

Recent Developments in Toxico-Cheminformatics: *Supporting a new paradigm for predictive toxicology*

8-9 December 2008

Lhasa Limited: New Horizons in Toxicity Prediction

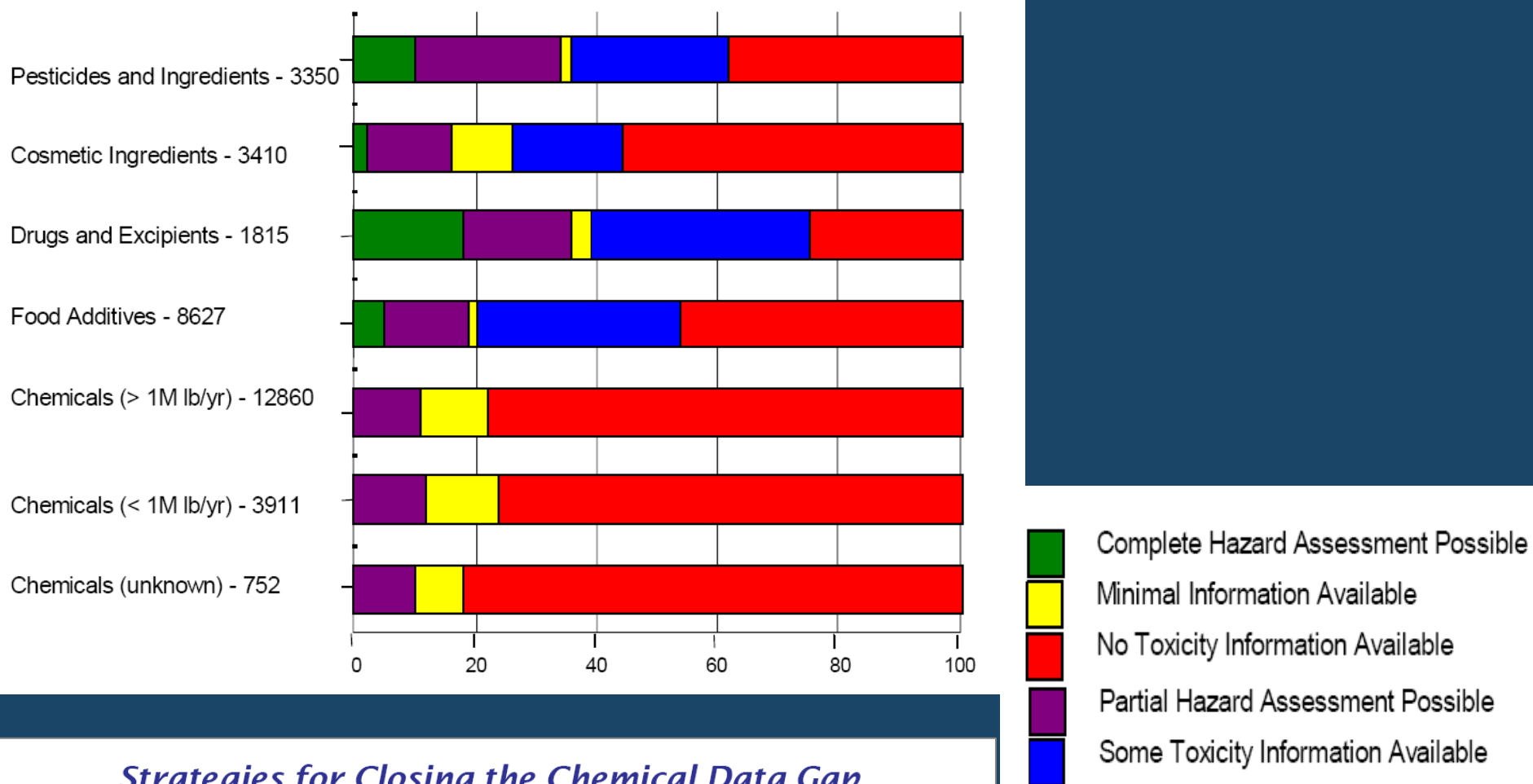
UNITED STATES ENVIRONMENTAL PROTECTION AGENCY



Ann Richard
richard.ann@epa.gov

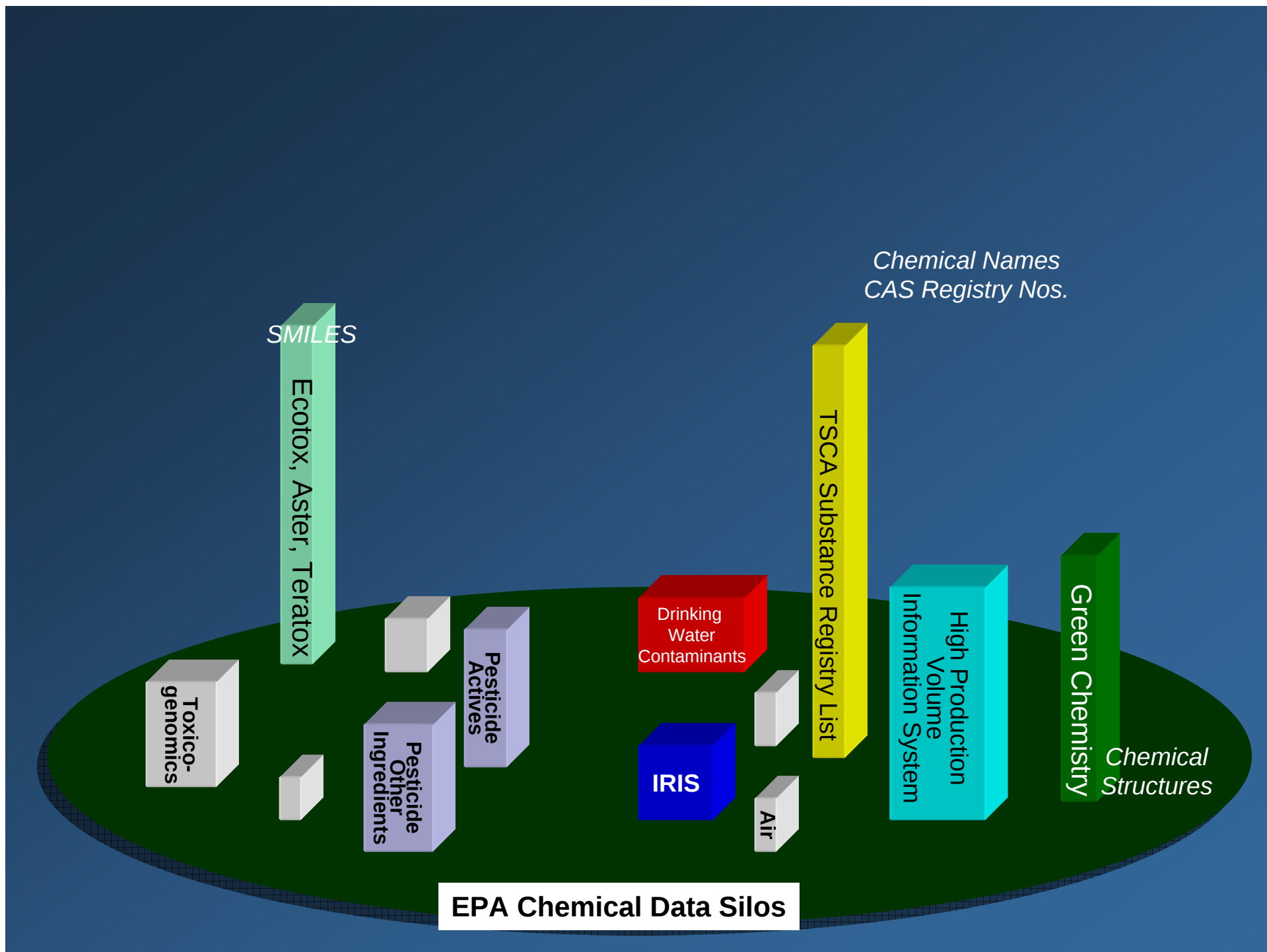
Environmental Chemicals: Toxicity Assessment Data Gaps

Estimated Mean Percent in Selected Universe



Strategies for Closing the Chemical Data Gap

by John S. Applegate and Katherine Baer



Chemical Research in Toxicology

JULY 2007
VOLUME 20, NUMBER 7

© Copyright 2007 by the American Chemical Society

Editorial

Aaron T. Jacobs and Lawrence J. Marnett*

*A.B. Hancock Jr. Memorial Laboratory for Cancer Research
Departments of Biochemistry, Chemistry, and Pharmacology
Vanderbilt Institute of Chemical Biology
Center in Molecular Toxicology
Vanderbilt-Ingram Cancer Center
Vanderbilt University School of Medicine
Nashville, Tennessee 37232*

TX7001564

The Future of Toxicology—Wrap Up

Wither toxicology? We have enjoyed a series of informative, occasionally provocative, commentaries on this subject from

a paucity of compelling problems to work on that would generate a broad mandate for large-scale investment in toxicol-

“A major focus for the future of computational toxicology will be integration and analysis of large data sets. The current state of toxicity databases is something of a mess. There are a number of databases, each with differing content, architecture, and searchability, that makes the task of integration extremely difficult.”

Part I

Data & Data Linkages



Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network



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

Search: ☐ All EPA ☒ This Area

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

Distributed Structure-Searchable Toxicity (DSSTox) Database Network is a project of [EPA's National Center for Computational Toxicology](#), 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>](#)



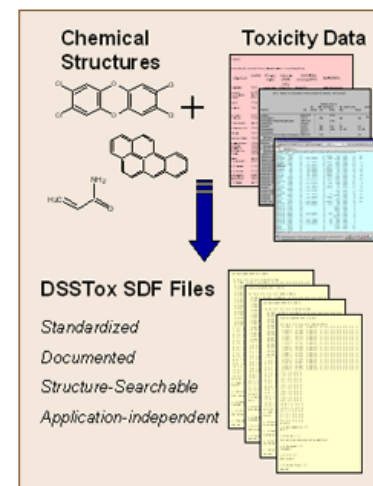
v2.0

[DSSTox Structure-Browser information Page](#)

27 October 2008

- [ARYEXP: European Bioinformatics Institute \(EBI\) ArrayExpress Repository for Gene Expression Experiments](#) **** New Files**
[Structure-index locator file](#) corresponds to chemical exposure-related experiments in ArrayExpress Repository with links to Experimental Accession numbers associated with each substance (data extracted 20Sept2008). Auxiliary data file (ARYEXP_Aux) indexed by chemical-experiment pair, includes 44 additional data fields.
- [GEOGSE: National Center for Biotechnology Information \(NCBI\) Gene Expression Omnibus \(GEO\) Series Experiments](#) **** New Files**
[Structure-index locator file](#) corresponds to chemical exposure-related experiments of GEO Series Repository with links to Series Accession numbers associated with each substance (data extracted 20Sept2008). Auxiliary data file (GEOGSE_Aux) indexed by chemical-experiment pair, includes 18 additional data fields.

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



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

DSSTox Data Files: [Details>](#)

[CPDBAS v5c 1547 29Apr2008](#)
[DBPCAN v4b 209 15Feb2008](#)
[EPAFHM v4b 617 15Feb2008](#)
[FDAMDD v3b 1216 15Feb2008](#)
[HPVCSI v2c 3548 15Feb2008](#)
[HPVISD v1b 1006 15Feb2008](#)
[IRISTR v1b 544 15Feb2008](#)
[NCTRER v4b 232 15Feb2008](#)
[NTPBSI v4a 2315 15Sep2008](#)
[NTPHTS v2b 1408 15Feb2008](#)
[TOXCST v2c 320 29Apr2008](#)

[ARYEXP v1a 887 21Oct2008](#) **** New**
[GEOGSE v1a 1014 23Oct2008](#) **** New**

[External Data Files \(ISSCAN\)](#)

- ▶ Chemical structure-indexing
- ▶ Quality review & data standardization
- ▶ Engage toxicologists in data representation
- ▶ Expand toxicological “endpoints” for modeling
- ▶ Facilitate structure-searching
- ▶ Improve linkages across data resources

(link-out to [EPA ACTON](#) coming soon, est. Sept. 2008)

▶ Structure “hit” resubmit feature: from either text or structure search, to facilitate structure

Published DSSTox Data Files:

- ~ 15K records
- >8000 unique chemicals
- 6 data files
- 6 files with links to Chemical Data web pages
- 3 HTS inventories
- 2 inventories of gene expression experiments
- EPA, NTP, FDA, NIH, EBI

ARYEXP	v1a 887 21Oct2008	ARYEXP: European Bioinformatics Institute (EBI) ArrayExpress Repository for Gene Expression Experiments Structure-Index Locator File : Index of chemical exposure-related experiments in ArrayExpress Repository with links to Experimental Accession numbers associated with each substance (data extracted 20Sept2008). Auxiliary data file (ARYEXP_Aux) indexed by chemical-experiment pair, includes 44 additional data fields.
CPDBAS	v5c 1547 29Apr2008	Carcinogenic Potency Database Summary Tables - All Species Database & Structure-Index Locator File : Tumor target site incidence, TD50 potencies, summary activity calls for rat, mouse, hamster, dog, and/or non-human primate; data reviewed and compiled from literature and NTP studies.
DBPCAN	v4b 209 15Feb2008	EPA Water Disinfection By-Products with Carcinogenicity Estimates Database : Carcinogenicity estimates (high, moderate, low concern) by EPA experts using a mechanism-based analog SAR approach on a set of 209 water disinfection by-products, mostly small halogenated organics.
EPAFHM	v4b 617 15Feb2008	EPA Fathead Minnow Acute Toxicity Database : Acute toxicities of 617 chemicals tested in common assay, with mode-of-action assessments and confirmatory measures.
FDAMDD ↑	v3b 1216 15Feb2008	FDA Center for Drug Evaluation & Research - Maximum (Recommended) Daily Dose Database : Maximum (recommended) daily dose (MRDD) values for 1216 pharmaceuticals in mg/kg-body weight (bw)/day, converted to mmol and normalized to dataset; MRDD values extracted from public literature sources.
GEOGSE	v1a 1014 23Oct2008	National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) Series Experiments Structure-Index Locator File : Index of chemical exposure-related experiments of GEO Series Repository with links to Series Accession numbers associated with each substance (data extracted 20Sept2008). Auxiliary data file (GEOGSE_Aux) indexed by chemical-experiment pair, includes 18 additional data fields.
HPVCSI	v2c 3548 15Feb2008	EPA High Production Volume Challenge Program Structure-Index File : Compiled structures for three chemical lists provided on EPA HPV Challenge Program website; each record includes reference index to dated list.
HPVISD	v1b 1006 15Feb2008	EPA High Production Volume Information System (HPV-IS) Data Structure-Index Locator File : Compiled structures for the chemical inventory of the on-line EPA HPV-IS with chemical-specific URLs linking to HPV-IS data pages containing chemical properties, fate properties and toxicity data.
IRISTR	v1b 544 15Feb2008	EPA Integrated Risk Information System (IRIS) Toxicity Review Database & Structure-Index Locator File : Compiled structures for EPA IRIS website with chemical-specific URLs linking to risk assessment summary data pages for 544 chemical substances.
NCTRER	v4b 232 15Feb2008	FDA National Center for Toxicological Research (NCTR) - Estrogen Receptor Binding Database : Estrogen receptor relative binding affinities tested in a common in vitro assay for 232 chemicals, listed with chemical class-based structure-activity features.
NTPBSI ↑	v4a 2315 15Sep2008	National Toxicology Program (NTP) On-line Chemical Bioassay Database Structure-Index Locator File : Compiled structures for the NTP On-line Database with chemical-specific URLs linking to NTP study summary pages; file includes fields for each of 4 main bioassay study areas with indicator values specifying presence or absence of study data for the chemical substance record.
NTPHTS	v2b 1408 15Feb2008	National Toxicology Program (NTP) High-Throughput Screening Project Structure-Index File : Compiled structures for set of 1408 NTP chemical substances provided to the NIH Chemical Genomics Center (NCGC) for HTS bioassay testing and to PubChem (NCGC PubChem_CIDs and PubChem_SIDs included in NTPHTS_v2a file); NCGC HTS bioassay data are deposited into PubChem by the NCGC Source and can be retrieved with these PubChem CID and SID lists.
TOXCST	v2c 320 29Apr2008	Research Chemical Inventory for EPA's ToxCast™ Program Structure-Index File : Compiled structures for 320 chemical substances that are candidates for Phase I High-Throughput screening (HTS) within the EPA ToxCast™ program.

CPDBAS_v5c_1547 (61 fields)

STRUCTURE
DSSTox_RID
DSSTox_CID
DSSTox_Generic_SID
DSSTox_FileID
STRUCTURE_Formula
STRUCTURE_MolecularWeight
STRUCTURE_ChemicalType
STRUCTURE_TestForm
STRUCTURE_TestForm_DefinedOrganic
STRUCTURE_Shown
TestSubstance_ChemicalName
TestSubstance_CASRN
TestSubstance_Description
ChemicalNote
STRUCTURE_ChemicalName_IUPAC
STRUCTURE_SMILES
STRUCTURE_Parent_SMILES
STRUCTURE_InChI
STRUCTURE_InChIKey
StudyType
Endpoint
Species

ActivityOutcome_CPDBAS_Mutagenicity

TD50_Rat_mg

TD50_Rat_mmol

TD50_Rat_Note

TargetSites_Rat_Male

TargetSites_Rat_Female

TargetSites_Rat_Both Sexes

ActivityOutcome_CPDBAS_Rat

ActivityScore_CPDBAS_Rat

TD50_Mouse_mg

...

TD50_Hamster_mg

...

TD50_Dog_mg

TargetSites_Dog

TD50_Rhesus_mg

TargetSites_Rhesus

TD50_Cynomolgus_mg

TargetSites_Cynomolgus

TD50_Dog_Rhesus_Cynomolgus_Note

ActivityOutcome_CPDBAS_Dog_Primates

ActivityOutcome_SingleCellCall

ActivityOutcome_MultiCellCall

ActivityOutcome_MultiCellCall_Details

Note_CPDBAS

NTP_TechnicalReport

ChemicalPage_URL

Potency Ranking
[1-100]

active
inactive
inconclusive

multisite
multisex
multispecies

adrenal gland;
bone;
clitoral gland;
esophagus;
ear/Zymbal's gland;
gall bladder;
harderian gland;
hematopoietic system;
kidney;
large intestine;
liver;
lung;
mesovarium;
mammary gland;
mixture;
myocardium;
nasal cavity
nervous system;
oral cavity
ovary;
pancreas;
peritoneal cavity;
pituitary gland;
preputial gland;
prostate;
skin;
small intestine;
spleen;
stomach;
subcutaneous tissue;
all tumor bearing animals;
testes;
thyroid gland;
urinary bladder;
uterus;
vagina;
vascular system.

CPDBAS_v5c_1547

Category & Species Totals by ActivityOutcome / ActivityScore

Category Field Name	ActivityOutcome_CPDBAS_Species (v5c)				ActivityScore_CPDBAS_Species [^]
	Active	Inactive	Inconclusive	Total # Entries	Total # Entries [^]
Mutagenicity (Salmonella)	403	457	0	860	NA
Rat	587	646	7	1240	1240
Mouse	445	558	5	1008	1008
Hamster	44	42	1	87	87
Dog_Primates	15	17	0	32	NA
SingleCellCall	806	738	3	1547	NA
MultiCellCall	582	570	0	1152	NA

Search PubChem Substance for epa dsstox

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

Tools: Links: Related Structures, BioAssays, Literature, Other Links

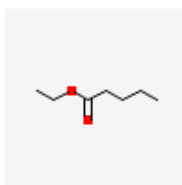
All: 16931 PubAssay: 3821 Protein3D: 0 Rule of 5: 10017

16931 DSSTox Substances

Items 1 - 20 of 16931

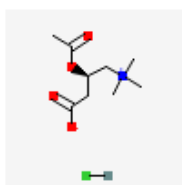
Page 1 of 847 Next
Related Structures

21: SID: 49693738



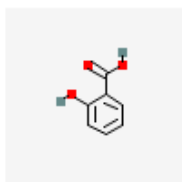
Ethyl pentanoate; ETHYL VALERATE; 539-82-2
Compound ID: 70882
Source: EPA DSSTox (45865)
IUPAC: ethyl pentanoate
MW: 130.184860 g/mol | MF: C₇H₁₄O₂

22: SID: 49693737



Acetyl-L-carnitine hydrochloride; (3R)-3-(acetyloxy)-4-(trimethylazaniumyl)butanoate hydrochloride; 5080-50-2
Compound ID: 6917959
Source: EPA DSSTox (45864)
IUPAC: (3R)-3-acetyloxy-4-trimethylazaniumylbutanoate hydrochloride
MW: 239.696520 g/mol | MF: C₉H₁₈ClNO₄

23: SID: 49693736



Alpha/Beta Hydroxy Acids (Glycolic Acid, Salicylic Acid)
Compound ID: 338
Source: EPA DSSTox (45863)
IUPAC: 2-hydroxybenzoic acid
MW: 138.120740 g/mol | MF: C₇H₆O₃

U.S. ENVIRONMENTAL PROTECTION AGENCY
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 | DSSTox | SDF Download Page | NTPBSI

SDF Download Page

NTPBSI: National Toxicology Program Bioassay On-line Database
Structure-Index Locator File

Version 4a, updated 15 September 2008
Update corresponds to chemical and Study Area content of NTP On-line Database Source Website as of 21 August 2008, with addition of 12 new records and corresponding studies in GeneTox or Carcinogenesis Study Areas.
NTP Bioassay On-line Database search page offers users the option of structure-searching the chemical portion of the NTP database through the DSSTox Structure-Browser.
NTPBSI substances have been deposited in PubChem and include URL links to chemical-specific NTP Study Pages.
Quick & Easy File Downloads: FTP Download Instructions

National Toxicology Program
Database Search Application
Search History: Search Results for 539-82-2
Found 1 Search Result for Search Term '539-82-2'
New Search | Clear History | Hide History

Table Instructions and Notes:
Selecting a CAS number will show you the studies conducted on this chemical. If applicable, matches on chemical synonyms are shown in parenthesis.

CAS NUMBER	TEST AGENT NAME
539-82-2	Ethyl valerate

Testing Status
Open All For This Test Agent

+ Genetic Toxicity Studies

Back to Top

Search PubChem BioAssay for dsstox Go Clear Save Search

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

Tool: Links: Related BioAssays, Compounds, Literature, Other Links

All: 74 Confirmatory: 63 MLSCN: 63 Protein Target: 0 Screening: 1 Summary: 0

Items 1 - 20 of 74

1: AID: 1204 Summary | Data (Active) Related BioAssays, Compounds, Literature, Other Links

DSSTox (NCTRER) National Center for Toxicological Research Estrogen Receptor Binding Database [Screening Method]

Source: EPA DSSTox

Substances Tested: 232; Active: 131

2: AID: 1195 Summary | Data (Active) Related BioAssays, Compounds, Literature, Other Links

DSSTox (FDAMDD) FDA Maximum (Recommended) Daily Dose Database [Other Method]

Source: EPA DSSTox

Substances Tested: 1240; Active: 587

3: AID: 1205

DSSTox (CPDBAS)

Source: EPA DSSTox

Substances Tested: 403; Active: 806

4: AID: 1189

DSSTox (CPDBAS)

Source: EPA DSSTox

Substances Tested: 131; Active: 131

5: AID: 1208

DSSTox (CPDBAS)

Source: EPA DSSTox

Substances Tested: 1240; Active: 587

11 DSSTox "Bioassays"

1. AID 1194: CPDBAS Salmonella Mutagenicity	403	/860 Active
2. AID 1189: CPDBAS SingleCellCall	806	/1547Active
3. AID 1205: CPDBAS MultiCellCall	582	/1152Active
4. AID 1208 CPDBAS Rat Bioassay (M/F/Both)	587	/1240 Active
5. AID 1199: CPDBAS Mouse Bioassay (M/F/Both)	445	/1007Active
7. AID 1190: CPDBAS Dog & Primates Bioassay	15	/32Active
8. AID 1195: FDAMDD – FDA Maximum Daily Dose	1216	/1216 Active
9. AID 1204: NCTRER – NCTR Estrogen Receptor Binding	131	/232 Active
10. AID 1188: EPA Fathead Minnow Acute Toxicity	580	/617 Active
11. AID 1201: EPA Disinfection By-Products Carcinogenicity Estimates	80	/209 Active



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

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

Search: ☐ All EPA ☒ This Area

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

DSSTox

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

Distributed Structure-Searchable Toxicity (DSSTox) Database Network is a project of [EPA's National Center for Computational Toxicology](#), 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>](#)



v2.0

[DSSTox Structure-Browser information Page](#)

27 October 2008

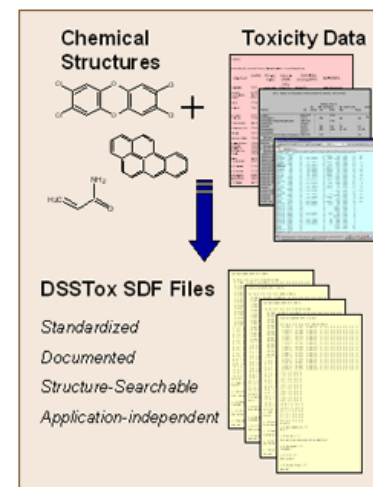
- [ARYEXP: European Bioinformatics Institute \(EBI\) ArrayExpress Repository for Gene Expression Experiments](#) **** New Files**
[Structure-index locator file](#) corresponds to chemical exposure-related experiments in ArrayExpress Repository with links to Experimental Accession numbers associated with each substance (data extracted 20Sept2008). Auxiliary data file (ARYEXP_Aux) indexed by chemical-experiment pair, includes 44 additional data fields.
- [GEOGSE: National Center for Biotechnology Information \(NCBI\) Gene Expression Omnibus \(GEO\) Series Experiments](#) **** New Files**
[Structure-index locator file](#) corresponds to chemical exposure-related experiments of GEO Series Repository with links to Series Accession numbers associated with each substance (data extracted 20Sept2008). Auxiliary data file (GEOGSE_Aux) indexed by chemical-experiment pair, includes 18 additional data fields.

15 September 2008

- [NTPBSI: National Toxicology Program Bioassay On-line Database Structure-Index Locator File](#) **** Updated to v4a**
Corresponds to updated content of NTP website, with addition of 12 new records and studies.

15 August 2008

- [Enhancements to DSSTox Structure Browser - Version 2.0](#)
 - Expanded "link-in" options and user-defined limits to allow external websites to access DSSTox Structure Browser Search Results pages
 - Minimum of 10 structure similarity hits displayed regardless of similarity search threshold
 - New **External Resources** feature provides structure-based "link-outs" to public on-line resources, such as [PubChem](#), [ChemSpider](#), and [Lazar In Silico Toxicology](#) [\[EXIT Disclaimer\]](#); (link-out to [EPA ACToR](#) coming soon, est. Sept. 2008)
 - Structure "hit" resubmit feature: from either text or structure search, to facilitate structure



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

DSSTox Data Files: [Details>](#)

[CPDBAS v5c 1547 29Apr2008](#)
[DBPCAN v4b 209 15Feb2008](#)
[EPAFHM v4b 617 15Feb2008](#)
[FDAMDD v3b 1216 15Feb2008](#)
[HPVCSI v2c 3548 15Feb2008](#)
[HPVISD v1b 1006 15Feb2008](#)
[IRISTR v1b 544 15Feb2008](#)
[NCTRER v4b 232 15Feb2008](#)
[NTPBSI v4a 2315 15Sep2008](#)
[NTPHTS v2b 1408 15Feb2008](#)
[TOXCST v2c 320 29Apr2008](#)

[ARYEXP v1a 887 21Oct2008](#) **** New**
[GEOGSE v1a 1014 23Oct2008](#) **** New**

➢ [External Data Files \(ISSCAN\)](#)

DSSTox Chemical Text Search

Choose search:

Auto-detect
Auto-detect
Chemical Name
CAS RN
InChI
Formula

Enter search text:

atrazine

Clear

Search

Data Files to Search

☒ All DSSTox Files
☐ Selected DSSTox Files

☒ ARYEXP_v1a

☒ EPBPAS_v5c

☒ AN_v4b

☒ EPAFHM_v4b

☒ FDAMDD_v3b

☒ GEOGSE_v1a

☒ HPVCSI_v2c

☒ HPVISD_v1b

☒ IRISTR_v1b

☒ NCTRER_v4b

☒ NTPBSI_v4a

☒ NTPHTS_v2b

☒ TOXCST_v2c

EBI ArrayExpress Repository for Gene Expression Experiments (887 records)

DSSTox Chemical Structure Search

Enter SMILES string:

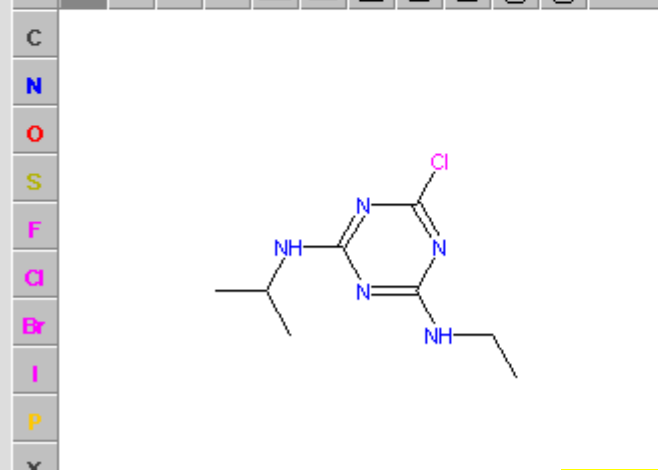
Preview below

Clear

Search

Or draw a molecule or substructure using the JME editor:

CLR NEW DEL D-R +/- UDO JME



Search Options

☒ Exact match

☒ Substructure

☒ Similarity

Threshold: 80 %

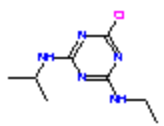
▲ ▼

Clear

Search

http://www.epa.gov/dsstox_structurebrowser

Search Results Summary for DSSTox Substances - File Breakdown Incidences

Query	Results Type	Hits	Display
Structure: 	Exact matches	1	Details
	Substructures	2	Details
	Similarity $\geq 51.2\%$	12	Details

PubChem CID

External Resources

[PubChem](#)
[EPA ACToR](#)
[ChemSpider](#)
[Lazar in silico tox](#)

[EXIT Disclaimer](#)

DSSTox File ?

ARYEXP_v1a
 CPDBAS_v5c
 DBPCAN_v4b
 EPAFHM_v4b
 FDAMDD_v3b
 GEOGSE_v1a
 HPVCSI_v2c
 HPVISD_v1b
 IRISTR_v1b
 NCTRER_v4b
 NTPBSI_v4a
 NTPHTS_v2b
 TOXCST_v2c

Total Unique Substance Hits

Total Substance Hits - All Files

?

NCBI **PubChem Compound**

PubChem Compound Summary

atrazine - Compound Summary

A selective triazine herbicide. In humans, apparent skin manifestations or poisoned sheep and cattle may suggest a stiff gait, increased respiratory rate, congestion of the lungs, liver, and 11th ed)

[Table of Contents](#)
[Drug and Chemical Info](#)
[BioActivity Results](#)
[Synonyms](#)
[Properties](#)
[Descriptors](#)
[Compound Info](#)
[Substance Info](#)
[Category](#)
[Exports](#)

Drug and Chemical Info for **Atrazine**

ChemSpider™
Building a Structure-Centric Community for Chemists

Home Search Services Resources About

Simple Search
Structure Search
LASSO Search
Advanced LASSO
Chemical Elements
Properties Search
Predicted Properties
Data Source Search
Literature Search
NEBI Entrez Search
PubChem Search
Advanced Search
Searches History

Lazar Toxicity Predictions

Input Validation Documentation

1. Draw a chemical structure (help)

or enter the SMILES string
CC1=NC(=NC2=C1N=CN2)C(=O)N

2. Select a toxic endpoint
 Rodent carcinogenicity (multiple sex/species/sites)

3. Predict

InChI

SMILES

CID

U.S. ENVIRONMENTAL PROTECTION AGENCY

ACToR: Aggregated Computational Toxicology Resource

Search: C All EPA R This Area

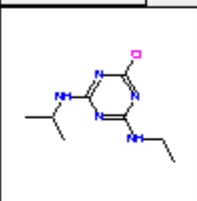
Chemical Summary : Atrazine

GOID	113
CASRN	1912-24-9
Formula	CBH14CN5
MW	215.6833
SMILES	CC1=NC(=NC2=C1N=CN2)C(=O)N
INCHI	InChI=1C8H14CN5/c14-10-7-12-6/9-13-8/14-7-11-5/2-3/6-4-2-1-3-9-3-10-11-12-13-14/9-10-11H

Details

Query Results Type Hits Display

Structure:



Exact matches	1
Substructures	2
Similarity ≥ 51.2%	12

Details

Details

Details

?

Output Options

Choose Format

?

External Resources

[Pubchem](#) [EPA ACToR](#)

[ChemSpider](#) [Lazar in silico tox](#)

EXIT Disclaimer

?
DSSTox

Substance ID

?
Similarity Score%

Structure Match

Substance Name

CASRN

Substance Description

Details
(Data Files)

20112	100	 ss	Atrazine	1912-24-9	single chemical compound	ARYEXP GEOGSE IRISTR NTPBSI TOXCST	CPDBAS HPVCSI NCTRER NTPHTS
21196	99	 ss	Propazine	139-40-2	single chemical compound	CPDBAS IRISTR	HPVCSI TOXCST
27608	96.3	 ss	1,3,5-Triazine-2,4-diamine, 6-chloro-N-(1,1-dimethylethyl)-N'-ethyl-	5915-41-3	single chemical compound	HPVCSI	
21268	96.2	 ss	Simazine	122-34-9	single chemical compound	CPDBAS IRISTR TOXCST	HPVCSI NCTRER

Click **ss** to submit displayed structure for structure search.

U.S. ENVIRONMENTAL PROTECTION AGENCY

Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

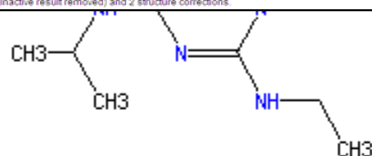
Recent Additions | Contact Us Search: All EPA ☐ This Area ☐ Go

You are here: EPA Home > Computational Toxicology Research > DSSTox > SDF Download Page: CPDBAS

SDF Download Page

CPDBAS: Carcinogenic Potency Database Summary Tables - All Species Database File

**** Version 5c, revised 29 April 2008:**
 > One Mouse bioassay correction (inactive result removed) and 2 structure corrections



CPDBAS: Carcinogenic Potency Database Summary Tables (1547 records)

CPDBAS_v5c_1547_29Apr2008

[CPDBAS Source Website](#)

[View CPDB Chemical Substance Data Page](#)

Search File Incidences Search Details Substance Results ?Help

The Carcinogenic Potency Project

The Carcinogenic Potency Database (CPDB)
 Lois Swirsky Gold, PhD, Director

Search This Web Site

Web Page on Each Individual Chemical	CPDB Summary Table by Chemical	CPDB Summary Table by Target Site	CPDB Plot Screen Version Format	CPDB Plot Published Format	CPDB Excel and Tab-separated Formats	CPDB Methods and Guide to Plots	Chemical Structures, SMILES, InChI
Dataset on Dosing and Survival: Excel and Tab-separated	Summary Table by Chemical or NC/NTP Bioassays	Summary Table by Target Site or NC/NTP Bioassays	Animal Cancer Tests and Human Cancer Risk	Macromolecules about the Causes of Cancer	Our Publications by Topic	Our Publications by Year	Aristolochic Acid

Acknowledgment and Support

DSSTox_RID
 DSSTox_Generic_SID
 TestSubstance_ChemicalName
 TestSubstance_CASRN
 TestSubstance_Description
 STRUCTURE_Shown
 StudyType
 Endpoint
 Species
 ActivityOutcome_CPDBAS_Mutagenicity
 TD50_Rat_mg
 TD50_Rat_mmol
 ActivityScore_CPDBAS_Rat
 TD50_Rat_Note
 TargetSites_Rat_Male
 TargetSites_Rat_Female
 ActivityOutcome_CPDBAS_Rat
 ActivityScore_CPDBAS_Mouse
 TD50_Mouse_Note



The Carcinogenic Potency Project <http://potency.berkeley.edu/>

Aniline (CAS 62-53-3)
 SMILES, InChI, and Structures are below

Rats and Mice: Cancer Test Summary

Rat Target Sites		Mouse Target Sites		TD ₅₀ (mg/kg/day)	
Male	Female	Male	Female	Rat	Mouse
no positive	no test	no test	no test	no positive	no test

Key to the Table Above

Positivity: For each chemical with a positive (carcinogenic) experiment in the Carcinogenic Potency Database (CPDB), results are included on carcinogenic potency (TD₅₀) in each species and target sites in males and females. Positivity is determined by an author's opinion in a published paper. If all experimental results in the CPDB are negative in a species group, "no positive" appears. If the CPDB has no experiments in the *res-species* group, "no test" appears. The summary presents the strongest evidence of carcinogenicity in each group. If there are both positive and negative experiments in a *res-species*, the negative results are ignored in the Summary Table.

Target Site Codes: Target sites are listed if any author of published experimental results concluded that tumors were induced in that organ by the test agent. If there is more than one positive experiment in a *res-species*, target sites listed may be from more than one experiment, e.g. if liver and lung are both listed, then liver may have been a target in one experiment and lung in another.

TD₅₀: Our standardized measure of carcinogenic potency. TD₅₀ is the daily dose rate in mg/kg body weight/day to induce tumors in half of test animals that would have remained tumor-free at zero dose. Whenever there is more than one positive experiment in a species, the reported TD₅₀ value is a **Harmonic Mean** calculated using the TD₅₀ value from the most potent target site in each positive experiment.

The Carcinogenic Potency Database (CPDB) is a unique and widely used international resource of the results of 6540 chronic, long-term animal cancer tests on 1547 chemicals. The CPDB provides easy access to the bioassay literature, with qualitative and quantitative analyses of both positive and negative experiments that have been published over the past 50 years in the general literature through 2001 and by the National Cancer Institute/National Toxicology Program through 2004. The CPDB standardizes the diverse literature of cancer bioassays that vary widely in protocol, histopathological examination and nomenclature, and in the published author's choices of what information to provide in their papers. Results are reported in the CPDB for tests in rats, mice, hamsters, dogs, and nonhuman primates.

For each experiment, information is included on species, strain, and sex of test animal, features of experimental protocol such as route of administration, duration of dosing, dose level(s) in mg/kg body weight/day, and duration of experiment; experimental results are provided on target organ, tumor type, and tumor incidence; carcinogenic potency (TD₅₀) and its statistical significance; shape of the dose-response; author's opinion as to carcinogenicity; and literature citation.

Only tests with dosing for at least 1/4 the standard lifespan of the species and experiment length at least 1/4 the lifespan are included in the CPDB. Only routes of administration

0.169693249315084 mmol/kg-bw/day

34

TD50 is harmonic mean of more than one positive test

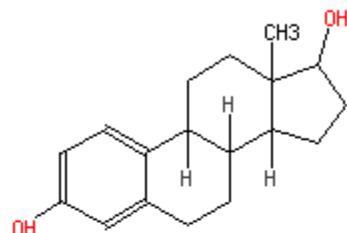
mammary gland

hematopoietic system; mammary gland; uterus

active

0

no positive results



GEOGSE:
NCBI Gene Expression Omnibus
Experiments (1014 records)

GEOGSE_v1a_1014_230ct2008

[GEOGSE Source Website](#)

EXIT

DSSTox_RID	55218
DSSTox_Generic_SID	20573
TestSubstance_ChemicalName	17 beta-Estradiol (E2)
TestSubstance_CASRN	50-28-2
TestSubstance_Description	single chemical compound
STRUCTURE_Shown	tested chemical
Chemical_StudyType	Treatment
StudyType	microarray

Experiment_Accession

GSE4664	GSE3013	GSE2251
GSE7844	GSE848	GSE5868
GSE1839	GSE1045	GSE5315
GSE1153	GSE2740	GSE4054
GSE4668	GSE5200	GSE6219
GSE7206	GSE2889	GSE10800
GSE10618	GSE1819	GSE2225
GSE7798	GSE1303	GSE2195
GSE9759	GSE11115	GSE1486
GSE2292	GSE11115	GSE8383
GSE11506	GSE11352	GSE11791

EXIT Disclaimer

NCBI
Gene Expression Omnibus

HOME | SEARCH | SITE MAP | Handout | NAR 2006 Paper | NAR 2002 Paper | FAQ | MIAME | Email GEO

NCBI > GEO

Not logged in | Login

Gene Expression Omnibus: a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

Public data

GPL Platforms	4954
GSM Samples	249197
GSE Series	9567
Total	263718

Site contents

Documentation

Overview | FAQ | Submission guide | Linking & citing | Journal citations | Programmatic access | DataSet clusters | GEO announce list

GEO navigation

QUERY

DataSets

Gene profiles

GEO accession

GEO BLAST

BROWS

DataSets Platforms

SUBMIT

NCBI
Gene Expression Omnibus

HOME | SEARCH | SITE MAP | Handout | NAR 2006 Paper | NAR 2002 Paper | FAQ | MIAME | Email GEO

NCBI > GEO > **Accession Display**

Not logged in | Login

GEO help: Mouse over screen elements for information.

Scope: Format: Amount: GEO accession:

Series GSE4664 [Query DataSets for GSE4664](#)

Status: Public on Apr 14, 2006

Title: 24 hour time course: regulation of uterine genes by estradiol in ovariectomized mice

Organism(s): *Mus musculus*

Experiment type: Expression profiling by array

Summary: Using microarray technology, we compared the global expression pattern of uterine RNA from ovariectomized control mice to those of ovariectomized mice treated with estradiol for various intervals between 30 minutes and 24 hours.
Keywords: estrogen, uterus, genomic, mouse

Overall design: Mice (C57 Bl/6) were treated with sesame oil or estradiol and uteri were collected at 30 minutes, two hours, 6 hours, 12 hours and 24 hours after treatment and snap frozen in liquid nitrogen. Three or four uteri from each treatment group were pooled and RNA was prepared using Trizol reagent and the RNeasy clean up protocol. Two arrays for each time point were used, incorporating a fluor reversal.

Contributor(s): Korach KS, Hewitt SC, Deroo BJ, Grissom SF, Collins JB

Submission date: Apr 13, 2006

Contact name: NIEHS Microarray Core

E-mail(s): microarray@niehs.nih.gov

Organization name: NIEHS

Department: DIR

Lab: Microarray Core

Street address: 111 T.W. Alexander Drive

City: RTP

State/province: NC

ZIP/Postal code: 27709

Country: USA

[http://www.epa.gov/dsstox_structurebrowser/...](http://www.epa.gov/dsstox_structurebrowser/)

?dbs=cpdbas

[?qtype=cas&qval=87-86-5](#)
[?q type=name&qval=atrazine](#)
[?q type=smiles&qval=CC](#)
[?qtype=inchikey&qval=...](#)
[?qtype=cid&qval=20238](#)

PubChem_CID
InChIKey
CAS
SMILES

Link-ins

?qtype=sid&qval=20112

?qtype=rid&qval=20112

Link-outs

DSSTox Download Page
Source Home URL
Source Chemical Data URL

EPA IRIS Summary

EPA HPV Information System

NTP

NCBI

EBI

DSSTox Structure-Browser

DSSTox SDF Files & Documentation

DSSTox

DSSTox

EPA ACToR

CPDB

The Carcinogenic Potency Project

PubChem

Lazar Toxicity Predictions

ChemSpider

U.S. Environmental Protection Agency

Integrated Risk Information System

Aniline Quickview (CASRN 62-53-3)

High Production Volume Information System (HPVIS)

National Toxicology Program

Gene Expression Omnibus

EMBL-EBI

ArrayExpress

Query for Experiments

DSSTox Chemical Structure Search

DSSTox Chemical Test Search

ACToR: Aggregated Computational Toxicology Resource

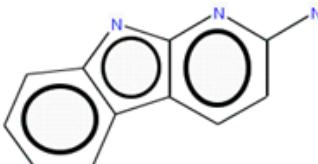
Chemical Summary: 1,3,5-Triazine-2,4-Diamine, 6-Chloro-N-Ethyl-N'-(1-Methylethyl)-

Chemical structures and IDs:

- 2,3,5-Tetrachlorobiphenyl (ID:14)
- 2,3,5-Tetrachlorobiphenyl (ID:15)
- 2,3,5-Tetrachlorobiphenyl (ID:16)
- 2,3,5-Tetrachlorobiphenyl (ID:17)
- 2,3,5-Tetrachlorobiphenyl (ID:18)
- 2,3,5-Tetrachlorobiphenyl (ID:19)
- 2,3,5-Tetrachlorobiphenyl (ID:20)
- 2,3,5-Tetrachlorobiphenyl (ID:21)
- 2,3,5-Tetrachlorobiphenyl (ID:22)
- 2,3,5-Tetrachlorobiphenyl (ID:23)
- 2,3,5-Tetrachlorobiphenyl (ID:24)
- 2,3,5-Tetrachlorobiphenyl (ID:25)
- 2,3,5-Tetrachlorobiphenyl (ID:26)
- 2,3,5-Tetrachlorobiphenyl (ID:27)
- 2,3,5-Tetrachlorobiphenyl (ID:28)
- 2,3,5-Tetrachlorobiphenyl (ID:29)
- 2,3,5-Tetrachlorobiphenyl (ID:30)
- 2,3,5-Tetrachlorobiphenyl (ID:31)
- 2,3,5-Tetrachlorobiphenyl (ID:32)
- 2,3,5-Tetrachlorobiphenyl (ID:33)
- 2,3,5-Tetrachlorobiphenyl (ID:34)
- 2,3,5-Tetrachlorobiphenyl (ID:35)
- 2,3,5-Tetrachlorobiphenyl (ID:36)
- 2,3,5-Tetrachlorobiphenyl (ID:37)
- 2,3,5-Tetrachlorobiphenyl (ID:38)
- 2,3,5-Tetrachlorobiphenyl (ID:39)
- 2,3,5-Tetrachlorobiphenyl (ID:40)
- 2,3,5-Tetrachlorobiphenyl (ID:41)
- 2,3,5-Tetrachlorobiphenyl (ID:42)
- 2,3,5-Tetrachlorobiphenyl (ID:43)
- 2,3,5-Tetrachlorobiphenyl (ID:44)
- 2,3,5-Tetrachlorobiphenyl (ID:45)
- 2,3,5-Tetrachlorobiphenyl (ID:46)
- 2,3,5-Tetrachlorobiphenyl (ID:47)
- 2,3,5-Tetrachlorobiphenyl (ID:48)
- 2,3,5-Tetrachlorobiphenyl (ID:49)
- 2,3,5-Tetrachlorobiphenyl (ID:50)

Generic Chemical

Preferences: Refresh Visibility Table Size Select Skin



GCID: 1
 CASRN: 26148-68-5
 CID: 11869
 CCID: 11869
 Formula: C11H9N3
 MW: 183.212
 SMILES: NC1C=CC2=C(N1)NC3=CC=CC=C23
 INCHI: InChI=1/C11H9N3/c12-10-6-5-8-7-3-1-2-4-9(7)13-11(8)
 QC Status: QUEUED

Substances

Chemical Id	Data Collection	Name
1	CPDBAS_DSSTOX	A-alpha-C
18819	CalEPA	A-alpha-C (2-Amino-9H-pyrido[2,3-b]indole)
35633	INCHEM_IARC	A-alpha-C (2-Amino-9H-Pyrido[2,3-b]Indole)
78935	Scorecard	2-AMINO-9H-PYRIDO (2,3-B) INDOLE (A-ALPHA-C)
95724	SRS_CAS-9C	2-Amino-9H-pyrido[2,3-b]indole
168920	SRS_ChemIDStd	2-Amino-9H-pyrido[2,3-b]indole
514465	ChemIDplus	1H-Pyrido[2,3-b]indol-2-amine
919306	EPA	26148-68-5
1413536	EPA_SRS_NTS	1H-Pyrido[2,3-b]indol-2-amine

Notes

Chemical Id	Data Collection	Note
1	CPDBAS_DSSTOX	blank
1413536	EPA_SRS_NTS	ACD/Name-to-Structure Batch v. 8.05

Synonyms

Name
1H-Pyrido[2,3-b]indol-2-amine
1H-Pyrido[2,3-b]indol-2-amine
2-AMINO-9H-PYRIDO (2,3-B) INDOLE (A-ALPHA-C)
2-AMINO-9H-PYRIDO(2,3-B)INDOLE
2-Amino-9H-pyrido[2,3-b]indole
2-Amino-alpha-carboline
26148-68-5
9H-Pyrido[2,3-b]indole, 2-amino-
9H-pyrido[2,3-b]indol-2-amine

Previous 1-10 of 17 Next 7

[ACToR Home](#)
[About ACToR](#)
[Data Collections](#)
[Search](#)

MSDS

MSDS Id
No rows yet.

Mesh

[Mesh Name](#)
[PubMed via MeSH: 2-amino-9H](#)

External Search

[External Search](#)
[TOXNET CCRIS](#)
[TOXNET DART ETIC](#)
[TOXNET EMIC](#)
[TOXNET GENETOX](#)
[TOXNET HazMap](#)
[TOXNET Household Products](#)
[TOXNET HSDB](#)
[TOXNET MESH Headings](#)
[TOXNET TOXLINE](#)
[TOXNET TOXMAP](#)

Where are the data??

ACToR: Aggregated Computational Toxicology Resource

[Bookmark](#)

Recent Additions | Contact Us Search: ☐ All EPA ☒ This Area

You are here: [EPA Home](#) » [Area Name](#) » Page Title

ACToR Home

Data Collection

Name: [IRISTR_DSSTOX](#)

Description	EPA Integrated Risk Information System http://www.epa.gov/iris/ . A structure index file with links to webpage for
Source Type	DSSTox Collection
Number of Substances	544
Number of Generic Chemicals	536
Compilation Date	9/27/2007
Compilation Instructions	Download SD and excel files from DSSTox site. Preprocess according to ToxCast SOP 004

Fetch Size: 100 Offset: 0 Fetch

Generic Chemical	CASRN	Structure	Substance	Source	SID	Name	Structure	PhysicoChemical	Biochemical	In vivo toxicology (tabular primary)	In vivo toxicology (study listing primary)	In vivo toxicology (tabular secondary)	In vivo toxicology (summary calls)	In vivo toxicology (summary report via URL)	Chemical Use Level	Carcinogenicity	GeneTox	DevTox	ReproTox	EcoTox	FoodSafe		
Details	83-32-9		Details	DSSTOX_1774	Acenaphthylene, 1,2-dihydro-		1687	10	7	1	1	4	3	5	5	5	6	3	3	1			
Details	208-96-8		Details	DSSTOX_3845	Acenaphthylene		5129				1	1	3	2		4	3	3	2	1			
the data??							DSSTOX_3846	Acephate	5130	9		6	1	10	5	1	9	4	6	5	3		
							DSSTOX_2	Acetaldehyde	2	9	40	3	1	4	16	8	9	14	9	5	3	2	1

Towards a Public “Toxico-chemogenomics” Capability

NCBI

Gene Expression Omnibus

HOME | SEARCH | SITE MAP | Handout | NAR 2006 Paper | NAR 2002 Paper | FAQ

NCBI > GEO ?

Gene Expression Omnibus: a gene expression/molecular abundance repository supporting MIAME compliant data submissions, and a curated, online resource for gene expression data browsing, query and retrieval.

GEO navigation

QUERY

BROWSE

SUBMIT

DataSets

GO

Public data

GPL	Platforms	4954
GSM	Samples	249197
GSE	Series	9567
Total		263718

Site contents

EMBL-EBI
European Bioinformatics Institute

You are logged in as guest Login »

ArrayExpress

Give an experiment accession number

or fill out some of the following fields to get a list of matching experiments

Species
« any species »

Experiment type
« any type »

Experimental Factors
« any factor »

Description contains the word

Author

Laboratory

Publication
« don't specify »

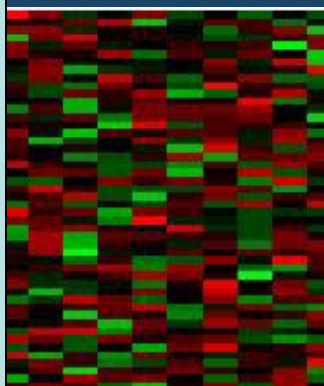
Array provider

- ▶ No chemical standards
- ▶ Difficult to identify chemical exposure-related experiments

- ▶ PERL scripts to filter web-accessed data content
- ▶ Manual review of Submitter textual descriptions
- ▶ Creation of initial chemical & experimental index
- ▶ QA & structure annotation



9957
Experiments



Experiments

Genes

Pathways

2381
Chemical
experiments

*Vehicle
Media
Reference
Combination*

2134
Chemical
treatment-
related
experiments

Treatment

1014
Unique
chemicals

*751 defined organics
71 inorganics
19 organometallics*

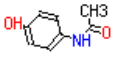
Pubchem
ChemSpider

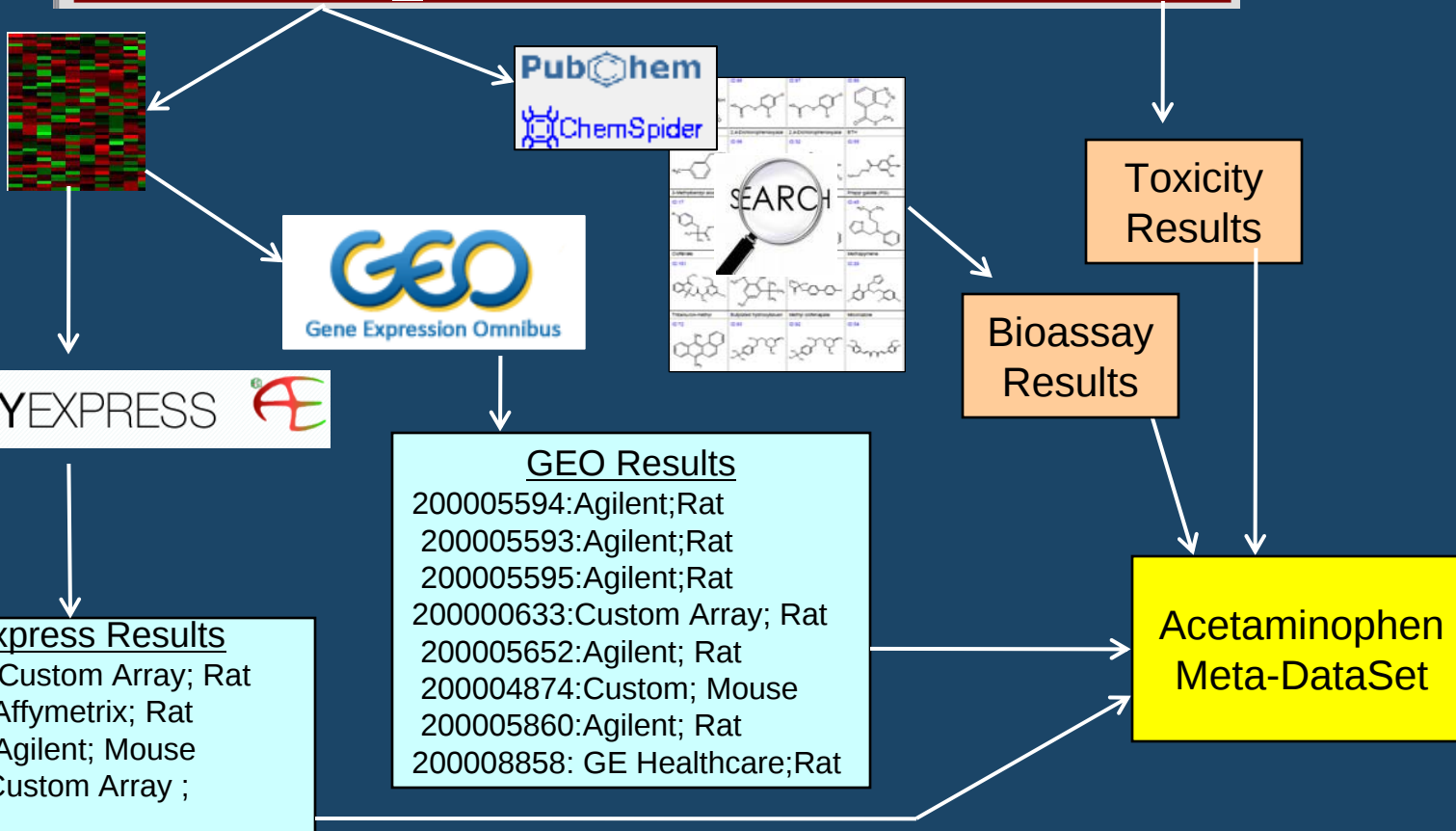
SEARCH

Chem 1
Chem 2
Chem 3
Chem 4
...
Chem 1014

DSSTox
GEOCSI

Chemical Search Paradigm

DSSTox Substance ID	Similarity Score%	Structure Match	Substance Name	CASRN	Substance Description	Details (Data Files)
20006	100		Acetaminophen	103-90-2	single chemical compound	CPDBAS FDAMDD NTPHTS EPAFHM NTPBSI



Part II

Toxicity Profiling

National Academy of Sciences Report (2007)
Toxicity Testing in the Twenty-first Century: A Vision and a Strategy

**NAS PANEL SEEKS MAJOR SHIFT IN HOW EPA ASSESSES
CHEMICALS' TOXICITY**

Date: June 22, 2007 -

Inside EPA

Online access provided by InsideEPA.com

A National Academy of Sciences (NAS) panel is calling for a major shift in how EPA assesses chemicals' toxicity, recommending that the agency base its toxicological research and regulatory processes on how substances affect biological pathways -- which send information within and between cells -- rather than so-called health endpoints, such as cancer.

POLICYFORUM

Science: Feb 15, 2008

TOXICOLOGY

**Transforming Environmental
Health Protection**

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

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.

In 2005, the U.S. Environmental Protection Agency (EPA), with support from the U.S. throughput screening (HTS) and other automated screening assays into its testing tion, usually between 2 and 10 μ M, and tolerate high false-negative rates. In contrast, in



Home

Basic Information

Organization
Post Doc Profiles

Framework

Databases and Models

Research Activities
ACToR
ToxCast™
Virtual Liver

Conferences and
Seminars

Publications

BOSC Information

EPA Community of
Practice

Jobs and Opportunities

Related Information

U.S. ENVIRONMENTAL PROTECTION AGENCY

National Center for Computational Toxicology

[Contact Us](#)

Search:

☐ All EPA ☒ This Area

Go

You are here: [EPA Home](#) » [National Center for Computational Toxicology](#) » ToxCast™ Program

ToxCast™ Program

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

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

Introduction

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. These hazard predictions will provide EPA regulatory programs with science-based information helpful in prioritizing chemicals for more detailed toxicological evaluations, and lead to more efficient use of animal testing.

In its first phase, ToxCast™ is profiling over 300 well-characterized chemicals (primarily pesticides) in over 400 HTS endpoints. These 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 compounds being examined in Phase 1 of ToxCast™ 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. ToxRefDB is integrated into a more comprehensive data management system developed by NCCT called ACToR (Aggregated Computational Toxicology Resource), that manages the large-

ToxCast™ Navigation

[Introduction](#)

[ToxCast™ Chemicals](#)

[ToxCast™ Assays](#)

[ToxCast™ Information
Management](#)

[ToxCast™ Partnerships](#)

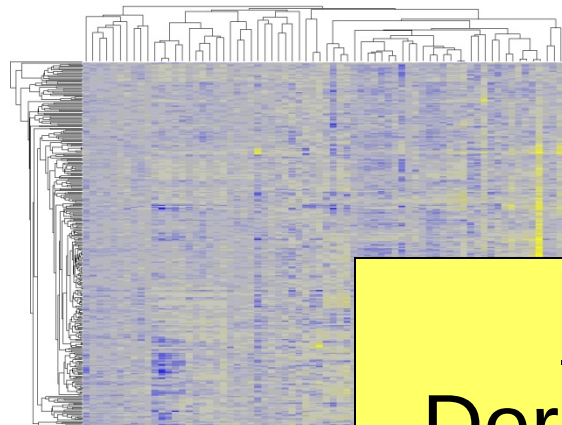
[ToxCast™ Contractors](#)

[ToxCast™ Presentations](#)

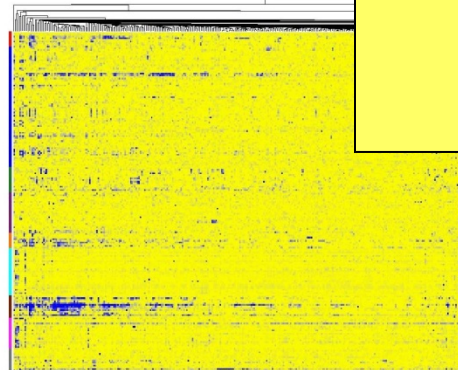
[ToxCast™ Publications](#)

[ToxCast™ News](#)

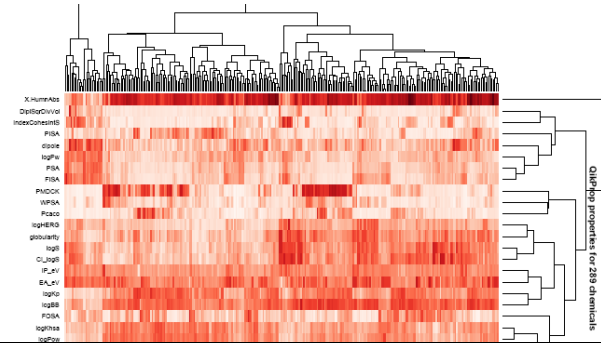
Correlating Domain Outputs



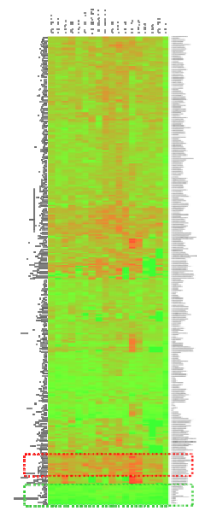
Cellular Assays



Biochemical Assays



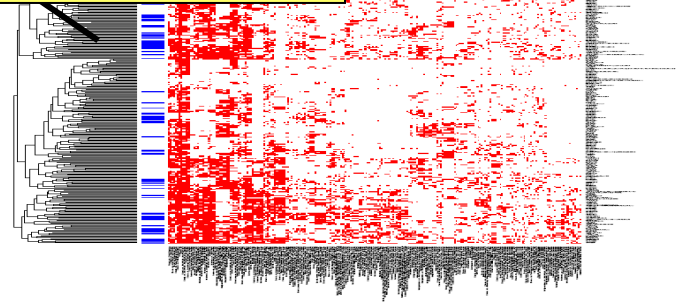
Genomic Signatures



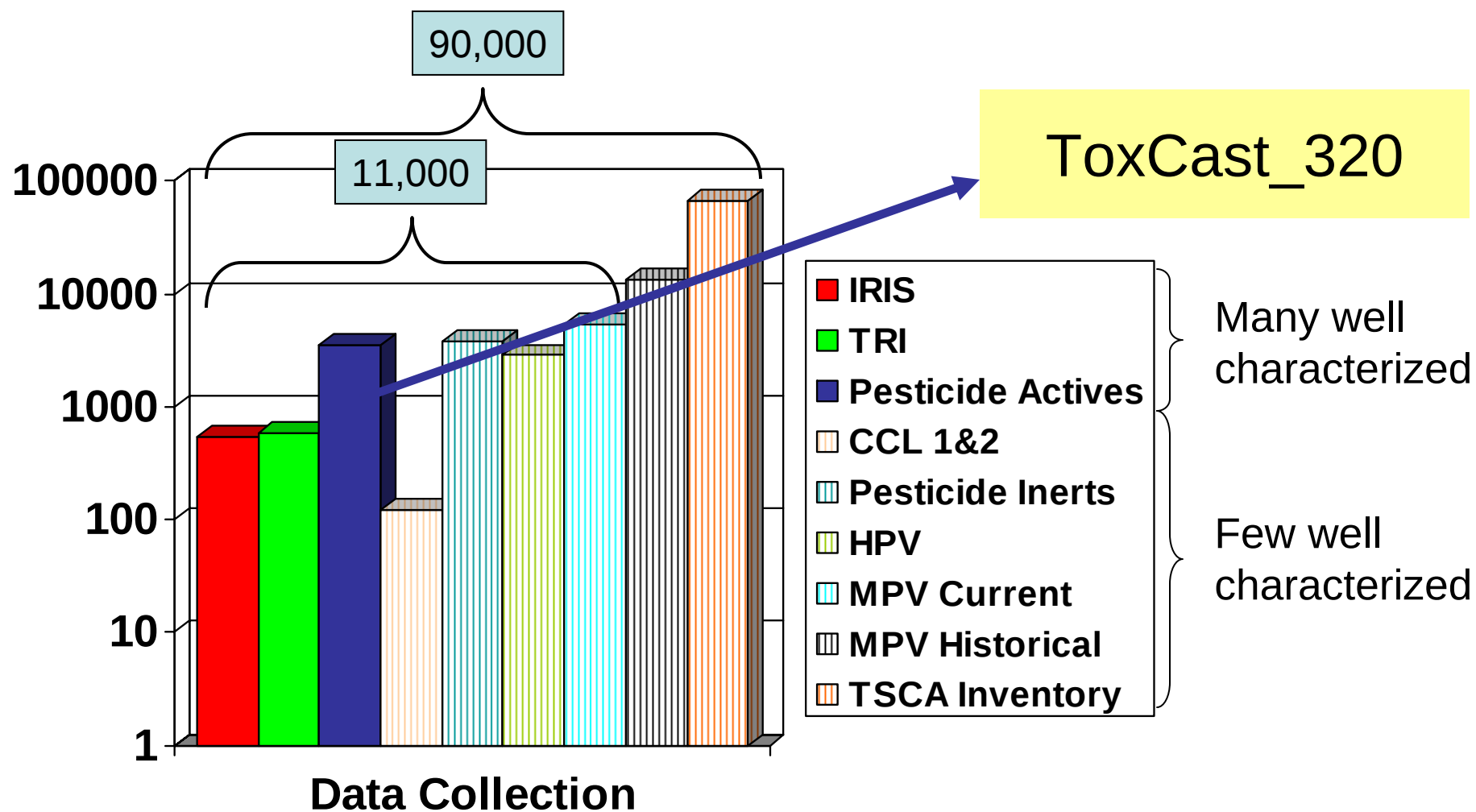
Toxicology Endpoints

EPA ToxCast Goal:
Derive “Signatures” from *in vitro* & *in silico* assays to predict *in vivo* endpoints

in silico Predictions



ToxCast Phase I Chemicals



Chemical Classes in ToxCast_320 (Phase I)

• 309 Unique Structures

• Replicates for QC

• 291 Pesticide Actives

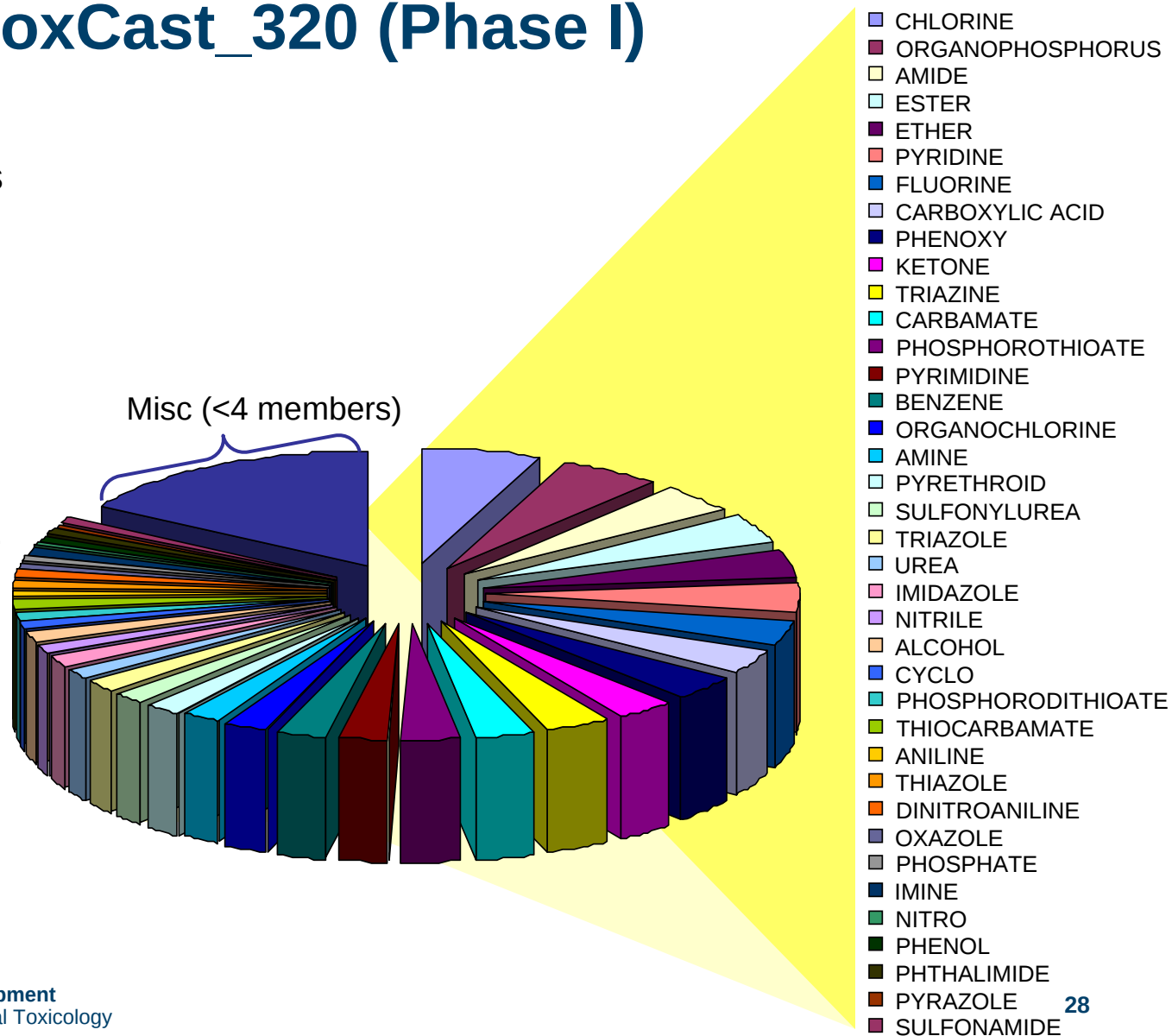
• 9 Industrial Chemicals

• 8 Metabolites

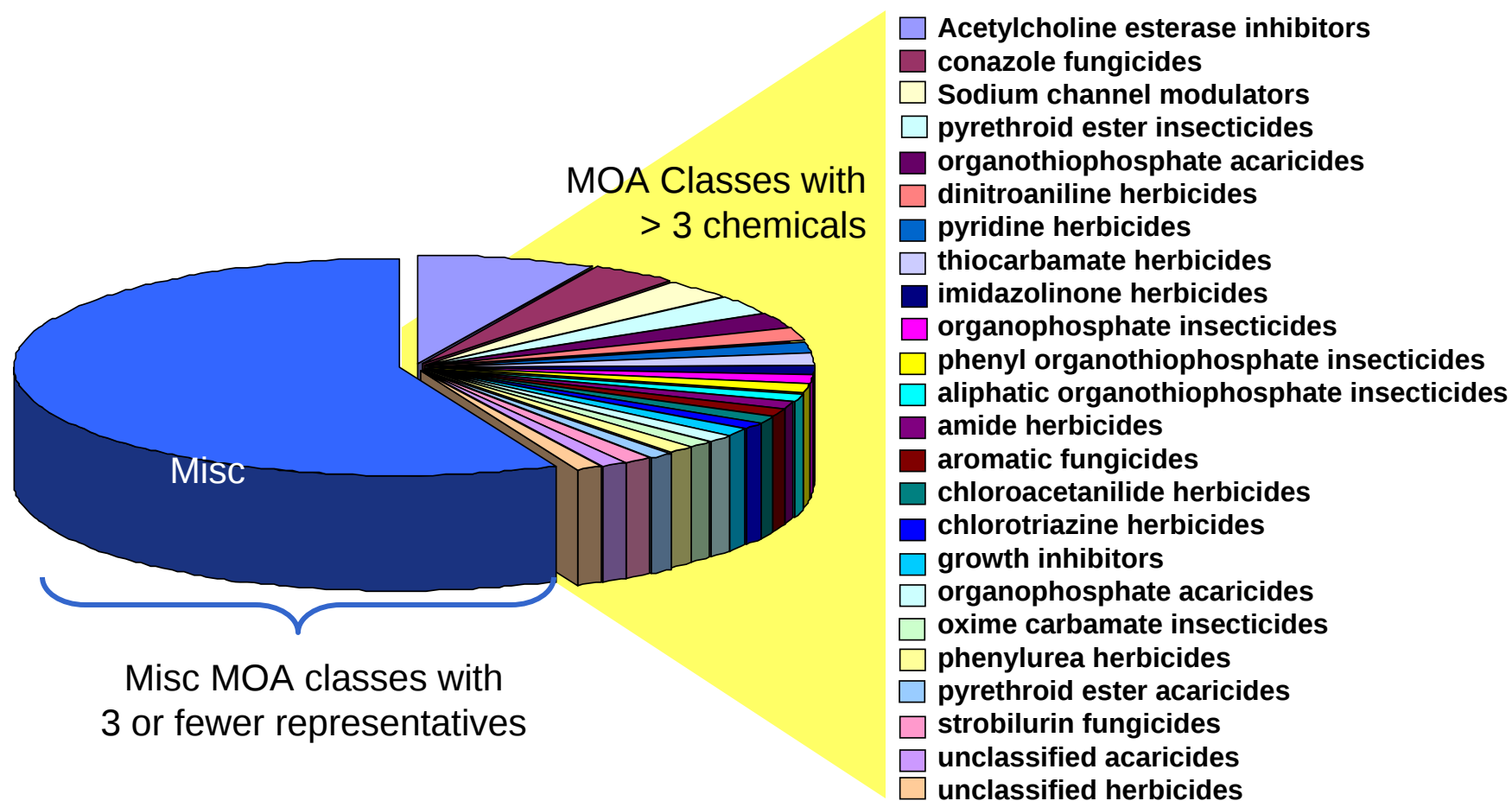
• 56/73 Proposed Tier 1
Endocrine Disruption
Screening Program

• 14 High Production
Volume Chemicals

• 11 HPV Challenge



Mode-of-Action Classes in ToxCast_320 (Phase I)



EPA Pesticide Programs: Data Evaluation Records (DERs)

- Used for hazard identification and characterization

- Study Types

- Chronic
- Cancer
- Subchronic
- Multigeneration
- Developmental
- Others: DNT, Neurotox, Immu

- Derive Endpoints (NOAEL/LOAEL)

- Systemic
- Parental
- Offspring
- Reproductive
- Maternal
- Developmental

- Critical Effects for Endpoints

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)

\$10,000,000

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

Data Entry Completeness Score

Partially Complete (Effect Data)

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

ToxRefDB
Input Form

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY



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

Additional Study Duration Information

Finish 104 week

Animal and Dose Information

Species rat

Method/Route of Administration

Strain [Other]

Feed

Animal and Dose Administration Comments (Including Not In List)

Strain: Hannover substrain (SPF) Wistar-derived

Study Effect List

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	# / Group	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

M

#/group

50

Dose Period Type

Initial-to-Terminal

Dose Units

2.7 mg/kg/day

Duration Units

104 week

Save

Delete

New



Navigation buttons: Previous, Next, First, Last

Show all
Effects
[Assign
LOAELs]

Study Design

Treatment Groups

Study Design Level Controls



Search



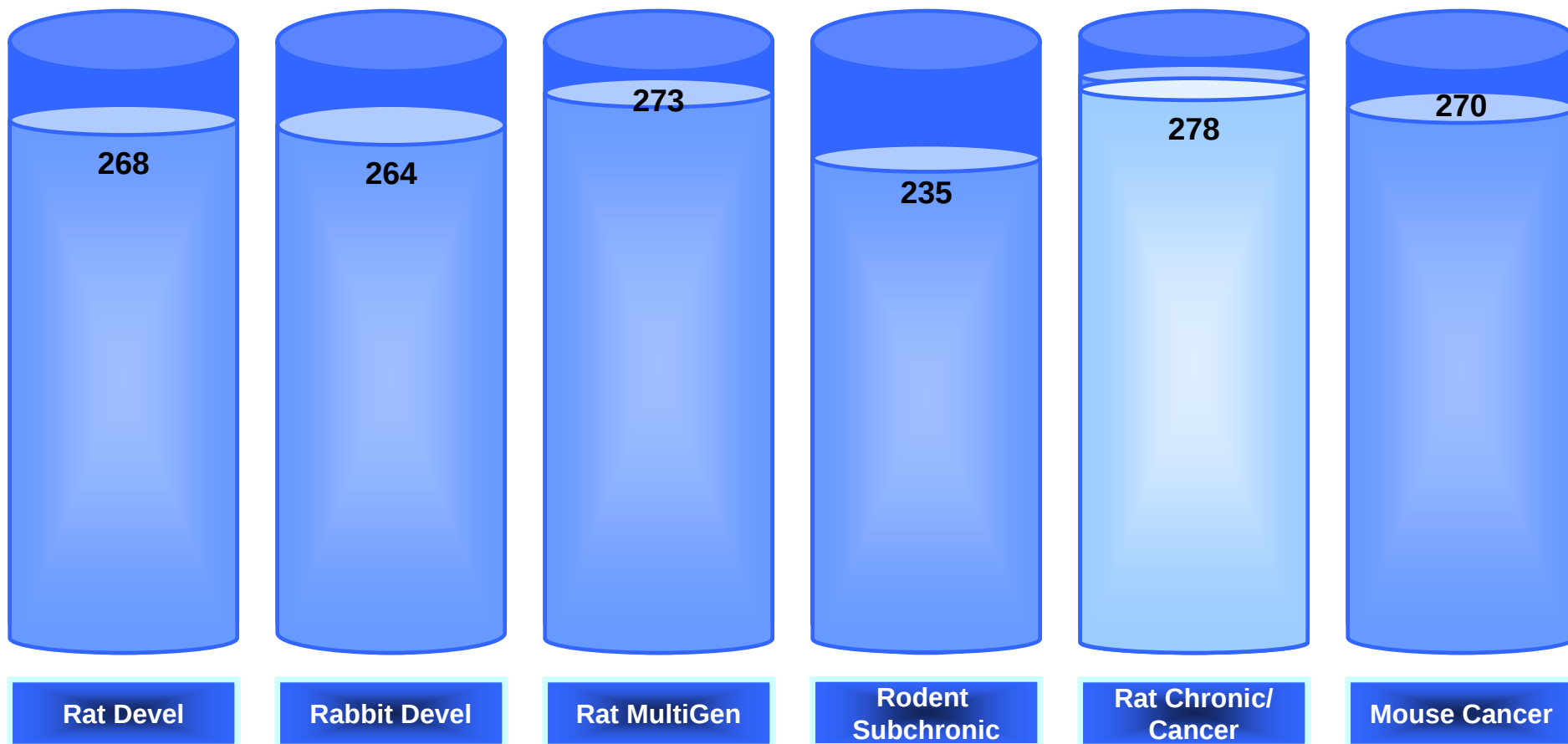
Enter New Study

Filename/
Citation

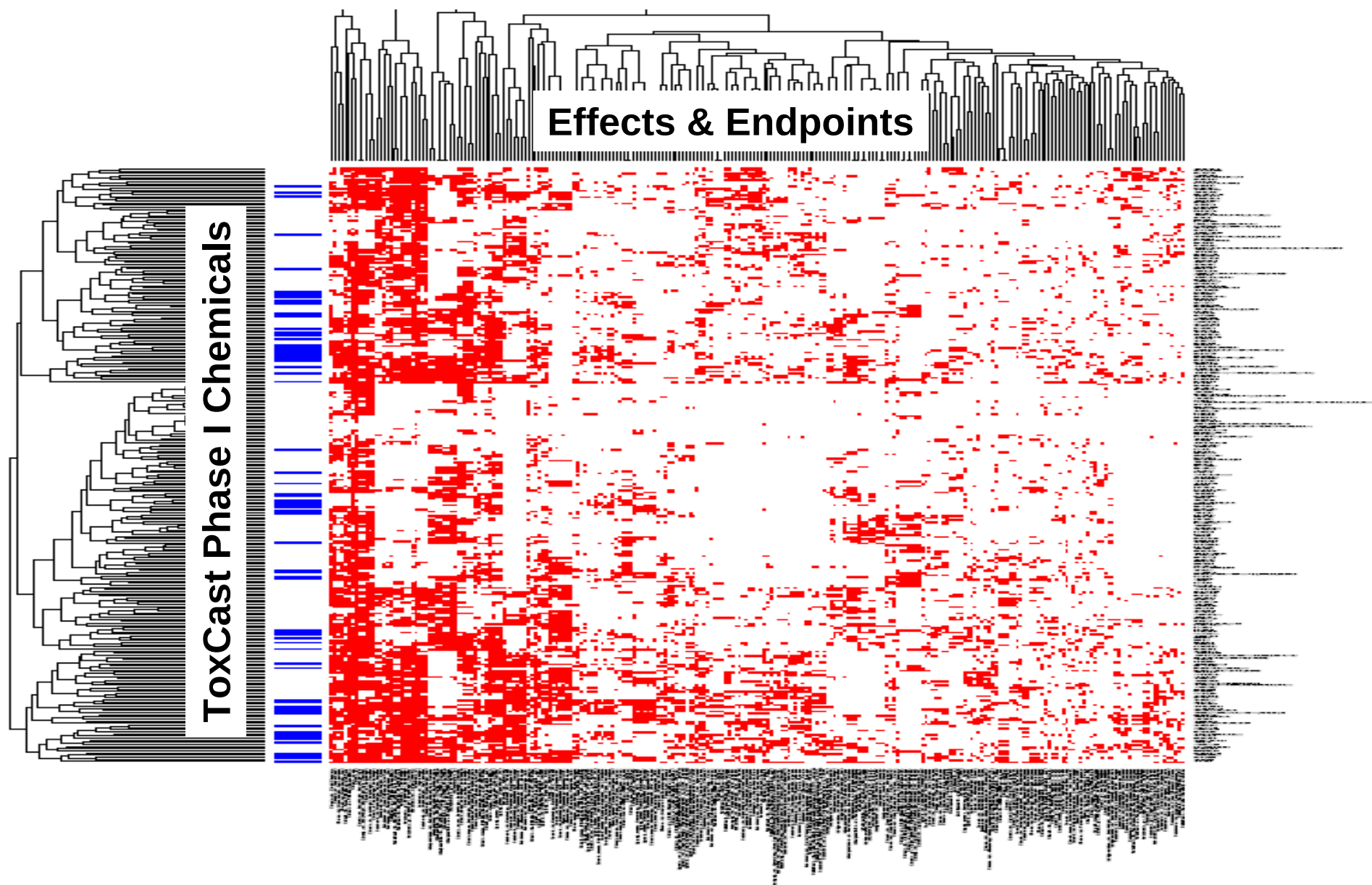
Toggle back to ToxRefDB Switchboard

ToxRefDB Data Entry: Phase I

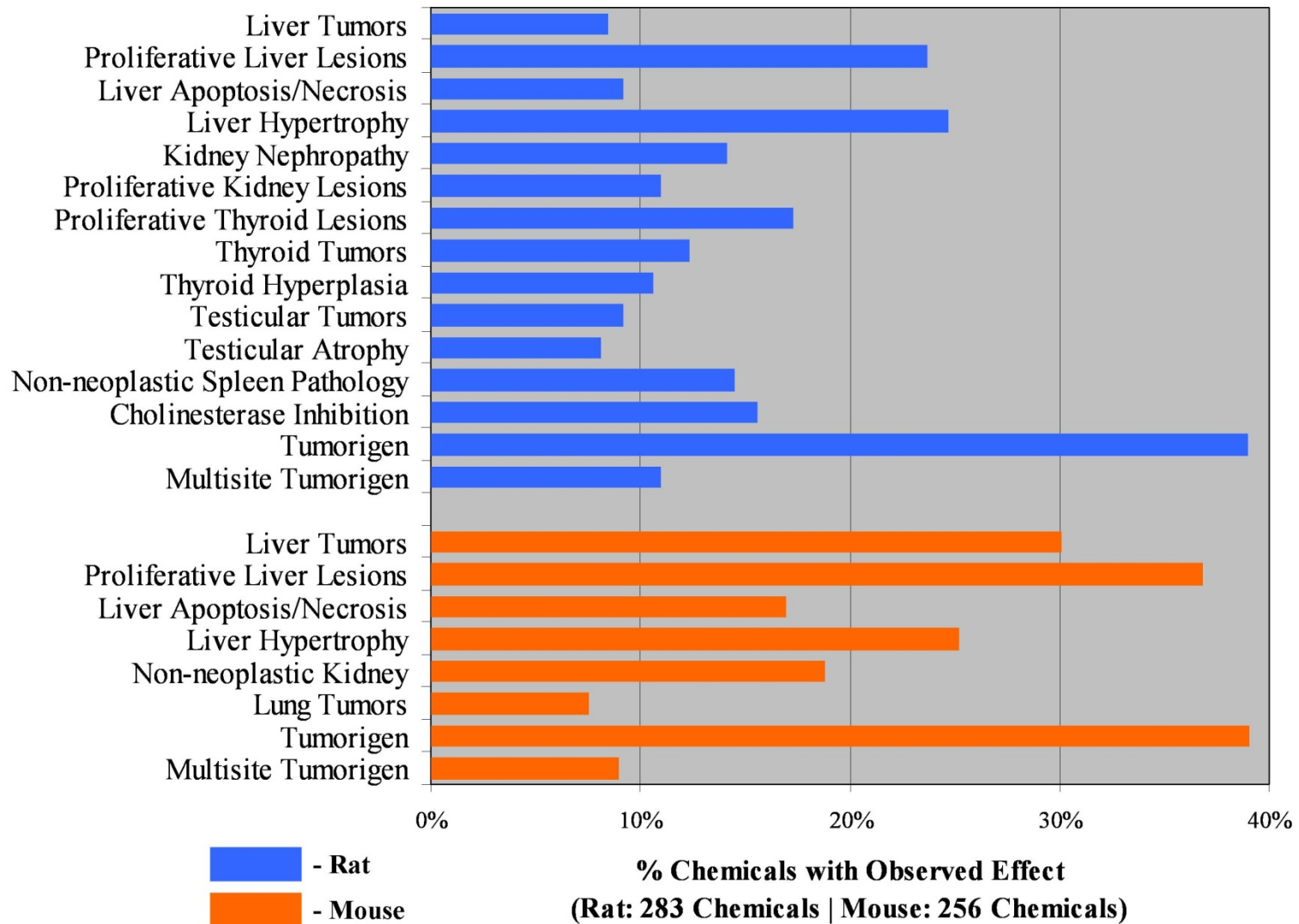
291 Pesticides



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



Common Phenotypes in Chronic Rodent Studies



ToxRefDB Profiling of Liver Effects for Pesticides

Liver non-neoplastic histopathology and increased organ weight are often associated with tumors and cancer

Effect Categories	ClinChem and Hematology						Organ Weight				NonNeoplastic Pathology				NeoPlastic Pathology			
Effects →	ALT	AST	WBC	RBC	Cholesterol	Body Weight	Liver	Kidney	Thyroid Gland	Testes	Liver	Kidney	Thyroid Gland	Testes	Liver Adenoma	Liver Carcinoma	Thyroid Gland	Testes
1H-1,2,4-Triazole, 1-((4-						-1	1	-1			1			1	1			1
2-Cyclohexen-1-one, 2-{1-						-1					1	1		1				
Ametryn						-1					1			1				
Bentazon							1	1	-1		1			1	1			
Cyproconazole	1	1			-1	-1	1				1			1	1	1		
Diclofop-methyl	1	1		-1	-1		1	1	1		1	1		1	1	1	1	1
Difenoconazole	1				1	-1	1	1			1				1	1		
Fenarimol			-1				1				1				1			
Fenbuconazole					1		1		1		1		1				1	
Imazalil	-1	-1		1		-1	1		-1		1		1		1	1	1	
Iprodione		1					1		1	1	1			1	1	1		1
Linuron				-1		-1					1	1			1			1
Mesotrione						-1	1	1			1	1						
Metaxyl	1	1					1				1							
Methoxyfenozide						-1	1	-1		1	1	1	1					
Myclobutanil	1			-1			1	-1		-1	1			1				
Norfurazon							1	1	1		1	1			1	1		
Novaluron							1				1	1						
Penoxsulam								1			1	1						
Permethrin							1	1			1	1						
Picloram	-1						1				1	1						
Primisulfuron-methyl						-1	1	-1			1			1	1	1		
Propiconazole					-1	-1	1				1				1			
						-1	1	-1			1							
						-1	1	1	-1		1							1
						-1	1			1	1		1		1		1	
						-1	1			1	1	1		1				
Acetochlor					1	-1		1				1						
Amdro						-1				-1				1				
Atrazine						-1						1						
Azinphosmethyl																		
Bifenazate			-1	-1	-1	-1		-1							1			
Cyfluthrin						-1												
Deltamethrin						-1												
Diethyl-1-(2,4-dichlorophenyl)-																		
Fenitrothion				-1		-1												
Imazapyr													1					
lambda-Cyhalothrin						-1												
Mesosulfuron-methyl			1	1														
Methyl parathion				-1	1		1	1										
Metolachlor						-1												
Molinate										-1				1				1
Oryzalin				-1		-1	1	1	1	1		1	1				1	
Oxamyl						-1												
Parathion		-1																

Activity Profile is refined "Endpoint" for SAR modeling

ToxCast Contracts for Generating HTS and Genomics Data

Receptors,
Enzymes



Metabolic
Transformation



In vitro
Genomics



Transcription
Factors



Cell
Function

Cell
Signaling



Cell
Function



Alternative
Species

Nine contracts provide for the procurement; hundreds of assays; biochemical, cell-based, and genomic assays; model organisms; and the capacity to screen up to 10,000 chemicals.

ToxCast Phase I Assays/Datasets

- ▶ *NR/transcription factors (Attagene, NCGC)*
- ▶ *Enzyme inhibition/receptor binding HTS (Novascreen)*
- ▶ *Cellular impedance (ACEA)*
- ▶ *Complex cell interactions (BioSeek)*
- ▶ *Hepatocellular HCS (Cellumen)*
- ▶ *Hepatic, renal and airway cytotoxicity (IVAL)*
- ▶ *In vitro hepatogenomics (IVAL, Expression Analysis)*
- ▶ *Zebrafish developmental toxicity (Phylonix)*
- ▶ *Neurite outgrowth HCS (NHEERL)*
- ▶ *Cell proliferation (NHEERL)*
- ▶ *Zebrafish developmental toxicity (NHEERL)*
- ▶ *NR Activation and translocation (CellzDirect)*
- ▶ *HTS Genotoxicity (Gentronix)*
- ▶ *Organ toxicity; dosimetry (Hamner Institutes)*
- ▶ *Toxicity and signaling pathways (Invitrogen)*
- ▶ *C. elegans WormTox (NIEHS)*
- ▶ *Gene markers from microscale cultured hepatocytes (MIT)*
- ▶ *D Cellular microarray with metabolism (Solidus)*
- ▶ *Zebrafish vascular/cardiotoxicity (Zygogen)*

20 Assay Sources
552 Endpoints

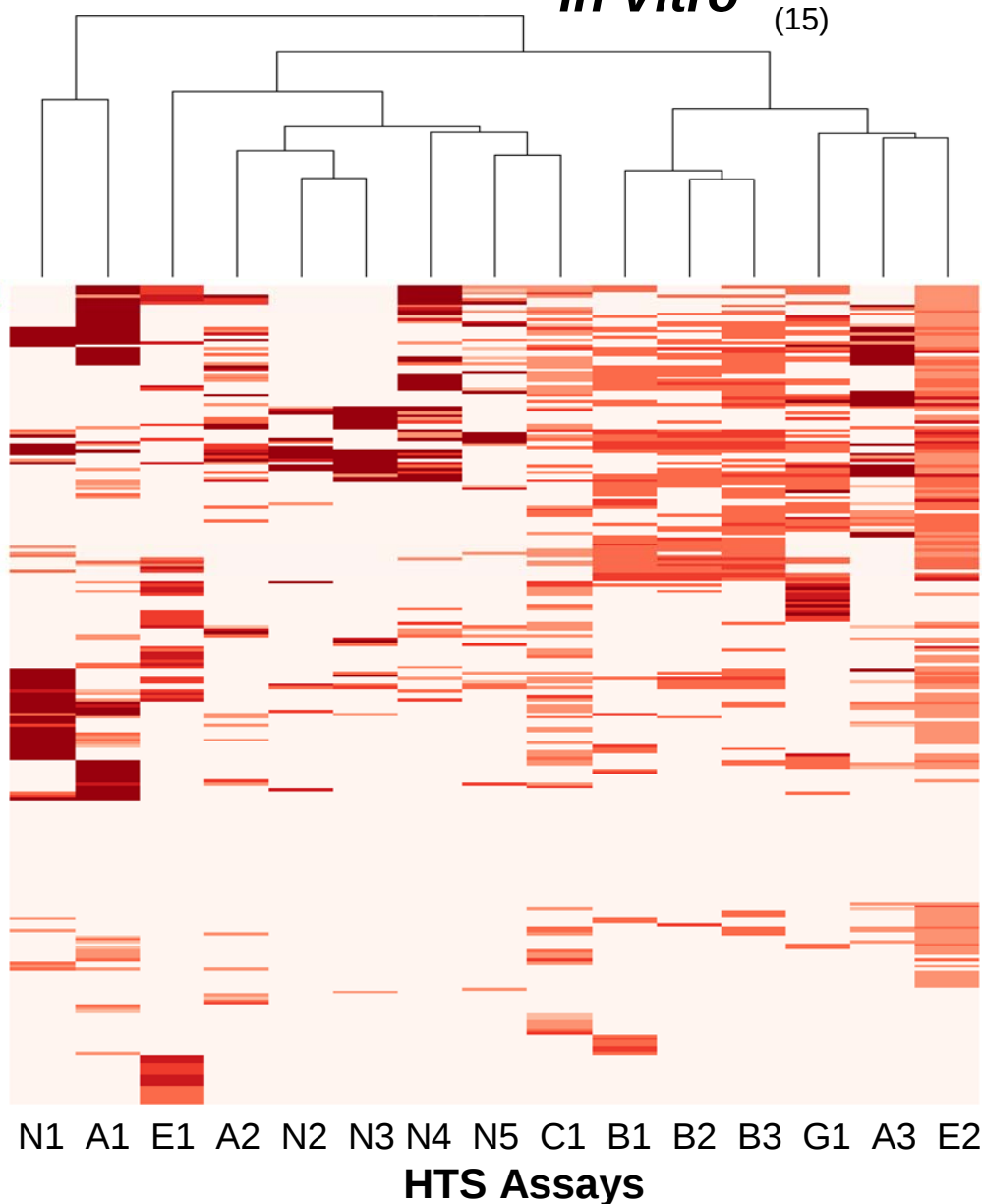
ToxCast Predictive Modeling of Chronic Rat Liver Apoptosis/Necrosis

In Vivo
(23)

Positive
cluster

Negative
cluster

In Vitro
(15)



Methods described in
Judson et al 2008

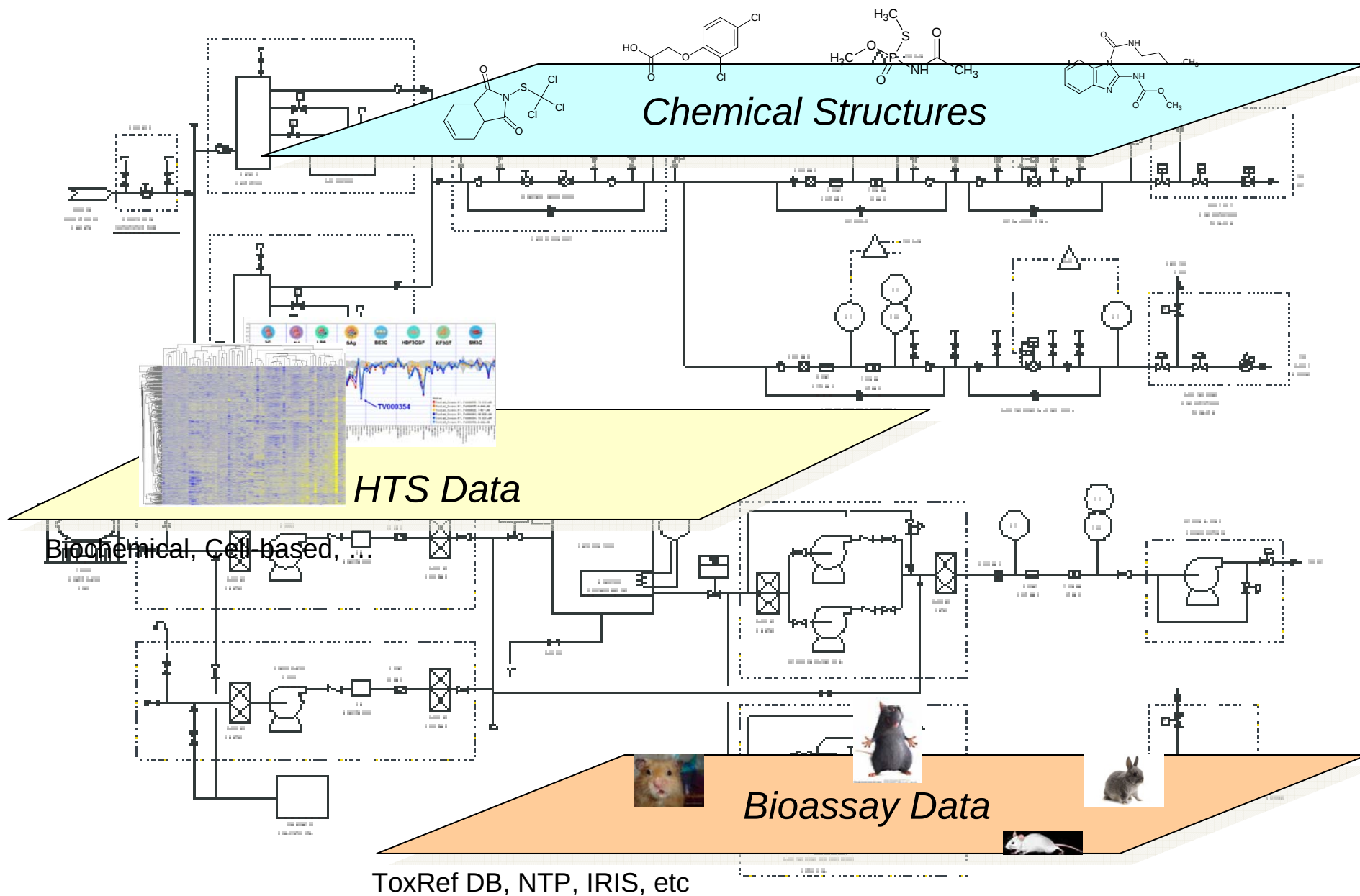
A comparison of machine learning
algorithms for chemical toxicity classification
using a simulated multi-scale data model.

BMC Bioinformatics 9:241

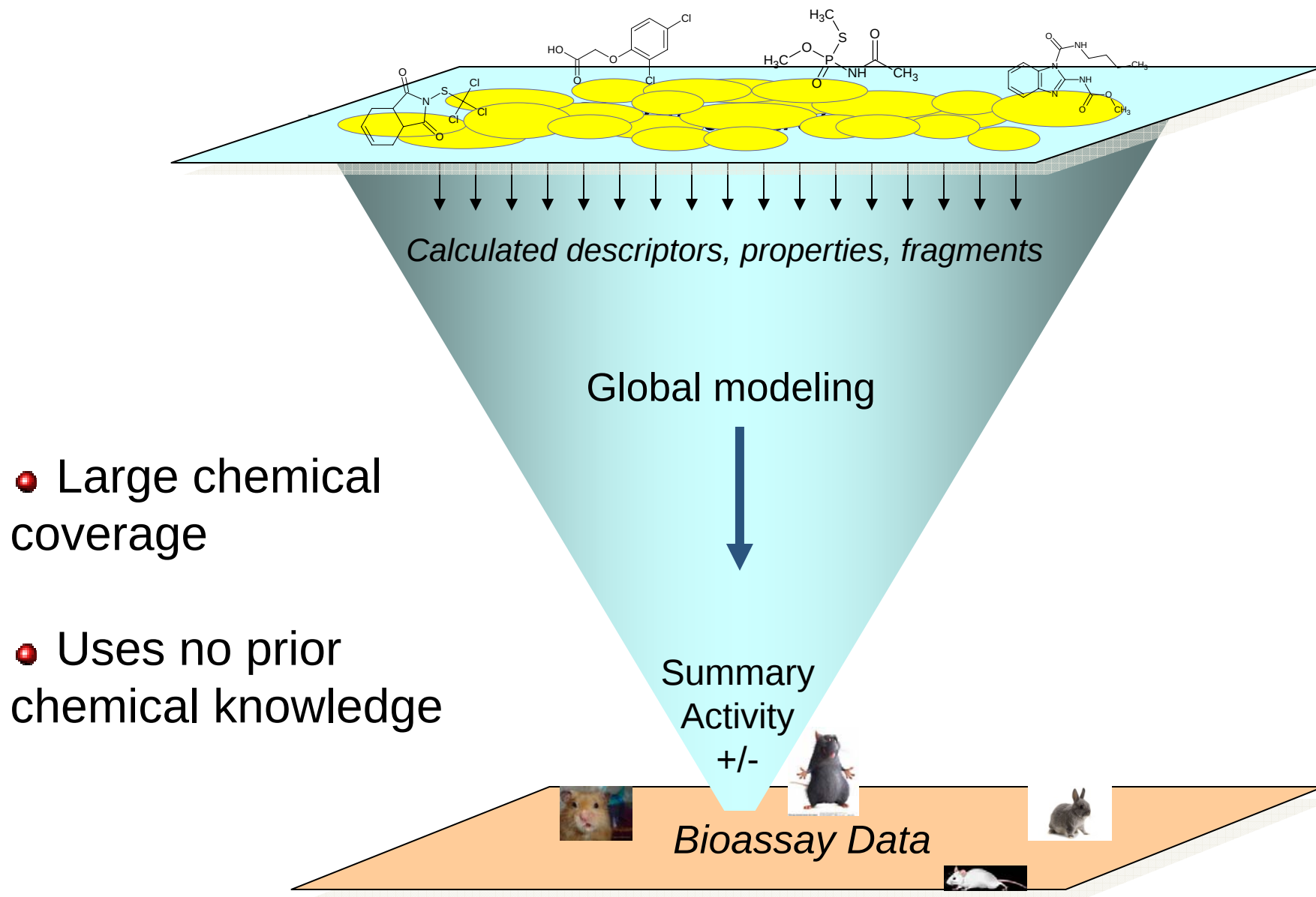
Part III

Incorporating SAR Concepts into ToxCast

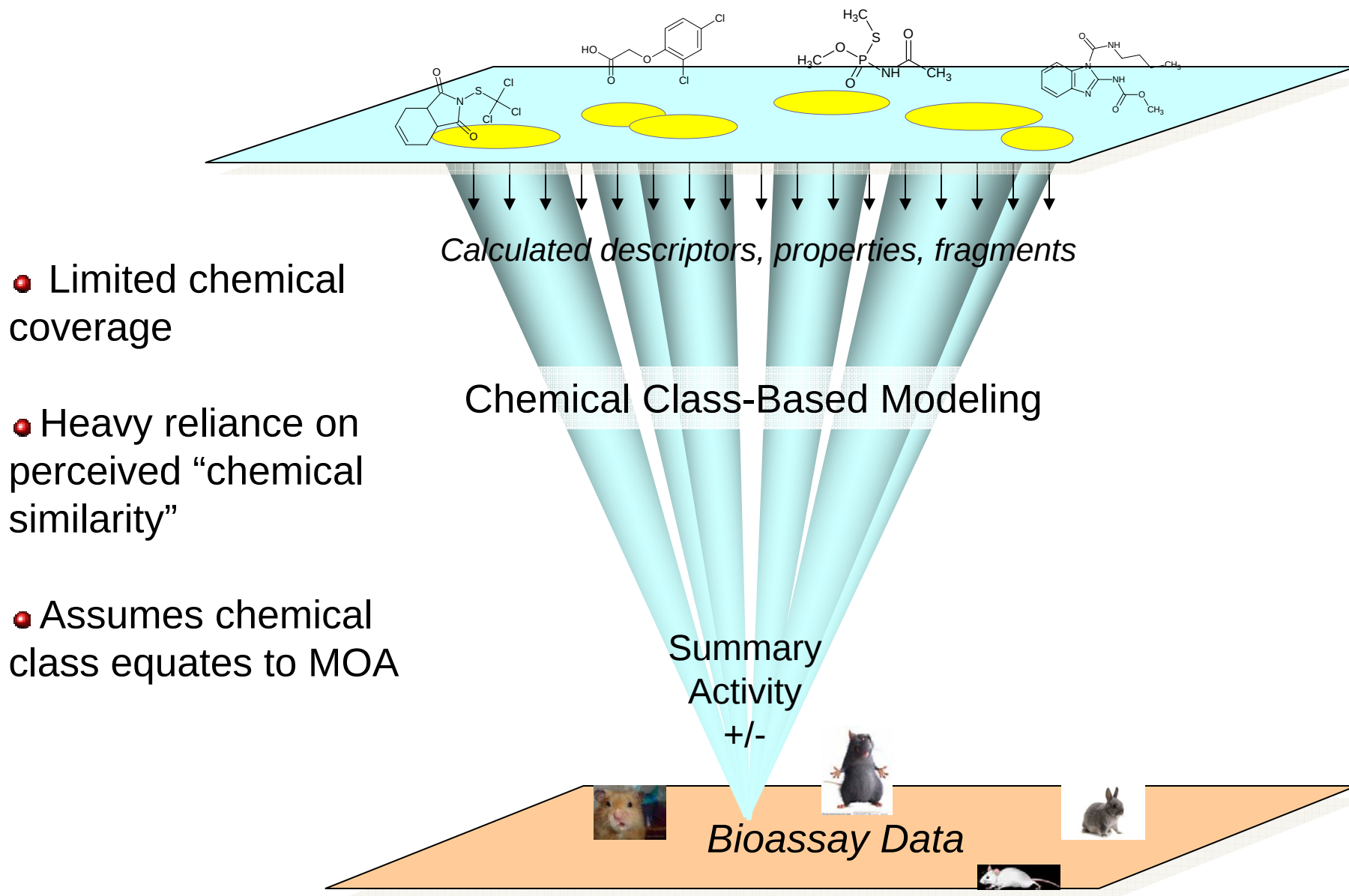
ToxCast: Multidimensional Data



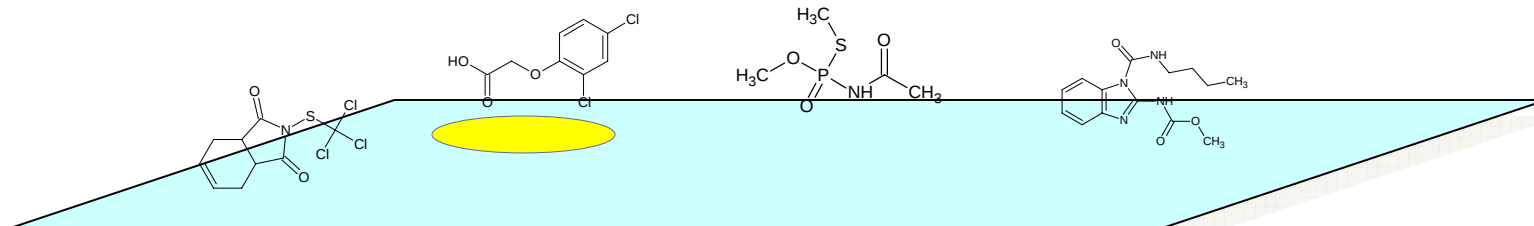
Structure-Activity Approaches to Toxicity Prediction



Structure-Activity Approaches to Toxicity Prediction

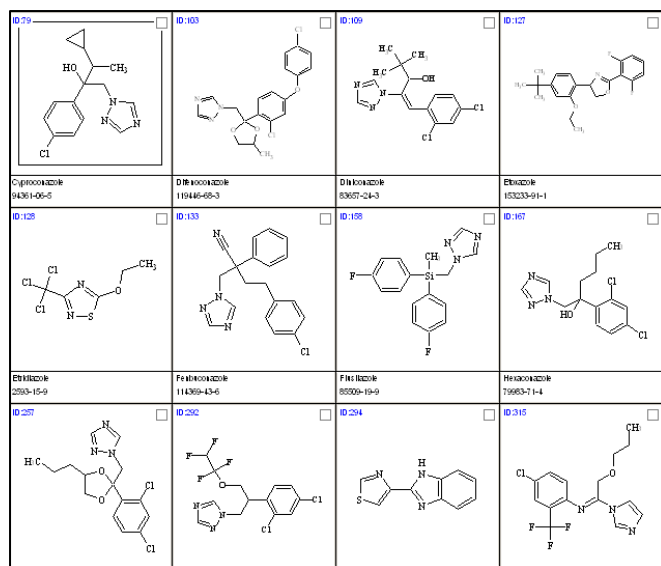


Bioactivity Profile of Structure Class



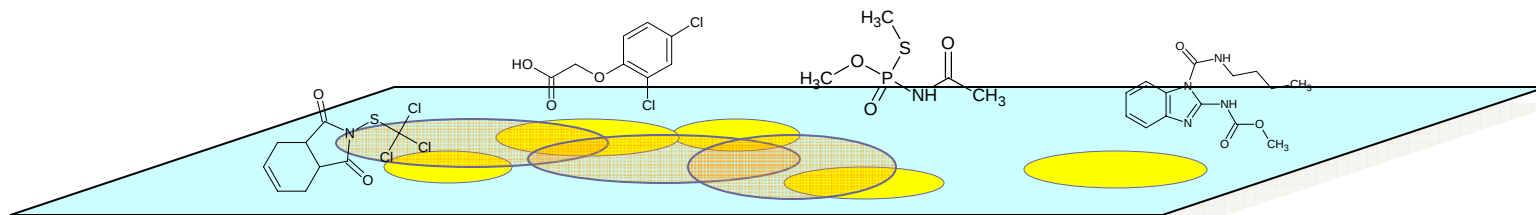
Sample HTS Results for Conazoles

- Patterns can inform SAR
- Chemical structures can suggest basis for activity differences



NAME	CYP2C19	CYP2C9	CYP3A1	Dopamine Transporter (Human)	CYP2D2	Androgen Receptor	Dopamine Transporter (Rat)	CYP2B6	CYP2D1	CYP3A4	Progesterone Receptor	Benzodiazepine Receptor
Cyproconazole	1	1	1	1	1	0	1	0	0	1	0	0
Difenconazole	1	1	1	1	1	0	0	1	1	0	0	0
Diniconazole	1	1	1	0	1	0	0	0	1	1	1	0
Fenbuconazole	1	1	0	0	0	0	0	0	0	1	0	0
Flusilazole	1	1	1	0	1	1	0	1	1	NA	1	1
Hexaconazole	1	1	1	1	1	0	1	1	1	NA	1	0
Imazalil	1	1	1	1	1	1	1	1	1	1	1	1
Myclobutanil	1	1	1	1	0	0	0	0	0	NA	0	0
Paclobutrazol	1	0	1	1	0	1	1	0	1	1	0	0
Prochloraz	1	1	1	1	1	1	1	1	1	NA	1	1
Propiconazole	1	1	1	0	0	0	0	1	0	NA	0	1
Tetraconazole	1	1	1	0	1	1	0	1	0	1	1	0
Triadimefon	1	1	0	1	1	1	1	0	0	1	0	1
Triadimenol	1	0	0	1	0	1	1	0	0	0	0	0
Triflumizole	1	1	1	1	1	1	0	1	1	1	1	1
Triticonazole	1	1	1	1	0	1	1	0	0	NA	0	0
Totals	16	14	13	11	10	9	8	8	8	8	7	6

Structure Class vs Bioactivity Class

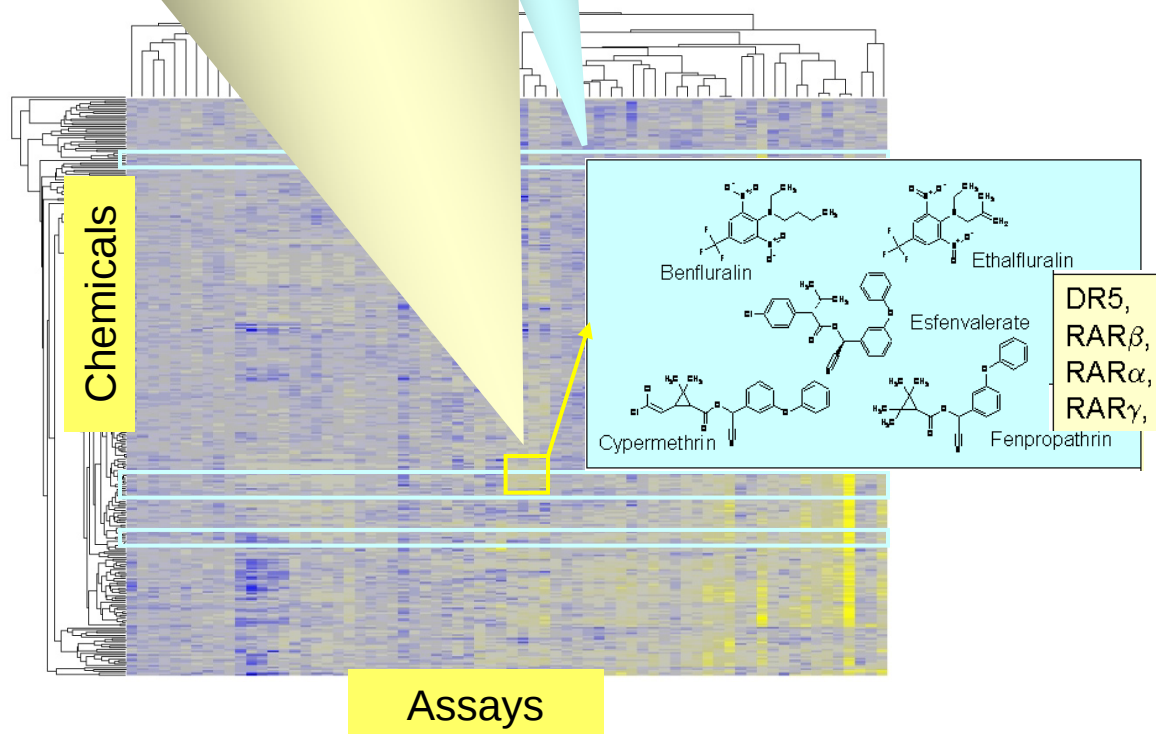


Chemical structure class:

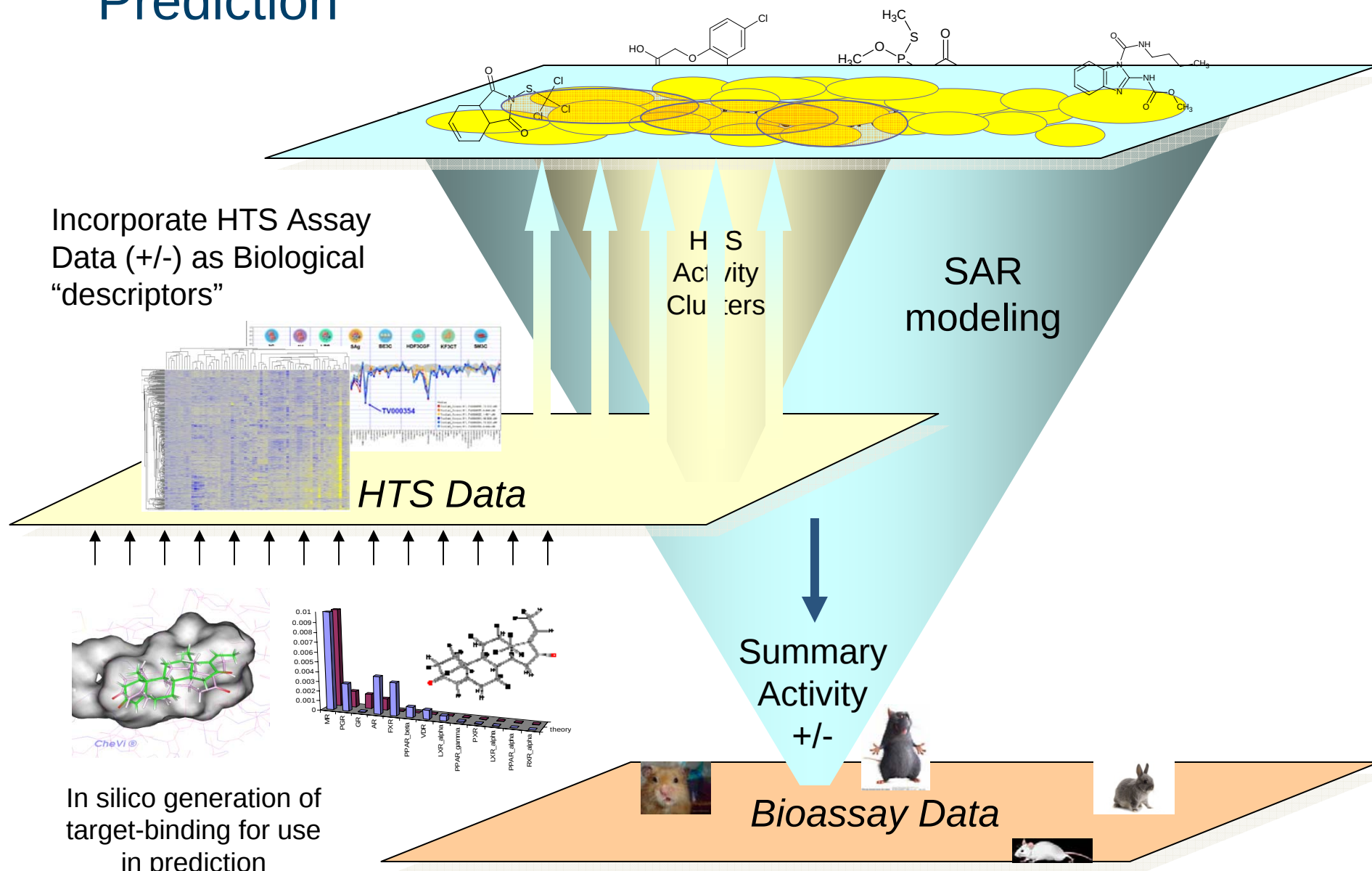
- Cluster according to activity and mechanism
- Differences in activity profiles can discriminate within structure class

Bioactivity profile class:

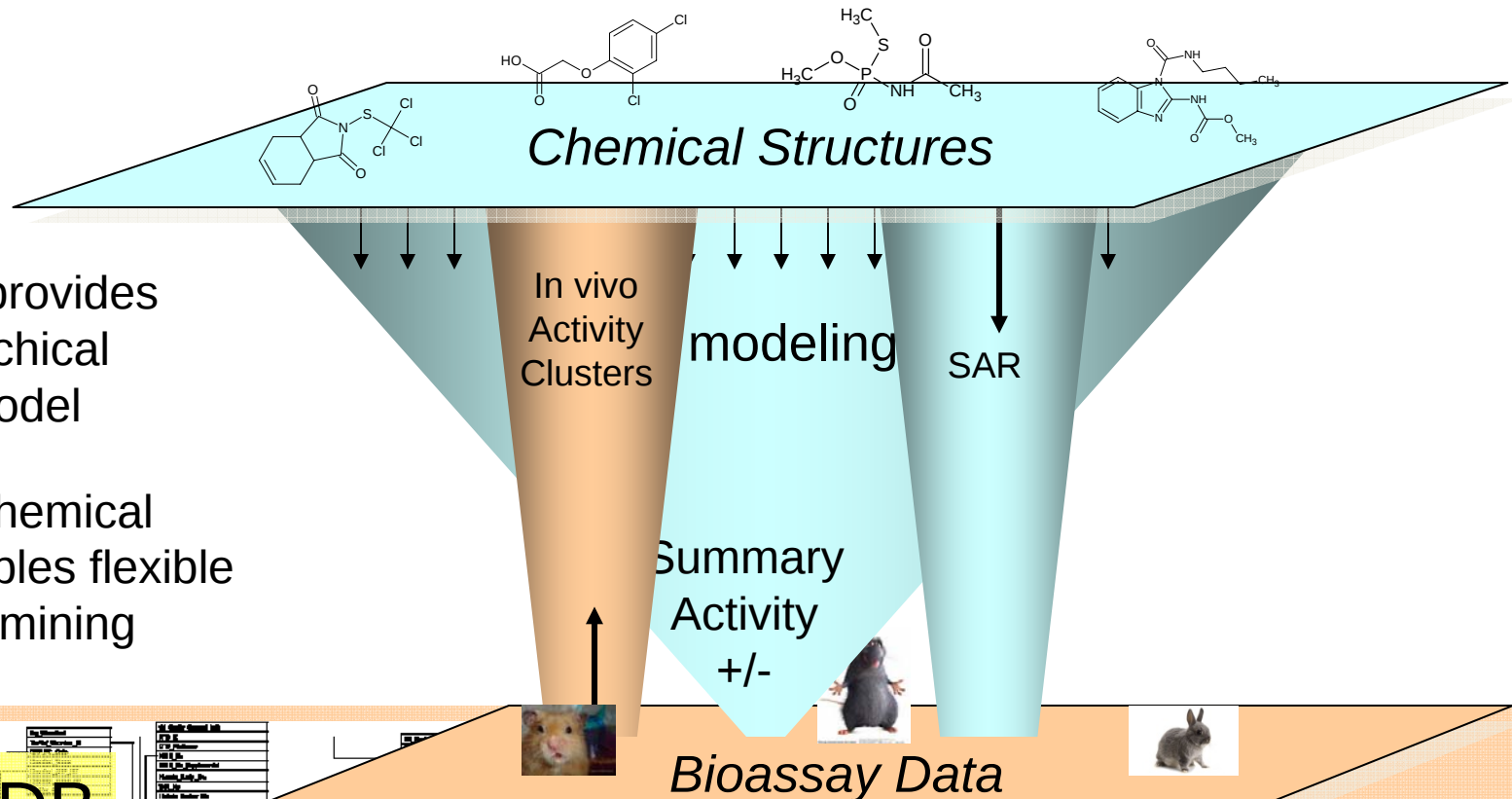
- Can project onto multiple chemical classes
- Potentially broader coverage of chemical space
- Implies mechanistic similarity



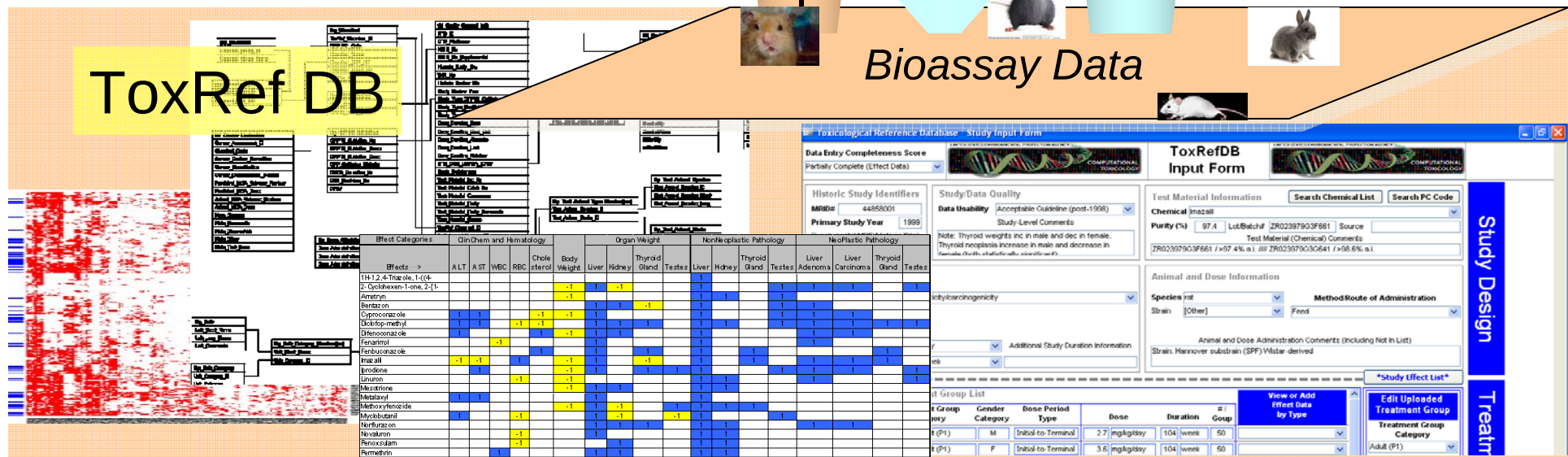
Structure-Activity Approaches to Toxicity Prediction



Use of Bioassay Activity Categories in SAR



- ToxRef DB provides detailed hierarchical toxicity data model
- Linkage to chemical structures enables flexible SAR and data mining

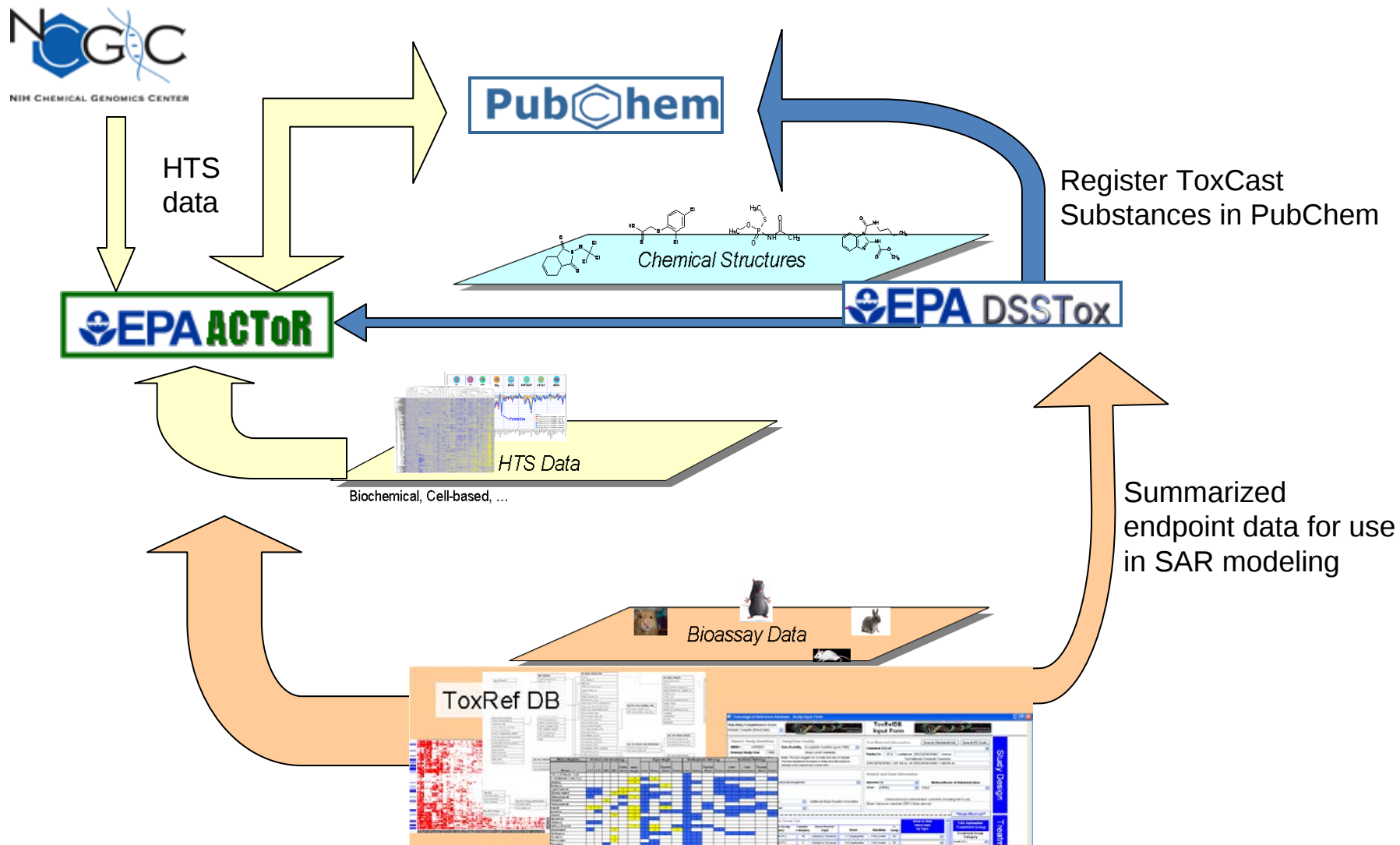


Structure-Activity Approaches

In a Brave New World -

- ▶ What in vivo “endpoints” do we use?
- ▶ How do we combine SAR & biological “descriptors”?
- ▶ How do we incorporate mechanism and pathway information?
- ▶ How do we assess biological/chemical similarity?
- ▶ How do we evaluate “applicability domain” for predictive profiles?
- ▶ How do we use biological information to improve upon SAR approaches?

ToxCast: Data Publication & Exploration



ToxCast_320

Bioactivity Analysis:

Retrieve all bioassay data in PubChem for ToxCast_320

482 Bioassays
45 Compounds

NCBI PubChem

PubChem BioAssay [GO]
PubMed | Entrez | Structure | PubChem | Help

PubChem » BioAssay Services » BioActivity Analysis: Summary

BioActivity Analysis: 482 Bioassays (473 Tested) and 45 Compounds

Summary | DataTable | Structure-Activity

Compounds: 45 (9 shown) [?]
Revise Compound Selection:
[Select Active](#)
[Add Active](#)
[Add Tested](#)
[Add Similar Compounds](#)

BioAssays: 482 [?]
Revise BioAssay Selection:
[Select Active](#)
[Select Tested](#)
[Add Active](#)
[Add Tested](#)
[Add Related BioAssays](#)
[Selected BioAssays](#)

Chemical structures of 9 compounds are displayed in a 3x3 grid.

Total Pages: 24 Display: 20 Go To Page 1 [Navigation buttons]

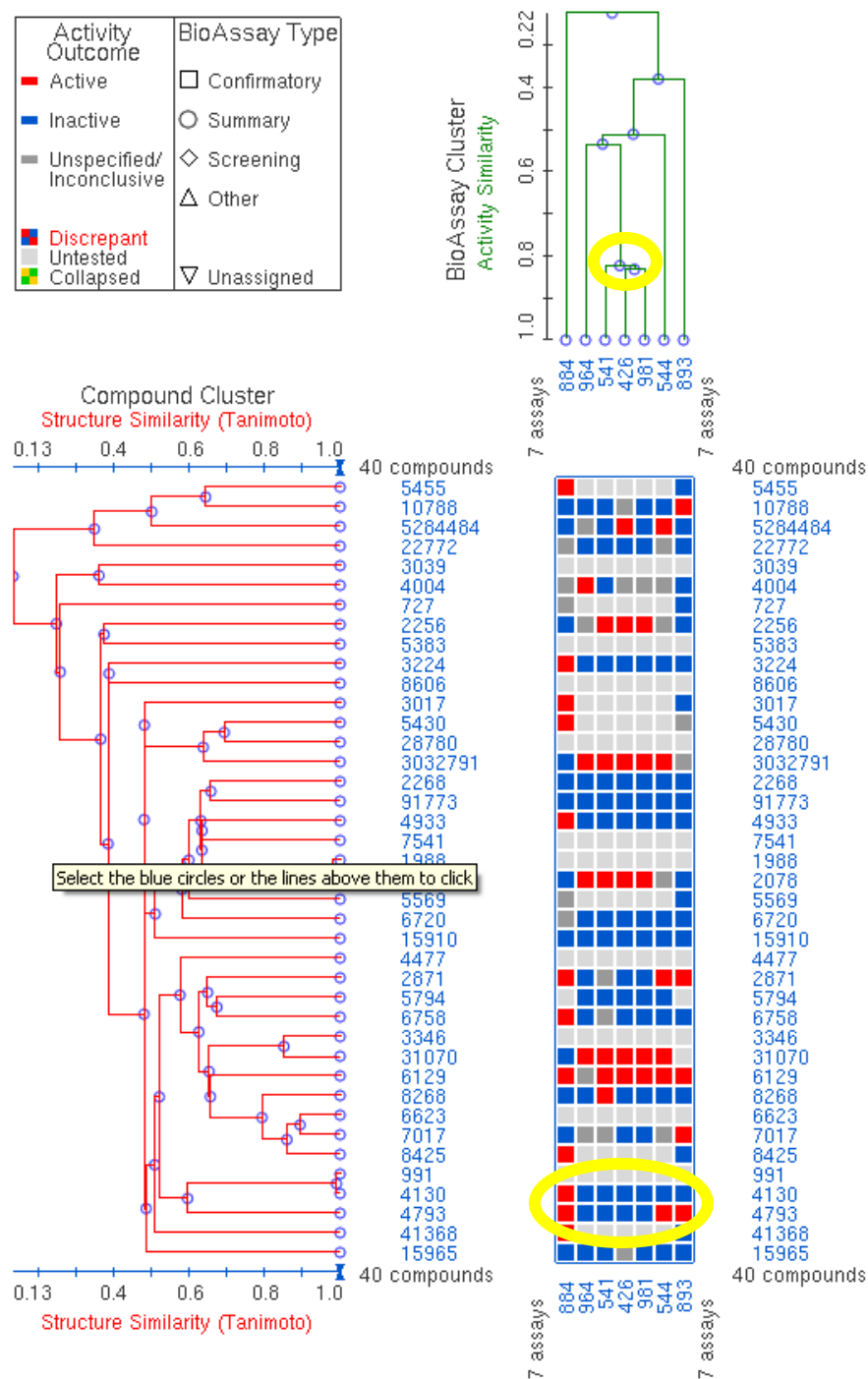
#	☐	AID	Active	Inactive	Discrepant	Tested	Outcome Method	Name
1	☒	884	16	15	3	39	Confirmatory	qHTS Assay for Inhibitors and Substrates of Cytochrome P450 3A4
2	☒	544	9	15	1	30	Confirmatory	Cell Viability - SH-SY5Y
3	☒	541	7	20		30	Confirmatory	Cell Viability - NIH 3T3
4	☒	426	7	20	1	30	Confirmatory	Cell Viability - Jurkat
5	☒	964	6	20	1	30	Confirmatory	Cell Viability - LYMP1-003 - Assay at 40 hr
6	☒	981	6	23	1	30	Confirmatory	Cell Viability - LYMP2-010
7	☒	893	6	20	1	30	Confirmatory	qHTS Assay for Inhibitors of HSD17B4, hydroxysteroid (17-beta) dehydrogenase 4
8	☐	167	6	20	1	30	Confirmatory	NCI Yeast Anticancer Drug Screen. Data for the bub3 strain
9	☐	165	5	6	1	10	Confirmatory	NCI Yeast Anticancer Drug Screen. Data for the cln2 rad14 strain

Selected bioassays

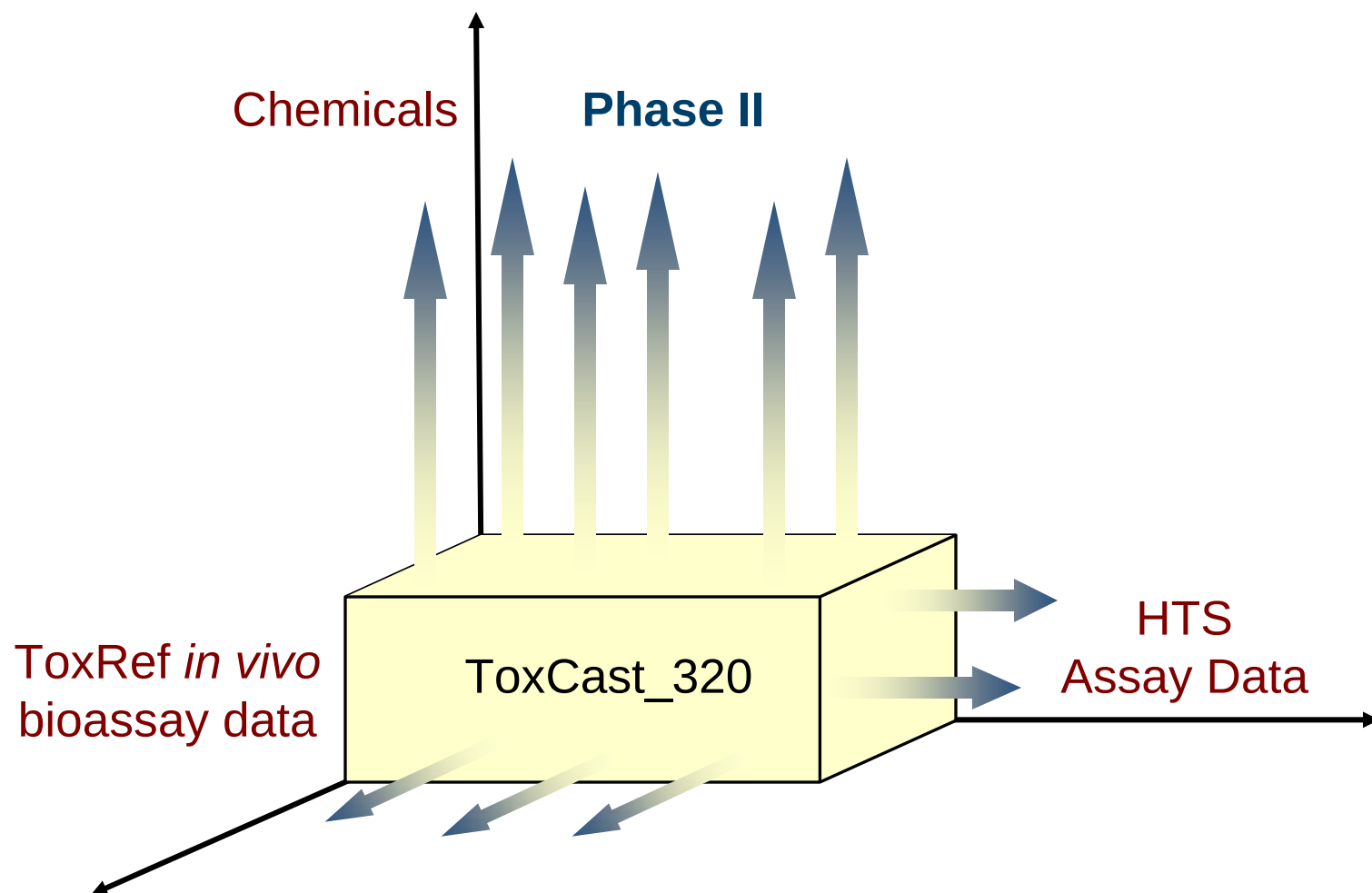
Structure-Activity Bioactivity Analysis:

7 bioassays, 45 Actives

View Bioassay Profile
by Structure Similarity
Cluster



ToxCast Phase I: Proof of Concept



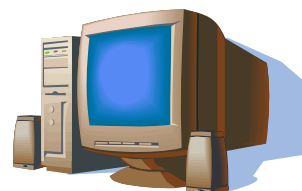
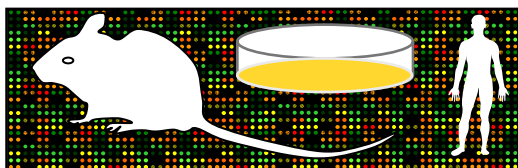
Phased Development of ToxCast

Phase	Number of Chemicals	Chemical Criteria	Purpose	Number of Assays	Cost per Chemical	Target Date
I	320	Data Rich (pesticides)	Signature Development	>400	\$20k	FY07-08
Ila	>300	Data Rich Chemicals	Validation	>400	\$15-20k	FY09
Ilb	>100	Known Human Toxicants	Extrapolation	>400	\$15-20k	FY09
Ilc	>300	Expanded Structure and Use Diversity	Extension	>400	\$15-20k	FY10
III	Thousands	Data poor	Prediction and Prioritization	???	\$10-15k	FY11-12

- Affordable science-based system for categorizing chemicals
- Increasing confidence as database grows
- Identifies potential mechanisms of action
- Refines and reduces animal use for hazard ID and risk assessment



Tox21 Collaboration



National Health and
Environmental Effects Research Laboratory

National Center for
Computational Toxicology

- Combined HTS plates (2x1408) high interest chemicals
- Joint assay development
- Use of NCGC HTS informatics capabilities



National Toxicology Program
Department of Health and Human Services



NIH CHEMICAL GENOMICS CENTER

Biomolecular Screening Branch

Toxicology Project Team




All Database

Search PubChem

About Entrez

Entrez Help

PubChem

Help | FAQ

Courses

PubChem Substance

Structures supplied by depositors

PubChem Compound

Unique structures with computed properties

PubChem BioAssay

Bioactivity assay results supplied by depositors

PubChem Structure Search

PubChem FTP

PubChem Text Search

PubChem BioAssay

PubChem Substance

PubChem Compound

PubChem BioAssay

ntp

GO

the biological activities of small molecules

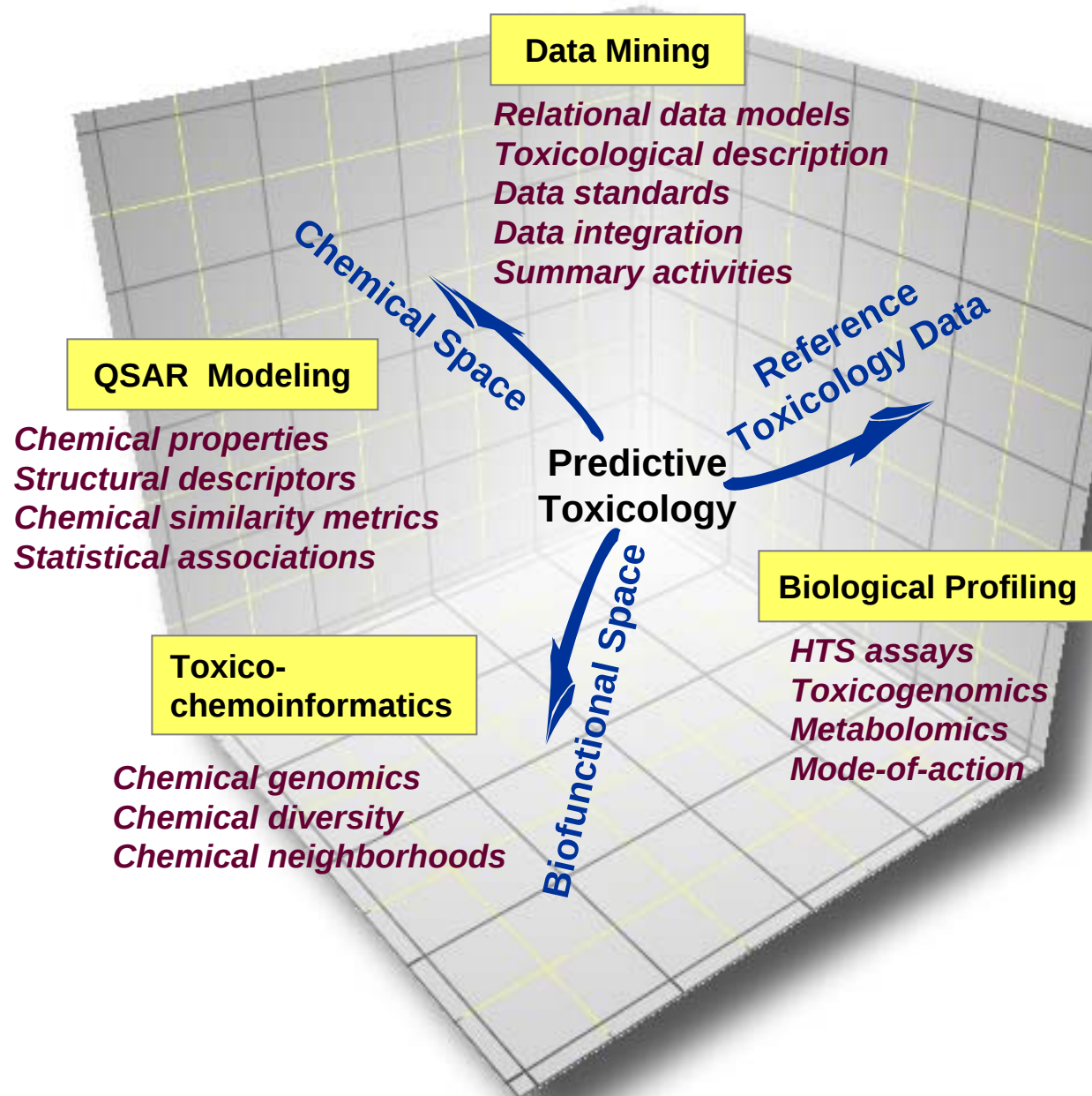
1. AID 434: Cell Viability – MRC5	102 /2816	Active
2. AID 421: Cell Viability – BJ	104 /2816	Active
3. AID 667: Cellular Toxicity (caspase-3) Renal Proximal Tubule	8 /1408	Active
4. AID 666: Cellular Toxicity (caspase-3) NIH 3T3	12 /1408	Active
5. AID 665: Cellular Toxicity (caspase-3) N2a	7 /1408	Active
6. AID 664: Cellular Toxicity (caspase-3) Hek293	18 /1408	Active
7. Cellular Toxicity (caspase-3) H-4-II-E	20 /1408	Active
8. Cellular Toxicity (caspase-3) SK-N-SH	20 /1408	Active
9. Cellular Toxicity (caspase-3) Mesangial	8 /1408	Active
10. AID 659: Cellular Toxicity (caspase-3) NIH 3T3	12 /1408	Active
11. AID 658: Cellular Toxicity (caspase-3) N2a	7 /1408	Active
12. Cellular Toxicity (caspase-3) SHSY5Y	10 /1408	Active
13. AID 656: Cellular Toxicity (caspase-3) HUV-EC-C		
14. AID 655: Cellular Toxicity (caspase-3) Jurkat		
15. AID 654: Cellular Toxicity (caspase-3) HepG2		
16. AID 544: Cell Viability – SH-SY5Y		
17. AID 435: Cell Viability – SK-N-SH		
18. AID 433: Cell Viability – HepG2		
19. AID 427: Cell Viability – Hek293		
20. AID 426: Cell Viability – Jurkat		
21. AID 541: Cell Viability – NIH 3T3		
22. AID 546: Cell Viability – Mesenchymal		
23. AID 545: Cell Viability – Renal Proximal Tubule		
24. AID 543: Cell Viability – H-4-II-E		
25. AID 542: Cell Viability – HUV-EC-C		
26. AID 540: Cell Viability – N2a		
27. AID 559: RNA polymerase		

As of Sept 20, 2008:

Data for 65 assays (1408+ chem) available for download

Keyword search: PubChem Bioassay> “ntp ncgc”

DSSTox CID.txt file with instructions available on NTPHTS download page



Acknowledgements:

✦ **EPA 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)

✦ **NIEHS/NTP – NTP HTS Program:**

Ray Tice, Cynthia Smith, Beth Bowden, Jennifer Fostel

✦ **Toxicogenomics:** ClarLynda Williams (EPA)

✦ **PubChem Project:** Steve Bryant, Yanli Wang

✦ **Carcinogenic Potency Project:** Lois Gold

✦ **Collaborators:** Alex Tropsha, Hao Zhu, Chihae Yang, Antony Williams