Advances in Toxico-Cheminformatics:
Supporting a new paradigm for predictive toxicology

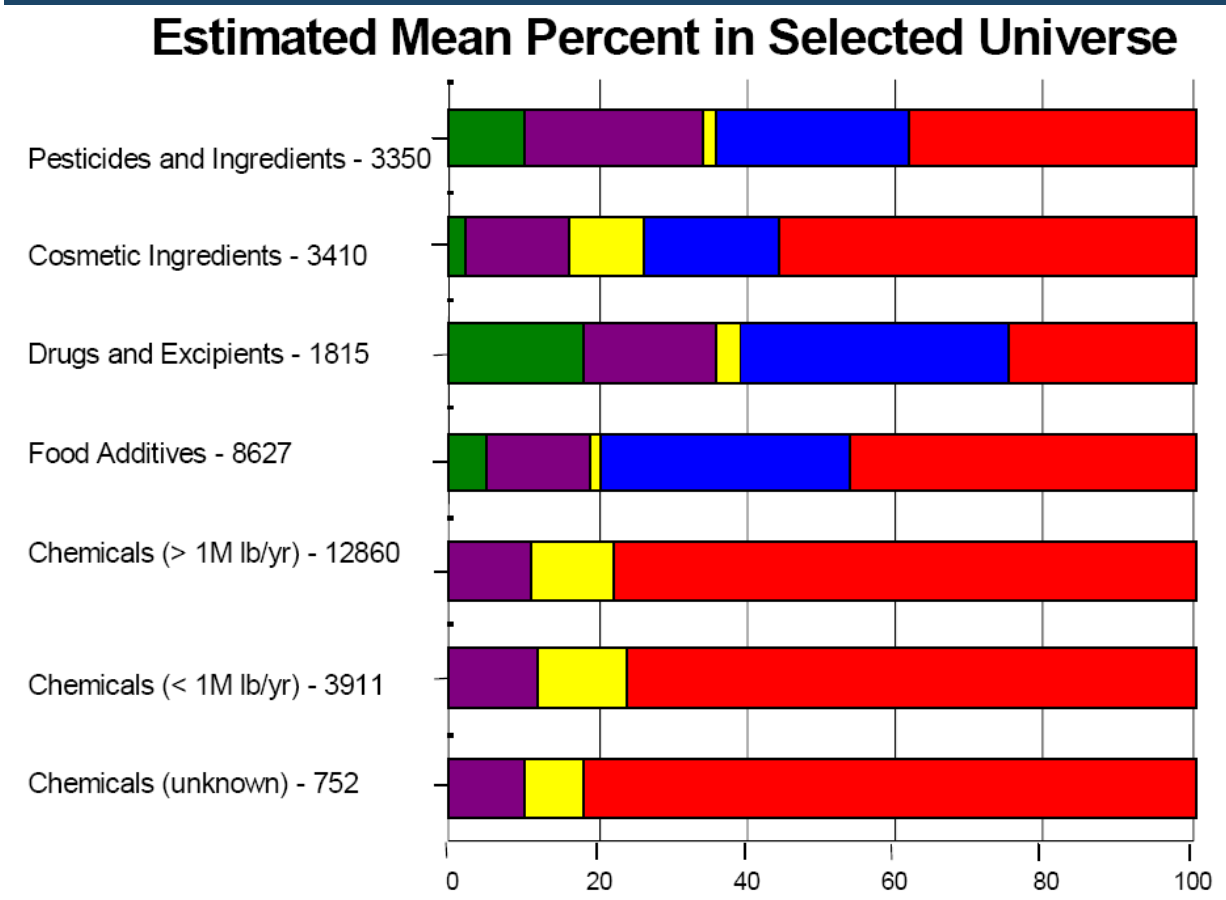
21-26 September 2008
Euro-QSAR 2008, Uppsala, Sweden

Ann Richard
richard.ann@epa.gov

United States Environmental Protection Agency

Office of Research and Development
National Center for Computational Toxicology

Environmental Chemicals: Toxicity Assessment Data Gaps



Complete Hazard Assessment Possible

Minimal Information Available

No Toxicity Information Available

Partial Hazard Assessment Possible

Some Toxicity Information Available

Strategies for Closing the Chemical Data Gap

by John S. Applegate and Katherine Baer

A Center for Progressive Reform Publication *April 2006*

Chemical Names
CAS Registry Nos.
SMILES
Ecotox, Aster, Teratox
TSCA Substance Registry List
High Production Volume Information System
Green Chemistry
Chemical Structures
Drinking Water Contaminants
Pesticide Actives
Toxico-genomics
Pesticide Other Ingredients
IRIS
Air
EPA Chemical Data Silos

Chemical Research in Toxicology

JULY 2007
VOLUME 20, NUMBER 7

Editorial

The Future of Toxicology—Wrap Up

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

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

U.S. ENVIRONMENTAL PROTECTION AGENCY

Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

Bookmark

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

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

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

- 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>

EPA DSSTox Structure-Browser GO v2.0

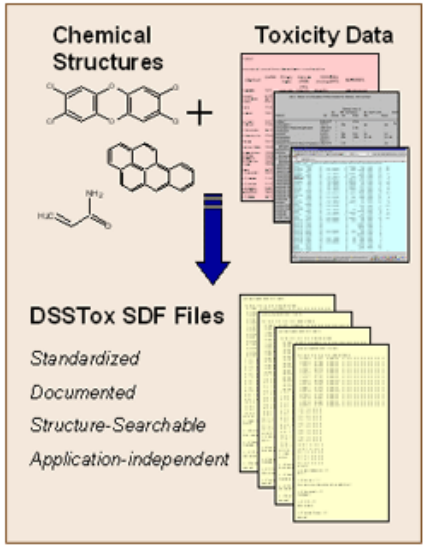
DSSTox Structure-Browser information Page

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



- DSSTox Graphic Flowchart
- DSSTox Project Goals
- DSSTox Publications

SSTox Data Files: Details>

PDBAS_v5c_1547_29Apr2008
BPCAN_v4b_209_15Feb2008
PAFHM_v4b_617_15Feb2008
DAMDD_v3b_1216_15Feb2008
PVCSI_v2c_3548_15Feb2008
PVISD_v1b_1006_15Feb2008
ISTR_v1b_544_15Feb2008
CTRER_v4b_232_15Feb2008
TPBSI_v4a_2315_15Sep2008 *update
TPHTS_v2b_1408_15Feb2008
DXCST_v2c_320_29Apr2008

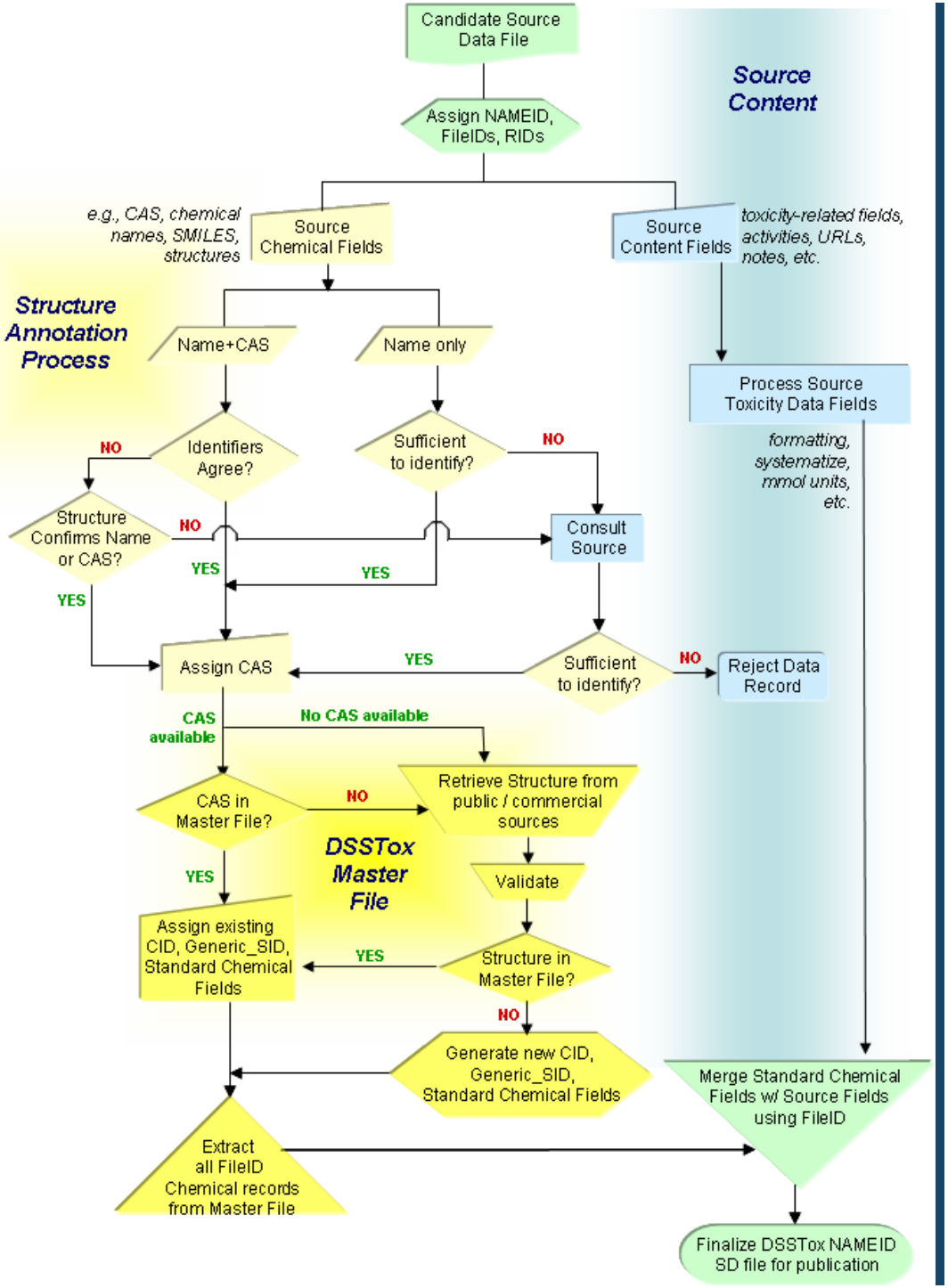
>Other (non-DSSTox) Data Files

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

NAMEID	version #records date	Expanded DSSTox Data File Title & Description
CPDBAS	v5c 1547 29Apr2008	**Carcinogenic Potency Database Summary Tables - All Species:** 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.
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 Data 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 for HTS bioassay testing and to PubChem (PubChem_CIDs and PubChem_SIDs included in NTPHTS_v2a file); NCGC HTS bioassay data are being deposited into PubChem and can be retrieved with these PubChem chemical CID and SID record listings.
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. File will be updated with links to PubChem CIDs and SIDs for retrieving assay data, and with updates to chemical inventory as Program moves to Phase II and beyond.

DSSTox Chemical Quality Assurance Procedures:

- Chemical identification
- Structure annotation
- Internal consistency
- Quality review



CPDBAS_v5b_1547 (61 fields)

STRUCTURE
DSSTox_RID
DSSTox_CID
DSSTox_Generic_SID
DSSTox_FileID
STRUCTURE_Formula
STRUCTURE_MolecularWeight
STRUCTURE_ChemicalType
STRUCTURE_TestedForm
_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

Mutagenicity_SAL_CPDB
TD50_Rat_mg
TD50_Rat_mmol
TD50_Rat_Note
TargetSites_Rat_Male, Female, 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
ToxicityNote
NTP_TechnicalReport
ChemicalPage_URL

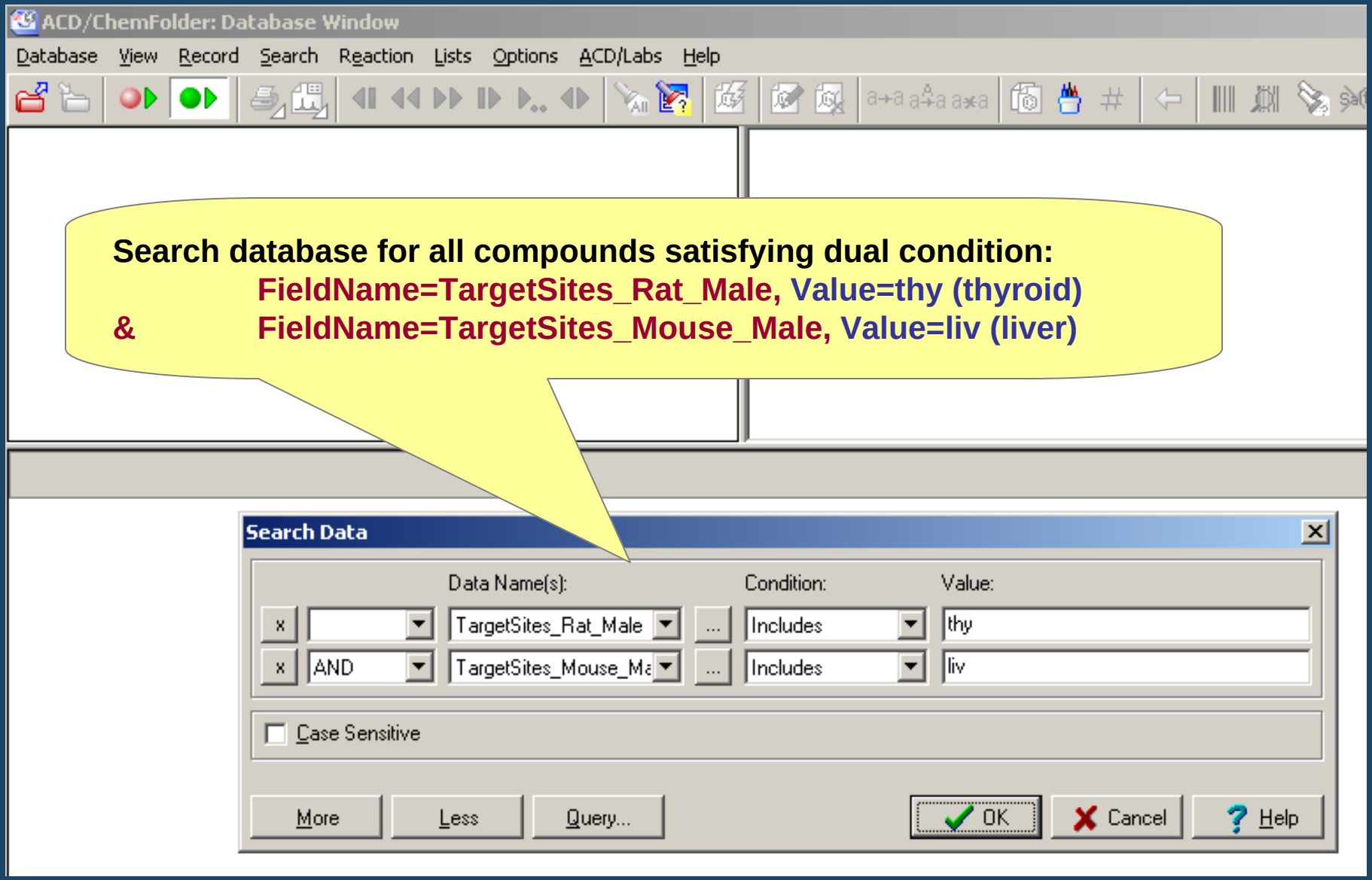
active
inactive
inconclusiv
e

Potency
Ranking [1-100]

multisite
multisex
multispecies

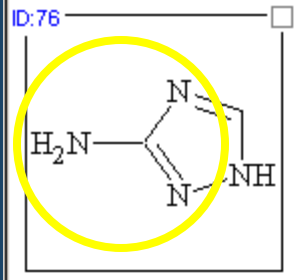
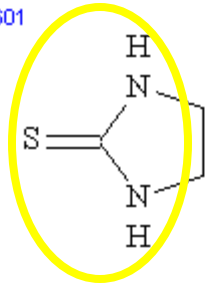
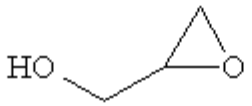
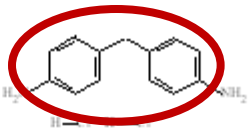
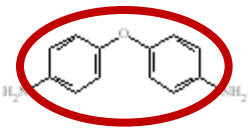
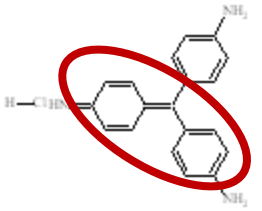
adrenal gland;
bone;
clitoral gland;
esophagus;
ear/Zymbal's gland;
gall bladder;
harderian gland;
hematopoietic system;
kidney;
large intestine;
liver;
lung;
mesovarium;
mammary gland;
mixture;
myocardium;
nasal cavity
nervous system;
oral cavity
ovary;
pancreas;
peritoneal cavity;
pituitary gland;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

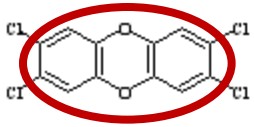
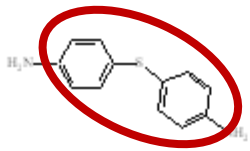
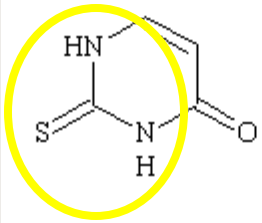
Relational Biological Content Searching: Carcinogenic Potency Database – All Species (CPDBAS_1481)



ACD/ChemFolder: Database Window - [D:\DSSTOX_MISC\ACD_CHEMFOLDERFILES_10APR2006\CPDBAS_V3B_1481_10APR2006.CFD]

Database View Record Search Reaction Lists Options ACD/Labs Help

ID:76 	ID:601 	ID:666 	ID:871 	ID:1094 	ID:1247 
76 CPDBAS_v3b_1481_10A	601 CPDBAS_v3b_1481_10	666 CPDBAS_v3b_1481_10	871 CPDBAS_v3b_1481_10	1094 CPDBAS_v3b_1481_1	1247 CPDBAS_v3b_1481_1
3-Aminotriazole	Ethylene thiourea (ETU)	Glycidol	4,4'-Methylenedianiline dihyd	4,4'-Oxydianiline	C.I. Basic red 9 monohydroc
negative	positive	positive	positive	positive	positive
thy	thy	ezy lgi mgl nrv per ski sto thy	liv thy	liv thy	ezy liv ski sub thy
pit thy	thy	cli hmo mgl nrv orc sto thy	thy	liv thy	ezy sub thy
liv	liv pit thy	hag liv lun ski sto	adr liv thy	hag liv	liv
liv	liv pit thy	hag mgl ski sub ute	hmo liv thy	hag liv thy	adr liv

ID:1315 	ID:1344 	ID:1347 
1315 CPDBAS_v3b_1481_1	1344 CPDBAS_v3b_1481_1	1347 CPDBAS_v3b_1481_1
2,3,7,8-Tetrachlorodibenzo-p	4,4'-Thiodianiline	Thiouracil
negative	positive	thy
orc thy	ezy lgi liv thy	thy
liv lun	ezy thy ute	liv
liv	liv thy	liv
liv thy	liv thy	

Search Results: 9 hits / 1481 total

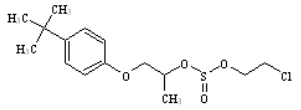
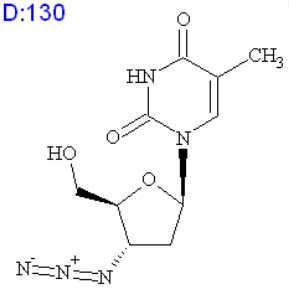
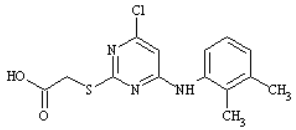
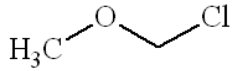
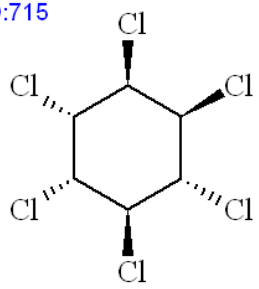
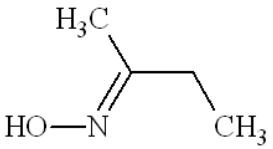
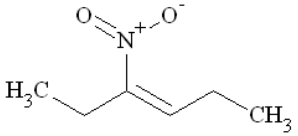
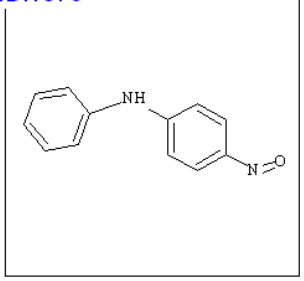
TargetSites_Rat_Male = thyroid

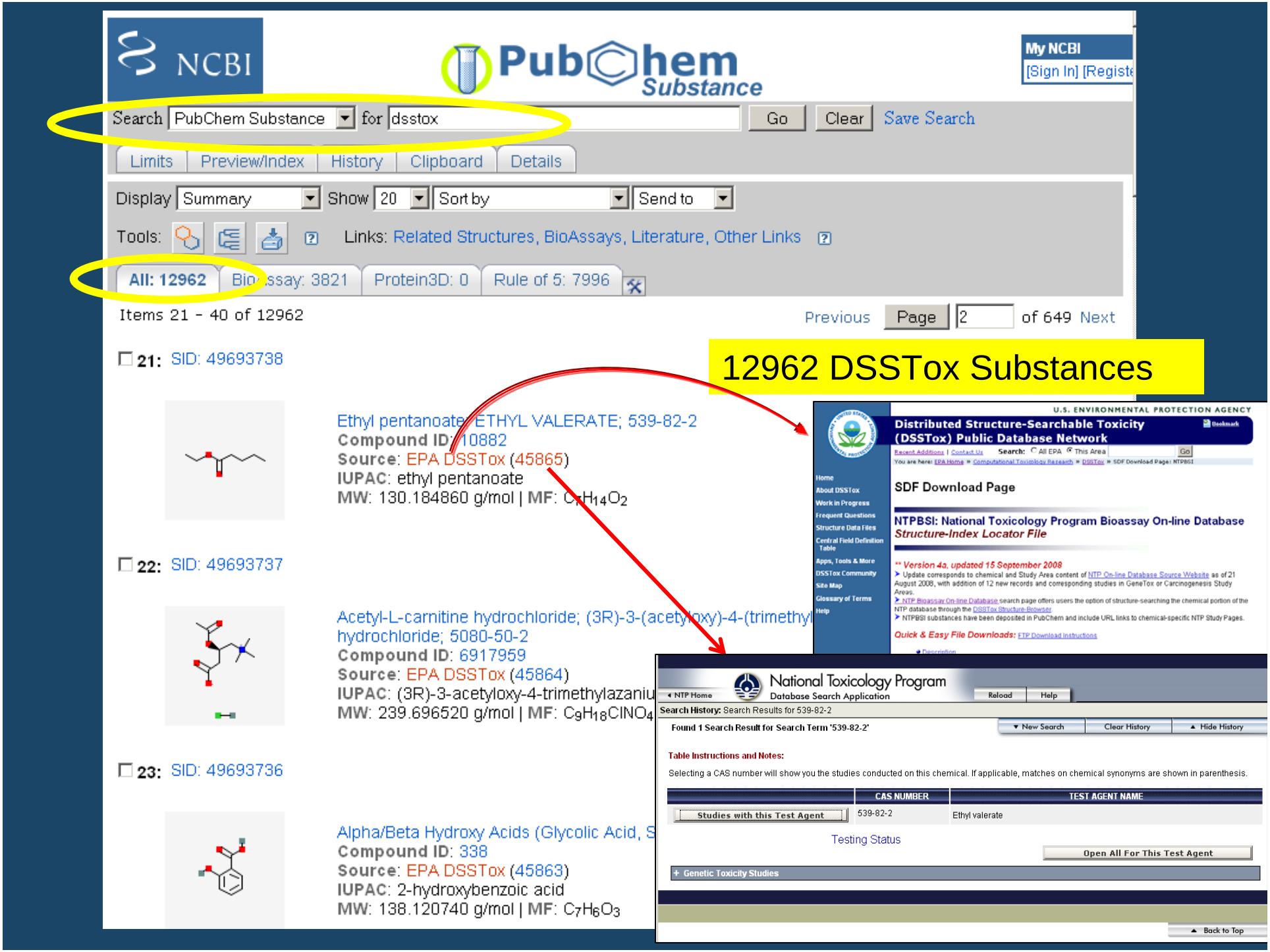
& TargetSites_Mouse_Male = liver

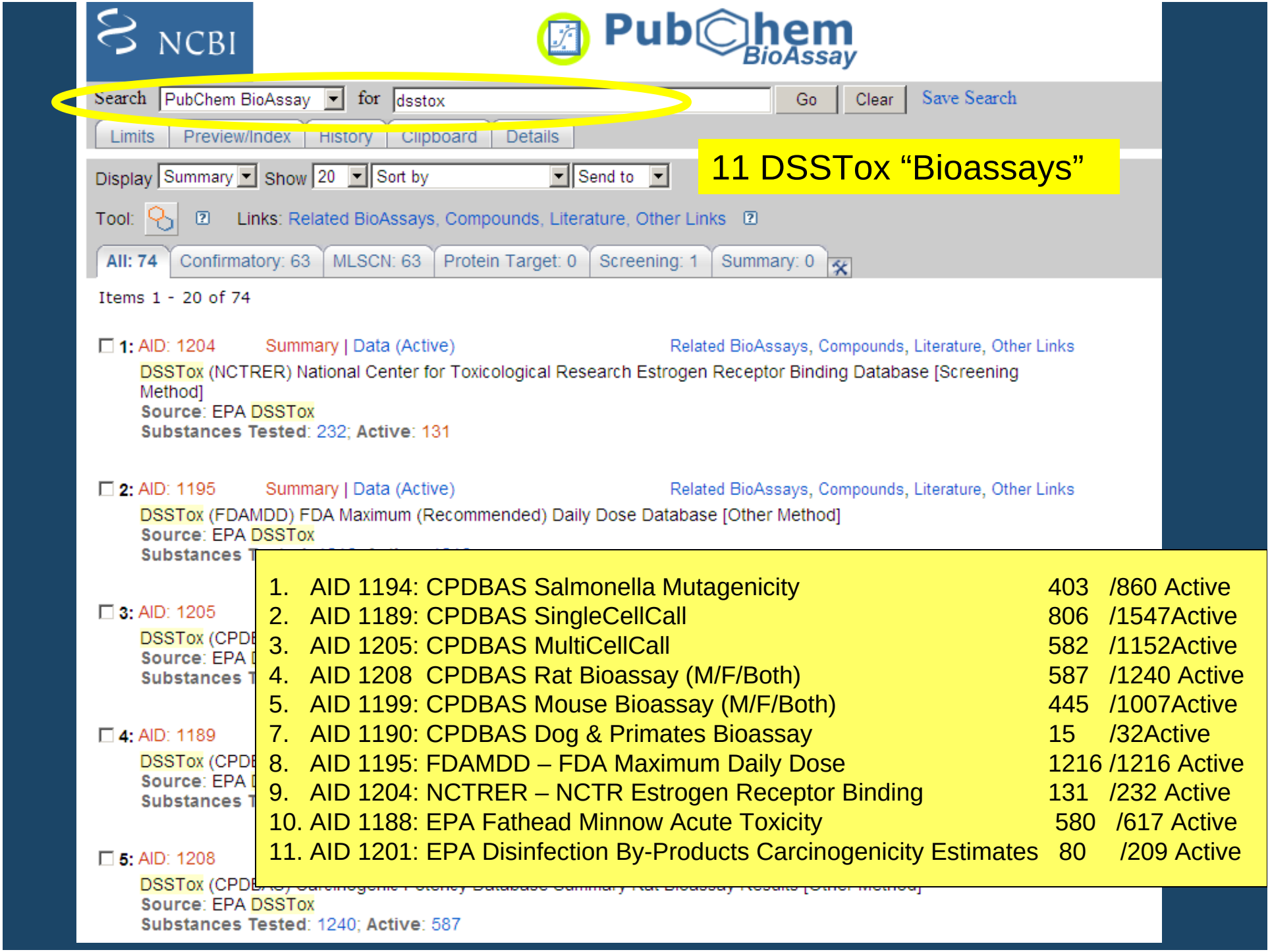
CPDBAS_v5b_1547

	ActivityCategory_		ActivityCategory_MultiCellCall_Details**			
Call	SingleCellCall	MultiCellCall**	multisite***	multisex	multispecies	Total Incidences* (CPDBAS_v5)
Active (1)	active					223
	active	active	✓			81
	active	active		✓		113
	active	active			✓	8
	active	active				
	active	active				
	active	active				
	active	active				
	active	active				
Inactive (0)	inactive					
	inactive	inactive				
	inactive	inactive				
	inactive	inactive				
	inactive	inactive				

If chemical is "**MultiCell Active**", it most likely crosses target site, sex and species.

ID:98 	ID:130 	ID:299 	ID:316 
Aramite 140-57-8	AZT 30516-87-1	[4-Chloro-6-(2,3-xylidino)-2-p 50892-23-4	Chloromethyl methyl ether 107-30-2
ID:715 	ID:912 	ID:1000 	ID:1076 
alpha-1,2,3,4,5,6-Hexachloro 319-84-6	Methylethylketoxime 96-29-7	3-Nitro-3-hexene 4812-22-0	p-Nitrosodiphenylamine 156-10-5


NCBI
PubChem Substance
My NCBI
[Sign In] [Regist
Search PubChem Substance for dsstox
Go
Clear
Save Search
Limits
Preview/Index
History
Clipboard
Details
Display Summary
Show 20
Sort by
Send to
Tools:
Links: Related Structures, BioAssays, Literature, Other Links
All: 12962
Bioassay: 3821
Protein3D: 0
Rule of 5: 7996
Items 21 - 40 of 12962
Previous
Page 2 of 649 Next
21: SID: 49693738
Ethyl pentanoate; ETHYL VALERATE; 539-82-2
Compound ID: 10882
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-(trimethyl
hydrochloride; 5080-50-2
Compound ID: 6917959
Source: EPA DSSTox (45864)
IUPAC: (3R)-3-acetyloxy-4-trimethylazaniu
MW: 239.696520 g/mol | MF: C9H18ClNO4
23: SID: 49693736
Alpha/Beta Hydroxy Acids (Glycolic Acid, S
Compound ID: 338
Source: EPA DSSTox (45863)
IUPAC: 2-hydroxybenzoic acid
MW: 138.120740 g/mol | MF: C7H6O3
12962 DSSTox Substances
U.S. ENVIRONMENTAL PROTECTION AGENCY
Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network
SDF Download Page
NTPBSI: National Toxicology Program Bioassay On-line Database
Structure-Index Locator File
** Version 4a, updated 15 September 2008
Quick & Easy File Downloads:
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
Studies with this Test Agent
539-82-2
Ethyl valerate
Testing Status
Open All For This Test Agent
Genetic Toxicity Studies
Back to Top


NCBI
PubChem BioAssay
Search PubChem BioAssay for dsstox Go Clear Save Search
Limits Preview/Index History Clipboard Details
11 DSSTox "Bioassays"
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
3: AID: 1205
4: AID: 1189
5: AID: 1208
Source: EPA DSSTox
Substances Tested: 1240; Active: 587
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

U.S. ENVIRONMENTAL PROTECTION AGENCY

Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

Bookmark

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

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

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

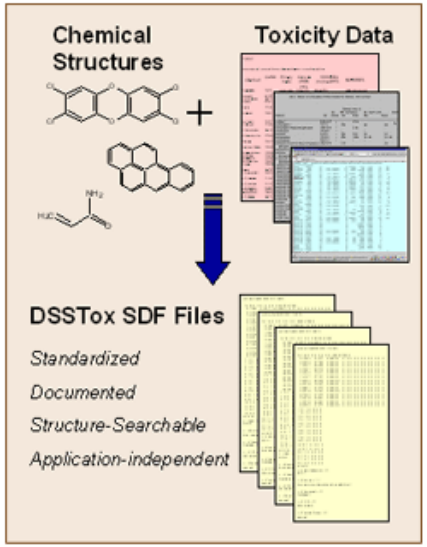
- About DSSTox
- Work in Progress
- Frequent Questions
- Structure Data Files
- Central Field Definition Table
- Apps, Tools & More
- DSSTox Community
- Site Map
- Glossary of Terms
- Help

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>

EPA DSSTox Structure-Browser GO v2.0

DSSTox Structure-Browser information Page



- DSSTox Graphic Flowchart
- DSSTox Project Goals
- DSSTox Publications

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.

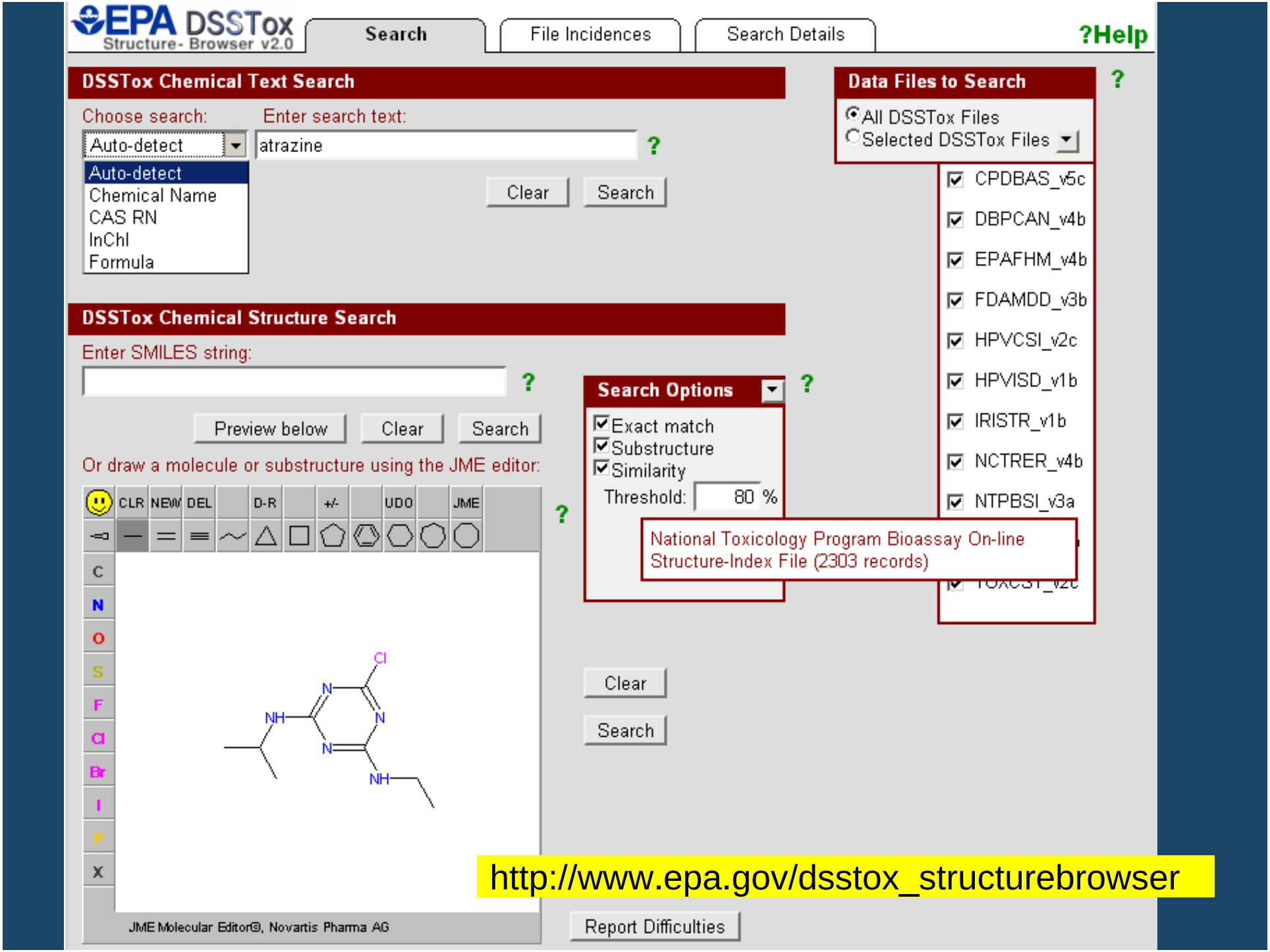
15 August 2008

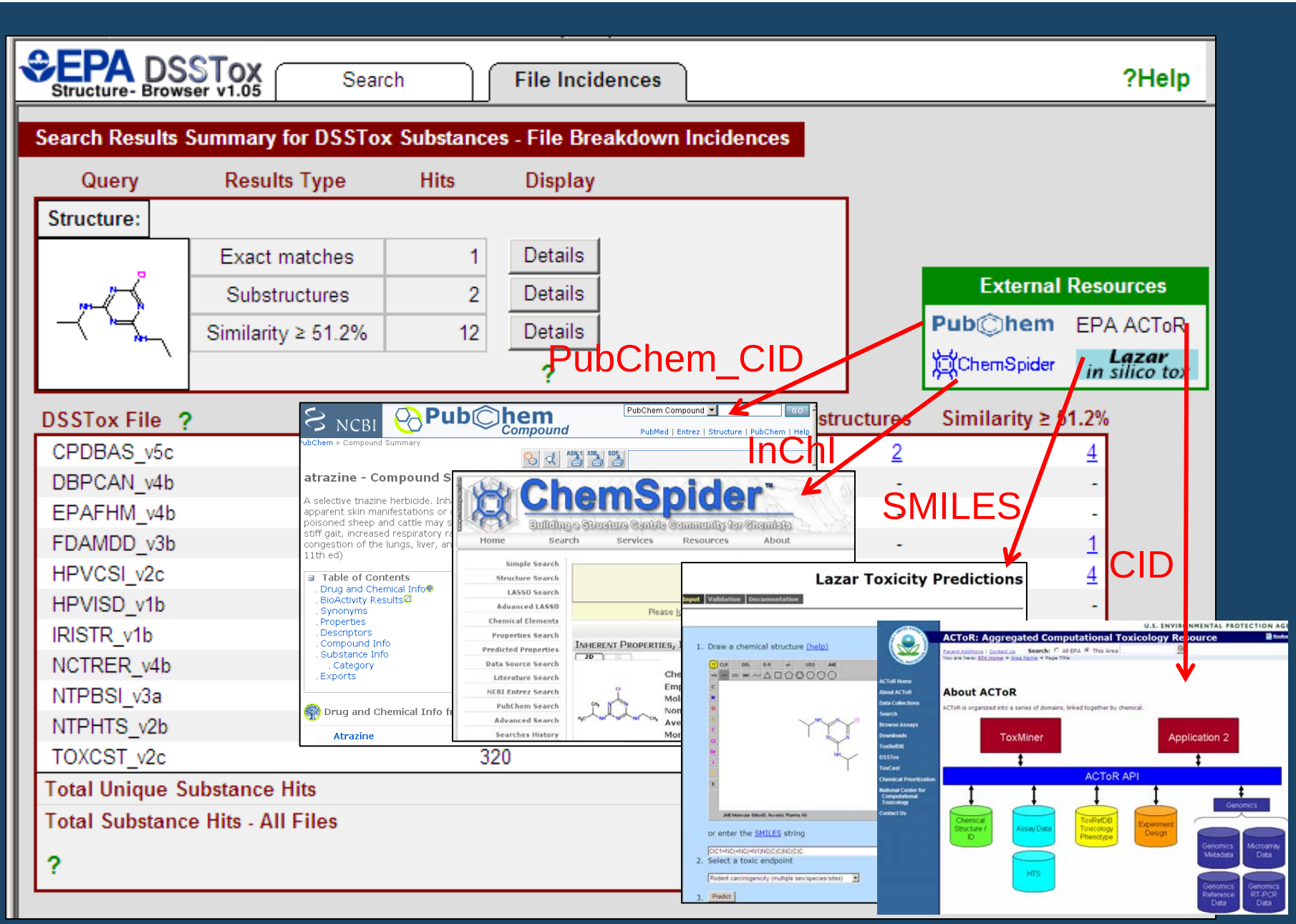
- Enhancements to DSSTox Structure Browser - Version 2.0
  - Expanded "link-in" options and user-defined limits to allow external websites to access DSSTox Structure Browser Search Results pages
  - Minimum of 10 structure similarity hits displayed regardless of similarity search threshold
  - New **External 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 faciliate structure look-up and searching
- Additional QA review, structure/CAS modifications, elimination of abbreviations in field entries, etc.

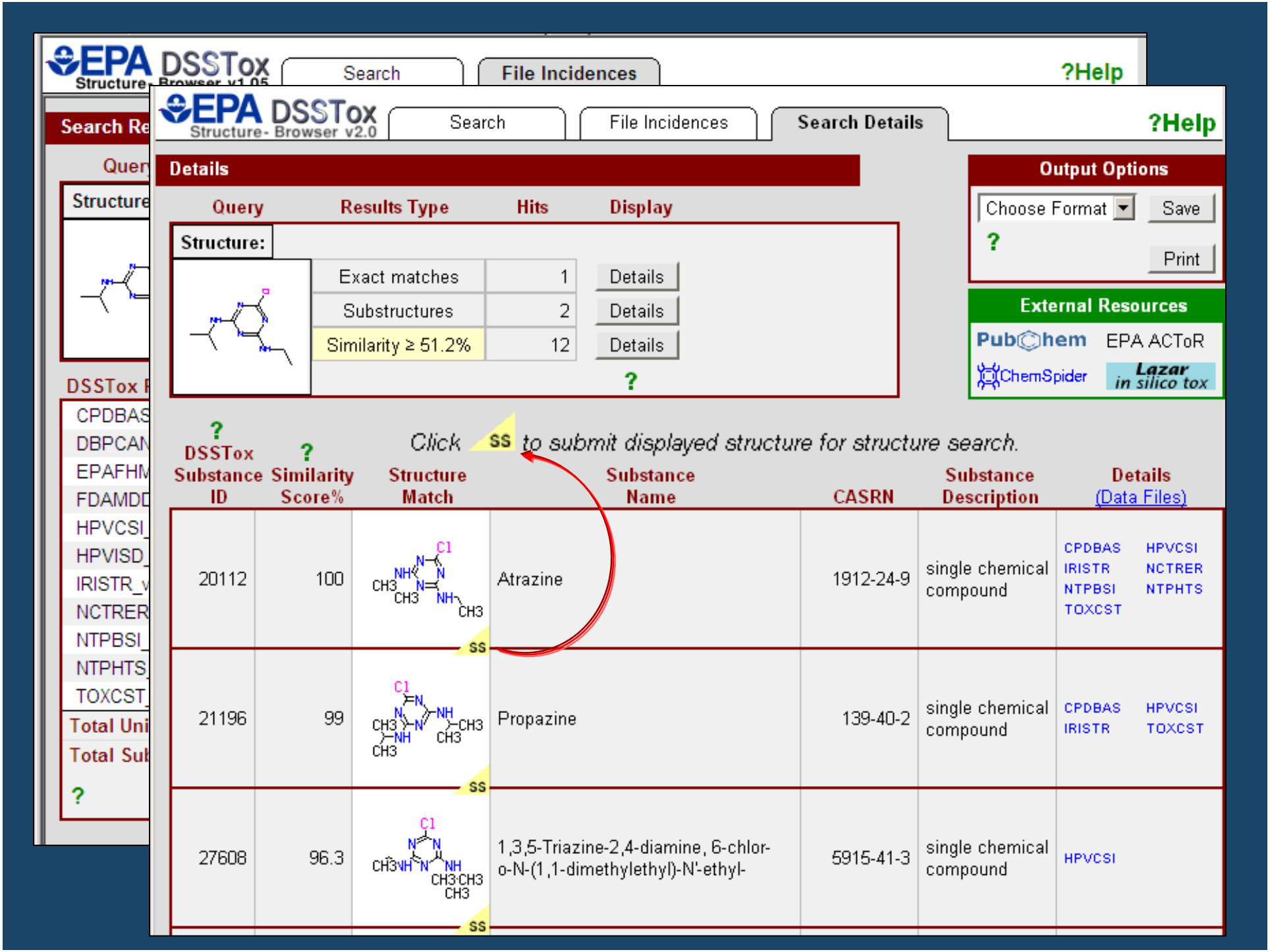
DSSTox Data Files: Details>

- 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 *update
- NTPHTS_v2b_1408_15Feb2008
- TOXCST_v2c_320_29Apr2008

>Other (non-DSSTox) Data Files


EPA DSSTox
Structure- Browser v2.0
Search
File Incidences
Search Details
?Help
DSSTox Chemical Text Search
Choose search:
Enter search text:
Auto-detect
atrazine
Auto-detect
Chemical Name
CAS RN
InChI
Formula
Clear
Search
Data Files to Search
All DSSTox Files
Selected DSSTox Files
CPDBAS_v5c
DBPCAN_v4b
EPAFHM_v4b
FDAMDD_v3b
HPVCSI_v2c
HPVISD_v1b
IRISTR_v1b
NCTRER_v4b
NTPBSI_v3a
DSSTox Chemical Structure Search
Enter SMILES string:
Preview below
Clear
Search
Search Options
Exact match
Substructure
Similarity
Threshold:
80 %
National Toxicology Program Bioassay On-line
Structure-Index File (2303 records)
Or draw a molecule or substructure using the JME editor:
CLR
NEW
DEL
D-R
+/-
UDO
JME
C
N
O
S
F
Cl
Br
I
P
X
JME Molecular Editor©, Novartis Pharma AG
Clear
Search
Report Difficulties
http://www.epa.gov/dsstox_structurebrowser


EPA DSSTox
Structure- Browser v1.05
Search
File Incidences
?Help
Search Results Summary for DSSTox Substances - File Breakdown Incidences
Query
Results Type
Hits
Display
Structure:
Exact matches 1 Details
Substructures 2 Details
Similarity ≥ 51.2% 12 Details
PubChem_CID
InChI
SMILES
CID
External Resources
PubChem
EPA ACToR
ChemSpider
Lazar in silico tox
DSSTox File ?
structures
Similarity ≥ 51.2%
CPDBAS_v5c 2 4
DBPCAN_v4b - -
EPAFHM_v4b - -
FDAMDD_v3b - 1
HPVCSI_v2c 4
HPVISD_v1b -
IRISTR_v1b
NCTRER_v4b
NTPBSI_v3a
NTPHTS_v2b
TOXCST_v2c 320
Total Unique Substance Hits
Total Substance Hits - All Files
?
NCBI
PubChem Compound
ChemSpider
Building a Structure Centric Community for Chemists
Lazar Toxicity Predictions
ACToR: Aggregated Computational Toxicology Resource
About ACToR
ToxMiner
Application 2
ACToR API


EPA DSSTox
Structure- Browser v2.0
Search
File Incidences
Search Details
?Help
Details
Query
Results Type
Hits
Display
Structure:
Exact matches
1
Details
Substructures
2
Details
Similarity ≥ 51.2%
12
Details
Output Options
Choose Format
Save
Print
External Resources
PubChem
EPA ACToR
ChemSpider
Lazar in silico tox
Click SS to submit displayed structure for structure search.
DSSTox Substance ID
Similarity Score%
Structure Match
Substance Name
CASRN
Substance Description
Details (Data Files)
20112
100
Atrazine
1912-24-9
single chemical compound
CPDBAS HPVCSI IRISTR NCTRER NTPBSI NTPHTS TOXCST
21196
99
Propazine
139-40-2
single chemical compound
CPDBAS HPVCSI IRISTR TOXCST
27608
96.3
1,3,5-Triazine-2,4-diamine, 6-chloro-N-(1,1-dimethylethyl)-N'-ethyl-
5915-41-3
single chemical compound
HPVCSI

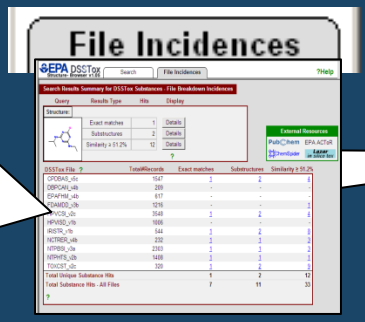

U.S. ENVIRONMENTAL PROTECTION AGEN
Distributed Structure-Searchable Toxicity (DSSTox)
Public Database Network
Home
About DSSTox
Work in Progress
Frequent Questions
Structure Data Files
Central Field Definition Table
Apps, Tools & More
DSSTox Community
Site Map
SDF Download Page
CPDBAS: Carcinogenic Potency Database Summary Tables - All Species
Database File
** Version 5c, revised 29 April 2008:
One Mouse bioassay correction (inactive result removed) and 2 structure corrections.
Search
File Incidences
Search Details
Substance Results
?Help
CPDBAS:
Carcinogenic Potency Database Summary T
Species (1547 records)
CPDBAS_v5c_1547_29Apr2008
CPDBAS Source Website
View CPDB Chemical Substance Data Page
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 of NCI/NTP Bioassays
Summary Table by Target Site of NCI/NTP Bioassays
Animal Cancer Tests and Human Cancer Risk
Misconceptions about the Causes of Cancer
Our Publications by Topic
Our Publications by Year
Aristolochic Acid
Acknowledgement and Support
The Carcinogenic Potency Project
http://potency.berkeley.edu/
Aniline (CAS 62-53-3)
Rats and Mice: Cancer Test Summary
Rat Target Sites
Mouse Target Sites
Male
Female
Rat
Mouse
no positive
no test
Key to the Table Above
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
0.169693249315084 mmol/kg-bw/day
ActivityScore_CPDBAS_Rat
34
TD50_Rat_Note
TD50 is harmonic mean of more than one positive test
TargetSites_Rat_Male
mammary gland
TargetSites_Rat_Female
hematopoietic system; mammary gland; uterus
ActivityOutcome_CPDBAS_Rat
active
ActivityScore_CPDBAS_Mouse
0
TD50_Mouse_Note
no positive results

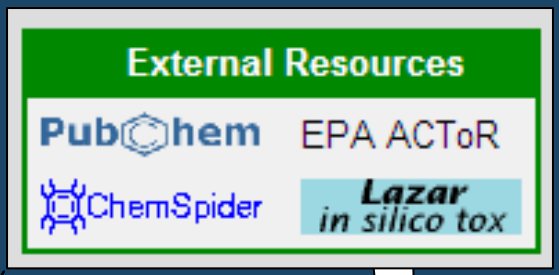
http://www.epa.gov/dsstox_structurebrowser/...

?dbs=cpdbas

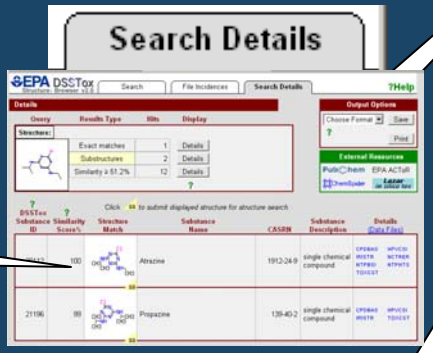
PubChem_CID
InChIKey
CAS
SMILES

?qtype=cas&qval=87-86-5
?q type=name&qval=atrazine
?q type=smiles&qval=CC
?qtype=inchikey&qval=…
?qtype=cid&qval=20238



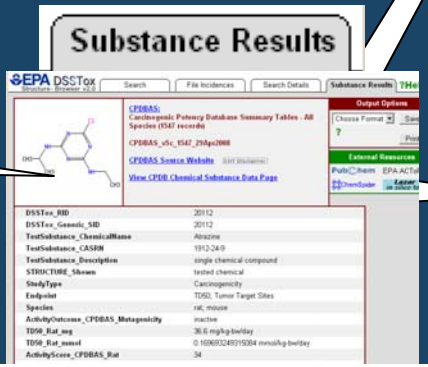


Link-ins



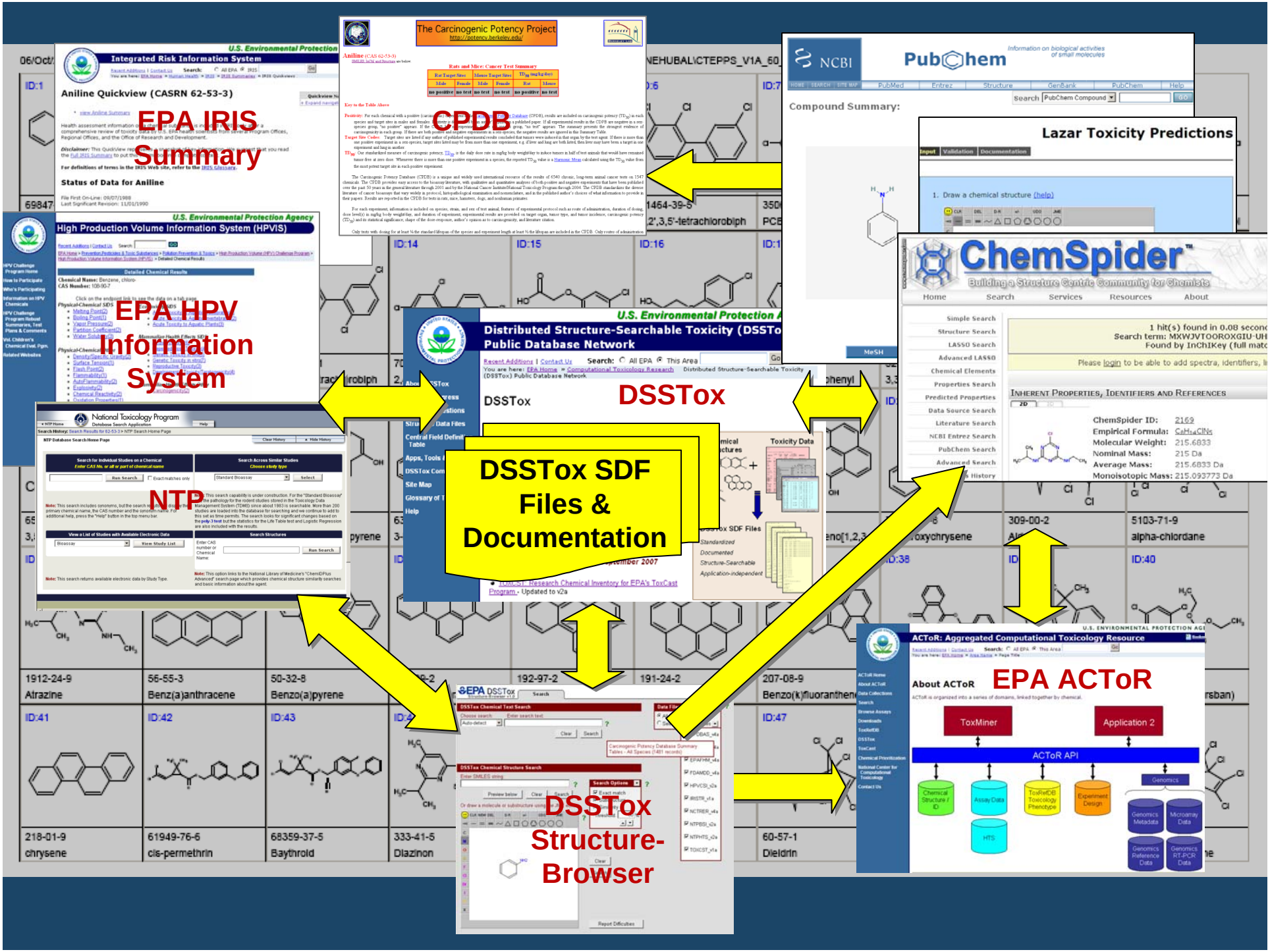
?qtype=sid&qval=20112

Link-outs



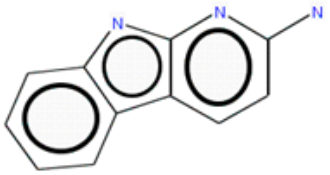
?qtype=rid&qval=20112

DSSTox Download Page
Source Home URL
Source Chemical Data URL


EPA IRIS Summary
CPDB
EPA HPV Information System
DSSTox
DSSTox SDF Files & Documentation
NTP
DSSTox Structure-Browser
EPA ACToR
Integrated Risk Information System
Aniline Quickview (CASRN 62-53-3)
The Carcinogenic Potency Project
High Production Volume Information System (HPVIS)
Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network
National Toxicology Program
ACToR: Aggregated Computational Toxicology Resource
PubChem
Lazar Toxicity Predictions
ChemSpider
1912-24-9 Atrazine
56-55-3 Benz(a)anthracene
50-32-8 Benzo(a)pyrene
192-97-2
191-24-2
207-08-9 Benzo(k)fluoranthene
218-01-9 chrysene
61949-76-6 cis-permethrin
68359-37-5 Baythroid
333-41-5 Diazinon
60-57-1 Dieldrin
309-00-2
5103-71-9 alpha-chlordane

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(N=1)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-9CI	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

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 Houshold Products
TOXNET HSDB
TOXNET MESH Headings
TOXNET TOXLINE
TOXNET TOXMAP

U.S. ENVIRONMENTAL PROTECTION AGENC

ACToR: Aggregated Computational Toxicology Resource

Bookmark

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

You are here: EPA Home » Area Name » Page Title

ACToR Home
About ACToR
Data Collections
Search

ACToR Home

Data Collection

Name: IRISTR_DSSTOX

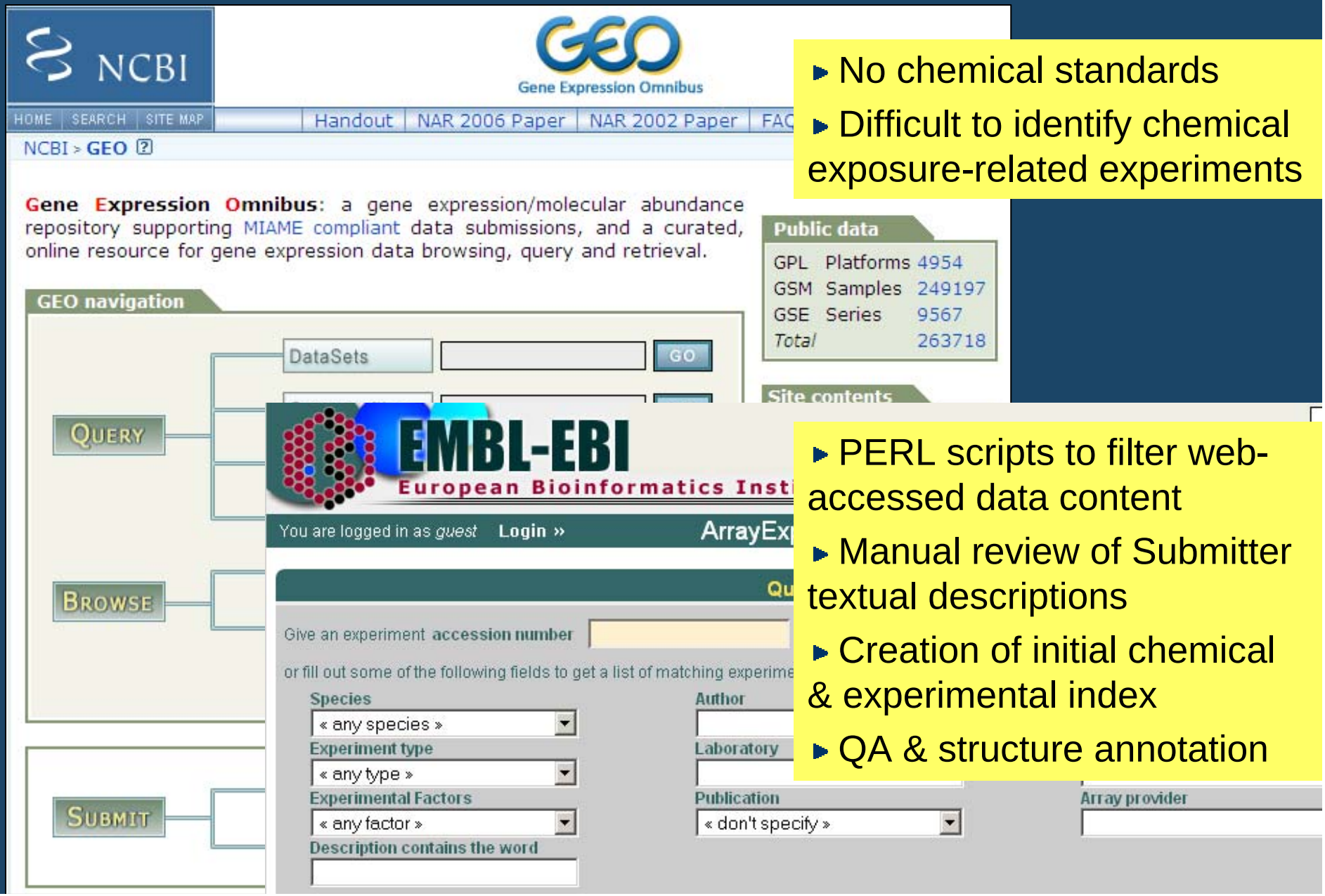
Description: EPA Integrated Risk Information System http://www.epa.gov/iris/. A structure index file with links to webpage for
Source Type: DSSTox Collection
Number of Substances: 544
Number of Generic Chemicals: 536
Compilation Date: 9/27/2007
Compilation Instructions: Download SD and excel files from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

Where are the data??

Towards a Public “Toxico-chemogenomics” Capability



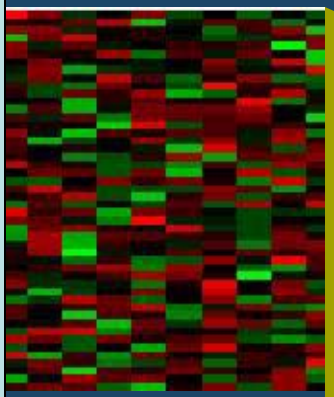
- No chemical standards
- Difficult to identify chemical exposure-related experiments

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

NCBI

GEO Gene Expression Omnibus

>9000
Experiments



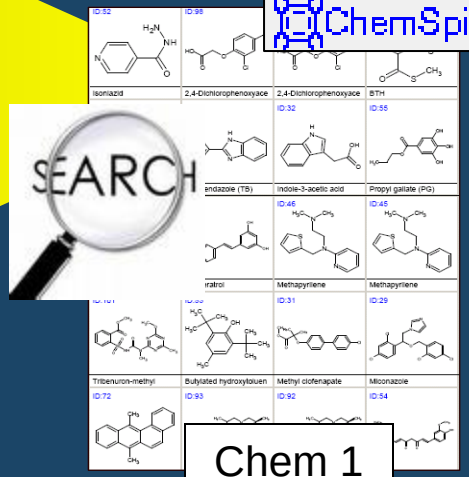
Experiments

Genes

Pathways

PubChem

ChemSpider



2220
Chemical experiments

Vehicle
Media
Reference
Combination

1293
Chemical treatment-related experiments

Treatment

943
Unique chemicals

837 defined organics
84 inorganics
22 organometallics

Chem 1
Chem 2
Chem 3
Chem 4
…
…
Chem 943

DSSTox
GEOCSI

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

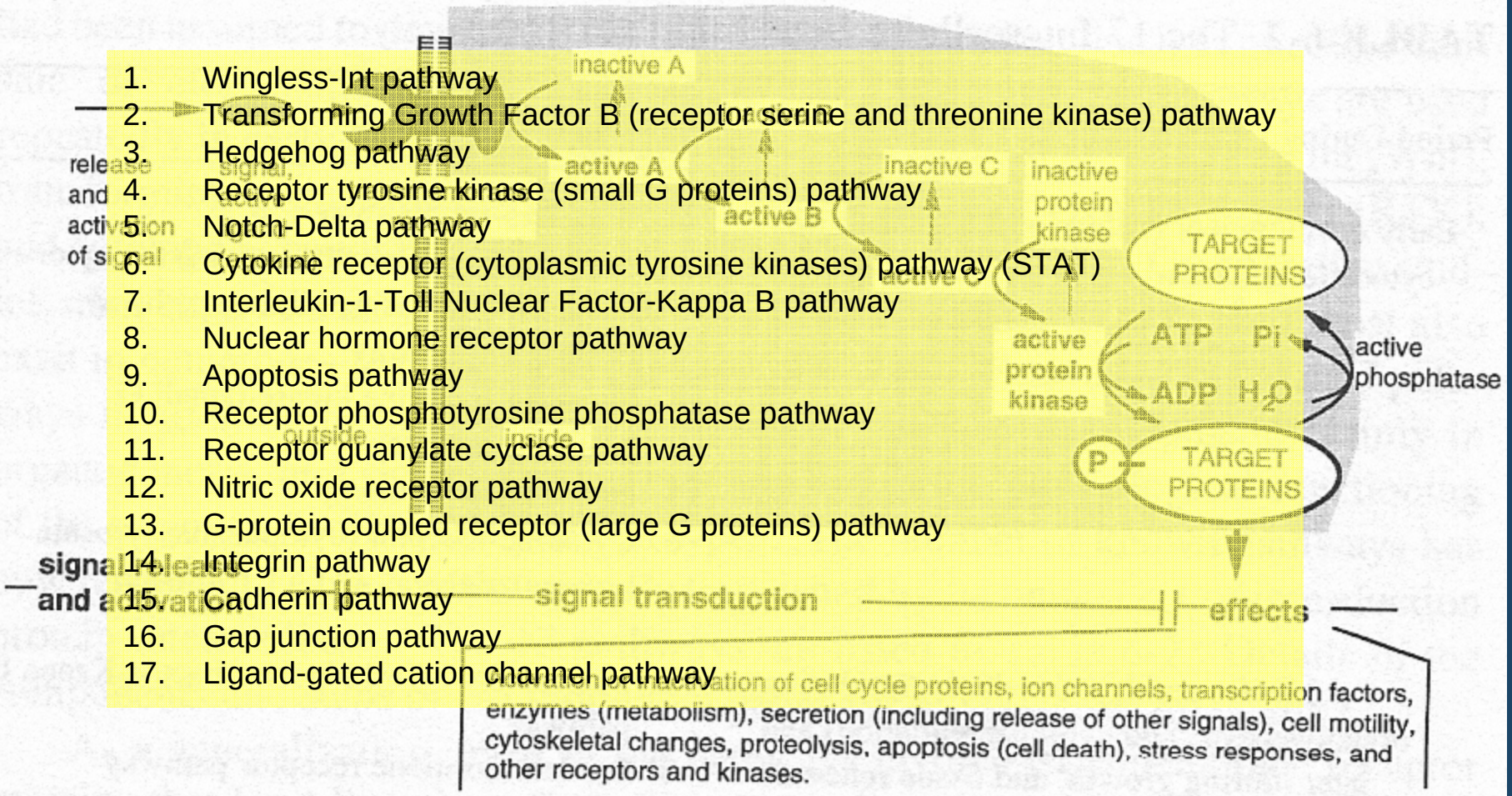
Francis S. Collins,[1*†] George M. Gray,[2*] John R. Bucher[3*]

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 μM, and tolerate high false-negative rates. In contrast, in

Scientific Frontiers in Developmental Toxicology & Risk Assessment - National Academy of Sciences, 2000

1. Wingless-Int pathway
2. Transforming Growth Factor B (receptor serine and threonine kinase) pathway
3. Hedgehog pathway
4. Receptor tyrosine kinase (small G proteins) pathway
5. Notch-Delta pathway
6. Cytokine receptor (cytoplasmic tyrosine kinases) pathway (STAT)
7. Interleukin-1-Toll Nuclear Factor-Kappa B pathway
8. Nuclear hormone receptor pathway
9. Apoptosis pathway
10. Receptor phosphotyrosine phosphatase pathway
11. Receptor guanylate cyclase pathway
12. Nitric oxide receptor pathway
13. G-protein coupled receptor (large G proteins) pathway
14. Integrin pathway
15. Cadherin pathway
16. Gap junction pathway
17. Ligand-gated cation channel pathway



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

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

- 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

ToxCast™ Program

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

ToxCast™ Navigation

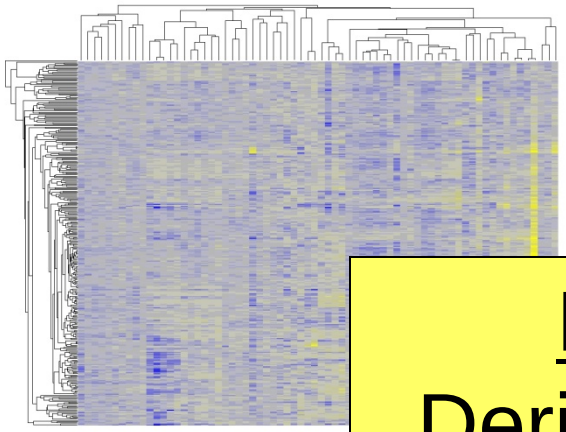
- Introduction
- ToxCast™ Chemicals
- ToxCast™ Assays
- ToxCast™ Information Management
- ToxCast™ Partnerships
- ToxCast™ Contractors
- ToxCast™ Presentations
- ToxCast™ Publications
- ToxCast™ News

Introduction

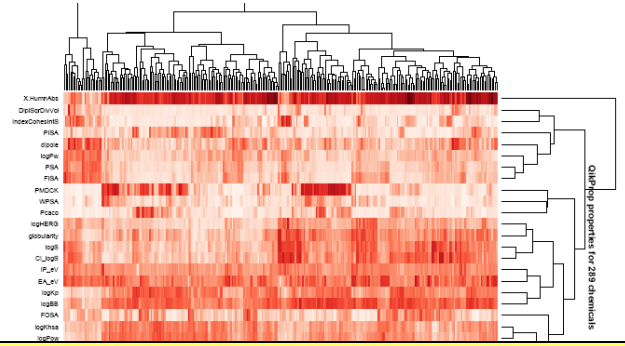
In 2007, EPA launched ToxCast™ in order to develop a cost-effective approach for prioritizing the toxicity testing of large numbers of chemicals in a short period of time. Using data from state-of-the-art high throughput screening (HTS) bioassays developed in the pharmaceutical industry, ToxCast™ is 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.

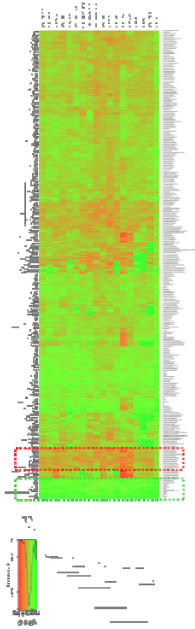
In its first phase, ToxCast™ is profiling over 300 well-characterized chemicals (primarily pesticides) in over 400 HTS endpoints. These endpoints include biochemical assays of protein function, cell-based transcriptional reporter assays, multi-cell interaction assays, transcriptomics on primary cell 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-

Correlating Domain Outputs



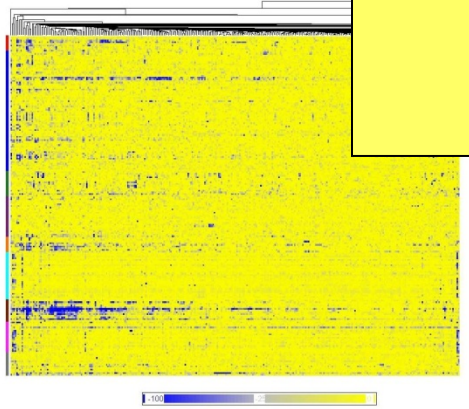
Cellular Assays



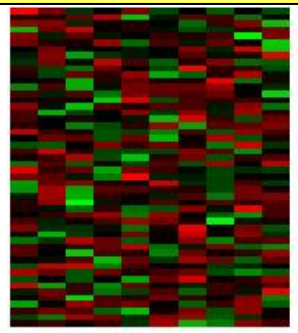


silico Predictions

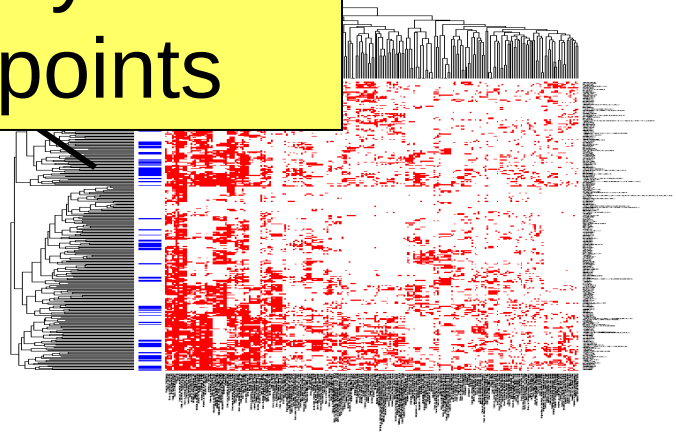
EPA ToxCast Goal:
Derive "Signatures" from *in vitro* & *in silico* assays to predict *in vivo* endpoints



Biochemical Assays

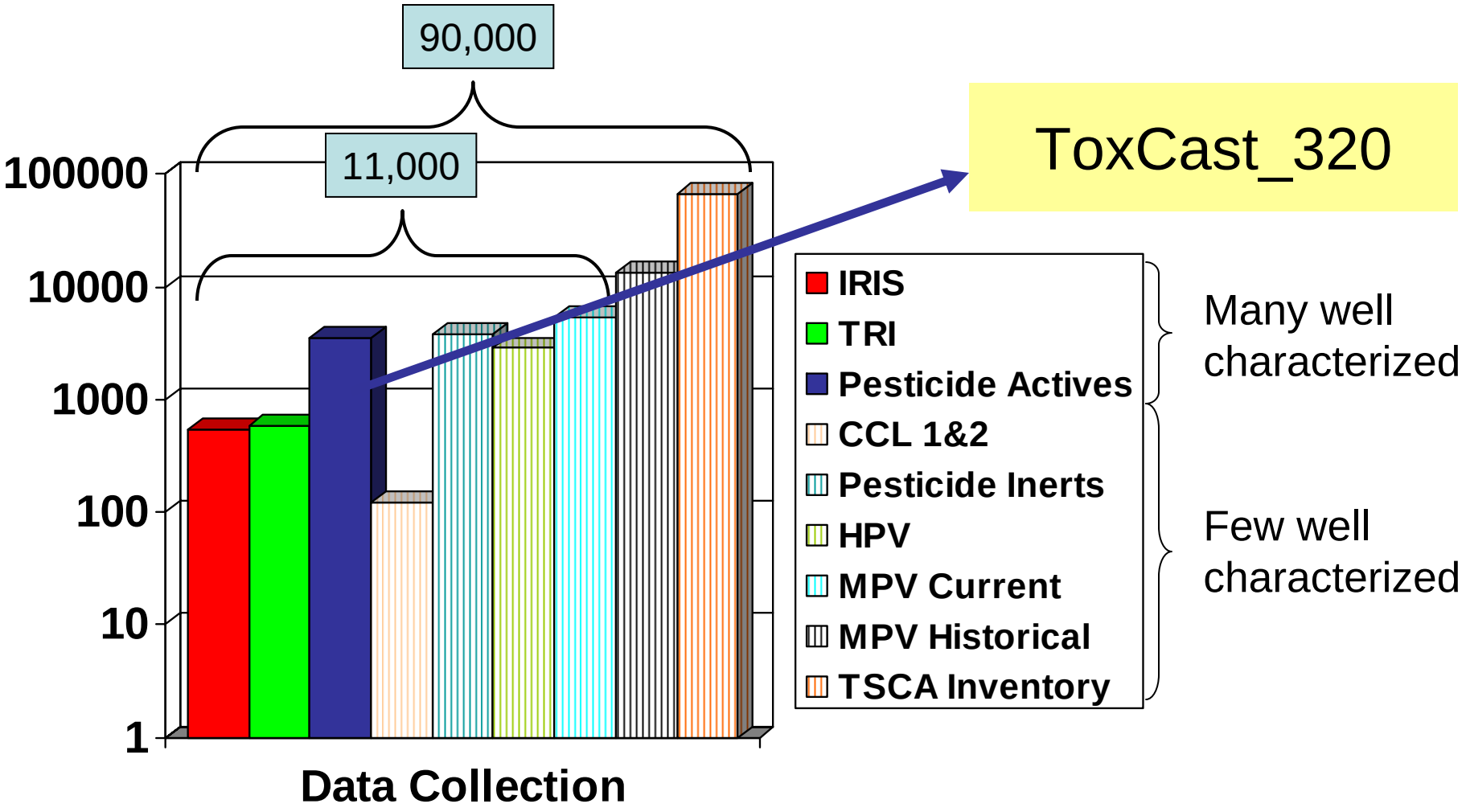


Genomic Signatures



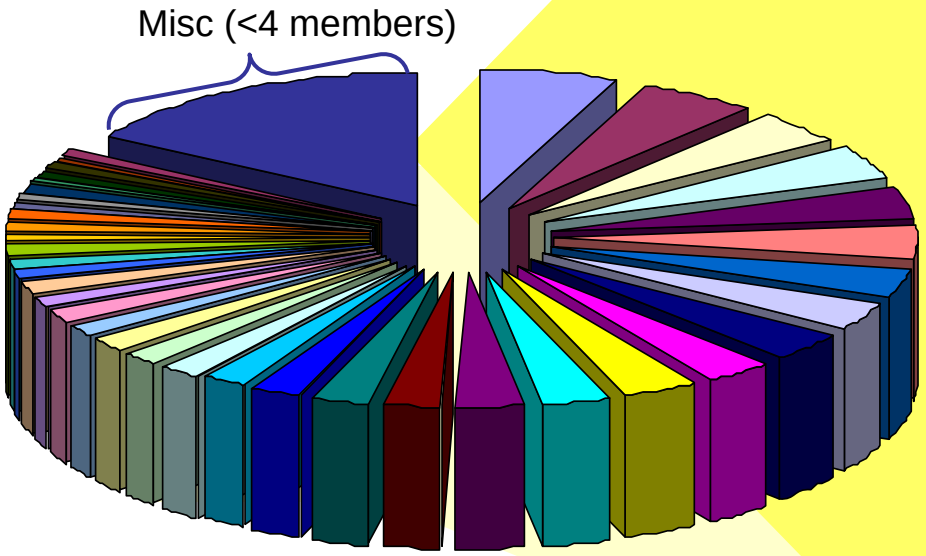
Toxicology Endpoints

ToxCast Phase I Chemicals



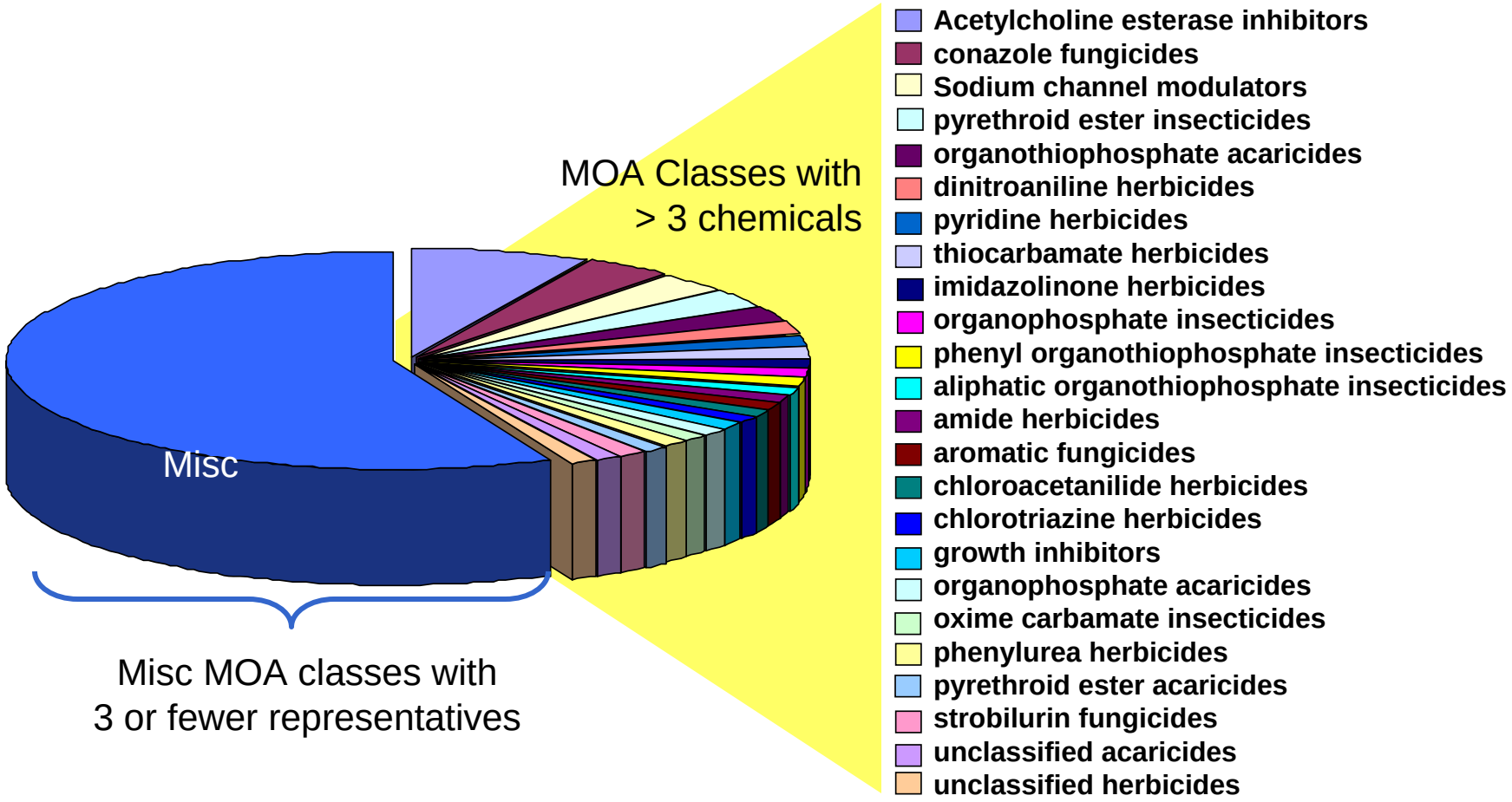
Chemical Classes in ToxCast_320 (Phase I)

- 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



- CHLORINE
- ORGANOPHOSPHORUS
- AMIDE
- ESTER
- ETHER
- PYRIDINE
- FLUORINE
- CARBOXYLIC ACID
- PHENOXY
- KETONE
- TRIAZINE
- CARBAMATE
- PHOSPHOROTHIOATE
- PYRIMIDINE
- BENZENE
- ORGANOCHLORINE
- AMINE
- PYRETHROID
- SULFONYLUREA
- TRIAZOLE
- UREA
- IMIDAZOLE
- NITRILE
- ALCOHOL
- CYCLO
- PHOSPHORODITHIOATE
- THIOCARBAMATE
- ANILINE
- THIAZOLE
- DINITROANILINE
- OXAZOLE
- PHOSPHATE
- IMINE
- NITRO
- PHENOL
- PHTHALIMIDE
- PYRAZOLE
- SULFONAMIDE

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



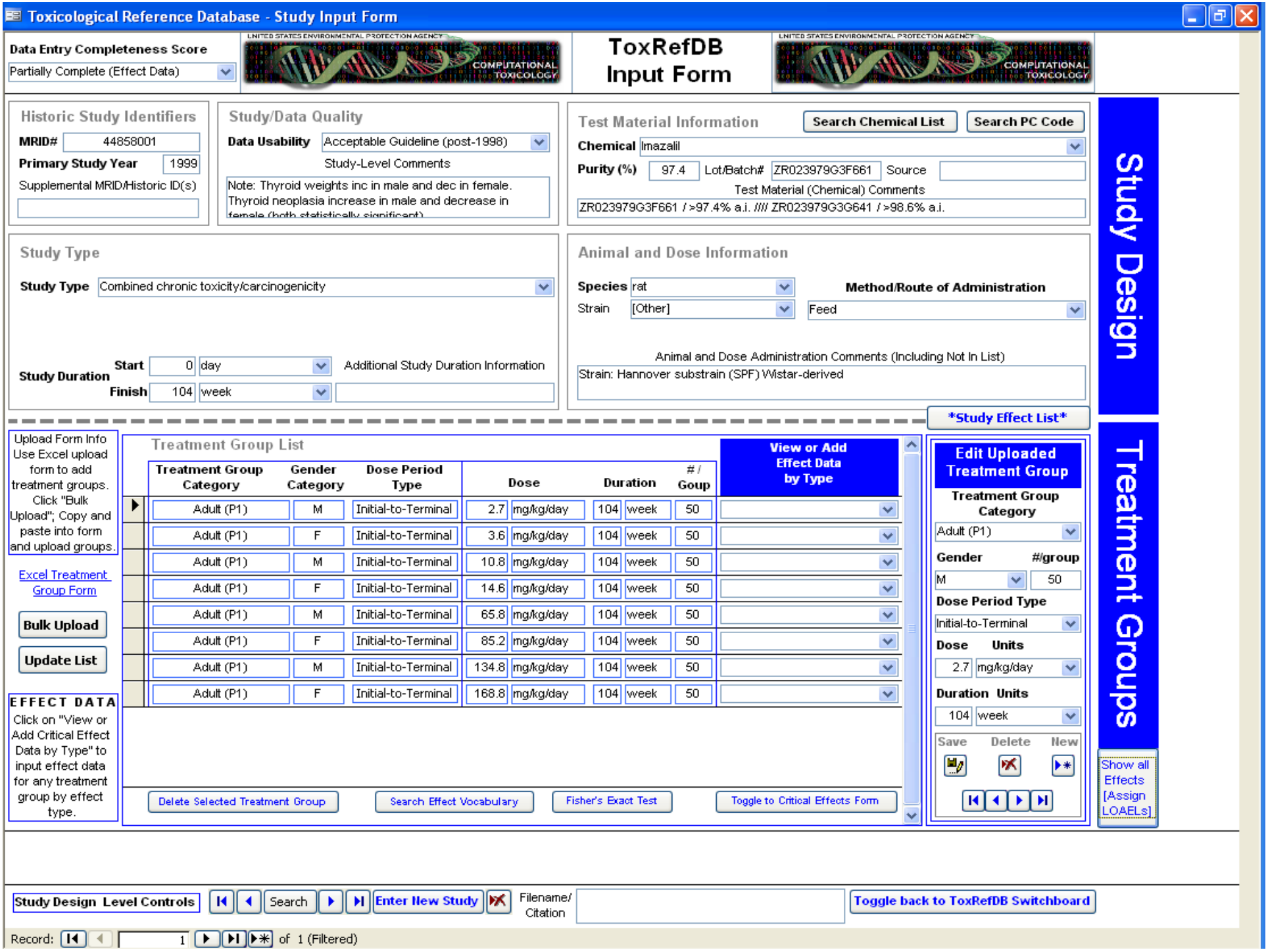
Classification based on OPPIN

EPA Pesticide Programs: Data Evaluation Records (DERs)

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

DER Format

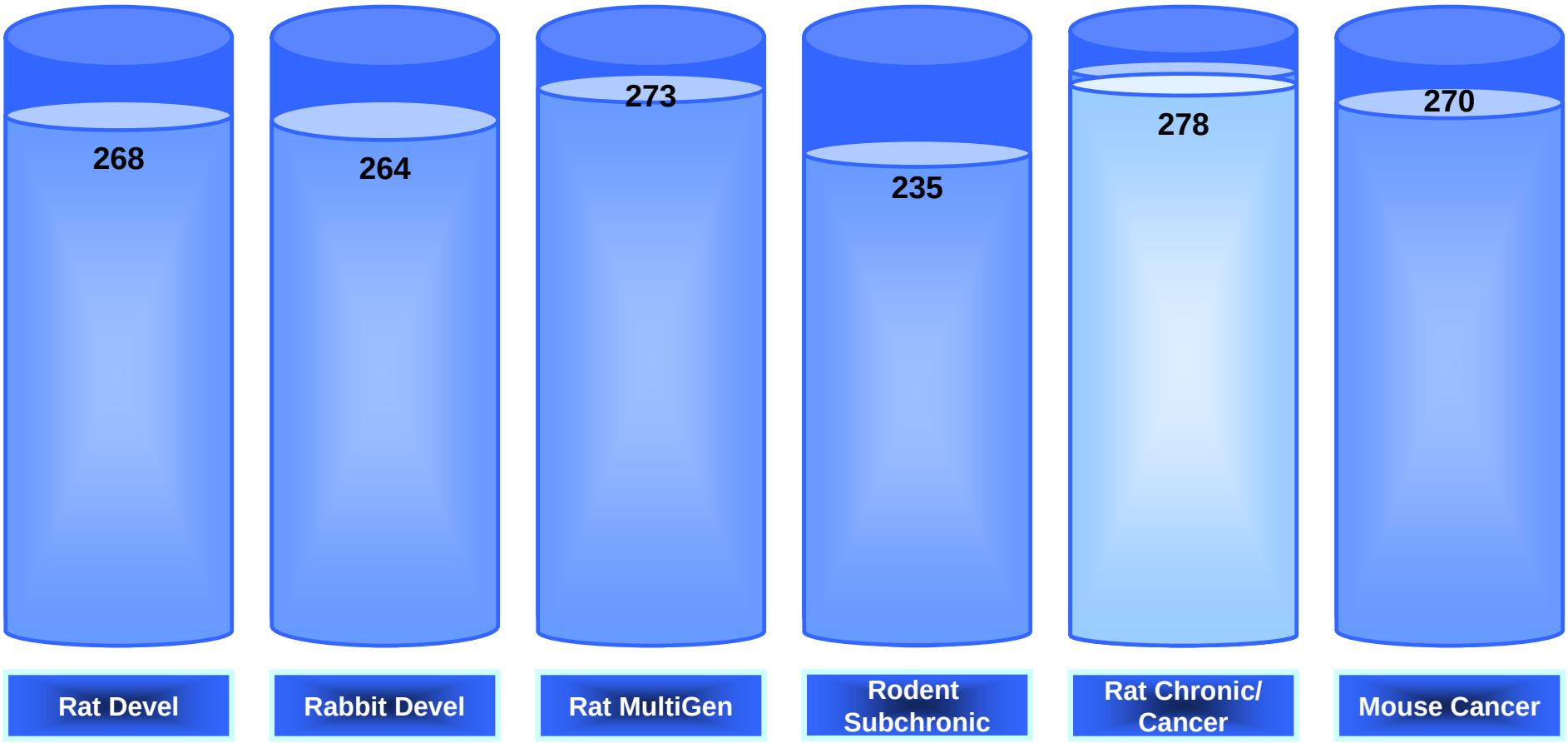
- Study Identifiers
  - Tested Chemical Information
    - IDs
    - Name
    - Purity
  - Study Type IDs
  - Reviewer Information
- Citation(s)
- Executive Summary
  - Summary Study Design
  - Summary Effects
  - Endpoints (NOAEL/LOAEL)
- Test Material
  - Purity
  - Source
  - Physical/Chemical Properties
- 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
M
#/group
50
Dose Period Type
Initial-to-Terminal
Dose
2.7
Units
mg/kg/day
Duration
104
Units
week
Save
Delete
New
Show all Effects [Assign LOAELs]
Study Design Level Controls
Search
Enter New Study
Filename/ Citation
Toggle back to ToxRefDB Switchboard
Record: 1 of 1 (Filtered)

ToxRefDB Data Entry: Phase I

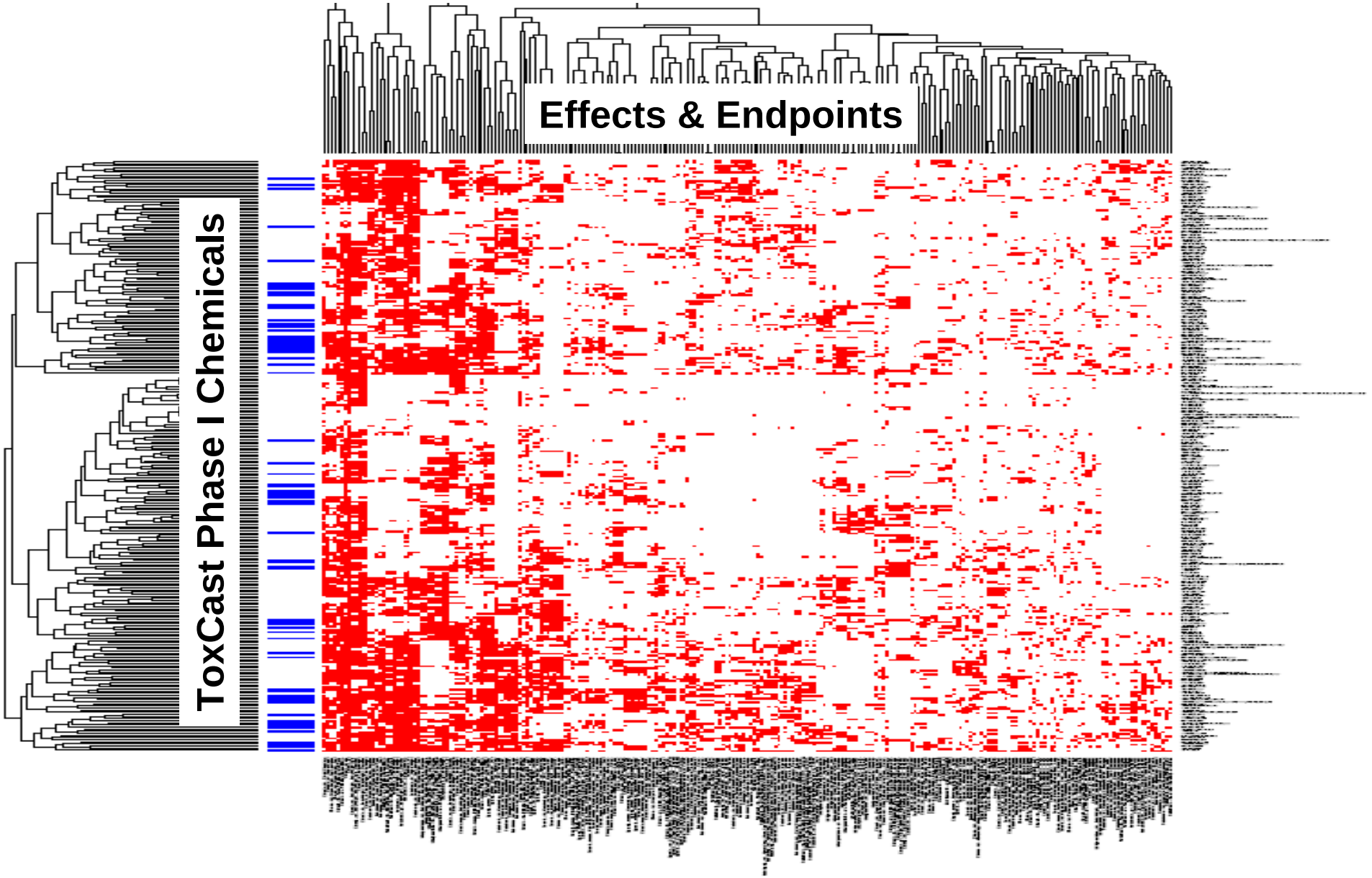
United States Environmental Protection Agency

291 Pesticides

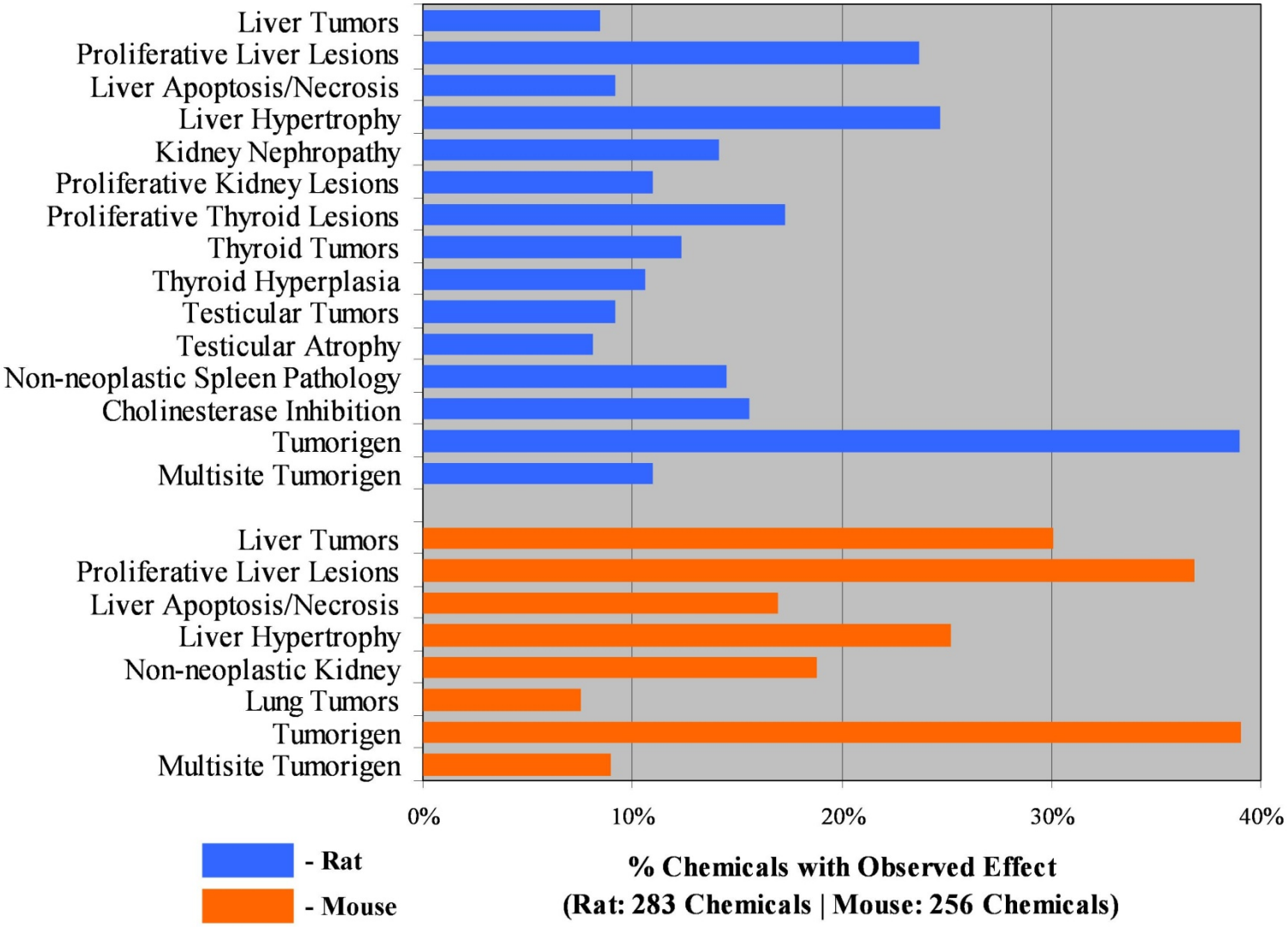


EPA
United States
Environmental Protection
Agency

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

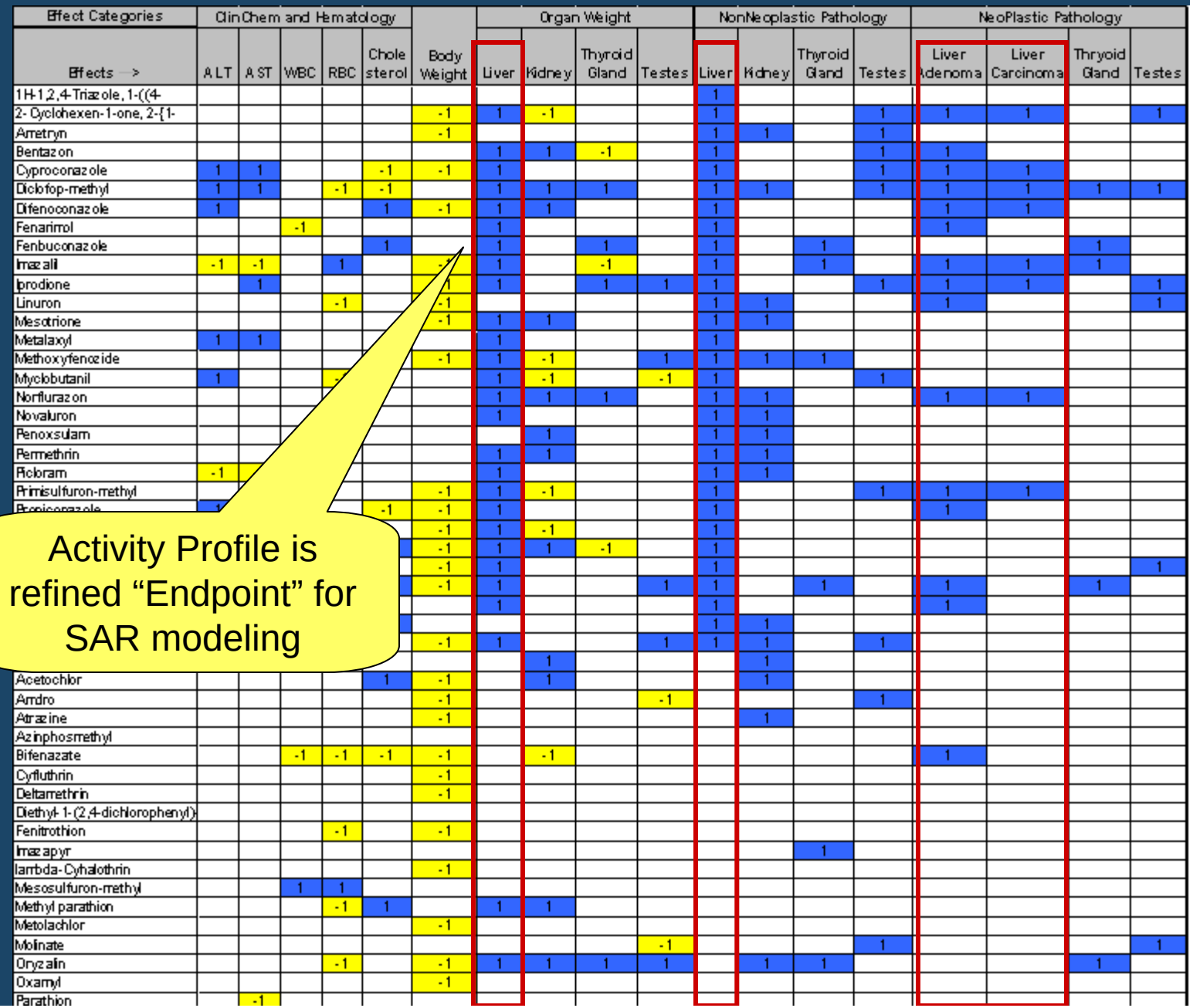


Common Phenotypes in Chronic Rodent Studies



ToxRefDB Profiling of Liver Effects for Pesticides

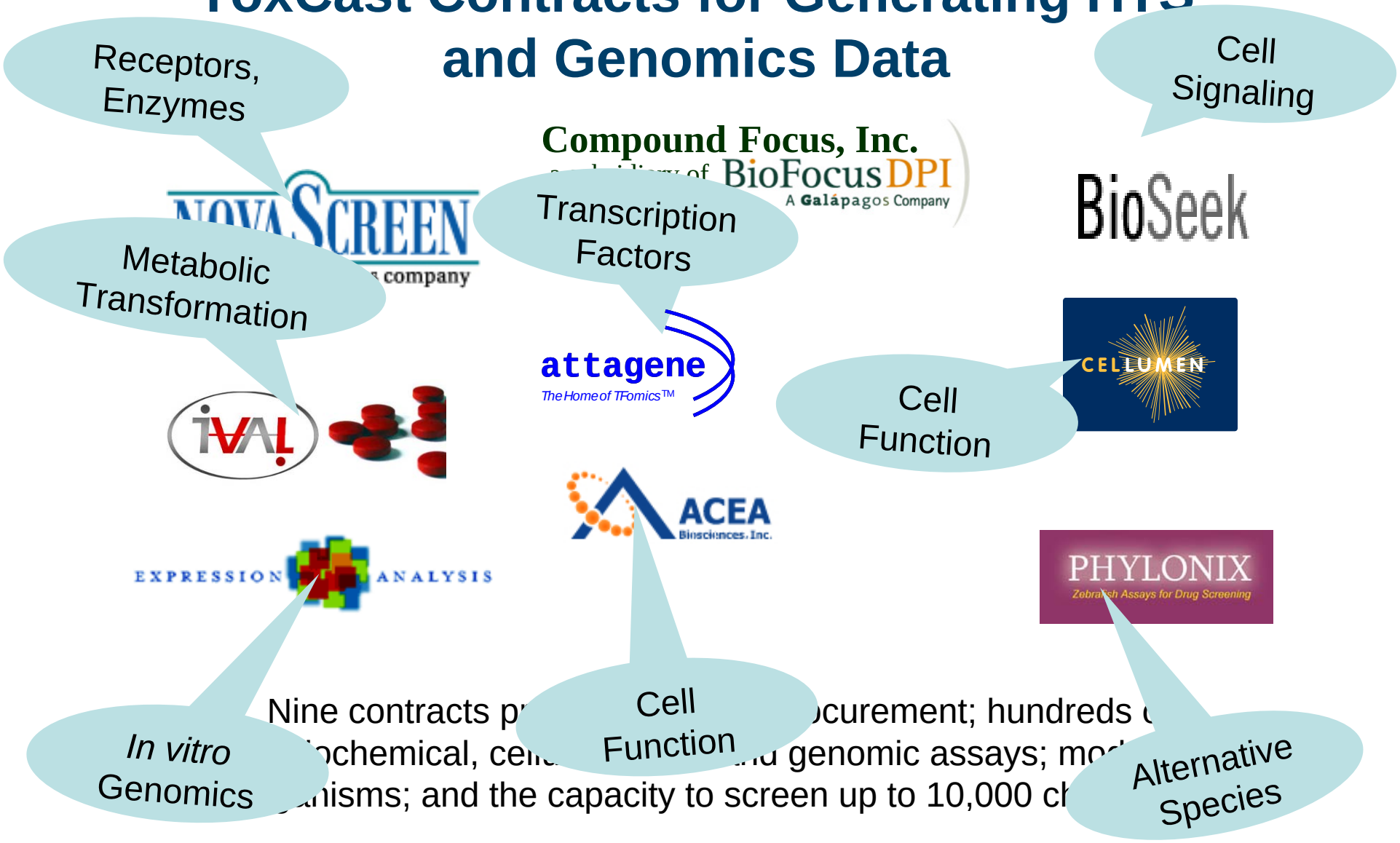
Liver non-neoplastic histopathology and increased organ weight are often associated with tumors and cancer



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	Thryoid Gland	Testes
1H-1,2,4-Triazole, 1-((4-											1							
2- Cyclohexen-1-one, 2-(1-						-1	1	-1			1			1	1	1		1
Arretryn						-1					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		1
Linuron				-1		-1					1	1			1			1
Mesotrione						-1	1	1			1	1						
Metalaxyl	1	1					1				1							
Methoxyfenozide						-1	1	-1		1	1	1	1					
Myclobutanil	1			-1			1	-1		-1	1			1				
Norflurazon							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				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



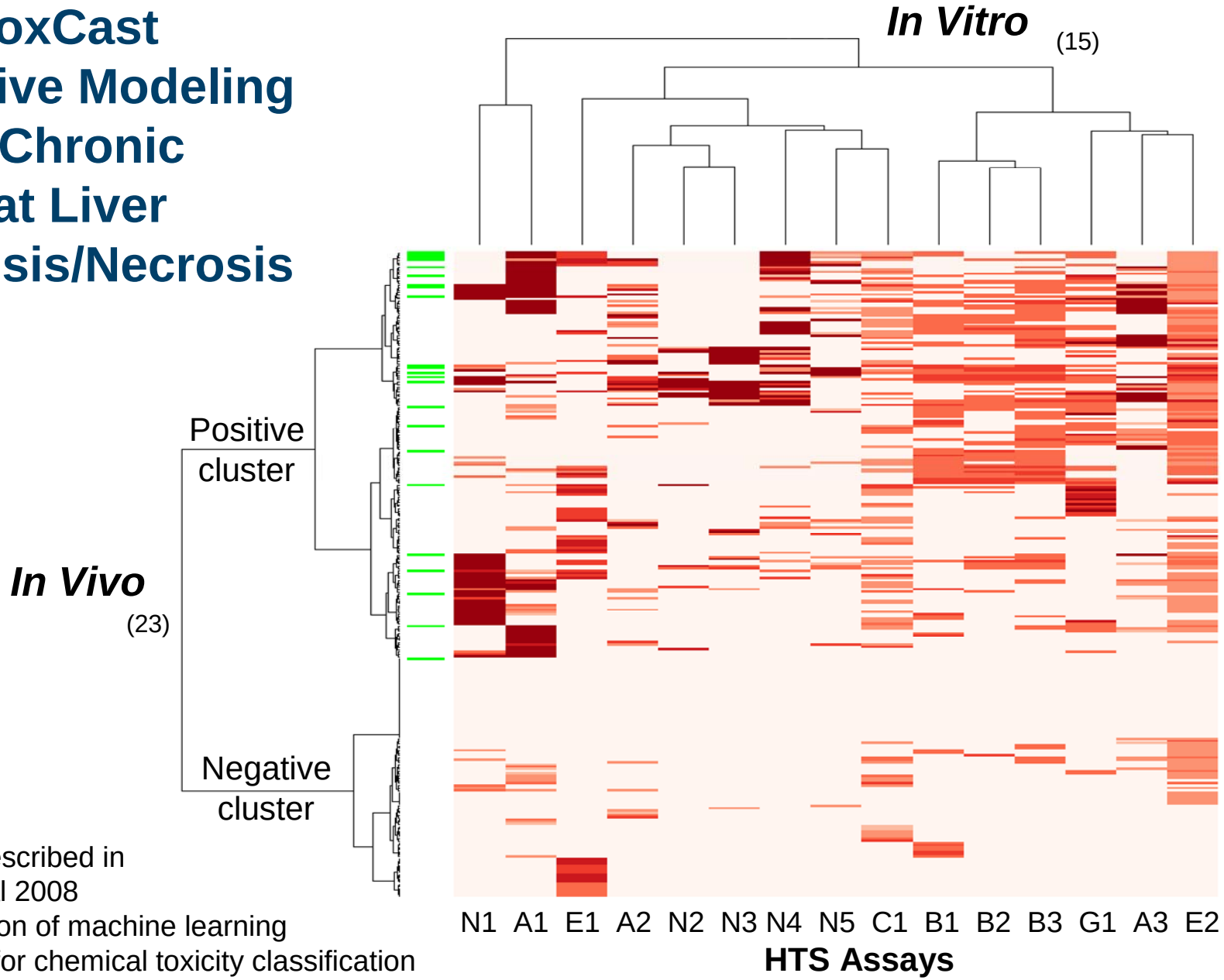
Nine contracts p ocurement; hundreds
ochemical, cel nd genomic assays; mo
anisms; and the capacity to screen up to 10,000 c

ToxCast Phase I Assays/Datasets

- *NR/transcription factors (Attagene, NCGC)*
- *Enzyme inhibition/receptor binding HTS (Novascreen)*
- *Cellular impedance (ACEA)*
- *Complex cell interactions (BioSeek)*
- *Hepatocelluar 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

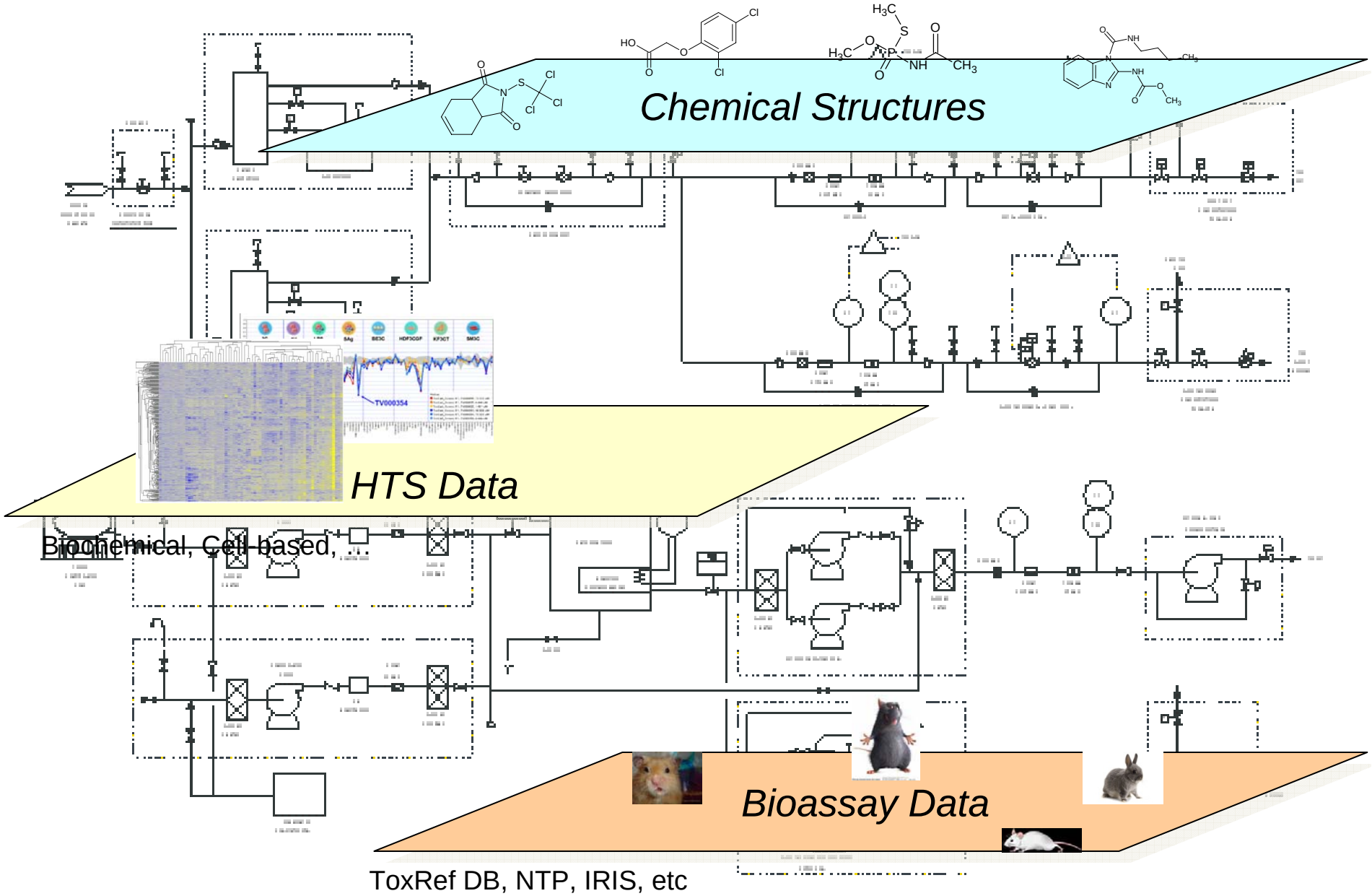


Methods described in
Judson et al 2008
A comparison of machine learning algorithms for chemical toxicity classification using a simulated multi-scale data model.
BMC Bioinformatics 9:241

Part III

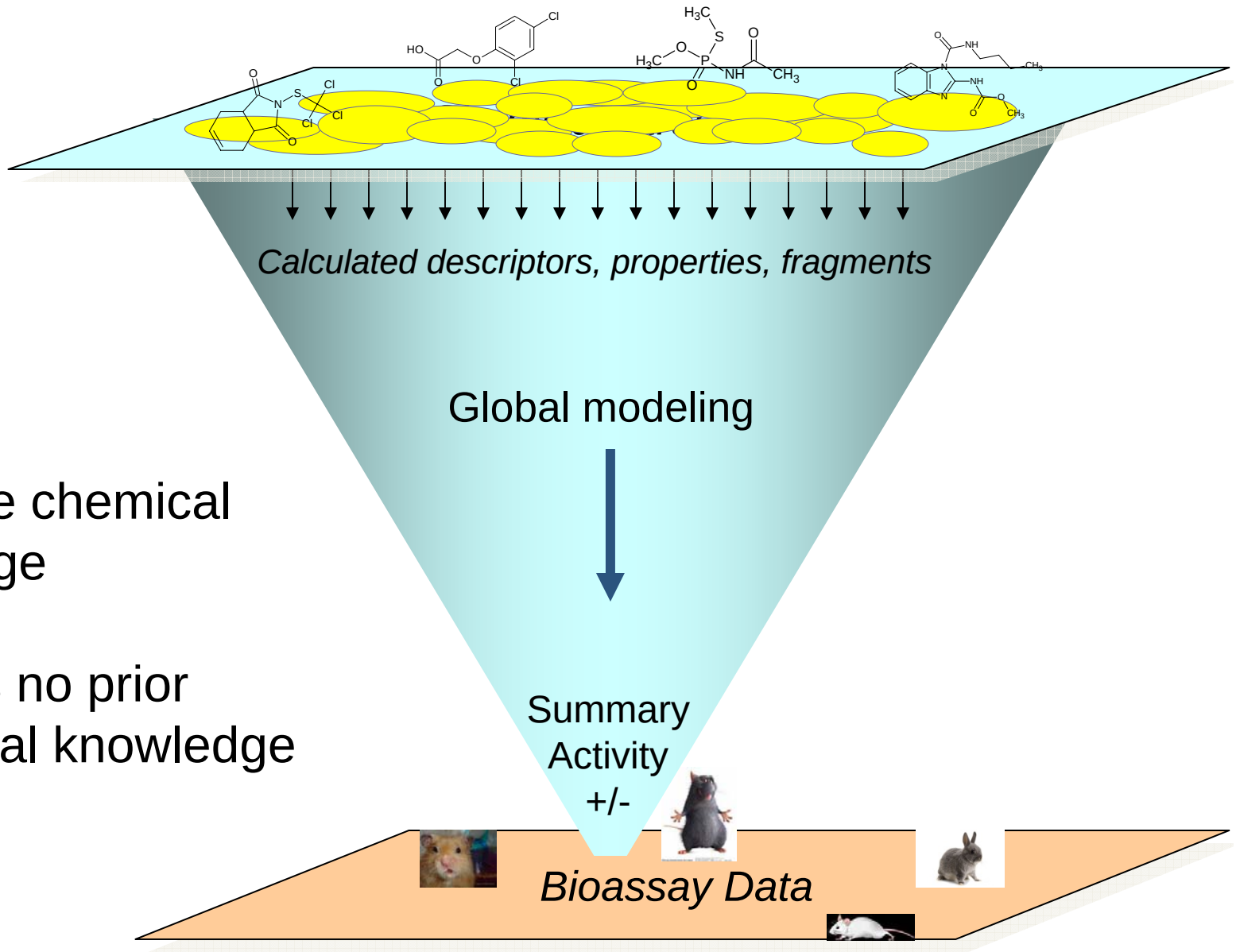
Incorporating SAR Concepts into ToxCast

ToxCast: Multidimensional Data



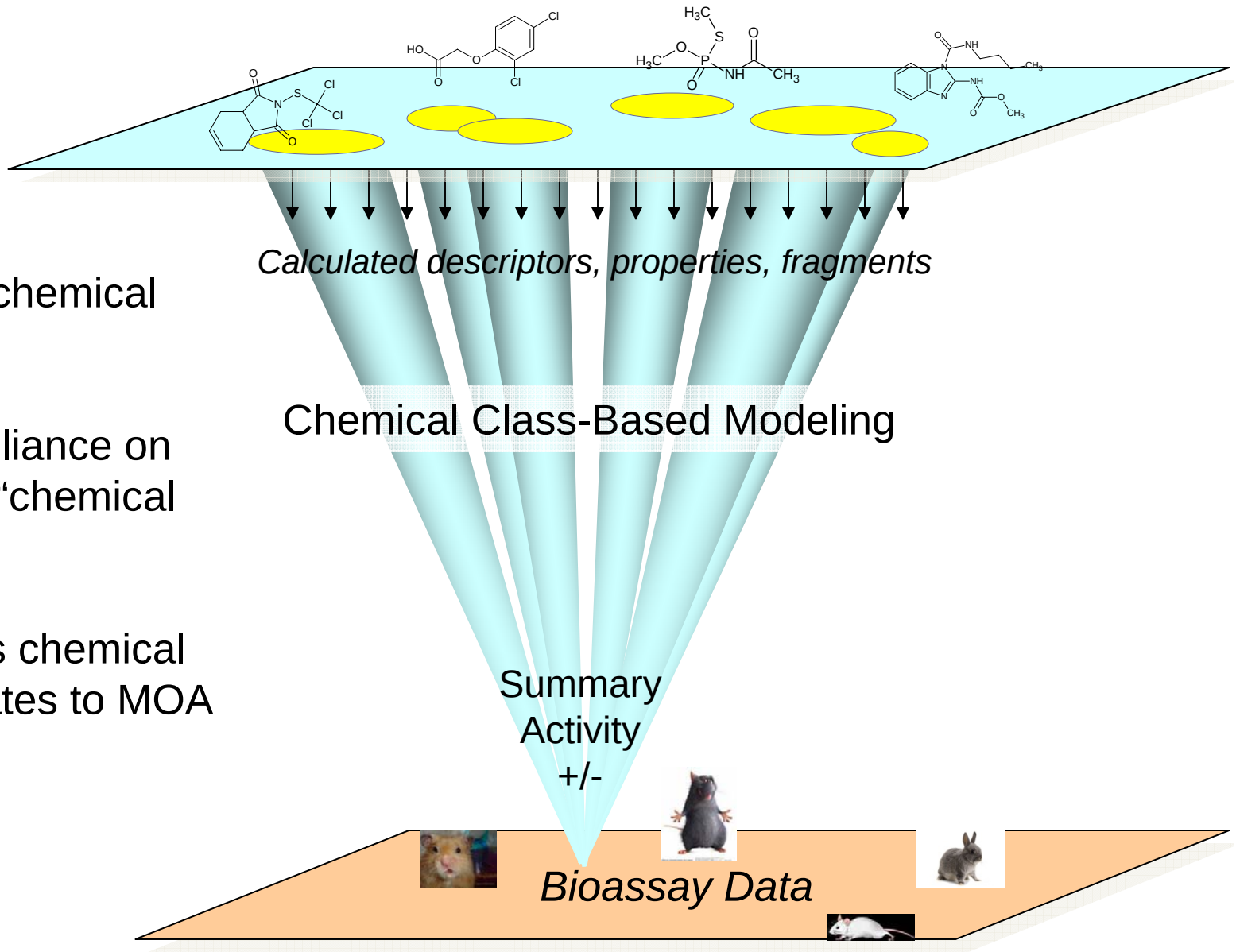
Structure-Activity Approaches to Toxicity Prediction

- Large chemical coverage
- Uses no prior chemical knowledge

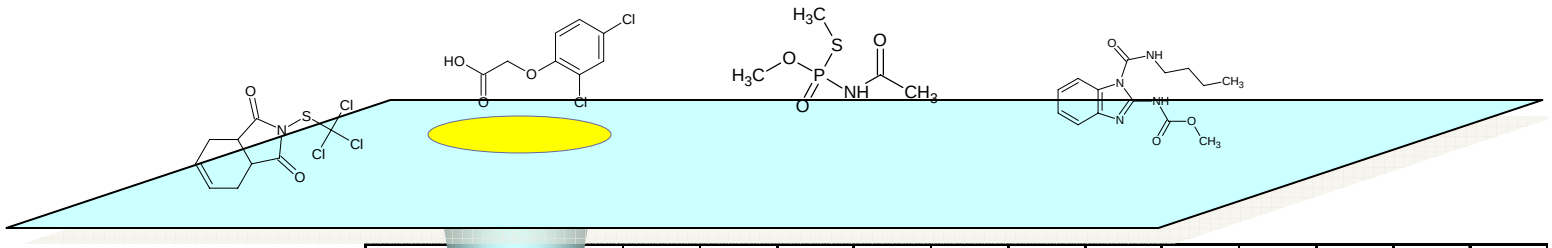


Structure-Activity Approaches to Toxicity Prediction

- Limited chemical coverage
- Heavy reliance on perceived “chemical similarity”
- Assumes chemical class equates to MOA

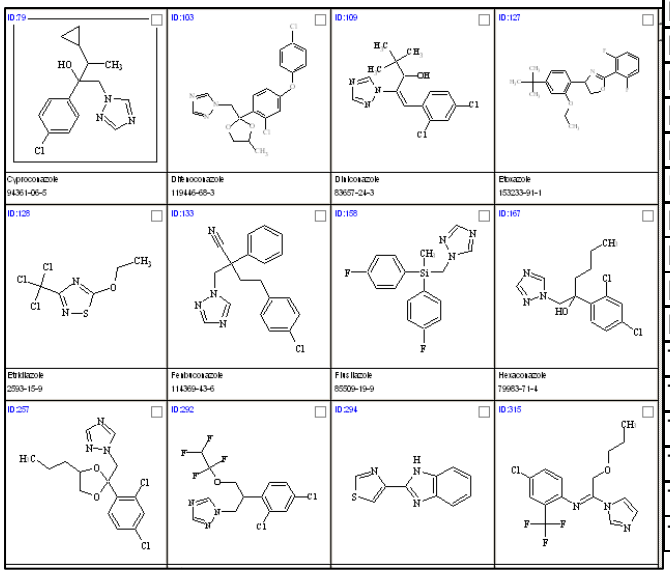


Bioactivity Profile of Structure Class



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
Difenoconazole	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**

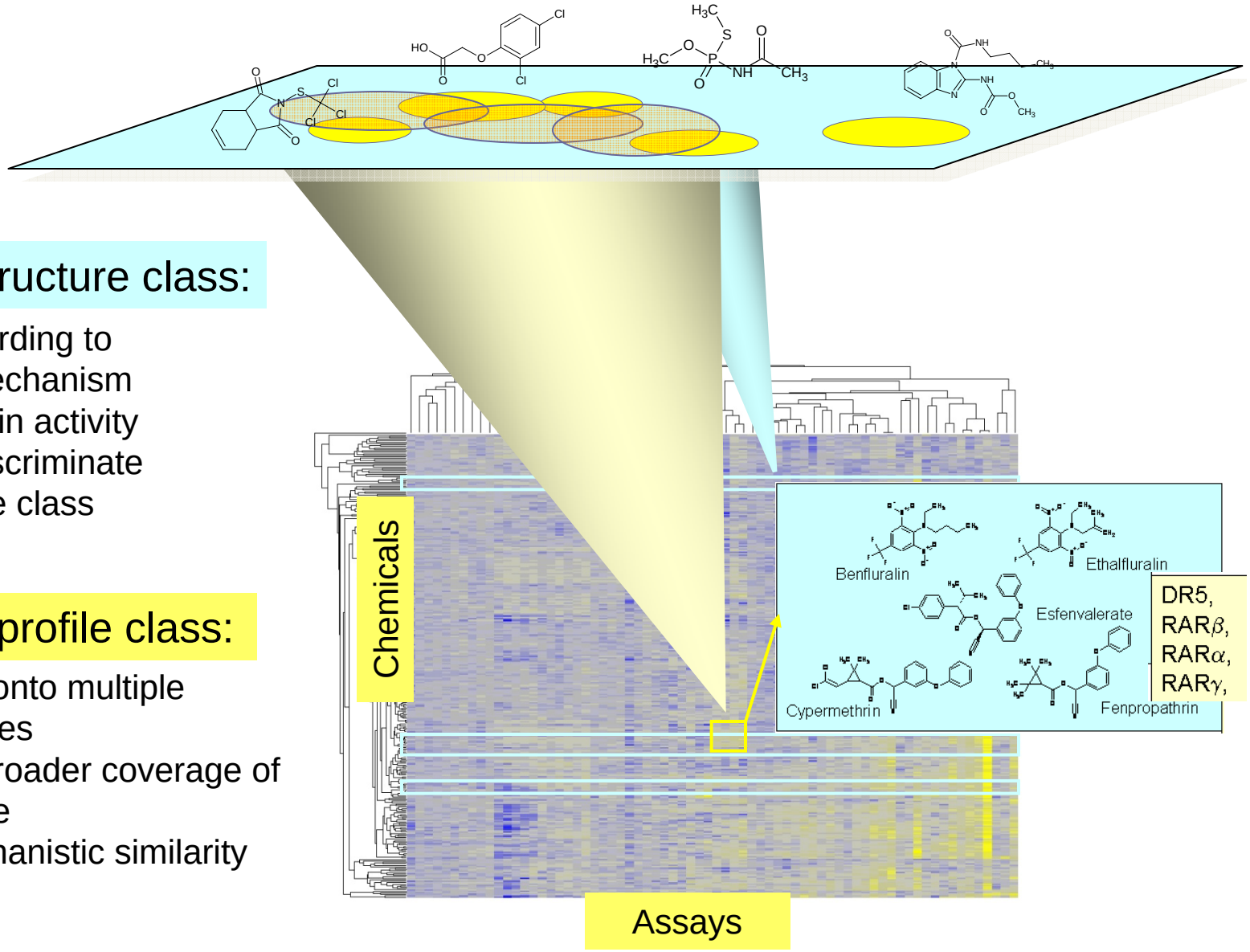
Structure Class vs Bioactivity Class

Chemical structure class:

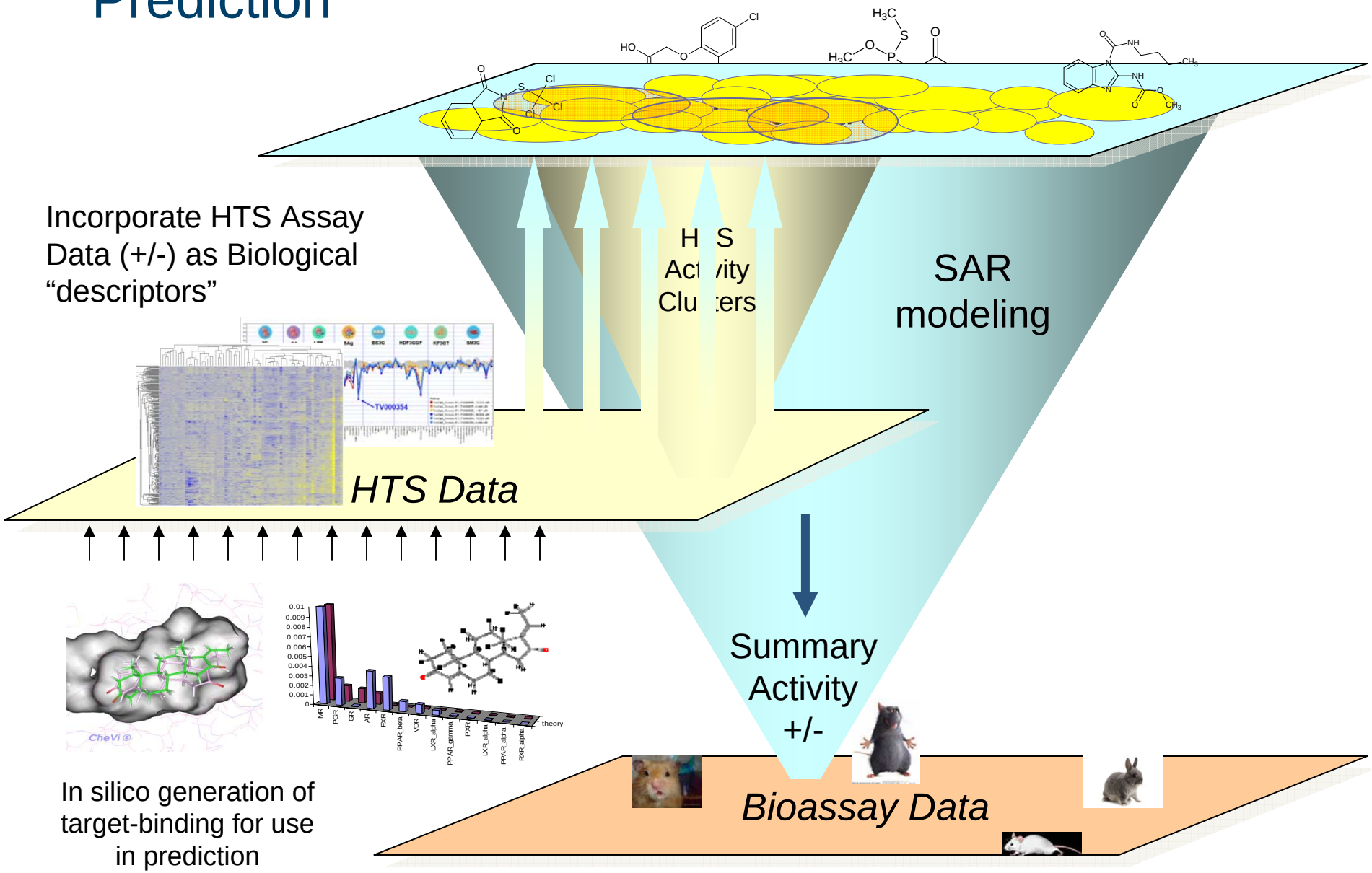
- Cluster according to activity and mechanism
- Differences in activity profiles can discriminate within structure class

Bioactivity profile class:

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

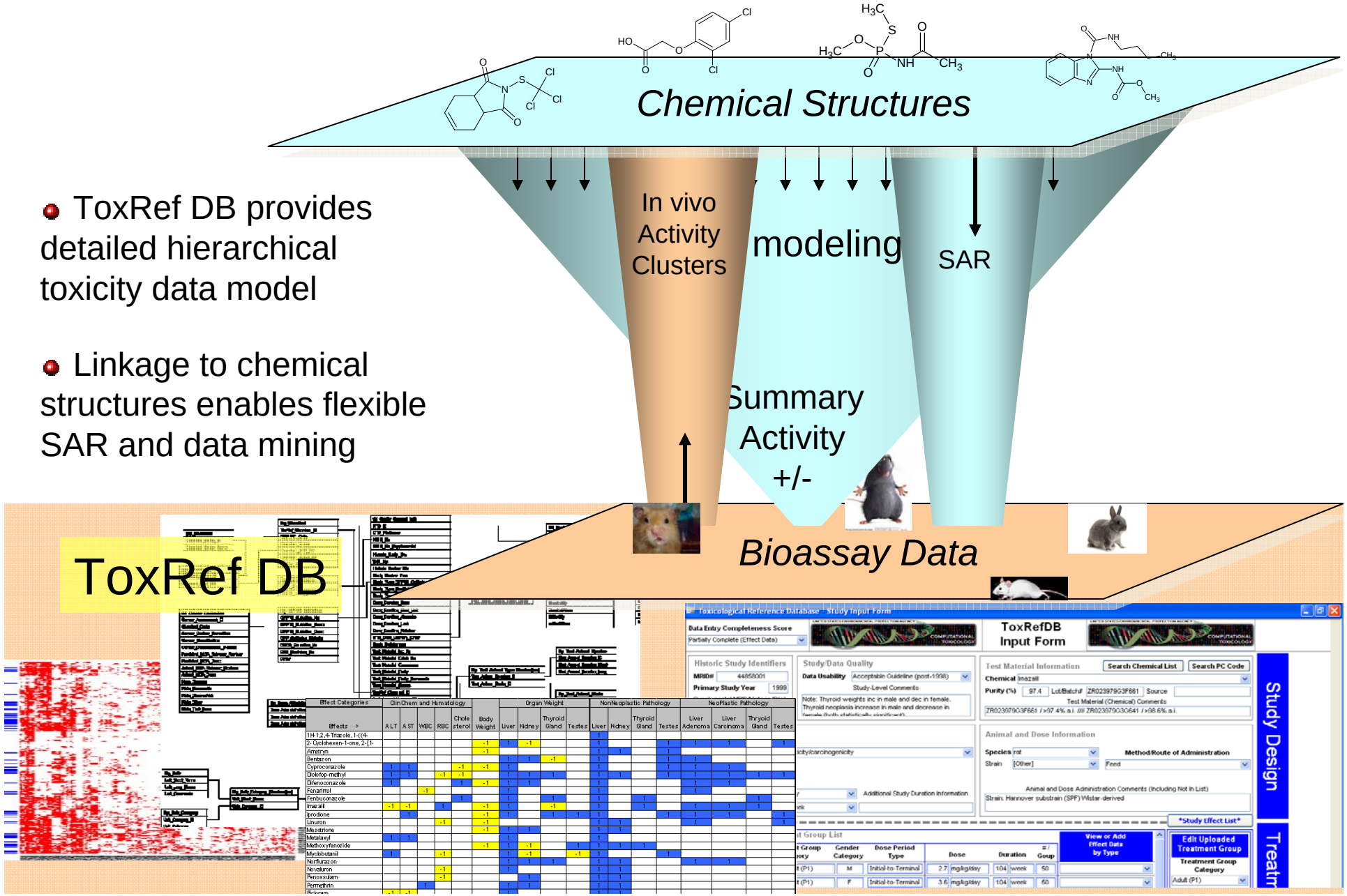


Structure-Activity Approaches to Toxicity Prediction



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

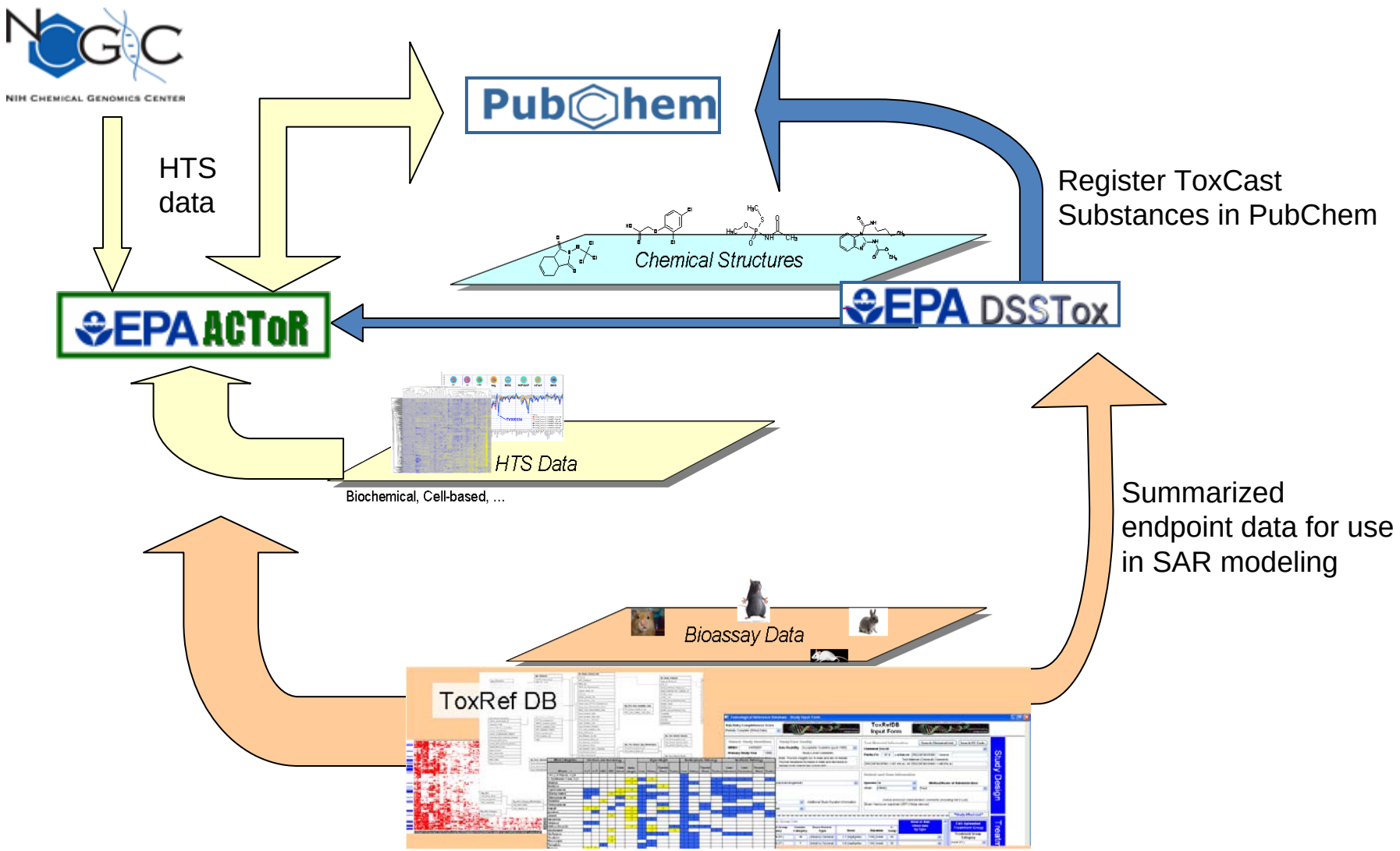
In a Brave New World -

- 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?

Bio

ToxRef DB, NTP, IRIS, etc

ToxCast: Data Publication & Exploration

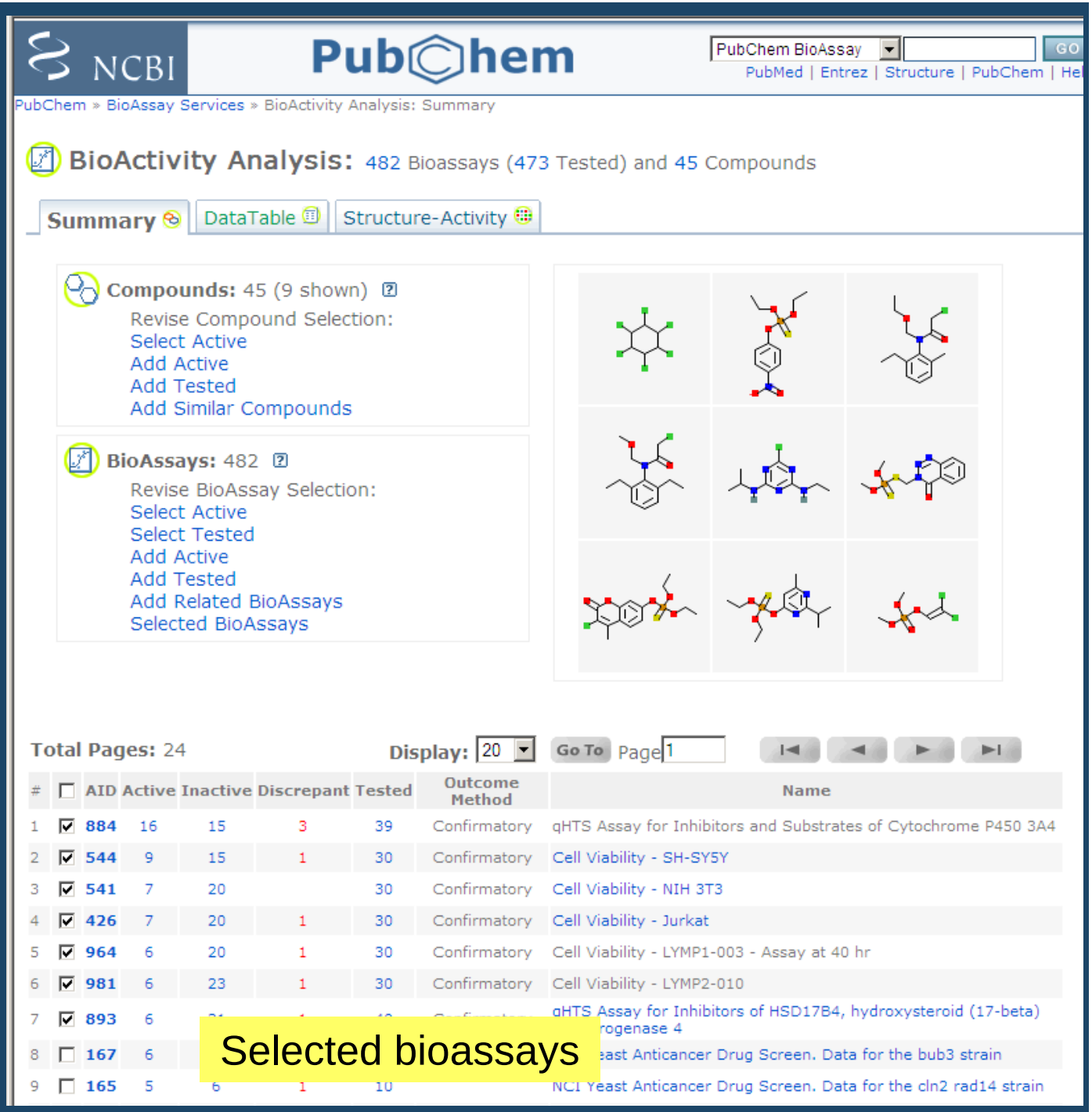


ToxCast_320

Bioactivity Analysis:

Retrieve all bioassay data in PubChem for ToxCast_320

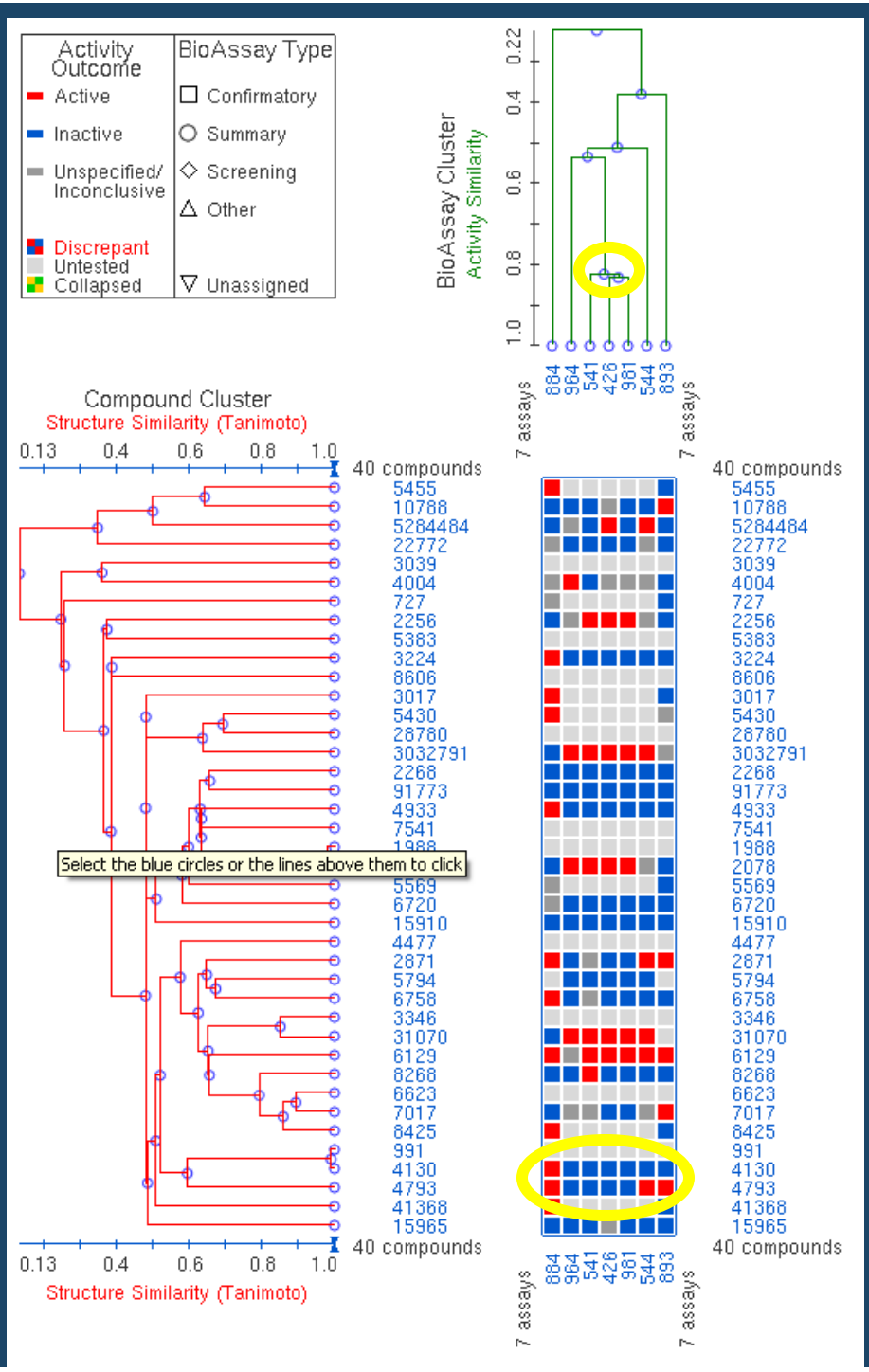
482 Bioassays
45 Compounds



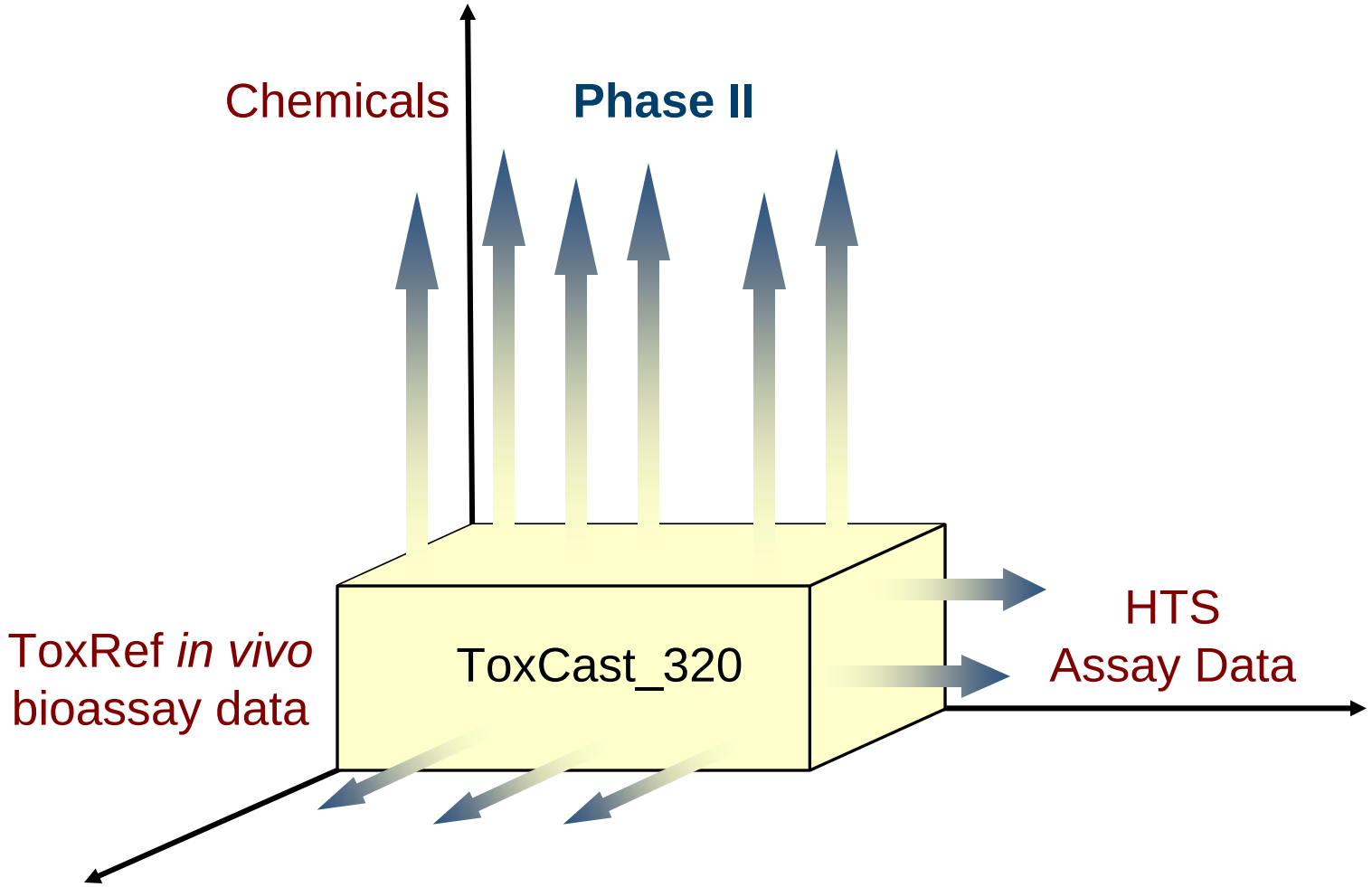
Structure-Activity *Bioactivity Analysis:*

7 bioassays, 45 Actives

View Bioassay Profile by Structure Similarity Cluster



ToxCast Phase I: Proof of Concept

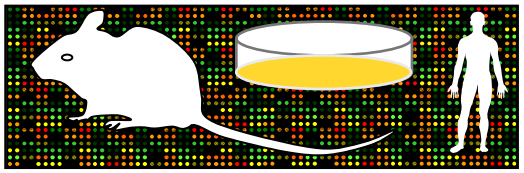


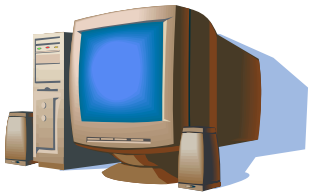
Phased Development of ToxCast

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
IIa	>300	Data Rich Chemicals	Validation	>400	$15-20k	FY09
IIb	>100	Known Human Toxicants	Extrapolation	>400	$15-20k	FY09
IIc	>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

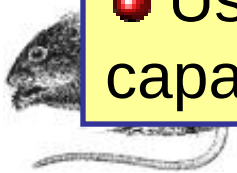




National Health and Enviro[illegible]

National Center for [illegible]cology

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

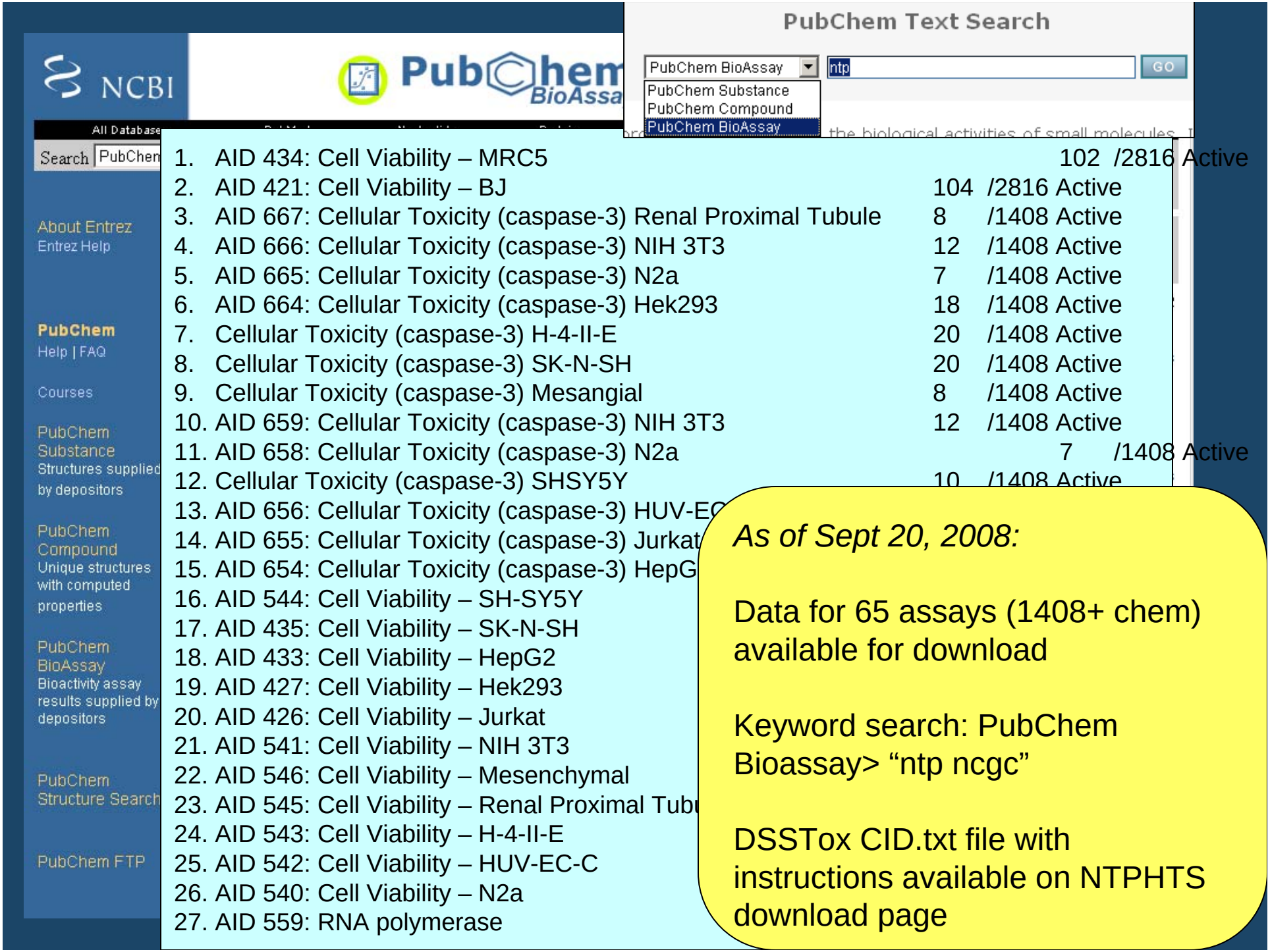


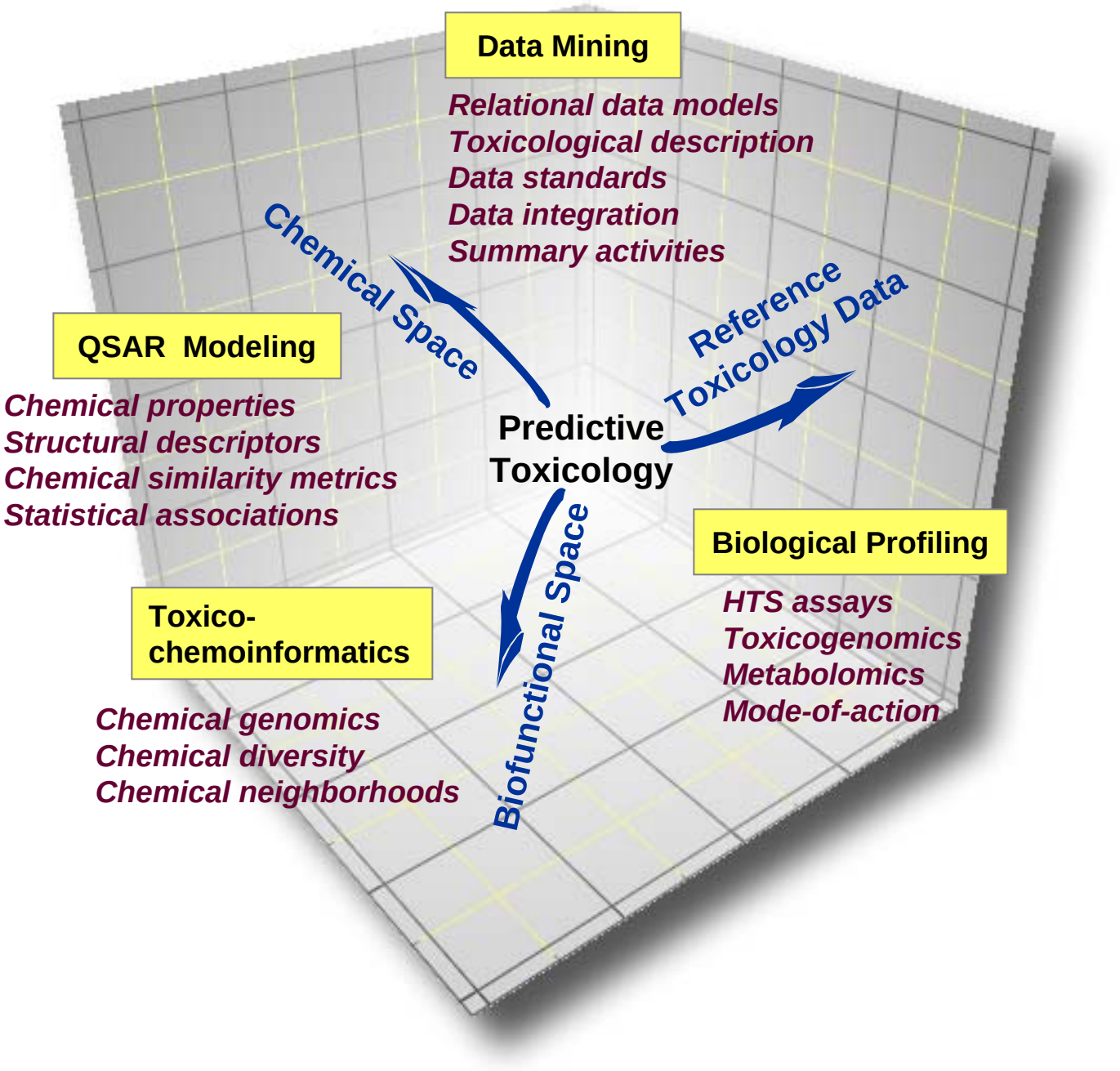
National Toxicology Program
Department of Health and Human Services

Biomolecular Screening Branch

NIH Chemical Genomics Center

Toxicology Project Team


NCBI
PubChem BioAssa
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
ntp
GO
PubChem Substance
PubChem Compound
PubChem BioAssay
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
14. AID 655: Cellular Toxicity (caspase-3) Jurkat
15. AID 654: Cellular Toxicity (caspase-3) HepG
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 Tubu
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


Data Mining
Relational data models
Toxicological description
Data standards
Data integration
Summary activities
Chemical Space
Reference Toxicology Data
QSAR Modeling
Chemical properties
Structural descriptors
Chemical similarity metrics
Statistical associations
Predictive Toxicology
Biofunctional Space
Biological Profiling
HTS assays
Toxicogenomics
Metabolomics
Mode-of-action
Toxico-chemoinformatics
Chemical genomics
Chemical diversity
Chemical neighborhoods

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), Jennifer Fostel (NIEHS)
- **PubChem Project:** Steve Bryant, Yanli Wang
- **Carcinogenic Potency Project:** Lois Gold
- **NIH Chemical Genomics Center:** Chris Austin, Noel Southerland