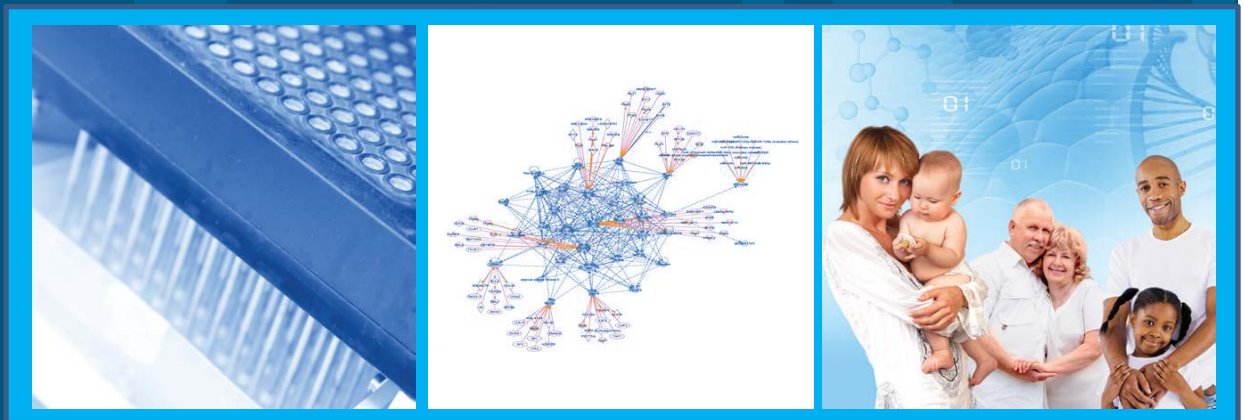
EPA/600/R-13/214A | September 2013 | www.epa.go/ncea

United States Environmental Protection Agency

Next Generation Risk Assessment:

Incorporation of Recent Advances in Molecular, Computational, and Systems Biology



External Review Draft

National Center for Environmental Assessment
Office of Research and Development, Washington, DC 20460

September 2013
EPA/600/R-13/214A

Next Generation Risk Assessment:

Incorporation of Recent Advances in Molecular, Computational, and Systems Biology

[External Review Draft]

National Center for Environmental Assessment

Office of Research and Development

U.S. Environmental Protection Agency

Washington, DC 20460

Disclaimer

This document is an external review draft. This information is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by EPA. It does not represent and should not be construed to represent any Agency determination or policy. It is being circulated for review of its technical accuracy and science policy implications. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Contents

Acronyms and Abbreviations

Acronym Abbreviation	Stands For
AC_{50}	concentration at 50% of maximum activity
AER	activity-to-exposure ratio
AhR	aryl hydrocarbon receptor
AML	acute myeloid leukemia
AOP	adverse outcome pathway
B[a]P	benzo[a]pyrene
BMD	benchmark dose
CDC	Centers for Disease Control and Prevention
C_{SS}	concentration, steady state (in blood)
CTD	Comparative Toxicogenomic Database
DEG	differentially expressed gene
DNA	deoxyribonucleic acid
EPA	U.S. Environmental Protection Agency
EWAS	environment-wide association studies
GEO	Gene Expression Omnibus
GWAS	genome-wide association studies
HCS	high-content screening
HPT	hypothalamus-pituitary-thyroid
HT	high-throughput
HTS	high-throughput screening
HTVMD	high-throughput virtual molecular docking
IC_{50}	concentration producing a 50% inhibition of response
IC_{10}	concentration producing a 10% inhibition of response
IVIVE	*in vitro* to *in vivo* extrapolation
KB	Knowledgebase
LEC	lowest effective concentration
MIE	molecular initiating event
MOA	mode of action
mRNA	messenger ribonucleic acid
NCEA	National Center for Environmental Assessment (EPA)
NexGen	Next Generation Risk Assessment
NHANES	National Health and Nutrition Examination Survey
NRC	National Research Council

Acronym Abbreviation	Stands For
NTP	National Toxicology Program
OECD	Organization of Economic Co-operation and Development
PAH	polycyclic aromatic hydrocarbon
PK	pharmacokinetic
POD	point of departure
ppb	part per billion
ppm	part per million
QSAR	quantitative structure-activity relationship
RNA	ribonucleic acid
ROS	reactive oxygen species
SNP	single nucleotide polymorphism
SOAR	Systematic Omics Analysis Review
Tox21	Toxicology in the 21st Century
VARIMED	VARiants Informing MEDicine
VT	virtual tissue

Acknowledgments

This document reflects contributions of many individuals whom we gratefully acknowledge. Without the generous contributions and collaborations of these scientists this report would not have been possible. Appendix A lists technical papers related to this report and provide an more complete list of contributors. We would also like to acknowledge Ms. Rebecca Clark, Dr. Kenneth Olden, and Ms. Debra Walsh for their continued support of this project.

Report Authors and Managing Editors

Ila Cote, Lyle Burgoon, Robert DeWoskin—EPA Office of Research and Development

Authors and Contributors

Executive Summary
Elaine Cohen Hubal, Rob DeWoskin, Lyle Burgoon, Ila Cote

Introduction, Preparation for Prototype Development and Consideration of Decision Context
Ila Cote, Paul Anastas, Stan Barone, Linda Birnbaum, Becki Clark, Kathleen Deener, David Dix, Stephen Edwards, and Peter Preuss

A Framework
Daniel Krewski, Margit Westphal, Greg Paoli, Maxine Croteau, Mustafa Al-Zoughool, Mel Andersen, Weihsueh Chiu, Lyle Burgoon, and Ila Cote

Science Community and Stakeholder Engagement
Kim Osborn, Gerald Poje, Ron White

Recurring Issues in Risk Assessment
Daniel Krewski, Melvin Andersen, Kim Boekelheide, Frederic Bois, Lyle Burgoon, Weihsueh Chiu, Michael DeVito, Hisham El-Masri, Lynn Flowers, Michael Goldsmith, Derek Knight, Thomas Knudsen, William Lefew, Greg Paoli, Edward Perkins, Ivan Rusyn, Cecilia Tan, Linda Teuschler, Russell Thomas, Maurice Whelan, Timothy Zacharewski, Lauren Zeise, and Ila Cote

The Prototypes

Major Scope Assessments- Benzene-Induced Leukemia
Ila Cote, Reuben Thomas, Alan Hubbard, Cliona McHale, Luoping Zhang, Stephen Rappaport, Qing Lan, Nathaniel Rothman, Jennifer Jinot, Babasaheb Sonawane, Martyn Smith and Kathryn Guyton

Major Scope Assessments - Ozone-induced Lung Inflammation and Injury
Robert Devlin, Kelly Duncan, James Crooks, David Miller, Lyle Burgoon, Michael Schmitt, Stephen Edwards, Shaun McCullough, and David Diaz-Sanchez

Major Scope Assessments Benzo[a]pyrene, Polycyclic Aromatic Hydrocarbons, and Cancer
Lyle Burgoon and Emma McConnell

Risk Assessment Implications across the Tier 3 Prototypes
Lyle Burgoon, Rob DeWoskin and Ila Cote

Tier 2: Limited Scope Assessments - Knowledge Mining – Diabetes/Obesity
Lyle Burgoon, Shannon Bell, Chirag Patel, Kristine Thayer, Stephen Edwards

Tier 2: Limited Scope Assessments Short-Term In Vivo Models – Alternative Species
Edward Perkins, Gerald Ankley, Stephanie Padilla, Dan Peterson, Daniel Villeneuve

Tier 2: Limited Scope Assessments Short-Term In Vivo Models – Mammalian Species
Michael DeVito, Jason Lambert, Scott Wesselkamper, Russell Thomas

Tier 1 QSAR and High-Throughput Virtual Molecular Docking (HTVMD)
Rob DeWoskin, Nina Wang, Jay Zhao, Scott Wesselkamper, Jason Lambert, Dan Petersen, Lyle Burgoon

Tier 1: High-Throughput and High-Content Assays, Toxicokinetics, High-Throughput Exposure Estimation, Virtual Tissue (VT) Modeling, Thyroid Example
Kevin Crofton and Richard Judson

Advanced Approaches to Issues in Risk Assessment

Human Variability including Genomic Variability
Lauren Zeise, Frederic Bois, Weihsueh Chiu, Ila Cote, Dale Hattis, Ivan Rusyn, Kathryn Guyton, and Lyle Burgoon

Early Life Exposures
Ila Cote, adapted from a paper by Boekelheide et al. 2012

Mixtures and Nonchemical Stressors
Timothy Zacharewski, Ila Cote, Linda Teuschler, and Lyle Burgoon

Inter-Species Extrapolation
Lyle Burgoon, Ila Cote, and Edward Perkins

Low Dose-Response Modeling
Weihsueh Chiu, Dan Krewski and Lyle Burgoon

Conclusions

Lessons Learned from Developing the Prototypes
Ila Cote, Rob DeWoskin, Lynn Flowers, John Vandenberg, Douglas-Crawford, Brown, Lyle Burgoon

Challenges and Next Steps
Elaine Cohn Hubel, Tina Bahadori, Ila Cote

Special Thanks to:

ICF International: Kim Osborn, Heather Dantzker, Jessica Wignall, William Mendez, Bruce Fowler, Codi Sharp, Deshira Wallace, and Pam Ross

Agency Partners who provided staff, data, ideas, and reviews

U.S. Army Corps of Engineers: Edward Perkins and Anita Meyer

U.S. Department of Defense, Office of the Secretary of Defense: Robert Boyd

California EPA: George Alexeeff, Martha Sandy, Lauren Zeise

Centers for Disease Control, National Center for Environmental Health, and the Agency for Toxic Substances and Disease Registry: Chris Portier (retired), Thomas Sinks, and Bruce Fowler (retired)

European Chemicals Agency: Derek Knight

European Joint Research Commission: Maurice Whelan

FDA's National Center Toxicological Research: Donna Mendrick and William Slikker

Health Canada: Carol Yauk

National Institutes of Health, National Center for Advancing Translational Science: Menghang Xia

NIH National Institute of Environmental Health Sciences: Linda Birnbaum, Scott Auerbach, John Balbus, Michael DeVito, Elizabeth Maull, Kristine Thayer, and Ray Tice

National Institute for Occupational Safety and Health: Christine Sofge, Paul Schulte, Ainsley Weston

EPA

Office of Research and Development: Michael Broder, David Bussard, Vincent Cogliano, Rory Conolly Dan Costa, Sally Darney, Elizabeth Erwin, Susan Euling. Annette Gatchett, Gary Hatch, Annie Jarabek, Robert Kavlock, Channa Keshava, Monica Linnenbrink, Matt Martin, Connie Meacham, Shaun McCullough. David Miller, Kenneth Olden, David Reif, James Samet, Rita Schoney, Imran Shah, Deborah Segal, Woodrow Setzer, Michael Slimak, Rong-Lin Wang, Debra Walsh, John Wambaugh, Paul White

Office of Air and Radiation: Souad Benromdhane, Bryan Hubbell, Kelly Rimer, Carl Mazza, Dierdre Murphy, Susan Stone and Lydia Wegman (retired)

- Office of Chemical Safety and Pollution Prevention: Stan Barone, Vicki Dellarco, Steven Knott, Mary Manibusan, Jennifer McLain, Jeff Morris, Anita Pease, Laura Parsons, and Jennifer Seed

Office of Superfund and Emergency Response: Michele Burgess, Rebecca Clark, Helen Dawson, Stiven Foster, and Kathleen Raffaele

Office of Water: Cynthia Dougherty, Elizabeth Doyle, and Elizabeth Southerland

Office of Children's Health Protection: Michael Firestone, Brenda Foos

Office of Environmental Justice: Charles Lee

Regional Liaisons: Carole Braverman, Bruce Duncan

Other Contributors

Ken Ramos, University of Louisville

Peter McClure, Heather Carlson-Lynch, and Julie Stickney, SRC

Catherine Blake, University of Illinois

William Pennie and Karen Leach, Pfizer

Executive Summary

The Next Generation (NexGen) of Risk Assessment program was initiated in 2010 as a multiyear, multi-organization effort to consider new molecular, computational, and systems biology approaches for use in risk assessments. The goal is to enable faster, less expensive, and more robust assessments for chemicals and other stressors that might adversely affect public health and the environment. Although this report is focused on human disease, the approaches described here are equally applicable to environmental risks. Specific aims of this initial phase of the NexGen program are to (1) demonstrate proof-of-concept that the data and methods from recent advances in biology can better inform risk assessment; (2) identify which of the information resources and practices are most useful for particular purposes (value of information); (3) articulate decision considerations for use of new types of data and methods to inform risk assessment; and (4) identify important data gaps.

To achieve the above, prototypes or case studies were designed to (1) implement the recommendations from workshops and experts on approaches to identifying and evaluating the available data in molecular, computational, and systems biology for use in risk assessment; (2) provide risk assessors, risk managers, and the general public with clear examples demonstrating how new data and advanced methods might support specific types of risk assessments; and (3) elicit interest, discussion, and participation from stakeholders to further improve risk assessments.

The assessment prototypes are broadly categorized into three groups based on the assessment's "fitness for intended use" given the decision context. Primary drivers of the decision context are the number of chemicals that must be addressed and the confidence needed in the scientific data to support a specific type of decision. The three categories or tiers have been defined as follows: Tier 3—major scope decision-making (considerable data indicating high hazard or widespread exposures); Tier 2—limited decision-making (limited exposure potential or limited hazard potential or data); and Tier 1—prioritization and screening (very little or no traditional data for chemicals known to be in commerce). Although the prototypes were designed for illustrative purposes to address these three types of decision context, the supporting data and methods can be deployed across all categories as available and as needed, and are arrayed as a continuum of approaches. Ideally, multiple data streams are brought to bear on consideration of potential risks.

The prototypes illustrate types of data and methods that are likely to be used in the near future, but are not intended to be exhaustive reviews. The primary intent of the first set of chemicals (Tier 3 prototypes) is to verify if and how new data types and approaches could be used to inform risk assessment by comparison to robust traditional "known" risk, thus verifying new approaches. The intent of the Tier 2 prototypes is to (1) explore new types of computational analyses and short-duration *in vivo* bioassays that are relatively uncommon in risk assessment but appear very promising for the near future; and (2) develop an assessment approach well suited to limited scope risk management decisions. In this case, limited generally means regional to local exposure potential, or limited hazard potential, or limited data to conduct more detailed assessments. The Tier 2 efforts fall between Tier 3 and Tier 1 in terms of resources required and amount of uncertainty in the assessment results. The Tier 1 prototypes explore entirely high-throughput approaches that could be applied to thousands or tens of thousands of chemicals, are the least resource intensive, and are likely to have the greatest uncertainty.

The following eight chemicals or chemical classes and their associated effects were chosen for prototype development:

- Tier 3:
  - Benzene and leukemia (molecular epidemiology),
  - Ozone and lung inflammation and injury (molecular clinical studies), and
  - Benzo[a]pyrene (B[a]P)/polycyclic aromatic hydrocarbons (PAHs) and liver cancer (molecular clinical studies meta-analyses and *in vivo* rodent bioassay).
- Tier 2:
  - Chemicals associated with diabetes and obesity ("big data" knowledge mining),
  - Chemicals associated with thyroid hormone disruption (short duration *in vivo* exposure bioassays-alternative species), and
  - Chemicals associated with cancer (short duration *in vivo* exposure bioassays-mammalian).
- Tier 1:
  - Chemicals associated with cancer and noncancer disorders especially developmental (QSAR) and
  - Chemicals associated with thyroid hormone disruption (high throughput *in vitro* assays).

Highlighted methods include molecular clinical and epidemiologic studies, *in vivo* molecular nonhuman studies, high-content *in vivo* assays (mammalian and nonmammalian species), bioinformatics, data mining, high-throughput *in vitro* screening assays, and quantitative structure activity modeling. NexGen methods and results were compared to robust traditional data set results.

Both bottom-up and top-down perspectives were used to evaluate the available data. The top-down approach focuses on higher system-level indicators of disease resulting from environmental exposures to known chemicals based on data from human clinical and epidemiologic studies. The bottom-up approach focuses on information describing chemically induced alterations in molecular and cellular components, as well as their network interactions. These data support the capability to develop risk assessments for chemicals with little or no traditional data. Additionally, these data can further inform assessments based on traditional data.

Data and insights from both bottom-up and top-down approaches are integrated to inform understanding of potential health risks associated with chemical exposures. The following summarizes lessons learned from development of the prototypes, as well as challenges for incorporating novel data streams to inform risk assessment:

- Advances in genomics, epigenomics, transcriptomics, metabolomics, and cell and systems biology, together with advanced analytical methods in biostatistics, bioinformatics, and computational biology, have the potential to increase dramatically understanding of the molecular basis of disease and environmental factors that alter disease risks.
- Of particular importance are the many new tools that facilitate testing and evaluation, on an unprecedented scale, of chemicals with limited or no traditional data. ToxCast™ and Toxicology in the 21st Century (Tox21) Programs provide examples.

- New approaches can be used to identify biological patterns or signatures that are associated with specific diseases, thus facilitating grouping and evaluating chemicals based on the mechanistic underpinnings of specific diseases. The Comparative Toxicogenomic Database provides a partial example.
- These signatures are best developed and understood as they relate to apical outcomes using systems biology. Conceptualization of these relationships among early molecular events, intermediate events, and apical outcomes are often termed mode of action or "adverse outcome pathways."[1]
- Signatures appear exposure-dose dependent (i.e., the magnitude of response changes with changes in exposure-dose) and hence, might be used to prioritize chemicals based on relative potencies, to serve as biomarkers of exposure and effect, and to inform quantitative risk assessment. Biological processes also are often time-dependent, which can complicate interpretation.
- The links between molecular perturbations and disease outcomes are influenced by a number of variables, that is, metabolism, cell type, genomic variants, cell and tissue interactions, and species. Thus, some test systems might better predict the potency of a chemical to disrupt normal biology than predict the specific adverse outcome resulting from that disruption.
- Historically, many controversial risk assessment issues lack data for substantive progress in understanding. NexGen approaches can provide new data types to improve the characterization of human variability and susceptibility, cross-species relevance, and low exposure-dose-response relationships via understanding mechanistic commonalities and differences. These issues are discussed in this report.

The prototype results presented in this report demonstrate proof-of-concept for an integrated approach to risk assessment based on molecular, computational, and systems biology. In addition, they explore which types of information appear most valuable for specific purposes and articulate some decision considerations for use of data. Based on lessons learned from this effort, near-term and longer term implications for risk assessment are also discussed.

Further advances in methods and knowledge undoubtedly will occur over the near term. Logistical and methodological challenges in interpreting and using newer data and methods in risk assessment, however, remain significant. Hence, incorporating new information into risk assessment will remain an ongoing opportunity.

[1]An adverse outcome pathway has been defined as the mechanistic or predictive relationship between initial chemical-biological interactions (i.e., molecular initiating event[s]; [MIE]) and subsequent perturbations to cellular functions sufficient to elicit disruptions at higher levels of organization, culminating in an adverse phenotypic outcome in an individual and population relevant to risk assessment (e.g., disease progression or organ dysfunction in humans) (Ankley, G. T. et al. 2010). Although commonly used, the term is something of a misnomer; pathways are not intrinsically adverse or nonadverse but rather pathways when perturbed in specific ways can lead to adverse outcomes. The same can be said of the commonly use term "toxicity" pathways.

[page intentionally left blank]

1. Introduction

Box 1. Next Generation Risk Assessment (NexGen)

This report describes the NexGen program, a multiyear, multi-organization effort to develop and evaluate new molecular, computational, and systems biology informed approaches to risk assessment. The goal of this effort is to advance risk assessment by facilitating faster, less expensive, and more robust assessments of public health risks by EPA's Office of Research and Development. The specific aims of the program are to:

- demonstrate proof of concept that recent advances in biology can better inform risk assessment;
- understand what information is most useful for particular purposes (value of information);
- articulate decision considerations for use of new types of data and methods to inform risk assessment; and
- identify important data gaps.

In recent years, public concern has grown about the number of chemicals in the environment and the ability to assess the risk to human health from potential exposures. Efforts by government agencies, including the U.S. Environmental Protection Agency (EPA), to protect public health from unreasonable chemical exposure have been hindered by limitations in current chemical testing methods and data. The European Commission has underscored this concern with recent initiatives to identify the many thousands of largely untested chemicals in use today and to increase the available toxicity information on those chemicals relative to the amount of chemical used and the potential for exposure in the environment (ECHA 2013a). As a result, significant efforts are underway throughout the world to redesign toxicity testing and understand how advances in biology, biotechnology, and computational science during the past two decades can be used in risk assessment. Specific goals are to increase dramatically our ability to test and assess chemicals more rapidly, understand disease processes and relationships to environmental factors, and facilitate the process from data acquisition to data analysis.

The technologies that have emerged from the sequencing of the human genome have ushered in a new era in biology (Collins, FS 2010) that supports the above goals. Advances in genomics, epigenomics, transcriptomics, metabolomics, proteomics, and cell and systems biology,[2] together with advanced analytical methods in biostatistics, bioinformatics, and computational biology, have dramatically increased our understanding of the molecular basis of disease—what causes disease and what exacerbates and ameliorates our risk of disease. Molecular signatures and other biomarkers are helping identify and define disease states and responses, and thousands of variations in previously unknown human health risk factors are being identified.

Researchers are generating massive amounts of biological data from the new "omics" technologies. Approximately 1.8 zettabytes (10^{21}) of new data are generated every year, roughly doubling the world's information resource every 2 years (Dearry 2013). More than 50,000 "genomics" papers are published each year (NCBI 2013). Large, publicly available data sets now support analyses of environmental health data on an unprecedented scale, driving further discovery of new knowledge (Dearry 2013, Abecasis et al. 2012, ENCODE Project Consortium 2012, Mechanic et al. 2012, Wang, I et al. 2012, Collins, MA 2009, Thomas, RS et al. 2009, Ramasamy et al. 2008). Concomitantly, powerful data mining, statistical, and bioinformatics methods have been developed to identify,

[2]**Systems biology** is defined as a "scientific approach that combines the principles of engineering, mathematics, physics, and computer science with extensive experimental data to develop a quantitative as well as a deep conceptual understanding of biological phenomena, permitting prediction and accurate simulation of complex (emergent) biological behaviors" (Wanjek 2013). See Wanjek's (2013) Web article *Systems as Biology as Defined by NIH* for more discussion of systems biology.

prioritize, and classify biomarkers with high discriminatory ability (Fang et al. 2012), and to store and manage the information in database libraries, including the Gene Expression Omnibus (GEO) (NCBI 2012a), the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa Laboratories 2013), the Comparative Toxicogenomic Database (CTD) (NIEHS 2013), and the Epigenomics Database (Chadwick 2012, NCBI 2009). As integration across differing types of data and levels of biological organization occurs (Birney 2012, ENCODE Project Consortium 2012), the degrees to which environmental risk assessment will be transformed and our understanding of disease at the individual and population level will be advanced are anticipated to be significant (Bhattacharya et al. 2011, Chiu et al. 2010).

Scientific discovery is now moving away from the traditional approach of individual scientists' conducting experiments in their laboratories to pooling of data into publicly available databases and broad collaborative participation in problem solving (Dearry 2013, Friend 2013, Derry et al. 2012). The magnitude of changes was highlighted in remarks by Frances Collins (Director of National Institutes of Health [NIH]) who stated that within the near future, most people in the United States will have a genome scan in their medical records as a tool for diagnosis, prognosis, and treatment of disease (Collins, FS 2010).

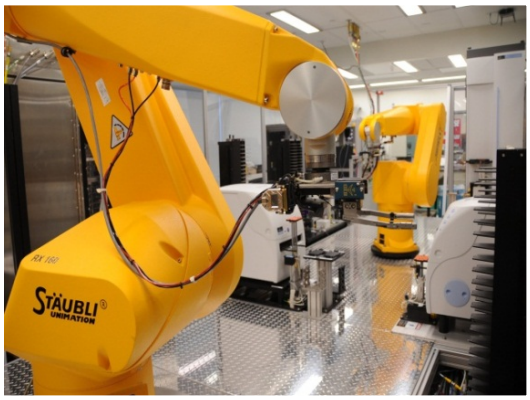


Figure 1. Toxicology in the 21st Century (Tox21) robot conducts bioassays on 10,000 chemicals. A robot arm (foreground) retrieves assay plates from incubators and places them at compound transfer stations or hands them off to another robot arm (background) that services liquid dispensers or plate readers. Photo by Maggie Bartlett (NHGRI 2012).

The impact of recent scientific advances on our ability to conduct risk assessments and protect public health cannot be overestimated. Particularly relevant to environmental risk assessment is that new data types and methods will result in much more rapid evaluation of chemicals, increase identification of causal mechanisms of disease, and provide a more profound understanding of the interrelated roles of genetics, epigenetics, and environmental factors. Experiments already can be conducted much more rapidly and efficiently using robotics and *in vitro* assays to measure molecular functions. Two examples are (1) Toxicology in the 21st Century (Tox21) testing of 10,000 chemicals within 3 years using approximately 150 assays (Figure 1) (Tice et al. 2013), and (2) the study of gene- and environment-wide associations with disease in tens of thousands of humans with multiple diseases—both unimaginable feats 15 years ago (Friend 2013, Mechanic et al. 2012). With the burgeoning amounts of data produced by high-volume testing and discovery, an effort in the European Union called "Safety Evaluation Ultimately Replacing Animal Testing," SEURAT-1 (http://www.seurat-1.eu/) has begun to develop a conceptual framework that can be used as a basis to combine information derived from predictive tools to support a safety assessment process. The overarching SEURAT-1 research strategy is to adopt a

toxicological mode-of-action[3] approach to describe how any substance might adversely affect human health, and to use this knowledge to develop complementary theoretical, computational, and experimental (*in vitro*) models.

In collaboration with its partners (see Acknowledgments), EPA initiated the NexGen program in 2010 to evaluate the use of these recent advances in biological and computational sciences for risk assessment (Text Box 1). We initially conducted workshops and solicited expert opinion to develop a framework and suggestions for prototype assessments that address the needs of the public and the risk assessment community. Federal, state, and other partners participated in the workshops and continue to provide advice, data, and review for NexGen reports. Text Box 2 lists related, ongoing legislation and government research activities in Europe and the United States.

[3]"Mode-of-action" is one term used to reference a mechanistic understanding of the impact of a chemical on human health. Other terms include "disease signature" and "network perturbations" from epidemiology for example, while toxicologists might reference the same concept using the terms "toxicity pathway," "mode-of-action," or "adverse outcome pathway." In general, this report uses the term "mechanism of action," in accordance with the National Research Council (NRC) report, *Science and Decisions: Advancing Risk Assessment* (2009); however, the exact term used in a specific section of this report is based on the references used and the context of the discussion.

Box 2. Current Legislation and Governmental Research Activities in Europe and the United States

EUROPEAN LEGISLATION AND ACTIVITIES

In response to environmental concerns, a desire for increased assessment efficiencies, and a desire to reduce reliance on *in vivo* animal testing, the European Union (EU) enacted an expansive new program called Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) in June 2007. This legislation places greater responsibility on industry to test and manage the risks posed by their chemicals. Under REACH, companies must develop detailed technical dossiers and chemical safety reports and submit them to the European Chemicals Agency (ECHA). Approximately 12,000 chemicals have been registered for consideration with ECHA. Many more chemicals are anticipated in the near future. Additionally, the 7th Amendment to the EU Cosmetics Directive prohibits putting animal-tested cosmetics on the market in Europe after 2013. Although current alternative methods more closely resemble traditional methods, the EU has invested 50M Euros in a research program to further next-generation methods (OECD 2012). Current ECHA guidance is available on the use of quantitative structure-activity relationships (QSARs), *in vitro* assays, and read-across (also known as near-analog structure-activity relationships) to support assessments.

REACH and the 7th amendment will significantly impact nearly all multinational companies and are important drivers for the development and use of new molecular-based methodologies. Europe's chemical trade accounts for about 40% of the global market, involving 27 countries and almost half a billion people.

The Joint Research Centre (JRC) is the scientific and technical arm of the European Commission. It provides scientific advice and technical support to EU policies. The JRC has seven scientific institutes (featuring laboratories and research facilities) located at five sites: Belgium, Germany, Italy, the Netherlands, and Spain. The JRC's Institute for Health and Consumer Protection's main research relevant to NexGen includes integrated risk and benefit assessments of chemical substances; fit-for-purpose analytical tools to help ensure the safety of food and consumer products; and optimization and validation of methods that reduce the reliance on animal tests in the safety assessment of chemicals.

U.S. ACTIVITIES

Several documents have guided the NexGen effort, including the Strategic Plan for the Future of Toxicity Testing and Risk Assessment at the U.S. Environmental Protection Agency (EPA 2009a), the Toxicology in the 21st Century (Tox21) strategy, and the National Institutes of Health Strategic Plan (NIEHS 2012c). Ongoing research activities of several federal agencies that have informed and continue to inform the NexGen effort are described below.

The Centers for Disease Control and Prevention (CDC) has several groups involved in systems biology and computational environmental health and occupational research. The **National Center for Environmental Health (NCEH)** and **Agency for Toxic Substances and Disease Registry (ATSDR)** scientists in the Computational Toxicology Laboratory have applied several new approaches for improving chemical risk assessments. They have mined the National Health and Nutrition Examination Survey (NHANES) data set to obtain high-quality analytical and human health information, which is representative of the general U.S. population, and used computer modeling to identify sensitive populations for health outcomes at environmental exposure levels. A second project involved use of NHANES public health genomics data to identify allelic differences in ALA dehydratase for susceptibility to lead-induced hypertension. Another concerned the development and application of QSAR, physiologically based pharmacokinetic (PBPK), and molecular docking approaches. These studies involved both data mining of the published scientific literature and collaborative laboratory studies with scientists at the Food and Drug Administration (FDA).

The National Institute for Occupational Safety and Health (NIOSH) is investigating susceptibility gene variants that contribute to the development and severity of occupational diseases using high-density and high-throughput (HT) genotyping platforms. Understanding the genetic contribution to the development, progression, and outcomes of complex occupational diseases will help improve the accuracy of risk assessment and improve safe exposure levels for genetically susceptible groups in the workforce.

The FDA National Center for Toxicological Research (NCTR) is conducting translational research to develop a scientifically sound basis for regulatory decisions and reduce risks associated with FDA-regulated products. NCTR research evaluates biological effects of potentially toxic chemicals, defines the complex mechanisms that govern their toxicity, identifies the critical biological events in the expression of toxicity, discovers biomarkers, and develops new scientific tools and methods to improve assessment of human exposure, susceptibility, and risk. Examples of tools created by NCTR include ArrayTrack™, Decision Forest, Endocrine Disruptor Knowledge Base (EDKB), Gene Ontology for Functional Analysis (GOFFA), and SNPTrack. Efforts include the MicroArray Quality Control (MAQC) consortia.

Box 2. Current Legislation and Governmental Research Activities in Europe and the United States (Continued)

U.S. ACTIVITIES (CONTINUED)

The National Institutes of Health (NIH) National Center for Advancing Translational Sciences (NCATS) conducts research to resolve scientific and technical challenges that might cause barriers to the efficient development of new treatments and tests to improve human health. The National Chemical Genomics Center (NCGC) at the National Center for Advancing Translational Sciences applies high-throughput screening (HTS) assay guidance, informatics, and chemistry resources for NCAT's Re-engineering Translational Sciences research projects. Specifically, NCGC research programs include assay development and HTS, and participation in Tox21. NCGC Assay Biology Teams are researching optimization of biochemical, cellular, and model organism-based assays submitted by the biomedical research community for HT small molecule screening. The results of these screens (probes) can be used to further examine protein and cell functions and biological processes relevant to physiology and disease (NIH 2012).

The National Human Genome Research Institute (NHGRI) was established by NIH in 1989 to implement the International Human Genome Project to map the human genome. NHGRI has developed programs for a variety of research projects including Encyclopedia of DNA Elements (ENCODE), Gene Expression Omnibus (GEO), and collaborative projects, including the Comparative Toxicogenomic Database (CTD), HapMap, and Gene. Through the application of these tools, NHGRI hopes to gain a greater understanding of human genetic disease, and develop better methods for the detection, prevention, and treatment of genetic disorders.

The National Institute of Environmental Health Science (NIEHS) and the National Toxicology Program (NTP) have played an integral role in the development and application of HTS data. Current research is focused on developing and validating Tox21 approaches to improve hazard identification, characterization, and risk assessment (Birnbaum 2012, Serafimova et al. 2007). The NTP HTS program has three specific goals: (1) prioritizing substances for in-depth toxicological evaluation, (2) identifying mechanisms of action for further investigation (e.g., disease-associated pathways), and (3) developing predictive models for *in vivo* biological response (i.e., predictive toxicology). NTP is developing innovative and flexible approaches to data integration, both across research programs and across different data types (e.g., HT, mechanistic, animal studies) (Bucher et al. 2011). These efforts seek to integrate results from new techniques with traditional toxicology data to provide a public health context.

The Engineer Research and Development Center (ERDC), the research organization of the U.S. Army Corps of Engineers, conducts research and development in support of warfighters, military installations, and civil works projects involving water resources and environmental missions. The ERDC Toxicogenomics research cluster focuses on using genomics to develop tools to rapidly assess toxicity of military chemicals in a wide range of animals, identifying gene biomarkers of exposure, understanding the mechanisms by which military chemicals cause toxicity, and extrapolating toxicity effects across multiple species. Capabilities of the team include advanced instrumentation to characterize impacts of chemicals on gene expression with high-density gene arrays, DNA sequencing, and real-time polymerase chain reaction (RT-PCR) assays. ERDC Toxicogenomic projects include development of rapid assays to assess whole genome impacts of munitions-related compounds, including gene arrays with short exposure screening in daphnia, rat cells, rat livers, and fish; comparison of genomic and behavioral responses of fathead minnows and zebrafish to chemical exposures; conservation of response to nitroaromatics across species; and support for a toxicogenomic assessment framework to integrate predictive toxicology of munitions-related compounds.

Several EPA Office of Research and Development laboratories and centers have been involved in NexGen. EPA's **National Center for Environmental Assessment (NCEA)** has assumed a leadership and coordination role for the NexGen effort. The **National Center for Computational Toxicology (NCCT)** is the largest component of EPA's Computational Toxicology Research Program. The Center coordinates computational toxicology research on chemical screening and prioritization, informatics, and systems modeling. NCCT research includes the (1) use of informatics, HTS technologies, and systems biology to develop accurate and flexible computational tools that can screen the thousands of chemicals for potential toxicity; and (2) application of mathematical and advanced computer models to help assess chemical hazards and risks. EPA's **National Center for Environmental Research (NCER)** supports extramural computational toxicology research. The **National Health and Environmental Effects Research Laboratory (NHEERL)** conducts toxicological, clinical, and epidemiological research to improve the process of human health risk assessments, including development of biological assays and toxicological assessment methods, predictive pharmacokinetic/pharmacodynamic models, and advanced extrapolation methods.

The initial NexGen prototypes were designed to provide concrete examples that illustrate the potential for various new methods and data to be used for specific risk assessments within a decision context[4] and to foster further discussion in the risk assessment and risk management communities to promote continual improvement.

This report presents and discusses the results of this effort, and is organized as follows:

- Section 2: Preparation for Prototype Development – describes the preliminary work and workshops conducted to characterize the decision context and conceptual framework and to identify the stakeholders and key issues so that the prototypes provide examples relevant to the needs of the risk assessment community.
- Section 3: The Prototypes – provides detailed examples of the use of various advanced methods and data in each of the three tiers, starting with chemicals for Tier 3 "Major Scope Assessments," which are data-rich chemicals, proceeding to Tier 2 and Tier 1 chemicals that have increasingly limited or no *in vivo* data sufficient to conduct a traditional (e.g., IRIS) risk assessment.
- Section 4: Advanced Approaches to Recurring Issues in Risk Assessment – describes how advanced methods are being used to address recurring and challenging issues, including characterizing variability in deriving toxicity values and assessing potential hazards from exposure to mixtures.
- Section 5: Lessons Learnedfrom Developing the Protypes – describes lessons learned in developing the Tier 1, 2, and 3 prototype examples.
- Section 6: Conclusions– outlines the major challenges and future direction for the NexGen program.
- Appendix A lists technical papers supporting this report.
- Appendix B provides a glossary.

2. Preparation for Prototype Development

2.1. Consideration of Decision Context

One of the first tasks undertaken in planning the NexGen effort was consideration of the various environmental situations of concern to EPA's Program Offices—in other words, the decision context [termed in Cote et al. (2012) and the National Research Council (NRC 2009) and National Academy of Science (NAS 2007) reports]. Decision context defines what environmental management decision is being made and why, as well as its relationship to other decisions previously made or anticipated. EPA Program Offices are generally organized around specific pieces of environmental legislation, such as the Clean Air Act and the Clean Water Act, and are responsible for administering those laws. Each major piece of legislation brings different responsibilities and nuances to problems faced by risk managers. In Figure 2, the decision context is represented in three categories for ease of description. The characteristics that define the three decision context categories and examples of specific problems faced by the Program Offices are shown. This figure elaborates on the decision

[4]See Section 2.1 for a definition of "decision context" and a discussion of its use.

context figure in the report, *Science and Decisions: Advancing Risk Assessment,* and is the result of discussions with EPA Program Offices (EPA 2011b).

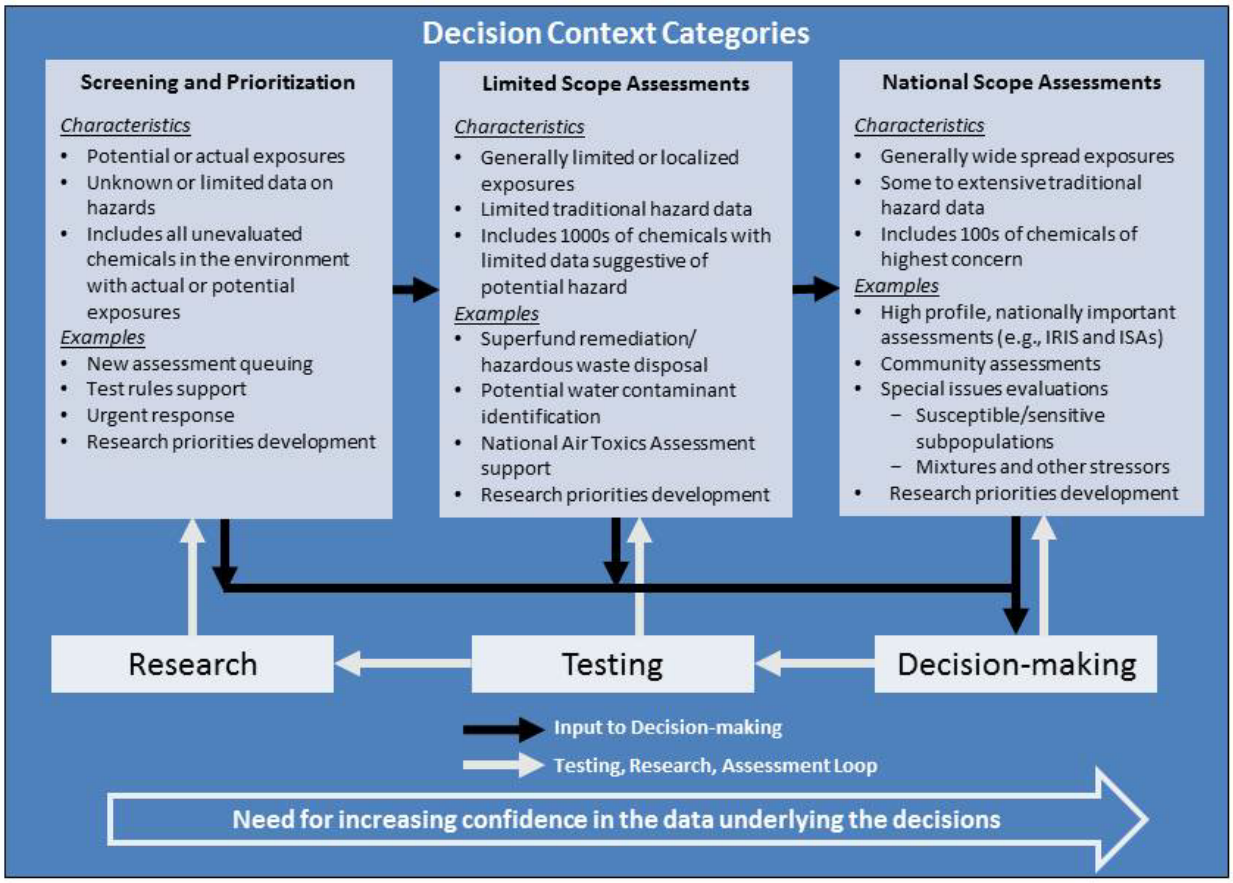


Figure 2. Description of decision context categories provided by EPA Program Offices. These decision context categories reflect the range of environmental problems to be addressed, from the need to screen many untested chemicals in the environment to national regulations for high profile chemicals. The flow from decision context through risk assessment to decision-making and the related roles of testing and research are also noted.

Three factors integral to the decision context for risk managers are the potential exposure, the number of chemicals that should be considered, and the weight of scientific evidence for supporting decision-making. Both legislative language and the history of specific regulatory programs influence the numbers of chemicals considered and the uncertainty in supporting data that can be tolerated. Tier 3 decision context focuses on nationally relevant chemicals with widespread exposures and established hazards and for which major regulatory evaluations are likely in progress. An example would include the International Agency for Research on Cancer (IARC) benzene assessment (2012) where molecular mechanistic information was used to support the causal link between benzene and hematopoietic cancers, particularly when the epidemiology data were somewhat limited. Tier 2 focuses on chemicals for which exposure or hazard appears limited or available data for detailed assessment are limited. An example includes evaluation of biological activity and cumulative risk potential of conazole fungicides (EPA 2011e) and potential endocrine disruptors (EPA 2011c), both of which are based on molecular biology data in combination with traditional data. Tier 1 decision context focuses on the tens of thousands of chemicals present in

commerce in significant amounts, but for which we have little knowledge of exposure levels or potential health effects. An example is the high-throughput (HT)-based evaluations of Deep Water Horizon Gulf oil spill dispersants (Judson et al. 2010).

2.2. A Framework

A second task that preceded finalizing plans for the NexGen prototypes was the development of a guiding framework. The framework draws together several important elements of earlier risk assessment frameworks and articulates guiding principles for risk assessment development informed by new data types and methods. A draft version of this framework was presented and discussed in October 2010 at a meeting with scientific experts (EPA 2010) and in February 2011 at a public meeting with stakeholders (EPA 2011a). The framework is described in a report by Krewski et al. (2013).

The NexGen framework is built on three cornerstones (as illustrated in Figure 3): (1) new risk assessment methodologies to better inform risk management decision-making; (2) new data types from advances in biology and toxicology on understanding perturbations of biological pathways; and (3) a population health perspective that recognizes that most adverse health outcomes involve multiple determinants. The NexGen framework integrates these three cornerstones into a framework for risk science that progresses in three phases: (1) Objectives, (2) Risk Assessment, and (3) Risk Management. Phase 1 (Objectives) focuses on problem formulation and scoping, taking into account the decision context and the range of available or admissible risk management decision-making options. Phase 2 (Risk Assessment) seeks to identify disease or outcome pathways using new toxicity testing tools and technologies and attempts to improve the characterization of risks and uncertainties using advanced risk assessment methodologies. Phase 3 (Risk Management) involves the development of evidence-based population health risk management strategies of a regulatory, economic, advisory, community, or technological nature, based on sound principles of risk management decision-making. Implementation of the NexGen framework is exemplified with a series of case-study prototypes, illustrating how aspects of the framework have been put into practice.

NRC provided a blueprint for pathway-based toxicity testing in its 2007 report, *Toxicity Testing in the 21st Century: A Vision and a Strategy* (NRC 2007). Guidance on some of the new risk assessment methods is provided by the 2009 report, *Science and Decisions, Advancing Risk Assessment* [NRC (2009)]. The integration of a population health approach was drawn from the McLaughlin Centre's integrated risk management and population health framework. Key elements of risk science and population health are combined to offer a multidisciplinary approach to the assessment and management of health risk issues (Krewski et al. 2007).

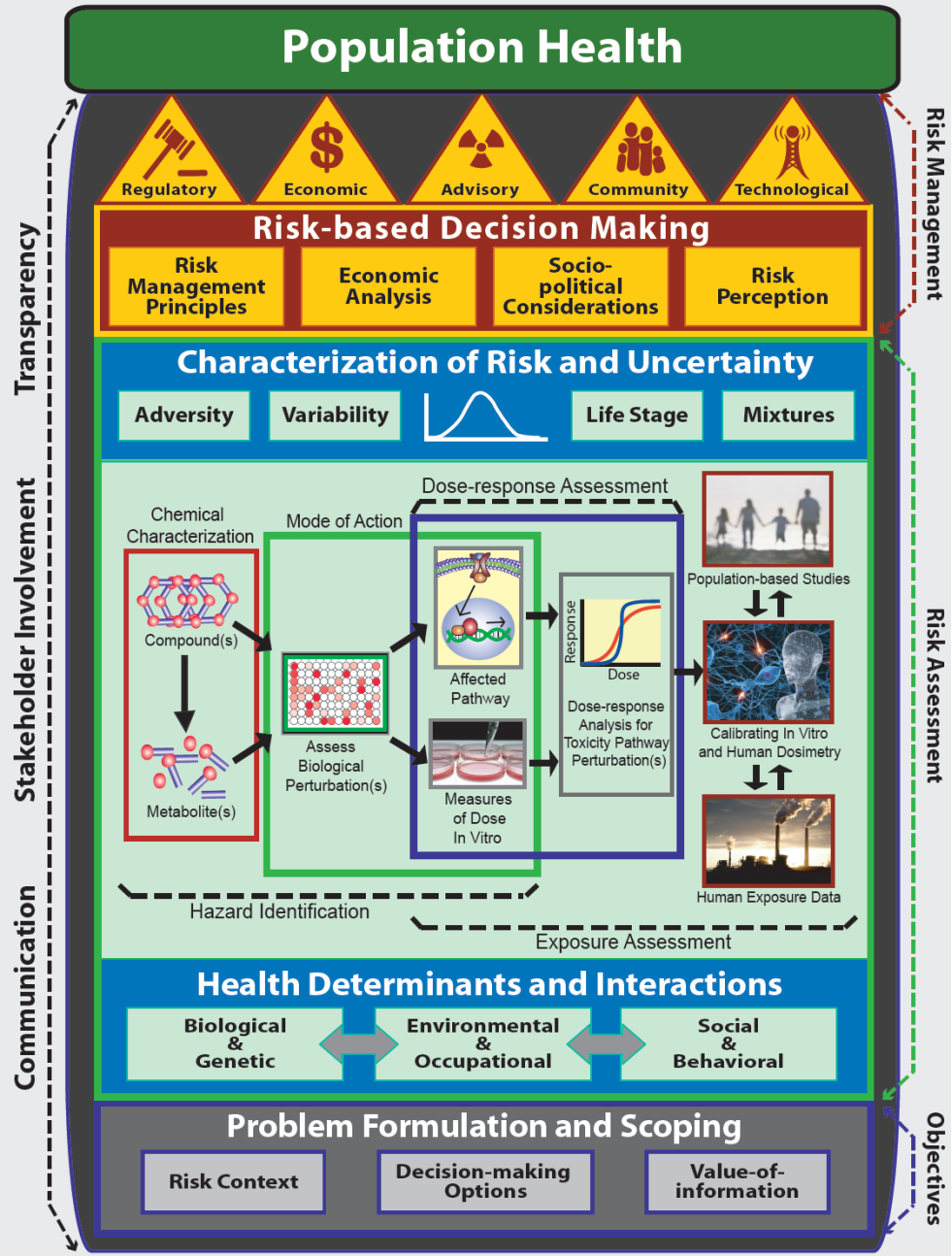


Figure 3. The Next Generation Framework for Risk Science. This framework is divided into three phases: (1) Objectives: Problem Formulation and Scoping takes into consideration Risk Context,[5] Decision-Making Options, and Value-of-Information; (2) Risk Assessment involves three sub-categories: (A) Health Determinants and Interactions, (B) New Scientific Tools and Technologies, and (C) New Risk Assessment Methodologies; and (3) Risk Management involves two categories: (A) Risk-Based Decision-Making that involves Risk Management Principles, Economic Analysis, Socio-Political Consideration, and Risk Perception, and (B) Risk Management Interventions with five possible categories: Regulatory, Economic, Advisory, Community, and Technical (Krewski et al. 2013).

[5]The term decision context is used elsewhere in this report.

2.3. Science Community and Stakeholder Engagement

Outreach to the science community and stakeholder groups was part of the NexGen strategy from its inception. This document in its final form is viewed as an interim step to implementation of new advances in risk assessment and is intended to promote further discussion with stakeholders toward continual improvement of risk assessments and prototypes informed by new data types and methods.

Given the technical complexity of the research, stakeholder engagement is a particular challenge and will necessitate ongoing outreach and discussion throughout the process. Our initial efforts are described below.

2.3.1. Expert Workshop

EPA convened a 3-day expert workshop on November 1–3, 2010, in Research Triangle Park, North Carolina, to discuss the draft framework, early draft prototypes, research, and other project elements. The workshop sought individual input, rather than consensus, in meeting its discussion goals. Days 1 and 2 of the workshop focused on deliberative drafts of data-rich prototype health assessments. The goals were to (1) refine health assessment case studies of data-rich chemicals informed by molecular biology (i.e., "prototypes"); (2) enhance "reverse engineering" from molecular system biology data, to "known" public health risk estimates based on *in vivo* human and animal bioassay data to demonstrate proof of concept, elucidate value of information, and characterize decision considerations; and (3) summarize options for expanded future work and research needs.

Day 3 focused on approaches applicable to assessing the potential risks posed by chemicals with limited or no traditional data. The goals were to (1) identify and discuss a wider variety of new data types, methods, and knowledge to help characterize data-limited chemicals; (2) consider how this information might augment, extend, or replace traditional data in health assessment; and (3) summarize options for expanded future work and research needs. Approximately 40 federal and nonfederal experts and 80 and partner organization staff members attended the workshop. A workshop report with the agenda and list of participants is available (EPA 2010).

In 2012, both the Science Advisory Board (SAB) and the Board of Scientific Counselors (BOSC) reviewed aspects of the NexGen program as part of their evaluations of EPA's computational toxicology research (SAB 2013, BOSC 2010). Both the SAB and BOSC commended the exceptional efforts of the Computational Toxicology Research Program to advance hazard/risk assessment and provided recommendations for the continued success of the program. The reviews emphasized further research on chemical exposure pathways resulting from human activity patterns (e.g., ExpoCast); engagement of the scientific community and stakeholders to foster future partnerships and promote information exchange; broader outreach for dissemination of scientific findings; gathering of user-feedback from the general public; improvements in data access through enhanced website navigation; and development of guidelines for data usage.

2.3.2. Stakeholder Involvement

Stakeholder Public Dialogue Conference

To engage stakeholders in the early stages of the NexGen program, EPA sponsored a public dialogue conference, "Advancing the Next Generation of Risk Assessment," on February 15 and 16, 2011, in Washington, DC. This conference presented stakeholders with an opportunity to learn about the NexGen program, and to provide their thoughts on the challenges the project faces and its path forward. Approximately 160 participants, representing 11 stakeholder groups (Figure 4), attended the conference. The conference report includes the agenda, list of participants, and recommendations of the group (EPA 2011a). In addition to this conference, "one-on-one" interviews (described below) were conducted with leaders of public-interest groups and the business community.

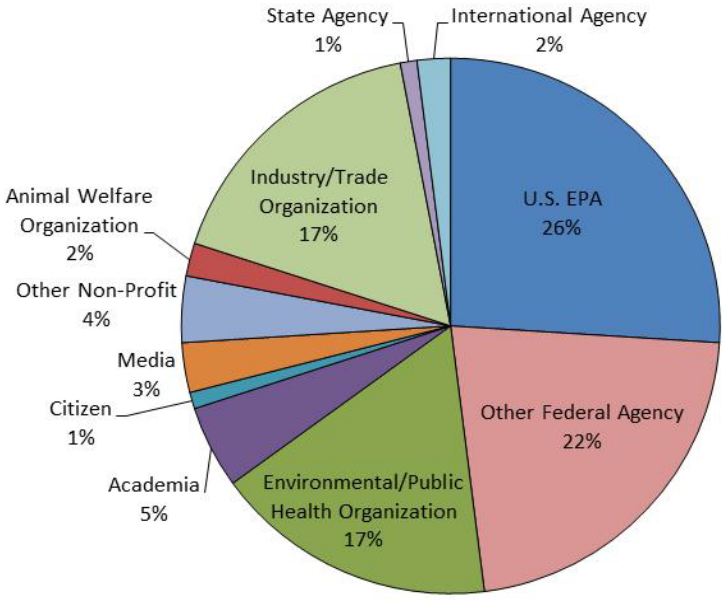


Figure 4. Categories of stakeholders that attended the February 2011 NexGen public dialogue conference (EPA 2011a).

Public Interest Group Perspectives

After the work shop, follow-up informal one-on-one interviews were conducted in mid-2010 with several Washington, DC-based representatives of national environmental, public health, and animal welfare public-interest organizations. Ronald White, a faculty member at Johns Hopkins Bloomberg School of Public Health, conducted these interviews and informational meetings as a component of his research on public engagement regarding emerging risk assessment methods. He also developed a web-based assessment in late 2010 to ascertain, from nongovernmental public-interest organizations, their knowledge and interest in emerging scientific approaches for chemical/pollutant risk assessment. Of the 24 organizations contacted, 8 (33%) responded to the assessment.

A key question raised in these interviews and web-based assessment was how relevant the NexGen program is to near-term EPA pollutant/chemical risk assessment procedures and control policies. The public-interest stakeholders interviewed and those who responded to the online assessment questions generally supported the concept of integrating the results from emerging biological science and analytic techniques into EPA's approach to conducting chemical health-based risk assessment. Significant concerns emerged, however, regarding the following:

- The potential for overstating the utility of NexGen approaches.
- How NexGen prototypes will address key risk assessment methodological issues, such as low-dose exposure assessment, population variability in response, and additivity to background exposures and disease processes.
- The transparency of the NexGen assessment development process and opportunities for early, meaningful engagement by public-interest organizations, and the application of NexGen approaches in risk management.

Business Community Perspectives

Industry or business perspectives on NexGen approaches also were of interest. Dr. Gerald Poje, (Environmental Health Consultant, Former Board Member of the U.S. Chemical Safety and Hazard Investigation Board) had follow-up discussions with nine individuals representing the high-production volume and smaller specialty chemical manufacturing industries, the pharmaceutical industry, the retail sector, and the energy sector. The participants were generally optimistic about potential advances in risk assessment and identified two potential advantages: (1) better prioritize the needs for more expensive and longer duration whole-animal testing, and (2) save time and money while rationalizing decisions in a tier-based manner using HT and other Tier 1 and Tier 2 tests. They also suggested that the success of NexGen effort depends on EPA's ability to prove the value of the newer, tiered approach within EPA's emerging risk assessment model, the level of EPA's investment in the long-term iterative NexGen research effort, and the timely and effective communication of the evidence to support science-based risk assessment.

Some in the business community expressed concern over whether EPA could match the pharmaceutical industry's growing infrastructure (needed to support and sustain a NexGen-like effort) such as EPA's ability to unite sufficient numbers of expert biologists, chemists, and bioinformatics to guide the program to a successful conclusion. The technical complexity of the NexGen program might also hinder its impact on current risk assessment, risk management, and business development practices, given the many unknowns that remain. Cultural challenges in winning over a larger community, who will welcome the use of more recent advances in risk assessment methods; however, was thought to be surmountable if EPA could be effective at capacity building and communicating how new data types and approaches could be used for risk assessment.

2.4. Recurring Issues in Risk Assessment

The fourth task that preceded the actual prototype development was identification of problematic issues that might be substantively informed by new methods and data. The issues included problem formation, adversity classifications and weight of evidence, dose-response modeling (especially at the low-dose end), variability in human response (due to a variety of factors), interspecies extrapolation, mixtures risk assessment, and characterization of uncertainty. These issues are explored in the prototypes to the extent feasible, and some are discussed in more detail in papers on human variability (Zeise et al. 2012), early-life exposure and later-life disease risks (Boekelheide et al. 2012), and multifactorial interactions of environment and genes (Patel et al. 2012a, Patel et al. 2012b, Zhuo et al. 2012, Shen et al. 2011, Smith, MT et al. 2011).

3. The Prototypes

EPA's Office of Research and Development, in conjunction with other federal, state, academic, public, and private partners (see Acknowledgments), developed prototype assessments to provide concrete illustrations of how new and emerging information could inform risk assessment. The prototypes used a variety of study types, methods, data, and risk assessment approaches, and are intended to (1) engender movement in the field of risk assessment from strategy to practical application of new approaches, and (2) foster discussion and refinement of approaches in the risk assessment and risk management communities, as well as with the public.

Box 3. Selection Criteria for Prototypes

- Decision context applicability (i.e., methods applicable to various types of risk management situations)
- Data availability (i.e., both NexGen and traditional data existed to allow for validation of new approaches)
- Illustration of a variety of methods
- Methods
  - ✓ Data quality
  - ✓ Multiple, high-quality studies
  - ✓ Consistent, coherent, and biologically plausible data
- Active collaborations with investigators to benefit from their knowledge, modify experiments, and conduct additional analyses as needed
- Cross-organizational collaborations fostered

The results presented in this report demonstrate proof of concept, provide insight on what types of information are valuable for specific purposes, and provide examples of the decision considerations for reasonable, consistent, and coherent use of the new types of information for specific applications. The prototypes also illustrate many of the challenges. Text Box 3 lists selection criteria used in choosing prototypes. Figure 5 broadly categorizes the types of methods aligned to decision context and evaluated in the prototypes. As noted earlier, the number of chemicals that need to be evaluated and the level of confidence required for decision-making are key components of designing fit-for-purpose assessments. The integration of knowledge from a wide variety of methods is likely to be most informative to risk assessment. Lessons learned from each prototype and group of similar prototypes will be noted as they arise and, then integrated and summarized in Section 5 "Lessons Learned from Developing the Prototypes."

Throughout this report, characterizing systems biology is greatly emphasized. Systems biology is a critical field in modern biology aimed at understanding the larger picture by integration across multiple levels of biology—for example, from the gene to the molecular intermediate phenotypes (e.g., gene expression), to alterations in molecular pathways and networks, and the propagation of effects from cells to tissues to organs and the whole body. Systems biology also can encompass subpopulation and population dynamics. Thoroughly understanding modern biology is difficult without understanding systems biology. Two basic approaches are used to develop systems understanding: bottom up and top down. The bottom-up approach focuses on altered molecular and cellular components, and seeks to understand how the altered components fit together. This approach is addressed most extensively in Tiers 1 and 2. The top-down approach focuses on larger scale network interactions and disease indicators based on human clinical and epidemiologic data, and associations between disease states and environmental factors (Friend 2013). This approach is addressed most extensively in Tiers 3 and 2. Both the bottom-up and top-down approaches can be informative, and are best used when integrated to support development of a comprehensive model.

Types of Methods and Characteristics		Tier 1: Screening and Prioritization	Tier 2: Limited Scope Assessments	Tier 3: Major Scope Assessments
Types of Methods	**Exposure Data:**	Surrogate	Limited	Extensive environmental
	Structure Information:	QSAR Models	QSAR Models and Read Across	Mechanistic Understanding
	Assay Types:	High-throughput (HT) Assays	High Content	All informative Traditional Data
	Extra Information Sources:	Computer (*In silico*) Toxicity Models	Database Mining	Biomarkers of Exposure and Effect
Characteristics	**Time to Conduct:**	Hours–Days	Hours–Weeks	Days–Years
	Cost:	$	$–$$	$$–$$$
	Exposures:	*In Vitro*	*In Vitro* and *In Vivo*	*In Vivo*
	Exposure Duration:	Short	Short	Longer
	Metabolism:	Little to None	Some	Substantial
	Endpoints:	Alterations in Key Biological Process	Alterations in Key Biological Process to Adverse Effects	Alterations in Key Biological Process, Intermediate Event to Adverse Effects/Disease
Increasing weight of evidence and resources to complete				

Figure 5. Shown are the types of methods used to generate data for the prototypes and characteristics of each method. Note that all methods can be used in each decision context as available.

3.1. Tier 3: Major Scope Assessments

Tier 3 prototypes focused on chemicals with robust traditional data sets, known public health outcomes, and high-confidence risk estimates. The purpose of studying these already well-characterized chemicals was to better understand how new data types and methods can be most effectively used in risk assessment situations where traditional data are absent or limited. In other words, by "reverse engineering" from known public health risks to new types of data, it was thought that potential advances in risk assessment using new types of data could be verified. Molecular epidemiology, molecular clinical, and molecular *in vivo* animal data were evaluated in the context of traditional information (Table 1). The Tier 3 prototypes aimed to: (1) demonstrate proof of concept that new data and methods can help identify hazards and inform exposure-dose-response relationships; (2) better characterize what information is most valuable for specific risk assessment purposes; and (3) articulate decision considerations for identifying, analyzing, and interpreting data, particularly for use in assessment of data-poor chemicals. Secondarily, this effort explored how new data types can augment robust traditional data sets, and brings new insights to the interpretation of traditional data.

Table 1. Tier 3 Prototypes Approach, Including Weight of Evidence, Pros, and Cons

Tier 3: Major Scope Assessments Environmentally-relevant *In Vivo* Exposure Studies with Molecular Characterization	
Approaches:	Focuses on human data from molecular epidemiology and molecular clinical studies. Includes molecularly augmented, traditional *in vivo* animal bioassays. Experimentally measures dose-dependent, chemically-induced alterations in biologic functions linked to traditional intermediate events and disease outcomes. Evaluates environmentally-relevant exposures. Characterizes sensitive subpopulations. Helps characterize impacts of various environmental factors.
Weight of evidence:	Determined by the quality and quantity of data, but can range from suggestive to known.
Pros:	Characterizes human population-associated or causal mechanisms. Can inform low-dose, species-to-species and inter-individual variability, and uncertainty with data. Allows extrapolation of molecular patterns to predict outcomes for less well studied chemicals.
Cons:	Are not faster or less expensive than traditional bioassays. Need to control for experimental variability.

The Tier 3 prototypes are benzene (and leukemia); ozone (and inflammation and lung injury); and benzo[a]pyrene (B[a]P), a polycyclic aromatic hydrocarbon (PAH) (and liver cancer). The prototypes focused on human data, both molecular epidemiology and molecular clinical data. Human environmental exposures for the benzene and ozone prototypes were very well characterized using a urinary biomarker and $^{18}O_2$ dosimetry, respectively. For B[a]P, we evaluated human environmental exposures and liver cancer omics data; this evaluation was qualitatively successful, but exposures were relatively poorly characterized for quantitative exposure-response assessment. Hence, experimental rodent data were evaluated in addition to the human data.

Overall, the prototypes evaluated the use of toxicogenomics to better characterize risks, including DNA transcription (transcriptomics), protein expression (proteomics), and genome-wide analyses of susceptibility genes (genomics analyses of human gene variants). Some limited discussion of epigenetic modification (epigenomics) in human populations is also included. Bioinformatics analyses were used to evaluate toxicogenomic profiles in the context of traditional knowledge of phenotypic endpoints. Each prototype:

- Describes a systems biology model suitable for informing hazard identification;
- Characterizes molecular biomarkers of exposure and effects suitable for characterizing exposure-response at environmental concentrations;
- Illustrates how multiple pathway alterations induced by environmental factors can lead to and modify risks, and notes how this information might be used to characterize data-limited chemicals and cumulative risks; and

- Identifies some gene variants that influence human susceptibility and alter risks for selected subpopulations, and notes how this information could be used to characterize population variability.

The results presented here are not intended to be a comprehensive review of all available data that might be used in a risk assessment, but rather provide examples of evaluation of new data types and to illustrate potential uses in risk assessment. In addition, toxicogenomics data must be interpreted carefully in this context (see Text Box 4).

Box 4. A Word of Caution in Interpreting Toxicogenomic Results

Technical variability in toxicogenomic results can be a substantial source of data misinterpretation. Rigorous study design and statistical techniques increase confidence in observed associations, and increase the power to detect associations, between exposure and gene expressions particularly at low exposure levels. More generally, without such considerations, variability may obscure actual outcomes or lead to specious associations. Studies without rigorous design, data collection, and analyses are less likely to be considered appropriate for use in risk assessment.

One caveat is that the studies used in the Tier 3 prototypes were chosen mainly because they had some of the most robust, concomitantly collected, traditional and new data types available. These data sets demonstrate partially what can be done with new data types; however, similar data are not likely to be available for many chemicals. This exercise clearly revealed that care must be given to the selection of studies for new types of risk assessment, as many are insufficient for the applications discussed below. The B[a]P prototype, in particular, highlights some of the challenges encountered. Additionally, the fields of molecular, computational, and systems biology are in their infancy in terms of application to human health risk assessment. Although results presented here are promising, robust understanding and full implementation of new methods in general practice, might take years, subject to the resources available for data generation and evaluation.

Implications for risk assessment identified by the Tier 3 prototypes are discussed at the end of this section and integrated with other lessons learned in Section 5, "Lessons Learned from Developing the Prototypes." It should be reiterated that the primary intention of the Tier 3 prototypes is to "ground truth" approaches that could be used in more data-limited situations.

3.1.1. Benzene-Induced Leukemia

Benzene is among the 20 most widely used chemicals in the United States and is among the most common environmental contaminants. A component of crude oil and gasoline, benzene is also used as an intermediate in the manufacture of resins, dyes, chemical solvents, waxes, paints, glues, plastics, and synthetic rubbers. The major sources of benzene exposure are anthropogenic and include fixed industrial sources, fuel evaporation from gasoline filling stations, and automobile exhaust. Benzene has been measured in outdoor air at various locations in the United States at concentrations ranging from 0.02 ppb (0.06 µg/m^3) in a rural area to 112 ppb (356 µg/m^3) in an urban area (IARC 2012). Personal monitoring of benzene exposure in Detroit, Michigan, reported a mean of 1.72 ppb (5.5 µg/m^3) (George et al. 2011). The maximum contaminant level (MCL) in drinking water is 5.0 µg/L or 5 ppb (EPA 2012b). The OSHA permissible exposure limit (PEL) for

benzene workers in the United States is 1 ppm (https://www.osha.gov/dts/chemicalsampling/data/CH_220100.html).

Benzene is a known human carcinogen (IARC 2012, ATSDR 2007, EPA 2000, NIOSH 1992). Epidemiologic studies have shown that benzene exposure leads to an increased risk of acute myeloid leukemia (AML), myelodysplastic syndrome (MDS), hematotoxicity (toxicity to the blood), and other blood disorders (IARC 2012, Schnatter et al. 2012, EPA 2000, Goldstein 1988). AML is characterized by uncontrolled proliferation of clonal neoplastic cells and accumulation in the bone marrow, with an impaired differentiation program. AML accounts for about 30% of all adult leukemias and is the most common cause of leukemia death (Howlader et al. 2013). Studies also indicate that benzene might cause lymphoma and childhood leukemia (Smith, MT et al. 2011).

The extensive molecular epidemiologic and molecular clinical data sets available for both benzene and leukemia are ideal to explore how new data types might be used to inform risk assessments. The work described here focuses on studies where traditional and molecular data were collected simultaneously using a variety of omic methods, including genome-wide analyses of susceptibility genes (using genomic methods), protein expression (proteomics), and epigenetic modification (epigenomics) (McHale et al., 2012). The studies also were conducted over a range of environmental exposure levels (<0.1 ppm to ≤ 10 ppm). The information was developed primarily by Martyn Smith and colleagues (University of California, Berkeley). Systems biology of benzene-induced leukemia is summarized in McHale et al. (2011) and Smith et al. (2011).

Systems Biology of Benzene-Induced Disease

Although benzene is among the most well-studied environmental chemicals, understanding the molecular mechanisms underlying hematopoietic cancer is somewhat recent (see Text Box 5 for a brief description). In 2009, McHale et al. (2012) identified exposure-dependent alterations in genes and pathways (in peripheral blood mononuclear cells using transcriptomics), and hematotoxicity associated with benzene exposure (>10 ppm) in occupationally exposed Chinese workers. McHale et al. (2011) extended these findings to lower exposure levels (<1 ppm to ≤ 10 ppm). (The current U.S. occupational standard is 1 ppm.) In subsequent work, Thomas et al. demonstrated changes in gene expression at current U.S. urban levels in Chinese workers exposed to levels <0.1 ppm. The exposure-response models used in these analyses were not selected *a priori*, but rather driven by the best fit of the data. Results are consistent with supralinear exposure-responses, which have also been reported in traditional epidemiology studies (Lan et al. 2004).

Box 5. Molecular Mechanism of Acute Myeloid Leukemia (AML)

The probable mechanism by which benzene induces leukemia involves the "targeting of critical genes and pathways" (McHale et al. 2012). Benzene has the potential to induce abnormalities in the genes, chromosomes, or epigenetic mechanisms of hematopoietic stem cells (HSC). It can also disrupt its normal cell cycle, leading to apoptosis, increased cell proliferation, and altered differentiation of the HSCs. Benzene causes these effects and ultimately leukemia through oxidative stress, dysregulating proteins that control normal functioning of HSCs, and reducing the ability of the body to detect and destroy cancerous cells (McHale et al. 2012).

For AML specifically, two events that are important for leukemic transformation have been identified. The first event is uncontrolled cell growth, which is mediated by upregulation of cell survival genes. The second event is alteration of transcription factors that control the HSC differentiation. That is, the transcription factor proteins can be mutated or can target certain genes in a way that interferes with the appropriate differentiation of HSCs (Kanehisa Laboratories 2013, Wang, I et al. 2012).

The systems biology of benzene-induced early effects have been articulated by McHale et al. (2012) and others (Smith, MT et al. 2011, Zhang, L et al. 2010). Benzene-induced leukemia is thought to be initiated when metabolites of benzene target genes or pathways that are critical to hematopoiesis in hematopoietic stem cells. Interactions among various cell types within the bone marrow and among various tissues also play a role in leukemia (e.g., immunosurveillance). The underlying mechanisms of benzene-induced leukemia, shown in Figure 6, center on exposure-dependent pathway alterations comprising 147 significantly altered genes (cross validated on two microarray test platforms [Illumina and Affymetrix]). The gene expression profile changes with dose, with some genes (and related biological processes) being expressed at all levels, while others are expressed only at higher concentrations. Of the 147 genes, the expression of 16 genes was significantly altered at all exposure levels. These 16 signature genes are involved in immune response, inflammatory response, cell adhesion, cell matrix adhesion, and blood coagulation, and are most strongly associated with AML disease pathways (McHale et al., 2011). This set of 16 genes forms a biomarker for exposure (and associated leukemia) for future work, particularly in augmenting traditional epidemiology studies and enabling new types of molecular epidemiology studies at lower concentrations. As will be discussed later in this section, understanding of the systems biology and molecular initiating events (MIEs) in leukemia can also potentially enable screening of relatively unstudied chemicals for similar signature events. Clinical studies of chemotherapeutic agents, which alter gene expression in these same pathways and are used in the treatment of leukemia add evidence to the causal relationships between specific gene/pathway alterations and leukemia (Hatzimichael and Crook 2013).

In addition to leukemia, a lymphoma disease signature is evident with benzene exposure (McHale et al. 2012, McHale et al. 2011, Smith, MT et al. 2011). The traditional epidemiology data on lymphoma are not conclusive. Characterization of a benzene-induced molecular mechanism for lymphoma adds considerably to the weight of evidence for benzene-induced lymphoma, highlighting the use of molecular mechanism or mode-of-action information to strengthen weight-of-evidence determinations (IARC 2012).

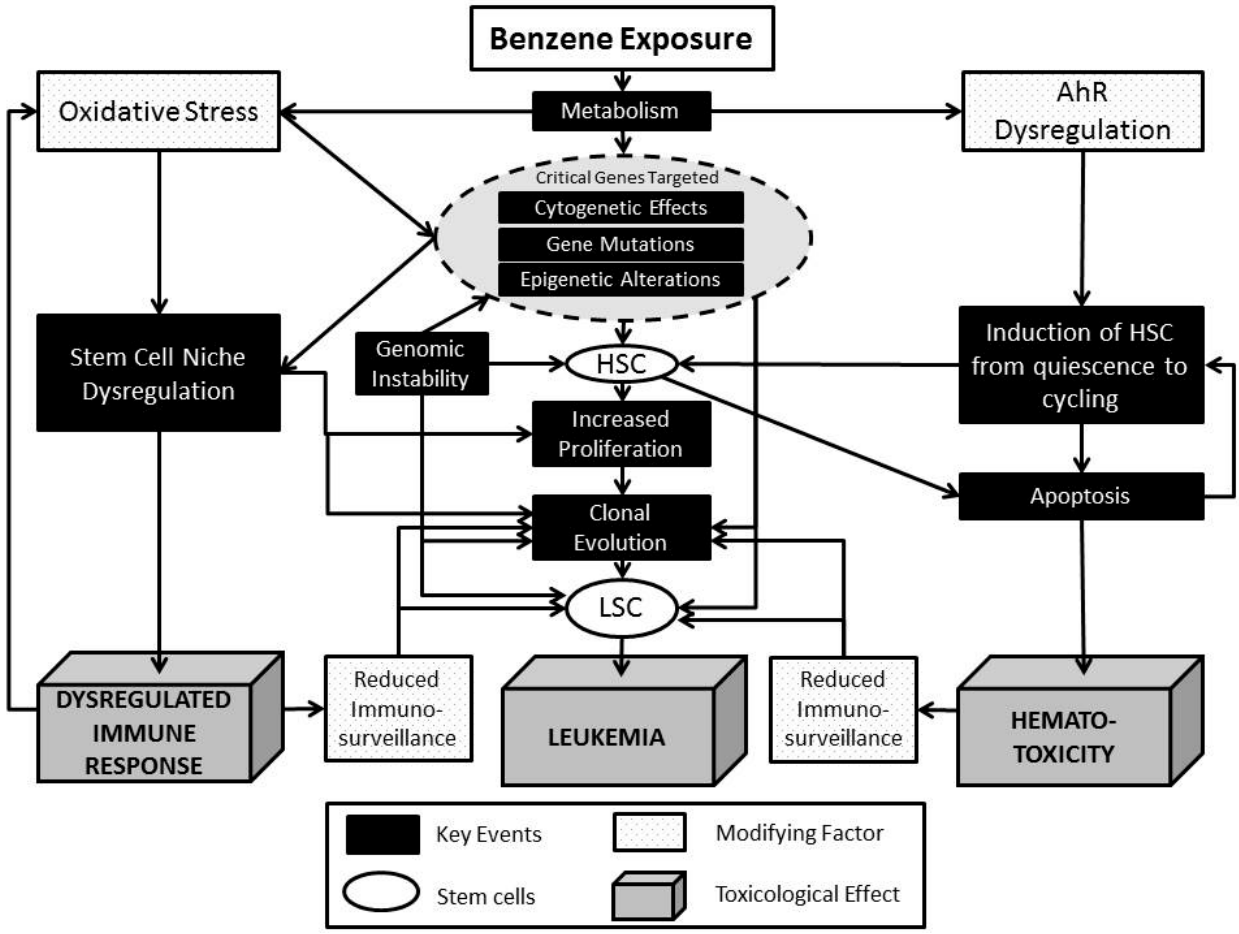


Figure 6. Multiple modes of action (MOAs) of benzene-induced leukemogenesis. Potential key events, modifying factors, and toxicological effects are depicted in the legend. Stem cells can be either HSCs (hematopoietic stem cells) or LSCs (leukemic stem cells) (Smith, MT et al. 2011), reproduced with permission from Elsevier.

De Novo and Other Chemical Leukemogen-Induced Disease

Interestingly, molecular mechanisms for benzene-induced leukemia appear similar to *de novo* (without an obvious cause) AML and AML induced by other environmental agents (e.g., alkylating agents, topoisomerase II inhibitors) (IARC 2012, McHale et al. 2012, Pedersen-Bjergaard et al. 2008). Figure 7a[6] shows a network of genes and pathways involved in *de novo* and chemically induced leukemia [Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa Laboratories 2013)]. The circles in the figure indicate some of the specific genes and pathways affected by leukemogenic agents and environmental modifiers (Kanehisa Laboratories 2013, IARC 2012, McHale et al. 2011, Pedersen-Bjergaard et al. 2008). Additional evidence for the causal role for these genes and pathways in AML is provided by the study of human genetic variants associated with altered risks and chemotherapeutics that reverse adverse alterations in some of these same

[6]The basic AML network figure used in Figures 7a and 7b is from the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa Laboratories 2013)). The added circles are the work of the report authors.

genes and pathways (discussed below). Although mechanistically similar, different agents can display specific characteristics; including origins in cells at different stages of hematopoiesis, distinct cytogenetic subtypes, and different latencies (Irons et al. 2013, McHale et al. 2012). Figure 7a highlights how a disease network can be modified at different points but still lead to a common disease outcome. These mechanistic commonalities and differences among *de novo* and chemically-induced health effects can be used to characterize chemicals with limited data, which have nevertheless been shown to induce mutations, chromosome changes, or specific changes in gene expression. Data-limited chemicals would be of elevated concern if they are shown to alter pathways similar to that observed in *de novo* disease or with well-studied leukemogens. For example, see the work in Thomas R et al. (2012) where the authors used existing information on gene and protein targets of 29 known leukemia-causing chemicals and 11 carcinogens that are not known to cause leukemia, the authors were able to develop a classification scheme that could distinguish a random leukemia-causing/nonleukemia-causing carcinogen pair with a 76% probability. Provided later in this section (in the ozone and B[a]P prototypes) is similar evidence for the importance of networks when considering chemical-related diseases, similarities of chemical-related and *de novo* diseases, and the role of mechanisms in improved understanding of cumulative risks.

Cumulative Risks from Environmental Factors

Evidence suggests that, in addition to environmental exposures, genetic variations and lifestyle factors such as smoking, obesity, diet, and alcohol use are risk factors for leukemia (Smith, MT et al. 2011, Pedersen-Bjergaard et al. 2008, Belson et al. 2007, Ilhan et al. 2006). Environmental exposures of the developing organism could also be a risk factor for disease later in life, given the potential of benzene and other environmental agents to alter epigenetics, the sensitivity of the developing organism to epigenomic changes, and the association of environmental exposures and childhood leukemias (Boekelheide et al. 2012). Figure 7a shows how multiple environmental factors can alter several molecular events in a manner that alters risks, and how mechanistic knowledge might be used to identify or exclude chemicals based on common mechanisms and impacts on cumulative risks.

Individuals exposed to known environmental and lifestyle risk factors are estimated to account for approximately 20% of acute leukemia incidences, indicating that host genetic susceptibility might be instrumental in the development of leukemia (Smith, MT et al. 2011). By identifying mechanistic commonalities, or the lack thereof, among chemicals, new omic approaches can provide tools for characterizing roles that intrinsic and extrinsic risk factors might play in individual and subpopulation risks. Below we discuss genetic variation more specifically and provide an example of altered subpopulation risks based on genetic variations.[7]

Genetic Variation and Susceptibility in the Human Population

Genetic susceptibility for developing AML, and how it relates to chemical risks, has been studied by several investigators (Zhuo et al. 2012, North et al. 2011, Shen et al. 2011, Smith, MT et al. 2011, Garte et al. 2008). Several genetic variations in individual genes appear to increase risks for

[7]Human genetic variation is the genetic differences among subpopulations. Multiple variants of any given gene might occur in the population. These differing DNA codings determine distinct traits or polymorphisms that can influence risks.

developing AML, while at least one decreases risks. Sillé et al. (2012) reported 12 independent risk loci (specific regions within the genome, which can be a single base, as in this case, or an entire gene) with the potential to alter gene expression. A significant number of variants (single nucleotide polymorphisms [SNPs] or single nucleotide variations) related to a tumor suppressor gene, signaling pathways, or residing in putative regulatory elements,[8] have been linked to various types of multiple hematological cancers. Figure 7b highlights genes that vary in the human population and are associated with altered leukemia risks (Hatzimichael and Crook 2013, Kanehisa Laboratories 2013). Figure 8 provides an example of differential risks resulting from one human variant.[9] The overall data indicated a significant variation in risk (42%) relative to the CYP1A1 genotype (Zhuo et al. 2012). The shift in odds ratio is also shown in Figure 8.

When one considers that many genes are associated with benzene-induced leukemia, the potential for variation in subpopulation risks via individual genes, combinations of genes, and gene variants becomes apparent. Other risk factors (e.g., lifestyle) would add to the human variability in response. As discussed in Section 4, NexGen approaches exist that can facilitate characterization of human variability as never before.

[8]Putative regulatory elements are areas of the gene that do not code for proteins but rather regulate DNA transcription into proteins.

[9]SNP leads to a base substitution of isoleucine with valine at codon 462 in exon7 (Ile462Val or CYP1A1*2C polymorphism, rs1048943). Thus, the exon7 restriction site polymorphism results in three genotypes: a predominant homozygous Ile/Ile, the heterozygote Ile/Val, and a rare homozygous Val/Val.

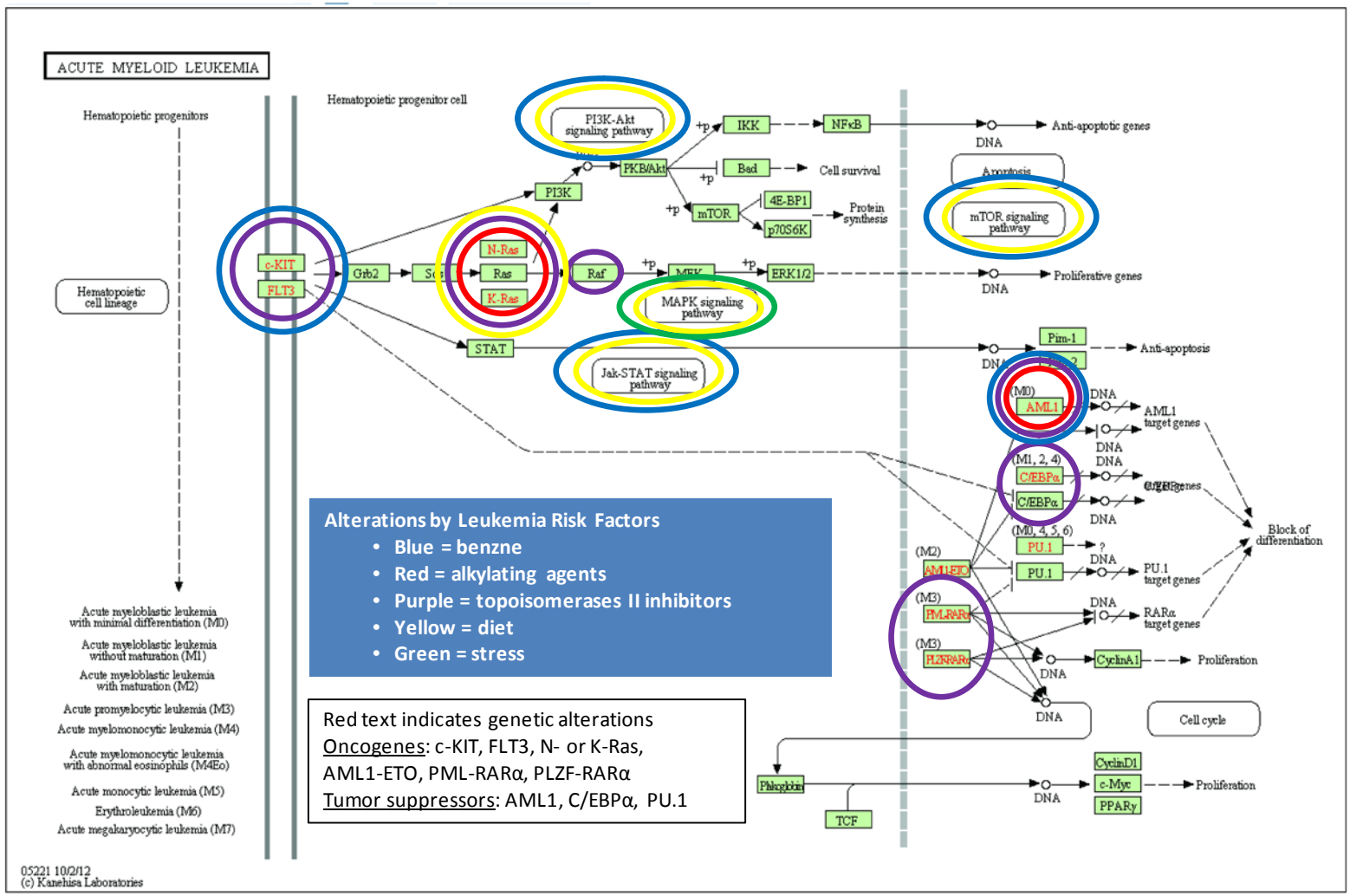


Figure 7a. The Kyoto Encyclopedia of Genes and Genomes (KEGG) diagram (http://www.genome.jp/kegg-bin/show_pathway?hsa05221) illustrates some of the currently understood molecular pathways involved in acute myeloid leukemia (AML). Altered oncogenes and tumor suppressor genes are noted in red type (Kanehisa Laboratories 2013). The circles (added by authors) note specific genes and pathways that are modified by benzene, other chemical leukemogens, and other risk factors. While intended to be illustrative rather than comprehensive, it can be seen how single or combinations of environmental factors could modify risks for leukemia, and how such knowledge could be used to evaluated joint effects of environmental factors (IARC 2012, McHale et al. 2012, Smith, MT et al. 2011, Pdersen-Bjergaard et al. 2008).

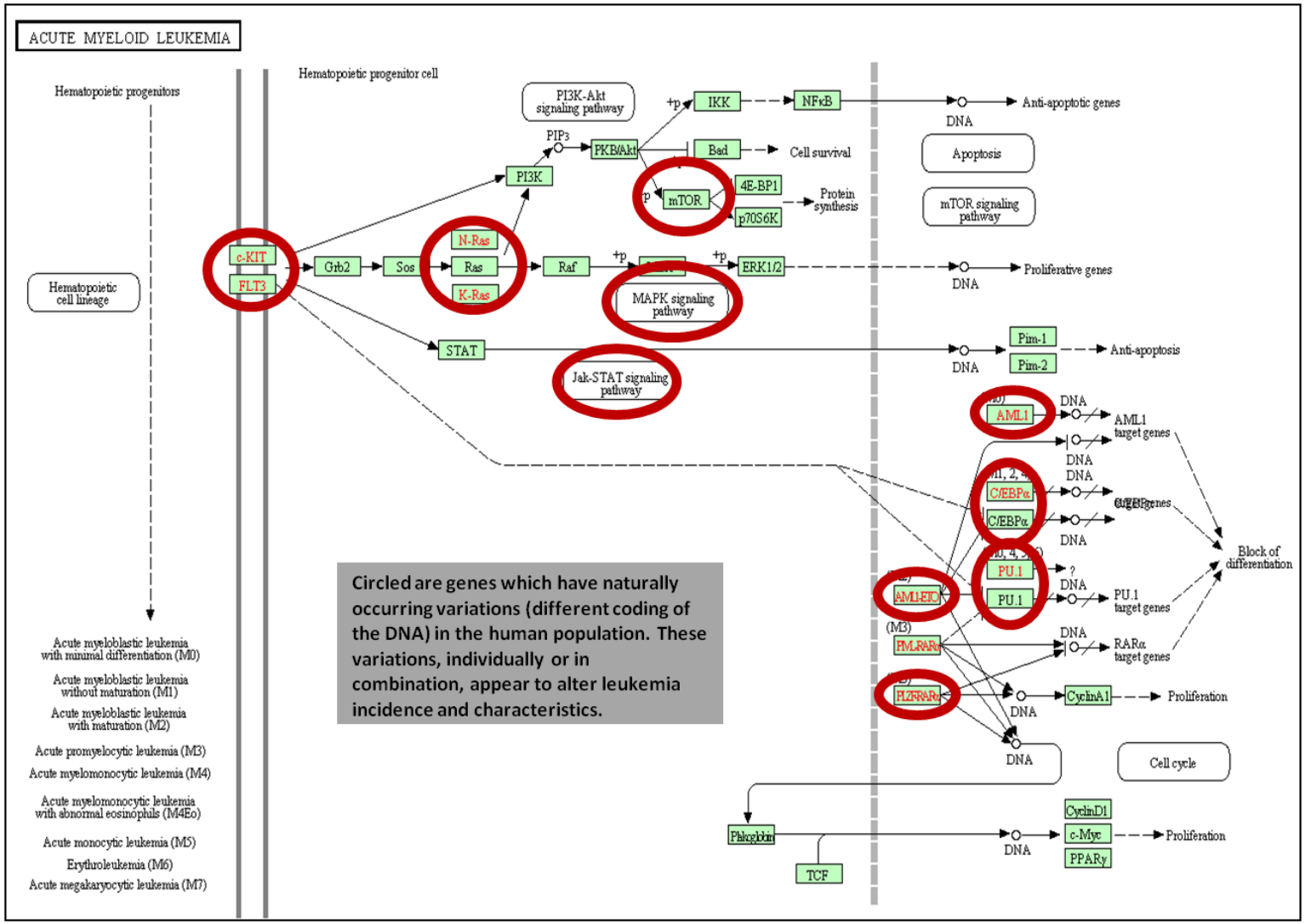


Figure 7b. This figure shows the same acute myeloid leukemia (AML) KEGG diagram (Kanehisa Laboratories 2013) as shown in Figure 7a, with circles added by authors. In this version, circled are the locations of naturally occurring human genomic variants that increase the risk of AML (Hatzimichael and Crook 2013, Sille et al. 2012). Characterizing genomic variant subpopulations and associated risks can help us to better describe human variability and susceptibility for specific diseases.

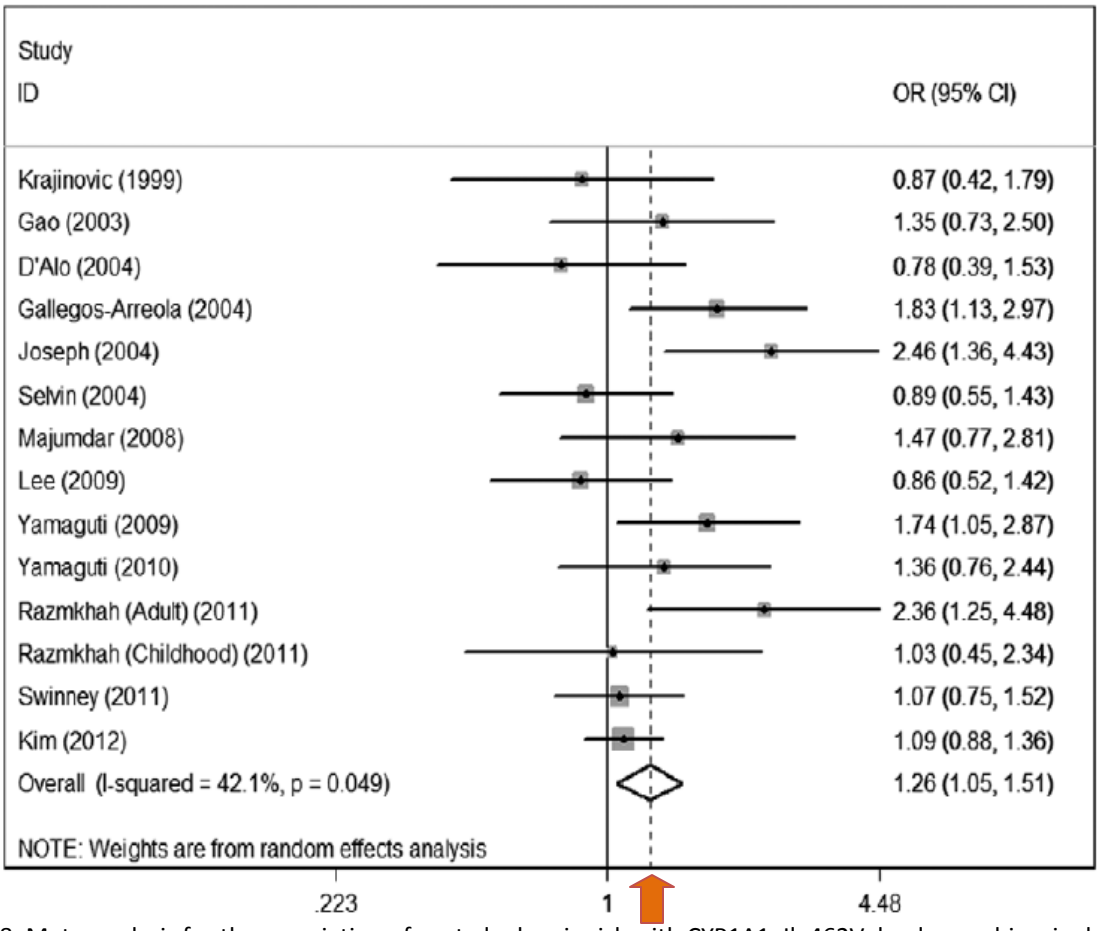


Figure 8. Meta-analysis for the association of acute leukemia risk with CYP1A1. Ile462Val polymorphism is shown (OR = odds ratio). The overall risk was 42% greater (95% CI = 1.11–1.98) for Val/Val+Val/Ile versus Ile/Ile (Zhuo et al. 2012). Reproduced with permission from PLoS One.

In Vitro Evaluation of Toxicogenomic Signatures

As has been previously noted, the primary function of the Tier 3 prototypes is to inform how we evaluate data-limited chemicals. Hence, a comparison of *in vivo* and *in vitro* benzene results is discussed here. Godderis et al. (2012) conducted an *in vitro* study in TK6 cells to detect gene signatures and biological pathway perturbations, using global gene expression analysis, resulting from exposure to 15 genotoxic carcinogens, including benzene and its metabolites. The goal was to determine if well-characterized chemicals could be used to characterize data-limited chemicals by comparing gene signatures. Although pathways altered by exposure to benzene and its metabolites were in general agreement with previous *in vivo* studies, the authors pointed out that several factors can complicate comparison of *in vivo* and *in vitro* data, for example, metabolism and cell types. The authors concluded that use of toxicogenomic signatures hold great promise for evaluation of data-limited chemicals. They noted that for the carcinogens in the study, some *in vitro* processes mapped against known or likely carcinogenic processes, but determining discriminatory mechanisms based on *in vitro* data alone was difficult. This observation suggests that the approach of developing putative mechanisms of action based on data-rich meta-analyses of human disease

and mapping *in vitro* data against these models might prove more successful than attempting to understand mechanisms of action based on *in vitro* data alone.

Risk Assessment Implications from the Benzene Prototype

The benzene prototype demonstrated how molecular biology data, particularly mechanistic signatures, can be used in hazard identification and exposure-dose-response assessment.

Hazard Identification – Specifically, genes and pathways altered by benzene exposures are strongly associated with a network of pathways associated with known (AML) and likely (lymphoma) outcomes. Additional evidence for a causal relationship between alterations in specific genes and pathways and increased leukemia risk is provided by observed similarities in pathway disruptions: (1) caused by other chemical leukemogens, (2) observed in leukemia of unknown origins, and (3) reversed by certain leukemia chemotherapeutic agents. Hence, observations from both molecular epidemiology and molecular clinical studies provide evidence that molecular signatures can predict specific diseases with some confidence. These data suggest that well-defined pathway and network disruptions strongly associated with a specific disease could be used to screen chemicals with limited molecular data for their potential to increase risks for the specified disease by causing similar mechanistic disruptions. Anchoring of the molecular patterns to apical outcomes, considerable systems biology knowledge, and high-quality data; however, appear necessary to define the disease signature against which data-limited chemicals could be compared.

Exposure-Dose-Response Assessment – A specific exposure-dose-dependent gene signature for leukemia was observed at all environmental exposure concentrations measured (<0.1 to >10 ppm); the magnitude of signature expression varied in a dose-dependent manner. This signature is a biomarker of both exposure and effect. Such signatures or biomarkers can extend the exposure range of traditional epidemiologic studies to lower exposures and reduce measurement error. This type of data can measure low-dose-response relationships and, potentially, mitigate a source of substantial controversy in chemical risk assessment, that is, low-dose extrapolation. In the future, one can envision routine replacement of low-dose extrapolation with measurements of molecular signatures. The established dose-response for specific gene signatures could be used to estimate the potency or relative potency for data-limited chemicals. In particular, ranking of chemicals is feasible when using similar protocols such as those characteristic of Toxicology in the 21st Century (Tox21) or ToxCast™.

The exposure-response models used in this prototype were not specified in advance, but the choice relied on the best fit from among multiple models. Hence, the model was "agnostic" on the issues of threshold/no threshold and the shape of the low-exposure-response relationship. Such an approach would mitigate another source of controversy in risk assessment, that of model choice.

Cumulative Risk Assessments – Understanding of a common mechanism of action for multiple environmental factors can allow for improved cumulative risk assessments. It should be noted that overly simplified descriptions of mode of action (MOA) or adverse outcome pathways (AOPs) could miss interactions among the environmental factors as shown in Figure 7a.

Variability and Susceptibility in Human Response – An example of risk characterization associated with different genetic variations is provided. With additional research and data evolving from personalized medicine, the understanding of population variation and distribution of

responses in the human population could be improved. These data also could help improve estimates of the size of sensitive subpopulations.

3.1.2. Ozone-induced Lung Inflammation and Injury

Use of Ozone as a Model Pollutant

Hundreds of controlled human exposure studies have described biological changes in volunteers exposed acutely (usually for 2–6 hours) to concentrations of ozone ranging from 0.06 to 0.4 ppm (EPA 2011d).[10] These studies show that exposure to ozone results in decrements in several indices of lung function, increases in markers of pulmonary inflammation, and alterations in host defenses against inhaled pathogens and lung injury. This database represents the single largest human database of any pollutant EPA has studied. As a consequence and because the mechanisms are well understood, the database provides an ideal opportunity to demonstrate proof of concept for use of molecular biology data to inform assessment of human risks, to develop decision considerations for use of such data, and to explore the value of various types of information.

The underpinning of an AOP-based paradigm in risk assessment methodology is the concept of studying biological pathways. The perturbation of a biological pathway initiates a set of key events that cause an adverse outcome associated with an environmental stressor. If such pathway responses are known and represented by a set of quantitative *in vitro* assays, the results of these assays can be used to build quantitative biological activity relationships. Coupling these results with appropriate physiologically based pharmacokinetic (PBPK) modeling and exposure estimates for estimating tissue doses can be useful for hazard identification and dose-response assessment. For *in vitro* pathway information to be used in risk assessment, the quantitative relationship between perturbation of a pathway following *in vitro* exposure and downstream endpoints (i.e., pathophysiological changes at the tissue or organism level following *in vivo* exposure of animals or preferably humans) must be established. This framework is not likely to be possible, however, as sufficient *in vivo* data are lacking for most of the toxicants that EPA is responsible for regulating (Crump et al. 2010). Therefore, using model systems in which both *in vitro* and *in vivo* data are available is necessary to validate how well pathway information from the former can predict human responses to toxicants. Ozone provides such a model system for lung inflammation and injury (see Text Box 6 for a description of inflammation). This model can be used for less well-studied chemicals to identify and characterize their potential to induce lung inflammation and injury. Figure 9 outlines physiological and cellular pathways by which ozone causes pathophysiological changes in humans via the lung response. This prototype focuses on pathways that lead to inflammation, which are shown in the open boxes. Several human studies characterize inflammation at multiple ozone concentrations during and after exposure, providing a rich data set

Box 6. Inflammation

Inflammation is the immune system's response to damage to cells and organs by pathogens, chemicals, or physical insult. Initially, inflammation involves changes in local blood flow and accumulation of various inflammatory cells (e.g., neutrophils, lymphocytes) at the site of injury. Pathogens and cell debris caused by the inflammatory response are then removed as tissues begin to repair. If the delicate balance between inflammation and resolution of the events leading to the inflammation is dysregulated, or tissue insult continues, inflammation can lead to disease pathology (Wang, I et al. 2012, Medzhitov 2008).

[10]The current ozone standard calls for limitation of the fourth highest daily maximal 8-hour ozone concentration in a year to 0.075 ppm, based on a 3-year average.

of human *in vivo* responses. Additional pathways based on neurological responses to ozone exposure that are not captured in this figure also might be possible.

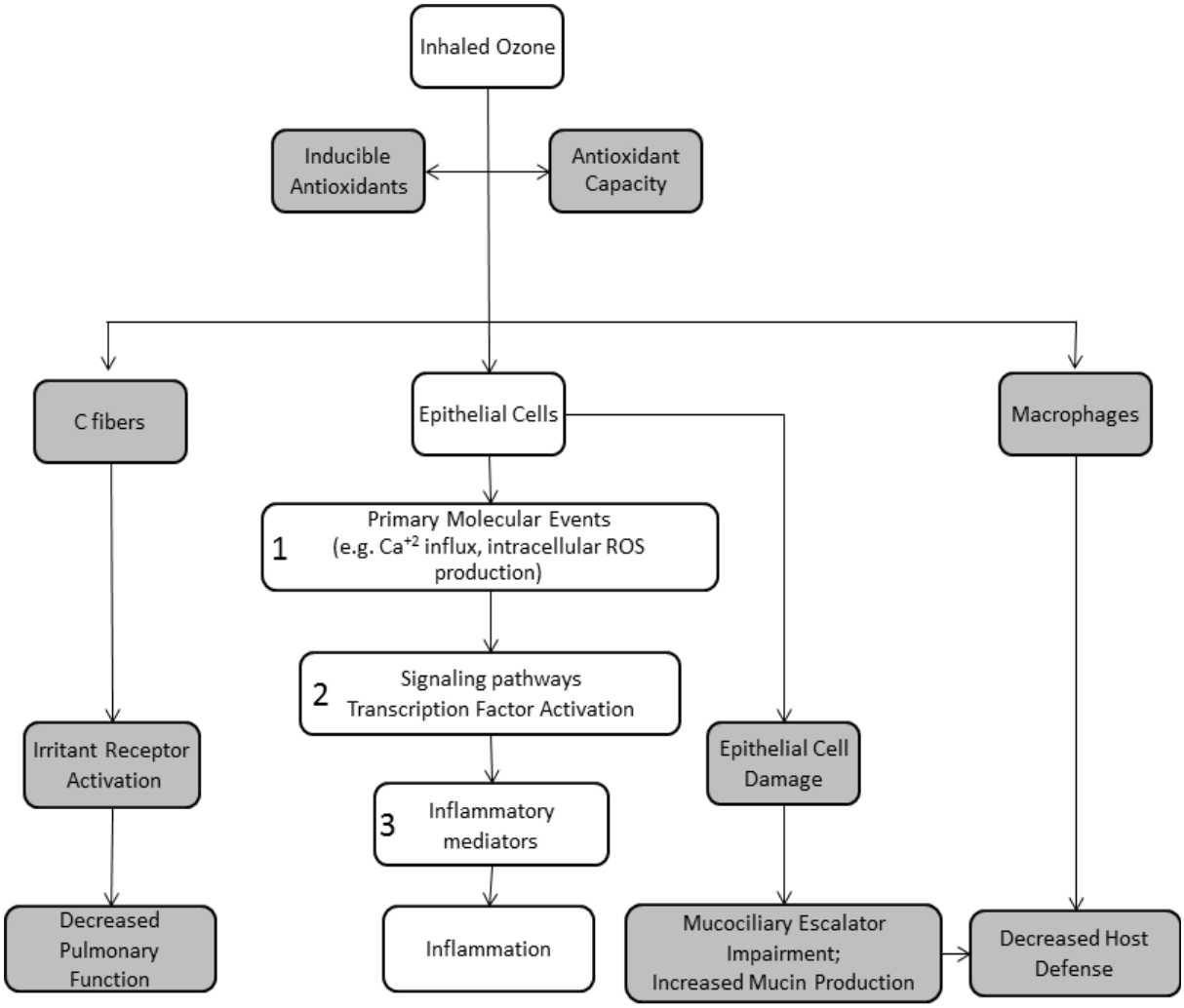


Figure 9. Framework diagram of ozone key events and modes of action (MOAs) related to lung injury occurring *in vivo*.

Challenges with Using an AOP Approach for Risk Assessment

Using model systems based on *in vitro* pathway information to predict human *in vivo* responses to toxicants for risk assessment purposes presents certain challenges. A major hurdle relates to extrapolation from *in vitro* to *in vivo* effects. Many *in vitro* approaches use animal cells or transformed cell lines derived from humans, which might not accurately reflect cell interactions or events in the pathway for human *in vivo* effects. For example, a parent toxicant might be biologically transformed into a more active form by cells that are not represented in the *in vitro* system (e.g., liver cells) before interacting with the target cells represented in the assay. In the lungs, epithelial cells that line the human airways are the first and primary targets of inhaled toxicants and are believed to be the cells that initiate lung inflammation. Studies have shown that pathways in the

cultured *in vitro* cells that have been activated by air pollutants are also altered in these same cells following *in vivo* exposure to the same pollutant (Selgrade et al. 1995). This ability to show concordance between *in vitro* and *in vivo* exposures thus is a major advantage of the modeled lung system discussed here.

A second challenge associated with *in vitro* approaches is ensuring that the dose of toxicant delivered to cultured cells is similar to that which these cells would encounter following an *in vivo* exposure. Frequently, cultured cells are exposed to toxicant levels that are orders of magnitude greater than they would be *in vivo.* There is no assurance that the same biological pathways are adversely affected in both situations. Ozone, however, can be prepared using the heavy oxygen isotope ($^{18}O_2$), which can be separated from $^{16}O_2$ and quantified by mass spectroscopy. When ozone attacks a target tissue, the $^{18}O_2$ tag is bound to that tissue. This approach has been used to normalize the dose of ozone delivered to rats and humans (Hatch et al. 1994) and to support estimates that target tissue doses in rats exposed to 2.0 ppm ozone are comparable to target tissue doses in humans exposed to 0.4 ppm ozone. This same approach can be used to normalize the dose of ozone delivered to cultured cells and humans.

Ozone is one of the few pollutants for which an extensive animal and human health effects database is available. Coupled with *in vitro* pathway data, this prototype pollutant can be used to illustrate both how a biologically based dose-response modeling approach can be used to provide this framework and how a systems biology model and genomics data can be used for risk assessment.

AOP Studies

In Vivo Studies – Young, healthy volunteers were exposed to filtered air and a relevant concentration of ozone (0.30 ppm) previously shown to induce a measurable inflammatory response. Bronchoscopy was used to obtain cells and lung fluid at 1 and 24 hours after exposure. To ensure that pathophysiological effects observed in this study were comparable to those reported in earlier studies, downstream biomarkers of inflammation such as the influx of neutrophils were measured (Devlin et al. 2012), as were markers of cell injury (lactate dehydrogenase) and leakage of plasma components across the damaged epithelial cell barrier (albumin) into the lung airways. Bronchial airway epithelial cells were obtained by brush scraping, and the microarray technology was used to define pathways affected by *in vivo* ozone exposure. In addition, quantitative proteomics was used to correlate changes in messenger ribonucleic acid (mRNA) measured by microarray with changes in their protein counterparts (see Figure 9, event 3).

In Vitro Studies – A subset of airway epithelial cells was collected from volunteers following exposure to filtered air and cultured at an air-liquid interface. These cells were exposed to concentrations of ozone that had been shown (from the results of $^{18}O_3$ experiments) to be comparable to the dose of ozone encountered by airway epithelial cells following a specified *in vivo* exposure. This approach allows comparison of an *in vitro* and *in vivo* response of cells from the same person for comparable exposures. Similar to the *in vivo* studies, microarray and proteomics were used to identify and define pathways affected by ozone in these cells.

Signaling Pathways – Upstream signaling events shown in Figure 9, event 2 (e.g., transcription factor activation, MAP kinase pathways, production of reactive oxygen species [ROS]) was assessed to determine the MOA by which ozone activates downstream batteries of pro-inflammatory genes. Pathways that are altered by exposure of cultured airway epithelial cells to ozone can be compared with those altered in airway epithelial cells of the same person exposed *in vivo* to ozone. A comparison can be made to determine the accuracy of the *in vitro* system in mimicking events

following exposure in the *in vivo* system and to assess differences in the variability of the response. Figure 10 illustrates potential upstream signaling pathways that could be induced by ozone and lead to activation of downstream batteries of pro-inflammatory genes.

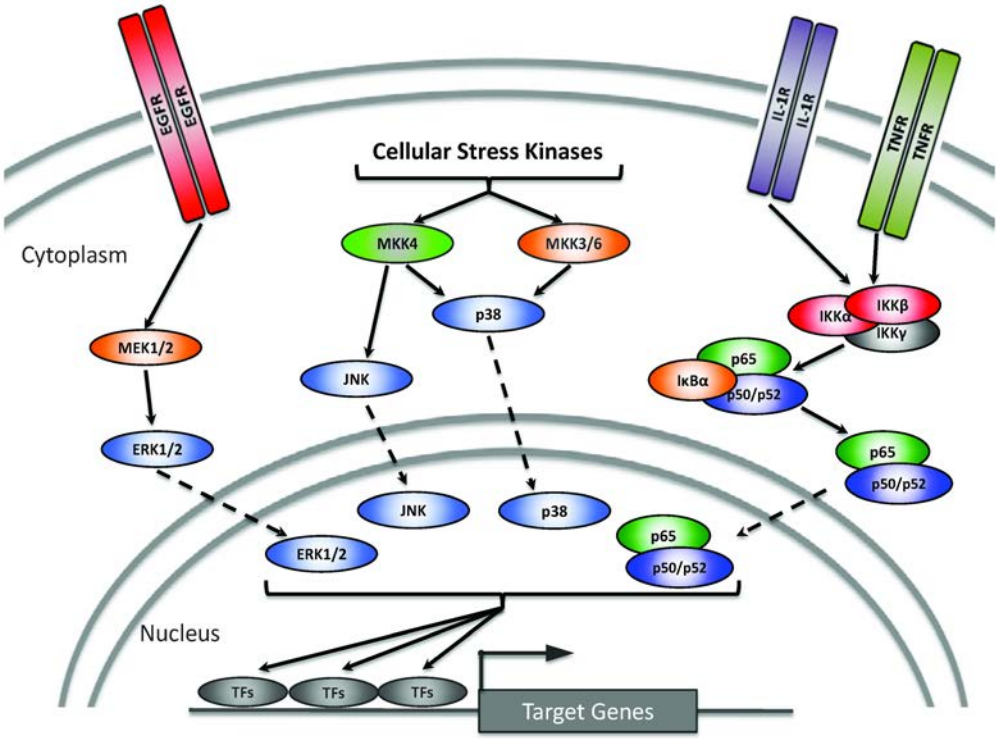


Figure 10. Potential pathways by which ozone causes production of pro-inflammatory mediators in epithelial cells.

Microarray technology was used to determine which of these pathways is most likely to be altered by ozone exposure. The two most highly scoring molecular networks following exposure of cultured airway epithelial cells to ozone *in vitro* are presented in Figure 11. The networks involve modulation of genes in NF-κB and extracellular signal-regulated kinase signaling pathways. The gene list input to the Ingenuity Pathway Analysis was generated by combining all genes found to be differentially expressed immediately following a 2-hour exposure of bronchial epithelial cells to 0.25, 0.50, 0.75, and 1.0 ppm ozone or clean air. Exposure-dose was normalized using $^{18}O_3$ dosimetry from *in vitro* and *in vivo* human studies. Networks are displayed with representative symbols for the protein products of the mRNA transcripts. Red represents putative upregulated transcripts induced by ozone, and green represents putative downregulated transcripts in response to ozone. Additional molecules from the Ingenuity Knowledge Base, which were not present in the differentially expressed gene (DEG) list, are uncolored in the networks. The same putative networks were also identified in epithelial cells removed from human airways 1 hour after *in vivo* exposure to ozone.

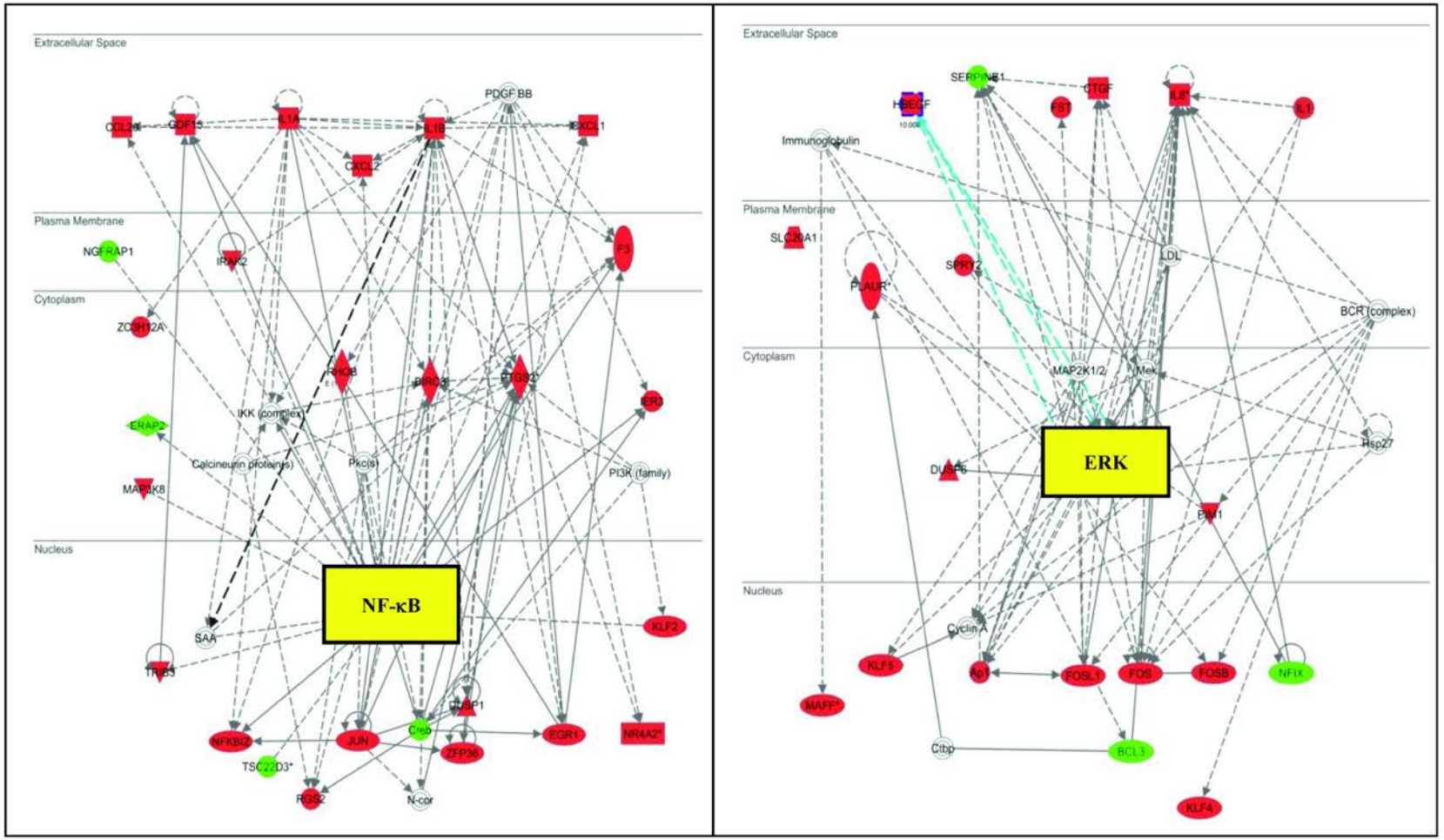


Figure 11. Molecular pathway analysis by Ingenuity Pathway Analysis.

Primary Molecular Events

Many pollutants induce intracellular oxidative stress, which can affect signaling pathways and ultimately lead to activation of batteries of pro-inflammatory genes. One pathway by which this might occur (Figure 12) is activated in cultured human airway epithelial cells exposed to particulate air pollution. Ozone is an inherently potent oxidant and is known to cause oxidative damage to lipids, proteins, and nucleic acids. Until recently, whether ozone also induced intracellular ROS was unknown. Figure 13 shows that ozone can induce a rapid dose- and time-dependent increase in cytosolic intracellular glutathione redox potential, a measure of ROS (Gibbs-Flournoy et al. 2013). Whether the ROS produced following ozone exposure actually activates downstream signaling pathways via the mechanism shown in Figure 12 is unknown.

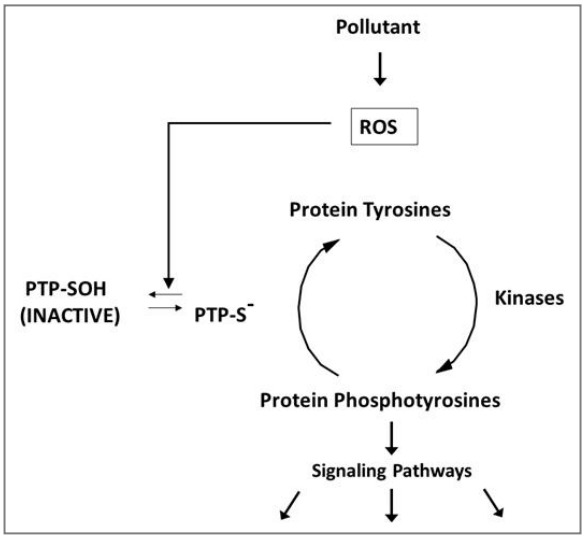


Figure 12. Role of reactive oxygen species (ROS) in mediating pollutant-induced inflammation.

System Biology Modeling

Quantitative systems biology models are translational, and their development is data driven, with model structure and dynamics parameterized using data on (1) basic biology, (2) how the biology is perturbed by toxicants, and (3) how and when adaptive and adverse responses develop. Sufficiently well-developed and well-validated models can be used to predict dose-response and time course behaviors for the perturbations, adaptive responses, and apical health effects, but the accuracy of these predictions depends on the extent and quality of the data used as inputs and on the technical quality of the model itself. Time-course and dose-response pathway data from *in vitro* exposure studies can be paired with pathway data from *in vivo* exposure studies and assembled into a nodes-and-edges graph encompassing mechanisms of action relevant to ozone toxicity, focusing on pathways most relevant to lung inflammation. This pairing and assembly will provide a framework for modeling ozone toxicity pathways to downstream pathophysiological changes (see Figure 9, event 3). At the intracellular level, upstream signaling pathways (e.g., NF-κB) that have been shown to mediate ozone-induced changes in gene expression will be represented, connecting the oxidative products of ozone formed in the cell to time-dependent changes in protein activity and RNA expression. For example, the canonical NF-κB signaling pathway shown in Figure 9 plays a role in ozone-induced inflammation. Finally, data on ROS production resulting from ozone exposure (see Figure 9, event 1) will be represented in the model, both as an input to ozone's perturbation of the molecular-level components and as drivers of downstream signaling pathways.

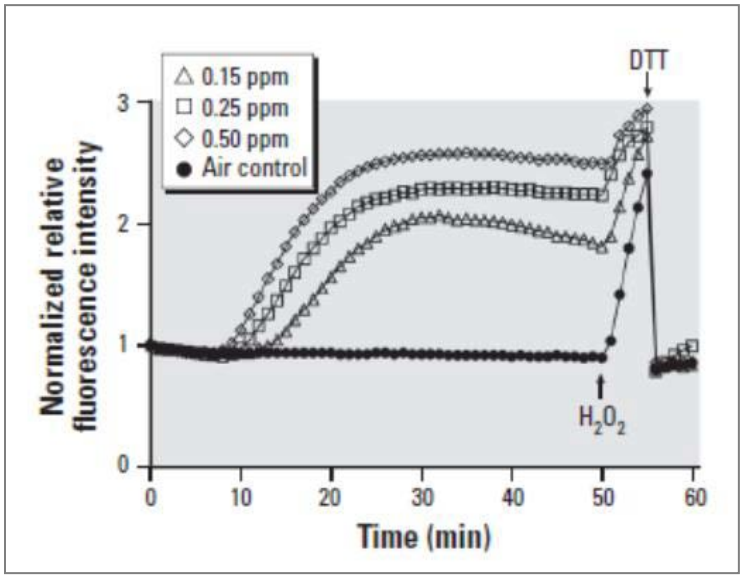


Figure 13. Exposure to ozone induces a rapid increase in intracellular reactive oxygen species (ROS). Addition of 0.1 mM H_2O_2 at the end of the ozone exposure produced a maximal response, which was fully reversible with the addition of 10 mM dithiothreitol (DTT), a strong reducing agent (Gibbs-Flournoy et al. 2013). Reproduced with permission from *Environmental Health Perspectives*.

Susceptibility

Not all individuals are equally responsive to toxicants; some are much more responsive because of age, gender, disease, lifestyle (e.g., obesity), or genetic/epigenetic factors. For example, the range of response in lung function decrements to ozone in young healthy individuals (McDonnell et al. 2012) is 10-fold. Individuals exposed to ozone a second time, many months later, retain their hierarchy on the response curve, implying that a long lasting factor, perhaps genetic or epigenetic, plays a role in ozone responsiveness. Asthmatics are known to have an enhanced inflammatory response to ozone (Bosson et al. 2003, Peden et al. 1997), as do individuals carrying the GSTM1 null allele (Kim et al. 2011). Understanding the MOA by which a person is more responsive to a pollutant should be a component of a systems biology approach to toxicity testing. Airway epithelial cells can be obtained from more-responsive and less-responsive individuals, and the pathways altered by ozone can be compared for both groups. Recently, cultured lung epithelial cells obtained from individuals carrying the GSTM1 null allele have been shown to be more responsive to air pollutants than cells obtained from individuals carrying the wild-type GSTM1 allele (Wu et al. 2011). Airway epithelial cells obtained from asthmatics appear to retain an asthma phenotype in culture and are more responsive to pollutants than cells obtained from nonasthmatics (Duncan et al. 2012). Readily obtaining bronchial airway cells can be difficult, so knowing that the response of cultured nasal epithelial cells to toxicants has recently been shown to be similar to that of bronchial cells (McDougall et al. 2008) can be instructive. These nasal cells can be readily and noninvasively obtained from most individuals, including children.

Involvement of the Inflammatory Network in Multiple Diseases

Chronic inflammation is implicated in the etiology of several diseases, including atherosclerosis, heart disease, obesity, diabetes, arthritis, cancer, and lung diseases (asthma, emphysema, pulmonary fibrosis). Both common and disease-specific inflammatory molecular patterns have been reported to underlie these diseases (Wang, I et al. 2012). Why a particular disease is expressed in an individual or subpopulation as the result of inflammation is likely the result of the site of injury, co-activation of other networks, genetic variation, or environmental factors. Such complicating factors therefore highlight several issues that might arise when using molecular patterns to predict disease risks: (1) observation of an inflammatory disease signature for a chemical that has not been well studied would raise concerns for inflammatory disease risks; (2) the specific inflammatory disease in question likely would be difficult to predict with a limited

systems biology context; (3) a network might be involved in multiple diseases; and (4) the specific disease expressed could involve multiple interactive pathways and networks.

Specifically, many air pollutants appear to induce cardiopulmonary inflammation, which likely plays a role in risks for asthma, emphysema, and pulmonary fibrosis. Molecular biology is likely to be a useful tool in sorting out the relative contributions of various air pollutant exposures to cardiopulmonary disease via inflammatory mechanisms.

Risk Assessment Implications Based on the Ozone Prototype

Hazard Identification – The pathway information, coupled with data about ozone-induced changes in upstream transcription factors, signaling pathways, and generation of ROS, can lead to the development of molecularly based dose-response system models that are predictive of downstream *in vivo* pathophysiological changes. These data suggest that ozone activates the NF-κB and ERK pathways, both known to modulate inflammation, *in vitro* and *in vivo*. This suggests that the *in vitro* airway epithelial cell model used here might be amenable to predicting *in vivo* inflammation. An HTS assay based on this cell model might be able to provide rapid hazard identification in the future.

Exposure-Dose-Response Assessment – We did not perform an analysis of transcriptional changes across a range of doses.

Cumulative Risk Assessment – This *in vitro* model could be used to make comparisons of the transcriptional response upstream of the inflammation process using complex mixtures of air pollutants. The comparison, however, might require specialized equipment and monitoring to ensure the mixture and dose of pollutants are proper and well controlled.

Variability and Susceptibility in Human Response – In the future, this and other similar models might identify pathways and mechanisms by which susceptible human populations respond to inhaled toxicants. Just as this *in vitro* model was derived from several young, healthy volunteers, performing a larger study of variability and susceptibility would be possible by recruiting and including specific populations. Such a study also would facilitate the creation of HTS assays for rapidly studying susceptible populations and variability in response.

3.1.3. Benzo[a]pyrene (a Polycyclic Aromatic Hydrocarbon), and Cancer

PAHs are produced from combustion or pyrolysis of carbon-containing material, exist in the environment almost exclusively as complex mixtures, are a major component of urban air pollution, and are a drinking water contaminant. Several PAH-containing complex mixtures are known to be carcinogenic in humans (e.g., coke oven emissions, diesel exhaust, and tobacco smoke). Many individual PAHs and PAH-containing mixtures have been tested in traditional bioassays; many, but not all, appear carcinogenic. Additionally, those that are carcinogenic vary in terms of potency. Given the universe of PAHs and potential PAH-containing mixtures, testing them all is not feasible. Hence, an alternative approach using molecular biology was explored in this prototype. See Text Box 7 for some challenges related to this prototype.

This effort focused on one PAH—B[a]P)—and liver cancer. Repeated B[a]P exposure has been associated with increased incidences of total tumors and of tumors at the site of exposure (dietary, gavage, inhalation, intratracheal instillation, and dermal and subcutaneous, in studies of numerous

strains and species of rodents and several nonhuman primates). Distant site tumors also have been observed after B[a]P administration by various routes, and B[a]P is frequently used as a positive control in carcinogenicity bioassays.

Box 7. Challenges Encountered With This Prototype

This prototype originally focused on identifying whether human transcriptomics data from PAH mixtures found in cigarette smoke could be associated with lung cancer. This prototype was envisioned as a real-world example of how data mining of existing data could be informatively performed. Unlike the other Tier 3 Prototypes, which were designed to have the best combination of data available, however, this prototype encountered numerous data access and experimental design challenges that we expect to be seen when applying these methods in the future. These challenges included:

- An inability to easily obtain the raw data required for re-analysis of the transcriptomics data.
- Lack of clear descriptions of the study design or analysis method.
- Different microarray platforms being used.
- Different analysis methods being employed within the same platform.
- Lack of a quantitative exposure estimate (especially common with human studies that lack a controlled exposure).

Together, these challenges make performing a quantitative meta-analysis difficult. For new types of data to be useful, improvements to data collection and concomitant exposure analyses are needed.

Systems Biology Model

EPA (2013), Burgoon (2011), have proposed a cellular systems model and pathways based on a systematic meta-analysis of transcriptomics data for B[a]P-mediated liver cancer (Figure 14 and Table 2). The core of the model is focused on induction of DNA adducts, mediation of p53 (a tumor suppressor gene) signaling, alterations of translesion synthesis,[11] and regulation of the G1/S-phase transition and the cell cycle. Based on this model, the DNA adducts are believed to be formed by reactive B[a]P metabolites through cytochrome P450 (CYP) enzyme induction, secondary to B[a]P activation of the aryl hydrocarbon receptor (AhR). Others have shown AhR-independent DNA adduct formation, raising questions about other non-CYP1A1- and CYP1A2-mediated B[a]P metabolism and adduct formation (Sagredo et al. 2006, Kondraganti et al. 2003).

The systematic meta-analysis started with a search for published, peer-reviewed transcriptomics data sets using B[a]P as the test substance. The Gene Expression Omnibus (GEO) and ArrayExpress databases were searched for microarray transcriptomic studies using the search terms in Table 3. The search focused on GEO and ArrayExpress as these databases store submitted data as raw data. The raw data are critical for performing meta-analyses, especially when different analysis methods might be used.

[11]Translesion synthesis is a mechanism that the cell uses to continue DNA replication/synthesis in the presence of a DNA lesion (e.g., DNA adduct).

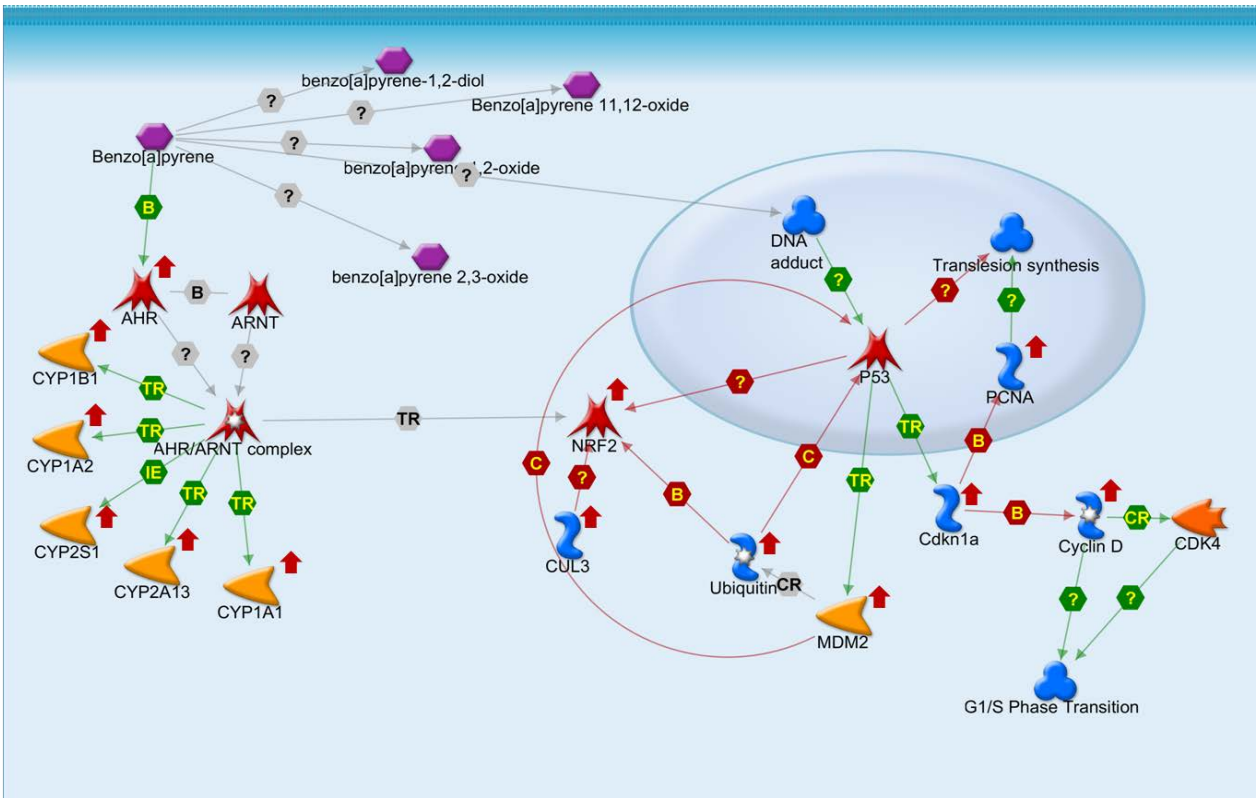


Figure 14. Consensus Outcome Pathway. This consensus pathway was synthesized by combining multiple pathway diagrams identified through analysis of the two data sets using GeneGo Metacore. The nodes (proteins or outcomes) are connected by lines. The green lines represent activation, while the red lines represent inhibition or repression. The thick red arrows near proteins represent increases in gene expression.

The search resulted in the identification of 26 peer-reviewed publications with 40 gene expression data sets. The adult mouse liver was chosen as the focus system based on the number of studies available across the species and tissues where B[a]P was used. Only 2 of the 26 publications focused on *in vivo* transcriptomic studies of the liver in the mouse. Study GSE24907 is a dose-response study where five male Muta mice (a LacZ transgenic mouse line) per group were gavaged with an olive oil vehicle and 25, 50, or 75 mg/kg B[a]P. Study GSE18789 is a time-course study where 27- to 30-day-old B6C3F1 mice were gavaged with 150 mg/kg B[a]P for 3 days and sacrificed at 4 or 24 hours after the final dose.

Table 2. Altered Genes/Functions and Their Relationship to Cancer (in this Model)

Altered Gene or Function	Relationship to Cancer in this Model
AhR/ARNT Complex	AhR regulated expression of several CYPs, including CYP1A1 and CYP1A2
CYPs (e.g., CYP1A1, CYP1A2)	Upregulation leads to production of oxidative radicals and B[a]P metabolites
NRF2	Regulates the expression of oxidative stress-protective genes
Ubiquitin	Protein that tags other proteins for destruction
CUL3	Regulates the inhibition of NRF2 signaling with ubiquitin
p53	Stops cell cycle by preventing G1/S phase transition; activated by DNA damage
MDM2	Regulates p53 through negative feedback mechanism with ubiquitin
Cdkn1a/p21	Upregulated by p53 activation; inhibits Cyclin D activation and prevents G1/S phase transition
Cyclin D	Activates G1/S phase transition, works with CDK4
CDK4	Activates G1/S phase transition, works with Cyclin D
G1/S Phase Transition	Starts cell cycle progression by allowing for DNA synthesis
Translesion Synthesis	DNA damage tolerance mechanism; allows DNA replication fork to progress beyond DNA damage sites
DNA Adduct	A piece of DNA covalently bound to a chemical that can modify expression of DNA

Table 3. Search Terms and the Number of Studies Retrieved from the Gene Expression Omnibus (GEO) and Array Express Microarray Repositories

Search Term	Number of Microarray Studies Retrieved
Coal tar	2
Polycyclic aromatic hydrocarbons or PAHs	13
Diesel	11
Smoke (NOT cigarette smoke)	16
Benzo[a]pyrene or B[a]P	53
Fuel oil	1
Cigarette smoke	63
Tobacco smoke	16

The Systematic Omics Analysis Review (SOAR) Tool was used to document and facilitate the evaluation of both studies (McConnell and Bell 2013). SOAR consists of 35 objective questions that help users determine if a study contains data of sufficient quality for use in a risk assessment context. SOAR was developed by toxicology and toxicogenomics experts, and based, in large part, on existing and published data standards such as the Minimum Information About a Microarray Experiment (MIAME) standard. Both studies (GSE24907 and GSE18789) met the SOAR screening threshold. Following a more in-depth scientific review, both studies were found to be of sufficient quality for use.

That DEG lists reported in the peer-reviewed literature are not reproducible across similar studies is well established (Shi et al. 2008, Chuang et al. 2007, Ein-Dor et al. 2005, Lossos et al. 2004, Fortunel et al. 2003). In one published example, three different studies aimed at identifying "stemness" genes[12] each yielded 230, 283, and 385 active genes, yet the overlap between them was only 1 gene (Fortunel et al. 2003). Therefore, a pathway-based meta-analysis approach was used, whereby fold change-based ranking, or more formal meta-analyses relying on raw data, along with a standardized analysis approach are considered to be more reproducible than published DEGs (Ramasamy et al. 2008, Shi et al. 2008, Chuang et al. 2007).

Both studies were reanalyzed independently at the feature level[13] using the same pre-processing, normalization, and analysis methods. GeneGo Metacore was used to identify pathways representing a large number of genes from both data sets.

The consensus systems model (Figure 14) was synthesized based on the results from GeneGo Metacore. The model conceptually describes the events that might occur when B[a]P enters the cell. Briefly, B[a]P binds to AhR, leading to upregulation of xenobiotic metabolizing enzymes and Nrf2, which might lead to additional B[a]P metabolism to epoxides and increased oxidative stress. B[a]P-mediated genotoxicity, evidenced by DNA adducts, will occur and will activate p53. Although Nrf2 is upregulated transcriptionally, p53 is expected to interfere with Nrf2 signaling, ensuring a pro-oxidant environment, which might perpetuate further DNA adduct formation. Upregulation of p21 (Cdkn1a) and MDM2 are most likely a result of p53. Upregulation of ubiquitin, while in the presence of p53-mediated MDM2 upregulation, is expected to destabilize p53. Destabilization of p53, in the presence of PCNA, is expected to allow translesion synthesis, which will allow mutations and adducts to perpetuate through DNA synthesis. Upregulation of Cyclin D could be sufficient to overcome p21 inhibitory competition, especially as p53 levels decrease, allowing for G1/S phase transition to occur. Thus, G1/S phase transition, combined with translesion synthesis, is expected to lead to propagation of mutations and DNA adducts into daughter cells. This loop might continue into a feed-forward situation until p53 signaling can be reinitiated.

[12]"Stemness" genes are those genes that are hypothesized to confer stem cell characteristics.

[13]A common misconception about microarrays is that they measure gene expression at the level of a gene. In reality, microarrays measure only a portion of a gene, typically anywhere from 20 to 100 nucleotide bases. This portion of the gene that is actually measured is called a "feature." Typically, only one feature exists per gene on a microarray. Some genes are represented more than once on a microarray, however, complicating downstream analyses (e.g., deciding how much a gene is expressed when the two features representing different parts of the same gene yield different numbers). Features could also be believed to map to a specific gene at one time, and the feature is later discovered to map to a completely different gene (this happens more frequently with lesser known or studied genes and lesser known or studied organisms where the genome might not be available). Thus, the gene associated with a feature can change over time, and most analysts will re-map their feature sequences against the genome periodically to ensure they have the latest annotation. This might result in reproducibility issues when comparing to studies performed at different times. Generally, when interpreting gene expression, analysts prefer to operate at the feature level for all analyses.

Using the gene expression changes and activating DNA adduct formation, the Boolean Network systems model (Figure 15-17)[14] predicts that cell cycle progression will be activated with translesion synthesis[15] (Figure 18). These data and the systems model support the notion that the high doses and acute durations used in the two mouse liver studies might initiate liver tumor progression through a genotoxic MOA, and promotion might occur through a cellular proliferation MOA. Due to the lack of data, speculating whether this system could be activated at low doses in the mouse is not possible. Due to genetic and epigenetic variability and potential species differences, these types of effects might occur at lower doses in humans than in mice.

The proposed model, however, provides a testable hypothesis for effects at lower doses, with other species, and other PAHs. For instance, transcriptomic studies with PAH mixtures, or other PAHs individually, can be analyzed to see if they might also impinge on this pathway. Further, the gene expression data from these other studies can be placed into this model, and an analysis can be performed to see how the cell might react, compared to B[a]P. This will give an indication of doses/exposures that could lead to DNA damage, activation of translesion synthesis, and G1/S-phase transition.

Human Susceptibility and Population Variability

Variations in human genetics will alter the susceptibility and population variability with respect to the tumorigenesis or carcinogenesis outcomes. For instance, SNPs are known to occur in p53, which might impact its ability to stop G1/S phase transition. In addition, the p53 gene has been shown to be mutated in many cancers (Vogelstein et al. 2000). A data mining approach can be taken to identify other relevant SNPs for the genes or proteins in the systems model.

---

[14]In a Boolean Network model, the system is represented as a series of connected nodes. Each node represents a gene/protein, and a connection represents some type of action/inhibition relationship. The connections are directed. For instance, p21 inhibits Cdk4, so the arrow originates at p21 and terminates at Cdk4. Some of the relationships are not as direct. For instance, Cyclin D interacts with Cdk4 to activate G1/S phase transition; however, in the model, this is best represented as a positive interaction between Cyclin D and Cdk4 given the relationship between Cdk4, Cyclin D, and p21. Each node has a state, either on (1) or off (0). Based on the state and the relationship to the other nodes, the Boolean Network can cycle through a series of states. To test the predicted outcomes (i.e., can the model sustain cell cycle progression and translesion synthesis once initiated?), this model was further simplified into just the DNA adduct/cellular proliferation part, and represented as a Boolean Network systems model. Specifically, we are looking for stable states or attractors—cycles of states that recur and self-perpetuate. States that lead to attractors are called the basin. The Boolean Network in Figure 15 has a single state attractor defined in Figure 16. This state can be defined as a cell cycle progression state with translesion synthesis turned on. If the cell were to enter this system state, it would be expected to self-perpetuate until a stimulus shuts it down. Important to note is that the systems model does not predict that all cells will enter this state or that this state is the default. Rather, the model is simply stating that if this state were entered, the cell would remain in this state until a stimulus occurs that forces it out. Such stimuli might include changes in gene expression, alterations of metabolic states, or a change in overall energy level. The Boolean Network model predicts that, with DNA adducts alone, the cell will enter into a five-state attractor (Figure 17). In this cycle, the cell is not predicted to enter into G1/S phase transition—which is expected because p53 should effectively shut down that pathway. Translesion synthesis is predicted to occur in this attractor cycle.

[15]Translesion synthesis is a mechanism for DNA damage tolerance that allows the DNA replication machinery to move beyond a DNA lesion or abasic site (i.e., a site that lacks a DNA base).

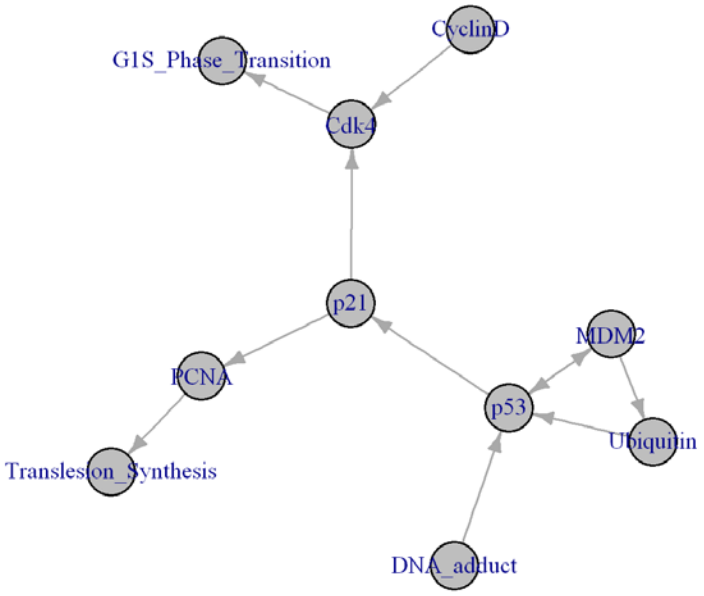


Figure 15. Liver Carcinogenesis Systems Model. The nodes represent proteins, and the lines are directional connections meaning activation or inhibition (activation and inhibition are not treated differently in the graphical depiction of the model). For instance, the arrow from PCNA to translesion synthesis means that PCNA activates translesion synthesis. The two major outcomes in this model are translesion synthesis and G1/S phase transition. The major external input is DNA adduct formation. DNA adducts cause structural damage to the DNA, which could become or lead to mutations and ultimately tumorigenesis and cancer.

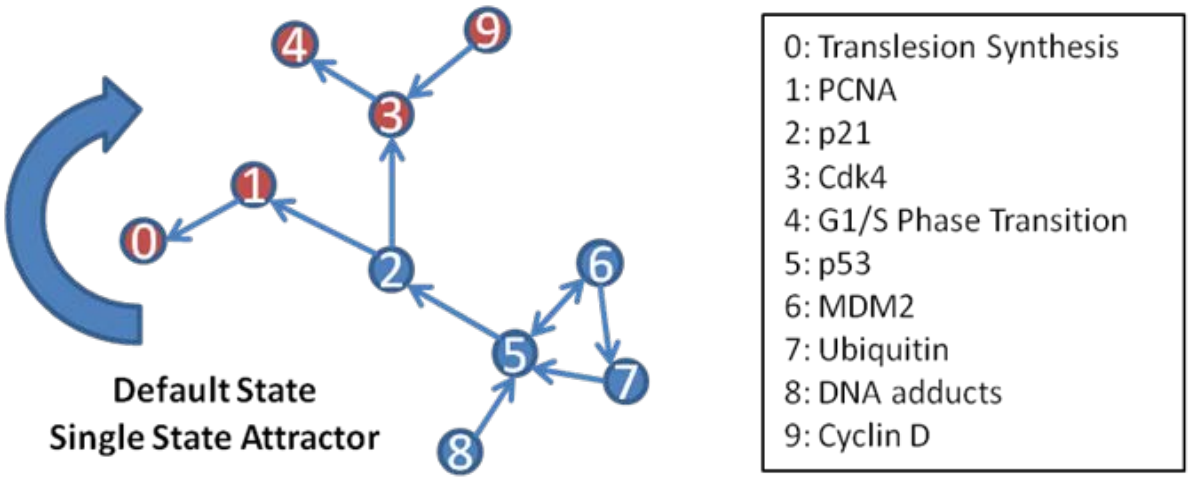


Figure 16. Default State, Single State Attractor**.** The systems model falls into a default state, single state attractor system. This is the same as the network represented in Figure 15. The names have been replaced by numbers, which are noted in the figure legend. Red nodes are those that are activated. Blue nodes are inactivated. The system here has not been perturbed by external forces. Of particular interest is that the "default" state for the system is one where the cell is actively proliferating and undergoing translesion synthesis.

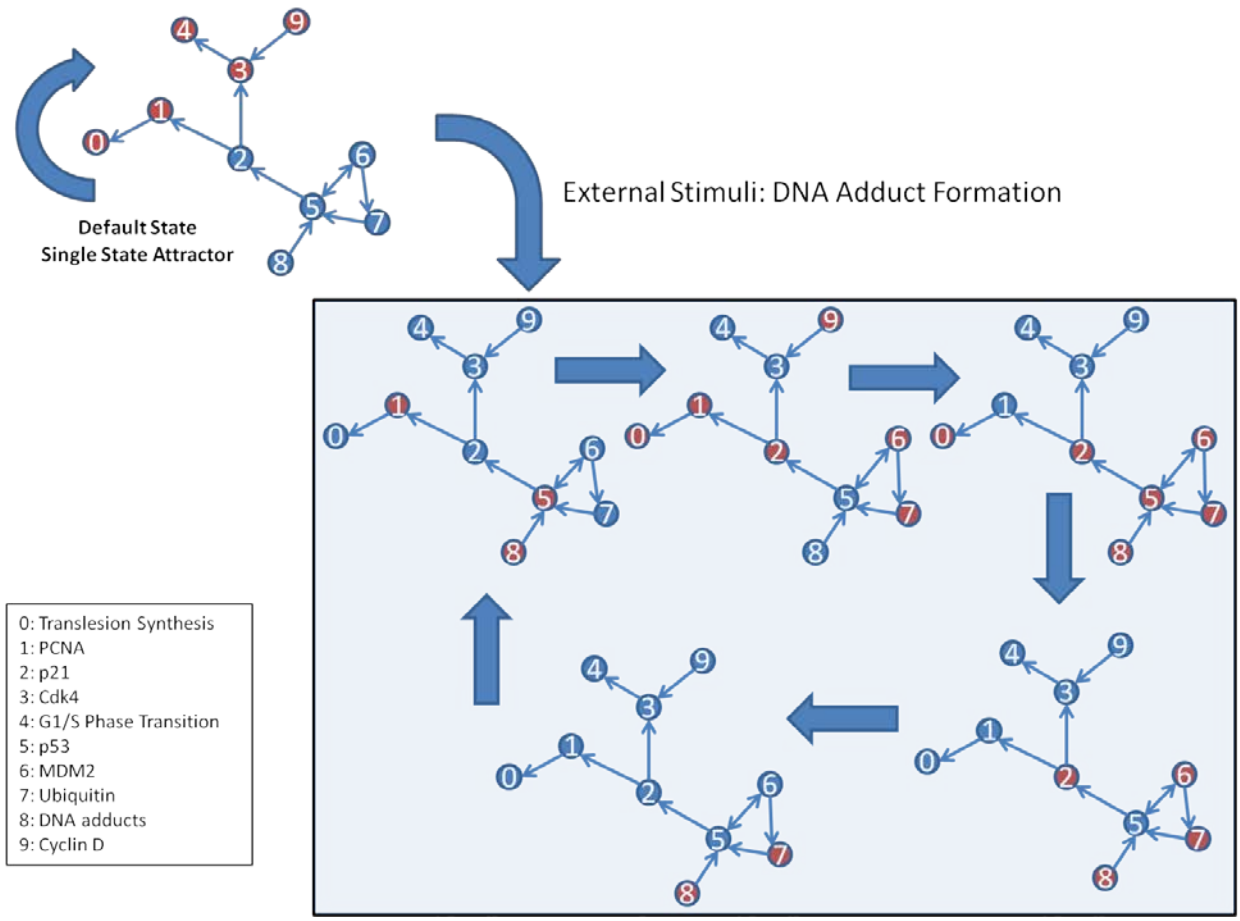


Figure 17. DNA Adduct Attractor System. When the systems model is perturbed through an external stimulus (DNA adduct formation), it transitions from the default stable starting state and moves to a new attractor (depicted in the inset). Once the system moves out of the basin for the default state attractor, it cannot return to that state without another significant stimulus. This multistability (the fact that a system can have multiple stable attractor states) is a characteristic of complex systems. Starting at the upper left of the inset, PCNA is activated, DNA adducts are activated, and p53 is activated. This leads to translesion synthesis and activation of p21, MDM2, and ubiquitin. Although Cyclin D gets activated, there is no activation of G1/S phase transition. The system then transitions to a state where translesion synthesis is primed and ready to go. If G1/S phase transition were to occur, p53 is activated, along with DNA adduct formation, MDM2, and ubiquitin. The next system state has continued p21 activation, loss of p53 activity presumably through ubiquitin and MDM2 activation in the prior system state, and DNA adduct formation. The system then transitions to only DNA adduct formation and ubiquitin activation, followed by restarting of the cycle.

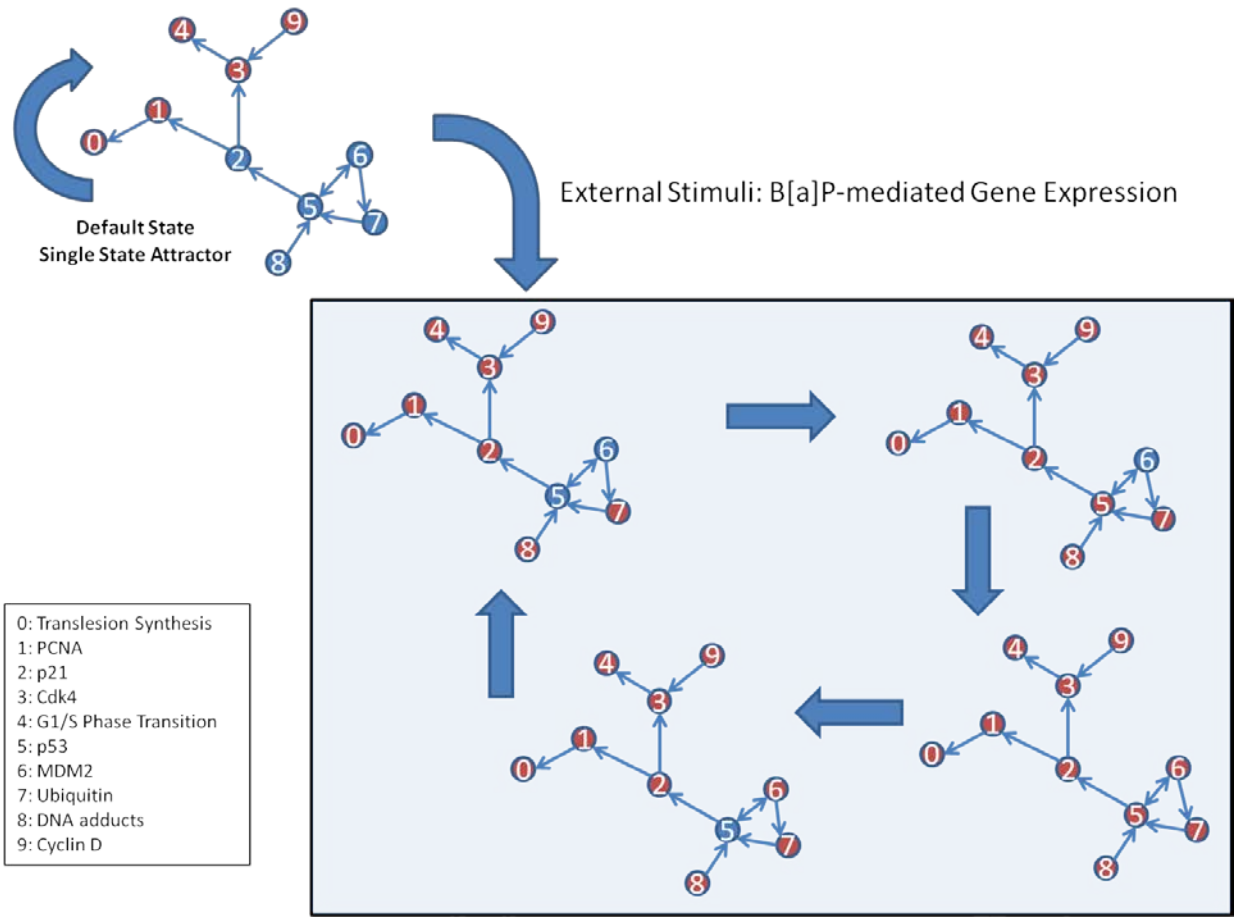


Figure 18. Gene Expression Data Attractor System. This four-system attractor is based on the gene expression data observed in both studies. This attractor system is notable as it shows DNA adduct formation, translesion synthesis, and G1/S phase transition occurring in all system states. This model predicts that DNA adducts and potential mutations are being passed forward to daughter cells through translesion synthesis as the cell cycle progresses at these doses and times in the mouse liver. This suggests that B[a]P at these doses and experimental time-points post exposure in the mouse liver could be an initiator and promoter of tumorigenesis. This adverse outcome pathway (AOP) might ultimately result in carcinogenesis.

Risk Assessment Implications Based on the B[a]P Prototype

Hazard Identification – These data suggest that B[a]P activates known human disease pathways associated with genotoxicity and tumor promotion/cell cycle progression. Similar pathway-based meta-analyses can be performed on transcriptomic data for other chemicals to screen for genotoxicity and tumor promotion, prior to the observation of tumors. For instance, using this specific Boolean systems model would inform risk assessors of the likelihood that other PAHs and PAH mixtures share a similar AOP. This type of chemical screening would need to be further validated with known or likely carcinogens and compared against chemicals that are believed not to be carcinogens (to establish performance of the screening method).

Disease-focused system models could be developed for a larger set of complex human diseases to expand the utility of this approach in the future. The pathway-based, diseased-focused, Boolean

systems model approach could be expanded to include emerging data streams, including metabolomics and proteomics, to create overall improvements in mechanistic understanding and hazard identification screens.

The genes in these Boolean systems models can be considered as those that might be tested together in a battery of assays to be used in Tox21 screening. HTS assay batteries based on these models can be implemented easily using current multiplex quantitative PCR assay systems.

Exposure-Dose-Response Assessment – Analyzing changes in the systems model and potential differences in adverse outcome across a range of doses was precluded due to the lack of sufficient dose-response data. The Boolean systems models approach used here, however, would allow for the prediction of adverse outcomes across a range of doses. In the B[a]P example, we examined the impacts of different scenarios. This same approach would be used to analyze different doses.

Cumulative Risk Assessment – Boolean systems models can be used to compare and integrate pathway-based results from multiple chemicals and nonchemical stressors. This approach would enable prediction of hazards from exposure to mixtures or cumulative stressors.

Variability and Susceptibility in Human Response – Human susceptibility can be modeled by using data from genome-wide association studies (GWAS), knock-out studies, or knock-down studies. In this instance, modeling the impacts on the adverse outcome predicted by the Boolean systems model is possible. For instance, the impacts of a gene knock-out generally can be modeled in the Boolean systems model as a constant inactivation of the protein.

Population variability would be modeled using a Monte Carlo simulation to estimate the risk of adverse outcomes across different genetic profiles. This would be accomplished by using the same types of models as in the human susceptibility context. The population variability scenario can be considered as creating a population of susceptibility Boolean systems models, where each model has a chance of being included in the overall analysis equal to its occurrence in the human population (or equal to its occurrence in a hypothesized human population if performing a what-if type of scenario). For instance, if 15% of the population is expected to have a loss of function polymorphism, the Monte Carlo model should have a 15% chance of choosing that type of Boolean systems model on each random draw from the population.

3.1.4. Risk Assessment Implications across the Tier 3 Prototypes

Looking across the Tier 3 prototypes:

- Benzene, ozone, and B[a]P displayed human molecular signatures that are strongly associated with specific human disorders and diseases.
- This type of molecular mechanistic understanding can be used to screen and predict an association between a chemical and a disease, or to augment the existing weight of evidence for an association between a chemical and a disease.
- With sufficient systems biology understanding and data, disease signatures also could be used to screen chemicals with no or limited traditional data for specific disease hazards.

- Meta-analyses that integrate pathway-based data across multiple studies yield the greatest evidence that associate chemical exposure to a disease, and are generally the most appropriate method for using transcriptomics data in a risk assessment. A pathway analysis from a single study will yield more evidence to associate a chemical exposure to a disease, and assuming the study design is adequate, might be appropriate for a risk assessment. An analysis built on a set of DEG lists is not reproducible or adequate for risk assessment purposes.
- Dynamic disease-based systems models will facilitate the understanding and prediction of chemical-disease associations in the near future. These models provide a nonbiased view of the underlying biology, and can facilitate making pathway-based predictions of adverse outcomes and disease when the interconnections within the pathway become complicated (e.g., the B[a]P case study).
- On an individual level, molecular signatures involve dynamic relationships among adaptive and nonadaptive processes that will require additional research to understand fully. At the population level, environmental factors can be thought of as shifting the population or subpopulation distributions toward (e.g., certain chemical exposures) or away from increased levels of risk (e.g., beneficial nutrients).
- *In vitro* responses appear to have commonalities with *in vivo* responses but also are affected by a number of variables, such as test system, metabolism, cell type, tissue type, time course of events (ozone data only), individual characteristics (intrinsic and extrinsic), and species.[16] These complexities make the identification of a specific disease hazard from *in vitro* only data difficult. Systems biology understanding, derived from *in vivo* data, increases confidence in the interpretation of *in vitro* data.
- For *in vitro* data, identifying hazards that occur at the organ or organismal level might be difficult. Thus, *in vitro* studies might be more appropriate for assessing the relative potencies of chemicals to alter biological processes (vs. induce disease) or to predict hazards that occur or are initiated at the tissue level (e.g., generalized inflammatory response). This is particularly true if relative potency is evaluated within a given protocol.
- Future research merging GWAS data and personalized medicine into organized data can help better characterize both intrinsic and extrinsic factors that contribute to human variability and susceptibility.
- The networks associated with a disease can apparently be disrupted in multiple places, all leading to altered risks of the specific disease. This is shown by mechanistic commonalities among diseases of unknown origins, other chemicals associated with the disease, and chemotherapeutics that can reverse or block components of the disease processes. This type of information can be a useful tool in characterizing cumulative risks. Overly narrow descriptions of mechanisms can miss interactions among environmental factors.

[16]Although not evaluated here, lifestage is also an important variable (Boekelheide et al. 2012).

- When searching for candidate Tier 3 prototypes, one important observation was that, even among the most well studied chemicals, very few chemicals had the type and quality of data needed for exploring the use of new data types in risk assessment. There are needs for systematic review criteria for new data types, adherence to standards of experimental and statistical practices in data generation and analyses, and thoughtful consideration of variability and uncertainty to improve the utility of new data types for risk assessment.

3.2. Tier 2: Limited Scope Assessments

The intent of the Tier 2 prototypes is to (1) explore new types of computational analyses and short-duration *in vivo* bioassay that are currently relatively uncommon in risk assessment but hold great promise for the near future; and (2) develop an assessment approach well suited to limited scope risk management decisions. In this case, "limited" generally means regional to local exposure potential, or limited hazard potential, or limited data to conduct more detailed assessments. Tier 2 efforts fall between Tier 3 and Tier 1 in terms of resources required and amount of uncertainty in the assessment results. The number of chemicals possibly identified in Tier 1 meriting further testing could overwhelm traditional or Tier 3 type evaluation, thus the need for an intermediate testing and assessment strategy as provided by Tier 2 (Thomas, RS et al. 2013a).

The hallmark of Tier 2 data in the NexGen program is integration across biological systems—molecule-to-cell(s)-to-tissue(s) and, in some systems, to-outcome(s)—to inform associations among environmental exposures, causal mechanisms, and outcomes, but generally using evaluations over relatively short time periods (hours to weeks). Tier 2 considers all information available from Tier 1 approaches, such as quantitative structure-activity relationship (QSAR) and HTS, along with other data derived from more complicated test systems that use intact tissues or organisms to provide a higher level of confidence in the assessment (Table 4). Limited scope assessment could include combining HTS with limited traditional data. Tier 2 data are commonly referred to as high-content data.

Table 4. Summary of Tier 2 NexGen Approaches, Including Weight of Evidence, Pros, and Cons

	Tier 2: Limited Scope Assessments Categories of Approaches Considered		
	Data Mining of Existing Databases	**Alternative Species *In Vivo* Assays**	**Mammalian Short-duration *In Vivo* Assays**
Approach:	Discovers or identifies associations among environmental exposures, omic patterns, and human disease. Often uses meta-analyses of large existing data sets. Suggests potential adverse outcomes based on existing knowledge of other chemical-induced molecular event and disease relationships.	Experimentally measure dose-dependent, chemically-induced alterations in biological functions in intact organisms using a range of specific and sensitive assays. Measures adverse outcomes that range from omics to phenotypic outcomes and population effects.	Experimentally measure dose-dependent, chemically-induced alterations in biological functions in intact animals using a range of specific and sensitive assays. Measures molecular or cellular changes; infers potential adverse outcomes based on existing knowledge of other chemical pathway or disease relationships.
Weight of evidence:	Determined by the quality and amount of underlying evidence (ranges from suggestive to likely) or is known with substantial complementary experimental data.	Determined by the quality and quantity of data, but generally suggestive to likely. Cross-species issues need consideration.	Determined by the quality and amount of underlying evidence, ranges from suggestive to likely when anchored to pathway and traditional data and some understanding of temporal progression.
Pros:	Significantly faster and less expensive than traditional bioassays. Can use combined data sets that include tens of thousands of humans. Includes tissue, organism, and life span-level integration, including metabolism	Significantly faster and less expensive than traditional bioassays. Can evaluate complex outcome birth defects and neurobehavioral outcomes.	Significantly faster and less expensive than traditional bioassays. Includes tissue and organism integration, including metabolism.
Cons:	Relationships generally associative; might be causal in certain circumstances (depending on data quality and amount of underlying evidence). Data on effects of early life exposures and effects generally lacking.	Species-to-species extrapolation is an issue as is the potential for parent compound not to be metabolized to toxicants that are active in humans. Omics information can be derived from organs, tissues, and multiple cell types versus only human-based target cells. Data on effects of early life exposures and effects generally lacking; an exception is the embryonic fish models.	Measure events early in disease initiation process; early events could be reversible; links to apical outcome can be an issue. Omic information is often derived from multiple cell types versus only target cells. Data on effects of early life exposures and effects generally lacking.

Two general approaches to Tier 2 data are discussed here:

- High-content knowledge mining (i.e., computer-driven surveys of the literature and large existing data libraries[17]) to retrieve data and conduct meta-analyses of existing systems biology data to construct mechanism-of-action models and establish associations between environmental exposure and disease. The diabetes/obesity prototype is provided as an example.
- Short-term *in vivo* or *in situ* exposures of intact organisms to enable incorporation of the intact metabolism in the toxicity evaluation and produce measures of biological change over a short time frame (i.e., ranging from hours to a few months) that are thought to be relevant to longer term outcomes. Two examples are provided using alternative and mammalian species. Considerable work is ongoing at various U.S. Federal Government agencies and elsewhere to refine assays where animals are exposed to chemicals *in vivo* for periods ranging from hours to a few months.

Implications for risk assessment identified by the Tier 2 prototypes are discussed at the end of this section and integrated with other lessons learned in Section 5 "Lessons Learned from Developing the Prototypes."

3.2.1. Knowledge Mining – Diabetes/Obesity

Knowledge mining[18] is explored in this prototype as a means to characterize the associative and potentially causal relationships among disease and exposures to environmental factors and intrinsic sources of human variability. The knowledge mining approach capitalizes on huge new databases that are being supplemented with each publication in the field of omics (>50,000 per year). These databases are generally oriented toward the omics of human disease but also include omics information on other species, as well as surveys and clinical assays measuring human exposure and health outcomes. The specific, related diseases explored here are diabetes and obesity and relationships to multiple environmental factors. Diabetes results from environmental and genetic factors and risk varies considerably in the population (Patel et al. 2013). Four interrelated efforts focusing on diabetes/obesity are reported here: (1) Comparison of Knowledge Mining Results and Expert Opinion; (2) Environment-wide Association Studies (EWAS); (3) Itemset Associations between Prediabetes/Diabetes and Chemical Exposures; and (4) Characterizing Human Susceptibility and Population Variability.

Comparison of Knowledge Mining Results and Expert Opinion

Thayer et al. (2012) reported on a recent National Toxicology Program (NTP) workshop that examined the possible causal relationships between environmental exposures and diabetes or obesity. At the workshop, results from an extensive information survey were evaluated by experts

[17]For example, the National Library of Medicine's Gene Expression Omnibus (GEO): a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.

[18]**Knowledge mining** is the computerized extraction of useful, often previously unknown, information from large databases or data sets using sophisticated data search capabilities and statistical algorithms to discover patterns and correlations and then interpret this new information in the context of systems biology to create new knowledge.

on the strength of the associations identified. The effort integrated both traditional and new types of data, including approximately 870 findings from more than 200 human studies; and the most useful and relevant endpoints in experimental animals and *in vitro* assays (e.g., ToxCast™ and Tox21). The environmental factors identified and discussed at the workshop included maternal smoking and nicotine, arsenic, persistent organic pollutants, organotins, phthalates, bisphenol A (BPA), and pesticides. Overall, the workshop results suggest that associations can be made between environmental factors and type 2 diabetes or obesity, but causality was more difficult to assign (Table 5).

Table 5. Summary of Literature Review Findings and Expert Judgments Concerning Causal Relationships

Chemical/ Environmental Factor	Outcome	Association/ Causality	Conclusions from Breakout Group
Maternal smoking and nicotine	Childhood obesity	Association, likely causal	Likely causal supported by epidemiology data and animal studies (Behl et al. 2013).
Arsenic	Diabetes	Association	Sufficient support for an association between arsenic and diabetes in populations with relatively high exposure levels (≥ 150 µg arsenic/L in drinking water) (Maull et al. 2012).
Organochlorine persistent organic pollutants	Diabetes	Association	Sufficient for a positive association of some organochlorine persistent organic pollutants with type 2 diabetes (Taylor et al. 2013).
Organotins	Obesity	Suggestive of an association in animal and *in vitro* models	Current data from human studies of exposure to organotins are nonexistent regarding an association with diabetes or obesity. Recent animal and mechanistic studies report stimulatory effects of tributyl tin on adipocyte differentiation (*in vitro* and *in vivo*) and an increased amount of fat tissue (i.e., larger epididymal fat pads) in adult animals exposed to TBT during fetal life. Although the organotin "obesogen" literature is relatively new, with few studies, the quality of the existing experimental studies was considered high by the breakout group (Thayer et al. 2012).
Bisphenol A (BPA)	Diabetes	Suggestive of an association	Overall, this breakout group concluded that the existing data, primarily based on animal and *in vitro* studies, are suggestive of an effect of BPA on glucose homeostasis, insulin release, cellular signaling in pancreatic β cells, and adipogenesis (Thayer et al. 2012).
Phthalates	Diabetes or obesity	Insufficient data to assess	Current data from human studies of exposure to phthalates provide insufficient evidence of an association with diabetes or obesity (Thayer et al. 2012).

Environment-Wide Association Studies[19]

Diabetes varies in the population due to both genetic and environmental factors but understanding these interactions has been difficult. Using an Environment-wide Association Study approach, Patel et al. (2012b) investigated the problem of many possible contributing factors by integrating genomic and toxicological data to obtain a candidate list of interacting genes, genetic variants, and environmental factors associated with type 2 diabetes. The method involved three steps. First, genetic and environmental data were summarized from VARIMED (VARiants Informing MEDicine; a genetic association database) and the National Health and Nutrition Examination Survey (NHANES, an environmental exposure and effects database). VARIMED contains data on 11,977 gene variants, 9,752 genes, and 2,053 individuals; NHANES includes 261 genotyped loci, 266 environmental factors measured in blood and urine, and clinical measures for the same individuals. They identified several environmental factors that positively or negatively affected risks for type 2 diabetes, including nutrients and persistent organic pollutants. They reported 18 human genetic variations (SNPs) and 5 serum-based environmental factors that interacted in association with type 2 diabetes. Thus Patel et al. (2013, 2012b) successfully identified association linking diabetes, genes, gene variants, and environmental factors.

This approach demonstrates a knowledge mining method that can be applied broadly to any number of common diseases to identify possible interactions between genetic and environmental factors and risks of disease. In *Genetic Variability in Molecular Response to Chemical Exposure*, Patel and Cullen (2012) review what has been learned to date with these types of efforts and discuss a more comprehensive representation of chemical exposures as the "envirome" and how we might use it to examine the interplay of genetics and the environment.

Itemset Associations between Prediabetes/Diabetes and Chemical Exposures

We followed up efforts by Thayer et al. (2012) and Patel et al. (2013, 2012b), using two independent frequent itemset mining analyses of the NHANES data. Frequent itemset mining is a data mining approach commonly used in business intelligence to derive marketing and pricing strategies or to identify credit risks. For example, grocery stores use frequent itemset mining to uncover products that are typically purchased together to determine pricing strategies (e.g., a grocer does not want to place items commonly purchased together on sale at the same time and might raise the price of an item commonly purchased with a sale item). Similarly, this technique can be used with the NHANES data to uncover a chemical or group of chemicals that tend to be associated with specific diseases.

We focused our analyses on the 2003–2004 NHANES cohort and evaluated associations between diabetes and individual chemicals. We also focused on the 2009–2010 NHANES cohort and evaluated associations among diabetes and a more complex lists of chemicals.[20] Both analyses focused on metals.

[19]This section is adapted largely from Patel et al. (2012b) and (2013) with the assistance of Dr. Patel.
[20]Both analyses use the Apriori algorithm (Borgelt 2013) to generate "rules" where X ≥ Y is read "X is associated with outcome Y." Our first study constrained the rule to read "prediabetes/diabetes is associated with chemical Y," or prediabetes/diabetes ≥ chemical Y. Our second study constrained Apriori to return prediabetes/diabetes as the outcome.

Prediabetes/Diabetes and Individual Chemical Exposures – These results suggest that type 2 prediabetes/diabetes is most often associated with lead and cadmium (blood or urine), with a suggestive association with arsenicals. Type 2 prediabetes/diabetes is not associated with cesium and uranium. Table 6 lists the resulting rules[21] showing the association or lack of association between diabetes and all of the metals monitored in NHANES.

When interpreting lift,[22] support,[23] and confidence,[24] we believe lift is the most informative to start with, followed by the other measures. If a rule has a lift value close to 1, the rule has little predictive value, regardless of the support and confidence. Once an analyst has identified models that are significantly different from random (lifts > 1), the analyst will typically then examine the support and confidence.

Support provides an indication of the percentage of people surveyed by the NHANES program that have both type 2 prediabetes/diabetes (the antecedent) and high blood lead, for instance (11% in this case). The support indicates what proportion of the population might be expected to have type 2 prediabetes/diabetes and high blood lead, assuming the NHANES sample is a truly representative sample of the U.S. population (in this case 11%).

The confidence tells the analyst how strong the rule is. In other words, confidence tells the analyst the percentage of people with type 2 prediabetes/diabetes (the antecedent) that have high blood lead, for instance (34% in this case). This is equivalent to the prevalence of Type 2 prediabetes/diabetes in individuals that have a high level of the particular metal, and is a potential indicator of risk.

[21]Ruleset is a collection of one or more rules used, in this case, to predict association between diabetes and chemical exposures (Oracle 2013a).

[22]Lift is a measure of how much better prediction results are using a model than could be obtained by chance (Oracle 2013b). A lift of 1 means the rule is no better at predicting the outcome than random chance. Thus, knowing that someone in this NHANES cohort is defined as prediabetic or diabetic provides a 1.44 times better chance to predict that the person has high blood lead, compared to random. The lift close to 1 provides no better indication of a person's urine uranium or cesium concentration compared to random guessing knowing that they are prediabetic or diabetic.

[23]Support is the percentage of subjects in the entire data set/database that have both the antecedent/condition and the predicted outcome. This can also be thought of as the number of subjects in the entire data set/database where the rule is true.

[24] Confidence is the percentage of the people who meet the antecedent/condition criteria that also meet the outcome criteria. For instance, 34% of the people in this NHANES cohort defined as being either prediabetic or diabetic also have high blood lead.

Table 6. Apriori Rule Associations between Type 2 Prediabetes/Diabetes and Chemical Exposures.

Antecedent/Condition	Predicted Outcome	Lift	Support	Confidence	Conclusion
Type 2 Prediabetes/Diabetes	High blood lead	1.44	0.11	0.34	Association
Type 2 Prediabetes/Diabetes	High urine cadmium	1.43	0.13	0.43	Association
Type 2 Prediabetes/Diabetes	High blood cadmium	1.26	0.09	0.30	Association
Type 2 Prediabetes/Diabetes	High urine arsenobetaine	1.25	0.10	0.33	Association
Type 2 Prediabetes/Diabetes	High urine lead	1.20	0.09	0.28	Association
Type 2 Prediabetes/Diabetes	High urine total arsenic	1.18	0.09	0.31	Association
Type 2 Prediabetes/Diabetes	High blood total mercury	1.12	0.09	0.30	Association
Type 2 Prediabetes/Diabetes	High urine cesium	1.03	0.08	0.25	No association
Type 2 Prediabetes/Diabetes	High urine uranium	1.01	0.07	0.24	No association

Prediabetes/Diabetes and Multiple Chemical Exposures – Table 7 lists the results showing associations between multiple chemicals and prediabetes/diabetes. The rule with the highest lift (1.46 times better than random) is where NHANES subjects had high urine cadmium, high blood lead, and high total urine arsenic. This rule is present in 11% of the 2009–2010 NHANES cohort, suggesting it might be true for 11% of the U.S. population at the time of study, assuming NHANES is a good random sample. Of all the subjects who had high urine cadmium, high blood lead, and high total urine arsenic, 59% also were either prediabetic or diabetic. Not surprisingly, the rule with the next highest lift is the same as the first, except these subjects also had high urine lead levels. This rule had a support of 10% and a confidence of 58%. Overall, this analysis supports strong associations between cadmium, lead, and total urine arsenic and type 2 prediabetes/diabetes due to their frequent occurrence in the top ranked rules. Cesium and cobalt occurred less frequently and would be expected to be less strongly associated.

Synthesis of Frequent Itemset Mining Results – Lead and cadmium exposure are highly likely to be associated with type 2 prediabetes/diabetes. High lead levels occurred in 9 of 10 and cadmium in 8 of 10 of the top-ranked rules in Burgoon's data set. Further evidence is provided by the results where blood lead, blood cadmium, and urine cadmium were the highest rated outcomes based on lift. Confirmatory evidence exists that these metals also might be elevated in other diabetic populations (Afridi et al. 2008). Low-dose mixtures of lead, cadmium, and arsenic might induce oxidative stress (Fowler et al. 2004), and evidence suggests that cadmium might induce hyperglycemia in rats (Bell, RR et al. 1990).

Table 7. Apriori Rule Associations between Type 2 Prediabetes/Diabetes and Exposure to Multiple Chemicals

Antecedent/Condition	Predicted Outcome	Lift	Support	Confidence	Conclusion
High urine cadmium High blood lead High total urine arsenic	Type 2 Prediabetes/Diabetes	1.46	0.11	0.59	Association
High urine cadmium High urine lead High blood lead High total urine arsenic	Type 2 Prediabetes/Diabetes	1.44	0.10	0.58	Association
High urine cadmium Low urine cobalt	Type 2 Prediabetes/Diabetes	1.40	0.11	0.56	Association
High urine cadmium High blood lead	Type 2 Prediabetes/Diabetes	1.38	0.17	0.56	Association
High urine cadmium High urine lead High blood lead	Type 2 Prediabetes/Diabetes	1.38	0.15	0.56	Association
High urine cadmium High urine cesium High blood lead	Type 2 Prediabetes/Diabetes	1.38	0.11	0.56	Association
High urine cadmium High blood cadmium High blood lead	Type 2 Prediabetes/Diabetes	1.37	0.13	0.55	Association
High urine lead High blood lead High total urine arsenic	Type 2 Prediabetes/Diabetes	1.37	0.12	0.55	Association
High urine cesium High blood lead High total urine arsenic	Type 2 Prediabetes/Diabetes	1.37	0.10	0.55	Association
High urine cadmium High urine lead High blood cadmium High blood lead	Type 2 Prediabetes/Diabetes	1.37	0.11	0.55	Association

Taking these results together, uranium and cesium are not likely to be associated with type 2 prediabetes/diabetes. Whether mercury is likely to be associated with type 2 prediabetes/diabetes remains unclear.

Extrapolating these results to the U.S. population suggests that a large proportion (>50%) of the population with elevated lead, cadmium, and arsenic levels might have type 2 prediabetes/diabetes. Although these data are not sufficient to demonstrate causality, they do suggest that mixtures of these metals are associated with type 2 prediabetes/diabetes. Possible explanations include (1) the mixture of these chemicals might cause type 2 prediabetes/diabetes; (2) prediabetic/diabetic phenotypes might alter the absorption, distribution, metabolism, and

excretion of these metals; or (3) only one of these chemicals might be associated with type 2 prediabetes/diabetes, while the absorption, distribution, metabolism, and excretion properties of the other chemicals are impacted by the first. Evidence exists that the three metals work together to induce oxidative stress, and cadmium itself might induce hyperglycemia in rats. These results suggest that further studies should be conducted to ascertain potential causality.

Further, these results demonstrate that Frequent Itemset Mining yields fruitful results and hypotheses that can be used to identify associations between chemical body burdens and potential disease endpoints. The results also illustrate ways that data mining methods developed for other fields can be implemented to identify predictive biomarkers of exposure and health outcomes.

Example: Characterizing Human Susceptibility and Population Variability

Risk managers can begin to identify populations with genetic susceptibility to Type 2 prediabetes/diabetes in their communities by combining the frequent itemset mining results above with data mining of human genetic variability data, health outcomes, and an understanding of disease processes and chemical MOAs. Combining this information with census demographic data, geographic information systems, and exposure models will further drive the possibilities of pinpointing specific geographic susceptible populations. In this prototype, we identify a potentially susceptible population to Type 2 prediabetes/diabetes by combining the cadmium-disease association, known gene-disease associations, and knowledge of risk allele frequencies in human ethnic groups.

Recently, a combination of EWAS and GWAS was performed that examined potential interactions between SNPs (i.e., a mutation of a single nucleotide within the DNA of a gene sequence), environmental chemical levels in blood and urine, and health outcomes—specifically type 2 diabetes—using data from NHANES (Patel et al. 2013). Although support for genotype and chemical interactions was limited, interesting interactions were noted between the nonsynonymous coding SNP rs13266634 in the SLC30A8 gene and cis- and trans-beta-carotene and gamma-tocopherol.

The SNP rs13266634 has been noted as being associated with type 2 diabetes previously (Rung et al. 2009, Takeuchi et al. 2009, Timpson et al. 2009, Pare et al. 2008, Diabetes Genetics Initiative of Broad Institute of Harvard et al. 2007, Scott et al. 2007, Sladek et al. 2007, Steinthorsdottir et al. 2007, Zeggini et al. 2007). SLC30A8 is a zinc transporter found in the pancreatic beta-cell secretory vesicles. Zinc has been associated with insulin biosynthesis (Emdin et al. 1980), and chronic decreased zinc intake has been associated with an increased risk of diabetes (Miao et al. 2013). The risk allele in rs13266634 is C (Sladek et al. 2007), while the minor allele is T (NCBI 2012b).

Risk managers might find the genotype and allele frequency data in dbSNP to be helpful in understanding population variance and to help identify susceptible populations. For instance, from a random sample of 100 individuals of Mexican descent in Los Angeles, 66% were homozygous for the risk allele, 30% were heterozygous, and 4% were homozygous for the nonrisk allele (NCBI 2012b). If we assume the sampling is representative of the entire population of Mexican-descended residents of Los Angeles, then approximately 66% of these individuals might be at an increased risk of developing diabetes, independent of their body mass index (OMIM 2012). Heterozygous individuals (30% of the population) might also carry some risk and might be more affected by their zinc intake (i.e., increased zinc intake might be protective). Likewise, the heterozygous individuals might be more sensitive to other metals, chemicals, or dietary factors that might compete with zinc for absorption, or they might be more sensitive to chemicals that could interfere with zinc

metabolism, transport, and insulin biosynthesis. Given the high rate of zinc deficiency in Mexican children that is not correlated with socioeconomic status, finding zinc deficiency in children of Mexican descent living in Los Angeles might not be surprising, especially if diet plays a significant role in the deficiency (Morales-Ruan Mdel et al. 2012).

Cadmium exposure will be of concern to individuals who are homozygous or heterozygous for the risk allele. Cadmium has been shown to compete with zinc transporters and might lead to beta-cell dysfunction, lack of insulin production, and ultimately hyperglycemia (El Muayed et al. 2012). Individuals with the rs13266634 risk allele could be more sensitive to cadmium exposures than the rest of the population.

Through database mining and an understanding of the allele disease pathway and a chemical's adverse outcome pathway, we can identify potentially susceptible populations more easily. This example could be extended by examining cadmium exposure data for the Los Angeles area and using a geographic information systems approach with census data to identify potentially susceptible individuals, based on the allele probabilities. This type of predictive modeling could help advance more targeted community-level responses in the future.

3.2.2. Short-Term *In Vivo* Models – Alternative Species

Alternative species (i.e., nonmammalian species) provide *in vivo* models for identifying hazards, integrating dose-response effects, and understanding pathways and apical effects useful for assessing chemical risks to humans and to other species. The shorter life spans of alternative species enable the evaluation of toxicity over the full life span of the intact organism, facilitating study of the entire etiology of disease from the MIE to apical outcomes, including more complex phenomena such as birth defects or neurobehavioral impairment.

Alternative species studies are playing a progressively more integral role in chemical testing, hazard identification, and dose-response assessment for both human and nonhuman species (ECHA 2013b, Perkins et al. 2013, EPA 2012d, EC 2011, Schug et al. 2011, OECD 2004). Both the European Chemicals Agency (ECHA) and EPA use alternative species tests as part of required tests for endocrine disruptors (EPA 2012e, 2009a). Alternative species studies can be used for prioritization and screening or as the basis for Tier 2 type assessments.[25]

Tier 2 Prototype: Using Alternative Species to Identify Thyroid Hormone Disruption

Endocrine disruptors are chemicals that interfere with the body's endocrine system and produce adverse effects in both humans and wildlife. In a state-of-the-science review, the World Health Organization (WHO) concluded that thyroid disruption-associated neurobehavioral disorders are

[25]The types of alternative or nonmammalian species can vary widely. Considerable work in toxicology has been done with fish, but work in very simple organisms such as yeast also provides insight into cellular regulation at multiple levels that control core biological processes and enable cells to respond to genetic and environmental changes (Yeung et al. 2011). Zhu et al. developed a network reconstruction approach that simultaneously integrates different types of data, and constructs a probabilistic causal network representing complex cell regulation: endogenous metabolite concentration, RNA expression, DNA variation, DNA–protein binding, protein–metabolite interaction, and protein–protein interaction data. Causal regulators of the resulting network were identified and provide insight into the mechanisms by which variations in network interactions affect gene expression and metabolite concentrations (Zhu et al. 2012).

occurring in children, and the incidence of these disorders has increased in recent decades (WHO 2012). Normal thyroid function is essential for normal brain development, particularly during pregnancy and after birth. Additionally, thyroid hormones are crucial to inner ear and bone development, and bone remodeling and physiological functions such as metabolism (De Coster and van Larebeke 2012). Internationally agreed-upon and validated test methods for identification of endocrine disruptors capture a limited range of the known endocrine disrupting effects (Miller, MD et al. 2009). In its state-of-the-science review, WHO advised that existing testing protocols do not characterize completely all essential functions and that adverse effects "are being overlooked" (WHO 2012). Identifying environmental factors that might disrupt normal thyroid function and impact public environmental health is needed, given the key role that thyroid hormone plays for normal development and physiologic functions in all vertebrates (Woodruff and Sutton 2011, Miller, MD et al. 2009).

In regulating development, the role of the thyroid hormone is of particular toxicological interest because thyroid hormone-dependent post-embryonic development is a common feature of vertebrate ontogeny (Paris and Laudet 2008). This period of development is typically characterized by transient elevations of thyroid hormone that elicit species-specific physiological and morphogenetic responses with lasting developmental consequences. Transitions from tadpoles to juvenile frogs and body plan reorganization in flatfish are two nonmammalian examples of thyroid hormone-controlled events. Human and vertebrate post-embryonic neurodevelopment is thyroid hormone-dependent and deviations from normal thyroid hormone concentrations at critical times are associated with a variety of neurological defects and deficits (Zoeller et al. 2002). The timing (or window) of exposure is critical as the impact of thyroid hormones changes as the brain develops (Zoeller and Rovet 2004). Thyroid hormone regulation is generally essential for normal development in vertebrates, thereby establishing the basis for cross-species extrapolation of developmental risks. Several methods using alternative species have been proposed to measure these outcomes for thyroid pathways (Makris et al. 2011, Nichols et al. 2011).

A key factor in thyroid hormone-related risk assessment is the ability to examine hormone disruption and the resultant developmental disruption at higher levels of tissue organization. Results from omics technologies and other thyroid hormone toxicity assessments, such as EPA's ToxCast™ chemical screening efforts (EPA 2008), can be linked to adverse outcome data from alternative species studies. Two examples are:

1. The construction of regulatory networks using time series data in genotyped populations and integration of multiple data types (i.e., endogenous metabolite concentrations, RNA expression, DNA variation, DNA-protein binding).
2. If a chemical is identified as a potential developmental disruptor in HTS, more information on *in vivo* effects might be required to establish dose-response relationships, windows of susceptibility, potential impacts of maternal exposure on progeny, and existence of subtle impacts on behavior, learning, and memory.

Systems and Pathway Models

As discussed throughout this document, understanding of systems biology and the events leading to an adverse effect are central features for the use of molecular biology data in risk assessment. Pathway analyses are useful to inform extrapolation across species and to aid in characterizing the variability within populations through identifying and describing both MIEs and key biological

events leading to adverse outcomes. They can also help identify how human-focused screening data can inform ecological risk assessment. Although making quantitative predictions of disease risks based on today's system biology or adverse outcome models is often very difficult, pathway assessments are critical to advancing dose-response assessment.

Figure 19 illustrates an example for thyroid hormone disruption. Disruption of the thyroid pathways can occur by thyroid peroxidase inhibition, iodine uptake (sodium iodide symporter) inhibition, enhanced phase II metabolism (glucuronosyltransferases or sulfotransferases) via alterations in specific genes (CAR/PRX [constitutive androstane receptor/prename x receptor]) or receptors (AhR), enhanced cellular transport, deiodinase inhibition, and interference with thyroid receptor function. In humans, this leads to birth defects, decreased IQ, and metabolic disorders. In rats, increased TSH leads to thyroid hyperplasia and cancer.

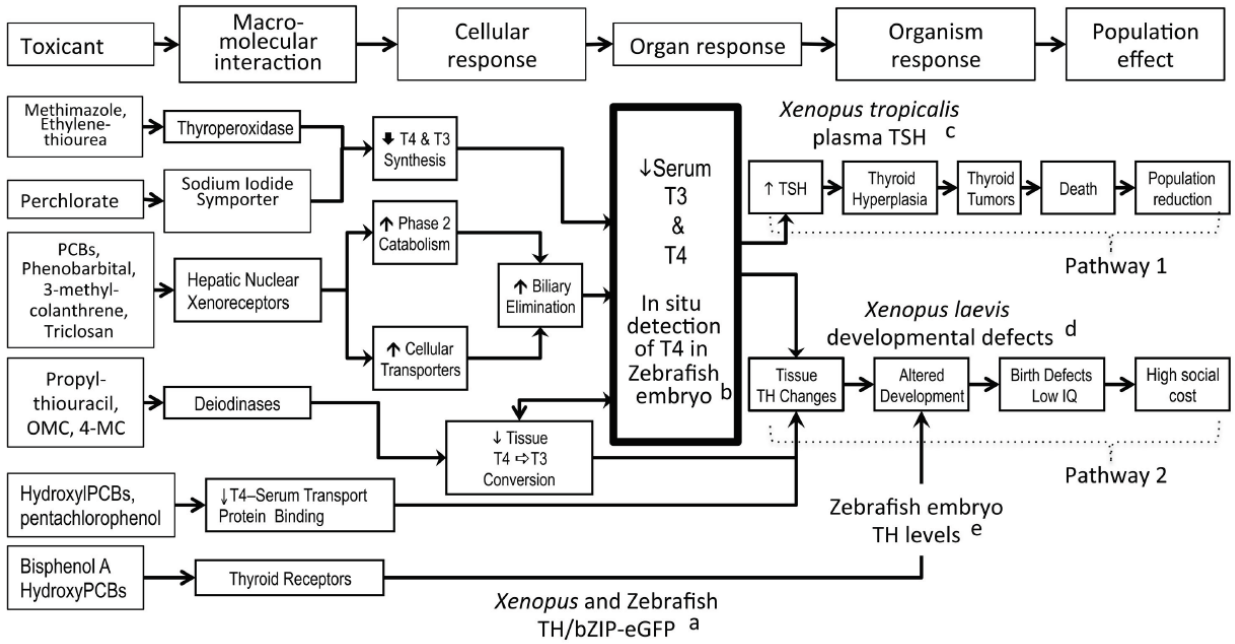


Figure 19. Major pathways involved in thyroid disruption with example toxicants and alternative models applicable to both human and ecological hazard assessment (Perkins et al. 2013). Reproduced with permission from *Environmental Health Perspectives*.[26]

[26] The thick black outlined box indicates the critical event of serum level concentrations of thyroid hormones. Pathway 1: rat pathway leading to tumors via thyroid hyperplasia. Pathway 2: principle pathway of concern affecting humans. Abbreviations: IQ, intelligence quotient; 4-MC, 4-methylbenzylidene camphor; OMC, octyl methoxycinnamate; T_3, triiodothyronine; T_4, thyroxine; TR, thyroid receptor. [a]Quantification of plasma TSH levels in *Xenopus tropicalis*. [b]Direct quantification of intrafollicular concentrations of T_4 in zebrafish embryos. [c]Detection of developmental defects with *X. laevis* metamorphosis assay. [d]Detection of developmental defects using zebrafish embryos. [e]Reporter gene (eGFP) detection of TR activity.

Informing Dose-Response Assessment

Understanding causal mechanisms and their relationships to adverse outcomes provides insights into both hazard identification and dose-response assessment. Although quantitatively predicting human disease risks based on knowledge of causal mechanisms is currently difficult, several approaches using alternative species data provide information on the potency of chemicals that cause effects: biomarkers of exposure and effect, relative potency to induce adverse effects, and understanding of the complexities of dose-response relationships.

Biomarkers of Exposure and Effects

Key events in the perturbed pathway can be represented with biomarkers of exposure and effect. In situations where considerable systems biology information links the event to outcomes, a biomarker might provide a measure of hazard for risk assessment. In the thyroid disruption example, upstream events converge on serum levels of the thyroid hormones, triiodothyronine (T3) and thyroxine (T4), and downstream events occur in peripheral tissues where a significant degree of species-specific effects are seen (Figure 20). As a result, serum T4 levels can be used as a biomarker of thyroid function across species. In the laboratory, researchers use T4 and thyroid stimulating hormone levels in fish and frogs to assess the thyroid disrupting potential of chemicals (Tietge et al. 2013, Thienpont et al. 2011). To assess human exposures, the Centers for Disease Control and Prevention (CDC) has used decreased serum levels of T4 (noted as key event in Figure 20) and increased levels of TSH measured in the U.S. population to infer increased potential risks for thyroid dysfunction-related disorders at low levels of perchlorate exposures (Lau et al. 2013, Blount et al. 2007).

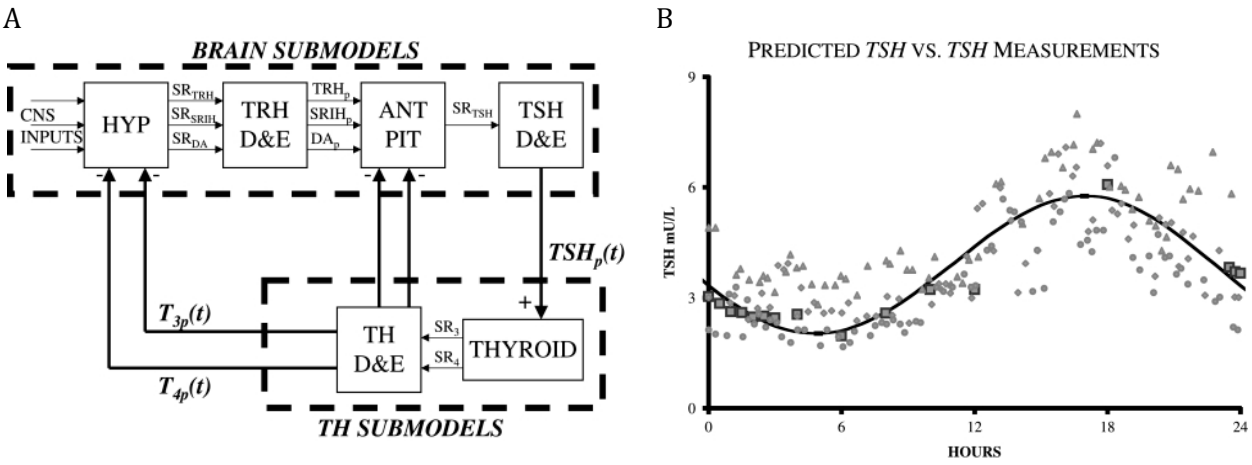


Figure 20. Dose-response relationships. Within species, significant advances are being made in quantitative systems biology modeling (Eisenberg et al. 2008). Panel A: Overall feedback control system model of thyroid hormone regulation with three source organ blocks (hypothalamus [HYP], anterior pituitary [ANT PIT], and thyroid glands [THYROID]); three sink blocks—for TRH, TSH, and T3 and T4 distribution; and elimination (elimination = metabolism and excretion) (D&E). TRH = thyrotropin-releasing hormone; TSH = thyroid-stimulating hormone; T3 = triiodothyronine; T4 = thyroxine; SR = secretion rate; p = plasma or portal plasma for TRH-related components; DA = dopamine; SRIH = somatastatin. Panel B: Feedback control system (FBCS) model validation study results. Predicted normal circadian TSH versus independent TSH data (not used in fitting the FBCS model) (triangles and diamonds represent data from Sarapura et al. (2002), circles represent data from Samuels et al. (1994). Also shown (squares) are the mean TSH data from the larger database used to fit the FBCS model of Blakesley et al. (2004). Reproduced with permission from Mary Ann Liebert, Inc.

Relative Potency

Identification of pathways and assays impacted by chemicals of known toxicity can be useful in initial prioritization of many compounds. These can be identified through development of predictive models built on relationships between *in vitro* ToxCast™ assay results and *in vivo* effects caused by known developmental toxicants (Sipes et al. 2011). A chemical's potency can be further refined using focused *in vivo* tests with alternative species to provide informative dose-response data and exposure window relationships. Alternative species provide easily manipulated model systems that can detect effects caused by mechanisms not assessed by *in vitro* HTS. For example, zebrafish were used as a screening model to assess the 309 EPA ToxCast™ Phase I chemicals for developmental toxicity to both humans and ecological species. In fish embryos or larvae, 191 (62%) chemicals were toxic (death or malformations) to the developing zebrafish. Both toxicity incidence and potency were correlated with chemical class and hydrophobicity. As an integrated model of the developing vertebrate, the zebrafish embryo screen provides information relative to overt and organismal toxicity. In 12 classes of chemicals, 100% of the chemicals induced developmental toxicity, 4 classes of which induced developmental toxicity with an average concentration at 50% of the maximum level (AC_{50})[27] below 4 µM. Translating such results directly into a dose-response for human risks is difficult, but results of Padilla et al. (2012) show that alternative species can be used

[27] In high-throughput screening (HTS) assay, AC_{50} is the concentration at which an assay is inhibited or activated by 50% when compared to control values. This value is useful in comparing assay results.

to build relative rankings of chemicals based on their potency to cause adverse effect. Such rankings can be used for ranking and prioritizing chemicals or classes of chemicals for additional evaluation.

Dose-Response Relationships

Chemical dose-response relationships characterized in one alternative species might be extrapolated to other species or to humans, using a pathway-based approach (Perkins et al. 2013). Because many biological functions and disease pathways are conserved across species, similarity of genes encoding those pathways can support direct comparisons of pathway or genomic effects between species. Where pathways are highly conserved between species, this information can be used to extrapolate dose-response relationships in alternative species to analogous relationships in mammals. For example, pathways in the hypothalamus-pituitary-gonad axis are highly conserved among vertebrates, which been used to show that chemical effects in fathead minnows are predictive of endocrine disrupting effects in rats (Ankley, G. T. and Gray 2013).

Pathway effects defined through gene expression changes can be used to define a benchmark dose or sensitivity of an animal to a chemical (Thomas, RS et al. 2011). Benchmark concentrations derived from aqueous exposures of alternative species can be translated to oral equivalents in other species, such as humans, by applying a dose scaling factor composed of a simple reverse toxicokinetics approach to estimate the blood dose and amount of metabolism in the target species (Wetmore et al. 2012). Using this approach, chemical concentration effects can be translated from alternative species to mammalian species. See Figure 20 for an example of how systems biology can inform dose-response extrapolation within species.

However, this type of an approach has added uncertainty, and may generally increase uncertainty to an unacceptable level, which precludes the calculation of a reference value, including a provisional value. There is uncertainty with respect to defining a benchmark dose based on gene expression changes and with respect to the pathway context and interpretation. For instance, changes in gene expression do not directly translate into changes in protein expression or activity. In addition, it is well known that signaling and metabolic pathways within the cell are intersecting and inter-related. There is uncertainty with respect to the dose-response changes at particular key events and how downstream key events may be altered by other intervening pathways. Thus, calculating a benchmark dose for a pathway itself is fraught with challenges and additional uncertainty that current reference value approaches do not take into account. In all likelihood, accounting for these additional sources of uncertainty may require new uncertainty factors to be developed, and increases the likelihood that an unacceptable level of uncertainty may be encountered.

Thus, it is more likely that, until better, less uncertain methods and techniques are developed and used, pathway-based effects based on gene expression are more suitable for screening level decisions and less suitable for reference value derivation.

Species to Species Extrapolation

For most species, qualitative predictions are likely to be tenable based on hypothalamus-pituitary-thyroid (HPT) dependent pathways, that is, iodine uptake. Altered iodine uptake hinders development, but the most sensitive outcome indicator might be different. In rats, thyroid hormone disruption can lead to thyroid tumor development (Hurley 1998), while in frogs, metamorphosis is

disrupted (Degitz et al. 2005). Quantitative predictions might not be feasible for many species due to limited data on downstream endpoints and key events (Perkins et al. 2013) .

Normal thyroid hormone-dependent post-embryonic development requires coordinated spatial-temporal control of thyroid hormone activity. Such activity is regulated not only through the classical features of the HPT axis, but also through peripheral mechanisms external to the hypothalamus, pituitary, and thyroid tissues, such as differential regulation of deiodinase activity, hepatic metabolism and excretion of thyroid hormones, thyroid hormone receptor regulation, and transmembrane thyroid hormone transport. Of these major controlling processes, the mechanisms of the central HPT axis are considered to be generally conserved across vertebrate species and useful for comparative efforts; however, those of the peripheral tissues are typically more divergent and must be used with care in cross-species analysis.

Population Variability

Understanding the variation of an individual relative to population variation can be key to identifying an adverse effect on a population. Polymorphisms affecting drug responses can vary widely in populations. In humans, 20–25% of prescription drugs are metabolized in the liver by cytochrome P450 CYP2D6 where variants confer widely different rates of drug metabolism, such that some people might respond with an onset of toxicity while others fail to experience efficacy (Ingelman-Sundberg 2005). Variants causing unanticipated results can comprise a significant portion of a population and that distribution can vary widely across populations (Sistonen et al. 2007, Ingelman-Sundberg 2005, Andersen et al. 2002, Wooding et al. 2002). Understanding the variation in adverse responses across a diverse testing population helps reduce the uncertainty of extrapolating laboratory data to real populations. Differential response to chemicals is an important consideration in ecological risk assessment where not only potentially sensitive subpopulations might exist, but also sensitive species.

Approaches have been developed to incorporate population diversity into toxicity testing through the use of large collections of different genetic lines of mice or cell cultures derived from them (O'Shea et al. 2011, Rusyn et al. 2010, Harrill et al. 2009). Alternative species could be especially useful for incorporating population variability into toxicity testing. The diversity in laboratory lines and outbred populations of fish can be high, especially if populations are collected from different areas impacted by pollutants (Williams and Oleksiak 2011, Guryev et al. 2006). Divergent lines of zebrafish can be used to examine variation in responses to chemicals in addition to determining possible genetic factors influencing adverse effects. Using this approach, Waits and Nebert (2011) crossed zebrafish lines displaying different levels of sensitivity to developmental cardiotoxicity caused by 3,3',4,4',5-pentachlorobiphenyl. The crosses were used in genome-wide quantitative trait loci mapping to identify several genes in addition to the AhR that contribute to the gene-gene and gene-environment interactions that drive developmental toxicity of dioxins and dioxin-like chemicals.

Although genetic diversity can be incorporated into testing using a panel of genetically inbred lines, unexpected results can occur. In a study comparing the responses of 19 inbred to 20 outbred zebrafish lines, Brown et al. (2011) found that effects of the endocrine disrupting chemical clotrimazole were dramatically different. Clotrimazole acts by inhibiting P450 activities involved in steroidogenesis production in fish. In inbred fish lines, 11-ketotestosterone production via steroidogenesis was significantly inhibited. In contrast, outbred lines responded with Leydig cell proliferation in testes and normal plasma concentrations of 11-ketotestosterone indicating that the

outbred lines could compensate for inhibition by clotrimazole. Here, inbreeding had a strong impact on the diversity and type of response to the endocrine disruptor. Ultimately, the combination of *in vivo* and *in vitro* data should enable development of a weight-of-evidence case as to the toxicity caused by the chemical and whether potential human health effects are likely.

Cumulative Risks

As has been described elsewhere in this document, correct identification of causal perturbations that lead to adverse outcomes will enable determination of which environmental factors are likely to contribute to the cumulative risk for specific outcomes and which are not. Additionally, testing of combinations of chemicals can perhaps be conducted most efficiently in alternative species. For example, alterations in neurosensory functions and intrafollicular thyroxine content of zebrafish exposed to potential disruptors have proven to be useful tools for evaluating multiple chemicals (Raldua et al. 2012, Thienpont et al. 2011, Froehlicher et al. 2009), as has the zebrafish developmental assay (Padilla et al. 2012).

3.2.3. Short-Term *In Vivo* Models – Mammalian Species

The use of new short-term *in vivo* exposure mammalian bioassays to support Tier 2 assessments are described here. The prototype is drawn from research described in papers by Thomas R.S. et al. (2011)and discussed further in Thomas R.S. et al (2013a, 2013b). The goal of this research was to describe what would be required for the application of short-term *in vivo* transcriptomic assays in predicting chemical toxicity.

Hazard Identification

Short-term *in vivo* transcriptomic assays provide the metabolic capability and systems-level integration of whole animal studies with a more rapid assessment of response to chemical treatment based on molecular-level data. As more data from short-term *in vivo* transcriptomic studies become publicly available, as study designs become standardized, and as gene expression patterns and network perturbations are identified, the ability to predict chemical toxicity comparable to longer term assays is expected to increase. See Text Box 8 for more about the transcriptome.

Box 8. What is the Transcriptome?

Ribonucleic acid (RNA) is the functional outcome of deoxyribonucleic acid (DNA) transcription by transcription factors. Researchers can study the transcriptome—the set of all RNA molecules in a given cell—and determine gene expression patterns, or signatures. Specifically, short-term transcriptomic assays in mammalian and alternative species enable observations of the effects of chemical exposure across multiple tissues.

For hazard identification, a host of previous studies has demonstrated that transcriptomic signatures from short-term *in vivo* studies can be used to predict both subchronic and chronic toxic responses. A transcriptomic "signature" is typically defined as a subset of genes for which the qualitative or quantitative expression pattern can be used to predict an *in vivo* adverse response with a defined accuracy.

To develop a broad-based repertoire of gene expression signatures for hazard prediction, several factors should be considered. First, the number of endpoints included should be sufficient to allow a comprehensive prediction of toxicological hazard. Previous studies that have used gene expression microarray analysis following short-term exposures of chemicals have been limited in the breadth of endpoints examined. These endpoints include the prediction of rat liver tumors (Fielden et al.

2011, Uehara et al. 2011, Auerbach et al. 2010, Ellinger-Ziegelbauer et al. 2008, Fielden et al. 2008, Fielden et al. 2007, Nie et al. 2006), mouse lung tumors (Thomas, RS et al. 2009), and rat renal tubular toxicity (Fielden et al. 2005). One strategy that could be employed would be the selection of a battery of tissues, which would include those most frequently positive in rodent cancer bioassays (i.e., liver, lung, mammary gland, stomach, vascular system, kidney, hematopoietic system, and urinary bladder) and tissues commonly affected by noncancer disease. In a previous analysis, these eight tissues cover 92 and 82% of all mouse and rat carcinogens, respectively (Gold et al. 2001). Additional tissues also would need to be added for developmental and reproductive effects, which could include the developing fetus and gonadal tissue.

Second, the number of positive and negative chemicals for each endpoint in the studies would need to be sufficient, and the chemical diversity must match the diversity in the chemicals that will ultimately be predicted. For complex toxicological responses such as tumor formation, a previous study estimated that at least 25 chemicals were necessary (Thomas, RS et al. 2009). Third, selection of the time point to perform the gene expression analysis is also a consideration. The time point selection is a balance between cost (i.e., the shorter the time point, the less expensive the study) and a more stable gene expression signature. Among the previous efforts, certain studies relied on much shorter time points (e.g., 5 days), but tended to increase the dose beyond that which would be tolerated in a chronic bioassay (Fielden et al. 2007). Other studies used the same doses as those in the chronic bioassay, but used exposures longer than 5 days (Thomas, RS et al. 2009). In one study that examined the effect of exposure duration, the overall conclusion was that increasing exposure duration increased the predictive performance of the gene expression signatures for genotoxicants (Auerbach et al. 2010).

Exposure/Dose-Response Assessment

As described by Thomas R.S. et al. (2013b, 2012, 2011, 2007), short-term *in vivo* transcriptomic assays have also been applied to dose-response assessment. In a NexGen collaborative effort between EPA and the Hamner Institute, female B6C3F1 mice were exposed to multiple concentrations of five chemicals that were positive for lung or liver tumor formation in a two-year rodent cancer bioassay (Thomas, RS et al. 2012, Thomas, RS et al. 2011). Histological and organ weight changes were evaluated and gene expression microarray analysis was performed on the liver or lung tissues. The histological changes, organ weight changes, and tumor incidences in traditional bioassays were analyzed using standard dose-response modeling methods to identify noncancer and cancer points-of-departure. The dose-related changes in gene expression were also analyzed using a modification of EPA's benchmark dose (BMD) approach (EPA 1995). The gene expression changes were grouped based on both biological processes and canonical signaling pathways. A comparison of the transcriptional BMD values with those for the traditional noncancer and cancer apical endpoints showed a high degree of correlation for specific biological processes (Thomas, RS et al. 2011) and signaling pathways (Thomas, RS et al. 2012). In addition, transcriptional changes in the most sensitive pathway were also highly correlated with the apical responses (see Figure 21). Further studies have also demonstrated the stability of the correlation between transcriptional changes and apical responses across different exposure periods (5 days to 13 weeks) (Thomas, RS et al. 2013b). Understanding of the effect of exposure duration on outcomes is a key issue in the design and use of new types of bioassays.

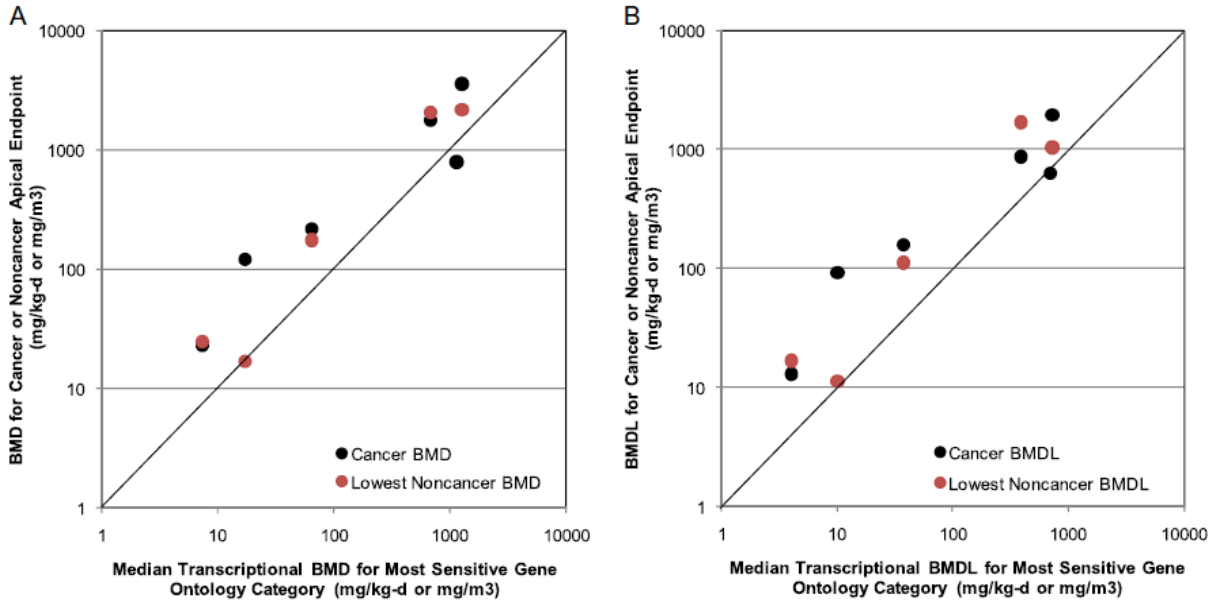


Figure 21. Scatter plot of the relationship between (A) benchmark dose (BMD) and (B) benchmark dose lower limit (BMDL) values for the cancer and noncancer apical endpoints and the transcriptional BMD and BMDL values for the most sensitive GO category. For each chemical and tissue, the BMD and BMDL values for tumor incidence and the lowest noncancer BMD and BMDL values were plotted. For MECL in the lung, no noncancer BMD or BMDL values were plotted because of the absence of histological changes (Thomas, RS et al. 2011). Reproduced with permission from Oxford Journals.

With the advent of quantitative high-throughput screening (qHTS), the potential to screen thousands of chemicals for biological activity presents as many challenges as promises. If qHTS can decrease the number of chemicals of interest by 90% (a 10% hit rate across chemicals and assays), this would still overwhelm the throughput of the traditional toxicity testing paradigm. Clearly, a multi-tiered approach to prioritization can lead to more effective applications of animal toxicity testing. As part of this tiered approach, short-term *in vivo* transcriptomic assays provide a tool that incorporates both metabolism and systems-level integration in response to chemical treatment. See also a description of cost savings in Text Box 9. The development of predictive gene expression signatures and dose-response studies would provide a relatively efficient and cost-effective method for both identifying chemicals of concern and estimating a point-of-departure for adverse responses. This information would help support large-scale prioritization and regulatory efforts in the United States and Europe. The gene expression data combined with other data types (e.g., toxicity data from similar chemicals, PK data) could provide sufficient information to replace the present chronic toxicity and carcinogenicity studies. It should be noted that expression changes can vary depending on dose, time, species, tissue life stage, and individual genetic profile; thus, increasing the complexity of identifying causal relationships between exposure, specific signatures, and outcomes.

Box 9. Short-Duration *In Vivo* Models Potential Cost Savings

We concluded that, assuming no overlap among chemicals across the battery of 10 tissues tested, a proof of concept for predicting hazard using short-term *in vivo* transcriptomic mammalian assays could be developed using approximately 250 chemicals at a single dose across 10 tissues. Additionally, they calculated that the costs associated with the proof of concept would be approximately $90,000 per chemical for a 28-day exposure for a total cost of $22.5 million. In comparison, each chronic exposure assay costs approximately $1.5 million resulting in a total of approximately $375 million.

Costs for similar efforts using alternative animals (fathead minnows, zebrafish, invertebrates) would be approximately $10,000 per chemical for definitive reproductive assays. Short-term, pathway-based benchmark dose (BMD) assays would be $10,000 per chemical for invertebrate or fish embryo assays and $48,000 per chemical for 21-day fish reproductive assays (5 tissues) or $2.5 million and $12 million for 250 chemicals.

3.3. Tier 1: Screening and Prioritization

This section summarizes new approaches that are available to develop data for screening and prioritizing large numbers of chemicals (i.e., greater than tens of thousands of chemicals) for more focused testing. The increasing maturity of these new approaches has led EPA and other organizations to plan on using these Tier 1 data to prioritize and screen chemicals for immediate regulatory decision, for further testing in Tiers 2 and 3, or in some cases, to add to the weight of evidence in Tier 2 and 3 assessments, especially with respect to identifying pathway or molecular signatures associated with chemical-induced disease.

Tier 1 risk assessments are based on *in vitro* assays (including use of human cells), statistical and systems models that focus on molecular molecular targets, QSAR models, and pathways considered relevant to adverse effects or clinical disease. One scientific rationale for using *in vitro* assays is that they probe key events or MIEs that can lead to adverse outcomes. Assay endpoints are designed to represent MIEs and predict subsequent adverse outcomes based on previous studies, both *in vitro* and *in vivo*. The analyses provide the anchoring information critical to characterizing relevance of a "hit" in an *in vitro* assay. Documenting the linkage from assay endpoint to MIE to potential for adversity is thus key to evaluating the relevance of each assay that might be used as part of a Tier 1 risk assessment. The evidence for this linkage can come from statistical modeling using *in vivo* and *in vitro* data on the same chemicals, or from detailed biological modeling (e.g., virtual tissue (VT) models or other types of systems biology models).

The modeling techniques used in Tier 1 (e.g., QSAR and HTS methods) are designed to assess hundreds to thousands of chemicals in parallel (see Figure 23). In addition to using high-throughput (HT) assays to generate hazard information, moderate- to HT toxicokinetics approaches (here called reverse toxicokinetics or RTK (Rotroff et al. 2010)) are also developed and applied. New approaches can now estimate doses that can activate particular relevant pathways in humans, using data from *in vitro* assays (Wetmore et al. 2012). Bayesian-based exposure models can also be used to generate exposure estimates for chemicals based on production volume and use patterns (Wambaugh and Shah 2010).

The data generated from the Tier 1 assays can be used to prioritize chemicals for further study or can simply augment the weight of evidence for chemicals that are already being considered in Tiers 2 or 3. For prioritization, from a large set of chemicals for which HTS data are available, one might

identify the subset that is likely to interact with known relevant pathways, or which demonstrate pathway disruptions similar to known diseases. Doses at which pathways may be activated or perturbed may be estimated. These estimates may be combined with exposure, occurrence, and other information to select chemicals which may advance into Tiers 2 or 3. Tier 1 data might also directly augment Tiers 2 and 3 weight-of-evidence determinations helping to identify or further characterize pathways alterations associated with disease for sensitive endpoints observed in higher level *in vivo* testing, providing good examples of the integration of the bottom-up and top-down perspectives advocated in the NexGen framework strategy. *In vitro* and modeling data can also be used to guide a next round of *in vivo* data generation.

Table 8 provides a brief description and critical review of the tools, methods, and models that could be used in Tier 1.

Table 8. Summary of Tier 1 NexGen Approaches, Including Weight of Evidence, Pros, and Cons

Tier 1: Screening and Prioritization Categories of Approaches Considered		
	New QSAR Models	Validated High-Throughput In Vitro Assays
Approach:	Uses structural characteristics and experimental data from chemical analogues to predict modes of action, metabolism, hazard, and fate and potency for data-poor chemicals.	Experimentally measures dose-dependent, chemically-induced alterations in biological functions using a range of specific and sensitive *in vitro* assays. Infer potential adverse outcomes based on existing knowledge of other chemical and potential importance of selected biological processes.
Weight of evidence:	Determined by quality and amount of existing data, but generally suggestive.	Determined by supporting traditional data and systems biology knowledge, but generally suggestive to likely.
Pros:	Data are readily and inexpensively available for all chemicals. If the basis for the QSAR model(s) matches the physical chemistry of the evaluated chemicals, the model(s) generally predicted potency within a factor of 100. Harmonized international approaches are available.	Rapid, inexpensive, multiple bioassay options are available. False negatives and positives for ToxCast™ evaluated assays are low (when testing directly acting chemicals, not toxic metabolites).
Cons:	If models do not match the physical chemistry of evaluated chemicals, results are unreliable. Models do not predict active metabolites.	Assay coverage of all important biological processes currently incomplete resulting in false negatives for chemicals that perturbed those processes. Similarly, disorders related to interactions among cell types or tissues cannot be evaluated, that is, reproductive/developmental effects. Limited ability to test for active metabolites or volatiles. False negative rates are of concern. Links to disease outcomes are variable.

3.3.1. QSAR and High-Throughput Virtual Molecular Docking (HTVMD)

(Q)SAR[28] models are regression or pattern recognition models that are used in risk assessment to classify or predict the potency of chemicals for toxicological activity, exposure potential, and the like as a function of one or more chemical descriptors. The descriptors are generally the inherent physiochemical properties of the chemical such as atomic composition, structure, substructures, hydrophobicity, surface area charge, and molecular volume. QSAR models require only the inherent properties of the 2-D or 3-D chemical structure as input parameters, and are thus considerably less costly and faster than hazard animal test results. With a variety of QSAR models to choose from (Hansen et al. 2011), and each model having a set of assumptions and a chemical domain of applicability, interpreting QSAR results for use in hazard and dose-response assessment requires expertise.

QSAR models have been most commonly used in classification of chemicals with unknown hazard or exposure potential by comparing the "query" chemical's inherent properties with similar properties for a set of chemicals that have known toxicological or exposure potential called the "training set" (Venkatapathy and Wang 2013, EPA 2012c, Goldsmith et al. 2012, OECD 2012, Wang, N et al. 2012b, EC 2010). SAR models provide a qualitative identification of specific hazards (e.g., suspected carcinogens, mutagens, and reprotoxicants). The commercially available TOPKAT model (TOPKAT, 2013) provides quantitative estimates that can be used to rank chemicals for potency (Venkatapathy and Wang 2013, Venkatapathy et al. 2004). In the European Community, QSAR results are used to prioritize chemicals for additional toxicity testing.

At EPA, (Q)SAR models are being used to screen, rank, and categorize chemicals for level of concern in a variety of EPA programs, including Superfund mitigation; the Office of Chemical Safety and Pollution Prevention (OCSPP) High Production Volume Challenge Program and Pre-Manufacture Notice review process; the OCSPP/Office of Water Endocrine Disruptors Screening Program (Weiss et al. 2012); and the Office of Water Candidate Contaminant List. The QSAR models used by EPA include the Sustainable Futures Initiative suite of models, the Organization of Economic Co-operation and Development (OECD) QSAR toolbox models (OECD 2012, 2004), High-throughput Virtual Molecular Docking (HTVMD) (Rabinowitz et al. 2008), MetaCore (Teschendorff and Widschwendter 2012, van Leeuwen et al. 2011), and the TOPKAT model (Rakyan et al. 2011, Venkatapathy et al. 2004).

HTVMD models use a ligand-based chemoinformatics strategy to predict relationships between various attributes of ligands and their binding to known targets. These models are increasingly being used in risk assessment and can screen thousands of chemicals for the potential affinity of their 3D structures to bind to active protein binding sites. HTVMD models have been used in the pharmaceutical industry for many years to identify candidate drugs. These models can also be used to estimate the likelihood that a chemical of toxicological interest would bind to a target protein, for example, the potential affinity as a direct agonist of the estrogen receptor.

[28]The parentheses around the "Q" in (Q)SAR indicate that the term refers to both qualitative predictive tools, i.e., structure-activity relationships (SARs) and quantitative predictive methods, i.e., quantitative structure-activity relationships (QSARs). Although the term (Q)SAR is often used to refer to predictive models, especially computer-based models, (Q)SAR actually includes a wide variety of computerized and noncomputerized tools and approaches (Hansen et al. 2011).

Recent advances in high-performance computing support simultaneous runs of QSAR and HTVMD models, dramatically decreasing the time to discovery. The U.S. Army Medical Research and Materiel Command, for example, has recently published their version of a Docking-based Virtual Screening pipeline that facilitates the usage of the AutoDock molecular docking software on high-performance computing systems (Jiang et al. 2008).

The OECD provides a free downloadable QSAR software package, the QSAR Toolbox, that is intended for use by governments, the chemical industry, and other stakeholders to assess potential chemical human and ecological toxicities for data-poor chemicals (OECD 2012). The QSAR Toolbox estimates the potential toxicity of a compound of interest based on the available information (e.g., mechanism, MOA, or toxicological effects) for structurally similar analogs, and uses read-across or trend analysis to construct categories of chemicals for screening purposes even if only a few of the members in the category have available test data. The popularity of the read-across method is driven by its relative simplicity and the availability of the QSAR Toolbox online. OECD has also developed guidance on the validation of QSAR models when used for regulatory purposes (OECD 2004). Assessments informed by new data types and methods will incorporate the results from data sources that can be automated (e.g., QSAR and molecular docking models, and HTS data), with the more traditional data (when available) to advance the speed and accuracy of chemical screening and to support the weight-of-evidence approach to toxicity prediction (Golbraikh et al. 2012, Lock et al. 2012, Rusyn et al. 2012, Wignall et al. 2012, Sedykh et al. 2011). Use of the above models and approaches will advance the ranking of chemicals currently being produced, as well as support the design of new products and chemical processes that increasingly minimize harm to health and the environment.

3.3.2. High-Throughput and High-Content Assays

HTS and high-content screening (HCS) assays are major tools used for early evaluation of chemicals and their ability to perturb molecular pathways (Judson et al. 2013, Sipes et al. 2013, Tice et al. 2013, Kavlock et al. 2012, Judson et al. 2011). Much of the HTS/HCS (for the remainder of this section use of the term HTS includes both HTS and HCS) methodology was developed to aid the pharmaceutical and biotechnology industries in the drug discovery process where one has a drug target of interest (e.g., a receptor or enzyme) and a need to screen up to millions of candidate compounds for leads (Mayr and Bojanic 2009, Bleicher et al. 2003). The technology has been used more broadly in approaches often called chemical genetics (or sometimes chemical biology) where small molecule screening is used to identify probes for biological signaling networks and cellular phenotypes (Schreiber 2003). These assays became of interest in toxicology because many targets of pharmaceutical and chemical biology interest could also be postulated to be involved in disease processes driven by unintentional exposures to environmental chemicals (Houck and Kavlock 2008). Generating a large data matrix of toxic chemicals and HTS assays against critical proteins and cellular phenotypic effects provides toxicologists an opportunity to discover novel MOAs that have long eluded the field.

The underlying technologies for HTS assays are well known, so a detailed discussion is not presented here. Instead, the discussion focuses on a broad description of the types of assays and some of the key issues to be considered when designing *in vitro* Tier 1 approaches. HTS assays are broadly divided into two types: cell-free/biochemical or cell-based. Cell-free assays typically test for the direct interaction of a test chemical with a specific protein such as a receptor or enzyme. Measures of interaction include binding or inhibition of enzyme activity. In cell-based assays, a

cellular readout can be molecular-based (e.g., changes in gene or protein expression) or phenotypic (cytotoxicity, changes in cell morphology). In a cell-based assay, the selection of the cell system is critical. Assays have been developed using a variety of primary cell types from various organs and species, immortalized cell lines, and stem cell types (Dick et al. 2010). The choices reflect the strengths and weaknesses of the different approaches. For example, immortalized cell lines generally produce very reproducible screening results over long periods of time due to the continuous growth and stability of the cell lines; however, this occurs at the cost of having significant differences from the *in vivo* physiology of the cell type from which the line was derived. The converse holds true for most primary cells, that is, better representation of true physiology but more challenging to work with in producing consistent, reproducible screening results. Co-culture systems combine different cells in an attempt to mimic *in vivo* systems requiring complex cell-cell signaling networks (Berg et al. 2010). Certain whole organisms, including *Caenorhabditis elegans* and zebrafish embryos, can also be used in HTS assays (Smith, MV et al. 2009, Parng et al. 2002).

3.3.3. Toxicokinetics

HTS assays provide toxicologists with an efficient and cost-effective tool to broadly screen chemicals for potential proximal biochemical and cellular interactions. As previously mentioned, the HTS assays are run in concentration-response format. The potency of each chemical in each assay can be summarized using AC_{50} or LEC (lowest effective concentration) values, depending on the type of dose-response data collected. The potency values among the *in vitro* assays, along with other chemical information, have been proposed for use in hazard identification (Martin et al. 2011, Sipes et al. 2011) and prioritization of chemicals for further testing (Reif et al. 2010). The relationship between the *in vitro* concentration of the chemical in the well to the concentration of the chemical in the blood or target tissue (*in vivo*), however, can be complex and dependent on variables that are not captured in the HTS assays. These variables include bioavailability, clearance, and protein binding (Wetmore et al. 2012).

In vitro to *in vivo* extrapolation (IVIVE) is a process that uses data generated within *in vitro* assays to estimate *in vivo* drug or chemical fate. In the past, IVIVE has been predominantly developed and applied in the pharmaceutical industry to estimate therapeutic blood concentrations for specific candidate drugs, and to identify potential drug-drug interactions (Chen, Y et al. 2012, Shaffer et al. 2012, Gibson and Rostami-Hodjegan 2007). Due to both legislative mandates and public pressure for increased toxicity testing, IVIVE is increasingly being used to predict the *in vivo* PK behavior of environmental and industrial chemicals (Basketter et al. 2012).

A combination of IVIVE and reverse dosimetry can be used to estimate the daily human oral dose (called the oral equivalent dose) necessary to produce steady-state *in vivo* blood concentrations (C_{SS}) that are considered equivalent (with respect to chemical concentration at potential targets) to the dose delivered *in vitro* at the AC_{50} or LEC values, and can be used for those values across the more than 600 *in vitro* assays (Wetmore et al. 2012, Rotroff et al. 2010).

3.3.4. High-Throughput Exposure Estimation: ExpoCast Prioritizations

The use of HT assays to characterize biological activity *in vitro* enables prioritization of potential environmental hazards once the results of *in vitro* assays have been anchored to, and found to be predictive of, *in vivo* effects. Without capabilities for HT assessment of potential for exposure, prioritization (with respect to potential risk) cannot be completed, as most chemicals have little or no exposure data (Wetmore et al. 2012, Arnot et al. 2010b, Arnot et al. 2010a, Cohen Hubal et al.

2010, Rotroff et al. 2010, Hubal 2009, Sheldon and Cohen Hubal 2009, Rosenbaum et al. 2008, Arnot and Mackay 2007, NRC 2006). Currently, few, if any, inexpensive *in vitro* assays are widely available to characterize those properties of chemicals relevant to exposure. Furthermore, the studies for assessing both the presence of environmental chemicals in the immediate vicinity of individuals (exposure potential) and any known biomarkers of actual exposure are expensive, labor intensive, and, with the notable exception of CDC's NHANES, typically difficult to extrapolate to the general population (Rudel et al. 2008, Angerer et al. 2006, Eskenazi et al. 2003). For these reasons, exposure prioritization must be drawn from mathematical models which, when parameterized by chemical-specific properties, provide a structured, consistent way to approach large numbers of unknown chemicals.

Physicochemical properties (e.g., water solubility, preference for binding in lipids) inherent to a given compound have been used to predict potential bioaccumulation, and even toxicity, within ecological species to make HT prioritizations of potential chemical exposure (Gangwal et al. 2012, Reuschenbach et al. 2008, Walker et al. 2002, Walker and Carlsen 2002). Beyond inherency, environmental fate and transport models have been developed to account for the accumulation of compounds in various environmental media (i.e., air, soil, water) and the degradation rates of those compounds in those media. These fate and transport models enable predictions of human exposure based on assumptions of human interaction with environmental media and derivation of food from the environment (Arnot et al. 2010b, Arnot et al. 2010a, Rosenbaum et al. 2008, Arnot and Mackay 2007). Parameterized using chemical structure and production volumes alone, these models can be used to make HT exposure prioritizations (Arnot and Mackay 2007).

EPA is developing the ExpoCast exposure model prioritization framework, which is flexible and expandable to incorporate new HT exposure models as they become available. Currently the framework relies on two quantitative fate and transport models amenable to HT operation: USEtox (Rosenbaum et al. 2008) and RAIDAR (Arnot and Mackay 2007). These models have been empirically assessed for their ability to predict exposures inferred from the NHANES data set. These "ground truth" biomonitoring data are used to calibrate the model predictions and estimate *de facto* uncertainty of the predictions for 41 chemicals where intake per unit emission, total production volume or volume applied, and actual exposures inferred from biomonitoring data were available. The calibration and uncertainty are then extrapolated to ~1,600 chemicals to make rank order predictions on a per unit emission basis, as well as a rank order prediction for ~600 chemicals adjusted using production volume (Wambaugh and Shah 2010).

NexGen efforts to incorporate exposure prioritization information could proceed along three fronts. First, efforts to evaluate the utility of the predictions must be undertaken to determine if the chemicals of highest priority are indeed present in the environment. Next, new models must be developed to address aspects of exposure currently underrepresented by fate and transport models—namely exposure from personal contact sources (i.e., consumer use). Finally, using the full uncertainty range of the absolute exposure predictions (mg/kg body weight/day), risk potentials could be calculated for risk-based prioritization.

3.3.5. Virtual Tissue (VT) Modeling

VT models provide an experimental and theoretical framework for the systematic and integrative analysis of complex multicellular systems. These models capture the flow of molecular information across cellular and biological networks, and process this information computationally into higher order responses that ideally simulate a potential adverse outcome(s). Responses to perturbation

depend on network topology, system state dynamics, and collective cellular behavior. A unique aspect is that these simulations are enabled from individual cellular behaviors in a multicellular field that can result in emergent properties, which are behaviors that arise from interactions of parts at the next level of a system (e.g., functions, phenotypes) that are not apparent from knowledge about the behavior of the parts alone.

The field of VT models is in the early stages of development but will become more prominent as the state of science develops. Jack et al. (2011) and Knudsen et al. (2010) provide examples of VT utility and the state of science. VT models are practical solutions for translating between biological data and individual and population-level health outcomes. They combine data and knowledge into computer models that predict behavior of a complex system, leading to adverse outcomes in hepatic toxicity, developmental toxicity, reproductive toxicity, cardiopulmonary toxicity, and more (EPA 2009b).

Virtual models are also briefly discussed in Section 4.4 as one of the new approaches that can address recurring issues in risk assessment, in this case, dose-response characterization.

3.3.6. Example: Thyroid Pathway Disrupting Chemicals and High-Throughput Systems

For EPA to base regulatory decisions on data from mechanistic-based evaluations, several issues must be addressed. EPA will need to develop criteria and approaches for translating data across the various types of testing and to identify the types of data and information to support the use of these data in a regulatory context. To this end, EPA's NexGen Thyroid Disrupting Chemical Workgroup (EPA 2012a) conducted a thyroid prototype case study that reviewed existing ToxCast™ assays and provided recommendations for how the data could be used to predict thyroid disruption-induced developmental neurotoxicity.

A major reason the workgroup selected the thyroid hormone system as its prototype is that the underlying biology of thyroid hormone homeostasis is well established, thus enabling the elucidation of the pathway(s) for thyroid hormone disruption (Zoeller and Crofton 2005). The workgroup identified three issues that should be addressed to use HT assays to predict which environmental chemicals would likely cause developmental neurotoxicity via disruption of thyroid hormone homeostasis. These issues are Assay Identification and Refinement; Algorithm Development for Toxicity and Hazard Prediction; and Standards Development for Assay Conduct, Data Analysis, and Data Reporting for Risk Assessment Needs. The following is a brief summary of the case study.

Assay Identification and Refinement

As a first step, the workgroup identified the HT assays in the ToxCast™ database that assess endpoints known to be relevant to disruption of thyroid function. The workgroup found that ToxCast™ contains multiple assays relevant to assessing the potential for a chemical to disrupt thyroid hormone homeostasis. Coverage of the effects of concern, however, is quite variable. Although five of the identified assays evaluate endpoints that directly affect the thyroid hormone pathway (e.g., thyroid hormone receptor binding and TRH receptor binding), the rest evaluate endpoints not specific to the thyroid hormone pathway. For example, of the 90 assays identified as thyroid-relevant, 85 are related to hepatic stimulation, metabolism, and clearance of thyroid hormones. Alteration of these pathways influences thyroid hormone homeostasis indirectly, and neurodevelopmental effects tied to thyroid disruption by this mechanism are thus secondary effects

of a chemical (inadequate hormone availability due to increased elimination). These secondary effects are in contrast to a primary effect, whereby a chemical interferes directly with the function of the thyroid gland itself or interacts at the site of thyroid hormone receptor in the brain of a developing organism.

Adequately assessing the potential of an environmental chemical to disrupt thyroid hormone homeostasis requires that appropriate endpoints be identified and assays be developed and incorporated into testing schemes. This process will involve identifying the specific endpoints in the pathways that need to be tested, additional assays that could be available but which are not currently part of ToxCast™, and additional assays that need to be developed. A recent workshop review by Murk et al. (2013) provides a state-of-the-science assessment of important MIEs for thyroid disruptors, potential and currently used assays for these MIEs, and recommendations for research priorities.

Algorithm Development for Toxicity and Hazard Prediction

The workgroup's second recommendation was to develop algorithms or decision logic flows that balance the potential adversity of the outcome with the uncertainty of the available data. Should assays evaluating endpoints directly affecting the thyroid-related brain changes be weighted more heavily in algorithms than those measuring upstream hepatic enzyme induction? How will algorithms incorporate the fact that multiple chemicals might interact with the same key event, and one chemical might interact with various MIEs, and thus lead to multiple adverse outcomes? Biological plausibility should be the driver in algorithm development.

Another aspect to consider is the methods used to incorporate assay results into analyses. Clearly, incorporating many sets of dose-response information into combinatorial analysis requires some simplification of assay results. Many current HT assay results are simplified via classification as either a positive or negative ("hit" or "no hit"), or are assigned a summary statistic such as an IC_{50} (the concentration producing a 50% inhibition of response) or lowest effective dose. Obviously, binary decisions such as hit/no hit determinations depend on the criteria chosen to define a hit. These criteria could be derived from statistical significance, biological significance, or an arbitrary, nominal level of change. Depending on the data set, the basis for the classification criteria might be difficult to determine, and might not be consistent across assays. Similarly, summary statistics depend on the model used to generate them or on the specific value chosen (such as IC_{50} versus IC_{10}). Relative potency ranks also might vary depending on the shape of the dose-response curve, such that within a given set of chemicals, Chemical A could have the lowest IC_{50} while Chemical B had the lowest IC_{10} value. Lack of such information will lead to greater uncertainty in its use.

Assay Conduct, Data Analysis, and Data Reporting for Risk Assessment Needs

Understanding the characteristics of the individual assays that will serve as the basis of these predictions is critical when using HTS data. Individual assay characteristics are key regardless of the ways in which the data are ultimately used, which might span the spectrum from combinatorial use in predictive algorithms, test batteries for hazard identification and prioritization, to supporting data for individual chemical risk assessments. Although these uses are potentially diverse, several common assay characteristics will be needed. Some of the specific types of information needs might vary depending on the type of risk assessment to be performed. Minimally, the data reporting should include sufficient information to document assay conduct and

reliability, the rationale for selection of exposure levels, data analysis techniques, and underlying assumptions regarding assay analysis, conduct, or conclusions.

Some advantages of the ToxCast™ data sets are the (1) availability of dose-response information for all assays, (2) availability of assay method details, and (3) availability of the source code for all computational models used in the data analyses. Reliable dose-response information is critical for these types of assays to be useful in risk assessments. Dose-response information is fundamental to understanding the many aspects of chemical toxicity, as it provides a means to evaluate the potency of the chemical and whether a threshold exists.

In conclusion, the current case study was complicated by the multitude of target sites at which the thyroid axis can be disrupted (Murk et al. 2013, Crofton and Zoeller 2005); the secondary, indirect nature of the insult produced; and the complexity of the endpoint of concern—neurodevelopment. By conducting this case study, however, the workgroup could identify not only the nodes in the thyroid toxicity pathway that still need coverage, but also the algorithm development and assay conduct issues that should be addressed if HTS assays are to be used in risk assessments.

4. Advanced Approaches to Recurring Issues in Risk Assessment

In addition to informing chemical specific assessments as discussed above, new data types and advanced approaches also can inform important, recurrent, cross-cutting risk assessment issues. These issues are often sources of controversy due to limited data specific to the issue. A number of these issues are discussed below: variability in human response (e.g., genetic variability, early life exposures; exposure to mixtures and nonchemical stressors); inter-species differences; and characterization of low-level chemical exposures likely to be encountered in the environment. This section discussed how new data types and approaches can inform these difficult issues, thus improving our understanding of public health risks.

4.1. Human Variability

Human response to environmental chemicals is influenced by both intrinsic (e.g., genetics, life stage) and extrinsic (e.g., chemical exposure, stress, nutrition) factors. New methods to examine gene-gene, gene-environment, and epigenome-gene-environment interactions are available (Patel et al. 2013, Lvovs et al. 2012, Meissner 2012, Patel et al. 2012a, Patel et al. 2012b, Baker 2010, Thomas D 2010, Cordell 2009). Zeise et al. (2012) explored how these factors can influence each of the series of biological and physiological steps (known as the source-to-outcome continuum) that ultimately manifests in variability with respect to adverse health outcomes (see Figure 22). The Zeise et al. (2012) review was informed by a National Research Council (NRC) workshop, "Biological Factors that Underlie Individual Susceptibility to Environmental Stressors and Their Implications for Decision-Making." The authors considered current and emerging data streams that are providing new types of information and models relevant for assessing interindividual variability.

Currently, human variability is usually accounted for by including an uncertainty factor of 1, 3, or 10 in the calculation of a reference dose for noncancer health effects. Variability is not explicitly accounted for in cancer health assessment with the exception of the incorporation of an age-specific adjustment factor of ≤ 10 for childhood exposures to genotoxic carcinogens. In a few cases, data on sensitive populations (e.g., asthmatics and those sensitive to air pollutants) might be specifically

incorporated into risk assessments. Figure 23 from Zeise et al. (2012) illustrates how different types of variability can influence dose-response relationships.

Several strategies have been developed to characterize variability in pharmacokinetics (PKs): (1) for data-rich chemicals (such as pharmaceuticals), a "population PK" approach is used to measure variability and discover the determinants; (2) "predictive PK" uses mechanistic models, assigns *a priori* distributions to specific parameters that can be measured experimentally, and uses Monte Carlo simulations to propagate distributions from model parameters to model predictions; and (3) reproduced "Bayesian PBPK" employs a synthesis of the two previous approaches (EPA 2008).

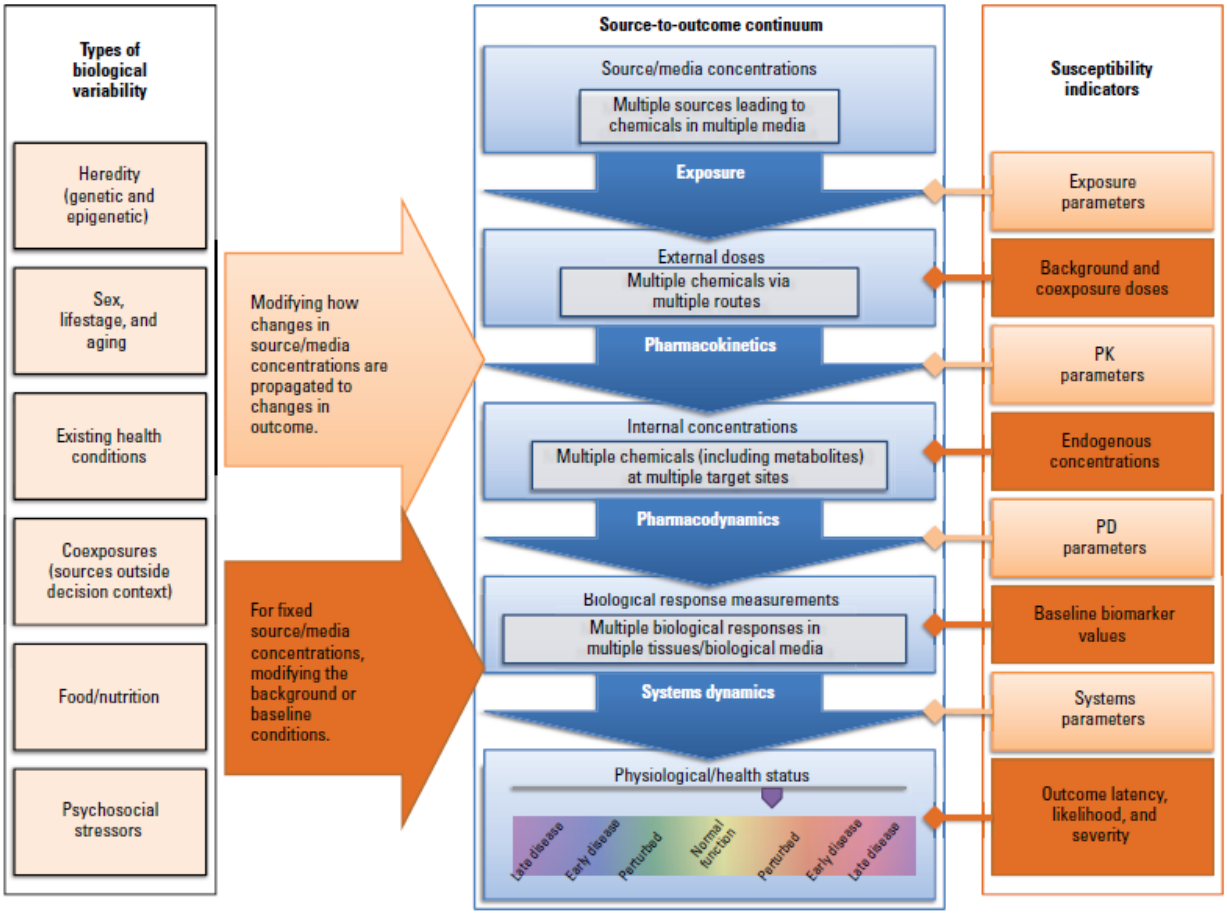


Figure 22. Framework illustration of how susceptibility arises from variability. Multiple types of biological variability intersect with the source-to-outcome continuum, either by modifying how changes to source/media concentrations propagate through to health outcomes, or by modifying the baseline conditions along the continuum. The aggregate result of these modifications is variability in how a risk management decision affects individual health outcomes. The parameters and initial conditions along the source-to-outcome continuum serve as indicators of differential susceptibility, some of which are more or less influential to the overall outcome (see Figure 25) (Zeise et al. 2012). Reproduced with permission from *Environmental Health Perspectives*.

4.1.1. Genomic Variability

An estimated 20%–50% of phenotypic variation is captured when all single nucleotide polymorphisms (SNPs) are considered simultaneously for several complex diseases and traits. The proportion of total variation explained by individual genome-wide-significant variants has reached 10%–20% for a number of diseases (Visscher et al. 2013). Environmental factors are thought to contribute the remaining variability. The interaction between genetic and environmental factors is a key concern in the description of public health risks.

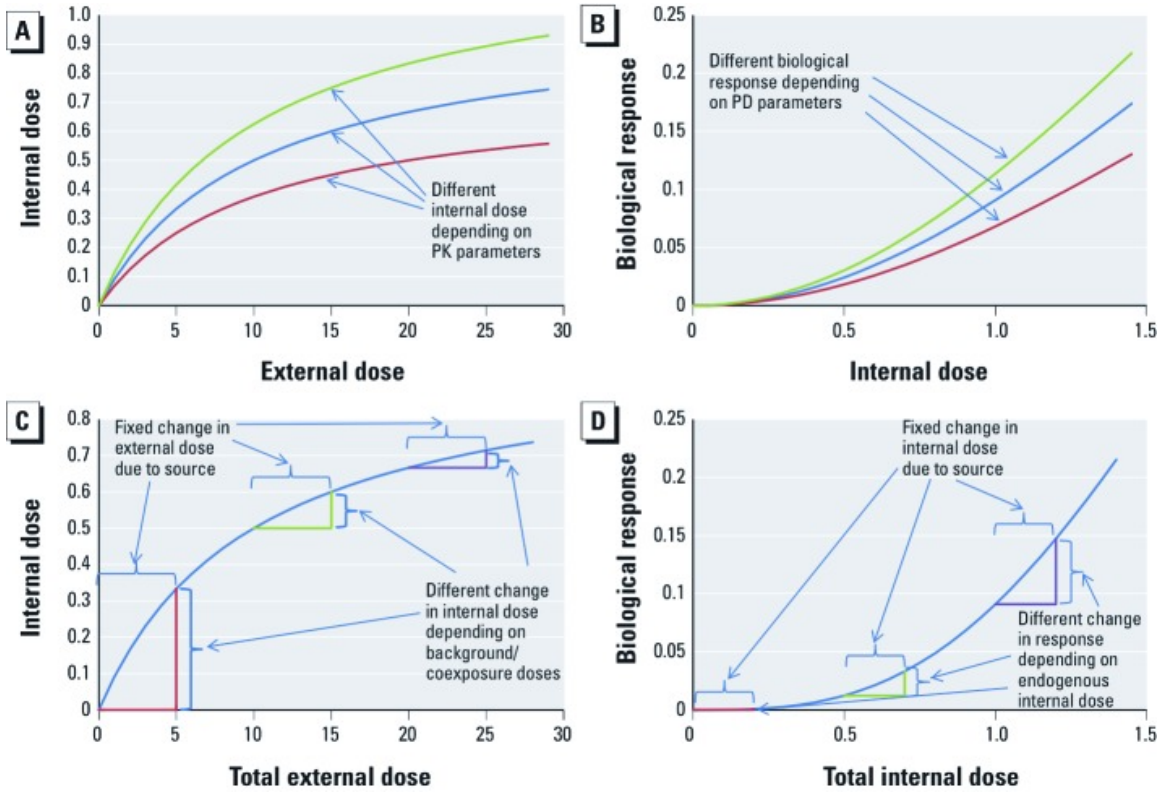


Figure 23. Effects of variability in pharmacokinetics (PK) (A), pharmacodynamics (PD) (B), background/exposures (C), and endogenous concentrations (D). In (A) and (B), individuals differ in PK or PD parameters. In (C) and (D), individuals have different initial baseline conditions (e.g., exposure to sources outside of the risk management decisions context; endogenously produced compounds) (Zeise et al. 2012). Reproduced with permission from *Environmental Health Perspectives*.

Several approaches to generating and evaluating genomic data are now emerging that can provide new insights into human variability (both PK and pharmacodynamic [PD]) including (1) *in silico* modeling approaches in which variability in parameter values is simulated, and differences among subpopulations explored (Shah et al. submitted, Knudsen et al. 2011, Knudsen and DeWoskin 2011, Shah and Wambaugh 2010); (2) high-throughput (HT) *in vitro* data generation using cells lines with different genetic backgrounds (Abdo et al. 2012, Lock et al. 2012, O'Shea et al. 2011); (3) *in vivo* studies in genetically diverse strains of rodents to identify genetic determinants of susceptibility (Harrill et al. 2012, NIEHS 2012a); (4) comprehensive scanning of gene coding regions in panels of diverse individuals to examine the relationships between environmental exposures, interindividual

sequence variation in human genes, and population disease risks (Mortensen and Euling 2013, NIEHS 2012b); (5) genome-wide association studies (GWAS) to uncover genomic loci that might contribute to human risk of disease (NHGRI 2013, Abecasis et al. 2012, Bush and Moore 2012); and (6) association studies that correlate measures of phenotypic differences among diverse populations with expression patterns for groupings of genes based on co-expression (Friend 2013, Patel et al. 2013, Patel et al. 2012a, Weiss et al. 2012). New understanding of the contribution of epigenomics to disease is rapidly advancing with evaluation of changes such as differential methylation of DNA (Teschendorff and Widschwendter 2012, Hansen et al. 2011, Rakyan et al. 2011). Risk assessments of the future will begin to incorporate these types of data as they become available.

Panel a) in Figure 24 illustrates one example of how new types of genetic variation data can be used in risk assessment, in this case, how a population concentration-response curve can be estimated for cycloheximide based on HT *in vitro* data using cell lines with different genetic backgrounds. The approach reported by Lock et al. 2012 is being used in Tox21 Phase II, (in collaboration with Rusyn and colleagues at the University of North Carolina) to expand the study of interindividual differential sensitivity to evaluate approximately 1,100 different human lymphoblastoid cell lines, with densely sequenced genomes representing 9 races of humankind, to 180 toxicants. Data will be collected on more chemicals in the future. The numbers of chemicals evaluated in the future in this manner will expand. The large number of human cell lines used allows for an analysis of genetic determinants associated with differential cytotoxicity *in vitro*. This approach will provide significant new insights into human variability in response and can better inform current and future risk assessments. Other examples of human variability data are discussed in the benzene prototype and in Text Box 10 using GWAS data.[29]

[29]One caveat: The differential risks conferred by human genetic variability are complex and might not be captured by analyses of small-scale gene variability alone. Hundreds to thousands of genes are likely to be involved in any disease, and multiple variations in genetic makeup might confer similar increased or decreased risk for the same disease. The occurrence of disease also could be influenced by emergent system properties that require analysis not only of how gene variations affect cellular components, but how effects on critical network interactions propagate up through higher levels of the biological system (Torkamani et al. 2008). Consequently, although incorporation of new types of data can better characterize human variability, the characterizations are likely to be incomplete.

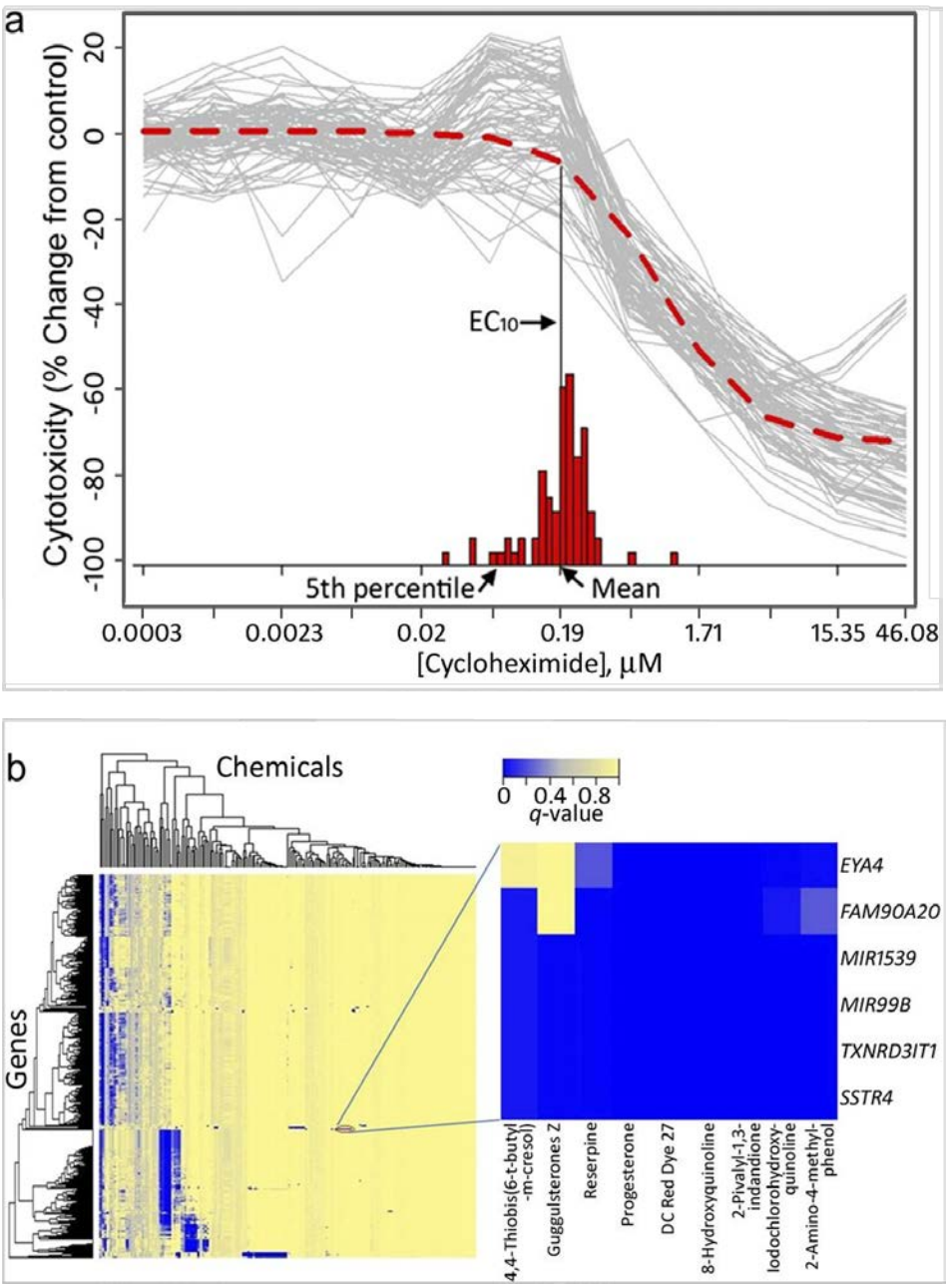


Figure 24. Panel a: A population concentration-response was modeled using *in vitro* quantitative high-throughput screening (qHTS) data using cycloheximide data (cytotoxicity assay) as an example. Logistic dose-response modeling was performed for each individual to the values shown in gray, providing individual 10% effect concentration values (EC_{10}). The EC_{10} values obtained by performing the modeling on average assay values for each concentration (see frequency distribution) are shown in the inset. Panel b: A heat map of clustered FDRs (q values, see color bar) for associations of the data from caspase-3/7 assay with publicly available RNA-Seq expression data on a subset of cell lines. A sample subcluster is shown (Lock et al. 2012). Reproduced with permission from Oxford Journals.

4.1.2. Early-Life Exposures

Early-life chemical exposures can invoke molecular effects that appear to result in increased susceptibility to disease or other morbidity later in life, often via epigenetic modifications (Boekelheide et al. 2012). Evidence from both humans and animals helped establish the influence of early-life exposure on later-life outcomes. For example, human observational data and animal studies report that arsenic exposure during prenatal and early postnatal life increase the risk of cancer, respiratory, and cardiovascular diseases, and neurobehavioral disorders, as supported by human observational data and animal models (Cronican et al. 2013, Boekelheide et al. 2012, Tokar et al. 2012, NRC 2011, Tokar et al. 2011). Later-in-life outcomes can be influenced by time of exposure, species' predisposition to a particular disease, an individual's genetic predilection to disease, or gender. Improved ability to predict disease risk associated with *in utero* or early postnatal exposures results from advances in identifying the targeted genomic region of chemicals/chemical mixtures, epigenetic alteration of gene expression, and the causal links between early-life chemical exposure and later-life outcomes (Boekelheide et al. 2012, NRC 2011).

Epigenetic biomarkers for early-life exposures (e.g., placental epigenetic biomarkers, plasma biomarkers) have the potential for use as early indicators of adverse health effects later in life. Development and interpretation of epigenomic biomarkers is in the early stages of development (Hansen et al. 2011, Rakyan et al. 2011); however, as understanding of the underlying epigenetic mechanisms (e.g., DNA methylation, histone modification, microRNA) advances, more will be known about the relationship between biomarkers of early-life exposure and later-life disease risk. A good example is the work that associated early-life exposure to arsenic and DNA hypomethylation with the development of arsenic-induced skin lesions (Boekelheide et al. 2012). The roles of environmental factors that positively and negatively influence health outcomes require study.

Box 10. Combining Genetics and Bioinformatics to Improve Estimates of Variability in Human Response

Variability in human response to chemical exposures is partly due to genetic influences. The National Center for Biotechnology Information at the National Institutes of Health National Library of Medicine has a vast array of databases devoted to human variability, especially genotype-to-phenotype associations. These resources include dbSNP (database of single nucleotide polymorphisms and estimates of their occurrence within the population), dbGaP (database of Genotypes and Phenotypes), GTEx database (Genotype-Tissue Expression), OMIM (Online Mendelian Inheritance in Man), and PheGenI (Phenotype-Genotype Integrator; aggregates information from many of the aforementioned resources).

In this example, genome-wide association study (GWAS) data were reviewed to examine the relationship between genotype and white blood cell count in benzene-exposed and non-benzene-exposed workers in China. This work has been used, in part, to describe a possible mode of action for benzene hematotoxicity. Lan et al. (2009) identified single nucleotide polymorphisms (SNPs) associated with four DNA repair and genomic maintenance genes that could be involved in carcinogenesis. These SNPs confer significant odds ratios from 1.4 to 5.7 of having a white blood cell count < 4000 cells/µL blood. This observation demonstrates a quantitative increased risk of hematotoxicity in people with any of these SNPs. Hematotoxicity is highly correlated with leukemia resulting from benzene exposure. Hence, these SNPs also might confer susceptibility to leukemia.

PheGenI provides links to dbSNP to view genetic diversity of SNPs within reported populations. For instance, rs12951053's A/C genotype is reported to occur in 51.1 % of Chinese and 31.1% of Japanese; and among Europeans and those of European decent, the A/C genotype occurs in approximately 9–17 % of the population (NCBI 2012a).

Overall, the minor allele (C), has a relatively low penetration within the global population at just 18.7% ± 2.2% (mean ± standard error of the mean), and an average heterozygosity of 30.0% ± 24.5 % (average ± standard error of the mean).

Using the global minor allele rate of 18.7% ± 2.2 %, we can construct a probability function and model that any given member of the population has the minor allele A for rs12951053 SNP. Using this probability function, we can estimate the number of people who might have a white blood cell count < 4,000, thus the potential for hematotoxicity, as well as the model uncertainty. This gives us a quantitative estimate of human health hazard.

In addition, this approach can help with environmental justice issues. For instance, by using census demographic data and the SNP occurrence data for people descended from specific groups, creating probabilistic models that might more accurately reflect the SNP pool of a population, and thus, human variability, is possible. With respect to at-risk populations, regulatory agencies could use this type of information to inform their site-specific risk assessments, such as a Superfund Site Risk Assessment in the United States.

4.1.3. Mixtures and Nonchemical Stressors

Cumulative risk is a function of the exposure to the combined threats from all intrinsic and extrinsic stressors (e.g., chemical exposure, pharmaceutical use, underlying susceptibility, socioeconomic status, work-life stress) and factors that improve health (e.g., good diet, exercise). The assessment of cumulative risk remains a challenging area for human health risk assessment. Only a few studies have examined the potential impact of exposure to environmental chemical mixtures, or to mixtures and nonchemical stressors; and innumerable combinations of chemical mixtures and nonchemical stressors occur in the environment. Conventional methods for risk assessment have made little progress in scaling this particularly mountainous cumulative risk challenge. New methodologies in systems biology, computational models, and data mining provide promise by taking a more comprehensive disease-oriented approach to identification and management of cumulative risk for chemical classes or structures. HTS and omics assay data can be combined with

bioinformatics data mining and computational cellular signaling simulations to predict possible disease outcomes (for screening-level assessments) that, combined with higher level systems data, can identify common patterns of significant pathway or network alterations associated with disease (for more quantitative risk assessments). As our molecular understanding of how nonchemical stressors modulate disease continues to evolve, we will also be able to leverage data from systems biology and network analyses to obtain a better understanding of potential cumulative chemical and nonchemical stressor interactions in biological systems and the resulting health impacts. Because epigenomic networks are more easily modulated by environmental factors than the genome, epigenomics should be considered an area of focus for identifying mechanisms that mediate cumulative risks imposed by exposures to environmental factors (Cortessis et al. 2012, Koturbash et al. 2011, Bollati and Baccarelli 2010).

4.2. Inter-Species Extrapolation

The traditional use of animal models in hazard identification and characterization of dose-response employs chemical testing in mammalian species, and application of an interspecies (animal-to-human) uncertainty factor (≤ 10) or body-weight conversion factor to derive an EPA reference value. Increased understanding of the toxicological or biological pathways and their similarity (or lack thereof) among species will improve the extrapolation of chemical effects across species, and the related challenge of selecting model organisms for testing, in contrast to solely comparing apical responses. As knowledge increases on the extent of pathway conservation among species, alternative test species, including nonmammalian vertebrates (adult and embryonic zebrafish) and invertebrate models, will be of greater use in chemical risk assessment. Regulatory toxicology as a whole will move toward increasing reliance on predictive approaches to assessing chemical risk, with a greater emphasis placed on understanding chemical perturbation(s) of conserved biological pathways at key junctures, including molecular initiating events (MIEs) (e.g., activation or inactivation of specific receptors, enzymes, or transport proteins).

Data from alternative mammalian species and *in vitro* models are valuable for both ecological and human health risk assessment when used in a pathway-based framework (Ankley, G. T. et al. 2010). The extrapolation between species can occur at different levels of biological organization, such as the MIE, the pathway, and the organ or individual levels. Based on the similarity of pathway-based values to standard toxicological values, this appears to be a useful approach for extrapolating hazard values across species, especially if a known pathway is involved.

That gene sequences are conserved—even between distantly related species—is well known and conservation across species is indicative of an essential function. DNA sequence similarity can, but does not always, reflect a functionally conserved role for the genes in question. Investigations of gene function homology can be approached through interspecies comparisons of various components that affect the phenotype in question. The implicated genes, their sequence variation, and the relevant signaling pathways and tissues (cells, organs, circuits) are all informative. Thus, new approaches to understanding the underlying molecular mechanism can improve our cross-species extrapolation (e.g., see Chen et al. (2007), Jubeaux et al. (2012), and Reaume and Sokolowski (2011)).

4.3. Low Dose-Response Modeling

Empirical dose-response models (e.g., benchmark dose [BMD] models) are widely used in environmental health risk assessment for screening and categorization of toxic substances; determination of toxic potency; determination of a point of departure (POD) for low-dose extrapolation; determination of human exposure guidelines; estimation of risk under specific exposure circumstances; and interpretation of human data. Models that are based on a robust understanding of biological processes, in contrast, are not common. Dose-response models could incorporate data from *in vitro* studies, human or test animal *in vivo* studies, or human epidemiology studies. For public and ecosystem health risk assessment, characterizing population-level responses is the goal.

Many risk assessments require models that can extrapolate beyond the data set used in developing the model to derive the toxicity values of interest. Such models are called biologically based models. To date, the main biologically based models used in risk assessment are physiologically based toxicokinetic (PBTK) models that simulate the toxicokinetic behavior of a chemical (i.e., the internal disposition of the chemical in the body following a given dosing regimen). Only a few examples of physiologically based toxicodynamic (PBTD) models are available to characterize the "response" side of the dose-response curve. Well-developed and adequately tested PBTK models are currently used in risk assessment to simulate the toxicokinetics of a chemical or chemicals across dosing regimens (duration, amounts, delivery rate, routes) and species, or from *in vitro* regimens to *in vivo* doses (IVIVE).

The establishment of human exposure guidelines for environmental agents involves determining a POD on the dose-response curve, such as a particular response level on a BMD model estimate of the dose-response, corresponding to a specified increase in risk usually in the 5% to 10% range, or signal-to-noise-crossover dose introduced by Sand et al. (2011). This POD is then further reduced by adjustment factors to derive a level of exposure that is considered to be protective of human health and the environment. The National Research Council (2009) suggests an integrated approach to the establishment of human exposure guidelines using adjustment factors applied to the POD, where the magnitude of the factor depends on the "expected" behavior of the exposure-response curve at low levels of exposure. The NRC also examined the influence of background exposures and background disease rates on the shape of the exposure-response curve at low levels of exposure.

Characterizing the expected response at low exposure levels (i.e., those that the public is most likely to encounter) is another of the great challenges to previous methods used in risk assessment, specifically the use of relatively high-dose *in vivo* animal assays as the source of data for apical endpoints because the spectrum of adverse effects might be quite different at lower doses. The NRC (2007) recommended developing new approaches and models to generate the data needed for characterizing dose-response curves and to improve estimates especially at doses applicable to likely human exposures. Examples of some new approaches to dose-response modeling are described in Burgoon and Zacharewski (2008), Parham et al. (2009), and Zhang et al. (2010). The application of sensitive HTS assays for pathway perturbations that directly measure biological effects at environmental exposure levels are described in Rotroff et al. (2010) and Wetmore et al. (2012). The reduced cost of HTS assays relative to mammalian toxicity tests might also permit the use of a much broader range of exposure levels, leading to a more detailed description of dose-

response relationships throughout the exposure range of interest. Figure 25 summarizes the automated dose-response modeling approach proposed by Burgoon and Zacharewski (2008).

A new class of biologically based models called "virtual models" is also being developed to simulate normal biology and to predict how chemical perturbations might lead to adverse effects (i.e., to predict a chemical's toxicodynamics) based on knowledge of potential mechanisms. Examples of virtual models being developed at various levels of biological organization or function include (1) the Physiome Project (Physiome Project 2013), a major resource and model repository of hundreds of physiology models (Hunter et al. 2002); (2) the European Virtual Physiological Human (VPH) project (Hunter et al. 2010); (3) HumMod, a whole-body integrated human physiology model (Hester et al. 2011); (4) Virtual Cell (V-Cell), a spatially realistic quantitative model of intracellular dynamics (Moraru et al. 2008); (5) EPA's Virtual Embryo™ (v-Embryo) project, a suite of models that simulate normal development leading to the formation of blood vessels, limb-buds, reproductive systems, and eye and neural differentiation (Knudsen et al. 2011, Knudsen and DeWoskin 2011); (6) EPA's Virtual Liver™ (v-Liver) model that simulates the dynamic interactions in the liver used to translate *in vitro* endpoints into predictions of low-dose chronic *in vivo* effects in humans (Shah and Wambaugh 2010); and (7) the Virtual Liver Network (German Federal Ministry for Education and Research 2013), a German initiative to develop a dynamic model of human liver physiology, morphology, and function integrating quantitative data from all levels of organization (Holzhutter et al. 2012).

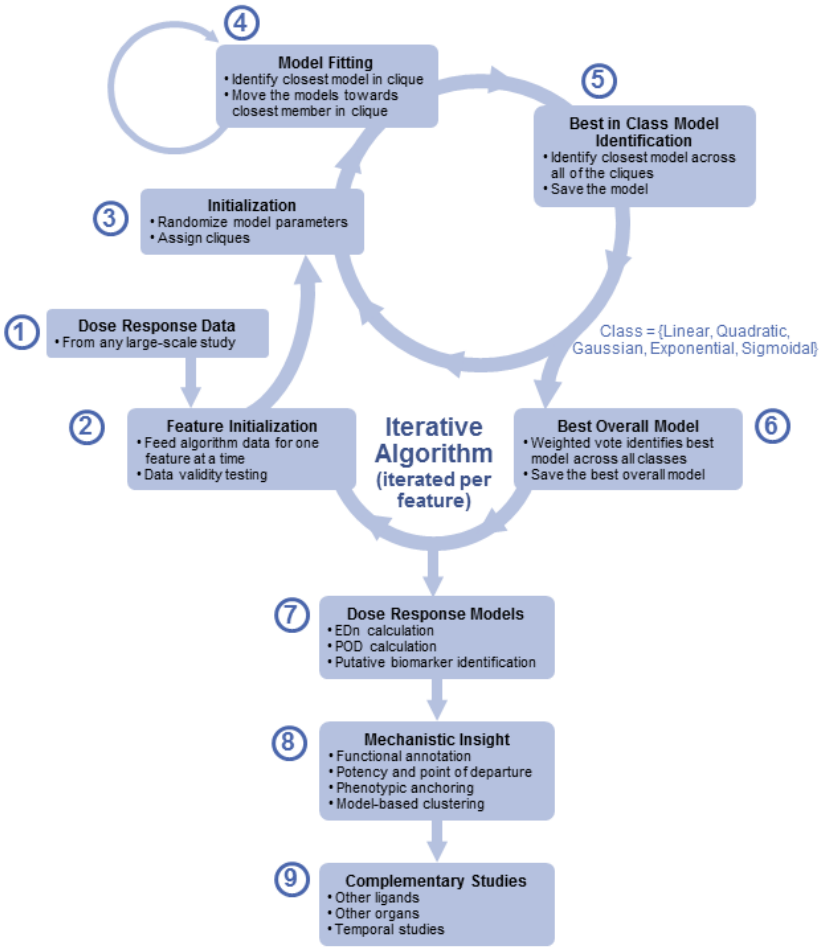


Figure 25. Overview of automated dose-response modeling from Burgoon and Zacharewski (2008). Step 1—Dose-response data from a large-scale study are loaded. Step 2—The application feeds dose-response data for one feature into the algorithm. Examples of feature data include mRNA, protein, or metabolite levels and enzyme or binding activities at each dose within a study. Step 3—The application initializes the particle swarm optimization (PSO) algorithm by randomizing model parameters and assigning cliques. Step 4—The PSO identifies the closest model in each clique at the end of an iteration, and moves the members of each clique toward that model. Step 5—This iterative process ends once a best-fit model has been identified, or when all of the iterations have been used. Steps 3 through 5 are repeated for each model class for the same feature, thus generating best-fit models for the exponential, Gaussian, quadratic, linear, and sigmoidal classes. Step 6—The best exponential, Gaussian, quadratic, sigmoidal, and linear models are compared with the best overall model using a weighted vote method. The model with the smallest Euclidean distance compared with the dose-response data receives the most votes. Step 7—The application uses the best overall model to calculate EDn and point of departure (POD) values, used to rank and prioritize putative biomarkers or chemical activities. Step 8—Model-based clusters can provide additional mechanistic insight by integrating potency and POD data with functional annotation and phenotypic anchoring. For example, EDn and POD data might generate model-based clusters for lipid metabolism and transport gene expression that could be associated with the occurrence of hepatic vacuolization and lipid accumulation. Step 9—Through complementary comparative studies using toxic and nontoxic congeners in responsive and nonresponsive species across time, data could emerge that differentiate biomarkers of exposure from toxicity-related responses that can support mechanistically based quantitative risk assessments. Reproduced with permission from Oxford Journals.

5. Lessons Learned from Developing the Prototypes

The NexGen prototypes presented in this report illustrate new data types and approaches applied to risk assessments for various decision contexts and provide concrete examples for discussion of new approaches in the risk assessment community. The prototypes show how large amounts of data can be synthesized in useful ways, and how the data can increase our understanding of the potential risk posed by chemical exposures, including hazard identification, and dose-response assessment. In addition, the outcomes of the prototype assessments have implications for current challenges in risk assessment, such as evaluating human variability and susceptibility, mixtures, and low-dose responses characterization. Overall, these new data types and approaches appear promising in terms of improved risk assessment and decision support. These new approaches are faster and less expensive than traditional approaches provide new insights. Each prototype considered hazard identification, exposure-dose-response, and mechanisms of action or adverse outcome pathways.

5.1. New Methods

Thousands of chemicals to which humans are exposed have inadequate data for predicting their potential for toxicological effects. Dramatic technological advances in molecular and systems biology, computational toxicology, and bioinformatics, however, have provided researchers and regulators with powerful new public health tools (NRC 2007, 2006). "High-throughput screening techniques are now routinely used in conjunction with computational methods and information technology to probe how chemicals interact with biological systems, both *in vitro* and *in vivo*. Progress is being made in recognizing the patterns of response in genes and pathways induced by certain chemicals or chemical classes that might be predictive of adverse health outcomes in humans. However, as with any new technology, both the reliability and the relevance of the approach need to be demonstrated in the context of current knowledge and practice" (Tice et al. 2013).

In general, two basic approaches are being taken to advance our understanding of the causes and modifiers of human disease risks: top-down and bottom-up (Friend 2013). In general, the top-down approach focuses on developing large-scale network models of disease by sifting through the substantial body of new human molecular clinical and epidemiologic data (e.g., >50,000 omics papers per year; zettabytes (10^{21}) of new data), looking for patterns associated with various disease states and environmental or genetic risk factors. In general, the bottom-up approach focuses on using *in vitro* high- and medium-throughput bioassays to understand alteration in molecular and cellular processes caused by chemical exposures. The top-down approach provides the human population biology context, while the bottom-up approach provides experimental support for associations identified in the top-down approach. The approaches are mutually supportive and, when integrated, provide a powerful means to advance risk assessment. The various prototypes presented in this document sought to illustrate both approaches. Due to the greater current availability of genomic data, the prototypes were heavily biased toward use of gene expression and transcriptomic data (i.e., gene expression levels and factors influencing transcription into proteins).

The results from large and essential areas of research, including epigenomics, are rapidly adding to our knowledge and will be incorporated in more detail in future efforts. [30]

The three sets or tiers of prototypes were intended to explore different aspects of decision context. The primary intent of the first set of chemicals (Tier 3 prototypes) was to verify if and how new data and approaches could be used to inform risk assessment by comparison to robust traditional assessments where risks are generally considered "known." In essence, we attempted to reverse engineer from known answers to verify new approaches, explore value information, and begin to characterize decision rules that could be reasonably applied to chemicals with limited or no traditional data. Secondarily, the Tier 3 prototypes explored how new types of data could expand our understanding of well-studied chemicals. The intent of the Tier 2 prototypes was to explore new types of computational analyses and short-duration *in vivo* bioassays that are intermediate in terms of required resources and confidence in the data between Tiers 3 and 1, and are suitable for evaluating hundreds to thousands of chemicals. These approaches are relatively uncommon in risk assessment to date but hold much promise. The intent of the Tier 1 prototypes was to explore entirely HT approaches that could be applied to tens of thousands of chemicals, might have the greatest uncertainties, but are the least resource intensive to use.

The following eight chemicals or chemical classes and their associated effects were chosen for prototype development:

- Tier 3:
    - Benzene and leukemia (molecular epidemiology),
    - Ozone and lung inflammation and injury (molecular clinical studies), and
    - Benzo[a]pyrene (B[a]P, a polycyclic aromatic hydrocarbon (PAH) and liver cancer (molecular clinical studies meta-analyses and *in vivo* rodent bioassay).

[30]In terms of top-down approaches, molecular, computational, and systems biology data have grown phenomenally in recent years, and have informed mechanisms of disease and factors that alter risks of disease. These data are generally stored in large databases such as ENCODE, Gene Expression Omnibus (GEO), and the Comparative Toxicogenomic Database (CTD) and are publicly available for further analyses. Analyses and meta-analyses of these data are providing new insights into environmental public health risks. Bioinformatics (computer-assisted approaches) are necessary to use these new data effectively due to the size of the relevant databases. The polycyclic aromatic hydrocarbon (PAH) and diabetes prototypes, in particular, illustrate bioinformatic "knowledge mining" to understand environmentally related disease.

In terms of bottom-up approaches, many new high- and medium-throughput methods have been and are being developed that facilitate testing and evaluation of chemicals on an unprecedented scale. In particular, the *in vitro* evaluations of chemicals with limited or no traditional data are being enabled. ToxCast™ and Toxicology in the 21st Century (Tox21) provide examples (see Section 3.3). Tox21 will test ~10,000 chemicals in a few years. New *in vivo* short-duration (hours to weeks) exposure paradigms also are emerging that provide new types of data to be used in health assessments. These paradigms use both nonmammalian (see Section 3.2.2) and mammalian species (see Section 3.2.3).

- Tier 2:
    - Chemicals associated with diabetes and obesity ("big data" knowledge mining),
    - Chemicals associated with thyroid hormone disruption (short-duration *in vivo* exposure bioassays – alternative species), and
    - Chemicals associated with cancer (short-duration *in vivo* exposure bioassays – mammalian).
- Tier 1:
    - Chemicals associated with cancer and noncancer disorders, especially developmental (QSAR) and
    - Chemicals associated with thyroid hormone disruption (high-throughput *in vitro* assays).

Table 9 provides more information on methods explored in each prototype (Krewski et al. 2013).

5.2. 5.2. Implications for Risk Assessment Derived from Prototypes

Based on the prototypes provided here and the work of others, new molecular, computational, and systems biology tools likely can better inform risk assessment. Substantial caution in interpretation and use of new information is warranted, however, in large part because our understanding of the science is still evolving, and appropriate data are still scarce. We propose initially to use new methods discussed in this document to: (1) generate hypotheses; (2) screen and rank chemicals for additional research and assessment; and (3) augment understanding of traditional data. Areas of particular promise include improved understanding of relative potency of chemicals to disrupt biologic processes, hazard identification, and mechanisms of disease and disorders; human variability and susceptibility; human relevancy of animal models; and low-dose-response relationships. These future risk assessments ideally would rely on the integration of a variety of new types of data and traditional data, as available. Additional discussion of the lessons learned from the prototypes follows.

Systems biology context is key to understanding these new data types and the relationship among various types of data. Network-level understanding is typically more informative than pathway-level understanding, which is usually more informative than individual genes. In general, information on individual genes, in the absence of systems biology-level of understanding, is likely to be inadequate for risk assessment purposes. Information that links molecular events to apical outcomes need not be chemical specific, but can be derived from mechanistic information on disease or from related chemicals. As with any risk assessment, the studies used should be well designed, conducted, and reported; systematic review criteria are necessary in study selection. Characterization of multisource variability is a substantial challenge with new data types because of the sheer amount of data being analyzed and, thus, must be carefully considered. Also, traditional weight-of-evidence criteria continue to be useful in considering new data types, for example, data from multiple, similar studies are preferred (Krauth et al. 2013). That environment-induced changes in biology are dynamic in nature also should be noted, and these dynamic changes are not well understood.

Highlights of the prototypes include:

- The effects of human chemical exposures at environmental levels on molecular events were linked to intermediate biological events and apical adverse outcomes using molecular epidemiology, molecular clinical, and environment-wide association studies (e.g., evaluation of NHANES) (EPA 2013, Patel et al. 2013, Devlin 2012, McHale et al. 2012, Patel et al. 2012a, Thayer et al. 2012, Burgoon 2011, McHale et al. 2011, Smith, MT et al. 2011). Chemicals evaluated included benzene, ozone, B[a]P (a PAH), metals, and persistent organic pollutants.
- The B[a]P and diabetes prototypes illustrated the use of "big data" knowledge mining to identify associations between environmental chemical exposures and disease (Patel et al. 2013, Patel et al. 2012a, Burgoon 2011). The chemicals of concern for diabetes identified using knowledge mining also were identified in a review of traditional literature by experts (Thayer et al. 2012). This powerful, relatively new technique has not been used extensively in environmental risk assessment, although it is commonly used in other areas of biology. Knowledge mining is particularly useful in developing a broad understanding of potential mechanisms of action, factors that may cause or modify disease risks, and human variability and susceptibility.
- Short-duration exposures coupled with new molecular and computational approaches appear to provide additional insights into potential environmental risks. Use of both alternative species and mammalian species in these new experimental models is explored. These models are faster and less expensive than the molecular epidemiology and molecular clinical studies noted above. Furthermore, unlike the QSAR and high-throughput (HT) models noted below, these models address intact metabolism and cell, and tissue interactions and can be used to study more complex outcomes such as developmental and neurobehavioral outcomes. In the case of alternative species, these models can detect effects over the entire lifespan of the organism and to population dynamics. These models have been used successfully to describe mechanisms, explore complex mechanistic behaviors, describe hazards, and evaluate chemical potency. Confidence in the data also generally lies between Tier 3 and Tier 2 approaches (Perkins et al. (2013), Thomas RS et al. (2013a), and Padilla et al. (2012).

Table 9. Prototype Use of New Scientific Tools and Techniques (Krewski et al. 2013)

Scientific Tools Used in Specific Prototypes	Tier 1: Screening and Prioritization		Tier 2: Limited Scope Assessments		Tier 3: Major Scope Assessments	
	Cancer & Hydrocarbon Mixtures (QSAR)	*Endocrine Disruptors & Deep Water Horizon Oil Spill Dispersants (In Vitro Bioassays)*	*Diabetes & Multiple Stressors (Knowledge Mining)*	*Cancer & Reproductive/ Developmental Hazards (Short-Duration In Vivo Exposure Bioassays)*	*Lung Injury & Ozone (Molecular Clinical)*	*Leukemia & Benzene (Molecular Epidemiology)*
Hazard Identification and Dose-Response Estimation Methods						
Quantitative structure-activity relationships	■	■		■		
High-throughput *in vitro* assays		■	■	■	■	■
High-content omic assays				■	■	■
Molecular and genetic population-based studies					■	■
Biomarkers of effect			■		■	■
Pathway/network analyses	■	■	■	■	■	■
Dosimetry and Exposure Assessment Methods						
In-vitro-to-*in-vivo* extrapolation	■	■		■		
Pharmacokinetic models and dosimetry		■		■	■	■
Biomarkers of exposure			■		■	■
Cross-cutting Disciplines						
Bioinformatics/ computational biology	■	■	■	■	■	■
Functional genomics			■	■	■	■
Systems biology			■	■	■	■

- QSAR models (Venkatapathy and Wang 2013, Goldsmith et al. 2012, 2012a, Wang, N et al. 2012b) and HT *in vitro* bioassays are being used to rapidly evaluate a wide array of chemicals (Judson et al. 2013, 2011, Sipes et al. 2013, Tice et al. 2013, Kavlock et al. 2012, Rusyn et al. 2012). "These tools can probe chemical–biological interactions at fundamental levels, focusing on the molecular and cellular pathways that are targets of chemical disruption" (Kavlock et al. 2012). Thousands of chemicals are currently being evaluated, particularly in the ToxCast and Tox21 programs. Both estimates of potency and insights into potential hazards are being generated. Additionally, tools exist to relate *in vitro* concentration to potential human exposure levels (reverse dosimetry) (Wetmore et al. 2013, Wetmore et al. 2012, Rotroff et al. 2010, Hubal 2009). Although directly correlating *in vitro* findings to risks of human disease is difficult, these QSAR and HT methods provide powerful new tools for screening and ranking large numbers of chemicals for further evaluation and assessment, as well as exploring underlying mechanisms of toxicity, and evaluating human variability in response to chemical exposures (Lock et al. 2012).

Thomas RS et al. (2013a) propose a framework for incorporating these new technologies into toxicity testing and risk assessment in an integrated fashion. The first steps proposed are to use *in vitro* assays to separate chemicals based on their relative selectivity in interacting with biological targets and to identify the concentration at which these interactions occur. Dosimetry modeling converts *in vitro* concentrations into external dose for calculation of the point-of-departure (POD) and comparisons to human exposure estimates to yield a margin of exposure (MOE). The second step involves short-term *in vivo* studies, expanded pharmacokinetic evaluations, and refined human exposure estimates, thus increasing confidence in the evaluation. The third step represents the traditional animal studies currently used to assess chemical risks. A significant percentage of chemicals evaluated in the first two tiers could be eliminated from further testing based on their MOE. Additionally, at each step, information might be suitable for supporting some types of Agency decision-making. The framework provides a risk-based and animal-sparing approach for evaluating chemicals using technological advances to increase efficiency.

In addition to informing hazard identification and dose-response, new data types and methods have the potential to inform recurrent, challenging risk assessment issues.

Experimental Low Dose Data vs. Low Dose Extrapolation – Dose-dependent molecular changes associated with adverse outcomes can be observed at environmental concentrations. Thus, these new approaches can provide experimental data to help characterize dose-response relationships at concentrations where responses, to date, have often been only inferred. Both assay methods and statistical analyses must demonstrate sufficient sensitivity to be considered informative. Observed molecular changes include changes in both magnitude and character, reflecting underlying alterations in biology with increasing dose and time. Biological processes that are consistently observed across the exposure range of interest are likely to be the most useful as biomarkers of exposure and effect. Elucidating the meaning of these dynamic changes in terms of risk will be challenging.

Variability and Susceptibility – New data and methods can enhance our ability to understand variability in response and the identification of potentially susceptible populations. Human cells from various individuals (e.g., 1000 Genome Project) evaluated in *in vitro* high-throughput models provide an avenue for understanding responses across subsets of the human population (Lock et al. 2012). Data mining and bioinformatics analyses will facilitate the identification of susceptible

populations and underlying sources of variability by combing existing molecular epidemiology and clinical databases. In all, this work can provide quantitative data, which to date have been generally lacking, to support more accurate estimates of human variability and identification of susceptible populations.

Evaluation of the Effects of Multiple Stressors – The ability to map mechanism of disease and adverse outcome pathways disrupted by various environmental agents gives us new tools for understanding the interactions of multiple environmental stressors, including chemical mixtures and lifestyle factors.

Certain caveats that apply generally to use of new data types in risk assessment deserve mention.

- Cell type, tissue, individual, subpopulation, species, and test system can alter how specific omics are expressed as traditional intermediate and apical outcomes, even when the molecular signature is the same. This is likely due, at least in part, to epigenomic differences and genomic plasticity. This issue should be considered, as feasible, in data interpretation.
- The metabolism of many chemicals often plays an important role in toxicity. That most HT test systems are not metabolically competent is important to consider. Various approaches to the issue of *in vitro* metabolism are being evaluated; however, this currently remains a complicating factor in most *in vitro* testing.
- Molecular profiles appear time-dependent, that is, they evolve over time with continued exposure and post-exposure. This can confound prediction of outcomes or disease outcomes based on "snapshots" in time of biological events. Fortunately, however, at least some signatures appear to stabilize over time and can serve as reliable indicators of chronic outcomes.
- Currently, studying multiple molecular processes (i.e., genomics, transcriptomics, proteomics, and epigenomics) in a single study is relatively rare, primarily due to expense. This lack of biological integration limits our understanding.
- Due primarily to experimental design and reporting issues (see B[a]P [a PAH] prototype), adequate data from the open literature to support risk assessment activities currently are available for few chemicals. This underscores the importance of high-quality research and testing programs like ToxCast™ and Tox21 and systematic review of data.
- Data reproducibility and false negative rates may remain a potential limitation of high throughput screening and high content assays (e.g., toxicogenomics). The false negative rate (i.e., calling a chemical non-toxic when it is) tends to decrease as an increasing number of independent replicates are used. Successful screening programs need low false negative rates, while balancing their efficiencies (i.e., cost, time, throughput).

The challenge is to use what we know today wisely, with the understanding that biological knowledge is evolving very rapidly, and likewise, risk assessment also will need to evolve.

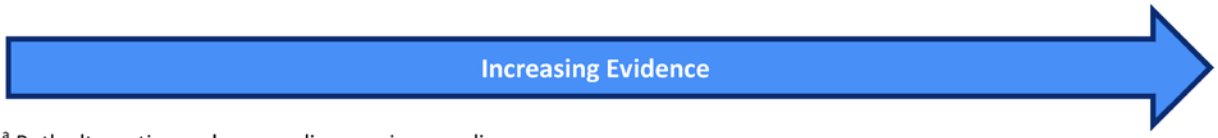
5.3. Summary

Throughout this report, examples are provided that illustrate how new types of data might be used to improve risk assessment. Table 10 summarizes (1) various decision context examples common at EPA; (2) a "toolbox" of various NexGen methodologies that could provide data to support each decision context; (3) types of "fit for purpose" toxicity values that might be derived from new data

types or traditional data; and (4) assessment products in which molecularly or computationally informed toxicity values could be used. Although all approaches can be used in any type of assessment, any one of the health data approaches listed in Tiers 1 and 2 could provide a minimum data set. In this scheme, Tier 1 is primarily QSAR or HT data-driven. Tier 2 is high-content or traditional data-driven (in addition to Tier 1 data, if available). Tier 3 will continue to be traditional data driven but could be augmented by molecular, computational, and systems biology data if the data are available, of sufficient quality, and substantively useful.

Table 10. Problem Formulation Table

PROBLEM FORMULATION			
	Tier1: Screening and Prioritization	**Tier 2: Limited Scope Assessments**	**Tier 3: Major Scope Assessments**
Decision Context Examples	• Emergency response • Unregulated drinking water chemicals identification • Potential emerging chemical problems or opportunities • Research directions	• National Air Toxics Assessment • Superfund listing/removal actions • Drinking Water Health Advisories	• National Regulatory Decisions • International, Tribal, State, & Local Technical Support
Toolbox of Possible Approaches	• QSAR • High-throughput (HT) Screening Assays • Computational Toxicology Models • No Traditional Data • Automated Data Integration	• High-content Assays ➢ Knowledge Mining ➢ Short Duration *In Vivo* Exposure Paradigms[a] • Limited Traditional Data[b] • Automated Data Integration	• Molecular Biology Data • Systems Biology Data • All Policy Relevant Data • Hand-Curated Data Integration
Possible Types of Toxicity Values	High-Throughput Toxicity Values	High-content Toxicity Values	Molecularly Informed Traditional Type Values
Health Assessment Categories	Prioritized Chemicals of Concern List; Screening Values	Provisional Toxicity Values	IRIS or ISA
Exposure Assessment	Physical-Chemical Surrogates	Limited Exposure Data	Extensive Exposure Data



[a] Both alternative and mammalian species paradigms.
[b] Potentially not chemical-specific data but rather disease or chemical class data.

Integration of information from multiple data types is preferred, but all types of data shown for any tier might not be available or of sufficient quality for inclusion in an assessment. Systematic review criteria are being established and are discussed in section 3.1.3 (McConnell and Bell 2013). Stakeholder input and external peer review will be solicited for new approaches to risk assessment.

Systems biology understanding is a fundamental aspect of the weight-of-evidence evaluation. As one progresses from Tier 1 to Tier 3 assessments, the weight of evidence increases; however, the

resources to generate the assessments also increases. For example, in Tier 1, toxicity values can be generated solely from extant QSAR data, a process that can be fully automated to be very quick and cost-efficient for a large number of chemicals. Wignall et al. (2013)(SOT poster abstract; manuscript in progress) describe an approach to generate toxicity values for chemicals with limited experimental data using a combination of QSAR, regression, and hybrid modeling (Rusyn et al. 2012), and incorporating Organization of Economic Co-operation and Development (OECD) principles for model building and external cross-validation. Tier 2 type assessments, ideally, enable the use of more types of data to inform our understanding of data-limited chemicals. For Tier 2, EPA is beginning to develop high-content toxicity values on a trial basis. High-content toxicity values can be developed using bioinformatic approaches based on data that can be machine read and, hence, readily mined and analyzed. Additional data integration and hand curation might be needed to use available data resources, such as high-throughput screening (HTS)/high-content screening (HCS) assays, alternative species testing, and study data compiled the European Union's Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulation. Tier 3 assessments will continue to be driven by traditional data, but new data types could provide new insights into difficult issues such as low dose-response, human variability and susceptibility, and the effects of multiple environmental stressors. These various "fit for purpose" assessment types can be used to develop hypotheses, screen chemicals, mechanistically fingerprint toxicants, set priorities, and inform hazards, relative potencies, and risks.

6. Conclusions

6.1. Challenges

Novel data streams and approaches are rapidly emerging that present opportunities for informing and supporting human health risk assessment, but challenges remain. Four key challenges are the need for (1) the ability to predict metabolism of test compounds, (2) improved understanding of the biology from a systems perspective; (3) evaluated methods to measure key aspects of biological space across multiple levels of organization; and (4) the knowledge infrastructure to ensure availability of relevant data. Future directions include filling these scientific gaps and continuing to build the framework for incorporating new information fit-for-purpose into assessments to support a range of decisions to promote health, protect the environment, and manage risks.

Arguably, the greatest challenge is posed by the need to consider and evaluate complex interactions of chemical and biological systems to predict potential for health risks. Systems biology provides an approach for investigating emergent properties in complex chemical-biological systems by probing how changes in one part affect the others, and the behavior of the whole. New data types are providing required information to develop these predictive models.

There is an imbalance, however, in the sophistication of methods available and the resolution of data being developed to evaluate impacts of chemical perturbations and to discover mechanistic commonalities. Large amounts of network or high-throughput screening/high-content data can be collected to measure effects at the molecular level. Substantial information is also available on disease outcomes, yet only very sparse data are being generated on intermediate events. A similar lack of exposure information commensurate with hazard data is also evident. Even given the rich data coming from implementation of the high-throughput (HT) toxicity testing schemes, gaps in coverage for key endpoints occur, and thus, developing and incorporating assays are needed to fill gaps in the biology required to assess potential for the full range of adverse outcomes required by

risk assessors. This discrepancy in available data across levels of biological organization should narrow over time, as methods continue to advance and as more metabolomics data for biomarkers of effects and exposure are made available. This will lead to development of models that predict disease outcomes with greater certainty from initiating events in individuals and populations relative to exposures likely to be experienced in the real environment.

6.2. Next Steps

Future plans for facilitating use of new data types and tools to support the full range of risk-based assessments and decisions include addressing needs for validated testing schemes and clearly articulating decision considerations for incorporating results of these analyses. In addition, further prototypes or case examples for incorporating HT toxicity data and other novel data types to inform risk assessment are required to demonstrate the added value of these advanced tools and to identify further the most significant scientific gaps.

Validation of HT toxicity testing schemes will be necessary if the data developed using these methods are to be used to inform risk-based decisions and to support efficient chemical risk assessments. The key to moving the wealth of information being generated through research efforts such as ToxCast™ and Toxicology in the 21st Century (Tox21) is to develop a framework for validating HT toxicity testing schemes to support specific chemical evaluation objectives. Traditional "validation" schemes designed to evaluate conventional assay and testing structures do not adequately address this gap and would take years to implement. As the technology for providing rapid, efficient, robust hazard and effects data continues to advance, the validation process for evaluating these new methods is also expected to undergo a transformation to provide fit-for-purpose confidence in results. Future incorporation of new types information to improve the scientific basis and efficiency of risk assessment requires clear articulation of decision considerations for using new types of data and methods. Some of these decision considerations might have standard principles supported by a broad range of risk managers and stakeholders, while others will need to be fit-for-purpose. Early consideration of these decision considerations has been initiated and plans are in place to develop criteria for systematic review of new types of data, disease signatures, adequate weight of evidence for use in risk assessment, and new approaches for risk assessment.

Demonstrating approaches for incorporating new molecular biology data and evaluating advanced methods might be facilitated by additional case examples and prototypes. Conducting a variety of case studies focused on using the HT toxicity data from ToxCast™ and Tox21, in combination with other chemical-specific information to improve efficiency of risk-based decisions where little traditional toxicity data are available, will be important for assessing the value added of these new data types.

Examples also will be identified where molecular biology data can be considered for Tier 3 assessments to augment traditional assessment methodologies. These will provide opportunities to solicit public comment and peer review.

Opportunities also exist for using new data types to guide development of NexGen approaches by considering prototypes for how this information could support some of the most challenging questions faced by risk managers. Population-level risks could be considered using both traditional and molecular biology data, with an additional emphasis on epigenomics and influences of broadly defined environmental factors. Additional insights for risk managers can be found in Crawford-

Brown (2013). Application of new methods might better inform our understanding of the combined effects of multiple stressors, such as multiple chemical exposures, diet, stress, and pre-existing disease. In recognition of the tremendous potential for these new methods and data types to support risk assessment, the EPA Office of Research and Development will continue to elaborate the NexGen framework, and begin to develop toxicity values informed by new biology for specific risk assessment purposes.

- EPA's Office of Research and Development will work with EPA's Program Offices using Tier 1 screening and prioritization approaches to queue up new assessments. Results from this work will be used to feed back into the testing paradigm for its refinement.
- Toxicity values informed by new types of knowledge will be developed in each tier to address needs from screening chemicals for future testing to assessment for potency or category of adverse effect.
- Levels of confidence in those values will be characterized depending on the types and quality of the supporting data.

EPA's Office of Research and Development will expand stakeholder discussion and the community of practice with regard to the use of new data types and methods in risk assessment, and the peer review of new methods. New assessments will receive public comment and peer review.

Finally, EPA's Office of Research and Development will continue working with other national and international agencies involved in assessment, testing, and research to coordinate and harmonize activities, and improve data collection, analyses, curation, sharing, and warehousing.

7. References

Abdo N, Xia M, Kosyk O, Huang R, Sakamuru S, Austin CP, et al. (2012). The 1000 genomes toxicity screening project: Utilizing the power of human genome variation for population-scale in vitro testing. *51st Annual Meeting of the Society of Toxicology* Poster Abstract #765. http://www.toxicology.org/AI/Pub/Tox/2012Tox.pdf.

Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, et al. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* 491: 56-65. http://www.ncbi.nlm.nih.gov/pubmed/23128226.

Afridi HI, Kazi TG, Kazi N, Jamali MK, Arain MB, Jalbani N, et al. (2008). Evaluation of status of toxic metals in biological samples of diabetes mellitus patients. *Diabetes Research & Clinical Practice* 80: 280-288. http://www.ncbi.nlm.nih.gov/pubmed/18276029.

Andersen S, Pedersen KM, Bruun NH, Laurberg P. (2002). Narrow individual variations in serum T(4) and T(3) in normal subjects: A clue to the understanding of subclinical thyroid disease. *Journal of Clinical Endocrinology & Metabolism* 87: 1068-1072. http://www.ncbi.nlm.nih.gov/pubmed/11889165.

Angerer J, Bird MG, Burke TA, Doerrer NG, Needham L, Robison SH, et al. (2006). Strategic biomonitoring initiatives: moving the science forward. *Toxicological Sciences* 93: 3-10. http://www.ncbi.nlm.nih.gov/pubmed/16785253.

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, et al. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. *Environmental Toxicology and Chemistry* 29: 730-741. http://service004.hpc.ncsu.edu/toxicology/websites/journalclub/linked_files/Fall10/Environ%20Toxicol%20Chem%202010%20Ankley.pdf.

Ankley GT, Gray LE. (2013). Cross-species conservation of endocrine pathways: A critical analysis of tier 1 fish and rat screening assays with 12 model chemicals. *Environmental Toxicology and Chemistry*. http://www.ncbi.nlm.nih.gov/pubmed/23401061.

Arnot JA, Mackay D TU Canadian Environmental Modelling Centre. (2007). *Risk Prioritization for a Subset of Domestic Substance List Chemicals Using the RAIDAR Model*.

Arnot JA, Mackay D, Parkerton TF, Zaleski RT, Warren CS. (2010a). Multimedia modeling of human exposure to chemical substances: The roles of food web biomagnification and biotransformation. *Environmental Toxicology and Chemistry* 29: 45-55.

Arnot JA, Mackay D, Sutcliffe R, Lo B. (2010b). Estimating farfield organic chemical exposures, intake rates and intake fractions to human age classes. *Environmental Modelling & Software* 25: 1166-1175.

ATSDR (Agency for Toxic Substances and Disease Registry). (2007). *Toxicological profile for benzene*. Retrieved from http://www.atsdr.cdc.gov/toxprofiles/tp3.pdf (accessed March 25, 2013).

Auerbach SS, Shah RR, Mav D, Smith CS, Walker NJ, Vallant MK, et al. (2010). Predicting the hepatocarcinogenic potential of alkenylbenzene flavoring agents using toxicogenomics and machine learning. *Toxicology and Applied Pharmacology* 243: 300-314. http://dx.doi.org/10.1016/j.taap.2009.11.021.

Baker M. (2010). Epigenome: Mapping in motion. *Nature Methods* 7: 181-186.

Basketter DA, Clewell H, Kimber I, Rossi A, Blaauboer B, Burrier R, et al. (2012). A roadmap for the development of alternative (non-animal) methods for systemic toxicity testing - t4 report*. *Alternatives to Animal Experiments* 29: 3-91. http://www.ncbi.nlm.nih.gov/pubmed/22307314.

Behl M, Rao D, Aagaard K, Davidson TL, Levin ED, Slotkin TA, et al. (2013). Evaluation of the association between maternal smoking, childhood obesity, and metabolic disorders: a national toxicology program workshop review. *Environmental Health Perspectives* 121: 170-180. http://www.ncbi.nlm.nih.gov/pubmed/23232494.

Bell RR, Early JL, Nonavinakere VK, Mallory Z. (1990). Effect of cadmium on blood glucose level in the rat. *Toxicology Letters* 54: 199-205. http://www.ncbi.nlm.nih.gov/pubmed/2260118.

Belson M, Kingsley B, Holmes A. (2007). Risk factors for acute leukemia in children: a review. *Environmental Health Perspectives* 115: 138-145. http://www.ncbi.nlm.nih.gov/pubmed/17366834.

Berg EL, Yang J, Melrose J, Nguyen D, Privat S, Rosler E, et al. (2010). Chemical target and pathway toxicity mechanisms defined in primary human cell systems. *Journal of Pharmacological and Toxicological Methods* 61: 3-15. http://www.ncbi.nlm.nih.gov/pubmed/19879948.

Bhattacharya S, Zhang Q, Carmichael PL, Boekelheide K, Andersen ME. (2011). Toxicity testing in the 21 century: defining new risk assessment approaches based on perturbation of intracellular toxicity pathways. *Public Library of Science One* 6: e20887. http://www.ncbi.nlm.nih.gov/pubmed/21701582.

Birnbaum LS. (2012). NIEHS and the future of toxicology. Presentation. *FutureTox Meeting of the Society of Toxicology* October 18-19. Arlington, VA.

Birney E. (2012). The making of ENCODE: Lessons for big-data projects. *Nature* 489: 49-51. http://www.ncbi.nlm.nih.gov/pubmed/22955613.

Blakesley V, Awni W, Locke C, Ludden T, Granneman GR, Braverman LE. (2004). Are bioequivalence studies of levothyroxine sodium formulations in euthyroid volunteers reliable? *Thyroid* 14: 191-200. http://www.ncbi.nlm.nih.gov/pubmed/15072701.

Bleicher KH, Bohm HJ, Muller K, Alanine AI. (2003). Hit and lead generation: beyond high-throughput screening. *Nature Reviews Drug Discovery* 2: 369-378. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12750740.

Blount BC, Valentin-Blasini L, Osterloh JD, Mauldin JP, Pirkle JL. (2007). Perchlorate exposure of the US Population, 2001-2002. *Journal of Exposure Science and Environmental Epidemiology* 17: 400-407. http://www.ncbi.nlm.nih.gov/pubmed/17051137.

Boekelheide K, Blumberg B, Chapin RE, Cote I, Graziano JH, Janesick A, et al. (2012). Predicting later-life outcomes of early-life exposures. *Environmental Health Perspectives* 120: 1353-1361. http://dx.doi.org/10.1289/ehp.1204934.

Bollati V, Baccarelli A. (2010). Environmental epigenetics. *Heredity (Edinb)* 105: 105-112. http://www.ncbi.nlm.nih.gov/pubmed/20179736.

Borgelt C. (2013). Software for frequent pattern mining. Retrieved from http://www.borgelt.net/fpm.html (accessed March 20, 2013).

BOSC (Board of Scientific Counselors). (2010). *Review of the EPA Office of Research and Development's (ORD) Computational Toxicology Research Program (CTRP)*. Washington, DC. Retrieved from http://www.epa.gov/osp/bosc/pdf/ctox1004rpt.pdf (accessed February 26, 2013).

Bosson J, Stenfors N, Bucht A, Helleday R, Pourazar J, Holgate ST, et al. (2003). Ozone-induced bronchial epithelial cytokine expression differs between healthy and asthmatic subjects. *Clinical and Experimental Allergy* 33: 777-782. http://www.ncbi.nlm.nih.gov/pubmed/12801312.

Brown AR, Bickley LK, Le Page G, Hosken DJ, Paull GC, Hamilton PB, et al. (2011). Are toxicological responses in laboratory (inbred) zebrafish representative of those in outbred (wild) populations? - A case study with an endocrine disrupting chemical. *Environmental Science and Technology* 45: 4166-4172. http://www.ncbi.nlm.nih.gov/pubmed/21469706.

Bucher JR, Thayer K, Birnbaum LS. (2011). The Office of Health assessment and translation: A problem-solving resource for the National Toxicology Program. *Environmental Health Perspectives* 119: A196-197. http://www.ncbi.nlm.nih.gov/pubmed/21531652.

Burgoon LD. (2011). Potential of Genomic Data on PAHs to Inform Cumulative Assessment. *National Academy of Sciences Meeting: Mixtures and Cumulative Risk Assessment*. http://nas-sites.org/emergingscience/meetings/mixtures/workshop-presentations-mixtures/.

Burgoon LD, Zacharewski TR. (2008). Automated quantitative dose-response modeling and point of departure determination for large toxicogenomic and high-throughput screening data sets. *Toxicological Sciences* 104: 412-418. http://www.ncbi.nlm.nih.gov/pubmed/18441342.

Bush WS, Moore JH. (2012). Chapter 11: Genome-wide association studies. *Public Library of Science Computational Biology* 8.

Chadwick LH. (2012). The NIH roadmap epigenomics program data resource. *Epigenomics* 4: 317-324. http://www.ncbi.nlm.nih.gov/pubmed/22690667.

Chen J, Blackwell TW, Fermin D, Menon R, Chen Y, Gao J, et al. (2007). Evolutionary-conserved gene expression response profiles across mammalian tissues. *OMICS* 11: 96-115. http://www.ncbi.nlm.nih.gov/pubmed/17411398.

Chen Y, Jin JY, Mukadam S, Malhi V, Kenny JR. (2012). Application of IVIVE and PBPK modeling in prospective prediction of clinical pharmacokinetics: Strategy and approach during the drug discovery phase with four case studies. *Biopharmaceutics & Drug Disposition* 33: 85-98. http://www.ncbi.nlm.nih.gov/pubmed/22228214.

Chiu WA, Euling SY, Scott CS, Subramaniam RP. (2010). Approaches to advancing quantitative human health risk assessment of environmental chemicals in the post-genomic era. *Toxicology and Applied Pharmacology*. http://www.ncbi.nlm.nih.gov/pubmed/20353796.

Chuang HY, Lee E, Liu YT, Lee D, Ideker T. (2007). Network-based classification of breast cancer metastasis. *Molecular Systems Biology* 3: 140. http://www.ncbi.nlm.nih.gov/pubmed/17940530.

Cohen Hubal EA, Richard A, Aylward L, Edwards S, Gallagher J, Goldsmith MR, et al. (2010). Advancing exposure characterization for chemical evaluation and risk assessment. *Journal of Toxicology and Environmental Health Part B: Critical Reviews* 13: 299-313. http://www.ncbi.nlm.nih.gov/pubmed/20574904.

Collins FS. (2010). Transcript from Newshour PBS interview with Frances Collins (June 24, 2010). http://www.pbs.org/newshour/bb/science/jan-june10/genome_06-24.html.

Collins MA. (2009). Generating 'omic knowledge': The role of informatics in high content screening. *Combinatorial Chemistry & High Throughput Screening* 12: 917-925. http://www.ncbi.nlm.nih.gov/pubmed/19531005.

Cordell HJ. (2009). Detecting gene-gene interactions that underlie human diseases. *Nature Reviews Genetics* 10: 392-404. http://www.ncbi.nlm.nih.gov/pubmed/19434077.

Cortessis VK, Thomas DC, Levine AJ, Breton CV, Mack TM, Siegmund KD, et al. (2012). Environmental epigenetics: Prospects for studying epigenetic mediation of exposure-response relationships. *Human genetics* 131: 1565-1589. http://www.ncbi.nlm.nih.gov/pubmed/22740325.

Cote I, Anastas PT, Birnbaum LS, Clark RM, Dix DJ, Edwards SW, et al. (2012). Advancing the next generation of health risk assessment. *Environmental Health Perspectives* 120: 1499-1502. http://www.ncbi.nlm.nih.gov/pubmed/22875311.

Crawford-Brown D. (2013). The role of advanced biological data in the rationality of risk-based regulatory decisions. *Journal of Environmental Protection* 4: 238-249. http://www.scirp.org/journal/PaperInformation.aspx?PaperID=28638.

Crofton KM, Zoeller RT. (2005). Mode of action: neurotoxicity induced by thyroid hormone disruption during development--hearing loss resulting from exposure to PHAHs. *Critical Reviews in Toxicology* 35: 757-769. http://www.ncbi.nlm.nih.gov/pubmed/16417043.

Cronican AA, Fitz NF, Carter A, Saleem M, Shiva S, Barchowsky A, et al. (2013). Genome-wide alteration of histone H3K9 acetylation pattern in mouse offspring prenatally exposed to arsenic. *Public Library of Science One* 8: e53478. http://www.ncbi.nlm.nih.gov/pubmed/23405071.

Crump KS, Chen C, Louis TA. (2010). The future use of in vitro data in risk assessment to set human exposure standards: challenging problems and familiar solutions. *Environmental Health Perspectives* 118: 1350-1354. http://www.ncbi.nlm.nih.gov/pubmed/20562051.

De Coster S, van Larebeke N. (2012). Endocrine-disrupting chemicals: Associated disorders and mechanisms of action. *Journal of Environmental Public Health* 2012: 713696. http://www.ncbi.nlm.nih.gov/pubmed/22991565.

Dearry A. (2013). Integrating environmental health data to advance discovery. Presentation. *National Academy of Sciences Meeting: Integrating Environmental Health Data to Advance Discovery* January 10-11. http://www.youtube.com/watch?v=5TIVa2x4x7c&list=PLzsdEyVNFvgyizsegxlcIbLz1glyOIHxJ&index=2.

Degitz SJ, Holcombe GW, Flynn KM, Kosian PA, Korte JJ, Tietge JE. (2005). Progress towards development of an amphibian-based thyroid screening assay using Xenopus laevis. Organismal and thyroidal responses to the model compounds 6-propylthiouracil, methimazole, and thyroxine. *Toxicological Sciences* 87: 353-364. http://www.ncbi.nlm.nih.gov/pubmed/16002479.

Derry JM, Mangravite LM, Suver C, Furia MD, Henderson D, Schildwachter X, et al. (2012). Developing predictive molecular maps of human disease through community-based modeling. *Nature Genetics* 44: 127-130. http://www.ncbi.nlm.nih.gov/pubmed/22281773.

Devlin RB. (2012). Using Ozone to Validate a Systems Biology Approach to Toxicity Testing. *National Academy of Sciences Meeting: Systems Biology-Informed Risk Assessment*. http://www.youtube.com/watch?v=5TIVa2x4x7c&list=PLzsdEyVNFvgyizsegxlcIbLz1glyOIHxJ&index=2.

Devlin RB, Duncan KE, Jardim M, Schmitt MT, Rappold AG, Diaz-Sanchez D. (2012). Controlled exposure of healthy young volunteers to ozone causes cardiovascular effects. *Circulation* 126: 104-111. http://www.ncbi.nlm.nih.gov/pubmed/22732313.

Diabetes Genetics Initiative of Broad Institute of Harvard, MIT LU, Novartis Institutes of BioMedical Research, Saxena R, Voight BF, Lyssenko V, et al. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316: 1331-1336. http://www.sciencemag.org/content/316/5829/1331.abstract.

Dick E, Rajamohan D, Ronksley J, Denning C. (2010). Evaluating the utility of cardiomyocytes from human pluripotent stem cells for drug screening. *Biochemical Society Transactions* 38: 1037-1045. http://www.ncbi.nlm.nih.gov/pubmed/20659000.

Duncan KE, Crooks J, Miller DJ, Burgoon L, Schmitt MT, Edwards S, et al. (2013). Temporal profile Of gene expression alterations in primary human bronchial epithelial cells following in vivo exposure to ozone *American Thoracic Society International Conference: D105 Genetics and Epigenetics of Lung Disease* Meeting Abstract A5866. http://www.atsjournals.org/doi/abs/10.1164/ajrccm-conference.2013.187.1_MeetingAbstracts.A5866.

Duncan KE, Dailey LA, Carson JL, Hernandez ML, Peden DB, Devlin RB. (2012). Cultured basal airway epithelial cells from asthmatics display baseline gene expression profiles that differ from normal healthy cells and exhibit differential responses to ambient air pollution particles *American Journal of Respiratory and Critical Care Medicine* 185: A4291.

EC (European Commission). (2010). *Main Findings of the Report: Alternative (Non-animal) Methods for Cosmetics Testing: Current Status and Future Prospects*. Institute for Health and Consumer Protection. Retrieved from http://www.iivs.org/workspace/assets/news-assets/findings_cosmetics_2011.pdf (accessed February 22, 2013).

EC (European Commission). (2011). *4th Report on the Implementation of the "Community Strategy for Endocrine Disruptors" a Range of Substances Suspected of Interfering with the Hormone Systems of Humans and Wildlife*. (COM (1999) 706). Brussels. Retrieved from http://ec.europa.eu/environment/endocrine/documents/sec_2011_1001_en.pdf (accessed March 4, 2013).

ECHA (European Chemicals Agency). (2013a). *Evaluation Under REACH: Progress Report 2012*. Helsinki. Retrieved from http://echa.europa.eu/documents/10162/13628/evaluation_report_2012_en.pdf (accessed March 29, 2013).

ECHA (European Chemicals Agency). (2013b). *Proposal for Identification of a Substance as a CMR 1A or 1B, Pbt, vPvB or a Substance of an Equivalent Level of Concern*. Retrieved from http://echa.europa.eu/documents/10162/13638/annex_xv_svhc_4nonylphenol_en.pdf (accessed March 28, 2013).

Ein-Dor L, Kela I, Getz G, Givol D, Domany E. (2005). Outcome signature genes in breast cancer: Is there a unique set? *Bioinformatics* 21: 171-178. http://www.ncbi.nlm.nih.gov/pubmed/15308542.

Eisenberg M, Samuels M, DiStefano JJ, 3rd. (2008). Extensions, validation, and clinical applications of a feedback control system simulator of the hypothalamo-pituitary-thyroid axis. *Thyroid* 18: 1071-1085. http://www.ncbi.nlm.nih.gov/pubmed/18844475.

El Muayed M, Raja MR, Zhang X, MacRenaris KW, Bhatt S, Chen X, et al. (2012). Accumulation of cadmium in insulin-producing β cells. *Islets* 4: 405-416.

Ellinger-Ziegelbauer H, Gmuender H, Bandenburg A, Ahr HJ. (2008). Prediction of a carcinogenic potential of rat hepatocarcinogens using toxicogenomics analysis of short-term in vivo studies. *Mutation Research* 637: 23-39. http://dx.doi.org/10.1016/j.mrfmmm.2007.06.010.

Emdin SO, Dodson GG, Cutfield JM, Cutfield SM. (1980). Role of zinc in insulin biosynthesis. Some possible zinc-insulin interactions in the pancreatic B-cell. *Diabetologia* 19: 174-182. http://www.ncbi.nlm.nih.gov/pubmed/6997118.

ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57-74. http://www.ncbi.nlm.nih.gov/pubmed/22955616.

EPA (U.S. Environmental Protection Agency). (1995). *The Use of the Benchmark Dose Approach in Health Risk Assessment.* (EPA/630/R-94/007). Washington, DC. Retrieved from http://www.epa.gov/raf/publications/pdfs/BENCHMARK.PDF (accessed March 8, 2013).

EPA (U.S. Environmental Protection Agency). (2000). Benzene (CASRN 71-43-2) IRIS Summary. II. Carcinogenicity Assessment for Lifetime Exposure. January 19. Retrieved from http://www.epa.gov/iris/subst/0276.htm (accessed March 25, 2013).

EPA (U.S. Environmental Protection Agency). (2008). *Uncertainty and Variability in Physiologically Based Pharmacokinetic Models: Key Issues and Case Studies*. Washington, DC. Retrieved from http://ofmpub.epa.gov/eims/eimscomm.getfile?p_download_id=477286.

EPA (U.S. Environmental Protection Agency). (2009a). *Strategic Plan for the Future of Toxicity Testing and Risk Assessment at the U.S. EPA*. Washington, DC: Office of the Science Advisor, Science Policy Council. Retrieved from http://www.epa.gov/spc/toxicitytesting/docs/toxtest_strategy_032309.pdf (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2009b). Virtual Tissues Project. April 21-22. Retrieved from http://www.epa.gov/ncct/virtual_tissues/ (accessed April 2, 2013).

EPA (U.S. Environmental Protection Agency). (2010). *Advancing the Next Generation (NexGen) of Risk Assessment: The Prototypes Workshop*. Research Triangle Park, NC. Retrieved from http://www.epa.gov/risk/nexgen/docs/NexGen-Prototypes-Workshop-Summary.pdf (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2011a). *Advancing the Next Generation (NexGen) of Risk Assessment: Public Dialogue Conference*. Washington, DC. Retrieved from http://www.epa.gov/risk/nexgen/docs/NexGen-Public-Conf-Summary.pdf (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2011b). *Framework for an EPA Chemical Safety for Sustainability Research Program*. Washington, DC: Office of Research and Development. Retrieved from http://www.epa.gov/ord/priorities/docs/CSSFramework.pdf (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2011c). *The Incorporation of In Silico Models and In Vitro High Throughput Assays in the Endocrine Disruptor Screening Program (EDSP) for Prioritization and Screening*. Washington, DC. Retrieved from http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2011d). *Integrated Science Assessment of Ozone and Related Photochemical Oxidants (Final Report)*. Washington, DC. Retrieved from http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492 (accessed March 28, 2013).

EPA (U.S. Environmental Protection Agency). (2011e). *Use of "Omic" Technology to Inform the Risk Assessment, Support Document for Case Study: Propiconazole. Appendices A and B*. Washington, DC: Federal Insecticide, Fungicide, and Rodenticide Act Scientific Advisory Panel (FIFRA SAP). Retrieved from http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2011-0284-0003 (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2012a). Advancing the Next Generation (NexGen) of Risk Assessment. July 31. Retrieved from http://www.epa.gov/risk/nexgen/ (accessed March 21, 2013).

EPA (U.S. Environmental Protection Agency). (2012b). Basic Information About Benzene in Drinking Water. Retrieved from http://water.epa.gov/drink/contaminants/basicinformation/benzene.cfm (accessed March 25, 2013).

EPA (U.S. Environmental Protection Agency). (2012c). Computational Toxicology Research. Retrieved from http://www.epa.gov/ncct/ (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2012d). *Computational Toxicology Research: Overview Materials for EPA Science Advisory Board Exposure & Human Health Committee*. Washington, DC. Retrieved from http://yosemite.epa.gov/sab/sabproduct.nsf/AEACDCDADEDFAAF3852579F2005A4C5A/$File/EPA+CompTox+SAB+Materials+4-26-2012+v4.pdf (accessed February 22, 2013).

EPA (U.S. Environmental Protection Agency). (2012e). Endocrine Disruptor Screening Program (EDSP). November 27. Retrieved from http://www.epa.gov/endo/ (accessed March 4, 2013).

EPA (U.S. Environmental Protection Agency). (2013). Sustainable futures: Models & methods. February 6. Retrieved from http://www.epa.gov/oppt/sf/tools/methods.htm (accessed March 22, 2013).

Eskenazi B, Bradman A, Gladstone EA, Jaramillo S, Birch K, Holland N. (2003). CHAMACOS, A longitudinal birth cohort study: Lessons from the fields. *Journal of Children's Health* 1: 3-27. http://ehs.sph.berkeley.edu/holland/publications/files/Eskenazi2003.pdf

Fang X, Bai C, Wang X. (2012). Bioinformatics insights into acute lung injury/acute respiratory distress syndrome. *Clinical and Translational Medicine* 1: 9. http://www.ncbi.nlm.nih.gov/pubmed/23369517.

Fielden MR, Adai A, Dunn RT, Olaharski A, Searfoss G, Sina J, et al. (2011). Development and evaluation of a genomic signature for the prediction and mechanistic assessment of nongenotoxic hepatocarcinogens in the rat. *Toxicological Sciences* 124: 54-74. http://dx.doi.org/10.1093/toxsci/kfr202.

Fielden MR, Brennan R, Gollub J. (2007). A gene expression biomarker provides early prediction and mechanistic assessment of hepatic tumor induction by nongenotoxic chemicals. *Toxicological Sciences* 99: 90-100. http://dx.doi.org/10.1093/toxsci/kfm156.

Fielden MR, Eynon BP, Natsoulis G, Jarnagin K, Banas D, Kolaja KL. (2005). A gene expression signature that predicts the future onset of drug-induced renal tubular toxicity. *Toxicologic Pathology* 33: 675-683. http://dx.doi.org/10.1080/01926230500321213.

Fielden MR, Nie A, McMillian M, Elangbam CS, Trela BA, Yang Y, et al. (2008). Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. *Toxicological Sciences* 103: 28-34. http://dx.doi.org/10.1093/toxsci/kfn022.

Fortunel NO, Otu HH, Ng HH, Chen J, Mu X, Chevassut T, et al. (2003). Comment on "'Stemness': Transcriptional profiling of embryonic and adult stem cells" and "a stem cell molecular signature". *Science* 302: 393; author reply 393. http://www.ncbi.nlm.nih.gov/pubmed/14563990.

Fowler BA, Whittaker MH, Lipsky M, Wang G, Chen XQ. (2004). Oxidative stress induced by lead, cadmium and arsenic mixtures: 30-day, 90-day, and 180-day drinking water studies in rats: an overview. *Biometals* 17: 567-568. http://www.ncbi.nlm.nih.gov/pubmed/15688865.

Friend S. (2013). Scientific opportunities from heterogeneous biological data analysis: Overcoming complexity. Presentation. *National Academy of Sciences Meeting: Integrating Environmental Health Data to Advance Discovery* January 10-11. http://nas-sites.org/emergingscience/files/2013/01/FRIEND-Jan10-final.pdf.

Froehlicher M, Liedtke A, Groh KJ, Neuhauss SC, Segner H, Eggen RI. (2009). Zebrafish (Danio rerio) neuromast: Promising biological endpoint linking developmental and toxicological studies. *Aquatic Toxicology* 95: 307-319. http://www.ncbi.nlm.nih.gov/pubmed/19467721.

Gangwal S, Reif DM, Mosher S, Egeghy PP, Wambaugh JF, Judson RS, et al. (2012). Incorporating exposure information into the toxicological prioritization index decision support framework. *Science of the Total Environment* 435-436: 316-325. http://www.ncbi.nlm.nih.gov/pubmed/22863807.

Garte S, Taioli E, Popov T, Bolognesi C, Farmer P, Merlo F. (2008). Genetic susceptibility to benzene toxicity in humans. *Journal of Toxicology and Environmental Health, Part A* 71: 1482-1489. http://www.ncbi.nlm.nih.gov/pubmed/18836923.

George BJ, Schultz BD, Palma T, Vette AF, Whitaker DA, Williams RW. (2011). *An Evaluation of EPA's National-Scale Air Toxics Assessment (NATA): Comparison with Benzene Measurements in Detroit, Michigan*. U.S. Environmental Protection Agency. Retrieved from http://www.google.com/url?sa=t&rct=j&q=an%20evaluation%20of%20epa's%20national-scale%20air%20toxics%20assessment&source=web&cd=2&ved=0CEEQFjAB&url=http%3A%2F%2Fcfpub.epa.gov%2Fsi%2Fsi_public_file_download.cfm%3Fp_download_id%3D500973&ei=zYEmUfj7L62r2AWMuoC4BQ&usg=AFQjCNGLT3m4187I8hvEnaY_fF6c24-uMA&bvm=bv.42661473,d.b2I (accessed February 22, 2013).

German Federal Ministry for Education and Research. (2013). Virtual Liver Network. Retrieved from http://www.virtual-liver.de/wordpress/en (accessed March 28, 2013).

Gibbs-Flournoy EA, Simmons SO, Bromberg PA, Dick TP, Samet JM. (2013). Monitoring intracellular redox changes in ozone-exposed airway epithelial cells. *Environmental Health Perspectives* 121: 312-317. http://www.ncbi.nlm.nih.gov/pubmed/23249900.

Gibson GG, Rostami-Hodjegan A. (2007). Modelling and simulation in prediction of human xenobiotic absorption, distribution, metabolism and excretion (ADME): In vitro-in vivo extrapolations (IVIVE). *Xenobiotica* 37: 1013-1014. http://www.ncbi.nlm.nih.gov/pubmed/17968734.

Godderis L, Thomas R, Hubbard AE, Tabish AM, Hoet P, Zhang L, et al. (2012). Effect of chemical mutagens and carcinogens on gene expression profiles in human TK6 cells. *Public Library of Science One* 7: e39205. http://europepmc.org/abstract/MED/22723965.

Golbraikh A, Wang XS, Zhu H, Tropsha A. (2012). Predictive QSAR modeling: Methods and applications in drug discovery and chemical risk assessment. *Handbook of Computational Chemistry*1309-1342. http://link.springer.com/referenceworkentry/10.1007%2F978-94-007-0711-5_37#.

Gold LS, Manley NB, Slone TH, Ward JM. (2001). Compendium of chemical carcinogens by target organ: Results of chronic bioassays in rats, mice, hamsters, dogs, and monkeys. *Toxicologic Pathology* 29: 639-652.

Goldsmith MR, Peterson SD, Chang DT, Transue TR, Tornero-Velez R, Tan YM, et al. (2012). Informing mechanistic toxicology with computational molecular models. *Methods in Molecular Biology* 929: 139-165. http://www.ncbi.nlm.nih.gov/pubmed/23007429.

Goldstein BD. (1988). Benzene toxicity. *Occupational Medicine* 3: 541-554. http://www.ncbi.nlm.nih.gov/pubmed/3043738.

Greenawalt DM, Sieberts SK, Cornelis MC, Girman CJ, Zhong H, Yang X, et al. (2012). Integrating genetic association, genetics of gene expression, and single nucleotide polymorphism set analysis to identify susceptibility Loci for type 2 diabetes mellitus. *American Journal of Epidemiology* 176: 423-430. http://www.ncbi.nlm.nih.gov/pubmed/22865700.

Guryev V, Koudijs MJ, Berezikov E, Johnson SL, Plasterk RH, van Eeden FJ, et al. (2006). Genetic variation in the zebrafish. *Genome Research* 16: 491-497. http://www.ncbi.nlm.nih.gov/pubmed/16533913.

Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, et al. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nature Genetics* 43: 768-775. http://www.ncbi.nlm.nih.gov/pubmed/21706001.

Harrill AH, Desmet KD, Wolf KK, Bridges AS, Eaddy JS, Kurtz CL, et al. (2012). A mouse diversity panel approach reveals the potential for clinical kidney injury due to DB289 not predicted by classical rodent models. *Toxicological Sciences* 130: 416-426. http://www.ncbi.nlm.nih.gov/pubmed/22940726.

Harrill AH, Ross PK, Gatti DM, Threadgill DW, Rusyn I. (2009). Population-based discovery of toxicogenomics biomarkers for hepatotoxicity using a laboratory strain diversity panel. *Toxicological Sciences* 110: 235-243. http://www.ncbi.nlm.nih.gov/pubmed/19420014.

Hatch GE, Slade R, Harris LP, McDonnell WF, Devlin RB, Koren HS, et al. (1994). Ozone dose and effect in humans and rats. A comparison using oxygen-18 labeling and bronchoalveolar lavage. *American Journal of Respiratory and Critical Care Medicine* 150: 676-683. http://www.ncbi.nlm.nih.gov/pubmed/8087337.

Hatzimichael E, Crook T. (2013). Cancer epigenetics: new therapies and new challenges. *Journal of Drug Delivery* 2013: 529312. http://www.ncbi.nlm.nih.gov/pubmed/23533770.

Hester RL, Brown AJ, Husband L, Illiescu R, Pruett D, Summers R, et al. (2011). HumMod: A modeling environment for the simulation of integrative human physiology. *Frontiers in Physiology* 2.

Holzhutter HG, Drasdo D, Preusser T, Lippert J, Henney AM. (2012). The virtual liver: A multidisciplinary, multilevel challenge for systems biology. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 4: 221-235. http://www.ncbi.nlm.nih.gov/pubmed/22246674.

Houck KA, Kavlock RJ. (2008). Understanding mechanisms of toxicity: Insights from drug discovery research. *Toxicology and Applied Pharmacology* 227: 163-178. http://www.ncbi.nlm.nih.gov/pubmed/18063003.

Howlader N, Noone AM, Krapcho M, Garshell J, Neyman N, Altekruse SF, et al. (2013). *SEER Cancer Statistics Review, 1975-2010*. Bethesda, MD: U.S. National Institutes of Health, National Cancer Institute. Retrieved from http://seer.cancer.gov/csr/1975_2010/, based on November 2012 SEER data submission, posted to the SEER web site, April 2013.

Hubal EA. (2009). Biologically relevant exposure science for 21st century toxicity testing. *Toxicological Sciences* 111: 226-232. http://www.ncbi.nlm.nih.gov/pubmed/19602574.

Hunter P, Coveney PV, de Bono B, Diaz V, Fenner J, Frangi AF, et al. (2010). A vision and strategy for the virtual physiological human in 2010 and beyond. *Philosophical Transactions Series A: Mathematical, Physical, and Engineering Sciences* 368: 2595-2614. http://www.ncbi.nlm.nih.gov/pubmed/20439264.

Hunter P, Robbins P, Noble D. (2002). The IUPS human Physiome Project. *Pflugers Archiv* 445: 1-9. http://www.ncbi.nlm.nih.gov/pubmed/12397380.

Hurley PM. (1998). Mode of carcinogenic action of pesticides inducing thyroid follicular cell tumors in rodents. *Environmental Health Perspectives* 106: 437-445. http://www.ncbi.nlm.nih.gov/pubmed/9681970.

IARC IAfRo Cancer). (2012). *Monograph 100F: Benzene*. Retrieved from http://monographs.iarc.fr/ENG/Monographs/vol100F/mono100F-24.pdf (accessed March 25, 2013).

Ilhan G, Karakus S, Andic N. (2006). Risk factors and primary prevention of acute leukemia. *Asian Pacific Journal of Cancer Prevention* 7: 515-517. http://www.ncbi.nlm.nih.gov/pubmed/17250419.

Ingelman-Sundberg M. (2005). Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics* 5: 6-13. http://www.ncbi.nlm.nih.gov/pubmed/15492763.

Irons RD, Chen Y, Wang X, Ryder J, Kerzic PJ. (2013). Acute myeloid leukemia following exposure to benzene more closely resembles de novo than therapy related-disease. *Genes Chromosomes Cancer* 52: 887-894. http://www.ncbi.nlm.nih.gov/pubmed/23840003.

Jack J, Wambaugh JF, Shah I. (2011). Simulating quantitative cellular responses using asynchronous threshold Boolean network ensembles. *BMC Systems Biology* 5: 109. http://www.ncbi.nlm.nih.gov/pubmed/21745399.

Jiang X, Kumar K, Hu X, Wallqvist A, Reifman J. (2008). DOVIS 2.0: An efficient and easy to use parallel virtual screening tool based on AutoDock 4.0. *Chemical Central Journal* 2: 18. http://www.ncbi.nlm.nih.gov/pubmed/18778471.

Jubeaux G, Audouard-Combe F, Simon R, Tutundjian R, Salvador A, Geffard O, et al. (2012). Vitellogenin-like proteins among invertebrate species diversity: potential of proteomic mass spectrometry for biomarker development. *Environmental Science and Technology* 46: 6315-6323. http://www.ncbi.nlm.nih.gov/pubmed/22578134.

Judson RS, Kavlock R, Martin M, Reif D, Houck K, Knudsen T, et al. (2013). Perspectives on validation of high-throughput assays supporting 21st century toxicity testing. *Alternatives to Animal Experiments* 30: 51-56. http://www.ncbi.nlm.nih.gov/pubmed/23338806.

Judson RS, Kavlock RJ, Setzer RW, Hubal EA, Martin MT, Knudsen TB, et al. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chemical Research in Toxicology* 24: 451-462. http://dx.doi.org/10.1021/tx100428e.

Judson RS, Martin MT, Reif DM, Houck KA, Knudsen TB, Rotroff DM, et al. (2010). Analysis of eight oil spill dispersants using rapid, in vitro tests for endocrine and other biological activity. *Environmental Science and Technology* 44: 5979-5985. http://dx.doi.org/10.1021/es102150z.

Kanehisa Laboratories. (2013). KEGG: Kyoto encyclopedia of genes and genomes. Retrieved from http://www.genome.jp/kegg/ (accessed February 22, 2013).

Kavlock R, Chandler K, Houck K, Hunter S, Judson R, Kleinstreuer N, et al. (2012). Update on EPA's ToxCast program: providing high throughput decision support tools for chemical risk management. *Chemical Research in Toxicology* 25: 1287-1302. http://www.ncbi.nlm.nih.gov/pubmed/22519603.

Kim CS, Alexis NE, Rappold AG, Kehrl H, Hazucha MJ, Lay JC, et al. (2011). Lung function and inflammatory responses in healthy young adults exposed to 0.06 ppm ozone for 6.6 hours. *American Journal of Respiratory and Critical Care Medicine* 183: 1215-1221. http://www.ncbi.nlm.nih.gov/pubmed/21216881.

Knudsen T, Daston GP. (2010). Virtual tissues and developmental systems biology (Chapter 23). In T Knudsen, G Daston (Eds.), *Second Edition Comprehensive Toxicology* (pp. 347-358). Oxford, UK: Elsevier Ltd.

Knudsen T, DeWoskin RS. (2011). Systems modeling in development toxicity. In *Handbook of Systems Toxicology*: John Wiley & Sons, Ltd.

Knudsen T, Kavlock RJ, Daston GP, Stedman D, Hixon M, Kim JH. (2011). Developmental toxicity testing for safety assessment: New approaches and technologies. *Birth Defects Research: Part B, Developmental and Reproductive Toxicology* 92: 413-420. http://www.ncbi.nlm.nih.gov/pubmed/21770025.

Kondraganti SR, Fernandez-Salguero P, Gonzalez FJ, Ramos KS, Jiang W, Moorthy B. (2003). Polycyclic aromatic hydrocarbon-inducible DNA adducts: Evidence by 32P-postlabeling and use of knockout mice for Ah receptor-independent mechanisms of metabolic activation in vivo. *International Journal of Cancer* 103: 5-11. http://www.ncbi.nlm.nih.gov/pubmed/12455047.

Koturbash I, Beland FA, Pogribny IP. (2011). Role of epigenetic events in chemical carcinogenesis--a justification for incorporating epigenetic evaluations in cancer risk assessment. *Toxicology Mechanisms and Methods* 21: 289-297. http://www.ncbi.nlm.nih.gov/pubmed/21495867.

Krauth D, Woodruff TJ, Bero L. (2013). Instruments for assessing risk of bias and other methodological criteria of published animal studies: a systematic review. *Environmental Health Perspectives* 121: 985-992. http://www.ncbi.nlm.nih.gov/pubmed/23771496.

Krewski D, Hogan V, Turner MC, Zeman PL, McDowell I, Edwards N, et al. (2007). An integrated framework for risk management and population health. *Human and Ecological Risk Assessment* 13: 1288-1312.

Krewski D, Westphal M, Paoli G, Croteau MC, Al-Zoughool M, Andersen ME, et al. (2013). A framework for the next generation of risk science.

Lan Q, Zhang L, Li G, Vermeulen R, Weinberg RS, Dosemeci M, et al. (2004). Hematotoxicity in workers exposed to low levels of benzene. *Science* 306: 1774-1776. http://www.ncbi.nlm.nih.gov/pubmed/15576619.

Lau FK, Decastro BR, Mills-Herring L, Tao L, Valentin-Blasini L, Alwis KU, et al. (2013). Urinary perchlorate as a measure of dietary and drinking water exposure in a representative sample of the United States population 2001-2008. *Journal of Exposure Science and Environmental Epidemiology* 23: 207-214. http://www.ncbi.nlm.nih.gov/pubmed/23188482.

Lock EF, Abdo N, Huang R, Xia M, Kosyk O, O'Shea SH, et al. (2012). Quantitative high-throughput screening for chemical toxicity in a population-based in vitro model. *Toxicological Sciences* 126: 578-588. http://dx.doi.org/10.1093/toxsci/kfs023.

Lossos IS, Czerwinski DK, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D, et al. (2004). Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. *New England Journal of Medicine* 350: 1828-1837. http://www.ncbi.nlm.nih.gov/pubmed/15115829.

Lvovs D, Favorova OO, Favorov AV. (2012). A polygenic approach to the study of polygenic diseases. *Acta Naturae* 4: 59-71. http://www.ncbi.nlm.nih.gov/pubmed/23150804.

Makris SL, Kim JH, Ellis A, Faber W, Harrouk W, Lewis JM, et al. (2011). Current and future needs for developmental toxicity testing. *Birth Defects Research: Part B, Developmental and Reproductive Toxicology* 92: 384-394. http://www.ncbi.nlm.nih.gov/pubmed/21922641.

Martin MT, Knudsen TB, Reif DM, Houck KA, Judson RS, Kavlock RJ, et al. (2011). Predictive Model of Rat Reproductive Toxicity from ToxCast High Throughput Screening. *Biology of Reproduction* 85: 327-339. http://www.ncbi.nlm.nih.gov/pubmed/21565999.

Maull EA, Ahsan H, Edwards J, Longnecker MP, Navas-Acien A, Pi J, et al. (2012). Evaluation of the association between arsenic and diabetes: a National Toxicology Program workshop review. *Environmental Health Perspectives* 120: 1658-1670. http://www.ncbi.nlm.nih.gov/pubmed/22889723.

Mayr LM, Bojanic D. (2009). Novel trends in high-throughput screening. *Current Opinion in Pharmacology* 9: 580-588. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19775937.

McConnell ER, Bell SM. (2013). Systematic Omics Analysis Review (SOAR) tool to support risk assessment. *Submitted*.

McDonnell WF, Stewart PW, Smith MV, Kim CS, Schelegle ES. (2012). Prediction of lung function response for populations exposed to a wide range of ozone conditions. *Inhalation Toxicology* 24: 619-633. http://www.ncbi.nlm.nih.gov/pubmed/22906168.

McDougall CM, Blaylock MG, Douglas JG, Brooker RJ, Helms PJ, Walsh GM. (2008). Nasal epithelial cells as surrogates for bronchial epithelial cells in airway inflammation studies. *American Journal of Respiratory Cell and Molecular Biology* 39: 560-568. http://www.ncbi.nlm.nih.gov/pubmed/18483420.

McHale CM, Zhang L, Lan Q, Vermeulen R, Li G, Hubbard AE, et al. (2011). Global gene expression profiling of a population exposed to a range of benzene levels. *Environmental Health Perspectives* 119: 628-634. http://dx.doi.org/10.1289/ehp.1002546.

McHale CM, Zhang L, Smith MT. (2012). Current understanding of the mechanism of benzene-induced leukemia in humans: Implications for risk assessment. *Carcinogenesis* 33: 240-252.

Mechanic LE, Chen HS, Amos CI, Chatterjee N, Cox NJ, Divi RL, et al. (2012). Next generation analytic tools for large scale genetic epidemiology studies of complex diseases. *Genetic epidemiology* 36: 22-35. http://www.ncbi.nlm.nih.gov/pubmed/22147673.

Medzhitov R. (2008). Origin and physiological roles of inflammation. *Nature* 454: 428-435. http://www.ncbi.nlm.nih.gov/pubmed/18650913.

Meissner A. (2012). What can epigenomics do for you? *Genome Biology* 13: 420. http://www.ncbi.nlm.nih.gov/pubmed/23095436.

Miao X, Sun W, Fu Y, Miao L, Cai L. (2013). Zinc homeostasis in the metabolic syndrome and diabetes. *Frontiers of Medicine* 7: 31-52. http://www.ncbi.nlm.nih.gov/pubmed/23385610.

Miller MD, Crofton KM, Rice DC, Zoeller RT. (2009). Thyroid-disrupting chemicals: Interpreting upstream biomarkers of adverse outcomes. *Environmental Health Perspectives* 117: 1033-1041. http://www.ncbi.nlm.nih.gov/pubmed/19654909.

Morales-Ruan Mdel C, Villalpando S, Garcia-Guerra A, Shamah-Levy T, Robledo-Perez R, Avila-Arcos MA, et al. (2012). Iron, zinc, copper and magnesium nutritional status in Mexican children aged 1 to 11 years. *Salud Publica de Mexico* 54: 125-134. http://www.ncbi.nlm.nih.gov/pubmed/22535171.

Moraru, II, Schaff JC, Slepchenko BM, Blinov ML, Morgan F, Lakshminarayana A, et al. (2008). Virtual cell modelling and simulation software environment. *IET Systems Biology* 2: 352-362. http://www.ncbi.nlm.nih.gov/pubmed/19045830.

Mortensen HM, Euling SY. (2013). Integrating mechanistic and polymorphism data to characterize human genetic susceptibility for environmental chemical risk assessment in the 21st century. *Toxicology and Applied Pharmacology* 271: 395-404. http://dx.doi.org/10.1016/j.taap.2011.01.015.

Murk AJ, Rijntjes E, Blaauboer BJ, Clewell R, Crofton KM, Dingemans MM, et al. (2013). Mechanism-based testing strategy using in vitro approaches for identification of thyroid hormone disrupting chemicals. *Toxicology In Vitro*. http://www.ncbi.nlm.nih.gov/pubmed/23453986.

NAS (National Academy of Sciences). (2007). *Scientific Review of the Proposed Risk Assessment Bulletin from the Office of Management and Budget*. (9780309104777). Washington, DC: The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=11811.

NCBI (National Center for Biotechnology Information). (2009). Epigenomics (database). Retrieved from http://www.ncbi.nlm.nih.gov/epigenomics (accessed March 4, 2013).

NCBI (National Center for Biotechnology Information). (2012a). Gene Expression Omnibus. Retrieved from http://www.ncbi.nlm.nih.gov/geo/ (accessed February 22, 2013).

NCBI (National Center for Biotechnology Information). (2012b). *Reference SNP(refSNP) Cluster Report: rs13266634*. dbSNP: Short Genetic Variations. Retrieved from http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=13266634 (accessed March 20, 2013).

NCBI (National Center for Biotechnology Information). (2013). PubMed Website. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed (accessed April 2, 2013).

NHGRI (National Human Genome Research Institute). (2012). Home Page. Retrieved from http://www.genome.gov/ (accessed February 22, 2013).

NHGRI (National Human Genome Research Institute). (2013). A Catalog of Published Genome-Wide Association Studies. Retrieved from http://www.genome.gov/gwastudies/ (accessed March 25, 2013).

Nichols JW, Breen M, Denver RJ, Distefano JJ, 3rd, Edwards JS, Hoke RA, et al. (2011). Predicting chemical impacts on vertebrate endocrine systems. *Environmental Toxicology and Chemistry* 30: 39-51. http://www.ncbi.nlm.nih.gov/pubmed/20963851.

Nie AY, Mcmillian M, Parker JB, Leone A, Bryant S, Yieh L, et al. (2006). Predictive toxicogenomics approaches reveal underlying molecular mechanisms of nongenotoxic carcinogenicity. *Molecular Carcinogenesis* 45: 914-933.

NIEHS (National Institute of Environmental Health Sciences). (2012a). Host Susceptibility Program. September 7. Retrieved from http://ntp.niehs.nih.gov/?objectid=B76D131B-F1F6-975E-71657BC3DC88C299 (accessed April 2, 2013).

NIEHS (National Institute of Environmental Health Sciences). (2012b). NIEHS SNPs Environmental Genome Project. Retrieved from http://egp.gs.washington.edu/ (accessed March 25, 2013).

NIEHS (National Institute of Environmental Health Sciences). (2012c). NIEHS Strategic Plan. Retrieved from http://niehs.nih.gov/about/strategicplan/index.cfm (accessed February 22, 2013).

NIEHS (National Institute of Environmental Health Sciences). (2013). Comparative Toxicogenomic Database (CTD)™. February 6. Retrieved from http://ctdbase.org/ (accessed March 4, 2013).

NIH (National Institutes of Health). (2012). NIH Chemical Genomics Center. Retrieved from http://www.ncats.nih.gov/research/reengineering/ncgc/ncgc.html (accessed February 22, 2013).

NIOSH (National Institute of Occupational Safety and Health). (1992). *Recommendations for Occupational Safety and Health: Compendium of Policy Documents and Statements*. Retrieved from http://www.cdc.gov/niosh/pdfs/92-100.pdf (accessed March 25, 2013).

North M, Tandon VJ, Thomas R, Loguinov A, Gerlovina I, Hubbard AE, et al. (2011). Genome-wide functional profiling reveals genes required for tolerance to benzene metabolites in yeast. *Public Library of Science One* 6: e24205. http://www.ncbi.nlm.nih.gov/pubmed/21912624.

NRC (National Research Council). (2006). *Human Biomonitoring for Environmental Chemicals*. The National Academies Press. Retrieved from http://www.nap.edu/openbook.php?record_id=11700.

NRC (National Research Council). (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. Washington DC. Retrieved from http://dels.nas.edu/resources/static-assets/materials-based-on-reports/reports-in-brief/Toxicity_Testing_final.pdf (accessed February 22, 2013).

NRC (National Research Council). (2009). *Science and Decisions: Advancing Risk Assessment*. Washington, DC. Retrieved from http://www.nap.edu/catalog/12209.html (accessed February 22, 2013).

NRC (National Research Council). (2011). *Predicting Later-Life Outcomes of Early-Life Exposures*. Retrieved from http://nas-sites.org/emergingscience/files/2011/05/inutero_final_April2011.pdf (accessed April 2, 2013).

O'Shea SH, Schwarz J, Kosyk O, Ross PK, Ha MJ, Wright FA, et al. (2011). In vitro screening for population variability in chemical toxicity. *Toxicological Sciences* 119: 398-407. http://www.ncbi.nlm.nih.gov/pubmed/20952501.

OECD (Organization for Economic Cooperation and Development). (2004). *OECD Principles for the Validation, for Regulatory Purposes, of Quantitative Structure Activity Relationship Models*. Paris, France. Retrieved from http://www.oecd.org/chemicalsafety/assessmentofchemicals/37849783.pdf (accessed February 22, 2013).

OECD (Organization for Economic Cooperation and Development). (2012). The OECD QSAR Toolbox. Version 3.0. Retrieved from http://www.oecd.org/env/chemicalsafetyandbiosafety/assessmentofchemicals/theoecdqsartoolbox.htm (accessed February 22, 2013).

OMIM (Online Mendelian Inheritance in Man). (2012). Solute Carrier Family 30 (Zinc Transporter), Member 8, SLC30A8 November 13. Retrieved from http://omim.org/entry/611145#0001 (accessed March 21, 2013).

Oracle. (2013a). 5 Configuring Rule Sets. Retrieved from http://docs.oracle.com/cd/B28359_01/server.111/b31222/cfrulset.htm#DVADM70150 (accessed March 20, 2013).

Oracle. (2013b). Glossary: "Lift". Retrieved from http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/glossary.htm (accessed March 20, 2013).

Padilla S, Corum D, Padnos B, Hunter DL, Beam A, Houck KA, et al. (2012). Zebrafish developmental screening of the ToxCast Phase I chemical library. *Reproductive Toxicology* 33: 174-187. http://www.ncbi.nlm.nih.gov/pubmed/22182468.

Pare G, Chasman DI, Parker AN, Nathan DM, Miletich JP, Zee RY, et al. (2008). Novel association of HK1 with glycated hemoglobin in a non-diabetic population: A genome-wide evaluation of 14,618 participants in the Women's Genome Health Study. *Public Library of Science Genetics* 4: e1000312. http://www.ncbi.nlm.nih.gov/pubmed/19096518.

Parham F, Austin C, Southall N, Huang R, Tice R, Portier C. (2009). Dose-response modeling of high-throughput screening data. *Journal of Biomolecular Screening* 14: 1216-1227. http://www.ncbi.nlm.nih.gov/pubmed/19828774.

Paris M, Laudet V. (2008). The history of a developmental stage: Metamorphosis in chordates. *Genesis* 46: 657-672. http://www.ncbi.nlm.nih.gov/pubmed/18932261.

Parng C, Seng WL, Semino C, McGrath P. (2002). Zebrafish: A preclinical model for drug screening. *Assay & Drug Development Technologies* 1: 41-48. http://www.ncbi.nlm.nih.gov/pubmed/15090155.

Patel CJ, Cullen MR. (2012). Genetic variability in molecular responses to chemical exposure. *Experientia* 101: 437-457. http://www.ncbi.nlm.nih.gov/pubmed/22945578.

Patel CJ, Chen R, Butte AJ. (2012a). Data-driven integration of epidemiological and toxicological data to select candidate interacting genes and environmental factors in association with disease. *Bioinformatics* 28: i121-126. http://www.ncbi.nlm.nih.gov/pubmed/22689751.

Patel CJ, Cullen MR, Loannidis JP, Butte AJ. (2012b). Systematic evaluation of environmental factors: Persistent pollutants and nutrients correlated with serum lipid levels. *International Journal of Epidemiology* 41: 828-843. http://dx.doi.org/10.1093/ije/dys003.

Patel CJ, Chen R, Kodama K, Loannidis JP, Butte AJ. (2013). Systematic identification of interaction effects between genome-and environment-wide associations in type 2 diabetes mellitus. *Human genetics*.

Peden DB, Boehlecke B, Horstman D, Devlin R. (1997). Prolonged acute exposure to 0.16 ppm ozone induces eosinophilic airway inflammation in asthmatic subjects with allergies. *Journal of Allergy and Clinical Immunology* 100: 802-808. http://www.ncbi.nlm.nih.gov/pubmed/9438490.

Pedersen-Bjergaard J, Andersen MK, Andersen MT, Christiansen DH. (2008). Genetics of therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia* 22: 240-248. http://www.ncbi.nlm.nih.gov/pubmed/18200041.

Perkins EJ, Ankley GT, Crofton KM, Garcia-Reyero N, Lalone CA, Johnson MS, et al. (2013). Current perspectives on the use of alternative species in human health and ecological risk assessments. *Environmental Health Perspectives*. http://www.ncbi.nlm.nih.gov/pubmed/23771518.

Physiome Project. (2013) Retrieved from http://physiomeproject.org/ (accessed March 28, 2013).

Rabinowitz JR, Goldsmith MR, Little SB, Pasquinelli MA. (2008). Computational molecular modeling for evaluating the toxicity of environmental chemicals: Prioritizing bioassay requirements. *Environmental Health Perspectives* 116: 573-577. http://www.ncbi.nlm.nih.gov/pubmed/18470285.

Rakyan VK, Down TA, Balding DJ, Beck S. (2011). Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics* 12: 529-541. http://www.ncbi.nlm.nih.gov/pubmed/21747404.

Raldua D, Thienpont B, Babin PJ. (2012). Zebrafish eleutheroembryos as an alternative system for screening chemicals disrupting the mammalian thyroid gland morphogenesis and function. *Reproductive Toxicology* 33: 188-197. http://www.ncbi.nlm.nih.gov/pubmed/21978863.

Ramasamy A, Mondry A, Holmes CC, Altman DG. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *Public Library of Science Medicine* 5: e184. http://www.ncbi.nlm.nih.gov/pubmed/18767902.

Reaume CJ, Sokolowski MB. (2011). Conservation of gene function in behaviour. *Philosophical Transactions of the Royal Society London B: Biological Science* 366: 2100-2110. http://www.ncbi.nlm.nih.gov/pubmed/21690128.

Reif DM, Martin MT, Tan SW, Houck KA, Judson RS, Richard AM, et al. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environmental Health Perspectives* 118: 1714-1720. http://www.ncbi.nlm.nih.gov/pubmed/20826373.

Reuschenbach P, Silvani M, Dammann M, Warnecke D, Knacker T. (2008). ECOSAR model performance with a large test set of industrial chemicals. *Chemosphere* 71: 1986-1995. http://www.ncbi.nlm.nih.gov/pubmed/18262586.

Rosenbaum RK, Bachmann TM, Swirsky Gold L, Huijbregts MAJ, Jolliet O, Juraske R. (2008). USEtox -- the UNEP-SETAC toxicity model: Recommended characterization factors for human toxicity and freshwater ecotoxicicty in life cycle impact assessment. *International Journal of Life Cycle Assessment* 13: 532-546.

Rotroff DM, Wetmore BA, Dix DJ, Ferguson SS, Clewell HJ, Houck KA, et al. (2010). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicological Sciences* 117: 348-358. http://www.ncbi.nlm.nih.gov/pubmed/20639261.

Rudel RA, Dodson RW, E. N, A.R. Z, J.G. B. (2008). Correlations between urinary phthalate metabolites and phthalates, estrogenic compounds 4-butyl phenol and o-phenyl phenol, and some pesticides in home indoor air and house dust. *Epidemiology* 19: S332.

Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proenca C, et al. (2009). Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nature Genetics* 41: 1110-1115. http://www.ncbi.nlm.nih.gov/pubmed/19734900.

Rusyn I, Gatti DM, Wiltshire T, Kleeberger SR, Threadgill DW. (2010). Toxicogenetics: Population-based testing of drug and chemical safety in mouse models. *Pharmacogenomics* 11: 1127-1136. http://www.ncbi.nlm.nih.gov/pubmed/20704464.

Rusyn I, Sedykh A, Low Y, Guyton KZ, Tropsha A. (2012). Predictive modeling of chemical hazard by integrating numerical descriptors of chemical structures and short-term toxicity assay data. *Toxicological Sciences* 127: 1-9. http://dx.doi.org/10.1093/toxsci/kfs095.

SAB (Science Advisory Board). (2013). *Draft SAB advice on advancing the application of computational toxicology research for human health risk assessment*. Washington, DC. Retrieved from http://yosemite.epa.gov/sab/sabproduct.nsf/46963ceebabd621905256cae0053d5c6/F3315D0EE2EDC11285257B0A005F5B7E/$File/CompTox-edited+1-29-13.pdf (accessed February 26, 2013).

Sagredo C, Ovrebo S, Haugen A, Fujii-Kuriyama Y, Baera R, Botnen IV, et al. (2006). Quantitative analysis of benzo[a]pyrene biotransformation and adduct formation in Ahr knockout mice. *Toxicology Letters* 167: 173-182. http://www.ncbi.nlm.nih.gov/pubmed/17049425.

Samuels MH, Luther M, Henry P, Ridgway EC. (1994). Effects of hydrocortisone on pulsatile pituitary glycoprotein secretion. *Journal of Clinical Endocrinology & Metabolism* 78: 211-215. http://www.ncbi.nlm.nih.gov/pubmed/8288706.

Sand S, Portier CJ, Krewski D. (2011). A signal-to-noise crossover dose as the point of departure for health risk assessment. *Environmental Health Perspectives* 119: 1766-1774. http://www.ncbi.nlm.nih.gov/pubmed/21813365.

Sarapura VD, Samuels MH, Ridgway EC. (2002). Thyroid-Stimulating Hormone. In S Melmed (Ed.), *The Pitutary* (Second ed., pp. 187-229). Malden, MA: Blackwell Science.

Schnatter AR, Glass DC, Tang G, Irons RD, Rushton L. (2012). Myelodysplastic syndrome and benzene exposure among petroleum workers: an international pooled analysis. *Journal of the National Cancer Institute* 104: 1724-1737. http://www.ncbi.nlm.nih.gov/pubmed/23111193.

Schreiber SL. (2003). The small-molecule approach to biology: Chemical genetics and diversity-oriented organic synthesis make possible the systematic exploration of biology. *Chemical & Engineering News* 81: 51-61.

Schug TT, Janesick A, Blumberg B, Heindel JJ. (2011). Endocrine disrupting chemicals and disease susceptibility. *Journal of Steroid Biochemistry and Molecular Biology* 127: 204-215. http://www.ncbi.nlm.nih.gov/pubmed/21899826.

Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, et al. (2007). A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 316: 1341-1345. http://www.ncbi.nlm.nih.gov/pubmed/17463248.

Sedykh A, Zhu H, Tang H, Zhang L, Richard A, Rusyn I, et al. (2011). Use of in vitro HTS-derived concentration-response data as biological descriptors improves the accuracy of QSAR models of in vivo toxicity. *Environmental Health Perspectives* 119: 364-370. http://www.ncbi.nlm.nih.gov/pubmed/20980217.

Selgrade MK, Cooper KD, Devlin RB, van Loveren H, Biagini RE, Luster MI. (1995). Immunotoxicity--bridging the gap between animal research and human health effects. *Fundamental and Applied Toxicology* 24: 13-21. http://www.ncbi.nlm.nih.gov/pubmed/7713335.

Serafimova R, Todorov M, Nedelcheva D, Pavlov T, Akahori Y, Nakai M, et al. (2007). QSAR and mechanistic interpretation of estrogen receptor binding. *SAR and QSAR in Environmental Research* 18: 389-421. http://www.ncbi.nlm.nih.gov/pubmed/17514577.

Shaffer CL, Scialis RJ, Rong H, Obach RS. (2012). Using Simcyp to project human oral pharmacokinetic variability in early drug research to mitigate mechanism-based adverse events. *Biopharmaceutics & Drug Disposition* 33: 72-84. http://www.ncbi.nlm.nih.gov/pubmed/22213407.

Shah et al. (submitted). OnToP: An ontology for toxicity pathways.

Shah I, Wambaugh J. (2010). Virtual tissues in toxicology. *Journal of Toxicology and Environmental Health Part B: Critical Reviews* 13: 314-328. http://www.ncbi.nlm.nih.gov/pubmed/20574905.

Sheldon LS, Cohen Hubal EA. (2009). Exposure as part of a systems approach for assessing risk. *Environmental Health Perspectives* 117: 119-1194. http://www.ncbi.nlm.nih.gov/pubmed/19672394.

Shen M, Zhang L, Lee KM, Vermeulen R, Hosgood HD, Li G, et al. (2011). Polymorphisms in genes involved in innate immunity and susceptibility to benzene-induced hematotoxicity. *Experimental & Molecular Medicine* 43: 374-378. http://www.ncbi.nlm.nih.gov/pubmed/21540635.

Shi L, Jones WD, Jensen RV, Harris SC, Perkins RG, Goodsaid FM, et al. (2008). The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. *BMC Bioinformatics* 9 Suppl 9: S10. http://www.ncbi.nlm.nih.gov/pubmed/18793455.

Sille FC, Thomas R, Smith MT, Conde L, Skibola CF. (2012). Post-GWAS functional characterization of susceptibility variants for chronic lymphocytic leukemia. *Public Library of Science One* 7: e29632. http://www.ncbi.nlm.nih.gov/pubmed/22235315.

Sipes NS, Martin MT, Kothiya P, Reif DM, Judson RS, Richard AM, et al. (2013). Profiling 976 ToxCast chemicals across 331 enzymatic and receptor signaling assays. *Chemical Research in Toxicology* 26: 878-895. http://www.ncbi.nlm.nih.gov/pubmed/23611293.

Sipes NS, Padilla S, Knudsen TB. (2011). Zebrafish: As an integrative model for twenty-first century toxicity testing. *Birth Defects Research: Part C, Embryo Today* 93: 256-267. http://www.ncbi.nlm.nih.gov/pubmed/21932434.

Sistonen J, Sajantila A, Lao O, Corander J, Barbujani G, Fuselli S. (2007). CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. *Pharmacogenetics and Genomics* 17: 93-101. http://www.ncbi.nlm.nih.gov/pubmed/17301689.

Skolness SY, Blanksma CA, Cavallin JE, Churchill JJ, Durhan EJ, Jensen KM, et al. (2013). Propiconazole inhibits steroidogenesis and reproduction in the fathead minnow (Pimephales promelas). *Toxicological Sciences*. http://www.ncbi.nlm.nih.gov/pubmed/23339182.

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. (2007). A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881-885. http://www.ncbi.nlm.nih.gov/pubmed/17293876.

Smith MT, Zhang L, McHale CM, Skibola CF, Rappaport SM. (2011). Benzene, the exposome and future investigations of leukemia etiology. *Chemico-Biological Interactions* 192: 155-159. http://www.ncbi.nlm.nih.gov/pubmed/21333640.

Smith MV, Boyd WA, Kissling GE, Rice JR, Snyder DW, Portier CJ, et al. (2009). A discrete time model for the analysis of medium-throughput C. elegans growth data. *Public Library of Science One* 4: e7018. http://www.ncbi.nlm.nih.gov/pubmed/19753303.

Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, et al. (2007). A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nature Genetics* 39: 770-775. http://www.ncbi.nlm.nih.gov/pubmed/17460697.

Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, Ohnaka K, et al. (2009). Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes* 58: 1690-1699. http://www.ncbi.nlm.nih.gov/pubmed/19401414.

Taylor KW, Novak RF, Anderson HA, Birnbaum LS, Blystone C, Devito M, et al. (2013). Evaluation of the Association between Persistent Organic Pollutants (POPs) and Diabetes in Epidemiological Studies: A National Toxicology Program Workshop Review. *Environmental Health Perspectives* 121: 774-783. http://www.ncbi.nlm.nih.gov/pubmed/23651634.

Teschendorff AE, Widschwendter M. (2012). Differential variability improves the identification of cancer risk markers in DNA methylation studies profiling precursor cancer lesions. *Bioinformatics* 28: 1487-1494. http://www.ncbi.nlm.nih.gov/pubmed/22492641.

Thayer KA, Heindel JJ, Bucher JR, Gallo MA. (2012). Role of environmental chemicals in diabetes and obesity: A National Toxicology Program Workshop report. *Environmental Health Perspectives* 120: 779-789. http://dx.doi.org/10.1289/ehp.1104597.

Thienpont B, Tingaud-Sequeira A, Prats E, Barata C, Babin PJ, Raldua D. (2011). Zebrafish eleutheroembryos provide a suitable vertebrate model for screening chemicals that impair thyroid hormone synthesis. *Environmental Science and Technology* 45: 7525-7532. http://www.ncbi.nlm.nih.gov/pubmed/21800831.

Thomas D. (2010). Gene--environment-wide association studies: Emerging approaches. *Nature Reviews Genetics* 11: 259-272. http://www.ncbi.nlm.nih.gov/pubmed/20212493.

Thomas R, Phuong J, McHale CM, Zhang L. (2012). Using bioinformatic approaches to identify pathways targeted by human leukemogens. *International Journal of Environmental Research and Public Health* 9: 2479-2503. http://www.ncbi.nlm.nih.gov/pubmed/22851955.

Thomas RS, Allen BC, Nong A, Yang L, Bermudez E, Clewell HJ, III, et al. (2007). A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. *Toxicological Sciences* 98: 240-248. http://dx.doi.org/10.1093/toxsci/kfm092.

Thomas RS, Bao W, Chu TM, Bessarabova M, Nikolskaya T, Nikolsky Y, et al. (2009). Use of short-term transcriptional profiles to assess the long-term cancer-related safety of environmental and industrial chemicals. *Toxicological Sciences* 112: 311-321. http://dx.doi.org/10.1093/toxsci/kfp233.

Thomas RS, Clewell HJ, 3rd, Allen BC, Yang L, Healy E, Andersen ME. (2012). Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: A five chemical case study. *Mutation Research* 746: 135-143. http://www.ncbi.nlm.nih.gov/pubmed/22305970.

Thomas RS, Clewell HJ, Allen BC, Wesselkamper SC, Wang NC, Lambert JC, et al. (2011). Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. *Toxicological Sciences* 120: 194-205. http://dx.doi.org/10.1093/toxsci/kfq355.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, et al. (2013a). Incorporating new technologies into toxicity testing and risk assessment: Moving from 21st century vision to a data-driven framework. *Toxicological Sciences*. http://www.ncbi.nlm.nih.gov/pubmed/23958734.

Thomas RS, Wesselkamper SC, Wang NC, Zhao QJ, Petersen DD, Lambert JC, et al. (2013b). Temporal concordance between apical and transcriptional points of departure for chemical risk assessment. *Toxicological Sciences* 134: 180-194. http://www.ncbi.nlm.nih.gov/pubmed/23596260.

Tice RR, Austin CP, Kavlock RJ, Bucher JR. (2013). Improving the human hazard characterization of chemicals: a tox21 update. *Environmental Health Perspectives* 121: 756-765. http://www.ncbi.nlm.nih.gov/pubmed/23603828.

Tietge JE, Degitz SJ, Haselman JT, Butterworth BC, Korte JJ, Kosian PA, et al. (2013). Inhibition of the thyroid hormone pathway in Xenopus laevis by 2-mercaptobenzothiazole. *Aquatic Toxicology* 126: 128-136. http://www.ncbi.nlm.nih.gov/pubmed/23178179.

Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, Strachan DP, et al. (2009). Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes* 58: 505-510. http://www.ncbi.nlm.nih.gov/pubmed/19056611.

Tokar EJ, Diwan BA, Thomas DJ, Waalkes MP. (2012). Tumors and proliferative lesions in adult offspring after maternal exposure to methylarsonous acid during gestation in CD1 mice. *Archives of Toxicology* 86: 975-982. http://www.ncbi.nlm.nih.gov/pubmed/22398986.

Tokar EJ, Qu W, Waalkes MP. (2011). Arsenic, stem cells, and the developmental basis of adult cancer. *Toxicological Sciences* 120 Suppl 1: S192-203. http://www.ncbi.nlm.nih.gov/pubmed/21071725.

Torkamani A, Topol EJ, Schork NJ. (2008). Pathway analysis of seven common diseases assessed by genome-wide association. *Genomics* 92: 265-272. http://www.ncbi.nlm.nih.gov/pubmed/18722519.

Uehara T, Minowa Y, Morikawa Y, Kondo C, Maruyama T, Kato I, et al. (2011). Prediction model of potential hepatocarcinogenicity of rat hepatocarcinogens using a large-scale toxicogenomics database. *Toxicology and Applied Pharmacology* 255: 297-306. http://dx.doi.org/10.1016/j.taap.2011.07.001.

van Leeuwen DM, Pedersen M, Knudsen LE, Bonassi S, Fenech M, Kleinjans JC, et al. (2011). Transcriptomic network analysis of micronuclei-related genes: A case study. *Mutagenesis* 26: 27-32. http://www.ncbi.nlm.nih.gov/pubmed/21164179.

Venkatapathy R, Moudgal CJ, Bruce RM. (2004). Assessment of the oral rat chronic lowest observed adverse effect level model in TOPKAT, a QSAR software package for toxicity prediction. *Journal of Chemical Information and Computer Sciences* 44: 1623-1629. http://www.ncbi.nlm.nih.gov/pubmed/15446819.

Venkatapathy R, Wang NC. (2013). Developmental toxicity prediction. *Methods in Molecular Biology* 930: 305-340. http://www.ncbi.nlm.nih.gov/pubmed/23086848.

Visscher H, Ross CJ, Rassekh SR, Sandor GS, Caron HN, van Dalen EC, et al. (2013). Validation of variants in SLC28A3 and UGT1A6 as genetic markers predictive of anthracycline-induced cardiotoxicity in children. *Pediatr Blood Cancer* 60: 1375-1381. http://www.ncbi.nlm.nih.gov/pubmed/23441093.

Vogelstein B, Lane D, Levine AJ. (2000). Surfing the p53 network. *Nature* 408: 307-310. http://www.ncbi.nlm.nih.gov/pubmed/11099028.

Waits ER, Nebert DW. (2011). Genetic architecture of susceptibility to PCB126-induced developmental cardiotoxicity in zebrafish. *Toxicological Sciences* 122: 466-475. http://www.ncbi.nlm.nih.gov/pubmed/21613231.

Walker JD, Carlsen L. (2002). QSARs for identifying and prioritizing substances with persistence and bioconcentration potential. *SAR and QSAR in Environmental Research* 13: 713-725. http://www.ncbi.nlm.nih.gov/pubmed/12570048.

Walker JD, Carlsen L, Hulzebos E, Simon-Hettich B. (2002). Global government applications of analogues, SARs and QSARs to predict aquatic toxicity, chemical or physical properties, environmental fate parameters and health effects of organic chemicals. *SAR and QSAR in Environmental Research* 13: 607-616. http://www.ncbi.nlm.nih.gov/pubmed/12479375.

Wambaugh J, Shah I. (2010). Simulating microdosimetry in a virtual hepatic lobule. *Public Library of Science Computational Biology* 6: e1000756. http://www.ncbi.nlm.nih.gov/pubmed/20421935.

Wang I, Zhang B, Yang X, Stepaniants S, Zhang C, Meng Q, et al. (2012). Systems analysis of eleven rodent disease models reveals an inflammatome signature and key drivers. *Molecular Systems Biology* 8: 594. http://www.ncbi.nlm.nih.gov/pubmed/22806142.

Wang N, Jay Zhao Q, Wesselkamper SC, Lambert JC, Petersen D, Hess-Wilson JK. (2012a). Application of computational toxicological approaches in human health risk assessment I. A tiered surrogate approach. *Regulatory Toxicology and Pharmacology* 63: 10-19. http://www.ncbi.nlm.nih.gov/pubmed/22369873.

Wang N, Rice GE, Teuschler LK, Colman J, Yang RS. (2012b). An in silico approach for evaluating a fraction-based, risk assessment method for total petroleum hydrocarbon mixtures. *Journal of Toxicology* 2012: 410143. http://www.ncbi.nlm.nih.gov/pubmed/22496687.

Wang N, Venkatapathy R, Bruce RM, Moudgal C. (2011). Development of quantitative structure-activity relationship (QSAR) models to predict the carcinogenic potency of chemicals. II. Using oral slope factor as a measure of carcinogenic potency. *Regulatory Toxicology and Pharmacology* 59: 215-226. http://www.ncbi.nlm.nih.gov/pubmed/20951756.

Wanjek C. (2013). Systems Biology as Defined by National Institute of Health (NIH). Retrieved from http://irp.nih.gov/catalyst/v19i6/systems-biology-as-defined-by-nih (accessed April 2, 2013).

Warner CM, Gust KA, Stanley JK, Habib T, Wilbanks MS, Garcia-Reyero N, et al. (2012). A systems toxicology approach to elucidate the mechanisms involved in RDX species-specific sensitivity. *Environmental Science and Technology* 46: 7790-7798. http://www.ncbi.nlm.nih.gov/pubmed/22697906.

Weiss JN, Karma A, MacLellan WR, Deng M, Rau CD, Rees CM, et al. (2012). "Good enough solutions" and the genetics of complex diseases. *Circulation Research* 111: 493-504. http://www.ncbi.nlm.nih.gov/pubmed/22859671.

Wetmore BA, Wambaugh JF, Ferguson SS, Li L, Clewell HJ, 3rd, Judson RS, et al. (2013). Relative impact of incorporating pharmacokinetics on predicting in vivo hazard and mode of action from high-throughput in vitro toxicity assays. *Toxicological Sciences* 132: 327-346. http://www.ncbi.nlm.nih.gov/pubmed/23358191.

Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM, Freeman K, et al. (2012). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicological Sciences* 125: 157-174. http://www.ncbi.nlm.nih.gov/pubmed/21948869.

WHO (World Health Organization). (2012). *State of the Science of Endocrine Disrupting Chemicals*. Retrieved from http://apps.who.int/iris/bitstream/10665/78102/1/WHO_HSE_PHE_IHE_2013.1_eng.pdf (accessed March 4, 2013).

Wignall J, Muratov E, Fourches D, Sedykh A, Tropsha A, Woodruff TJ, et al. (2012). Modeling toxicity values using chemical structure, in vitro screening, and in vivo toxicity data *51st Annual Meetign of the Society of Toxicology* Poster Abstract # 299. http://www.toxicology.org/AI/Pub/Tox/2012Tox.pdf.

Wignall J, Muratov E, Fourches D, Tropsha A, Woodruff TJ, Zeise L, et al. (2013). Conditional toxicity value (CTV) predictor for generating toxicity values for data sparse chemicals. *52nd Annual Meeting of the Society of Toxicology* Poster Abstract #2454. http://www.toxicology.org/AI/PUB/Tox/2013Tox.pdf.

Williams LM, Oleksiak MF. (2011). Ecologically and evolutionarily important SNPs identified in natural populations. *Molecular Biology and Evolution* 28: 1817-1826. http://www.ncbi.nlm.nih.gov/pubmed/21220761.

Wooding SP, Watkins WS, Bamshad MJ, Dunn DM, Weiss RB, Jorde LB. (2002). DNA sequence variation in a 3.7-kb noncoding sequence 5' of the CYP1A2 gene: Implications for human population history and natural selection. *American Journal of Human Genetics* 71: 528-542. http://www.ncbi.nlm.nih.gov/pubmed/12181774.

Woodruff TJ, Sutton P. (2011). An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. *Health Affairs (Millwood)* 30: 931-937. http://www.ncbi.nlm.nih.gov/pubmed/21555477.

Wu W, Doreswamy V, Diaz-Sanchez D, Samet JM, Kesic M, Dailey L, et al. (2011). GSTM1 modulation of IL-8 expression in human bronchial epithelial cells exposed to ozone. *Free Radical Biology and Medicine* 51: 522-529. http://www.ncbi.nlm.nih.gov/pubmed/21621609.

Yeung KY, Dombek KM, Lo K, Mittler JE, Zhu J, Schadt EE, et al. (2011). Construction of regulatory networks using expression time-series data of a genotyped population. *Proceedings of the National Academy of Sciences USA* 108: 19436-19441. http://www.ncbi.nlm.nih.gov/pubmed/22084118.

Zacharewski T, Teuschler L, Cote I, Burgoon L. (submitted). Improving cumulative risk assessment through systems and network biology driven data mining.

Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. (2007). Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316: 1336-1341. http://www.ncbi.nlm.nih.gov/pubmed/17463249.

Zeise L, Bois FY, Chiu WA, Hattis D, Rusyn I, Guyton KZ. (2012). Addressing human variability in next-generation human health risk assessments of environmental chemicals. *Environmental Health Perspectives* Epub doi: 10.1289/ehp.1205687.

Zhang L, McHale CM, Rothman N, Li G, Ji Z, Vermeulen R, et al. (2010). Systems biology of human benzene exposure. *Chemico-Biological Interactions* 184: 86-93. http://www.ncbi.nlm.nih.gov/pubmed/20026094.

Zhang Q, Bhattacharya S, Andersen ME, Conolly RB. (2010). Computational systems biology and dose-response modeling in relation to new directions in toxicity testing. *Journal of Toxicology and Environmental Health Part B: Critical Reviews* 13: 253-276. http://www.ncbi.nlm.nih.gov/pubmed/20574901.

Zhu H, Rao RS, Zeng T, Chen L. (2012). Reconstructing dynamic gene regulatory networks from sample-based transcriptional data. *Nucleic Acids Research* 40: 10657-10667. http://www.ncbi.nlm.nih.gov/pubmed/23002138.

Zhuo W, Zhang L, Zhu B, Qiu Z, Chen Z. (2012). Association between CYP1A1 Ile462Val variation and acute leukemia risk: Meta-analyses including 2164 cases and 4160 controls. *Public Library of Science One* 7: e46974. http://www.ncbi.nlm.nih.gov/pubmed/23056546.

Zoeller RT, Crofton KM. (2005). Mode of action: Developmental thyroid hormone insufficiency--neurological abnormalities resulting from exposure to propylthiouracil. *Critical Reviews in Toxicology* 35: 771-781. http://www.ncbi.nlm.nih.gov/pubmed/16417044.

Zoeller RT, Dowling AL, Herzig CT, Iannacone EA, Gauger KJ, Bansal R. (2002). Thyroid hormone, brain development, and the environment. *Environmental Health Perspectives* 110 Suppl 3: 355-361. http://www.ncbi.nlm.nih.gov/pubmed/12060829.

Zoeller RT, Rovet J. (2004). Timing of thyroid hormone action in the developing brain: Clinical observations and experimental findings. *Journal of Neuroendocrinology* 16: 809-818. http://www.ncbi.nlm.nih.gov/pubmed/15500540.

[page intentionally left blank]

Appendix A. Technical Papers Supporting the NexGen Report

Technical Papers Supporting the Report	
Preparation for Prototype Development	*Advancing the Next Generation of Risk Assessment* by Ila Cote, Paul Anastas, Linda Birnbaum, Becki Clark, David Dix, Stephen Edwards, and Peter Preuss (2012)
	Advancing the Next Generation (NexGen) of Risk Assessment: The Prototypes Workshop by EPA (2010)
	Summary Report of Advancing the Next Generation of Risk Assessment Public Dialogue Conference by EPA (2011a)
	A Framework for the Next Generation of Risk Assessment by Daniel Krewski, Margit Westphal, Greg Paoli, Maxine Croteau, Mustafa Al-Zoughool, Mel Andersen, Weihsueh Chiu, Lyle Burgoon, and Ila Cote (2013)
	Reconsideration of Important Risk Assessment Issues Informed by Molecular Systems Biology by Daniel Krewski, Melvin Andersen, Kim Boekelheide, Frederic Bois, Lyle Burgoon, Weihsueh Chiu, Michael DeVito, Hisham El-Masri, Lynn Flowers, Michael Goldsmith, Derek Knight, Thomas Knudsen, William Lefew, Greg Paoli, Edward Perkins, Ivan Rusyn, Cecilia Tan, Linda Teuschler, Russell Thomas, Maurice Whelan, Timothy Zacharewski, Lauren Zeise, and Ila Cote (in preparation)
Tier 3 Prototypes: *Leukemia & Benzene, Lung Injury & Ozone, Liver Cancer & B[a]P/PAHs*	*Characterization of Changes in Gene Expression and Biochemical Pathways at Low Levels of Benzene Exposure* by Reuben Thomas, Alan Hubbard, Cliona McHale, Luoping Zhang, Stephen Rappaport, Qing Lan, Nathaniel Rothman, Kathryn Guyton, Jennifer Jinot, Babasaheb Sonawane, and Martyn Smith (in preparation)
	Current Understanding of the Mechanism of Benzene-Induced Leukemia in Humans: Implications for Risk Assessment by Cliona McHale, Luoping Zhang, and Martyn Smith (2012)
	Benzene, the Exposome and Future Investigations of Leukemia Etiology by Martyn Smith, Luoping Zhang, Cliona McHale, Christine Skibola, and Stephen Rappaport (2011)
	Global Gene Expression Profiling of a Population Exposed to a Range of Benzene Levels by Cliona McHale, Luoping Zhang, Qing Lan, Roel Vermeulen, Guilan Li, Alan Hubbard, Kristin Porter, Reuben Thomas, Christopher Portier, Min Shen, Stephen Rappaport, Songnian Yin, Martyn Smith, and Nathaniel Rothman (2011)
	Temporal Profile of Gene Expression Alterations In Primary Human Bronchial Epithelial Cells Following In Vivo Exposure to 0.3 ppm Ozone (meeting abstract) by Kelly Duncan, James Crooks, David Miller, Lyle Burgoon, Michael Schmitt, Stephen Edwards, David Diaz-Sanchez, and Robert Devlin (2013)
	Transcriptional Profiling of Ozone-Induced Stress Responses in Primary Human Bronchial Epithelial Cells Cultured at an Air-Liquid Interface by Kelly Duncan et al. (in preparation)
	Ozone-induced Inflammation Is not Mediated via the Canonical NF-κB Pathway in Humans by David Miller, Stephen Edwards, Lyle Burgoon, Rory Conolly, William Lefew, Kelly Duncan, Robert Devlin, and James Samet (in preparation)
	Systems Biology Informed Assessment of Benzo[a]pyrene/Polycyclic Aromatic Hydrocarbons and Liver Cancer by Lyle Burgoon and Emma McConnell (in preparation)
	IRIS Toxicological Review of Benzo[a]pyrene (Public Comment Draft). U.S. Environmental Protection Agency, Washington, DC, EPA/635/R-13/138a-b (2013).

Technical Papers Supporting the Report	
Tier 2 Prototypes *Knowledge Mining Diabetes/Obesity Example*	*Data Mining Informed Risk Analysis of Environmental and Genetic Factors Associated with Type 2 Diabetes Mellitus* by Lyle Burgoon (in preparation)
	Data Mining NHANES to Identify Environmental Chemical and Disease Associations by Shannon Bell and Stephen Edwards (in preparation)
	Systematic Identification of Interaction Effects Between Genome- and Environment-Wide Associations in Type 2 Diabetes Mellitus by Chirag Patel, Rong Chen, Keiichi Kodama, John Ioannidis, and Atul Butte (2013)
	Data-Driven Integration of Epidemiological and Toxicological Data to Select Candidate Interacting Genes and Environmental Factors in Association with Disease by Chirag Patel, Rong Chen, and Atul Butte (2012a)
	Genetic Variability in Molecular Responses to Chemical Exposure by Chirag Patel and Mark Cullen (2012)
Tier 2 Prototypes *Short-term in Vivo Nonmammalian*	*Role of Environmental Chemicals in Diabetes and Obesity: An NTP Workshop Review* by Kristina Thayer, Jerrold Heindel, John Bucher, and Michael Gallo (2012)
	Current Perspectives on the Use of Alternative Species in Human Health and Ecological Hazard Assessments by Edward Perkins, Gerald Ankley, Kevin Crofton, Natàlia Garcia-Reyero, Carlie LaLone, Mark Johnson, Joseph Tietge, and Daniel Villeneuve (2013)
	Propiconazole Inhibits Steroidogenesis and Reproduction in the Fathead Minnow (Pimephales promelas) by Sarah Skolness, Chad Blanksma, Jenna Cavallin, Jessica Churchill, Elizabeth Durhan, Kathleen Jensen, Rodney Johnson, Michael Kahl, Elizabeth Makynen, Daniel Villeneuve, and Gerald Ankley (2013)
	Zebrafish Developmental Screening of the ToxCast™ Phase I Chemical Library by Stephanie Padilla, Daniel Corum, Beth Padnos, Deborah Hunter, Andrew Beam, Keith Houck, Nisha Sipes, Nicole Kleinstreuer, Thomas Knudsen, David Dix, and David Reif (2012)
	A Systems Toxicology Approach to Elucidate the Mechanisms Involved in RDX Species-Specific Sensitivity by Christopher Warner, Kurt Gust, Jacob Stanley, Tanwir Habib, Mitchell Wilbanks, Natàlia Garcia-Reyero, and Edward Perkins (2012)
Tier 2 Prototypes Short-term In Vivo Mammalian	*Development of a Paradigm for the Next Generation of Chemical Risk Assessment: Short-term In Vivo Models for Tier 2 Assessments* by Michael DeVito, Russell Thomas, and Jason Lambert (in preparation)
	Incorporating New Technologies into Toxicity Testing and Risk Assessment: Moving from 21st Century Vision to a Data-Driven Framework by Russell Thomas, Martin Philbert Scott Auerbach, Barbara Wetmore, Michael Devito, Ila Cote, Craig Rowlands, Maurice Whelan, Sean Hays, Melvin Andersen, Bette Meek, Lawrence Reiter, Jason Lambert, Harvey Clewell III, Martin Stephens, Jay Zhao, Scott Wesselkamper, Lynn Flowers, Edward Carney, Timothy Pastoora, Dan Petersen, Carole Yauk, and Andy Nong (2013a)
	Temporal Concordance Between Apical and Transcriptional Points of Departure for Chemical Risk Assessment by Russell Thomas, Scott Wesselkamper, Nina Wang, Jay Zhao, Dan Peterson, Jason Lambert, Ila Cote, Yang Longlong, Eric Healy, Michael Black, Harvey Clewell, Bruce Allen, and Melvin Andersen (2013b)
	Integrating Pathway-Based Transcriptomic Data into Quantitative Chemical Risk Assessment: A Five Chemical Case Study by Russell Thomas, Harvey Clewell III, Bruce Allen, Longlong Yang, Eric Healy, and Melvin Andersen (2012)
	Application of Transcriptional Benchmark Dose Values in Quantitative Cancer and Noncancer Risk Assessment by Russell Thomas, Harvey Clewell III, Bruce Allen, Scott Wesselkamper, Nina Ching Wang, Jason Lambert, Janet Hess-Wilson, Jay Zhao, and Melvin Andersen (2011)

Technical Papers Supporting the Report	
Tier 1 Prototypes ***Integration of QSAR and Various Biological Data Streams***	*Predictive QSAR Modeling: Methods and Applications in Drug Discovery and Chemical Risk Assessment* by Alexander Golbraikh, Xiang Simon Wang, Hao Zhu, and Alexander Tropsha (2012)
	Developmental Toxicity Prediction by Raghuraman Venkatapathy and Nina Wang (2013)
	Predictive Modeling of Chemical Hazard by Integrating Numerical Descriptors of Chemical Structures and Short-term Toxicity Assay Data by Ivan Rusyn, Alexander Sedykh, Yen Low, KZ Guyton, and Alexander Tropsha (2012)
	An In Silico Approach for Evaluating a Fraction-Based, Risk Assessment Method for Total Petroleum Hydrocarbon Mixtures by Nina Ching Wang, Glenn Rice, Linda Teuschler, Joan Colman, and Raymond Yang (2012b)
	Application of Computational Toxicological Approaches in Human Health Risk Assessment I. A Tiered Surrogate Approach by Nina Ching Yi Wang, Jay Zhao, Scott Wesselkamper, Jason Lambert, Dan Petersen, and Janet Hess-Wilson (2012a)
	Development of Quantitative Structure-Activity Relationship (QSAR) Models to Predict the Carcinogenic Potency of Chemicals. II. Using Oral Slope Factor as a Measure of Carcinogenic Potency by Nina Ching Yi Wang, Raghuraman Venkatapathy, Robert Mark Bruce, and Chandrika Moudgal (2011)
Tier 1 Prototypes ***High-throughput Screening***	*Perspectives on Validation of High-Throughput Assays Supporting 21st Century Toxicity Testing* by Richard Judson, Robert Kavlock, Matthew Martin, David Reif, Keith Houck, Thomas Knudsen, Ann Richard, Raymond Tice, Maurice Whelan, Menghang Xia, Ruili Huang, Christopher Austin, George Daston, Thomas Hartung, John Fowle III, William Wooge, Weida Tong, and David Dix (2013)
	Estimating Toxicity-Related Biological Pathway Altering Doses for High-Throughput Chemical Risk Assessment by Richard Judson, Robert Kavlock, Woodrow Setzer, Elaine Cohen Hubal, Matthew Martin, Thomas Knudsen, Keith Houck, Russell Thomas, Barbara Wetmore, and David Dix (2011)
Key Risk Assessment Issues	*Addressing Human Variability in Next Generation Health Assessments of Environmental Chemicals* by Lauren Zeise, Frederic Bois, Weihsueh Chiu, Dale Hattis, Ivan Rusyn, and Kathryn Guyton (2012)
	Quantitative High-Throughput Screening for Chemical Toxicity in a Population-Based In Vitro Model by Eric Lock, Nour Abdo, Ruili Huang, Menghang Xia, Oksana Kosyk, Shannon O'Shea, Yi-Hui Zhou, Alexander Sedykh, Alexander Tropsha, Christopher Austin, Raymond Tice, Fred Wright, and Ivan Rusyn (2012)
	Predicting Later-Life Outcomes of Early-Life Exposures by Kim Boekelheide, Bruce Blumberg, Robert Chapin, Ila Cote, Joseph Graziano, Amanda Janesick, Robert Lane, Karen Lillycrop, Leslie Myatt, Christopher States, Kristina Thayer, Michael Waalkes, and John Rogers (2012)
	In Vitro Screening for Population Variability in Chemical Toxicity by Shannon O'Shea, John Schwarz, Oksana Kosyk, Pamela Ross, Min Jin Ha, Fred Wright, and Ivan Rusyn (2011)
	Improving Cumulative Risk Assessment Through Systems and Network Biology Driven Data Mining by Timothy Zacharewski, Ila Cote, Linda Teuschler, and Lyle Burgoon (submitted)
	The Role of Advanced Biological Methods and Data in Regulatory Rationality by Douglas Crawford-Brown (2013)
	Incorporating New Technologies into Toxicity Testing and Risk Assessment: Moving from 21st Century Vision to a Data-Driven Framework T by Russell S. Thomas, Martin Philbert, Scott Auerbach, Barbara Wetmore, Michael Devito, Ila Cote, et al. (2013)

Note: EPA also thanks Christine Sofge, Paul Schulte, and Ainsley Weston for sharing their pre-publication draft manuscript.

Appendix B. Glossary

Glossary Term	Description
adverse outcome pathway (AOP)	An adverse outcome pathway is the mechanistic or predictive relationship between an initial chemical-biological interaction (i.e., molecular initiating event[s]) and subsequent perturbation to cellular functions sufficient to elicit disruptions at higher levels of organization, culminating in an adverse phenotypic outcome in an individual and population relevant to risk assessment (i.e., disease progression or organ dysfunction in humans). Ankley GT; Bennett RS; Erickson RJ; Hoff DJ; Hornung MW; Johnson RD; Mount DR; Nichols JW; Russom CL; Schmieder PK; Serrrano JA; Tietge JE; Villeneuve DL. (2010). Adverse outcome pathways: A conceptual framework to support ecotoxicology research and risk assessment. Environmental Toxicology and Chemistry 29 (3): 730-741. http://service004.hpc.ncsu.edu/toxicology/websites/journalclub/linked_files/Fall10/Environ%20Toxicol%20Chem%202010%20Ankley.pdf.
ArrayTrack™	Publicly available toxicogenomics software for DNA microarrays. It contains three integrated components: (1) a database (MicroarrayDB) that stores microarray data and associated toxicological information; (2) tools (TOOL) for data visualization and analysis; and (3) libraries (LIB) that provide curated functional data from public databases for data interpretation. Using ArrayTrack™, an analysis method can be selected from TOOL and applied to selected microarray data stored in the MicroarrayDB. Analysis results can be linked directly to pathways, gene ontology, and other functional information stored in LIB. Food and Drug Administration (FDA). (2012). ArrayTrack™ FAQs. Available online at http://www.fda.gov/ScienceResearch/BioinformaticsTools/Arraytrack/ucm135070.htm (accessed September 27, 2012).
assay	1. The process of quantitative or qualitative analysis of a component of a sample; or 2. Results of a quantitative or qualitative analysis of a component of a sample. National Library of Medicine. (2012). IUPAC Glossary of Terms Used in Toxicology, 2nd Ed. Available online at http://sis.nlm.nih.gov/enviro/iupacglossary/frontmatter.html (accessed September 27, 2012).
Bayesian Network	A graph-based model of joint multivariate probability distributions that captures properties of conditional independence between variables. Friedman N; Linial M; Nachman I; Pe'er D. (2000). Using Bayesian networks to analyze expression data. Journal of Computational Biology 7 (3-4): 601-620.

Glossary Term	Description
Bayesian Network reconstruction	The process of integrating Bayesian Network data using a software program to generate gene causal networks predictive of complex phenotypes. Wang I.; Zhang B; Yang X; Stepaniants S; Zhang C; Meng Q; Peters M; He Y; Ni C; Slipetz D; Crackower MA; Houshyar H; Tan CM; Asante-Appiah E; O'Neill G; Luo MJ; Theiringer R; Yuan J; Chiu C; Lum PY; Lamb J; Boie Y; Wilkinson HA; Schadt E; Dai H; Roberts C. (2012). Systems analysis of eleven rodent disease models reveals an inflammatome signature and key drivers. Molecular Systems Biology 8 594.
bioinformatics	A field of biology in which complex multivariable data from high-throughput screening and genomic assays are interpreted in relation to target identification and effects of sustained perturbations on organs and tissues to make biological discoveries or predictions. This field encompasses all computational methods and theories applicable to molecular biology and areas of computer-based techniques for solving biological problems, including manipulation of models and data sets. National Center for Biotechnology Information (NCBI). (2012). Bioinformatics. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=bioinformatics (accessed September 27, 2012).
biological assay (bioassay)	A method of measuring the effects of a biologically active substance using an intermediate *in vivo* or *in vitro* tissue or cell model under controlled conditions. It includes virulence studies in animal fetuses *in utero*, mouse convulsion bioassay of insulin, quantitation of tumor-initiator systems in mouse skin, calculation of potentiating effects of a hormonal factor in an isolated strip of contracting stomach muscle, etc. National Center for Biotechnology Information (NCBI). (2012). Biological Assay. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=bioassay (accessed September 27, 2012).
biological pathway altering dose (BPAD)	The provisional acceptable exposure level at the low end of the distribution of the external dose required to perturb a biological pathway, accounting for uncertainty and variability. Judson RS; Kavlock RJ; Setzer RW; Hubal EA; Martin MT; Knudsen TB; Houck KA; Thomas RS; Wetmore BA; Dix DJ. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. Chem Res Toxicol 24 (4): 451-462. http://dx.doi.org/10.1021/tx100428e
biomarkers	Measurable and quantifiable biological parameters (e.g., specific enzyme concentrations, specific hormone concentrations, a specific gene phenotype distribution in a population, presence of biological substances) that serve as indices for health- and physiology-related assessments, such as disease risk, psychiatric disorders, environmental exposure and its effects, disease diagnosis, metabolic processes, substance abuse, pregnancy, cell line development, epidemiologic studies. National Center for Biotechnology Information (NCBI). (2012). Biological Markers. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=biological%20markers (accessed September 27, 2012).

Glossary Term	Description
cell biology	The study of the structure, behavior, growth, reproduction, and pathology of cells; and the function and chemistry of cellular components. National Center for Biotechnology Information (NCBI). (2012). Cell Biology Available online at http://www.ncbi.nlm.nih.gov/mesh?term=cell%20biology (accessed September 27, 2012).
Chemical Effects in Biological Systems (CEBS) database	An NIH/NIEHS publicly available toxicogenomic database that houses data of interest to environmental health scientists. CEBS has received depositions of data from academic, industrial, and governmental laboratories. CEBS is designed to display data in the context of biology and study design, and to permit data integration across studies for novel meta-analysis. National Institute for Environmental Health Sciences (NIEHS). (2012). Chemical Effects in Biological Systems (CEBS). Available online at http://www.niehs.nih.gov/research/resources/databases/cebs/index.cfm (accessed September 27, 2012).
Comparative Toxicogenomic Database (CTD)™	A publicly available toxicogenomic database on the National Library of Medicine's (NLM) Toxicology Data Network (TOXNET®). The CTD™ elucidates molecular mechanisms by which environmental chemicals affect human disease. It contains manually curated data describing cross-species chemical-gene/protein interactions and chemical- and gene-disease relationships. The results provide insight into the molecular mechanisms underlying variable susceptibility and environmentally influenced diseases. These data also will provide insights into complex chemical-gene and protein interaction networks. National Library of Medicine (NLM). (2012). Fact Sheet. Comparative Toxicogenomics Database (CTD)™. Available online at http://www.nlm.nih.gov/pubs/factsheets/ctdfs.html (accessed September 27, 2012).
computational models	Computerized predictive tools. Sometimes referred to as “in silico” models. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).

Glossary Term	Description
decision context	Decision context seeks to understand and describe what management decisions are being made, why these decisions are made, and the relationship of these decisions to previous and anticipated decisions. For example, decision context tries to answer some of the following questions: Are risks being ranked; if so, why? How will risk information be used in future decisions? Is a change in policy or management under consideration; and if so, what is driving the change and what are the underlying policy objectives? What is the general scope of alternatives under consideration and why? Decision context defines the roles and responsibilities of the ultimate decision maker, stakeholders, and key technical experts in relation to the decision process. Decision context also identifies the constraints within which a decision must be made and outputs that will result from the decision. Structured Decision Making (SDM). (2008). Steps in the Decision Process: Introduction. Available online at http://www.structureddecisionmaking.org/DecisionContext.htm (accessed March 19, 2013).
DNA microarray	A grid of nucleic acid molecules of known sequence linked to a solid substrate, which can be probed with a sample containing either messenger RNA or complementary DNA from a cell or tissue to reveal changes in gene expression relative to a control sample. Microarray technology, also known as "DNA gene chip" technology, enables the expression of many thousands of genes to be assessed in a single experiment. DNA microarrays exploit the ability of complementary strands of nucleic acids to base-pair with each other and bind. For example, ATATGCGC will bind to its complement (TATACGCG) with a certain affinity. DNA copies (cDNAs) are melted, or denatured, to single strands, which then can be used to bind to, or hybridize with, fluorescently labeled nucleic acid samples from cancerous or normal cells. After washing away the unbound molecules, bound fluorescent nucleic acid samples can be identified by laser microscopy. Fluorescent dots indicate expressed genes, and differences in microarray patterns between normal and cancerous cells can be quickly identified. National Library of Medicine. (2012). IUPAC Glossary of Terms Used in Toxicology, 2nd Ed. Available online at http://sis.nlm.nih.gov/enviro/iupacglossary/frontmatter.html (accessed September 28, 2012).
Enzyme-Linked Immunosorbent Assay (ELISA)	An immunoassay utilizing an antibody labeled with an enzyme marker such as horseradish peroxidase. Although either the enzyme or the antibody is bound to an immunosorbent substrate, they both retain their biologic activity; the change in enzyme activity as a result of the enzyme-antibody-antigen reaction is proportional to the concentration of the antigen and can be measured spectrophotometrically or with the naked eye. Many variations of the method have been developed. National Center for Biotechnology Information (NCBI). (2012). Enzyme-Linked Immunosorbent Assay. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=elisa (accessed September 27, 2012).

Glossary Term	Description
epigenetics	An emerging field of science that studies heritable changes caused by the activation and deactivation of genes with no change in the underlying DNA sequence of the organism. The word is Greek in origin and literally means over and above (epi) the genome. National Human Genome Research Institute (NHGRI). (2012). Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=528&textonly=true (accessed September 27, 2012).
functional genomics	The study of dynamic cellular processes such as gene transcription, translation, and gene product interactions that define an organism. The National Institutes of Health (NIH). (2009). Genomics and Advanced Technologies. Available online at http://www.niaid.nih.gov/topics/pathogengenomics/Pages/definitions.aspx (accessed September 28, 2012).
gene-environment interaction	The combined effects of genotypes and environmental factors on phenotypic characteristics. National Center for Biotechnology Information (NCBI). (2012). Gene-Environment Interaction. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=gene%20environment%20interaction (accessed September 28, 2012).
gene expression	The phenotypic manifestation of a gene or genes by the processes of genetic transcription and genetic translation. National Center for Biotechnology Information (NCBI). (2012). Gene Expression. Available online at http://www.ncbi.nlm.nih.gov/mesh/68015870 (accessed September 28, 2012).
Gene Expression Omnibus (GEO)	A public repository that archives and freely distributes microarray, next-generation sequencing, and other forms of high-throughput functional genomic data submitted by the scientific community. In addition to data storage, a collection of Web-based interfaces and applications is available to help users query and download the studies and gene expression patterns stored in GEO. National Center for Biotechnology Information (NCBI). (2012). Gene Expression Omnibus. Frequently Asked Questions. Available online at http://www.ncbi.nlm.nih.gov/geo/info/faq.html (accessed September 27, 2012).

Glossary Term	Description
Gene Ontology (GO) database	A product of the Gene Ontology (GO) project. The GO project provides structured, controlled vocabularies and classifications that cover several domains of molecular and cellular biology and are freely available for community use in the annotation of genes, gene products, and sequences. Many model organism databases and genome annotation groups use the GO database and contribute their annotation sets to the GO resource. The GO database integrates the vocabularies and contributed annotations and provides full access to this information in several formats. Members of the GO Consortium continuously work collectively, involving outside experts as needed, to expand and update the GO vocabularies. The GO Web resource also provides access to extensive documentation about the GO project and links to applications that use GO data for functional analyses. Gene Ontology Consortium. (2004). The Gene Ontology (GO) database and informatics resource. Nucleic Acids Research 32: Database issue D258-261.
genetics	The branch of science concerned with the means and consequences of transmission and generation of the components of biological inheritance. Used for mechanisms of heredity and the genetics of organisms, for the genetic basis of normal and pathologic states, and for the genetic aspects of endogenous chemicals. It includes biochemical and molecular influence on genetic material. National Center for Biotechnology Information (NCBI). (2012). Genetics. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=genetics (accessed September 27, 2012).
genome-wide association study (GWAS)	An approach used in genetics research to associate specific genetic variations with particular diseases. The method involves scanning the genomes from many different people and looking for genetic markers that can be used to predict the presence of a disease. Once such genetic markers are identified, they can be used to understand how genes contribute to the disease and develop better prevention and treatment strategies. National Institutes of Health (NIH). (2012). Talking Glossary of Genetic Terms: Genome-wide Association Studies (GWAS). National Human Genome Research Institute. Available online at http://www.genome.gov/glossary/index.cfm?id=91&textonly=true (accessed September 27, 2012).
green chemistry	The design of chemical products and processes to reduce or eliminate the use and generation of hazardous substances. Green Chemistry framework includes three main principles: (1) to incorporate sustainable designs across all stages of the chemical lifecycle, (2) to reduce the hazard of chemical products and processes by design, and (3) to work as a cohesive set of design criteria. Twelve design criteria have been developed to fulfill these three principles (prevention, atom economy, less hazardous chemical synthesis, designing safer chemicals, safer solvents and auxiliaries, design for energy efficiency, use of renewable feedstocks, reduce derivatives, catalysis, design for degradation, real-time analysis for pollution prevention, and inherently safer chemistry for accident prevention). Anastas, P, Eghbali, N. (2010). Green chemistry: Principles and practice. Chem Soc Rev 39 (1): 301-312.

Glossary Term	Description
high-throughput screening (HTS)	A rapid method of measuring the effect of an agent in a biological or chemical assay. The assay usually involves some form of automation or a way to conduct multiple assays at the same time using sample arrays. National Center for Biotechnology Information (NCBI). (2012). High-Throughput Screening Assays. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=high%20throughput%20screening%20method (accessed September 27, 2012).
in silico	Referring to or describing data generated and analyzed using computer modeling and information technology. National Library of Medicine. (2012). IUPAC Glossary of Terms Used in Toxicology, 2nd Ed. Available online at http://sis.nlm.nih.gov/enviro/iupacglossary/frontmatter.html (accessed September 27, 2012).
IVIV extrapolation (IVIVE)	A method that uses determinations of protein binding, liver/kidney clearance, and oral uptake to estimate ranges of oral human exposures leading to tissue/plasma concentrations similar to *in vitro* point-of-departure concentrations. Krewski D; Westphal M; Paoli G; Croteau M; Al-Zoughool M; Andersen M; Chiu W; Cote I. (in preparation). A framework for the next generation of risk science.
knowledgebases	Provide an alternative approach for storing and searching the complete networks of highly interconnected information produced by linking bioassays and pathways. Developed decades ago to codify human knowledge so that they could be used to efficiently support decisions, knowledgebases are finding practical applications in meaningfully organizing vast amounts of linked biological data using ontologies.
Kyoto Encyclopedia of Genes and Genomes (KEGG)	A database resource that integrates genomic, chemical, and systemic functional information. In particular, gene catalogs from completely sequenced genomes are linked to higher level systemic functions of the cell, the organism, and the ecosystem. KEGG is a reference knowledgebase for integration and interpretation of large-scale data sets generated by genome sequencing and other high-throughput experimental technologies. Kanehisa Laboratories. (2012). KEGG: Kyoto encyclopedia of genes and genomes. Available online at http://www.genome.jp/kegg/ (accessed February 22, 2013).
lift	Lift is a measure of how much better prediction results are using a model than could be obtained by chance. For example, say 2% of customers who receive a catalog in the mail make a purchase, and when a model is used to select catalog recipients, 10% make a purchase. The lift for the model would be 10/2 or 5. Oracle. (2013). Glossary: "Lift". Available online at http://docs.oracle.com/cd/B28359_01/datamine.111/b28129/glossary.htm (accessed March 20, 2013).

Glossary Term	Description
Meta Data Viewer	A publicly available graphical display software program that can be used to graph animal and human data. Meta Data Viewer can display up to 15 text columns and to graph 1–5 numerical values. Users can sort, group, and filter data and examine patterns of findings across studies. Users can use the program and any associated National Toxicology Program (NTP) data files for their own purposes, including for use in publications. National Toxicology Program (NTP). (2012). Meta Data Viewer. Available online at http://ntp.niehs.nih.gov/?objectid=1DF7D40E-A957-9727-733C9B89E243634B (accessed September 27, 2012).
microarray analysis	The simultaneous analysis, on a microchip, of multiple samples or targets arranged in an array format. National Center for Biotechnology Information (NCBI). (2012). Microarray Analysis. Available online at http://www.ncbi.nlm.nih.gov/mesh/?term=microarray%20analysis (accessed September 27, 2012).
microarray technology	A developing technology used to study the expression of many genes at once. It involves placing thousands of gene sequences in known locations on a glass slide called a gene chip. A sample containing DNA or RNA is placed in contact with the gene chip. Complementary base pairing between the sample and the gene sequences on the chip produces light that is measured. Areas on the chip producing light identify genes that are expressed in the sample. National Human Genome Research Institute (NHGRI). (2012). Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=125&textonly=true (accessed September 27, 2012).
mode of action	The key steps in the toxic response after chemical interaction at the target site that is responsible for the physiological outcome or pathology of the chemical; how chemicals perturb normal biological function. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).
mode-of-action-based *in vitro* toxicity pathway assays	Fit-for-purpose assays using human cells to assess biological pathway perturbations based on specific or generic modes of action. The suite of these assays would form the test battery for safety assessment. Krewski D; Westphal M; Paoli G; Croteau M; Al-Zoughool M; Andersen M; Chiu W; Cote I. (in preparation). A framework for the next generation of risk science.

Glossary Term	Description
molecular epidemiology	Referring to the application of molecular biology to answer epidemiological questions. The examination of patterns of changes in DNA to implicate particular carcinogens and the use of molecular markers to predict which individuals are at highest risk for a disease are common examples. Molecular epidemiology incorporates molecular markers of exposure and biological change into population-based studies; integrates knowledge of the human genome into epidemiological studies to understand genetic susceptibility and gene-environment interaction in disease causation. National Center for Biotechnology Information (NCBI). (2012). Molecular Epidemiology. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=molecular%20epidemiology (accessed September 27, 2012); Krewski D; Westphal M; Paoli G; Croteau M; Al-Zoughool M; Andersen M; Chiu W; Cote I. (in preparation). A framework for the next generation of risk science.
omics	Refers to a broad field of study in biology, ending in the suffix ''-omics'' such as genomics, proteomics, transcriptomics. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).
ontology	Defines types of data (e.g., chemicals, genes, assays, interactions, pathways, cells, species) and their interrelationships (chemicals “activate” proteins; assays “measure” changes in proteins; genes are “part of” pathways, etc.).
phenotype	An individual's observable traits, such as height, eye color, and blood type. The genetic contribution to the phenotype is called the genotype. Some traits are largely determined by the genotype, while other traits are largely determined by environmental factors. National Human Genome Research Institute (NHGRI). (2012). Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=152&textonly=true (accessed September 27, 2012).
polymerase chain reaction (PCR)	A method for amplifying a DNA base sequence using a heat-stable polymerase and two 20-base primers, one complementary to the (+) strand at one end of the sequence to be amplified and one complementary to the (-) strand at the other end. Because the newly synthesized DNA strands can subsequently serve as additional templates for the same primer sequences, successive rounds of primer annealing, strand elongation, and dissociation produce rapid and highly specific amplification of the desired sequence. PCR also can be used to detect the existence of the defined sequence in a DNA sample. Department of Energy (DOE). (2010). Human Genome Project Information: Genome Glossary. Available online at http://www.ornl.gov/sci/techresources/Human_Genome/glossary/glossary_p.shtml (accessed September 27, 2012).

Glossary Term	Description
principal components analysis (PCA)	A mathematical procedure that transforms several possibly correlated variables into a smaller number of uncorrelated variables called principal components. National Center for Biotechnology Information (NCBI). (2012). Principal Components Analysis. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=principal%20component%20analysis (accessed September 27, 2012).
probe	Single-stranded DNA or RNA molecules of specific base sequence, labeled either radioactively or immunologically, that are used to detect the complementary base sequence by hybridization. Department of Energy (DOE). (2010). Human Genome Project Information: Genome Glossary. Available online at http://www.ornl.gov/sci/techresources/Human_Genome/glossary/glossary_p.shtml (accessed September 27, 2012).
proteomics	The study of the function of all expressed proteins. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).
quantitative structure activity relationship (QSAR)	A mathematical relationship between a quantifiable aspect of chemical structure and a chemical property or reactivity or a well-defined biological activity, such as toxicity. Using a sample set of chemicals, a relationship is established between one or many physical-chemical properties a chemical possesses due to its structure and a chemical property or biological activity of concern. This mathematical expression is then used to predict the chemical property or biological response expected from other chemicals with similar structures. It is based on the presumption that similar molecules or chemical structures have similar properties or biological activities or toxicity potential. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).
QSAR Toolbox	A software application intended for use by government, the chemical industry, and other stakeholders in filling gaps in (eco)toxicity data needed for assessing the hazards of chemicals. The Toolbox incorporates information and tools from various sources into a logical workflow. Crucial to this workflow is grouping chemicals into chemical categories. The seminal features of the Toolbox are identification of relevant structural characteristics and the potential mechanism or mode of action of a target chemical, identification of other chemicals that have the same structural characteristics or mechanism/mode of action (or both), and use of existing experimental data to fill the data gap(s). QSAR Toolbox. (2012). About: What does the QSAR Toolbox do? Available online at http://www.qsartoolbox.org/ (accessed September 28, 2012).

Glossary Term	Description
reverse toxicokinetics (RTK)	Also known as reverse dosimetry, refers to the use of a pharmacokinetic model to estimate external dose (exposure) from a known internal concentration. The method uses a one-compartment model and makes default assumptions such as chemicals are eliminated wholly through metabolism and renal excretion; renal excretion is a function of the glomerular filtration rate and the fraction of unbound chemical in the blood (i.e., no active transport); and oral absorption is 100%. Using these assumptions, the plasma concentration of the chemical at steady state per unit dose then can be estimated. The two experimental chemical-specific parameters required to generate an estimate are the rate of disappearance of parent via hepatic metabolism (intrinsic clearance) and fraction bound (or conversely unbound) to plasma proteins. Both parameters can be measured experimentally in a relatively high-throughput manner. Judson RS; Kavlock RJ; Setzer RW; Hubal EA; Martin MT; Knudsen TB; Houck KA; Thomas RS; Wetmore BA; Dix DJ. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. Chem Res Toxicol Chem Res Toxicol 24 (4): 451-462. http://dx.doi.org/10.1021/tx100428e
rule	A rule describes an association between elements on the left-hand side of the rule and items on the right-hand side of the rule. For instance, the rule [diapers, cola] => [milk] in a supermarket database might mean that when customers bought diapers and cola, they also purchased milk.
ruleset	A ruleset is a collection of one or more rules that can be associated with a realm authorization, factor assignment, command rule, or secure application role. The ruleset will be “true” or “false” based on evaluation of each rule in the ruleset and the evaluation type for the ruleset, which can be “all true” or “any true.” Oracle. (2013). 5 Configuring Rule Sets. Available online at http://docs.oracle.com/cd/B28359_01/server.111/b31222/cfrulset.htm#DVADM70150 (accessed March 20, 2013).
SNPs	Refers to single nucleotide polymorphisms, which are single nucleotide variations in a genetic sequence that occur at appreciable frequency in the population. National Center for Biotechnology Information (NCBI). (2012). SNPs. Available online at http://www.ncbi.nlm.nih.gov/mesh?term=SNPS (accessed September 28, 2012).
stem cell biology	A branch of biology that studies and develops stem cells, which are cells with the ability to divide for indefinite periods in culture and to give rise to specialized cells. The National Institutes of Health (NIH). (2009). Stem Cell Basics. Available online at http://irp.nih.gov/catalyst/v19i6/systems-biology-as-defined-by-nih (accessed September 28, 2012).

Glossary Term	Description
systems biology	A scientific approach that combines the principles of engineering, mathematics, physics, and computer science with extensive experimental data to develop a quantitative as well as a deep conceptual understanding of biological phenomena, permitting prediction and accurate simulation of complex (emergent) biological behaviors. Wanjek, C. (2011). Systems biology as defined by NIH. The NIH Catalyst 19 (6): November-December. http://irp.nih.gov/catalyst/v19i6/systems-biology-as-defined-by-nih.
TOM (topological overlap matrix) heat map	A graphical representation in which the rows and columns represent genes in a symmetric manner; the color intensity represents the interaction strength between genes. Wang I.; Zhang B; Yang X; Stepaniants S; Zhang C; Meng Q; Peters M; He Y; Ni C; Slipetz D; Crackower MA; Houshyar H; Tan CM; Asante-Appiah E; O'Neill G; Luo MJ; Theiringer R; Yuan J; Chiu C; Lum PY; Lamb J; Boie Y; Wilkinson HA; Schadt E; Dai H; Roberts C. (2012). Systems analysis of eleven rodent disease models reveals an inflammatome signature and key drivers. Molecular Systems Biology 8 594.
toxicity pathways	The 2007 NRC report on Toxicity Testing in the 21st Century envisioned that new technologies will help us better understand how chemicals perturb normal biological function, and thus identify toxicity pathways. Potential toxic effects of chemicals would be predicted based on *in vitro* bioactivity profiles derived from a chemical's effects on cellular molecules and processes. The interpretation of chemically induced perturbations in toxicity pathways depends on linking *in vitro* effects with adverse outcomes *in vivo*, and on computer modeling that extrapolates to predicted responses in whole tissues, organisms, and populations based on realistic human or environmental exposures. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).
toxicogenomics	Study of the roles that genes play in the biological responses to environmental toxicants and stressors by the collection, interpretation, and storage of information about gene and protein activity. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).
transcription	The biosynthesis of RNA carried out on a template of DNA. The biosynthesis of DNA from an RNA template is called reverse transcription. National Center for Biotechnology Information (NCBI). (2012). Transcription. Available online at http://www.ncbi.nlm.nih.gov/mesh/68014158 (accessed September 27, 2012).

Glossary Term	Description
transcriptome	The pattern of gene expression, at the level of genetic transcription, in a specific organism or under specific circumstances in specific cells. National Center for Biotechnology Information (NCBI). (2012). Transcriptome. Available online at http://www.ncbi.nlm.nih.gov/mesh/68059467 (accessed September 27, 2012).
transcriptomics	The study of gene expression at the RNA level. U.S. Environmental Protection Agency (EPA). (2012). Glossary of Terms: Methods of Toxicity Testing and Risk Assessment. Available online at http://www.epa.gov/opp00001/science/comptox-glossary.html (accessed April 2, 2013).
transgenic	Produced from a genetically manipulated egg or embryo; containing genes from another species. National Center for Biotechnology Information (NCBI). (2012). Transgenic. Available online at http://www.ncbi.nlm.nih.gov/mesh/?term=transgenic (accessed September 27, 2012).
translation	The process of translating the sequence of a messenger RNA (mRNA) molecule to a sequence of amino acids during protein synthesis. The genetic code describes the relationship between the sequence of base pairs in a gene and the corresponding amino acid sequence that it encodes. In the cell cytoplasm, the ribosome reads the sequence of the mRNA in groups of three bases to assemble the protein. National Human Genome Research Institute (NHGRI). 2012. Talking Glossary of Genetic Terms. Available online at http://www.genome.gov/glossary/index.cfm?id=200&textonly=true (accessed September 28, 2012).
translesion synthesis	A mechanism for DNA damage tolerance that allows the DNA replication machinery to move beyond a DNA lesion or abasic site (i.e., a site that lacks a DNA base).
Virtual Tissue (v-Tissues™) Models	In silico cross-scale models of cellular organization and emergent functions used to better understand disease progression. Tissues are the clinically relevant level for diagnosing and treating the transition from normal to adverse states in chemical-induced toxicities leading to cancer, immune dysfunction, developmental defects, and more. Currently, *in vivo* rodent experiments are used to evaluate tissue-level effects of altered molecular and cellular function; however, the extrapolation of animal models to humans is often uncertain. v-Tissues™ aim to simulate key molecular and cellular processes computationally in the context of normal tissue biology to: (1) help understand complex physiological relationships, and (2) predict adverse effects due to chemicals. As the number of chemicals in consumer products, the workplace, and the environment continues to rise, v-Tissues™ offers the promise of a more efficient, effective, and humane approach for evaluating their impact on human health. U.S. Environmental Protection Agency (EPA), Computational Toxicology Research Program. What are Virtual Tissues? (2012). Available online at http://www.epa.gov/ncct/virtual_tissues/what.html (accessed September 27, 2012).