

Learning by Doing – Phase I of ToxCast

Robert Kavlock
Director, NCCT

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY



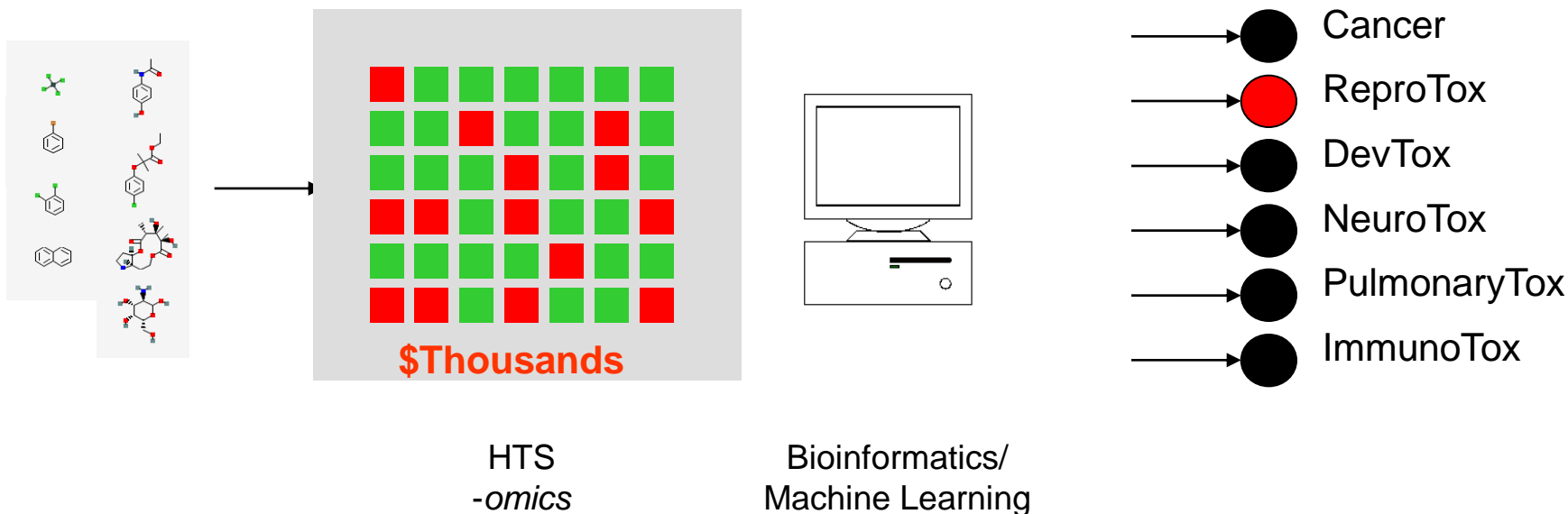


“...to integrate modern computing and information technology with molecular biology to improve Agency prioritization of data requirements and risk assessment of chemicals”

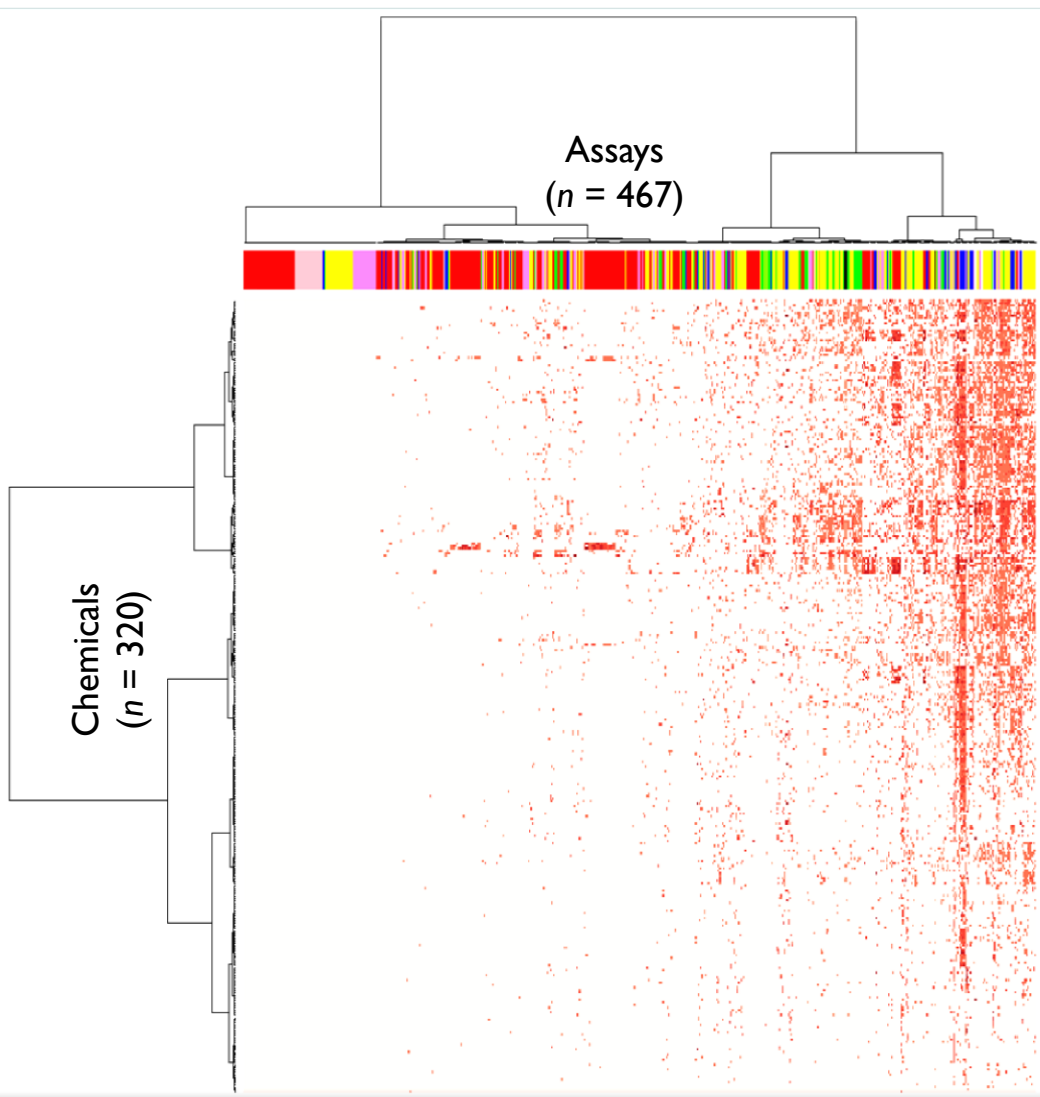
***Providing Decision Support Tools for
High-Throughput Screening, Risk Assessment and Risk
Management***

ToxCast Bioactivity Profiling

in vitro testing *in silico* analysis



ToxCast Phase I HTS Results



Cellular Assays

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

Biochemical Assays

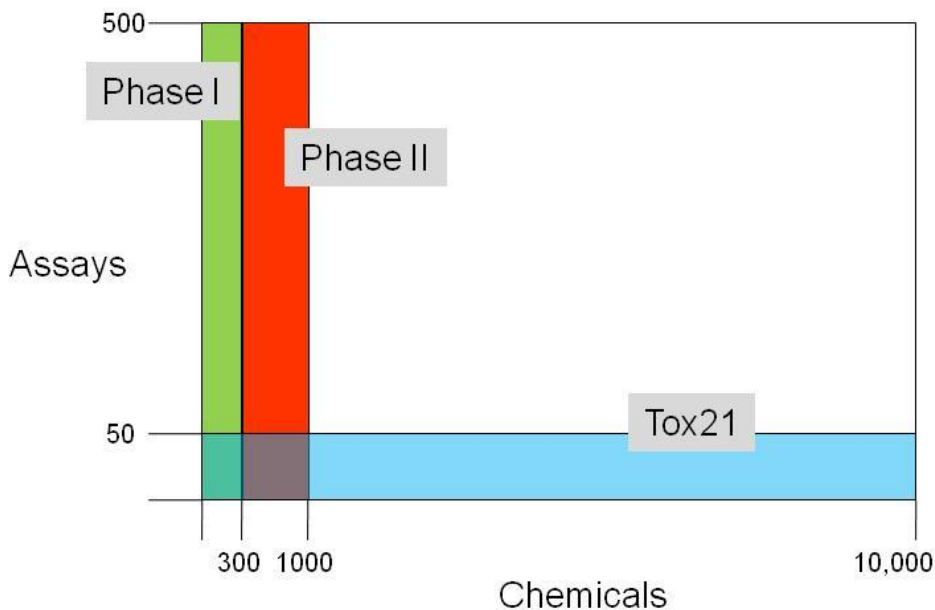
- **Protein families**
 - GPCR
 - NR
 - Kinase
 - Phosphatase
 - Protease
 - Other enzyme
 - Ion channel
 - Transporter
- **Assay formats**
 - Radioligand binding
 - Enzyme activity
 - Co-activator recruitment

Judson *et al* EHP 118:485-492 (2010)

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

From Phase I to Phase II and Tox21

	Phase I	Phase II	Tox21
Actives	272	120	700
Inerts	24	100	1000
Antimicrobials	33	100	500
HPV	35	170	1300
MPV	7	60	1500
Green	4	60	500
PCCL	73	150	500
Pharmaceuticals	0	100	2500
Consumer Products /Food additives	0	0	1500
Total	309	~700	~10000



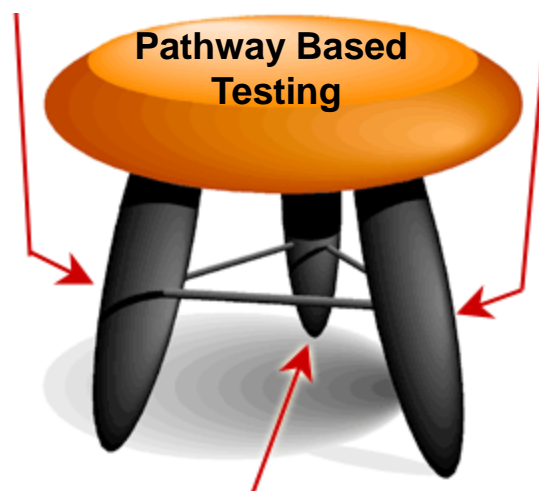
Overview

- Information Management
- Transparency
- Unexpected Results
- Team Effort

Enabling the Transformation in Toxicology

High Throughput Screening

Information Technology and Management



Molecular, Cellular and Systems Biology

Databases & Tools

- DSSTox – Distributed Structure Toxicity Data
- ACToR – Aggregated Computational Toxicology Database
- ToxRefDB – In vivo guideline study data
- ToxMiner – ToxCast data and analyses
- EDB – NCCT Endocrine Database
- VTKB – Virtual Tissue Knowledgebase
- Chemical Registry

ToxCast/Tox21: Chemical Annotation & Data Publication



HTS
data



Cheminformatics Needs:

- ▶ Select chemicals for testing
 - ✓ high environmental / tox interest
 - ✓ suitable for testing
 - ✓ chemically diverse
 - ✓ analogs, metabolites, etc
- ▶ Standardized chemical structure/substance annotation across all testing & reference inventories
- ▶ Chemical annotation QC
- ▶ Analytical QC
- ▶ Public reporting of data & QC

Annotation
C
king

Standardized
data for
SAR
g

DSSTox: Distributed Structure-Searchable Database Network Project...& ToxCast™

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 Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network

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

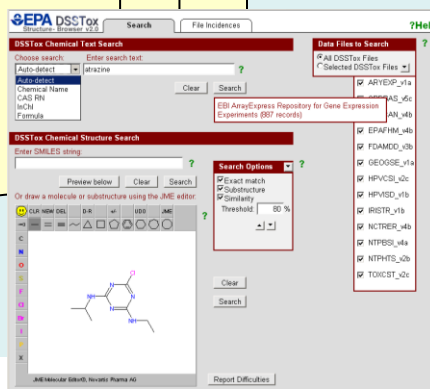
Distributed Structure-Searchable Toxicity (DSSTox) Database Network is a project of EPA's Computational Toxicology Program, helping to build a public data foundation for improved structure-activity and predictive toxicology capabilities. The DSSTox website provides a public forum for publishing downloadable,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

Chemical Structures + **Toxicity Data**

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

External links: PubChem, ChemSpider, Lazar, ACToR

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



- Chemical information QA and structure annotation
 - ToxCast Phase I & II
 - ToxRefDB
 - Tox21
- Primary source of QC'd structures for ACToR
- Facilitate external linkages and data publication
- Publish summary activities and chemical classifiers for modeling
- PubChem Source depositor for structures & "assays"

Current Guidelines & Chemical Standards in Toxicology & Biological Effects Publishing:

Archives of Toxicology
Brain Research Bulletin
Ecotoxicol. Environ. Saf.
Environ Health Perspect
Environ.Sci. Technol.
Environ.Toxicol.Pharmacol.
Environ & Molec Mutagenesis
Environmental Toxicology
International J. of Toxicology
J.Appl.Toxicol.
J Toxicology & Enviro.Health Part A
Neurotoxicology & Teratology
Reproductive Toxicology
Toxicol. Appl. Pharmacol.
Toxicological Sciences
Toxicologic Pathology
Toxicology Letters
Toxicology

*Independent
Chemical QC*

*Chemical
structures &
hyperlinks*

*Systematic
names
encouraged*

Nature
Chem. Biol.

Chem. Res.Toxicol.

Author responsibility

“Authors should provide sufficient detail to allow the work to be reproduced.”

Nature Chemical Biology **3**, 297 (2007)
doi:10.1038/nchembio0607-297

A new look for chemical information

Nature Chemical Biology is committed to enhancing interdisciplinary communication and features online content to increase the accessibility of chemical information for our readers.

“Chemical compound information

For all significant compounds included in Nature Chemical Biology original research papers, a compound data page, linked directly from the compound reference in the full text, appears in the online journal.

Compound Data Index

From the following article

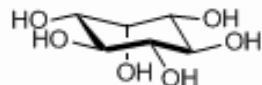
[Small molecules enhance autophagy and reduce toxicity in Huntington's disease models](#)

Sovan Sarkar, Ethan O Perlstein, Sara Imarisio, Sandra Pineau, Axelle Cordenier, Rebecca L Maglathlin, John A Webster, Timothy A Lewis, Cahir J O'Kane, Stuart L Schreiber & David C Rubinsztein

Nature Chemical Biology **3**, 331-338 (2007) Published online: 7 May 2007

doi:10.1038/nchembio883

A PubChem link allows users of Nature Chemical Biology to go, in a single click, from the mention of a molecule in a paper to a rich and growing collection of information about chemical structures and their biological assay results, hosted by the NCBI.



[Compound 1](#)

Inositol

[View in PubChem](#)

[View compound page \(2 KB\)](#) | [View in 3D \(2 KB\)](#) | [Download ChemDraw file of structure \(2 KB\)](#)

□
“Authors ensure that the chemical compound information within their papers is complete, scientifically accurate and appropriately formatted.”

ACToR

Aggregated Computational Toxicology Resource

U.S. ENVIRONMENTAL PROTECTION AGENCY

ACToR: Aggregated Computational Toxicology Resource

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

You are here: [EPA Home](#) > [ACToR](#) > Data Collection

Data Collection: EPA CCL3

Name: [EPA CCL3 List.pdf](#)

Description: EPA has drinking water regulations for more than 90 contaminants. The Safe Drinking Water Act (SDWA) includes a process that we must follow to identify and list unregulated contaminants which may require a national drinking water regulation in the future. EPA must periodically publish this list of contaminants (called the Contaminant Candidate List or CCL). In February 2008 we announced the draft CCL 3.

ID: 139

Institutional Source: EPA



Source Type: Chemicals

Number of Substances: 93

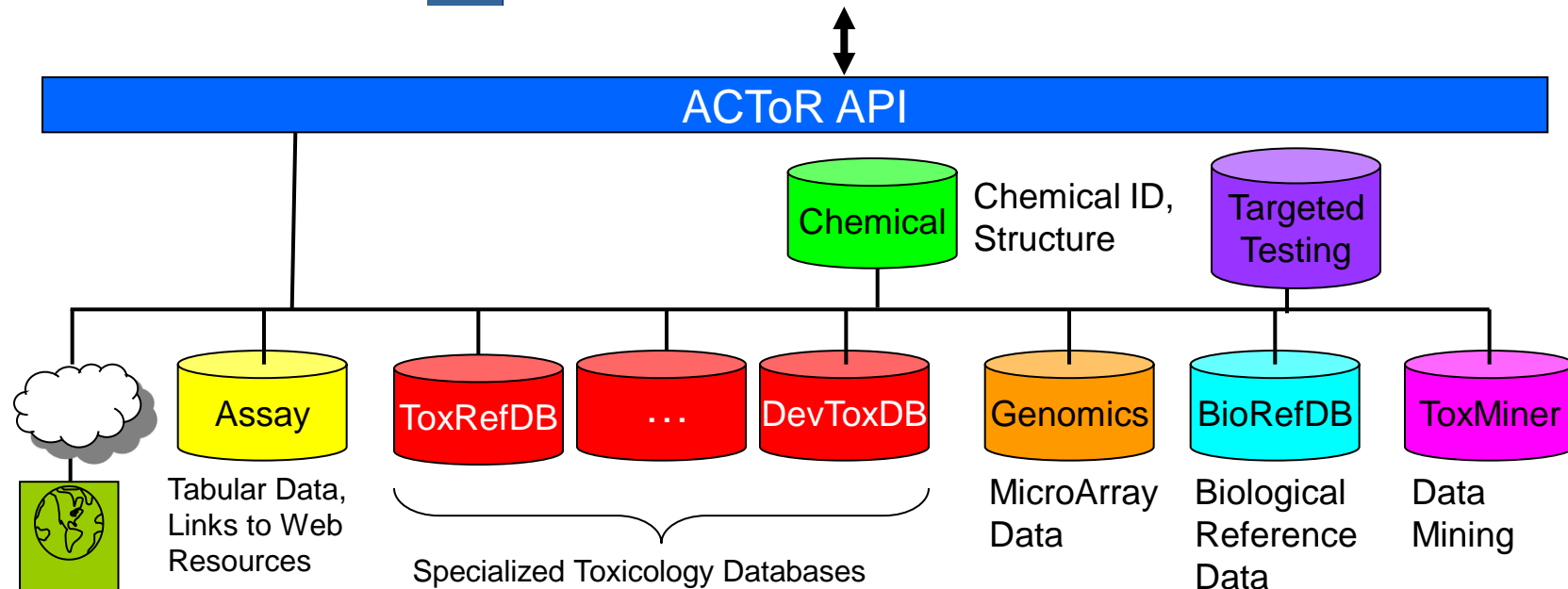
Number of Generic Chemicals: 92

Chemical Table

Page 1 of 2: [List](#)

Structure	Name	CASRN	Generic Chemical Details	Hazard	Contaminant	Developmental Toxicity	Reproductive Toxicity	Chronic Toxicity	Food Safety
	1,1,1,2-Tetrachloroethane	630-20-6	Details	Ha	Ca	G	D	R	Cr
	1,1-Dichloroethane	75-34-3	Details	Ha	Ca	G	D	R	Cr

ACToR Web
Browser



<http://actor.epa.gov/>

The Reach of ACToR

Category	Count
Data Collections	580
Substances	2,139,297
Compounds	1,044,834
Generic Chemicals	546,956
Generic Chemicals with Structure	438,756
Assays	3,213
Assay Components	7,221
Assay Results	6,662,296

Chemical List View

Red box indicates that data is available for that phenotype, not that chemical causes that phenotype

Home | ACToR | US EPA - Mozilla Firefox

http://actorpreprod.epa.gov/actor/faces/ACToRHome.jsp

U.S. ENVIRONMENTAL PROTECTION AGENCY

ACToR: Aggregated Computational Toxicology Resource

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

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

Home

ACToR is one of many EPA tools available for those interested in chemical toxicity to find data about potential chemical risks to human health and the environment. The data warehouse:

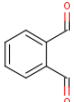
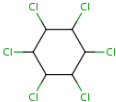
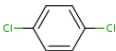
- Is a collection of databases collated or developed by the US EPA National Center for Computational Toxicology (NCCT).
- Aggregates data from over 500 public sources on over 500,000 environmental chemicals searchable by chemical name, other identifiers, and by chemical structure.
- Data includes chemical structure, physico-chemical values, in vitro assay data and in vivo toxicology data.
- Chemicals include, but are not limited to, high and medium production volume industrial chemicals, pesticides (active and inert ingredients), and potential ground and drinking water contaminants.
- Provides a connection to another EPA chemical screening tool called [ToxCast](#), a multi-year, multi-million dollar effort that uses advanced science tools to help efficiently understand biological processes impacted by chemicals that may lead to adverse health effects.

Chemical Name Parameters ☒ Search on Chemical Names ☐ Search on CAS Numbers

Match by ☐ Exact ☒ Any

Enter Chemical Name:

Search Results

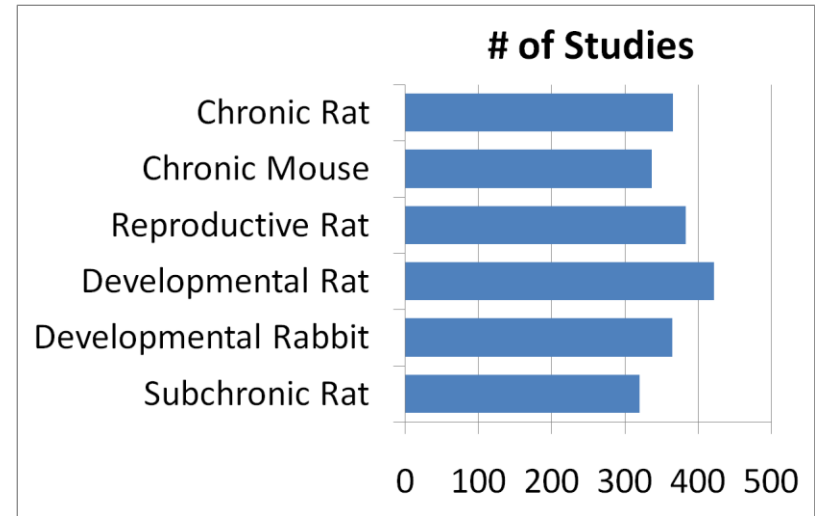
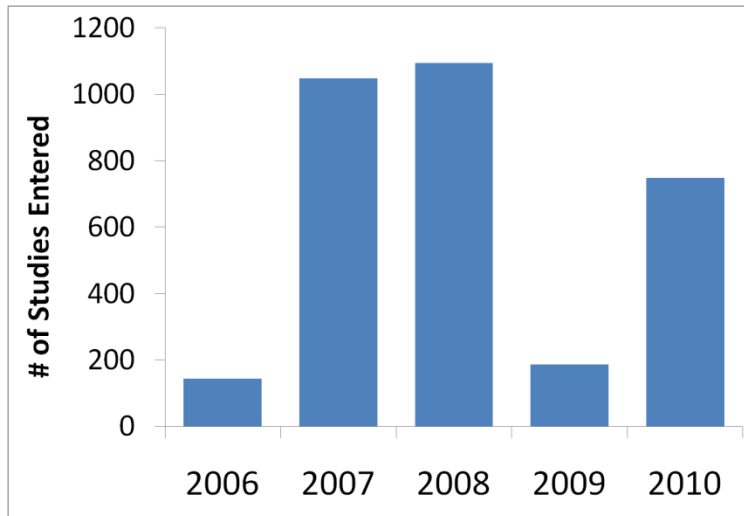
Details	Image	CASRN	Preferred Name	Chronic Hazard	Toxicity	Carcinogenicity	Genotoxicity	Developmental Toxicity	Reproductive Toxicity	Food Safety	Exposure
details		643-79-8	1,2-Benzenedicarboxaldehyde	Ha				D			Ex
details		608-73-1	Hexachlorocyclohexane	Ha	Cr	Ca	G	D		FS	Ex
details		106-46-7	1,4-Dichlorobenzene	Ha	Cr	Ca	G	D	R	FS	Ex

Find: dsstox ☐ Match case

Done

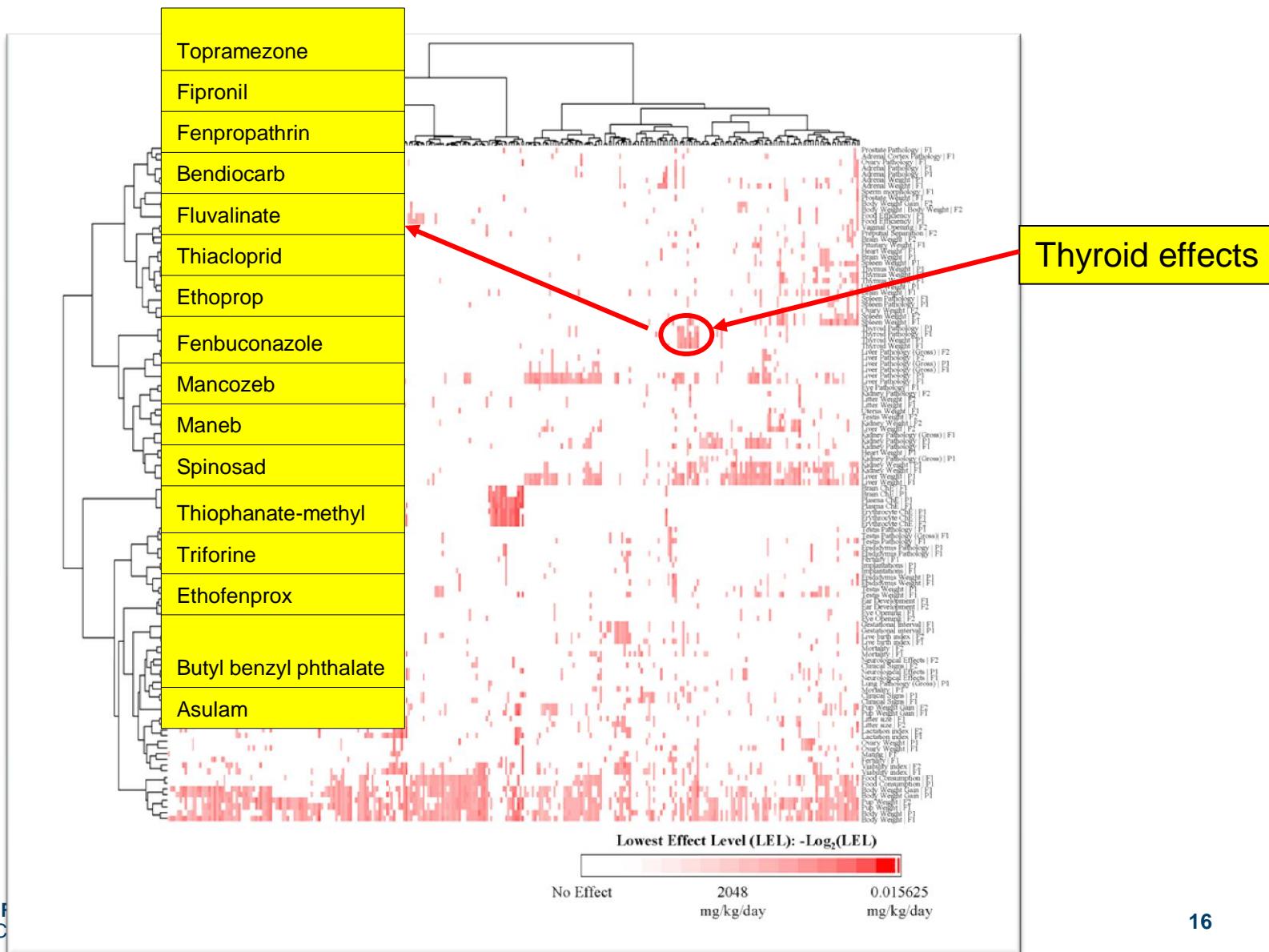
ToxRefDB

- Developed in 2006
 - Database Structure
 - Database Vocabulary
- Manual entry of >3200 studies for > 500 chemicals



- Datasets released publically in 2009 (3 publications)
- Full database & Web Interface released publically in 2010

ToxRefDB: A Computable Database



NCCT Endocrine Database

ER Summary | NCCT EDB | US EPA - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://nitan.epa.gov:22722/servlet/EDB

Most Visited Getting Started Latest Headlines PubMed home Google Scholar EPA Desktop Library Home | ACToR | US E... Oxford Journals | Life... ER Summary | NCCT ... Data Collection List | ...

Data Collection List | ToxMiner | US EPA ER Summary | NCCT EDB | US EPA About | ACToR | US EPA richard brennan microarray carcinogen ... EWG || Human Toxome Project

ToxCast Phase I Endocrine Activity Database [Share](#)

[Recent Additions](#) | [Contact Us](#) Search: ☐ All EPA ☒ This Area

You are here: [EPA Home](#) » [NCCT EDB](#) » ER Summary

Summary of ToxCast Phase I Estrogen Receptor Assays vs. Literature Results

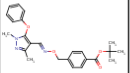
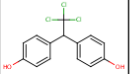
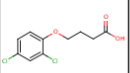
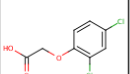
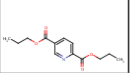
Legend:

Green: Consistent positive findings

Red: Consistent negative findings

Yellow: Some positive, some neagative findings

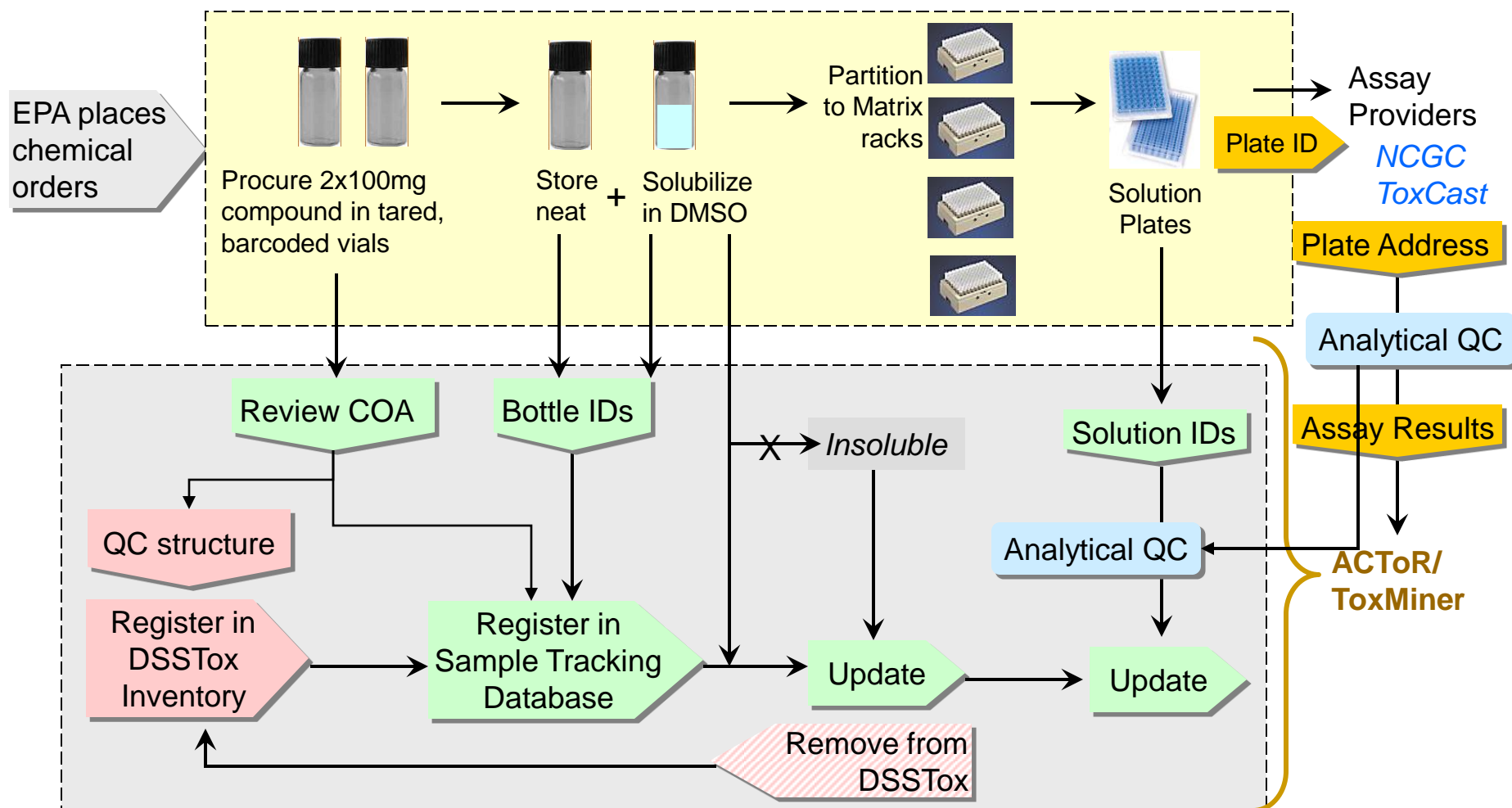
White: No literature findings

Details	Structure	CASRN	Name	source_name_sid	NVS Binding ER Human	NVS Binding ER Bovine	ATG ERa TRANS	ATG ERc CIS	NCGC ER Agonist	NCGC ER Antagonist	Lit. Binding	Lit. Reporter Gene	Lit. Proliferation	Lit. Uterotrophic
Results		111812-58-9	(Z,E)-Fenpyroximate	DSSTOX_40463	Bnd	Bnd	RG	RG	RG	RG				
Results		2971-36-0	2,2-Bis(4-hydroxyphenyl)-1,1,1-trichloroethane (HPTE)	DSSTOX_40307	Bnd	Bnd	RG	RG	RG	RG	Bnd	RG		Ut
Results		94-82-6	2,4-DB	DSSTOX_40326	Bnd	Bnd	RG	RG	RG	RG			Pr	
Results		94-75-7	2,4-Dichlorophenoxyacetic acid (2,4-D)	DSSTOX_40308	Bnd	Bnd	RG	RG	RG	RG	Bnd	RG	Pr	
Results		136-45-8	2,5-Pyridinedicarboxylic acid, dipropyl ester	DSSTOX_40310	Bnd	Bnd	RG	RG	RG	RG				

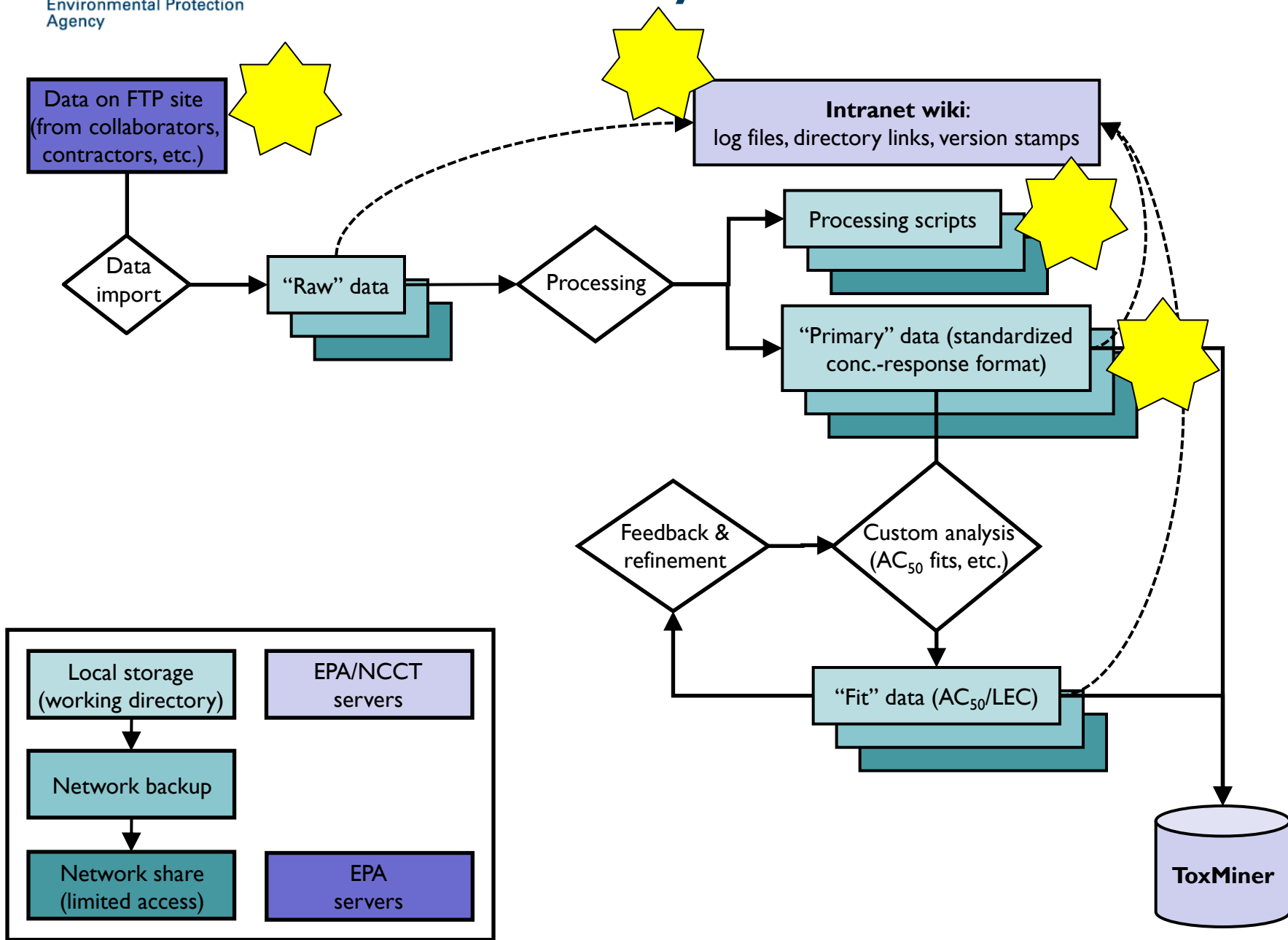
Find: dsstox ☐ Match case

Done

Chemical Sample Registration Workflow



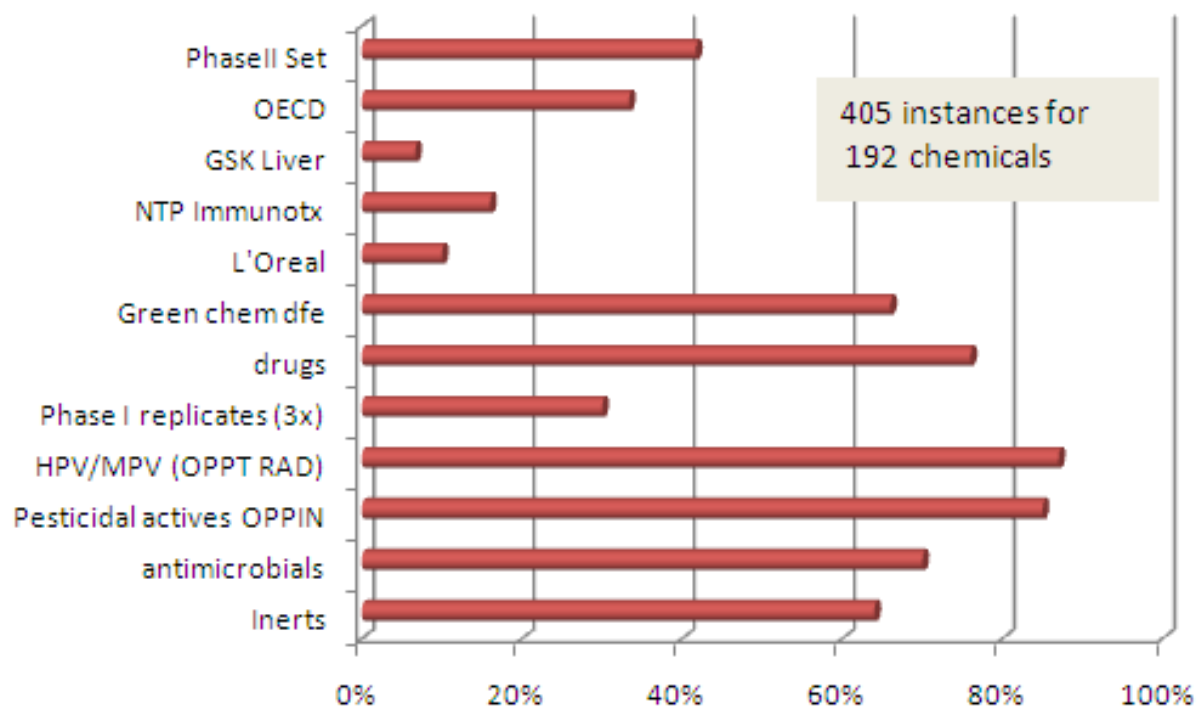
ToxCast Assay Data Workflow: Phase-II



Tox21 & ToxCast Phase II Chemicals

- 3800 chemicals “On hand”, registered or in process
- Solubilize_1 set (1616 compounds): **135 DMSO insolubles**
- Solubilize_2 set (1300 compounds): submitted 5/17/2010
- 114 pharma compounds (failed drugs) received and solubilized from Merck, GSK, Pfizer, Sanofi-Aventis
- 192 chemicals from Solubilize_1 set + 114 pharma compounds plated for Phase II
- 10 Phase II triplicates “On hand”, 3 included in 192

ToxCast Phase II Chemical Selection



- Phase II chemicals being selected from Tox21 library
- High-priority EPA lists + In vivo Data
- Pfizer, Merck, GSK chem w/clinical data (approx 110)
- Defined structures

Transparency

•Website

•Full data release

•Use of open source software

•R, MySQL, PERL

•Chemical Prioritization Community of Practice

•ToxCast Data Analysis Summit

•Communications

•Presentations, presentations, presentations

•Media releases

National Center for Computational Toxicology
U.S. ENVIRONMENTAL PROTECTION AGENCY

United States
Environmental Protection Agency

Home
Basic Information
Organization
Post Doc Profiles
Framework
Research Activities
ACTox
DSSTox
ExpoCast™
ToxCast™
ToxRefDB
V-Liver™
V-Embryo™
Conferences and Seminars
Products
Journal Articles
Book Chapters
Presentations
BOSC Information
EPA Communities of Practice
Jobs and Opportunities
Related Information

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

In 2007, EPA launched ToxCast™ to develop a cost-effective approach for efficiently prioritizing the toxicity testing of thousands of chemicals.

- Uses data from state-of-the-art high throughput screening (HTS) bioassays.
- Builds computational models to forecast potential chemical toxicity in humans.
- Provides EPA regulatory programs with science-based information helpful in prioritizing chemicals for more detailed toxicological evaluations and more efficient use of animal testing.
- Phase I profiled over 300 well-characterized chemicals (primarily pesticides) in over 400 HTS endpoints. Endpoints include biochemical assays of protein function, cell-based transcriptional reporter and gene expression, cell line and primary cell functional, and developmental endpoints in zebrafish embryos and embryonic stem cells.
- Phase I chemicals have already been tested using traditional toxicology methods including developmental toxicity, multi-generation reproductive studies, and sub-chronic and chronic rodent bioassays. [ToxCRefDB](#) is the relational database storing this information- nearly \$2 billion worth of animal toxicity studies.
- Phase II of ToxCast will screen additional chemical compounds representing broader chemical structure and use classes to evaluate the predictive toxicity signatures developed in Phase I.
- Toxicity signatures from ToxCast will be defined and evaluated by how well they predict outcomes from mammalian toxicity tests and identify toxicity pathways relevant to human health effects.
- Data sets and published research are listed in the table below. If a free electronic copy of a research publication is not available on the table below, please contact Monica.Linnenbrink@epa.gov.

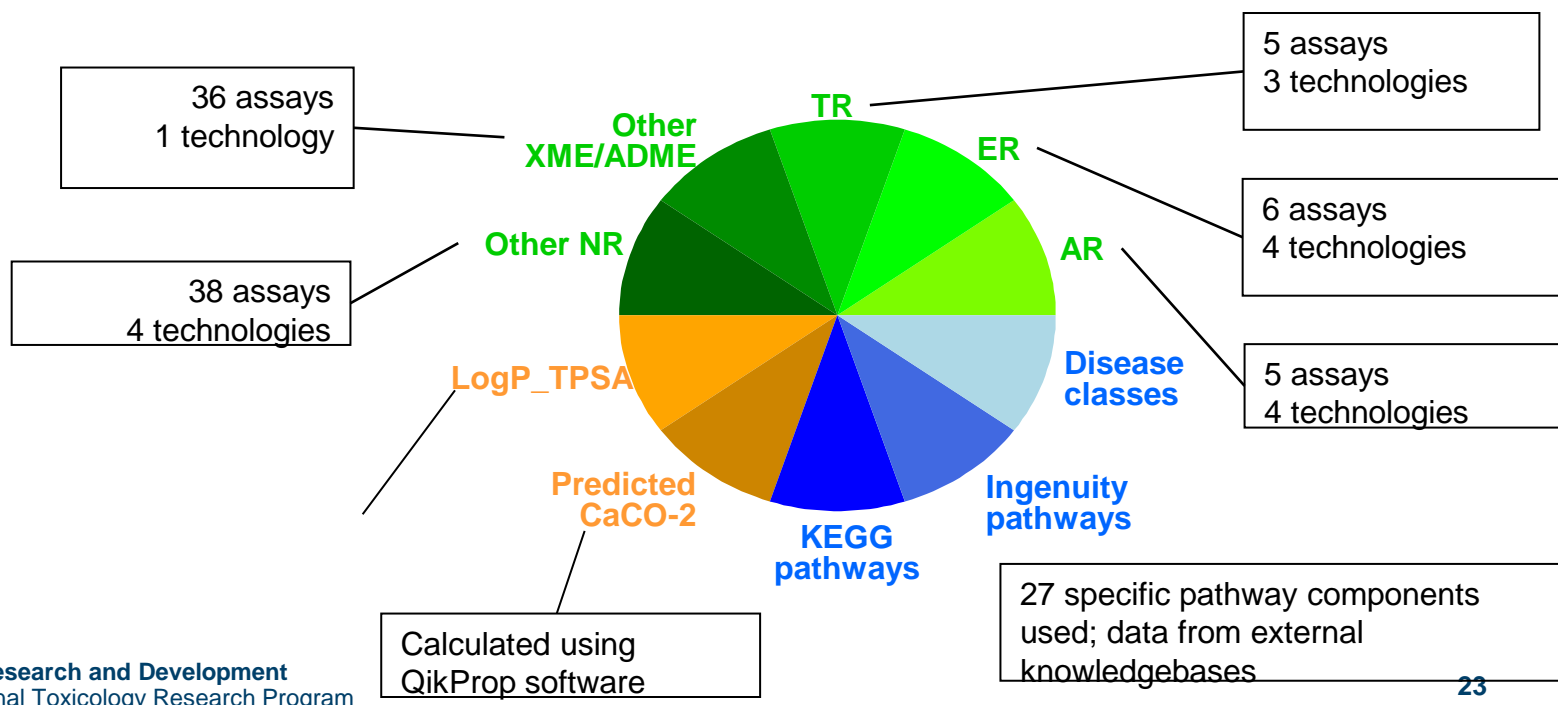
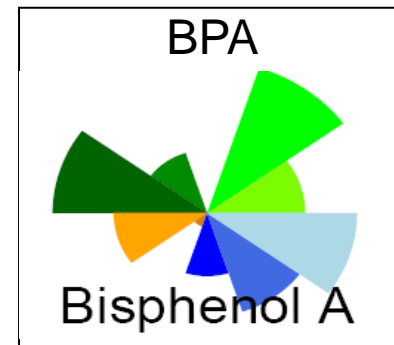
Data Set	Description	Download	Publication
ToxCast Phase I Data (Concentration Response Data)	All concentration response data from ToxCast Phase I (used to derive AC50/LEC data below)	Download (22.6 MB, ZIP)	Judson et al. (2010) "In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization - The ToxCast Project" Environmental Health Perspectives DOI:10.1289/ehp.0901392
ToxCast Phase I Data	Supplemental data and figures for Judson et al	Download (1.01 MB, ZIP)	Judson et al. (2010) "In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization - The ToxCast Project" Environmental Health Perspectives DOI:10.1289/ehp.0901392

In Vitro to In Vivo Associations
Cytb (purple), other enzymes (blue), GPCRs (orange), NR (dark green), ion channels (grey), transporters (light green)

ToxCast™ Navigation
Introduction
Chemicals
Assays
Information Management
Partnerships
Contractors
Presentations
Publications
News
Data Analysis Summit
E-mail Alerts

Unexpected Results

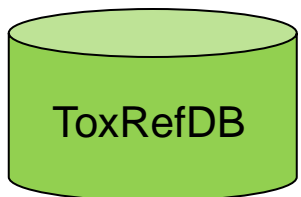
- Breakdown of materials in DMSO
 - Need for constant quality assurance
- Need to orthogonal methods
 - False positives and false negatives
 - Advantage of multiple platforms



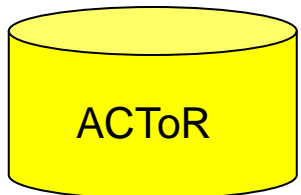


Data and Tool Interoperability

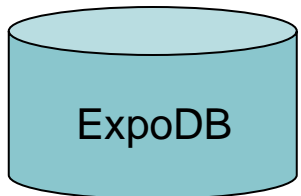
Data and Knowledgebases



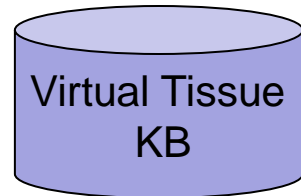
ToxRefDB



ACToR



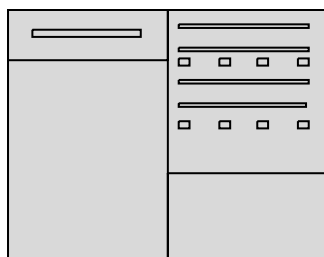
ExpoDB



Virtual Tissue
KB

Common Ontologies

Disease, ADR
Assay
Physiology



Analysis Tools

Web Pages
Web Services

OpenTox API

Open Ontologies / OBO

Open Source Tools

