United States Environmental Protection Agency

New Chemical/Biological Profiling and Informatics Approaches for Exploring Mutagenicity & Carcinogenicity: *Updates of EPA ToxCast & Tox21 Programs*

ICEM, Florence, Italy, August 23, 2009

Ann Richard
richard.ann@epa.gov

Office of Research and Development
National Center for Computational Toxicology

EPA
United States
Environmental Protection
Agency

"…to integrate modern computing and information technology with molecular biology to improve Agency prioritization of data requirements and risk assessment of chemicals"

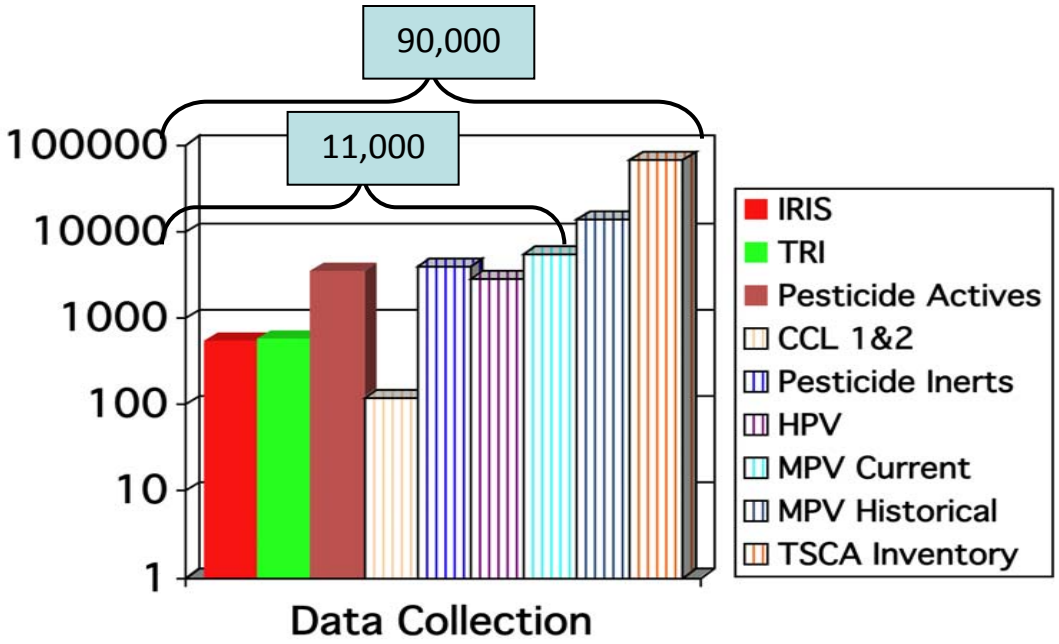
Decision Support Tools for High-Throughput Risk Assessment

Office of Research and Development
National Center for Computational Toxicology

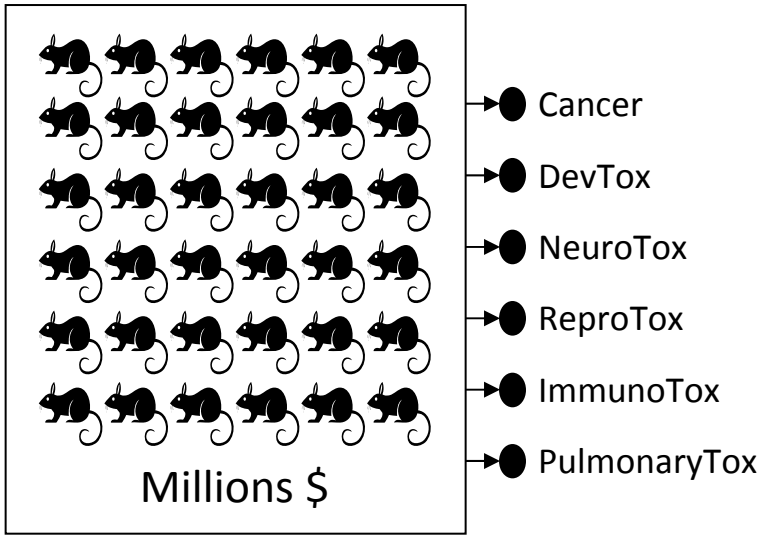
http://www.epa.gov/ncct

Change Needed Because …..

Too Many Chemicals



Too High a Cost



…and not enough data.

Judson, et al *EHP, 2008*

Future of Chemical Toxicity Testing

NAS report, July 2007

THE NATIONAL ACADE

REPORT IN BRIEF

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.

Science, February 2008

POLICYFORUM

TOXICOLOGY

Transforming Environmental Health Protection

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.

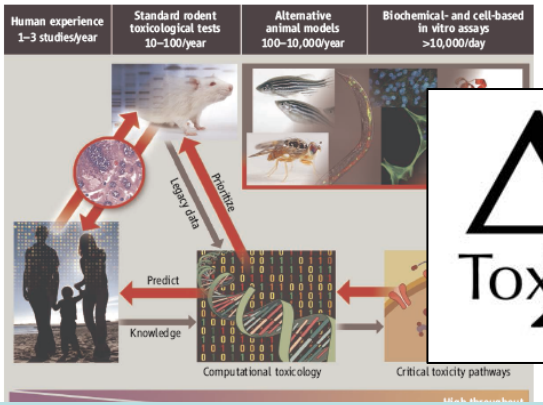
Francis S. Collins,[1*†] George M. Gray,[2*] John R. Bucher[3*]

...n 2005, the U.S. Environmental Protection ...throughput screening (HTS) and other auto...ssays into its testing ...he EPA established the ... Computational Toxi...cology (NCGC). 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).

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 sub-

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-

tion, usually between 2 and 10 μM, and 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 μM, 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://mli.nih.gov/), are being made publicly available through Web-based databases [e.g., PubChem (http://pubchem.ncbi.nlm.nih.gov)]. In addition,



[1]Director, National Human Genome Research Institute (NHGRI), National Institutes of Health, Bethesda, MD 20892; [2]Assistant Administrator for the Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC 20460; [3]Associate Director, U.S. National Toxicology Program, National Institute of Environmental

ed from www.sciencemag.org on February 15, 2008

Science, August 21, 2009

Industrial Chemistry

NEWS

Putting Chemicals on a Path To Better Risk Assessment

Industry and regulators are embracing new technologies to move beyond slow, expensive, and perplexing animal tests

PRACTICALLY EVERY BOTTLE OF SUNSCREEN contains ethylhexyl methoxycinnamate, a compound that blocks ultraviolet rays. But there's a slight risk that it could pose a health hazard of its own, because the compound is

issued a 20-year strategic plan. It incorporates much of the advice of a major report* issued in 2007 by the National Academies' National Research Council (NRC): Use computers to predict toxicity and gather

mines what kinds of tests companies have to conduct, says Dan Newton of the Society of Chemical Manufacturers and Affiliates in Washington, D.C. "For the first time in a very long time, the prominence of attention being paid to chemical issues within the agency has risen significantly," says Richard Denison of the Environmental Defense Fund, an advocacy group in Washington, D.C.

Costly cornerstones

The current system of toxicity testing in the United States dates back to a 1937 tragedy, when a company advertised an antimicrobial drug called "Elixir of Sulfanilamide." More than 73 people died from toxic side effects. In

than would be expected for typical human exposures, requiring assumptions about

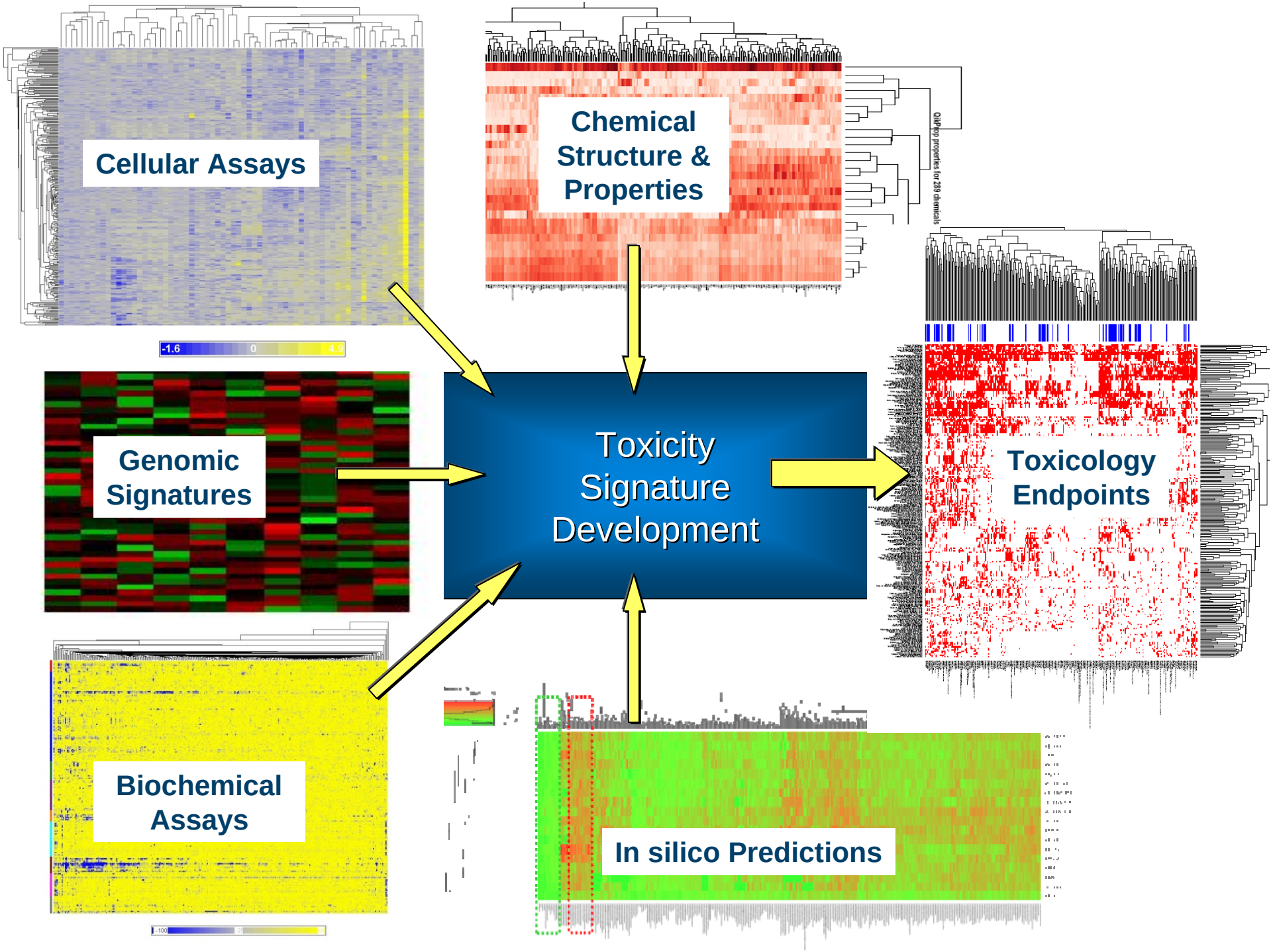
These powerful new approaches should help to address a number of challenges facing the

National Acade

Tox21

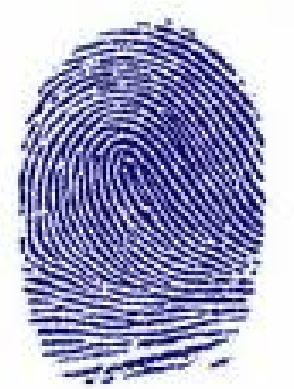
EPAs Contribution: The ToxCast Research Program

www.epa.gov/ncct/toxcast

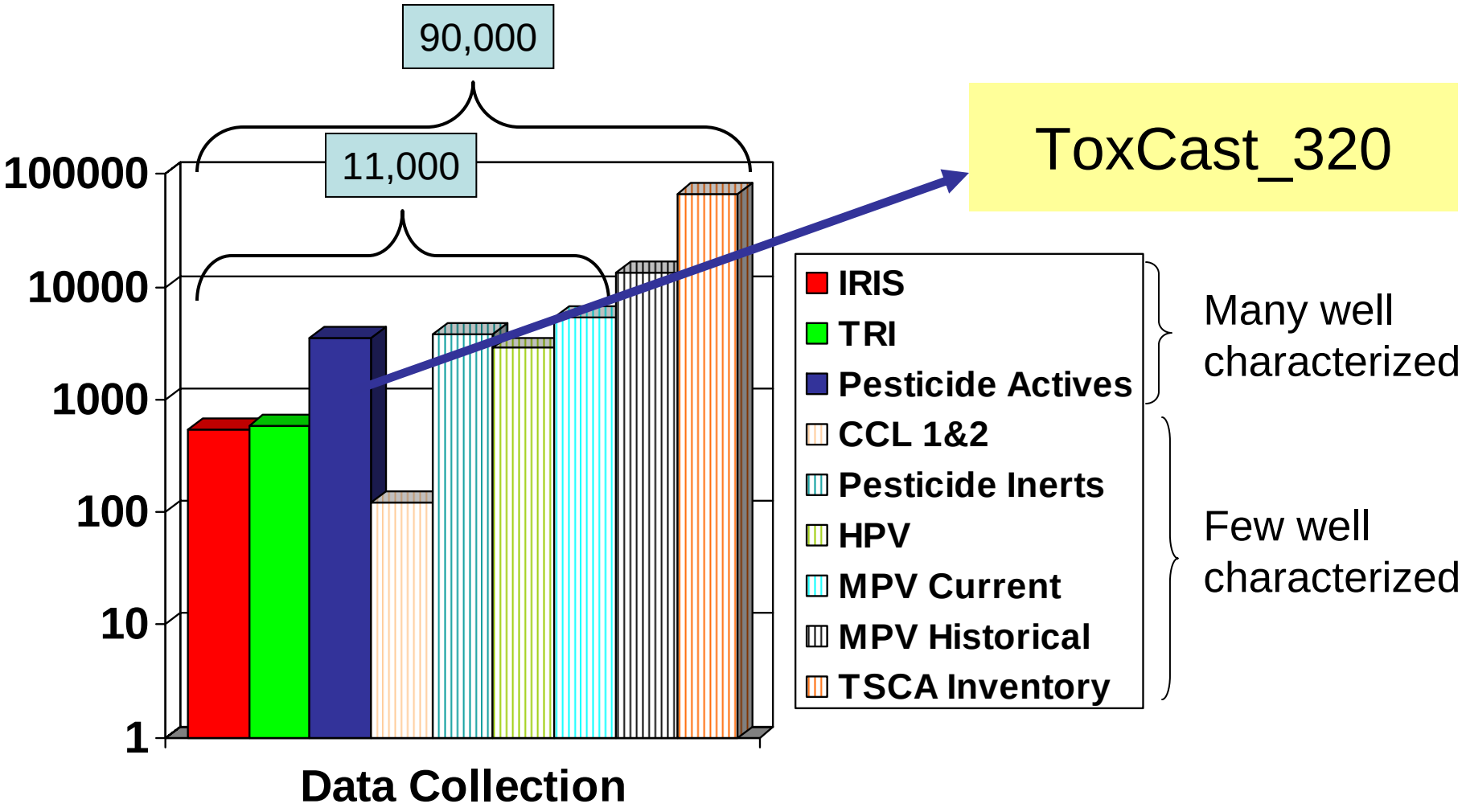

Cellular Assays
Chemical Structure & Properties
QikProp properties for 289 chemicals
Genomic Signatures
Toxicity Signature Development
Toxicology Endpoints
Biochemical Assays
In silico Predictions

ToxCast™ Background

- Addresses chemical screening and prioritization needs for pesticidal inerts, anti-microbials, CCLs, HPVs and MPVs
- Comprehensive use of HTS technologies to generate biological fingerprints and predictive signatures
- Coordinated with NTP and NIH/NCGC via Tox21
- Committed to stakeholder involvement and public release of data & tools
  - Communities of Practice- Chemical Prioritization; Exposure
  - NCCT website http://www.epa.gov/ncct/toxcast
    - *ACToR* *http://www.epa.gov/actor/*
    - *ToxRef DB* *http://www.epa.gov/ncct/toxrefdb/*
    - *DSSTox (PubChem)* *http://www.epa.gov/ncct/dsstox/*

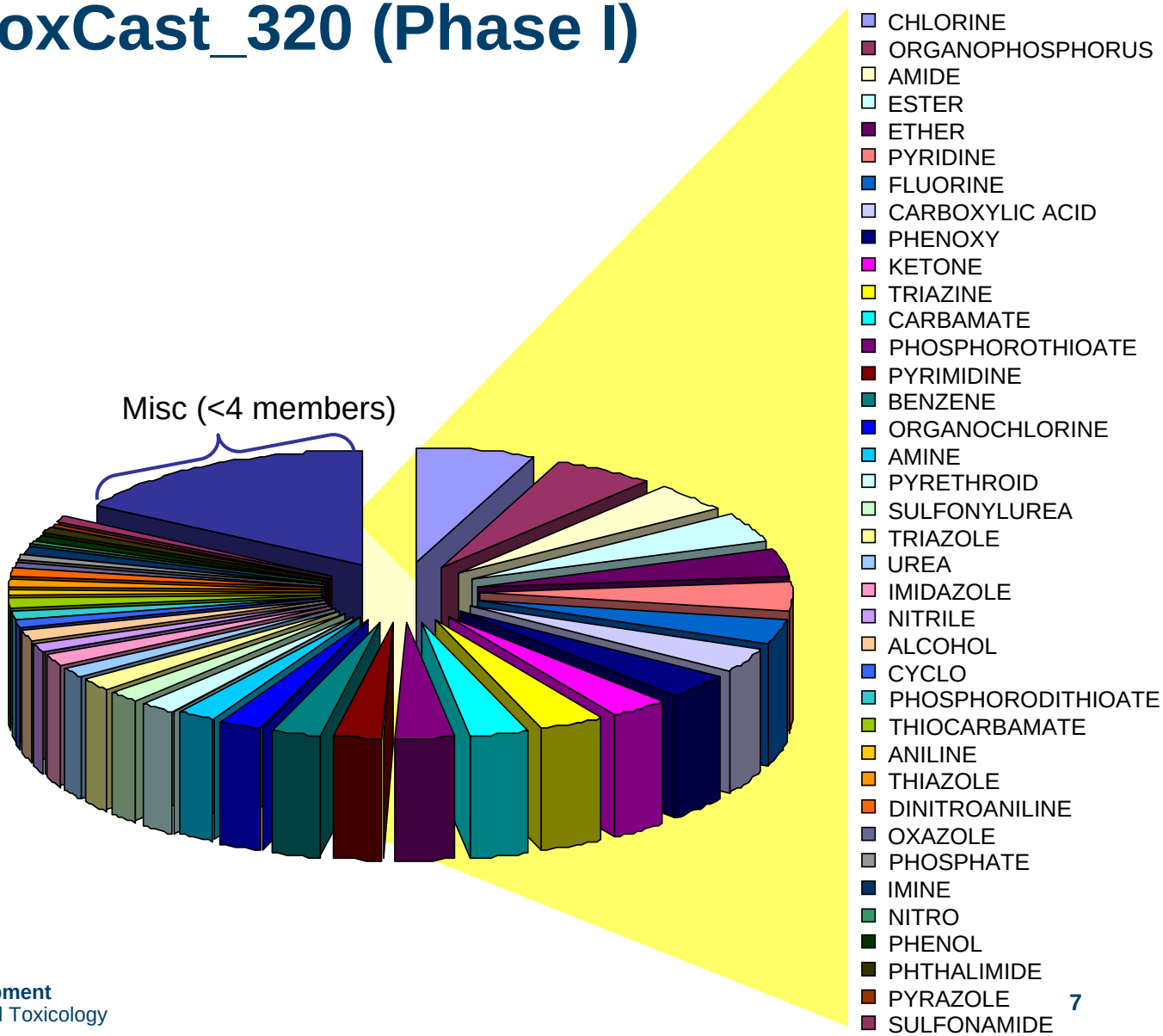


ToxCast Phase I Chemicals



Chemical Classes in ToxCast_320 (Phase I)

- 309 Unique Structures
- Replicates for QC
- 291 Pesticide Actives
- 9 Industrial Chemicals
- 13 Parent/Metablolite pairs
- 56/73 Proposed Tier 1 Endocrine Disruption Screening Program
- 14 High Production Volume Chemicals
- 11 HPV Challenge



EPA Pesticide Programs: Data Evaluation Records (DERs)

- Used for hazard identification and characterization
- Study Types
  - Chronic
  - Cancer
  - Subchronic
  - Multigeneration
  - Developmental
  - Others: DNT, Neurotox, Immunotox, Mutagenicity
- Derive Endpoints (NOAEL/LOAEL)
  - Systemic
  - Parental
  - Offspring
  - Reproductive
  - Maternal
  - Developmental
- Critical Effects for Endpoints

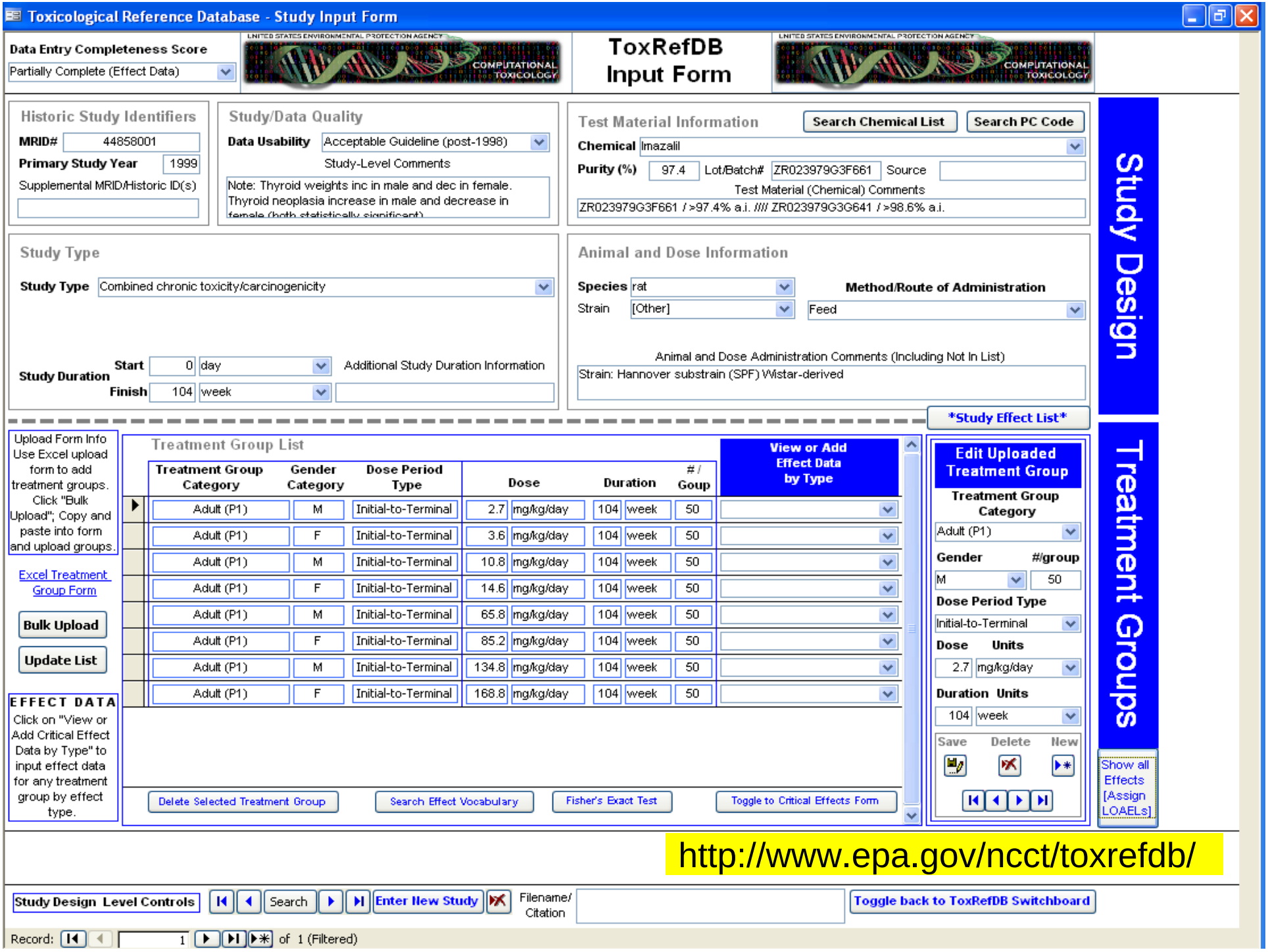
$10,000,000

DER Format

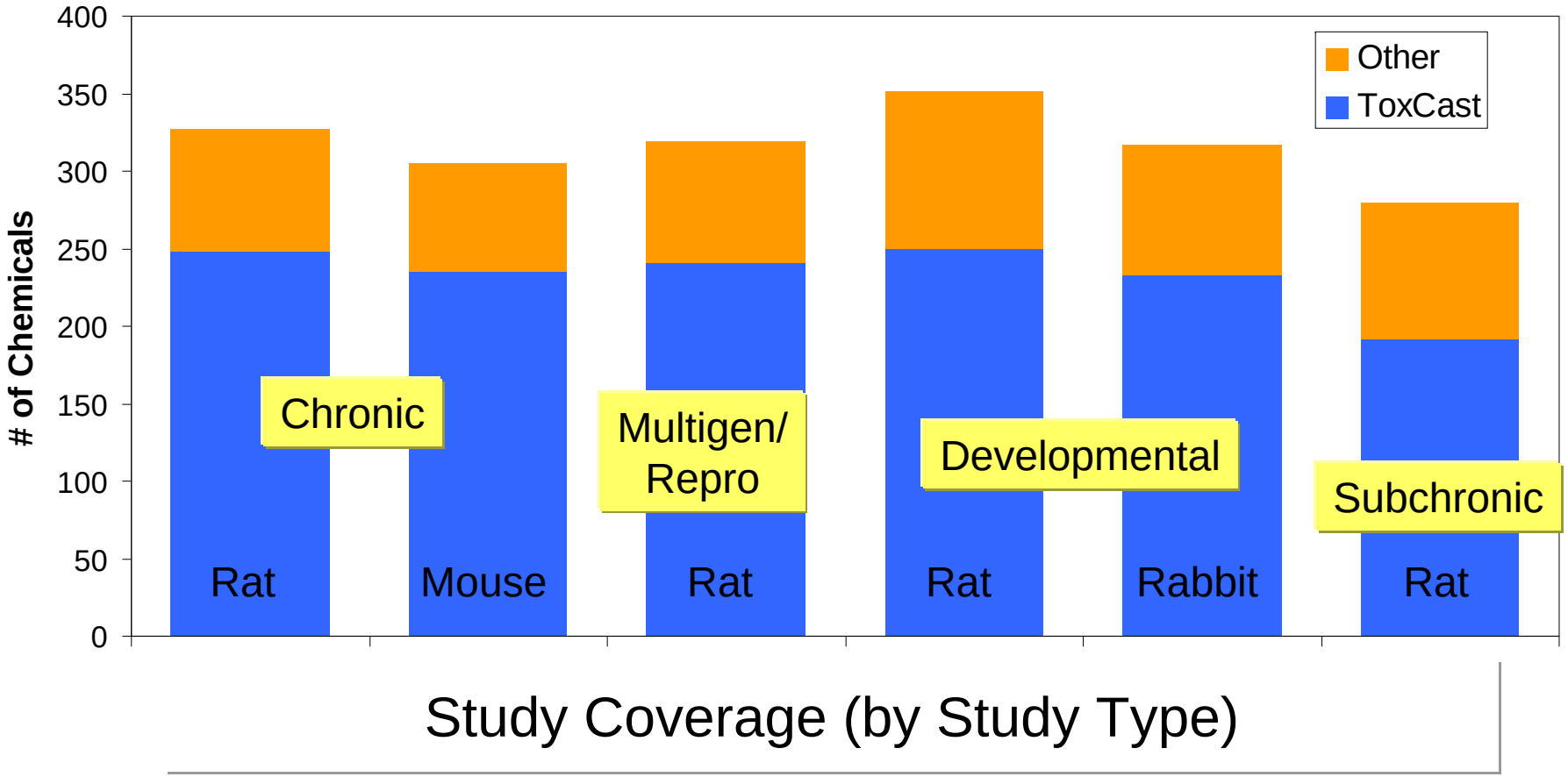
- Study Identifiers
  - Tested Chemical Information
    - IDs
    - Name
    - Purity
  - Study Type IDs
  - Reviewer Information
- Citation(s)
- Executive Summary
  - Summary Study Design
  - Summary Effects
  - Endpoints (NOAEL/LOAEL)

Chemical Properties

- Animal Information
  - Species
  - Strain
  - Husbandry
- Results (full dose-response)
  - Clinical signs
  - Body weight
  - Clinical Chemistry/ Hematology
  - Gross Pathology
  - Non-neoplastic Pathology
  - Neoplastic Pathology
  - Parental vs. Offspring
  - Maternal vs. Fetal

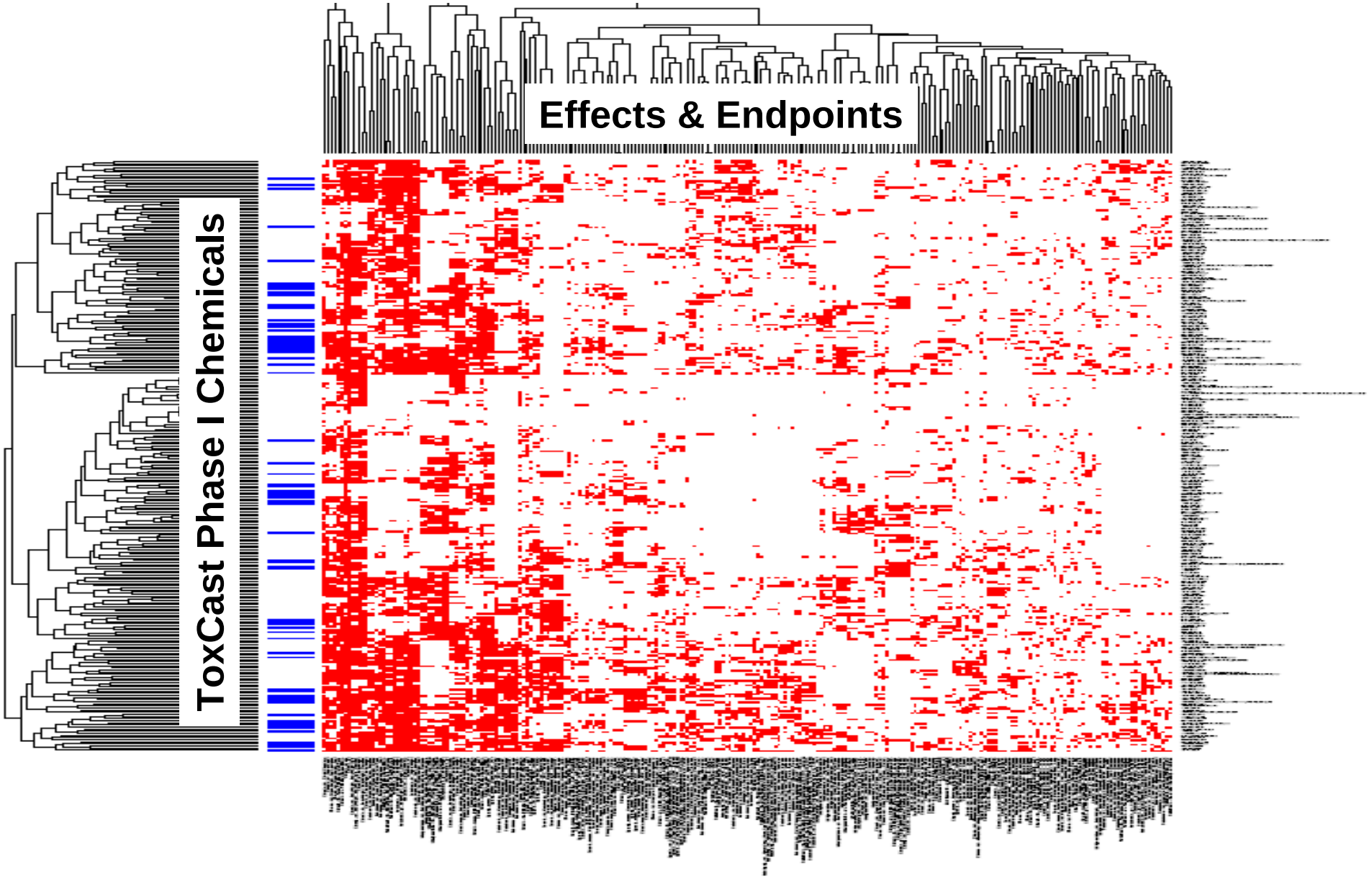

Toxicological Reference Database - Study Input Form
Data Entry Completeness Score
Partially Complete (Effect Data)
ToxRefDB Input Form
Historic Study Identifiers
MRID# 44858001
Primary Study Year 1999
Supplemental MRID/Historic ID(s)
Study/Data Quality
Data Usability Acceptable Guideline (post-1998)
Study-Level Comments
Note: Thyroid weights inc in male and dec in female. Thyroid neoplasia increase in male and decrease in female (both statistically significant)
Test Material Information
Search Chemical List
Search PC Code
Chemical Imazalil
Purity (%) 97.4
Lot/Batch# ZR023979G3F661
Source
Test Material (Chemical) Comments
ZR023979G3F661 / >97.4% a.i. /// ZR023979G3G641 / >98.6% a.i.
Study Type
Study Type Combined chronic toxicity/carcinogenicity
Study Duration Start 0 day
Finish 104 week
Additional Study Duration Information
Animal and Dose Information
Species rat
Strain [Other]
Method/Route of Administration
Feed
Animal and Dose Administration Comments (Including Not In List)
Strain: Hannover substrain (SPF) Wistar-derived
Study Effect List
Study Design
Treatment Groups
Upload Form Info
Use Excel upload form to add treatment groups. Click "Bulk Upload"; Copy and paste into form and upload groups.
Excel Treatment Group Form
Bulk Upload
Update List
EFFECT DATA
Click on "View or Add Critical Effect Data by Type" to input effect data for any treatment group by effect type.
Treatment Group List
Treatment Group Category | Gender Category | Dose Period Type | Dose | Duration | # / Goup | View or Add Effect Data by Type
Adult (P1) | M | Initial-to-Terminal | 2.7 mg/kg/day | 104 week | 50
Adult (P1) | F | Initial-to-Terminal | 3.6 mg/kg/day | 104 week | 50
Adult (P1) | M | Initial-to-Terminal | 10.8 mg/kg/day | 104 week | 50
Adult (P1) | F | Initial-to-Terminal | 14.6 mg/kg/day | 104 week | 50
Adult (P1) | M | Initial-to-Terminal | 65.8 mg/kg/day | 104 week | 50
Adult (P1) | F | Initial-to-Terminal | 85.2 mg/kg/day | 104 week | 50
Adult (P1) | M | Initial-to-Terminal | 134.8 mg/kg/day | 104 week | 50
Adult (P1) | F | Initial-to-Terminal | 168.8 mg/kg/day | 104 week | 50
Delete Selected Treatment Group
Search Effect Vocabulary
Fisher's Exact Test
Toggle to Critical Effects Form
Edit Uploaded Treatment Group
Treatment Group Category
Adult (P1)
Gender #/group
M 50
Dose Period Type
Initial-to-Terminal
Dose Units
2.7 mg/kg/day
Duration Units
104 week
Save Delete New
Show all Effects [Assign LOAELs]
http://www.epa.gov/ncct/toxrefdb/
Study Design Level Controls
Search
Enter New Study
Filename/ Citation
Toggle back to ToxRefDB Switchboard
Record: 1 of 1 (Filtered)

ToxRefDB: 2073 Studies Entered for 480 Chemicals

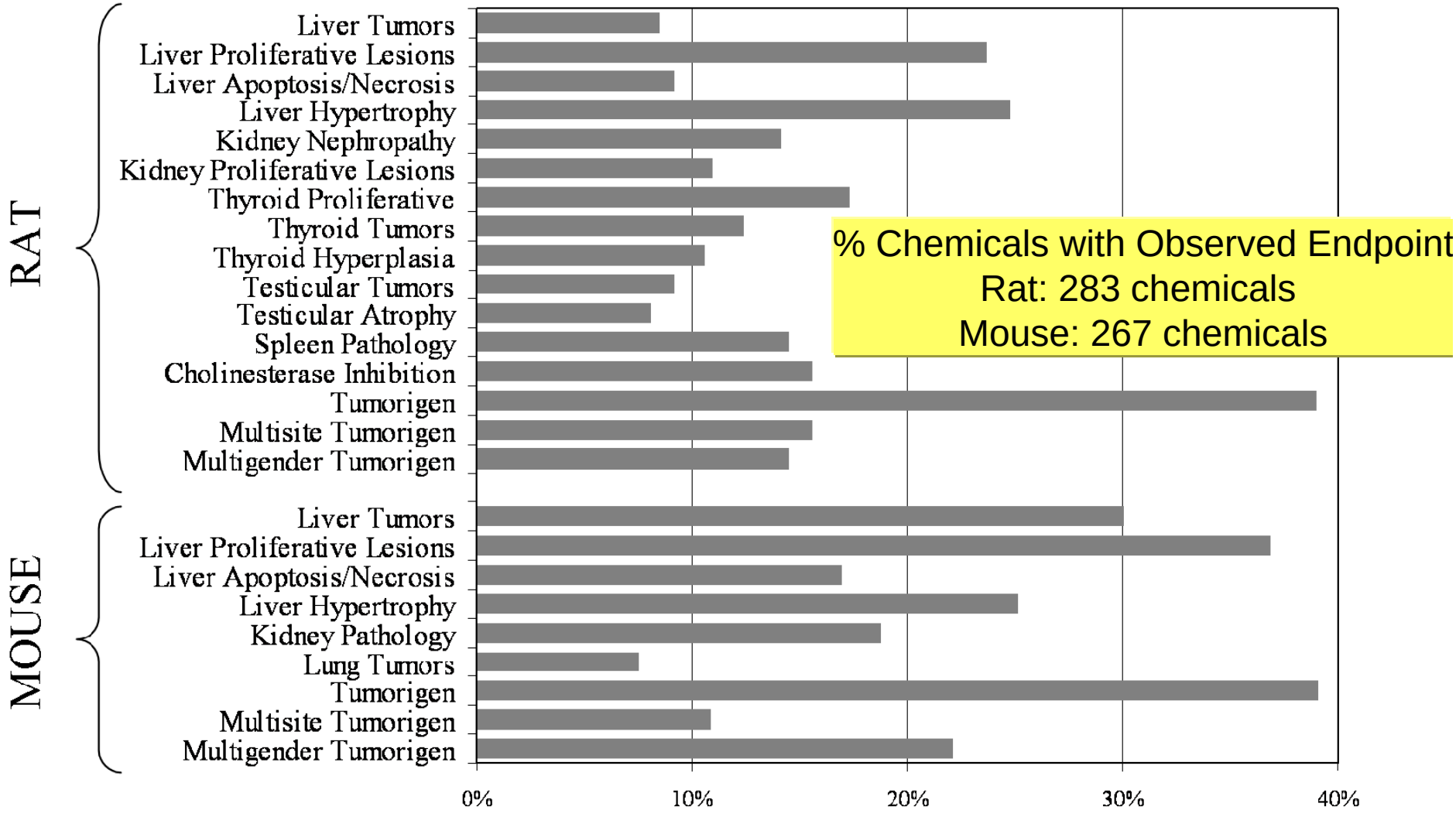


>$1Billion Dollars Worth of *In Vivo* Chronic/Cancer Bioassay Effects and Endpoints

United States Environmental Protection Agency



Initial Chronic Rat & Mouse Endpoints for Predictive Modeling



% Chemicals with Observed Endpoint
Rat: 283 chemicals
Mouse: 267 chemicals

Martin et al. (2009) Environ Health Perspec 117:392-399

ToxCast *In vitro* HTS Assays

United States Environmental Protection Agency

Biochemical Assays

- **Protein families**
  - GPCR
  - NR
  - Kinase
  - Phosphatase
  - Protease
  - Other enzyme
  - Ion channel
  - Transporter
- **Assay formats**
  - Radioligand binding
  - Enzyme activity
  - Co-activator recruitment

467 Endpoints

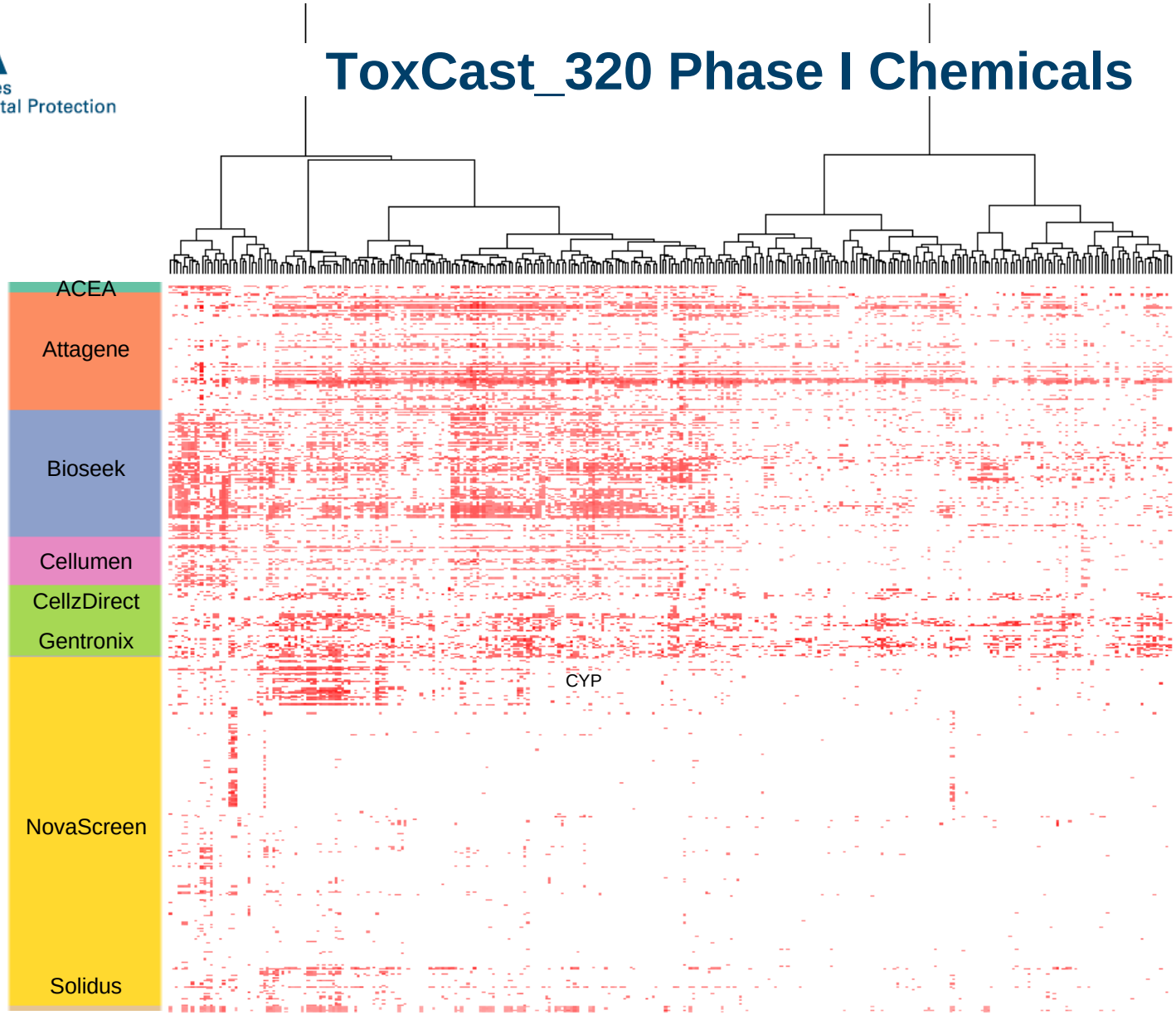
Cellular Assays

- **Cell lines**
  - HepG2 human hepatoblastoma
  - A549 human lung carcinoma
  - HEK 293 human embryonic kidney
- **Primary cells**
  - Human endothelial cells
  - Human monocytes
  - Human keratinocytes
  - Human fibroblasts
  - Human proximal tubule kidney cells
  - Human small airway epithelial cells
- **Biotransformation competent cells**
  - Primary rat hepatocytes
  - Primary human hepatocytes
- **Assay formats**
  - Cytotoxicity
  - Reporter gene
  - Gene expression
  - Biomarker production
  - High-content imaging for cellular phenotype

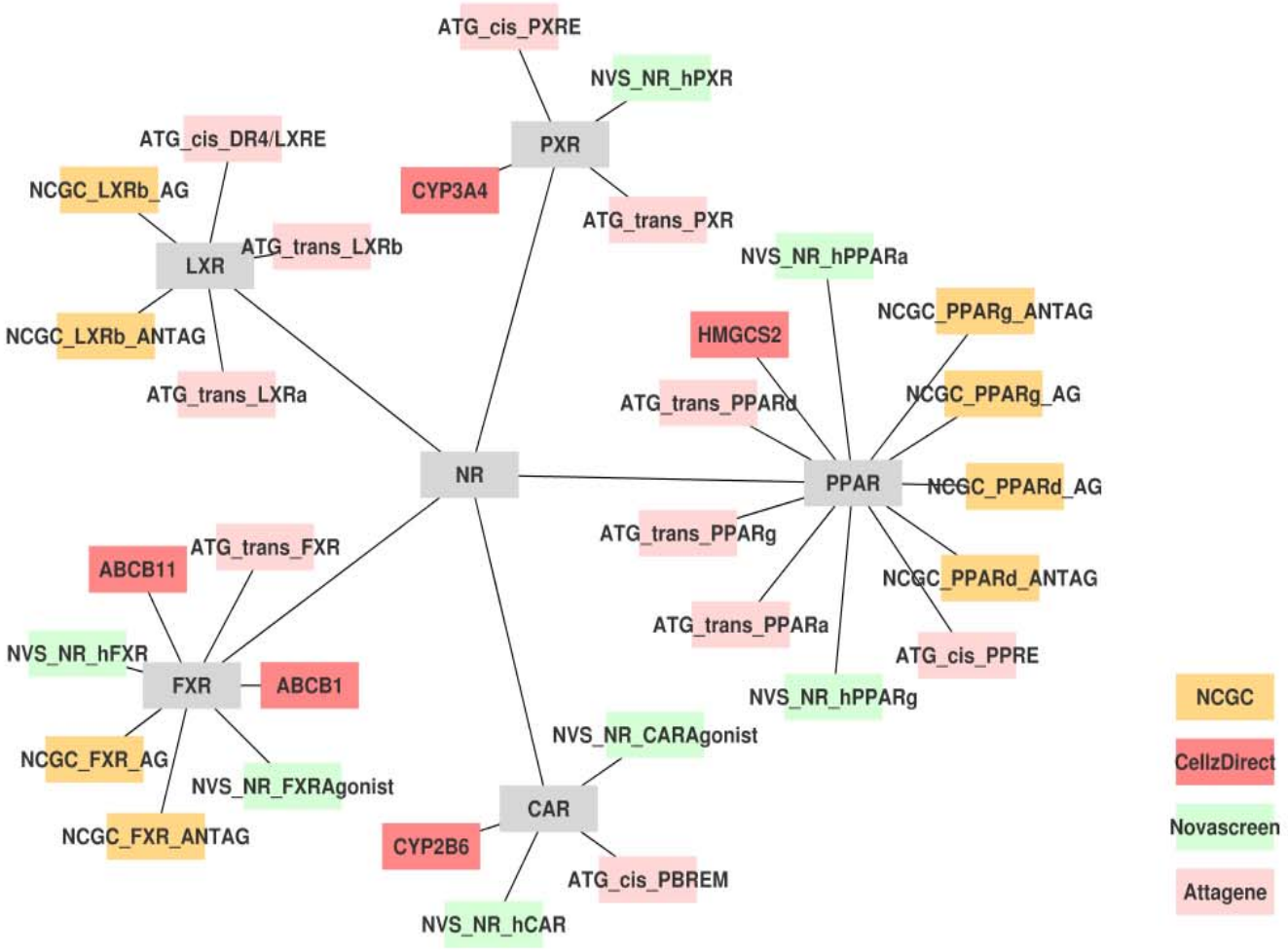
United States Environmental Protection Agency

ToxCast_320 Phase I Chemicals

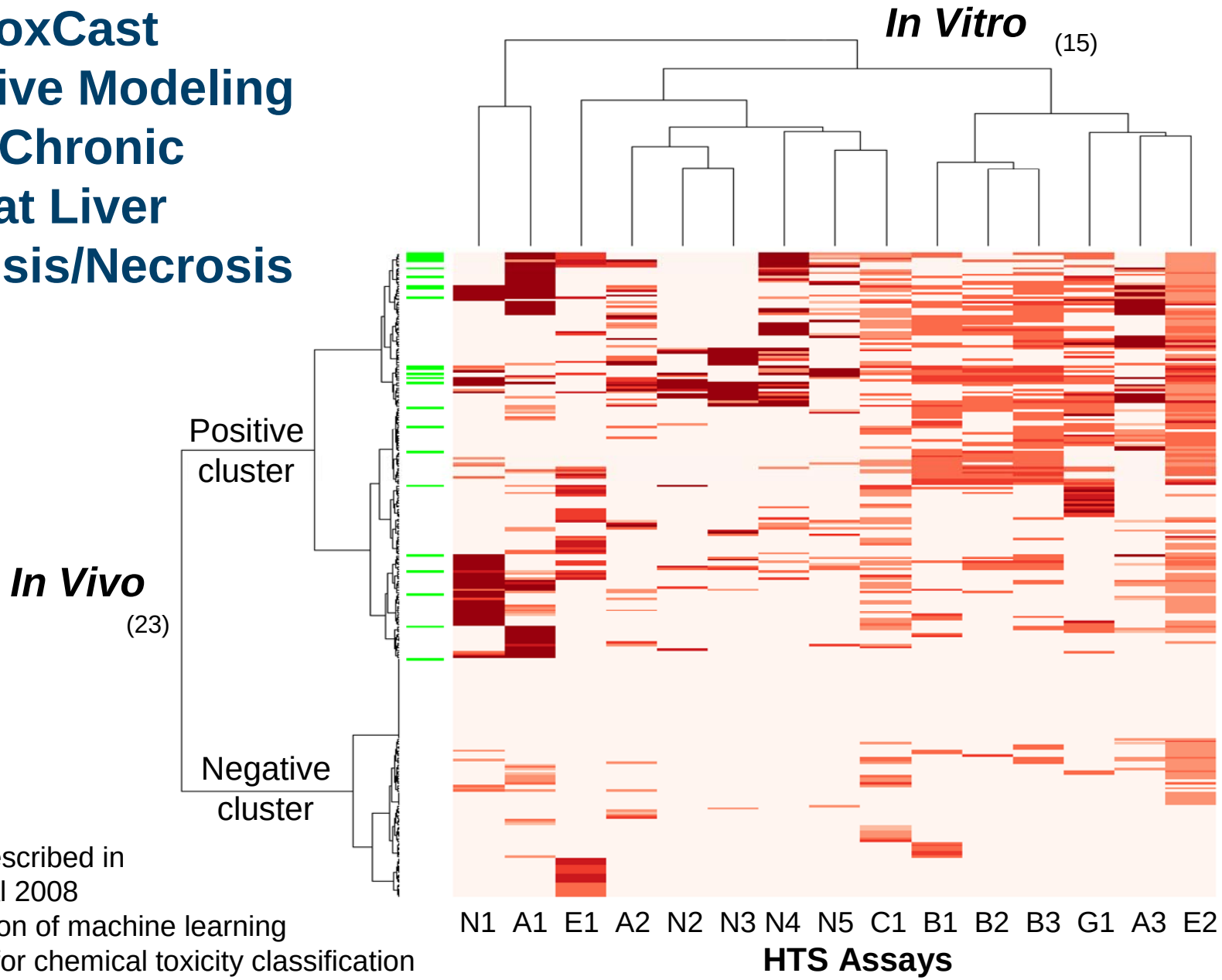
ToxCast Phase I HTS (467 assays)



Multiple Assays per Endpoint



ToxCast Predictive Modeling of Chronic Rat Liver Apoptosis/Necrosis



Methods described in
Judson et al 2008
A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model.
BMC Bioinformatics 9:241

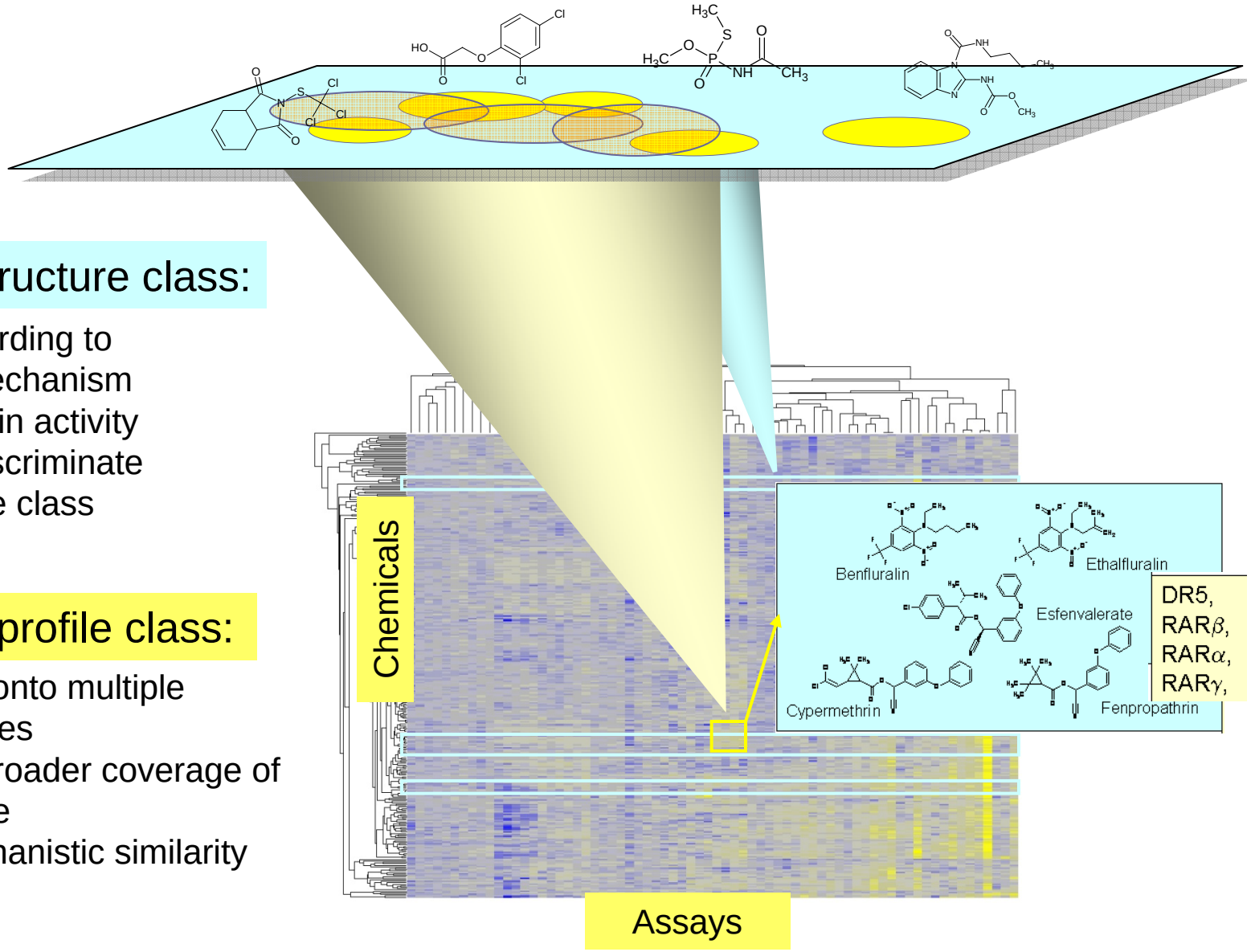
Structure Class vs Bioactivity Class

Chemical structure class:

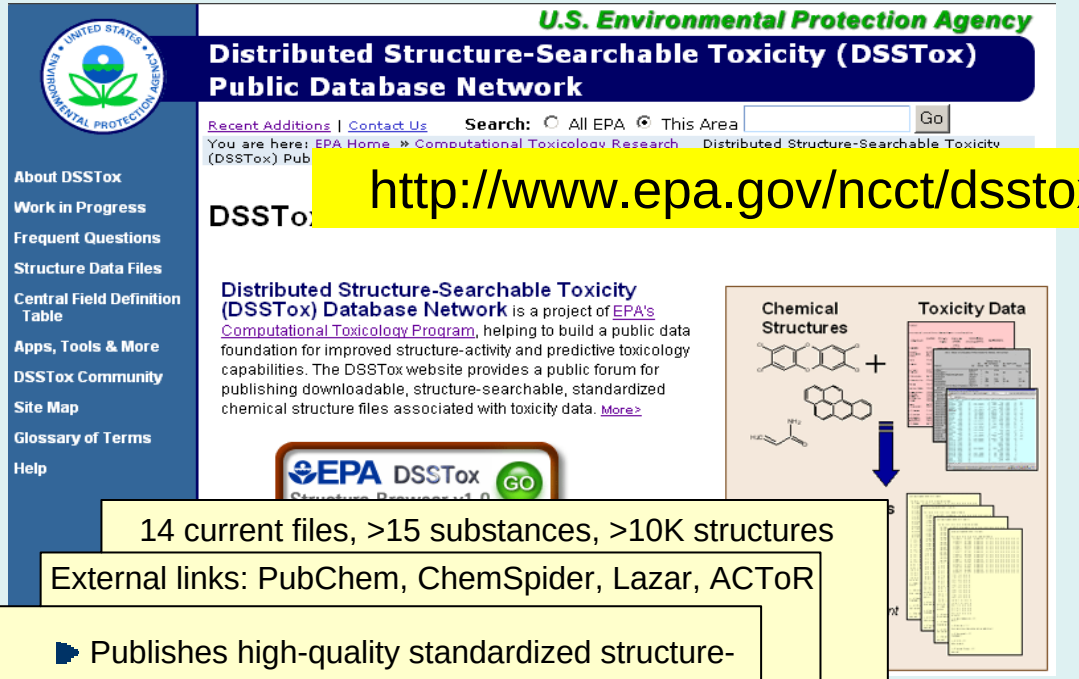
- Cluster according to activity and mechanism
- Differences in activity profiles can discriminate within structure class

Bioactivity profile class:

- Can project onto multiple chemical classes
- Potentially broader coverage of chemical space
- Implies mechanistic similarity



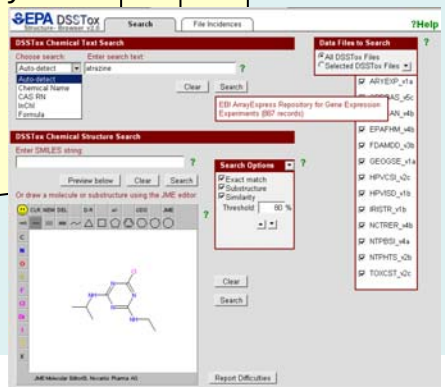
DSSTox: Distributed Structure-Searchable Database Network Project…& ToxCast



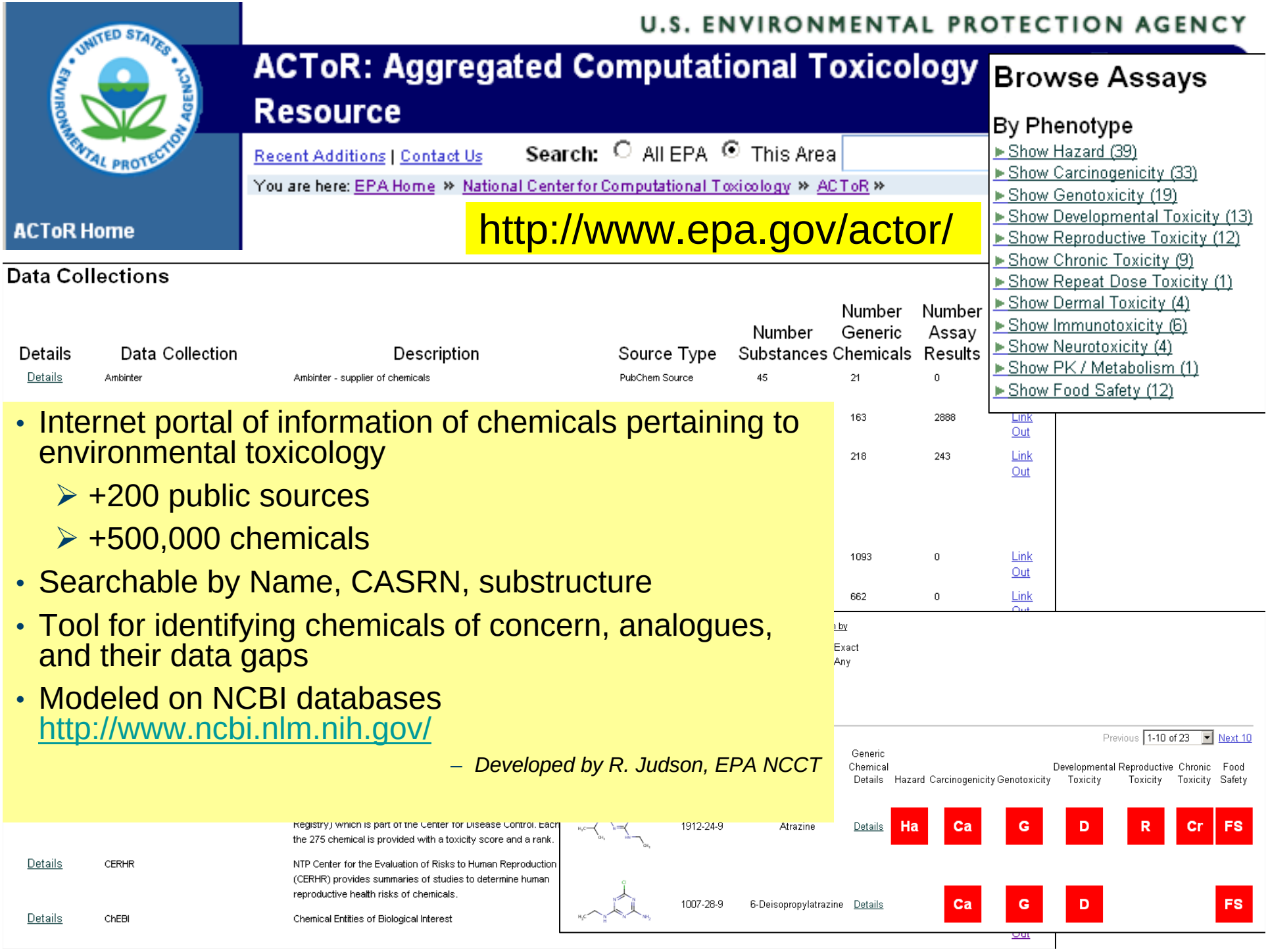
14 current files, >15 substances, >10K structures

External links: PubChem, ChemSpider, Lazar, ACToR

- Publishes high-quality standardized structure-data (SD) files pertaining to toxicology:
  - *EPA, HPV-IS, IRIS, NTP, FDA, NCBI, EBI, …*
  - Strict chemical QA procedures
- Public substance/structure ID registry system
  - DSSTox Structure-Browser
    - *Search by structure, name, CAS,*
    - *Similarity & fragment search*
    - *Search within a DSSTox file*
    - *Link-in and link out of browser*



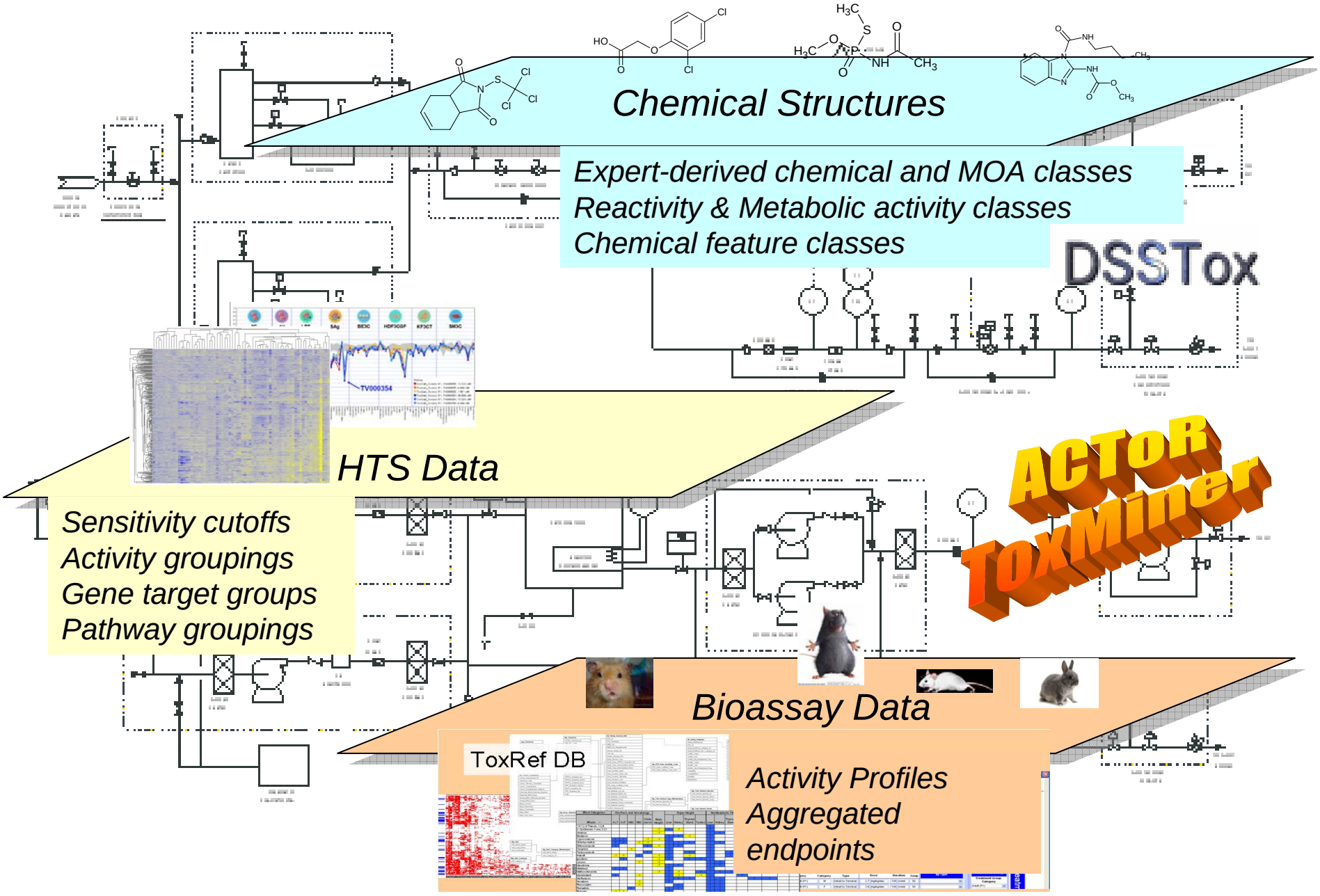
- Chemical information QA and structure annotation
  - *ToxCast Phase I & II*
  - *ToxRefDB*
  - *Tox21*
  - *ACToR*
- Facilitate external linkages and data publication
- Publish summary activities and chemical classifiers for modeling
- PubChem Source depositor for structures & “assays”



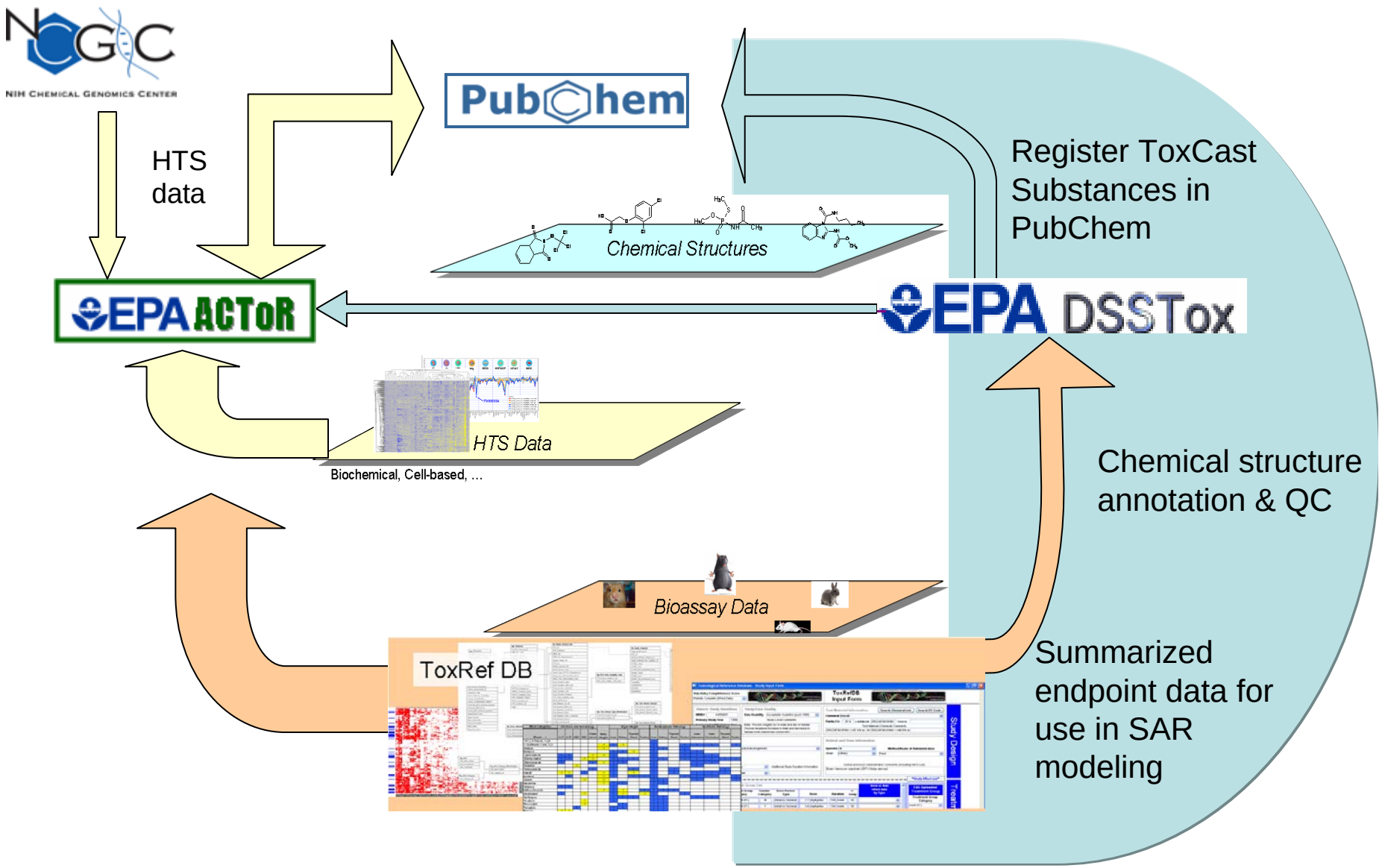
- Internet portal of information of chemicals pertaining to environmental toxicology
  - ➢ +200 public sources
  - ➢ +500,000 chemicals
- Searchable by Name, CASRN, substructure
- Tool for identifying chemicals of concern, analogues, and their data gaps
- Modeled on NCBI databases http://www.ncbi.nlm.nih.gov/

– Developed by R. Judson, EPA NCCT

ToxCast: High-Multi-Dimensional Data

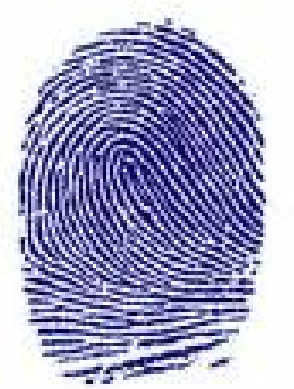


ToxCast: Data Publication & Exploration

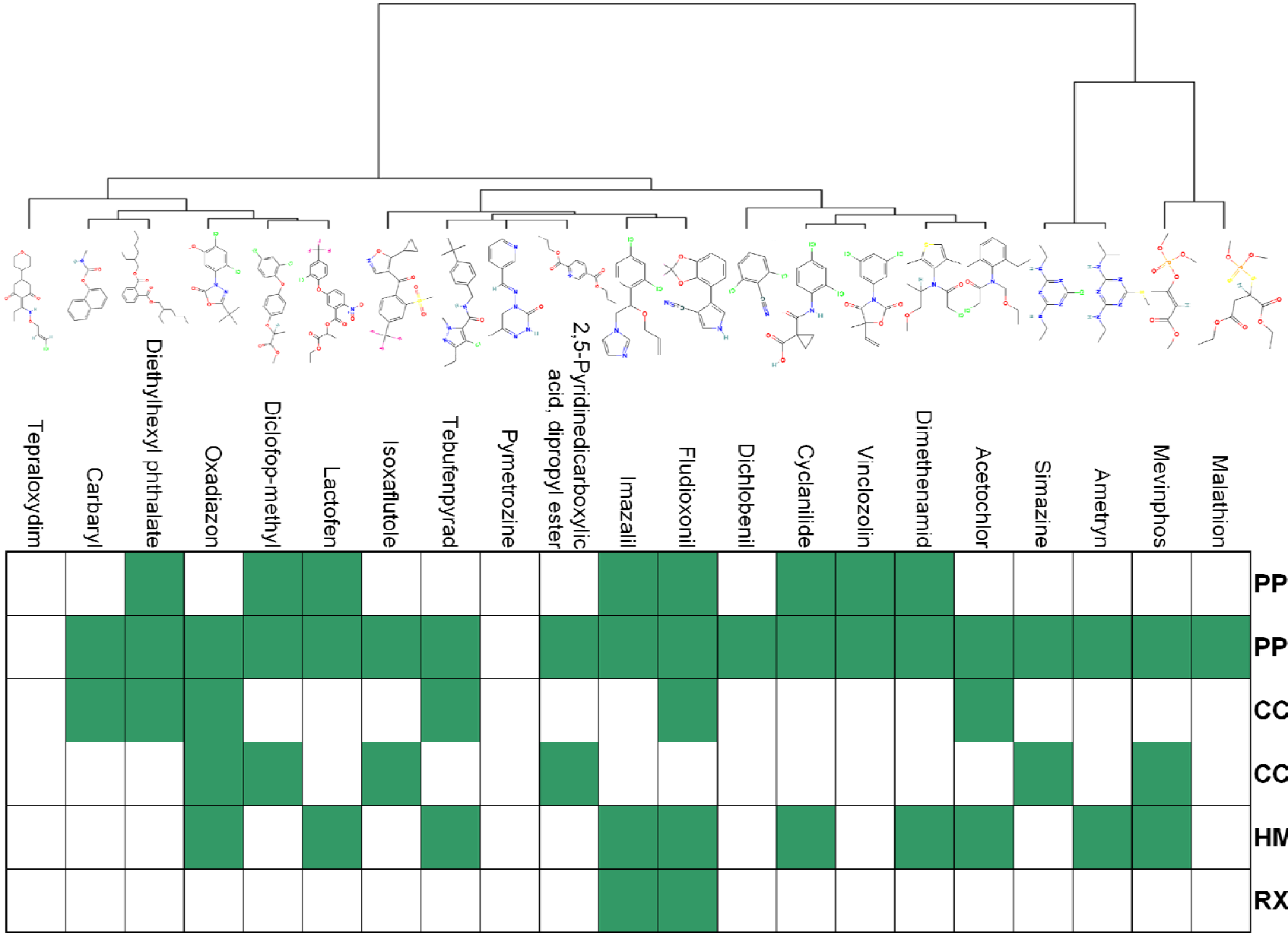


ToxCast™ Data Analysis Summit, May 14-15, 2009

- Phase 1 ToxCast data made available to analysis partners prior to full public release
  - >500 HTS assays – categorical (1/0)
  - 76 “bioassay” endpoints from ToxRefDB for modeling
  - Chemical structure SD file (DSSTox), chemical information files (descriptors)
- Over 200 registered attendees, 60 presenters
- Wide variety of prediction schemes
  - *In vitro* → *In vivo*
  - Chemical descriptors → *In vivo* (SAR)
  - Chemical descriptors + *In vitro* → *In vivo*
- Wide variety of approaches
  - Statistics, clustering, machine learning, particle swarm, etc.



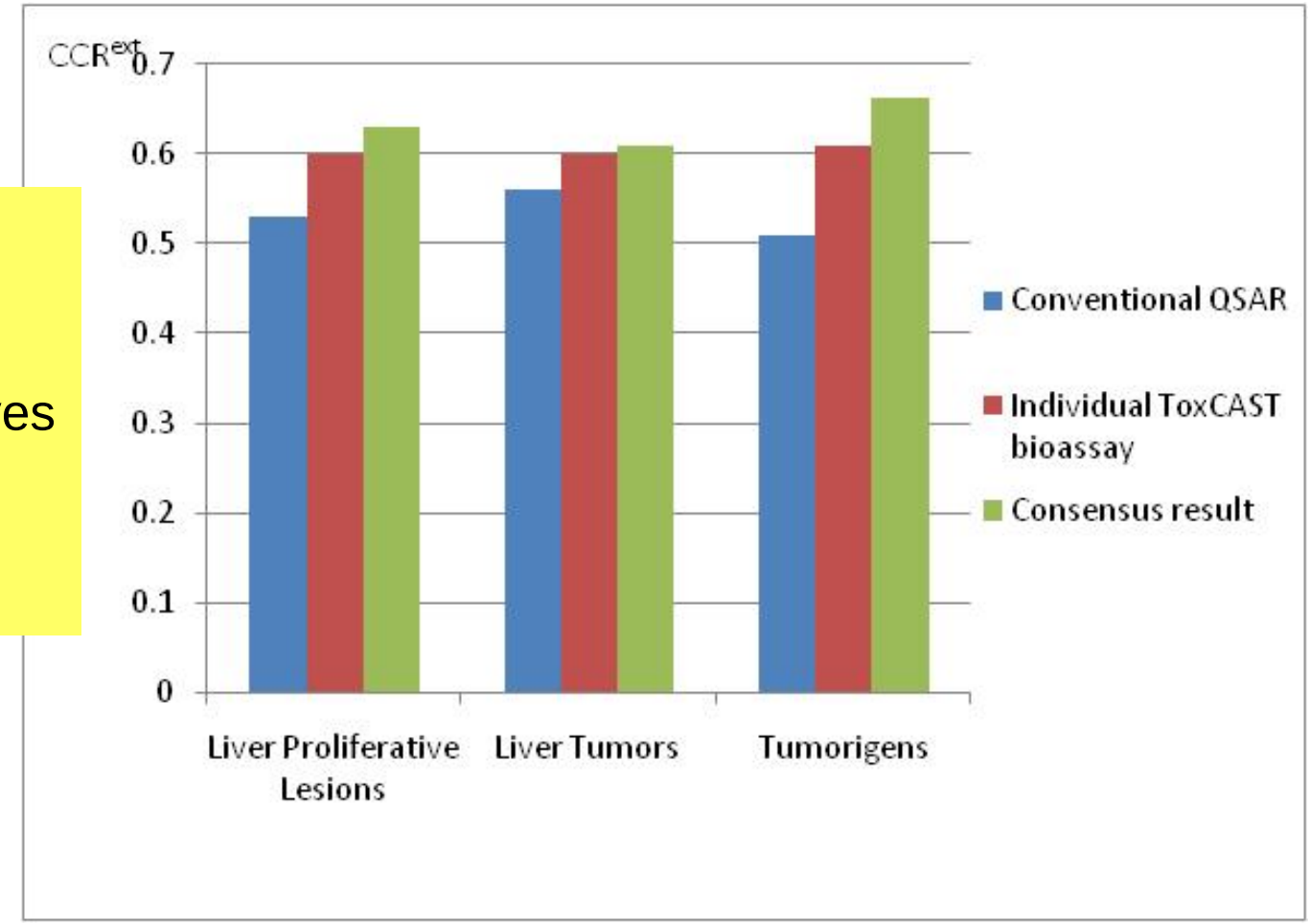
Rat Liver Tumorigens are diverse in chemical structure and *in vitro* Signature



Prediction of animal toxicity endpoints of ToxCast Phase I compounds using a combination of chemical and biological in vitro descriptors

✓Need for external validation

✓ Predictivity improves when HTS data combined with chemical descriptors

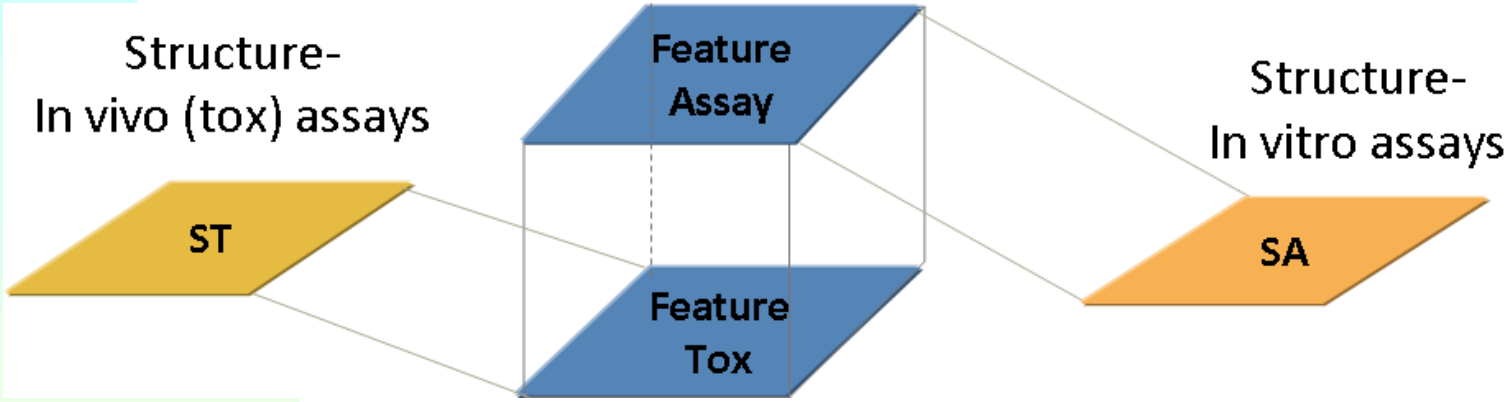


Slide results courtesy of Alex Tropsha, UNC School of Pharmacy, ToxCast Data Analysis Summit, May 2009

UNC ESHELMAN SCHOOL OF PHARMACY

Navigating through the domains of biology and chemistry

Features In vivo → Features In vitro

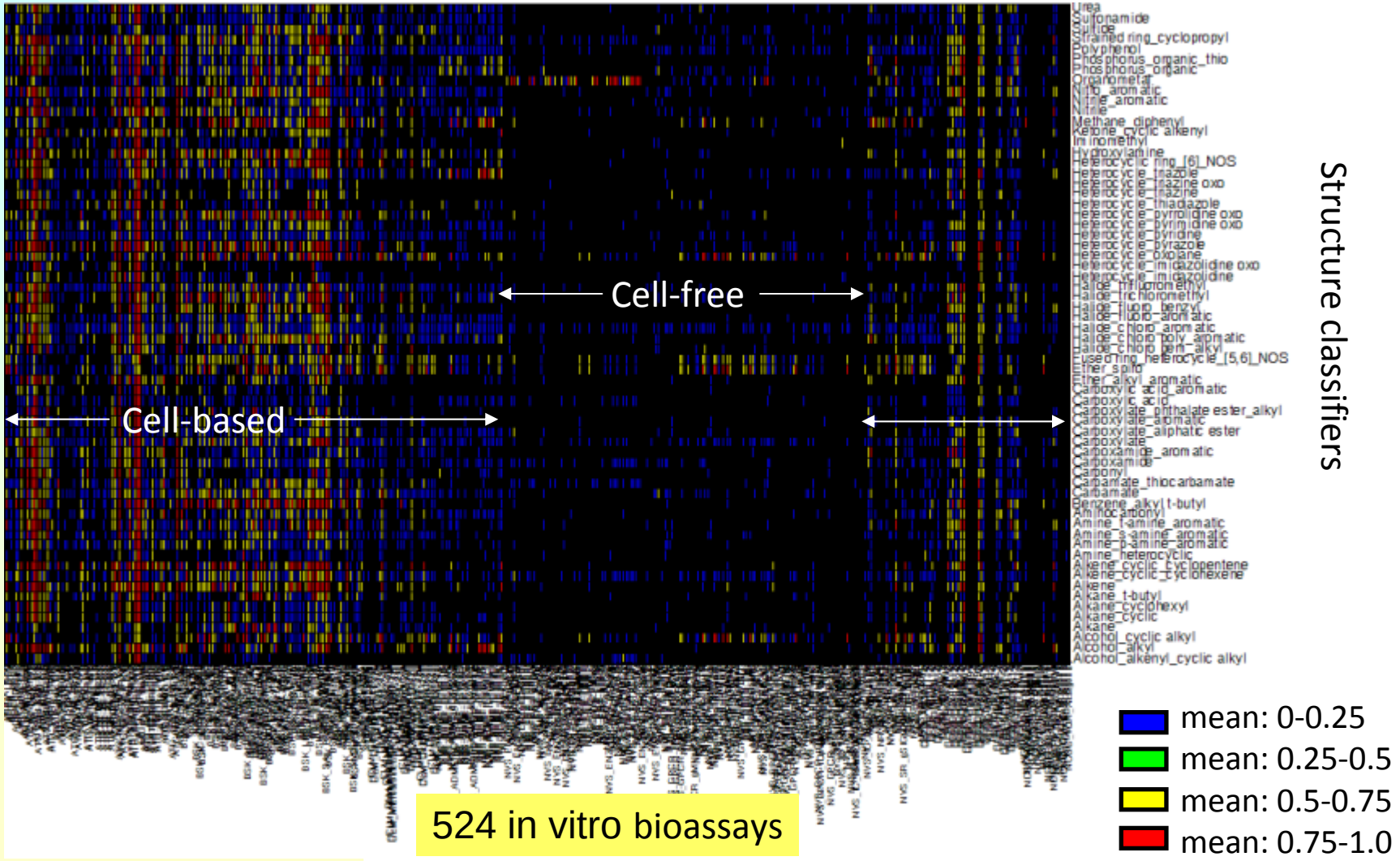


Chemical features (i.e., chemical substructures):

- ✓ greater representation than compounds across datasets
- ✓ more statistically robust in associating with activity
- ✓ intuitive and chemically meaningful

Slide results courtesy of ChihaeYang, FDA CFSAN, ToxCast Data Analysis Summit, May 2009

All in vitro assays against selected “Structural Classifiers”



Slide results courtesy of ChihaeYang, FDA CFSAN, ToxCast Data Analysis Summit, May 2009

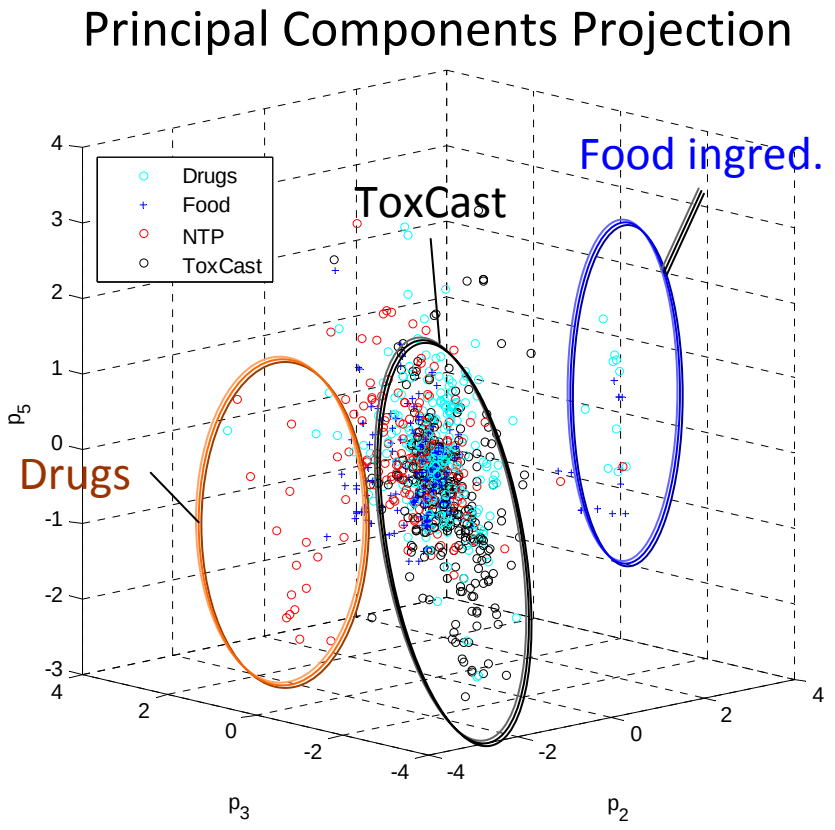
ToxCast™ Data Analysis Summit, May 14-15, 2009

Impressions, Conclusions, Lessons…

- Global associations (*in vitro* to *in vivo*), trends not apparent → *local models possible in cross sections of chemical feature/biology space*
- Statistical means for dealing with highly dimensional, sparse, unbalanced data needed → *new methods proposed*
- Chemical descriptors and features improve model performance when combined with HTS
- Existing SAR carcinogenicity prediction models (LAZAR, ToxTree, PASS) built from public data performed poorly → *point to lack of coverage of non-genotoxic mechanisms*
- Public data availability and transparency successful in engaging wide range of researchers and capabilities in analysis

ToxCast Data Landscape: Implications for genetox

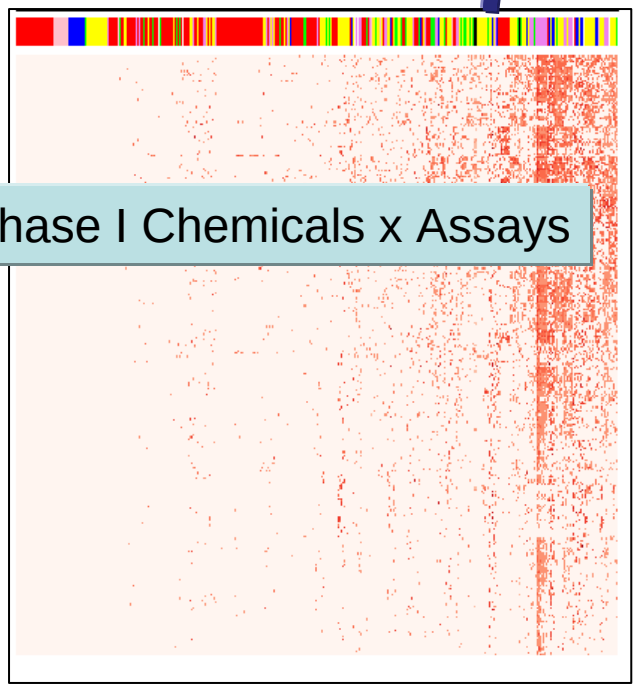
- 309 ToxCast Phase I chemicals
  - Mostly pesticides
  - Different chemical space
- High proportion of non-genotoxic carcinogens
- No genetox data provided
- SAL data collected from public sources for approx 108 ToxCast chemicals to assess genetox-related assays *



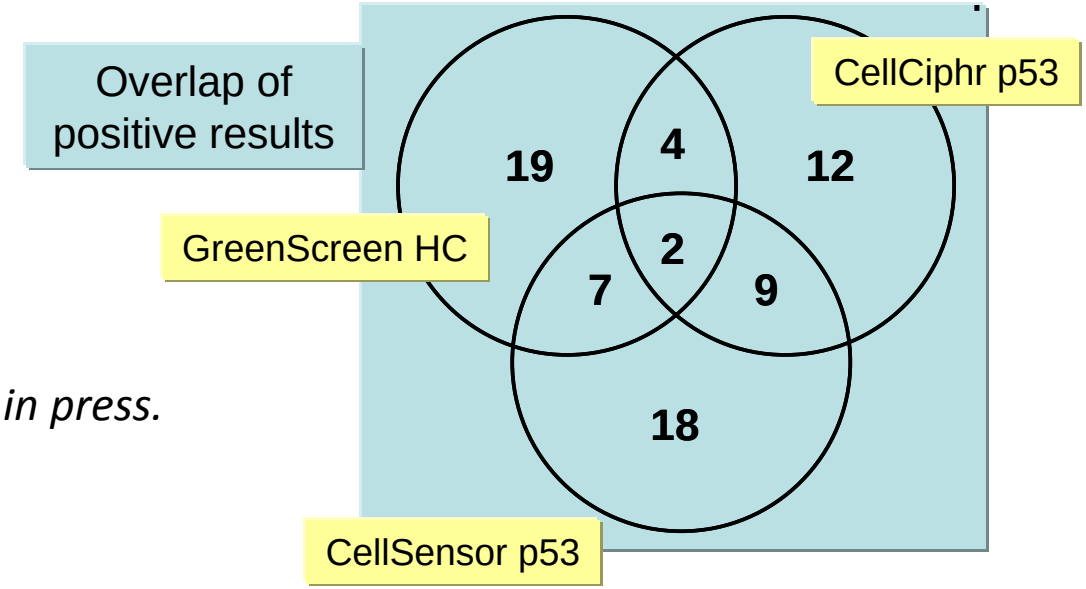
Knight et. al, Reg Tox Pharmacol, 2009, in press.

Slide results courtesy of ChihaeYang, FDA CFSAN, ToxCast Data Analysis Summit, May 2009

ToxCast/Tox21: GeneTox

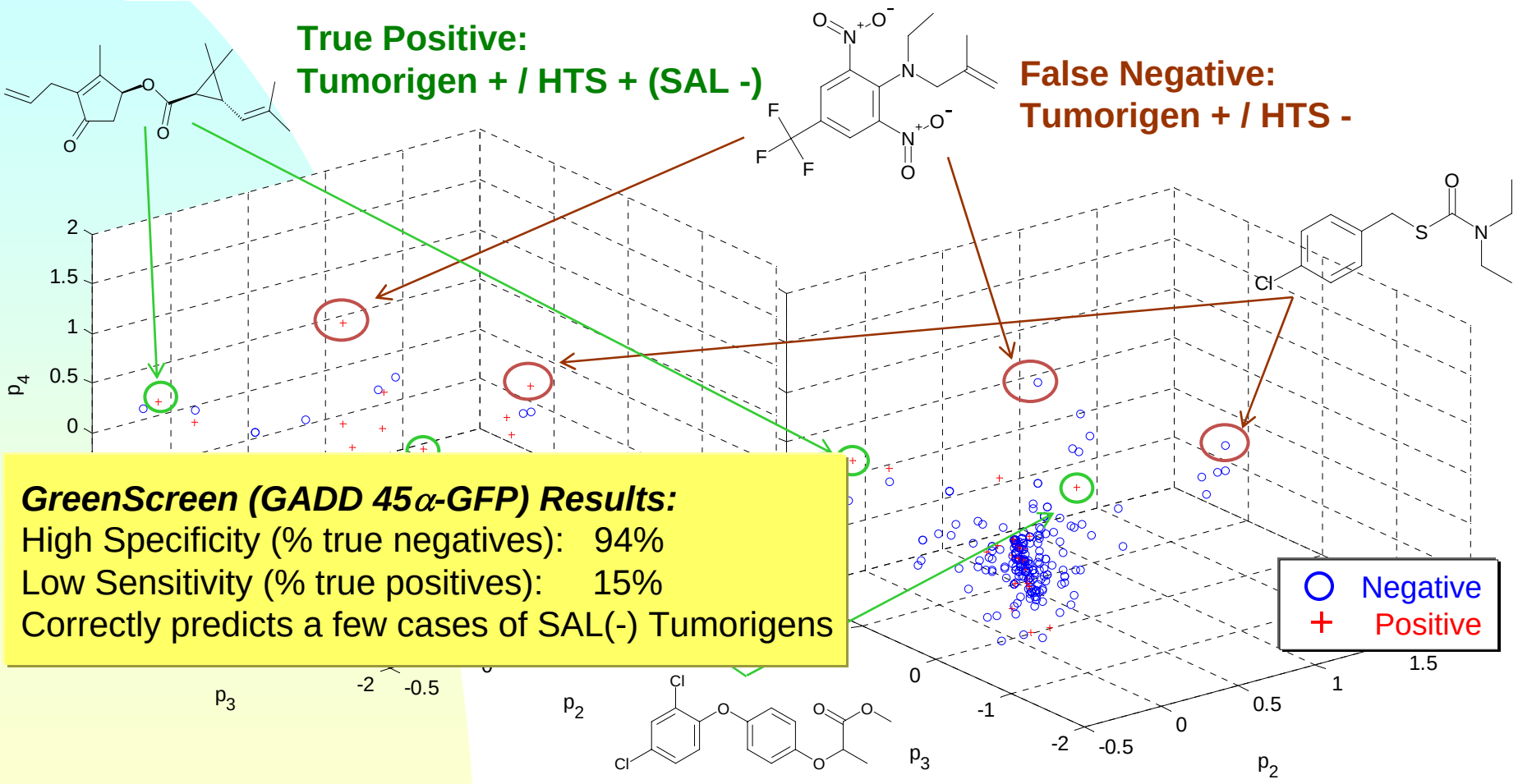


- GreenScreen HC GADD45α-GFP Reporter Assay (p53 competent) – (*Gentronix, Ltd.)*
- CellCiphr p53 *(Cellumen Inc.)*
- CellSensor p53RE-bla *(Invitrogen Corp., provided by NCGC)*



Knight et. al, Reg Tox Pharmacol, 2009, in press.

A rodent bioassay vs. an HTS genetox assay



Knight et. al, Reg Tox Pharma, 2009, in press.

Slide results courtesy of ChihaeYang, FDA CFSAN, ToxCast Data Analysis Summit, May 2009

ToxCast & Tox21

Chemical Sample Annotation & QC

Generic Chemical Annotation

ToxCast: Phase I

ToxCast: Phase II

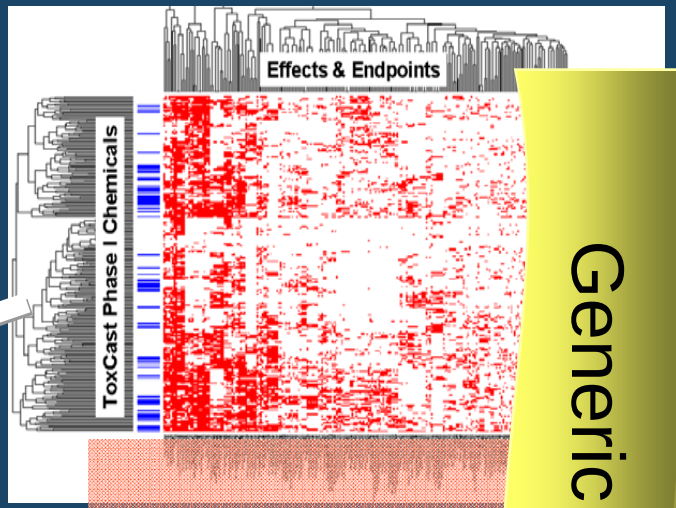
Tox21

EPA

NTP

NCGC

320 chemicals x >450 HTS assays



Methods development Preliminary signatures

~1000 chemicals >400 HTS assays

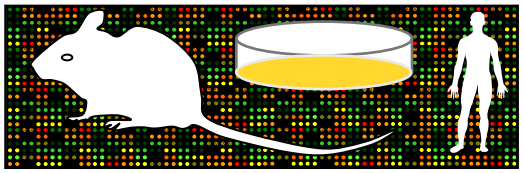
+ Public Tox Data

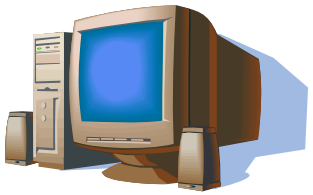
Refined predictive signatures

~10,000 chemicals, 100s -1000s of HTS assays

Sparse Public Tox Data

Tox21 Collaboration





National Health and

National Center for

- Combined HTS plates (7x1408) high interest chemicals
- Joint assay development
- Use of NCGC HTS testing capabilities
- EPA informatics (ACToR/DSSTox)

National Toxicology Program
Department of Health and Human Services

NIH Chemical Genomics Center

Biomolecular Screening Branch

Toxicology Project Team

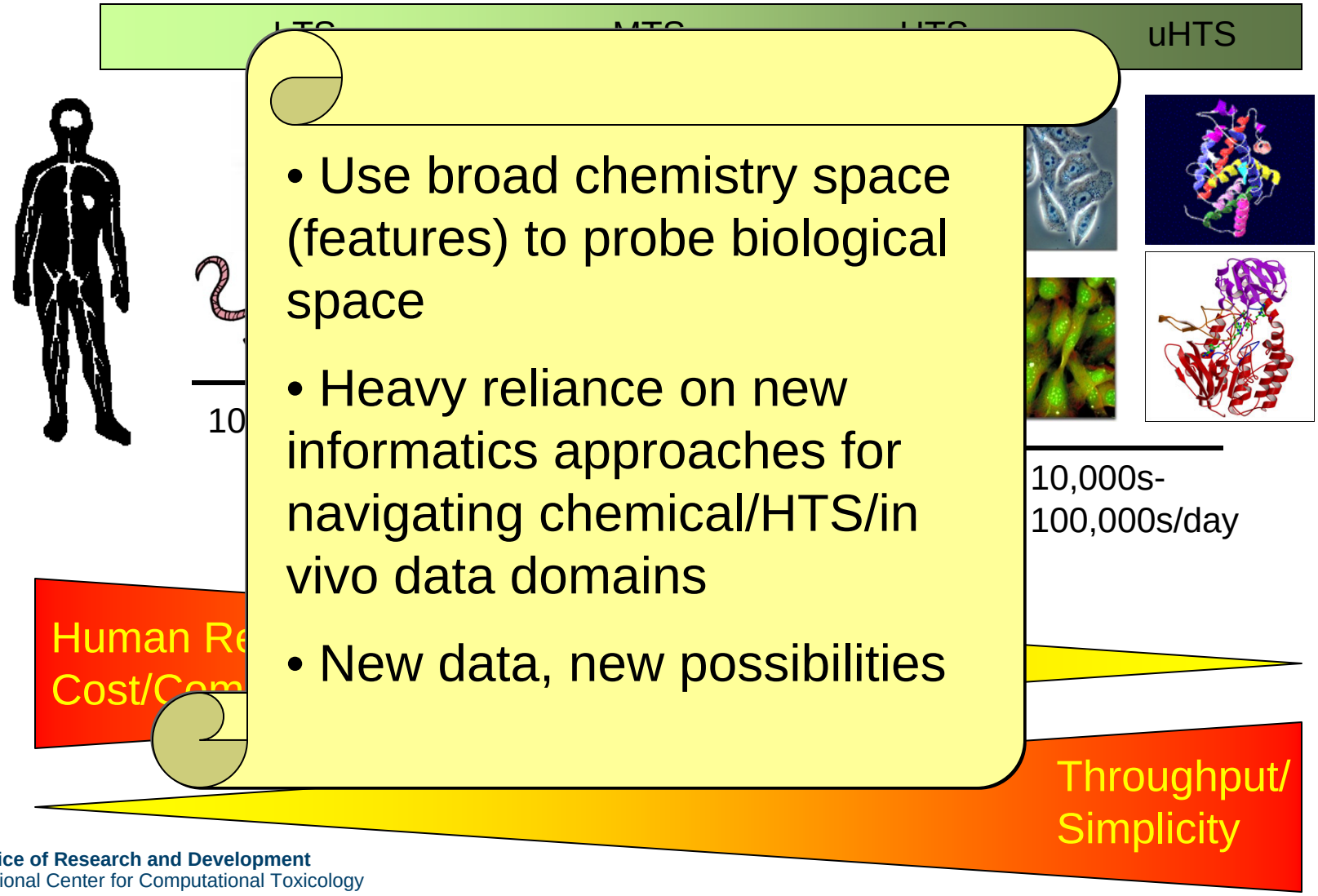
Tox21 & ToxCast Phase II

- HTS chemical space will extend to approx 10,000 chemicals:
  - *EPA Pesticides, high interest EPA and stakeholder inventories, data rich chemicals*
  - *NTP data rich & high interest chemicals, DSSTox inventories*
  - *NCGC - marketed drugs*
  - *EPA HPV classes, metabolite/parent pairs, Green chemistry pairs (toxic, safe)*
- Pfizer: ~ 100 failed drugs with pre-clinical/clinical tox data
- Glaxo: liver toxicity data for approx 150 drugs
- L'Oreal: sponsoring 10 chemicals for Phase II
- FDA CFSAN data rich chemicals to be included
- Model organisms: *c. elegans* (NTP), whole embryo zebrafish (EPA)
- Expanded toxicity data models and databases to include:
  - *Developmental Neurotox, Immunotox, Genetox*

ToxCast/Tox21: GeneTox

- Available HTS GeneTox tests under consideration:
  - GreenScreen Human Cell Assay (GADD45α –GFP reporter)
  - Ames II assay
  - *In vitro* Micronucleus assay
  - *In vitro* Comet assay
  - *In vitro* Caspase 3/7 Cytotoxicity assay
- ToxRefDB Genetox data from DERs for pesticides
- Inclusion of hundreds of additional chemicals having rich profile of both in vivo chronic (cancer) data & genetox data
  - NTP chronic bioassay studies
  - FDA CFSAN chronic studies

New Approaches to Toxicity Screening



Acknowledgements:

- **EPA NCCT DSSTox Team:**
  Maritja Wolf (DSSTox) and Tom Transue (Structure-browser) – Lockheed Martin, Contractors to the US EPA
- **EPA NCCT ToxCast Team:**
  Robert Kavlock (Director, NCCT)
  David Dix (ToxRefDB, HTS, Genomics)
  Keith Houck (HTS)
  Matt Martin (ToxRefDB)
  Richard Judson (ACToR, ToxMiner)
  Thomas Knudsen (ToxRefDB, v-Embryo)
  David Reif (ToxMiner)
  Stephen Little (Genetox)
- **External Collaborators:**
  Chihae Yang, FDA/CFSAN
  Alex Tropsha, UNC-Chapel Hill
  Andrew Knight & colleagues, Gentronix Ltd.
  Chris Austin & colleagues, NCGC

This work was reviewed by EPA and approved for publication but does not necessarily reflect official Agency policy.