EPA United States Environmental Protection Agency

Molecular Point-of-Departure Values from High-Throughput Profiling Assays: Potential Applications for Chemical Safety Assessment

Joshua A. Harrill

USEPA Center for Computational Toxicology and Exposure (CCTE)

EPA United States Environmental Protection Agency

Disclaimer

The views expressed in this presentation are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency, nor does mention of trade names or products represent endorsement for use.

EPA United States Environmental Protection Agency

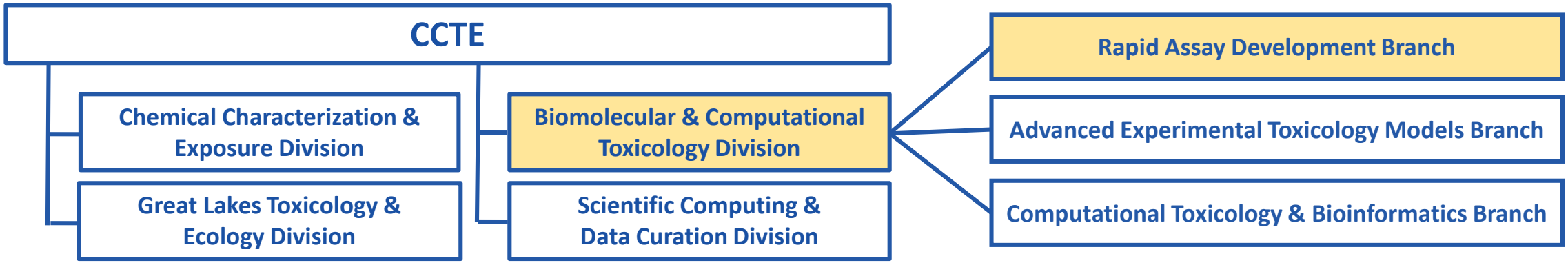
Outline

- **Background**
  - Who is CCTE?
  - What Does CCTE Do?
  - Blueprint for Computational Toxicology at USEPA
- **High Throughput Transcriptomics (HTTr)**
- **High Throughput Phenotypic Profiling (HTPP)**
- **Potential Applications for HTTr- and HTPP-derived Molecular PODs**

Who is CCTE?

Center for Computational Toxicology and Exposure (CCTE)

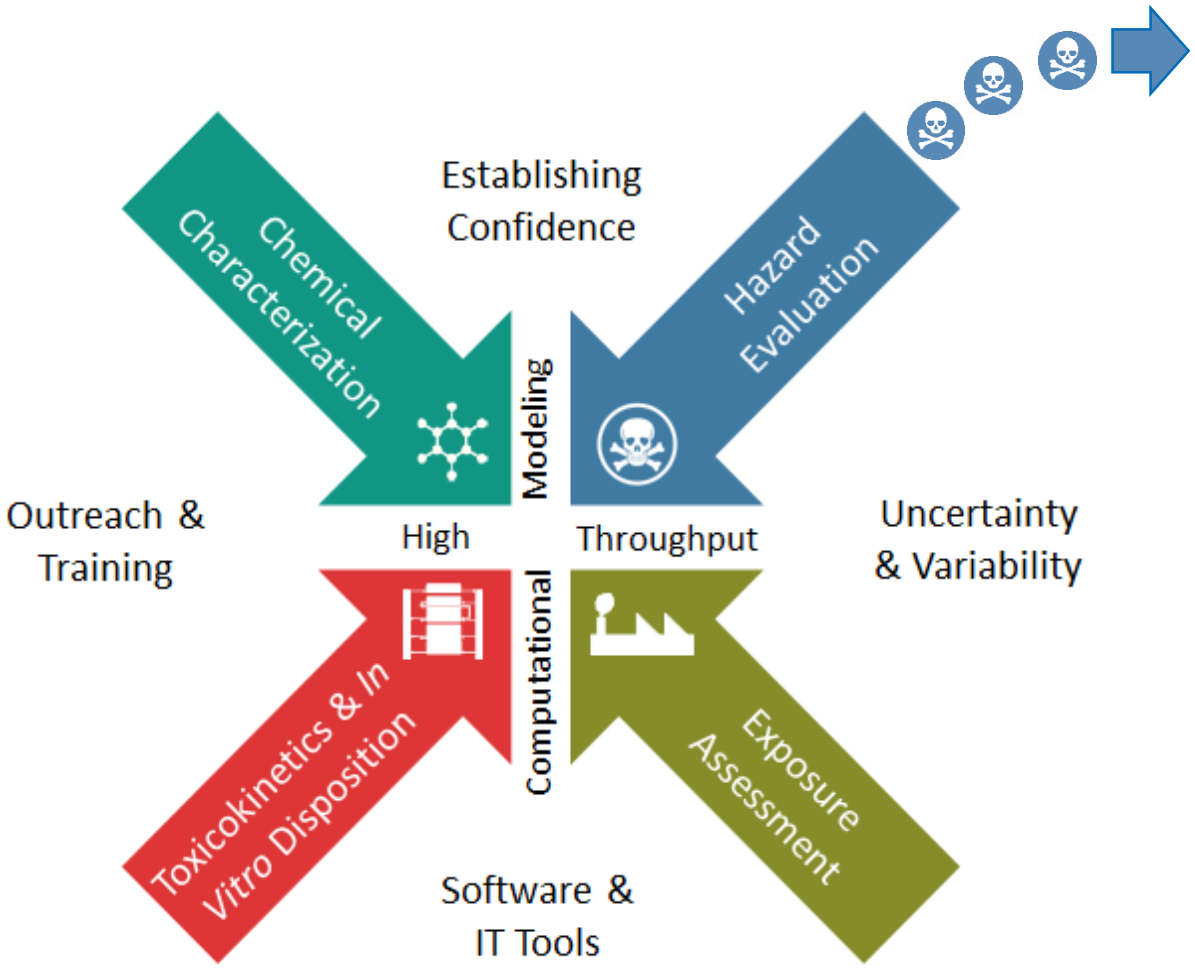
A research organization at US EPA Office of Research and Development tasked with **developing** and **applying** cutting edge innovations in methods to rapidly evaluate chemical toxicity, transport and exposure to people and environments.



Rapid Assay Development Branch (RADB)

Develops the next generation of **high-throughput toxicity assays** to comprehensively cover the potential **molecular and phenotypic responses** resulting from chemical exposure and **fill gaps** in biological pathways and processes not addressed using existing assays.

Computational Toxicology Research Areas



The NexGen Blueprint of CompTox at USEPA, Tox. Sci. 2019; 169(2):317-322

- **ToxCast:** Used targeted high-throughput screening (HTS) assays to expose living cells or isolated proteins to chemicals and assess bioactivity and potential toxic effects.

	# of assays	# of chemicals	Types of chemicals
Phase 1 *(2007 – 2009)*	500	300	Mostly pesticides
Phase 2 *(2009 – 2013)*	700	2,000	Industrial, consumer product, food use, "green"

- Mostly targeted assays (*chemical X → target Y*)
- Incomplete coverage of biological space.

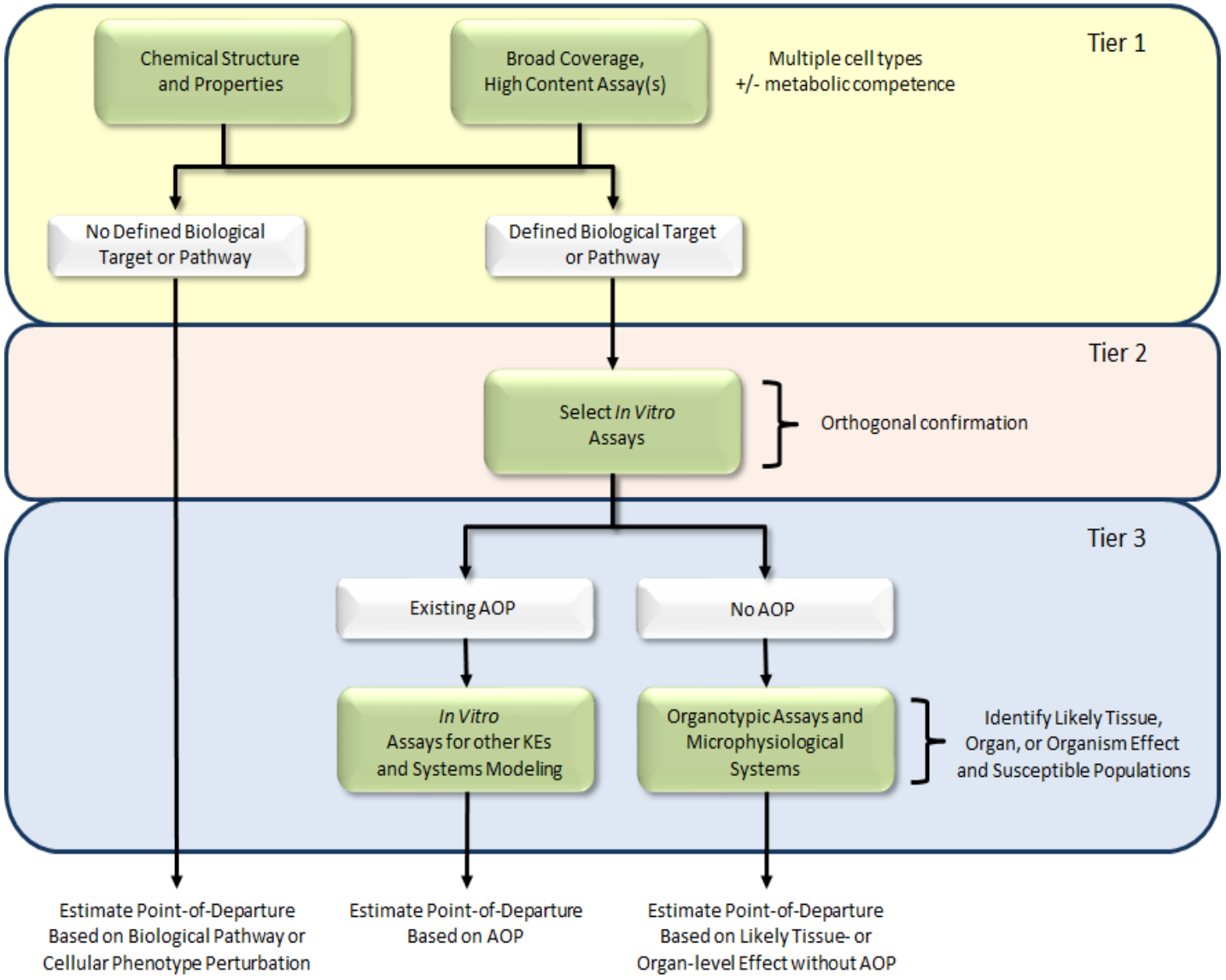
- **New Strategy for Hazard Evaluation:** Improve efficiency and increase biological coverage by using broad-based (i.e. non-targeted) **profiling assays** that cast the broadest net possible for capturing the potential molecular and phenotypic responses of human cells in response to chemical exposures.

Tiered Hazard Evaluation Approach (1)

- **New Approach Methodologies (NAMs)** are any technology, methodology, approach or combination thereof that can be used to provide information on chemical hazard and risk that avoids the use of intact animals.

- US EPA CompTox Blueprint advocates the use of **high throughput profiling (HTP) assays** as the first tier in a NAMs-based hazard evaluation approach.

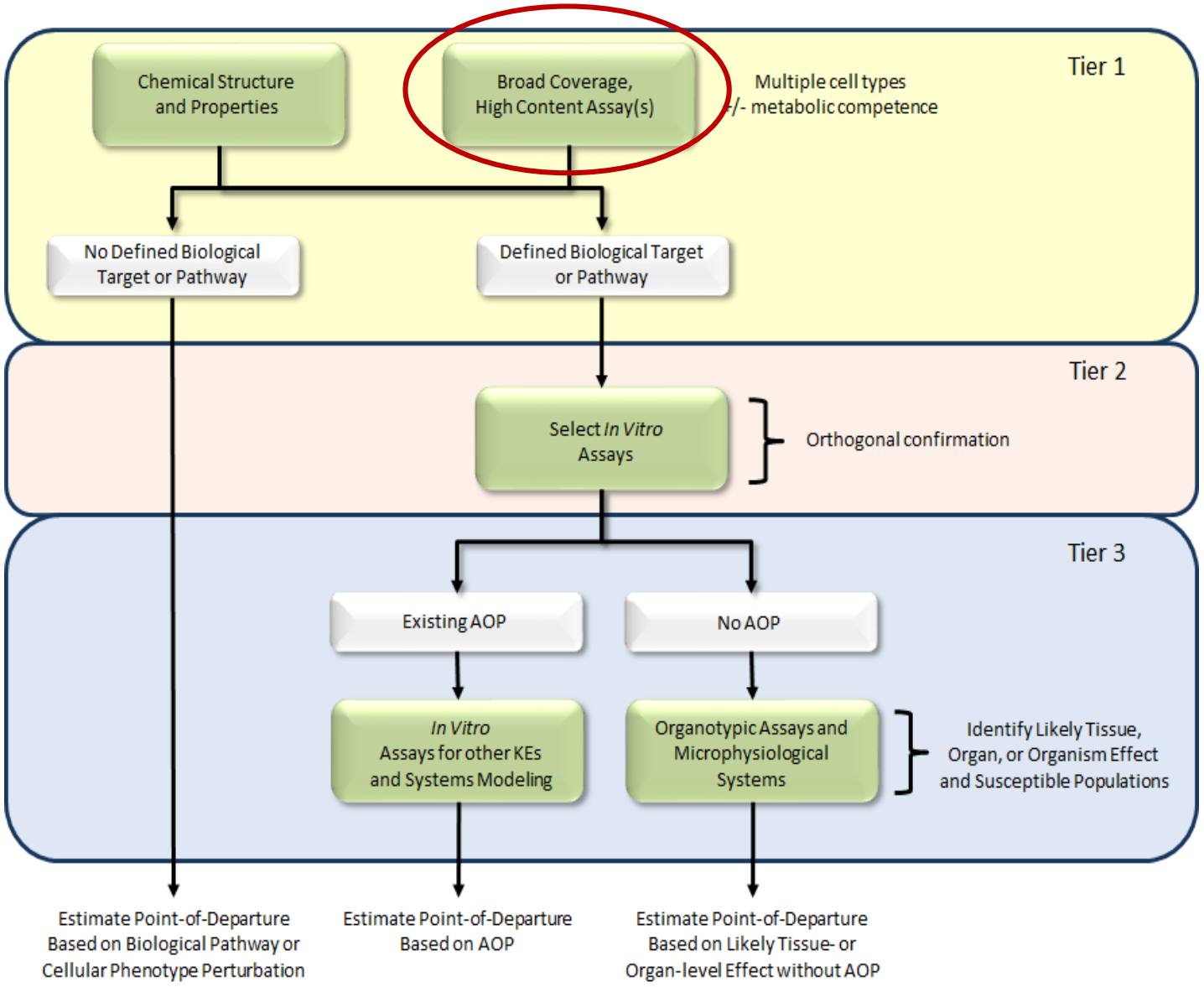
- **HTP assay criteria:**
  1. Yield bioactivity profiles that can be used for **potency estimation**, **mechanistic prediction** and evaluation of **chemical similarity**.
  2. Compatible with multiple human-derived culture models.
  3. Concentration-response screening mode.
  4. Cost-effective.



The NexGen Blueprint of CompTox as USEPA Tox. Sci. 2019; 169(2):317-322

Tiered Hazard Evaluation Approach (2)

- To date, identified and implemented two assays that meet this criteria.
- Increasing efficiency and declining cost of generating whole transcriptome profiles has made **high-throughput transcriptomics (HTTr)** a practical option for *in vitro* chemical screening.
  - **Whole Transcriptome TempO-Seq**
- Imaging-based **high-throughput phenotypic profiling (HTPP)** provides a cost-effective means for characterizing the effects of chemicals on apical cellular morphology (i.e. cellular pathology).
  - **Cell Painting**
- Both methods are **complementary** to each other and can be used in many different human-derived cell types.



The NexGen Blueprint of CompTox as USEPA Tox. Sci. 2019; 169(2):317-322

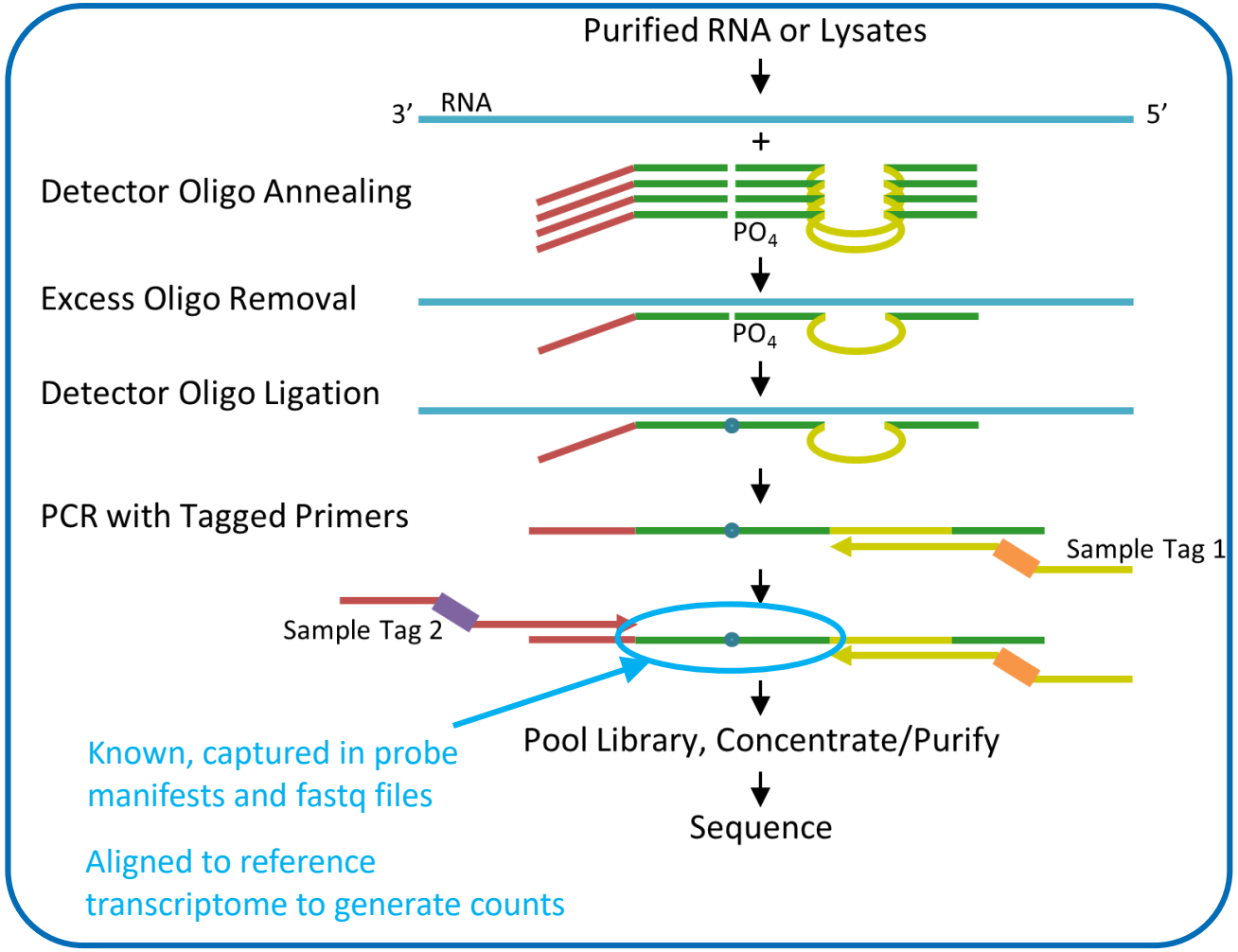
EPA
United States
Environmental Protection
Agency

High-Throughput Transcriptomics

Templated Oligo with Sequencing Readout (TempO-Seq)

- The **TempO-Seq** human whole transcriptome assay measures the expression of greater than 20,000 transcripts.
- Requires only picogram amounts of total RNA per sample.
- Compatible with purified RNA samples or **cell lysates**.
- Lysates are barcoded according to sample identity and combined in a single library for sequencing using industry standard instrumentation.
- Scalable, targeted assay:
  - 1) specifically measures transcripts of interest
  - 2) ~50-bp reads for all genes
  - 3) requires less flow cell capacity than RNA-Seq

TempO-Seq Assay Illustration



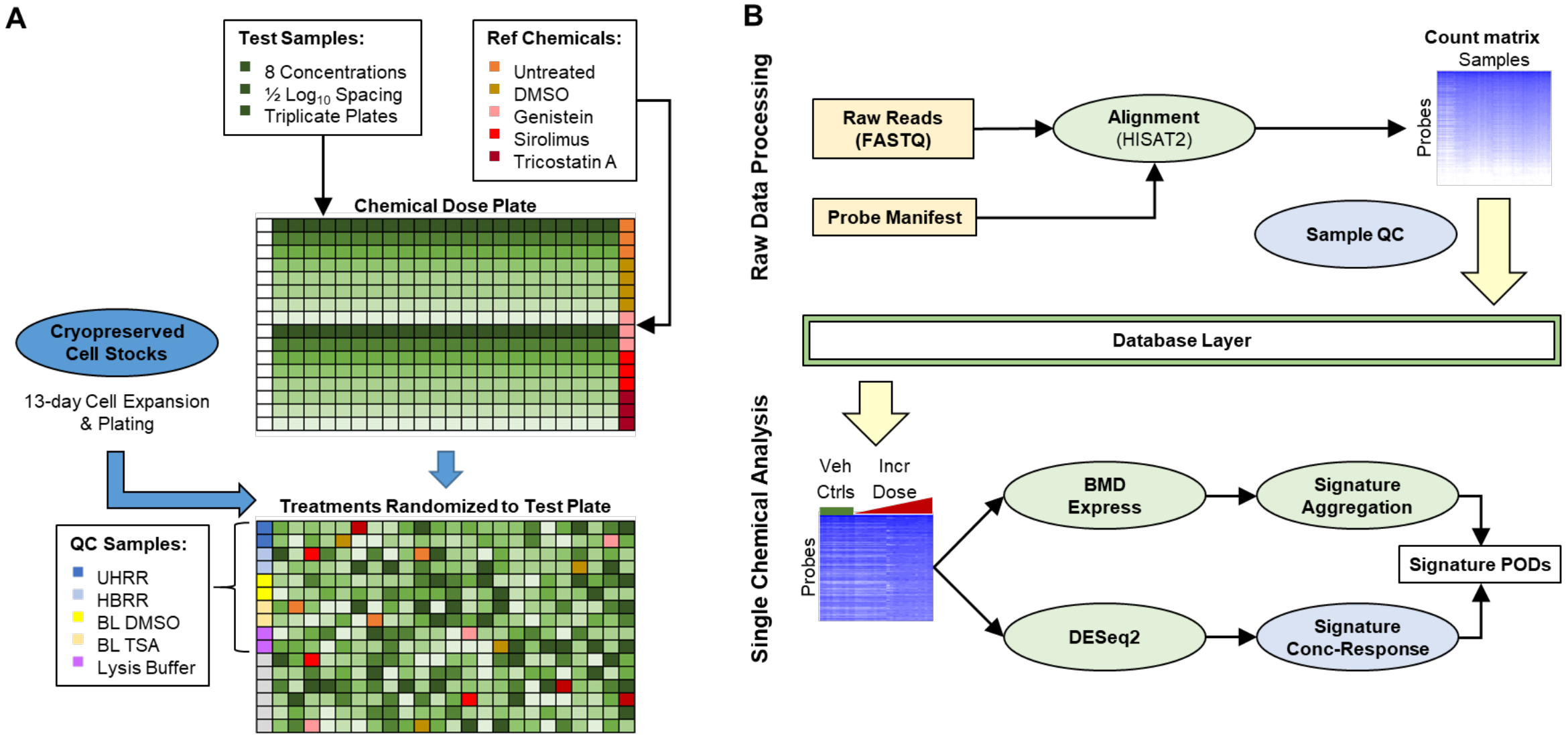
Yeakley, et al. PLoS ONE 2017

EPA United States Environmental Protection Agency

MCF-7 Pilot Experimental Design

Parameter	Multiplier	Notes
Cell Type(s)	1	MCF7
Culture Condition	1	DMEM + 10% HI-FBS
Chemicals	44	ToxCast chemicals with mechanistic variety and some redundancy.
Time Points:	1	6 hours
Assay Formats:	2	Cell Painting Cell Viability
Concentrations:	8	3.5 $\log_{10}$ units; semi $\log_{10}$ spacing
Biological Replicates:		Independent cultures

HTTr Experimental Design and Bioinformatics Workflow

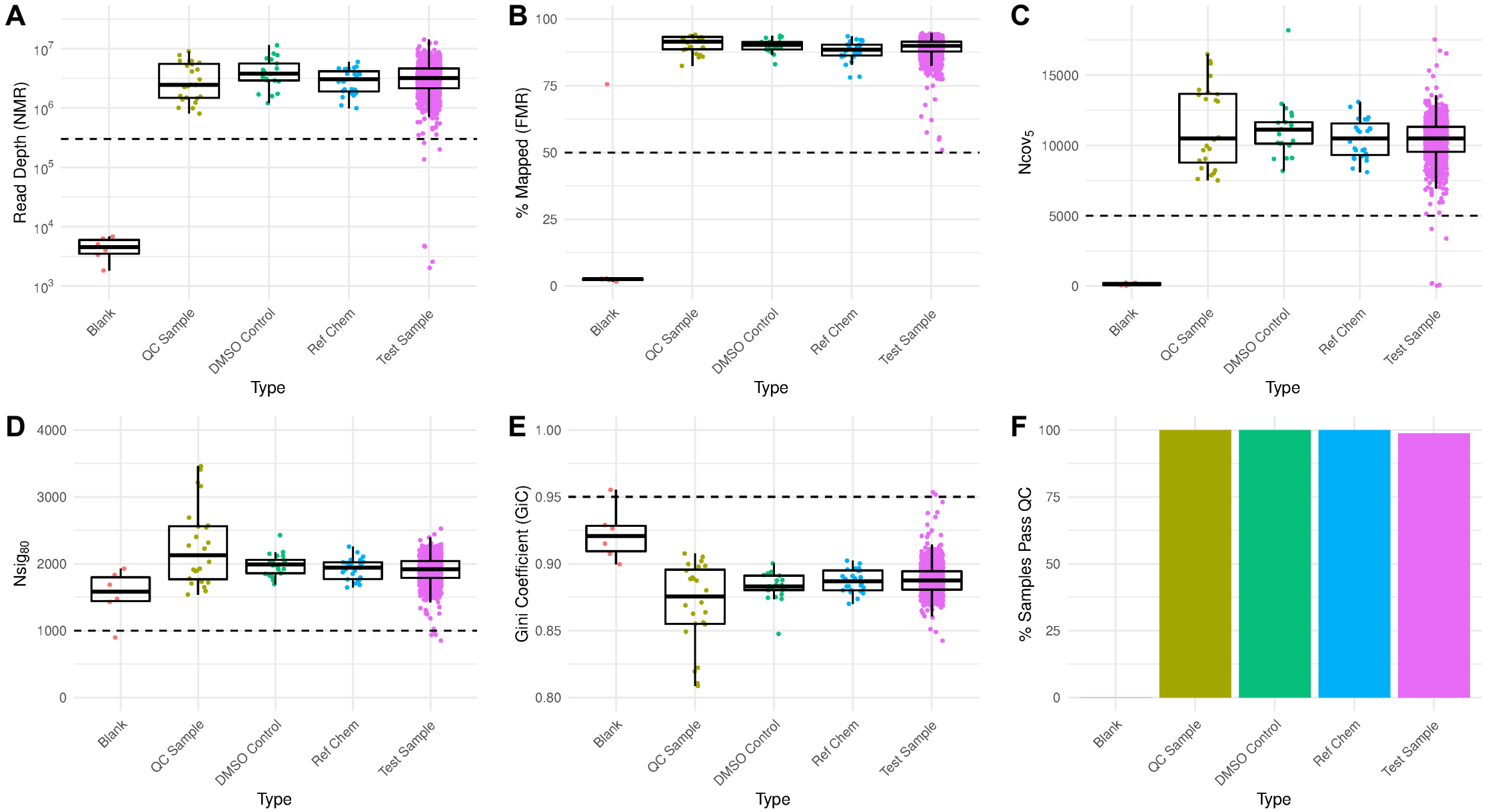


HTTr Quality Control Criteria

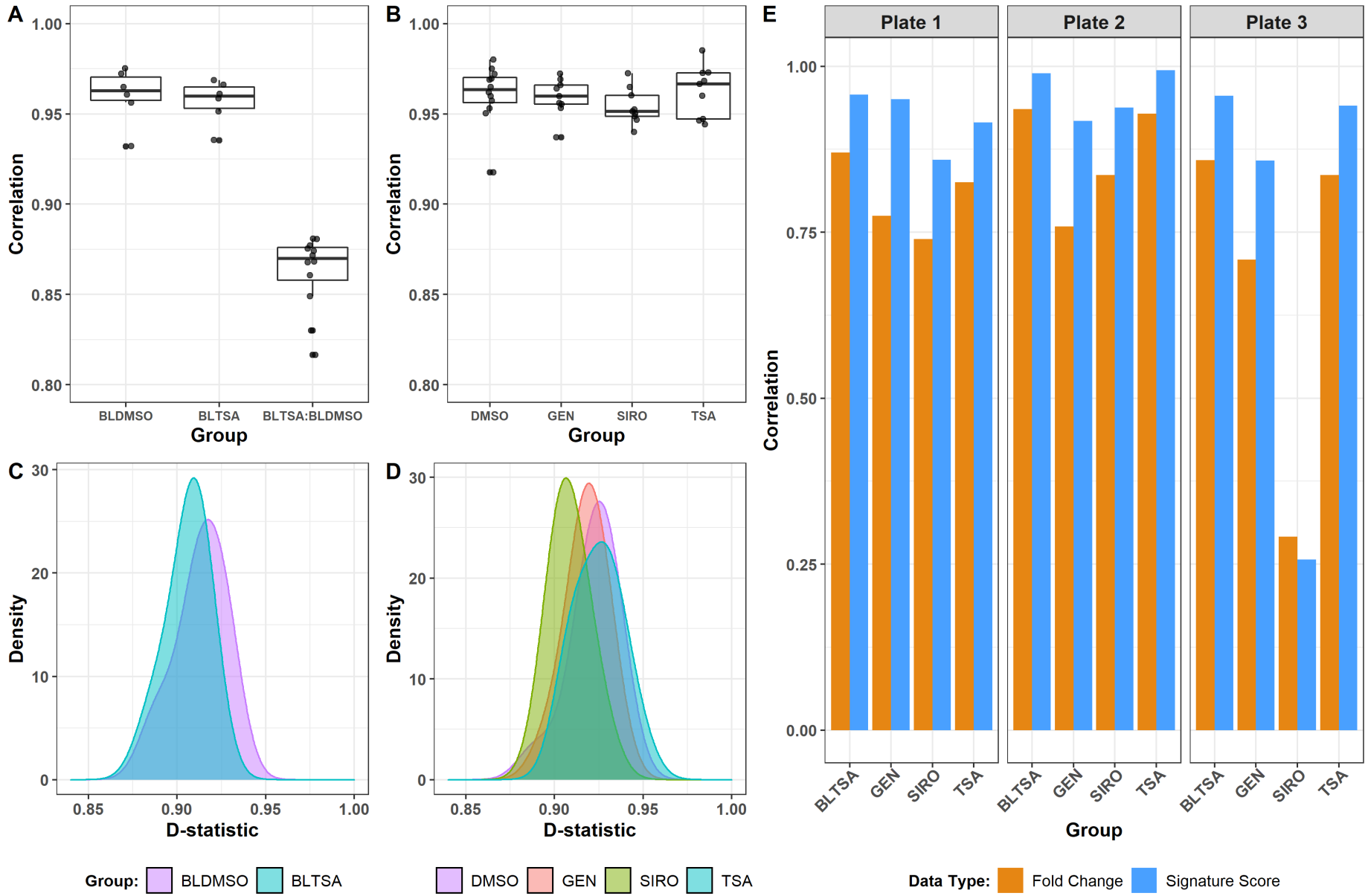
Abbreviation	Description	Threshold	Additional Information
FrVC	Fraction of viable cells (PI-negative or Casp3/7-negative)	Reject < 50%	Highly cytotoxic conditions no longer represent molecular initiating event
NMR	Number of mapped reads, defined as sum of total read counts summed over all detected probes	Reject < 300,000	Threshold =10% of target depth
FMR	Fraction of uniquely mapped reads	Reject < 50%	Majority of reads must align to a single probe sequence
$Ncov_5$	The number of probes with at least 5 uniquely mapped reads	Reject < 5,000	Based on Tukey's Outer Fence (3*IQR) of all viable samples cultured on each plate (test samples, vehicle controls, and reference chemical treatments)
$Nsig_{80}$	The number of probes capturing the top 80% of signal in a sample	Reject < 1,000	
GiC	Gini coefficient computed for each sample based on the distribution of raw counts for all probes including those with 0 aligned reads	Reject > 0.95	

EPA United States Environmental Protection Agency

HTTr Sample Quality Assessment (1)

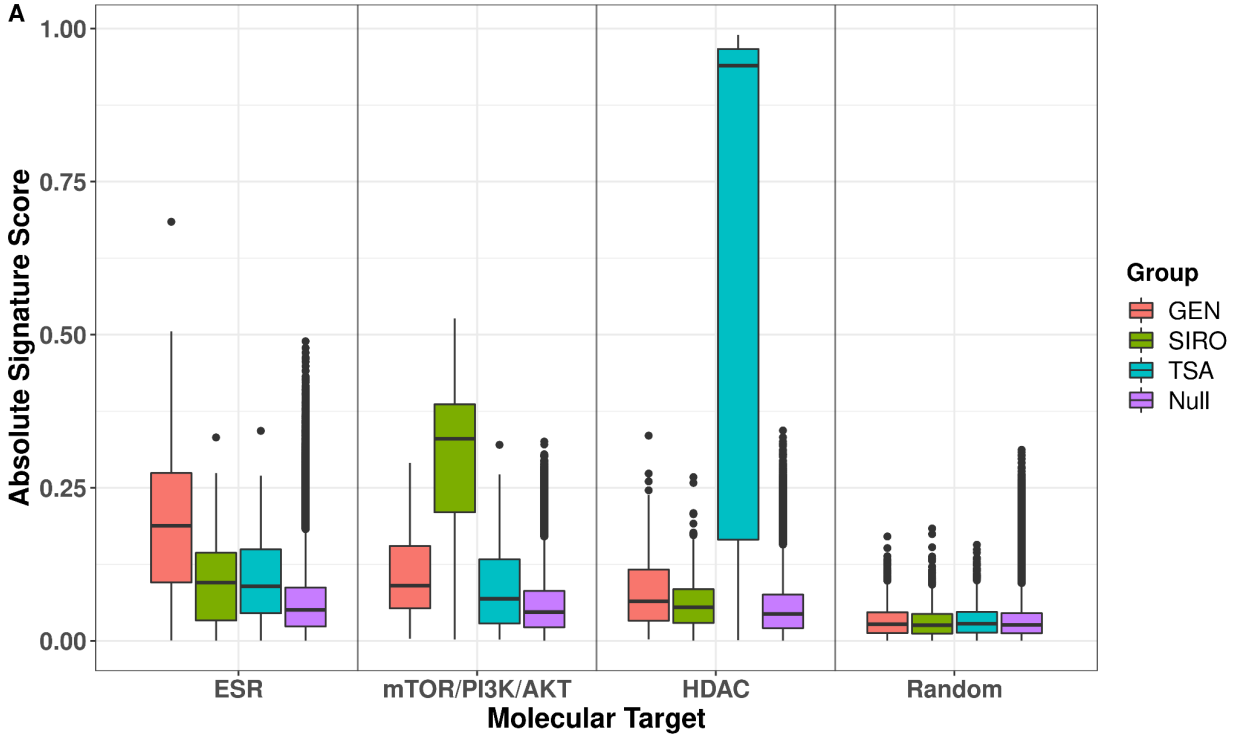


HTTr Sample Quality Assessment (2)



HTTr Sample Performance Assessment

A

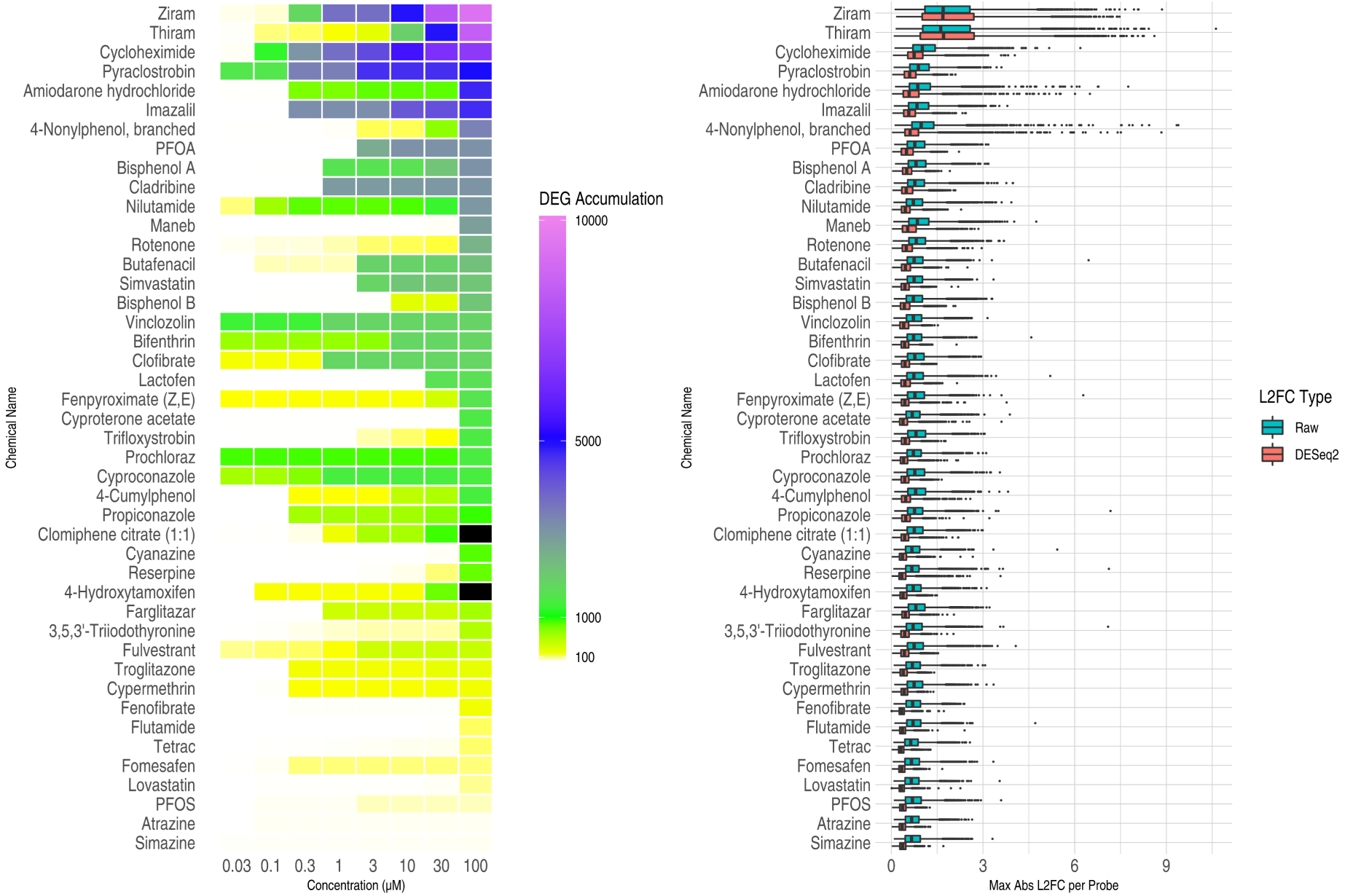


B

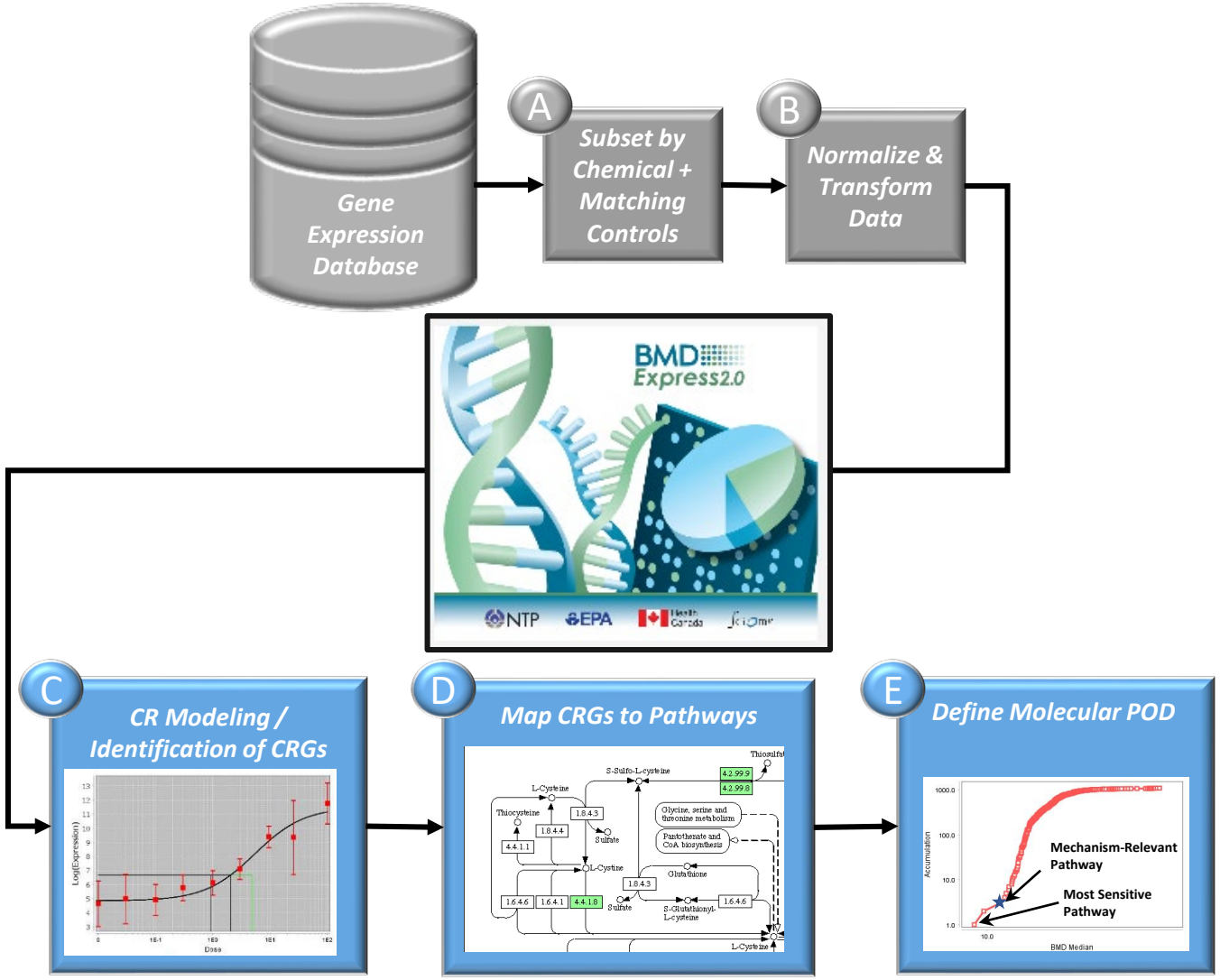
	Signature	Absolute Score	Direction
Genistein	RYAN_ESTROGEN_RECEPTOR_ALPHA	0.6844617	+
	CMAP equilin 1.5e-05 100 5542 100	0.5054424	+
	CMAP fulvestrant 1e-08 100 1417 100	0.4857079	-
	CMAP fulvestrant 1e-06 100 6811 100	0.4519683	-
	BHAT_ESR1_TARGETS_NOT_VIA_AKT1	0.4270213	+
Sirolimus	CMAP sirolimus 1e-07 100 5437 100	0.5265263	+
	CMAP sirolimus 1e-07 100 6409 100	0.5166319	+
	CMAP sirolimus 1e-07 100 8359 100	0.5001286	+
	CMAP wortmannin 1e-06 100 577 100	0.4990501	+
	CMAP wortmannin 1e-08 100 1423 100	0.4971645	+
Trichostatin A	CMAP vorinostat 1e-05 100 817 100	0.9899307	+
	CMAP trichostatin A 1e-07 100 8791 100	0.9869679	+
	CMAP trichostatin A 1e-07 100 6547 100	0.9854099	+
	CMAP trichostatin A 1e-06 100 8056 100	0.9852016	+
	CMAP trichostatin A 1e-07 100 7972 100	0.9849593	+

- Signature scoring using the single sample Gene Set Enrichment Analysis (ssGSEA) approach (Barbie et al. 2009)
- The “correct” target classes were identified for reference chemical treatments.

HTTr Signal Strength



Concentration Response Modeling: BMDExpress



Adapted from Harrill et al. (2019)

Based on National Toxicology Program Approach to Genomic Dose-Response Modeling (NTP RR 5)

BMDExpress Parameter	Criteria
Pre-filter:	|FC| > 2 at any test concentration
Models	Hill, Power, Linear, Poly2, Exponential 2|3|4|5
BMR Factor:	1.349*SD of controls (10%)
Best Model Selection:	Lowest AIC
Hill Model Flagging:	'k' < 1/3 Lowest Positive Dose Exclude Flagged Hill Models from Best Model Selection
Conc-Response Hit Criteria	(0.1*lowest conc. < BMC < highest conc.) BMC fit p-value > 0.1 BMCL / BMCU < 40
Gene Set Analysis:	≥ 3 Concentration-responsive genes ≥ 5% Gene Set Coverage
Gene Set Collections:	MSigDB (Liberzon et al. 2015) BioPlanet (Huang et al. 2019) CMAP (Subramanian et al. 2005)

Concentration-Response Modeling of Signature Scores (1)

Step 1: Inputs

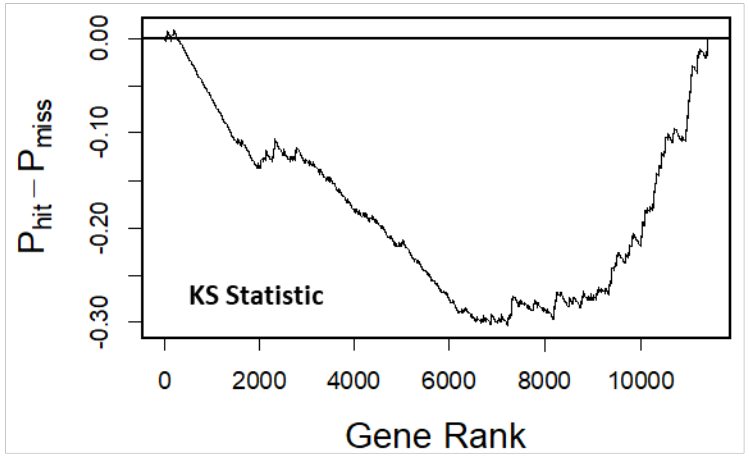
Experimental Data: Chemical_Conc × Gene matrix of $\log_2$ (fold-change) (l2fc) values.

Signature Collections: MSigDB (*Liberzon et al. 2015*), BioPlanet (*Huang et al. 2019*), CMAP (*Subramanian et al. 2005*)

Step 2: Pathway Scoring

Scores based on single sample GSEA method (Barbie et al. 2009)

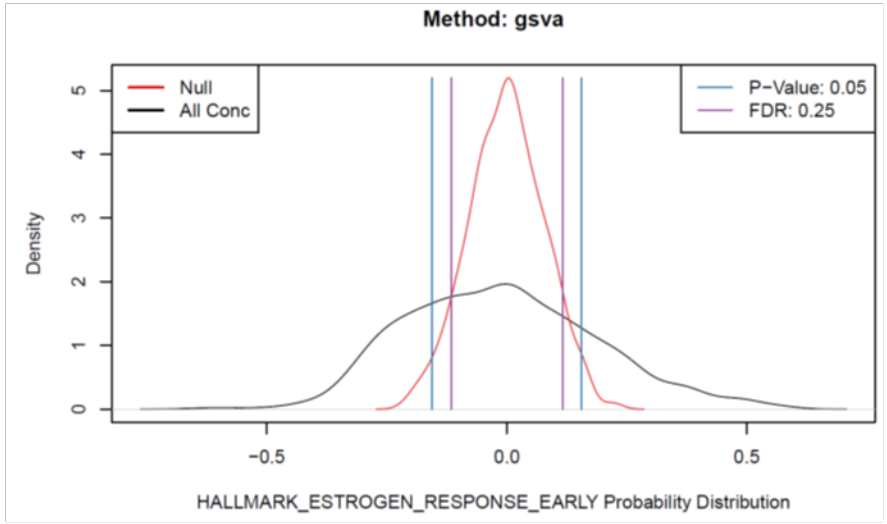
K-S statistic



Chemical_Conc × Pathway matrix of scores.

Step 3: Cut-off Estimation via NULL Modeling

- For each gene, **resample** l2fc based on the cross-sample gene distribution → breaks gene correlation
- Calculate **pathway scores for "null" data**
  - One null distribution (n = 1000 scores) / pathway

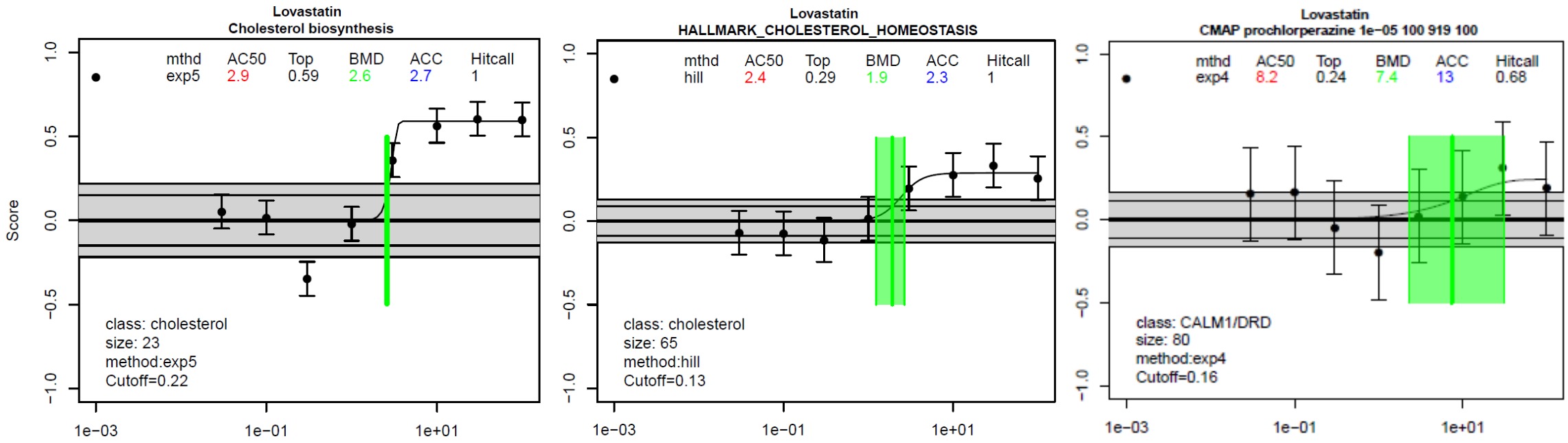


Analysis by Thomas Sheffield and Richard Judson

Concentration-Response Modeling of Signature Scores (2)

Step 4: CR Modeling

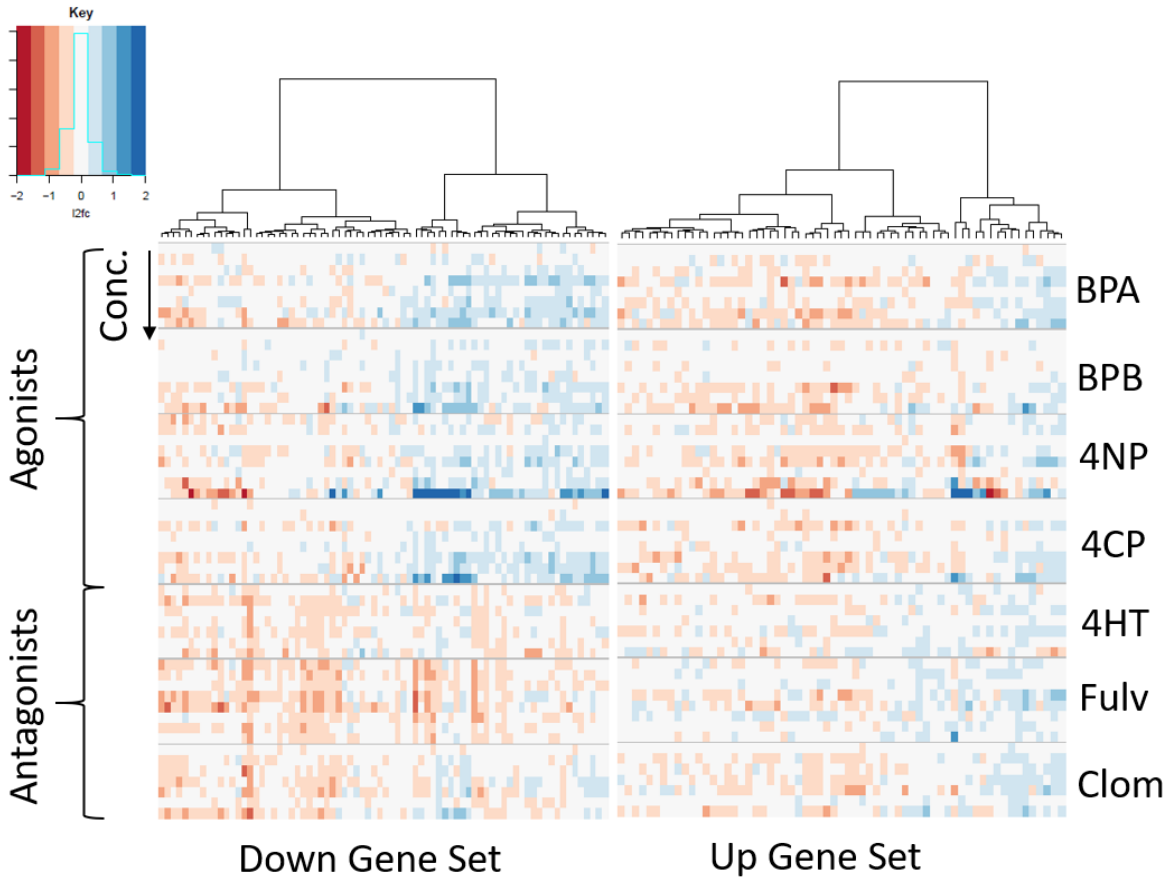
Concentration response modeling of signature scores using *tcplfit2* (Sheffield et al. *submitted*)



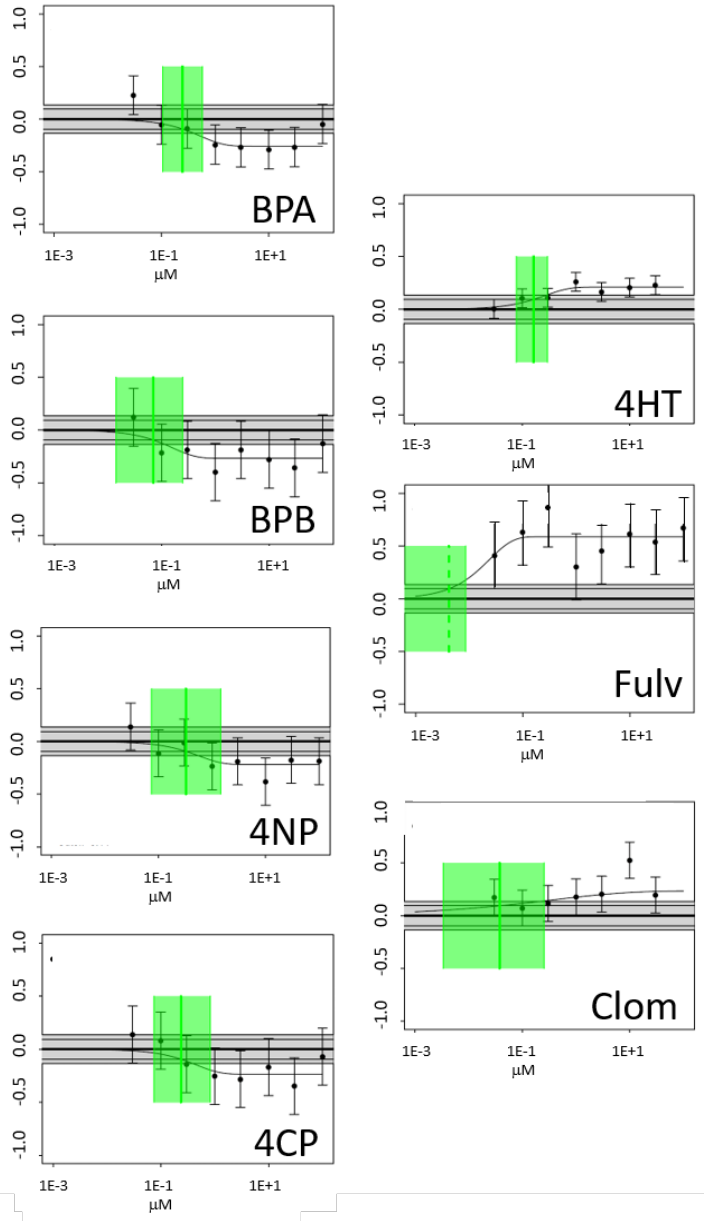
- Takes into account coordinated changes in gene expression that may not be identified using gene level fitting approaches.
- All curve forms from BMDExpress, plus constant model.
- Provides continuous hit calls for identifying high confidence and low confidence hits.

Concentration-Response Modeling of Signature Scores (3)

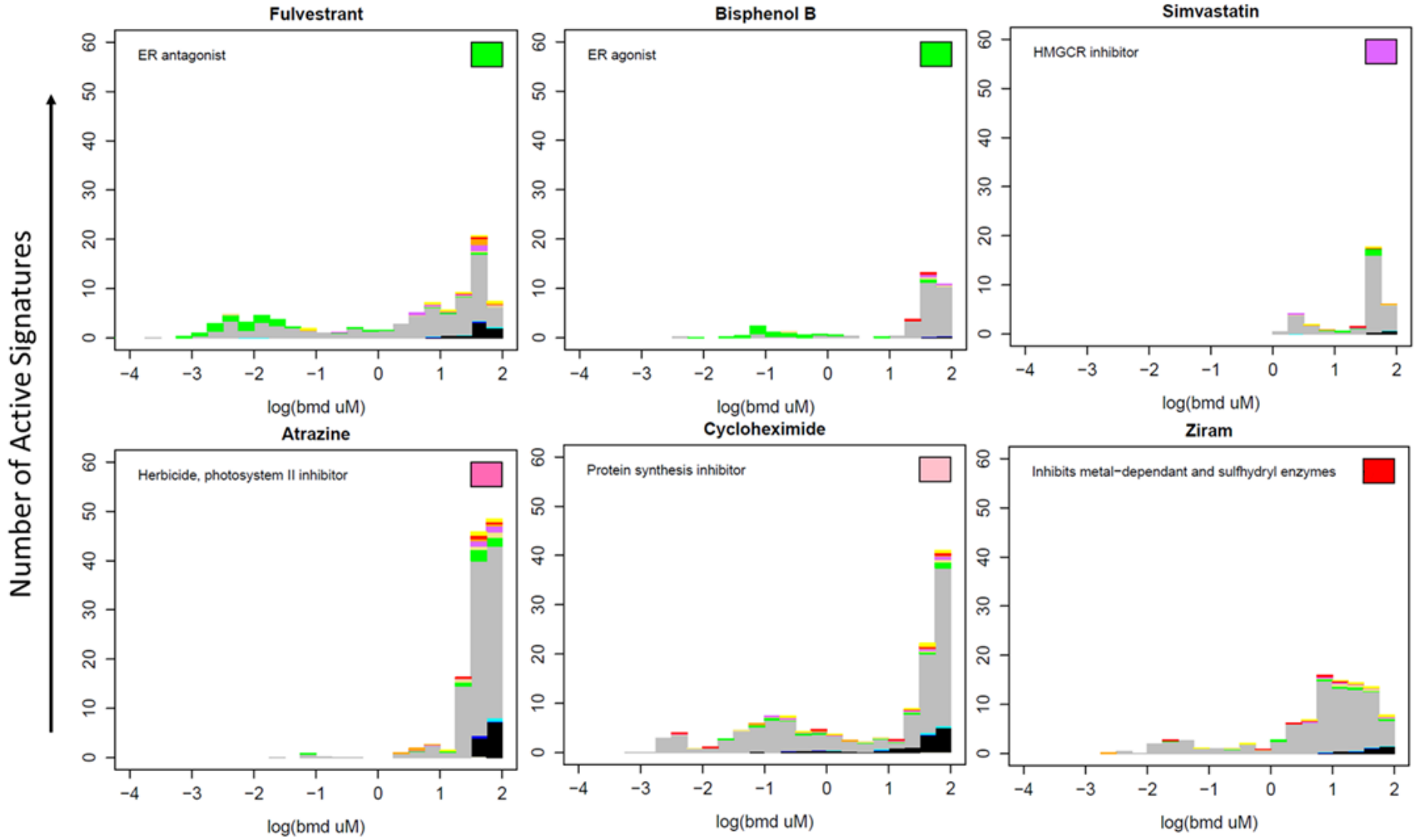
A



B

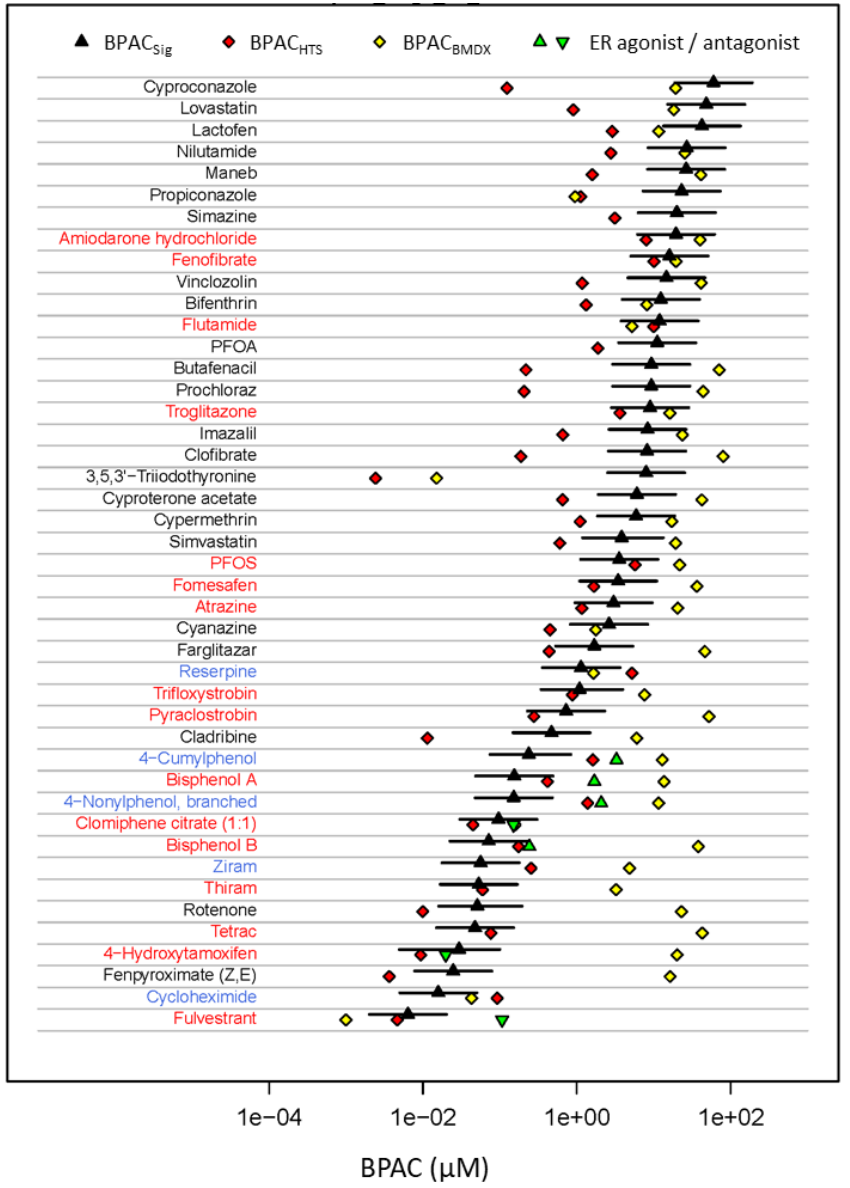


Signature Modeling Reveals Biologically Relevant Targets as Most Sensitive



Concentration-Response Modeling of Signature Scores (3)

- $BPAC_{Sig}$ → 5^{th} lowest BPAC of active signatures
- $BPAC_{BMDX}$ → Most sensitive signature / pathway
- $BPAC_{HTS}$ → Lower 5^{th} percentile of active AC50 values for assays that pass a series of quality filters.
- $BPAC_{HTS}$ and $BPAC_{Sig}$ are in better agreement than $BPAC_{HTS}$ and BPACBMDX
- In most of these cases, $BPAC_{HTS}$ is also more potent than $BPAC_{BMDX}$.
- The majority of these cases can be explained by the use of ToxCast assays for the specific target of the chemical that are not active/expressed in MCF7 cells.
  - THRA / THRB
  - CYP Assays
  - PTPN Assays



EPA United States Environmental Protection Agency

High-Throughput Phenotypic Profiling

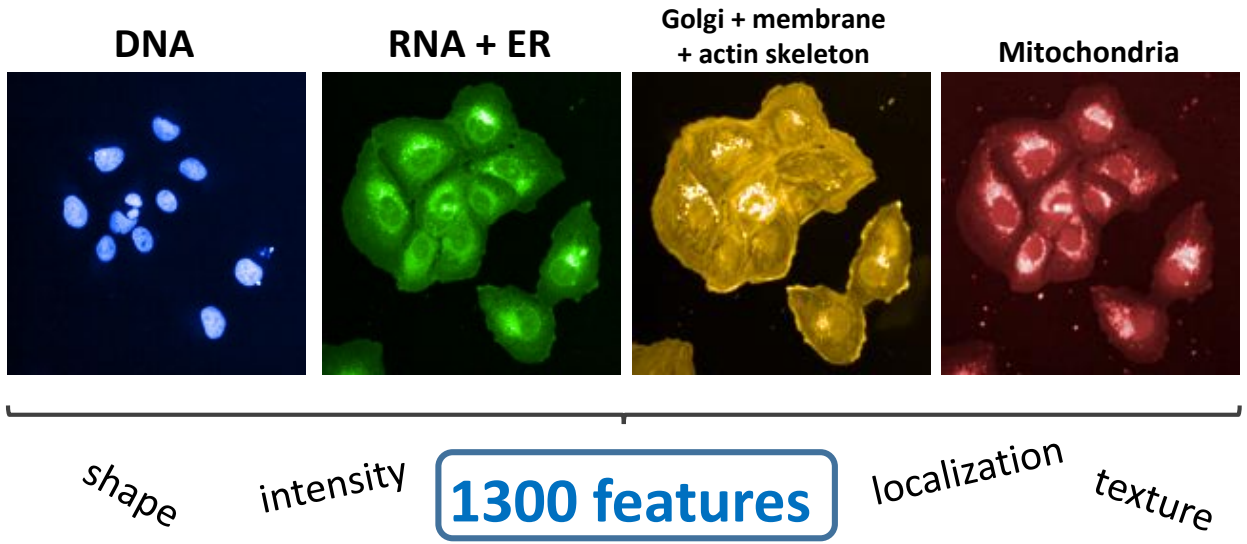
High Throughput Phenotypic Profiling with Cell Painting

- **Cell Painting** is a profiling method that measures a large variety of phenotypic features in fluoroprobe labeled cells *in vitro*.

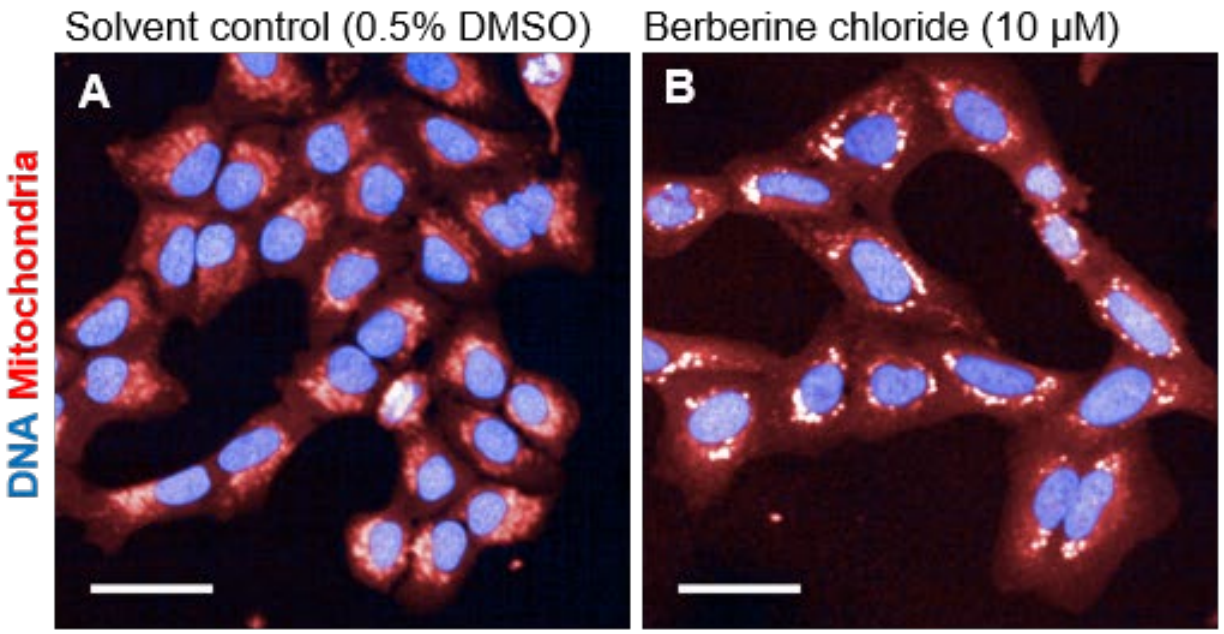
- Previous Uses:
  - Functional genomics
  - Drug discovery
  - Compound efficacy and toxicity screening
  - Mechanism-of-action identification
  - Chemical grouping

- Efficient and cost-effective method for evaluating the bioactivity of environmental chemicals.

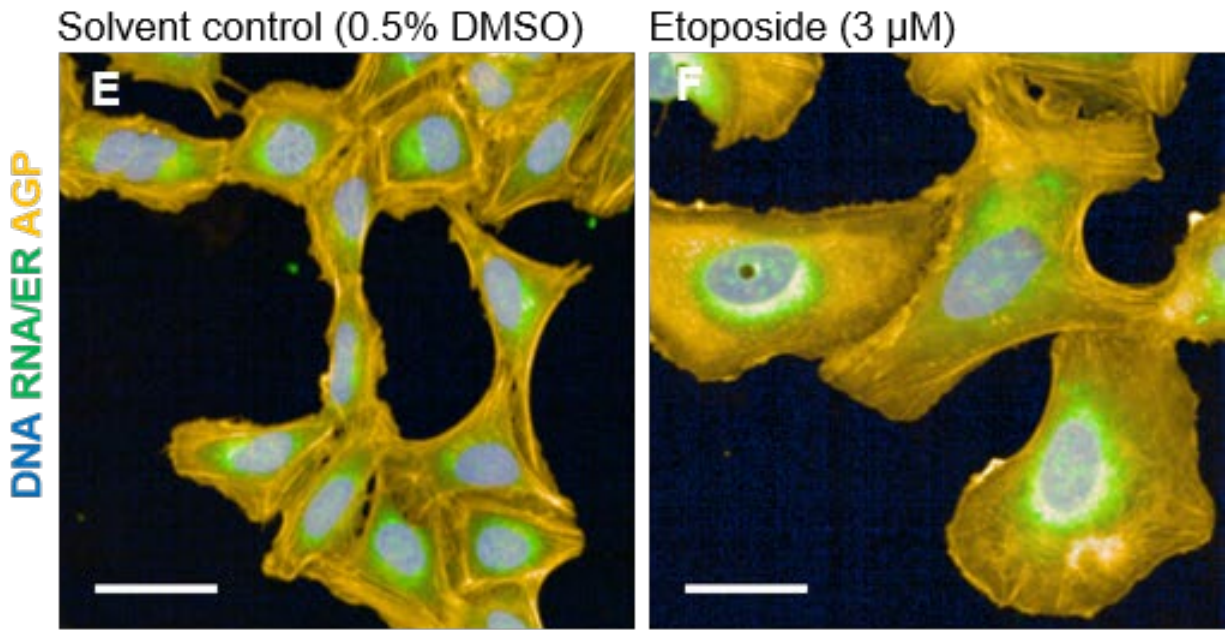
Marker	Cellular Component	Labeling Chemistry	Labeling Phase	Opera Phenix	
				Ex.	Em.
Hoechst 33342	Nucleus	Bisbenzamide probe that binds to dsDNA	Fixed	405	480
Concanavalin A – AlexaFluor 488	Endoplasmic reticulum	Lectin that selectively binds to α-mannopyranosyl and α-glucopyranosyl residues enriched in rough endoplasmic reticulum		435	550
SYTO 14 nucleic acid stain	Nucleoli	Cyanine probe that binds to ssRNA		435	550
Wheat germ agglutinin (WGA) – AlexaFluor 555	Golgi Apparatus and Plasma Membrane	Lectin that selectively binds to sialic acid and N-acetylglucosaminyl residues enriched in the trans-Golgi network and plasma membrane		570	630
Phalloidin –AlexaFluor 568	F-actin (cytoskeleton)	Phallotoxin (bicyclic heptapeptide) that binds filamentous actin			
MitoTracker Deep Red	Mitochondria	Accumulates in active mitochondria	Live	650	760



Example Chemicals



→ Mitochondrial compactness/texture



→ Cells are larger

- Strong phenotypes are observable qualitatively and can be measured quantitatively using Cell Painting

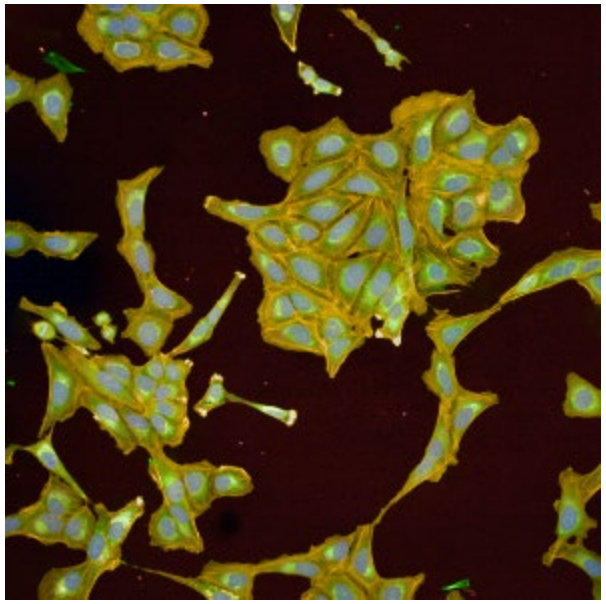
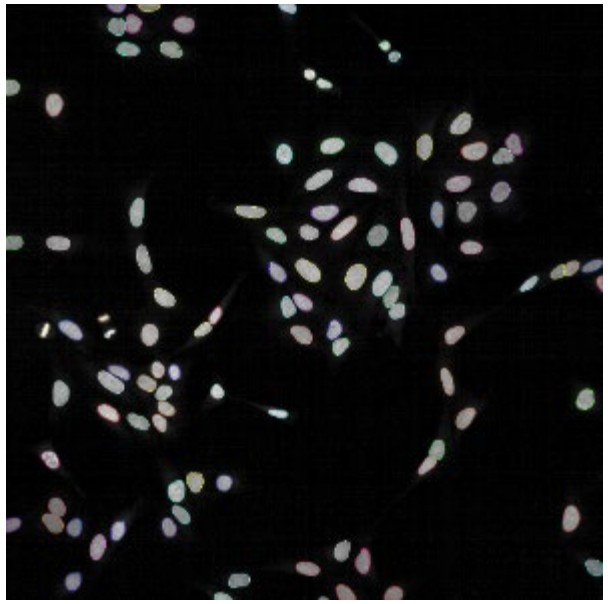
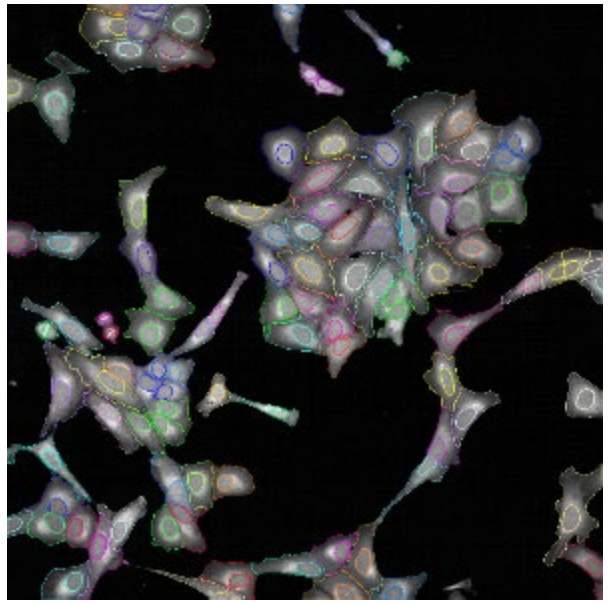
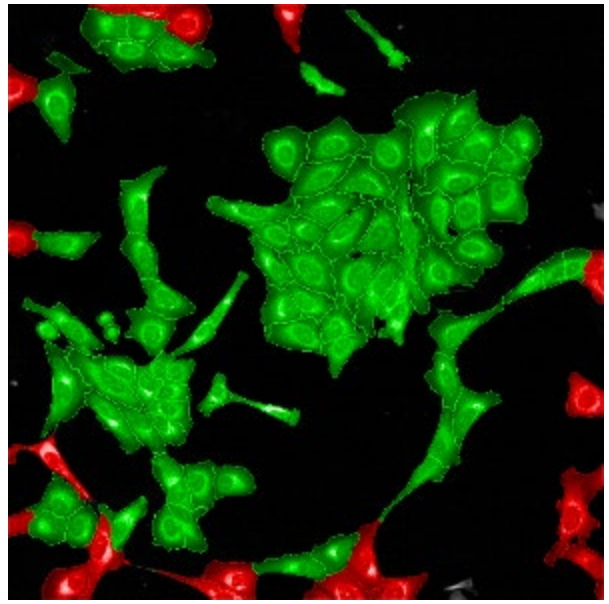
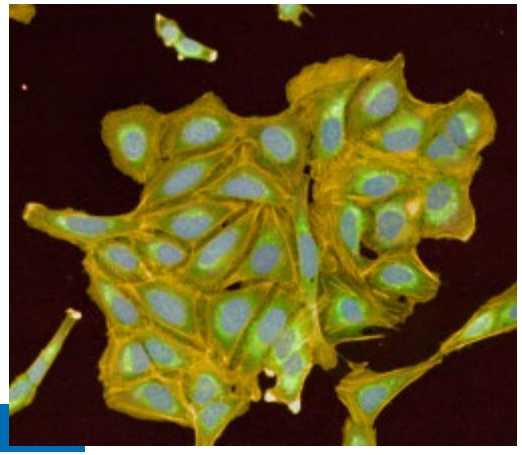
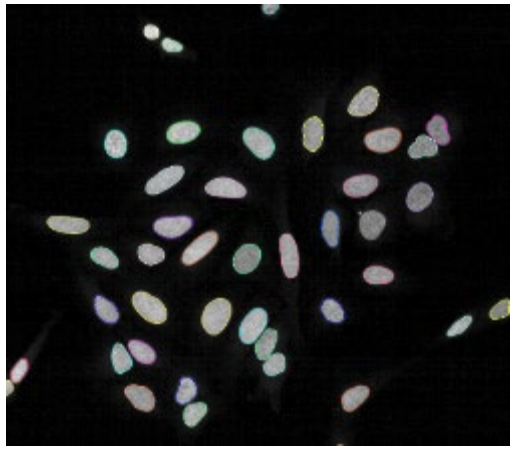
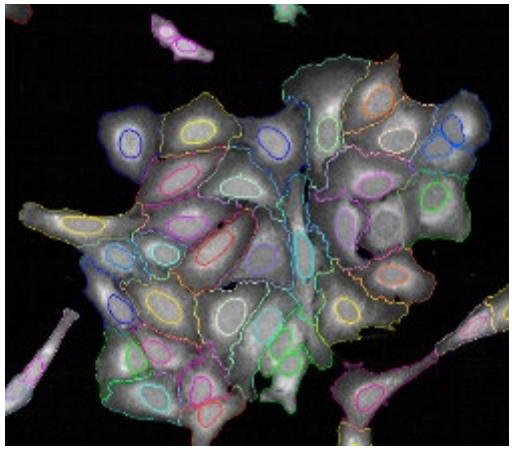
adapted from Nyffeler et al. 2020a

Image Analysis Workflow → Image Segmentation

1. find nuclei

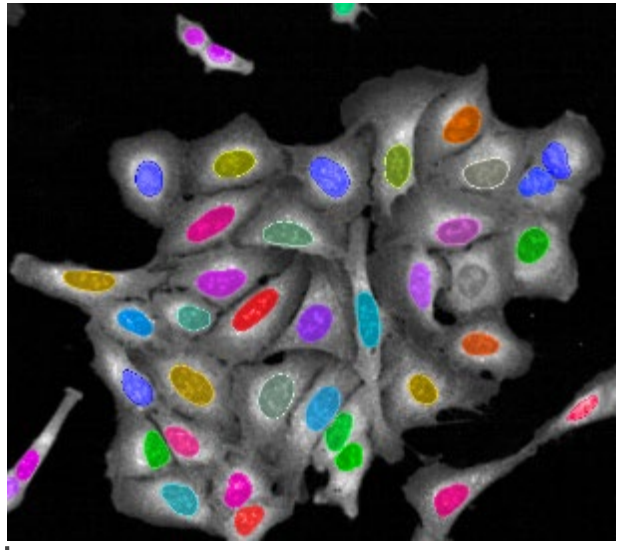
2. find cell outline

3. reject border objects

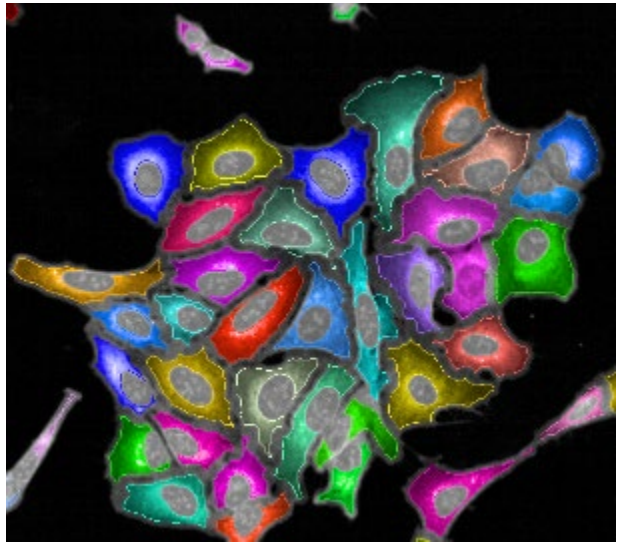








Define Cellular Compartments

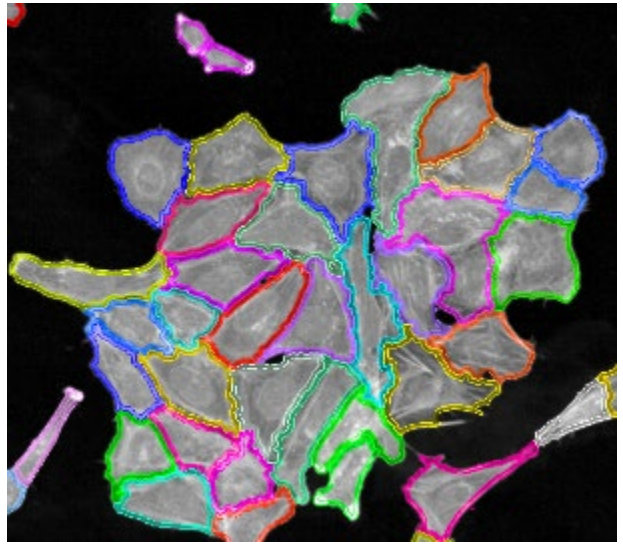
nuclei



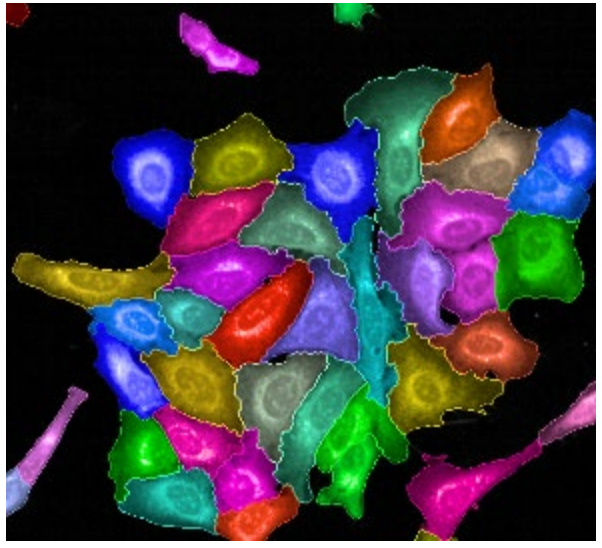
cytoplasm



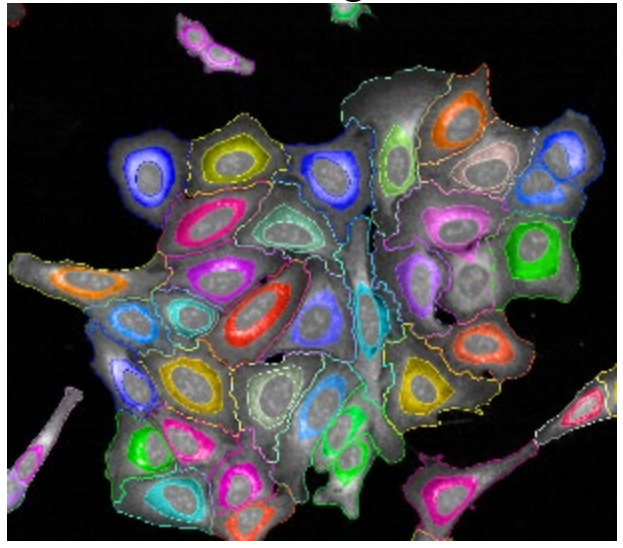
membrane



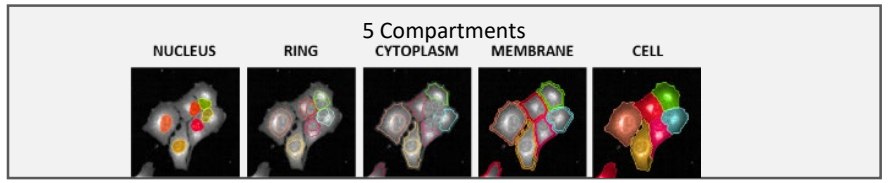
cell

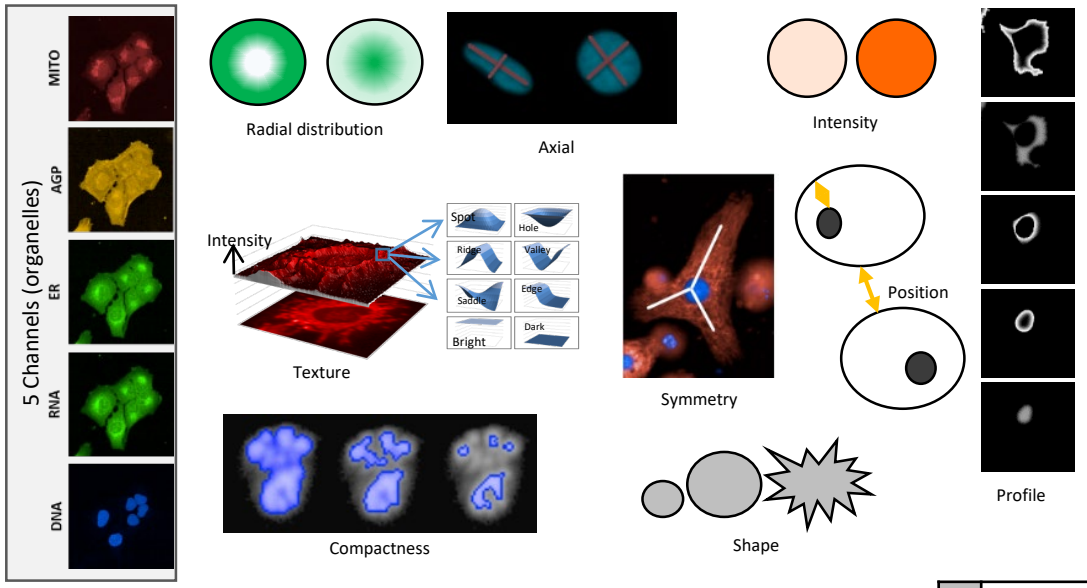


ring



Phenotypic Feature Extraction





49 Feature Categories
(ex. MITO_Texture_Cytoplasm)

1300 features / cell

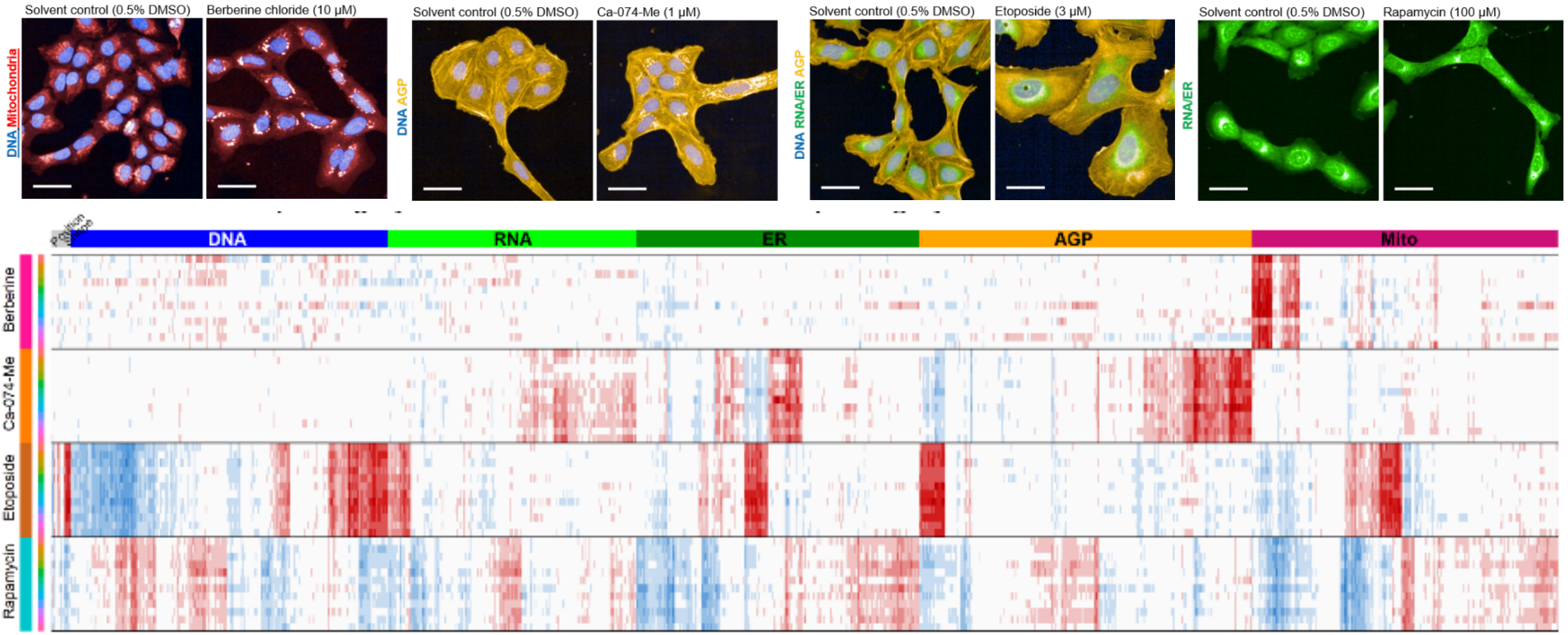
		Module								
		Position [7]	Basic morph-ology [5]	SCARP morphology					Intensity [9]	Texture [14]
				Symmetry [80]	Compactness [40]	Axial [20]	Radial [28]	Profile [20-30]		
Channel	DNA			Nuclei	Nuclei	Nuclei	Nuclei Cell	Nuclei Cytoplasm	Nuclei	Nuclei
	RNA			Nuclei	Nuclei	Nuclei	Nuclei	Nuclei	Nuclei	Nuclei
	ER			Cell	Cell	Cell	Cell	Cytoplasm	Ring Cytoplasm	Ring Cytoplasm
	AGP			Cell	Cell	Cell	Cell	Nuclei Cytoplasm	Ring Cytoplasm Membrane	Ring Cytoplasm Membrane
	Mito			Cell	Cell	Cell	Cell	Nuclei Cytoplasm	Ring Cytoplasm	Ring Cytoplasm
	Not associated with a channel	Nuclei Cell	Nuclei Cell							

PerkinElmer Opera Phenix

Modality: Confocal (single *z*)
Objective: 20X Water
Plate: CellCarrier-384 Ultra
Fields: 5 or 9

With illustrations from Perkin Elmer

Reference Chemical Phenotypes



- Reference chemicals produce distinct, but reproducible phenotypes in U-2 OS cells.

HTPP Data Analysis Pipeline

Data reduction

cell-level data

Normalization
MAD normalization

$$\frac{\text{cell value} - \text{median}_{DMSO}}{1.4826\ \text{MAD}_{DMSO}}$$

normalized cell-level data

Aggregation
median

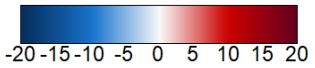
well-level data

Standardization
Z transformation

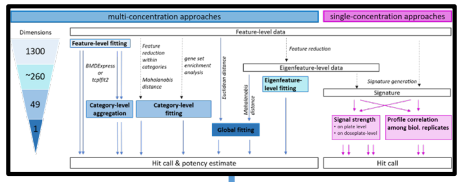
Cell Count Info
Conc. > 50% cell loss

scaled well-level data

clipped well-level data



Concentration Response Modeling



See Next Slide

Fit Multiple Curve Shapes

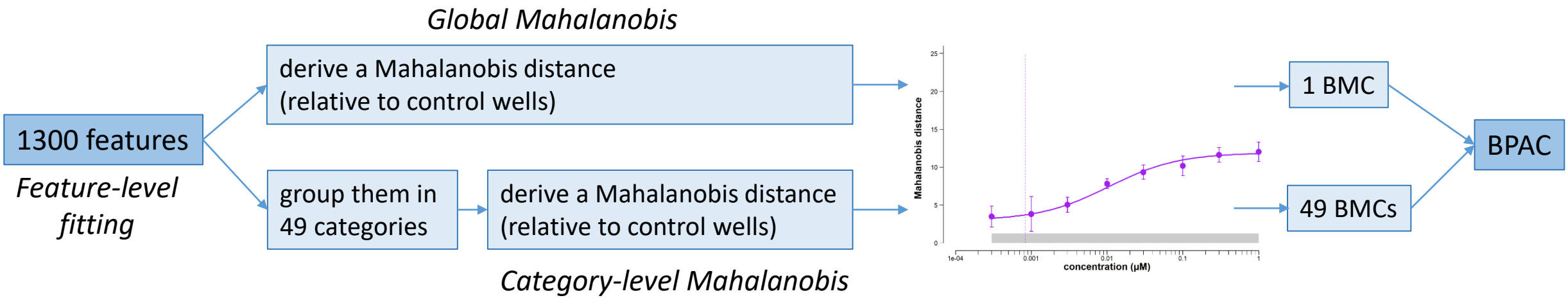
Best Model Selection

BMC

Berberine chloride
Mito_Cells_Morph_STAR

Mahalanobis Distance Modeling of HTPP Data

Mahalanobis Distance (D_M): A multivariate distance metric that measures the distance between a point (vector) and a distribution.

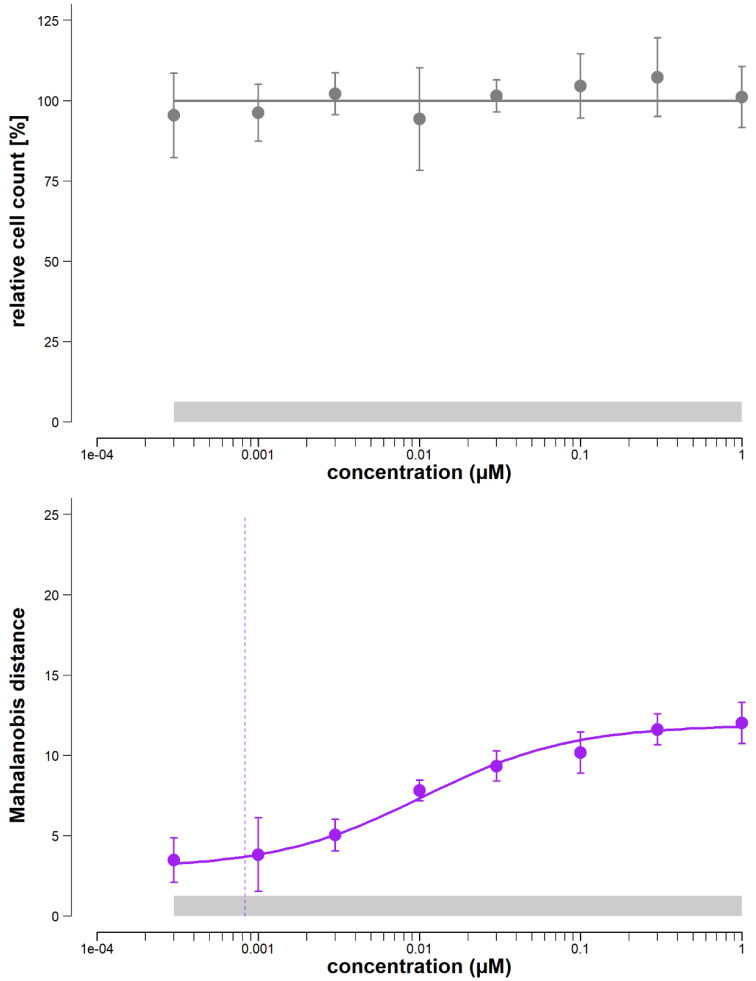


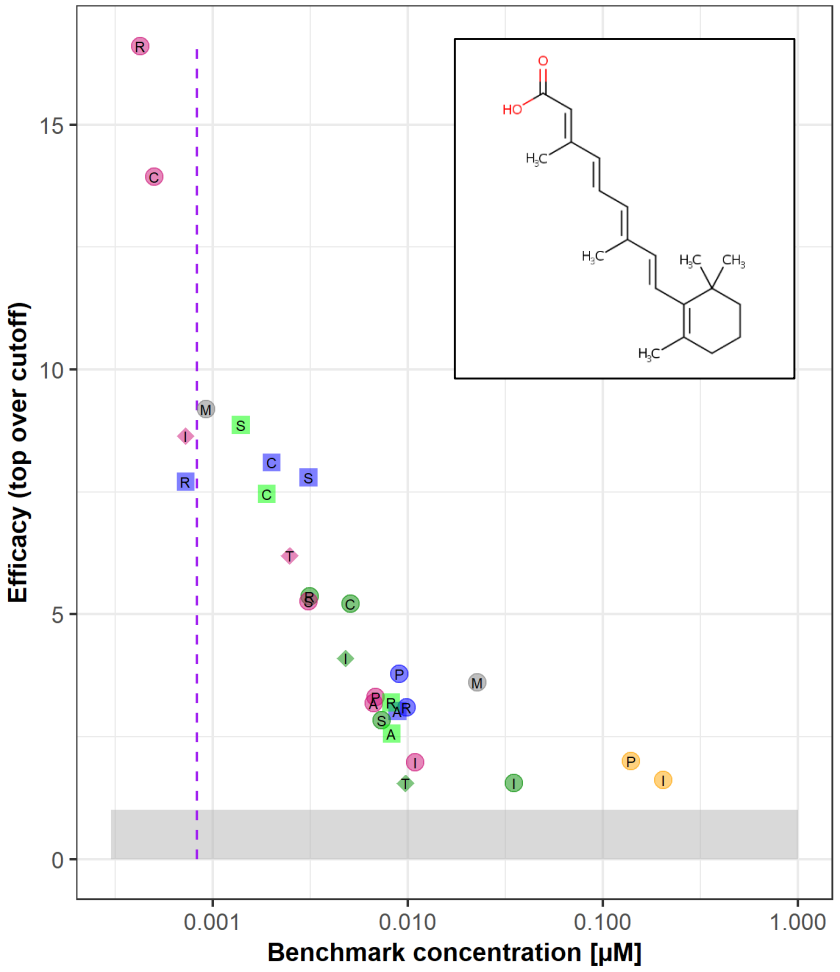
- Chemicals where a BMC can be determined using either the global or category D_M approach are considered active.
- The minimum of the global or most sensitive category BMC is the **Biological Phenotype Altering Concentration** (BPAC)

Concentration Response Modeling Example Chemical

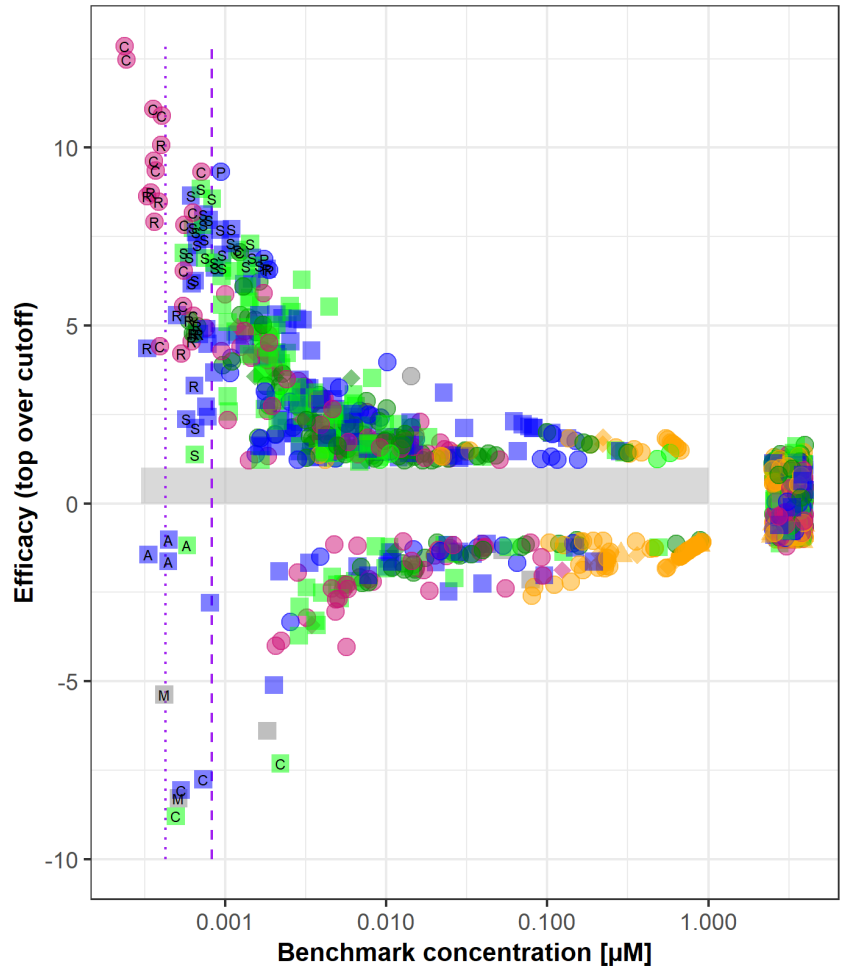
all-trans-Retinoic acid

DTXSID7021239 | 302-79-4 | RA





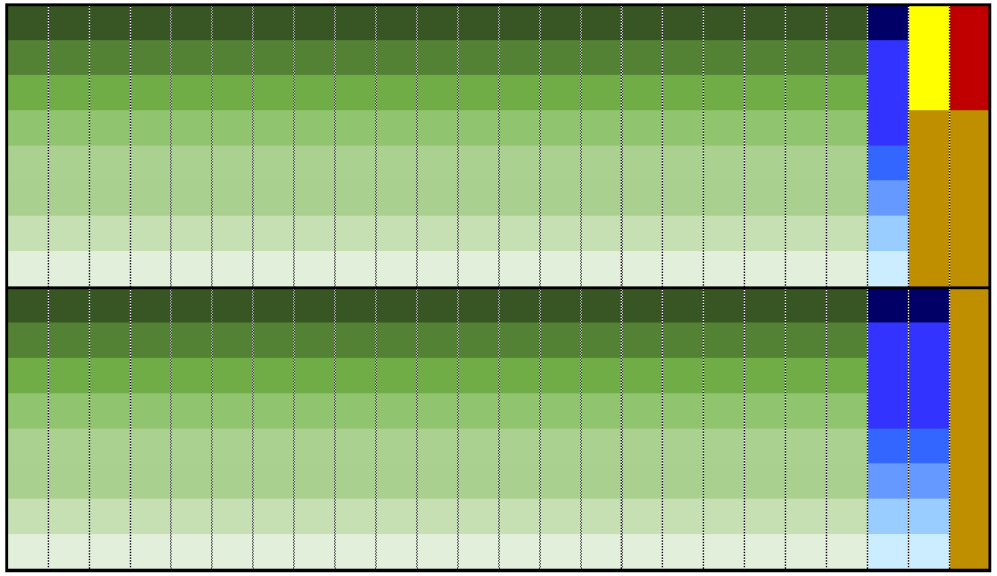
2020-07-27

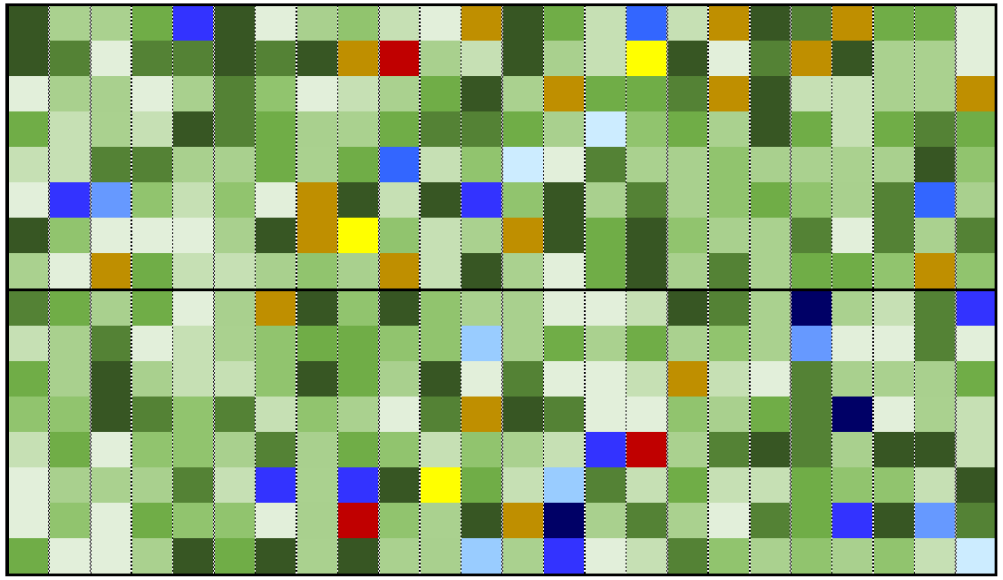


U-2 OS ToxCast Screen Experimental Design

Parameter	Multiplier	Notes
Cell Type(s)	1	U-2 OS
Culture Condition	1	DMEM + 10% HI-FBS
Chemicals	1,202	TSCA Chemicals of interest to USEPA Includes 462 APCRA case study chemicals Includes 179 chemicals with annotated molecular targets
Time Points:	1	24 hours
Assay Formats:	2	High Throughput Phenotypic Profiling (Cell Painting) *High Throughput Transcriptomics (TempO-Seq)*
Concentrations:	8	3.5 $\log_{10}$ units; ~half-$\log_{10}$ spacing
Biological Replicates:	4	--

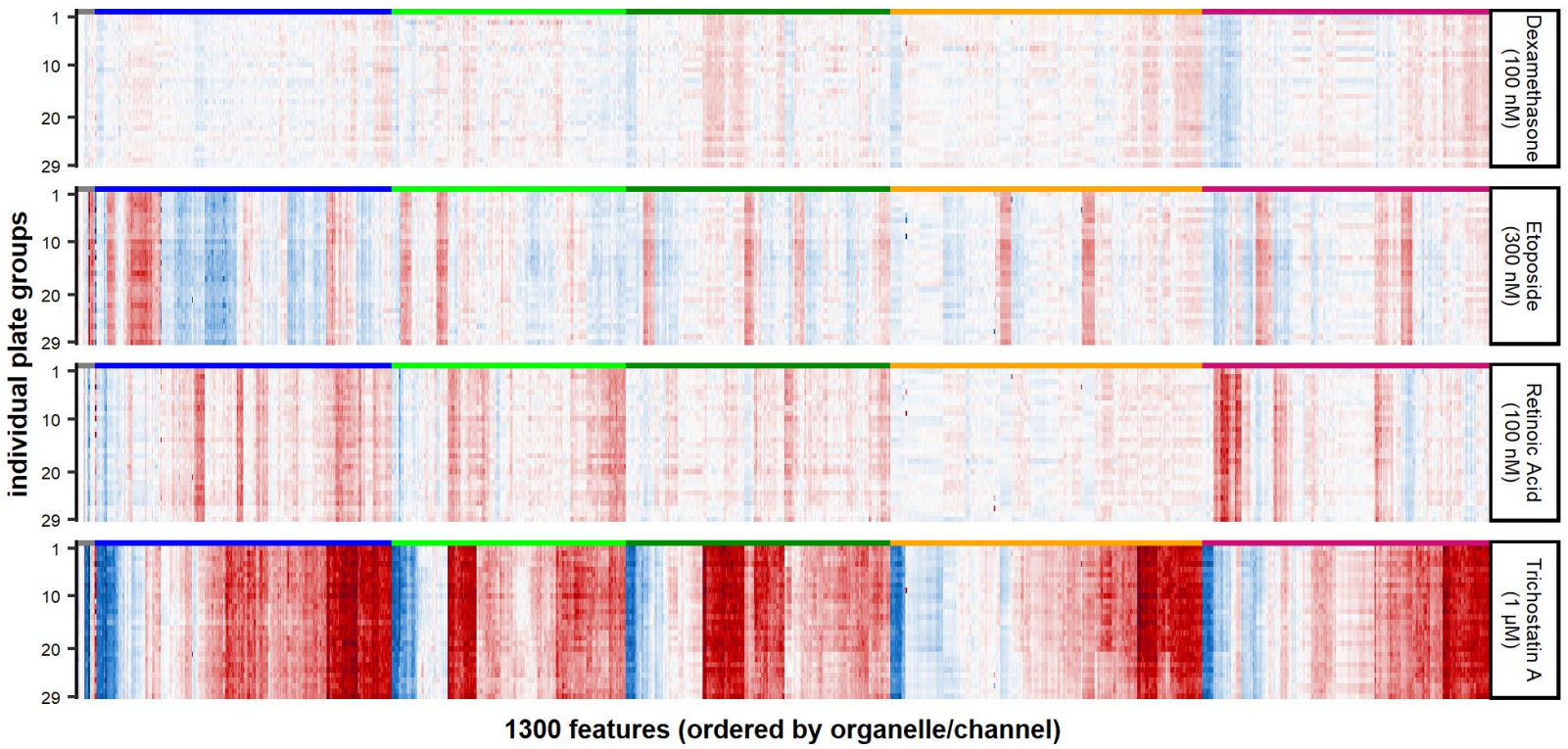
U-2 OS ToxCast HTPP Screen Dose Plate Design

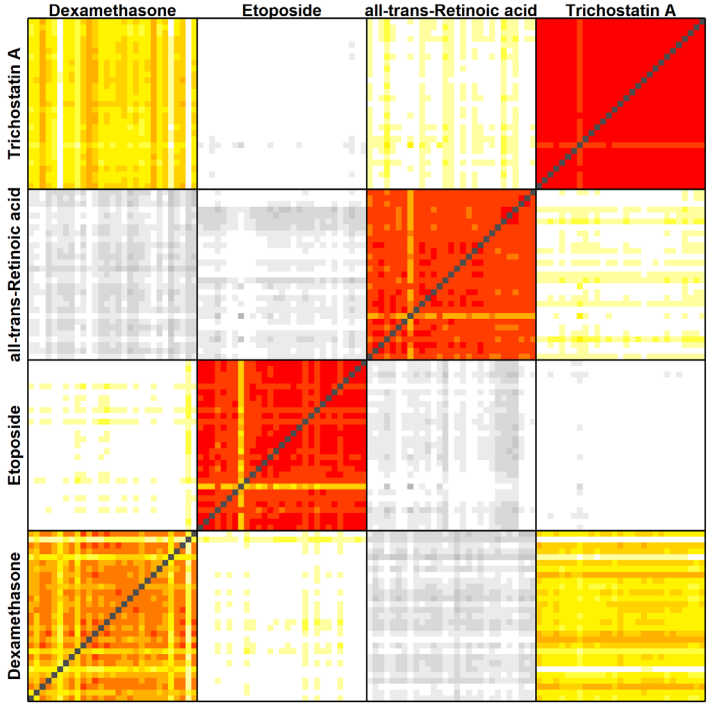




Label	Reference Chemicals:	Molecular Mechanism-of-Action	Test Concentrations
A	Etoposide	DNA topoisomerase inhibitor	0.03 - 10 µM
B	all-trans-Retinoic Acid	Retinoic acid receptor agonist	0.0003 – 1 µM
C	Dexamethasone	Glucocorticoid receptor agonist	0.001 – 3 µM
D	Trichostatin A	Histone deacetylase inhibitor	1 µM
E	Staurosporine	Cytotoxicity control	1 µM
F	DMSO	Vehicle control	0.5 %

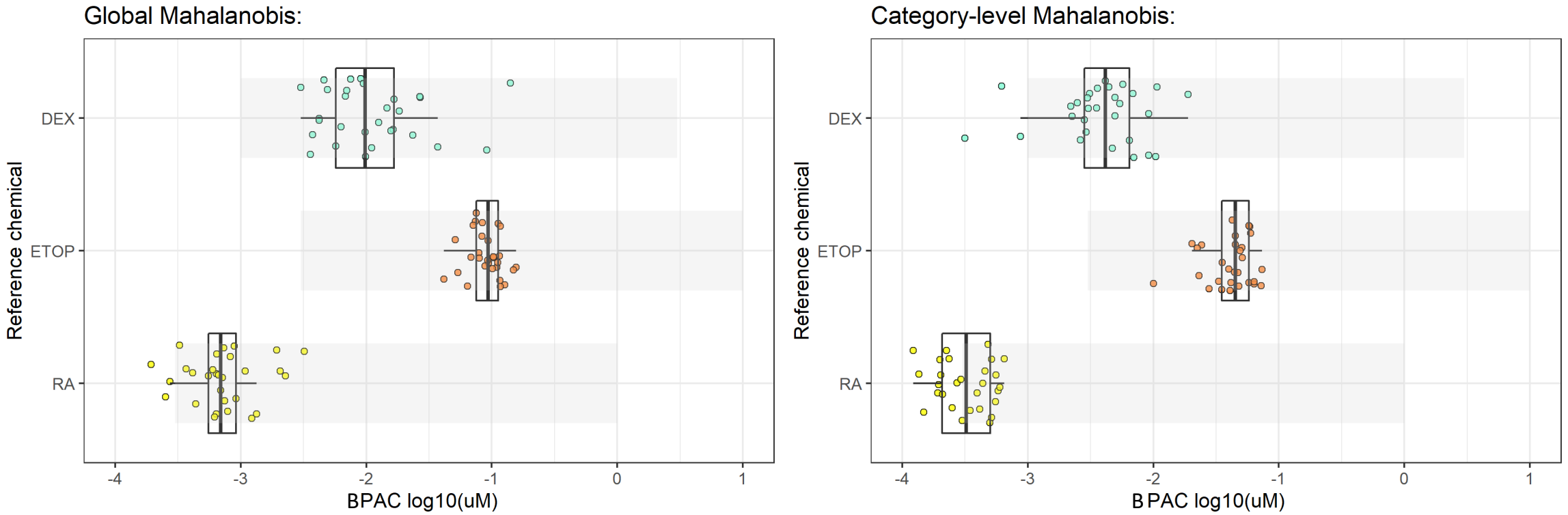
Assay Performance / Reproducibility (1)





⇨ **Reference chemicals produce <u>reproducible</u> and <u>distinct</u> profiles.**

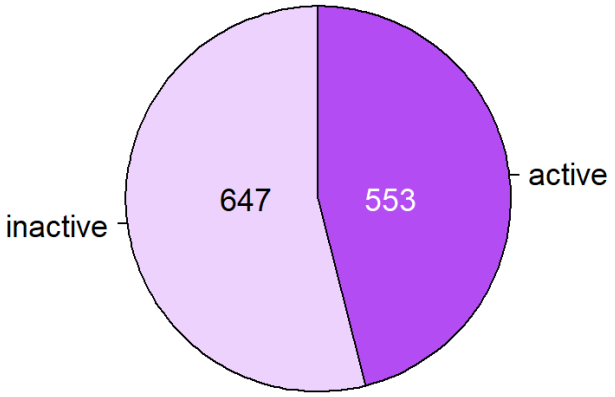
Assay Performance / Reproducibility (2)

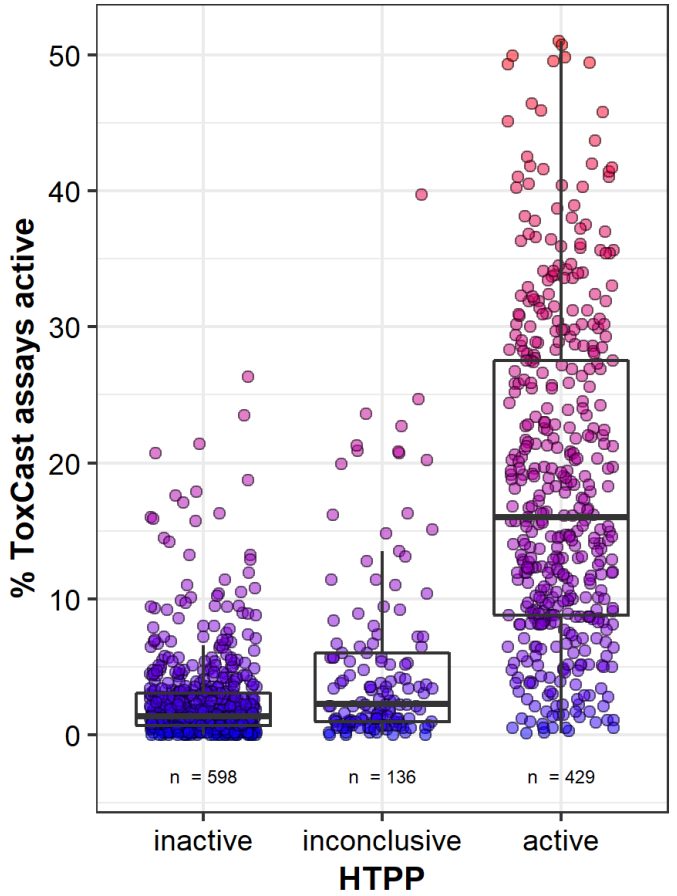


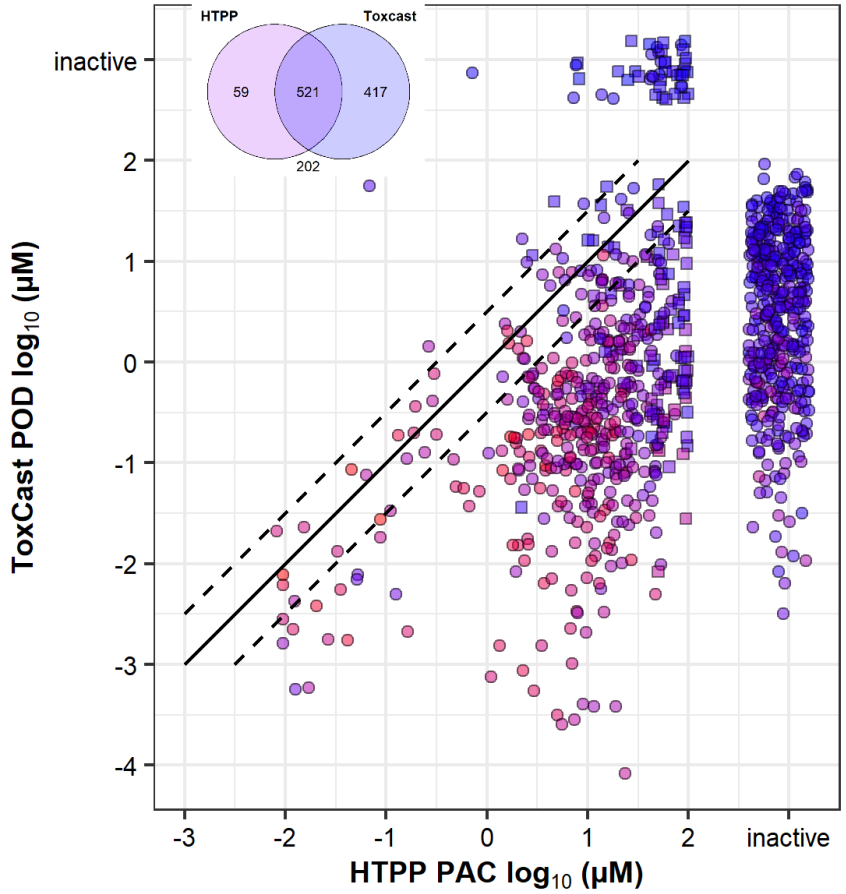
- Comparable results between global and category Mahalanobis distances, but BPACs for the latter are roughly ½ an order of magnitude lower.
- The SD for a BPAC is < ½ an order of magnitude

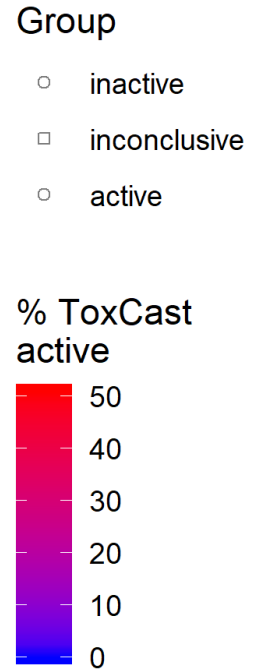
BPAC	
Retinoic Acid:	~ 0.3 nM
Dexamethasone:	~ 3 nM
Etoposide:	~ 30 nM

HTPP ToxCast Screening Results Summary









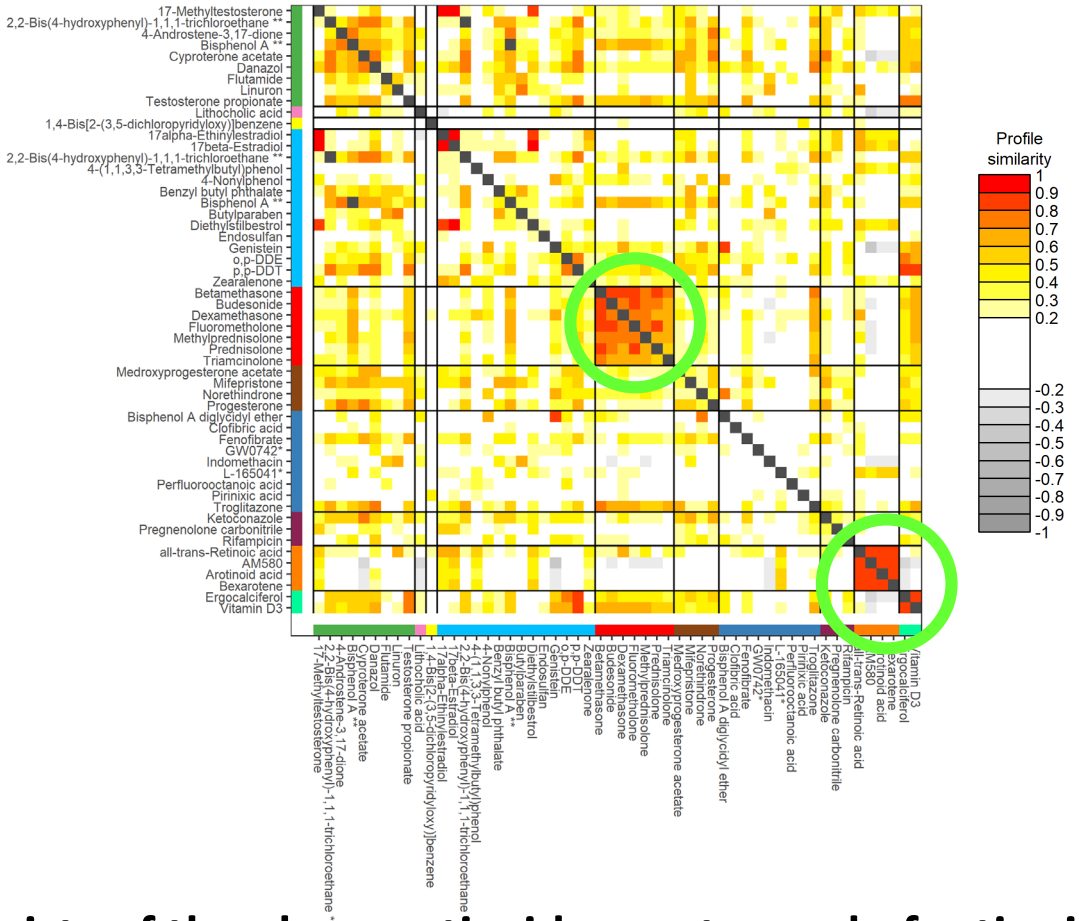
⇨ Chemicals active in the HTPP assay tend to have more hits in the ToxCast assay collection.

Phenotypic Profile Similarity with Nuclear Receptor Modulators

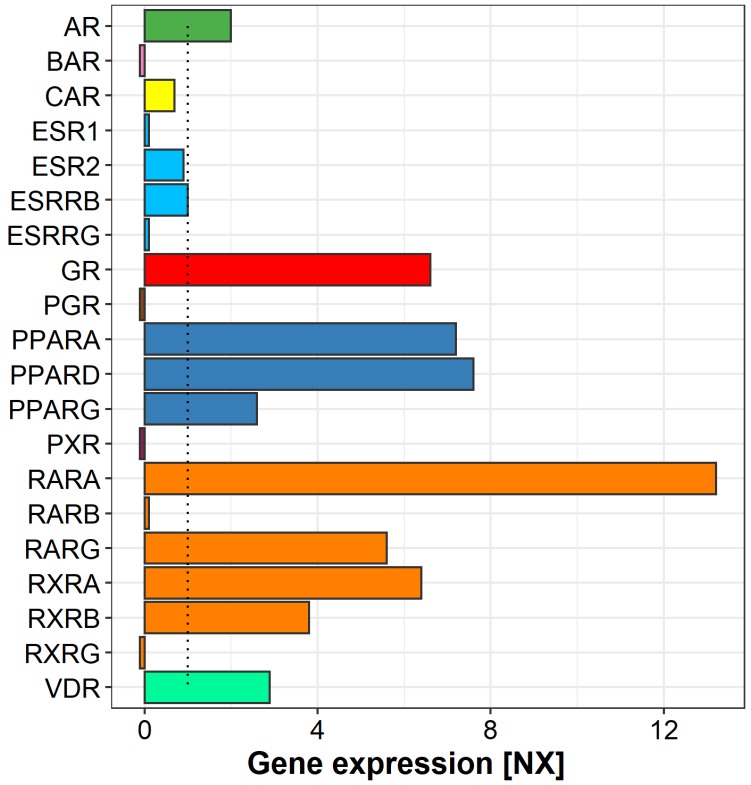
target

- AR
- BAR
- CAR
- ESR
- GR
- PGR
- PPAR
- PXR
- RAR
- VDR

Biological similarity in HTPP



Gene expression in U-2 OS

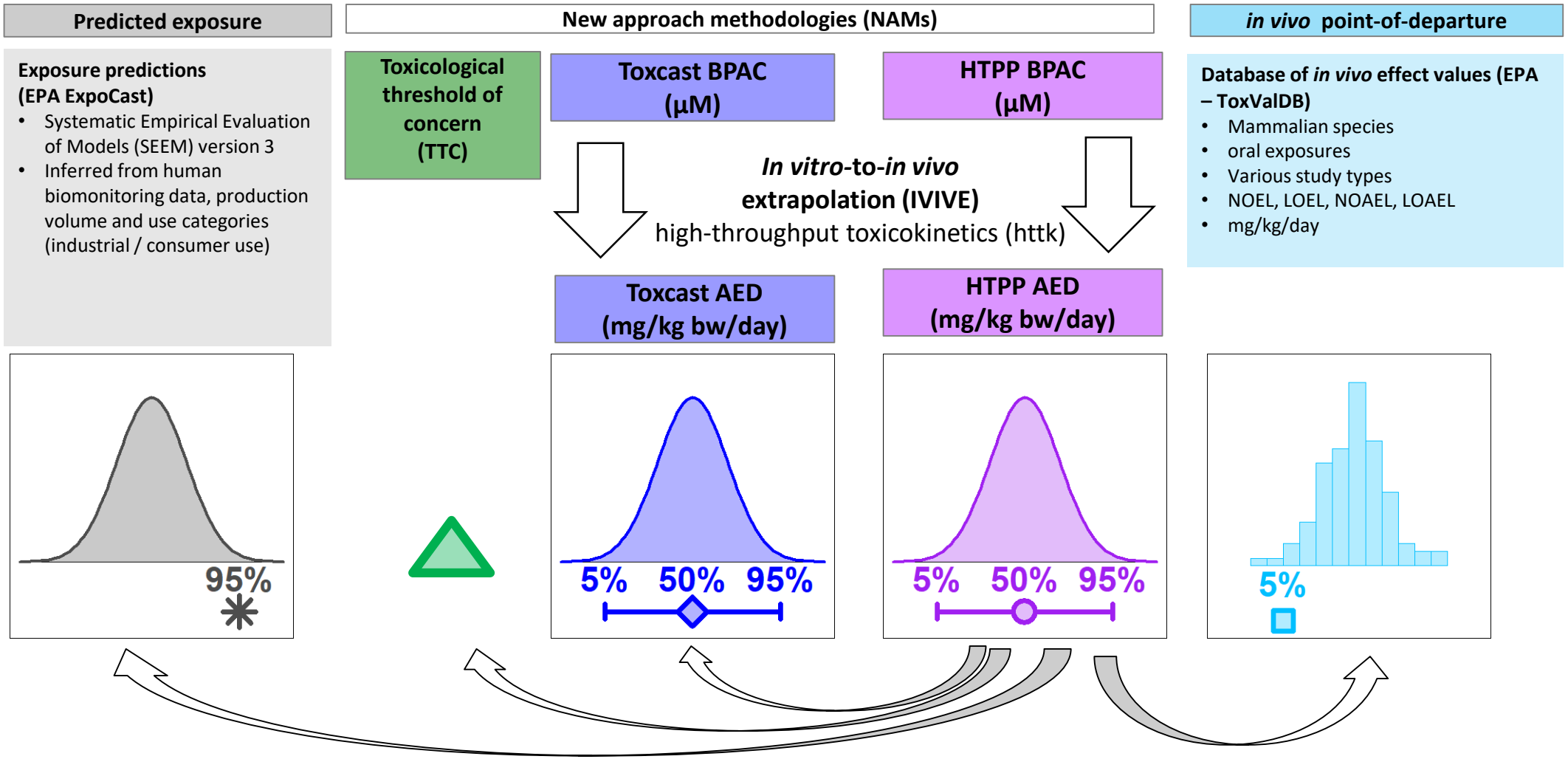


- ⇨ Agonists of the glucocorticoid receptor and of retinoic acid receptors display characteristic profiles
- ⇨ Expression of a target does not guarantee that characteristic profiles are observed (e.g. PPAR)

EPA United States Environmental Protection Agency

In Vitro to *In Vivo* Extrapolation (IVIVE) & Bioactivity to Exposure Ratio (BER) Analysis

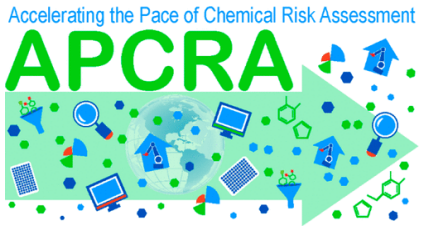
In Vitro to *In Vivo* Extrapolation (IVIVE) Using httk



POD: point-of-departure
AED: administered equivalent dose

High-Throughput Transcriptomics (HTTr) Screens

Parameter	Multiplier	Notes
Cell Types & Exposure Durations	3	MCF7 (6 HR) U-2 OS (24 HR) HepaRG_2D (24 HR)
Chemicals	~ 1,200	TSCA Chemicals of interest to USEPA **Includes 462 APCRA case study chemicals**
Assay Formats:	1	High Throughput Transcriptomics (TempO-Seq)
Concentrations:	8	3.5 log_{10} units; ~half-log_{10} spacing
Biological Replicates:	3	--

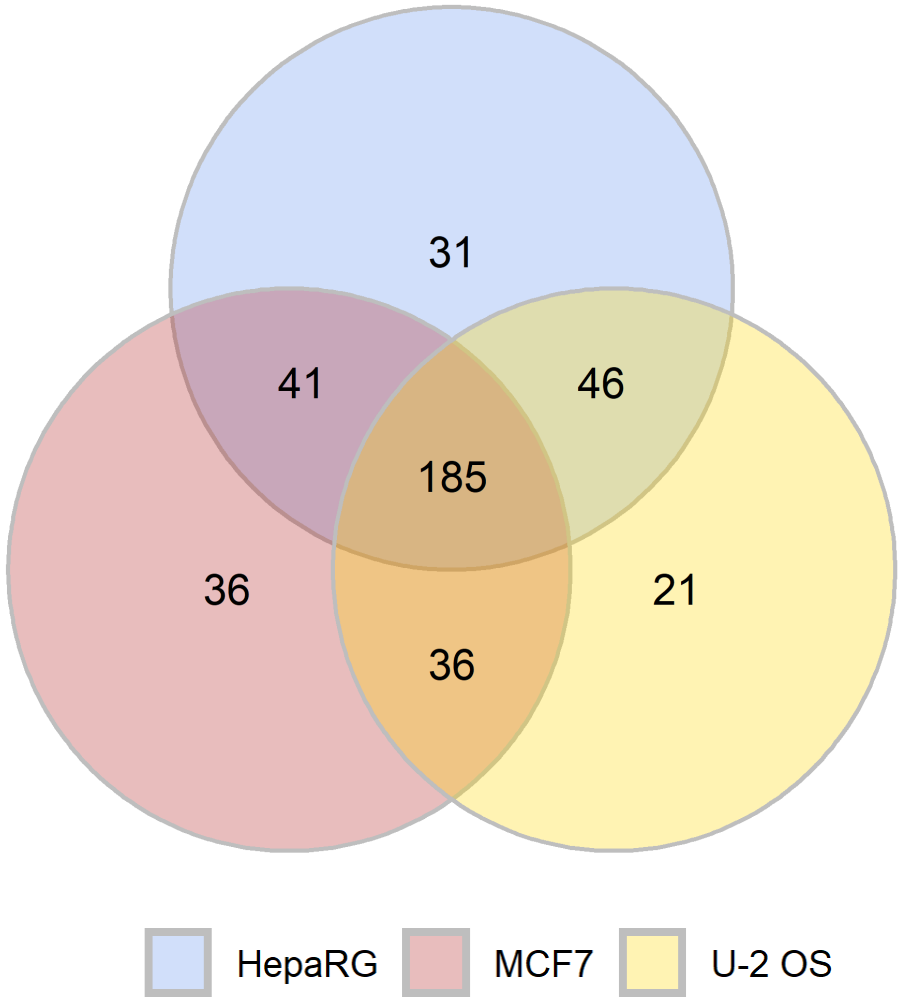


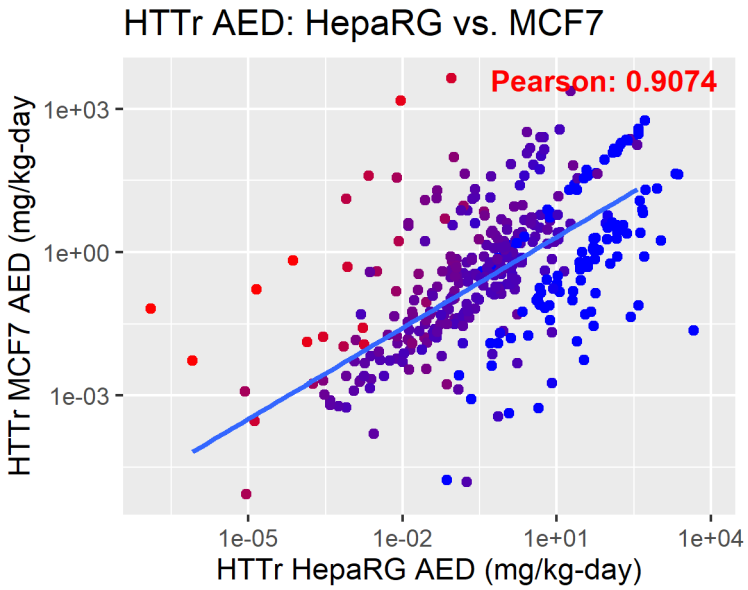
Kavlock et al. (2018)
Chem. Res. Tox; 31(5): 287-290

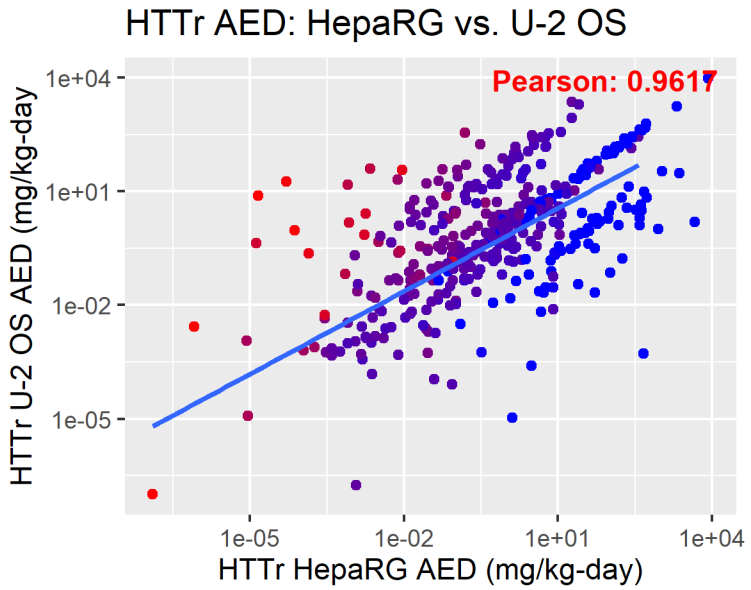
International collaboration of regulatory scientists focused on next generation chemical risk assessment including **deriving quantitative estimates of risk based on NAM-derived potency information and computational exposure estimates.**

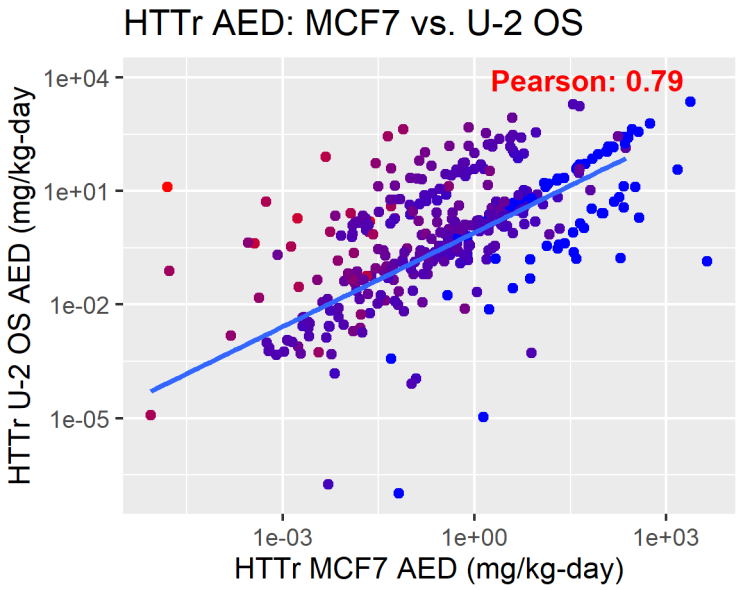
APCRA Chemicals → PK parameters necessary for *in vitro* to *in vivo* extrapolation (IVIVE)
in vivo toxicity data

HTTr Screening Hit Tally & AED Correlation

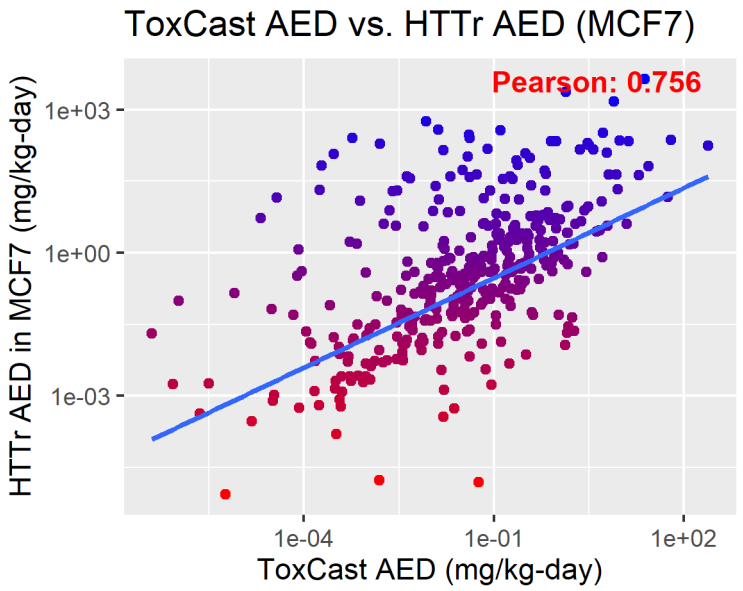


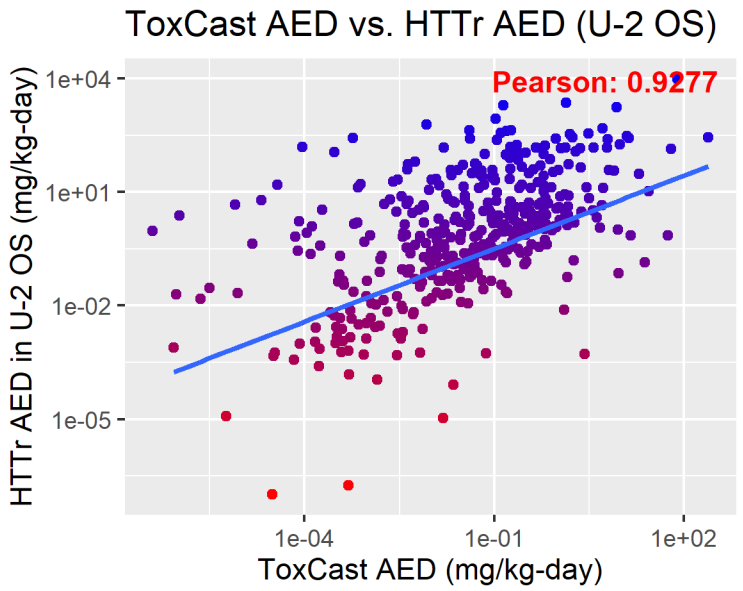


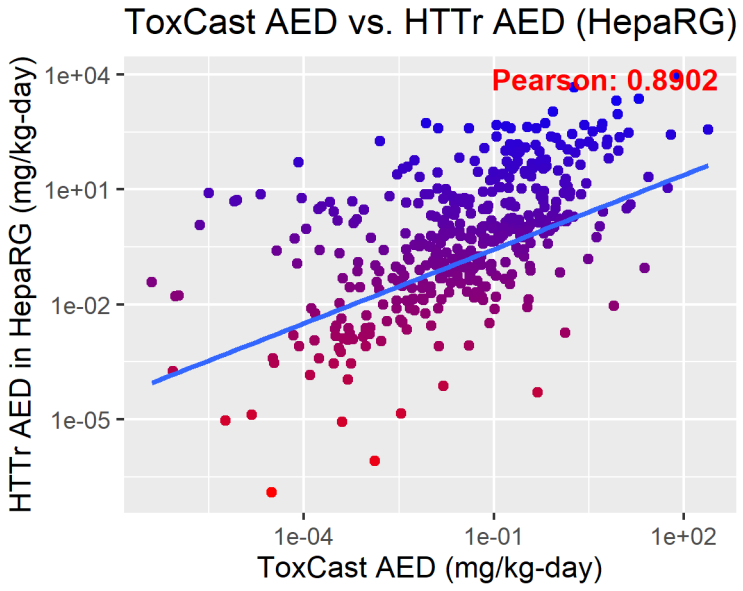


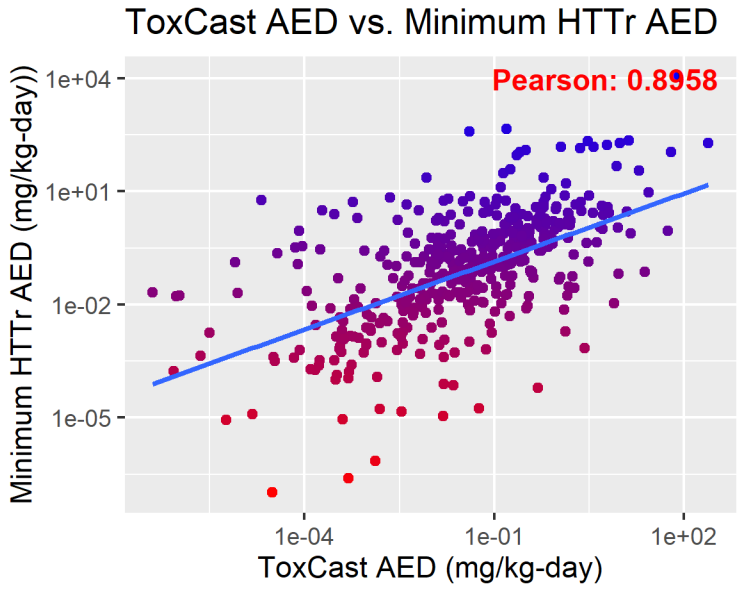


Correlation Between HTTr AED vs. ToxCast AED

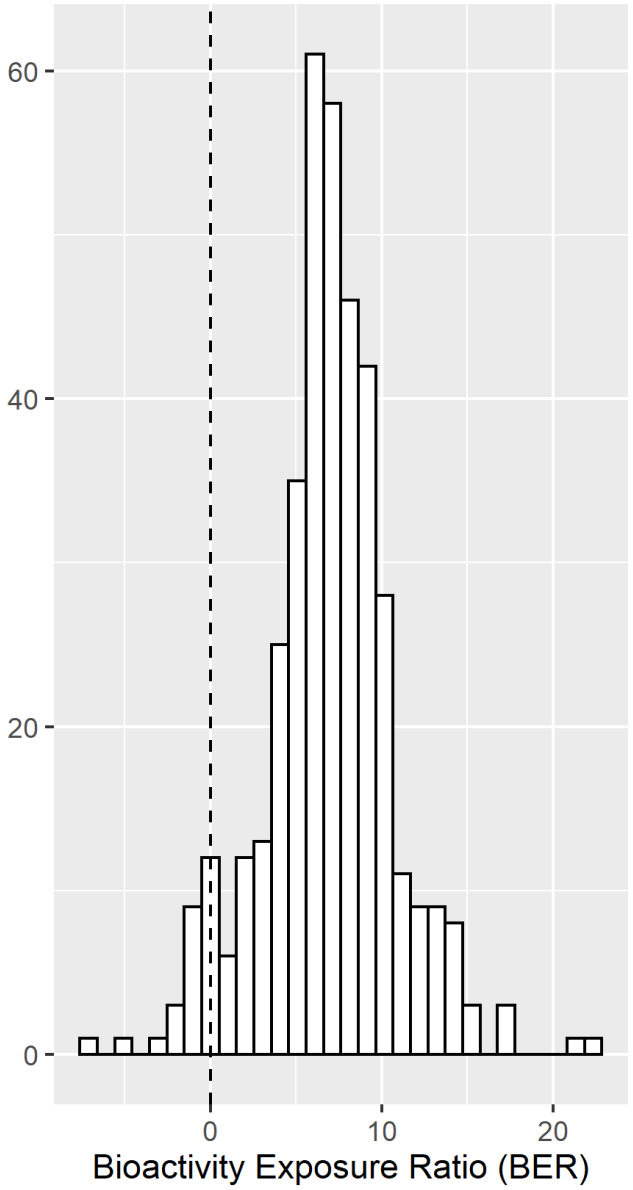


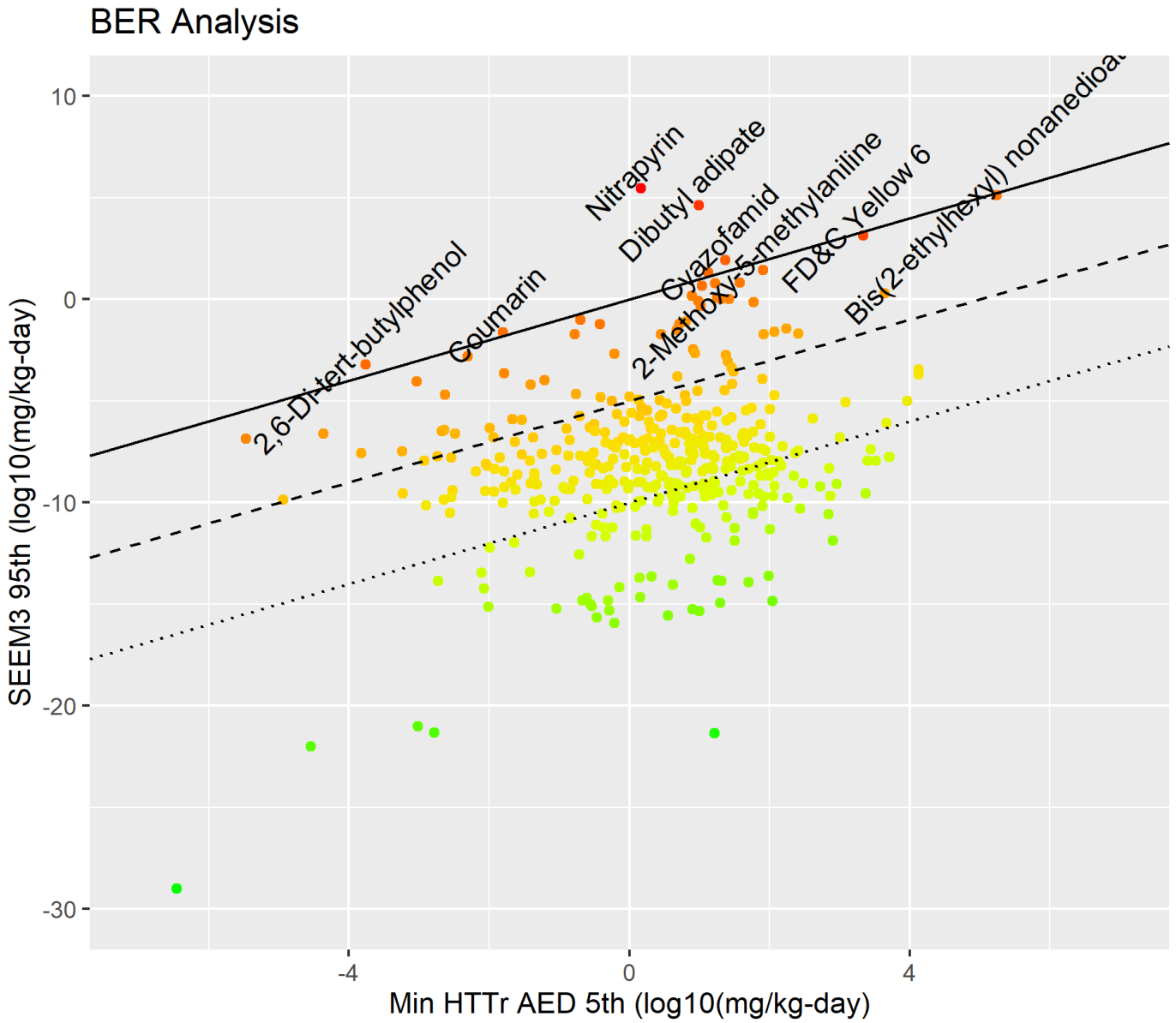




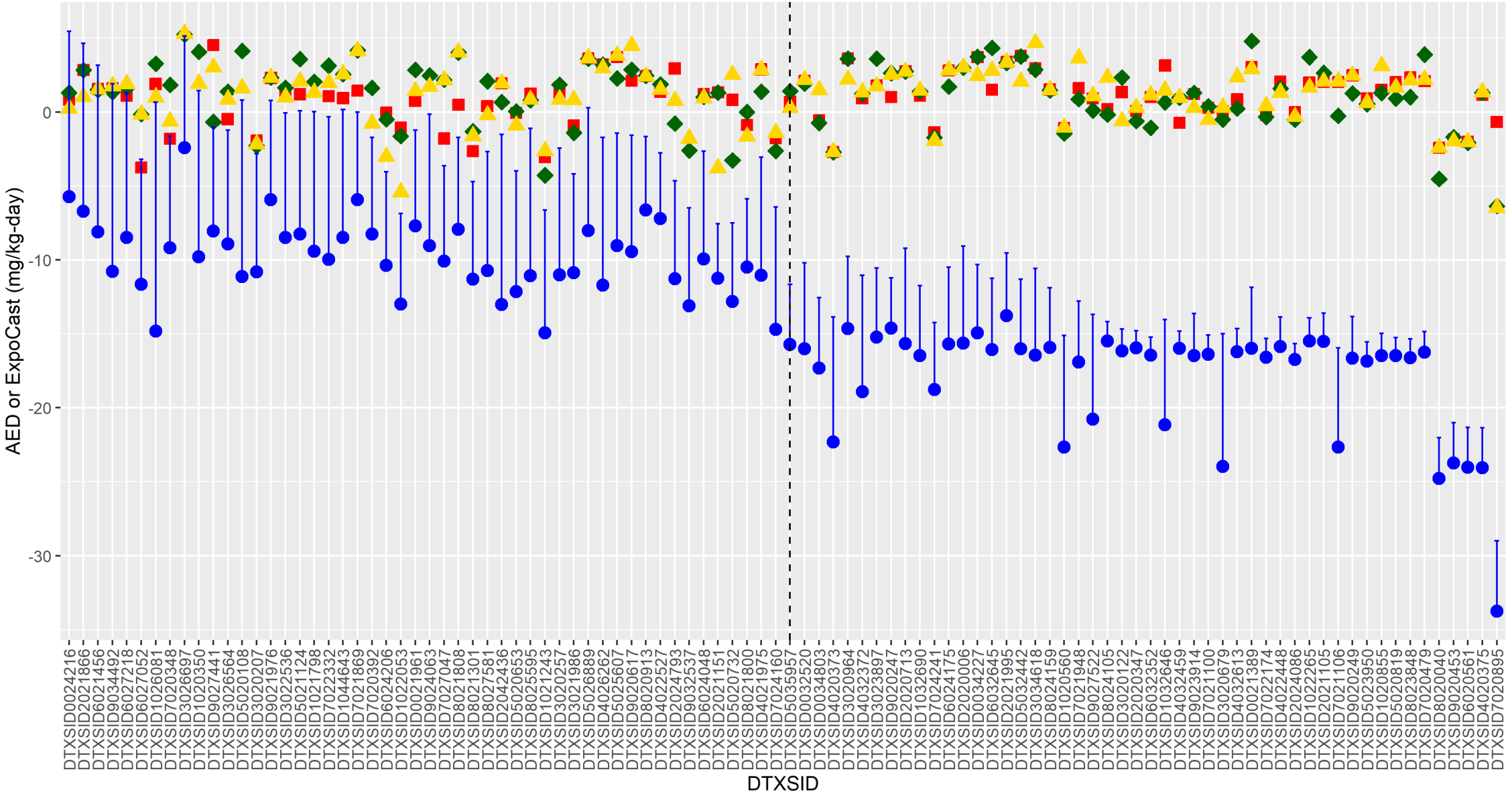


Bioactivity Exposure Ratio (BER) Analysis [1]

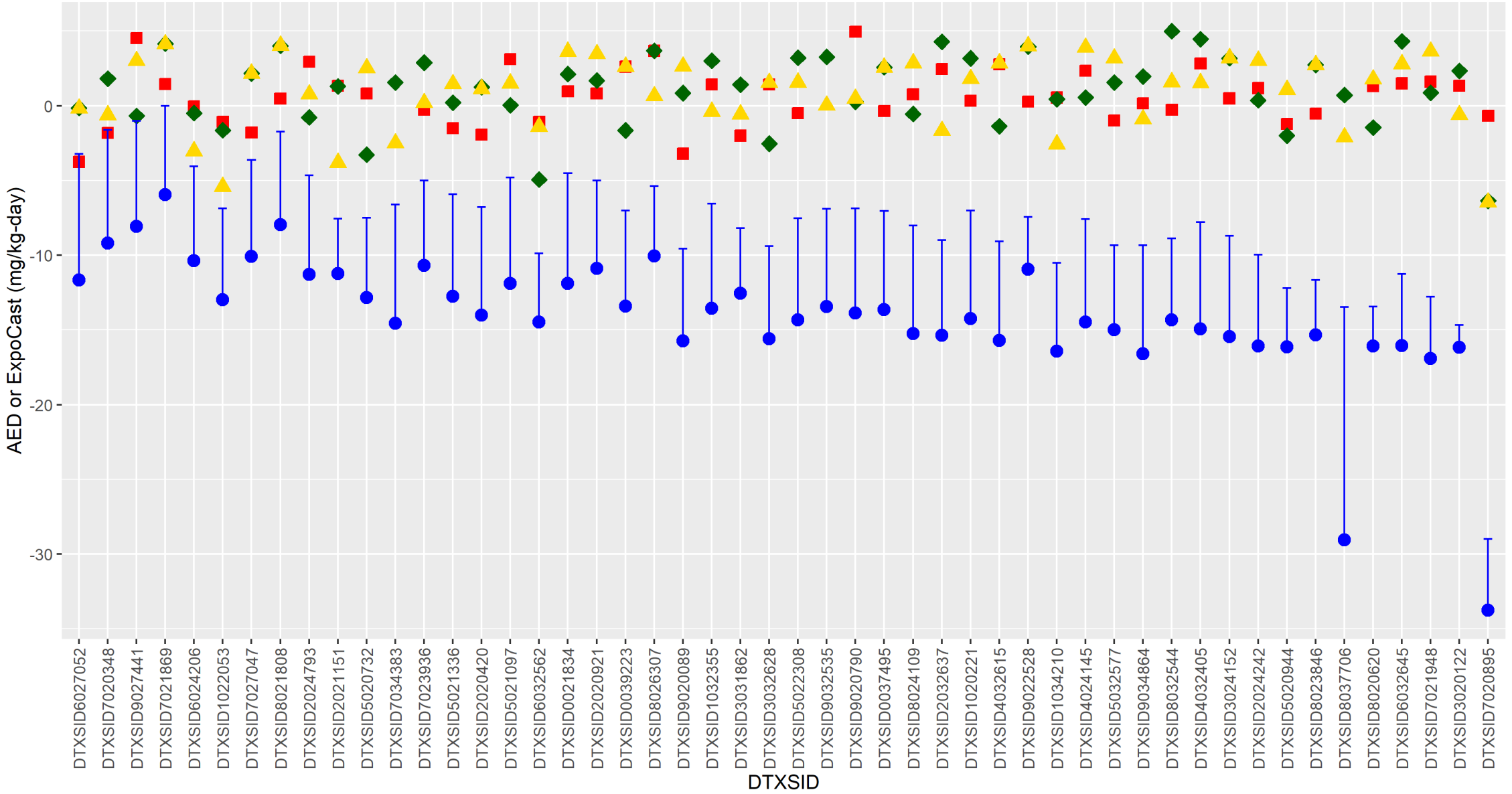




Bioactivity Exposure Ratio (BER) Analysis [2]



Bioactivity Exposure Ratio (BER) Analysis [3]



Summary and Conclusions

- **HTTr & HTPP Screening:** We have established robust and scalable laboratory and bioinformatics workflow for transcriptomics and phenotypic screening of environmental chemicals in human-derived cell lines.
- **Assay Reproducibility:** We have demonstrated a high degree of assay reproducibility for both HTTr and HTPP screening assays through the use of reference chemicals and standardized reference materials.
- **Bioactivity to Exposure Ratio:** Biological pathway/phenotype altering concentrations (BPACs) can be converted to administered equivalent doses (AEDs) and compared to human exposure predictions for chemical ranking and prioritization.
- **Comparability to ToxCast:** The AEDs derived from HTTr and HTPP assays are positive correlated with those that are derived from ToxCast HTS assays.
- **Future Work:** Expand the amount of biological space evaluated for environmental chemicals by screening in additional, complementary cell types.

Acknowledgements

Office of Research and Development (ORD)
Center for Computational Toxicology and Exposure (CCTE)

- **Johanna Nyffeler**
- Clinton Willis
- Rick Brockway
- Megan Culbreth
- Dan Hallinger
- Terri Fairley
- Ann Richard
- Kathy Coutros
- Maureen Gwinn
- Russell Thomas
- Katie Paul-Friedman
- Logan Everett
- Imran Shah
- Richard Judson
- Woody Setzer
- Grace Patlewicz
- Derik Haggard

PerkinElmer

- Joe Trask
- Dana Hanes
- Jim Hostetter

National Toxicology Program
U.S. Department of Health and Human Services

- Scott Auerbach
- Nisha Sipes

Ciome

- Ruchir Shah
- Jason Phillips

BioSpyder™

- Jo Yeakley
- Bruce Seligmann
- Joel McComb
- Pete Shepherd
- Milos Babic
- Dalia Gonzalez
- Kyle LeBlanc
- Garrett McComb