

EPA 100/B-04/002  
December 2004



# **U.S. Environmental Protection Agency**

## **Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA**

**Prepared for the U.S. Environmental Protection Agency  
by members of the Genomics Task Force Workgroup, a group of  
EPA's Science Policy Council**

**Science Policy Council  
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 and distribution. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

---

---

# **Genomics Task Force Workgroup A Group of the Science Policy Council**

## **Workgroup Co-Chairs**

William Benson  
NHEERL  
Office of Research and Development

Kerry Dearfield  
Office of the Science Advisor

## **Workgroup Members**

Michael Brody, OCFO  
Anne Fairbrother, ORD  
Kathryn Gallagher, OSA  
Jafrul Hasan, OW  
Lee Hofmann, OSWER  
Jack Jones, ORD  
Rebecca Klaper, AAAS Fellow, ORD  
David Lattier, ORD  
Susan Lundquist, OEI  
Nancy McCarroll, OPPTS  
Elizabeth Mendez, OPPTS

Gregory Miller, OPEI  
Ines Pagan, ORD  
Maria Pimentel, OAR  
Julian Preston, ORD  
Philip Sayre, OPPTS  
Rita Schoeny, OW  
Jennifer Seed, OPPTS  
Bobbie Smith, Region 9 RSL  
Anita Street, ORD  
Richard Troast, OSWER

## **Workgroup Leads for the Science Policy Council**

Lawrence Reiter  
NHEERL  
Office of Research and Development

Vanessa Vu  
Science Advisory Board  
Office of the Administrator

---

**TABLE OF CONTENTS**

FOREWORD .....	Page vii
ACKNOWLEDGMENTS .....	Page viii
ACRONYMS .....	Page ix
EXECUTIVE SUMMARY .....	Page 1
I. Introduction .....	Page 3
A. Background .....	Page 3
B. Emerging Impacts of Genomics Technologies .....	Page 4
C. Overview of Genomic Science .....	Page 5
D. Purpose and Intent of this Document .....	Page 7
II. Regulatory Applications .....	Page 10
A. Prioritization of Contaminants (Chemicals and Microbes) and Contaminated Sites .....	Page 10
1. Introduction .....	Page 10
2. Regulatory and Voluntary Activities Potentially Affected by Genomics Information .....	Page 11
a. Representative Activities .....	Page 11
b. Additional Activities .....	Page 11
B. Monitoring .....	Page 12
1. Introduction .....	Page 12
2. Regulatory Activities Potentially Affected by Genomics Information .	Page 13
a. Representative Activities .....	Page 13
b. Additional Activities .....	Page 14
C. Reporting Provisions .....	Page 14
1. Introduction .....	Page 14
a. Reporting on Adverse Effects of Commercialized Chemicals and Pesticides .....	Page 14
b. Toxics Release Inventory Program .....	Page 15
2. Regulatory Activities Potentially Affected by Genomics Information .	Page 15
III. Risk Assessment .....	Page 17
A. Introduction .....	Page 17
B. Mode of Action .....	Page 17
1. Overview .....	Page 17
2. Risk Assessment Activities Potentially Affected by Genomics Information .....	Page 18

---

a. Representative Activities .....	Page 18
i. Improving Hazard Identification .....	Page 18
ii. Improving Dose-Response Assessment .....	Page 19
iii. Improving Extrapolations .....	Page 20
iv. Improving Exposure Assessment .....	Page 21
C. Susceptible Populations and Sensitive Life Stages .....	Page 22
1. Overview .....	Page 22
a. Susceptible Human Populations and Sensitive Life Stages ...	Page 22
b. Susceptible Wildlife Populations and Sensitive Life Stages ..	Page 23
2. Risk Assessment Activities Potentially Affected by Genomics Information .....	Page 24
a. Susceptible Human Populations and Sensitive Life Stages ...	Page 24
i. Representative Activities .....	Page 24
b. Susceptible Wildlife Populations and Sensitive Life Stages ..	Page 26
i. Representative Activities .....	Page 26
D. Mixtures .....	Page 27
1. Overview .....	Page 27
2. Risk Assessment Activities Potentially Affected by Genomics Information .....	Page 28
a. Representative Activities .....	Page 29
b. Additional Activities .....	Page 29
IV. Research Needs and Activities .....	Page 30
A. Introduction .....	Page 30
B. Research Needs and Activities for Regulatory Applications .....	Page 31
1. Prioritization .....	Page 31
2. Monitoring .....	Page 32
C. Research Needs and Activities for Risk Assessment .....	Page 34
1. Mode of Action .....	Page 34
2. Susceptible Populations and Life Stages .....	Page 35
3. Mixtures .....	Page 37
V. Challenges and Recommendations .....	Page 39
A. Research Challenges .....	Page 40
1. Linking Genomics Information to Adverse Outcomes .....	Page 40
2. Interpretation of Genomics Information for Risk Assessment .....	Page 40
3. Recommendations to Address Research Challenges .....	Page 41
B. Technical Development Challenges .....	Page 42
1. Framework for Analysis and Acceptance Criteria for Genomics Information .....	Page 42
2. Recommendations to Address Technical Development Challenges ...	Page 42
C. Capacity/Human Capital Challenges .....	Page 43
1. Applying Strategic Hiring Practices to Recruit Individuals Who	

---

Possess “Genomics Core Competencies” .....	Page 43
2. Training EPA Risk Assessors and Managers to Interpret and Understand Genomics Data in the Context of a Risk Assessment .....	Page 43
3. Recommendations to Address Capacity/Human Capital Challenges ..	Page 43
D. Summary Recommendation .....	Page 44
REFERENCES .....	Page 46
Appendix A: Interim Policy on Genomics .....	Page A-1
Appendix B: Details on Additional Activities Potentially Affected by Genomics Information .....	Page B-1
I. Prioritization .....	Page B-1
A. Program Offices .....	Page B-1
OPPT - Premanufacture Notices .....	Page B-1
OPPT - Voluntary Children’s Chemical Evaluation Program ...	Page B-1
OPPTS/OSCP - Endocrine Disruptors Screening Program .....	Page B-1
OPP - Prioritization of Pesticides and Inert Chemicals .....	Page B-1
OW - Prioritizing Streams or Wetlands for Study or Clean Up ..	Page B-1
OAR - Hazardous Air Pollutants .....	Page B-1
OSWER - Superfund .....	Page B-1
ORD- Research Planning .....	Page B-2
B. Regions, States, and Tribes .....	Page B-2
Resource Prioritization for Site Remediations and Chemical Evaluations .....	Page B-2
II. Monitoring .....	Page B-2
A. Program Offices .....	Page B-2
OPP - Exposure Monitoring .....	Page B-2
OAR/Office of Air Quality Planning and Standards (OAQPS) - Stationary Source Monitoring .....	Page B-2
OSWER - Superfund Monitoring .....	Page B-2
OSWER/OSW - RCRA-Required Monitoring .....	Page B-2
OEI, ORD - Bioindicator Development .....	Page B-3
B. Regions, States, and Tribes .....	Page B-3
State and Local Beach Closures .....	Page B-3
Air Quality Monitoring Program .....	Page B-3
Endocrine-Disruptor Monitoring .....	Page B-3
Regional Pesticide Program Decisions .....	Page B-3
III. Risk Assessment .....	Page B-4
A. Program Offices .....	Page B-4
OSWER, OAR, OW - Site Assessments/Monitoring .....	Page B-4
B. Regions, States, and Tribes .....	Page B-4
Site Assessments/Monitoring .....	Page B-4

---

Appendix C: Glossary ..... Page C-1  
    Glossary References ..... Page C-5

## FOREWORD

Genomics information has great potential to enhance assessment of risks to human health and the environment and will have significant implications for EPA's risk assessment practice and regulatory decision making. EPA's Interim Policy on Genomics (2002) states that while genomics data may be considered in the decision making process at this time, these data alone are insufficient as a basis for decisions, and will be considered for assessment purposes on a case-by-case basis only.

Following the release of the Interim Policy, at the request of the Science Policy Council (SPC), a cross-Agency Genomics Task Force Workgroup was formed. The workgroup was charged with examining the broader implications genomics is likely to have on EPA programs and policies, and with developing scenarios to describe various circumstances under which EPA might receive these data and the resulting implications for EPA policies and programs. Thus, the purpose of this document is to present exemplary applications and resultant implications of the use of genomics technologies in EPA practice for the consideration of Agency managers. Although, as the Interim Policy notes, understanding genomic responses with respect to adverse ecological and/or human health outcomes is far from established, it is important for managers to begin to consider the likely future impacts of genomics technologies on their programs.

It is clear that genomics technologies have great potential to enhance assessment of risks to human health and the environment, and the Agency must be proactive in preparing itself to address the oncoming challenges associated with interpreting and applying genomics information. It is essential for EPA to continue to collaborate with other federal agencies, academia, the regulated community, public interest groups, and other stakeholders in this endeavor in order to benefit from ongoing advances in genomics in the wider scientific and regulatory communities.

I want to acknowledge and thank the Genomics Task Force Workgroup for their efforts in assembling this document. I particularly appreciate the efforts of the SPC co-chairs, Larry Reiter and Vanessa Vu, and the workgroup co-chairs, Kerry Dearfield and Bill Benson in leading the workgroup.

It is with great pleasure that I present the *Potential Implications of Genomics for Regulatory and Risk Assessment Applications at EPA*.

Paul Gilman, Ph.D.  
EPA Science Advisor  
Chair, Science Policy Council

---



## **ACKNOWLEDGMENTS**

The Genomics Task Force Workgroup would like to acknowledge the Science Policy Council and its Steering Committee for their recommendations and contributions to this document. In particular, we would like to acknowledge the following individuals who provided comments through the internal peer review process: Roland Hemmett (Region 2), Carl Mazza (OAR), Michael Firestone (OCHP), Claudia Walters (ORD), Marian Olsen (Region 2), and Craig Annear (OGC). We would also like to acknowledge program office and regional staff who provided comments on the document to workgroup members: including Nicole Paquette (OEI), Stephen Devito (OEI), Edward Bender (OSA), and Karen Hammerstrom (ORD). We also appreciate the comments from the external peer reviewers, whose comments and suggestions enhanced the final document.

---

**ACRONYMS**

ATSDR	Agency for Toxic Substances and Disease Registry
BBDR	Biologically-Based Dose Response
BST	Bacteriological Source Tracking
CAA	Clean Air Act
CCL	Contaminant Candidate List
CERCLA	Comprehensive Environmental Response, Compensation and Liability Act
CIIT	Chemical Industry Institute of Technology
CWA	Clean Water Act
CYP	Cytochrome P-450
DNA	Deoxyribonucleic Acid
EDC	Endocrine Disrupting Chemical
EDSP	Endocrine Disruptors Screening Program
EPA	Environmental Protection Agency
ELSI	Ethical, Legal, and Social Implications
EPCRA	Emergency Planning and Community Right-to-Know Act
EUP	Experimental Use Permit
FDA	Food and Drug Administration
FIFRA	Federal Insecticide, Fungicide and Rodenticide Act
FQPA	Food Quality Protection Act
GM	Genetically Modified
HAPs	Hazardous Air Pollutants
HPV	High Production Volume
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Test Methods
ILSI	International Life Sciences Institute
IRIS	Integrated Risk Information System
LOAEL	Lowest Observed Adverse Effect Level
MOA	Mode of Action
MRA	Microbial Risk Assessment
MST	Microbial Source Tracking
NHEERL	National Health and Environmental Effects Research Laboratory
NIEHS	National Institute of Environmental Health Sciences
NPDES	National Pollutant Discharge Elimination System
NOAEL	No Observed Adverse Effect Level
OAQPS	Office of Air Quality Planning and Standards
OAR	Office of Air and Radiation
OCFO	Office of the Chief Financial Officer
OCHP	Office of Children's Health Protection
OECD	Organization for Economic Cooperation and Development
OEI	Office of Environmental Information

---

---

OGWDW	Office of Ground Water and Drinking Water
OPEI	Office of Policy, Economics and Innovation
OPP	Office of Pesticide Programs
OPPT	Office of Pollution Prevention and Toxics
OPPTS	Office of Prevention, Pesticides and Toxic Substances
ORD	Office of Research and Development
OSA	Office of the Science Advisor
OSWER	Office of Solid Waste and Emergency Response
OW	Office of Water
PBPK	Physiologically-Based Pharmacokinetic
PCR	Polymerase Chain Reaction
PMN	Premanufacture Notice
POD	Point of Departure
QSAR	Quantitative Structure Activity Relationship
RCRA	Resource Conservation and Recovery Act
RED	Reregistration Eligibility Decision
RfC	Inhalation Reference Concentration
RfD	Oral Reference Dose
RNA	Ribonucleic Acid
RT-PCR	Reverse-Transcription Polymerase Chain Reaction
SAR	Structure Activity Relationship
SDWA	Safe Drinking Water Act
SIDS	Screening Information Data Set
SNP	Single Nucleotide Polymorphism
SPC	Science Policy Council
TMDL	Total Maximum Daily Load
TRI	Toxics Release Inventory
TSCA	Toxic Substances Control Act
UF	Uncertainty Factor
VCCEP	Voluntary Children's Chemical Evaluation Program

---

## EXECUTIVE SUMMARY

Advances in genomics will have significant implications for risk assessment practice and regulatory decision making. The Environmental Protection Agency's (EPA's) Interim Policy on Genomics (Appendix A), issued in 2002, appropriately acknowledges that genomics technologies have the potential to improve our understanding of an organism's response to stressors (USEPA, 2002a). The Interim Policy describes genomics as the study of all the genes of a cell or tissue, at the DNA, mRNA, or protein level. This policy states that while genomics data may be considered in the decision making process at this time, these data alone are insufficient as a basis for decisions. EPA will consider genomics information for assessment purposes on a case-by-case basis only. Following release of the Interim Policy, EPA held internal discussions regarding the potential of genomics approaches to improve our understanding of the effects of environmental stressors on cells. It was concluded that genomics information may lead to the development of predictive biomarkers of effect, thereby allowing for the identification of potentially sensitive populations and earlier predictions of adverse outcomes and, ultimately, leading to better intervention strategies. Enhancing understanding of the molecular mechanisms of toxicity may greatly reduce the uncertainty of extrapolations used in the current risk assessment process. Further, genomics technologies may enhance the development of more sensitive and cost-effective methods for toxicity screens and tests and may ultimately lead to the reduction, refinement, or replacement of more complex and costly standard tests for human and wildlife species.

“Genomics analysis is the study of all the genes of a cell or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) level.” Interim Policy on Genomics

Following these internal discussions, at the request of EPA's Science Policy Council (SPC), a Genomics Task Force was formed. The Task Force was charged to examine the broader implications genomics is likely to have for EPA programs and policies, to attempt to gain further understanding of the appropriate usage of these data and the potential consequences of their use, as well as to identify possible infrastructure needs. The Task Force was also charged with developing scenarios to describe various circumstances under which EPA might receive these data. The resulting document is intended to present implications of the use of genomics technologies in EPA practice for the consideration of Agency managers. It is the intent of the Genomics Task Force to initiate a discussion of the scientific issues regarding the incorporation of genomics information into human health and ecological risk assessments and of how these data will likely affect regulatory policy and decision making in the future. Although, as the Interim Policy notes, understanding genomic responses with respect to adverse ecological and/or human health outcomes is far from established, it is important for managers to begin to consider the likely future impacts of genomics technologies on their programs. Four areas have been identified as those very likely to be influenced by the generation of genomics information within EPA and the submission of such information to EPA: (1) prioritization of contaminants and

---

contaminated sites, (2) monitoring, (3) reporting provisions, and (4) risk assessment. This document briefly addresses ongoing research within the Agency for each of these four areas and identifies remaining research needs. It should be noted that genomics will not fundamentally alter the risk assessment process, but is expected to serve as a new, more powerful tool for evaluating the exposure to and effects of environmental stressors.

The Task Force identified several overarching challenges associated with genomics that fall into three categories: research, technical development, and capacity. These challenges are defined as critical needs for the Agency to strengthen its capability to use genomics information in a meaningful way, and to enable the Agency to address potential regulatory applications that are likely to arise with respect to genomics, such as those outlined in this paper. For research, the critical needs are identified as (1) linking genomics information to adverse outcomes; and (2) interpreting genomics information for risk and hazard assessment. It is important to note that significant research by EPA and other agencies and researchers will be necessary to fully understand and apply genomics technologies to human health and ecological risk assessment. One critical need in the area of technical development was identified as the need to establish a framework for analysis and acceptance criteria for genomics information for scientific and regulatory purposes (including data quality standards based on genomic assay performance). Two critical needs were identified with respect to capacity, including human capital: (1) applying strategic hiring practices to recruit individuals who possess “genomics core competencies” essential for crucial areas of research, analysis, systems biology, bioinformatics, and risk assessment; and (2) training EPA risk assessors and managers to interpret and understand genomics data in the context of a risk assessment.

Though advances in genomics and chemical development will present the Agency with new challenges, it is likely that genomics approaches will greatly assist in advancing EPA’s risk assessment and regulatory policy and decision making processes. The Agency must be proactive in identifying, developing, and standardizing applicable genomics approaches. Additionally, many scientific, policy, ethical, and legal concerns are developing along with the emergence of this science and will need to be addressed. The Genomics Task Force recommends that EPA begin taking steps to address the identified research, technical development, and capacity challenges in order to strengthen its capability to effectively use genomics information in the future. Recommendations for initial steps to address these challenges are presented in the final section of this paper. It is essential for EPA to continue to collaborate with other federal agencies, academia, the regulated community, public interest groups, and other stakeholders in this endeavor in order to benefit from ongoing advances in genomics in the wider scientific and regulatory communities.

---

## I. Introduction

### A. Background

The mapping of the genomes of diverse animal, plant, and microbial species, and related technologies are already significantly affecting research across all areas of the life sciences and will continue to do so for decades to come. The current understanding of biological systems is rapidly changing in ways previously unimagined, and novel applications of this technology are already being commercialized. These scientific and technological advances have spurred many federal agencies to consider the far-reaching implications for policy, regulation, and society as a whole.

On June 25, 2002, EPA released the Interim Policy on Genomics (Appendix A, USEPA, 2002a) communicating its initial approach to using genomics information in risk assessment and decision making. The policy describes genomics as the study of all the genes of a cell or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) level. The Interim Policy notes that, while genomics offers the opportunity to understand how an organism responds at the gene expression level to stressors in the environment, understanding such molecular events with respect to adverse ecological and/or human health outcomes is far from established. It concludes that while genomics data may be considered in the decision making process at this time, these data alone are insufficient as a basis for decisions. Therefore, EPA will consider genomics information for assessment purposes on a case-by-case basis only.

Following release of the Interim Policy, EPA held internal discussions to consider the potential genomics technologies have to improve our understanding of the effects of environmental stressors on cells. It was concluded that genomics information may lead to the development of predictive biomarkers of effect, thereby allowing for the identification of potentially sensitive populations and earlier predictions of adverse outcomes and, ultimately, leading to better intervention strategies. Enhancing understanding of the molecular mechanisms of toxicity could greatly reduce the uncertainty of extrapolations used in the current risk assessment process. The potential results may be the development of more sensitive and cost-effective methods for toxicity screens and tests and significant reductions in, or eventual elimination of, conventional animal testing.

Following these discussions, at the request of the Science Policy Council (SPC), a Genomics Task Force was formed. The Task Force was charged with the task of examining the broader implications genomics is likely to have on EPA programs and policies, to attempt to gain further understanding of the appropriate usage of these data and the potential consequences of their use, as well as to identify possible infrastructure needs. The Task Force was also charged with developing scenarios to describe various circumstances under which EPA might receive these data and the resulting implications (e.g., interpretation, relevance, evaluation, analytical, and research needs) for EPA policies and programs.

---

---

## B. Emerging Impacts of Genomics Technologies

While these are new technologies and most are not as yet ready for application in risk assessment and decision making, it is important for Agency managers to begin to consider the likely future impacts of genomics technologies on their programs. It should be noted that genomics will not fundamentally alter the risk assessment process, but is expected to serve as a more powerful tool for evaluating the exposure to and effects of environmental stressors and will offer a means to simultaneously examine a number of response pathways. EPA and other regulatory agencies are beginning to address the use of genomics data for various risk assessment applications, including the need to establish a link between genomic alterations and adverse outcomes of regulatory concern (Klaper et al., 2003). EPA must soon develop an explicit prescriptive strategy for accepting “omics” data submissions because such information (Genter et al., 2002) has already been referenced in a submission for a pesticide reregistration. Given the rapidly evolving nature of genomics technologies, care must be taken to develop an acceptable scheme to simplify and refine the risk-related information and to distinguish it from the large amount of complex scientific and statistical data available. This strategy must remain dynamic in anticipation of continuing technical evolution at the molecular levels (e.g., DNA, RNA, and protein). Furthermore, bioinformatic approaches for data acquisition and analysis, including technologies designed to store and analyze the profusion of data generated from microarray analysis, must be considered in parallel with the data-generating methods. Additionally, many scientific, policy, ethical, and legal concerns are developing along with the emergence of this science and will need to be addressed.

The Interim Policy on Genomics provides guidance concerning how and when genomics information should be used to assess the risks of environmental contaminants under the various regulatory programs implemented by the Agency at the present time. The standardization of experimental design and data analysis for genomics is important for the utility of genomics information in future risk assessment and regulatory decisions. Such standardization will enhance the reproducibility of results obtained and the reliability of conclusions drawn from these data. Furthermore, EPA should consider the development of data quality standards based on performance of microarrays, as well as other genomics technologies. This in turn will help to ensure the integrity of EPA’s approach to assessing the genomics information submitted to the Agency.

Genomics issues have already arisen in environmental decision making. For example, a pesticide registrant has cited several published genomic articles as part of their data package submission for product registration to EPA’s Office of Pesticide Programs. The data were submitted to propose an alternative mode of action that would affect human health assessment conclusions. Additional, similar submissions may soon be made by other pesticide registrants.

There are a number of other regulatory areas where genomics information will start having an impact. For example, a research consortium including State of California regulatory

---

agencies, public utilities, and EPA recently participated in a study comparing the performance of various genomics-based methods designed to identify the source of fecal material in ambient waters in an approach called microbial source tracking (Griffith et al., 2003). These methods are being evaluated to assist dischargers in complying with Clean Water Act (CWA) requirements to develop Total Maximum Daily Loads (TMDLs) for water bodies that are listed as impaired due to the presence of fecal coliforms. This work will also address the issue of beach closures; current microbial methods require several days to complete and do not distinguish between bacteria from humans and other sources such as sea gulls or seals. In another application, the State of California, as part of an ongoing ambient water quality monitoring program, is initiating an effort to evaluate surface waters for the presence of estrogenic endocrine effects using a reverse-transcription polymerase chain reaction (RT-PCR) assay for vitellogenin gene expression in livers of exposed male rainbow trout. If results show that some surface waters exhibit estrogenic effects, the Regional Water Quality Controls Boards in California, which issue National Pollutant Discharge Elimination System (NPDES) permits and perform ambient water quality monitoring, may begin to consider including this bioassay in their monitoring program for wastewater treatment facilities even though it is not yet an approved "EPA method." Additionally, one group of tribes in Northern California and Southern Washington proposes to use a series of molecular-biology-based assays to assess exposure to hormonally active compounds using either a multiplex RT-PCR approach or a multigene array. The information could ultimately be used to establish Tribal Water Quality Standards.

These examples indicate the emerging need to make proactive policy decisions and to develop processes to address how genomics data will be used in Agency decision making.

### C. Overview of Genomic Science

Genomics tools provide the observer with a means to examine changes in gene expression and protein and metabolite profiles within the cells of any organism, in contrast to older methods of analyses which restrict observers to looking only at whole organism effects or changes in single biochemical pathways. Genomics tools can provide detailed data about the underlying biochemical mechanisms of disease or toxicity (i.e., disease etiology), sensitive measures of exposures to chemicals, new approaches to detecting effects of such exposures, and methods for predicting genetic predispositions that may lead to disease or higher sensitivity to particular stressors in the environment.

Rapid advances in genomics may have significant implications for risk assessment and regulatory decision making.

As a means of introduction to genomics and its potential impact on regulatory decision making, it is important to understand the basic principles behind the technology. Only about 1-2% of the human DNA actually codes for RNA message that can be translated into a protein.

---



This 1-2% is considered the theoretical *functional genome*. Any particular cell type (i.e., from various organs or species) will have its own practical functional genome, which is a subset of the entire functional genome that encodes the proteins actually functioning in that cell. The functional genome for any cell type can be assessed by measuring its messenger RNA (mRNA) profile. The mRNA copies the necessary portion of the cell's DNA code and takes the information to the place in the cell where proteins are manufactured. Thus, the assessment of mRNA profiles is called *functional genomics*. Such profiles are constructed using *microarrays* that contain all (or a sampling) of a cell's functional genome. Hybridization of the mRNA that is being actively produced by the cell to these microarrays demonstrates which genes are currently active in that cell. Within the 98-99% of DNA not coding for RNA message is information that affects the activity of the functional genome by influencing where and when genes are active in an organism. Thus both coding and noncoding DNA are important in organismal function and response to perturbations. The oft-repeated statement that no two humans are alike (with the exception of identical twins) is valid at the genomics level as well. There is a wide range of DNA among individuals, even within the same family. Some of these differences arise spontaneously (but rarely) as mutations. Others are more frequent and represent very small DNA alterations that might or might not affect gene function; these are called *single nucleotide polymorphisms* (SNPs). While measurement of SNPs is not difficult, the need remains to associate these mutations or polymorphisms with specific genetic traits or cellular activities that could lead to adverse health outcomes.

The study of a cell's protein composition is called *proteomics*. Currently, it is possible to analyze only a fraction of a cell's proteins, but rapid advances in this field should allow more complete profiling in the near future. Another discipline of biology analyzes biofluids and tissues to determine the profiles of endogenous metabolites present under normal conditions or when the organism has been affected by factors such as exposure to environmental chemicals. This type of whole cell analysis is called *metabonomics* (or metabolic profiling). In order to understand how a cell functions under normal or stressed circumstances, it is necessary to characterize the proteins that are manufactured by the cell, as well as endogenous metabolites. This facilitates an understanding of global metabolism and how proteins interact along cellular activity pathways. This approach describes the area of *systems biology*, in which the cell, tissue, or organism is considered as a complete, albeit complex, system.

For the purposes of this document it is important to note that all of these so-called "omic" technologies can be used to compare functional genomes (mRNA) and proteomic and metabonomic profiles in normal cells and tissues with responses in stressed cells and tissues such as those exposed to environmental agents. Analysis of the large data sets generated for these type of analyses requires the development of new *bioinformatic* and

***Bioinformatics*** is data acquisition and processing technologies designed to store and analyze data generated from genomic analyses.

and

*computational* tools. An integrated analysis and understanding of biological systems and their responses to perturbation, from genes to adverse effects, and the capacity to collect and evaluate data supportive of such a view would be expected to greatly enhance the risk assessment process and, thus, aid in formulating regulatory policy and making regulatory decisions.

As the Agency considers the significance of the current state of genomics technologies, it is critical to realize that these technologies continue to advance at very rapid rates. Some of the technology “laws” that have been developed to describe this advance include Monsanto’s Law, “the amount of useful genetic information doubles every 18 - 24 months,” and Dawkin’s Law, “the cost of sequencing DNA base pairs halves every 27 months.” As an example, a commercial producer of gene chips has reported that the information content of their chips has been growing exponentially from 16,000 cDNA probes per chip in 1994 to over 500,000 in 2002. While commercial enterprises have recently developed arrays capable of handling large portions of the human genome, the cost of such microarray technology is still high enough that its use in toxicity-based approaches is relatively limited. However, these costs are falling rapidly as the broad utility of the technology becomes apparent. In this regard, it is estimated that in approximately two years, costs will decline to where clinical use will be feasible (e.g., assist in selecting treatments, intervene with disease before overt symptoms occur, offer genetically personalized nutrition and lifestyle advice, customize drug prescriptions), and the resulting economies of scale will contribute to a continuing decline in costs to the research community (Personal communication, from interviews by Robert Olson, Research Director, Institute for Alternative Futures, September 2003).

## **D. Purpose and Intent of this Document**

This is an unprecedented time in the history of science because of the rapid development of genomics and associated technologies. Genomics technologies are becoming highly sophisticated and have great potential for contributing to the assessment and management of environmental risks. The challenge lies in understanding how this information is likely to change current Agency approaches to human and ecological risk assessment and decision making. Although EPA recognizes the inherent issues currently associated with genomics studies in general and microarray experiments in particular and that these issues will need to be addressed before this technology can be fully accepted in risk assessment, the Agency also recognizes that genomics information will likely become an integral part of risk analysis in the future. The purpose of this document, consequently, is to present implications of the use of genomics technologies in Agency practice. It is also the intent of the Genomics Task Force to inform, invite discussion, and shed light on issues that will need to be addressed now or in the near future.

The purpose of this document is to present implications of the use of genomics and associated technologies in Agency decisions.

---

A key charge for the Task Force was identifying various exemplary circumstances under which EPA might receive genomics data and the resulting implications for EPA policies and programs. Sections II and III outline these circumstances or applications.

Section II identifies examples of regulatory applications in which genomics will likely affect regulatory decision making:

- a) Prioritization of Contaminants and Contaminated Sites
- b) Monitoring
- c) Reporting Provisions

Section III addresses areas where genomics will likely have applications for risk assessment practices. The risk assessment applications will also serve as tools for regulatory applications and decision making. For each of the regulatory and risk assessment applications, select representative activities are presented to illustrate the application. Additional activities are identified and described in Appendix B.

Section IV identifies genomics research needs and provides an overview of current EPA genomics research that may aid in addressing the regulatory and risk assessment applications outlined in Sections II and III.

Section V describes three categories of challenges EPA faces in applying genomics information to risk assessment and decision making and provides recommendations for addressing these challenges.

- a) Research
    - 1) Linking genomics information to adverse outcomes
    - 2) Interpreting genomics information for risk assessment
  - b) Technical Development
    - 1) Establishing a technical framework for analysis and acceptance criteria for genomics information (including data quality standards based on genomic assay performance)
  - c) Capacity/Human Capital
-

- 1) Applying strategic hiring practices to recruit individuals who possess genomics core competencies essential for crucial areas of research, analysis, systems biology, bioinformatics, and risk assessment
- 2) Training EPA risk assessors and managers to interpret and understand genomics data in the context of risk assessment

---

## II. Regulatory Applications

### A. Prioritization of Contaminants (Chemicals and Microbes) and Contaminated Sites

#### 1. Introduction

There are over 80,000 chemicals currently listed in the Toxic Substances Control Act (TSCA) Inventory (Personal communication, Dr. Henry Lau, USEPA, March 2004); a portion of these chemicals may no longer be in commerce. Sufficient information to allow a thorough evaluation of risk exists 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 variety of approaches has been developed to assist in prioritization decisions. In the current approach, chemical prioritization may be determined by several 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 a better knowledge-based approach is needed.

Genomics, combined with modern computing and information technologies, will advance predictive toxicology and improve the efficiency and reliability of prioritization and risk assessments within the Agency.

As an example, microorganisms have the potential to be spread through drinking water supplies and distribution systems. The Office of Water's 1998 final Contaminant Candidate List (CCL) comprises 60 contaminants and contaminant classes, including 10 microbial contaminants and groups of related microorganisms. Computer model results or expert judgment are currently used for CCL hazard estimation and prioritization activities.

Thus, there is a large number of stressors that the Agency must prioritize for further evaluation. Currently, there 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 hold the promise of providing more mechanistic, molecular-based data for risk-based prioritization of these stressors. In addition, these technologies are likely to offer more efficient, potentially high throughput, and low cost alternatives to the 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. Which model system(s) would one employ to test chemicals and what endpoint(s) would one look for are important questions to ask when using genomics information for prioritization.

---

## 2. Regulatory and Voluntary Activities Potentially Affected by Genomics Information

### a. Representative Activities

#### **Office of Pollution Prevention and Toxics (OPPT) - High Production Volume (HPV)**

**Challenge:** Genomics data could be applied to the voluntary HPV screening process. For example, specific gene expression data could be used to predict the relevance of an endpoint evaluated in an animal model Screening Information Data Set (SIDS) test to an adverse health response in humans. Further, these types of genomics data could, in the future, potentially supplement or supplant animal testing needed to complete the SIDS data set. However, given the timeframe of the HPV program, genomics data will unlikely supplant part of the SIDS dataset in the near term. More likely for the near term is the use of genomics data to validate category groupings in HPV and possibly future high volume chemical screening processes.

**Office of Water - Contaminant Candidate List:** Genomics data generated for CCL chemicals may be able to supplement computer model results or expert judgment in hazard estimation and prioritization activities. For example, computerized analysis and the growing use of automated polymerase chain reaction (PCR) techniques have allowed tremendous gains in the study of microbial genomics, as well as of whole organisms. A number of microorganism genomes have already been studied, many of which are associated with waterborne disease. Genomics databases may play a role in prioritizing pathogens based on the availability of virulence genes of concern and their corresponding gene products.

### b. Additional Activities

Genomics may also have regulatory implications for prioritization in other program offices as well as in regions, states, and tribes. Further details on the following activities are found in Appendix B.

#### Program Offices

- Office of Pollution Prevention and Toxics (OPPT): Premanufacture Notices (PMNs), Voluntary Children's Chemical Evaluation Program (VCCEP), Endocrine Disruptors Screening Program (EDSP)
- Office of Pesticide Programs (OPP): Pesticides, Inerts, EDSP
- Office of Water (OW): Prioritizing stream or wetlands for study or cleanup
- Office of Air and Radiation (OAR): Hazardous air pollutants (HAPs)
- Office of Solid Waste and Emergency Response (OSWER): Superfund sites
- Office of Research and Development (ORD): Future research on chemicals

#### Regions, States, and Tribes

- Site remediations and chemical evaluations
-

---

## B. Monitoring

### 1. Introduction

The term “monitoring” in the present context refers to any activity by which environmental samples are taken and used for regulatory or prioritization decisions and for developing environmental status and trends information. Many programs have either site-specific or media-specific data requests that are used to make regulatory decisions, monitor compliance, and/or to prioritize the

Monitoring and compliance applications may be some of the quickest Agency uses of genomics-based information.

use of EPA’s human and economic resources. In addition, EPA is charged with determining the state of the environment, as well as assessing the status and trends of ecological condition. In many instances, entities other than EPA generate the data EPA uses (e.g., other federal agencies, states, tribes, the regulated community, and interested stakeholders in volunteer monitoring programs). These data are generated through various headquarters and regional programs through contracts, grants, and cooperative agreements. In fact, a large portion of EPA regions’ budgets are directed to programs (e.g., in states and tribes) that generate information that could fall into the category of monitoring data or information.

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 (e.g., fish and/or invertebrate IBIs [index of biotic integrity], algal and plant community structure, invasive species evaluations in terrestrial and aquatic ecosystems); and microbial community and pathogenic microorganism analyses of air, water, soil, and sediment.

Many of these types of environmental data could be generated using genomics-based techniques, and some applications are already being tested. One example is in the area of microbial source tracking to determine the sources of fecal contamination that may be causing impairment of a water body resulting in a beach closure. Several state agencies and public utilities are evaluating molecular-biology-based and genomics-based techniques to determine whether these approaches can distinguish among fecal sources in order to develop TMDLs for impaired water bodies (Griffith et al., 2003). A second example is the area of site clean up in which changes in microbial community response to a stressor such as an oil spill may be characterized using these techniques. One genomic approach to evaluating changes in microbial community is to use total DNA, representing all of the microbial community, rather than the one-to-two percent of the microbes that can be cultured. This genomic information could be used to differentiate and evaluate the feasibility among remedial alternatives, such as active remediation

---

(i.e., adding nutrients or microbial cultures) versus monitored natural attenuation. The Department of Energy is exploring how this type of microbial community “fingerprinting” can be used to distinguish the conditions that promote effective bioremediation of petroleum-contaminated soils or sediments ([http://www.sc.doe.gov/ober/ERSD/ersd\\_nabir.html](http://www.sc.doe.gov/ober/ERSD/ersd_nabir.html)). A third example is in the area of development of multi-gene arrays of model animals (e.g., “fish-on-a-chip”). Future toxicity testing for compliance with discharge requirements could involve using gene chips for species such as fathead minnow to determine whether a water sample resulted in a toxic pattern of response in exposed fish.

Thus, there is a wide range of monitoring information that the Agency considers and a wide range of potential applications of genomics technologies. The cost and time required to collect and analyze 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.

## **2. Regulatory Activities Potentially Affected by Genomics Information**

### **a. Representative Activities**

**Office of Water (OW)/Office of Ground Water and Drinking Water (OGWDW):** OGWDW anticipates that monitoring for chemicals or microbial pathogens will use genomics-based data in five to ten years for the following purposes:

- Compliance monitoring by federal, state, and tribal agencies to determine if surface waters meet designated uses standards under CWA (TMDLs, Criteria, Standards)
  - Monitoring finished or source drinking waters for contaminants
  - Real-time monitoring of ambient surface water, e.g., for beach or shellfish bed closures (Beaches Environmental Assessment and Coastal Health [BEACH] Act, Standards)
  - Monitoring classes of compounds based on biological activity or mode of action (e.g., cholinesterase inhibition)
  - Developing occurrence data as a basis for Safe Drinking Water Act (SDWA) regulation or listing on CCL
  - Developing future drinking water regulations (6 Year Review)
-



**Regions, States, and Tribes - State NPDES permits:** The State of California, as part of an ongoing ambient water quality monitoring program, is initiating an effort to evaluate surface waters for the presence of estrogenic endocrine effects, using a RT-PCR assay for vitellogenin gene expression in livers of exposed male rainbow trout. If results show that some surface waters exhibit estrogenic effects, the Regional Water Quality Controls Boards in California, which issue NPDES permits as well as perform ambient water quality monitoring, may consider including this type of bioassay into the monitoring program for waste water treatment facilities even though it is not yet an "EPA method." Currently, NPDES permits contain only chemical and toxicity-based limits, require chemical analyses and toxicity testing of effluents to demonstrate compliance with permit limits, and relate waste loads to watershed TMDLs. Genomics technologies could provide new and more sensitive monitoring tools to develop discharge limits for NPDES permits.

### **b. Additional Activities**

Genomics may also have regulatory implications for monitoring in other program offices, as well as in regions, states, and tribes. Further details on the following activities are found in Appendix B.

#### Program Offices

- OPP: Pesticide monitoring for registrations and reregistrations
- OAR: Stationary source monitoring
- OSWER: Superfund and Resource Conservation and Recovery Act (RCRA)-required monitoring
- Office of Environmental Information (OEI) and ORD: Biomarker development

#### Regions, States, and Tribes

- State and local beach closure and TMDL issues associated with pathogens
- State and local air quality monitoring
- Tribal issues (e.g., monitoring for endocrine disruptors)
- Regional pesticide program inspections

## **C. Reporting Provisions**

### **1. Introduction**

#### **a. Reporting on Adverse Effects of Commercialized Chemicals and Pesticides**

Reporting of certain adverse effects/risks for industrial chemicals and pesticides already on the market is mandated under both the Toxic Substances Control Act (TSCA) and the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). TSCA section 8(e) requires that "[a]ny person who manufactures, processes, or distributes in commerce a chemical substance or mixture

---

and who obtains information which reasonably supports the conclusion that such substance or mixture presents a substantial risk of injury to health or the environment shall immediately inform [EPA] of such information” (15 U.S.C. 2607(e)). FIFRA section 6(a)(2) states “If at any time after the registration of a pesticide the registrant has additional factual information regarding unreasonable adverse effects on the environment of the pesticide, the registrant shall submit such information to the Administrator.” There is a need to interpret how these TSCA and FIFRA 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. This issue may be best approached in a multi-stakeholder process to ensure scientific consensus around the understanding of adverse effects based on genomics data.

### **b. Toxics Release Inventory Program**

The Toxics Release Inventory (TRI) database was established under the Emergency Planning and Community Right-to-Know (EPCRA) Act of 1986. Section 313 of 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 of EPCRA section 313 (d)(2) are primarily based on hazard, not on 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. A chemical may be added to the TRI list if (a) the chemical is known to cause, or can reasonably be anticipated to cause, significant adverse acute human health effects at concentration levels that are reasonably likely to exist beyond facility site boundaries as a result of continuous, or frequently recurring, releases; (b) the chemical is known to cause or can reasonably be anticipated to cause, in humans, cancer or teratogenic effects or serious or irreversible reproductive dysfunctions, neurological disorders, heritable genetic mutations, or other chronic health effects; or (c) if the chemical is known to cause, or can reasonably be anticipated to cause, because of its toxicity, its toxicity and persistence in the environment, or its toxicity and tendency to bioaccumulate in the environment, a significant adverse effect on the environment of sufficient seriousness, in the judgement of the Administrator, to warrant reporting.

## **2. Regulatory Activities Potentially Affected by Genomics Information**

In order for genomics technologies to have an effect on reporting provisions, the issue of linking genomics changes to adverse effects or response pathways needs to be addressed. Once genomic changes are linked to adverse effects, the Agency will need to make decisions regarding whether genomic changes apply to reporting provisions.

---

Genomics technologies could affect reporting requirements under TSCA 8(a) and FIFRA 6(a)(2) if genomic changes detected are linked with substantial risks or adverse effects. If genomics data do, in the future, become a reporting requirement, this could also affect the number of reports received under these reporting provisions and the resources required to evaluate the reports.

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 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. Practical application of genomics-derived information will improve the quality of hazard assessments conducted by EPA, including those conducted by EPA's TRI Program.

---

### III. Risk Assessment

#### A. Introduction

Genomics technologies present an opportunity to greatly enhance Agency risk assessments. Specifically, genomics technologies are likely to contribute significantly to improvements in defining a chemical's mode of action, evaluating effects on susceptible populations and life stages, and assessing exposure to and effects of chemical mixtures, as outlined below. For example, in collaboration with the International Life Sciences Institute (ILSI), EPA has developed a framework for microbial risk assessment (MRA), and OW is working to expand the framework into a full MRA protocol which may include consideration of genomics data. OW is also working with EPA's Risk Assessment Forum to develop MRA guidelines. It is important that any change, refinement, or addition to a risk assessment practice be vetted by interested stakeholders and ultimately peer reviewed to ensure scientific consensus around the practice.

Genomics information will not fundamentally alter risk assessment practices, but genomics will provide powerful new tools and insights to use for risk assessment.

#### B. Mode of Action (MOA)

##### 1. Overview

The term "mode of action" (MOA) is defined in the Draft Guidelines for Carcinogen Risk Assessment (USEPA, 2003a) as a sequence of key events and processes, starting with the interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in an adverse outcome. A "key event" is an empirically observable precursor step that is in itself a necessary element of the MOA or is a biomarker for such an element. Genomics technologies can be used to better understand the MOA of a chemical agent and, thus, can lead to advances in human and ecological risk assessments of chemicals. As genomics information contributes to our understanding of MOAs, the validity of using this information as an indicator of both adverse effects and exposure is enhanced.

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 MOA. There are many examples of possible modes of carcinogenic action, such as mutagenicity, mitogenesis, inhibition of cell death, cytotoxicity with regenerative cell proliferation, and immune suppression. MOAs have been identified for other adverse outcomes, both human health and ecological (e.g., cytotoxicity, endocrine disruption, loss of homeostasis). Such approaches will significantly enhance the Agency's ability to harmonize risk assessment approaches for different outcomes for which the development of a list of common MOAs is essential.

---

---

Understanding the MOA of environmental agents that induce toxic effects other than cancer or that induce carcinogenicity in animal models should facilitate the assessment of the relevance of these findings in protecting human health and safeguarding the environment. An important issue for extrapolation of responses in animal models to humans or environmental endpoints is to establish whether the MOA in the test species is relevant in the target species.

## **2. Risk Assessment Activities Potentially Affected by Genomics Information**

Many program and regional offices need to make judgments about chemicals with little or no data available. Others have extensive datasets, but still struggle with accuracy and precision as well as extrapolations to nontested species or scenarios. Genomics approaches are envisioned to provide improvements in (a) hazard identification, (b) dose-response assessment, (c) extrapolations, and (d) exposure assessment.

### **a. Representative Activities**

#### **i. Improving Hazard Identification**

*Evaluating chemicals for genotoxic or other MOAs.* New approaches using tissue microarrays can enhance throughput and the linking of genomic and cellular outcomes. Further, combining the findings of gene expression studies with data from chemical exposures of genetically altered animal models (e.g., knockout or null mice) is a powerful tool to link specific genes to specific detrimental outcomes. Such data will allow the development of gene profile “fingerprints” of genomic characteristics for specific MOAs. The development of genomic “fingerprints” will provide a rapid screening method to categorize chemicals with unknown MOAs for both human health and ecological assessments.

*Predicting or Defining Metabolic Pathways.* The chemical evaluation process includes consideration of the parent compound and its potentially active metabolites. Genomics approaches, particularly at the proteome level, will aid in the characterization of metabolic pathways and the identification of toxicologically active metabolites. Computational toxicology approaches will further enhance the prediction of metabolic pathways and metabolites for chemicals that have not been investigated experimentally and potentially will reduce the use of test animals and the cost of data generated to support risk assessments. Metabolic pathways and the genes associated with those pathways need to be linked to adverse effects of concern.

*Replacement of standard toxicity tests in regulatory batteries with rapid, pathway-specific response tests.* It has been envisioned that relevant genes and gene products for specific toxicities such as genotoxicity can be formatted on arrays to provide a more comprehensive analysis than currently available assays (Aardema and MacGregor, 2002). Because many of the toxicological testing procedures and strategies required by EPA have remained largely unchanged for 20 years, it is reasonable to assume that many of the current assay systems used may be replaced by more sensitive, rapid, and predictive genomic assays able to identify specific

---

pathways of response. Acceptance of these genomic protocols for both human health and ecological assessments will lead to time and cost savings and may also lead to more accurate risk assessments.

### **ii. Improving Dose-Response Assessment**

***Linear versus nonlinear.*** The Agency's traditional approach to cancer risk assessment for agents that are known mutagens and carcinogens employs a linear, low dose extrapolation to quantify possible human cancer risks. The underlying premise for this linear default is that electrophilic compounds are presumed to form single DNA modifications in single cells that could potentially lead to cancer. Due largely to the successful use of genetic toxicological testing schemes for screening, however, the number of new genotoxic carcinogens entering the environment is likely to be small. The recent Final Draft Guidelines for Carcinogen Risk Assessment addresses the issue of nongenotoxic carcinogens and encourages the use of mechanistic data to identify whether a nonlinear extrapolation is appropriate for nongenotoxic carcinogens. For this purpose, biomarkers of response for genotoxic carcinogens are available at least at the cellular response level if the general cancer MOA is known. The challenge will be to develop biomarkers of response that can be used for predicting specific outcomes for nongenotoxic chemicals (Bartosiewicz et al., 2001a). This goal can be realized through gene expression pattern recognition that parallels histological changes in tissues and the eventual progress to tumor formation. An example would be the CIIT (Chemical Industry Institute of Technology) Centers for Health Research's efforts to identify genes associated with peroxisome proliferators, such as the PPAR- $\alpha$  (peroxisome proliferator-activated receptor alpha) that are linked to alterations in mouse hepatocellular growth following peroxisome proliferator exposure.

The regulatory impact of genomics on possible nonlinear extrapolation for nongenotoxic carcinogens is significant. Within the Agency, nongenotoxic carcinogens without plausible MOA data are currently subjected to the same linear low dose extrapolation applied to genotoxic carcinogens. It is a reasonable assumption that a collaboration among industry and EPA scientists will occur in the area of MOA-based cancer risk assessment to ascertain if a nonlinear low dose extrapolation is appropriate for nongenotoxic carcinogens. The same MOA approach can be used to help more clearly discern dose-response relationships for chemicals that affect other health endpoints as well.

***Lowering of Points of Departure (PODs) based on genomic responses.*** Genomics technologies have the potential to affect dose-response analyses for nonlinear assessments of adverse toxicological outcomes. In traditional toxicology, doses used to determine adverse effects are generally high to ensure that tissue level or whole animal toxic responses are demonstrated. This permits the selection of toxicity endpoints and establishment of doses at which no adverse effects are seen (NOAEL) and the lowest doses at which an adverse effect is seen (LOAEL). Thus, most toxic substances currently are regulated on frank toxicity rather than on a molecular level response, and the association between a molecular level change and an adverse outcome has only rarely been established. A few substances are regulated based on biochemical changes with

---

known relationships to adverse outcomes such as cholinesterase-inhibiting pesticides and lead. Organophosphate pesticides are regulated at NOAELs which are generally much lower than many chemical classes because the endpoint, cholinesterase inhibition, is determined biochemically, and inhibition can generally be detected well below the levels showing overt toxicity. In contrast, many fungicides or herbicides have relatively high NOAELs because clear pathological alterations only occur in animals at high doses. Regardless of the chemical class or use, however, with the advent of molecular technologies including genomics, chemically-induced changes in gene expression are likely to result in the identification of simple, sensitive, and relevant biomarkers of effect that can be used in dose-response studies to more readily identify effects in the low dose range (i.e., below doses causing frank pathology) for humans and wildlife species.

If EPA chooses to establish regulatory limits (e.g., NOAELs) based on changes in gene expression (e.g., identifying upstream precursor effects or markers of adverse effects), the POD used to set the regulatory limit could be higher or lower in both the human health and ecological arenas. EPA needs to examine whether a lower effect level based on a molecular effect is “safer” than a level based on the currently used toxic effect. How this will affect risk assessment practices will need to be addressed via a multi-stakeholder process and peer review.

### **iii. Improving Extrapolations**

***High to low dose extrapolations; Route-to-route extrapolations.*** Reduction of uncertainty is one of the primary ways to improve the risk assessment process. Reduction in uncertainty in dose-response assessments can be enhanced by the use of predictive models such as Physiologically-Based Pharmacokinetic (PBPK) and Biologically-Based Dose Response (BBDR) models. These models can provide better methods for calculating dose metrics (e.g., target tissue doses) that are more flexible and relevant for extrapolation across exposure routes, between species, and from high to low doses. The potential of molecular indicators to define the shape of the dose-response relationship at low exposures suggests the possibility that alteration or elimination of some uncertainty factors (UFs) may be justified; data-derived factors based on genomics information may be determined and applied in the future. Similarly, molecular-based pharmacokinetic data that describes the distribution of biologically effective doses of active ingredients to target organs via other portals of entry offers the possibility of reducing uncertainties associated with route-to-route extrapolations.

### ***Interspecies extrapolations.***

**Relevance to humans.** Further improvements in human health assessments can be realized through the use of genomics data that support an evaluation of whether or not MOAs determined in test animals are similar and feasible in humans (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 humans and rodents would indicate interspecies differences and support a possible conclusion of non-

relevance to humans. Conversely, data showing good agreement in key genes or expression patterns between humans and rodents would provide higher confidence in the relevance of the findings to human health. 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 looks promising for linking cellular responses to a specific environmental chemical or mixture in laboratory animals to responses in humans or human cells *in vitro*.

Key biological systems have fundamental genomics processes, some of which, if altered, are universally deleterious. Demonstration of a common interspecies genomic response linked to an adverse effect and evaluation of the dose-response relationships in the lower animal (e.g., invertebrates) and humans could permit the extrapolation of genomic responses in lower order animals to adverse effects in humans. The increasing development of genomic information in lower organisms may provide a means to evaluate potential effects in humans that will extend the use of lower organisms beyond current mutagenicity testing.

**Relevance to wildlife species of concern.** In ecological risk assessment, it is necessary to extrapolate results from a very limited set of test species to a wide range (potentially hundreds to thousands) of species present in the environment. 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 present in the ecosystem. Use of genomics tools for the development of quantifiable pharmacodynamic models and applicable molecular markers will also significantly enhance species-species extrapolations and reduce the current reliance on the application of uncertainty factors.

#### **iv. Improving Exposure Assessment**

Genomics technologies are likely to lead to the development of simple, sensitive, and informative 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. Current methods rely on residue analyses or modeled scenarios and a few well-documented biomarkers of exposure (e.g., CYP1A, cholinesterase, delta-aminolevulinic acid dehydratase, metallothionein). Molecular techniques such as the use of microarrays or RT-PCR are already providing tools for documenting actual exposures to humans and 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. In the near future, genomics may aid in the identification of new biomarkers that can identify exposure to more stressors or more MOAs, and potentially enhance

---



the quantitation of these. When pharmacodynamic and MOA studies become sufficiently robust to relate exposure endpoints to whole organism adverse effects, risk assessment predictions will become significantly more accurate.

## C. Susceptible Populations and Sensitive Life Stages

### 1. Overview

#### a. Susceptible Human Populations and Sensitive Life Stages

Genomics and related technologies offer a tremendous opportunity to define and identify people with enhanced susceptibility to many environmental contaminants. The human genome consists of 30,000 or so genes that build cellular structures, control the cell cycle, execute metabolic functions, and mediate the information flow within and between cells. Small differences in gene sequence, known as single nucleotide polymorphisms (SNPs), can have a dramatic or inconsequential effect on protein function and activity depending on the particular polymorphism. Genomics technologies have the potential to yield information about the distribution of SNPs within the human population and their potential effects on genes that are responsive to various environmental contaminants. The interaction of genetic variants with environmental conditions can affect individual susceptibility to a variety of diseases such as cancer, diabetes, and heart disease and can promote sensitivity to disease from exposure (Bishop et al., 2001).

Genomics technologies offer a tremendous opportunity to define and identify people with enhanced susceptibility to many environmental contaminants.

Delineation of the frequency of occurrence of these polymorphisms within racial or ethnic groups may raise ethical, legal, and social implications in the area of environmental justice (Marchant, 2002). For example, genomics technologies might result in further categorization of individuals, and consideration must be given to how these newly identified at-risk groups will be included in environmental policy. Ethical practices by which genomics studies are conducted on human test subjects in specific ethnic or racial communities will also need to be carefully considered. The Agency will need to be proactive to ensure that these issues are handled in a just manner.

For a susceptible human population, an increased risk of illness could result from exposure to environmental chemicals or microbial pathogens at any age. In contrast, an exposure during a susceptible life stage could result in higher risk during a specific portion of an individual's lifetime or could influence an outcome at another life stage and the potential adverse effect incurred may be irreversible. For example, an exposure during early childhood development might yield a specific form of cancer that would not have been induced if the

exposure occurred later in life. Similarly, an exposure during a sensitive life stage could result in an increased probability of disease later in life. The primary difference between susceptible populations and susceptible life stages is that a susceptible population's toxic exposure generally will yield an adverse outcome regardless of the age at which the exposure occurred. Nevertheless, many of the ethical, legal, and social implications (ELSI) that apply to susceptible populations will also apply to susceptible life stages. As a consequence, the Agency will need to develop a policy regarding the collection and use of human genomics information from individuals to provide safeguards regarding privacy.

The National Institute of Environmental Health Sciences (NIEHS) has funded and supported research to identify gene variations that affect susceptibility to environmental agents. Approximately 500 genes were identified as part of the Environmental Genome Project of 1997. These genes affect metabolism, DNA repair, cell cycle control, receptors, and immune function. Although scientific progress has been made in understanding genetic variations and susceptibility to toxic chemicals and pathogens, research efforts have yielded inconsistent results. In spite of this, researchers continue to characterize genetic variations of susceptibility and provide insights about individuals who are more or less susceptible to disease from exposure to toxic substances (Marchant, 2002).

#### **b. Susceptible Wildlife Populations and Sensitive Life Stages**

The term "susceptible population" is most frequently used in reference to at-risk human populations; however, the term can also be used to describe wildlife populations that may be at risk due to exposure to environmental contaminants. Certain taxonomic subgroups of plants and animals may be more susceptible than others although this varies with contaminant and life stage. Currently, EPA examines species sensitivity through toxicity testing of environmental contaminants on representatives from various classes of organisms. Genomics technologies offer a powerful tool with which to examine toxicological responses across species for prediction of sensitivities or tolerances in untested organisms. In addition, genomics technologies will allow for the examination of long-term ecosystem health and the potential irreversibility of toxic effects. Examining the genetic diversity of organisms inhabiting a given ecosystem would allow risk assessors to determine if exposure to environmental contaminants might cause an evolutionary change in ecosystem structure or function.

Genomics technologies offer a powerful tool to examine toxicological responses across species for prediction of sensitivities or tolerances in untested organisms.

Threatened and endangered species represent a subset of all species faced with the possibility of extinction and are afforded extra protection under the Endangered Species Act. A species, however, may be sensitive to exposure to an environmental contaminant without being endangered. Comparisons of the sensitivities of endangered and common fish test species to

---

toxicants have been made. This research indicated that the sensitivities of the endangered species tested were generally not greatly different from that of the more common species (Sappington, et al., 2001). However, because the population size of endangered species is, by definition, quite small, reduced genetic variation may result in reduced tolerance to multiple stressors such as combinations of contaminants and climate stress (Porter et al., 1984). Alternatively, continuous exposure of a widespread, susceptible species to an environmental contaminant might result in the species becoming endangered.

While human behaviors such as over-hunting or habitat destruction are the main stressors that threaten species survival, contamination of an ecosystem with a pesticide, industrial chemical, or pathogen can also harm ecosystem health. Genomics technologies will provide significant insight into the biochemical mechanisms by which environmental contaminants might adversely affect certain species and may provide insights into which species are more likely to be headed down the path towards extinction. Cross-species extrapolations of sensitivity coupled with the ability to measure contaminant-induced reduction in genetic diversity within specific populations will provide valuable input to population viability models.

## **2. Risk Assessment Activities Potentially Affected by Genomics Information**

### **a. Susceptible Human Populations and Sensitive Life Stages**

EPA anticipates genomics research will be used to assess hazards and risks of chemicals and pathogens to specific human populations. Currently, the Agency does not generally take into account genetic factors when assessing the risks posed by chemical or biological substances although life stage and, to some extent, gender are considered. Additionally, the Agency rarely considers the genetic predisposition of a specific individual, race, or ethnic background when determining the toxic effects of chemicals. The Food Quality Protection Act (FQPA) of 1996 does, however, direct the Agency to determine the potential of increased susceptibility of infants and children from exposure to toxic substances such as pesticides.

#### **i. Representative Activities**

**OPPTS, OW, OAR, ORD, OSWER.** In any situation where a “safe” exposure level is designated (FIFRA, Clean Water Act, Safe Drinking Water Act, Clean Air Act, CERCLA), the identification of a susceptible population or life stage has the potential to force a change in the health standards employed. For example, the health standards used during hazardous waste remediation projects are often first reviewed and published in the Integrated Risk Information System (IRIS) database maintained by ORD. One of the goals of the IRIS database is to provide oral Reference Doses (RfDs) and inhalation Reference Concentrations (RfCs). Both the RfD and RfC are based on the assumption that nonlinear dose-response curves exist for certain toxic effects such as cellular necrosis (USEPA, 2002b). In general, these values provide an estimate of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime (USEPA, 1999). Genomics

technologies will provide a powerful tool for the identification of sensitive subgroups and allow specific decisions to be made that address specific sensitivities. As the use of genomics technologies becomes widespread, the number of susceptible populations identified will likely grow in number. Consequently, the Agency will need to be vigilant when revisiting health standards in all media to identify and protect susceptible populations.

The impact of genomics technologies on EPA's understanding of life stage sensitivity also will become an important issue. For example, different human age groups may express varying levels of some metabolic enzymes (Hakkola et al., 1998). Enzyme over- or under-expression could play a part in determining the severity of a toxic exposure. Additionally, the rapid growth taking place early in life is largely dependent on gene-environment interactions. Perturbation of this interaction through toxic exposure has the potential to significantly influence development. Genomics will provide tools to identify life stages that need separate assessments based on their unique susceptibilities.

**OPPTS.** The extent of enzyme activation is partially responsible for determining the severity of response to a chemical exposure. In addition, individuals may respond to metabolic stimuli to varying degrees depending on their genetic composition. Ethanol exposure, for example, is known to induce the metabolic enzyme CYP2E1; however, the amount and activity of CYP2E1 produced may vary among individuals (Haber et al., 2002; Snawder and Lipscomb, 2000). If genomics technologies are successful in identifying populations susceptible to specific pesticides or industrial chemicals, product labeling will probably be necessary. For example, labels might include warnings for particular populations known to exhibit higher frequencies of an at-risk genetic polymorphism. The pharmaceutical industry already includes warnings to susceptible populations on drug labels. The Agency has the ability to follow similar practices for pesticides because Section 3 of FIFRA provides EPA the authority to regulate labels (USEPA, 2003c).

**OSWER, Regions, States, and Tribes.** Genomics data might be used in the future to help identify susceptible populations or life stages when assessing risks at hazardous waste sites or for remediation of contaminated areas such as Superfund sites or other scenarios in which relatively small, identifiable populations are exposed. Genomics data may provide information on potential exposure patterns and also might be useful in developing site-specific remediation goals. If, for example, a genomics study was to identify a susceptible population at risk due to exposure to a contaminant at a Superfund site through a correlation of genomic analysis of local populations and measured or expected exposure levels, the Agency might choose to reduce the RfD/RfC value and propose more strict remediation measures. This, of course, presupposes an established linkage of the genomic endpoint and an adverse effect. Use of new genomics tools could, however, limit the extent of remediation measures by more accurately predicting the potential for exposure of the sensitive population. Thus, genomics tools may play a key role in determining intensity and extent of clean up practices and have large implications for time and cost of such procedures.

---

---

## **b. Susceptible Wildlife Populations and Sensitive Life Stages**

EPA currently does not use genomics technologies in ecological assessments or criteria development. However, numerous potential applications for the identification of sensitive species, populations, or life stages may become available in the near term. Longer-term applications are also under development.

### **i. Representative Activities**

**OPPTS, OSWER, OW, OAR, ORD.** Due to time and resource limitations, as well as ethical considerations, all species cannot be tested for responses to contaminants either for site-specific mitigation needs, for product registration, nor for criteria development. Therefore, the development of reliable methods for extrapolating toxicity information from tested species to those that are of concern but cannot be directly tested is necessary. This need is particularly acute for chemicals that may target sensitive life stages (e.g., metamorphosis in amphibians) or vulnerable species (those with small population sizes or that may have greater sensitivities to particular chemicals). Genomics technologies will provide the potential for extrapolating between test species and sensitive wildlife species or life stages in a rapid, cost-effective manner.

Because of the smaller population sizes of endangered species, it may be possible to identify biomarkers of effect can be evaluated non-invasively using genomics techniques (e.g., metabonomic analysis of urine or feces). Additionally, genomic techniques could potentially be used to evaluate the differences between populations of species, to determine whether they are unique species or are geographically separated populations of the same species. This could affect the way potential impacts on threatened or endangered species are managed.

The most immediate application is likely to be with aquatic organisms and the application to quantitative structure activity relationships (QSARs) or computational toxicology. This will be particularly useful in the Premanufacture Notice (PMN) process of OPPTS, but may find applications in other offices as well (e.g., Office of Water). The next application likely will be in the development of methods for screening chemicals that are potential endocrine disruptors in aquatic ecosystems.

While the protection of individual plants and animals from clinical disease caused by xenobiotic compounds and pathogens is an achievable goal, how chemicals combine with other environmental stressors to change the genetic properties of a population over time is less clear. If genetic diversity is reduced or if particular genes are suppressed or expressed at abnormal rates, it is possible that (1) the population may become less fit over time, (2) response to additional stressors may not be adequate; and (3) the reduced diversity may lead to a bottleneck and/or eventual extinction. Because genetic diversity is the fundamental basis for adaptation and evolution, it is increasingly being recognized as an important endpoint for risk assessment. To date, the methodology has not been available to address this issue, but newly emerging genomics techniques will allow such assessments in the future.

---

**OSWER, Regions, States, and Tribes.** The current Agency practice in ecological risk assessment and clean up of contaminated sites (Superfund, Brownfields) is to focus on the most sensitive species when determining the effects of a chemical contaminant on an ecosystem. Genomics will allow risk assessors a greater ability to focus the ecological risk assessment on the mechanistic level. Although the Agency does not usually take into account genetic factors when assessing the ecological risks posed by chemical or biological substances, genomics data gained from the examination of a single species might prove useful in preventing harm to other species with similar genetic characteristics.

## D. Mixtures

### 1. Overview

Human health and ecological risk assessments of toxic substances often are incomplete because most toxicological screening is performed for single chemicals. However, human and wildlife exposures to chemicals are rarely limited to a single chemical, but instead are usually to complex mixtures of chemicals during a lifetime. Environmental exposures from point and nonpoint sources, irrespective of medium, generally occur as simultaneous or sequential exposures to multiple chemicals. In addition to environmental exposures, the majority of the human population engages in intentional exposure to a variety of pharmacologically active chemical compounds such as those in recreational drugs (alcohol and tobacco), medicinal products, and foods and is inadvertently exposed to other chemicals, such as those in vehicle exhaust, drinking water, indoor air, and workplace environments.

Genomics technologies may aid in the identification of unique patterns of gene expression in the tissues of aquatic and terrestrial organisms, and human cell-based models, induced by exposure to multiple environmental stressors.

Chemical mixtures in the environment are, in general, a complex group of active and inert parent compounds, transformation products, and/or residues of which composition is qualitatively and quantitatively not fully known (Feron and Groten, 2002). Because mixtures change with time and distance from the original release site due to the differential fate and transport of their components (Pohl et al., 1997), regulations established on toxicological data for the original mixture may have little bearing on the actual exposures resulting from a release. It is estimated that approximately 275 million tons of hazardous waste are produced annually in the United States, and more than 2000 distinctive toxicants in site-specific media have been identified by EPA, with hazardous substances in given mixtures numbering in hundreds (Suk et al., 2002).

Specific environmental chemicals have been demonstrated to adversely affect the health of humans and wildlife. These concerns are amplified by the awareness that exposure to

---

chemicals often occurs in mixtures of chemicals that might exhibit complex interactions. The various types of toxicological interaction associated with complex chemical mixtures can be sorted into three reference categories: greater-than-additive (synergism, potentiation), additive (the sum is equal to the parts, no apparent interaction), and less-than-additive (antagonism, inhibition, or masking). Of particular importance is whether a mixture of components, each of which is present at concentrations below the level of concern, may be hazardous due to additivity, specific interactions, or both (Hertzberg and MacDonell, 2002). Synergistic toxicity resulting from co-exposure to pesticides has been observed, and greater than expected toxicity has been noted for pesticide mixtures of certain cholinesterase inhibiting insecticides and some fungicides.

The Agency has been directed by FQPA to consider the combined effects on human health that can result from exposure to toxic substances that share a common mechanism of toxicity (e.g., organophosphates). This cumulative risk assessment approach is based on an evaluation of the potential for people to be exposed to more than one member of a group of chemicals at a time and considers exposures from food, drinking water, and residential sources. It is important to note that for any group of toxic substances with a common mechanism of action, agents within that group that have low toxicity but the potential for high exposure can present a risk similar to a toxic substance with higher toxicity and lower exposure potential. It seems likely that the hazard, dose-response, and exposure assessment components of the cumulative risk assessment process will be greatly improved by the elucidation of mechanisms or modes of action made possible by genomics data. While genomics information will be useful in demonstrating whether two or more compounds share a MOA, there may be a need to establish a logical framework for applying genomics information to the interpretation of the effects of mixtures.

The Agency for Toxic Substances and Disease Registry (ATSDR) of the Department for Health and Human Services and EPA, in addition to other Federal research and regulatory agencies, support *in vitro* and *in vivo* research to further understand the chemical characterization, molecular mechanisms of action, and toxicity of chemical mixtures and their relationship with human health effects and other biological systems. Genomics can provide the tools for accomplishing this work that has to date been very expensive and required a large commitment to animal testing.

## **2. Risk Assessment Activities Potentially Affected by Genomics Information**

Genomics technologies will aid in the identification of unique patterns of gene expression in aquatic and terrestrial organisms and human cell-based models induced by exposure to multiple environmental stressors. Several studies already have described tissue-specific transcriptional patterns and have begun to address the concept of “fingerprinting” for chemical mixtures in laboratory animals (Bartosiewicz et al., 2001a, 2001b) and in specific cell lines (Mumtaz et al., 2002).

---

**a. Representative Activities**

**OPPTS.** There are several specific applications of genomics technologies that may improve risk assessment of mixtures in the future.

- The toxicity of a test mixture could be evaluated through effects on genomic biomarkers that have been linked to adverse effects.
- Genomics technologies could be applied to the evaluation of constructed test mixtures to examine how chemicals may interact (additivity, synergism, antagonism).
- Genomic diagnostic indicators may help identify components of chemical mixtures with unidentified constituents through the use of a genomic “fingerprint” database. For example, genomics technologies could be applied to evaluating the differences in gene expression between a mixture of known chemical constituents and a test mixture containing unknown chemical constituents. Differences in response to the two mixtures could then be evaluated and attributed to unknowns in the test mixture. This approach could be applied to exposure assessments.

**b. Additional Activities**

Genomics may have additional regulatory implications for evaluating chemical mixtures for other offices, as well as regions, states, and tribes. Further details on the following activities are found in Appendix B.

*Program Offices*

- OSWER, OAR, OW: identifying unknowns in mixtures at contaminated sites and in media, or during monitoring

*Regions, States, and Tribes*

- Identifying unknowns in mixtures at contaminated sites during monitoring
-



---

## IV. Research Needs and Activities

### A. Introduction

The Genomics Task Force identified a number of research needs that should be addressed in order to fully utilize genomics data in decision making, including:

- a) the link between molecular indicator, exposure, and adverse outcome needs to be established
- b) the genomic dose-response curve needs to be delineated
- c) inter- and intraspecies variations in response need to be quantified
- d) detection limits and variability for the genomic indicator need to be established
- e) the normal variability of gene expression, protein, and metabolite profiles needs to be understood to evaluate any changes induced by stressors
- f) baseline genomic responses of species to stressors other than chemical challenges need to be developed in establishing the utility of specific genomic indicators as markers of responses
- g) the complexity of the toxicological data bases that are likely to be developed by EPA's ORD will require new computational approaches for their analysis

It should be recognized though that there should not be a "higher standard" for use of genomics data in decision making than for other more traditional types of data typically used.

Within ORD, much research activity is directed toward the identification of gene expression changes in cells and, to some extent, in tissues in response to environmental chemical exposure. To date, the majority of this research involves the assessment of changes at the transcriptional level using mRNA microarrays rather than at the translational level using proteomic approaches. At this relatively early stage of the use of molecular profiling techniques, the majority of the effort in Agency research is to establish reproducibility and consistency of data for single molecular profiles for a time, concentration, species, cell type, or tissue set of parameters. The functional aspects of genomics (i.e., relating gene expression changes to cellular perturbations) will be addressed in the next phase of the research program. The overall goal of ORD's overarching Computational Toxicology Research Program is to use emerging technologies to improve quantitative risk assessment. Readers are encouraged to review the Framework (USEPA, 2003b) for additional details on ORD's Computational Toxicology activities at: [http://www.epa.gov/comptox/publications/comptoxframework06\\_02\\_04.pdf](http://www.epa.gov/comptox/publications/comptoxframework06_02_04.pdf).

---

The following section outlines genomics research needs and the current or planned Agency genomics research activities. In addition, it delineates how the Agency plans to use the data generated to address the types of regulatory and risk assessment activities described in Sections II and III.

## **B. Research Needs and Activities for Regulatory Applications**

### **1. Prioritization Research Needs and Activities**

The overall aim of EPA in prioritization efforts is to establish which chemicals (or chemical classes) and microbial contaminants warrant a more rigorous scientific investigation leading up to a risk assessment. To this end, the gaps in the Agency's genomics initiative are more in what is actually being conducted, rather than in what is planned. For example, to date there has been a limited research effort with wildlife species because of the paucity of genome sequence and genetics information for the majority of such species. Attempts are being made to help rectify this, in part through a collaboration with the Department of Energy's Joint Genome Institute to provide cDNA libraries of the fathead minnow and a frog species (e.g., *Xenopus tropicalis*, the Western clawed frog). Similar efforts with other species will be needed so that a set of sentinel species can be made available for environmental assessment.

Research is needed to develop a functional approach to establishing a linkage between genomic changes at the mRNA level and protein changes and cellular and tissue changes. This type of approach can provide information necessary for eventually developing a systems biology approach for defining pathways to disease or other adverse outcomes. These research needs will require a strategic hiring process so that the necessary expertise is available within the Agency. Collaborations will also be required to address the complex issues associated with a systems biology approach. Analysis of the large data sets generated through genomic assays will require the development of bioinformatic and computational methods. In addition, enhanced QSAR methods are required because the judicious use of QSAR approaches can greatly reduce the reliance on experimental approaches for establishing chemical priorities for additional research.

The initial steps in developing approaches for prioritizing chemicals for additional research that could lead to the development of a risk assessment are described in the Agency's Computational Toxicology initiative. The overall approach requires the development of toxicological pathways for candidate chemicals such that the key events leading to specific adverse outcomes can be identified. Chemicals can be designated as requiring further research if they are predicted to initiate the key events for a particular adverse outcome.

The Agency's Computational Toxicology initiative stems from an FY02 Congressional mandate to explore alternatives to the use of animals in toxicological studies. To address this mandate, ORD is using endocrine disrupting chemicals (EDCs) as model compounds in research that includes *in silico*, *in vitro*, and *in vivo* approaches as proof-of-concept for the overall

---

---

approach. The objectives of this effort are to determine the feasibility of using genomics and computational toxicology to facilitate the prioritization of chemicals for screening. Another goal is to reduce the need for some *in vivo* assays while providing a greater breadth of coverage of endocrine alterations and a better predictiveness of potential adverse health outcomes.

Similar approaches are being pursued for other chemical classes and other adverse outcomes (e.g., disinfection by-products and cancer, and conazoles and cancer, reproductive, developmental, and neurological effects). The aim is to establish a priority for chemicals that require further study for development of risk assessments.

Another example of ongoing research in support of prioritization involves determining the effects of contaminants on aquatic animals using protein profiling. This will result in the ability to rapidly screen chemicals based on specific modes of action. Protein profiles can be used as specific biomarkers of effects of bioactive compounds. The sheepshead minnow is being used as a model species for proof-of-concept studies.

A necessary component for the development of toxicological pathways in support of prioritization (and for MOA and risk assessment approaches) is the establishment of standardized approaches for conducting genomic and proteomic studies including data acquisition, data storage, and bioinformatics approaches. Efforts to achieve these goals are underway within the Agency.

## **2. Monitoring Research Needs and Activities**

The research needs to support EPA's monitoring programs cover a broad range of issues, including environmental monitoring and public health monitoring, specifically in the areas of epidemiology and molecular epidemiology. It is likely that genomics technologies will prove productive in each of these fields. The following examples provide an indication of the types of research needs that the Agency faces in an expanded monitoring program and highlight Agency activities that may help to address these needs.

If environmental and human health assessments are to employ biological indicators, reliable and informative markers of exposure, dose, and response must be selected. The more that is known about the mechanisms involved in the pathway from exposure to adverse outcome or response, the more readily informative biomarkers can be identified. The research need, therefore, is to develop these mechanistic data, select proposed informative biomarkers, and utilize these in field conditions or in molecular epidemiological studies. Validation of the biomarkers can be achieved through human or environmental health assessments, to establish how accurately the various biomarkers predicted outcomes. This is a massive effort that will require considerable collaboration within the Agency and outside the Agency.

EPA is conducting research to address the need for environmental biomarkers. For example, Agency scientists are conducting a comparison of the sensitivity of cellular indicators

---

of genetic damage in model stream fish using controlled laboratory exposures and subsequent field validation. These indicators of cellular damage are likely to have an application in ecological monitoring projects. Another example of ongoing research in this area involves the development of methods for measuring the induction of the vitellogenin gene during water monitoring studies. This indicator provides information on exposure of an organism to an endocrine disrupting chemical. The next step is to conduct research that can help establish the relationship between measured exposure and adverse effect. Such information is required prior to the use of the method in a regulatory setting.

Similarly, for microbial source tracking, the relationship between existing indicators (e.g., total coliforms, enterococci, etc.) and genomics-based indicators must be established. The relationship between occurrence and disease response in humans from human and non-human sources of pathogens, especially bacteria, must be defined. These types of information will be used by state and local agency decision makers to determine which methods are "acceptable." This is a significant step because most agencies have limited resources and are often reluctant to change to new technologies because of the associated high capital and human resource investments that must be made. The Agency is evaluating ways to apply DNA-based technology to detect and track fecal contamination to its source in complex environmental matrices, including recreational and drinking water sources. A microarray method to identify potential waterborne pathogens is also under development by the Agency. The Office of Water is currently supporting research investigating the effects of specific gene combinations that are associated with waterborne pathogen virulence. These projects could be applied to ambient and drinking water monitoring.

The development of genetic markers and/or proteomic markers of plant responses to herbicides and other xenobiotics is needed by the Regions to enhance monitoring capability for assessing effects of spray drift and determining which plant groups are most likely to be at risk from xenobiotics. Researchers are currently studying proteomic responses of plants to high potency, low-dose herbicides as a method of monitoring exposure. Further, markers for monitoring gene transgression from genetically modified (GM) crops to non-crop plants as well as for detecting contamination of non-GM seed shipments or foodstuffs with GM material, are being developed for use by the Regions.

Thus, the aim of a genomics research program in support of the Agency's monitoring efforts is to develop informative biomarkers of response for the assessment of the impact of environmental exposures on human and environmental health. In parallel, such bioindicators can be used to assess the impact of regulatory actions on human and environmental health.

---

---

## C. Research Needs and Activities for Risk Assessment

### 1. Mode of Action Research Needs and Activities

There are numerous issues associated with MOAs that require additional research before genomics technologies can be fully utilized in risk assessments. An overriding issue that affects more than just the MOA is the need to relate changes in gene expression to adverse effects. To establish the linkages between genomic changes and adverse outcomes, reliable data sets for gene expression at the RNA and protein levels are required. These data need to include a range of sample times and exposure concentrations and be repeatable. A parallel need is the development of expertise in analyzing these data so that profiles of responses at the molecular level can be produced and linked to specific chemicals or mixtures. This type of correlation then has to be extended to adverse outcomes at the organ, tissue, and whole animal level. This approach could be applied in both human health and ecological risk assessments.

Metabolic pathways for chemicals need to be defined, and the active metabolites that cause cellular responses need to be identified. The use of mRNA or protein arrays would enhance the Agency's ability to address this issue in a timely fashion. The initial requirement is to establish a set of profiles that are *in toto* representative for each known metabolic pathway.

Genomics-based approaches, including proteomic and metabonomic tests, will need to be developed to reduce, refine, or replace more complex and costly standard tests for human and wildlife species. Public pressure to reduce reliance on animal testing, particularly for toxicological studies, will continue to increase, making this a relatively high Agency priority. The overriding research need is the development of molecular profiles for *in vitro* cellular models and for a suite of animal species exposed to chemicals. The aim is to identify key components of MOAs from such profiles for comparison with similar profiles in humans and wildlife species. The long-term goal is to determine whether molecular profiles can be used to evaluate risk levels for chemicals with little toxicological information and for nontested species. A high priority, long-term research goal of the Interagency Coordinating Committee on the Validation of Alternative Test Methods (ICCVAM) is to investigate the utility of genomics for the assessment of acute toxicity, especially for the prediction of NOAELs and LOAELs. EPA's membership in ICCVAM will promote collaboration with other federal agencies to achieve this goal.

The following are descriptions of some of the research activities that are underway in ORD that address MOA. The list is not exhaustive, but does provide an idea of the breadth of the activities under the MOA umbrella. A significant fraction of the genomics research currently ongoing in the Agency is directed towards identifying MOAs for a range of chemicals for a number of different adverse outcomes, including cancer, endocrine disruption, reproductive and developmental effects, and neurotoxicity. These research activities are briefly presented to show their range.

---

In an attempt to use MOAs in the harmonization of risk assessment approaches, a research team is comparing the effects of a group of conazole pesticides in different tissues and for different endpoints. The aim is to establish if a common MOA is able to explain the range of different endpoints and the specificity of tissue responses.

Genomics technologies are also being used to characterize the MOA of selected drinking water disinfection by-products for use in BBDR models. In the same studies, genomics tools are being applied to develop markers of response that will provide information for predicting adverse outcomes at low doses.

Gene array technologies are being used to identify biomarkers that will be informative of responses specific to the human testis. The MOA that is being developed for rodents will be used to establish whether responses to particular chemicals have relevance to humans. Other studies are developing various biomarkers of response for environmental monitoring with relevance to humans.

Additional efforts are underway to establish if readily available cells in humans, such as peripheral lymphocytes or buccal cells, can be used as predictors of adverse responses in tissues that are targets for adverse outcomes such as cancer, and reproductive and developmental effects. Initial approaches involve the use of microarrays to study gene expression patterns in lymphocytes and in germ cells using appropriate animal models and selected chemical stressors. In a similar way, genomics approaches are being used to establish whether markers of susceptibility identified in readily available cells in humans, such as peripheral lymphocytes or skin fibroblasts, can predict sensitivity to adverse health outcomes. These initial gene expression studies will be expanded to include protein changes and functional associations with exposures. Currently, these types of approaches are in the early stages of development within the Agency.

As a part of the current effort in genomics, several groups are attempting to identify informative biomarkers of response in laboratory animals as well as in sentinel species for use in ecological assessments. It is proposed that MOA is a viable way of conducting interspecies comparisons of outcome.

The current and proposed Agency genomics research directed towards enhancing our knowledge of the various MOAs whereby chemicals can induce adverse outcomes is focused on identifying key events along toxicological pathways from exposure to response. The identification of key events will not only aid risk assessment approaches for single chemicals, but will enhance efforts to harmonize risk assessment, to predict responses to chemical mixtures, and to identify susceptible populations.

## **2. Susceptible Populations and Life Stages Research Needs and Activities**

Before the issue of how to incorporate susceptible populations into human health or ecological risk assessment can be addressed through EPA policy, the methods for identifying

---

---

susceptible populations must first be developed along with quantitative methods for assessing the magnitude of the sensitivity. To accomplish this, the Agency needs to develop knowledge of the MOA for a chemical(s) of concern as well as the prevalence of this MOA in the population. For example, this will require establishing the frequency of SNPs and their effects within human populations as part of the identification of a susceptible population. An important proviso to these types of studies is that ethical, social, and legal issues need to be addressed before starting such work.

There is a significant research effort in ORD to address the issue of children as a susceptible lifestage. Specific focus regards the induction of diseases in children and the effects of early life exposures on the development of adult diseases. The aims are to determine the magnitude of any sensitivities and the underlying mechanisms that might account for increased sensitivity. The genomics research component is directed towards developing informative biomarkers of response that can eventually be used in animal model systems to predict adverse outcomes from specific exposure scenarios and in human epidemiological studies. These informative biomarkers can also encompass specific genetic markers such as SNPs.

Genomic methods, including proteomic approaches, may also be useful in more accurately estimating exposures of individuals to contaminants in the environment, thereby identifying susceptible populations at the exposure level. These emerging technologies could lead to the development of personal dosimeters for a wide range of chemicals such that exposure would be assessed at the individual level. Similarly, genomic-level biomarkers (e.g., enhanced personal microarray technologies) could provide a real-time, high throughput method for screening potentially exposed individuals for incipient effects. While these approaches are technologically feasible, the Agency has no definitive plans to develop research programs along these lines in the near-term.

Sensitive fish and/or wildlife species might serve as early indicators of overall ecosystem health and as sentinels for risks to human health. In cases where chemical or pathogen contamination reduces species fitness, genomics technologies could be used to examine the genetic makeup of a species in order to determine the biochemical mechanism of the adverse effect(s). Like humans, plants and animals possess genetic polymorphisms that code for multiple metabolic enzyme variants. In addition, levels and forms of the same enzyme (e.g., the cytochrome P450 family of enzymes) vary between species and between life stages within species. Thus, as the genomes of species are sequenced, genomics can be used to identify the most sensitive species and sensitive life stages. This will significantly enhance our ability to set scientifically defensible water quality criteria or sediment and soil protection values under the Clean Water and Safe Drinking Water Acts.

Similarly, understanding genetic-based differences among plants and wildlife species in terms of the MOAs of chemicals is a fundamental step towards understanding which one(s) will be responsive. In the long term, this will enable more accurate cross-species extrapolations and will significantly reduce the need for animal testing. The Computational Toxicology initiative in

---

ORD is directed towards utilizing genomics to identify toxicity pathways. The current focus is on humans and fish. In the future, genomic-based approaches need to be developed for other wildlife species, as well as for aquatic and terrestrial plant species.

ORD is currently developing tools to incorporate genomics technologies into population dynamics models to enhance their predictive and explanatory power for assessing risks to populations of wildlife and aquatic life. Genetic processes include the distribution and dynamics of neutral and fitness-linked genetic markers. Depending upon their sophistication and data requirements, the resulting population models can be used in screening to definitive tiers in the ecological risk assessment process. In addition, a genetic dissection of the mechanisms of resistance to anthropogenic contaminants is underway in zebrafish and fathead minnows. Both of these research efforts will yield information regarding the sensitivities of various fish species and will likely be helpful in projecting the potential impacts of environmental contaminants on ecosystem health.

Similar approaches need to be considered for human populations. Here the need is to establish the overall effect of environmental exposures on human health. This will require knowledge of susceptible populations in terms of both the frequency and magnitude of sensitivity. The use of genomics to aid in the development of informative bioindicators for this effort is essential.

In the context of assessing the impact of chemical exposures to overall human and ecological health, the influence of susceptible populations is of critical importance. The needs are to consider the roles of both life stages and genetic variation in the etiology of susceptibility. While there is ongoing research addressing these issues, it is currently relatively limited.

### **3. Mixtures Research Needs and Activities**

Exposures of human and wildlife populations to environmental contaminants generally involve complex mixtures of chemicals, rarely individual chemicals. Although there have been some efforts to address responses to both simple and complex mixtures, much of the past and current research of the Agency has addressed the risk from exposures to single chemicals. Clearly, addressing the overall toxicological responses to mixtures is a complex problem that may require approaches different from those used for single chemicals. Given the charge to the Agency to increase its focus on research into the effects of mixtures, it is important to assess how genomics techniques might aid in meeting this need.

A range of research is needed to assess the risks of chemical mixtures. For example, it is necessary to determine if the MOA approach discussed above can be used to determine whether a mixture can induce a qualitatively different set of key events than any of the individual chemicals constituting the mixture. The next step would be to determine whether there are quantitative differences in the induction of key events by mixtures as compared to the individual chemicals and to use genomic measures to assess the magnitude of these definitive key events.

---



Because both human health and ecological risk assessment could benefit from a genomics approach, discussion is underway concerning how to incorporate this type of research into the Computational Toxicology initiative.

The research needs for mixtures overlap considerably with those of prioritization, MOA, and susceptible populations. Thus, mixtures assessment is an issue that will need to be addressed in concert with these other Agency priority regulatory needs.

---

## V. Challenges and Recommendations

As noted throughout this document, advances in genomics have significant implications for risk assessment practice and regulatory decision making. The use of genomics technologies generates a large volume of data and the field of bioinformatics is evolving rapidly to meet data analysis needs. The Agency's Interim Policy on Genomics (USEPA, 2002a) appropriately acknowledges that genomics technologies have the potential to improve our understanding of an organisms response to stressors. This information, in turn, can lead to the development of predictive biomarkers of effect and allow the identification of potentially sensitive populations and earlier predictions of adverse outcomes. Early detections can be converted into more effective intervention strategies. Genomics technologies will also enhance the understanding of the molecular mechanisms of toxicity. This will significantly reduce the uncertainty of extrapolations used in the risk assessment process. The result will be the development of more sensitive and cost-effective methods for toxicity screens and tests. Although, as the Interim Policy states, understanding genomic responses with respect to adverse ecological and/or human health outcomes is far from established, it is important for managers to begin to consider the likely future impacts of genomics technologies on their programs.

Though not detailed in this paper, bioinformatics (including data analysis and interpretation) is one of the biggest challenges to the effective use of genomics information.

Chemical production is highest in the Organization for Economic Cooperation and Development (OECD) countries and that growth is fastest in specialty chemicals and the life science sectors (OECD, 2001). Moreover, innovation in new chemical development and manufacturing practices is extremely high due to advances in combinatorial chemistry, nanotechnology, and biotechnology. These changes have raised concerns about the Agency's ability to sustain its current approaches to prioritization, monitoring, and risk assessment activities.

This paper outlined the potential of genomics technologies to improve and refine the current approach to regulatory applications and risk assessment and identified genomics as a means to alleviate the above concerns. There are, however, a number of challenges that must be overcome in order for genomics technologies to be fully applied to regulatory decision making. In this regard, the Genomics Task Force identified three categories of overarching scientific and resource challenges: research, technical development, and capacity. To address the regulatory and risk assessment applications outlined in this paper and to most effectively use genomics information, the Agency needs to address these challenges.

---

---

## A. Research Challenges

### 1. Linking Genomics Information to Adverse Outcomes

Linking genomics changes to adverse outcomes represents a significant research challenge that must be addressed before genomics data can provide information essential to the support of risk assessment and regulatory decision making. Additionally, establishing a quantitative relationship between gene expression changes and adverse responses will provide essential information. As noted throughout this paper, changes in gene expression at the mRNA and protein levels need to be related to cellular effects and, ultimately, to adverse outcomes. In many ways, the detection of gene expression changes is the easiest part of a genomic assessment. However, in a risk assessment framework, it is necessary to link a function to genes whose expression is altered. Gene expression changes that encompass defense mechanisms, which may be adaptive or beneficial and bear no causal relationship with the development of pathologies, must be separated from those that damage key cell functions (e.g., cell cycle control, structural integrity of proteins, control of free radicals, or loss of homeostasis and DNA repair mechanisms). Combining the findings of gene expression studies with data from *in vivo* chemical exposure of genetically altered animal models (e.g., knockout or null mice) is a powerful way to link specific genes to specific detrimental outcomes. Simpler whole organ systems may also offer powerful means to link genomic response to adverse effect. Key biological systems have fundamental genomic processes, some of which, if altered, are universally deleterious.

### 2. Interpretation of Genomics Information for Risk Assessment

Genomics information can be very relevant, and at times critical, to Agency risk assessments by providing mechanistically oriented insight into the hazard identification, dose-response, and exposure portions of risk assessments. This paper outlined specific areas where genomics data may aid in risk assessment including MOA assessment, identification of individual and population susceptibility, application of biomarkers of exposure, evaluation of effects of mixtures, understanding gene-environmental interactions, and application of a systems-wide examination of responses to stressors.

As a major example of how genomics information can provide insight for risk assessment, the mode of action of a stressor has been discussed. Mechanism or mode of action information can help identify potential hazards and help interpret dose-responses and extrapolations. Genomics technologies can be used to better understand the MOA of a chemical agent, and thus lead to advances in human and ecological risk assessments of chemicals. As genomics information contributes to our understanding of MOAs, the validity of using genomics information is in turn enhanced as an indicator of both adverse effects and exposure.

---

Providing links between genomics changes and phenotypic changes at the cell and tissue levels requires the use of a number of rapidly evolving cellular and molecular techniques (e.g., immunocytochemistry, gene silencing) and bioinformatic technologies. New approaches using tissue microarrays will enhance throughput and the linking of genomic and cellular outcomes. However, approaches to unraveling a profile of gene expression linked to a significant toxicological event present a number of challenges (Fielden and Zacharewski, 2001) in part because of the magnitude of the data sets developed and the potential variations in the level of expression for a single parameter (e.g., expression of a single gene). Analysis of the large data sets generated via genomics assays will require the development of new bioinformatic and computational tools. An integrated analysis and understanding of biological systems and their responses to perturbation, from genes to adverse effects, and the capacity to collect and evaluate data supportive of such a view, would be expected to greatly enhance the risk assessment process, and thus aid in formulating regulatory policy and making regulatory decisions.

Understanding the MOA of environmental agents that induce toxic effects other than cancer or induce carcinogenicity in animal models should facilitate the assessment of the relevance of these findings in protecting human health and safe guarding the environment. An important issue for extrapolation of responses in animal models to humans or environmental endpoints is to establish whether the MOA in the test species is relevant in the target species. A range of different types of data can be used to establish a MOA, but the endpoints for cross-species extrapolation generally are more limited. Such approaches will aid in addressing EPA's challenge to harmonize risk assessment approaches for different outcomes.

### **3. Recommendations to Address Research Challenges**

In order to contribute to the development of linkages between genomic changes and adverse outcomes and to the interpretation of genomics information for risk assessment, the Agency should aggressively support and build its own genomics research through the ORD Computational Toxicology initiative and support external research through competitive grants and contracts. Research plans and timing should be guided by developments in the genomics field. Through appropriate direction of its research, EPA can support important regulatory applications that are more likely to arise in the near future. These include priority-setting activities such as high throughput screening of chemicals, and monitoring activities such as source tracking of pollutants and pathogens in water. EPA should also encourage industry efforts to conduct genomics research necessary for application in risk assessment, when developing its data. It is critical that the Agency coordinate research with other agencies and institutions to link genomic changes and adverse outcomes

---

---

## **B. Technical Development Challenges**

### **1. Framework for Analysis and Acceptance Criteria for Genomics Information**

EPA acknowledges that genomics technologies will eventually contribute to risk assessment through a better understanding of mechanisms of chemical toxicity, dose-response relationships, identification of susceptible populations, and estimates of uncertainty factors. However, to date, EPA has had limited access to relevant genomics data to begin examining its potential influence. Even without specific cases relevant to the Agency, it is clear that a plan is needed to develop methods for incorporating these types of information into the decision making process.

EPA and other regulatory agencies recognize the requirement to develop acceptance criteria for genomics data. The Food and Drug Administration's (FDA) Center for Drug Evaluation and Research is in the process of identifying which pharmacogenetic data developed by companies for the evaluation of human drugs will be required by the FDA due to the regulatory implications of the data (FDA, 2003). The FDA is consulting with its Federal Advisory Committee and stakeholders to develop its policies and guidance. The efforts at the FDA are aimed at developing a framework for submission, storage, analysis, and regulatory review of genomics data. This is driven, in part, by the high level of use of genomics information for human drug development and evaluation.

The use of genomics information for the analysis of risks of industrial chemicals, agrochemicals, and microbial contaminants is not as widespread at the current time. However, as the understanding of genomics data and its relevance and applicability to chemical and microbial risk analysis increases, EPA will need to continue developing its own technical framework for the consideration of genomics information for scientific and regulatory purposes. The full range of ethical and legal issues (Marchant, 2003) will also have to be considered as genomics is incorporated into the Agency's risk assessment process.

### **2. Recommendations to Address Technical Development Challenges**

The Genomics Task Force recommends that the Agency charge a workgroup with developing a technical framework for analysis and acceptance criteria for genomics information for scientific and regulatory purposes. This framework should build upon EPA's Interim Policy on Genomics. Issues that need to be considered in developing such a framework include consideration of the performance of assays across genomic platforms (e.g., reproducibility, sensitivity) and the criteria for accepting genomics data for use in a risk assessment (e.g., assay validity, biologically meaningful response). It is essential for the Agency to continue to engage other federal agencies, such as the FDA, NIEHS, and Department of Energy, as well as other stakeholders, including industry, academia, and public interest groups when developing this framework. Such a framework, once established, can be used by the EPA program offices to

---

determine the applicability of specific genomics information to the evaluation of chemical risks under various statutes. The Interim Policy of considering genomics data on a case-by-case basis should continue to be applied until the technical framework is completed.

## **C. Capacity/Human Capital Challenges**

### **1. Applying Strategic Hiring Practices to Recruit Individuals Who Possess “Genomics Core Competencies”**

An important undertaking will be to identify the skills needed to establish “genomics core competencies” and to apply strategic hiring practices to recruit individuals who possess these skills. It will be essential to have technical specialists in genomics on staff in the crucial areas of research, analysis, systems biology, bioinformatics, and risk assessment to enhance the Agency’s expertise in genomics and related technologies.

### **2. Training EPA Risk Assessors and Managers to Interpret and Understand Genomics Data in the Context of a Risk Assessment**

It will be essential to train current EPA risk assessors so that they will be prepared to interpret and apply genomics data in the context of a risk assessment, including consideration of genomics data uncertainties. Risk assessors must be able to communicate both the underlying science and the interpretative tools and models used to develop the risk assessment to risk managers and stakeholders. Along similar lines, it will be important to provide training to risk managers regarding the use of genomics information in risk assessments and the strengths and limitations of such data.

A related concern is the capacity of regions, states, tribes, and local agencies to implement genomics tools and to evaluate genomics data, particularly with respect to their responsibilities under delegated programs. They will need resources, technical support, and targeted training to bring the scientists and managers within their organizations to a point where they will be able to effectively use genomics tools in their regulatory decision making, especially with respect to risk characterizations. Regional, state, tribal and local agencies' use of genomics tools will require both capital investment for analytical equipment, plus ongoing expenses for disposables such as the microarrays and associated supplies.

### **3. Recommendations to Address Capacity/Human Capital Challenges**

EPA programs and regions should apply strategic hiring practices to recruit individuals who possess genomics skills and should consider existing guidance from other agencies, such as the Centers for Disease Control, regarding recommended “genomics core competencies.”

---

The Genomics Task Force also recommends that the Agency convene a workgroup tasked with developing training modules for the interpretation and application of genomics data for risk assessments for both risk assessors and risk managers. It would be useful to develop and initiate training in the near future to prepare risk assessors and risk managers, because genomics issues have begun to arise in environmental decision making. The initial training could address basic genomics concepts, technologies and potential applications including consideration of the basic steps necessary to interpret and apply genomics data to a risk assessment. Regions 9 and 10, in collaboration with ORD, have already developed and conducted basic genomics training on a pilot basis. Development of Agency-wide training materials could build upon these efforts and those of external sources. It is expected that the training would need to be revised and expanded as our understanding of genomics improves over time. The training could also be offered on a limited basis to tribes, states, and local governments to assist in the development of their capability to effectively use genomics tools in regulatory decision making. Finally, the Genomics Task Force recommends that the development of training tools and workshops be conducted in collaboration with other agencies and institutions.

It is recognized that it may be difficult to allocate additional resources (e.g., funds, people, and training programs) required to efficiently incorporate and effectively understand these new types of genomics data. It is important for EPA not to lose focus on these important needs and ensure that adequate funds and people are brought to bear on this need.

**D. Summary Recommendation**

The Genomics Task Force recommends that EPA begin taking steps to address the identified research, technical development, and capacity challenges to strengthen its capability to effectively use genomics information (see table below for summary).

<b>Recommendations to Meet Challenges</b>	
<b><u>Research</u></b>	
a)	Encourage work on linking genomics changes to adverse effects
b)	Build EPA research capacity (e.g., Computational Toxicology program)
c)	Consider genomics capabilities in developing research plans
d)	Encourage others (e.g., industry) to conduct research applicable to risk assessment
<b><u>Technical Development</u></b>	
a)	Form technical framework workgroup(s), to develop: <ol style="list-style-type: none"> <li>1) acceptance criteria (e.g., data quality, experimental design)</li> <li>2) analysis guidance</li> <li>3) performance standards</li> </ol>
b)	Engage all stakeholders

**Recommendations to Meet Challenges****Capacity/Human Capital**

- a) Apply strategic hiring practices to recruit talented individuals with genomics skills
- b) Develop training tools, modules, etc. for different audiences (e.g., risk assessors, risk managers, non-technical users)
- c) Collaborate with others in workshops
- d) Work via budget process so adequate resources are available for these needs

It is essential for the Agency to continue to collaborate with other federal agencies, academia, the regulated community, public interest groups, and other stakeholders in this endeavor, in order to benefit from ongoing advances in genomics in the wider scientific and regulatory communities. Such collaborations can include development of training tools, workshops to ensure interaction of active parties, development of common approaches for use, storage, acceptance, and interpretation of genomics related information, and leveraging research resources to increase our research capabilities.

---



---

## REFERENCES

- Aardema, M.J., MacGregor, J.T. 2002. Toxicology and genetic toxicology in the new era of “toxicogenomics”: impact of “-omics” technologies. *Mutation Research* 499(1):13-25.
- Bartosiewicz, M.J., Jenkins, D., Penn, S., Emery, J., Buckpitt, A. 2001a. Unique gene expression patterns in liver and kidney associated with exposure to chemical toxicants. *Journal of Pharmacology and Experimental Therapeutics* 297:895-905.
- Bartosiewicz, M., Penn, S., Buckpitt, A. 2001b. Applications of gene arrays in environmental toxicology: fingerprints of gene regulation associated with cadmium chloride, benzo(a)pyrene, and trichloroethylene. *Environmental Health Perspectives* 109:71-74.
- Bishop, W.E., Clarke, D.P., Travis, C.C. 2001. The genomic revolution: what does it mean for risk assessment? *Risk Analysis* 21(6):985-987.
- Feron, V.J., Groten, J.P. 2002. Toxicological evaluation of chemical mixtures. *Food and Chemical Toxicology* 40:825-839.
- Fielden, M.R., Zacharewski, T.R. 2001. Challenges and limitations of gene expression profiling in mechanistic and predictive toxicology. *Toxicological Sciences* 60:6-10.
- Genter, M.B., Burman, D.M., Soundarapandian, V., Ebert, C.L., Aronow, B.J. 2002. Genomic analysis of alachlor-induced oncogenesis in rat olfactory mucosa. *Physiological Genomics* 12:35-45.
- Griffith, J.F., Weisberg, S.B., McGee, C.D. 2003. Evaluation of microbial source tracking methods using mixed fecal sources in aqueous test samples. *Journal of Water and Health* 1(4):141-152.
- Haber, L.T., Maier, A., Gentry, P. R., Clewell, H.J., Dourson, M.L. 2002. Genetic polymorphisms in assessing interindividual variability in delivered dose. *Regulatory Toxicology and Pharmacology* 35:177-197.
- Hakkola, J., Pelkonen, O., Pasanen, M., Raunio, H. 1998. Xenobiotic-metabolizing cytochrome P450 enzymes in the human fetoplacental unit: role in intrauterine toxicity. *Critical Reviews in Toxicology* 28(1):35-72.
- Hertzberg, R.C., MacDonell, M.M. 2002. Synergy and other ineffective mixture risk definitions. *Science of the Total Environment* 288:31-42.
-

Klaper, R., Euling, S., Frederick, R., Sonawane, R. 2003. Current use and future needs of genomics in ecological and human health risk assessment. Internal Workshop Report, May 8, Alexandria, VA. EPA/600/R-04/039.

Marchant, G.E. 2002. Genomics and environmental regulations: scenarios and implications. Report under Foresight and Governance Project at the Woodrow Wilson Center for Scholars. 29 pp. <http://www.environmentalfutures.org/Images/marchantwhitepaper.pdf>

Marchant, G.E. 2003. Genomics and toxic substances: Part I - Toxicogenomics. *Environmental Law Reporter* 33:10071-10093.

Mumtaz, M.M., Tully, D.B., El-Masri, H.A., De Rosa, C.T. 2002. Gene induction studies and toxicity of chemical mixtures. *Environmental Health Perspectives* 110 (Supplement 6):947-956.

Organization for Economic Cooperation and Development. 2001. The Chemicals Industry. In: Environmental Outlook, OECD Publications, Paris. pp. 224-225.

Pohl, H.R., Hansen, H., Chou, C.H.S.J. 1997. Public health guidance values for chemical mixtures: current practice and future directions. *Regulatory Toxicology and Pharmacology* 26:322-329.

Porter, W.P., Hinsdill, R.D., Fairbrother, A., Olson, L.J., Jaeger, J., Yuill, T.M., Bisgaard, S., Hunter, W.G., Nolan, K. 1984. Toxicant-disease-environment interactions associated with suppression of immune system, growth and reproduction. *Science* 224:1014-1017.

Sappington, L.C., Mayer, F.L., Dwyer, F.J., Buckler, D.R., Jones, J.R., Ellersieck, M.R. 2001. Contaminant sensitivity of threatend and endangered fishes compared to standard surrogate species. *Environmental Toxicology and Chemistry* 20(12):2869-2876.

Simpson, J., Santo Domingo, J.W., Reasoner, D.J. 2002. Microbial source tracking: state of the science. *Environmental Science and Technology* 36(24):5279-5288.

Snawder, J.E., Lipscomb, J.C. 2000. Interindividual variance of cytochrome P450 forms in human hepatic microsomes: correlation of individual forms with xenobiotic metabolism and implications in risk assessment. *Regulatory Toxicology and Pharmacology* 32:200-209.

Suk, W.A., Olden, K., Yang, R.S.H. 2002. Chemical mixtures research: significance and future perspectives. *Environmental Health Perspectives* 110 (Supplement 6):891-892.

U.S. Department of Energy, Office of Science, Biological and Environmental Research Division. [http://www.sc.doe.gov/ober/ERSD/ersd\\_nabir.html](http://www.sc.doe.gov/ober/ERSD/ersd_nabir.html)

---

U.S. Environmental Protection Agency. 1999. Glossary of Integrated Risk Information System. <http://www.epa.gov/iris/gloss8.htm>

U.S. Environmental Protection Agency, Science Policy Council. 2002a. Interim Policy on Genomics. <http://www.epa.gov/OSP/spc/genomics.pdf>

U.S. Environmental Protection Agency. 2002b. A Review of the Reference Dose and Reference Concentration Processes. EPA/630/P-02/002F. <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=55365>

U.S. Environmental Protection Agency. 2003a. Draft Final Guidelines for Carcinogen Risk Assessment. EPA/630/P-03/001A. <http://www.epa.gov/ncea/raf/cancer2003.htm>

U.S. Environmental Protection Agency. 2003b. A Framework for a Computational Toxicology Research Program in ORD. EPA/600/R-03/065. [http://www.epa.gov/comptox/publications/comptoxframework06\\_02\\_04.pdf](http://www.epa.gov/comptox/publications/comptoxframework06_02_04.pdf)

U.S. Environmental Protection Agency. 2003c. Pesticide product label system (PPLS). <http://www.epa.gov/pesticides/pestlabels/>

U.S. Food and Drug Administration. 2003. Guidance for Industry: Pharmacogenomic Data Submissions, Draft Guidance. <http://www.fda.gov/OHRMS/DOCKETS/98fr/2003d-0497-gd10001.pdf>

---

---

## Appendix A: Interim Policy on Genomics

### Science Policy Council

June 25, 2002

#### Background

The Environmental Protection Agency (EPA) is aware of the rapidly advancing field of genomics since completion of the initial sequencing of the human genome. EPA expects that genomics data may be received as supporting information for various assessment and regulatory purposes, e.g., identifying an environmental stressor's mode or mechanism of action. Genomic research tools now permit the study of gene and protein expression changes in various organisms and their cells, or tissues, with specificity to the level of molecular function. While genomics offers the opportunity to understand how an organism responds at the gene expression level to stressors in the environment, understanding of such molecular events with respect to adverse human or ecological health outcomes is far from established. Understanding these relationships becomes increasingly more complex as the number of sequenced genomes for an ever increasing variety of organisms becomes available. Thus, there is a need for the Agency to provide an interim policy on the current interpretation and utility of genomics data in the context of risk assessment and risk management and the implications this has for EPA's infrastructure needs.

Genomics approaches have the long term promise to aid in the understanding of an organism's response to stressors and to guide the selection of informative bioindicators for monitoring the impact of stressors on human and ecological health. Thus, EPA believes that genomics will have an enormous impact on our ability to assess the risk from exposure to stressors and ultimately to improve our risk assessments.

#### Science Policy Council's Interim Position

EPA believes that genomic data and analyses will significantly impact many areas of scientific research and human and ecological health assessments. Genomics data may allow EPA to enhance its assessments and better inform the decision-making process. Accordingly, EPA must understand how to develop and use the research tools made possible from genomics and understand the appropriate uses of genomics data to inform Agency decisions. For EPA, the term "genomics" encompasses a broader scope of scientific inquiry and associated technologies than when genomics was initially considered. A genome is the sum total of all an individual organism's genes. Thus, genomics is the study of all the genes of a cell, or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) levels.

Genomics methodologies are expected to provide valuable insights for evaluating how environmental stressors affect cellular/tissue functions and how changes in gene expression may relate to adverse effects. However, the relationships between changes in gene expression and adverse effects are unclear at this time and may likely be difficult to elucidate. Nonetheless, EPA believes that some of these changes may prove to be predictive of subsequent adverse effects. Changes in gene expression can be informative when a weight-of-evidence approach for human

---

and ecological health assessments is performed, particularly when used to explore the possible link between exposure, mechanism(s) of action, and adverse effects. In addition, genomics information may be useful to EPA in setting priorities, in ranking of chemicals for further testing, and in supporting possible regulatory actions. While genomics data may be considered in decision-making at this time, these data alone are insufficient as a basis for decisions. For assessment purposes, EPA will consider genomics information on a case-by-case basis. Before such information can be accepted and used, agency review will be needed to determine adequacy regarding the quality, representativeness, and reproducibility of the data.

EPA believes that genomics will ultimately improve the quality of information used in the risk assessment process. For example, genomics shows promise to identify variability and susceptibilities in individuals from exposed populations or among different species. It will also likely provide a better understanding of the mechanism or mode of action of a stressor and thus assist in predictive toxicology, in the screening of stressors, and in the design of monitoring activities and exposure studies. Application of genomics methodologies may help reduce or eliminate traditional types of toxicity testing as well as improve the scientific rationale for when such testing is needed. Genomic analysis also holds promise to evaluate the cumulative impacts resulting from the interplay of factors such as genetic diversity, health status, and life stage in responding to exposure(s) to multiple stressors.

EPA encourages further research on methods development, methods evaluation, and data collection to address existing gaps in knowledge concerning the consequences of genomic changes. EPA's goal is to develop knowledge that will ultimately reduce the uncertainties in the assessment of hazard, exposure, and risk from stressors. In parallel with data generation, there is an equal need for developing information technologies, for research on the analysis of data, and for applications of genomics data in computational toxicology. As EPA gains experience in applying genomics information and refines its understanding of the use of such information, it will develop guidance to explain how genomics data can be better utilized in informing decision-making and related ethical, legal, and social implications. EPA is working with other Federal, state, and Tribal organizations, as well as with academic, international, and industry groups to facilitate scientifically sound applications of genomics data. In addition, EPA will continue to build partnerships and communicate with all interested stakeholders as an essential component of the Agency's future activities in genomics.

---

---

## Appendix B: Details on Additional Activities Potentially Affected by Genomics Information

### I. Prioritization

#### A. Program Offices

**OPPT - Premanufacture Notices (PMN):** Genomics data may be useful for evaluating PMNs. Genomics data generated for PMNs may be able to supplement, or potentially supplant, computer model results or expert judgment in hazard estimation and prioritization activities.

**OPPT - Voluntary Children's Chemical Evaluation Program (VCCEP):** Genomics data may be useful in the VCCEP program. These data could support, or potentially supplant, current toxicity tests.

**OPPTS/OSCP - Endocrine Disruptors Screening Program (EDSP):** OSCP has been developing a tiered approach for testing under the EDSP. Currently, prioritization is primarily based upon exposure information, but it is anticipated that chemical prioritization will be greatly facilitated by the use of high throughput screening. Genomics could be used in the high throughput screening process.

**OPP - Prioritization of Pesticides and Inert Chemicals:** Genomics data could be useful for prioritizing pesticide products for testing procedures. For example, data demonstrating that an agent does not elicit differential gene expression that is predictive of toxicologically relevant responses may indicate that a pesticide is non-toxic in a particular test species. The chemical might then be slated to receive expedited review under the reduced risk chemicals program, and waiver requests for associated standard toxicity tests might be considered. Conversely, data indicating that a pesticide produces an altered gene expression profile for a toxicological pathway that is relevant for an adverse outcome may potentially signal an alert. The pesticide may then be assigned to an evaluation pathway involving a second level of genomics and standard toxicity testing.

**OW - Prioritizing Streams or Wetlands for Study or Clean Up:** Genomics data could be applied to prioritizing streams and wetlands for additional study or clean up activities.

**OAR - Hazardous Air Pollutants:** Genomics data may be useful in prioritizing hazardous air pollutants for chemical testing.

**OSWER - Superfund:** Superfund site prioritization may be enhanced through the use of genomics data.

---

**ORD- Research Planning:** Genomics data may provide useful information to ORD researchers and managers in prioritizing chemicals for future research.

## **B. Regions, States, and Tribes**

**Resource Prioritization for Site Remediations and Chemical Evaluations:** Genomics technologies could provide regions, states, and tribes with a fast, relatively low cost method to prioritize their areas of focus and deployment of resources for delegated program site remediations and chemical evaluations.

## **II. Monitoring**

### **A. Program Offices**

**OPP - Exposure Monitoring:** OPP uses exposure monitoring data for new chemicals generated via Experimental Use Permits (EUPs) in registration decisions, and exposure monitoring data for currently registered chemicals for reregistration eligibility decisions (REDs). For these EUPs and REDs, genomics data could be used to track the movement of pesticides off-site via spray drift or into ground or surface water. Biological monitoring data for human and wildlife exposures and potential effects could contribute to regulatory decisions, and genomics technologies could provide information about occupational exposures and wildlife incident data for reregistration decisions.

**OAR/Office of Air Quality Planning and Standards (OAQPS) - Stationary Source Monitoring:** Genomics technologies could contribute to stationary source monitoring conducted under the Clean Air Act (CAA).

**OSWER - Superfund Monitoring:** Genomics technologies could be applied to monitor movement of contaminants off-site from Superfund sites prior to remedial actions. In evaluating contaminated sites, near term benefits could be derived by targeting toxic chemical remediation using biomarkers present in lower organisms (e.g., metallothionein expression). Bioavailability, a key parameter in determining the toxicity of a chemical, can be effectively determined using genomics tools; thus greater precision could be achieved in remedial activities. That is, the truly hazardous compounds could be identified and removed with more precision, and other materials need not be disturbed unnecessarily. Further, monitoring the operation and maintenance of remedial actions and residual contaminants at Superfund sites that have undergone cleanup could be enhanced through genomics technologies.

**OSWER/OSW - RCRA-Required Monitoring:** Genomics could contribute to post-clean up monitoring activities conducted under RCRA.

---

---

**OEI, ORD - Bioindicator Development:** OEI and ORD expect to use genomics approaches to develop a selection of informative bioindicators for monitoring the exposures and effects of stressors on human and ecological health.

## **B. Regions, States, and Tribes**

**State and Local Beach Closures - TMDL Issues Associated with Pathogens:** A possible near-term scenario is the use of genomics technologies to detect microbial pathogens and to determine their sources (so-called microbial or bacteriological source tracking, MST or BST). The Southern California Coastal Water Research Program, a State of California research agency that receives partial funding from waste water dischargers, has sponsored the first round of research addressing the feasibility of molecular-based MST techniques. Several Regions are working with this group and with ORD scientists to identify the "best performers" among the various MST techniques and to produce guidance on the effective use of use these methods. This guidance could be used by State and local agencies to make decisions about testing ambient surface waters for beach closures and establishing pathogen/bacterial TMDLs. Recently, however, EPA researchers evaluating MST methods have concluded these methods will require further development before they can be widely applied (Simpson et al., 2002).

**Air Quality Monitoring Program:** EPA's ambient air quality monitoring program for criteria pollutants is carried out by State and local agencies. Genomics approaches could be used in the state and local air monitoring programs.

**Endocrine-Disruptor Monitoring:** Another example of near-term decision making using genomics data comes from a group of tribes in Northern California and Southern Washington that is interested in the potential impact of exposure to pharmaceuticals and personal care products and the potential for endocrine disruptor effects from municipal waste treatment facilities and/or cattle grazing on the health of wild salmon populations that are part of tribal cultural and economic resources. The Tribe proposes to use a series of molecular-biology-based assays for exposure to hormonally active compounds, either a multiplex RT-PCR approach or a multigene array. The information could be used to develop NPDES permit limits and establish Tribal Water Quality Standards.

**Regional Pesticide Program Decisions:** Another use of genomics data may be in the area of the regions' pesticides programs. If genomics data show that a particular unregistered pesticide is of concern in certain populations of humans, animals, or plants, the region could work with the state to assign a higher fine to an incident of use and/or offering for sale if the violation was in an area where such a sensitive population existed. Alternatively, the "harm value" might be lower if the non-target groups in an area were not considered to be part of a sensitive population.

---



### **III. Risk Assessment - Mixtures**

#### **A. Program Offices**

**OSWER, OAR, OW - Site Assessments/Monitoring:** Genomics technologies could be used as diagnostic indicators to help identify components of chemical mixtures with unidentified constituents. This approach could be used in initial assessments of contaminated sites or in general monitoring applications.

#### **B. Regions, States, and Tribes**

**Site Assessments/Monitoring:** Regions, states, and tribes could apply genomics diagnostic indicators to identify components of chemical mixtures with unidentified constituents for initial assessments of contaminated sites or in general monitoring applications.

---

---

## Appendix C: Glossary

**Allele:** an alternative form of a gene or any other segment of a chromosome<sup>1</sup>.

**Bioinformatics:** the analysis of biological information using computers and statistical techniques; the science of developing and utilizing computer databases and algorithms to accelerate and enhance biological research<sup>1</sup>.

**Biomarker:** a molecular indicator of a specific biological property; a biochemical feature or facet that can be used to measure the progress of disease or the effects of treatment<sup>1</sup>.

**Biotechnology:** the set of biological techniques developed through basic research and now applied to research and product development. In particular, biotechnology refers to the use by industry of recombinant DNA, cell fusion, and new bioprocessing techniques<sup>2</sup>.

**Complementary DNA (cDNA):** DNA made from a messenger RNA (mRNA) template. The single-stranded form of cDNA is often used as a probe in physical mapping<sup>1</sup>.

**Computational Toxicology - Comp Tox:** the application of mathematical and computer models and molecular biological approaches to improve the Agency's prioritization of data requirements and risk assessments<sup>3</sup>.

**Deoxyribonucleic acid (DNA):** the substance of heredity; a large molecule that carries the genetic information that cells need to replicate and to produce proteins<sup>4</sup>. The nucleic acid that constitutes the genetic material of all cellular organisms and DNA viruses. The genetic information is used in the synthesis of ribonucleic acids (RNAs) from DNA templates (transcription) and in the synthesis of proteins from messenger RNA (mRNA) templates (translation).

**Expressed sequence tag:** a unique stretch of DNA within a coding region of a gene that is useful for identifying full-length genes and serves as a landmark for mapping<sup>1</sup>.

**Gene:** the fundamental physical and functional unit of heredity. A gene is an ordered sequence of nucleotides located in a particular position on a particular chromosome that encodes a specific functional product (i.e., a protein or RNA molecule)<sup>2</sup>.

**Gene chip technology:** development of cDNA microarrays from a large number of genes; used to monitor and measure changes in gene expression for each gene represented on the chip<sup>2</sup>.

**Gene expression:** process by which a gene's coded information is converted into the structures present and operating in the cell. **Expressed genes** include those that are transcribed into mRNA

---

and then translated into protein and those that are transcribed into RNA but not translated into protein (e.g., transfer and ribosomal RNAs)<sup>2</sup>.

**Genetics:** the study of inheritance patterns of specific traits<sup>2</sup>.

**Genetic testing:** analyzing an individual's genetic material to determine predisposition to a particular health condition or to confirm a diagnosis of genetic disease<sup>2</sup>.

**Genomics:** the study of all the genes of a cell or tissue, at the DNA (genotype), mRNA (transcriptome), or protein (proteome) level<sup>5</sup>.

**Genome:** all the genetic material in the chromosomes of a particular organism; its size is generally given as its total number of base pairs<sup>2</sup>.

**Genotype:** the genetic composition of an organism or a group of organisms; a group or class of organisms having the same genetic constitution<sup>1</sup>.

**Hazard Assessment:** the process of determining whether exposure to an agent can cause an increase in the incidence of a particular adverse health effect (e.g., cancer, birth defect) and whether the adverse health effect is likely to occur in humans<sup>6</sup>.

**In Silico:** literally “within silicon ”; refers to modeling research conducted with computers only<sup>7</sup>.

**In Vitro:** literally, “in glass,” i.e., in a test tube or in the laboratory; the opposite of in vivo (in a living organism)<sup>1</sup>.

**In Vivo:** in a living organism, as opposed to in vitro (in the laboratory)<sup>1</sup>.

**Knockout:** inactivation of specific genes. Knockouts are often created in laboratory organisms such as yeast or mice so that scientists can study the knockout organism as a model for a particular disease<sup>1</sup>.

**Mapping:** charting the location of genes on chromosomes<sup>1</sup>.

**Mass spectrometry:** a method used to determine the masses of atoms or molecules in which an electrical charge is placed on the molecule and the resulting ions are separated by their mass to charge ratio<sup>1</sup>.

**Messenger RNA (mRNA):** a type of RNA that reflects the exact nucleotide sequence of the genetically active DNA. mRNA carries the "message" of the DNA to the cytoplasm of cells where protein is made in amino acid sequences specified by the mRNA<sup>1</sup>.

---

---

**Metabonomics:** the evaluation of tissues and biological fluids for changes in metabolite levels that result from toxicant-induced exposure<sup>1</sup>.

**Microarray:** a tool used to sift through and analyze the information contained within a genome. A microarray consists of different nucleic acid probes that are chemically attached to a substrate, which can be a microchip, a glass slide, or a microsphere-sized bead<sup>1</sup>.

**Northern blot:** a technique used to separate and identify pieces of RNA<sup>1</sup>.

**Nucleotide:** a subunit of DNA or RNA. To form a DNA or RNA molecule, thousands of nucleotides are joined in a long chain<sup>1</sup>.

**“Omics”:** term including genomics, proteomics, metabonomics (some differentiate this term from metabolomics), transcriptomics, and associated bioinformatics<sup>8</sup>.

**Phenotype:** the observable physical or biochemical traits of an organism as determined by genetics and the environment; the expression of a given trait based on phenotype; an individual or group of organisms with a particular phenotype<sup>1</sup>.

**Polymorphism:** the quality or character of occurring in several different forms<sup>1</sup>.

**Proteome:** all of the proteins produced by a given species just as the genome is the totality of the genetic information possessed by that species<sup>1</sup>.

**Proteomics:** study of the full set of proteins encoded by a genome<sup>2</sup>.

**Risk Assessment:** a qualitative or quantitative evaluation of the risk posed to human health and the environment by the actual or potential presence of pollutants<sup>9</sup>.

**RNA:** a chemical found in the nucleus and cytoplasm of cells; it plays an important role in protein synthesis and other chemical activities of the cell. The structure of RNA is similar to that of DNA. There are several classes of RNA molecules, including messenger RNA, transfer RNA, ribosomal RNA, and other small RNAs, each serving a different purpose<sup>2</sup>.

**Signal transduction pathway:** the course by which a signal from outside a cell is converted to a functional change within the cell<sup>1</sup>.

**Single nucleotide polymorphism (SNP):** a change in which a single base in the DNA differs from the usual base at that position<sup>1</sup>.

---

**Susceptibility:** the increased likelihood of an adverse effect, often discussed in terms of relationship to a factor that can be used to describe a human subpopulation (e.g. life stage, demographic feature, or genetic characteristic)<sup>6</sup>.

**Susceptible Subgroups:** may refer to life stages, for example, children or the elderly, or to other segments of the population, for example, asthmatics or the immune-compromised, but are likely to be somewhat chemical-specific and may not be consistently defined in all cases<sup>6</sup>.

**Systems Biology:** a holistic approach to the study of biology with the objective of simultaneously monitoring all biological processes operating as an integrated system<sup>10</sup>.

**Systems Toxicology:** the study of perturbation of organisms by chemicals and stressors, monitoring changes in molecular expression and conventional toxicological parameters, and iteratively integrating biological response data to describe the functioning organism<sup>11</sup>.

**Throughput:** output or production, as of a computer program or a biological assay, over a period of time<sup>1</sup>.

**Toxicity:** deleterious or adverse biological effects elicited by a chemical, physical, or biological agent<sup>6</sup>.

**Toxicology:** the study of harmful interactions between chemical, physical, or biological agents and biological systems<sup>6</sup>.

**Toxicogenomics:** the study of how genomes respond to environmental stressors or toxicants. Combines genome-wide mRNA expression profiling with protein expression patterns using bioinformatics to understand the role of gene-environment interactions in disease and dysfunction<sup>2</sup>.

**Transgenic:** having genetic material (DNA) from another species<sup>1</sup>.

---

---

## Glossary References

1. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Environmental Health Sciences, National Center for Toxicogenomics, Glossary. <http://www.niehs.nih.gov/nct/glossary.htm>
  2. U.S. Department of Energy, Oak Ridge National Laboratory, Human Genome Project Information, Genome Glossary. [http://www.ornl.gov/sci/techresources/Human\\_Genome/glossary](http://www.ornl.gov/sci/techresources/Human_Genome/glossary)
  3. U.S. Environmental Protection Agency. A Framework for a Computational Toxicology Research Program in ORD. EPA/600/R-03/065. [http://www.epa.gov/nheerl/comptoxframework/comptoxframeworkfinaldraft7\\_17\\_03.pdf](http://www.epa.gov/nheerl/comptoxframework/comptoxframeworkfinaldraft7_17_03.pdf)
  4. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute, Understanding Gene Testing, Glossary. <http://www.accessexcellence.org/AE/AEPC/NIH/gene27.html>
  5. U.S. Environmental Protection Agency, Science Policy Council. 2002a. Interim Policy on Genomics. <http://www.epa.gov/OSP/spc/genomics.pdf>
  6. U.S. Environmental Protection Agency. 1999. Glossary of Integrated Risk Information System. <http://www.epa.gov/iris/gloss8.htm>
  7. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of General Medical Sciences, The Chemistry of Health Glossary [http://www.nigms.nih.gov/news/science\\_ed/chemhealth/glossary.html](http://www.nigms.nih.gov/news/science_ed/chemhealth/glossary.html)
  8. Henry, C.J., Phillips, R., Carpanini F., Corton, J.C., Craig, K., Igarashi, K., Leboeuf, R., Marchant, G., Osborn, K., Pennie, W.D., Smith, L.L., Teta, M.J., Vu, V. 2002. Use of genomics in toxicology and epidemiology: findings and recommendations of a workshop. *Environmental Health Perspectives* 110(10):1047-1050.
  9. U.S. Environmental Protection Agency, Risk Assessment Web Page. <http://www.epa.gov/ebtpages/enviriskassessment.html>
  10. Sumner, L.W., Mendes, P., Dixon, R.A. 2003. Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* 62:817-836.
  11. Waters, M., Boorman, G., Bushel, P., Cunningham, M., Irwin R., Merrick, A., Olden, K., Paules, R., Selkirk, J., Stasiewicz, S., Weis, B., Van Houten, B., Walker, N., Tennant, R. 2003. Systems toxicology and the Chemicals Effects in Biological Systems (CEBS) knowledge base. *Environmental Health Perspectives Toxicogenomics* 111(6):811-824.
-