROUP

Moderators: Vicki Dellarco, Ines Pagan

Participants: Linda Birnbaum, David Bussard, Chao Chen, David Dix, Karen Hamernik, Oscar Hernandez, Robert McGaughy, Julian Preston, Vickie Wilson

Question 1. Where in the overall risk assessment process (e.g., hazard identification, dose-response, exposure assessment, risk characterization) do you think omics data are more likely to play an important role? How can the data from omics be potentially used in risk assessment? Consider the mechanism of toxicity as well as treatment conditions (e.g., route, duration, magnitude of exposure) that are important for expression of the toxic effect.

Response:

Several areas were mentioned, including the following:

- Identifying hazards.
- Defining the type of toxicity of various chemicals (e.g., genotoxic versus hepatotoxic chemicals).
- Acquiring MOA and mechanism of action information of toxicants.

- Using gene expression patterns in the future to prioritize chemicals for screening and testing.

Question 2. To use omics data in human health risk assessment, what issues need to be addressed (e.g., handling the breadth and scope of data, interpreting biological and statistical meaning, training Agency risk assessors)? What criteria should be considered for those data to be useful to risk assessors?

Response:

It is unlikely that gene expression will be used as the sole indicator of an adverse effect, but it will be used in conjunction with other endpoints. To use omics data in human health risk assessment, issues that need to be addressed include the following:

- Validation of methods, including validation of data analysis methods such as Minimum Information about a Microarray Experiment (MIAME) (see www.mged.org/Workgroups/MIAME/miame.html for MIAME=compliant data/study methods).
- Development of interpretation tools, including computer software, statistics, and bioinformatics tools for risk assessors.
- Development of criteria for use of omics data (e.g., criteria for extrapolation of information for model organisms to humans).
- Determination of whether there is a link between histopathology data and traditional endpoints of toxicity to omics data.
- Identification of sentinel genes (i.e., biomarkers of effect) that are good predictors of toxic response. Case studies to serve as examples of the use of omics in risk assessment and development of "lessons learned" across agencies and within EPA (across offices). Recommendations for case studies included
    1. Use a simple case study with few confounding factors as a proof of principle.
    2. Use existing research from the Office of Research and Development to develop "lessons learned" and research needs; then develop a second phase of research.
    3. In the end, develop a quality assurance filter (i.e., a practice set of guidelines/standards/criteria and guidance for reviewers to interpret data).
- Statistical approaches to analyze genomics data for use within risk assessment.
- Overall guidance for methods to use this type of information in risk assessments.

- Risk communication issues: How can genomics data be translated into information that is easy to understand for the general public and applicable for a regional assessment?
- EPA involvement in partnerships to address the preceding issues and to develop a set of guidelines that will actively support the development of both criteria (see below) and tools to enable EPA risk assessors to display and analyze omics data.

Criteria that need to be developed include

- A framework for use of genomics data in risk assessment that is similar in scope to other EPA frameworks, including the framework for evaluating a hypothesized carcinogenic MOA within the guidelines for carcinogen risk assessment (U.S. EPA, 2005); the framework for human health risk assessment (U.S. EPA, 1998); the framework for application of the toxicity equivalence methodology for polychlorinated dioxins, furans, and biphenyls in ecological risk assessment (U.S. EPA, 2003a); and the framework for cumulative risk assessment (U.S. EPA, 2003b). As for any new technologies, test validation needs to be done with genomics, including several levels of quality assurance (QA) and quality control (QC).
- For dose-response data, the level of exposure that will induce a genomics response and determination of what that response indicates about the effects of that chemical on the organism are needed. Genomics data currently provide only a "snapshot" in time; thus, criteria will need to be developed as to how many snapshots, and at what time periods and intervals, are needed to provide the data necessary to link an exposure to an effect.
- Other questions that were raised included
    1. Can biomarkers be found using omics?
    2. What is the state of metabonomics technologies?
    3. How will metabonomics data be used?
    4. Can omics be an appropriate tool for the identification of surrogate tissues to test toxicity endpoints? If so, what types of cells are needed to identify a specific response?
- To use these types of data, genomics data need to be linked to an MOA that is relevant to humans in the proper time course and duration, and the data need to be accurate in extrapolations from high to low doses. In addition, research efforts need to establish correlations between omic response and adverse effect. It will be critical to determine the normal, or unperturbed, biological variability to establish the gene expression patterns for the normal versus the disease or toxic response state (i.e., validation).

3.3. RISK ASSESSMENTS OF SUSCEPTIBLE/SENSITIVE POPULATIONS (WILDLIFE AND HUMANS) BREAKOUT GROUP

Moderators: Margaret Chu, Les Touart

Participants: Ross Highsmith, Elizabeth Mendez, Marian Olsen, Brenda Percovich Foos, Chris Saint, Bob Sonawane, Ravi Subramaniam, Larry Valcovic, Vanessa Vu

The group recognized that the definitions for susceptibility, sensitivity, and omics technologies could affect the responses to the questions addressed. In general, the group thinks that omics data have great potential for identifying susceptible and/or sensitive individuals and/or species. Currently, omics data are being applied in clinical medicine. For example, broadly defined genomic technologies are used clinically in determining susceptibility in complex diseases such as cancer. As omics technologies develop, they may be used in determining susceptibility to, and reducing uncertainty in, assessing environmental and health risk assessment.

This summary should be viewed as the breakout group's discussion of how current omics has or can be applied to ecological and human health risk assessments only, not to other areas of biomedical applications.

Question 1. How have omics been applied to identify sensitive/susceptible subpopulations and/or species?

Response:

Overall, very few examples have been identified where genomics data have been used to identify susceptible populations. Genomic techniques are currently used to define polymorphisms in humans (mostly through animal models for humans, such as those for mice and rats) by looking for genetic variation associated with diseases (e.g., genes associated with diabetes and cancer). Identification of these polymorphisms is currently aiding in the development of drugs to counteract the effects of genetic susceptibility or in the development of alternative therapies for susceptible genotypes. Genomics is being used to improve the effectiveness and specificity of pharmaceuticals by directing the action to a molecular target known to play a role in susceptibility. However, to use this information in risk assessment, the

risks and levels of susceptibility will need to be quantitative, whereas currently they are merely a qualitative identification of genes.

Genomic technologies at present are not reproducibly quantitative. In applying these new tools, it is important to understand which genes are being affected and by what exposures. These tools offer promise in screening for patterns of gene expression. Understanding the underlying disease mechanisms more fully will allow investigation and identification of relevant gene patterns. To elucidate which genes cause susceptibility, one can look at patterns to see whether connections appear. Different genes are turned on or off at different life stages, and this complicates interpretation. Many genes activated in the cancer process, for example, are very active in early development but not in mature individuals. For risk assessment, it is expected that the critical pathways can be identified by comparing treated to untreated states.

Question 2. What types or combinations of omics technologies are most likely to have the greatest impact on developing biomarkers of susceptibility?

Response:

To date, the science is not at the point where a recommendation can be made as to the type of technology that will have the greatest impact on development of biomarkers of susceptibility. However, compared with conventional biomarkers, omic-based measures may be more advantageous. In the best case, they could provide sequential connections between the different levels of physiology, from the gene, to gene transcription, to protein formation, and to physiological function. Genomics may provide a link between a specific biomarker and the cause of a particular susceptibility, and subsequently the biomarker may be an mRNA, protein, or metabolite. The type of genomics techniques that will be the most useful will be determined by which marker provides the appropriate information.

A critical question that arises for the application of omics as biomarkers is how to effectively deal with the inherent increased sensitivity of the technology. Changes in gene expression may not necessarily translate into an adverse consequence, but understanding the linkage is important. Perhaps the greatest benefit is in identifying, under certain conditions, which populations or subpopulations are the most susceptible.

Question 3. How can omics help reduce uncertainty?

Response:

In the future, genomics will be important for interspecies extrapolations, helping to determine the physiological relationship among species and, therefore, the extent of uncertainty when extrapolating results of toxicological tests from model organisms to organisms of interest. Genomics will likely be more useful for identifying patterns of response (i.e., qualitatively) than for defining a dose-response curve. This technology can be used to identify critical gene(s) linked to an effect. Genomics may also be useful for determining the factors associated with low-dose responses to compounds, providing evidence to support or reject hypotheses surrounding hormetic effects. Genomics is a powerful tool, but to be useful it has to be applied to a reasonable and tractable question.

3.4. HAZARD IDENTIFICATION: SCREENING AND PRIORITIZATION BREAKOUT GROUP

Moderators: Susan Euling, Jennifer Seed

Participants: Nancy McCarroll, Cynthia Nolt-Helms, Robin Oshiro, Devon Payne-Sturges, Phil Sayre, Rita Schoeny, Deborah Segal

Question 1a. What types of genomics technologies might be useful for chemical screening purposes in the future (e.g., for the Endocrine Disruptor Screening Program [EDSP], pesticide inerts, Toxic Substances Control Act's High Production Volume [HPV] chemicals, and Office of Water's Candidate Contaminant List [CCL])?

Response:

A large number of chemicals need to be prioritized for testing in a number of screening programs, such as the EDSP (http://www.epa.gov/scipoly/oscpendo/edspoverview/index.htm). Prioritization is a separate, complex exercise and consists of a number of different criteria. For example, prioritization could be based on a chemical's MOA or quantitative structure-activity relationship (QSAR). Genomics technologies could help to determine the MOA of a chemical and therefore assist in prioritization. Genomics technologies could also be used to gain information about structure-activity relationships and could then be incorporated into the dataset

for QSAR modeling. If a gene expression profile has been linked to an adverse effect after chemical exposure, then chemicals could be screened by their gene expression profile. Chemicals without this link to an adverse effect could be assigned a lower toxicity testing priority. In addition, genomics could be used to develop a gene expression fingerprint for the response of an organism to chemical mixtures.

Genomics technologies need to be validated before routine use in a screening program. Specifically, microarray or proteomic data need to be linked to an adverse effect (in the case of human health risk assessment) or endpoint of concern (in the case of ecological risk assessment). Verifying this link is complicated by a number of factors, including high inter-experiment and intra-experiment variability of response, that have made it difficult to replicate genomics experiments in some cases. The group members thought that microarray analysis is probably the furthest along for liver-mediated toxicity and estrogen receptor-mediated toxicity.

Microarrays may not be the best tool for chemical screening purposes if proven enzyme screening methods already exist (e.g., cholinesterase activity). Eventually, this technology may lead to more rapid screens and to a decrease in the use of animals in testing, but initially, as the techniques are being developed and validated, an increase in cost and animal usage may occur. In the future, genomics will most likely provide a more sensitive and specific means with which to measure effects from chemicals than current methods. Proteomics, in particular, may be the most useful technology for chemical screening because proteins are, in most known cases, the actual functional component within the cell and organism (e.g., small RNAs have been found to be the functional molecule for some genes [Lee et al., 1993]). Therefore, measuring the global proteomic response may be the optimum indicator of the physiological response of the organism after chemical exposure.

Question 1b. How do these genomics technologies compare with other screening technologies currently used? Identify the strengths and weaknesses.

Response:

Current screening technologies include the following:

- QSARs, which are currently limited to use in screening for ecotoxicology and mutagenic carcinogens.
- Various in vitro tests (e.g., receptor binding assays, Ames test).

- Short-term in vivo assays that focus on a specific response (e.g., uterotrophic).
- Short-term in vivo assays that have a broad focus (e.g., The Organisation for Economic Co-operation and Development [OECD] combined repeat/reproduction, acute toxicity study).

The group discussed the question, "What are the requirements for a screening assay?" The group recommended that the assay must be:

1. Sensitive (erring on the side of false positives so that problem chemicals are identified).
2. Somewhat specific to the endpoint of concern.
3. Fast and efficient (relative measure).
4. Inexpensive (relative measure).

Currently, genomics technologies are not as rapid, inexpensive, or efficient as the currently used screening technologies. In some cases, they may be more sensitive and specific than other assays, but some current screening assays are quite selective and specific to the MOA of interest and rapid (e.g., cholinesterase activity for organophosphate pesticides). Therefore, the state of the technology is not ready for screening or prioritization because the omics technologies are currently not as rapid or inexpensive as the other currently available methods.

Question 2. What types of genomics technologies/experimental design could be useful for microbial screening for drinking water quality?

Response:

Current microbial screening techniques include culturing microorganisms from water samples and using polymerase chain reaction (PCR) techniques to examine a water sample for a genetic component of a bacterium of interest (e.g., 16sRNA). Culturing is the gold standard for identification and quantification of bacteria, but the technique does have limitations. It may be slow and the cultures may contain both virulent and nonvirulent forms of the bacteria of interest. Additionally, under certain conditions, viable bacteria may be present but nonculture able. For most viruses, there are few or no culture techniques available.

PCR techniques rely on known differences in segments of genetic material within each strain to identify the presence of a species of interest. Microarray technologies, like PCR

techniques, rely on identifying genetic components of known pathogens or the virulence factors within those pathogens. Microarray technology has an added benefit in that it will also screen for emerging pathogens that happen to have the same virulence factor. However, both of these techniques are limited because they may not identify a specific pathogen due to genetic similarities between strains, leading to false positive results and small mutations that do not change the virulence but do change the rate of detection. In addition, Genbank (http://ncbi.nlm.nih.gov/Genbank/index.html), the repository for sequence information on many species, including bacterial pathogens, is subject to scientific and clerical errors. Sequenced samples may be impure or slightly inaccurate, which could lead to a greater variability when developing a diagnostic test. Finally, water samples need to be large enough to detect the bacteria, and samples may need to be amplified using PCR before detection methods can be effective.

Proteomics may be useful in the future but may not be as sensitive as genomics for water sampling because the production of proteins may not have occurred by the time of sampling. EPA's regulations at 40 CFR 136.4, 136.5, and 40 CFR 141.27 allow one to apply for permission by EPA to use an alternate test procedure instead of an EPA-approved reference method. The Alternative Test Procedure Program (http://www.epa.gov/waterscience/methods) in the Office of Water has received inquiries regarding submissions for the use of genetic techniques for detecting microbes in recreational waters, but these data cannot be accepted yet because there are no approved methods for genetic-based tests. Such methods cannot be accepted for review until methods are published in 40 CFR Part 136 (ambient water, wastewater, or biosolids methods) or 40 CFR Part 141 (drinking water methods) as approved methods.

Question 3. What are the current limitations of the technology for use in screening assays?

Response:

The following limitations to the use of omics in chemical screening were noted:

- Linkage to adverse effect of concern is needed. There are not many cases where this has been established.
- Sensitivity of the technologies (erring on the side of false positives so problem chemicals do not slip through) is not well established.
- Specificity/selectivity is not established.

- Speed/efficiency.
- Cost.
- Animal usage compared with some in vitro screens.

The following limitations to the use of omics in microbial screening were noted:

- Sensitivity is not well established. Microarray analysis may not identify a specific pathogen owing to similarities in strains leading to false positive results and small mutations that do not change the virulence but do change the rate of detection.
- Specificity/selectivity is not established.
- Cost.
- Currently a need for a large water sample using PCR.
- Procedures are less well established for viruses.
- Microarrays may be preferable to proteomics.

4. COLLOQUIUM DISCUSSION CONCLUSIONS

In the future, genomics data could play a significant role in several aspects of the risk assessment process, including hazard identification, definition of mode(s) and mechanism(s) of toxicity, identification of genetic susceptibilities, and prioritization for screening and testing of environmental chemicals. However, it is unlikely that gene expression data will be used as the sole indicator of an adverse effect; rather, such data will be used in conjunction with in vivo endpoints. To use omics data in risk assessment, several issues will need to be addressed, including

1. Validation of methodology and data analysis, including several levels of QA/QC (similar to validation of other new technologies).
2. Development of interpretation tools for risk assessors.
3. Development of criteria for cross-species extrapolation from model organisms to humans.
4. Linkage of traditionally used in vivo endpoints to genomics data.
5. Development of a method to communicate this information both within and outside the Agency.
6. Development of criteria for the inclusion of genomics data in risk assessment (similar to criteria that have been developed for new technologies in the past).

It will be important to develop case studies to demonstrate linkage (i.e., proof of concept) between exposure, gene expression, and adverse outcomes. Studies need to start by characterizing and supporting the links between a well-defined stressor (e.g., toxic agent) and a well-established environmental effect. For example, one can build a case for the links between estrogen exposure, vitellogenin protein expression, and male feminization effects in fish. Then, gene or protein expression data will need to be assessed for whether they are linked to exposure to the stressor as well as the response or effect in the organism or population.

Eventually, genomics will also be useful for screening and testing, but there will be limitations and issues similar to those described above. Genomics technologies could be used to gain information about QSARs and could then be incorporated into the dataset for QSAR modeling. This technology may lead to more rapid screening assays and to decreased use of animals in testing. But initially, as the techniques are being developed and validated, animal usage may increase.

In the future, genomics will most likely provide a more sensitive and specific means to measure effects after chemical exposure than those offered by current methods. Proteomics in particular may be the most useful technology for chemical screening because proteins are typically the functional component within the cell and organism. Currently, the state of the technology has not been optimized for screening and/or prioritization purposes because genomics is not as consistent in response (i.e., high variability), efficient, or low in cost as some of the other currently available methods. For example, cholinesterase activity assays are rapid and specific to the MOA of interest.

Genomics techniques may be particularly useful for microbial screening because culture techniques are limited. For example, under certain conditions viable bacteria may be present but nonculture able. Thus, gene expression products of the species of interest may be a method to circumvent this problem. An added benefit to the microarray approach is that it can simultaneously screen for emerging pathogens that happen to have a similar virulence factor. The Alternative Test Procedure Program has received inquiries regarding submissions for the use of genetic techniques for detecting microbes, but these data cannot be accepted until methods for genetic-based tests have been approved. A current Agency focus is to begin developing these genetic methods.

An additional point was raised for the incorporation of genomics into ecological risk assessment (ERA). ERA focuses on the population rather than on the individual, so one of the

primary goals will be to extend genomics data at the individual level to the population level with regard to both detection and effects. Thus, exposure and effects for many different species in an environment need to be determined. Because it is impossible to determine the exposure and effects for every species in an ecosystem, species extrapolation is important. Therefore, it is necessary to determine the genomic (i.e., global gene expression profiles) homologies, degree of conservation of genes whose expression is altered, and similarities in biochemical mechanisms and metabolism among species to be able to extend genomic technologies developed for one species to another.

This colloquium and future EPA activities will provide opportunities for risk assessors to learn more about this field and to exchange ideas about the use of genomics in risk assessment. Participants recognized the need for further interactions among laboratory scientists and risk assessors to inform the development and possible uses of this technology for risk assessment. It is clear that risk assessors must be included in discussions of genomics within the Agency so that data will be designed and presented in a form that will be useful for risk assessment purposes.

REFERENCES

Lee, RC; Feinbaum, RL; Ambros, V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5):843–54.

U.S. EPA (Environmental Protection Agency). (1998) Summary of the U.S. EPA colloquium on a framework for human health risk assessment (Volume 2, 1998). Risk Assessment Forum, Washington, DC.

U.S. EPA (Environmental Protection Agency). (2002) Interim policy on genomics. Prepared by the Science Policy Council, Washington, DC. Available online at http://www.epa.gov/osa/spc/genomics.htm.

U.S. EPA (Environmental Protection Agency). (2003a) Framework for application of the toxicity equivalence methodology for polychlorinated dioxins, furans and biphenyls in ecological risk assessment, external review draft. Risk Assessment Forum, Washington, DC. EPA/630/P-03/002A.

U.S. EPA (Environmental Protection Agency). (2003b) Framework for cumulative risk assessment. Office of Research and Development, National Center for Environmental Assessment, Washington, DC. EPA/600/P-02/001F.

U.S. EPA (Environmental Protection Agency). (2005) Guidelines for carcinogen risk assessment. Federal Register 70(66)17765-17817. Available online at http://www.epa.gov/cancerguidelines.

APPENDIX A: INFORMATION GATHERED PRIOR TO THE COLLOQUIUM ABOUT THE USE AND KNOWLEDGE OF GENOMICS ACROSS EPA OFFICES

Rebecca Klaper of the National Environmental Assessment (NCEA) designed seven questions to assess the level of understanding and knowledge of genomics technologies as they relate to risk assessment. Various EPA office contacts were asked in January of 2003 to comment on each question and were encouraged to get input and have discussions about the questions from others in their offices. In addition, contacts were asked to inform their office about the upcoming colloquium.

The peer consultants, listed below, responded to the questions.

NATIONAL CENTER FOR ENVIRONMENTAL ASSESSMENT
Susan Euling
Bob Frederick
Inez Pagan
Bob Sonawane
Cindy Sonich-Mullin
Michel Stevens

OFFICE OF WATER
Joyce Donohue
Jafrul Hasan
Tala Henry
Tony Maciorowski
Edward Ohanian
Rita Schoeny

INTEGRATED RISK INFORMATION SYSTEM
Mike Broder
Lynn Flowers

OFFICE OF AIR
Tom Curran
Carl Mazza
Maria Pimentel

OFFICE OF POLLUTION, PREVENTION AND TOXICS
Phil Sayre
Jennifer Seed

OFFICE OF PREVENTION AND TOXIC SUBSTANCES
Gary Timm
Les Touart
Maurice Zeeman

EPA REGION 9
Bobbye Smith (RSL)

EPA REGION 4
Thomas Baugh (RSL)

OFFICE OF SOLID WASTE AND EMERGENCY REPSONSE
Lee Hoffman

Summary of Responses to Questions about Knowledge and Experience with Genomics from Peer Consultants

The questions were sent to the contacts via e-mail. The summaries of all responses received are described below. The intention was to present the range of responses received. Individual responses are not presented here.

Question 1. Do you think that people in your office understand the definition and types of data associated with genomics and proteomics?

Response:

Comments ranged from offices with some staff having a basic awareness of the technology of genomics and the potential use in risk assessment to offices reporting that no one had heard of the term "genomics." Among the offices with a basic awareness of genomics, it was commented that very few staff had a detailed understanding of the "omics" technologies.

Question 2. Are people in your office familiar with some of the limitations in analyzing this data?

Summary of Responses:

Most offices stated that they were not familiar with limitations. Among the offices that reported being familiar with limitations, the following limitations were noted:

- New, not validated technologies.
- Little data to understand the patterns of gene expression after exposure to chemicals (i.e., little toxicogenomics data).

Question 3. Would your office be most interested in data to enhance information on effects, exposure, or identifying susceptible populations?

Summary of Responses:

- Effects were mentioned most often, and the areas of dose-response, reduction of uncertainty, and risk assessment were highlighted. Exposure and susceptible populations were mentioned equally. It was noted that all three areas, effects, exposure, and susceptible population identification, were of interest.

Question 4. What are your office's training needs in the area of omics?

Summary of Responses:

The following course topics were suggested:

- Basic "Genomics 101" training course (mentioned most frequently).
- Data analysis, use in risk assessment (when this becomes available).
- Statistical analysis employed in the analysis of these data.
- Ethical implications and legal issues.
- Training for risk communicators.

Other needs mentioned:

- Consider the needs of states and tribes.
- Discussion of when these data will come to fruition.
- Statistical analysis employed in the analysis of these data.
- Case-study: Risk assessment that supports regulatory decision making.

Question 5. Has your office received any type of these data?

Summary of Responses:

All queried offices stated they had not received genomics data. However, some offices stated that they had received or used single gene expression data (the Science Policy Council's definition of genomics).

Question 6. What impact do you see this having on the work in your office?

Summary of Responses:

It was noted that genomics research has the potential to improve human and ecological risk assessments. Areas that genomics will contribute to were noted:

- Mechanisms of chemical toxicity.
- Biological interaction of chemicals and chemical mixtures.
- Signal transduction pathways.
- Induced gene expression and, therefore, development of biomarkers of human exposure.
- Understanding mechanisms for genetic damage and/or DNA repair, among other mechanisms.

It was noted that the application of this new technology could assist EPA risk assessors to:

- Characterize the shape of dose-response curve for a number of pollutants in a timely and cost-effective manner.

- Allow comparisons between animal mode species and humans (i.e., interspecies extrapolation).
- Reduce uncertainty factors through the understanding of individual susceptibility to environmental stressors.
- Incorporate genomics data into both human health and ecological risk assessments and monitoring programs. Additional comments: Exposure assays for endocrine disruptors have been the single most requested molecular biology tool from California and tribes. Genomics assays may eventually replace some of the current screens used in EPA's Endocrine Screening and Testing Program. However, with any new technology there is considerable training and new infrastructure costs.
- Analyze qualitative information on mode of action and susceptible populations (probably from genetic polymorphisms) to support rulemaking efforts. Additional comments: Such data/tools will require clear definition of programmatic problems in order to evaluate when and where such data could be incorporated into risk assessments and decision–making processes.

Question 7. Other interests/comments?

Summary of Responses:

Two comments were received:

- The Office of Pollution Prevention and Toxics (OPPT) is currently working with the Office of Research and Development laboratories to look at the feasibility of fingerprinting for certain chemical classes, and OPPT has a Cooperative agreement with International Life Sciences Institute.
- The Office of Water is using genomics data within its Chemical Contaminants List program.

APPENDIX B: COLLOQUIM PARTICIPANTS

Name	Office/Division	Breakout Group	Role(s)
Thomas Baugh	Region 4	ERA	Participant
Linda Birnbaum	NHEERL-RTP	HHRA	Participant
Michael Brody	OCFO/OPPA	ERA	Participant
David Bussard	NCEA-W	HHRA	Participant
Chao Chen	NCEA-W	S/SP	Participant
Margaret Chu	NCEA-W	S/SP	Co-chair
Sig Degitz	NHEERL-Duluth	ERA	Co-chair
Vicky Dellarco	OPPTS	HHRA	Co-chair
David Dix	NHEERL-RTP	HHRA	Participant/Speaker
Susan Euling	NCEA-W	SC	Co-chair/Organizer
Bob Fredrick	NCEA-W	ERA	Co-chair/Organizer
Karen Hamernik	OPPTS/OPPT	HHRA	Participant
Jafrul Hansan	OW/OST	ERA	Participant
Tala Henry	OW/OST	ERA	Participant
Oscar Hernandez	OPPTS/OPPT	HHRA	Participant
Ross Highsmith	NERL	S/SP	Participant
Rebecca Klaper	AAAS fellow at NCEA-W	FLOAT	Organizer/Speaker
Nancy McCarroll	OPPTS/OPP	SC	Participant
Robert McGaughy	NCEA-W	HHRA	Participant
Elizabeth Mendez	OPPTS/OPP	S/SP	Participant
Ann Miracle	NERL-Cincinnati	ERA	Participant
Cynthia Nolt-Helms	NCER	SC	Participant
Marian Olsen	Region 2	S/SP	Participant
Robin Oshiro	OW/OST	SC	Participant
Ines Pagan	NCEA-RTP	HHRA	Co-Chair
Devon Payne-Sturges	OPEI	SC	Participant

Brenda Percovich Foos	OA/OCHP	S/SP	Participant
Julian Preston	NHEERL-RTP	HHRA	Participant
Chris Saint	NCER	S/SP	Participant
Phil Sayre	OPPTS/OPPT	SC	Participant
Rita Schoeny	OW/OST	SC	Participant
Jennifer Seed	OPPTS/OPPT	SC	Co-Chair
Deborah Segal	NCER	SC	Participant
Bob Sonawane	NCEA-W	S/SP	Organizer
Ravi Subramaniam	NCEA-W	S/SP	Participant
Greg Susanke	OAA/OSP	ERA	Participant
Shirlee Tan	AAAS fellow at OPPTS/OSCP	ERA	Participant
Greg Toth	NERL-Cinci	ERA	Participant/Speaker
Les Touart	OPPTS/OSCP	S/SP	Co-Chair
Larry Valcovic	NCEA-W	S/SP	Participant
Vanessa Vu	OA/SAB	S/SP	Participant/Speaker
Vickie Wilson	NHEERL-RTP	HHRA	Participant/ Speaker

ERA, Ecological Risk Assessment; HHRA, Human Health Risk Assessment; S/SP, Risk Assessments of Susceptible/Sensitive Populations; SC, Hazard Identification: Screening and Prioritization

APPENDIX C: COLLOQUIUM AGENDA

9:00 –9:15 a.m.	David Bussard (Director, NCEA–W), Rebecca Klaper (AAAS fellow at NCEA) Welcome and Introduction to the Purpose of the Colloquium
9:15–9:50 a.m.	Vickie Wilson (NHEERL–RTP) The Nuts and Bolts of Genomics Research; Microarrays and Proteomics
9:50–10:25 a.m.	David Dix (NHEERL–RTP) Integration of Toxicogenomics and Risk Assessment: Common Modes of Action and Biomarkers
10:25–10:35 a.m.	Break
10:35–11:40 a.m.	Greg Toth (NERL–Cincinnati) Ecological Risk Assessment Example of Genomics Data
11:40 a.m.–12:15 p.m.	Vanessa Vu (OA/SAB) Development and Implementation of EPA Genomics Action Plan
12:15–1:15 p.m.	Break for Lunch
1:15–1:30 p.m.	Rebecca Klaper (AAAS fellow at NCEA) Information Gathered Regarding the Use of Genomics Data Across the EPA Offices
1:30–1:40 p.m.	Susan Euling (NCEA-W) Charge to the Breakout Groups
1:40–3:40 p.m.	Breakout Group Discussions
3:40–3:50 p.m.	Break
3:50–4:50 p.m.	Reports from Breakout Groups
4:50–5:00 p.m.	Bob Frederick (NCEA-W) Summary of the Day and Close of the Meeting

APPENDIX D: ABSTRACTS FOR SOME OF THE PRESENTATIONS

David Dix (NHEERL-RTP)
Integration of Toxicogenomics and Risk Assessment: Common Modes of Action and Biomarkers

Genomics and proteomics will greatly improve the accuracy of risk assessments by informing dose and species extrapolations, guiding cumulative assessments based on common modes of action, and identifying sensitive subpopulations. Integration of toxicogenomic data into risk assessments will require applicable genomic methods developed by regulatory agencies and their partners, knowledge of the laboratory and bioinformatic methods that affect outcomes, and the ability to evaluate the quality of genomic data. Critical questions to be addressed include the following: Can toxicogenomic data identify NOAELs and LOAELs for risk assessment purposes? Can toxicogenomics be applied to human epidemiology investigations? What problems/methods in toxicogenomic data analysis have the greatest effect on use in risk assessments? EPA should consider establishing ORD-OPPTS working groups to develop the necessary tools for risk assessors to use toxicogenomics. This proposed tool kit would include a regulatory toxicogenomics database, toxicogenomics data quality evaluation software, and prototype data sets and risk assessments centered on ORD's strengths in carcinogenesis and reproductive toxicology.

Greg Toth (NERL-Cincinnati)
Ecological Risk Assessment Example of Genomics Data

Diagnostic and prognostic risk assessments of chemical and biological stressors in aquatic ecosystems stand to be improved significantly by application of data from the omic technologies. Aquatic organisms especially offer the potential to serve as models for the linkage of exposure and effects models for the prediction of adverse outcomes all the way to the population level. EPA/ORD omics research with aquatic organisms, structured significantly by the emerging framework for computational toxicology, incorporates all of the elements of the source-to-outcome paradigm. Integration of metabonomics, proteomics, and genomics data to more completely test hypotheses has become a realistic goal for molecular ecologists in the immediate future. EPA/ORD approaches this goal with the potential for huge sequence resources, instrumentation for advanced proteomics and metabonomics, and an awareness of the complexities being revealed at the systems biology level. This presentation lays out several hypotheses to address quantitative risk assessment in this overall context.