



EPA 100/B-04/002
March, 2004
External Review Draft

U.S. Environmental Protection Agency

DRAFT

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**

NOTICE

This document is an **external review draft**. It has not been formally released by the U.S. Environmental Protection Agency and should not at this stage be construed to represent Agency position.

External Review Draft

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

Workgroup Co-Chairs

William Benson
NHEERL
Office of Research and Development

Kerry Dearfield
Office of the Science Advisor
Office of Research and Development

Workgroup Members

Michael Brody, OCFO
Anne Fairbrother, ORD
Kathryn Gallagher, ORD
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

DISCLAIMER

(To be added)

Table of Contents

FOREWORD Page vi

ACKNOWLEDGMENTS Page vii

ACRONYMS Page viii

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 10

 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 15

 1. Introduction Page 15

 a. Reporting on Adverse Effects of Commercialized Chemicals and
 Pesticides Page 15

 b. Toxics Release Inventory Program Page 15

 2. Regulatory Activities Potentially Affected by Genomics Information
 Page 16

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 23 |
| 1. Overview | Page 23 |
| a. Susceptible Human Populations and Sensitive Life Stages | Page 23 |
| b. Susceptible Wildlife Populations and Sensitive Life Stages ... | Page 24 |
| 2. Risk Assessment Activities Potentially Affected by Genomics Information | |
| | Page 25 |
| a. Susceptible Human Populations and Sensitive Life Stages | Page 25 |
| i. Representative Activities | Page 25 |
| b. Susceptible Wildlife Populations and Sensitive Life Stages .. | Page 27 |
| i. Representative Activities | Page 27 |
| D. Mixtures | Page 29 |
| 1. Overview | Page 29 |
| 2. Risk Assessment Activities Potentially Affected by Genomics Information | |
| | Page 30 |
| a. Representative Activities | Page 30 |
| b. Additional Activities | Page 31 |
| IV. Research Needs and Activities | Page 32 |
| A. Introduction | Page 32 |
| B. Research Needs and Activities for Regulatory Applications | Page 32 |
| 1. Prioritization | Page 32 |
| 2. Monitoring | Page 34 |
| B. Research Needs and Activities for Risk Assessment | Page 35 |
| 1. Mode of Action | Page 35 |
| 2. Susceptible Populations and Life Stages | Page 37 |
| 3. Mixtures | Page 39 |
| V. Challenges and Recommendations | Page 40 |
| A. Research Challenges | Page 40 |
| 1. Linking Genomics Information to Adverse Outcomes | Page 40 |
| 2. Interpretation of Genomics Information for Risk Assessment | Page 41 |
| 3. Recommendations to Address Research Challenges | Page 42 |
| 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 43 |
| 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 44

3. Recommendations to Address Capacity/Human Capital Challenges .. Page 44

References Page 46

Appendix I: Details on Additional Activities Potentially Affected by Genomics Information
..... Page 50

I. Prioritization Page 50

II. Monitoring Page 51

III. Risk Assessment Page 52

Appendix II: Glossary Page 54

FOREWORD

(To be added)

ACKNOWLEDGMENTS

(To be added)

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 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 Adverse Effect Level
OAQPS - Office of Air Quality Planning and Standards
OAR - Office of Air and Radiation
OCFO - Office of the Chief Financial Officer
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
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

1
2
3 Advances in genomics will have significant implications for risk assessment practice and
4 regulatory decision making. The Environmental Protection Agency's (EPA's) Interim Policy on
5 Genomics, issued in 2002, appropriately acknowledges that genomics technologies have the
6 potential to improve our understanding of an organism's response to stressors (USEPA, 2002a).
7 The Interim Policy describes genomics as the
8 study of all the genes of a cell or tissue, at the
9 DNA, mRNA, or protein level. This policy
10 states that while genomics data may be
11 considered in the decision making process at
12 this time, these data alone are insufficient as a
13 basis for decisions; EPA will consider
14 genomics information for assessment
15 purposes on a case-by-case basis only. Following release of the Interim Policy, EPA held
16 internal discussions regarding the potential of genomics approaches to improve our
17 understanding of the effects of environmental stressors on cells. It was concluded that genomics
18 information may lead to the development of predictive biomarkers of effect, thereby allowing for
19 the identification of potentially sensitive populations and earlier predictions of adverse outcomes
20 and, ultimately, leading to better intervention strategies. Enhancing understanding of the
21 molecular mechanisms of toxicity may greatly reduce the uncertainty of extrapolations used in
22 the current risk assessment process. Further, genomics technologies may enhance the
23 development of more sensitive and cost-effective methods for toxicity screens and tests and may
24 ultimately lead to the reduction, refinement, or replacement of more complex and costly standard
25 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

26
27 Following these internal discussions, at the request of EPA's Science Policy Council
28 (SPC) a Genomics Task Force was formed. The group was charged with the task of examining
29 the broader implications genomics is likely to have for EPA programs and policies, to attempt to
30 gain further understanding of the appropriate usage of these data and the potential consequences
31 of their use, as well as to identify possible infrastructure needs. The group was also charged
32 with developing scenarios to describe various circumstances under which EPA might receive
33 these data. The resulting document is intended to present implications of the use of genomics
34 technologies in EPA practice for the consideration of Agency managers. It is the intent of the
35 Genomics Task Force to initiate a discussion of the scientific issues regarding the incorporation
36 of genomics information into human health and ecological risk assessments and of how these
37 data will likely affect regulatory policy and decision making in the future. Although, as the
38 Interim Policy notes, understanding genomic responses with respect to adverse ecological and/or
39 human health outcomes is far from established, it is important for managers to begin to consider
40 the likely future impacts of genomics technologies on their programs. Four areas have been
41 identified as those very likely to be influenced by the generation of genomics information within
42 EPA and the submission of such information to EPA: (1) prioritization of contaminants and
43 contaminated sites, (2) monitoring, (3) reporting provisions; and (4) risk assessment. This

1 document also briefly addresses ongoing research within the Agency for each of these four areas
2 and identifies remaining research needs. It should be noted that genomics will not
3 fundamentally alter the risk assessment process, but is expected to serve as a new, more powerful
4 tool for evaluating the exposure to and effects of environmental stressors.
5

6 The Task Force identified the following overarching challenges associated with genomics
7 that fall into three categories: research, technical development, and capacity. These challenges
8 are defined as critical needs for the Agency to strengthen its capability to use genomics
9 information in a meaningful way, and to enable the Agency to address potential regulatory
10 applications that are likely to arise with respect to genomics, such as those outlined in this paper.
11 For research, the critical needs are identified as (1) linking genomics information to adverse
12 outcomes; and (2) interpreting genomics information for risk and hazard assessment. It is
13 important to note that significant research by EPA and other agencies and researchers will be
14 necessary to fully understand and apply genomics technologies to human health and ecological
15 risk assessment. One critical need in the area of technical development was identified as the
16 need to establish a framework for analysis and acceptance criteria for genomics information for
17 scientific and regulatory purposes (including data quality standards based on genomic assay
18 performance). Two critical needs were identified with respect to capacity, including human
19 capital: (1) applying strategic hiring practices to recruit individuals who possess “genomics core
20 competencies” essential for crucial areas of research, analysis, systems biology, bioinformatics,
21 and risk assessment; and (2) training EPA risk assessors and managers to interpret and
22 understand genomics data in the context of a risk assessment.
23

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

1 **I. Introduction**

2
3 **A. Background**

4
5 The mapping of the genomes of diverse animal, plant, and microbial species, and related
6 technologies are already significantly affecting research across all areas of the life sciences and
7 will continue to do so for decades to come. The current understanding of biological systems is
8 rapidly changing in ways previously unimagined, and novel applications of this technology are
9 already being commercialized. These scientific and technological advances have spurred many
10 federal agencies to consider the far-reaching implications for policy, regulation, and society as a
11 whole.

12
13 On June 25, 2002, EPA released the Interim Policy on Genomics (USEPA, 2002a)
14 communicating its initial approach to using genomics information in risk assessment and
15 decision making (<http://www.epa.gov/osp/spc/genomics.htm>). This policy describes genomics
16 as the study of all the genes of a cell or tissue, at the DNA (genotype), mRNA (transcriptome), or
17 protein (proteome) level. The Interim Policy notes that, while genomics offers the opportunity to
18 understand how an organism responds at the gene expression level to stressors in the
19 environment, understanding such molecular events with respect to adverse ecological and/or
20 human health outcomes far from established. It concludes that while genomics data may be
21 considered in the decision making process at this time, these data alone are insufficient as a basis
22 for decisions. Therefore, EPA will consider genomics information for assessment purposes on a
23 case-by-case basis only.

24
25 Following release of the Interim Policy, EPA held internal discussions to consider the
26 potential genomics technologies have to improve our understanding of the effects of
27 environmental stressors on cells. It was concluded that genomics information may lead to the
28 development of predictive biomarkers of effect, thereby allowing for the identification of
29 potentially sensitive populations and earlier predictions of adverse outcomes and, ultimately,
30 leading to better intervention strategies. Enhancing understanding of the molecular mechanisms
31 of toxicity could greatly reduce the uncertainty of extrapolations used in the current risk
32 assessment process. The potential results may be the development of more sensitive and cost-
33 effective methods for toxicity screens and tests and significant reductions in, or eventual
34 elimination of, conventional animal testing.

35
36 Following these discussions, at the request of the SPC a Genomics Task Force was
37 formed. The group was charged with the task of examining the broader implications genomics is
38 likely to have on EPA programs and policies, to attempt to gain further understanding of the
39 appropriate usage of these data and the potential consequences of their use, as well as to identify
40 possible infrastructure needs. The group was also charged with developing scenarios to describe
41 various circumstances under which EPA might receive these data and the resulting implications
42 (e.g., interpretation, relevance, evaluation, analytical, and research needs) for EPA policies and
43 programs.

B. Emerging Impacts of Genomics Technologies

1
2
3 While these are new technologies and most are not as yet ready for application in risk
4 assessment and decision making, it is important for Agency managers to begin to consider the
5 likely future impacts of genomics technologies on their programs. It should be noted that
6 genomics will not fundamentally alter the risk assessment process, but is expected to serve as a
7 more powerful tool for evaluating the exposure to and effects of environmental stressors and will
8 offer a means to simultaneously examine a number of response pathways. EPA and other
9 regulatory agencies are beginning to address the use of genomics data for various risk
10 assessment applications, including the need to establish a link between genomic alterations and
11 adverse outcomes of regulatory concern (Klaper et al., 2003). EPA must soon develop an
12 explicit prescriptive strategy for accepting “omics” data submissions because such information is
13 already being submitted. Given the rapidly evolving nature of genomics technologies, care must
14 be taken to develop an acceptable scheme to simplify and refine the risk-related information and
15 to distinguish it from the large amount of complex scientific and statistical data available. This
16 strategy must remain dynamic in anticipation of continuing technical evolution at the molecular
17 levels (e.g., DNA, RNA, and protein). Furthermore, bioinformatic approaches for data
18 acquisition and analysis, including technologies designed to store and analyze the profusion of
19 data generated from microarray analysis, must be considered in parallel with the data-generating
20 methods. Additionally, many scientific, policy, ethical, and legal concerns are developing along
21 with the emergence of this science and will need to be addressed.

22
23 The Interim Policy on Genomics provides guidance concerning how and when genomics
24 information should be used to assess the risks of environmental contaminants under the various
25 regulatory programs implemented by the Agency at the present time. The standardization of
26 experimental design and data analysis for genomics is important for the utility of genomics
27 information in future risk assessment and regulatory decisions. Such standardization will
28 enhance the reproducibility of results obtained and the reliability of conclusions drawn from
29 these data. Furthermore, EPA should consider the development of data quality standards based
30 on performance of microarrays, as well as other genomics technologies. This in turn will help to
31 ensure the integrity of EPA’s approach to assessing the genomics information submitted to the
32 Agency.

33
34 Genomics issues have already arisen in environmental decision making. For example, a
35 pesticide registrant has cited several published genomic articles as part of their data package
36 submission for product registration to EPA’s Office of Pesticide Programs. The data were
37 submitted to propose an alternative mode of action that would affect human health assessment
38 conclusions. Additional, similar submissions may soon be made by other pesticide registrants.

39
40 There are a number of other regulatory areas where genomics information will start
41 having an impact. For example, a research consortium including State of California regulatory
42 agencies, public utilities, and EPA recently participated in a study comparing the performance of
43 various genomics-based methods designed to identify the source of fecal material in ambient

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

19
20 These examples indicate the emerging need to make proactive policy decisions and to
21 develop processes to address how genomics data will be used in Agency decision making.

22 23 **C. Overview of Genomic Science**

24
25 Genomics tools provide the observer with a means to examine changes in gene
26 expression and protein and metabolite profiles within the cells of any organism, in contrast to
27 older methods of analyses which restrict
28 observers to looking only at whole organism
29 effects or changes in single biochemical
30 pathways. Genomics tools can provide
31 detailed data about the underlying
32 biochemical mechanisms of disease or
33 toxicity (i.e., disease etiology), sensitive
34 measures of exposures to chemicals, new approaches to detecting effects of such exposures, and
35 methods for predicting genetic predispositions that may lead to disease or higher sensitivity to
36 particular stressors in the environment.

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

37
38 As a means of introduction to genomics and its potential impact on regulatory decision
39 making, it is important to understand the basic principles behind the technology. Only about 1-
40 2% of the human DNA actually codes for RNA message that can be translated into a protein.
41 This 1-2% is considered the theoretical *functional genome*. Any particular cell type (i.e., from
42 various organs or species) will have its own practical functional genome, which is a subset of the
43 entire functional genome that encodes the proteins actually functioning in that cell. The

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

18
19 The study of a cell's protein composition is called **proteomics**. Currently, it is possible to
20 analyze only a fraction of a cell's proteins, but rapid advances in this field should allow more
21 complete profiling in the near future. Another discipline of biology analyzes biofluids and
22 tissues to determine the profiles of endogenous metabolites present under normal conditions or
23 when the organism has been affected by factors such as exposure to environmental chemicals.
24 This type of whole cell analysis is called **metabonomics** (or metabolic profiling). In order to
25 understand how a cell functions under normal or stressed circumstances, it is necessary to
26 characterize the proteins that are manufactured by the cell, as well as endogenous metabolites.
27 This facilitates an understanding of global metabolism and how proteins interact along cellular
28 activity pathways. This approach describes the area of **systems biology**, in which the cell, tissue,
29 or organism is considered as a complete, albeit complex, system.

30
31 For the purposes of this document it is important to note that all of these so-called "omic"
32 technologies can be used to compare
33 functional genomes (mRNA) and proteomic
34 and metabonomic profiles in normal cells and
35 tissues with responses in stressed cells and
36 tissues such as those exposed to
37 environmental agents. Analysis of the large
38 data sets generated for these type of analyses
39 requires the development of new
40 **bioinformatic** and **computational** tools. An integrated analysis and understanding of biological
41 systems and their responses to perturbation, from genes to adverse effects, and the capacity to
42 collect and evaluate data supportive of such a view would be expected to greatly enhance the risk
43 assessment process and, thus, aid in formulating regulatory policy and making regulatory

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

1 decisions.
2

3 As the Agency considers the significance of the current state of genomics technologies, it
4 is critical to realize that these technologies continue to advance at very rapid rates. Some of the
5 technology “laws” that have been developed to describe this advance include Monsanto’s Law,
6 “the amount of useful genetic information doubles every 18 - 24 months,” and Dawkin’s Law,
7 “the cost of sequencing DNA base pairs halves every 27 months.” As an example, a commercial
8 producer of gene chips has reported that the information content of their chips has been growing
9 exponentially from 16,000 cDNA probes per chip in 1994 to over 500,000 in 2002. While
10 commercial enterprises have recently developed arrays capable of handling large portions of the
11 human genome, the cost of such microarray technology is still high enough that their use in
12 toxicity-based approaches is relatively limited. However, these costs are falling rapidly as the
13 broad utility of the technology becomes apparent. In this regard, it is estimated that in
14 approximately two years, costs will decline to where clinical use will be feasible (e.g., assist in
15 selecting treatments, intervene with disease before overt symptoms occur, offer genetically
16 personalized nutrition and lifestyle advice, customize drug prescriptions), and the resulting
17 economies of scale will contribute to a continuing decline in costs to the research community
18 (Personal communication, from interviews by Robert Olson, Research Director, Institute for
19 Alternative Futures, September 2003).
20

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

22

23 This is an unprecedented time in the history of science because of the rapid development
24 of genomics and associated technologies. Genomics technologies are becoming highly
25 sophisticated and have great potential for contributing to the assessment and management of
26 environmental risks. The challenge lies in
27 understanding how this information is likely
28 to change current Agency approaches to
29 human and ecological risk assessment and
30 decision making. Although EPA recognizes
31 the inherent issues currently associated with
32 genomics studies in general and microarray
33 experiments in particular and that these issues will need to be addressed before this technology
34 can be fully accepted in risk assessment, the Agency also recognizes that genomics information
35 will likely become an integral part of risk analysis in the future. The purpose of this document,
36 consequently, is to present implications of the use of genomics technologies in Agency practice.
37 It is also the intent of the Genomics Task Force to inform, invite discussion, and shed light on
38 issues that will need to be addressed now or in the near future.
39

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

40 A key charge for the Task Force was identifying various exemplary circumstances under
41 which EPA might receive genomics data and the resulting implications for EPA policies and
42 programs. Sections II and III outline these circumstances or applications.
43

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

- 3
- 4 a) Prioritization of Contaminants and Contaminated Sites
- 5
- 6 b) Monitoring
- 7
- 8 c) Reporting Provisions
- 9

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

15

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

19

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

- 23
 - 24 a) Research
 - 25
 - 26 1) Linking genomics information to adverse outcomes
 - 27
 - 28 2) Interpreting genomics information for risk assessment
 - 29
 - 30 b) Technical Development
 - 31
 - 32 1) Establishing a technical framework for analysis and acceptance
33 criteria for genomics information (including data quality standards
34 based on genomic assay performance)
 - 35
 - 36 c) Capacity/Human Capital
 - 37
 - 38 1) Applying strategic hiring practices to recruit individuals who
39 possess genomics core competencies essential for crucial areas of
40 research, analysis, systems biology, bioinformatics, and risk
41 assessment
 - 42
 - 43 2) Training EPA risk assessors and managers to interpret and
-

1
2

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. Most of these chemicals have not undergone extensive toxicological testing, and there is sufficient information to allow a thorough evaluation of risk for only a fraction of them. 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 another 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. However, there currently is no rapid, comprehensive method for prioritizing which chemicals or microbes should be tested, and it is recognized that it is not possible to test all stressors. Further, there is currently no scientific consensus concerning which tests would be most appropriate for the Agency's different prioritization needs. 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 and low cost alternatives to the tests EPA currently relies on for prioritization.

2. Regulatory and Voluntary Activities Potentially Affected by Genomics Information

a. Representative Activities

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

2 **Challenge:** Genomics data could be applied to the voluntary HPV screening process.
3 For example, specific gene expression data could be used to predict the relevance of an
4 endpoint evaluated in an animal model Screening Information Data Set (SIDS) test to an
5 adverse health response in humans. This type of genomics data could, in the future,
6 potentially supplement or supplant animal testing needed to complete the SIDS data set.
7

8 **Office of Water - Contaminant Candidate List:** Genomics data generated for CCL
9 chemicals may be able to supplement computer model results or expert judgment in
10 hazard estimation and prioritization activities. For example, computerized analysis and
11 the growing use of automated polymerase chain reaction (PCR) techniques have allowed
12 tremendous gains in the study of microbial genomics, as well as of whole organisms. A
13 number of microorganism genomes have already been studied, many of which are
14 associated with waterborne disease. Genomics databases may play a role in prioritizing
15 pathogens based on the availability of virulence genes of concern and their corresponding
16 gene products.
17

18 **b. Additional Activities**

19
20 Genomics may also have regulatory implications for prioritization in other program
21 offices as well as in regions, states, and tribes (Further details on the following activities
22 are found in Appendix I.)
23

24 Program Offices

- 25 • Office of Pollution Prevention and Toxics (OPPT): Premanufacture Notices
26 (PMNs), Voluntary Children's Chemical Evaluation Program (VCCEP),
27 Endocrine Disruptors Screening Program (EDSP)
- 28 • Office of Pesticide Programs (OPP): Pesticides, Inerts, EDSP
- 29 • Office of Water (OW): Prioritizing stream or wetlands for study or cleanup
- 30 • Office of Air and Radiation (OAR): Hazardous air pollutants (HAPs)
- 31 • Office of Solid Waste and Emergency Response (OSWER): Superfund sites
- 32 • Office of Research and Development (ORD): Future research on chemicals
33

34 Regions, States, and Tribes

- 35 • Site remediations and chemical evaluations
36
37
38
39
-

1 **B. Monitoring**

2
3 **1. Introduction**

4
5 The term “monitoring” in the present context refers to any activity by which environmental
6 samples are taken and used for regulatory or
7 prioritization decisions and for developing
8 environmental status and trends information.
9 Many programs have either site-specific or
10 media-specific data requests that are used to
11 make regulatory decisions, monitor
12 compliance, and/or to prioritize the use of
13 EPA’s human and economic resources. In

There are many potential applications of genomics-based data that could come into the Agency under the category of ‘monitoring data.’

14 addition, EPA is charged with determining the state of the environment, as well as assessing the
15 status and trends of ecological condition. In many instances, entities other than EPA generate
16 the data EPA uses (e.g., other federal agencies, states, tribes, the regulated community, and
17 interested stakeholders in volunteer monitoring programs). These data are generated through
18 various headquarters and regional programs through contracts, grants, and cooperative
19 agreements. In fact, a large portion of EPA Regions’ budgets are directed to programs (e.g., in
20 states and tribes) that generate information that could fall into the category of monitoring data or
21 information.
22

23 EPA obtains, requests, and receives many types of environmental data for both assessment
24 and compliance purposes, including but not limited to the following: chemical and physical
25 analyses of air, water, soil, and sediment; toxicity testing of various environmental media or
26 chemicals; plant, animal, and human tissue residues of various chemicals or their breakdown
27 products; community structure analyses (e.g., fish and/or invertebrate IBIs [index of biotic
28 integrity], algal and plant community structure, invasive species evaluations in terrestrial and
29 aquatic ecosystems); and microbial community and pathogenic microorganism analyses of air,
30 water, soil, and sediment.
31

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

1 (i.e., adding nutrients or microbial cultures) versus monitored natural attenuation. The
2 Department of Energy is exploring how this type of microbial community “fingerprinting” can
3 be used to distinguish the conditions that promote effective bioremediation of petroleum-
4 contaminated soils or sediments (http://www.sc.doe.gov/ober/ERSD/ersd_nabir.html). A third
5 example is in the area of development of multi-gene arrays of model animals (e.g., “fish-on-a-
6 chip”). Future toxicity testing for compliance with discharge requirements could involve
7 exposing chips containing the genome of a fathead minnow to determine whether a sample of
8 water demonstrated a toxic pattern of response.
9

10 Thus, there is a wide range of monitoring information that the Agency considers and a wide
11 range of potential applications of genomics technologies. The cost and time required to collect
12 and analyze the large number of conventional environmental samples needed to make sound
13 regulatory decisions and to evaluate environmental status is enormous. Some genomics tools,
14 such as gene chips, offer the advantage of portability (i.e., they can be taken into the field for
15 direct testing) and consequently reduce the need to preserve and transport bulky samples (e.g.,
16 sediment and water) for laboratory analysis. Genomics technologies are likely to offer many
17 rapid, efficient, and cost-effective methods for environmental monitoring.
18

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

20 **a. Representative Activities**

21
22
23 **Office of Water (OW)/Office of Ground Water and Drinking Water (OGWDW):**
24 OGWDW anticipates that monitoring for chemicals or microbial pathogens will use
25 genomics-based data in five to ten years for the following purposes:
26

- 27 • Compliance monitoring by federal, state, and tribal agencies to determine if
28 surface waters meet designated uses standards under CWA (TMDLs, Criteria,
29 Standards)
30
 - 31 • Monitoring finished or source drinking waters for contaminants
32
 - 33 • Real-time monitoring of ambient surface water, e.g., for beach or shellfish bed
34 closures (Beaches Environmental Assessment and Coastal Health [BEACH] Act,
35 Standards)
36
 - 37 • Monitoring classes of compounds based on biological activity or mode of action
38 (e.g., cholinesterase inhibition)
39
 - 40 • Developing occurrence data as a basis for Safe Drinking Water Act (SDWA)
41 regulation or listing on CCL
42
 - 43 • Developing future drinking water regulations (6 Year Review)
-

1
2 **Regions, States, and Tribes - State NPDES permits:** The State of California, as part
3 of an ongoing ambient water quality monitoring program, is initiating an effort to
4 evaluate surface waters for the presence of estrogenic endocrine effects, using a RT-
5 PCR assay for vitellogenin gene expression in livers of exposed male rainbow trout. If
6 results show that some surface waters exhibit estrogenic effects, the Regional Water
7 Quality Controls Boards in California, which issue NPDES permits as well as perform
8 ambient water quality monitoring, may consider including this type of bioassay into the
9 monitoring program for waste water treatment facilities even though it is not yet an
10 "EPA method." Currently, NPDES permits contain only chemical and toxicity-based
11 limits, require chemical analyses and toxicity testing of effluents to demonstrate
12 compliance with permit limits, and relate waste loads to watershed TMDLs. Genomics
13 technologies could provide new and more sensitive monitoring tools to develop
14 discharge limits for NPDES permits.

15
16 **b. Additional Activities**

17
18 Genomics may also have regulatory implications for monitoring in other program
19 offices, as well as in regions, states, and tribes (Further details on the following
20 activities are found in Appendix I.)

21
22 Program Offices

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

28
29 Regions, States, and Tribes

- 30 • State and local beach closure and TMDL issues associated with pathogens
 - 31 • State and local air quality monitoring
 - 32 • Tribal issues (e.g., monitoring for endocrine disruptors)
 - 33 • Regional pesticide program inspections
- 34
-

1 **C. Reporting Provisions**

2
3 **1. Introduction**

4
5 **a. Reporting on Adverse Effects of Commercialized Chemicals and Pesticides**

6
7 Reporting of certain adverse effects/risks for industrial chemicals and pesticides
8 already on the market is mandated under both the Toxic Substances Control Act (TSCA) and
9 the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). TSCA section 8(e)
10 requires that “[a]ny person who manufactures, processes, or distributes in commerce a
11 chemical substance or mixture and who obtains information which reasonably supports the
12 conclusion that such substance or mixture presents a substantial risk of injury to health or the
13 environment shall immediately inform [EPA] of such information” (15 U.S.C. 2607(e)).
14 FIFRA section 6(a)(2) states “If at any time after the registration of a pesticide the registrant
15 has additional factual information regarding unreasonable adverse effects on the environment
16 of the pesticide, the registrant shall submit such information to the Administrator.” There is
17 a need to interpret how these TSCA and FIFRA provisions apply to genomics data. There
18 are already certain types of conventional tests whose data are not considered to present
19 indication of substantial risk to health or the environment and are not required by the Agency
20 as stand-alone submissions. As the predictability and validity of genomics methods increase,
21 EPA may need to re-evaluate its stance on these reporting provisions. Because these
22 provisions address the reporting of adverse effects, the issue of what genomic changes mean
23 in terms of adversity must be addressed before reporting for genomic responses may be
24 required.

25
26 **b. Toxics Release Inventory Program**

27
28 The Toxics Release Inventory (TRI) database was established under the Emergency
29 Planning and Community Right-to-Know (EPCRA) Act of 1986. Section 313 of EPCRA
30 requires certain industrial facilities to annually report information on toxic chemical releases
31 and other waste management activities to EPA and the states to inform communities of
32 chemical hazards in their area.

33
34 The statutory chemical listing/delisting criteria of EPCRA section 313 (d)(2) are
35 primarily based on hazard, not on risk. The emphasis of EPA’s hazard assessment is on a
36 chemical’s inherent toxicity rather than the potential risks from exposure to the chemical. A
37 chemical may be added to the TRI list if (a) the chemical is known to cause, or can
38 reasonably be anticipated to cause, significant adverse acute human health effects at
39 concentration levels that are reasonably likely to exist beyond facility site boundaries as a
40 result of continuous, or frequently recurring, releases; (b) the chemical is known to cause or
41 can reasonably be anticipated to cause, in humans, cancer or teratogenic effects or serious or
42 irreversible reproductive dysfunctions, neurological disorders, heritable genetic mutations, or
43 other chronic health effects; or (c) if the chemical is known to cause, or can reasonably be

1 anticipated to cause, because of its toxicity, its toxicity and persistence in the environment, or
2 its toxicity and tendency to bioaccumulate in the environment, a significant adverse effect on
3 the environment of sufficient seriousness, in the judgement of the Administrator, to warrant
4 reporting.
5

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

7

8 In order for genomics technologies to have an effect on reporting provisions, the issue of
9 linking genomics changes to adverse effects or response pathways needs to be addressed. Once
10 genomic changes are linked to adverse effects, the Agency will need to make decisions regarding
11 whether genomic changes apply to reporting provisions.
12

13 Genomics technologies could affect reporting requirements under TSCA 8(a) and FIFRA
14 6(a)(2) if genomic changes detected are linked with substantial risks or adverse effects. If
15 genomics data do, in the future, become a reporting requirement, this could also affect the
16 number of reports received under these reporting provisions and the resources required to
17 evaluate the reports.
18

19 If there is a linkage to adverse effects in humans or on the environment, genomics data may
20 be considered in the hazard assessment when determining whether or not a chemical meets the
21 TRI chemical listing/delisting criteria. The Interim Policy on Genomics allows such information
22 to be used in the overall assessment on a case-by-case basis, but genomics information alone
23 currently cannot be used to determine hazard at this time. Practical application of genomics-
24 derived information will improve the quality of hazard assessments conducted by EPA, including
25 those conducted by EPA's TRI Program.
26
27

1 **III. Risk Assessment**

2
3 **A. Introduction**

4
5 Genomics technologies present an opportunity to greatly enhance Agency risk assessments.
6 Specifically, genomics technologies are likely
7 to contribute significantly to improvements in
8 defining a chemical’s mode of action,
9 evaluating effects on susceptible populations
10 and life stages, and assessing exposure to and
11 effects of chemical mixtures, as outlined
12 below. In collaboration with the International
13 Life Sciences Institute (ILSI), EPA has developed a framework for microbial risk assessment
14 (MRA), and OW is working to expand the framework into a full MRA protocol which may
15 include consideration of genomics data. OW is also working with EPA’s Risk Assessment
16 Forum to develop MRA guidelines.
17

Genomics technologies are likely to contribute significantly to improvements in risk assessment

18 **B. Mode of Action (MOA)**

19
20 **1. Overview**

21
22 The term “mode of action” (MOA) is defined in the Draft Guidelines for Carcinogen Risk
23 Assessment (USEPA, 2003a) as a sequence of key events and processes, starting with the
24 interaction of an agent with a cell, proceeding through operational and anatomical changes, and
25 resulting in an adverse outcome. A "key event" is an empirically observable precursor step that
26 is in itself a necessary element of the MOA or is a biomarker for such an element. Genomics
27 technologies can be used to better understand the MOA of a chemical agent and, thus, can lead
28 to advances in human and ecological risk assessments of chemicals. As genomics information
29 contributes to our understanding of MOAs, the validity of using this information as an indicator
30 of both adverse effects and exposure is enhanced.
31

32 Genomics data may allow the development of gene, protein, or metabolite profiles that can
33 advance the screening of individual chemicals and allow faster and more accurate categorization
34 into defined classes according to their MOA. There are many examples of possible modes of
35 carcinogenic action, such as mutagenicity, mitogenesis, inhibition of cell death, cytotoxicity with
36 regenerative cell proliferation, and immune suppression. MOAs have been identified for other
37 adverse outcomes, both human health and ecological (e.g., cytotoxicity, endocrine disruption,
38 loss of homeostasis). Such approaches will significantly enhance the Agency’s ability to
39 harmonize risk assessment approaches for different outcomes for which the development of a list
40 of common MOAs is essential.
41

42 Understanding the MOA of environmental agents that induce toxic effects other than cancer
43 or that induce carcinogenicity in animal models should facilitate the assessment of the relevance

1 of these findings in protecting human health and safeguarding the environment. An important
2 issue for extrapolation of responses in animal models to humans or environmental endpoints is to
3 establish whether the MOA in the test species is relevant in the target species.

4 5 **2. Risk Assessment Activities Potentially Affected by Genomics Information**

6
7 Many program and regional offices need to make judgments about chemicals with little or no
8 data available. Others have extensive datasets, but still struggle with accuracy and precision as
9 well as extrapolations to nontested species or scenarios. Genomics approaches are envisioned to
10 provide improvements in (a) hazard identification, (b) dose-response assessment, (c)
11 extrapolations; and (d) exposure assessment.

12 13 **a. Representative Activities - OPPTS, ORD, OW**

14 15 **i. Improving Hazard Identification**

16
17 ***Evaluating chemicals for genotoxic or other MOAs.*** New approaches using tissue
18 microarrays can enhance throughput and the linking of genomic and cellular outcomes.
19 Further, combining the findings of gene expression studies with data from chemical
20 exposures of genetically altered animal models (e.g., knockout or null mice) is a
21 powerful tool to link specific genes to specific detrimental outcomes. Such data will
22 allow the development of gene profile “fingerprints” of genomic characteristics for
23 specific MOAs. The development of genomic “fingerprints” will provide a rapid
24 screening method to categorize chemicals with unknown MOAs for both human health
25 and ecological assessments.

26
27 ***Predicting or Defining Metabolic Pathways.*** The chemical evaluation process
28 includes consideration of the parent compound and its potentially active metabolites.
29 Genomics approaches, particularly at the proteome level, will aid in the
30 characterization of metabolic pathways and the identification of toxicologically active
31 metabolites. Computational toxicology approaches will further enhance the prediction
32 of metabolic pathways and metabolites for chemicals that have not been investigated
33 experimentally and potentially will reduce the use of test animals and the cost of data
34 generated to support risk assessments. Metabolic pathways and the genes associated
35 with those pathways need to be linked to adverse effects of concern.

36
37 ***Replacement of standard toxicity tests in regulatory batteries with rapid, pathway-***
38 ***specific response tests.*** It has been envisioned that relevant genes and gene products
39 for specific toxicities such as genotoxicity can be formatted on arrays to provide a more
40 comprehensive analysis than currently available assays (Aardema and MacGregor,
41 2002). Because many of the toxicological testing procedures and strategies required by
42 EPA have remained largely unchanged for 20 years, it is reasonable to assume that
43 many of the current assay systems used may be replaced by more sensitive, rapid, and

1 predictive genomic assays able to identify specific pathways of response. Acceptance
2 of these genomic protocols for both human health and ecological assessments will lead
3 to time and cost savings and may also lead to more accurate risk assessments.
4

5 **ii. Improving Dose-Response Assessment**
6

7 ***Linear versus nonlinear.*** The Agency's traditional approach to cancer risk assessment
8 for agents that are known mutagens and carcinogens employs a linear, low dose
9 extrapolation to quantify possible human cancer risks. The underlying premise for this
10 linear default is that electrophilic compounds are presumed to form single DNA
11 modifications in single cells that could potentially lead to cancer. Due largely to the
12 successful use of genetic toxicological testing schemes for screening, however, the
13 number of new genotoxic carcinogens entering the environment is likely to be small.
14 The recent Final Draft Guidelines for Carcinogen Risk Assessment addresses the issue
15 of nongenotoxic carcinogens and encourages the use of mechanistic data to identify
16 whether a nonlinear extrapolation is appropriate for nongenotoxic carcinogens. For this
17 purpose, biomarkers of response for genotoxic carcinogens are available at least at the
18 cellular response level if the general cancer MOA is known. The challenge will be to
19 develop biomarkers of response that can be used for predicting specific outcomes for
20 nongenotoxic chemicals (Bartosiewicz et al., 2001a). This goal can be realized through
21 gene expression pattern recognition that parallels histological changes in tissues and the
22 eventual progress to tumor formation. An example would be the CIIT (Chemical
23 Industry Institute of Technology) Centers for Health Research's efforts to identify
24 genes associated with peroxisome proliferators, such as the PPAR- α (peroxisome
25 proliferator-activated receptor alpha) that are linked to alterations in mouse
26 hepatocellular growth following peroxisome proliferator exposure.
27

28 The regulatory impact of genomics on possible nonlinear extrapolation for
29 nongenotoxic carcinogens is significant. Within the Agency, nongenotoxic carcinogens
30 without plausible MOA data are currently subjected to the same linear low dose
31 extrapolation applied to genotoxic carcinogens. It is a reasonable assumption that a
32 collaboration among industry and EPA scientists will occur in the area of MOA-based
33 cancer risk assessment to ascertain if a nonlinear low dose extrapolation is appropriate
34 for nongenotoxic carcinogens. The same MOA approach can be used to help more
35 clearly discern dose-response relationships for chemicals that affect other health
36 endpoints as well.
37

38 ***Lowering of Points of Departure (PODs) based on genomic responses.*** Genomics
39 technologies have the potential to affect dose-response analyses for nonlinear
40 assessments of adverse toxicological outcomes. In traditional toxicology, doses used to
41 determine adverse effects are generally high to ensure that tissue level or whole animal
42 toxic responses are demonstrated. This permits the selection of toxicity endpoints and
43 establishment of doses at which no adverse effects are seen (NOAEL) and the lowest

1 doses at which an adverse effect is seen (LOAEL). Thus, most toxic substances
2 currently are regulated on frank toxicity rather than on a molecular level response, and
3 the association between a molecular level change and an adverse outcome has only
4 rarely been established. A few substances are regulated based on biochemical changes
5 with known relationships to adverse outcomes such as cholinesterase-inhibiting
6 pesticides and lead. Organophosphate pesticides are regulated at NOAELs which are
7 generally much lower than many chemical classes because the endpoint, cholinesterase
8 inhibition, is determined biochemically, and inhibition can generally be detected well
9 below the levels showing overt toxicity. In contrast, many fungicides or herbicides
10 have relatively high NOAELs because clear pathological alterations only occur in
11 animals at high doses. Regardless of the chemical class or use, however, with the
12 advent of molecular technologies including genomics, chemically-induced changes in
13 gene expression are likely to result in the identification of simple, sensitive, and
14 relevant biomarkers of effect that can be used in dose-response studies to more readily
15 identify effects in the low dose range (i.e., below doses causing frank pathology) for
16 humans and wildlife species.

17
18 If EPA chooses to establish regulatory limits (e.g., NOAELs) based on changes in gene
19 expression, the POD used to set the regulatory limit could be higher or lower in both
20 the human health and ecological arenas. EPA needs to examine whether a lower effect
21 level based on a molecular effect is “safer” than a level based on the currently used
22 frank effect.

23 24 **iii. Improving Extrapolations**

25 26 *High to low dose extrapolations*

27
28 Reduction of uncertainty is one of the primary ways to improve the risk assessment
29 process. Reduction in uncertainty in dose-response assessments can be enhanced by
30 the use of predictive models such as Physiologically-Based Pharmacokinetic (PBPK)
31 and Biologically-Based Dose Response (BBDR) models. These models can provide
32 better methods for calculating dose metrics (e.g., target tissue doses) that are more
33 flexible and relevant for extrapolation across exposure routes, between species, and
34 from high to low doses. The potential of molecular indicators to define the shape of the
35 dose-response relationship at low exposures suggests the possibility that alteration or
36 elimination of some uncertainty factors (UFs) may be justified; data-derived factors
37 based on genomics information may be determined and applied in the future.
38 Similarly, molecular-based pharmacokinetic data that describes the distribution of
39 biologically effective doses of active ingredients to target organs via other portals of
40 entry offers the possibility of reducing uncertainties associated with route-to-route
41 extrapolations.

42 43 *Interspecies extrapolations*

1 **Relevance to humans.** Further improvements in human health assessments can be
2 realized through the use of genomics data that support an evaluation of whether or not
3 MOAs determined in test animals are similar and feasible in humans (i.e., whether the
4 target genes are conserved and operative across species). Genomics data that show
5 little or no similarity in key genes or patterns of gene expression between humans and
6 rodents would indicate interspecies differences and support a possible conclusion of
7 non-relevance to humans. Conversely, data showing good agreement in key genes or
8 expression patterns between humans and rodents would provide higher confidence in
9 the relevance of the findings to human health. Similarly, interspecies comparison of
10 pharmacodynamic responses enhanced by the use of genomics data could be used to
11 define toxicological pathways in a quantitative sense. This information could be
12 compared across species by the choice of appropriate molecular markers. Gene
13 expression profiling is one approach that looks promising for linking cellular responses
14 to a specific environmental chemical or mixture in laboratory animals to responses in
15 humans or human cells *in vitro*.

16
17 Key biological systems have fundamental genomics processes, some of which, if
18 altered, are universally deleterious. Demonstration of a common interspecies genomic
19 response linked to an adverse effect and evaluation of the dose-response relationships
20 in the lower animal (e.g., invertebrates) and humans could permit the extrapolation of
21 genomic responses in lower order animals to adverse effects in humans. The increasing
22 development of genomic information in lower organisms may provide a means to
23 evaluate potential effects in humans that will extend the use of lower organisms beyond
24 current mutagenicity testing.

25
26 **Relevance to wildlife species of concern.** In ecological risk assessment, it is
27 necessary to extrapolate results from a very limited set of test species to a wide range
28 (potentially hundreds to thousands) of species present in the environment. The
29 development of reliable methods for extrapolating toxicity information from test
30 species to those that are of concern but cannot be directly tested is necessary. As in
31 human health assessments, an important issue is determining whether the MOA in the
32 test species is feasible for other species present in the ecosystem. Use of genomics
33 tools for the development of quantifiable pharmacodynamic models and applicable
34 molecular markers will also significantly enhance species-species extrapolations and
35 reduce the current reliance on the application of uncertainty factors.

36 37 **iv. Improving Exposure Assessment**

38
39 Genomics technologies are likely to lead to the development of simple, sensitive, and
40 informative biomarkers of exposure that can be used in exposure assessments,
41 particularly in the evaluation of potential occupational exposures for human health
42 assessments and for environmental exposures for both human health and ecological risk
43 assessments. Current methods rely on residue analyses or modeled scenarios and a few

1 well-documented biomarkers of exposure (e.g., CYP1A, cholinesterase, delta-
2 aminolevulinic acid dehydratase, metallothionein). Molecular techniques such as the
3 use of microarrays or RT-PCR are providing new tools for documenting actual
4 exposures to humans and ecological species of concern. When pharmacodynamic and
5 MOA studies become sufficiently robust to relate exposure endpoints to whole
6 organism adverse effects, risk assessment predictions will become significantly more
7 accurate.

1 **C. Susceptible Populations and Sensitive Life Stages**

2
3 **1. Overview**

4
5 **a. Susceptible Human Populations and Sensitive Life Stages**

6
7 Genomics and related technologies offer a tremendous opportunity to define and identify
8 people with enhanced susceptibility to many environmental contaminants. The human genome
9 consists of 30,000 or so genes that build
10 cellular structures, control the cell cycle,
11 execute metabolic functions, and mediate the
12 information flow within and between cells.
13 Small differences in gene sequence, known as
14 Single Nucleotide Polymorphisms (SNPs),
15 can have a dramatic or inconsequential effect
16 on protein function and activity depending on
17 the particular polymorphism. Genomics technologies have the potential to yield information
18 about the distribution of SNPs within the human population and their potential effects on genes
19 that are responsive to various environmental contaminants. The interaction of genetic variants
20 with environmental conditions can affect individual susceptibility to a variety of diseases such as
21 cancer, diabetes, and heart disease and can promote sensitivity to disease from exposure (Bishop
22 at al., 2001).

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

23
24 Delineation of the frequency of occurrence of these polymorphisms within racial or ethnic
25 groups may raise ethical, legal, and social implications in the area of environmental justice
26 (Marchant, 2002). For example, genomics technologies might result in further categorization of
27 individuals, and consideration must be given to how these newly identified at-risk groups will be
28 included in environmental policy. Ethical practices by which genomics studies are conducted on
29 human test subjects in specific ethnic or racial communities will also need to be carefully
30 considered. The Agency will need to be proactive to ensure that these issues are handled in a
31 just manner.

32
33 For a susceptible human population, an increased risk of illness could result from exposure to
34 environmental chemicals or microbial pathogens at any age. In contrast, an exposure during a
35 susceptible life stage could result in higher risk during a specific portion of an individual's
36 lifetime or could influence an outcome at another life stage and the potential adverse effect
37 incurred may be irreversible. For example, an exposure during early childhood development
38 might yield a specific form of cancer that would not have been induced if the exposure occurred
39 later in life. Similarly, an exposure during a sensitive life stage could result in an increased
40 probability of disease later in life. The primary difference between susceptible populations and
41 susceptible life stages is that a susceptible population's toxic exposure generally will yield an
42 adverse outcome regardless of the age at which the exposure occurred. Nevertheless, many of
43 the ethical, legal, and social implications (ELSI) that apply to susceptible populations will also

1 apply to susceptible life stages. As a consequence, the Agency will need to develop a policy
2 regarding the collection and use of human genomics information from individuals to provide
3 safeguards regarding privacy.
4

5 The National Institute of Environmental Health Sciences (NIEHS) has funded and supported
6 research to identify gene variations that affect susceptibility to environmental agents.
7 Approximately 500 genes were identified as part of the Environmental Genome Project of 1997.
8 These genes affect metabolism, DNA repair, cell cycle control, receptors, and immune function.
9 Although scientific progress has been made in understanding genetic variations and
10 susceptibility to toxic chemicals and pathogens, research efforts have yielded inconsistent
11 results. In spite of this, researchers continue to characterize genetic variations of susceptibility
12 and provide insights about individuals who are more or less susceptible to disease from exposure
13 to toxic substances (Marchant, 2002).
14

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

16
17 The term “susceptible population” is most frequently used in reference to at-risk human
18 populations; however, the term can also be used to describe wildlife populations that may be at
19 risk due to exposure to environmental
20 contaminants. Certain taxonomic subgroups
21 of plants and animals may be more
22 susceptible than others although this varies
23 with contaminant and life stage. Currently,
24 EPA examines species sensitivity through
25 toxicity testing of environmental
26 contaminants on representatives from various
27 classes of organisms. Genomics technologies offer a powerful tool with which to examine
28 toxicological responses across species for prediction of sensitivities or tolerances in untested
29 organisms. In addition, genomics technologies will allow for the examination of long-term
30 ecosystem health and the potential irreversibility of toxic effects. Examining the genetic
31 diversity of organisms inhabiting a given ecosystem would allow risk assessors to determine if
32 exposure to environmental contaminants might cause an evolutionary change in ecosystem
33 structure or function.
34

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

35 Threatened and endangered species represent a subset of all species faced with the possibility
36 of extinction and are afforded extra protection under the Endangered Species Act. A species,
37 however, may be sensitive to exposure to an environmental contaminant without being
38 endangered. Comparisons of the sensitivities of endangered and common fish test species to
39 toxicants have been made. This research indicated that the sensitivities of the endangered
40 species tested were generally not greatly different from that of the more common species
41 (Sappington, et al., 2001). However, because the population size of endangered species is, by
42 definition, quite small, reduced genetic variation may result in reduced tolerance to multiple
43 stressors such as combinations of contaminants and climate stress (Porter et al., 1984).

1 Alternatively, continuous exposure of a widespread, susceptible species to an environmental
2 contaminant might result in the species becoming endangered.

3
4 While human behaviors such as over-hunting or habitat destruction are the main stressors
5 that threaten species survival, contamination of an ecosystem with a pesticide, industrial
6 chemical, or pathogen can also harm ecosystem health. Genomics technologies will provide
7 significant insight into the biochemical mechanisms by which environmental contaminants might
8 adversely affect certain species and may provide insights into which species are more likely to
9 be headed down the path towards extinction. Cross-species extrapolations of sensitivity coupled
10 with the ability to measure contaminant-induced reduction in genetic diversity within specific
11 populations will provide valuable input to population viability models.

12 13 **2. Risk Assessment Activities Potentially Affected by Genomics Information**

14 15 **a. Susceptible Human Populations and Sensitive Life Stages**

16
17 EPA anticipates genomics research will be used to assess hazards and risks of chemicals and
18 pathogens to specific human populations. Currently, the Agency does not generally take into
19 account genetic factors when assessing the risks posed by chemical or biological substances
20 although life stage and, to some extent, gender are considered. Additionally, the Agency rarely
21 considers the genetic predisposition of a specific individual, race, or ethnic background when
22 determining the toxic effects of chemicals. The Food Quality Protection Act (FQPA) of 1996
23 does, however, direct the Agency to determine the potential of increased susceptibility of infants
24 and children from exposure to toxic substances such as pesticides.

25 26 **i. Representative Activities**

27
28 **OPPTS, OW, OAR, ORD, OSWER.** In any situation where a “safe” exposure level
29 is designated (FIFRA, Clean Water Act, Safe Drinking Water Act, Clean Air Act,
30 CERCLA), the identification of a susceptible population or life stage has the potential
31 to force a change in the health standards employed. For example, the health standards
32 used during hazardous waste remediation projects are often first reviewed and
33 published in the Integrated Risk Information System (IRIS) database maintained by
34 ORD. One of the goals of the IRIS database is to provide oral Reference Doses (RfDs)
35 and inhalation Reference Concentrations (RfCs). Both the RfD and RfC are based on
36 the assumption that nonlinear dose-response curves exist for certain toxic effects such
37 as cellular necrosis (USEPA, 2002b). In general, these values provide an estimate of a
38 daily exposure to the human population (including sensitive subgroups) that is likely to
39 be without an appreciable risk of deleterious effects during a lifetime
40 (<http://www.epa.gov/iris/gloss8.htm#r>). Genomics technologies will provide a
41 powerful tool for the identification of sensitive subgroups and allow specific decisions
42 to be made that address specific sensitivities. As the use of genomics technologies
43 becomes widespread, the number of susceptible populations identified will likely grow

1 in number. Consequently, the Agency will need to be vigilant when revisiting health
2 standards in all media to identify and protect susceptible populations.

3
4 The impact of genomics technologies on EPA's understanding of life stage
5 sensitivity also will become an important issue. For example, different human age
6 groups may express varying levels of some metabolic enzymes (Hakkola et al., 1998).
7 Enzyme over- or under-expression could play a part in determining the severity of a
8 toxic exposure. Additionally, the rapid growth taking place early in life is largely
9 dependent on gene-environment interactions. Perturbation of this interaction through
10 toxic exposure has the potential to significantly influence development. Genomics will
11 provide tools to identify life stages that need separate assessments based on their
12 unique susceptibilities.

13
14 **OPPTS.** The extent of enzyme activation is partially responsible for determining the
15 severity of response to a chemical exposure. In addition, individuals may respond to
16 metabolic stimuli to varying degrees depending on their genetic composition. Ethanol
17 exposure, for example, is known to induce the metabolic enzyme CYP2E1; however,
18 the amount and activity of CYP2E1 produced may vary among individuals (Haber et
19 al., 2002; Snawder and Lipscomb, 2000). If genomics technologies are successful in
20 identifying populations susceptible to specific pesticides or industrial chemicals,
21 product labeling will probably be necessary. For example, labels might include
22 warnings for particular populations known to exhibit higher frequencies of an at-risk
23 genetic polymorphism. The pharmaceutical industry already includes warnings to
24 susceptible populations on drug labels. The Agency has the ability to follow similar
25 practices for pesticides because Section 3 of FIFRA
26 (<http://www.epa.gov/pesticides/pestlabels/>) provides EPA the authority to regulate
27 labels.

28
29 **OSWER, Regions, States, and Tribes.** Genomics data might be used in the future to
30 help identify susceptible populations or life stages when assessing risks at hazardous
31 waste sites or for remediation of contaminated areas such as Superfund sites or other
32 scenarios in which relatively small, identifiable populations are exposed. Genomics
33 data may provide information on potential exposure patterns and also might be useful
34 in developing site-specific remediation goals. If, for example, a genomics study were
35 to identify a susceptible population at risk due to exposure to a contaminant at a
36 Superfund site through a correlation of genomic analysis of local populations and
37 measured or expected exposure levels, the Agency might choose to reduce the RfD/RfC
38 value and propose more strict remediation measures. This, of course, presupposes an
39 established linkage of the genomic endpoint and an adverse effect. Use of new
40 genomics tools could, however, limit the extent of remediation measures by more
41 accurately predicting the potential for exposure of the sensitive population. Thus,
42 genomics tools may play a key role in determining intensity and extent of clean up
43 practices and have large implications for time and cost of such procedures.

1
2 **b. Susceptible Wildlife Populations and Sensitive Life Stages**
3

4 EPA currently does not use genomics technologies in ecological assessments or criteria
5 development. However, numerous potential applications for the identification of sensitive
6 species, populations, or life stages may become available in the near term. Longer-term
7 applications are also under development.
8

9 **i. Representative Activities**

10
11 **OPPTS, OSWER, OW, OAR, ORD.** Due to time and resource limitations, as well as
12 ethical considerations, all species cannot be tested for responses to contaminants either
13 for site-specific mitigation needs, for product registration, nor for criteria development.
14 Therefore, the development of reliable methods for extrapolating toxicity information
15 from tested species to those that are of concern but cannot be directly tested is
16 necessary. This need is particularly acute for chemicals that may target sensitive life
17 stages (e.g., metamorphosis in amphibians) or vulnerable species (those with small
18 population sizes or that may have greater sensitivities to particular chemicals).
19 Genomics technologies will provide the potential for extrapolating between test species
20 and sensitive wildlife species or life stages in a rapid, cost-effective manner.
21

22 The most immediate application is likely to be with aquatic organisms and the
23 application to Quantitative Structure Activity Relationships (QSARs) or computational
24 toxicology. This will be particularly useful in the Premanufacture Notice (PMN)
25 process of OPPTS, but may find applications in other Offices as well (e.g., Office of
26 Water). The next application likely will be in the development of methods for
27 screening chemicals that are potential endocrine disruptors in aquatic ecosystems.
28

29 While the protection of individual plants and animals from clinical disease
30 caused by xenobiotic compounds and pathogens is an achievable goal, how chemicals
31 combine with other environmental stressors to change the genetic properties of a
32 population over time is less clear. If genetic diversity is reduced or if particular genes
33 are suppressed or expressed at abnormal rates, it is possible that (1) the population may
34 become less fit over time, (2) response to additional stressors may not be adequate; and
35 (3) the reduced diversity may lead to a bottleneck and/or eventual extinction. Because
36 genetic diversity is the fundamental basis for adaptation and evolution, it is increasingly
37 being recognized as an important endpoint for risk assessment. To date, the
38 methodology has not been available to address this issue, but newly emerging genomics
39 techniques will allow such assessments in the future.
40

41 **OSWER, Regions, States, and Tribes.** The current Agency practice in ecological risk
42 assessment and clean up of contaminated sites (Superfund, Brownfields) is to focus on
43 the most sensitive species when determining the effects of a chemical contaminant on

1 an ecosystem. Genomics will allow risk assessors a greater ability to focus the
2 ecological risk assessment on the mechanistic level. Although the Agency does not
3 usually take into account genetic factors when assessing the ecological risks posed by
4 chemical or biological substances, genomics data gained from the examination of a
5 single species might prove useful in preventing harm to other species with similar
6 genetic characteristics.
7
8

1 **D. Mixtures**

2
3 **1. Overview**

4
5 Human health and ecological risk assessments of toxic substances often are incomplete
6 because most toxicological screening is performed for single chemicals. However, human and
7 wildlife exposures to chemicals are rarely limited to a single chemical, but instead are usually to
8 complex mixtures of chemicals during a
9 lifetime. Environmental exposures from
10 point and nonpoint sources, irrespective of
11 medium, generally occur as simultaneous or
12 sequential exposures to multiple chemicals.
13 In addition to environmental exposures, the
14 majority of the human population engages in
15 intentional exposure to a variety of
16 pharmacologically active chemical
17 compounds such as those in recreational drugs (alcohol and tobacco), medicinal products, and
18 foods and is inadvertently exposed to other chemicals, such as those in vehicle exhaust, drinking
19 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.

20
21 Chemical mixtures in the environment are, in general, a complex group of active and inert
22 parent compounds, transformation products, and/or residues of which composition is
23 qualitatively and quantitatively not fully known (Feron and Groten, 2002). Because mixtures
24 change with time and distance from the original release site due to the differential fate and
25 transport of their components (Pohl et al., 1997), regulations established on toxicological data for
26 the original mixture may have little bearing on the actual exposures resulting from a release. It is
27 estimated that approximately 275 million tons of hazardous waste are produced annually in the
28 United States, and more than 2000 distinctive toxicants in site-specific media have been
29 identified by EPA, with hazardous substances in given mixtures numbering in hundreds (Suk et
30 al., 2002).

31
32 Specific environmental chemicals have been demonstrated to adversely affect the health of
33 humans and wildlife. These concerns are amplified by the awareness that exposure to chemicals
34 often occurs in mixtures of chemicals that might exhibit complex interactions. The various types
35 of toxicological interaction associated with complex chemical mixtures can be sorted into three
36 reference categories: greater-than-additive (synergism, potentiation), additive (the sum is equal
37 to the parts, no apparent interaction), and less-than-additive (antagonism, inhibition, or
38 masking). Of particular importance is whether a mixture of components, each of which is
39 present at concentrations below the level of concern, may be hazardous due to additivity, specific
40 interactions, or both (Hertzberg and MacDonell, 2002). Synergistic toxicity resulting from
41 co-exposure to pesticides has been observed, and greater than expected toxicity has been noted
42 for pesticide mixtures of certain cholinesterase inhibiting insecticides and some fungicides.
43

1 The Agency has been directed by FQPA to consider the combined effects on human health
2 that can result from exposure to toxic substances that share a common mechanism of toxicity
3 (e.g., organophosphates). This cumulative risk assessment approach is based on an evaluation of
4 the potential for people to be exposed to more than one member of a group of chemicals at a time
5 and considers exposures from food, drinking water, and residential sources. It is important to
6 note that for any group of toxic substances with a common mechanism of action, agents within
7 that group that have low toxicity but the potential for high exposure can present a risk similar to
8 a toxic substance with higher toxicity and lower exposure potential. It seems likely that the
9 hazard, dose-response, and exposure assessment components of the cumulative risk assessment
10 process will be greatly improved by the elucidation of mechanisms or modes of action made
11 possible by genomics data.

12
13 The Agency for Toxic Substances and Disease Registry (ATSDR) of the Department for
14 Health and Human Services, and EPA, in addition to other Federal research and regulatory
15 agencies, supports *in vitro* and *in vivo* research to further understand the chemical
16 characterization, molecular mechanisms of action, and toxicity of chemical mixtures and their
17 relationship with human health effects and other biological systems. Genomics can provide the
18 tools for accomplishing this work that has to date been very expensive and required a large
19 commitment to animal testing.

2. Risk Assessment Activities Potentially Affected by Genomics Information

23 Genomics technologies will aid in the identification of unique patterns of gene expression in
24 aquatic and terrestrial organisms and human cell-based models induced by exposure to multiple
25 environmental stressors. Several studies already have described tissue-specific transcriptional
26 patterns and have begun to address the concept of “fingerprinting” for chemical mixtures in
27 laboratory animals (Bartosiewicz et al., 2001a, 2001b) and in specific cell lines (Mumtaz et al.,
28 2002).

a. Representative Activities

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

- 35 • The toxicity of a test mixture could be evaluated through effects on genomic
36 biomarkers that have been linked to adverse effects.
 - 37
 - 38 • Genomics technologies could be applied to the evaluation of constructed test
39 mixtures to examine how chemicals may interact (additivity, synergism,
40 antagonism).
 - 41
 - 42 • Genomic diagnostic indicators may help identify components of chemical
43 mixtures with unidentified constituents through the use of a genomic
-

1 “fingerprint” database. For example, genomics technologies could be applied to
2 evaluating the differences in gene expression between a mixture of known
3 chemical constituents and a test mixture containing unknown chemical
4 constituents. Differences in response to the two mixtures could then be
5 evaluated and attributed to unknowns in the test mixture. This approach could
6 be applied to exposure assessments.
7

8 **b. Additional Activities**
9

10 Genomics may have additional regulatory implications for evaluating chemical
11 mixtures for other offices, as well as regions, states, and tribes. (Further details on the
12 following activities are found in Appendix I.)
13

14 Program Offices

- 15 • OSWER, OAR, OW: identifying unknowns in mixtures at contaminated sites
16 and in media, or during monitoring
17

18 Regions, States, and Tribes

- 19 • Identifying unknowns in mixtures at contaminated sites during monitoring
20
21
22
23
24
-

IV. Research Needs and Activities

A. Introduction

In order to incorporate any particular type of genomics data into decision making, the link between molecular indicator, exposure, and adverse outcome has to be established. In addition, the dose-response curve needs to be delineated, inter- and intraspecies variations in response need to be quantified, and detection limits and variability for the genomic indicator need to be established. Understanding the normal variability of gene expression, protein, and metabolite profiles is critical for evaluating any changes induced by stressors. Developing baseline genomic responses of species to stressors other than chemical challenges will also be crucial in establishing the utility of specific genomic indicators as markers of response. 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.

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/nheerl/comptoxframework/comptoxframeworkfinaldraft7_17_03.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

1 genetics information for the majority of such species. Attempts are being made to help rectify
2 this, in part through a collaboration with the Department of Energy's Joint Genome Institute to
3 provide cDNA libraries of the fathead minnow and a frog species (e.g., *Xenopus tropicalis*, the
4 Western clawed frog). Similar efforts with other species will be needed so that a set of sentinel
5 species can be made available for environmental assessment.

6
7 Research is needed to develop a functional approach to establishing a linkage between
8 genomic changes at the mRNA level and protein changes and cellular and tissue changes. This
9 type of approach can provide information necessary for eventually developing a systems biology
10 approach for defining pathways to disease or other adverse outcomes. These research needs will
11 require a strategic hiring process so that the necessary expertise is available within the Agency.
12 Collaborations will also be required to address the complex issues associated with a systems
13 biology approach. Analysis of the large data sets generated through genomic assays will require
14 the development of bioinformatic and computational methods. In addition, enhanced QSAR
15 methods are required because the judicious use of QSAR approaches can greatly reduce the
16 reliance on experimental approaches for establishing chemical priorities for additional research.

17
18 The initial steps in developing approaches for prioritizing chemicals for additional research
19 that could lead to the development of a risk assessment are described in the Agency's
20 Computational Toxicology initiative. The overall approach requires the development of
21 toxicological pathways for candidate chemicals such that the key events leading to specific
22 adverse outcomes can be identified. Chemicals can be designated as requiring further research if
23 they are predicted to initiate the key events for a particular adverse outcome.

24
25 The Agency's Computational Toxicology initiative stems from an FY02 Congressional
26 mandate to explore alternatives to the use of animals in toxicological studies. To address this
27 mandate, ORD is using endocrine disrupting chemicals (EDCs) as model compounds in research
28 that includes *in silico*, *in vitro*, and *in vivo* approaches as proof-of-concept for the overall
29 approach. The objectives of this effort are to determine the feasibility of using genomics and
30 computational toxicology to facilitate the prioritization of chemicals for screening. Another goal
31 is to reduce the need for some *in vivo* assays while providing a greater breadth of coverage of
32 endocrine alterations and a better predictiveness of potential adverse health outcomes.

33
34 Similar approaches are being pursued for other chemical classes and other adverse outcomes
35 (e.g., disinfection by-products and cancer, and conazoles and cancer, reproductive,
36 developmental, and neurological effects). The aim is to establish a priority for chemicals that
37 require further study for development of risk assessments.

38
39 Another example of ongoing research in support of prioritization involves determining the
40 effects of contaminants on aquatic animals using protein profiling. This will result in the ability
41 to rapidly screen chemicals based on specific modes of action. Protein profiles can be used as
42 specific biomarkers of effects of bioactive compounds. The sheepshead minnow is being used as
43 a model species for proof-of-concept studies.

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

6 7 **2. Monitoring Research Needs and Activities**

8
9 The research needs to support EPA's monitoring programs cover a broad range of issues,
10 including environmental monitoring and public health monitoring, specifically in the areas of
11 epidemiology and molecular epidemiology. It is likely that genomics technologies will prove
12 productive in each of these fields. The following examples provide an indication of the types of
13 research needs that the Agency faces in an expanded monitoring program and highlight Agency
14 activities that may help to address these needs.

15
16 If environmental and human health assessments are to employ biological indicators, reliable
17 and informative markers of exposure, dose, and response must be selected. The more that is
18 known about the mechanisms involved in the pathway from exposure to adverse outcome or
19 response, the more readily informative biomarkers can be identified. The research need,
20 therefore, is to develop these mechanistic data, select proposed informative biomarkers, and
21 utilize these in field conditions or in molecular epidemiological studies. Validation of the
22 biomarkers can be achieved through human or environmental health assessments, to establish
23 how accurately the various biomarkers predicted outcomes. This is a massive effort that will
24 require considerable collaboration within the Agency and outside the Agency.

25
26 EPA is conducting research to address the need for environmental biomarkers. For example,
27 Agency scientists are conducting a comparison of the sensitivity of cellular indicators of genetic
28 damage in model stream fish using controlled laboratory exposures and subsequent field
29 validation. These indicators of cellular damage are likely to have an application in ecological
30 monitoring projects. Another example of ongoing research in this area involves the development
31 of methods for measuring the induction of the vitellogenin gene during water monitoring studies.
32 This indicator provides information on exposure of an organism to an endocrine disrupting
33 chemical. The next step is to conduct research that can help establish the relationship between
34 measured exposure and adverse effect. Such information is required prior to the use of the
35 method in a regulatory setting.

36
37 Similarly, for microbial source tracking, the relationship between existing indicators (e.g.,
38 total coliforms, enterococci, etc.) and genomics-based indicators must be established. The
39 relationship between occurrence and disease response in humans from human and non-human
40 sources of pathogens, especially bacteria, must be defined. These types of information will be
41 used by state and local agency decision makers to determine which methods are "acceptable."
42 This is a significant step because most agencies have limited resources and are often reluctant to
43 change to new technologies because of the associated high capital and human resource

1 investments that must be made. The Agency is evaluating ways to apply DNA-based technology
2 to detect and track fecal contamination to its source in complex environmental matrices,
3 including recreational and drinking water sources. A microarray method to identify potential
4 waterborne pathogens is also under development by the Agency. The Office of Water is
5 currently supporting research investigating the effects of specific gene combinations that are
6 associated with waterborne pathogen virulence. These projects could be applied to ambient and
7 drinking water monitoring.

8
9 The development of genetic markers and/or proteomic markers of plant responses to
10 herbicides and other xenobiotics is needed by the Regions to enhance monitoring capability for
11 assessing effects of spray drift and determining which plant groups are most likely to be at risk
12 from xenobiotics. Researchers are currently studying proteomic responses of plants to high
13 potency, low-dose herbicides as a method of monitoring exposure. Further, markers for
14 monitoring gene transgression from Genetically Modified (GM) crops to non-crop plants as well
15 as for detecting contamination of non-GM seed shipments or foodstuffs with GM material, are
16 being developed for use by the Regions.

17
18 Thus, the aim of a genomics research program in support of the Agency's monitoring efforts
19 is to develop informative biomarkers of response for the assessment of the impact of
20 environmental exposures on human and environmental health. In parallel, such bioindicators can
21 be used to assess the impact of regulatory actions on human and environmental health.

22 23 24 **B. Research Needs and Activities for Risk Assessment**

25 26 **1. Mode of Action Research Needs and Activities**

27
28 There are numerous issues associated with MOAs that require additional research before
29 genomics technologies can be fully utilized in risk assessments. An overriding issue that affects
30 more than just the MOA is the need to relate changes in gene expression to adverse effects. To
31 establish the linkages between genomic changes and adverse outcomes, reliable data sets for
32 gene expression at the RNA and protein levels are required. These data need to include a range
33 of sample times and exposure concentrations and be repeatable. A parallel need is the
34 development of expertise in analyzing these data so that profiles of responses at the molecular
35 level can be produced and linked to specific chemicals or mixtures. This type of correlation then
36 has to be extended to adverse outcomes at the organ, tissue, and whole animal level. This
37 approach could be applied in both human health and ecological risk assessments.

38
39 Metabolic pathways for chemicals need to be defined, and the active metabolites that cause
40 cellular responses need to be identified. The use of mRNA or protein arrays would enhance the
41 Agency's ability to address this issue in a timely fashion. The initial requirement is to establish
42 a set of profiles that are *in toto* representative for each known metabolic pathway.
43

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

15
16 The following are descriptions of some of the research activities that are underway in ORD
17 that address MOA. The list is not exhaustive, but does provide an idea of the breadth of the
18 activities under the MOA umbrella. A significant fraction of the genomics research currently
19 ongoing in the Agency is directed towards identifying MOAs for a range of chemicals for a
20 number of different adverse outcomes, including cancer, endocrine disruption, reproductive and
21 developmental effects, and neurotoxicity. These research activities are briefly presented to show
22 their range.

23
24 In an attempt to use MOAs in the harmonization of risk assessment approaches, a research
25 team is comparing the effects of a group of conazole pesticides in different tissues and for
26 different endpoints. The aim is to establish if a common MOA is able to explain the range of
27 different endpoints and the specificity of tissue responses.

28
29 Genomics technologies are also being used to characterize the MOA of selected drinking
30 water disinfection by-products for use in BBDR models. In the same studies, genomics tools are
31 being applied to develop markers of response that will provide information for predicting
32 adverse outcomes at low doses.

33
34 Gene array technologies are being used to identify biomarkers that will be informative of
35 responses specific to the human testis. The MOA that is being developed for rodents will be
36 used to establish whether responses to particular chemicals have relevance to humans. Other
37 studies are developing various biomarkers of response for environmental monitoring with
38 relevance to humans.

39
40 Additional efforts are underway to establish if readily available cells in humans, such as
41 peripheral lymphocytes or buccal cells, can be used as predictors of adverse responses in tissues
42 that are targets for adverse outcomes such as cancer, and reproductive and developmental
43 effects. Initial approaches involve the use of microarrays to study gene expression patterns in

1 lymphocytes and in germ cells using appropriate animal models and selected chemical stressors.
2 In a similar way, genomics approaches are being used to establish whether markers of
3 susceptibility identified in readily available cells in humans, such as peripheral lymphocytes or
4 skin fibroblasts, can predict sensitivity to adverse health outcomes. These initial gene expression
5 studies will be expanded to include protein changes and functional associations with exposures.
6 Currently, these types of approaches are in the early stages of development within the Agency.
7

8 As a part of the current effort in genomics, several groups are attempting to identify
9 informative biomarkers of response in laboratory animals as well as in sentinel species for use in
10 ecological assessments. It is proposed that MOA is a viable way of conducting interspecies
11 comparisons of outcome.
12

13 The current and proposed Agency genomics research directed towards enhancing our
14 knowledge of the various MOAs whereby chemicals can induce adverse outcomes is focused on
15 identifying key events along toxicological pathways from exposure to response. The
16 identification of key events will not only aid risk assessment approaches for single chemicals,
17 but will enhance efforts to harmonize risk assessment, to predict responses to chemical mixtures,
18 and to identify susceptible populations.
19

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

21
22 Before the issue of how to incorporate susceptible populations into human health or
23 ecological risk assessment can be addressed through EPA policy, the methods for identifying
24 susceptible populations must first be developed along with quantitative methods for assessing the
25 magnitude of the sensitivity. To accomplish this, the Agency must develop knowledge of the
26 MOA for a chemical(s) of concern as well as the prevalence of this MOA in the population. For
27 example, this will require establishing the frequency of SNPs and their effects within human
28 populations as part of the identification of a susceptible population. An important proviso to
29 these types of studies is that ethical, social, and legal issues need to be addressed before starting
30 such work.
31

32 There is a significant research effort in ORD to address the issue of children as a susceptible
33 population. Specific focus regards the induction of diseases in children and the effects of early
34 life exposures on the development of adult diseases. The aims are to determine the magnitude of
35 any sensitivities and the underlying mechanisms that might account for increased sensitivity.
36 The genomics research component is directed towards developing informative biomarkers of
37 response that can eventually be used in animal model systems to predict adverse outcomes from
38 specific exposure scenarios and in human epidemiological studies. These informative
39 biomarkers can also encompass specific genetic markers such as SNPs.
40

41 Genomic methods, including proteomic approaches, may also be useful in more accurately
42 estimating exposures of individuals to contaminants in the environment, thereby identifying
43 susceptible populations at the exposure level. These emerging technologies could lead to the

1 development of personal dosimeters for a wide range of chemicals such that exposure would be
2 assessed at the individual level. Similarly, genomic-level biomarkers (e.g., enhanced personal
3 microarray technologies) could provide a real-time, high throughput method for screening
4 potentially exposed individuals for incipient effects. While these approaches are technologically
5 feasible, the Agency has no definitive plans to develop research programs along these lines in the
6 near-term.
7

8 Sensitive fish and/or wildlife species might serve as early indicators of overall ecosystem
9 health and as sentinels for risks to human health. In cases where chemical or pathogen
10 contamination reduces species fitness, genomics technologies could be used to examine the
11 genetic makeup of a species in order to determine the biochemical mechanism of the adverse
12 effect(s). Like humans, plants and animals possess genetic polymorphisms that code for multiple
13 metabolic enzyme variants. In addition, levels and forms of the same enzyme (e.g., the
14 cytochrome P450 family of enzymes) vary between species and between life stages within
15 species. Thus, as the genomes of species are sequenced, genomics can be used to identify the
16 most sensitive species and sensitive life stages. This will significantly enhance our ability to set
17 scientifically defensible water quality criteria or sediment and soil protection values under the
18 Clean Water and Safe Drinking Water Acts.
19

20 Similarly, understanding genetic-based differences among plants and wildlife species in
21 terms of the MOAs of chemicals is a fundamental step towards understanding which one(s) will
22 be responsive. In the long term, this will enable more accurate cross-species extrapolations and
23 will significantly reduce the need for animal testing. The Computational Toxicology initiative in
24 ORD is directed towards utilizing genomics to identify toxicity pathways. The current focus is
25 on humans and fish. In the future, genomic-based approaches need to be developed for other
26 wildlife species, as well as for aquatic and terrestrial plant species.
27

28 ORD is currently developing tools to incorporate genomics technologies into population
29 dynamics models to enhance their predictive and explanatory power for assessing risks to
30 populations of wildlife and aquatic life. Genetic processes include the distribution and dynamics
31 of neutral and fitness-linked genetic markers. Depending upon their sophistication and data
32 requirements, the resulting population models can be used in screening to definitive tiers in the
33 ecological risk assessment process. In addition, a genetic dissection of the mechanisms of
34 resistance to anthropogenic contaminants is underway in zebrafish and fathead minnows. Both
35 of these research efforts will yield information regarding the sensitivities of various fish species
36 and will likely be helpful in projecting the potential impacts of environmental contaminants on
37 ecosystem health.
38

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

1 In the context of assessing the impact of chemical exposures to overall human and ecological
2 health, the influence of susceptible populations is of critical importance. The needs are to
3 consider the roles of both life stages and genetic variation in the etiology of susceptibility. While
4 there is ongoing research addressing these issues, it is currently relatively limited.
5

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

7

8 Exposures of human and wildlife populations to environmental contaminants generally
9 involve complex mixtures of chemicals, rarely individual chemicals. Although there have been
10 some efforts to address responses to both simple and complex mixtures, much of the past and
11 current research of the Agency has addressed the risk from exposures to single chemicals.
12 Clearly, addressing the overall toxicological responses to mixtures is a complex problem that
13 may require approaches different from those used for single chemicals. Given the charge to the
14 Agency to increase its focus on research into the effects of mixtures, it is important to assess how
15 genomics techniques might aid in meeting this need.
16

17 A range of research is needed to assess the risks of chemical mixtures. For example, it is
18 necessary to determine if the MOA approach discussed above can be used to determine whether
19 a mixture can induce a qualitatively different set of key events than any of the individual
20 chemicals constituting the mixture. The next step would be to determine whether there are
21 quantitative differences in the induction of key events by mixtures as compared to the individual
22 chemicals and to use genomic measures to assess the magnitude of these definitive key events.
23 Because both human health and ecological risk assessment could benefit from a genomics
24 approach, discussion is underway concerning how to incorporate this type of research into the
25 Computational Toxicology initiative.
26

27 The research needs for mixtures overlap considerably with those of prioritization, MOA, and
28 susceptible populations. Thus, mixtures assessment is an issue that will need to be addressed in
29 concert with these other Agency priority regulatory needs.
30
31
32

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, 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.

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 has outlined the potential of genomics technologies to improve and refine the current approach to regulatory applications and risk assessment and has 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 must meet 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. As noted throughout this paper, changes in gene expression at the mRNA and protein levels need to be related to cellular effects and,

1 ultimately, to adverse outcomes. In many ways, the detection of gene expression changes is the
2 easiest part of a genomic assessment. However, in a risk assessment framework, it is necessary
3 to link a function to genes whose expression is altered. Gene expression changes that encompass
4 defense mechanisms, which may be adaptive or beneficial and bear no causal relationship with
5 the development of pathologies, must be separated from those that damage key cell functions
6 (e.g., cell cycle control, structural integrity of proteins, control of free radicals, or loss of
7 homeostasis and DNA repair mechanisms). Combining the findings of gene expression studies
8 with data from *in vivo* chemical exposure of genetically altered animal models (e.g., knockout or
9 null mice) is a powerful way to link specific genes to specific detrimental outcomes. Simpler
10 whole organ systems may also offer powerful means to link genomic response to adverse effect.
11 Key biological systems have fundamental genomic processes, some of which, if altered, are
12 universally deleterious.

14 2. Interpretation of Genomics Information for Risk Assessment

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

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

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

1 process, and thus aid in formulating regulatory policy and making regulatory decisions.

2
3 Understanding the MOA of environmental agents that induce toxic effects other than cancer
4 or induce carcinogenicity in animal models should facilitate the assessment of the relevance of
5 these findings in protecting human health and safe guarding the environment. An important
6 issue for extrapolation of responses in animal models to humans or environmental endpoints is to
7 establish whether the MOA in the test species is relevant in the target species. A range of
8 different types of data can be used to establish a MOA, but the endpoints for cross-species
9 extrapolation generally are more limited. Such approaches will aid in addressing EPA's
10 challenge to harmonize risk assessment approaches for different outcomes.

11 12 **3. Recommendations to Address Research Challenges**

13
14 In order to contribute to the development of linkages between genomic changes and adverse
15 outcomes and to the interpretation of genomics information for risk assessment, the Agency
16 should aggressively support and build its own genomics research through the ORD
17 Computational Toxicology initiative and support external research through competitive grants
18 and contracts. Research plans and timing should be guided by developments in the genomics
19 field. Through appropriate direction of its research, EPA can support important regulatory
20 applications that are more likely to arise in the near future. These include priority-setting
21 activities such as high throughput screening of chemicals, and monitoring activities such as
22 source tracking of pollutants and pathogens in water. EPA should also encourage industry
23 efforts to conduct genomics research necessary for application in risk assessment, when
24 developing its data. It is critical that the Agency coordinate research with other agencies and
25 institutions to link genomic changes and adverse outcomes

26 27 **B. Technical Development Challenges**

28 29 **1. Framework for Analysis and Acceptance Criteria for Genomics Information**

30
31 EPA acknowledges that genomics technologies will eventually contribute to risk assessment
32 through a better understanding of mechanisms of chemical toxicity, dose-response relationships,
33 identification of susceptible populations, and estimates of uncertainty factors. However, to date,
34 EPA has had limited access to relevant genomics data to begin examining its potential influence.
35 Even without specific cases relevant to the Agency, it is clear that a plan is needed to develop
36 methods for incorporating these types of information into the decision making process.

37
38 EPA and other regulatory agencies recognize the requirement to develop acceptance criteria
39 for genomics data. The Food and Drug Administration's (FDA) Center for Drug Evaluation and
40 Research is in the process of identifying which pharmacogenetic data developed by companies
41 for the evaluation of human drugs will be required by the FDA due to the regulatory implications
42 of the data (FDA, 2003)

43 (<http://www.fda.gov/OHRMS/DOCKETS/98fr/2003d-0497-gdl0001.pdf>). The FDA is

1 consulting with its Federal Advisory Committee and stakeholders to develop its policies and
2 guidance. The efforts at the FDA are aimed at developing a framework for submission, storage,
3 analysis, and regulatory review of genomics data. This is driven, in part, by the high level of use
4 of genomics information for human drug development and evaluation.
5

6 The use of genomics information for the analysis of risks of industrial chemicals,
7 agrochemicals, and microbial contaminants is not as widespread at the current time. However,
8 as the understanding of genomics data and its relevance and applicability to chemical and
9 microbial risk analysis increases, EPA will need to continue developing its own technical
10 framework for the consideration of genomics information for scientific and regulatory purposes.
11 The full range of ethical and legal issues (Marchant, 2003) will also have to be considered as
12 genomics is incorporated into the Agency’s risk assessment process.
13

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

15
16 The Genomics Task Force recommends that the Agency charge a workgroup with
17 developing a technical framework for analysis and acceptance criteria for genomics information
18 for scientific and regulatory purposes. This framework should build upon EPA’s Interim Policy
19 on Genomics. Issues that need to be considered in developing such a framework include
20 consideration of the performance of assays across genomic platforms (e.g., reproducibility,
21 sensitivity) and the criteria for accepting genomics data for use in a risk assessment (e.g., assay
22 validity, biologically meaningful response). It is essential for the Agency to continue to engage
23 other federal agencies, such as the FDA, NIEHS, and Department of Energy, as well as other
24 stakeholders, including industry and academia, when developing this framework. Such a
25 framework, once established, can be used by the EPA program offices to determine the
26 applicability of specific genomics information to the evaluation of chemical risks under various
27 statutes. The Interim Policy of considering genomics data on a case-by-case basis should
28 continue to be applied until the technical framework is completed.
29

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

31 **1. Applying Strategic Hiring Practices to Recruit Individuals who Possess “Genomics Core 32 Competencies”**

33
34
35 An important undertaking to will be to identify the skills needed to establish “genomics core
36 competencies” and to apply strategic hiring practices to recruit individuals who possess these
37 skills. It will be essential to have technical specialists in genomics on staff in the crucial areas of
38 research, analysis, systems biology, bioinformatics, and risk assessment to enhance the Agency’s
39 expertise in genomics and related technologies.
40

41 **2. Training EPA Risk Assessors and Managers to Interpret and Understand Genomics 42 Data in the Context of a Risk Assessment**

43

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

8
9 A related concern is the capacity of regions, states, tribes, and local agencies to implement
10 genomics tools and to evaluate genomics data, particularly with respect to their responsibilities
11 under delegated programs. They will need resources, technical support, and targeted training to
12 bring the scientists and managers within their organizations to a point where they will be able to
13 effectively use genomics tools in their regulatory decision making, especially with respect to risk
14 characterizations. Regional, state, tribal and local agencies' use of genomics tools will require
15 both capital investment for analytical equipment, plus ongoing expenses for disposables such as
16 the microarrays and associated supplies.

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

18
19
20 EPA programs and regions should apply strategic hiring practices to recruit individuals who
21 possess genomics skills and should consider existing guidance from other agencies, such as the
22 Centers for Disease Control, regarding recommended “genomics core competencies.”

23
24 The Genomics Task Force also recommends that the Agency convene a workgroup tasked
25 with developing training modules for the interpretation and application of genomics data for risk
26 assessments for both risk assessors and risk managers. It would be useful to develop and initiate
27 training in the near future to prepare risk assessors and risk managers, because genomics issues
28 have begun to arise in environmental decision making. The initial training could address basic
29 genomics concepts, technologies and potential applications including consideration of the basic
30 steps necessary to interpret and apply genomics data to a risk assessment. Regions 9 and 10, in
31 collaboration with ORD, have already developed and conducted basic genomics training on a
32 pilot basis. Development of Agency-wide training materials could build upon these efforts and
33 those of external sources. It is expected that the training would need to be revised and expanded
34 as our understanding of genomics improves over time. The training could also be offered on a
35 limited basis to tribes, states, and local governments to assist in the development of their
36 capability to effectively use genomics tools in regulatory decision making. Finally, the
37 Genomics Task Force recommends that the development of training tools and workshops be
38 conducted in collaboration with other agencies and institutions.

39
40
41 In summary, the Genomics Task Force recommends that EPA begin taking steps to address
42 the identified research, technical development, and capacity challenges to strengthen its
43 capability to effectively use genomics information. It is essential for the Agency to continue to

1 collaborate with other federal agencies, academia, the regulated community, and other
2 stakeholders in this endeavor, in order to benefit from ongoing advances in genomics in the
3 wider scientific and regulatory communities.
4
5

References

- 1
2
3 Aardema, M.J., MacGregor, J.T. 2002. Toxicology and genetic toxicology in the new era of
4 “toxicogenomics”: impact of “-omics” technologies. *Mutation Research* 499(1) 13-25.
5
6 Bartosiewicz, MJ, Jenkins, D., Penn, S., Emery, J., Buckpitt, A. 2001a. Unique gene expression
7 patterns in liver and kidney associated with exposure to chemical toxicants. *Journal of*
8 *Pharmacology and Experimental Therapeutics* 297:895-905.
9
10 Bartosiewicz, M., Penn, S., Buckpitt, A. 2001b. Applications of gene arrays in environmental
11 toxicology: Fingerprints of gene regulation associated with cadmium chloride, benzo(a)pyrene,
12 and trichloroethylene. *Environmental Health Perspectives* 109:71-74.
13
14 Bishop, W. E., Clarke, D. P., Travis, C. C. 2001. The Genomic Revolution: What Does It Mean
15 for Risk Assessment? *Risk Analysis* 21(6): 985-987.
16
17 Feron, V. J., Groten, J. P. 2002. Toxicological evaluation of chemical mixtures. *Food and*
18 *Chemical Toxicology* 40:825-839.
19
20 Fielden, M.R., Zacharewski, T.R. 2001. Challenges and limitations of gene expression profiling
21 in mechanistic and predictive toxicology. *Toxicological Sciences* 60:6-10.
22
23 Griffith, J.F., Weisberg, S.B., McGee, C.D. 2003. Evaluation of microbial source tracking
24 methods using mixed fecal sources in aqueous test samples. *Journal of Water and Health*
25 1(4):141-152.
26
27 Haber, L.T., Maier, A., Gentry, P. R., Clewell, H. J. Dourson., M. L. 2002. Genetic
28 polymorphisms in assessing interindividual variability in delivered dose. *Regulatory Toxicology*
29 *and Pharmacology* 35:177-197.
30
31 Hakkola, J., Pelkonen, O., Pasanen, M., Raunio, H. 1998. Xenobiotic-metabolizing cytochrome
32 P450 enzymes in the human fetoplacental unit: Role in intrauterine toxicity. *Critical Reviews in*
33 *Toxicology* 28(1):35-72.
34
35 Henry, C.J., Phillips, R., Carpanini F., Corton, J.C., Craig, K., Igarashi, K., Leboeuf, R.,
36 Marchant, G., Osborn, K., Pennie, W.D., Smith, L.L., Teta, M.J., Vu, V. 2002. Use of genomics
37 in toxicology and epidemiology: Findings and recommendations of a workshop. *Environmental*
38 *Health Perspectives* 110(10):1047-1050.
39
40 Hertzberg, R. C., MacDonell, M. M. 2002. Synergy and other ineffective mixture risk
41 definitions. *Science of the Total Environment* 288:31-42.
42
43 Klaper, R., Euling, S., Frederick, R., Sonawane, R. 2003. Current Use and Future Needs of
-

1 Genomics in Ecological and Human Health Risk Assessment. Internal Workshop Report, May
2 8, Alexandria, VA.

3
4 Marchant, G.E. 2003. Genomics and Toxic Substances: Part I- Toxicogenomics.
5 *Environmental Law Reporter* 33: 10071-10093.

6
7 Marchant, G.E. 2002. Genomics and Environmental Regulations: Scenarios and Implications.
8 Report under Foresight and Governance Project at the Woodrow Wilson Center for Scholars. 29
9 pp.

10
11 Mumtaz, M. M., Tully, D. B., El-Masri, H. A., and De Rosa, C. T. 2002. Gene induction studies
12 and toxicity of chemical mixtures. *Environmental Health Perspectives* 110 (Supplement 6): 947-
13 956.

14
15 Organization for Economic Cooperation and Development. 2001. The Chemicals Industry. In:
16 Environmental Outlook, OECD Publications, Paris. pp. 224-225.

17
18 Pohl, H. R., Hansen, H., and Chou, C. H. S. J. 1997. Public health guidance values for chemical
19 mixtures: Current practice and future directions. *Regulatory Toxicology and Pharmacology*
20 26:322-329.

21
22 Porter, W. P., Hindsill, R.D., Fairbrother, A., Olson, L. J. Jaeger, J., Yuill, T. M. Bisgaard, S.
23 Hunter, W. G., Nolan. K. 1984. Toxicant-disease-environment interactions associated with
24 suppression of immune system, growth and reproduction. *Science* 224:1014-1017.

25
26 Sappington, L.C., Mayer, F.L., Dwyer, F.J., Buckler, D.R., Jones, J.R., Ellersieck, M.R. 2001.
27 Contaminant sensitivity of threatend and endangered fishes compared to standard surrogate
28 species. *Environmental Toxicology and Chemistry* 20(12):2869-2876.

29
30 Simpson, J., Santo Domingo, J.W, Reasoner, D.J. 2002. Microbial source tracking: State of the
31 science. *Environmental Science and Technology* 36(24): 5279-5288.

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

36
37 Suk, W. A., Olden, K., and Yang, R. S. H. 2002. Chemical mixtures research: Significance and
38 future perspectives. *Environmental Health Perspectives* 110 (Supplement 6):891-892.

39
40 Sumner, L.W., Mendes, P., Dixon, R.A. 2003. Plant metabolomics: large-scale phytochemistry
41 in the functional genomics era. *Phytochemistry* 62:817-836.

42
43 U.S. Department of Energy, Oak Ridge National Laboratory, Human Genome Project

1 Information, Genome Glossary. http://www.ornl.gov/sci/techresources/Human_Genome/glossary

2
3 U.S. Department of Energy, Office of Science, Biological and Environmental Research Division.
4 http://www.sc.doe.gov/ober/ERSD/ersd_nabir.html

5
6 U.S. Department of Health and Human Services, Public Health Service, National Institutes of
7 Health, National Institute of Environmental Health Sciences, National Center for
8 Toxicogenomics, Glossary. <http://www.niehs.nih.gov/nct/glossary.htm>

9
10 U.S. Department of Health and Human Services, Public Health Service, National Institutes of
11 Health, National Cancer Institute, Understanding Gene Testing, Glossary.
12 <http://www.accessexcellence.org/AE/AEPC/NIH/gene27.html>

13
14 U.S. Department of Health and Human Services, Public Health Service, National Institutes of
15 Health, National Institute of General Medical Sciences, The Chemistry of Health, Glossary.
16 http://www.nigms.nih.gov/news/science_ed/chemhealth/glossary.html

17
18 U.S. Environmental Protection Agency. 2003a. Draft Final Guidelines for Carcinogen Risk
19 Assessment. EPA/630/P-03/001A. <http://www.epa.gov/ncea/raf/cancer2003.htm>

20
21 U.S. Environmental Protection Agency. 2003b. A Framework for a Computational Toxicology
22 Research Program in ORD. EPA/600/R-03/065.
23 http://www.epa.gov/nheerl/comptoxframework/comptoxframeworkfinaldraft7_17_03.pdf

24
25 U.S. Environmental Protection Agency. 2003c. Pesticide product label system (PPLS).
26 <http://www.epa.gov/pesticides/pestlabels/>.

27
28 U.S. Environmental Protection Agency, Science Policy Council. 2002a. Interim Policy on
29 Genomics. <http://www.epa.gov/OSP/spc/genomics.pdf>

30
31 U.S. Environmental Protection Agency. 2002b. A Review of the Reference Dose and Reference
32 Concentration Processes. EPA/630/P-02/002F.
33 <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=55365>

34
35 U.S. Environmental Protection Agency. 1999. Glossary of Integrated Risk Information System
36 (IRIS) terms. <http://www.epa.gov/iris/gloss8.htm>

37
38 U.S. Environmental Protection Agency, Risk Assessment Web Page.
39 <http://www.epa.gov/ebtpages/enviriskassessment.html>

40
41 U.S. Food and Drug Administration. 2003. Guidance for Industry: Pharmacogenomic Data
42 Submissions, Draft Guidance.
43 <http://www.fda.gov/OHRMS/DOCKETS/98fr/2003d-0497-gd10001.pdf>

- 1 Waters, M., Boorman, G., Bushel, P., Cunningham, M., Irwin R., Merrick, A., Olden, K., Paules,
- 2 R., Selkirk, J., Stasiewicz, S., Weis, B., Van Houten, B., Walker, N., Tennant, R. 2003.
- 3 Systems toxicology and the Chemicals Effects in Biological Systems (CEBS) Knowledge Base.
- 4 Environmental Health Perspectives Toxicogenomics 111(6): 811-824.

Appendix I: Details on Additional Activities Potentially Affected by Genomics Information

I. Prioritization

A. Program Offices

OPPT - PMNs. 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 - VCCEP: Genomics data may be useful in the VCCEP program. These data could support, or potentially supplant, current toxicity tests.

OPPTS/OSCP - EDSP: OSCP has been developing a tiered approach for testing under the Endocrine Disruptors Screening Program (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: Genomics data may provide useful information to ORD researchers and managers in prioritizing chemicals for future research.

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

2
3 Genomics technologies could provide regions, states, and tribes with a fast, relatively low
4 cost method to prioritize their areas of focus and deployment of resources for delegated
5 program site remediations and chemical evaluations.
6

7 **II. Monitoring**

8
9 **A. Program Offices**

10
11 **OPP:** OPP uses exposure monitoring data for new chemicals generated via Experimental
12 Use Permits (EUPs) in registration decisions, and exposure monitoring data for currently
13 registered chemicals for reregistration eligibility decisions (REDs). For these EUPs and
14 REDs, genomics data could be used to track the movement of pesticides off-site via spray
15 drift or into ground or surface water. Biological monitoring data for human and wildlife
16 exposures and potential effects could contribute to regulatory decisions, and genomics
17 technologies could provide information about occupational exposures and wildlife incident
18 data for reregistration decisions.
19

20 **OAR/Office of Air Quality Planning and Standards (OAQPS):** Genomics technologies
21 could contribute to stationary source monitoring conducted under the Clean Air Act (CAA).
22

23 **OSWER - Superfund monitoring:** Genomics technologies could be applied to monitor
24 movement of contaminants off-site from Superfund sites prior to remedial actions. In
25 evaluating contaminated sites, near term benefits could be derived by targeting toxic
26 chemical remediation using biomarkers present in lower organisms (e.g., metallothionein
27 expression). Bioavailability, a key parameter in determining the toxicity of a chemical, can
28 be effectively determined using genomics tools; thus greater precision could be achieved in
29 remedial activities. That is, the truly hazardous compounds could be identified and removed
30 with more precision, and other materials need not be disturbed unnecessarily. Further,
31 monitoring the operation and maintenance of remedial actions and residual contaminants at
32 Superfund sites that have undergone cleanup could be enhanced through genomics
33 technologies.
34

35 **OSWER/OSW - RCRA-required monitoring:** Genomics could contribute to post-clean up
36 monitoring activities conducted under RCRA.
37

38 **OEI and ORD:** OEI and ORD expect to use genomics approaches to develop a selection of
39 informative bioindicators for monitoring the exposures and effects of stressors on human and
40 ecological health.
41

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

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

14
15 **Air Quality Monitoring Program:** EPA's ambient air quality monitoring program for
16 criteria pollutants is carried out by State and local agencies. Genomics approaches could be
17 used in the state and local air monitoring programs.

18
19 **Endocrine-Disruptor Monitoring:** Another example of near-term decision making using
20 genomics data comes from a group of tribes in Northern California and Southern Washington
21 that is interested in the potential impact of exposure to pharmaceuticals and personal care
22 products and the potential for endocrine disruptor effects from municipal waste treatment
23 facilities and/or cattle grazing on the health of wild salmon populations that are part of tribal
24 cultural and economic resources. The Tribe proposes to use a series of molecular-biology-
25 based assays for exposure to hormonally active compounds, either a multiplex RT-PCR
26 approach or a multigene array. The information could be used to develop NPDES permit
27 limits and establish Tribal Water Quality Standards.

28
29 **Regional Pesticide Program Decisions:** Another use of genomics data may be in the area
30 of the Regions' Pesticides Programs. If genomics data show that a particular unregistered
31 pesticide is of concern in certain populations of humans, animals, or plants, the Region could
32 work with the State to assign a higher fine to a incident of use and/or offering for sale if the
33 violation was in an area where such a sensitive population existed. Alternatively, the "harm
34 value" might be lower if the non-target groups in an area were not considered to be part of a
35 sensitive population.

36 37 **III. Risk Assessment - Mixtures**

38 39 **1. Program Offices**

40
41 **OSWER, OAR, OW:** Genomics technologies could be used as diagnostic indicators to help
42 identify components of chemical mixtures with unidentified constituents. This approach
43 could be used in initial assessments of contaminated sites or in general monitoring

1 applications.
2

3 **2. Regions, States, and Tribes**
4

5 Regions, states, and tribes could apply genomics diagnostic indicators to identify
6 components of chemical mixtures with unidentified constituents for initial assessments of
7 contaminated sites or in general monitoring applications.
8
9

Appendix II: Glossary

1
2
3 **Allele:** an alternative form of a gene or any other segment of a chromosome¹.

4
5 **Bioinformatics:** the analysis of biological information using computers and statistical
6 techniques; the science of developing and utilizing computer databases and algorithms to
7 accelerate and enhance biological research¹.

8
9 **Biomarker:** a molecular indicator of a specific biological property; a biochemical feature or
10 facet that can be used to measure the progress of disease or the effects of treatment¹.

11
12 **Biotechnology:** the set of biological techniques developed through basic research and now
13 applied to research and product development. In particular, biotechnology refers to the use by
14 industry of recombinant DNA, cell fusion, and new bioprocessing techniques².

15
16 **Complementary DNA (cDNA):** DNA made from a messenger RNA (mRNA) template. The
17 single-stranded form of cDNA is often used as a probe in physical mapping¹.

18
19 **Computational Toxicology - Comp Tox:** the application of mathematical and computer
20 models and molecular biological approaches to improve the Agency's prioritization of data
21 requirements and risk assessments. (USEPA, 2003b)

22
23 **Deoxyribonucleic acid (DNA):** the substance of heredity; a large molecule that carries the
24 genetic information that cells need to replicate and to produce proteins³. The nucleic acid that
25 constitutes the genetic material of all cellular organisms and DNA viruses. The genetic
26 information is used in the synthesis of ribonucleic acids (RNAs) from DNA templates
27 (transcription) and in the synthesis of proteins from messenger RNA (mRNA) templates
28 (translation).

29
30 **Expressed sequence tag:** a unique stretch of DNA within a coding region of a gene that is
31 useful for identifying full-length genes and serves as a landmark for mapping¹.

32
33 **Gene:** the fundamental physical and functional unit of heredity. A gene is an ordered sequence
34 of nucleotides located in a particular position on a particular chromosome that encodes a specific
35 functional product (i.e., a protein or RNA molecule)².

36
37 **Gene chip technology:** development of cDNA microarrays from a large number of genes; used
38 to monitor and measure changes in gene expression for each gene represented on the chip².

39
40 **Gene expression:** process by which a gene's coded information is converted into the structures
41 present and operating in the cell. **Expressed genes** include those that are transcribed into mRNA
42 and then translated into protein and those that are transcribed into RNA but not translated into
43 protein (e.g., transfer and ribosomal RNAs)².

1 **Genetics:** the study of inheritance patterns of specific traits².

2
3 **Genetic testing:** analyzing an individual's genetic material to determine predisposition to a
4 particular health condition or to confirm a diagnosis of genetic disease².

5
6 **Genomics:** the study of all the genes of a cell or tissue, at the DNA (genotype), mRNA
7 (transcriptome), or protein (proteome) level (USEPA, 2002a).

8
9 **Genome:** all the genetic material in the chromosomes of a particular organism; its size is
10 generally given as its total number of base pairs².

11
12 **Genotype:** the genetic composition of an organism or a group of organisms; a group or class of
13 organisms having the same genetic constitution¹.

14
15 **Hazard Assessment:** the process of determining whether exposure to an agent can cause an
16 increase in the incidence of a particular adverse health effect (e.g., cancer, birth defect) and
17 whether the adverse health effect is likely to occur in humans⁴.

18
19 **In Silico:** literally “within silicon ”; refers to modeling research conducted with computers
20 only⁵.

21
22 **In Vitro:** literally, “in glass,” i.e., in a test tube or in the laboratory; the opposite of in vivo (in a
23 living organism)¹.

24
25 **In Vivo:** in a living organism, as opposed to in vitro (in the laboratory)¹.

26
27 **Knockout:** inactivation of specific genes. Knockouts are often created in laboratory organisms
28 such as yeast or mice so that scientists can study the knockout organism as a model for a
29 particular disease¹.

30
31 **Mapping:** charting the location of genes on chromosomes¹.

32
33 **Mass spectrometry:** a method used to determine the masses of atoms or molecules in which an
34 electrical charge is placed on the molecule and the resulting ions are separated by their mass to
35 charge ratio¹.

36
37 **Messenger RNA (mRNA):** a type of RNA that reflects the exact nucleotide sequence of the
38 genetically active DNA. mRNA carries the "message" of the DNA to the cytoplasm of cells
39 where protein is made in amino acid sequences specified by the mRNA¹.

40
41 **Metabonomics:** the evaluation of tissues and biological fluids for changes in metabolite levels
42 that result from toxicant-induced exposure¹.

43

1 **Microarray:** a tool used to sift through and analyze the information contained within a genome.
2 A microarray consists of different nucleic acid probes that are chemically attached to a substrate,
3 which can be a microchip, a glass slide, or a microsphere-sized bead¹.

4
5 **Northern blot:** a technique used to separate and identify pieces of RNA¹.

6
7 **Nucleotide:** a subunit of DNA or RNA. To form a DNA or RNA molecule, thousands of
8 nucleotides are joined in a long chain¹.

9
10 **“Omics”:** term including genomics, proteomics, metabonomics (some differentiate this term
11 from metabolomics), transcriptomics, and associated bioinformatics (Henry, C.J. et.al, 2002).

12
13 **Phenotype:** the observable physical or biochemical traits of an organism as determined by
14 genetics and the environment; the expression of a given trait based on phenotype; an individual
15 or group of organisms with a particular phenotype¹.

16
17 **Polymorphism:** the quality or character of occurring in several different forms¹.

18
19 **Proteome:** all of the proteins produced by a given species just as the genome is the totality of
20 the genetic information possessed by that species¹.

21
22 **Proteomics:** study of the full set of proteins encoded by a genome².

23
24 **Risk Assessment:** a qualitative or quantitative evaluation of the risk posed to human health and
25 the environment by the actual or potential presence of pollutants⁶.

26
27 **RNA:** a chemical found in the nucleus and cytoplasm of cells; it plays an important role in
28 protein synthesis and other chemical activities of the cell. The structure of RNA is similar to that
29 of DNA. There are several classes of RNA molecules, including messenger RNA, transfer RNA,
30 ribosomal RNA, and other small RNAs, each serving a different purpose².

31
32 **Signal transduction pathway:** the course by which a signal from outside a cell is converted to
33 a functional change within the cell¹.

34
35 **Single nucleotide polymorphism (SNP):** a change in which a single base in the DNA differs
36 from the usual base at that position¹.

37
38 **Susceptibility:** the increased likelihood of an adverse effect, often discussed in terms of
39 relationship to a factor that can be used to describe a human subpopulation (e.g. life stage,
40 demographic feature, or genetic characteristic)⁴.

41
42 **Susceptible Subgroups:** may refer to life stages, for example, children or the elderly, or to
43 other segments of the population, for example, asthmatics or the immune-compromised, but are

1 likely to be somewhat chemical-specific and may not be consistently defined in all cases⁴.

2
3 **Systems Biology:** a holistic approach to the study of biology with the objective of
4 simultaneously monitoring all biological processes operating as an integrated system (Sumner et
5 al., 2003).

6
7 **Systems Toxicology:** the study of perturbation of organisms by chemicals and stressors,
8 monitoring changes in molecular expression and conventional toxicological parameters, and
9 iteratively integrating biological response data to describe the functioning organism (Waters et
10 al, 2003).

11
12 **Throughput:** output or production, as of a computer program or a biological assay, over a
13 period of time¹.

14
15 **Toxicity:** deleterious or adverse biological effects elicited by a chemical, physical, or biological
16 agent⁴.

17
18 **Toxicology:** the study of harmful interactions between chemical, physical, or biological agents
19 and biological systems⁴.

20
21 **Toxicogenomics:** the study of how genomes respond to environmental stressors or toxicants.
22 Combines genome-wide mRNA expression profiling with protein expression patterns using
23 bioinformatics to understand the role of gene-environment interactions in disease and
24 dysfunction².

25
26 **Transgenic:** having genetic material (DNA) from another species¹.

27
28
29

Web-based Glossary Sources

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

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
3. 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>
4. U.S. Environmental Protection Agency. 1999. Glossary of Integrated Risk Information System. <http://www.epa.gov/iris/gloss8.htm>
5. 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
6. U.S. Environmental Protection Agency, Risk Assessment Web Page. <http://www.epa.gov/ebtpages/enviriskassessment.html>