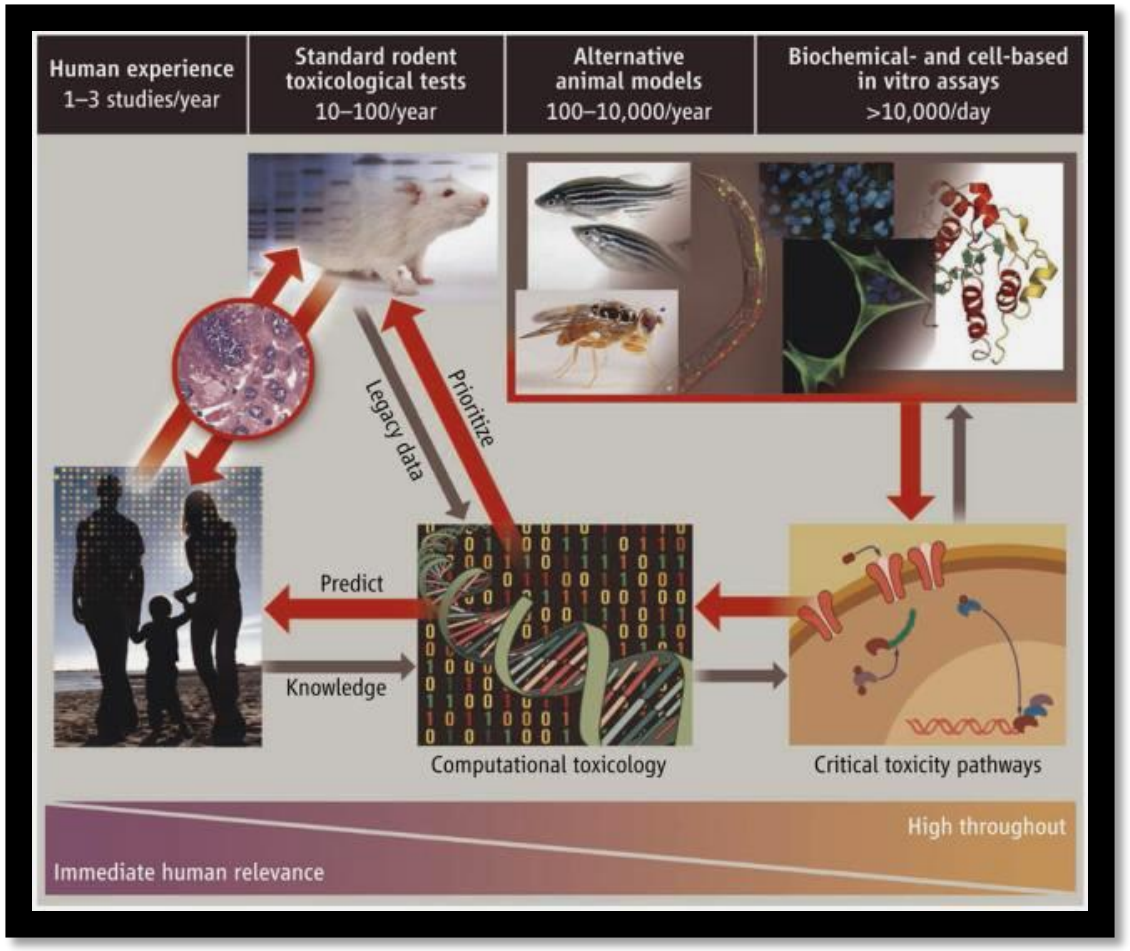
Overview of the ToxCast Research Program: Applications to Predictive Toxicology and Chemical Prioritization

Keith Houck
U.S. EPA, National Center for Computational Toxicology
Office of Research and Development

SETAC, Nashville, TN 2013

Tox21 Vision: Transforming Toxicity Testing



National Center for Advancing Translational Sciences (NCATS)
http://www.ncats.nih.gov/

SOURCE: Collins, Gray and Bucher (2008) Toxicology. Transforming environmental health protection. Science 319: 906 [2]

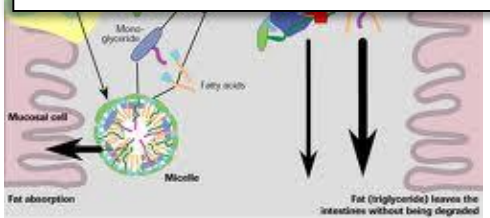
Problem Statement

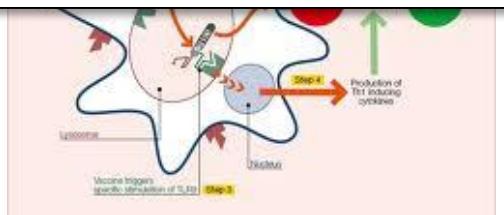
Too many chemicals to test with standard animal-based methods

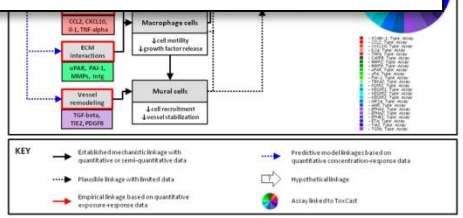
- Cost, time, animal welfare

Need for better mechanistic data

- Determine human relevance
- What is the Mode of Action (MOA) or Adverse Outcome Pathway (AOP)?





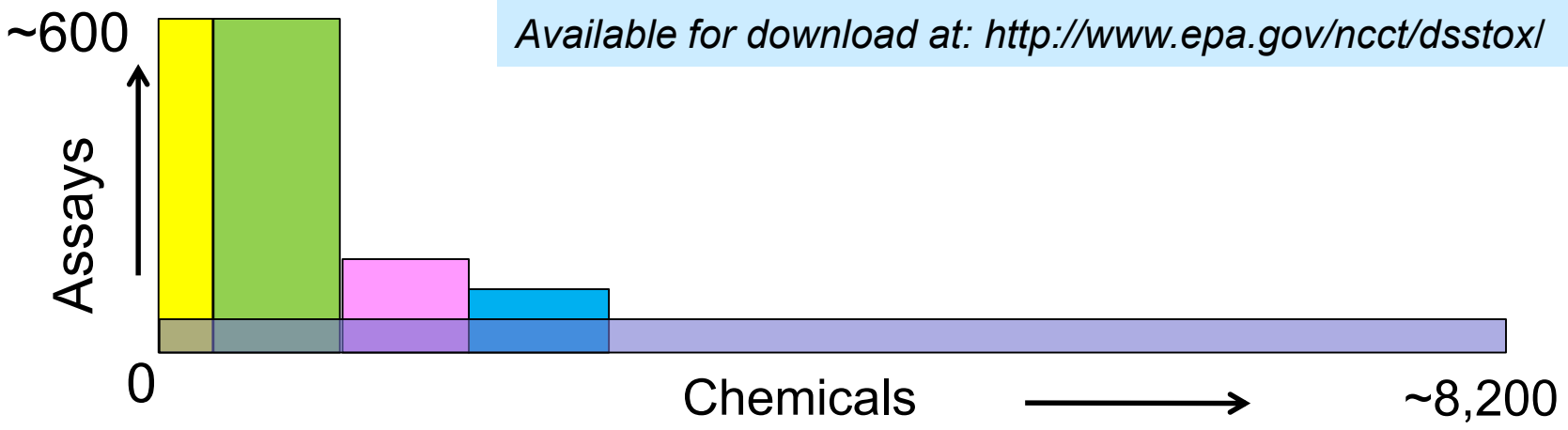


ToxCast /Tox21 Overall Strategy

- Identify targets or pathways linked to toxicity (AOP focus)
- Identify/develop high-throughput assays for these targets or pathways
- Develop predictive systems models
  - *in vitro → in vivo*
  - *in vitro → in silico*
- Use predictive models (qualitative):
  - Prioritize chemicals for targeted testing
  - Suggest / distinguish possible AOP / MOA for chemicals
- *High-throughput Exposure Predictions*
- *High-throughput Risk Assessments (quantitative)*

Testing under ToxCast and Tox21 Chemicals, Data and Release Timelines

Set	Chemicals	Assays	Endpoints	Completion	Available
ToxCast Phase I	293	~600	~700	2011	Now
ToxCast Phase II	767	~600	~700	03/2013	12/2013
ToxCast Phase IIIa	1001	~100	~100	Just starting	2014
E1K (endocrine)	880	~50	~120	03/2013	11/2013
Tox21	8,193	~25	~50	Ongoing	Ongoing

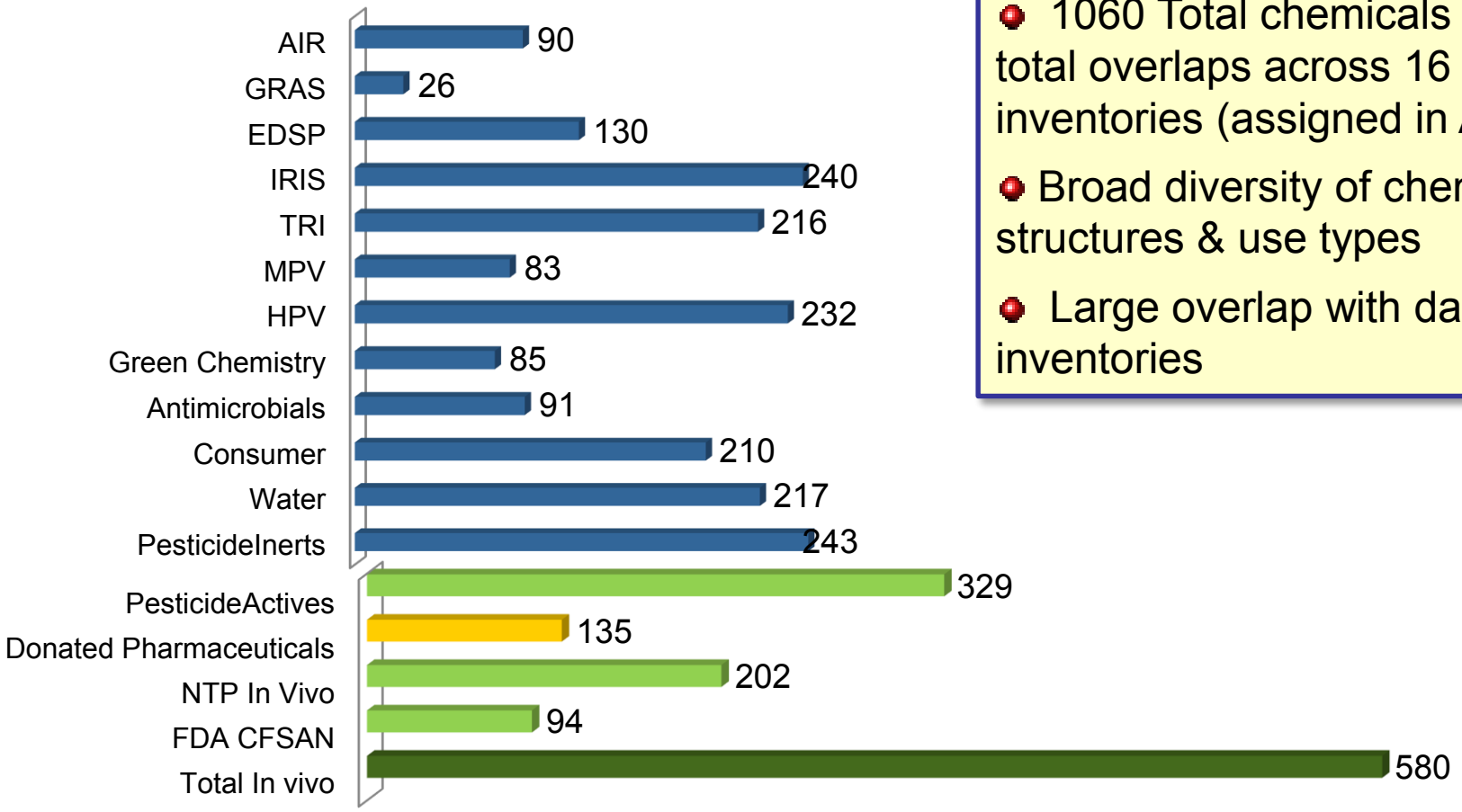


Available for download at: http://www.epa.gov/ncct/dsstox/

Pesticides , antimicrobials, food additives, green alternatives, HPV, MPV, endocrine reference cmpds, other tox reference cmpds, failed drugs, NTP in vivo, EPA high interest compounds, industrial, marketed drugs, fragrances, …

ToxCast PhI&PhII chemicals:

Spanning diverse inventories of EPA interest

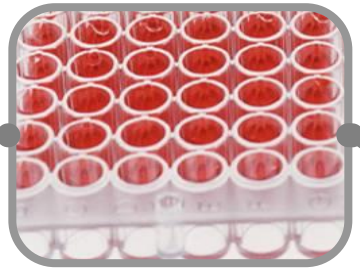


- 1060 Total chemicals → 2806 total overlaps across 16 diverse inventories (assigned in ACToR)
- Broad diversity of chemical structures & use types
- Large overlap with data-rich inventories

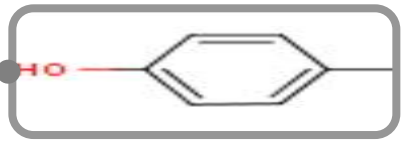
High Throughput Screening 101 (HTS)

United States Environmental Protection Agency

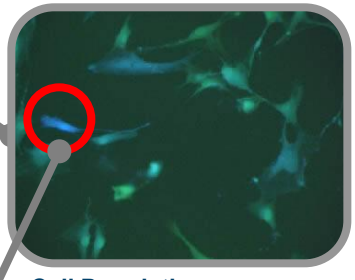
Robots



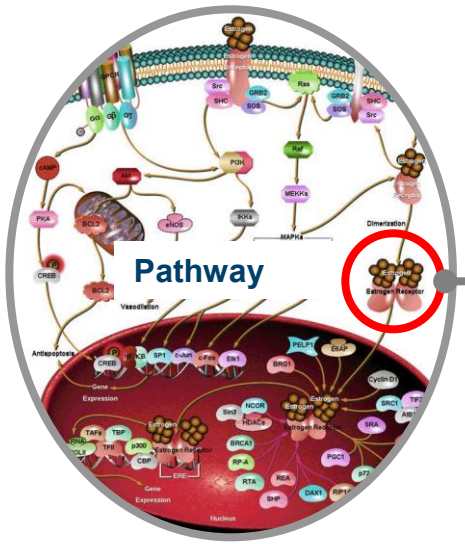
96-, 384-, 1536 Well Plates

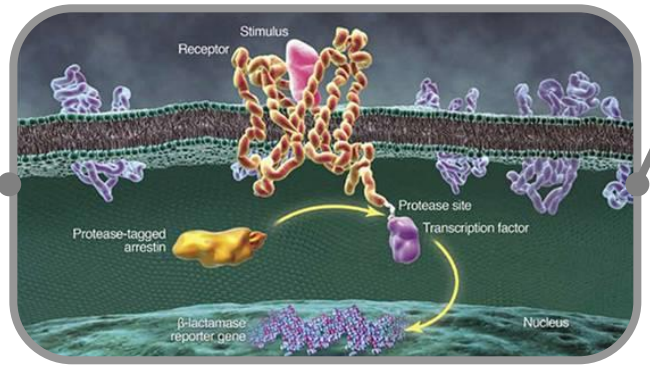


Chemical Exposure

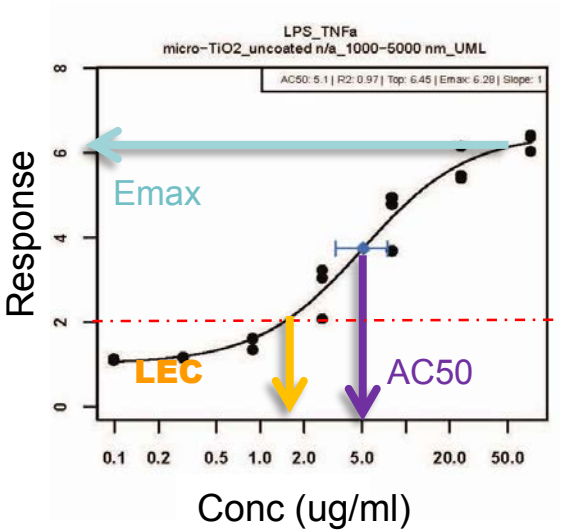


Cell Population

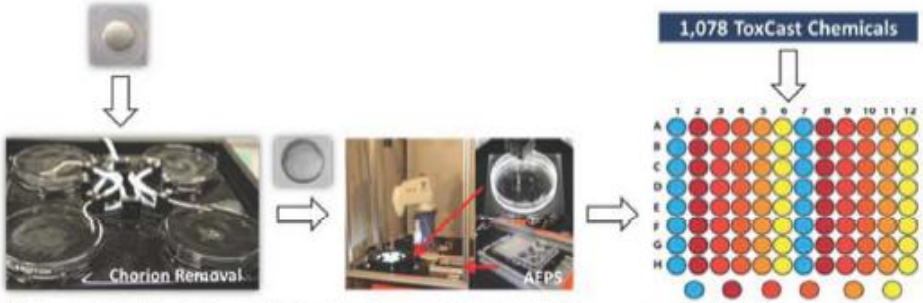




Target Biology (e.g., Estrogen Receptor)



Zebrafish Development Screen



24 hours later...

Morphology Assessment

24 hpf End Points

End Point	Abbreviation
Mortality	MO24
Developmental Delay	DP24
Spontaneous Movement	SM24
Notochord	NC24

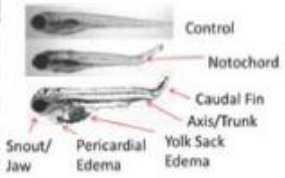
24hpf embryonic behavior is assessed by the Photo-motor Response Assessment Tool (PRAT). PRAT quantifies an individual embryo's photo-motor response to two distinct pulses of light

At 5 Days (120 hpf)

Morphology Assessment

120 hpf End Points

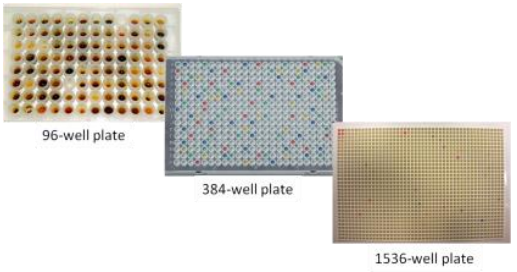
End Point	Abbreviation
Mortality	MORT
Yolk Sac Edema	YSE_
Body Axis	AXIS
Eye Defect	Eye
Snout	SNOU_
Jaw	JAW_
Otic Vesicle	OTIC
Pericardial Edema	PE_
Brain	BRAI
Somite	SOMI
Pectoral Fin	PFIN
Caudal Fin	CFIN
Pigment	PIG_
Circulation	CIRC
Truncated Body	TRUN
Swim Bladder	SWIM
Notochord & Bent Tail	NC_
Touch Response	TR_



Using the Viewpoint Zebrabox, light-induced larval locomotor activity was measured. Locomotor activity is tracked for 25 minutes

Truong et al., Tox. Sci, *in press*

ToxCast Assays (>700 endpoints)



Assay Provider
ACEA
Apredica
Attagene
BioReliance
BioSeek
CeeTox
CellzDirect
Tox21/NCATS
NHEERL MESC
NHEERL Zebrafish
NovaScreen (Perkin Elmer)
Odyssey Thera
Vala Sciences

Biological Response
cell proliferation and death
cell differentiation
Enzymatic activity
mitochondrial depolarization
protein stabilization
oxidative phosphorylation
reporter gene activation
gene expression (qNPA)
receptor binding
receptor activity
steroidogenesis

Target Family
response Element
transporter
cytokines
kinases
nuclear receptor
CYP450 / ADME
cholinesterase
phosphatases
proteases
XME metabolism
GPCRs
ion channels

Assay Design
viability reporter
morphology reporter
conformation reporter
enzyme reporter
membrane potential reporter
binding reporter
inducible reporter

Readout Type
single
multiplexed
multiparametric

Cell Format
cell free
cell lines
primary cells
complex cultures
free embryos

Species
human
rat
mouse
zebrafish
sheep
boar
rabbit
cattle
guinea pig

Tissue Source

Lung	Breast
Liver	Vascular
Skin	Kidney
Cervix	Testis
Uterus	Brain
Intestinal	Spleen
Bladder	Ovary
Pancreas	Prostate
Inflammatory	Bone

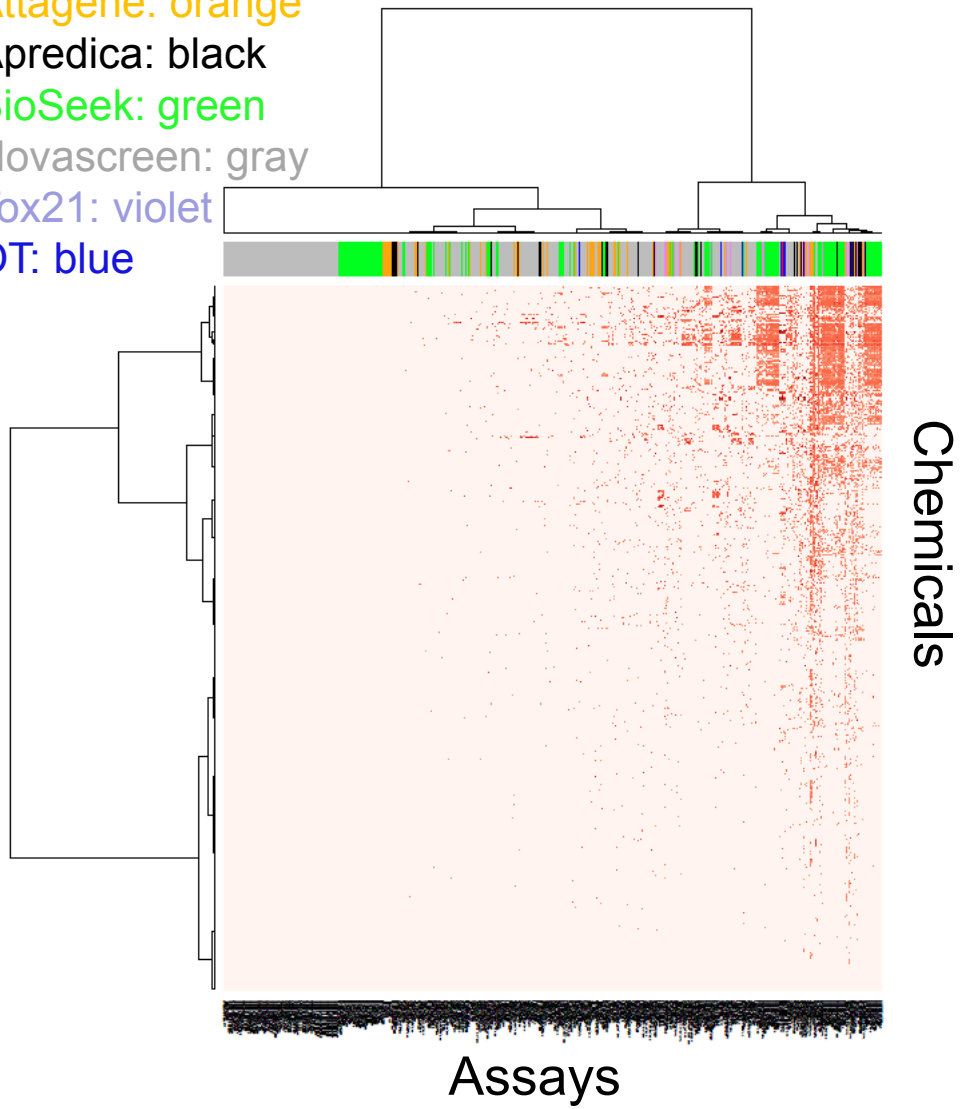
Detection Technology
qNPA and ELISA
Fluorescence & Luminescence
Alamar Blue Reduction
Arrayscan / Microscopy
Reporter gene activation
Spectrophotometry
Radioactivity
HPLC and HPEC
TR-FRET

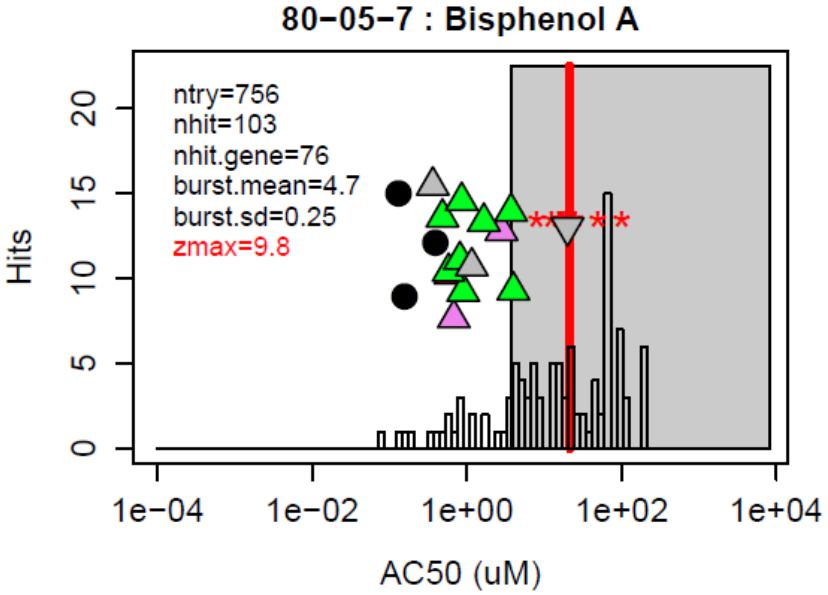
9

List of assays and related information at: **http://www.epa.gov/ncct/**

ToxCast Phase II: 1051 Chemicals x 791 Assay Readouts

ACEA: red
Attagene: orange
Apredica: black
BioSeek: green
Novascreen: gray
Tox21: violet
OT: blue





Most chemicals cause activity in many assays near the cytotoxicity threshold

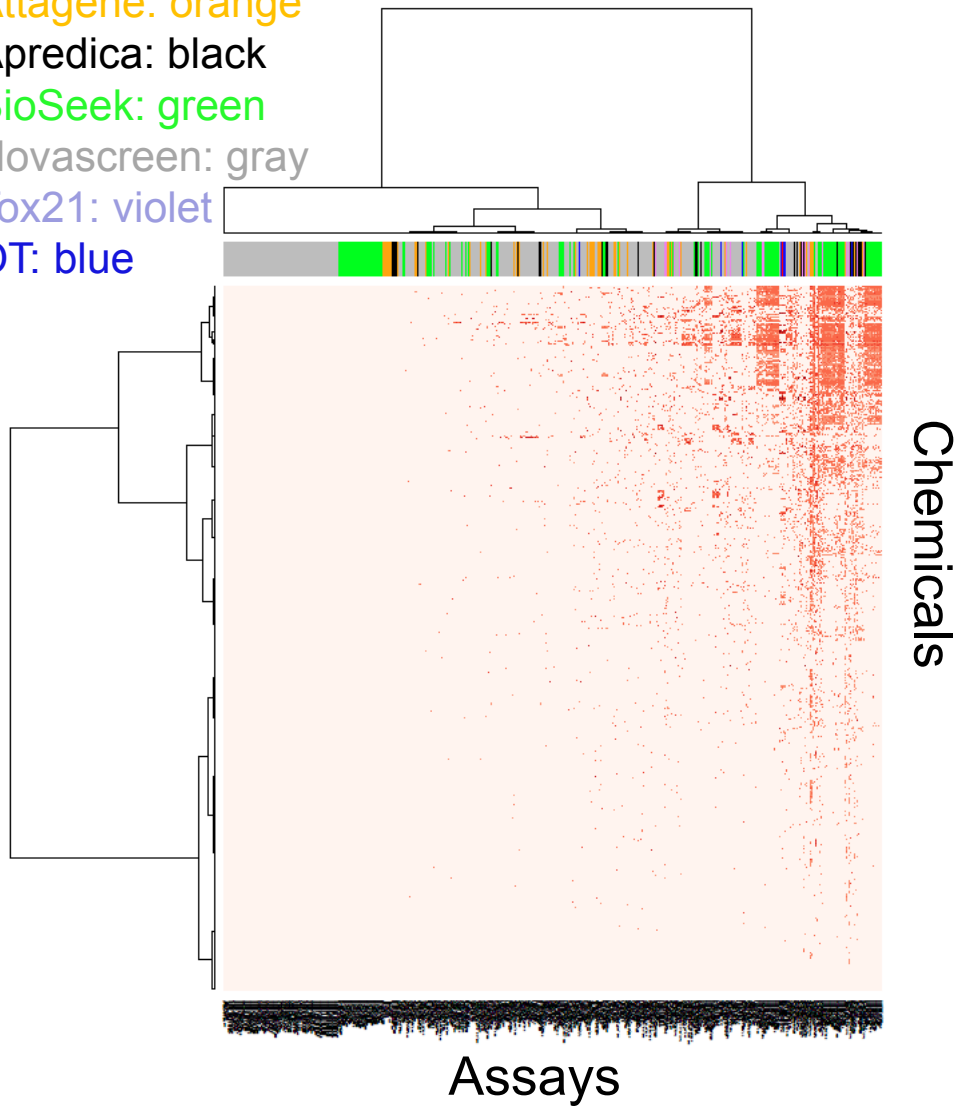
Cell stress-related assay activity

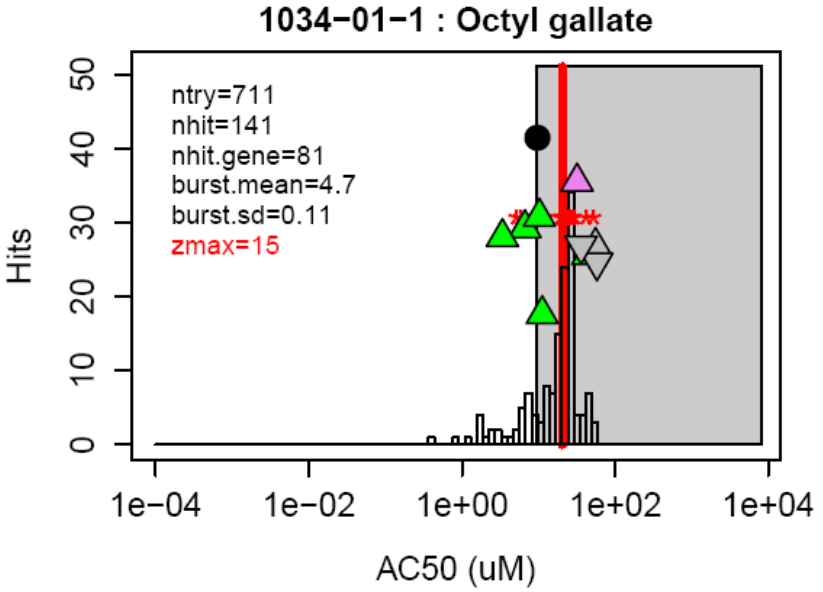
"Hit" (AC50) in burst region is less likely to result from specific activity
(e.g. binding to receptor or enzyme)

Z-score: # of SD from burst center
-High Z: more likely to be specific
-Low Z: less likely to be specific

ToxCast Phase II: 1051 Chemicals x 791 Assay Readouts

ACEA: red
Attagene: orange
Apredica: black
BioSeek: green
Novascreen: gray
Tox21: violet
OT: blue





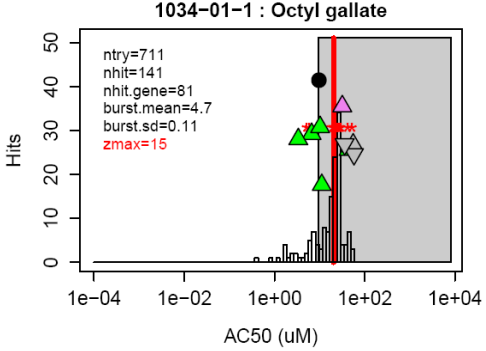
Most chemicals cause activity in many assays near the cytotoxicity threshold

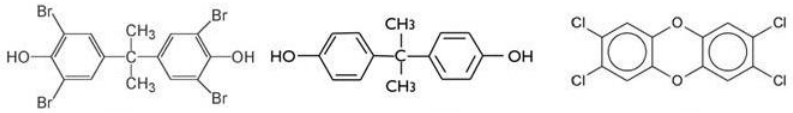
Cell stress-related assay activity

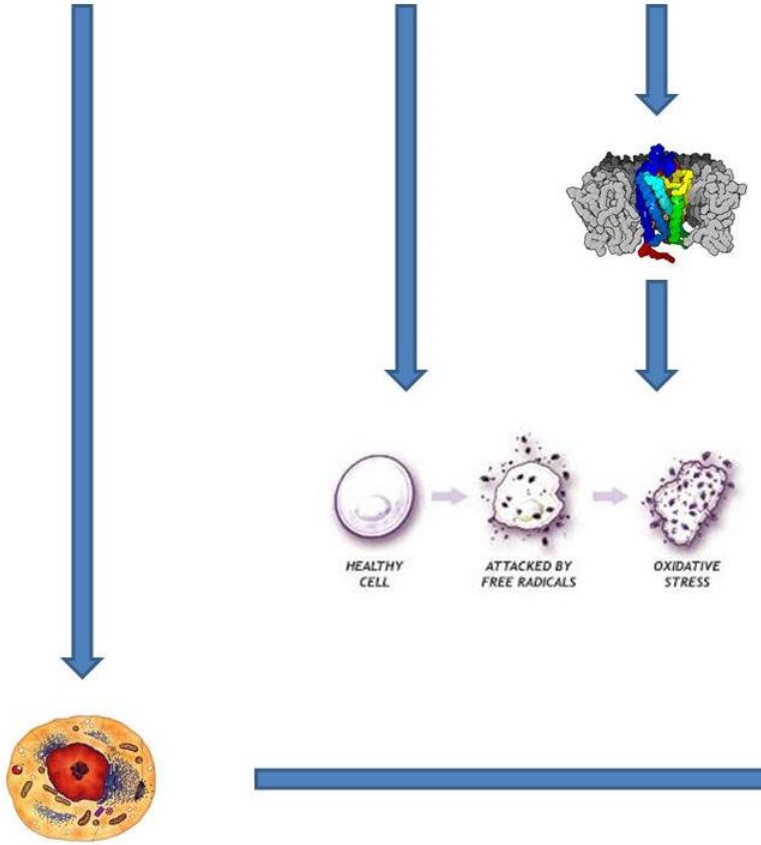
"Hit" (AC50) in burst region is less likely to result from specific activity
(e.g. binding to receptor or enzyme)

Z-score: # of SD from burst center
-High Z: more likely to be specific
-Low Z: less likely to be specific

Significance of In Vitro Effects







Assay Target Class		Assessment
Molecular Target	EDC Acetylcholinesterase Inhibition Ion channel blocker Genotoxicity	AOP Assessment Targeted testing
Cell Stress Mediated	Oxidative stress Membrane disruption Nucleophiles Electrophiles Energy depletion	Estimate MTD Estimate NOEL
No Effect	Non-reactive chemical Not bioactive Effects would require high doses	Estimate NOEL

Gene Score: Combine potency and specificity

- How to summarize 1000s of chemicals x 100s of assays?
  - Potency: -log(AC50)
  - Specificity: Z-score
  - Gene score = Potency + Specificity
    - average over assays for gene [-log(AC50) + Z-score]
  - Can be used to get quick ranking of chemicals
  - Gene Score > 7 are most interesting
    - Z-score=2 and AC50=10 μM
    - 5670 chemical-gene combinations >7 (~1%)
    - 281 Genes (out of 330)
    - 1231 Chemicals (out of 1877)

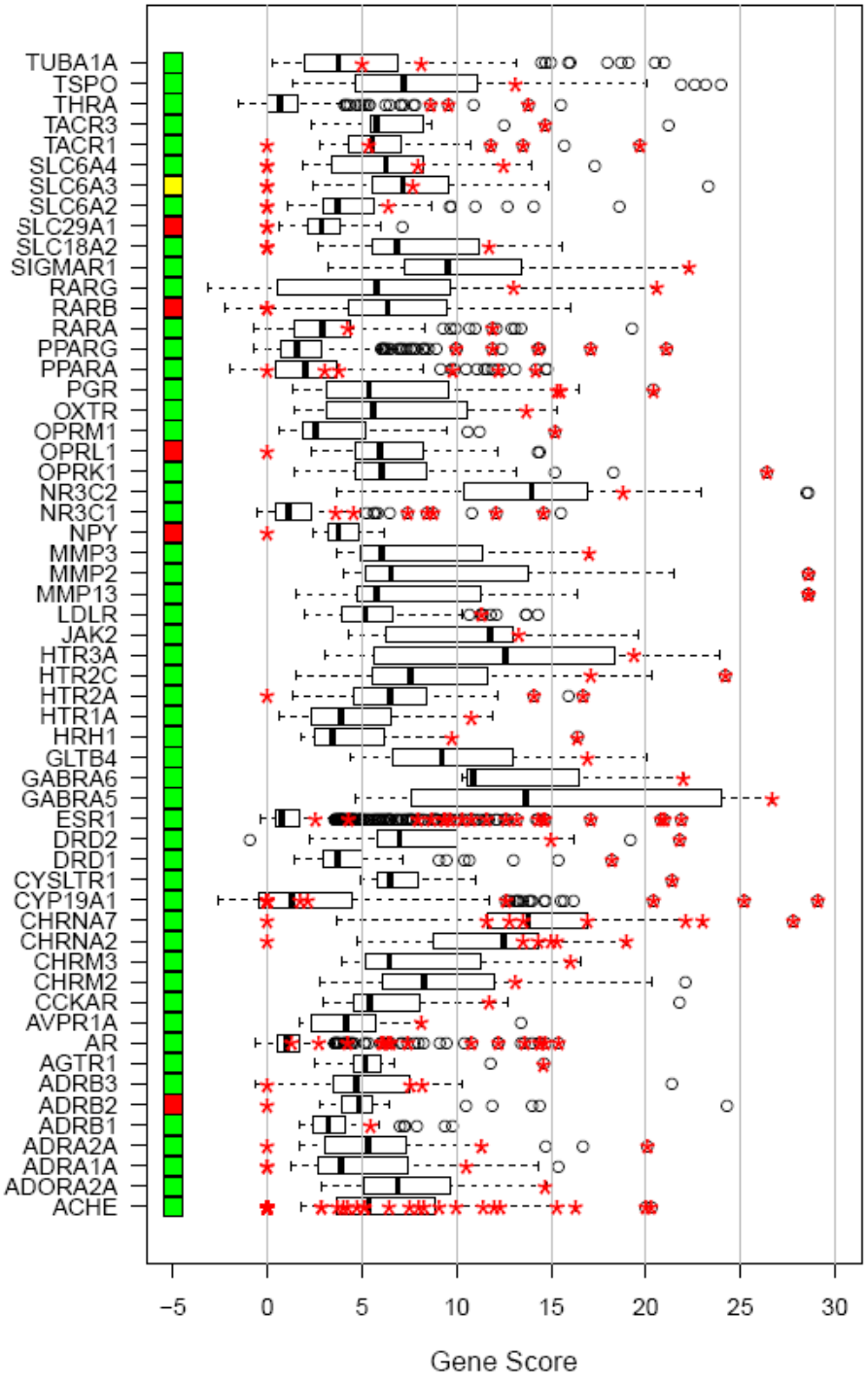
Do Assays Detect Potent Reference Chemicals?

* =Reference chemicals

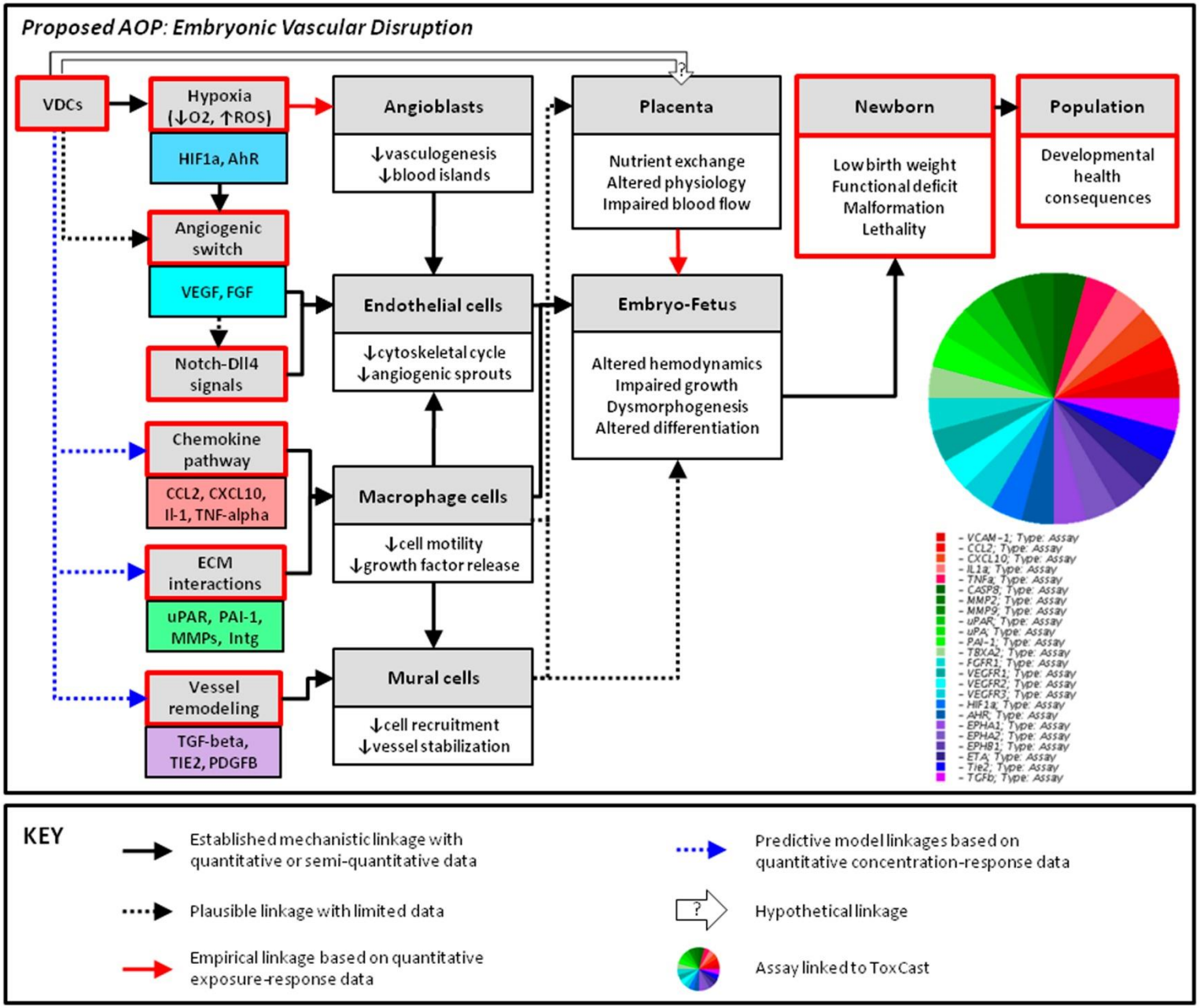
- These chemicals should be near the right of the gene score distribution

- Most assays show reference chemicals to be potent and specific

- Gives confidence that novel chemicals active in the assay are perturbing that pathway



Use of HTS Results in an Adverse Outcome Pathway (AOP)

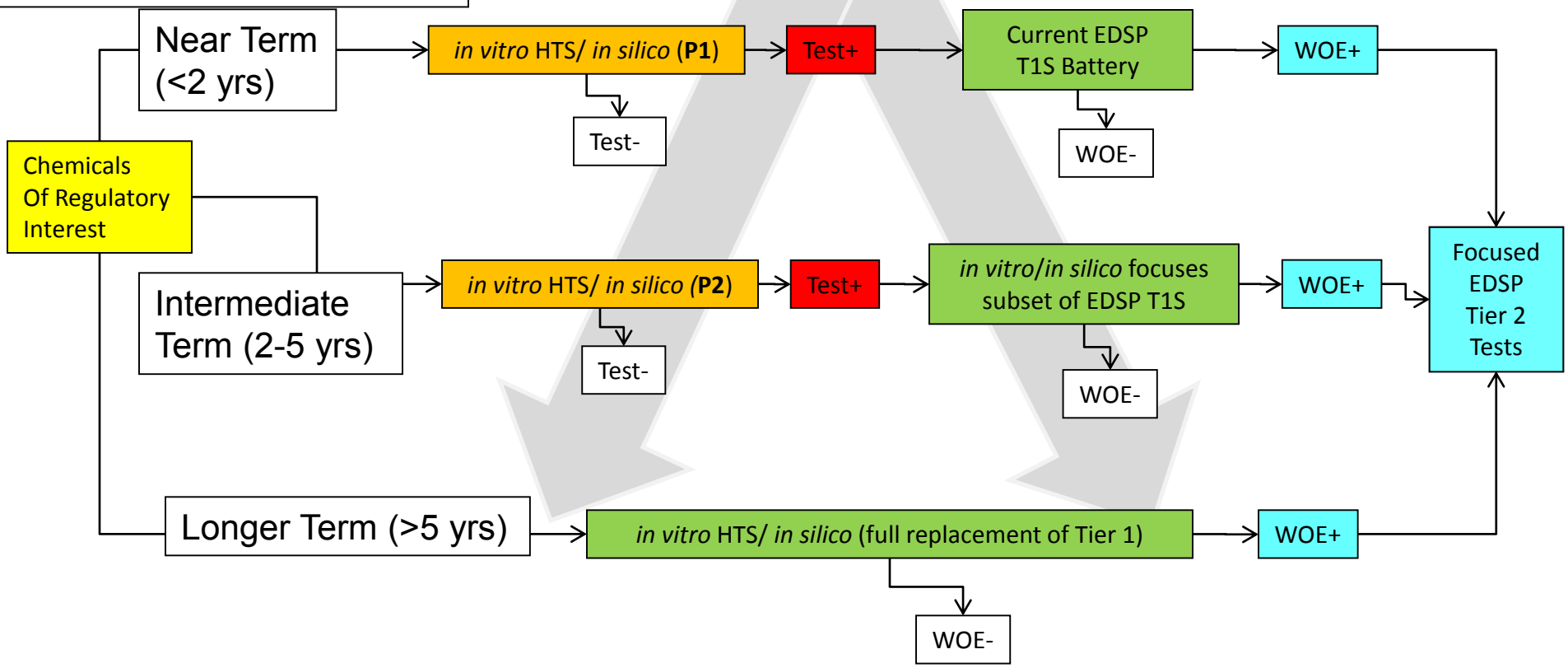


Knudsen and Kleinstreuer. Birth Def Res C. 2012

ToxCast and the Endocrine Disruptor Screening Program

The universe of chemicals passes through each version of the HTS/*in silico* pipeline to evaluate chemicals in refined tests, or for new pathways, to evaluate improve and validate methods.

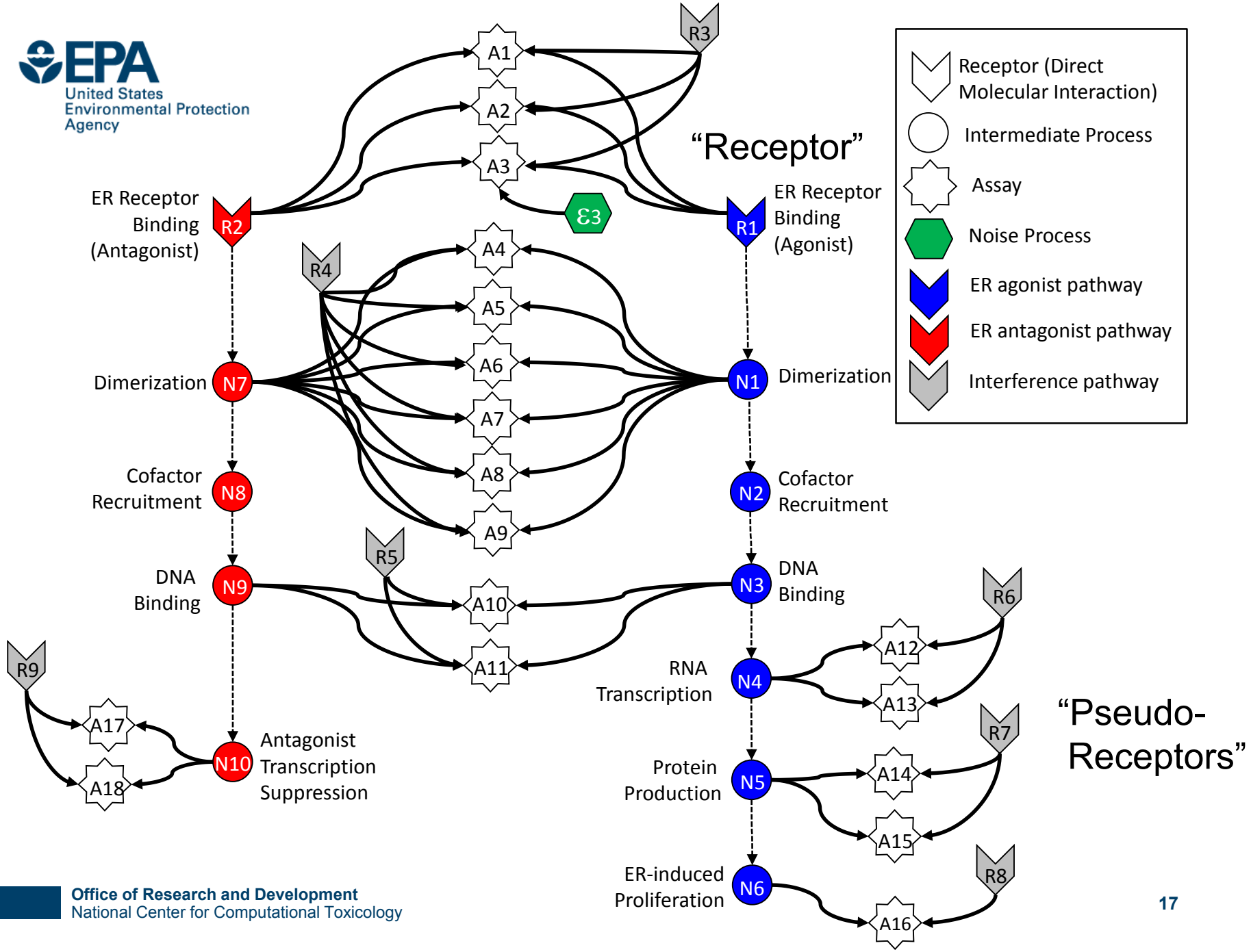
EPA Research provides basis for improving the suite of assays and models to advance chemical prioritization and screening



Chemical Prioritization
Includes registration review timeline, physico-chemical properties, exposure estimates, *in vitro* assays and computer models (QSAR, expert systems, systems biology models).

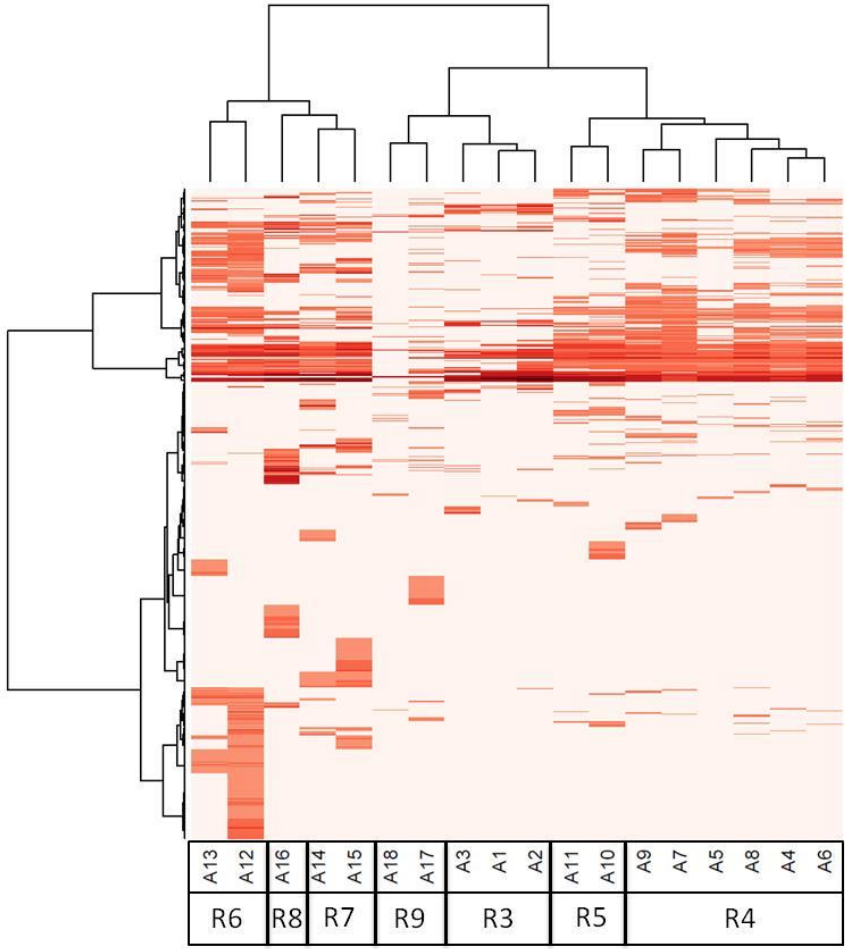
Screening Decisions
Near Term = Incorporates HTS/*in silico* prioritization methods for post EDSP List 2
Intermediate = Run subset of T1S assays indicated by HTS and *in silico* predictions
Long Term = Full replacement of EDSP T1S Battery

http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf


EPA
United States
Environmental Protection
Agency
"Receptor"
"Pseudo-
Receptors"
Receptor (Direct
Molecular Interaction)
Intermediate Process
Assay
Noise Process
ER agonist pathway
ER antagonist pathway
Interference pathway
ER Receptor
Binding
(Antagonist)
ER Receptor
Binding
(Agonist)
Dimerization
Dimerization
Cofactor
Recruitment
Cofactor
Recruitment
DNA
Binding
DNA
Binding
RNA
Transcription
Protein
Production
ER-induced
Proliferation
Antagonist
Transcription
Suppression
R1
R2
R3
R4
R5
R6
R7
R8
R9
A1
A2
A3
ε3
A4
A5
A6
A7
A8
A9
A10
A11
A12
A13
A14
A15
A16
A17
A18
N1
N2
N3
N4
N5
N6
N7
N8
N9
N10

Major theme – all assays have false positives and negative

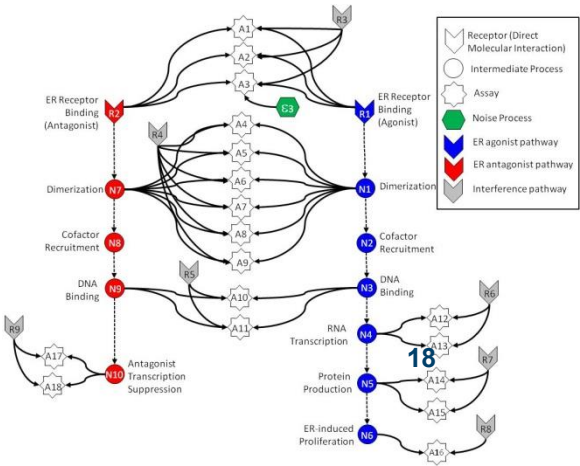
Assays cluster by technology, suggesting technology-specific non-ER activity



Much of this “noise” is reproducible, i.e. it is “assay interference”

Result of interaction of chemical with complex biology in the assay

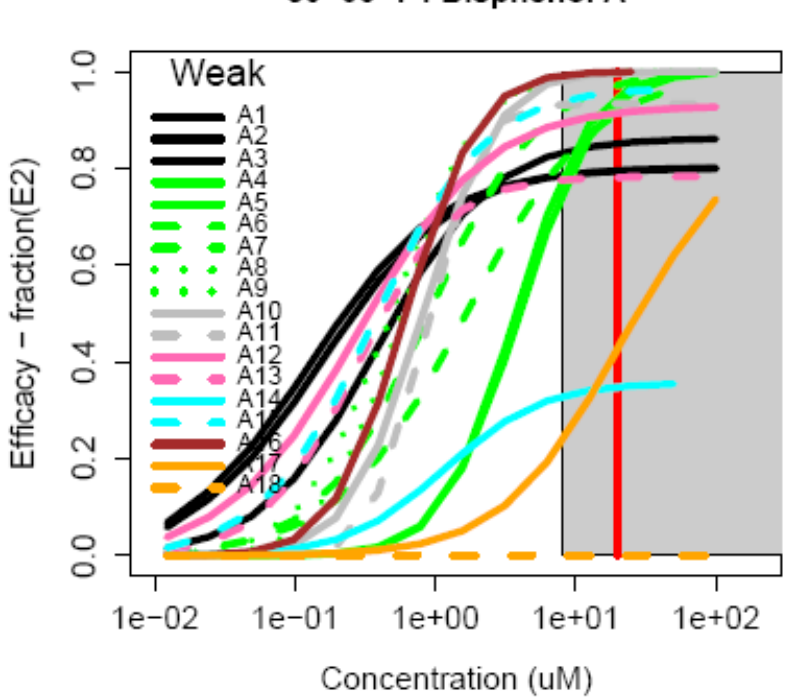
Our chemical library is only partially “drug-like”
-Solvents
-Surfactants
-Intentionally cytotoxic compounds
-Metals
-Inorganics



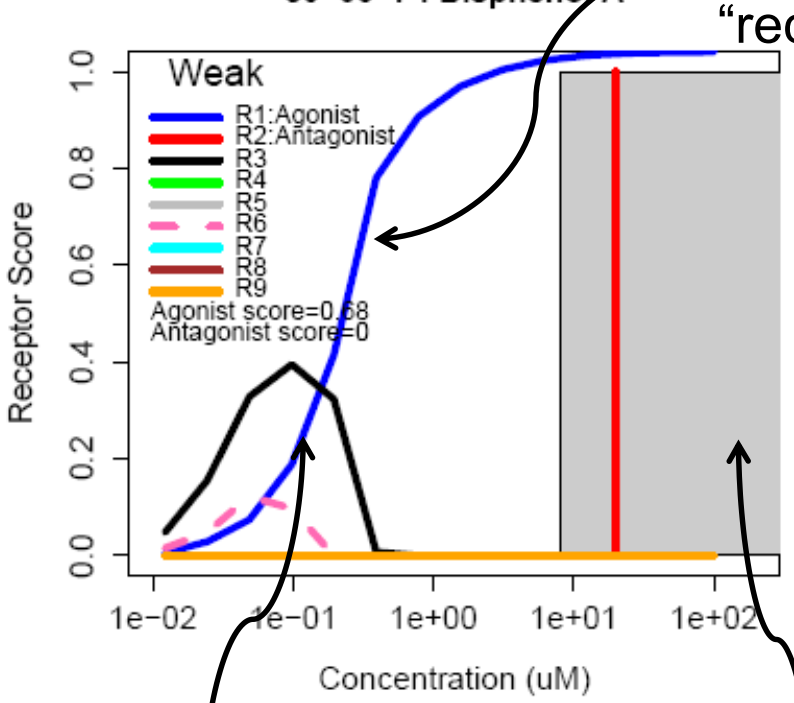
Example 1 – BPA – true agonist (AUC=0.66)

United States Environmental Protection Agency

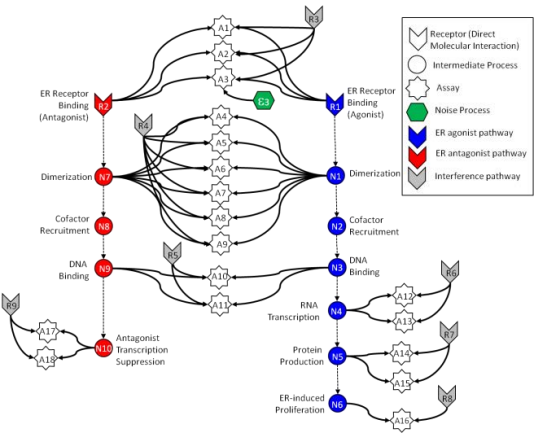
Assays



"Receptors"



Blue: agonist "receptor"



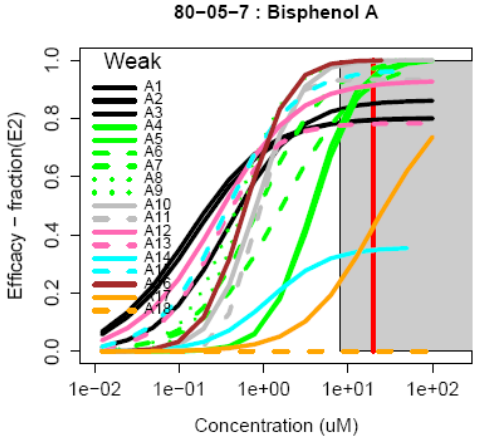
Binding assays active at lowest concentration

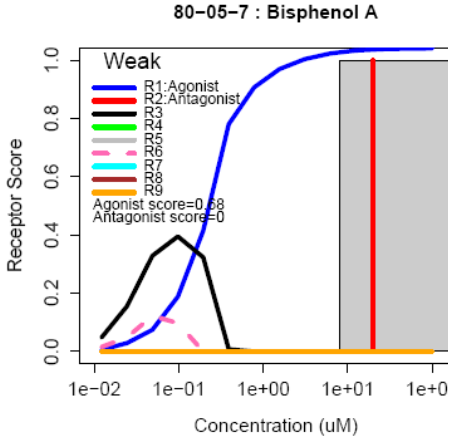
AUC "sign" feature will discount this

Cytotoxicity Region: red line is median cytotox AC50

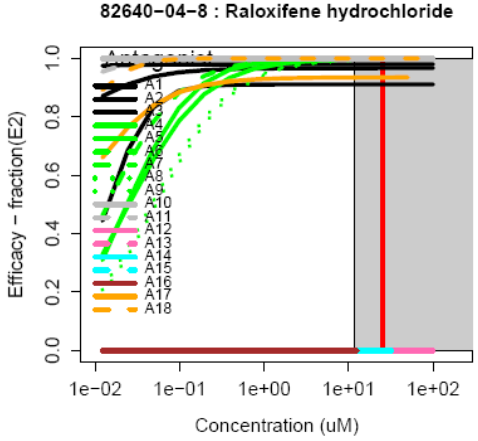
Example curves

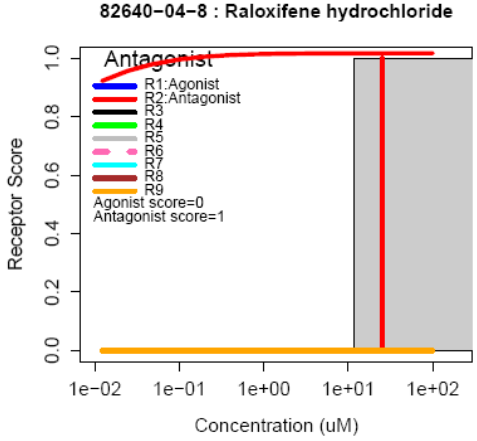
True Agonist



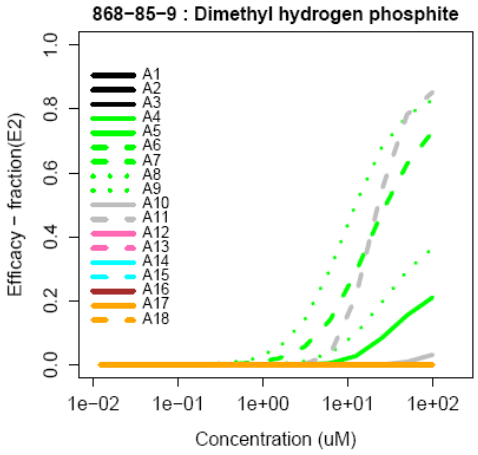


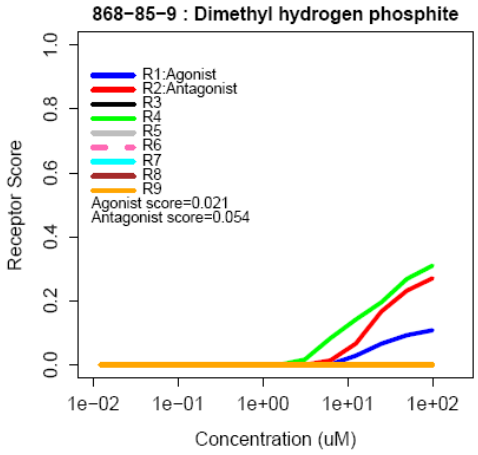
True Antagonist



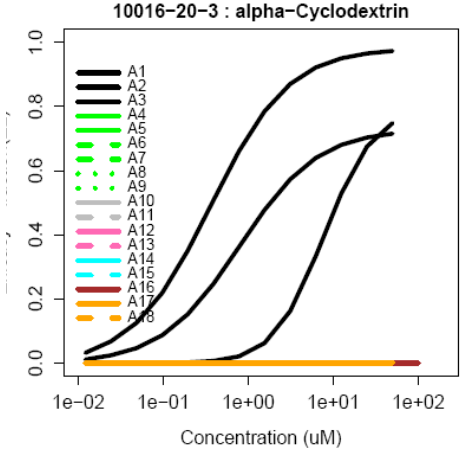


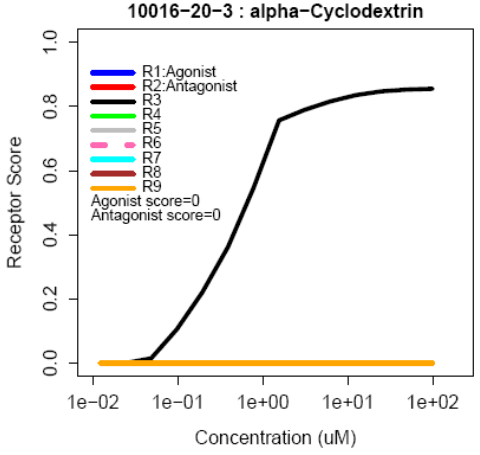
Negative-Broad Assay Interference



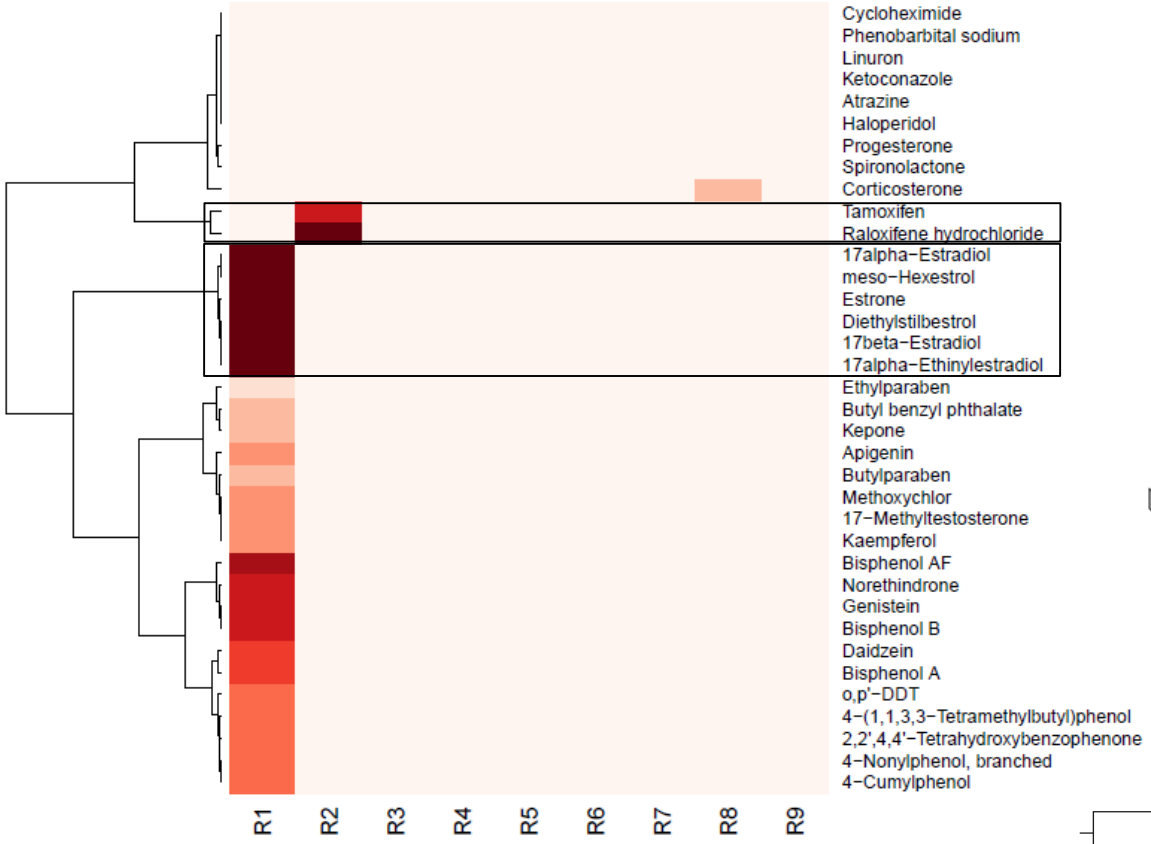


Negative-Narrow Assay Interference

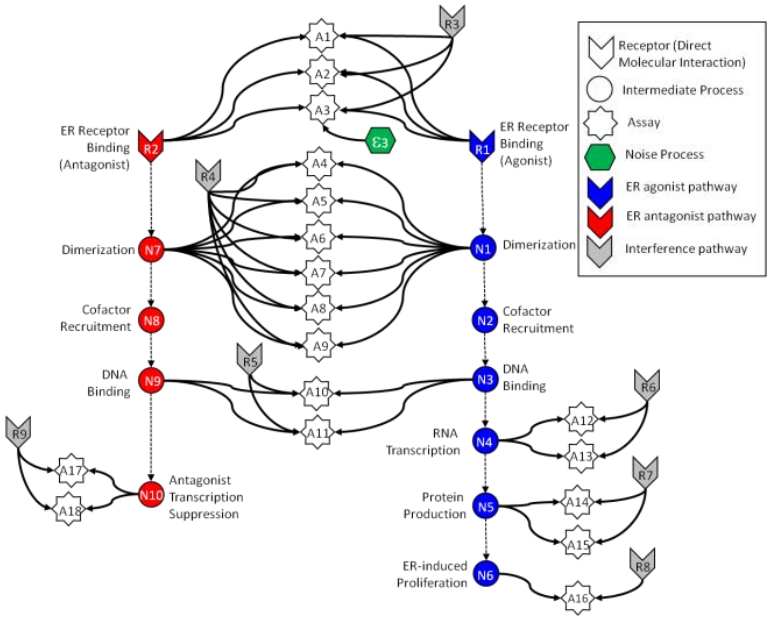


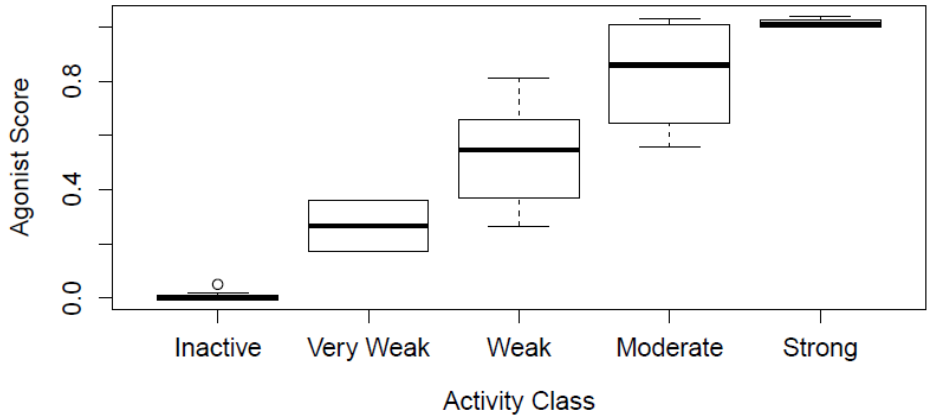


Reference Chemical Classification



AUC heat map for Reference chemicals



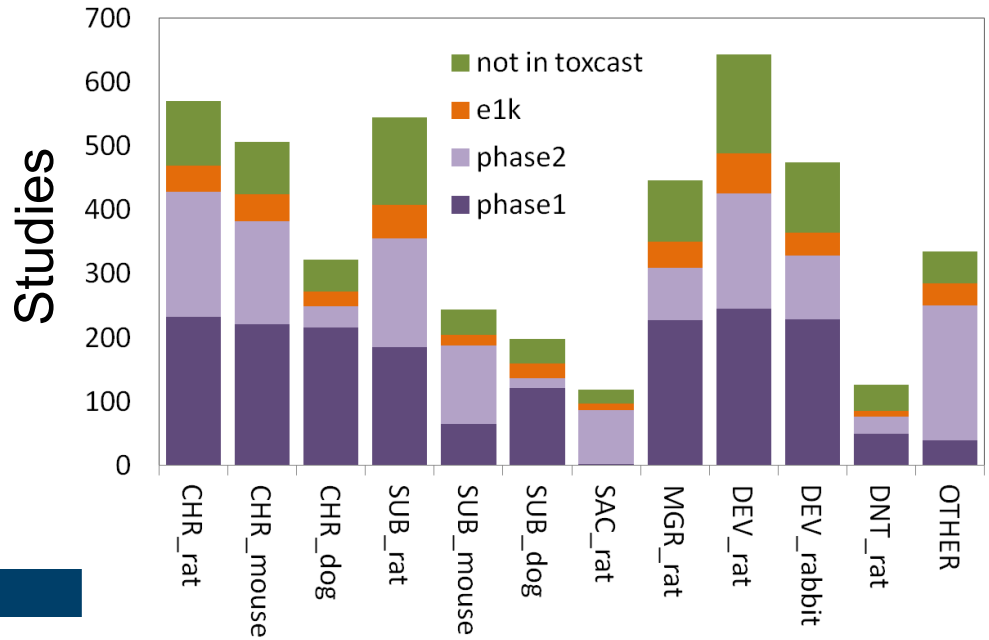


Predictive Models/Signatures

- Need to anchor to in vivo
- Guideline toxicity studies useful
  - EPA has extensive reports in support of registrations (pesticides)
  - Standardized
  - EPA regulates using these
- Recent incorporation of failed human drugs will provide more human-relevant in vivo

Toxicity Reference Database (ToxRefDB)

- ToxRefDB holds in vivo endpoint data from animal toxicology studies (DERs, NTP, open literature, pharma)
- Currently at 5567 studies on 1049 unique chemicals
- Used by:
  - ORD in predictive modeling (prospective)
    - e.g., multigen reproductive effects Martin et al., 2009)
  - OPP & OECD for assessing the impact of guideline studies on risk assessments (retrospective)
  - Public as a general chemical toxicity data resource

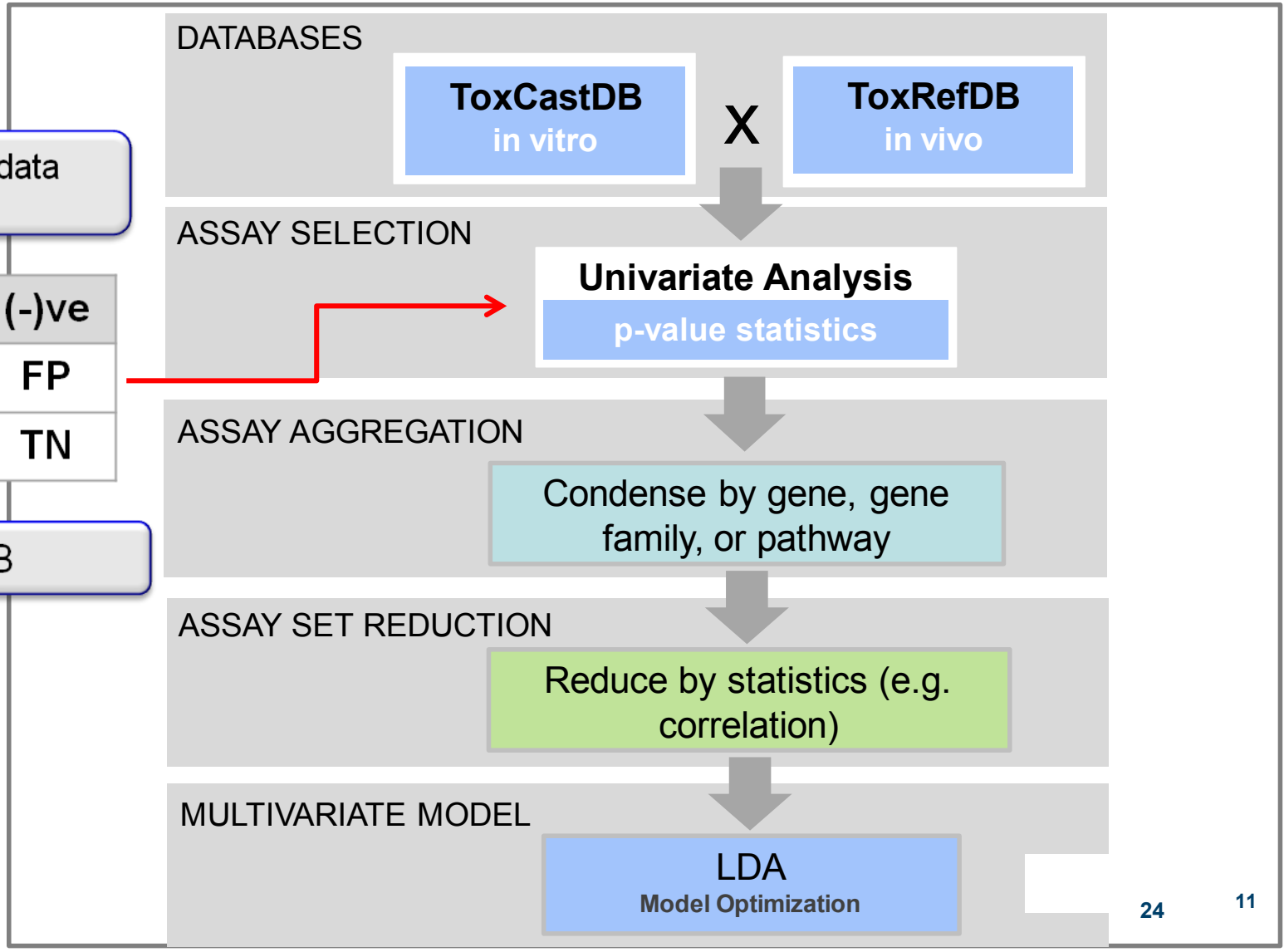


Data Source	Study Count
EPA OPP_der	3279
Open Literature	731
National Toxicol Program	666
Sanofi_pharma	222
Unpublished_submissions	50
GSK_pharma	38
Health Canada PMRA_der	23

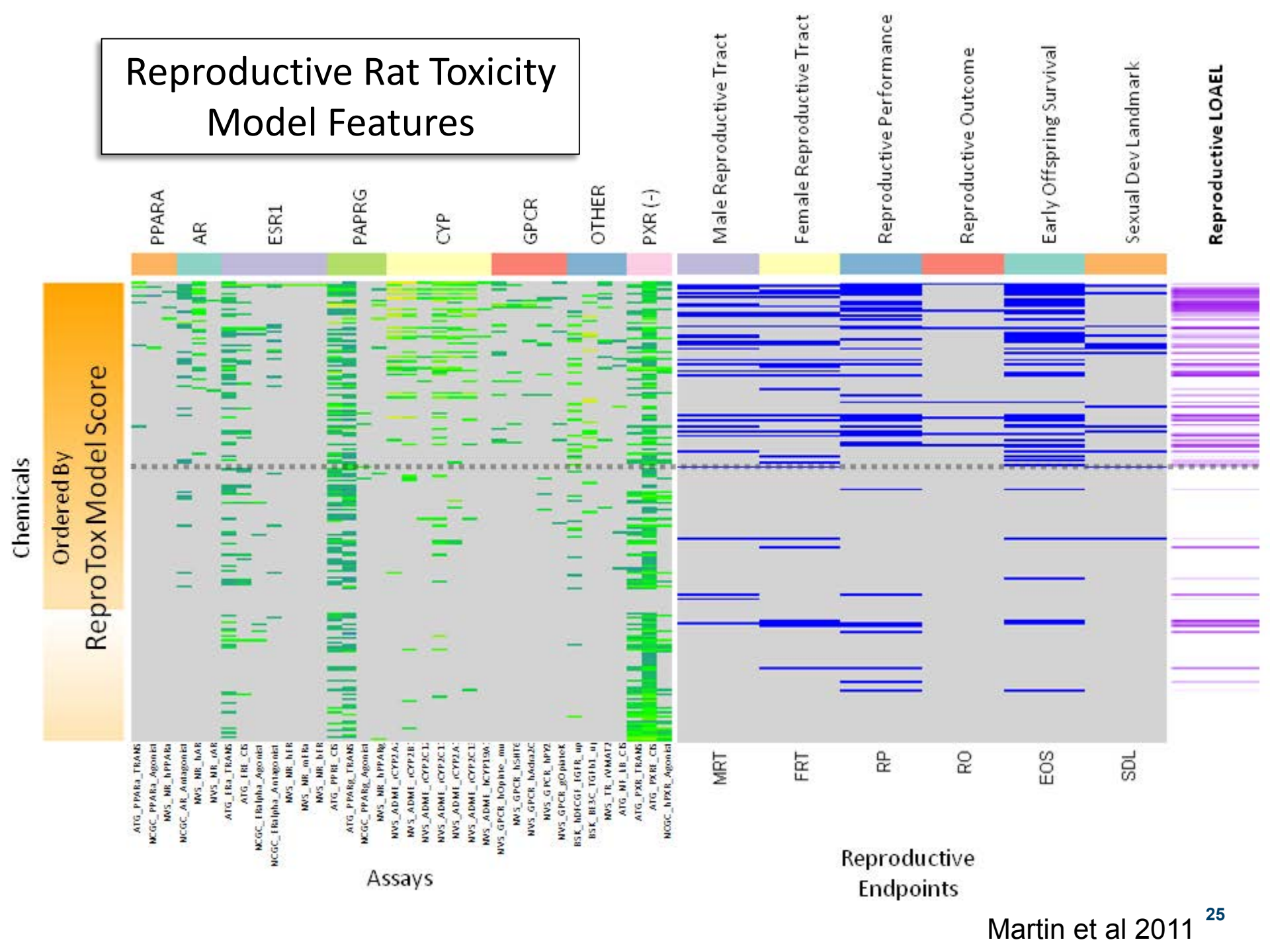
http://actor.epa.gov/toxrefdb/

Predictive Model Development from ToxCast and Other Data

EPA United States Environmental Protection Agency

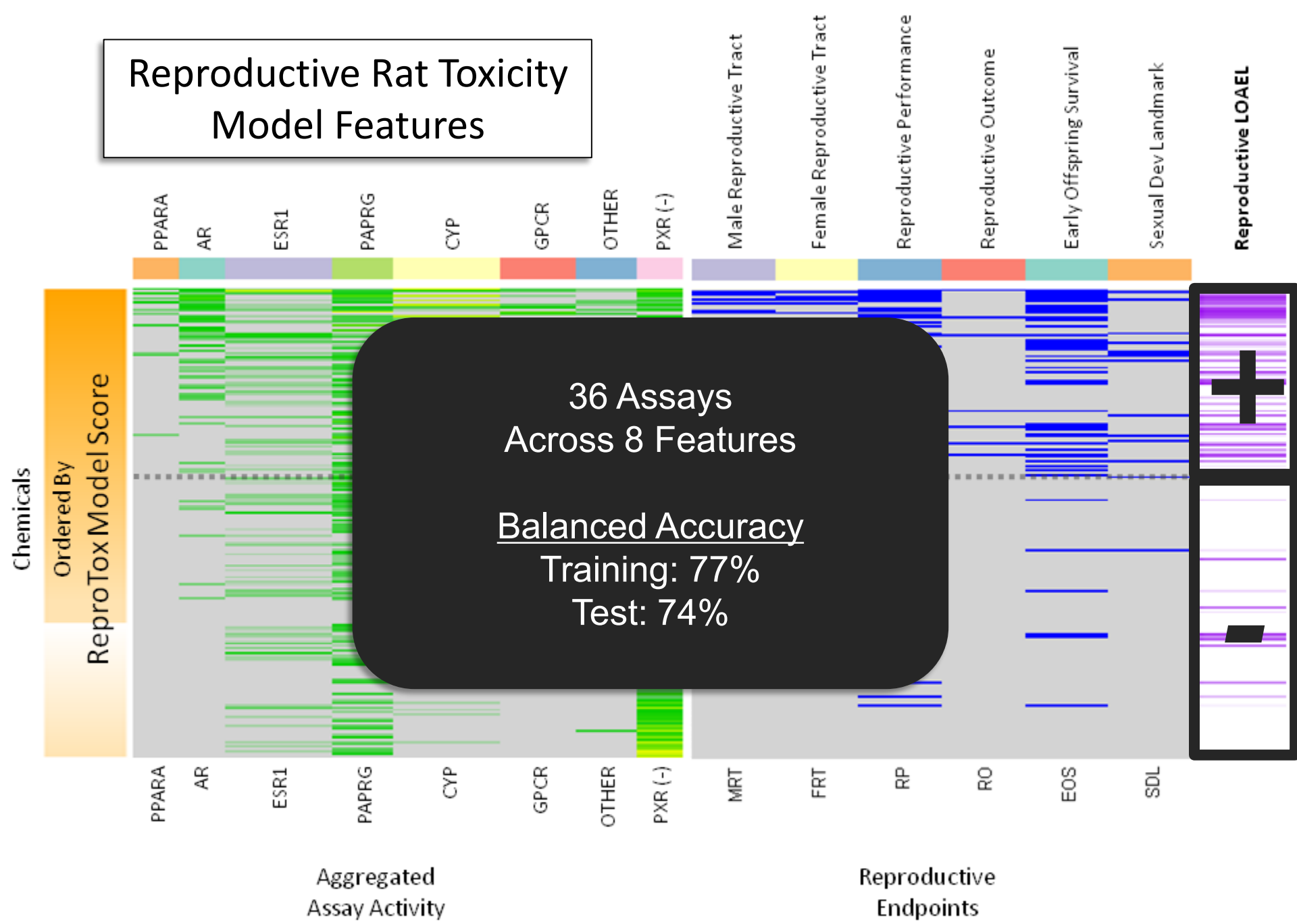


Reproductive Rat Toxicity Model Features



Martin et al 2011

Reproductive Rat Toxicity Model Features



Martin et al 2011

Predictive Toxicity Modeling Based on ToxCast Data

- **Predictive models: endpoints**
  - liver tumors: Judson et al. 2010, Env Hlth Persp 118: 485-492
  - hepatocarcinogenesis: Shah et al. 2011, PLoS One 6(2): e14584
  - cancer: Kleinstreuer et al. 2012, submitted
  - rat fertility: Martin et al. 2011, Biol Reprod 85: 327-339
  - rat-rabbit prenatal devtox: Sipes et al. 2011, Toxicol Sci 124: 109-127
  - zebrafish vs ToxRefDB: Sipes et al. 2011, Birth Defects Res C 93: 256-267

- **Predictive models: pathways**
  - endocrine disruption: Reif et al. 2010, Env Hlth Persp 118: 1714-1720
  - microdosimetry: Wambaugh and Shah 2010, PLoS Comp Biol 6: e1000756
  - mESC differentiation: Chandler et al. 2011, PLoS One 6(6): e18540
  - HTP risk assessment: Judson et al. 2011, Chem Res Toxicol 24: 451-462
  - angiogenesis: Kleinstreuer et al. 2011, Env Hlth Persp 119: 1596-1603
- **Continuing To Expand & Validate Prediction Models**
- **Generally moving towards more mechanistic/AOP-based models**

Understanding Success and Failure

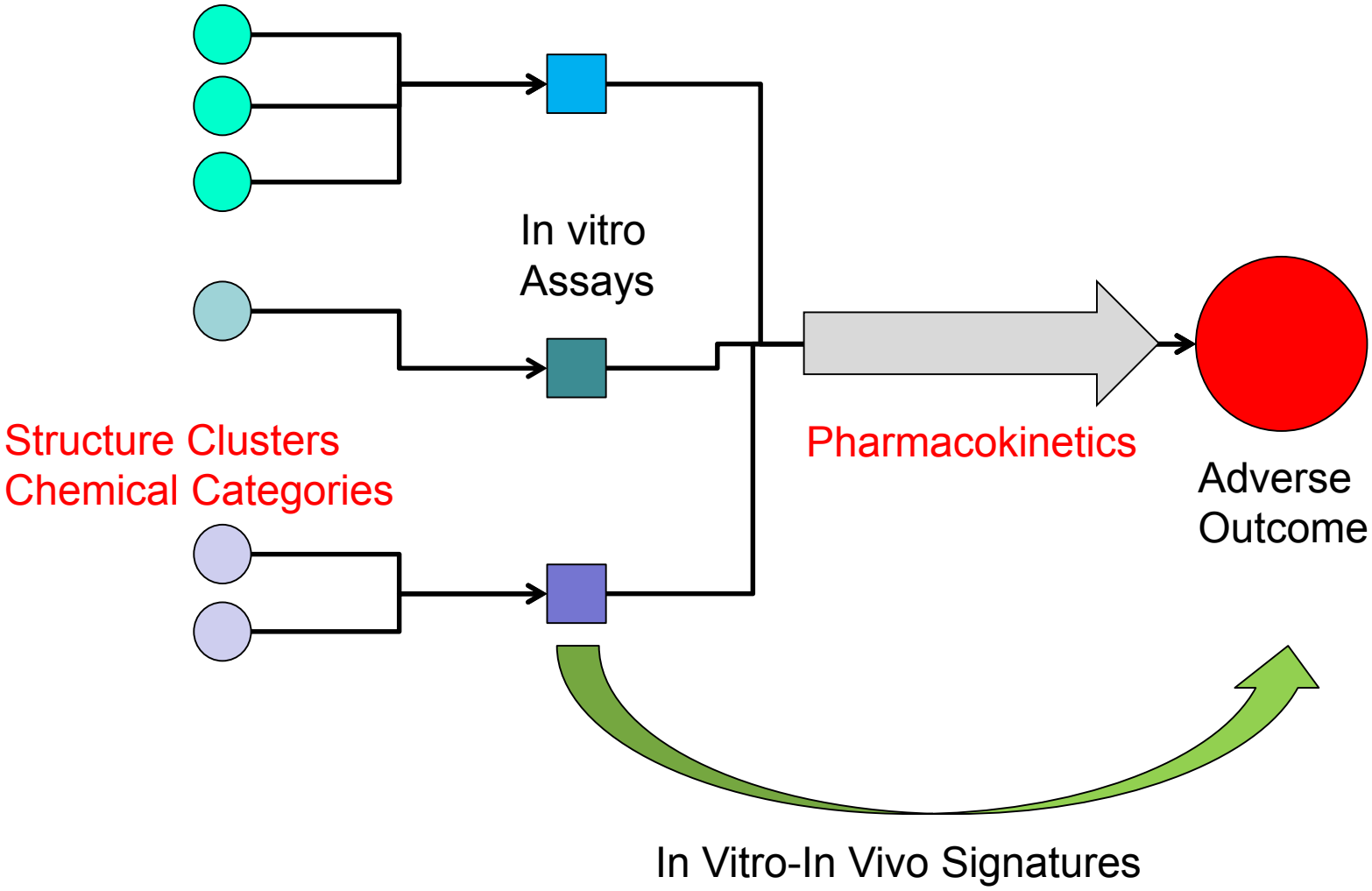
- Why *In vitro* to *in vivo* can work:
  - Chemicals cause effects through direct molecular interactions that we can measure with *in vitro* assays

- Why *in vitro* to *in vivo* does not always work:

Systems Models

- ★ Pharmacokinetics issues: biotransformation, clearance (**FP**, **FN**)
- ★ Assay issues: don't have all the right assays (**FN**)
- ★ Tissue issues: may need multi-tissue signaling networks (**FN**)
- ★ Statistical power issues: need enough chemicals acting through a given MOA to be able to build and test model (**FN**)
- ★ Compensation: system may adapt to initial insult (**FP**)
- *In vitro* assays are not perfect! (**FP**, **FN**)
- *In vivo* rodent data is not perfect! (**FP**, **FN**)

Beyond *in vitro* to *in vivo* signatures



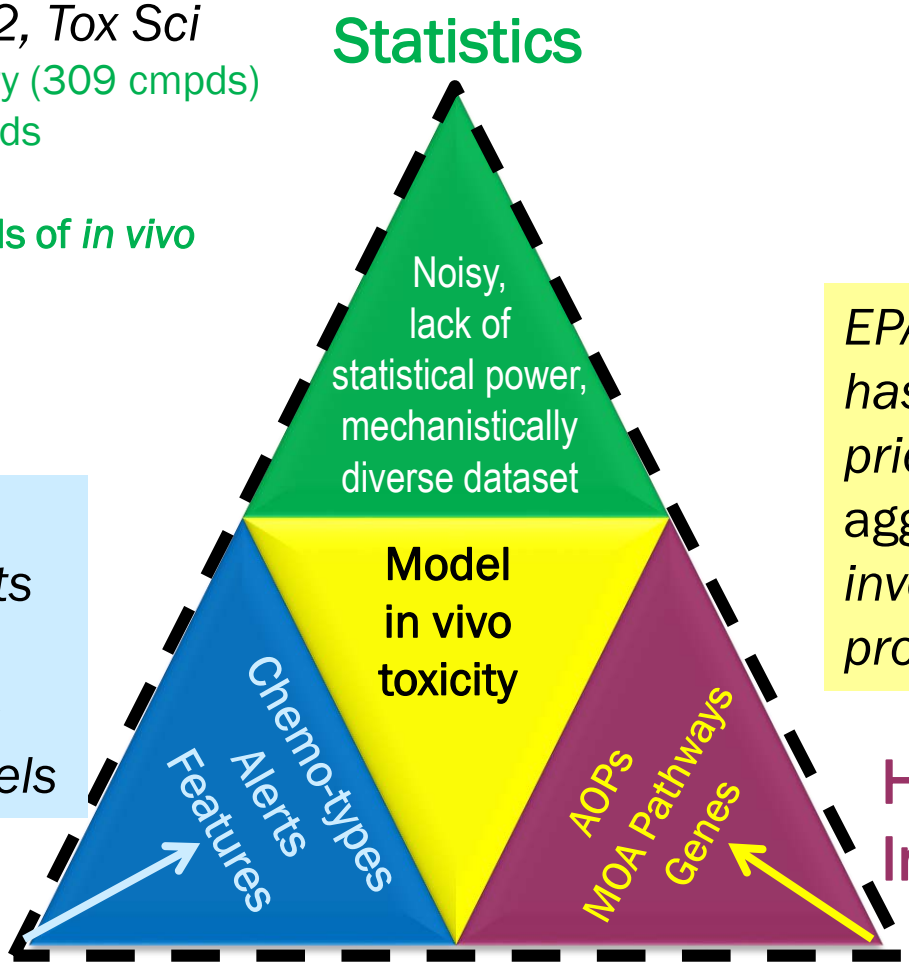
Why is Chemistry needed?

Thomas et al, 2012, Tox Sci

- ToxCast Phase I library (309 cmpds)
- >80 statistical methods
- ToxRef DB endpoints

→ No successful models of *in vivo* endpoints

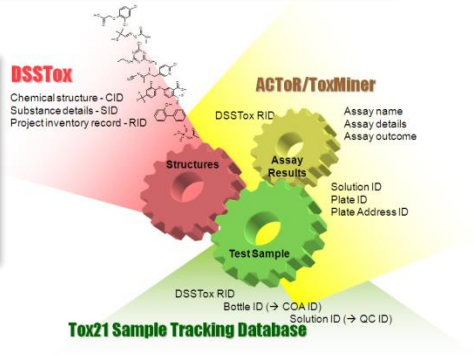
Prior knowledge of chemical determinants of reactivity & toxicity are needed to build & refine predictive models

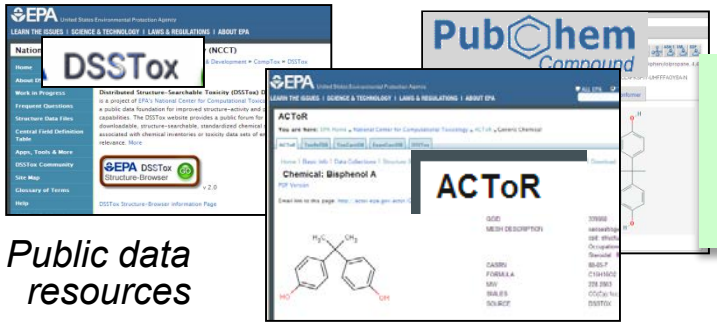


EPA's modeling success has relied upon use of prior knowledge & aggregation to focus investigations into productive areas

Chemistry: What's needed?

- Accurate chemical annotations of testing libraries (e.g., ToxCast & Tox21), transparency, & reporting of error sources
- Cheminformatics foundation to enable structure modeling

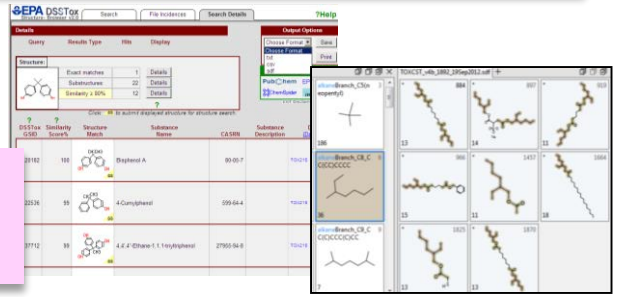




Public data resources

Public data release: ability of non-chemists (biologists, statisticians) to access & utilize chemical information

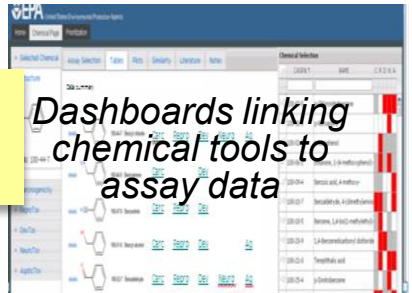
Structure-similarity searching



Use all available data (HTS+chemistry) to guide & inform analog selection & model predictions

Informed feature-grouping & highlighting

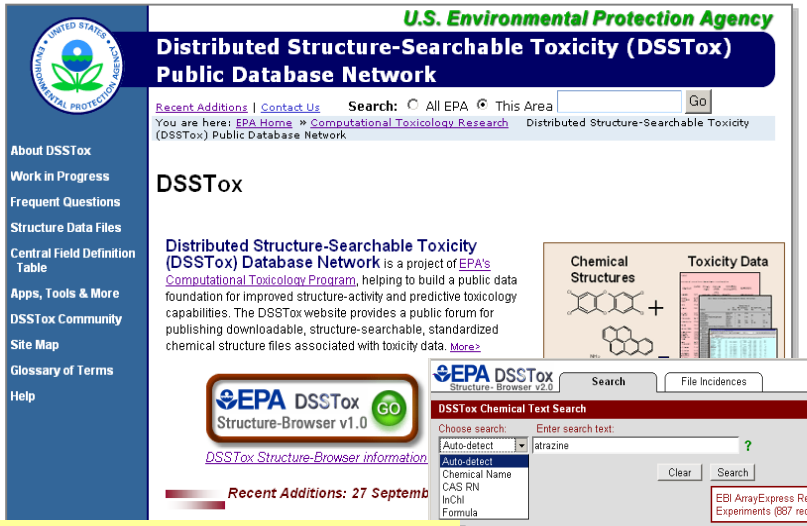
Incorporate chemical information into usable tools for chemical prioritization & safety assessments

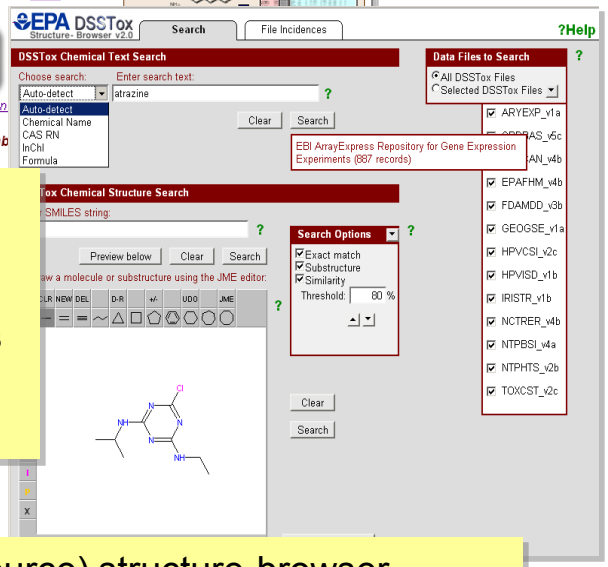


Dashboards linking chemical tools to assay data

CSS: Chemical Data Resources & Databases

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





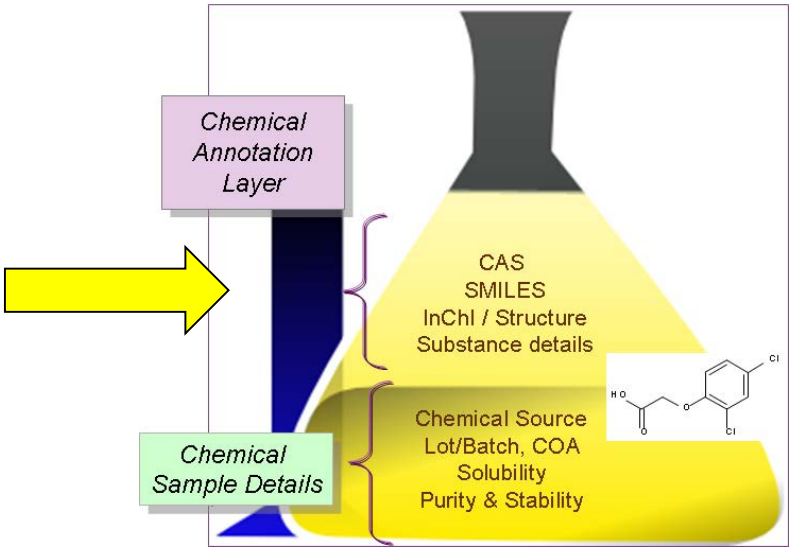
- Public chemical-toxicity (SAR) data resource
- Source of high quality chemical structures across high-interest EPA inventories

- Public (open source) structure-browser provides inventory-specific structure-similarity searches & external linkages

DSSTox Master DB (Access → MySQL)

- ✓ 16K structures (mol, SMILES, InChI)
- ✓ 18K registered substances
- ✓ 22K CAS (6K STN-checked, 3K deleted)
- ✓ 36 inventories (45K record IDs)

ToxCast & Tox21 chemical QC & data management



Toxicity Prediction Challenge:

Bringing all knowledge & data to bear on problem

Data-mining

Biologically-based QSAR & Cheminformatics

Aggregation

Adverse Outcomes:
> Pathways
> Genes
> Assays
+ Statistical associations

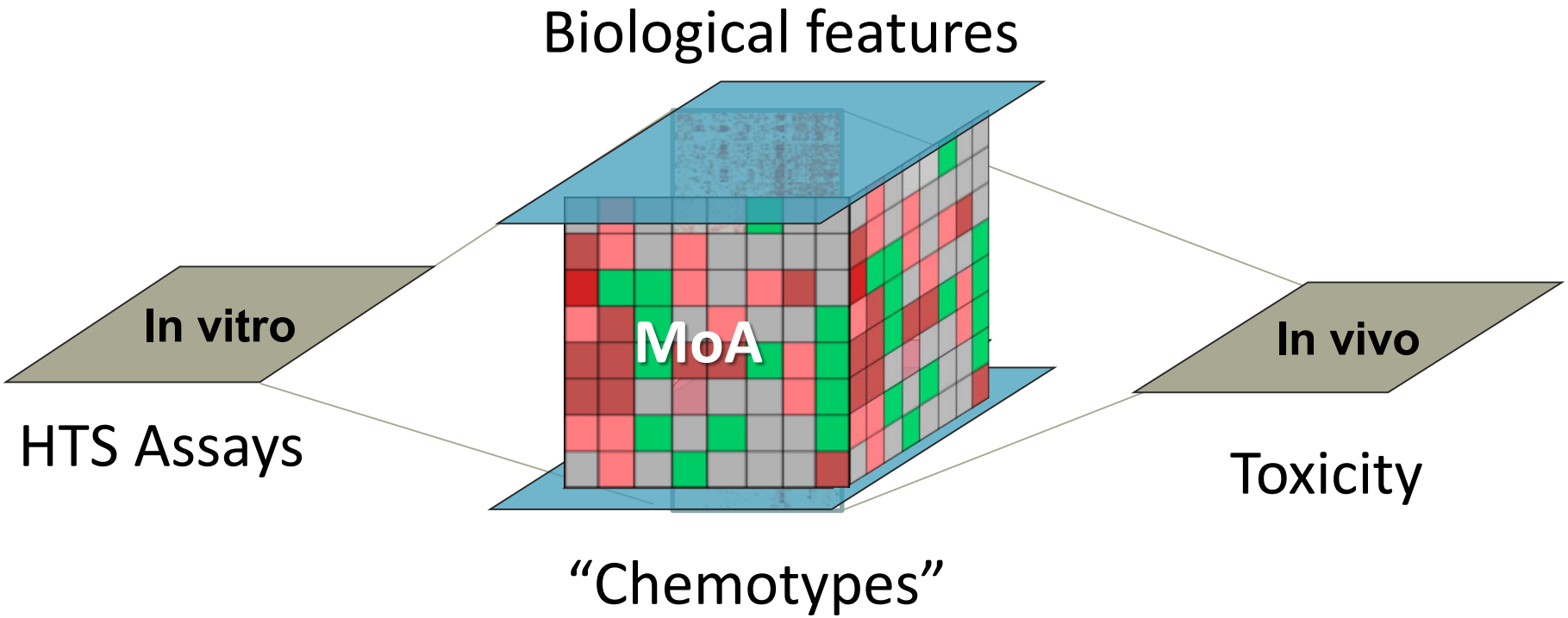
Reactivity & toxicity-informed features & classes

Mechanistically well-defined toxicity endpoint

Existing knowledge

QSAR using biologically informed chemical features

"MoA QSAR"



HTS results are used to inform feature selection, linking chemical features to putative toxicity Mode of Action (MoA) of toxicity

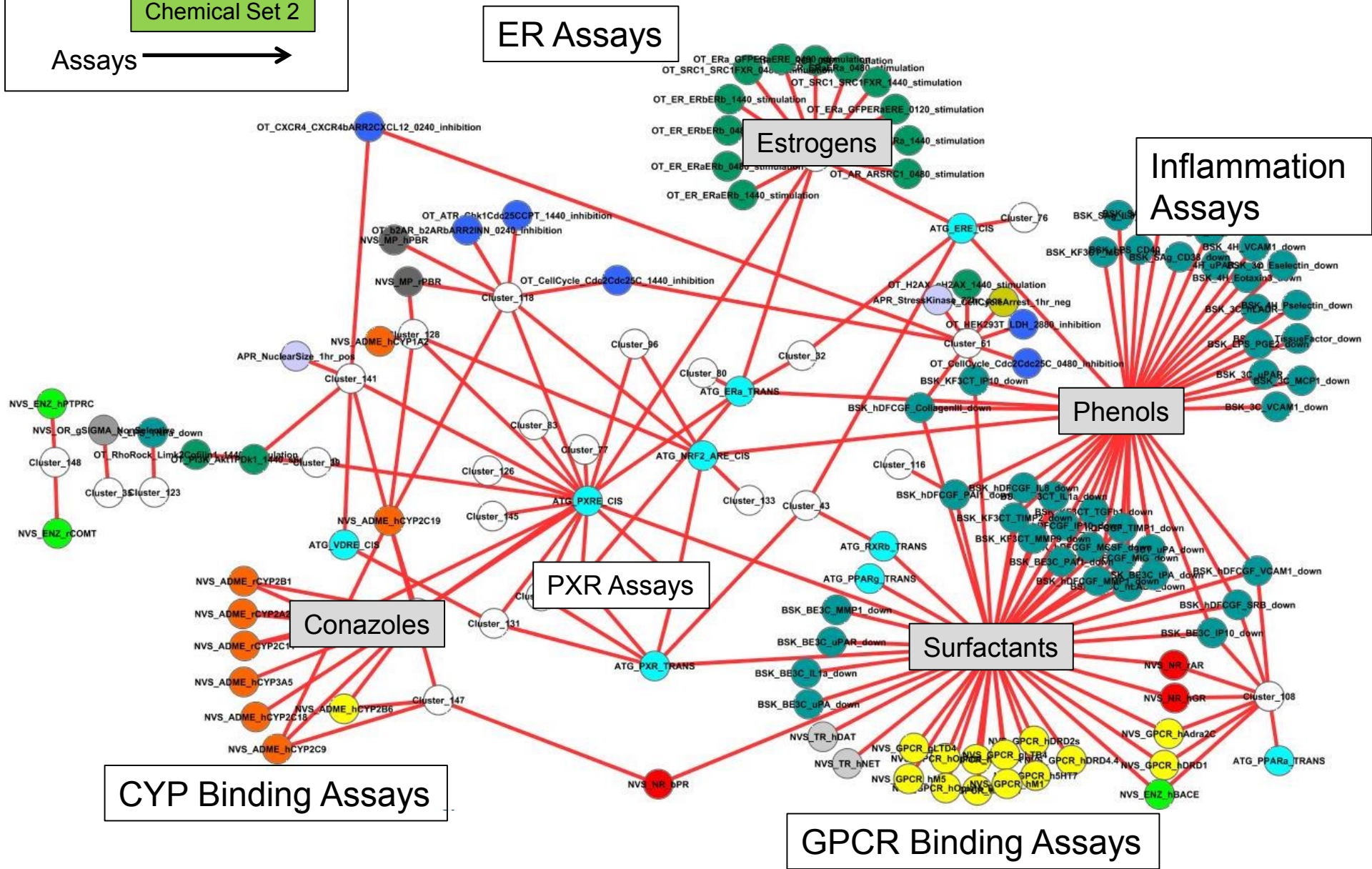
Data Set Incomplete

Chemical Set 1

Chemical Set 2

Assays →

Clusters 80% predictive of assay hit



Understanding Success and Failure

- Why *In vitro* to *in vivo* can work:
  - Chemicals cause effects through direct molecular interactions that we can measure with *in vitro* assays

- Why *in vitro* to *in vivo* does not always work:

Systems Models

- Pharmacokinetics issues: biotransformation, clearance (**FP**, **FN**)
- Assay coverage: don't have all the right assays (**FN**)
- Tissue issues: may need multi-cellular networks and physiological signaling (**FN**)
- Statistical power issues: need enough chemicals acting through a given MOA to be able to build and test model (**FN**)
- Homeostasis: A multi-cellular system may adapt to initial insult (**FP**)
- *In vitro* assays are not perfect! (**FP**, **FN**)
- *In vivo* rodent data is not perfect! (**FP**, **FN**)

ToxCast Phase II Data Release

- ToxCast Assay Summary Activity Files (toxminer_v19b)
  - Rows of Chemicals, Columns of Assays, Intersection of AC50, EMAX
- ToxCast Assay Annotation Files (toxcast_assay_annotation_v1)
  - Assignment of assay design information
  - Assignment of target information (gene target)
- ToxCast Chemical Library & Structure Files (dsstox)
- ToxCast Concentration Response Data Files (toxminer_v19b)
  - Detailed files of normalized data (>50M Rows)
- ToxRefDB Effect & Endpoint Data Files (toxrefdb)
  - Flattened version of ToxRefDB with all effect information (>6000 studies & 1000 chemicals)
  - Endpoint summary file that has NEL/LEL and NOAEL/LOAEL across all studies

ToxCast Data Overview Video

The Interactive Chemical Safety for Sustainability (iCSS) web application is releasing its first dashboard, called the ToxCast Dashboard. The ToxCast Dashboard is intended to provide an interactive data exploration tool. We are currently releasing ToxCast Dashboard version 0.5, a beta version of the application.

If you would like to provide feedback or be on a mailing list that provides updates on new releases of the ToxCast Dashboard as well as ToxCast data

ToxCast Dashboard v0.5 provides users with the ability to perform basic data and chemical selection as well as simple data exploration in a seemless environment. We will be striving to continuously add functionality and improve overall utility and performance. The initial release also intends to convey the conceptual framework and design of the iCSS web application with the intention of producing updated versions of the ToxCast Dashboard as well as additional Dashboards.

The ToxCast Dashboard contains the results of over 800 Assay Endpoints (High Throughput Screening (HTS) Data) across over 1800 chemicals from 7 primary HTS assay sources. The release of the ToxCast Dashboard coincides with the release of the ToxCast Phase II data, which is available below.

ToxCast Phase II Data Release:

ToxCast_Assay Summary Activity Files (toxminer_v19b)

ToxCast Assay Annotation Files (toxcast_assay_annotation_v1)

ToxCast Chemical Library & Structure Files (dsstox)

ToxCast Concentration Response Data Files (toxminer_v19b)

ToxRefDB Effect & Endpoint Data Files (toxrefdb)

Disclaimer:

ToxCast Data will change over time as our understanding of

EPA United States Environmental Agency

EPA iCSS

TOXCAST HOME | DATA SELECTION | CHEMICAL SELECTION | DATA EXPLORER | CHEMICAL EXPLORER | PRIORITIZATION COMING SOON!

DATA EXPLORER

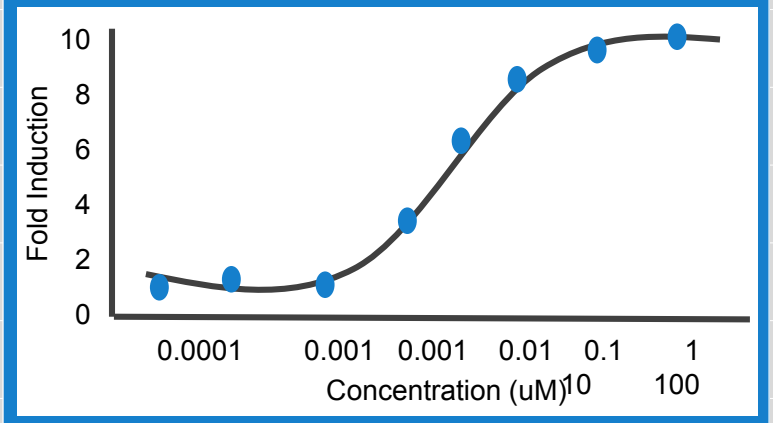
Assay List

Data Class	Assay Endpoint
ATG	ATG_ERa_TRANS
ATG	ATG_ERE_CIS
ATG	ATG_PPARg_TRANS
ATG	ATG_PPRE_CIS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS
ATG	ATG_ERa_TRANS

Summary Activity Table

CASRN	Chemical Name	Q	AC50	T	B	W	Emax
80-05-7	Bisphenol A						
80-05-1	Bisphenol B						
80-05-2	Bisphenol C						
80-05-3	Bisphenol D						
80-05-4	Bisphenol E						
80-05-5	Bisphenol F						
80-05-6	Bisphenol G						
80-05-8	Bisphenol H						
80-05-9	Bisphenol I						

CONC RESPONSE PLOT



Summary

- Goal: use *in vitro* assays to screen and prioritize many data-poor chemicals
- Signature generation uses combination of biological insight and statistics
- Initial models point the way to real-world applications
- Further refinements are in the works
  - More chemicals and assays
  - Use of chemoinformatics
  - Systems-level models
  - Targeted testing approaches

EPA National Center for Computational Toxicology Staff
February 5, 2013