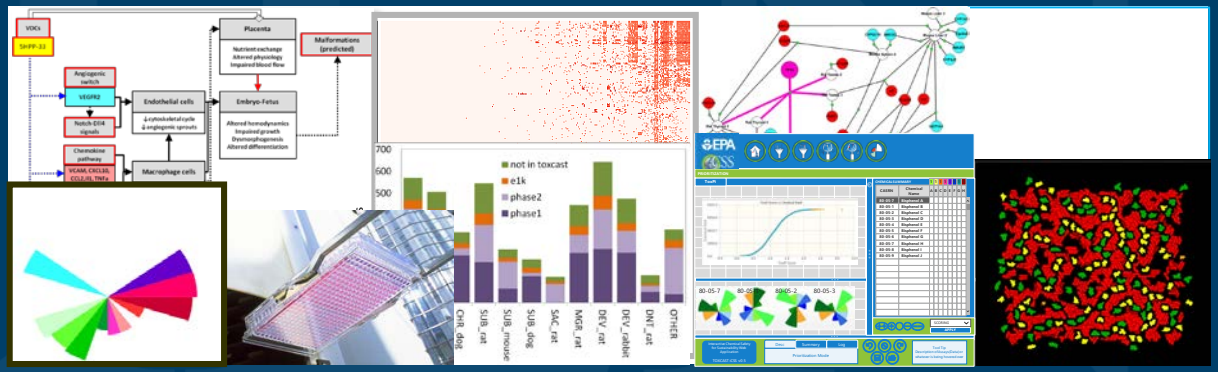
United States
Environmental Protection
Agency

Developing a Roadmap for Integrating Computational and *In Vitro* Approaches in Risk-Based Chemical Safety Decisions



SSCT-SweTox Workshop
October 14, 2015

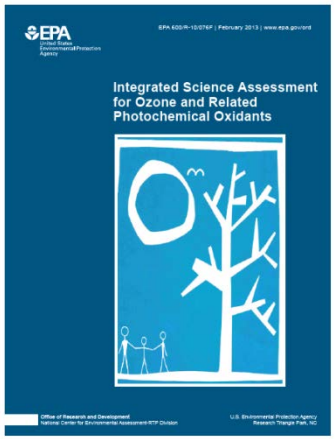
Rusty Thomas
Director
National Center for Computational Toxicology

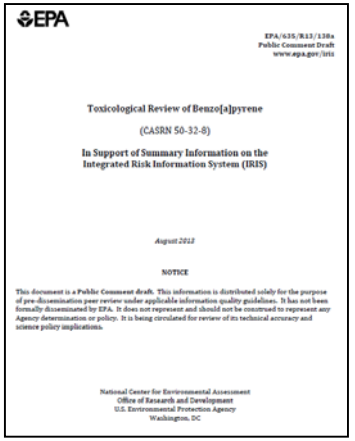
The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the U.S. EPA

Multiple Destinations Necessary to Address Diverse Regulatory Needs…

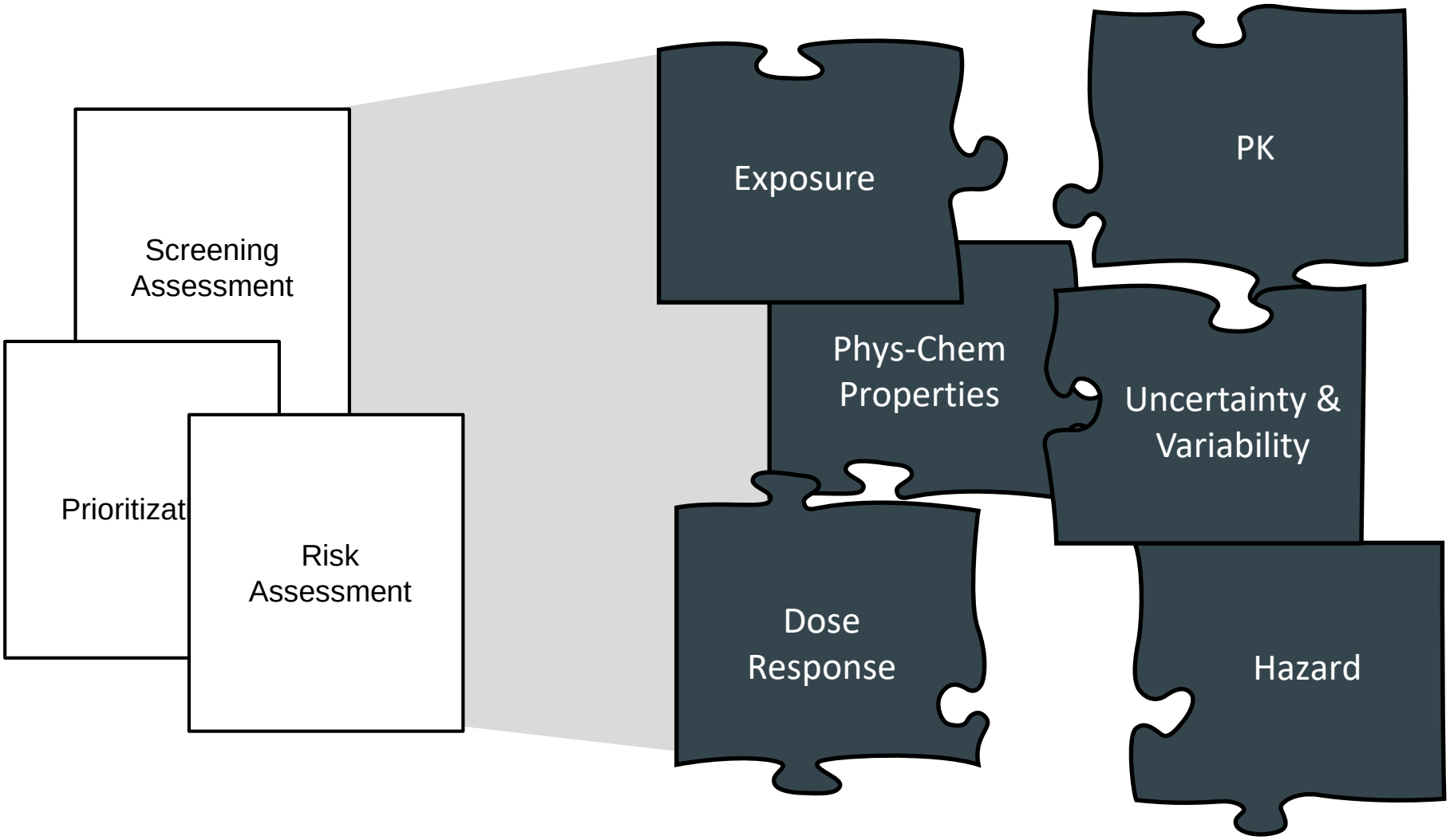
- Multiple drivers shape assessment needs
  - Regulatory demands
  - Economic considerations
  - Multiple applications
- Chemical assessments are “fit-for-purpose”
  - Prioritization (e.g., EDSP, PMN, SNUR)
  - Screening-level assessments (e.g., CCL, GreenChem)
  - Provisional assessments (e.g., PPRTVs)
  - Toxicity assessments (e.g., IRIS)
  - Risk assessments (e.g., MCLs, pesticides)







Common Elements are Considered Across Many Decision Contexts



But, the Current Infrastructure is Antiquated and Inefficient

Acute toxicity studies (LD_{50}) developed to standardized batches of pharmaceuticals

Genotoxicity assays developed

Continuous breeding studies for reproductive toxicity

1920 | | 2010

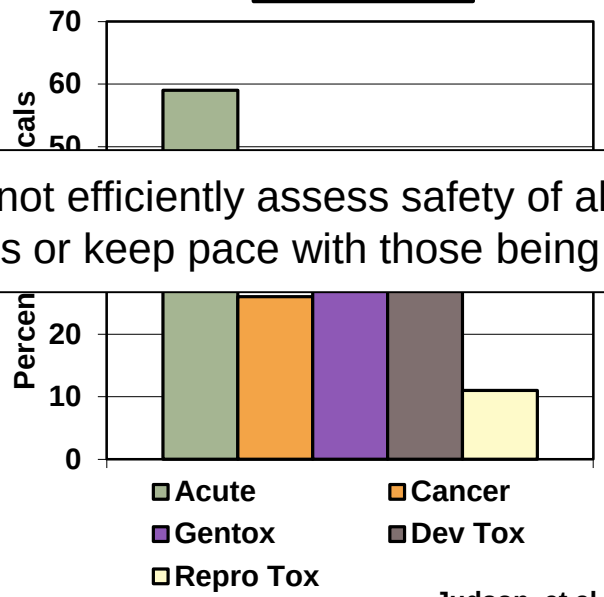
Current testing paradigm does not incorporate advances in technology and does not provide mechanistic data

Draize test introduced for eye irritants

Thalidomide led to testing of pharmaceuticals for developmental toxicity

Rodent cancer bioassay introduced

... and cannot efficiently assess safety of all the existing chemicals or keep pace with those being developed



Judson, et al *EHP* (2010)

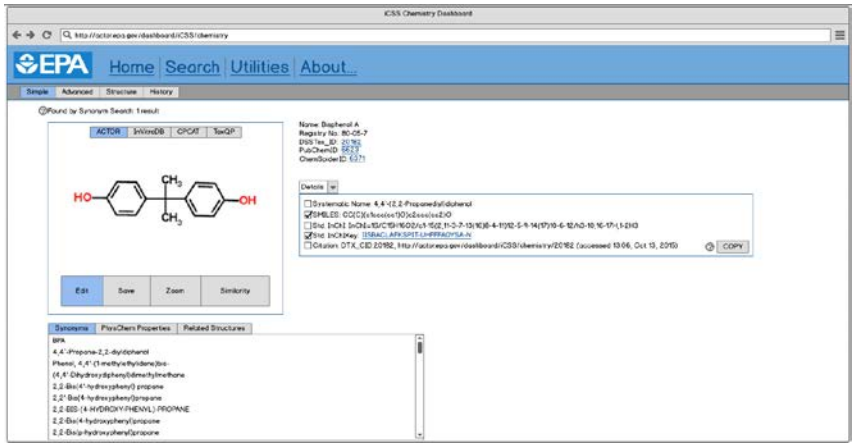
Infrastructure Must Be Flexible and On a Scale Not Yet Employed in Tox



Need to Start With High Quality Chemical Foundation

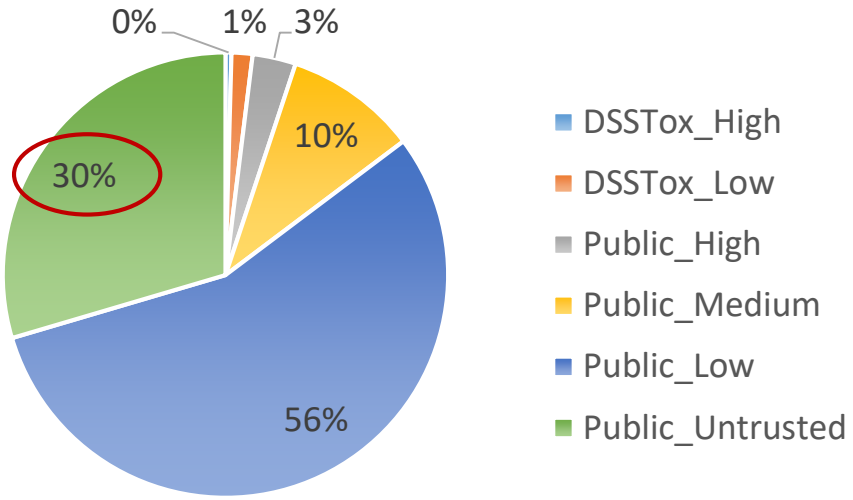
United States Environmental Protection Agency

Chemical Structure & Property Database



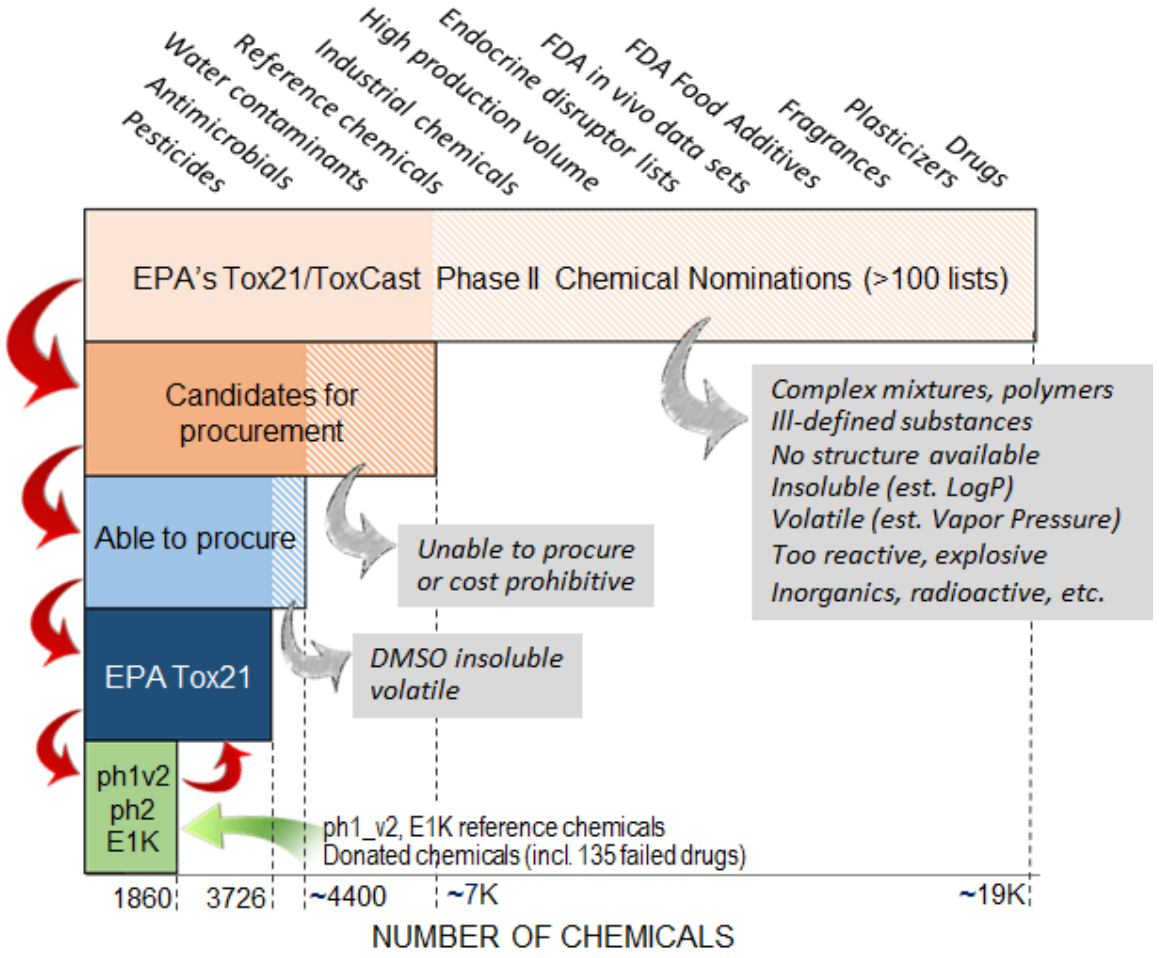
- DSSTox database of over 1 million chemical structure-ID mappings
- DSSTox QC flags indicate confidence in structural associations
- Database of phys-chem property predictions and experimental measurements with QC flags (coming soon)

Quality of Structure-CAS Mappings

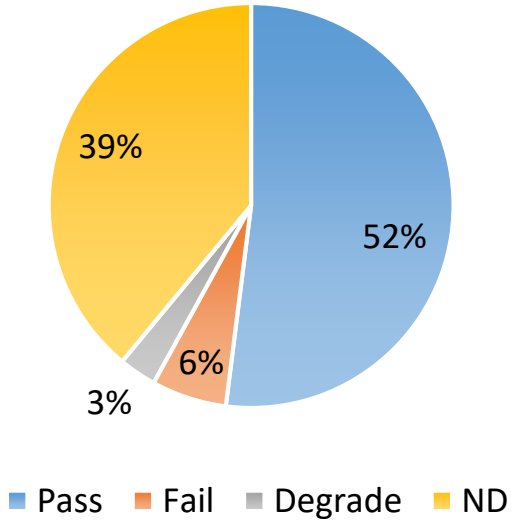




Need to Start With High Quality Chemical Foundation



Analytical QC of Chemical Library

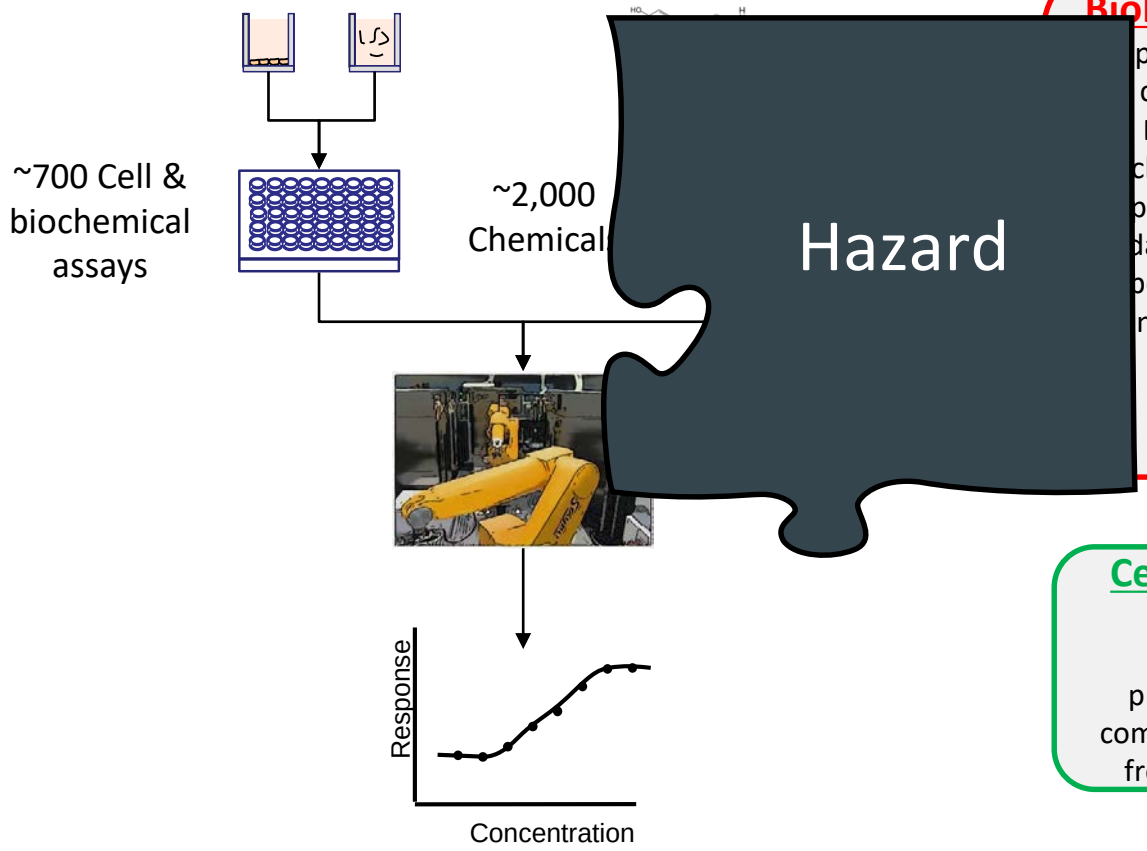


- EPA ToxCast chemical screening library of 3,729 unique chemicals

Pass = C (75%) or greater
Fail = D, F, Ac, Bc, Cc

High-Throughput Bioactivity Screening as an Indicator of Hazard

ToxCast



Biological Response

proliferation and death
cell differentiation
Enzymatic activity
chondrial depolarization
protein stabilization
dative phosphorylation
porter gene activation
ne 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

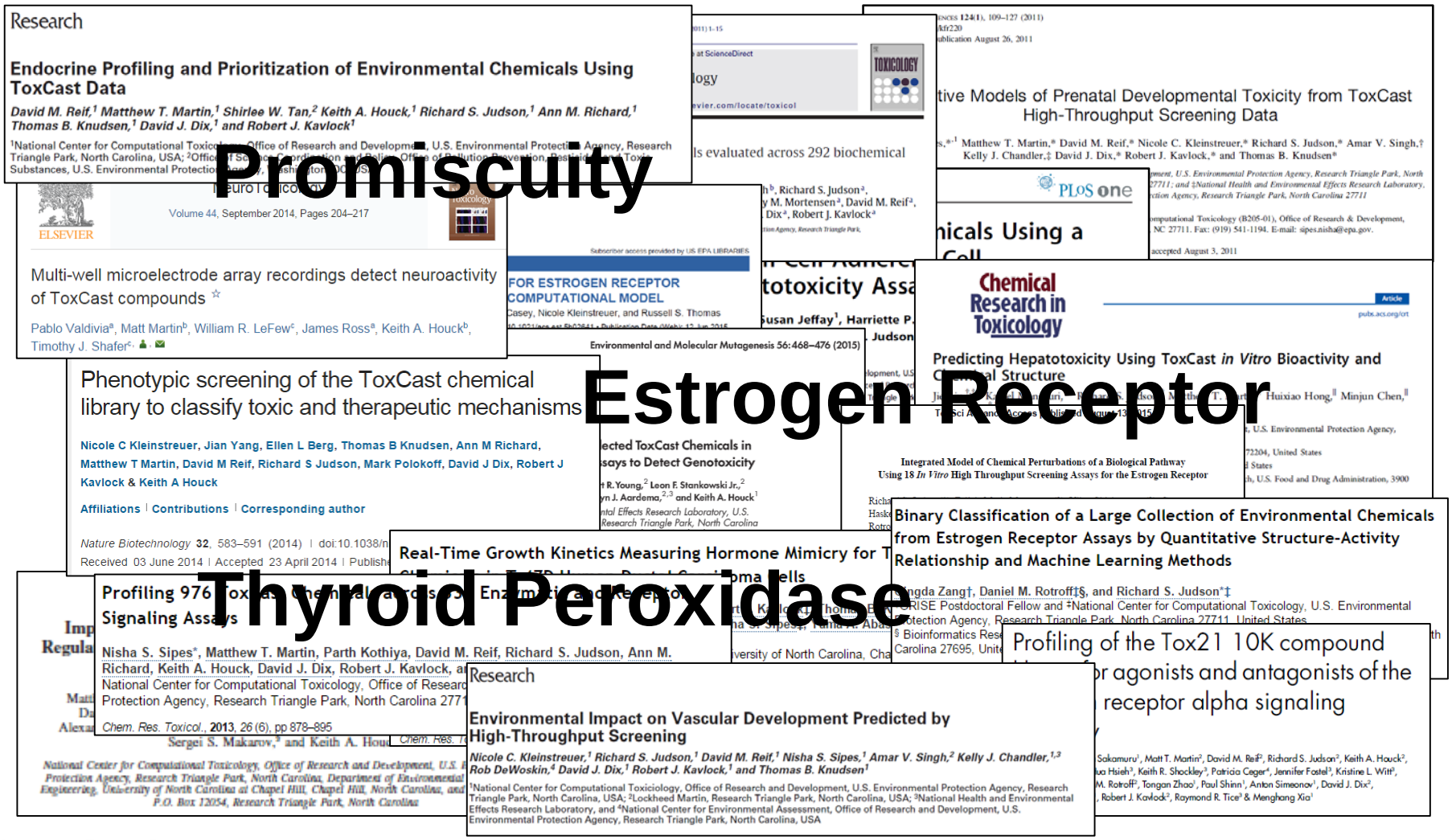
Cell Format

cell free
cell lines
primary cells
complex cultures
free embryos

Assay Design

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

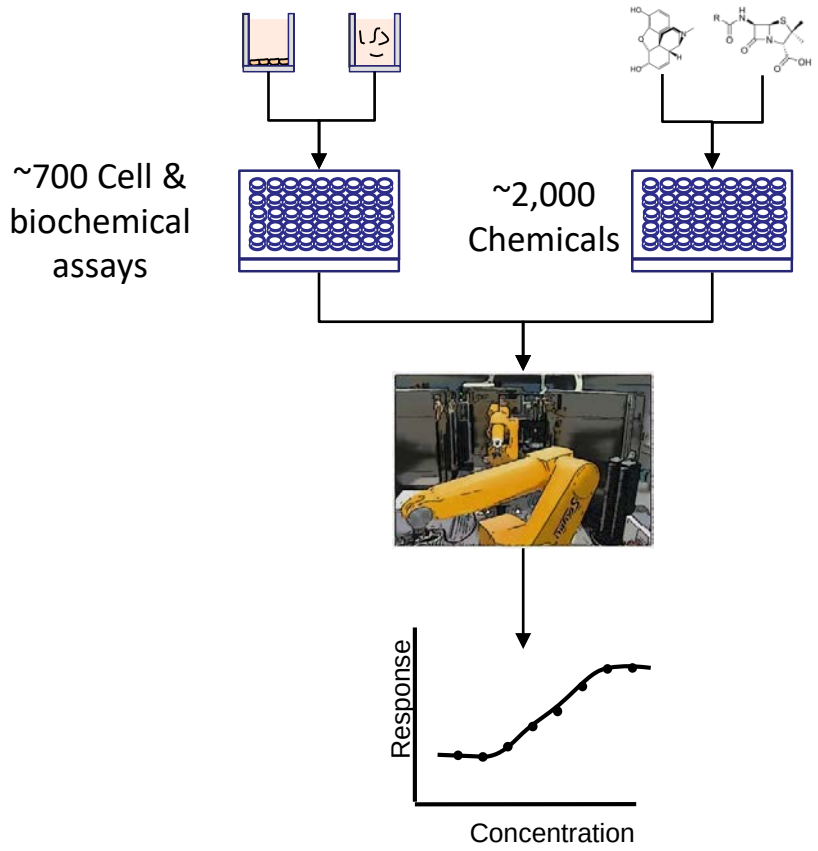
Lots of Possible Research to Highlight

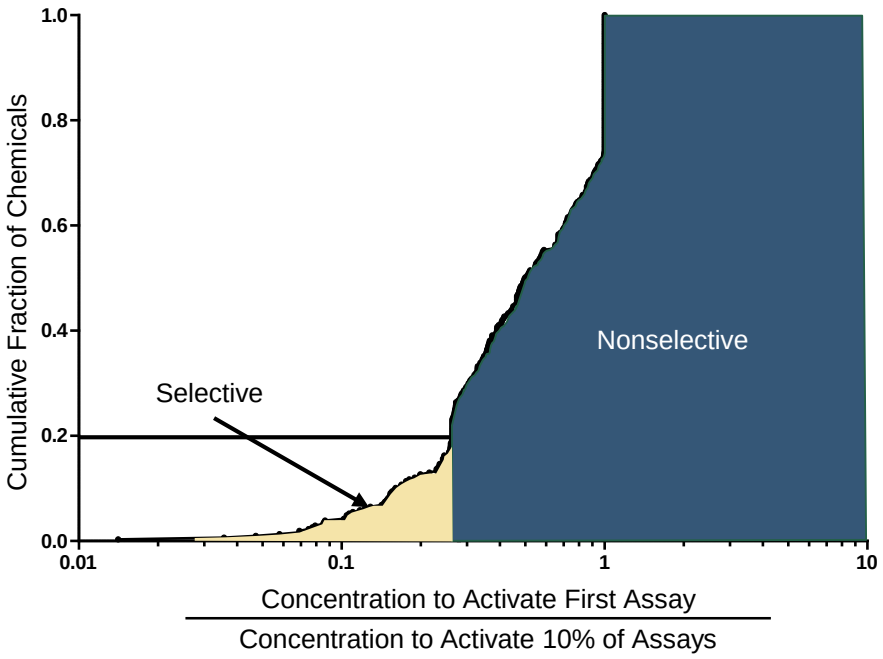


Most Environmental Chemicals are Nonselective for Biological Targets

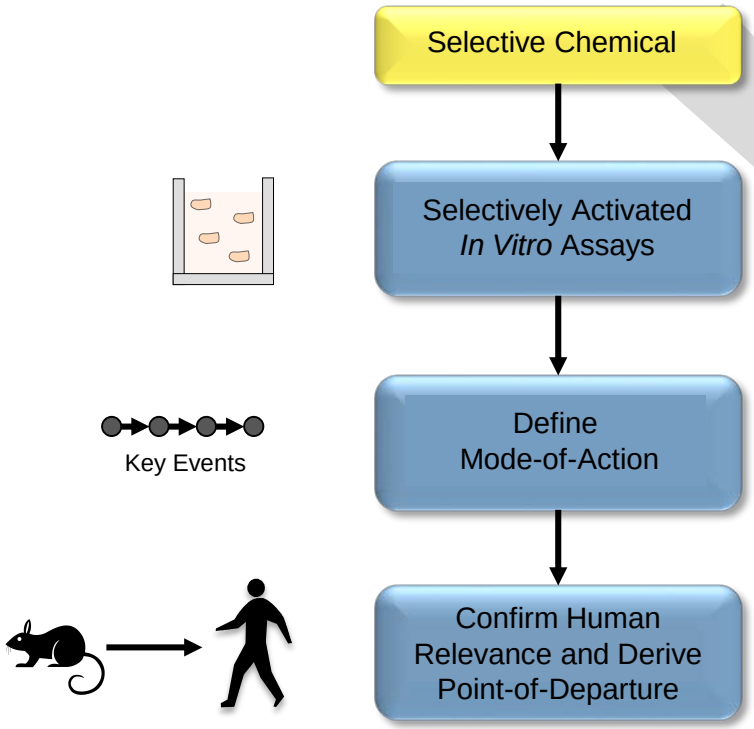
United States Environmental Protection Agency

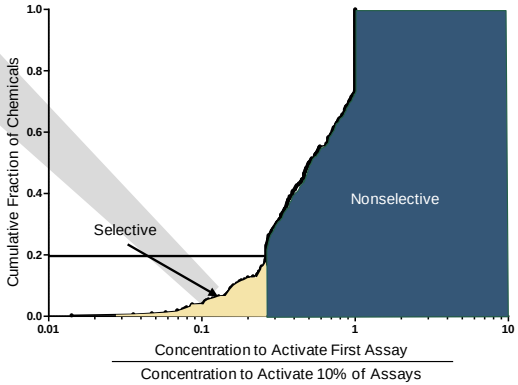
ToxCast



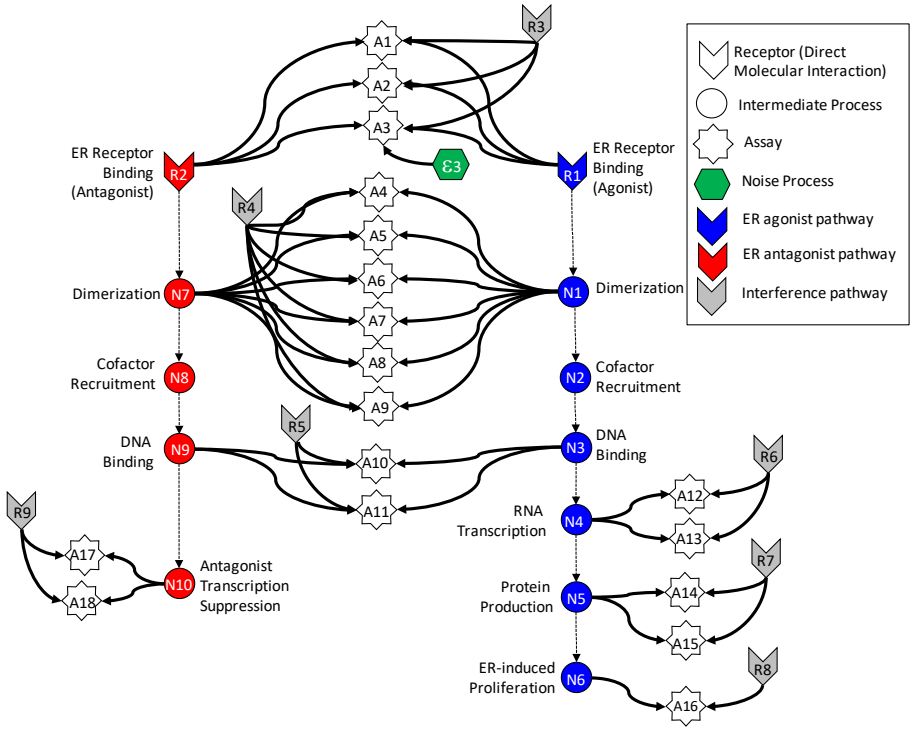


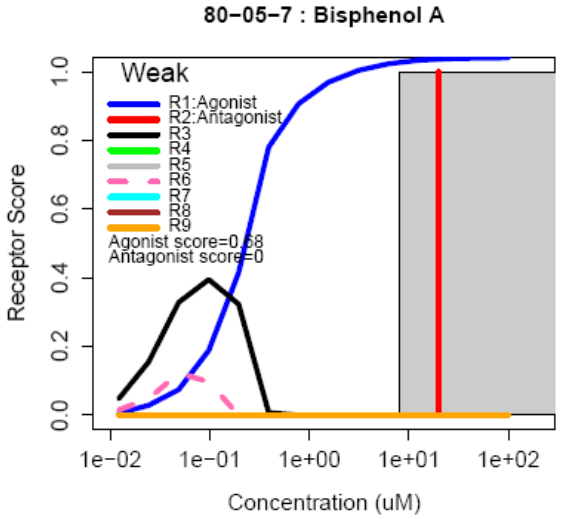
In Vitro Assay Selectivity as a Starting Point for AOPs

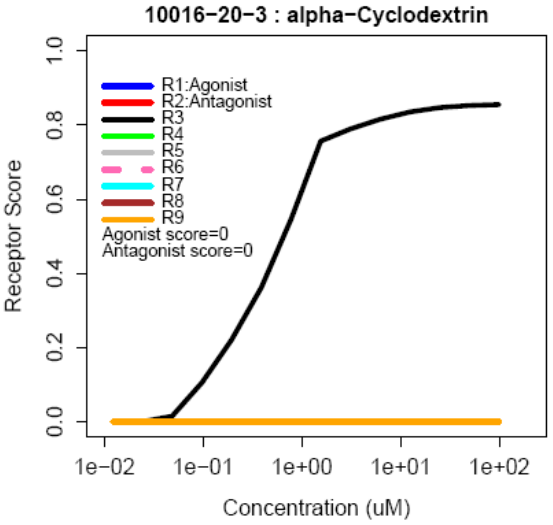




Computational Modeling to Integrate Upstream Events in ER AOP





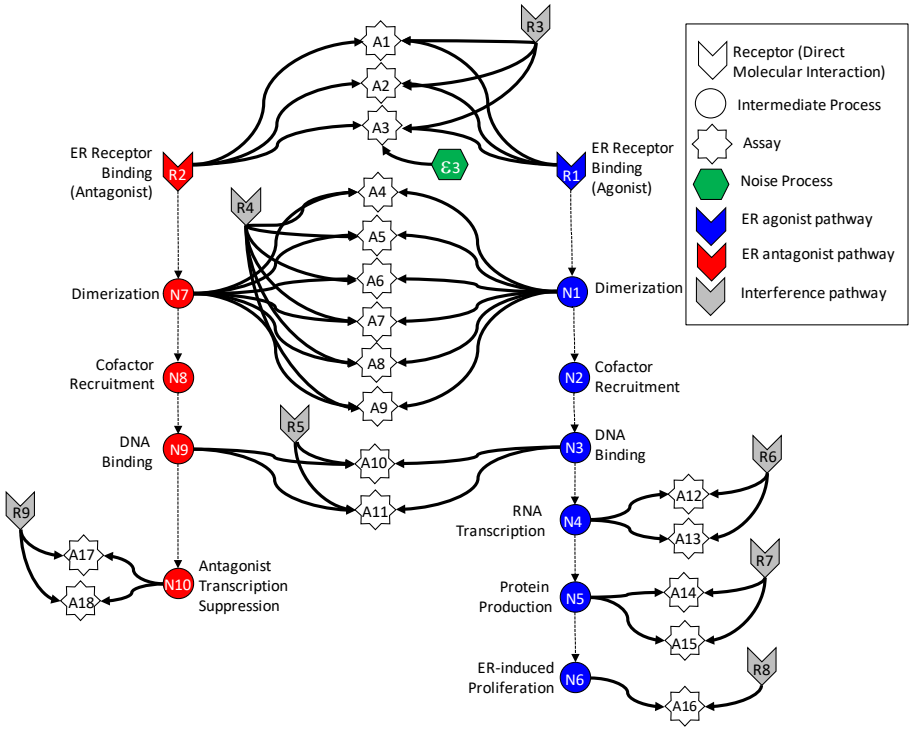


- 18 ToxCast/Tox21 *in vitro* assays measure ER-related activity
- Screened ~2,000 total chemicals
- 38 *in vitro* reference chemicals
- 39 *in vivo* reference chemicals

Judson *et al., Tox Sci.* In Press

Computational Modeling to Integrate Upstream Events in ER AOP

18 ToxCast/Tox21 *In Vitro* Assays Measure ER-Related Activity



Judson *et al., Tox Sci.* In Press
Browne *et al., ES&T.* 2015

In Vitro Reference Chemicals*

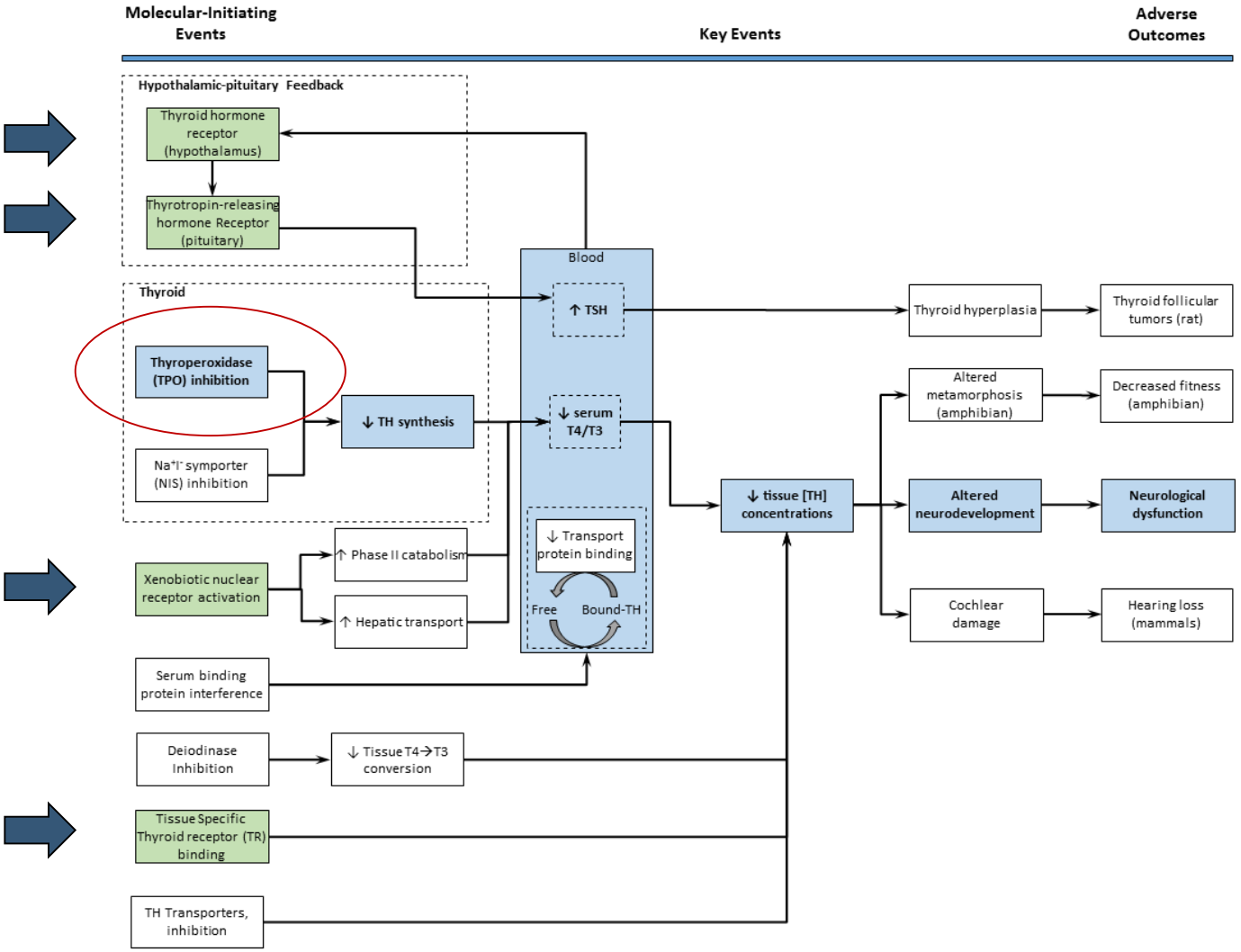
True Positive	26 (25)
True Negative	11 (11)
False Positive	1 (0)
False Negative	2 (2)
Accuracy	**0.93 (0.95)**
Sensitivity	**0.93 (0.93)**
Specificity	**0.92 (1.0)**

In Vivo Reference Chemicals*

True Positive	29 (29)
True Negative	8 (8)
False Positive	5 (1)
False Negative	1 (1)
Accuracy	**0.86 (0.95)**
Sensitivity	**0.97 (0.97)**
Specificity	**0.67 (0.89)**

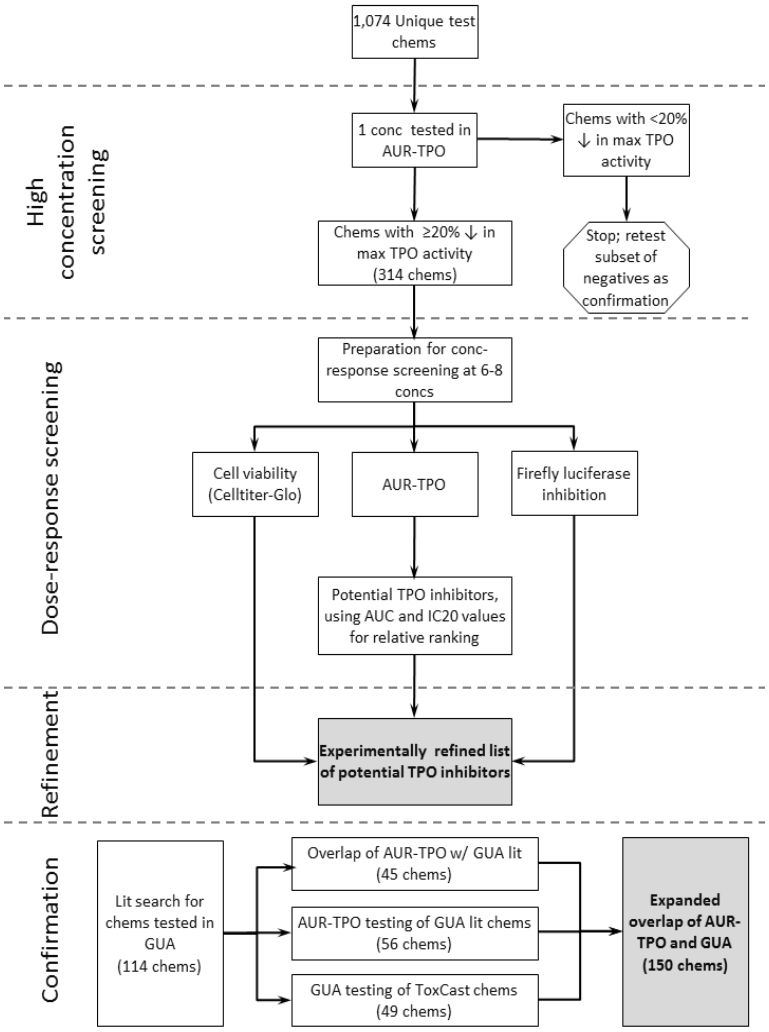
*Values in parentheses exclude inconclusive chemicals

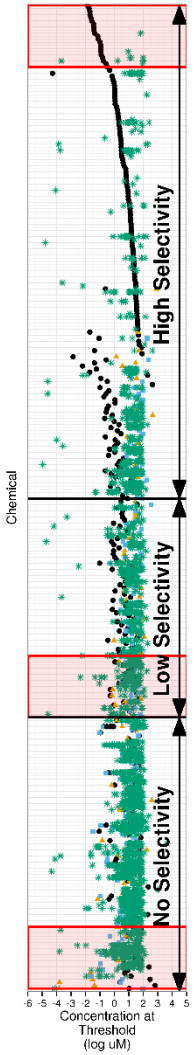
AOP Network for Thyroid Hormone Disruption

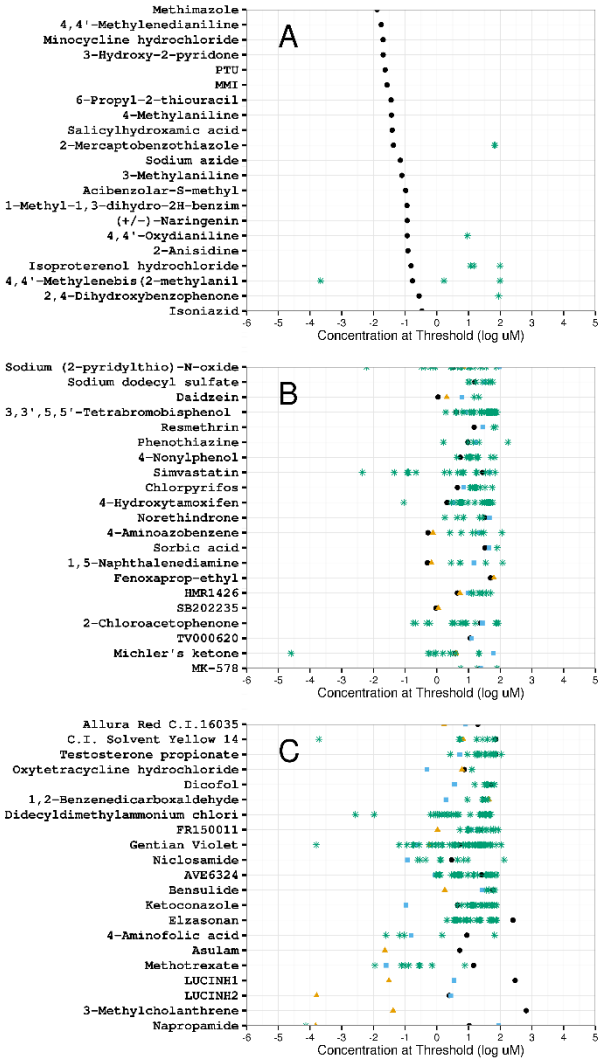


Paul *et al.*, In Review

Identification of Selective TPO Inhibitors

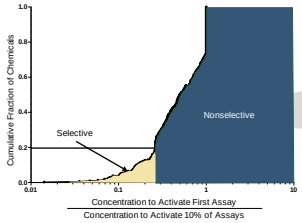




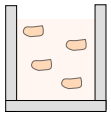


Paul *et al.*, In Review

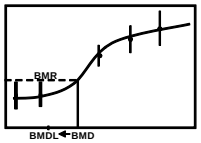
What About the Non-Selective Chemicals?



Nonselective Chemical



?????????



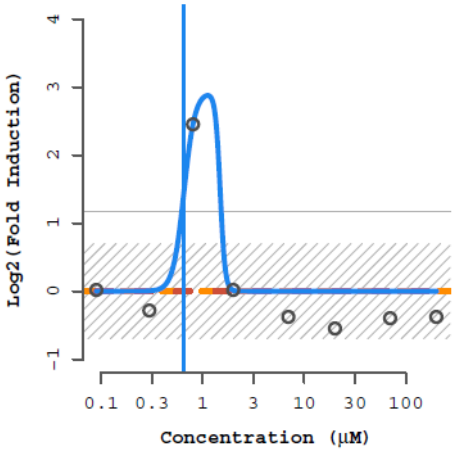
Define Point-of-Departure

Most sensitive response generally protective on a dose level

Absence of activity difficult to interpret

Specific adverse outcome not reliably predicted

Efforts to Ensure HTS Data Quality and Increase Transparency




```
ASSAY:    AEID117 (ATG_ERa_TRANS)

NAME:     Thioglycolic acid
CHID:     26141    CASRN: 68-11-1
SPID(S): TX007664
L4ID:     420385

HILL MODEL (in red):
       tp        ga        gw
val:  3.1e-11  -2.15     0.416
sd:   NaN      NaN       NaN

GAIN-LOSS MODEL (in blue):
       tp        ga        gw        la        lw
val:  2.93      -0.184    8         0.173     18
sd:   3.56      0.334     9.48      5.82      814

       CNST         HILL         GNLS
AIC:  20.14        26.14        17.79
PROB: 0.23         0.01         0.76
RMSE: 0.92         0.92         0.32

MAX_MEAN: 2.45       MAX_MED: 2.45       BMAD: 0.233

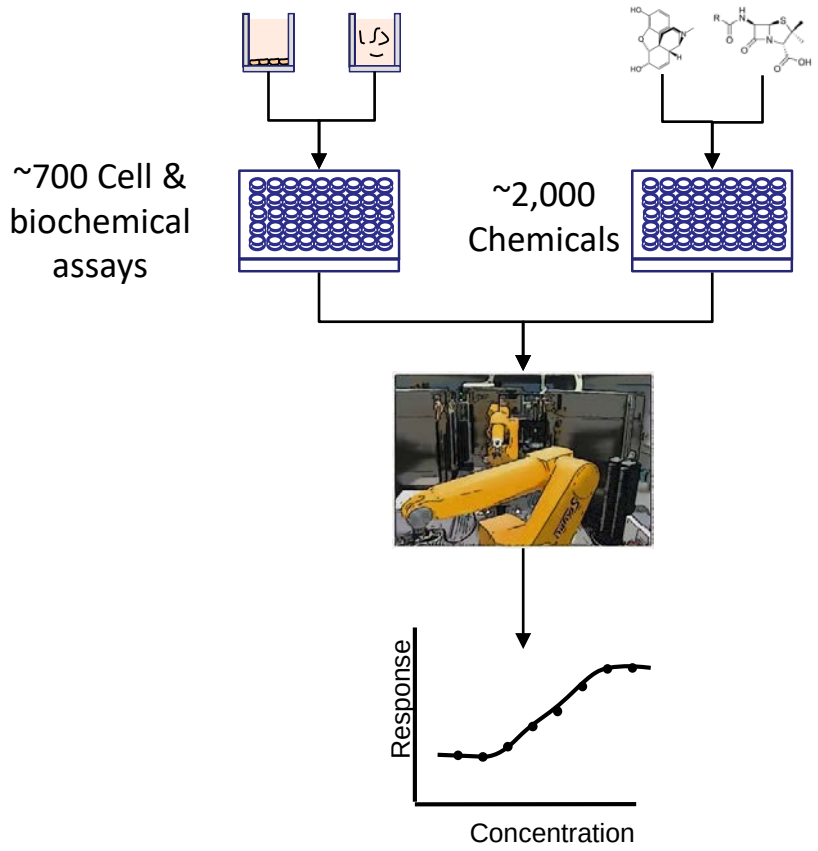
COFF: 1.17    HIT-CALL: 1     FITC: 50    ACTP: 0.77
```

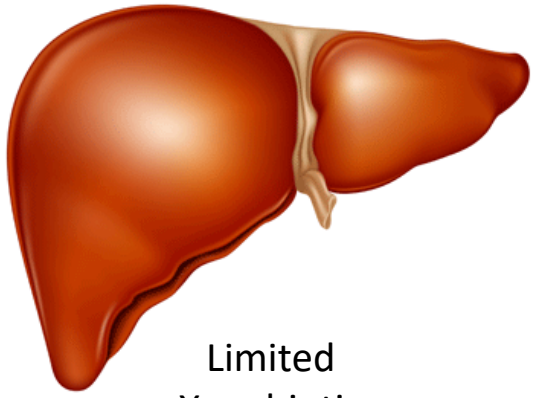
```
FLAGS:
Only one conc above baseline, active
Borderline active
```

- Public release of Tox21 and ToxCast data on PubChem and EPA web site (raw and processed data)
- ToxCast data analysis pipeline has been completely revamped
  - More statistically rigorous and less prone to outliers
  - Data quality flags to indicate concerns with chemical purity and identity, noisy data, systematic assay errors, and activity in range of cytotoxicity
  - Available for download as an R package
- Release of ToxCast “Owner’s Manual”
  - Chemical Procurement and QC
  - Data Analysis
  - Assay Characteristics and Performance
- External audit on ToxCast data and data analysis pipeline
- Continued offering of webinars and workshops to educate stakeholders on high-throughput screening data analysis and interpretation

Efforts to Address Metabolism Challenge

ToxCast

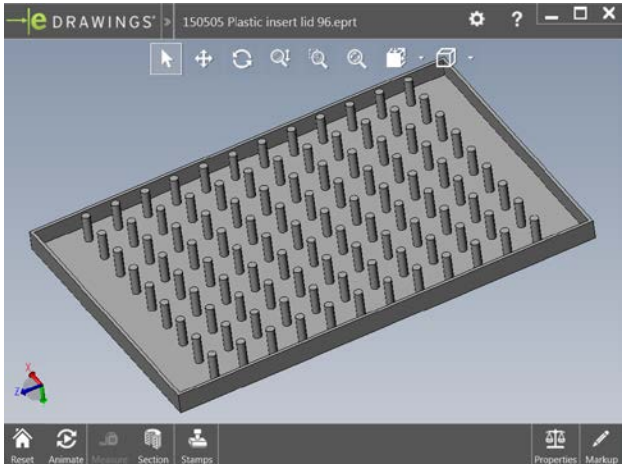


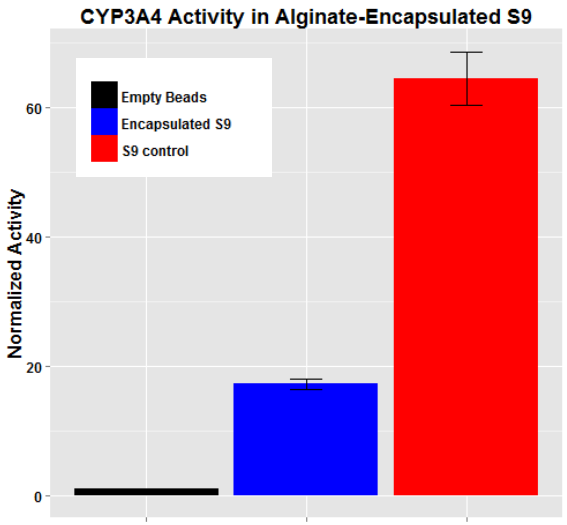


Limited Xenobiotic Metabolism

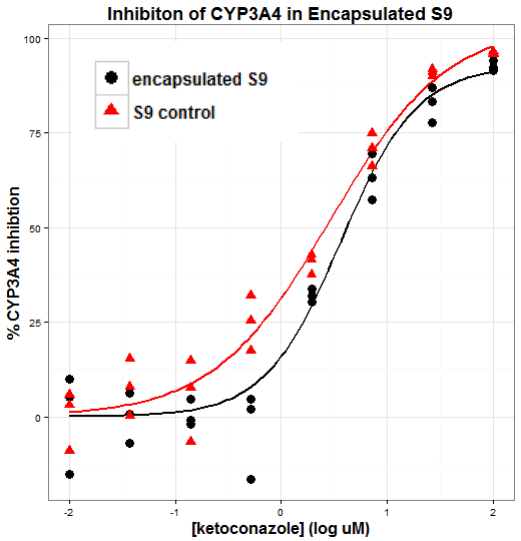
Efforts to Address Metabolism Challenge

In Development





P450 Glo IPA Assay



Efforts to Address Limited Biological Coverage

United States Environmental Protection Agency

ToxCast

~700 Cell & biochemical assays

~2,000 Chemicals



Response

Concentration

Gene Coverage in Development

Multiple cell lines/types

~2,000 Chemicals

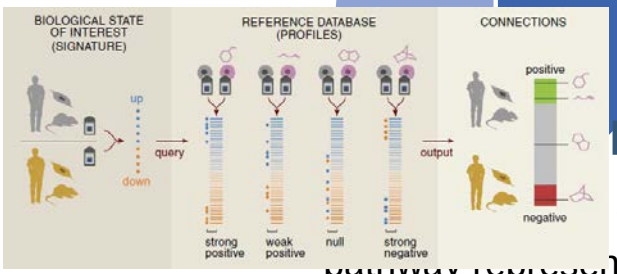
ToxCast

Not in ToxCast

Pathway Coverage

High-Throughput Transcriptomics

- Low-cost
- Whole genome
- 384-well
- Automatable



Response

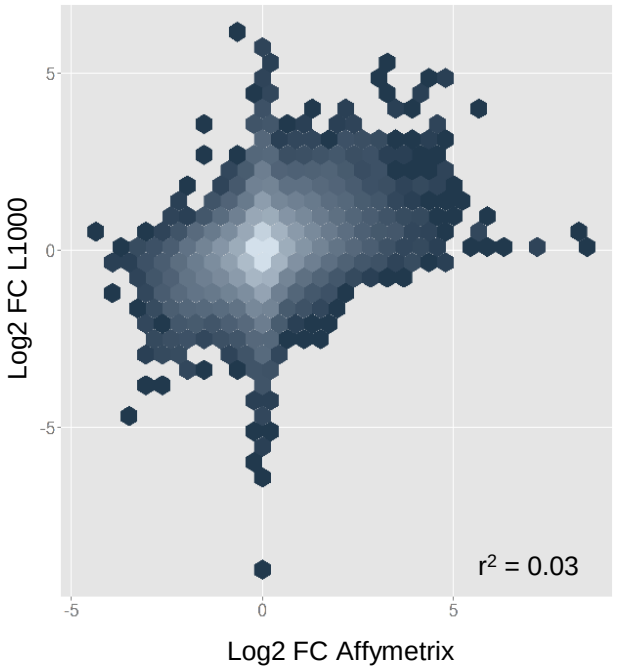
Concentration

from pathway represented

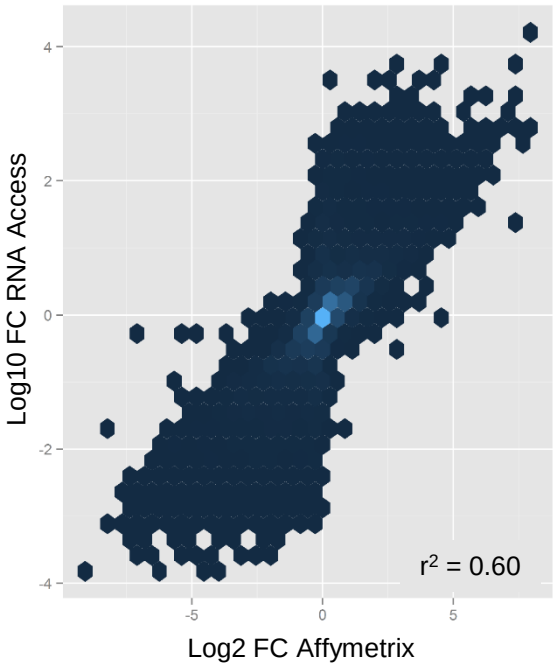
Lamb et al. *Science* (2006)

Efforts to Address Limited Biological Coverage

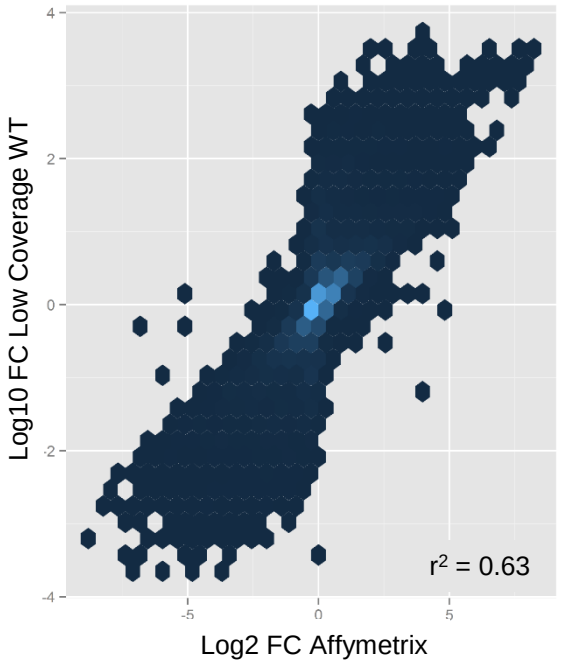
L1000

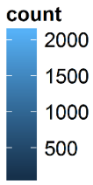


Illumina RNA Access

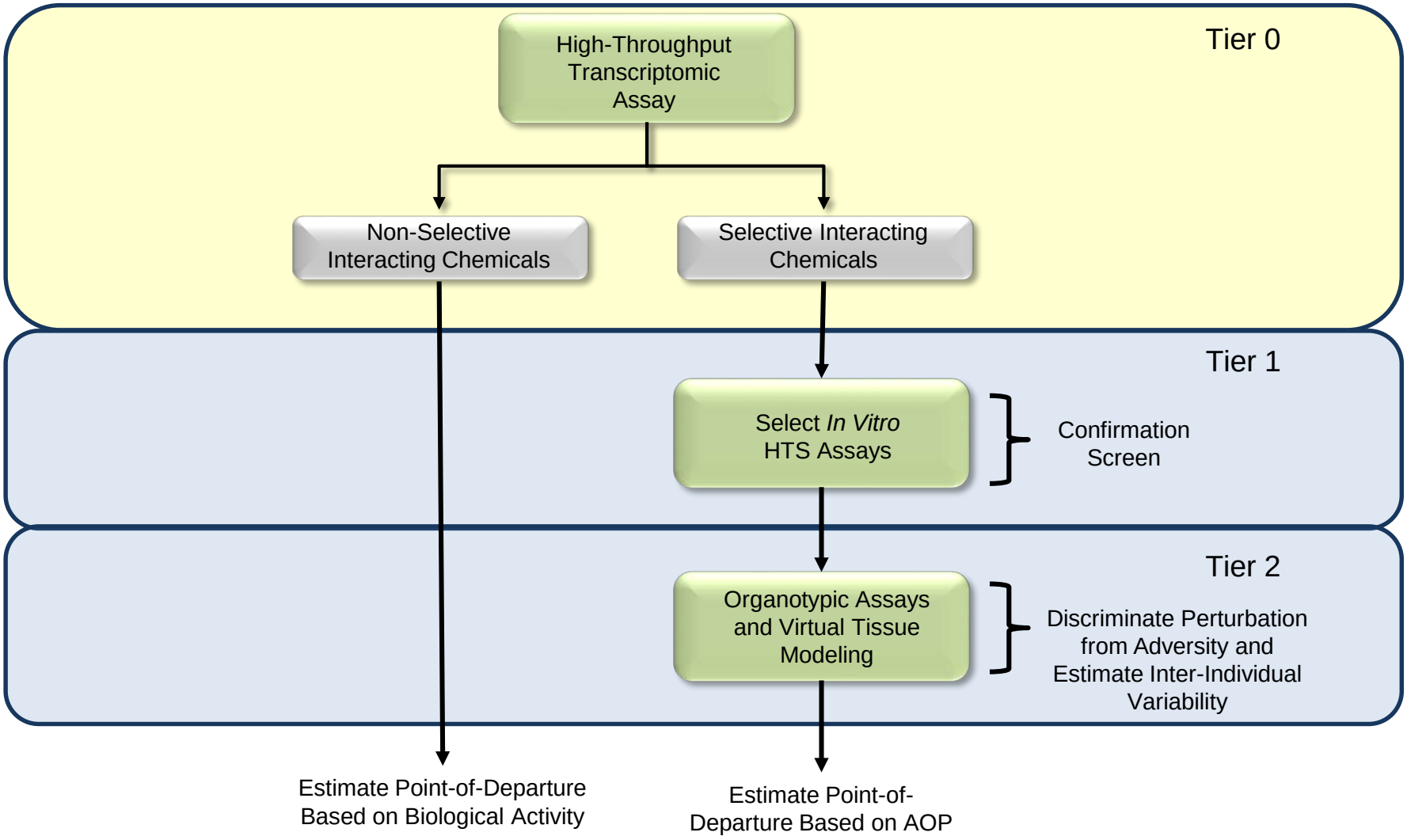


Illumina Low Coverage WT

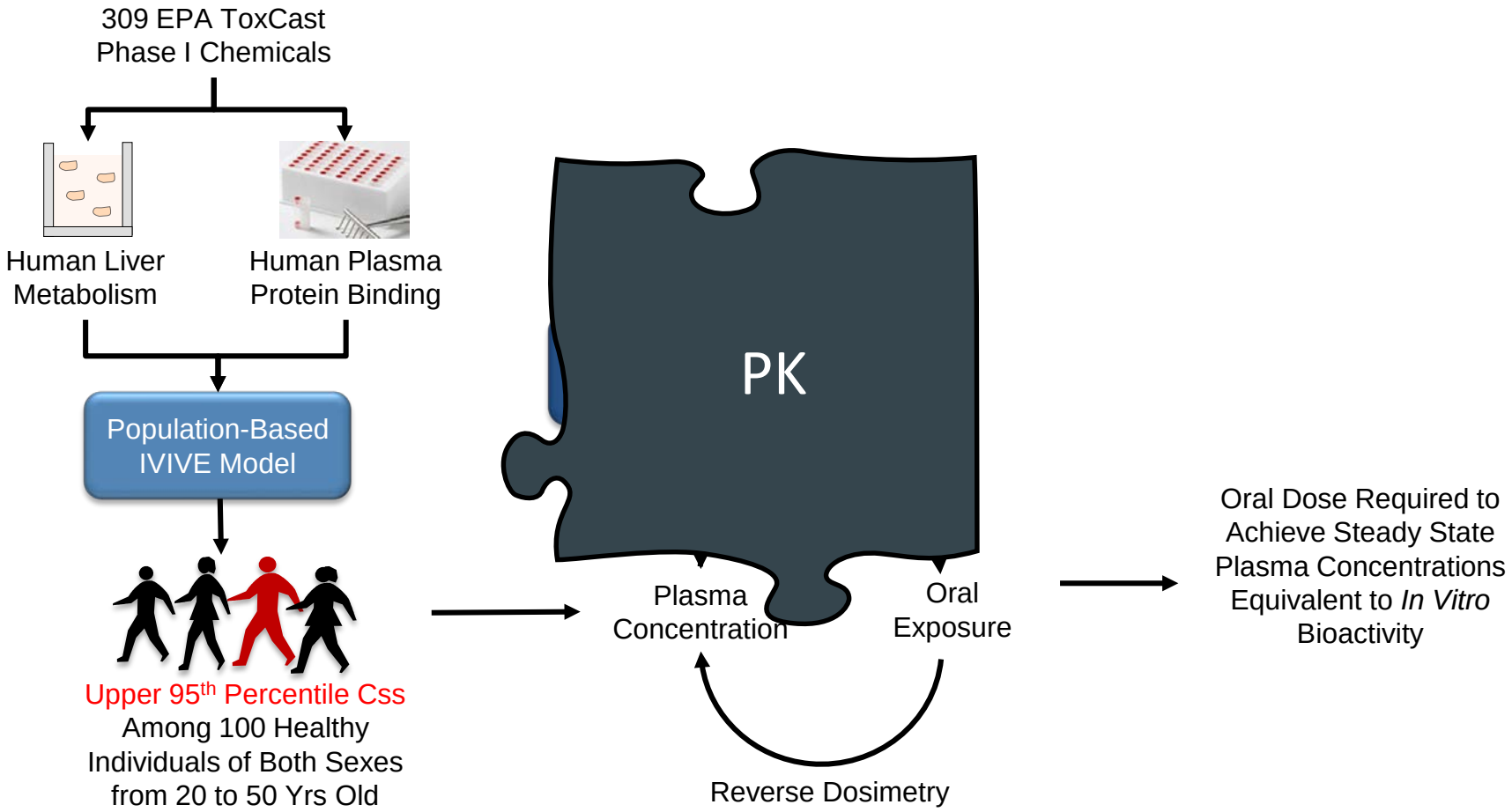




Developing a Broad Hazard Screening Platform

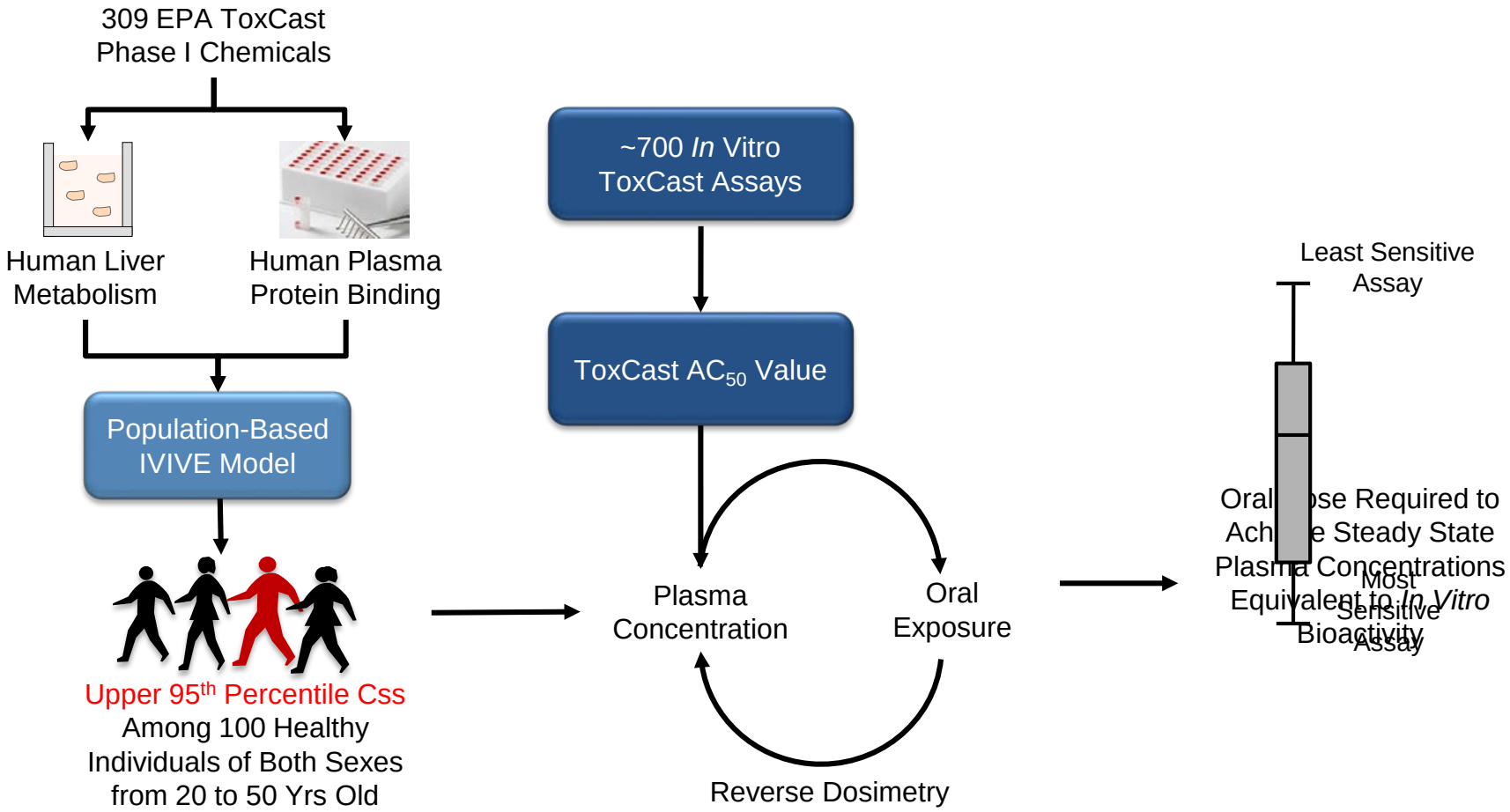


Integrating High-Throughput Pharmacokinetic Approach



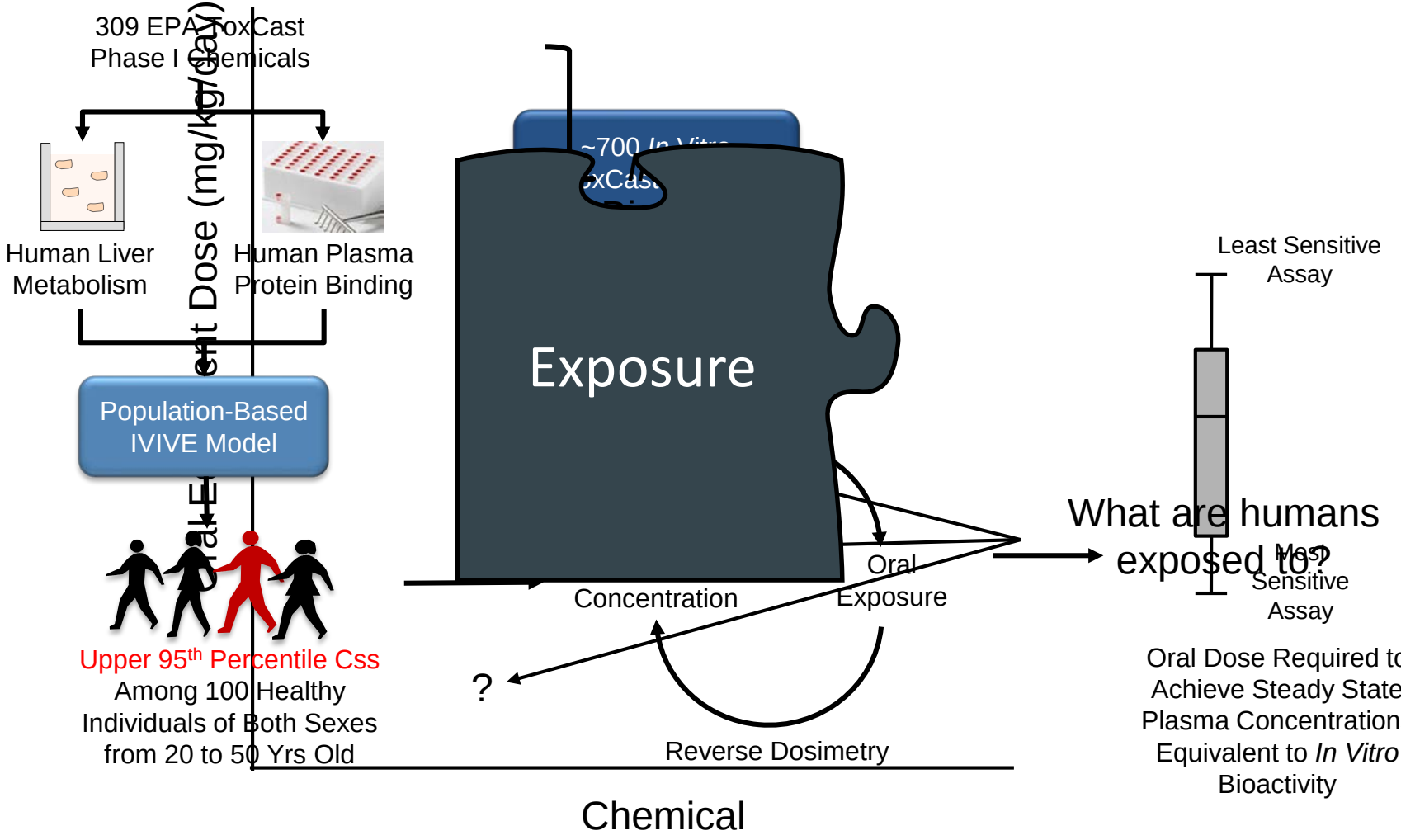
Rotroff *et al., Tox Sci.*, 2010
Wetmore *et al., Tox Sci.,* 2012

Integrating High-Throughput Pharmacokinetic Approach



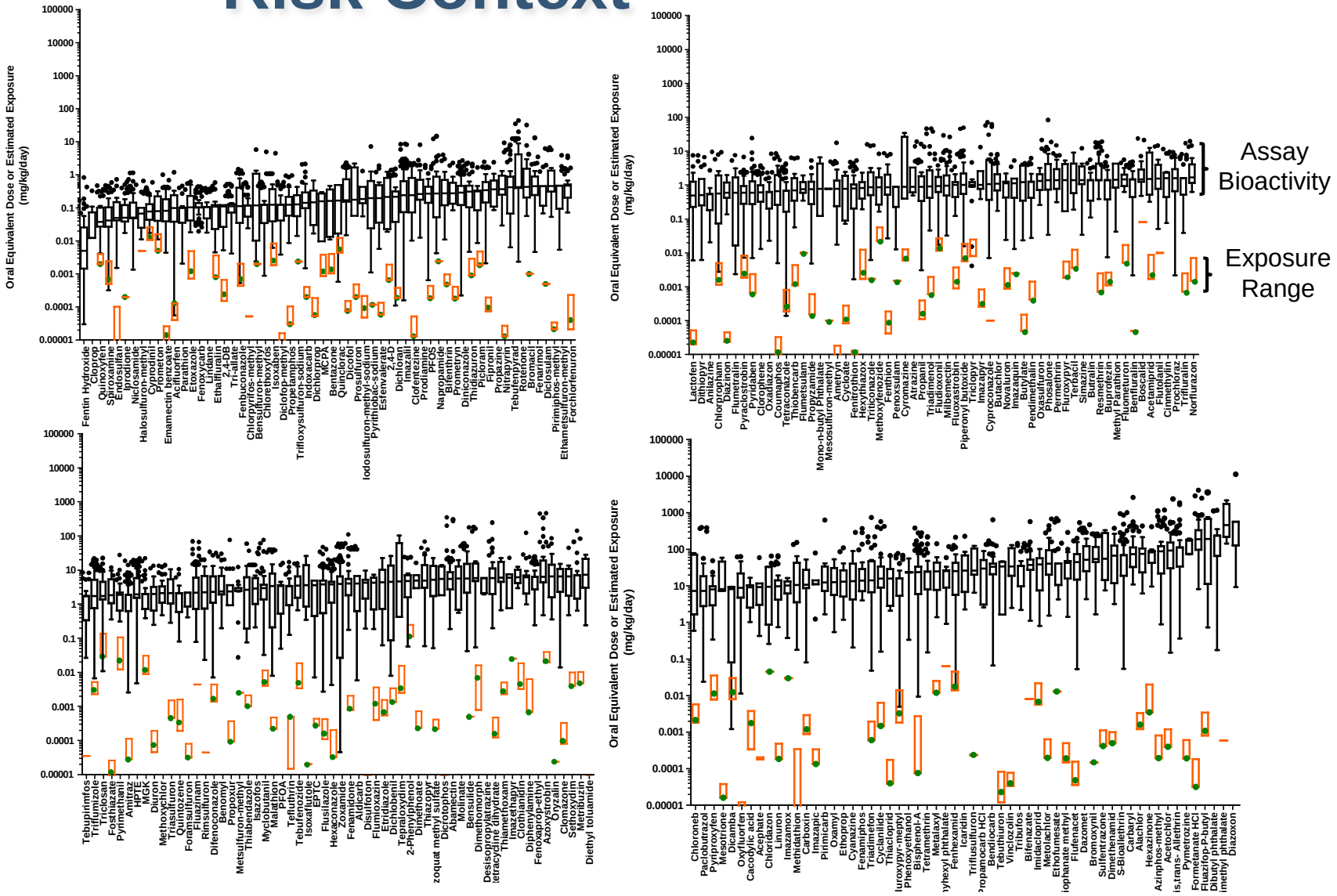
Rotroff *et al., Tox Sci.*, 2010
Wetmore *et al., Tox Sci.,* 2012

Comparing with Exposure for Risk Context



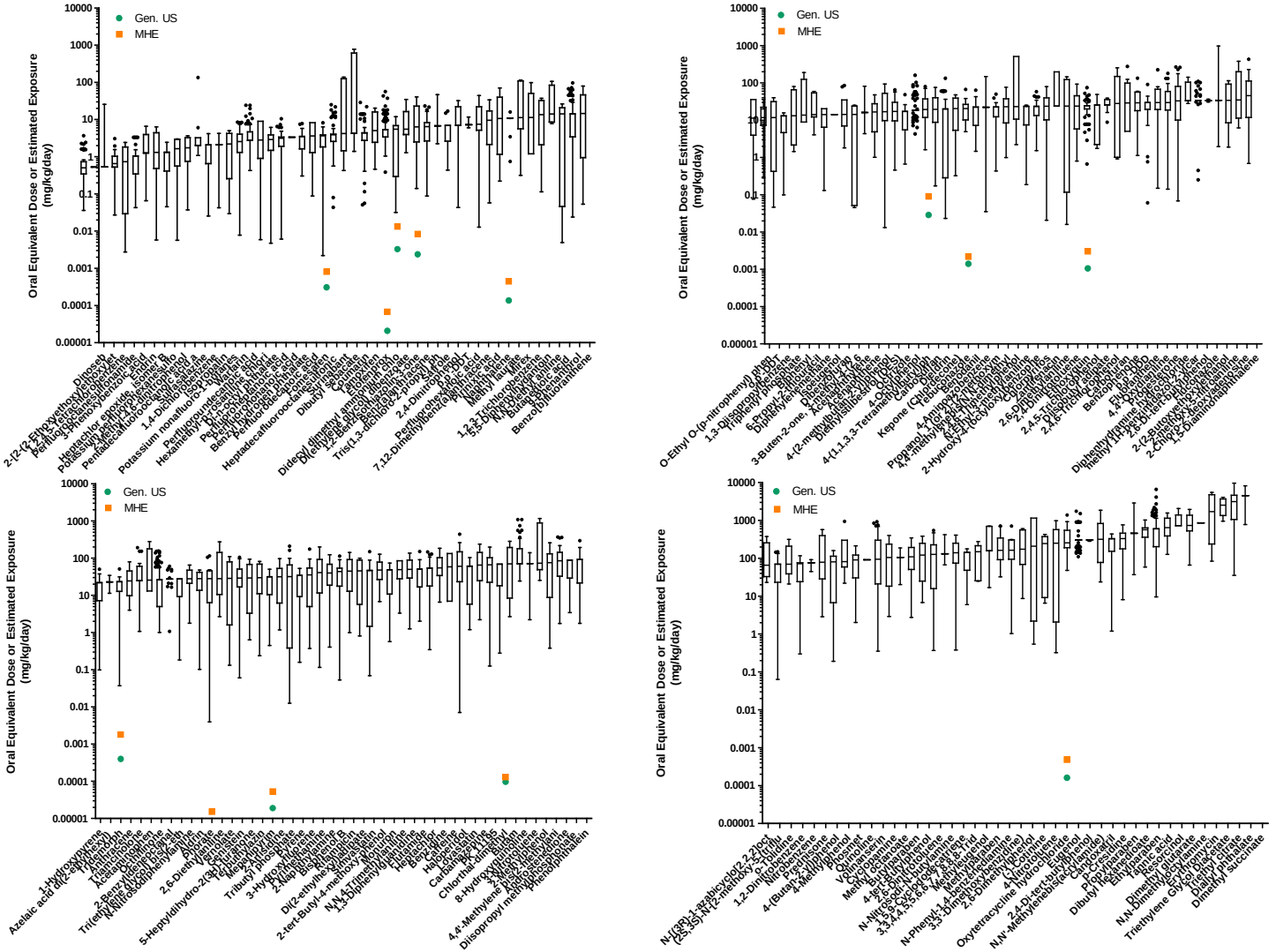
Rotroff *et al., Tox Sci.*, 2010
Wetmore *et al., Tox Sci.,* 2012

Comparing with Exposure for Risk Context



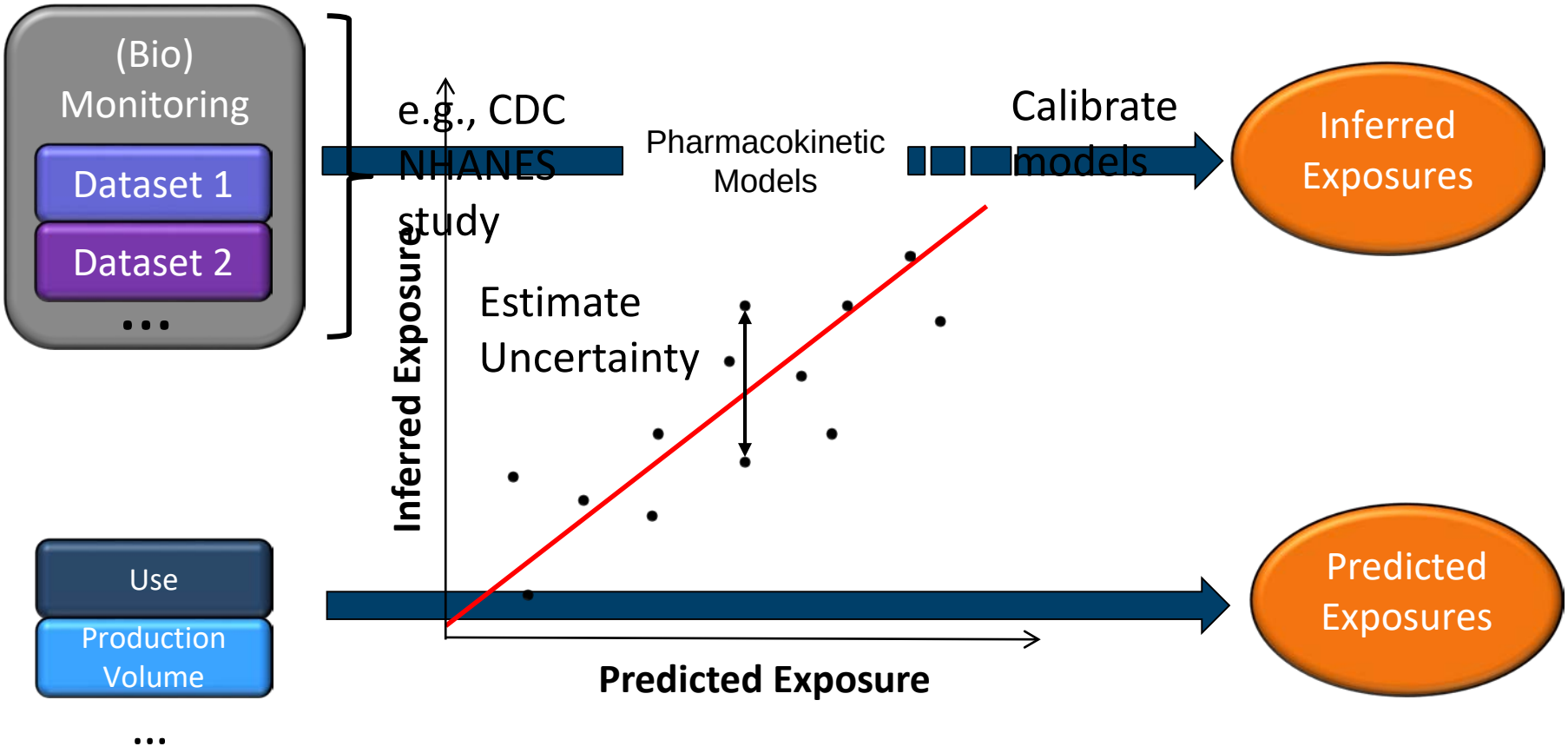
Wetmore *et al.*, *Tox Sci.*, 2012

Certainly the Road Less Traveled…



Bioactivity from Wetmore *et al.*, *Tox Sci.*, 2015

Developing High-Throughput Exposure Models



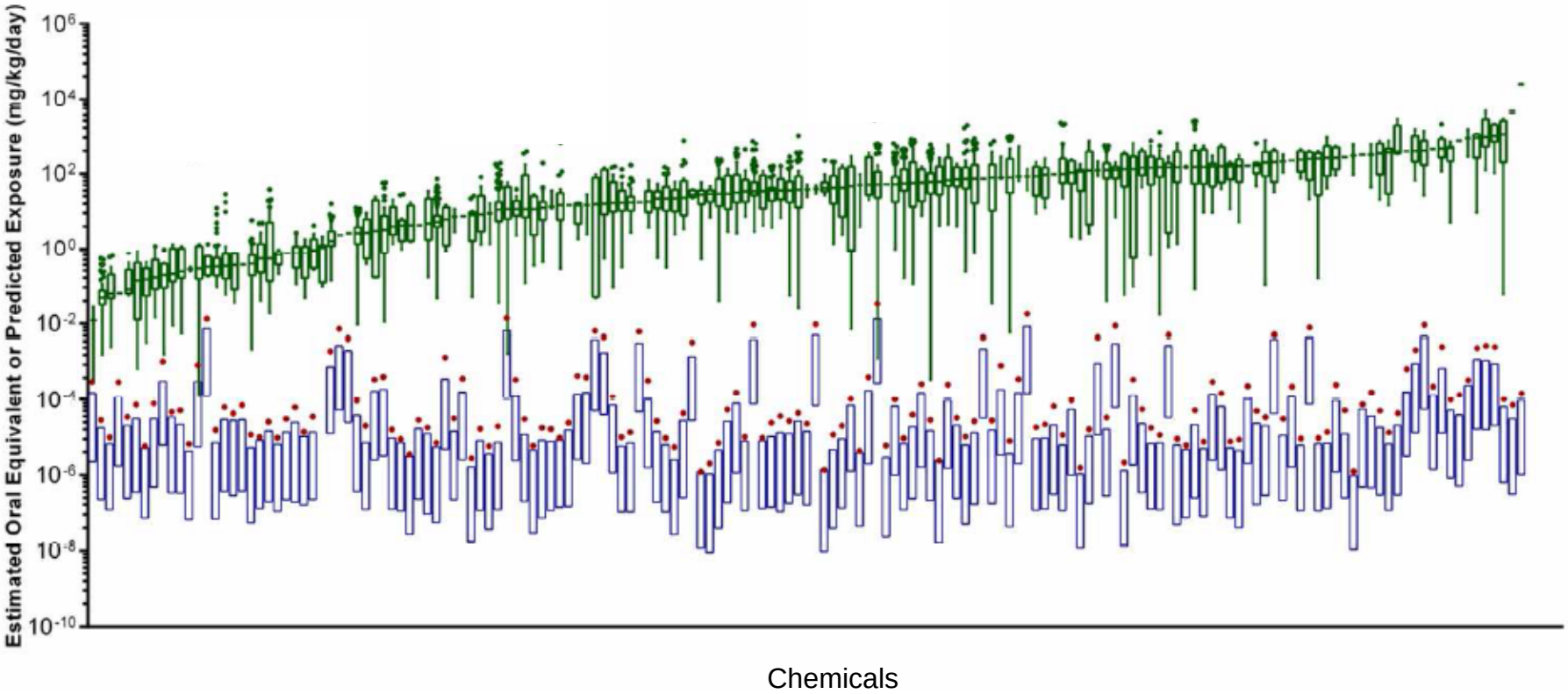
Wambaugh *et al.*, *Environ Sci Technol.*, 2014

High-Throughput Exposure Heuristics

Heuristic	Description
ACToR "Consumer use & Chemical/Industrial Process use"	Chemical substances in consumer products (*e.g.*, toys, personal care products, clothes, furniture, and home-care products) that are also used in industrial manufacturing processes. Does not include food or pharmaceuticals.
ACToR "Chemical/Industrial Process use with no Consumer use"	Chemical substances and products in industrial manufacturing processes that are not used in consumer products. Does not include food or pharmaceuticals
ACToR UseDB "Pesticide Inert use"	Secondary (*i.e.*, non-active) ingredients in a pesticide which serve a purpose other than repelling pests. Pesticide use of these ingredients is known due to more stringent reporting standards for pesticide ingredients, but many of these chemicals appear to be also used in consumer products
ACToR "Pesticide Active use"	Active ingredients in products designed to prevent, destroy, repel, or reduce pests (*e.g.*, insect repellants, weed killers, and disinfectants).
TSCA IUR 2006 Total Production Volume	Sum total (kg/year) of production of the chemical from all sites that produced the chemical in quantities of 25,000 pounds or more per year. If information for a chemical is not available, it is assumed to be produced at <25,000 pounds per year.

Wambaugh *et al.*, *Environ Sci Technol.*, 2014

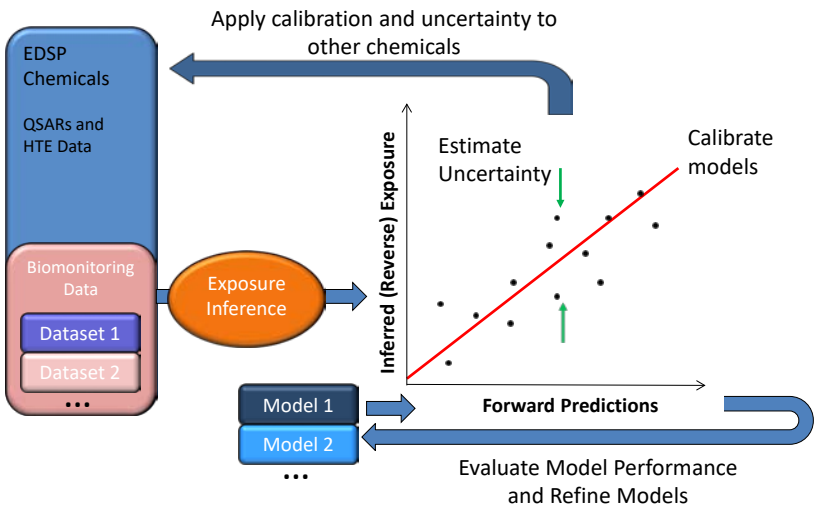
Comparing Bioactivity with Exposure Predictions for Risk Context



Wetmore *et al., Tox Sci.,* 2015

Work In Progress to Address Exposure Data Gaps

ExpoCast



- Limited domain of applicability of existing biomonitoring data
- Limited knowledge of qualitative and quantitative composition of chemical ingredients and emissivity of consumer products and home goods
- Limited experimental measurements of physical chemical properties





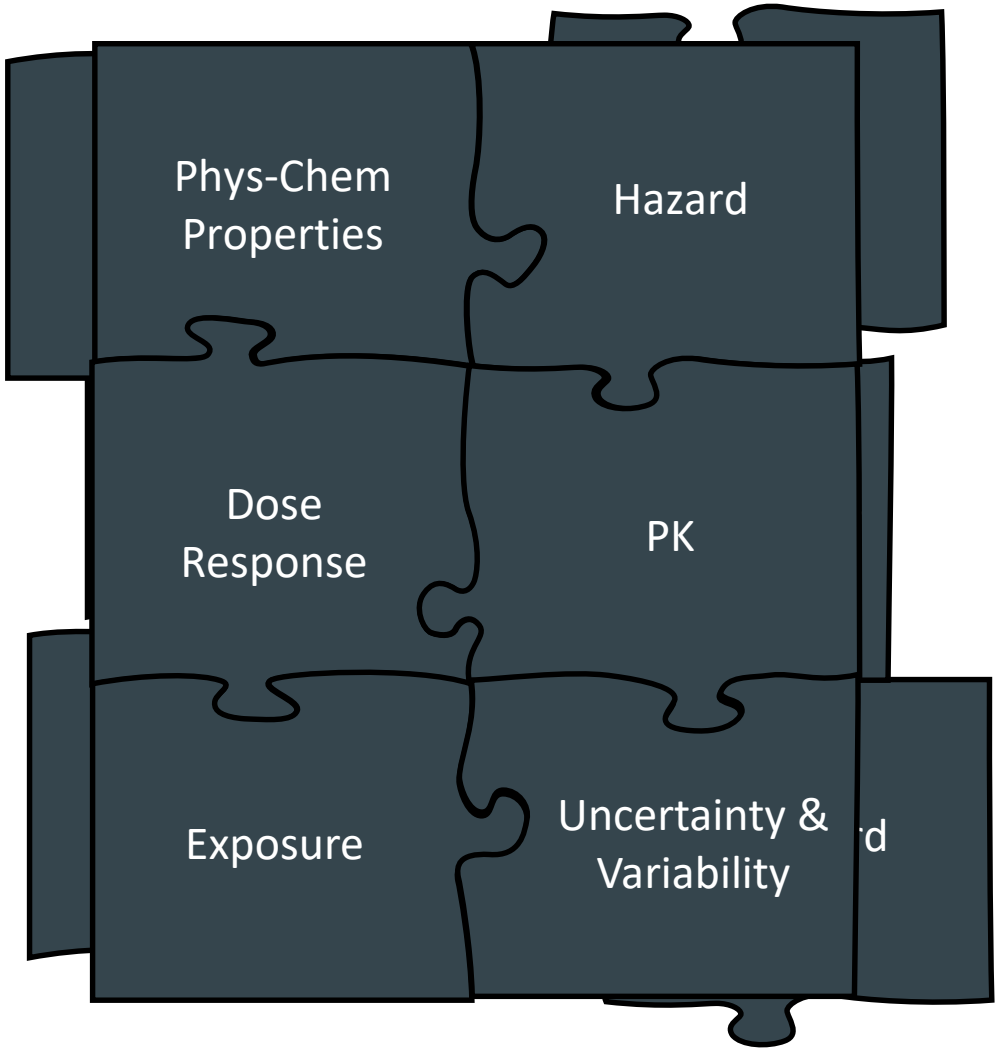
Non-Targeted Analysis of Consumer Products (Baby Products Pilot)

	Brand #1		Brand #2		Brand #3		Brand #4		Brand #5	
Category	**Number**	**Peak Area**	**Number**	**Peak Area**	**Number**	**Peak Area**	**Number**	**Peak Area**	**Number**	**Peak Area**
Reported peaks with reviewed library matches	98	430366619	106	123796416	114	1439691153	67	189430603	56	88784398
Reported unknowns (>500,000)	11	138064492	40	67604884	27	473487828	0	0	1	961833
Confirmed Hydrocarbons (n-alkanes)	7	4549995	5	85130984	20	14500452	20	28668397	21	30176895
Unconfirmed Hydrocarbons C10-C16	171	96701978	245	992103366	141	116097791	117	50621285	109	83257076
Unconfirmed Hydrocarbons C17-C32	--	--	--	--	181	107558101	261	155810354	243	142323751
Unresolved C17-C32	2457	116556867334	1934	67979297371	--	--	--	--	--	--
Excluded unknowns (<500,000)	37	5917014	120	21963344	66	10205424	52	7493668	48	6471715
Excluded non-specific (<500,000)	1	391747	0	0	11	425378	17	1038180	14	1169202
Excluded trace (<100,000) and similarity < 850	36	2259514	32	1795899	118	7218428	116	7015017	123	6698879
Excluded artifacts	1	393278	4	2172274	34	162097507	16	10001825	7	11311153
Total	2819	117235511971	2486	69273864537	712	2331282062	666	450079329	622	371154902

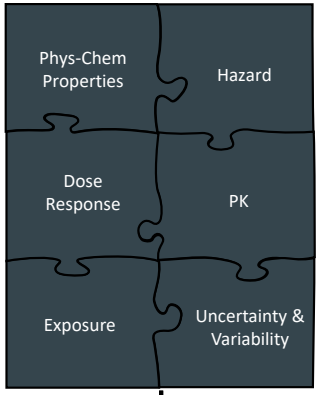
- Solvent extracted followed by GCXGC-TOFMS
- A total of 306 unique compounds identified across all toys, including 102 Tox21 chemicals.
- Bisphenol A found in “BPA free product”

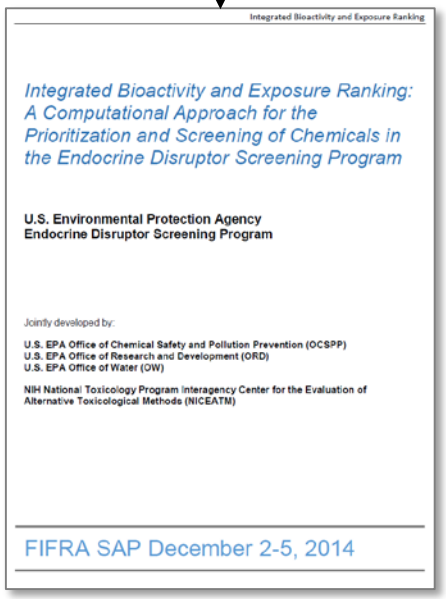
EPA United States Environmental Protection Agency

Pulling the Pieces Together for Regulatory Application



Application to Endocrine Disruptor Screening Program




Integrated Bioactivity and Exposure Ranking

Integrated Bioactivity and Exposure Ranking: A Computational Approach for the Prioritization and Screening of Chemicals in the Endocrine Disruptor Screening Program

U.S. Environmental Protection Agency
Endocrine Disruptor Screening Program

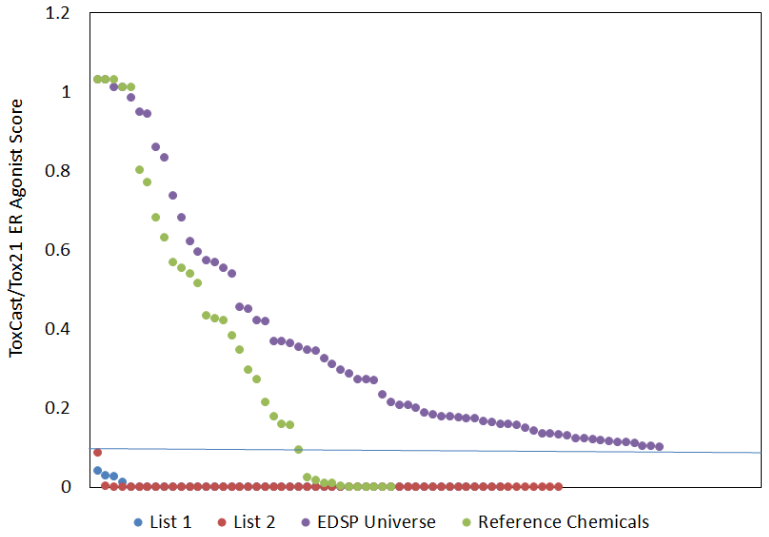
Jointly developed by:

U.S. EPA Office of Chemical Safety and Pollution Prevention (OCSPP)
U.S. EPA Office of Research and Development (ORD)
U.S. EPA Office of Water (OW)

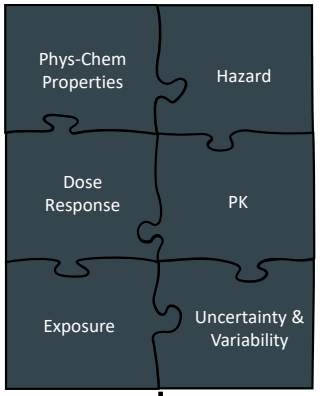
NIH National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)

FIFRA SAP December 2-5, 2014

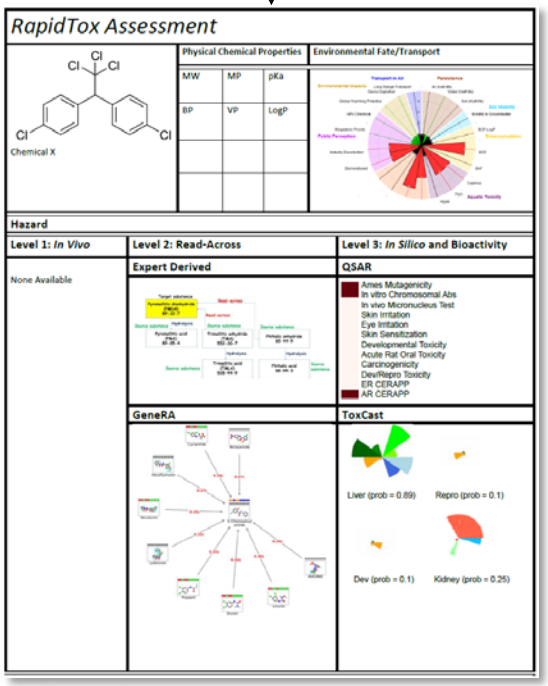
- ER, AR, and exposure modeling approaches underwent FIFRA SAP review
- FR Notice published for using ER assays and models to replace subset of Tier 1 assays
- Work continues on AR, steroidogenesis, and thyroid

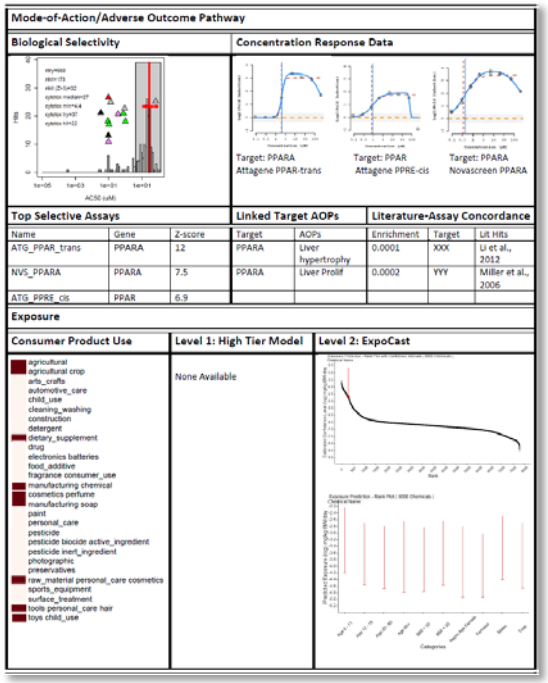


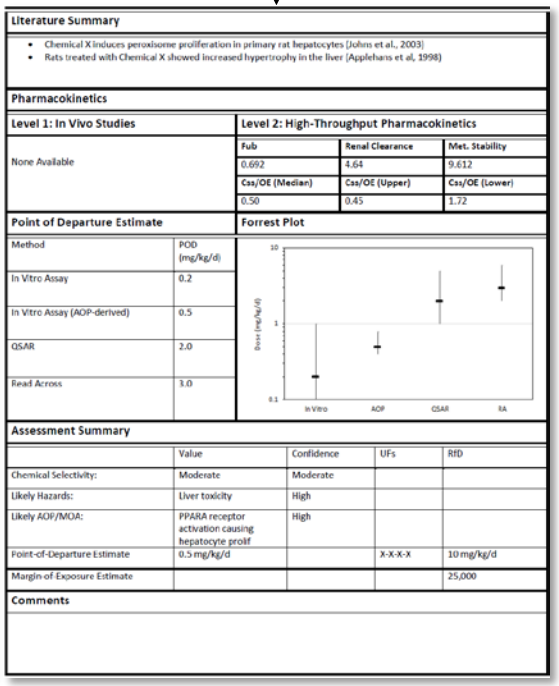
Application to Risk Assessment



- Semi-automated assessment workflow
- Dashboard interface
- Two case studies





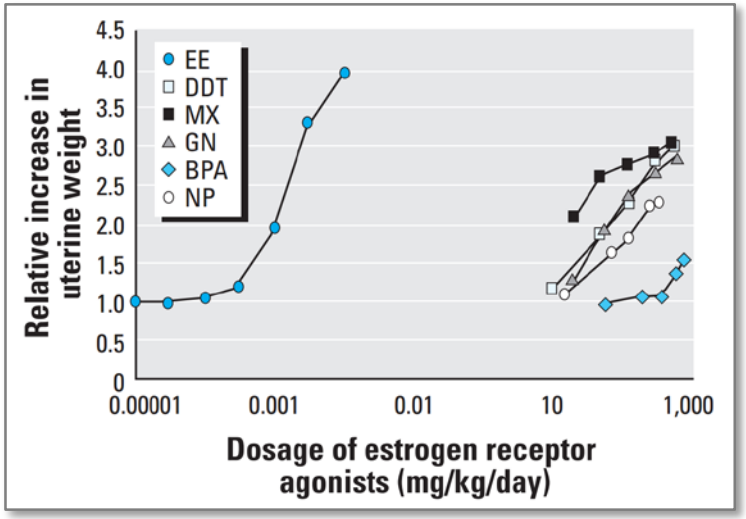


Challenges

- Dealing with the “V” word
  - Defining a fit-for-purpose framework(s) that is time and resource efficient (e.g., Judson *et al., ALTEX* 30(1):51-6, 2013)
  - Role of *in vivo* rodent studies
  - Accepting and incorporating the inherent uncertainty of the *in vivo* studies

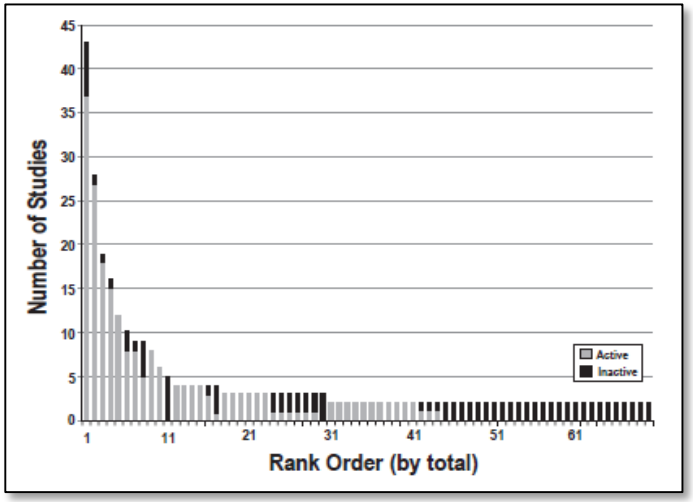
Dealing with the “V” Word

Original OECD TG 440 Validation



Owens and Koëter, *Environ Health Perspect,* 2003

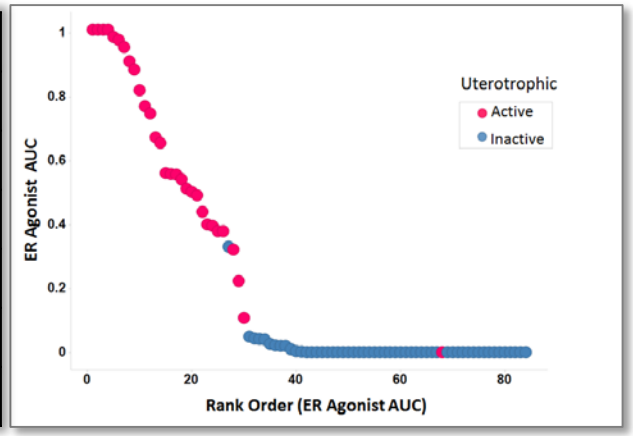
Concordance of *In Vivo* Uterotrophic Studies



Kleinstreuer et al., EHP In Press

Predictive Performance of *In Vitro* Assays for *In Vivo* Uterotrophic Studies

True Positive	29 (29)
True Negative	8 (8)
False Positive	5 (1)
False Negative	1 (1)
Accuracy	**0.86 (0.95)**
Sensitivity	**0.97 (0.97)**
Specificity	**0.67 (0.89)**



Browne *et al., ES&T.* 2015

Challenges

- Dealing with the "V" word
  - Defining a fit-for-purpose framework(s) that is time and resource efficient (e.g., Judson *et al., ALTEX* 30(1):51-6, 2013)
  - Role of *in vivo* rodent studies
  - Accepting and incorporating the inherent uncertainty of the *in vivo* studies
- Moving from an apical to a molecular paradigm and defining adversity
- Legal defensibility of new methods and assessment products
- Predicting human safety vs. toxicity
- Systematically integrating multiple data streams from the new approaches in a risk-based, weight of evidence assessment
- Ensuring a comprehensive screening and testing paradigm
- Quantifying and incorporating uncertainty and variability
- Application to cumulative risk/mixtures

Acknowledgements

Tox21 Colleagues:
- NTP Crew
- FDA Collaborators
- NCATS Collaborators

ORD Colleagues:
- NERL
- NHEERL
- NCEA

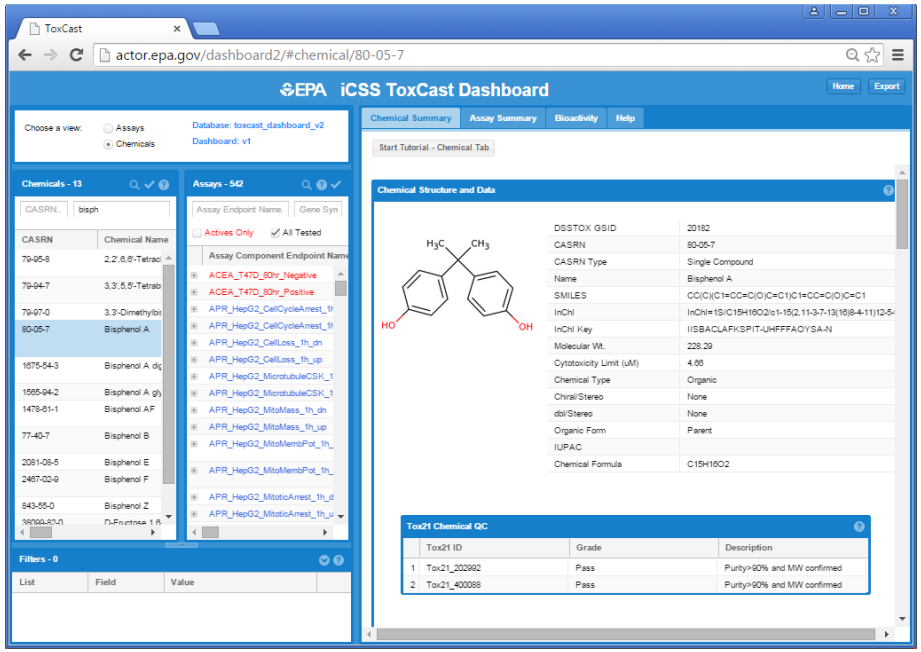
Alice Yau - SWRI



EPA's National Center for Computational Toxicology

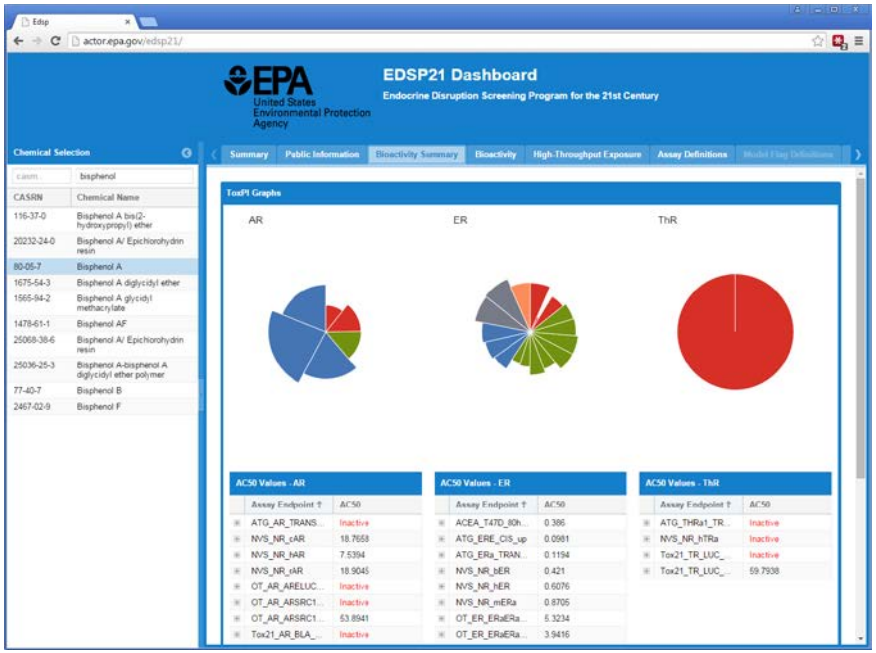
New Dashboards Provide Improved Access to the Data

iCSS Dashboard



http://actor.epa.gov/dashboard2/

EDSP21 Dashboard



http://actor.epa.gov/edsp21/

EPA United States Environmental Protection Agency

Extra Slides

Eigenvalues from Principle Component Analysis

United States Environmental Protection Agency

Cell Lines

Genes

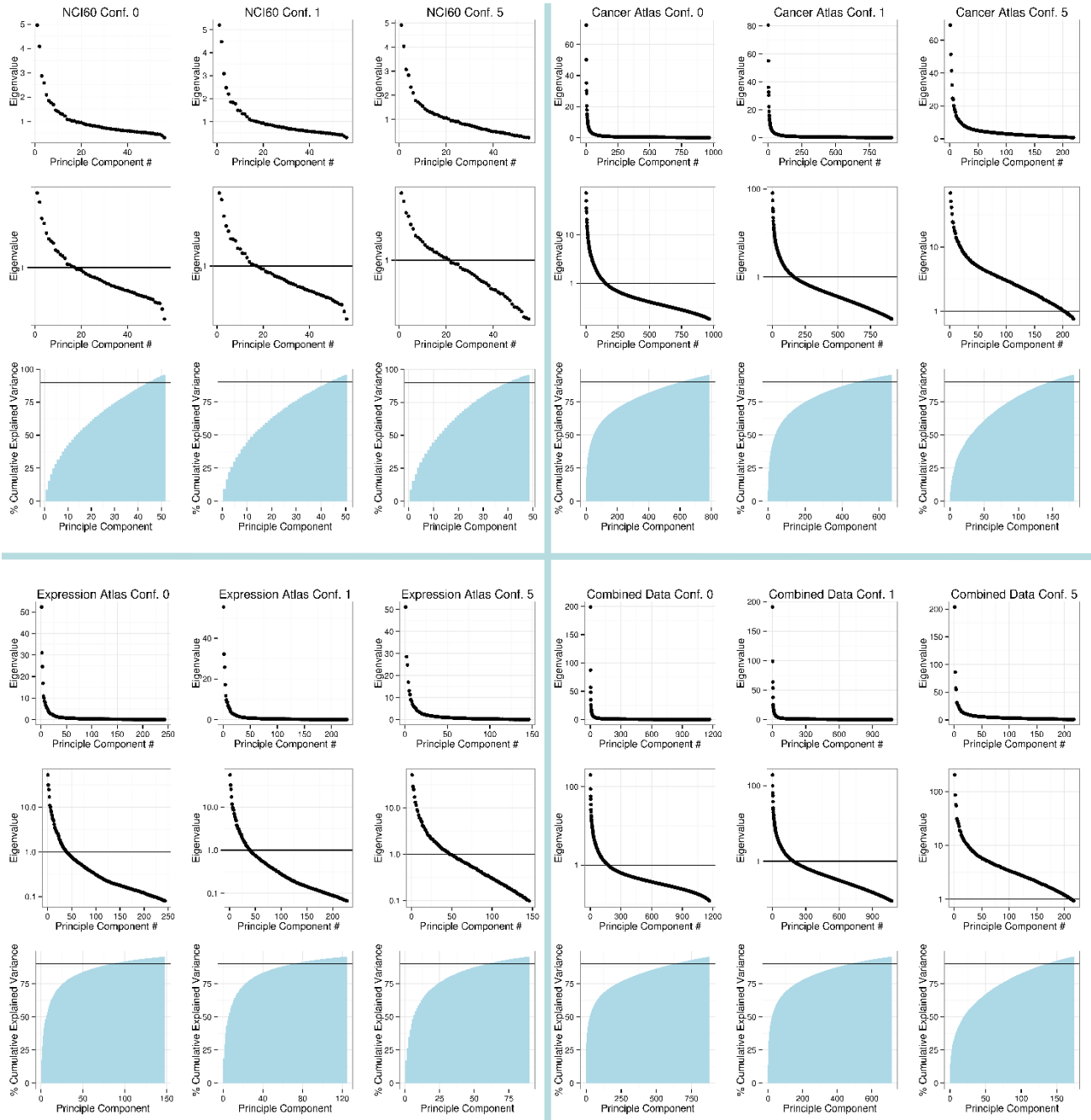
Data Matrix

OR

Cell Lines

Genes

Data Matrix



Dimensionality and Fraction of Variance Explained

	All Genes (20,960)	Druggable 1 (3845)	Druggable 5 (237)
NCI60 (59)	17 55%	17 57%	21 66%
Cancer Atlas (1036)	154 66%	165 71%	202 98%
Primary Cells (303)	39 83%	40 83%	46 86%
Combined (1398)	172 72%	187 77%	211 99%

cell lines

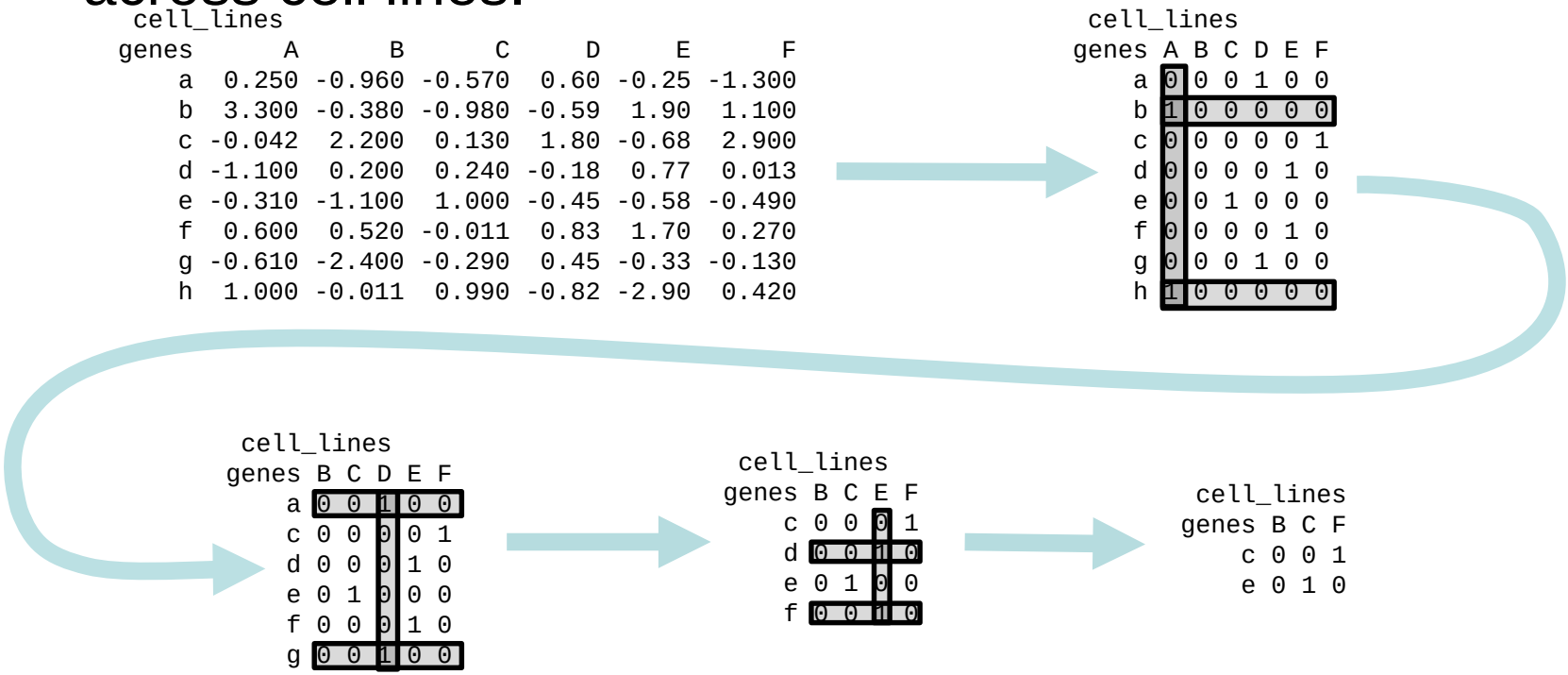
genes

eigenvalues ≥ 1

- Dimensionality is relatively constant within cell lines
- Dimensionality *increases* (slightly) as # genes decrease

As Many Genes 'On' as Possible

Toy Example: Call a gene 'on' if its expression is in the upper 20th %-ile of its distribution across cell lines.



Result:
Cell Lines A, D, and E cover 6 of 8 genes.

Fraction of Genes ‘On’ in At Least One Cell Line

