

Computational Toxicology

Elaine Cohen Hubal

*Workshop on Toxicogenomics in Risk Assessment.
Toxicology And Risk Assessment Conference.
Cincinnati, OH April 14, 2008*



Definition: Computational Toxicology

The application of mathematical and computer models to predict adverse effects and to better understand the mechanism(s) through which a given chemical induces harm.

National Center for Computational Toxicology

- The emerging field of computational toxicology applies mathematical and computer models and molecular biological and chemical approaches to explore both qualitative and quantitative relationships between sources of environmental pollutant exposure and adverse health outcomes.
- The integration of modern computing with molecular biology and chemistry will allow scientists to better prioritize data, inform decision makers on chemical risk assessments, and understand a chemical's progression from the environment to the target tissue within an organism and ultimately to the key steps that trigger an adverse health effect.

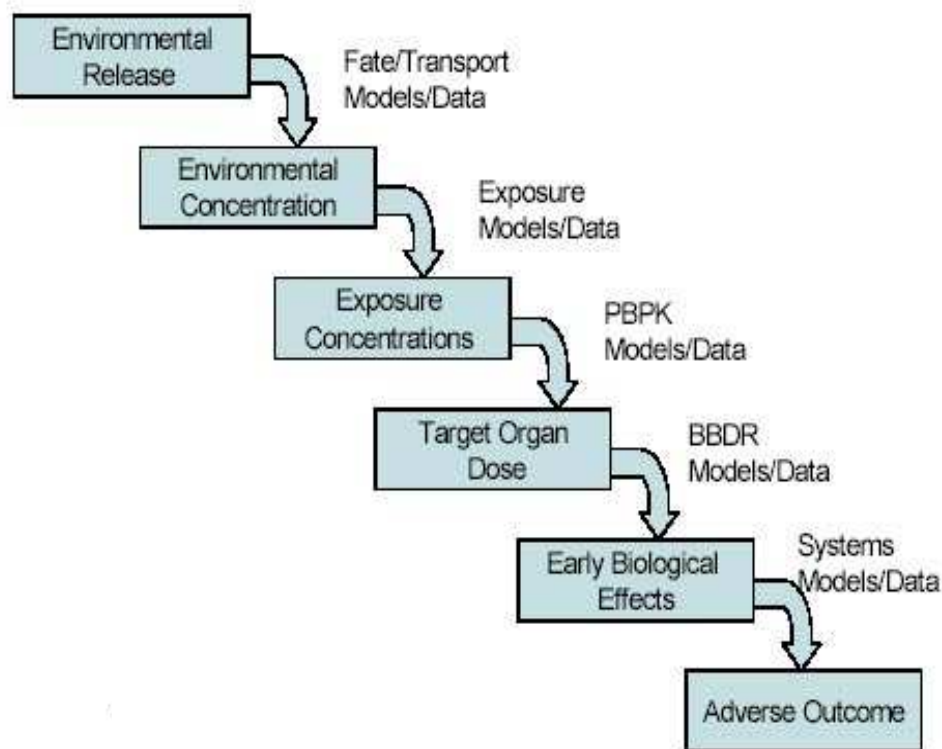
<http://www.epa.gov/comptox/>

Strategic Objectives of Computational Toxicology Initiative

- Improve understanding of the linkages in the continuum between the source of a chemical in the environment and adverse outcomes
- Provide predictive models for screening and testing
- Improve quantitative risk assessment

<http://www.epa.gov/comptox/>

Improving Links in Source-to-Outcome Prediction



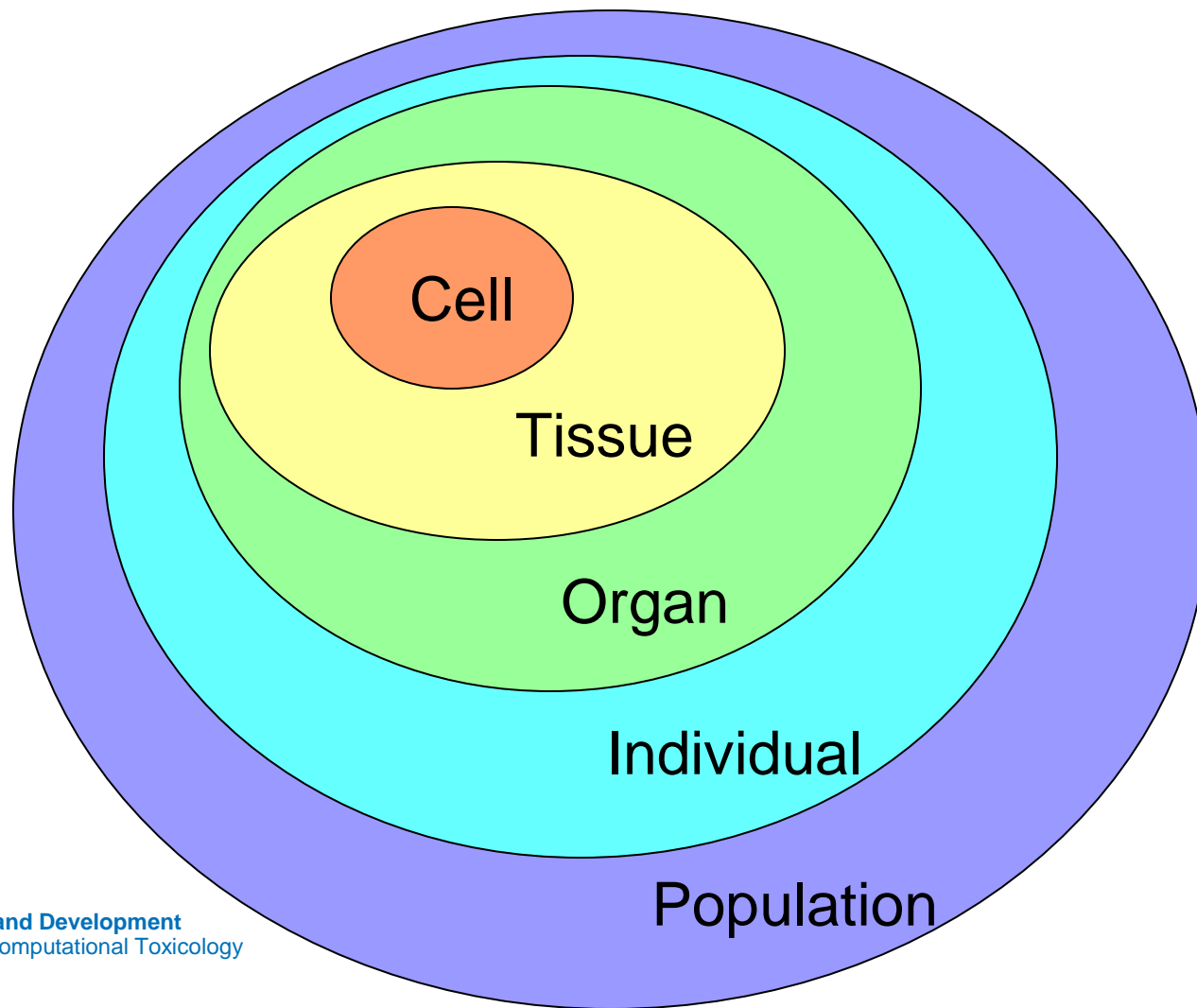
Risk Assessment Challenges

Elucidating the mechanisms of chemical-induced toxicity

Understanding species-specific nature of toxicity

Predicting toxicity at low doses

Levels of Biological Organization



Transforming Toxicology

July 2007

Toxicity Testing in the 21st Century: A Vision and a Strategy

Advances in molecular biology, biotechnology, and other fields are paving the way for major improvements in how scientists evaluate the health risks posed by potentially toxic chemicals found at low levels in the environment. These advances would make toxicity testing quicker, less expensive, and more directly relevant to human exposures. They could also reduce the need for animal testing by substituting more laboratory tests based on human cells. This National Research Council report creates a far-reaching vision for the future of toxicity testing.

Toxicity tests on laboratory animals are conducted to evaluate chemicals—including medicines, food additives, and industrial, consumer, and agricultural chemicals—for their potential to cause cancer, birth defects, and other adverse health effects. Information from toxicity testing serves as an important part of the basis for public health and regulatory decisions concerning toxic chemicals. Current test methods were developed incrementally over the past 50 to 60 years and are conducted using laboratory animals, such as rats and mice. Using the results of animal tests to predict human health effects involves a number of assumptions and extrapolations that remain controversial. Test animals are often exposed to higher doses than would be expected for typical human exposures, requiring assumptions about

effects at lower doses or exposures. Test animals are typically observed for overt signs of adverse health effects, which provide little information about biological changes leading to such health effects. Often controversial uncertainty factors must be applied to account for differences between test animals and humans. Finally, use of animals in testing is expensive and time consuming, and it sometimes raises ethical issues.

Today, toxicological evaluation of chemicals is poised to take advantage of the on-going revolution in biology and biotechnology. This revolution is making it increasingly possible to study the effects of chemicals using cells, cellular components, and tissues—preferably of human origin—rather than whole animals. These powerful new approaches should help to address a number of challenges facing the



THE NATIONAL ACADEMIES
Advisors to the Nation on Science, Engineering, and Medicine

National Academy of Sciences • National Academy of Engineering • Institute of Medicine • National Research Council

REPORT
IN BRIEF

THE NATIONAL ACADEMIES

POLICY FORUM

TOXICOLOGY

Transforming Environmental Health Protection

Francis S. Collins,^{1*} George M. Gray,^{2*} John R. Bucher^{3*}

In 2005, the U.S. Environmental Protection Agency (EPA), with support from the U.S. National Toxicology Program (NTP), funded a project at the National Research Council (NRC) to develop a long-range vision for toxicity testing and a strategic plan for implementing that vision. Both agencies wanted future toxicity testing and assessment paradigms to meet evolving regulatory needs. Challenges include the large numbers of substances that need to be tested and how to incorporate recent advances in molecular toxicology, computational sciences, and information technology, to rely increasingly on human as opposed to animal data; and to offer increased efficiency in design and costs (1–5). In response, the NRC Committee on Toxicity Testing and Assessment of Environmental Agents produced two reports that reviewed current toxicity testing, identified key issues, and developed a vision and implementation strategy to create a major shift in the assessment of chemical hazard and risk (6, 7). Although the NRC reports have laid out a solid theoretical rationale, comprehensive and rigorously gathered data (and comparisons with historical animal data) will determine whether the hypothesized improvements will be realized in practice. For this purpose, NTP, EPA, and the National Institutes of Health Chemical Genomics Center (NCGC) (organizations with expertise in experimental toxicology, computational toxicology, and high-throughput technologies, respectively) have established a collaborative research program.

EPA, NCGC, and NTP Joint Activities

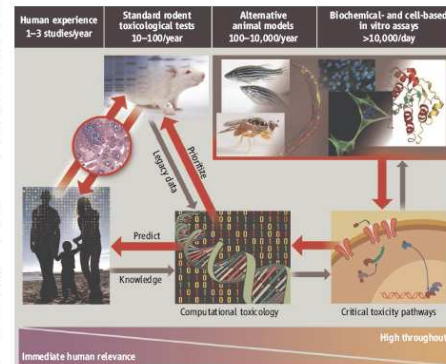
In 2004, the NTP released its vision and roadmap for the 21st century (1), which established initiatives to integrate high-

throughput screening (HTS) and other automated screening assays into its testing program. In 2005, the EPA established the National Center for Computational Toxicology (NCCT). Through these initiatives, NTP and EPA, with the NCGC, are promoting the evolution of toxicology from a predominantly observational science at the level of disease-specific models in vivo to a predominantly predictive science focused on broad inclusion of target-specific, mechanism-based, biological observations in vitro (1, 4) (see figure, below).

Toxicity pathways. In vitro and in vivo tools are being used to identify cellular responses after chemical exposure expected to result in adverse health effects (7). HTS methods are a primary means of discovery for drug development, and screening of >100,000 compounds per day is routine (8). However, drug-discovery HTS methods traditionally test compounds at one concentra-

We propose a shift from primarily in vivo animal studies to in vitro assays, in vivo assays with lower organisms, and computational modeling for toxicity assessments.

tion, usually between 2 and 10 μ M, and to tolerate high false-negative rates. In contrast, in the EPA, NCGC, and NTP combined effort, all compounds are tested at as many as 15 concentrations, generally ranging from ~5 nM to ~100 μ M, to generate a concentration-response curve (9). This approach is highly reproducible, produces significantly lower false-positive and false-negative rates than the traditional HTS methods (9), and facilitates multiassay comparisons. Finally, an informatics platform has been built to compare results among HTS screens; this is being expanded to allow comparisons with historical toxicologic NTP and EPA data (<http://ncgc.nih.gov/pub/openhts>). HTS data collected by EPA and NTP, as well as by the NCGC and other Molecular Libraries Initiative centers (<http://mln.nih.gov/>), are being made publicly available through Web-based databases [e.g., PubChem (<http://pubchem.ncbi.nlm.nih.gov/>)]. In addition,



Transforming toxicology. The studies we propose will test whether high-throughput and computational toxicology approaches can yield data predictive of results from animal toxicity studies, will allow prioritization of chemicals for further testing, and can assist in prediction of risk to humans.

906

15 FEBRUARY 2008 VOL 319 SCIENCE www.sciencemag.org

Science, Feb 15, 2008

Robert Kavlock

6

Toxicity Testing in the Twenty-first Century: A Vision and a Strategy

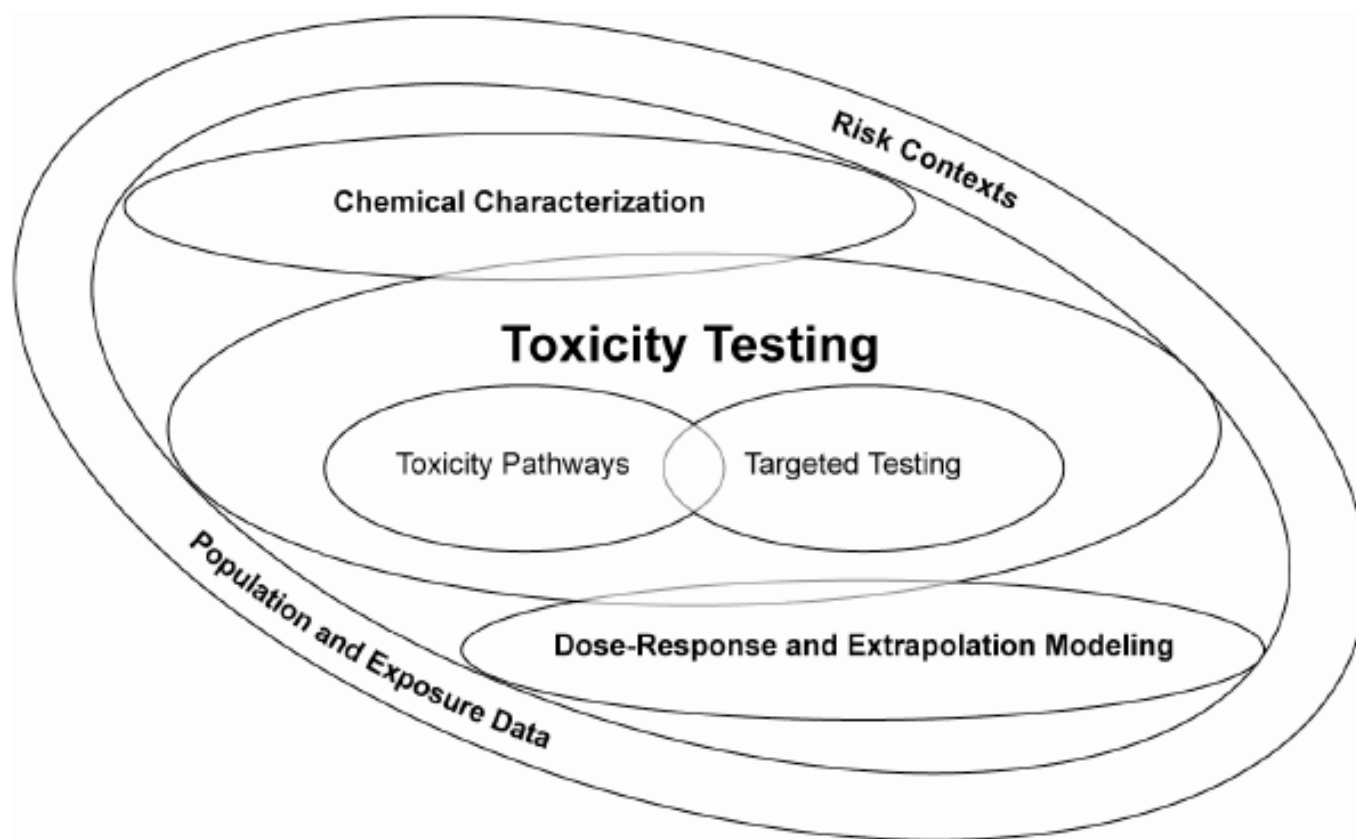


FIGURE 2-3 The committee's vision is a process that includes chemical characterization, toxicity testing, and dose-response and extrapolation modeling. At each step, population-based data and human exposure information are considered, as is the question of what data are needed for decision-making.

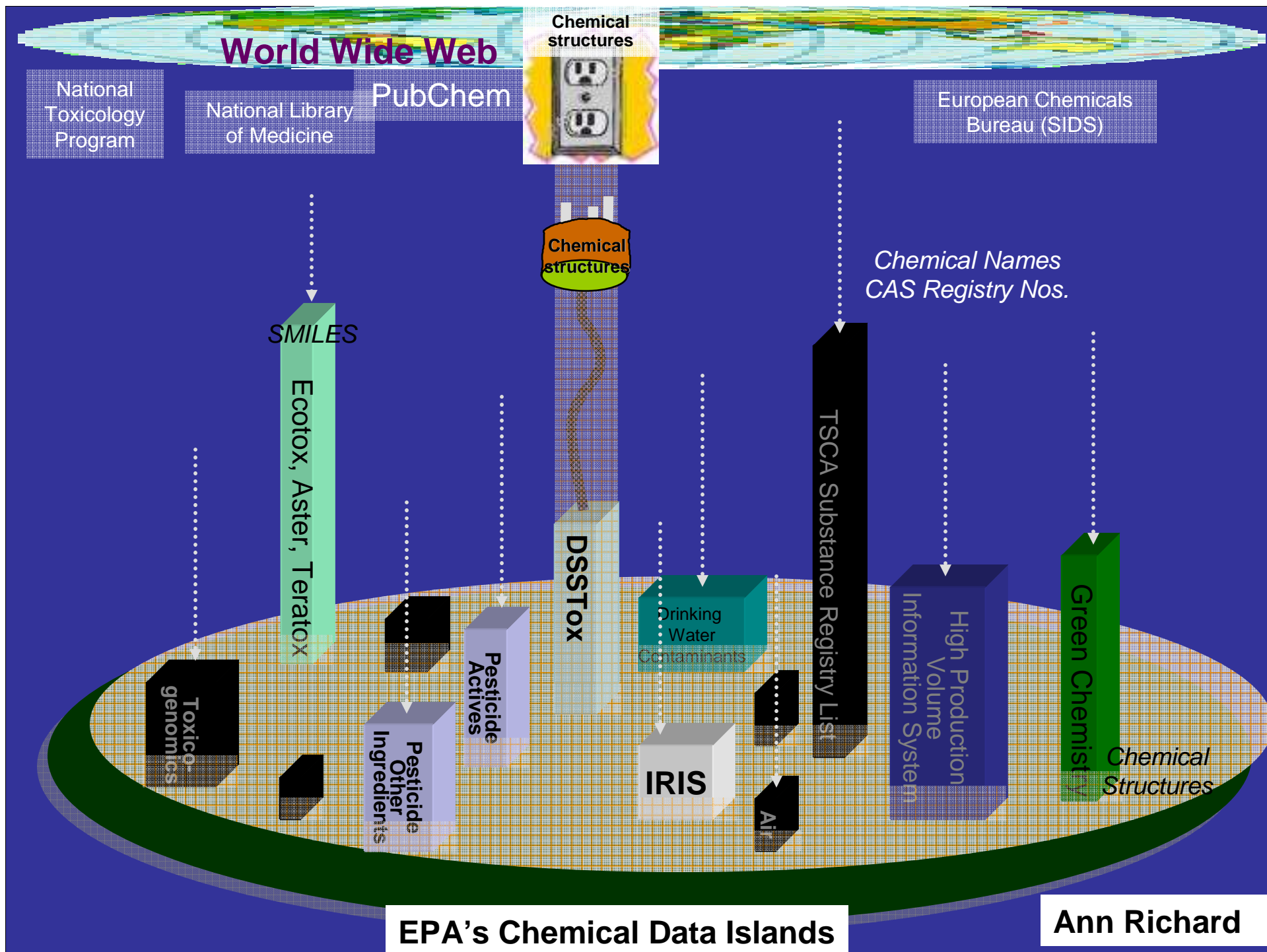
Outline

- Distributed Structure-Searchable Toxicity (DSSTox) Database Network (Ann Richard)
- ToxCast TM Program for Prioritizing Toxicity Testing of Environmental Chemicals (David Dix and Keith Houck)
- Virtual Liver Project (Imran Shah)
- Mechanistic Indicators of Childhood Asthma (MICA) Study (Jane Gallagher)
- Exposure Science for Toxicity Testing (Elaine Cohen Hubal)

Distributed Structure-Searchable Toxicity (DSSTox) Database Network (Ann Richard)

The Research Problem

- Computational toxicology requires integration and analysis of large data sets



Distributed Structure-Searchable Toxicity (DSSTox) Database Network

- Creating a chemical data foundation for improved structure-activity and predictive toxicology capabilities across and outside of EPA.
- DSSTox website publishes downloadable, standardized chemical structure files associated with toxicity data in a variety of formats, along with documentation and links to source information.
- Website hosts a large amount of additional information on the DSSTox project, as well as information on data standards, EPA and outside collaborations, quality review procedures, guidance for users, etc.

<http://epa.gov/ncct/dsstox/>



Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

[Recent Additions](#) | [Contact Us](#)

Search: ☐ All EPA ☒ This Area

Go

You are here: [EPA Home](#) » [Computational Toxicology Research](#) Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

[About DSSTox](#)

[Work in Progress](#)

[Frequent Questions](#)

[Structure Data Files](#)

[Central Field Definition Table](#)

[Apps, Tools & More](#)

[DSSTox Community](#)

[Site Map](#)

[Glossary of Terms](#)

[Help](#)

DSSTox

<http://www.epa.gov/ncct/dsstox/>

Distributed Structure-Searchable Toxicity (DSSTox) Database Network is a project of EPA's [Computational Toxicology Program](#), helping to build a public data foundation for improved structure-activity and predictive toxicology capabilities. The DSSTox website provides a public forum for publishing downloadable, structure-searchable, standardized chemical structure files associated with toxicity data. [More>](#)



[DSSTox Structure-Browser information Page](#)

Recent Additions: 27 September 2007

- [TOXCST: Research Chemical Inventory for EPA's ToxCast Program](#) - Updated to v2a

Recent Additions: 28 August 2007

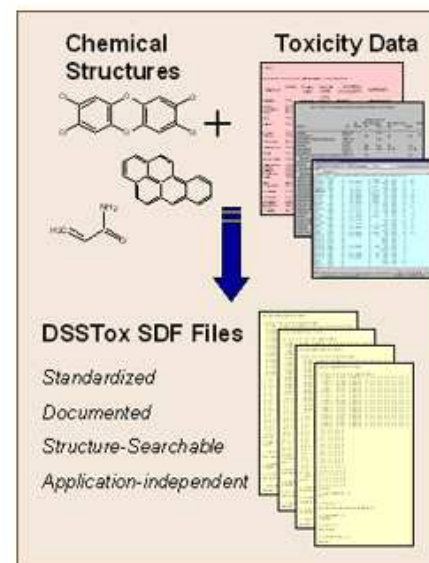
***Launch of DSSTox Structure-Browser v1.0:

- A new structure-search capability for published [DSSTox Data Files](#), allows users to search by [DSSTox Standard Chemical Fields](#) and includes options for:

- **Text Search:** Chemical Name, CAS RN, InChI, Formula
- **Structure Search (Exact, Substructure, Similarity):** SMILES or Structure Drawing Tool entry

***Revised Standard ID Fields for all DSSTox files:

- Modified [Record, File, Chemical, and Substance ID fields](#) to index all unique DSSTox structures and substances, also with respect to file record and version

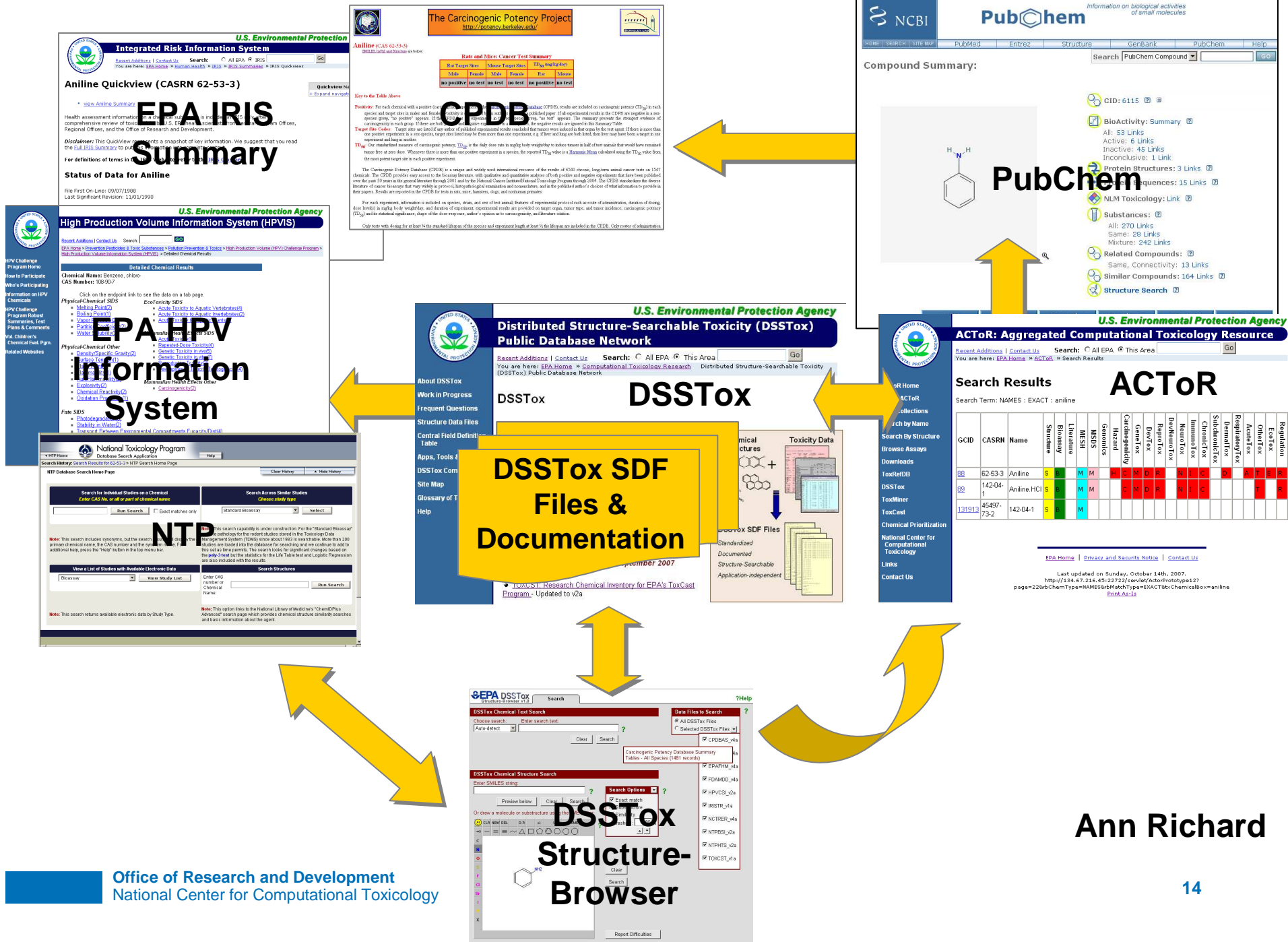


- [DSSTox Graphic Flowchart](#)
- [DSSTox Project Goals](#)
- [DSSTox Publications](#)

DSSTox Data Files: [Details>](#)

CPDBAS v4a	1481	15Jun2007	**New content
DBPCAN v4a	209	15Jun2007	
EPAFHM v4a	617	15Jun2007	
FDAMDD v3a	1216	25Jul2007	
HPVCSI v2a	3548	15Aug2007	**New content
IRISTR v1a	544	28Jul2007	**New file
NCTRER v4a	232	15Jun2007	
NTPBSI v2a	2293	24Aug2007	**Updated content
NTPHTS v1a	1408	25Jul2007	
TOXCST v2a	320	25Sep2007	**Updated

[More on Data File Types](#)



Ann Richard

Distributed Structure-Searchable Toxicity (DSSTox) Database Network

- The DSSTox website provides a public forum for publishing downloadable, structure-searchable, standardized chemical structure files associated with toxicity data
- Quality controlled, linked to physicochemical and toxicological data
- This approach and platform is an example of how to link HTS data to historical toxicological test results

<http://epa.gov/ncct/dsstox/>



ToxCast™ Program for Prioritizing Toxicity Testing of Environmental Chemicals (David Dix and Keith Houck)

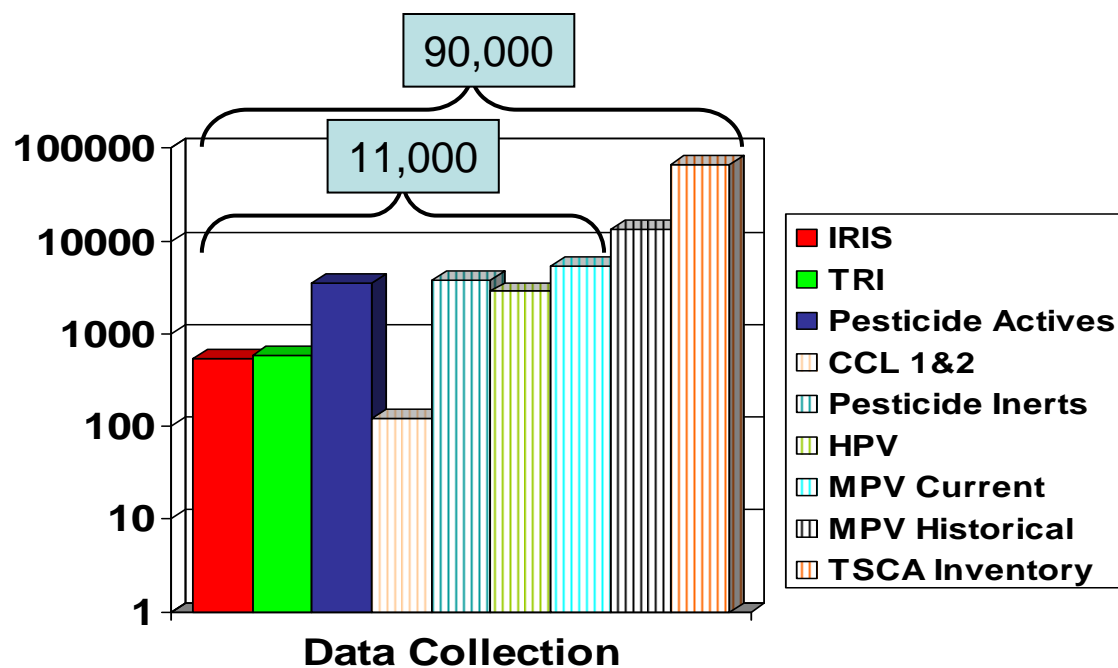
Predicting Hazard, Characterizing Toxicity Pathways, and Prioritizing the Toxicity Testing of Environmental Chemicals

- The U.S. EPA has identified a clear need to develop methods to evaluate a large number of environmental chemicals for their potential toxicity.
- Doing so will enable EPA to prioritize the use of its limited testing resources on those chemicals that present the greatest likelihood of risk.
- In 2007, EPA launched ToxCast™ in order to develop a cost-effective approach for prioritizing the toxicity testing of large numbers of chemicals in a short period of time.
- Using data from state-of-the-art high throughput screening (HTS) bioassays developed in the pharmaceutical industry, ToxCast™ is building computational models to forecast the potential human toxicity of chemicals.

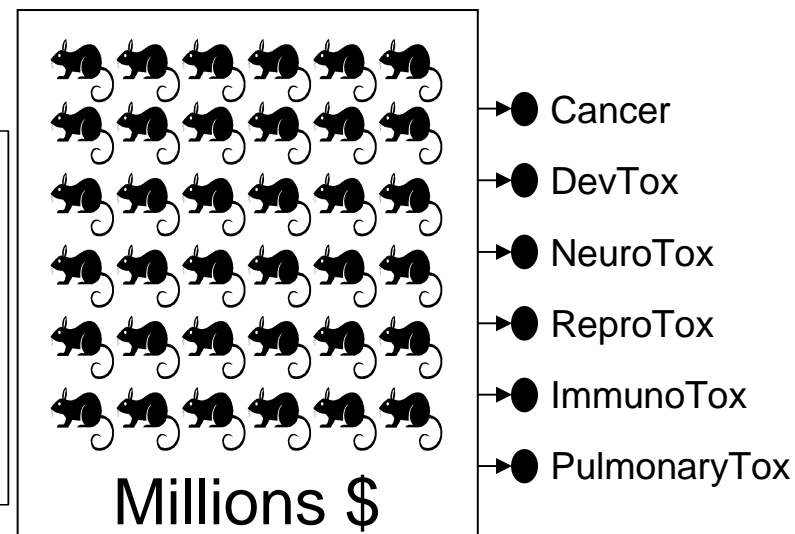
<http://epa.gov/ncct/toxcast/>

Putting Numbers on the Problem

Too Many Chemicals



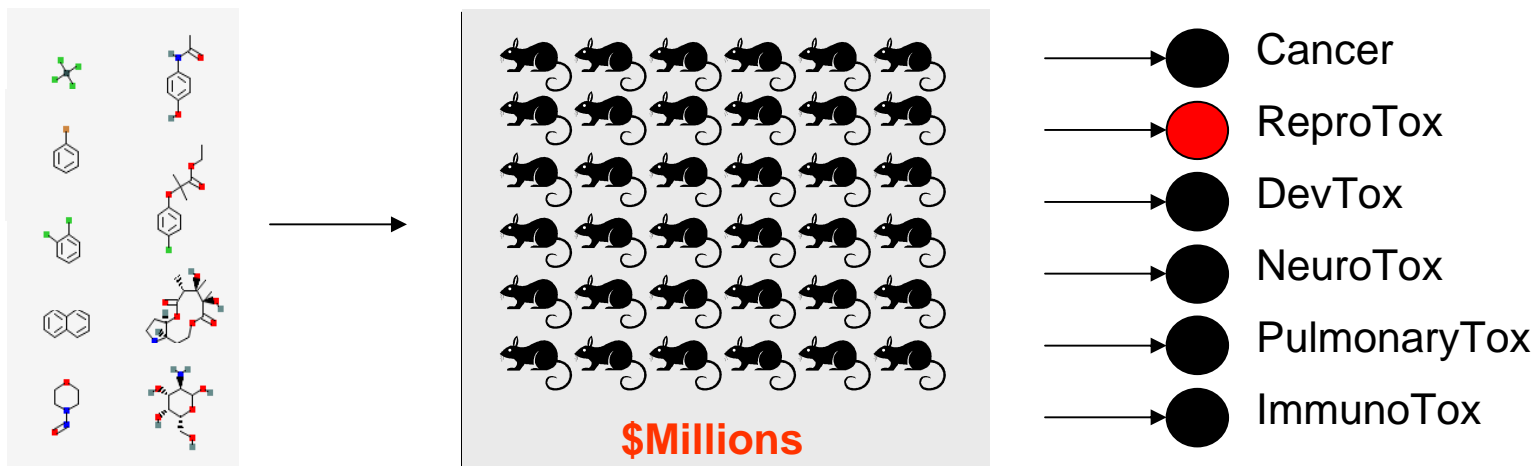
Too High a Cost



...and not enough data.

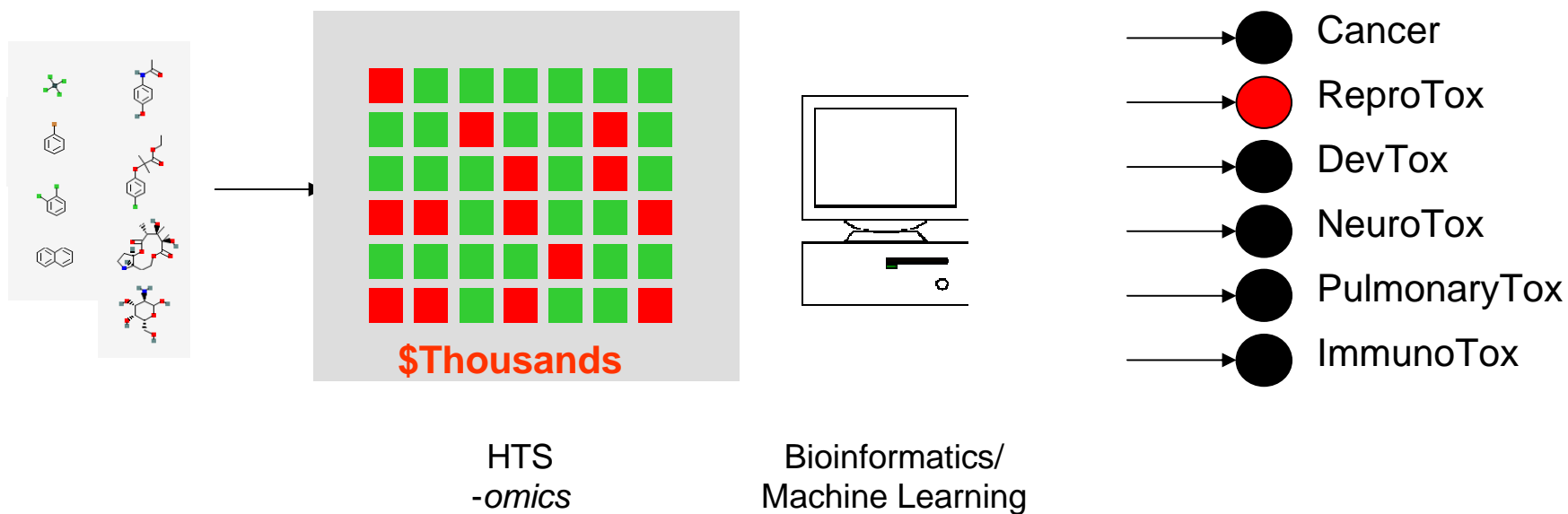
Current Approach for Toxicity Testing

in vivo testing



Future of Toxicity Testing

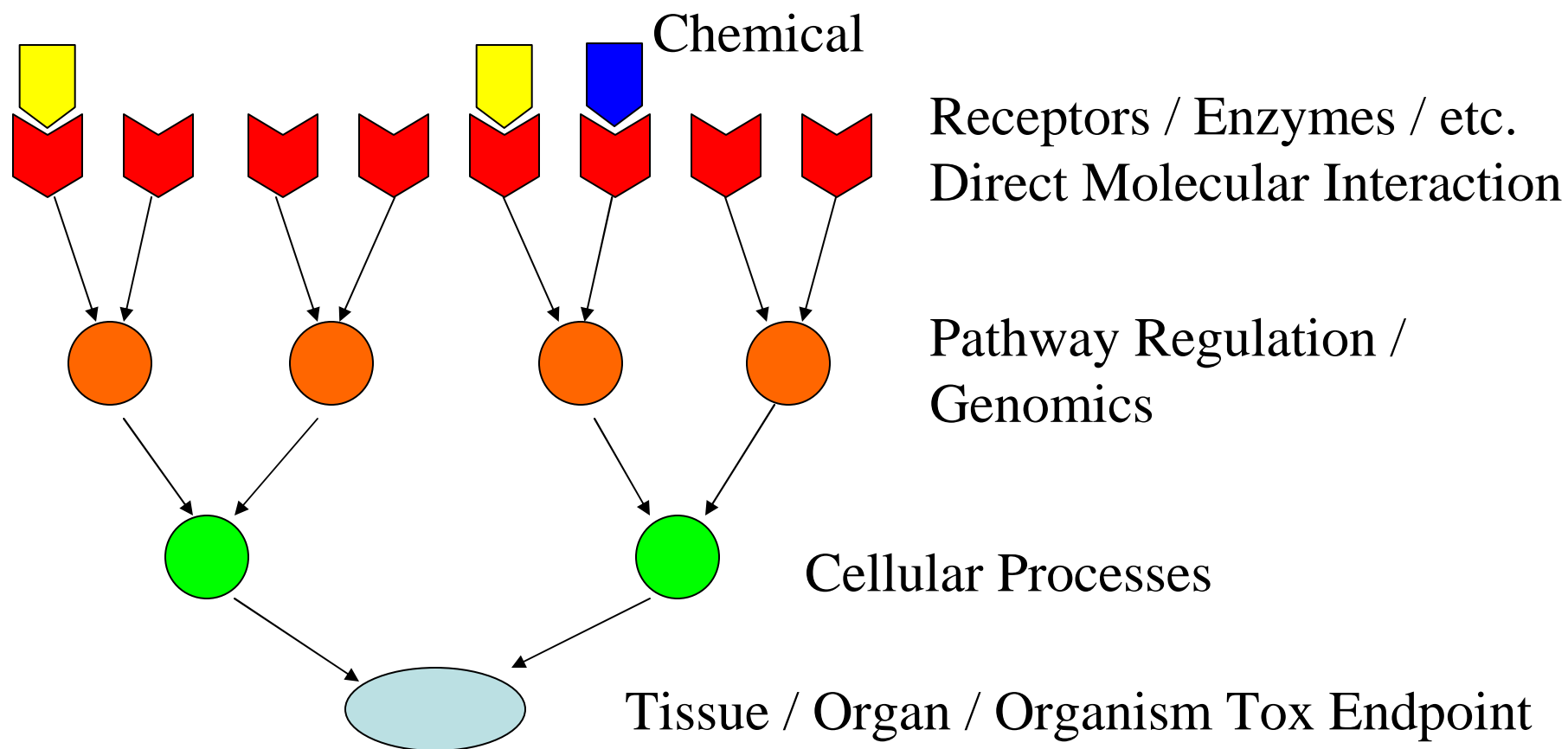
in vitro testing *in silico* analysis



ToxCast Data Analysis Goals

- Find “Signatures” that use *in vitro* & *in silico* assays to predict *in vivo* endpoints
 - Focus on current guideline-defined protocols and endpoints
 - Accurate enough to be used for prioritization
 - Inexpensive and fast enough to be used for screening
 - The predictive power of each signature must be characterized
- Understand mechanism of toxicity
 - Help characterize toxicity pathways
 - Facilitate cross-species extrapolation
 - Provide input to models for low-dose extrapolation

Data Spans Many Levels of Biological Organization



ToxCast™ - Phase I

- Profiling over 300 well-characterized chemicals (primarily pesticides) in over 400 HTS endpoints.
- Endpoints include biochemical assays of protein function, cell-based transcriptional reporter assays, multi-cell interaction assays, transcriptomics on primary cell cultures, and developmental assays in zebrafish embryos.
- Almost all of the Phase 1 compounds have been tested in traditional toxicology tests, including developmental toxicity, multi-generation studies, and sub-chronic and chronic rodent bioassays.
- ToxRefDB, a relational database being created to house this information, will contain nearly \$1B worth of toxicity studies in animals when completed.

<http://epa.gov/ncct/toxcast/>

ToxCast™ - Data management and Informatics

- ToxRefDB is integrated into a more comprehensive data management system developed by NCCT called ACToR (Aggregated Computational Toxicology Resource), that manages the large-scale datasets of ToxCast™.
- ACToR is comprised of several independent data repositories linked to a common database of chemical structures and properties, and to tools for development of predictive HTS and genomic bioactivity signatures that strongly correlate with specific toxicity endpoints from ToxRefDB.
- These ToxCast™ signatures will be defined and evaluated by their ability to predict outcomes from existing mammalian toxicity testing, and identify toxicity pathways that are relevant to human health effects.

<http://epa.gov/ncct/databases.html>

ToxCast™ - Phase II

- ToxCast™ will screen additional compounds representing broader chemical structure and use classes, in order to evaluate the predictive bioactivity signatures developed in Phase I.
- Following successful conclusion of Phases I and II, ToxCast™ will provide EPA regulatory programs an efficient tool for rapidly and efficiently screening compounds and prioritizing further toxicity testing.

<http://epa.gov/ncct/toxcast/>



Virtual Liver Project (Imran Shah)

Virtual Liver – Assessing for Risk

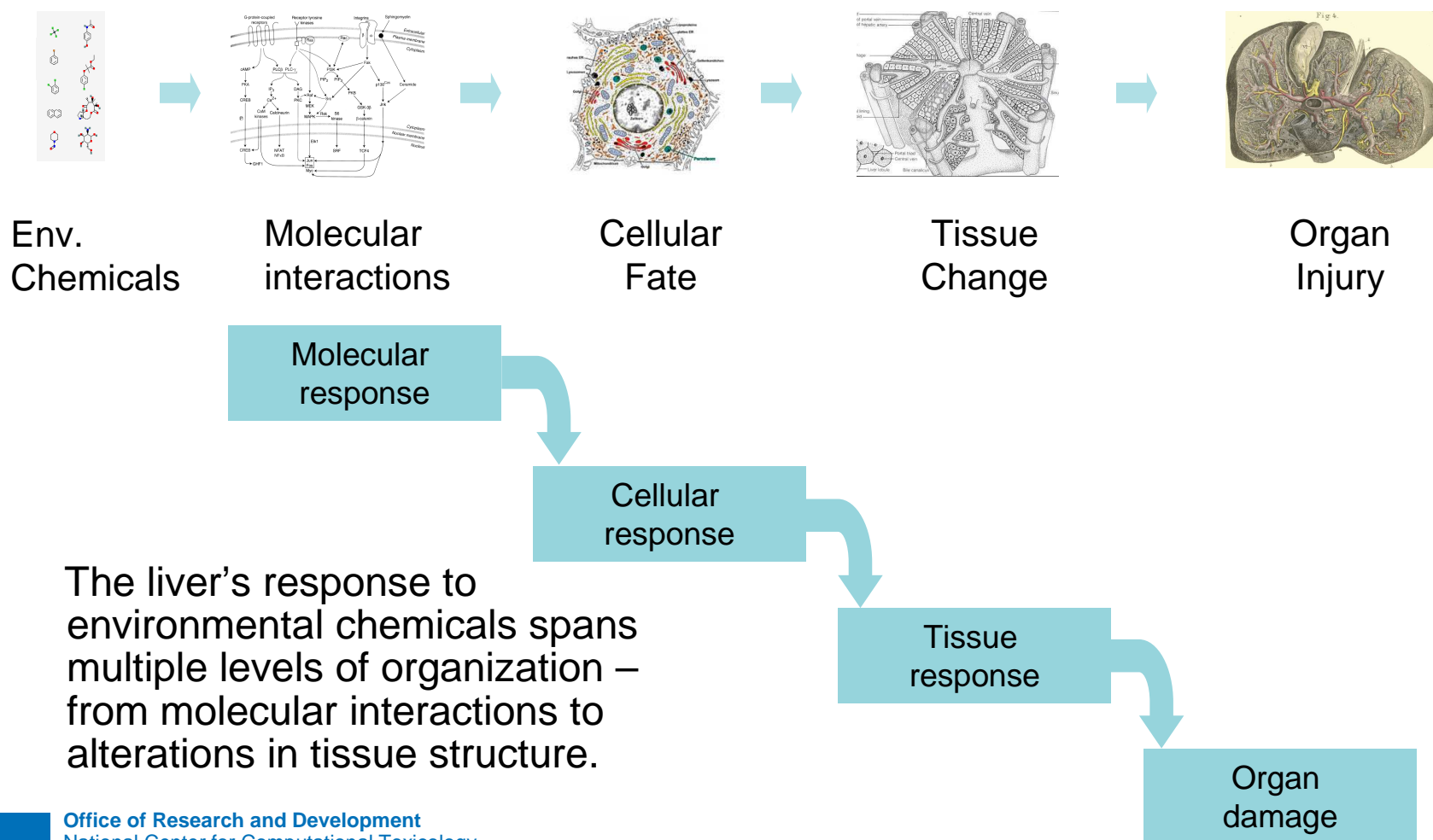
- The liver plays a key role in removing chemicals from the body and frequently shows the earliest signs of their harmful effects.
- Animal testing gives useful information about the effects of environmental chemicals, however, the relevance of these findings to humans is not always clear.
- To address this challenge scientists at The National Center for Computational Toxicology are leading the development of a "Virtual Liver": a large-scale biologically-based computer model of the organ.
- The long-term objective of the Virtual Liver is to simulate the effect of chemicals accurately, efficiently and more humanely.

http://epa.gov/ncct/virtual_liver/

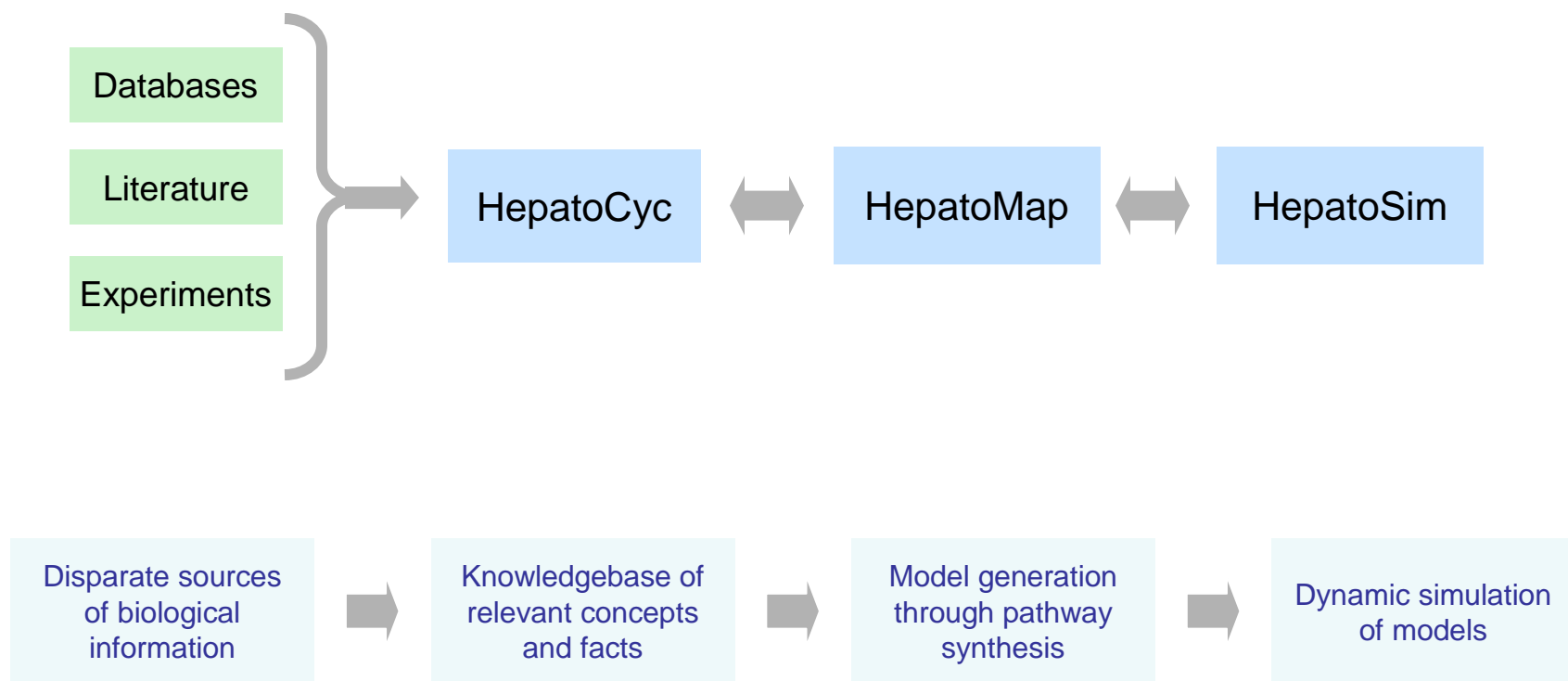
Virtual Liver – Goal

- The ultimate goal for applying HTP technologies in toxicity testing is to establish *in vitro* “signatures” of *in vivo* rodent and human toxicity.
- To achieve this objective, computational methods are required that can simulate *in silico* the biology of a given organ system.
- The goal of the Virtual Liver project is to develop models for predicting liver injury due to chronic chemical exposure by simulating
 - Dynamics of perturbed molecular pathways
 - Linkage with adaptive or adverse processes leading to alterations of cell state
 - Integration of responses

Linking Chemicals to Organ Injury



Virtual Liver: Computational Framework for Multiscale Biological Modeling



Virtual Liver Project - Approach

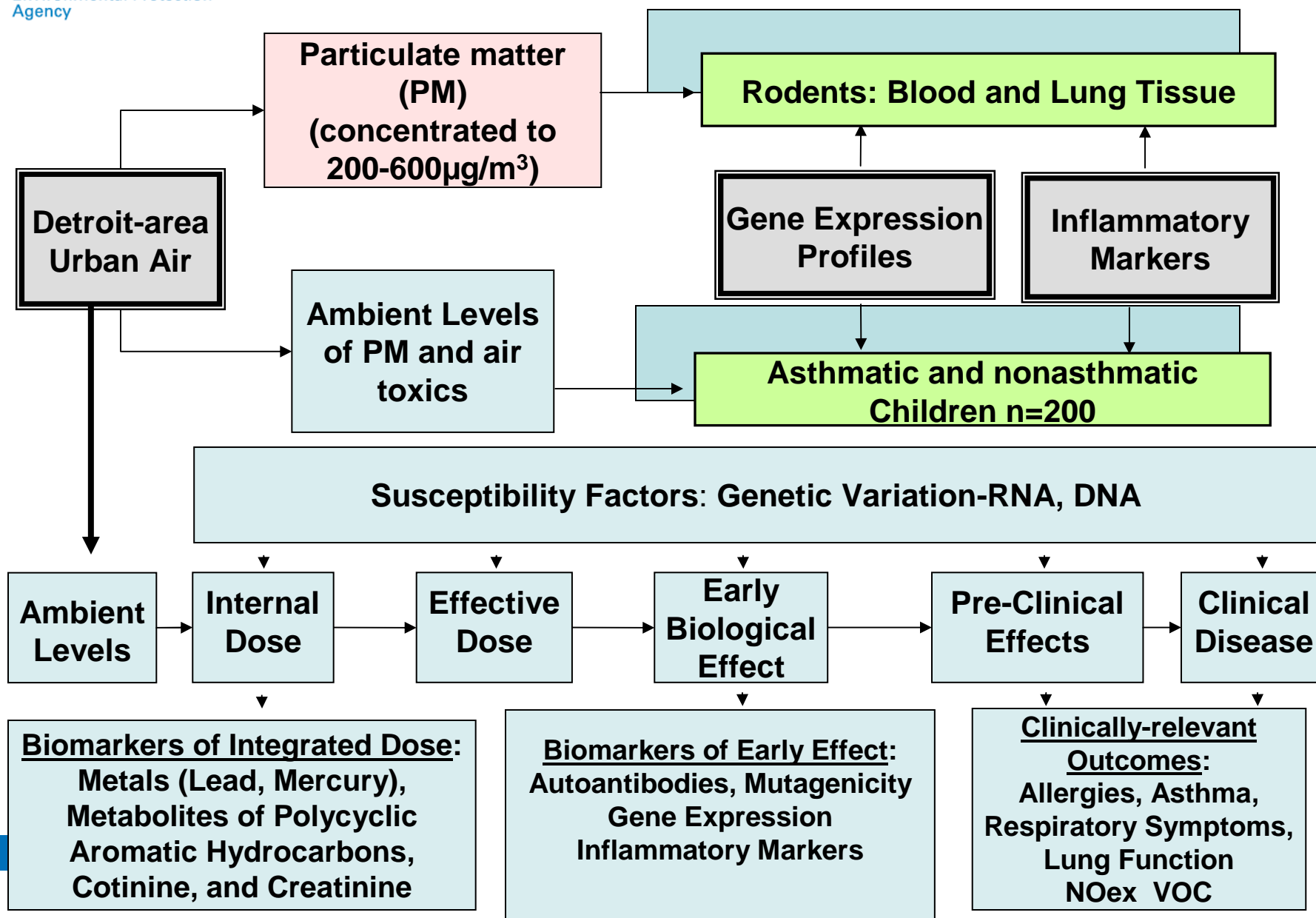
- Develop a knowledgebase for qualitatively describing species-specific toxicity pathways due to exposure to chemicals
- Develop a virtual liver tissue that lays the foundation for quantitatively predicting the risk of non-genotoxic neoplastic lesions due to activation of certain genetic regulatory elements (ie, nuclear receptors and other transcription factors) in humans.

http://epa.gov/ncct/virtual_liver/

Mechanistic Indicators of Childhood Asthma (MICA) Study (Jane Gallagher)

Mechanistic Indicators of Childhood Asthma (MICA)

- Apply state-of-the-art technologies to examine the interplay between environmental and genetic factors affecting asthma
- 150 asthmatic and 50 non-asthmatic children, ages 9-12 years (subset of Detroit Children's Health Study cohort)
- Collected multiple types of clinical, demographic, exposure, and gene expression data
- Consider markers of susceptibility, exposure, and effects to analyze and characterize combined risk factors that relate to asthma severity



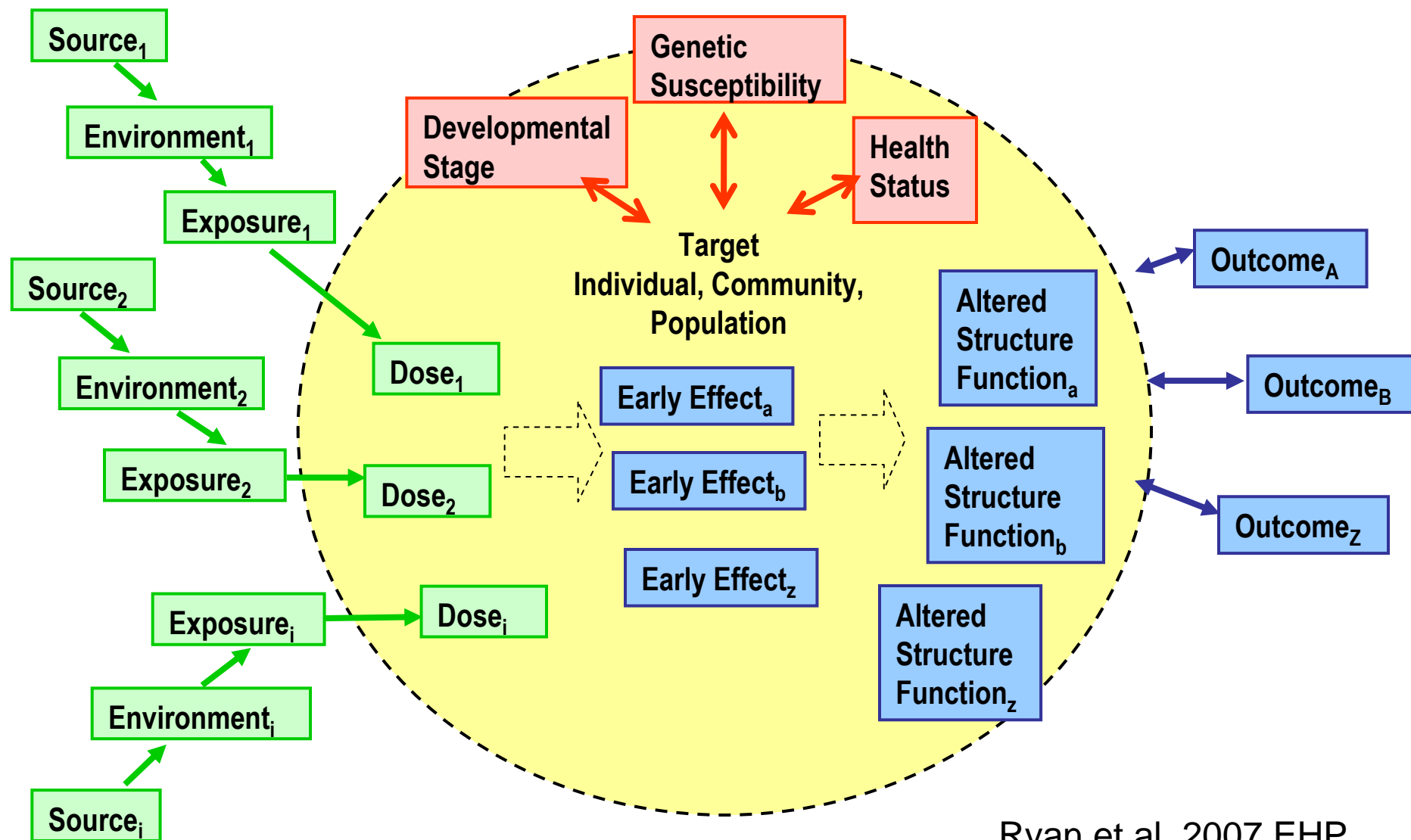
Collaborative NCCT-MICA Goals

- The objectives of the NCCT include advancement of a systems approach to evaluate complex relationships between
 - environmental factors
 - physiological biomarkers
 - health outcomes.
- NCCT collaborating to apply advanced statistical and machine learning methods to evaluate biomarker data collected in MICA
- Contribution of the NCCT component:
 - analyze genetic and gene expression data
 - use a systems biology approach to put data into framework for evaluating ecogenetics

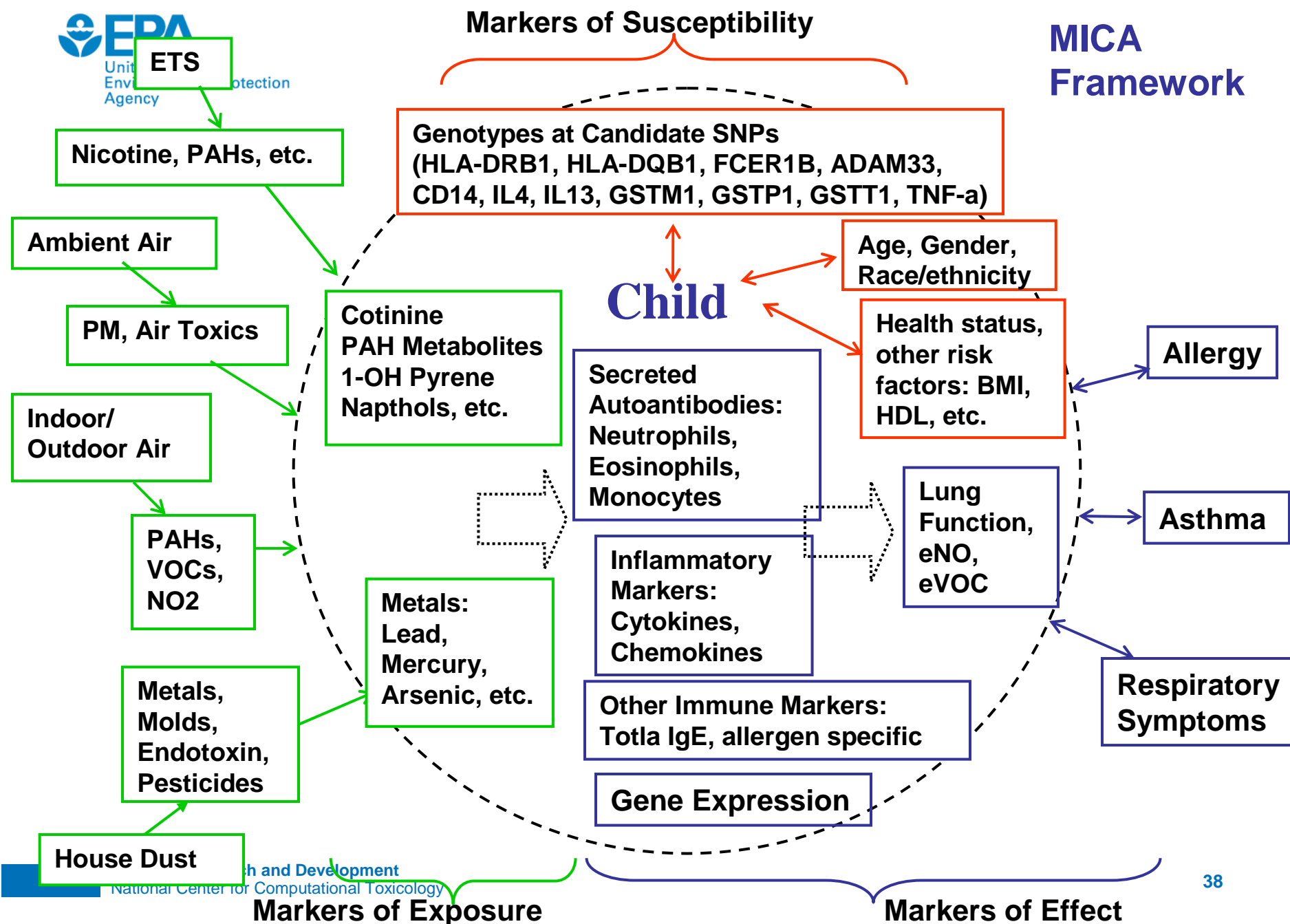
Data Analysis Methods

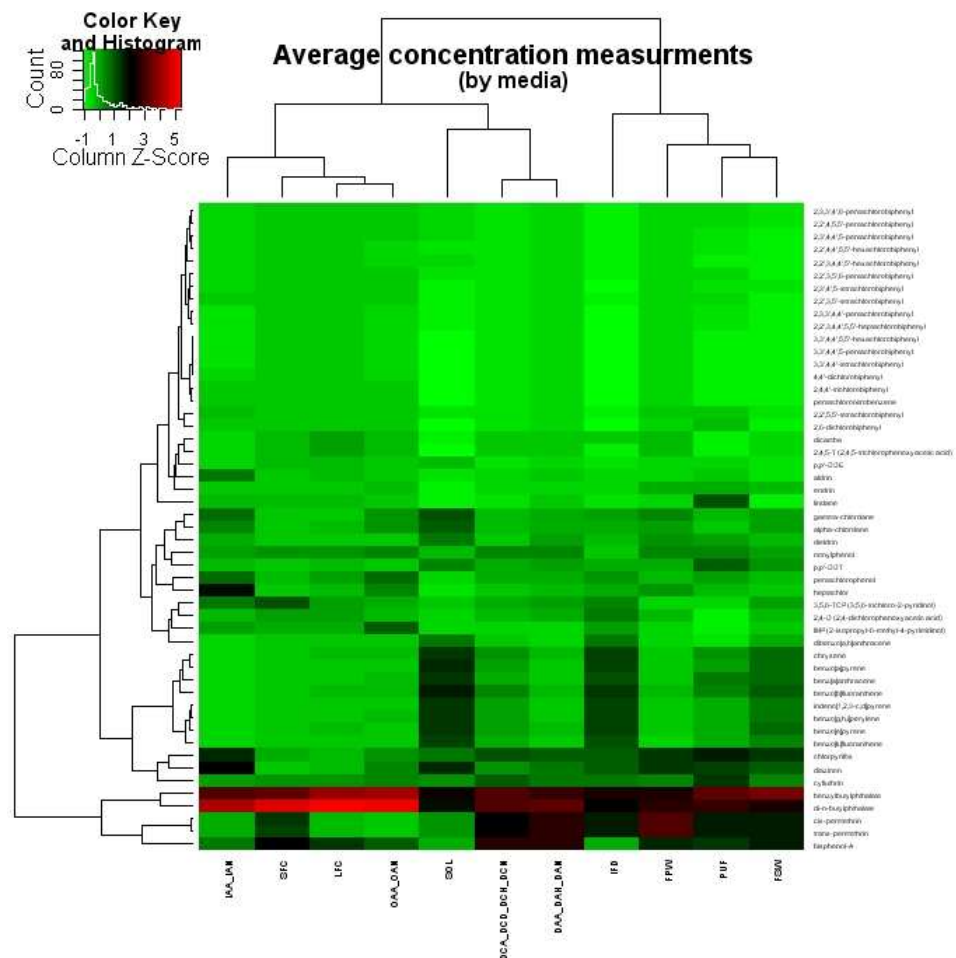
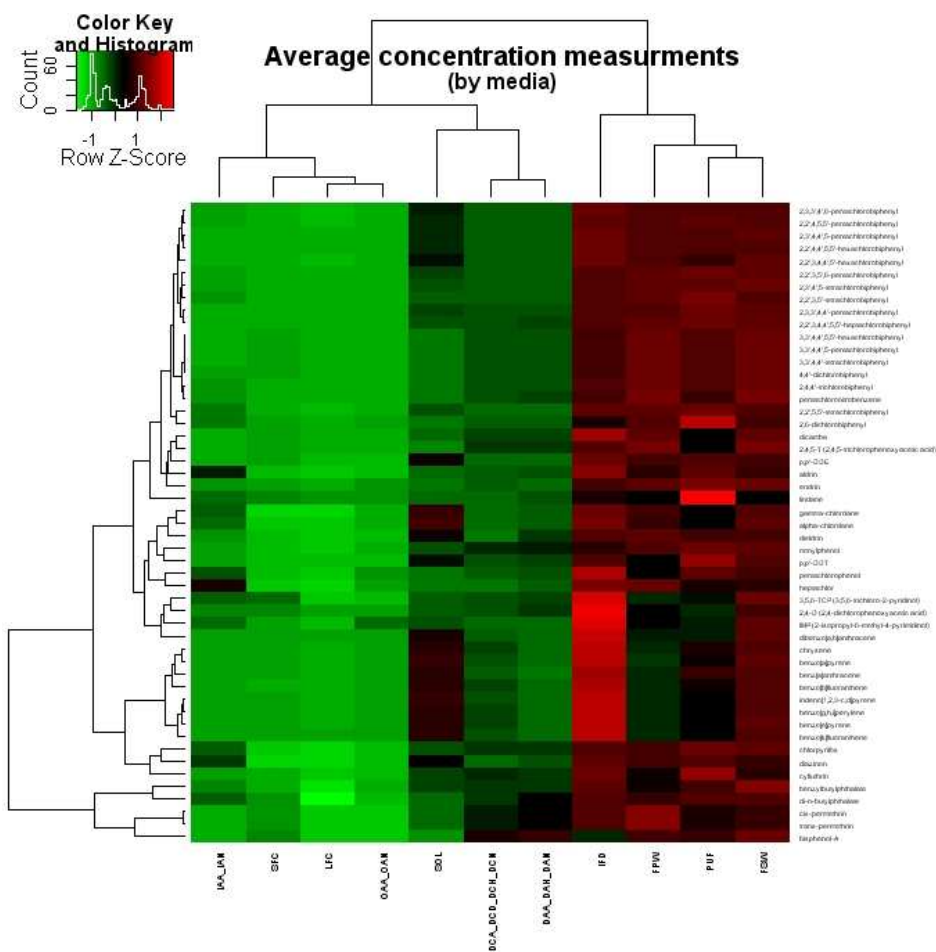
- Traditional statistics: linear regression, logistic regression, ANOVA, linear discriminant analysis.
- Machine learning: recursive partitioning trees, bootstrap aggregation (bagging) techniques, evolutionary computation-optimized classifiers, multifactor dimensionality reduction, random forests.
- Bioinformatics: protein interaction databases, knowledge (literature) mining tools, biological pathway database and inference software.
- Graphical approaches: cluster diagrams, expression “heat” maps, dendrograms, overlaid scatter plots (both exploratory and summary), distributional “violin” plots, regression plots.

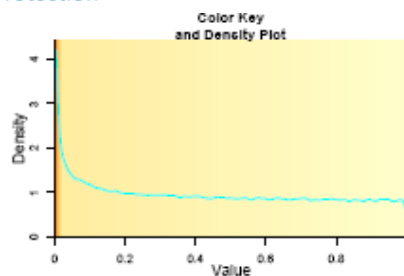
Framework



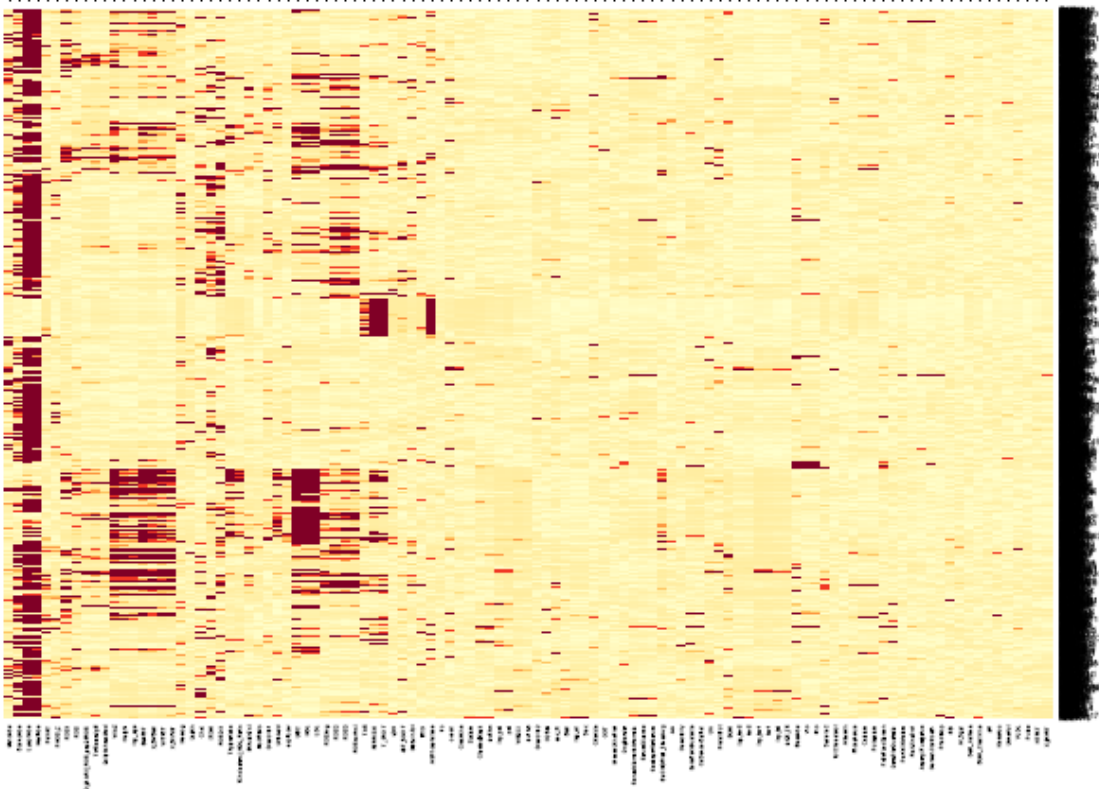
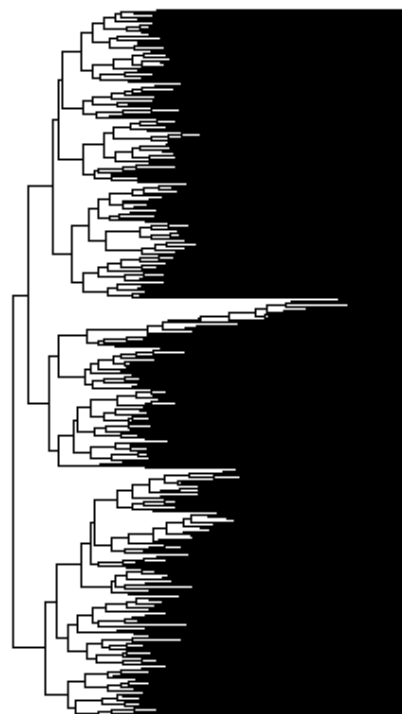
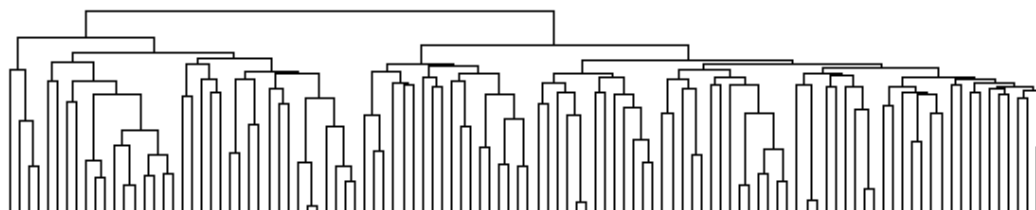
Ryan et al. 2007 EHP







REDI_MAS5_20080204.Rdata
(8737 genes)



Can Gene Expression Distinguish Subtypes of Asthmatics

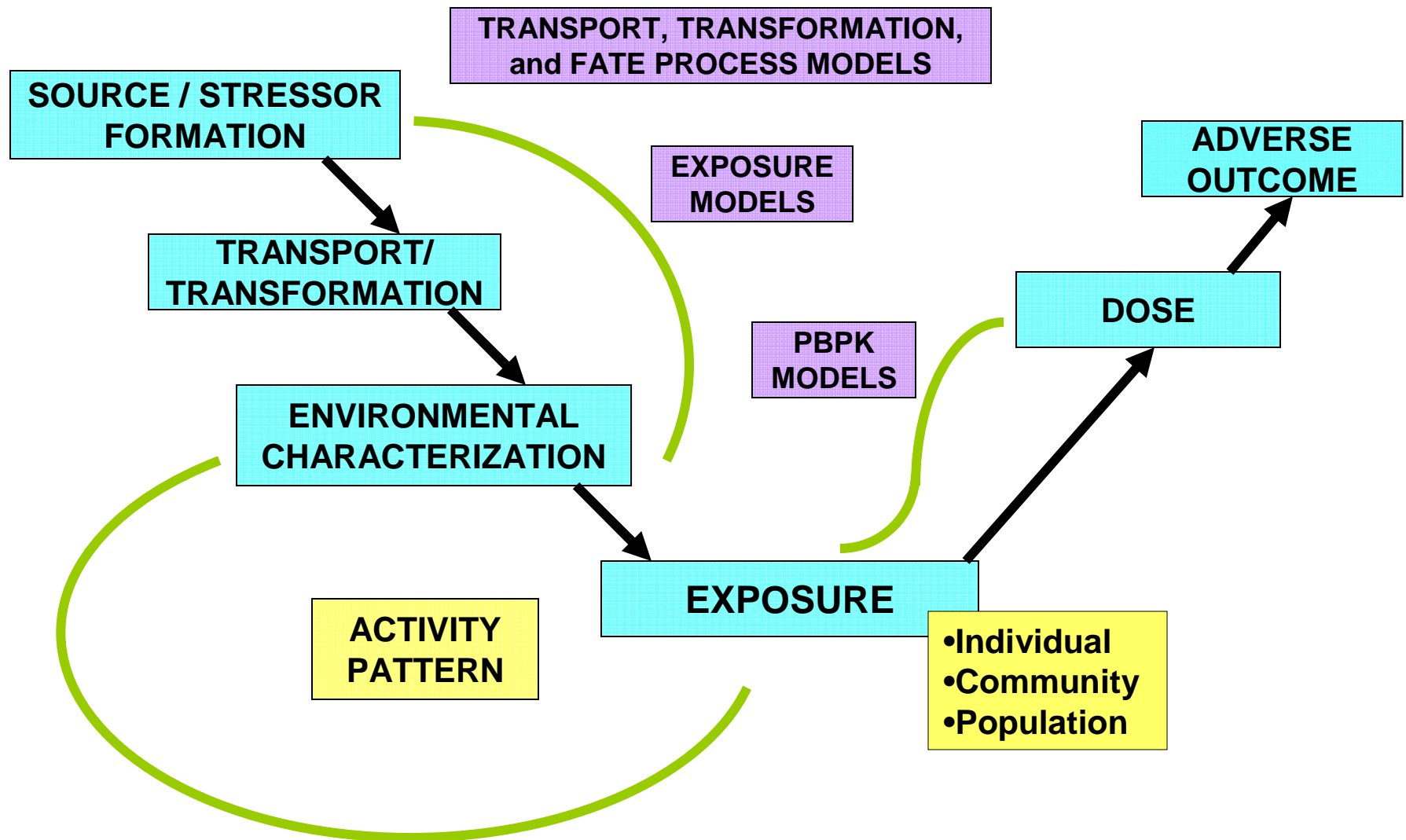
- Clustering of subjects via “global” gene expression suggests different mechanisms among the asthmatics
- By investigating the discriminating genes for these groups, it may be possible to understand the mechanistic basis for phenotypic differences
- Can we link gene expression modules to other biomarkers?
 - Inform mechanism
- Can we relate blood expression to lung expression using the results from Phase I study?
- We can assemble gene expression networks to investigate mechanism in more detail.
 - Role for genetics?

Stephen Edwards



Exposure Science for Toxicity Testing (Elaine Cohen Hubal)

Purpose of Exposure Assessment



Toxicity Testing in the Twenty-first Century: A Vision and a Strategy

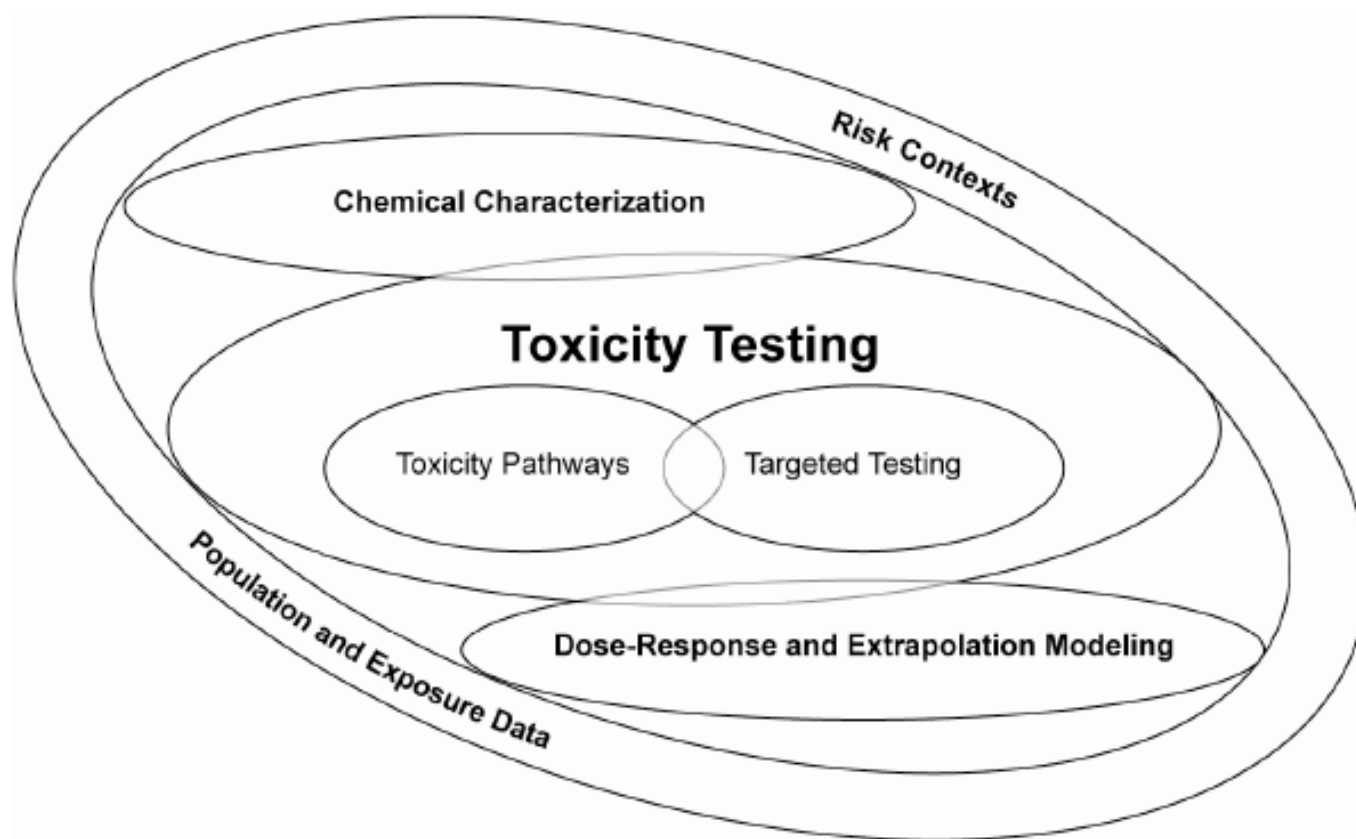


FIGURE 2-3 The committee's vision is a process that includes chemical characterization, toxicity testing, and dose-response and extrapolation modeling. At each step, population-based data and human exposure information are considered, as is the question of what data are needed for decision-making.

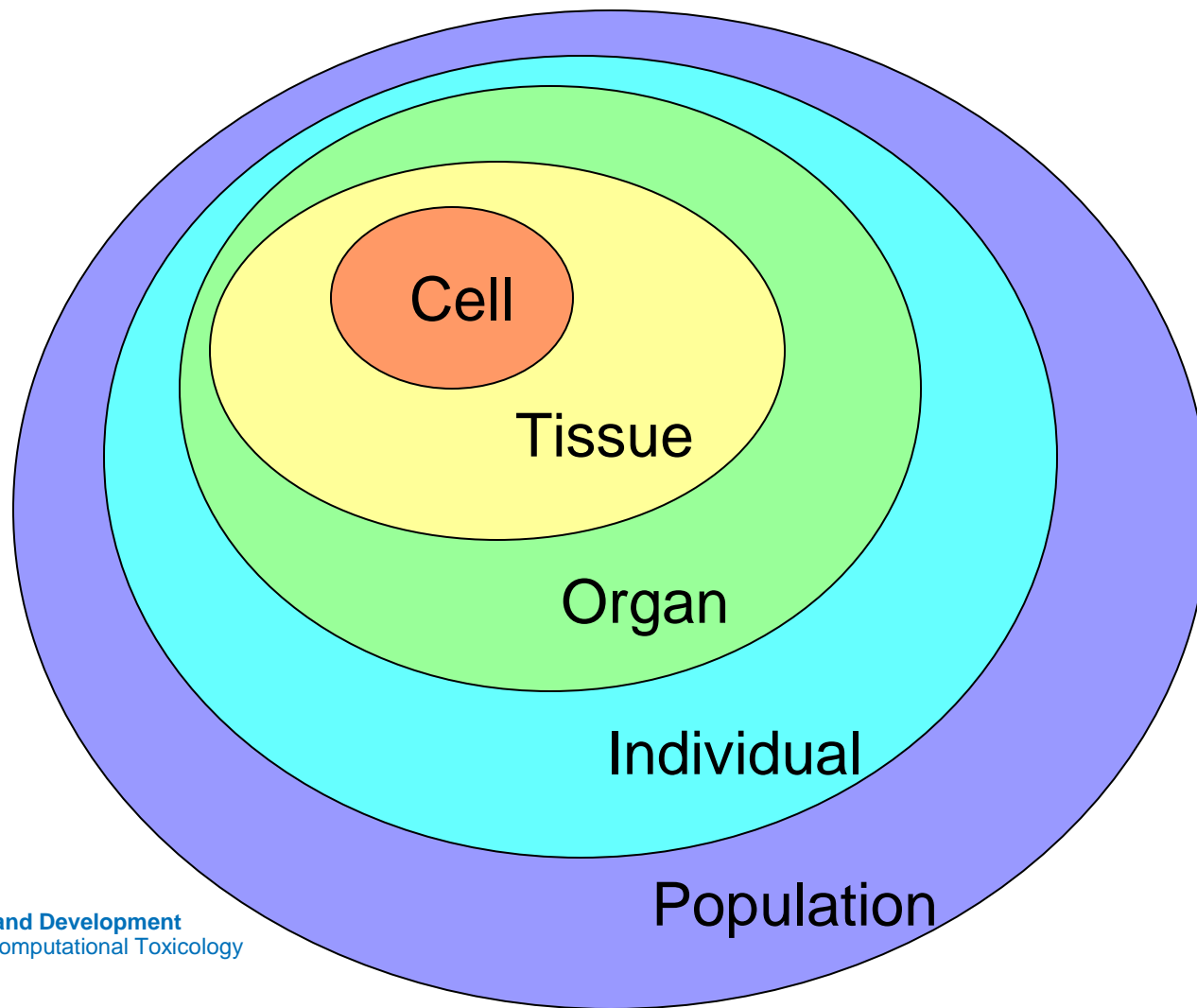
Toxicity Testing in the Twenty-first Century: A Vision and a Strategy

“The vision emphasizes the generation and use of population-based and human exposure data where possible for interpreting test results and encourages the collection of such data on important chemicals with biomonitoring, surveillance, and epidemiologic studies. Population-based and human exposure data, along with the risk context, will play a role in both guiding and using the toxicity information that is produced.”

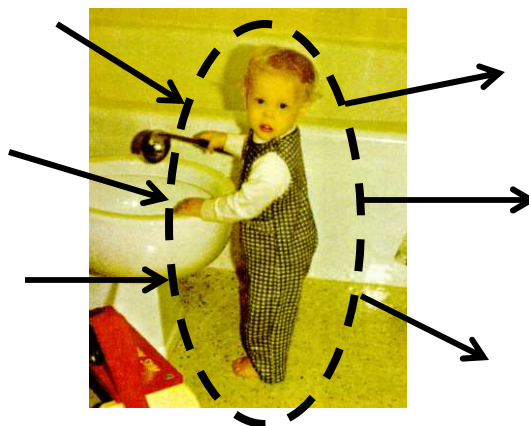
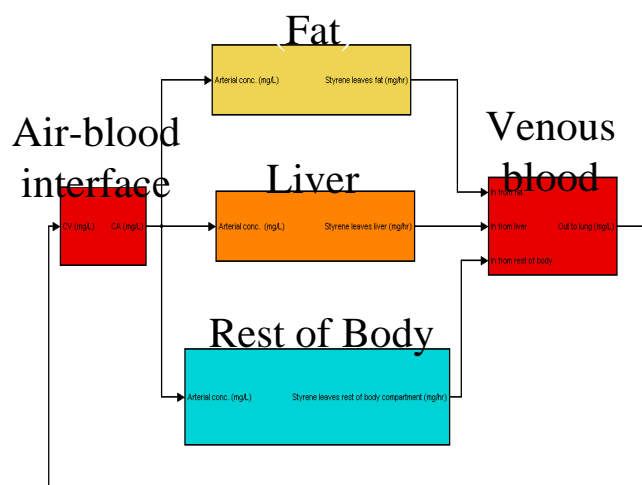
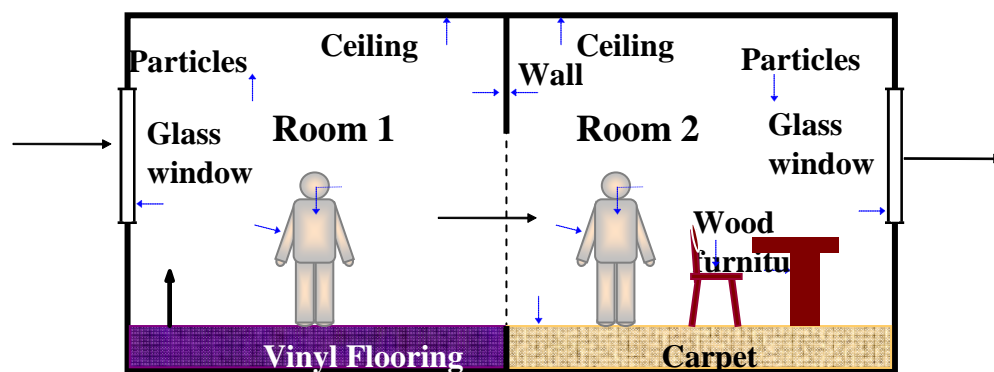
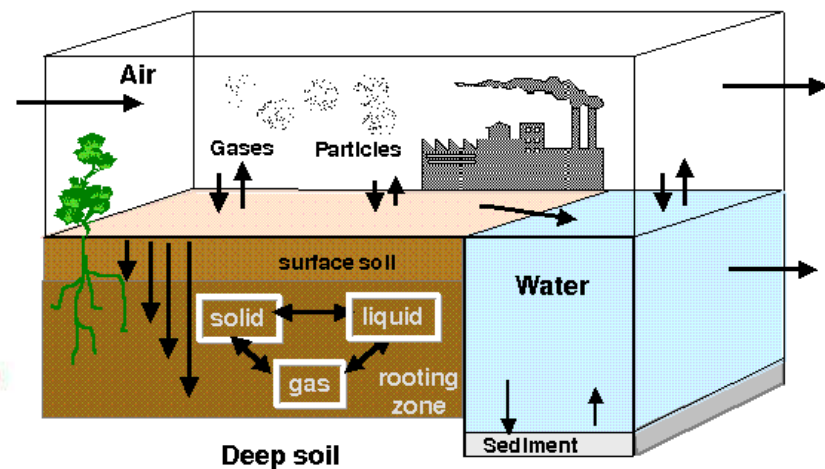
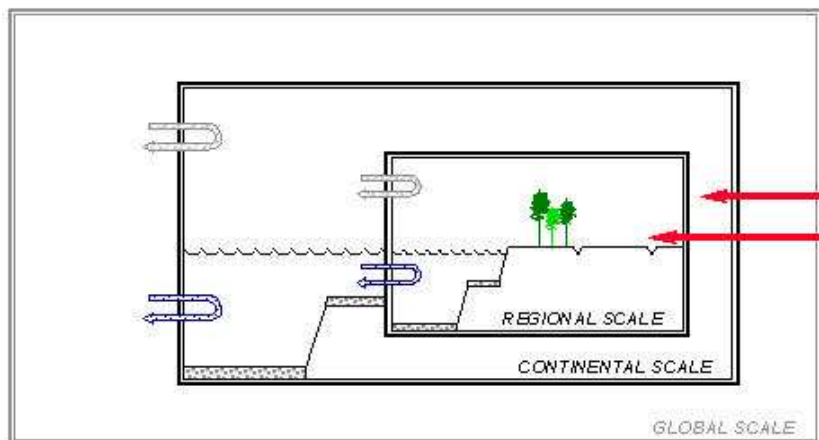
Research Questions

- How do we use information on host susceptibility and background exposures to interpret and extrapolate in vitro test results?
- How do we use human exposure data to select doses for toxicity testing so we develop information on biological effects at environmentally relevant exposures?
- How can we relate human exposure data from biomonitoring surveys to concentrations that perturb toxicity pathways to identify potentially important exposures?

Levels of Biological Organization



Mass Balance Models



Exposure Science for Toxicity Testing

- Relevance for Real World
 - Real Individuals, Populations
 - Real Doses
 - Real Patterns (Time Frames)
- Challenge to Link Across Scales
 - System Definition
 - Model Resolution
 - Exposure Data

Acknowledgements

- NCCT – Bob Kavlock, Jerry Blancato
- DSSTox - Ann Richard
- ToxCast TM - David Dix and Keith Houck, plus many
- Virtual Liver - Imran Shah
- MICA Study - Jane Gallagher, Stephen Edwards, David Reif, plus many

Disclaimer

Although this work was reviewed by EPA and approved for presentation, it may not necessarily reflect official Agency policy.