High-Throughput Transcriptomics (HTTr) Screening as a Component of NAMs-based Tiered Hazard Evaluation

Joshua A. Harrill

USEPA Center for Computational Toxicology and Exposure (CCTE)

U.S. Environmental Protection Agency – Japan Ministry of the Environment
14th Bilateral Meeting on Endocrine Disruption Test Methods Development
**June 9th, 2021**

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.

Outline

- **NAMs-Based Tiered Hazard Evaluation Approach**

- **High-Throughput Transcriptomics (HTTr)**
  - Assay Concept
  - Data Analysis Pipeline
  - Concentration Response Modeling of Transcriptomic Signatures

- **Results**
  - Comparison to ToxCast
  - Comparison to ER Model
  - Bioactivity / *In Vivo* Effect Value Ratio Analysis
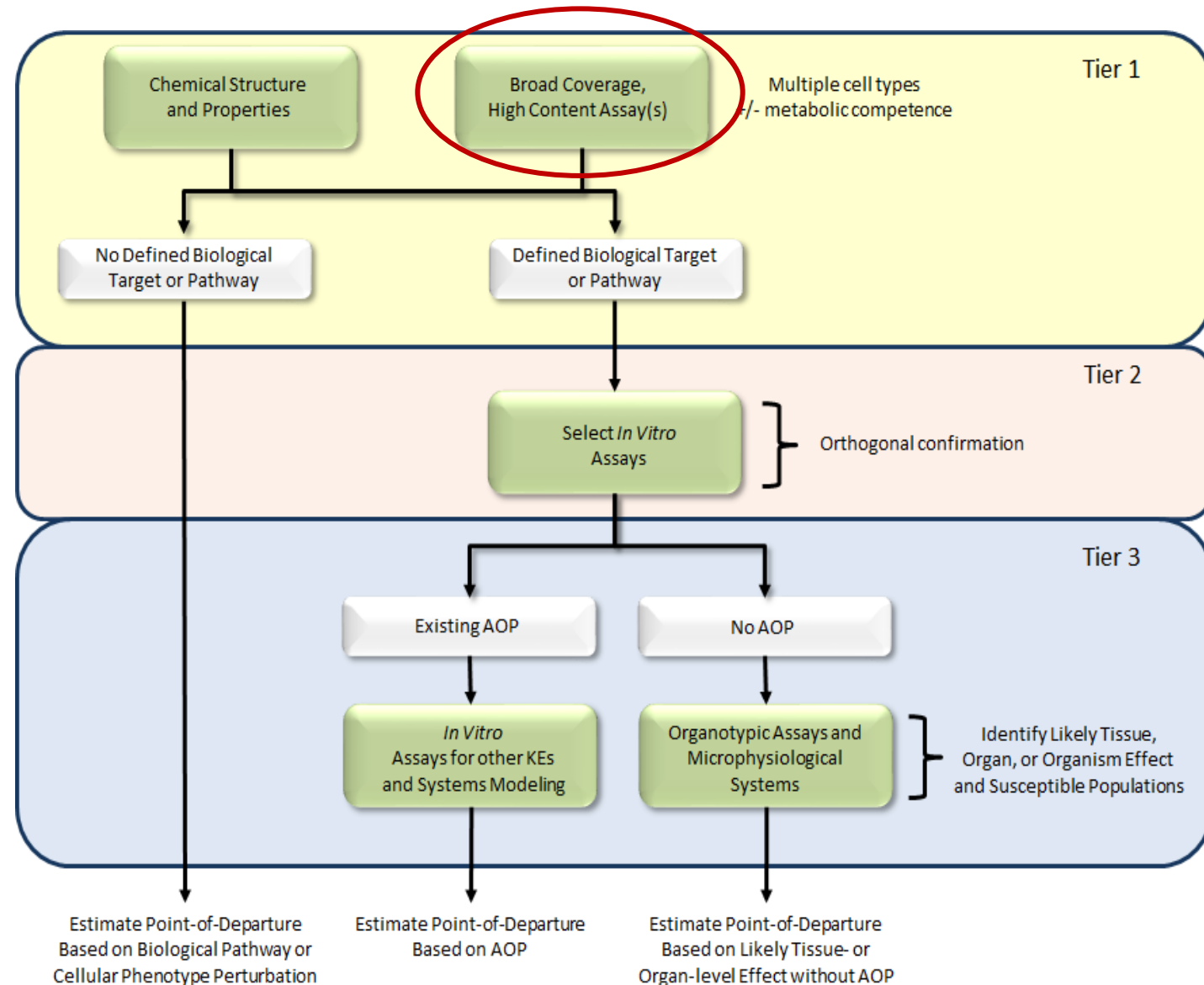
- **Current and Future Directions**

- **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.

- NAMs are a potential means to **reduce** the use of animals in toxicity testing and **accelerate** the pace of chemical risk assessment.

- 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
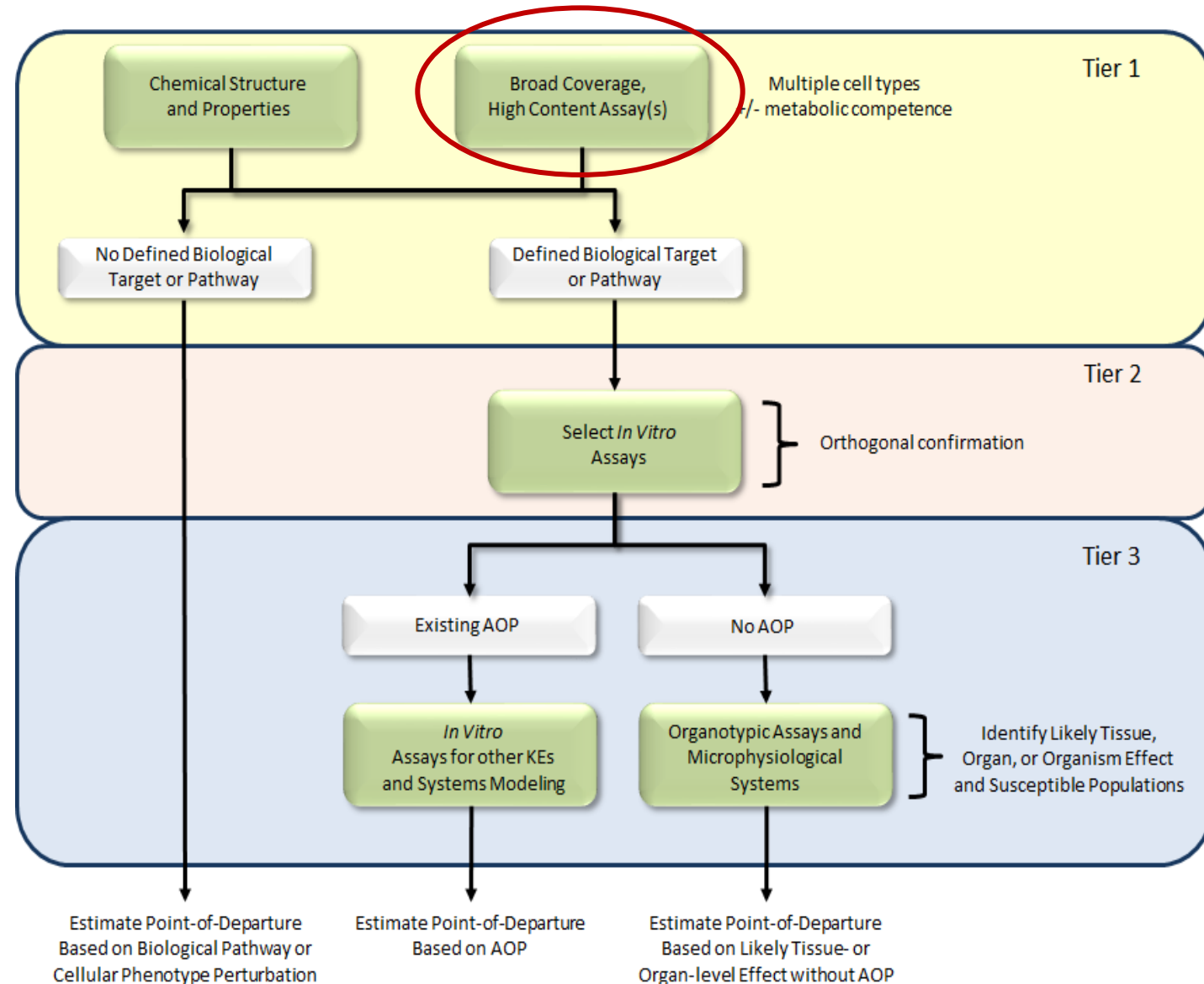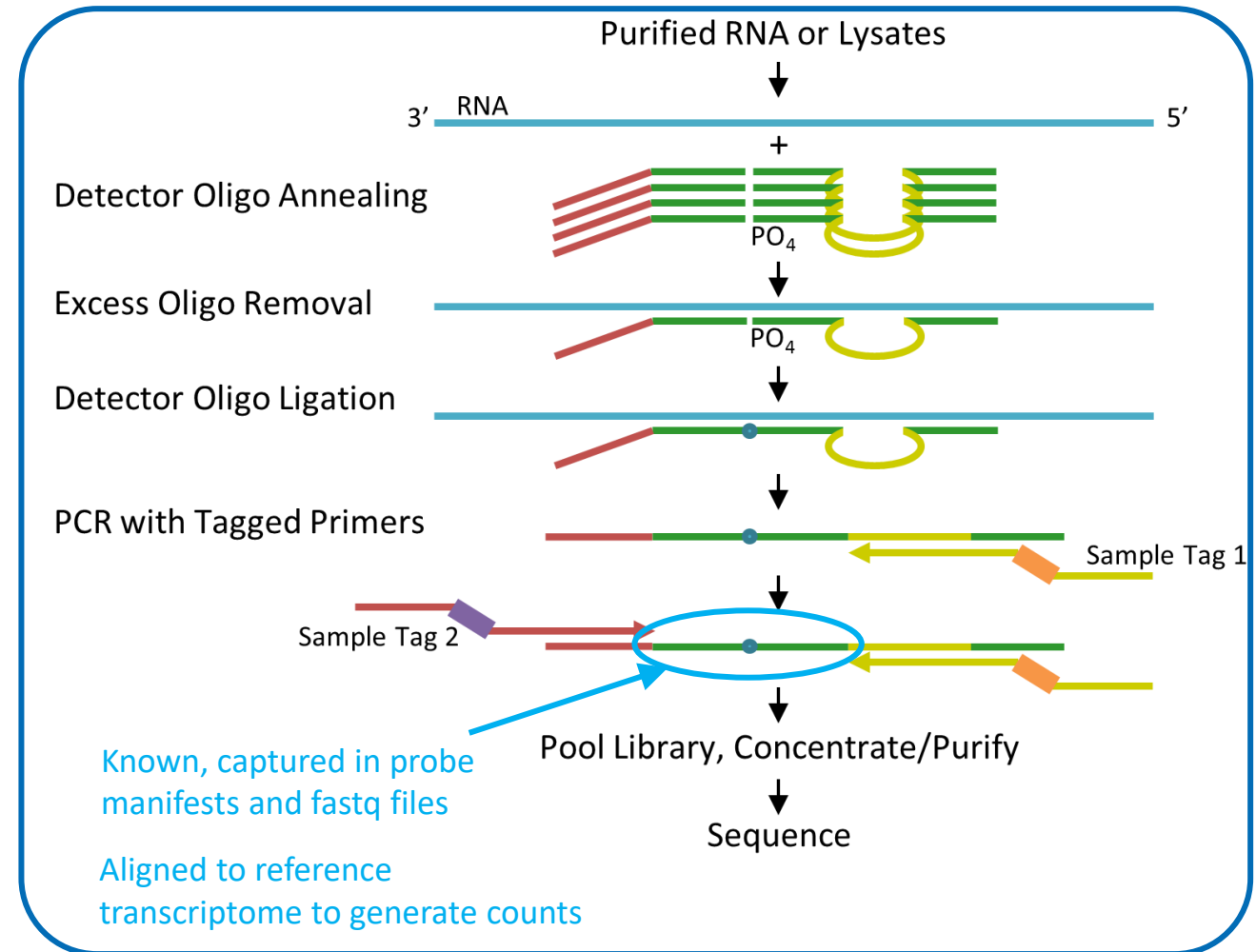
Tiered Hazard Evaluation Approach (2)

- To date, EPA has identified and implemented two HTP assays that meet this criteria.

- **High-Throughput Transcriptomics [HTTr]**

  - Whole Transcriptome TempO-Seq

- **High-Throughput Phenotypic Profiling [HTPP]**

  - Cell Painting

- Both methods are **complementary** to each other and can be used in many different human-derived cell types.

- EPA has established scalable laboratory and bioinformatics workflows for each assay.

Tier 1

Chemical Structure and Properties

Broad Coverage, High Content Assay(s) — Multiple cell types +/- metabolic competence

No Defined Biological Target or Pathway

Defined Biological Target or Pathway

Tier 2

Select *In Vitro* Assays — Orthogonal confirmation

Tier 3

Existing AOP

No AOP

In Vitro Assays for other KEs and Systems Modeling

Organotypic Assays and Microphysiological Systems — Identify Likely Tissue, Organ, or Organism Effect and Susceptible Populations

Estimate Point-of-Departure Based on Biological Pathway or Cellular Phenotype Perturbation

Estimate Point-of-Departure Based on AOP

Estimate Point-of-Departure Based on Likely Tissue- or Organ-level Effect without AOP

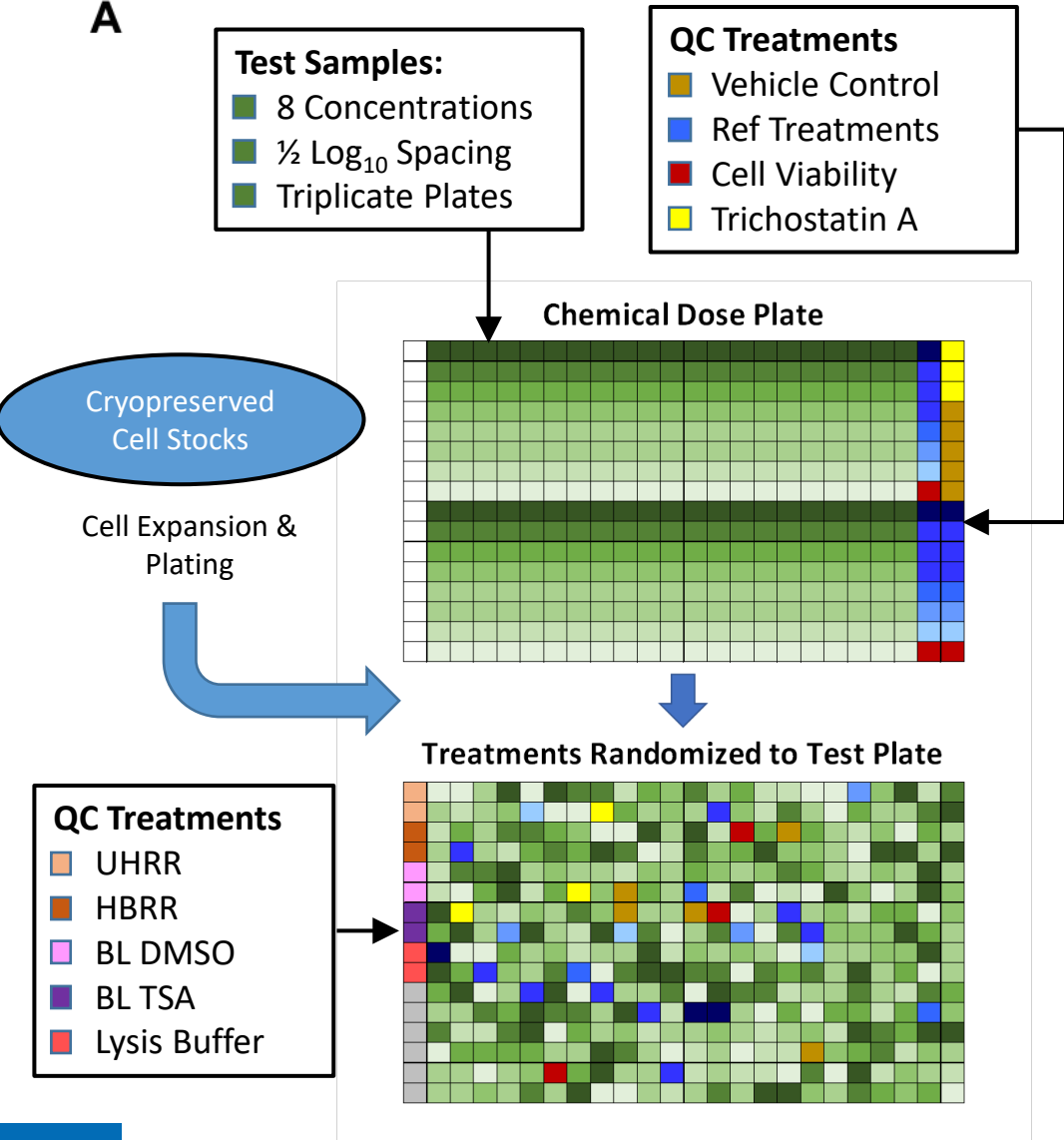# Templated Oligo with Sequencing Readout (TempO-Seq)

- 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

Purified RNA or Lysates

3' RNA 5'
+
Detector Oligo Annealing
PO_4

Excess Oligo Removal
PO_4

Detector Oligo Ligation

PCR with Tagged Primers
Sample Tag 1

Sample Tag 2

Known, captured in probe manifests and fastq files

Aligned to reference transcriptome to generate counts

Pool Library, Concentrate/Purify

Sequence
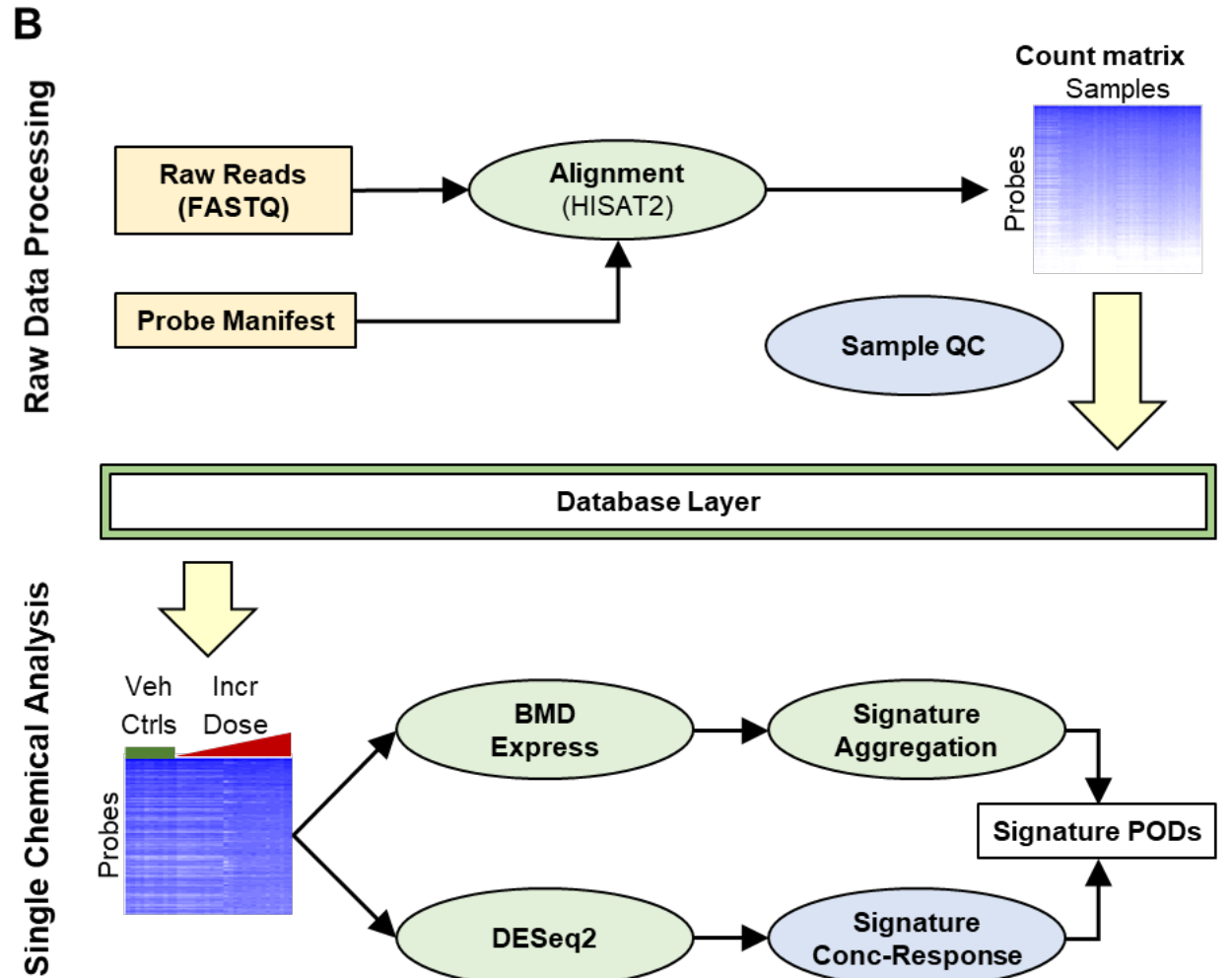
Yeakley et al., PLoS One. 2017 May 25;12(5):e0178302

HTTr Experimental Design and Bioinformatics Workflow

Concentration-Response Modeling of Gene Signatures (1)

- Understanding the results of change in gene expression for **10,000 – 20,000 genes** is difficult.

- Grouping genes into gene signatures aids in **data interpretation**.

- **Examples of signature types**:
  - Genes that are perturbed in diseased tissue vs. healthy tissue
  - Genes perturbed in individuals with congenital diseases vs. those without
  - Genes perturbed by gene knockdowns / knockouts
  - Genes perturbed by drugs or other chemicals with known (or unknown) mechanisms

- **Example use:**
  - If an unknown *chemical X* perturbs genes that are also perturbed by a well-characterized chemical with a specific mechanism of action, then one can infer the *chemical X* may affect the same molecular target(s).

- **CCTE signature collection:**
  - Compiled from many public sources (MSigDB[1], BioPlanet[2], CMAP[3] , DisGeNET[4]) → ~10,000 signatures
  - Each signature has been manually-assigned a "super target" class to aid in interpretation → ~1000 super targets
    - Disease groups (Immune, Cancer, etc.)
    - Biological organization groups (molecule, pathway, cell, tissue, organ, etc.)

[1] Liberzon et al., Bioinformatics. 2011 Jun 15;27(12):1739-40
[2] Huang et al., Front Pharmacol. 2019 Apr 26;10:445
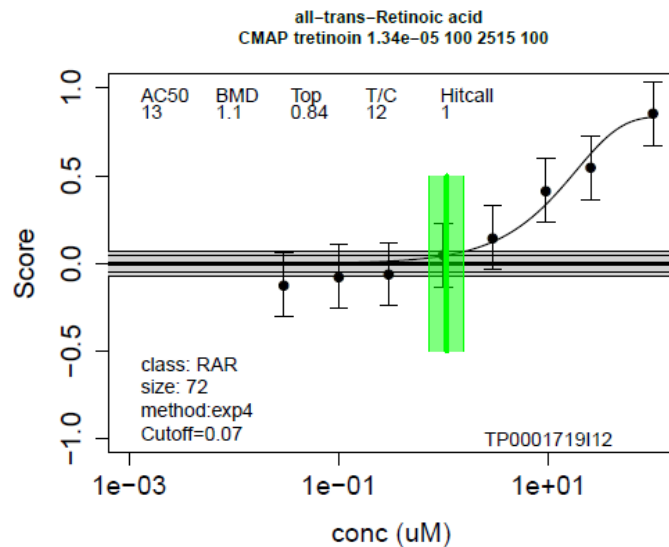[3] Subramanian et al., Science. 2006 Sep 29;313(5795):1929-35.
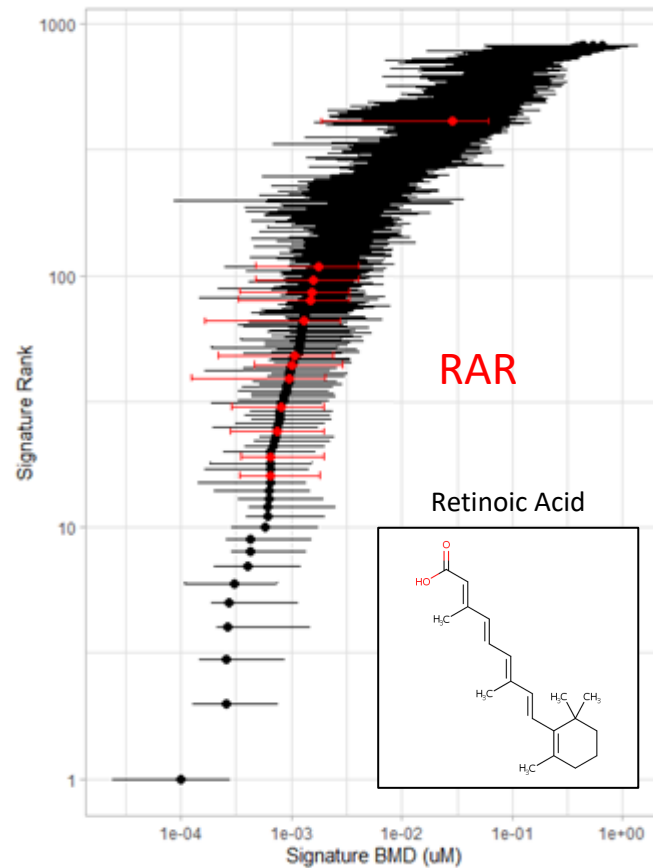[4] Pinero et al., Database (Oxford). 2015 Apr 15;2015:bav028

Concentration-Response Modeling (*tcplfit2*)

Signature-Level:
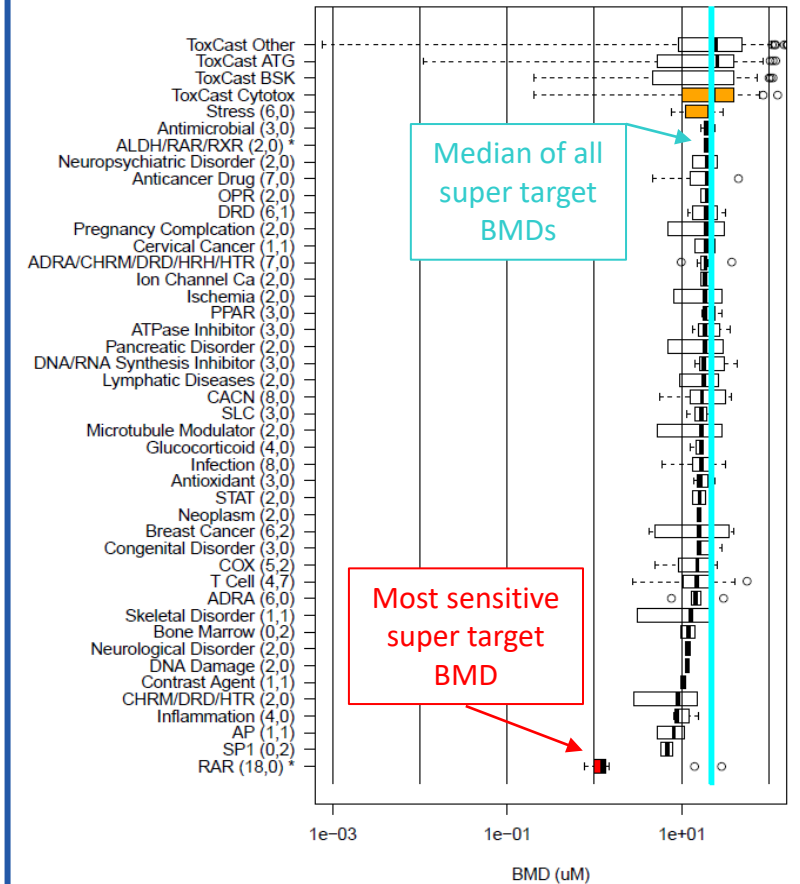- Benchmark Dose (BMD)
- Confidence Interval on BMD
- Hit Call Probability

Ranking of Signatures

RAR

Retinoic Acid

Signature Aggregation

Median of all super target BMDs

Most sensitive super target BMD

- Molecular PODs based on **biological pathway altering concentrations (BPACs)** may be derived in several ways.
- Most sensitive signature **OR** statistic based on distribution of active signatures (5th %ile) **OR** by target class.

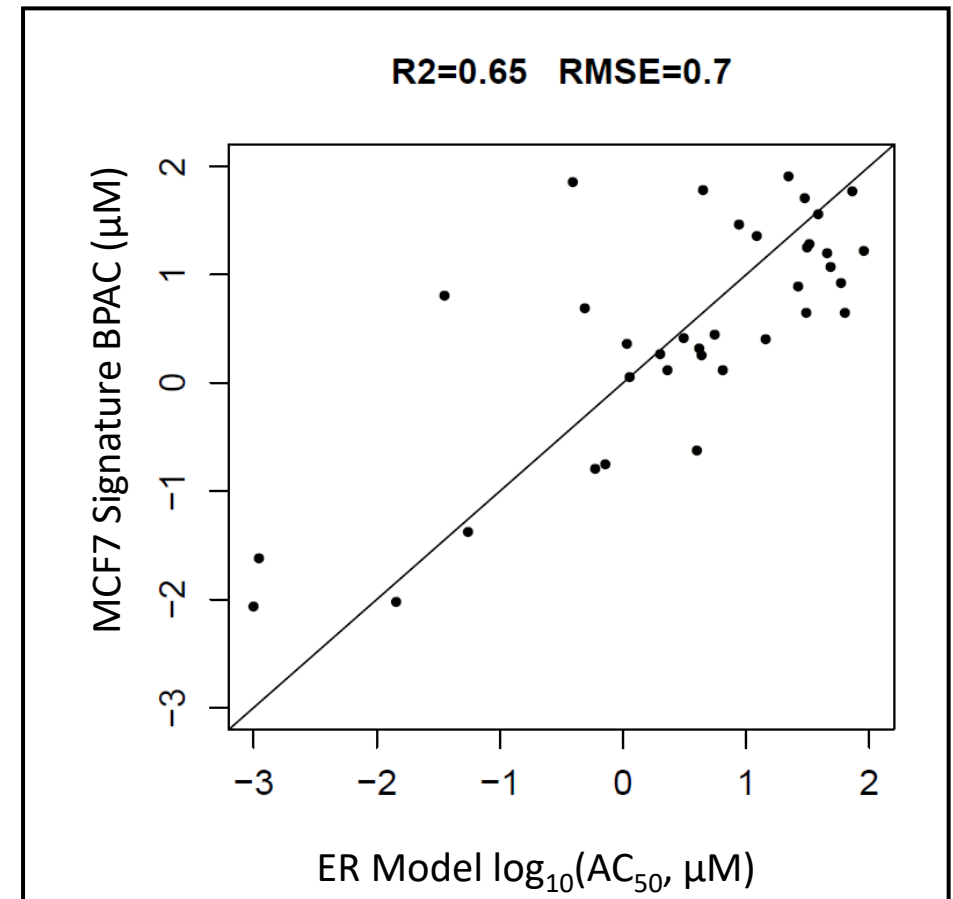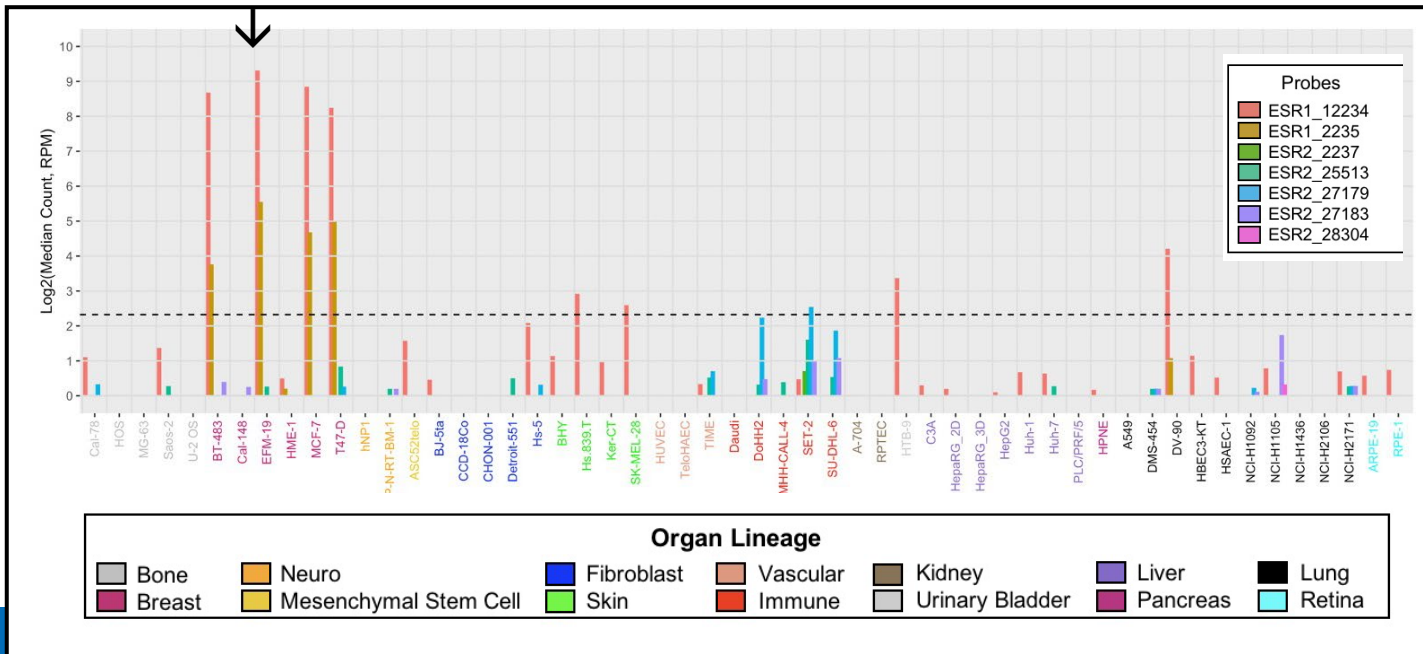# Comparison of Transcriptional BPACs to ToxCast BPACs
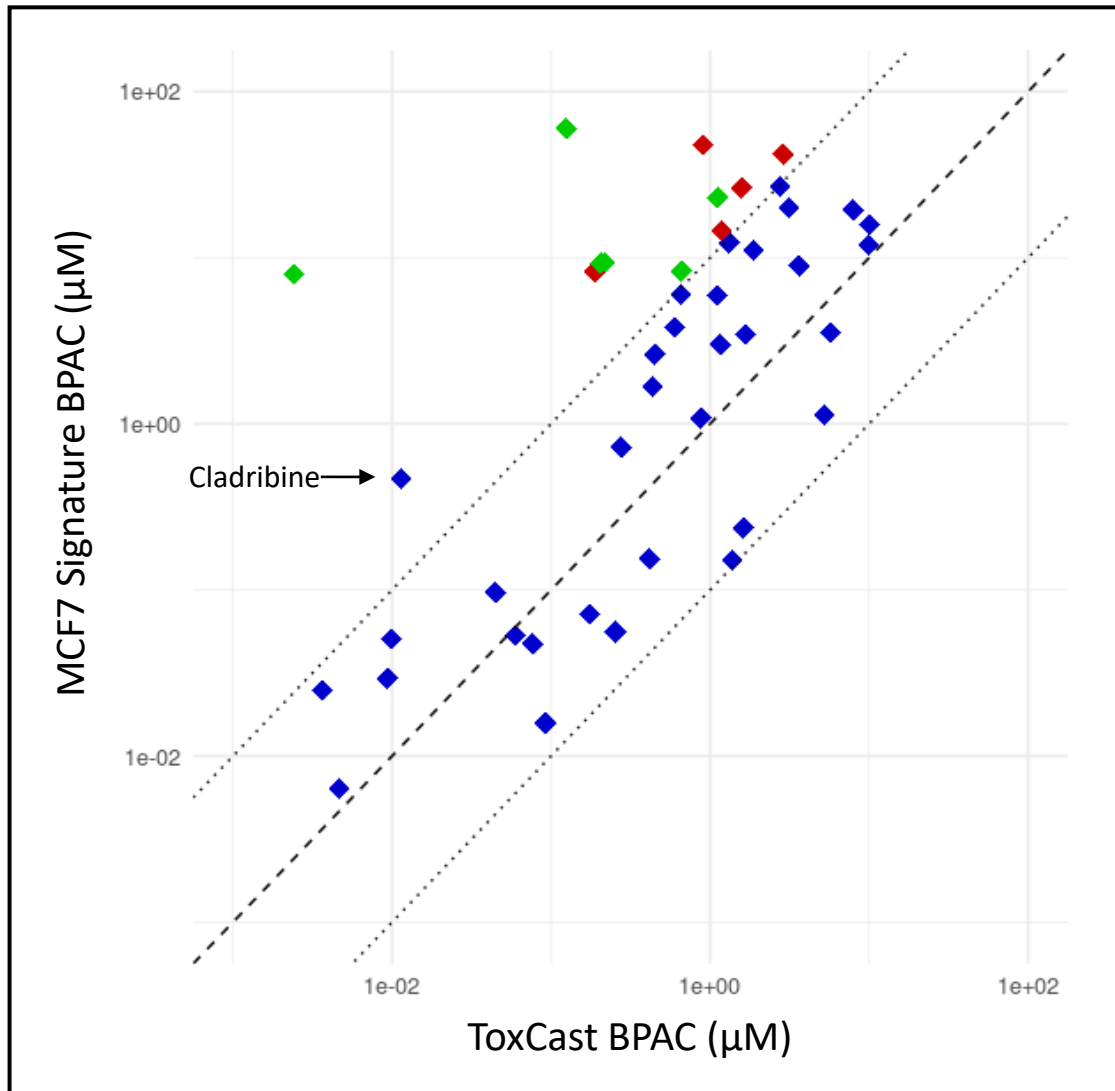
- Pilot study of 44 well-characterized ToxCast chemicals in MCF-7 cells *(Harrill et al., Toxicol Sci. 2021 Apr 27;181(1):68-89)*.

- Compare HTTr-derived PODs from MCF-7 cells to previous ToxCast HTS assay results *(Paul-Friedman et al., Toxicol Sci. 2020 Jan 1;173(1):202-225)*.

- Signature-based BPACs are highly concordant with ToxCast results for a majority of test chemicals in the pilot study.

  - 6 chemicals with targets that low/absent expression in MCF-7 cells.
  - 5 chemicals with off-target hit as most potent assay in ToxCast
  - Cladribine is a non-specific DNA synthesis inhibitor.

Harrill et al., Toxicol Sci. 2021 Apr 27;181(1):68-89.

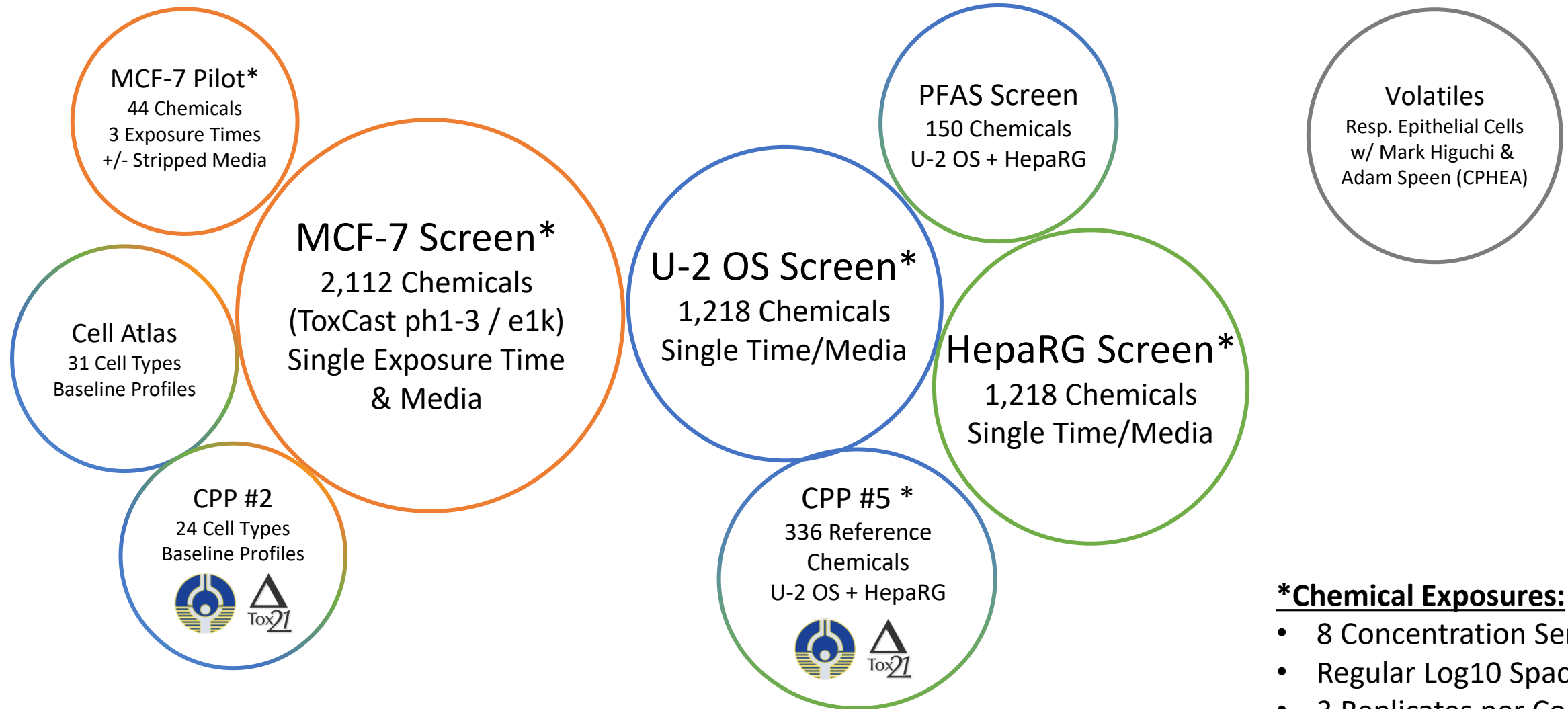# Signature Modeling of Estrogenic Chemicals
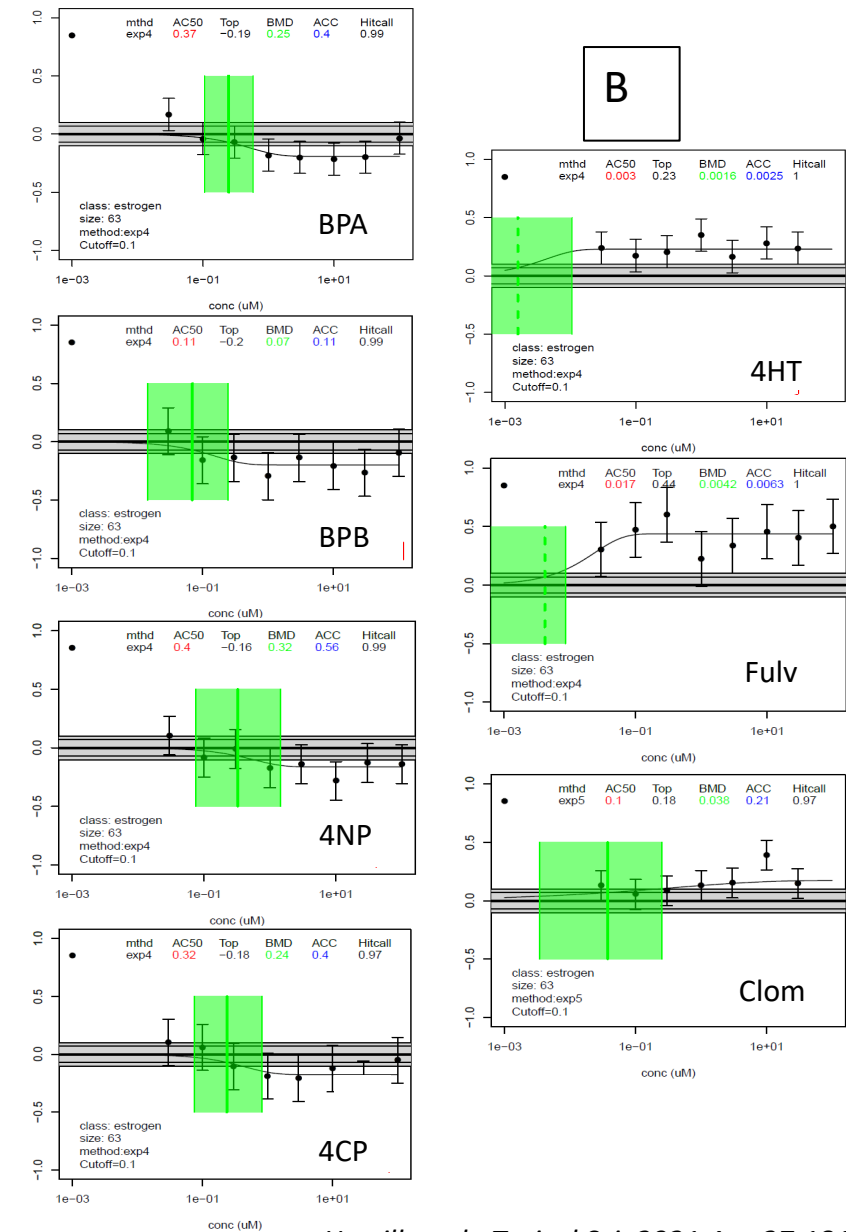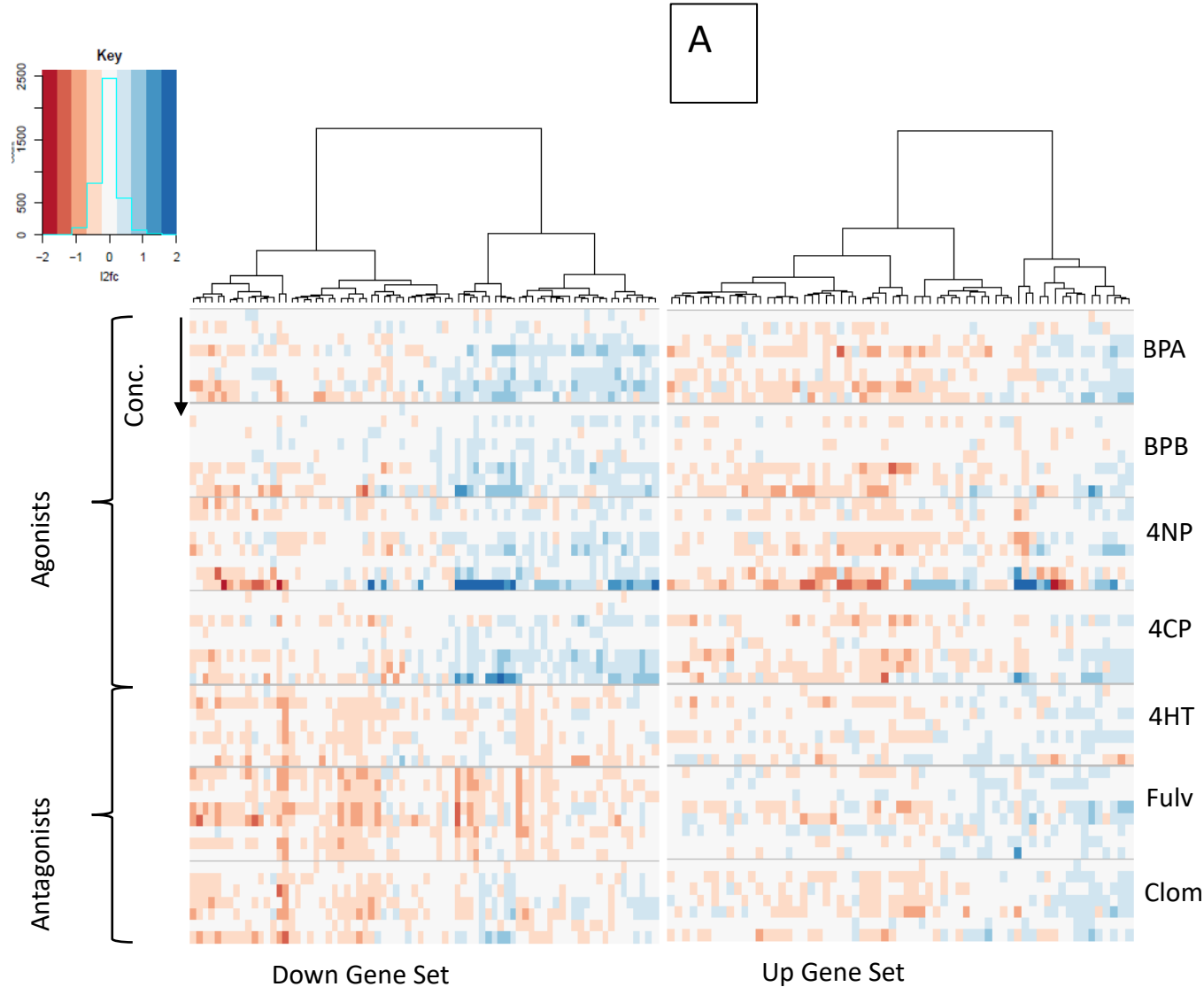
Comparison of Transcriptional BPACs to ER Model

- US EPA has developed a battery of 18 ToxCast assays to predict activity at the estrogen receptor *(Browne et al., Environ Sci Technol. 2015 Jul 21;49(14):8804-14)*

- Log_{10} AC_{50} values from the ToxCast ER model assays were compared to transcriptomic signature BPACs in MCF-7 cells for a collection of 37 estrogenic chemicals.

- Signature-based BPACs are concordant with ER model predictions. →

- Estrogen receptor is also abundantly expressed in MCF-7 cells (and other breast-derived cell lines)
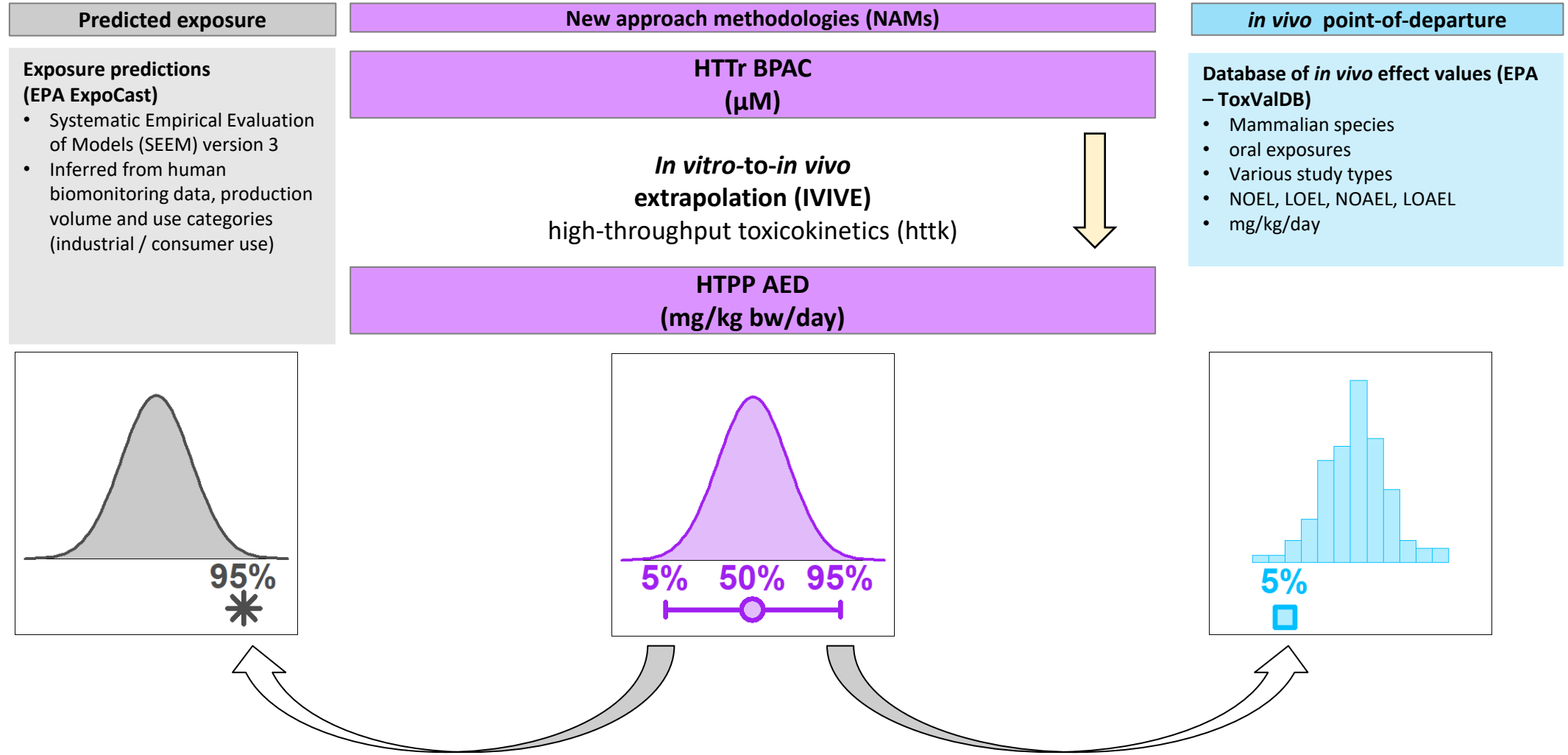

Harrill et al., (unpublished). DO NOT CITE OR QUOTE

High Throughput Transcriptomics (HTTr) Data Landscape

MCF-7 Pilot*
44 Chemicals
3 Exposure Times
+/- Stripped Media

MCF-7 Screen*
2,112 Chemicals
(ToxCast ph1-3 / e1k)
Single Exposure Time
& Media

PFAS Screen
150 Chemicals
U-2 OS + HepaRG

Volatiles
Resp. Epithelial Cells
w/ Mark Higuchi &
Adam Speen (CPHEA)

Cell Atlas
31 Cell Types
Baseline Profiles

U-2 OS Screen*
1,218 Chemicals
Single Time/Media

HepaRG Screen*
1,218 Chemicals
Single Time/Media

CPP #2
24 Cell Types
Baseline Profiles

CPP #5 *
336 Reference
Chemicals
U-2 OS + HepaRG

***Chemical Exposures:**
- 8 Concentration Series
- Regular Log10 Spacing
- 3 Replicates per Conc
- Randomized Plate Layout

Slide courtesy of Logan Everett

In Vitro to *In Vivo* Extrapolation (IVIVE) Using High-Throughput Toxicokinetic (httk) Modeling

BPAC: Biological Pathway Altering Concentration
AED: Administered Equivalent Dose

Bioactivity / *In Vivo* Effect Value Ratio Analysis

APCRA Chemicals

- **Negative ratios** indicate that AEDs derived from HTP NAMs molecular PODs are **conservative** surrogates for traditional *in vivo* PODs.

- When cell lines are considered individually, **~66-68%** of chemicals had negative ratios.

- When considered in combination, the number and percentage of chemicals with negative ratios **increased (82.3 %)**.

- Paul-Friedman et al. (2020):
  - Using ToxCast, **89 %** of APCRA chemicals had negative ratios.

- Positive ratios observed for several organophosphate and carbamate pesticides.

Harrill et al., (unpublished). DO NOT CITE OR QUOTE
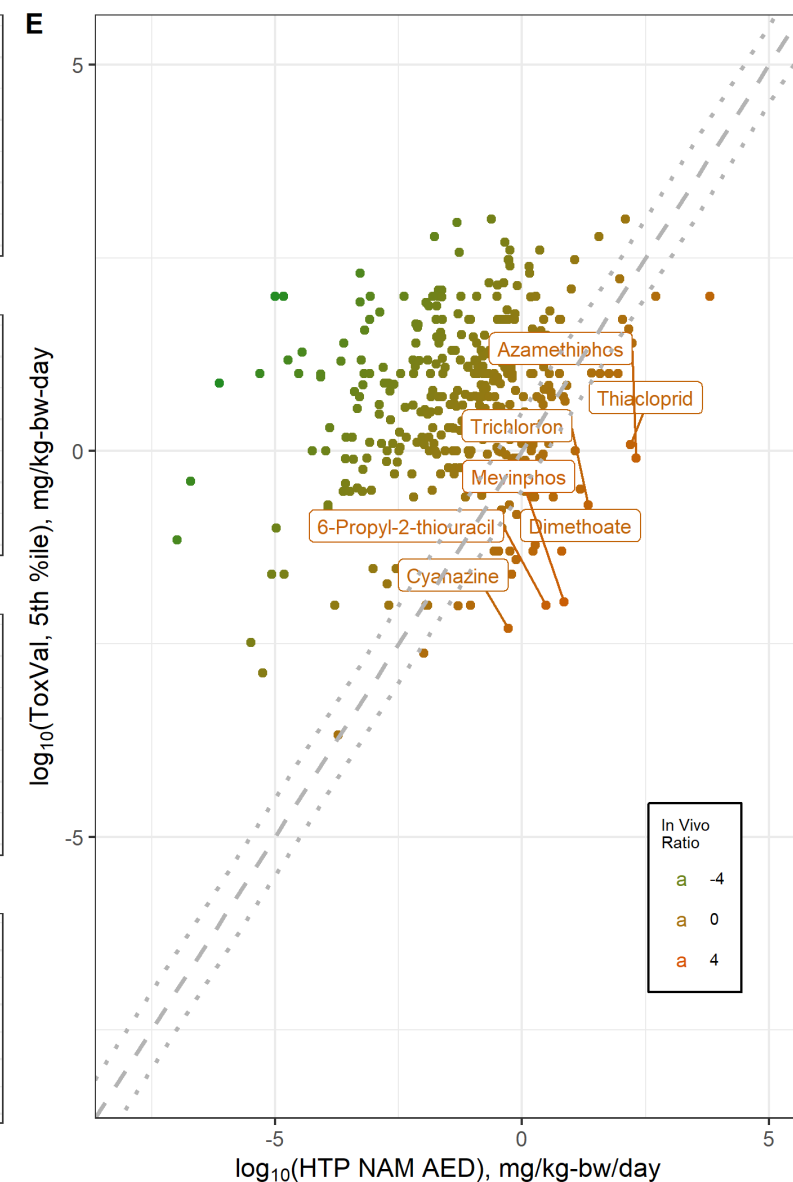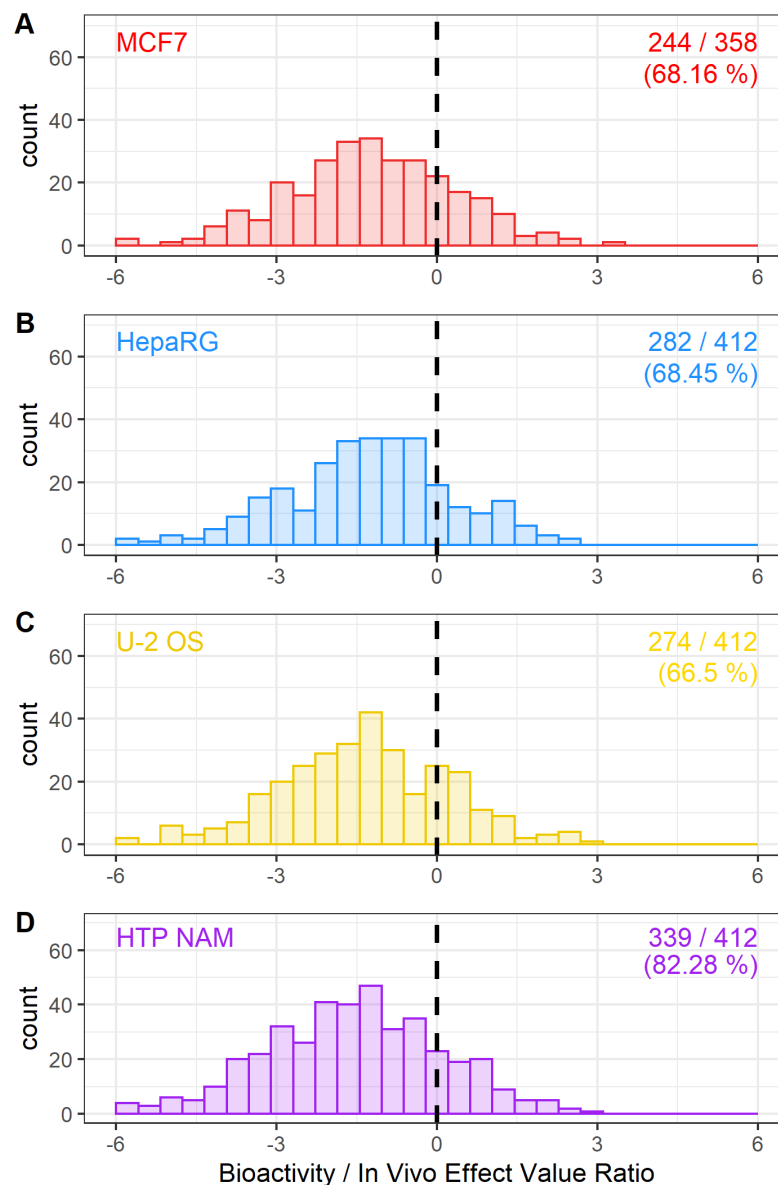
Future Directions

- **Expand chemical space**
  - Screen additional chemicals in accordance with programmatic needs / goals

- **Expand biological space**
  - Continue screening a subset of chemicals through many biologically diverse cell lines

- **Refine signature concentration-response modeling approach**
  - Reduce redundancy in signature collection
  - Continued curation of target annotation

- **Refine methods for putative target prediction & confirmation**
  - Integration with other NAM's data streams
  - Machine learning approaches
  - Bioactivity confirmation within tiered hazard evaluation framework

Acknowledgements

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

- Clinton Willis
- Johanna Nyffeler
- Rick Brockway
- Megan Culbreth
- Dan Hallinger
- Terri Fairley

- Ann Richard
- Kathy Coutros
- Maureen Gwinn
- Russell Thomas

- Logan Everett
- Imran Shah
- Richard Judson
- Derik Haggard
- Thomas Sheffield
- Joseph Bundy
- Woody Setzer

- Katie Paul Friedman
- John Wambaugh

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

- Scott Auerbach

Bio Spyder™

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