EPA United States Environmental Protection Agency

Integrating High Throughput Transcriptomics into a Tiered Framework to Prioritize Chemicals for Toxicity Testing

Jesse Rogers, PhD
ORISE Postdoctoral Fellow, Biomolecular and Computational Toxicology Division
rogers.jesse@epa.gov

Office of Research and Development
Center for Computational Toxicology and Exposure

February 22, 2023
ASCCT-ESTIV Webinar

United States Environmental Protection Agency

Disclaimer

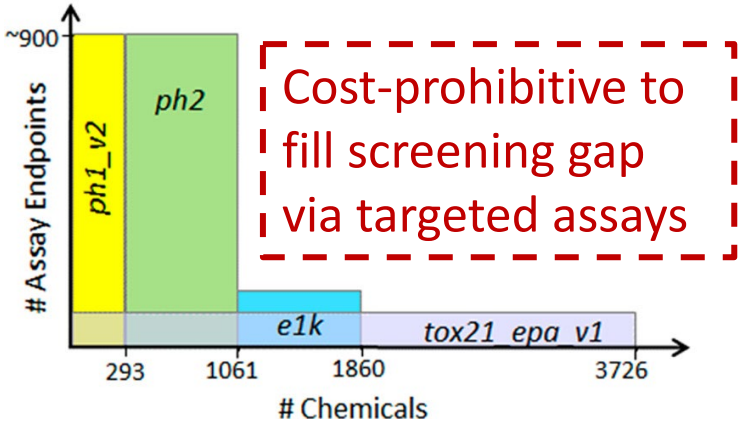
The views expressed in this presentation are those of the authors and do not necessarily represent the views or policies of US EPA.

Addressing Gaps in Chemical Toxicity Testing

Toxicity Testing in the 21st Century (NRC 2007)

- Shift from traditional animal-based toxicity testing to New Approach Methodologies (NAMs) and predictive toxicology

Testing Phase	Chemical Set	Unique Chemicals	Assay Endpoints
ToxCast Phase I	*ph1_v1*	310	~700
ToxCast Phase II	*ph1_v2*	293	~200
	ph2	768	~900
	e1k	799	~50
Tox21	*tox21_epa_v1*	3726	~80



Cost-prohibitive to fill screening gap via targeted assays

US EPA ToxCast program (Dix et al. *Toxicol Sci* 2007)

- Broad bioactivity profiling of chemicals via high-throughput screening (HTS) assays
- Limited biological target coverage, reduced xenobiotic metabolism *in vitro* (Rice et al. *Environ Health Perspect* 2013)

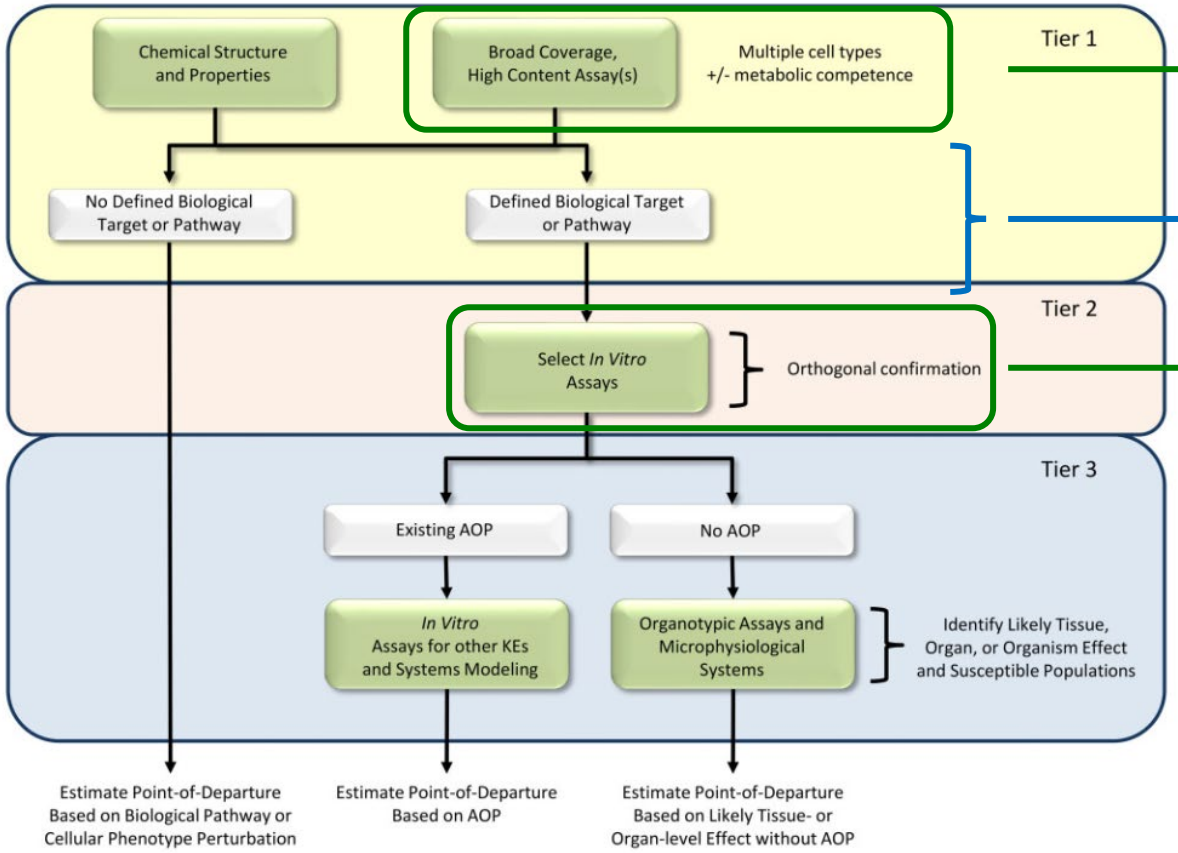
Next Generation Blueprint for Hazard Evaluation: Integrate multiple assay technologies into a single framework for efficient hazard screening (Thomas et al. *Toxicol Sci* 2019)

Slide adapted from Logan Everett

Richard *Chem Res Toxicol* 2016

Integrating Data Streams to Improve Scientific Confidence in NAMs

Tiered hazard evaluation framework: investigate potential mechanisms-of-action (MoAs) via high-throughput screening platforms and link verified chemicals to likely adverse outcomes (Thomas et al. *Toxicol Sci* 2019)



High-Throughput Transcriptomics (HTTr)
High-Throughput Phenotypic Profiling (HTPP)

?

Targeted High-Throughput Screening Assays (ToxCast)

Computational Needs:

- Derive high-confidence MoAs from transcriptomic NAMs
- Develop criteria to prioritize chemicals for key hazards based on Tier 1-2 NAMs

Thomas *Toxicol Sci* 2019

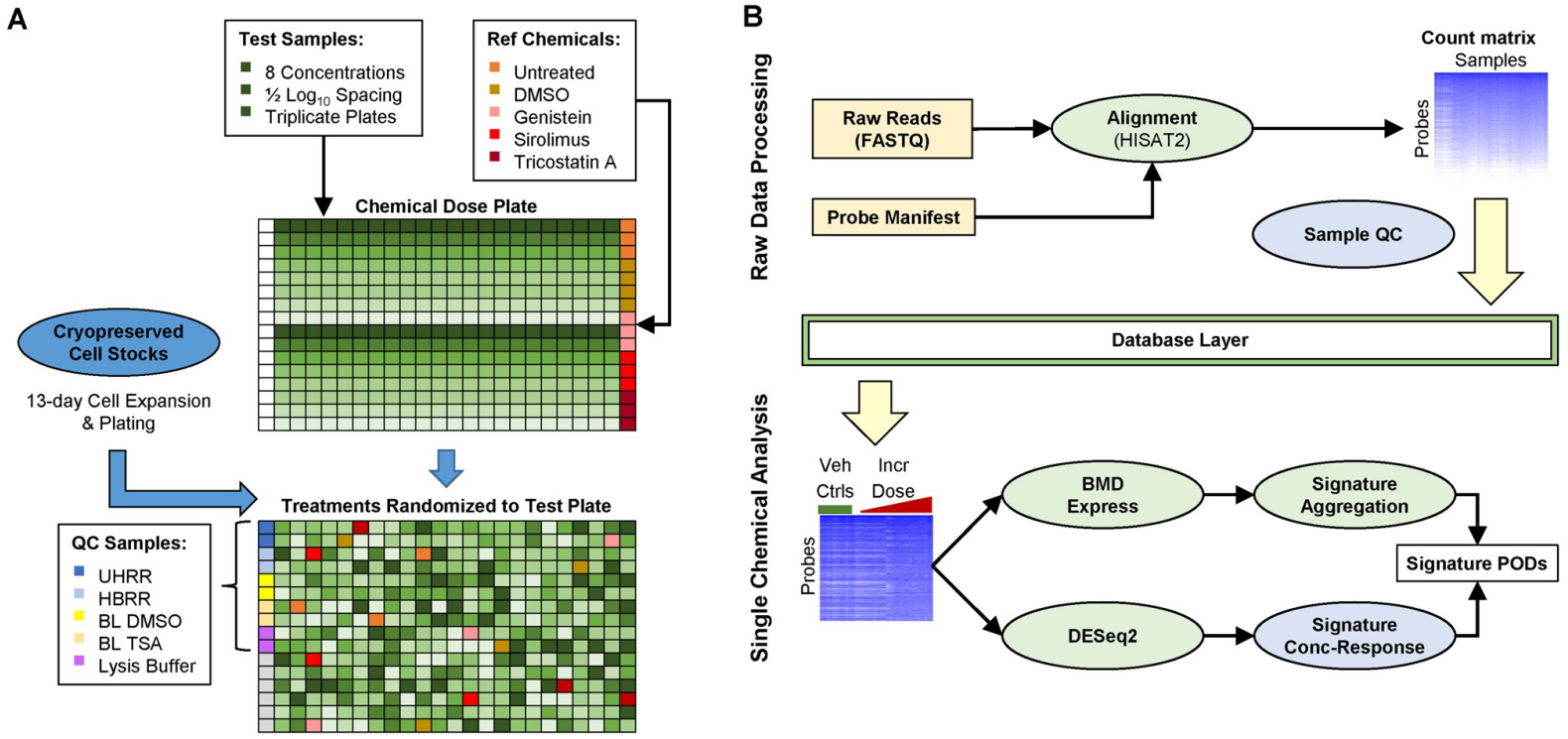
EPA United States Environmental Protection Agency

Project Outline

- Develop HTTr Signatures for MoA-Specific Activity
- Define Tiered Framework for Chemical Prioritization
- Apply Framework to Retrospective Tier 1-2 Screening Data
- Identify Candidates for Prospective Tier 2 Assessment
- Conclusions and Next Steps

High-Throughput Transcriptomics for Chemical Screening

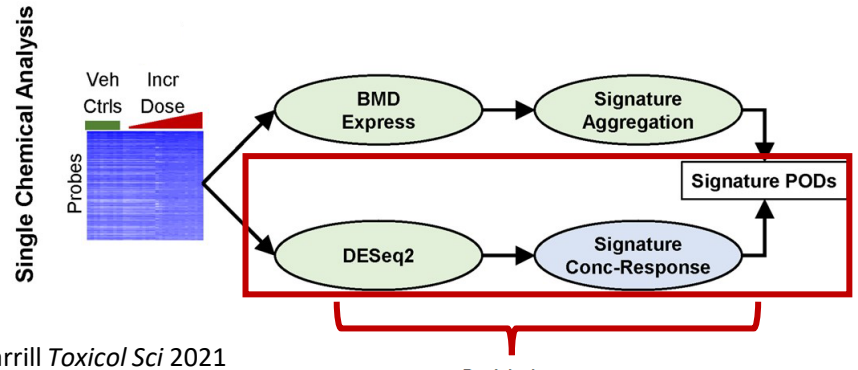
- TempO-Seq : Next-gen sequencing of >20,000 probes hybridized to expressed transcripts (Yeakley *et al. PLoS One* 2017)
- Up to 1,387 chemicals screened in multi-concentration format for multiple cell lines:
  - MCF7 Breast Carcinoma Cells (Harrill *et al. Toxicol Sci* 2021)
  - **U-2 OS Osteosarcoma Epithelial Cells** (Bundy *et al.* In Prep)
  - **HepaRG Hepatic Progenitor Cells** (Shah *et al.* In Prep)



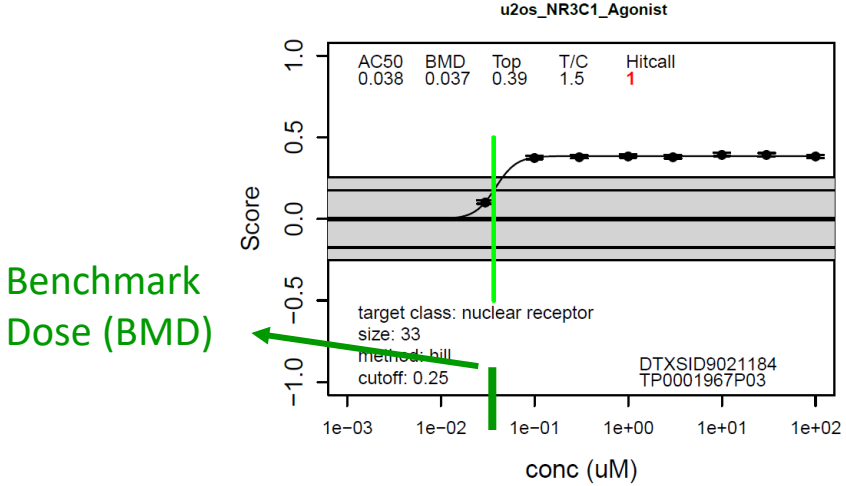
Harrill *Toxicol Sci* 2021

MoA Identification from Transcriptomic Data Streams

- Single-sample gene set enrichment analysis of compiled signatures (Barbie *et al. Nature* 2009)
- Concentration-response profiling of enrichment scores via *tcplfit2* (Sheffield *et. al. Bioinformatics* 2022)



Harrill *Toxicol Sci* 2021



Catalog of >11,000 public gene set signatures with toxicological relevance, annotated for known molecular targets:

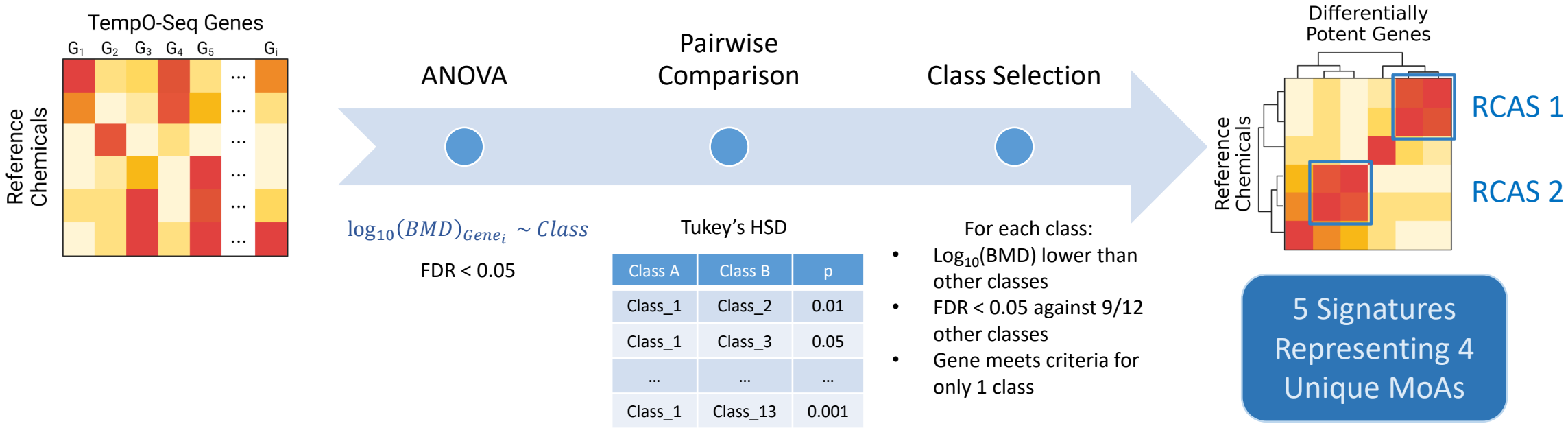
- **Bioplanet** (Huang, *et al. Front Pharmacol* 2019)
- **CMap** (Subramanian, *et al. Cell* 2017)
- **DisGeNET** (Pinero, *et al. Database* 2015)
- **MSigDB** (Liberzon, *et al. Cell Syst* 2015)

- Some public signatures may not be well-suited for probing MoAs in current assay
  - Cell lines used for derivation
  - Methods used for development, e.g. KEGG/Reactome
- ***Data-driven signatures may improve assay translation*** by profiling gene expression related to molecular initiating events

Data-Driven Signature Development Identifies Uniquely-Potent Features

Reference Class Associated Signatures (RCAS): gene sets uniquely potent for individual MoAs identified via univariate strategy

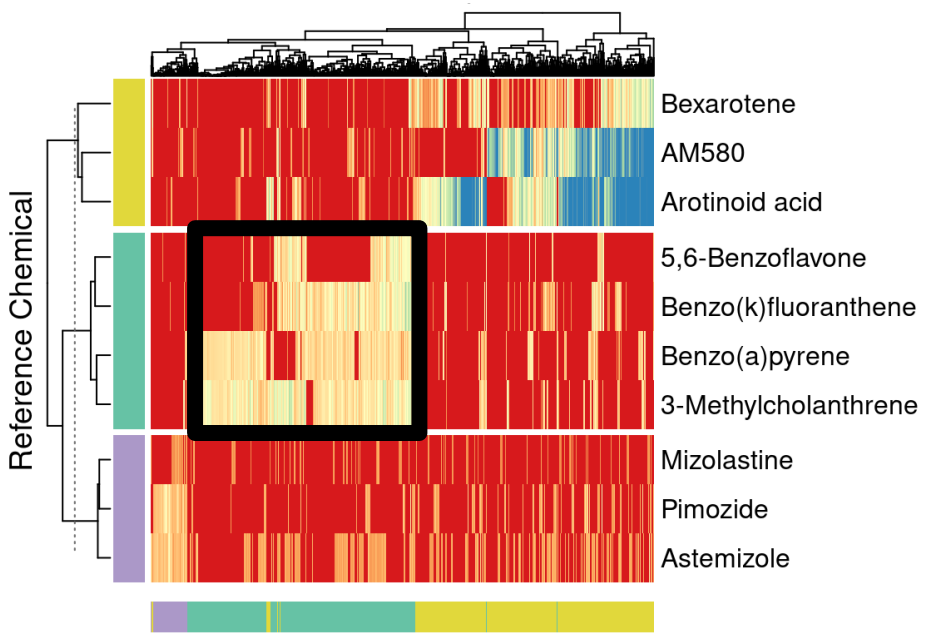
- Reference chemicals identified via *RefChemDB*: automated mining of literature databases for chemical-molecular target interactions (Judson et. al. *ALTEX* 2019)



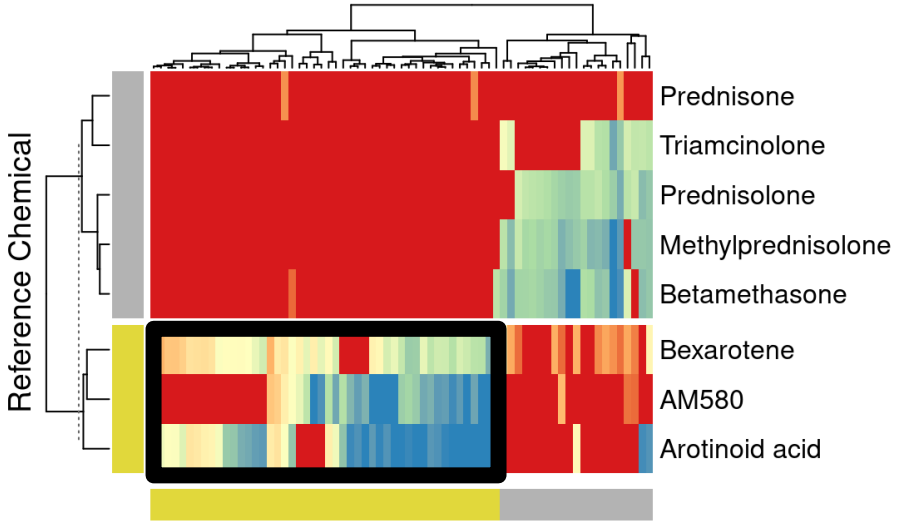
RCAS Gene Potencies Reveal Distinct Patterns by MoA

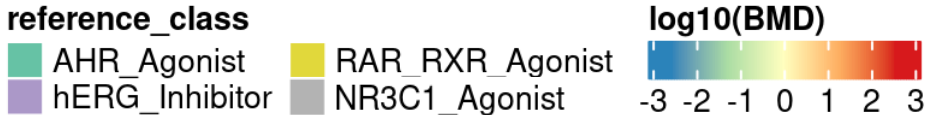
- Reference chemicals **annotated for same MoA** as signature demonstrate activity at low concentrations
- Reference chemicals annotated for other MoAs compared to signature show activity at high concentrations or no concentration-responsiveness

HepaRG Gene Clustering: 1173 genes identified



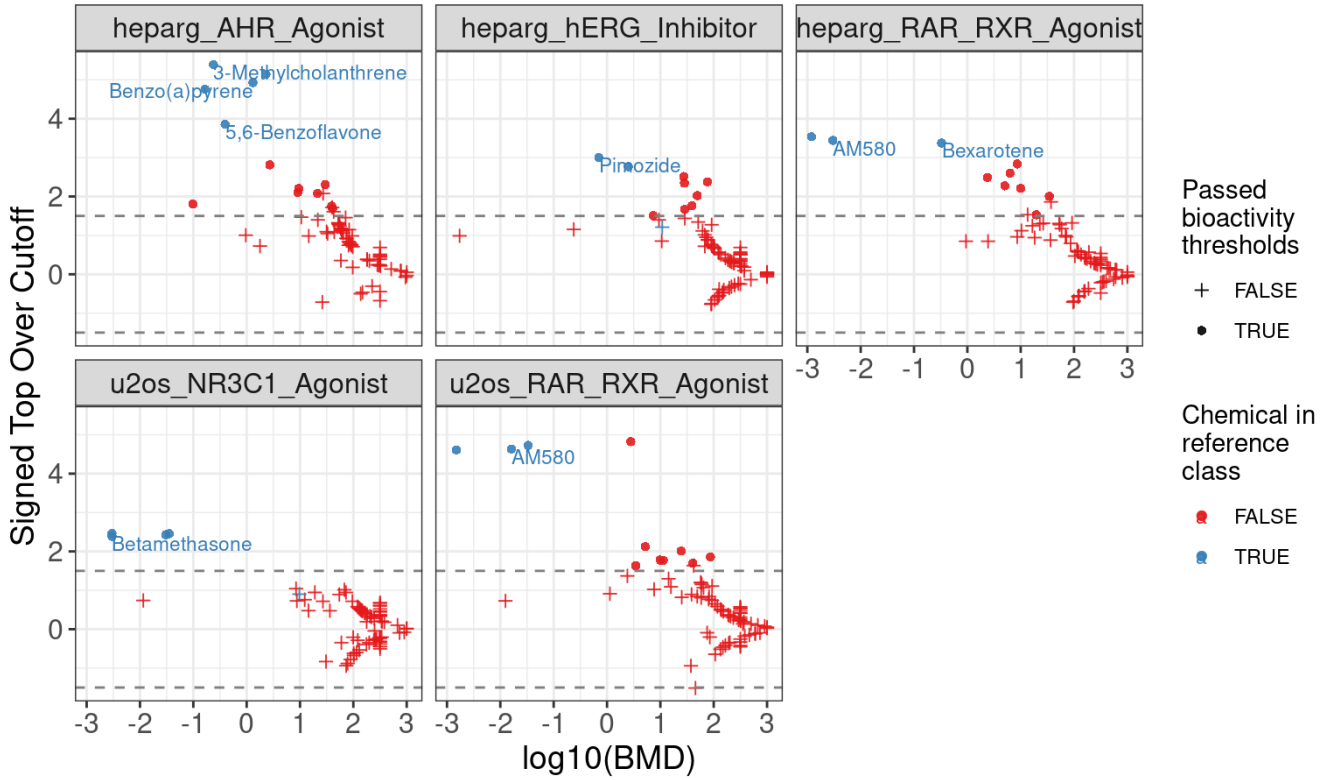
U-2 OS Gene Clustering: 69 genes identified





Efficacy and Potency for RCAS are Greatest for Matching Reference Chemicals

- Concentration-response modeling of reference signatures via CompTox-httrpathway package (https://github.com/USEPA/CompTox-httrpathway)
  - Enrichment scores estimated via ssGSEA (Barbie *et. al. Nature* 2009)
  - BMDs estimated from normalized enrichment scores via tcplfit2 (Sheffield *et. al. Bioinformatics* 2022)
- Signature bioactivity determined via thresholding of confidence and efficacy metrics:
  - Curve-fit confidence: hitcall ≥ 0.9
  - Efficacy: top over cutoff ≥ 1.5



In-class chemicals: low BMD, high efficacy

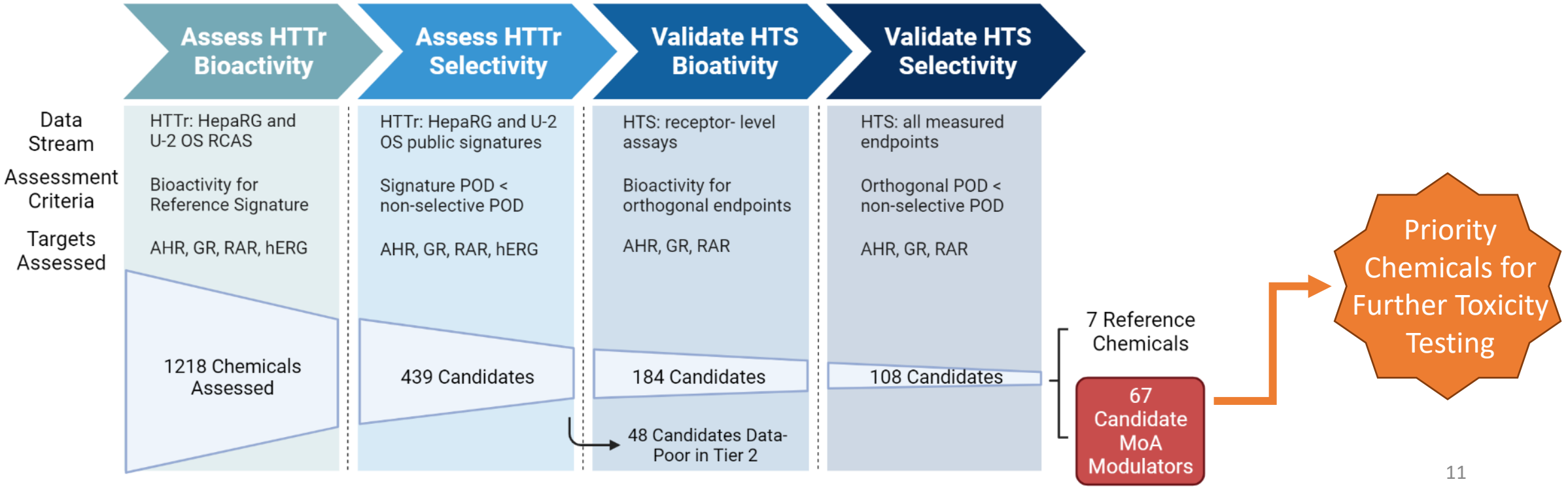
Out-of-class chemicals: high BMD, low efficacy

Chemicals annotated for each target passed threshold criteria for related signature, and few chemicals negative for each target passed criteria (except U2OS-NR3C1, in which none passed)

Integration of Transcriptomics into Chemical Prioritization Framework

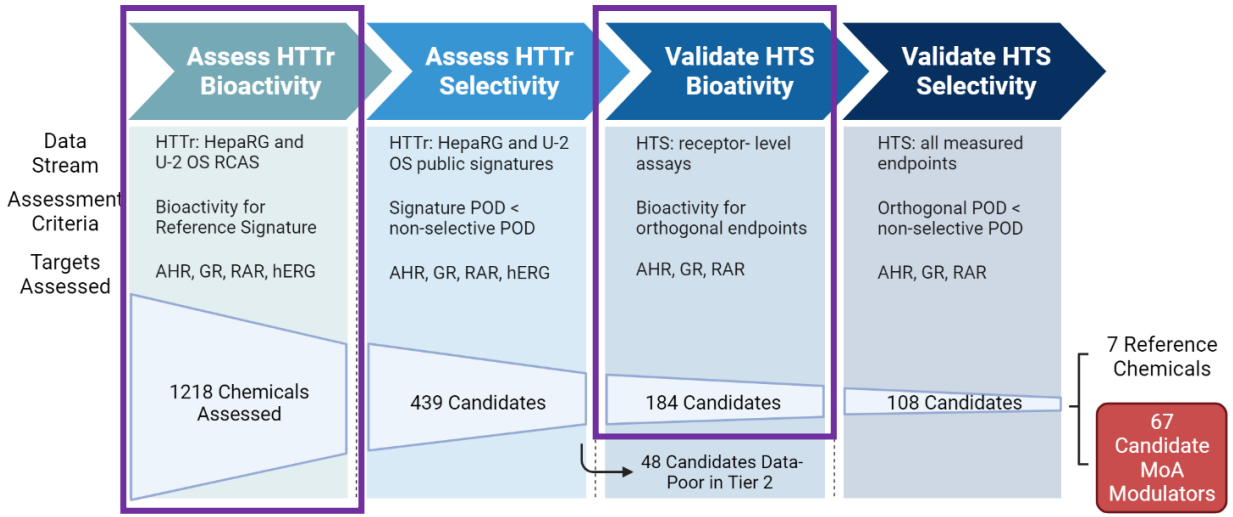
Primary Assessment Aim: identify chemicals with selective effects on molecular targets using transcriptional and receptor-level Points of Departure (PODs)

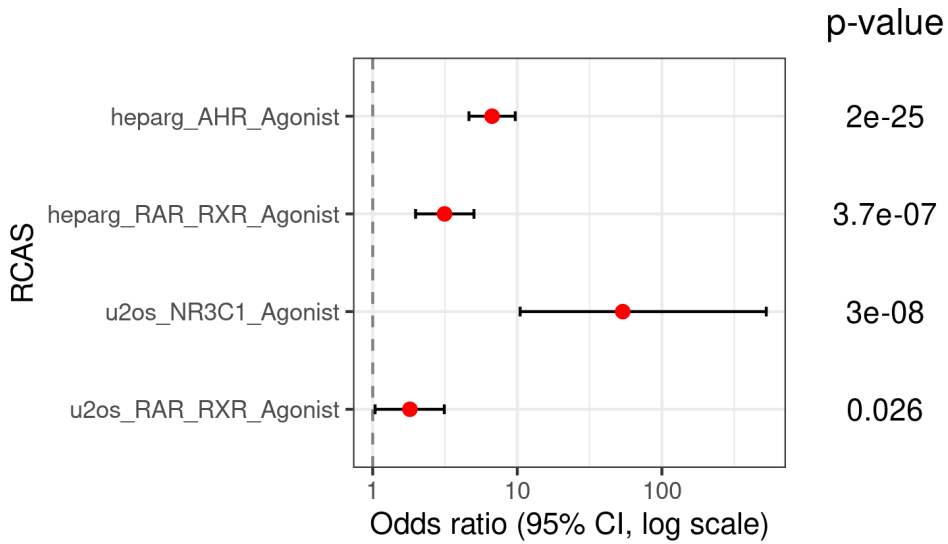
- Reference signature potencies compared to non-selective PODs estimated from distribution of >10,000 publicly-sourced signatures (Judson *et. al. Tox Sci 2016)*



Tier 1 Assessment Pre-Filters for Tier 2-Positive Chemicals

Association between Individual Tier Outcomes: Determine likelihood that Tier 1-bioactive chemicals are bioactive in at least one orthogonal Tier 2 assay



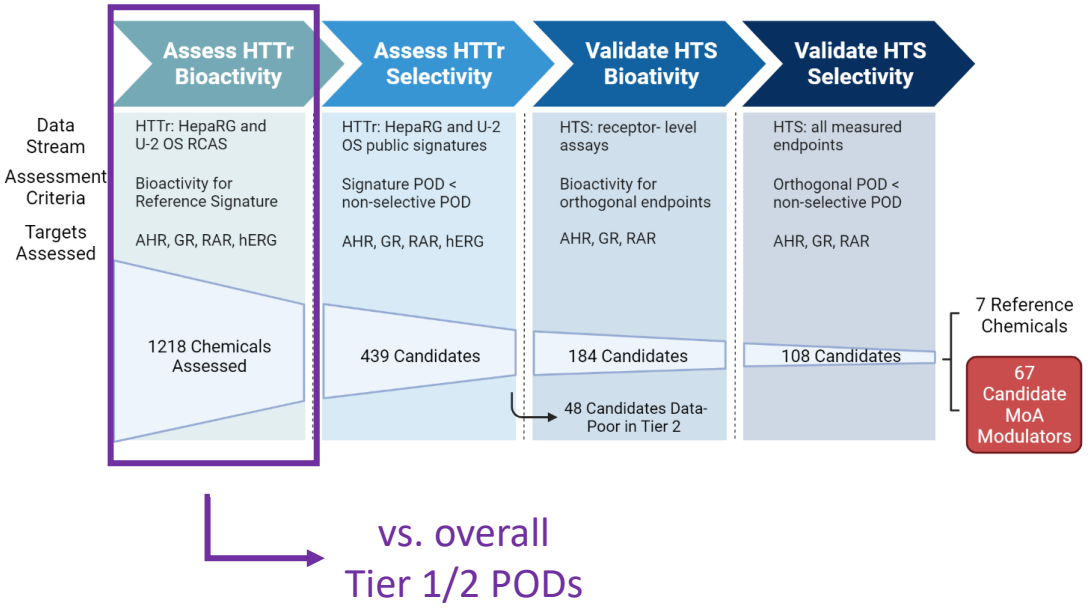


Chemicals positive for HTTr signatures were significantly more likely to show bioactivity in an orthogonal Tier 2 endpoint via Fisher's exact tests

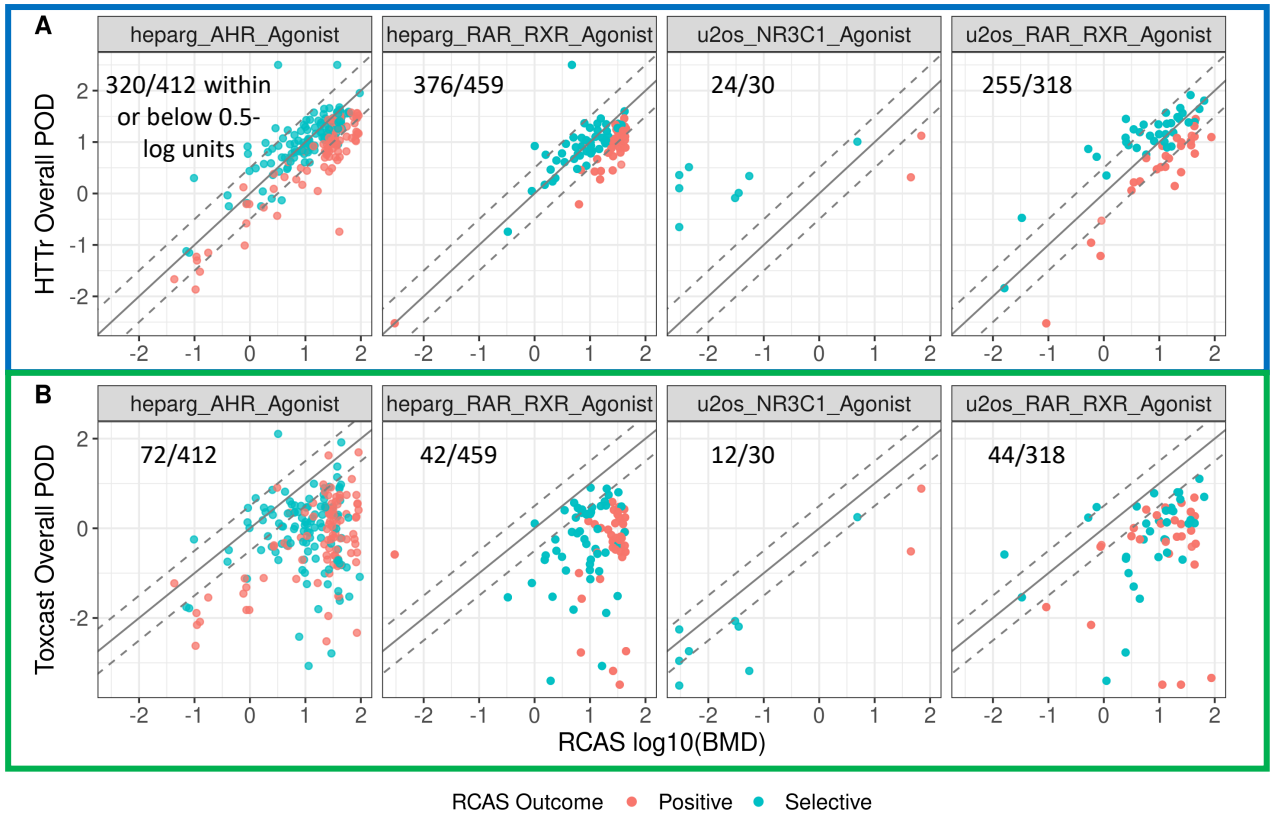
Target-Specific Potencies Reflect Overall Transcriptomic PODs

Comparison to Previous PODs: Determine difference between Tier 1 potency estimates and overall PODs from Tier 1-2 Assays

- Tier 1: 5th percentile BMD from >10,000 publicly-sourced signatures
- Tier 2: 5th percentile ACC from all measured ToxCast endpoints



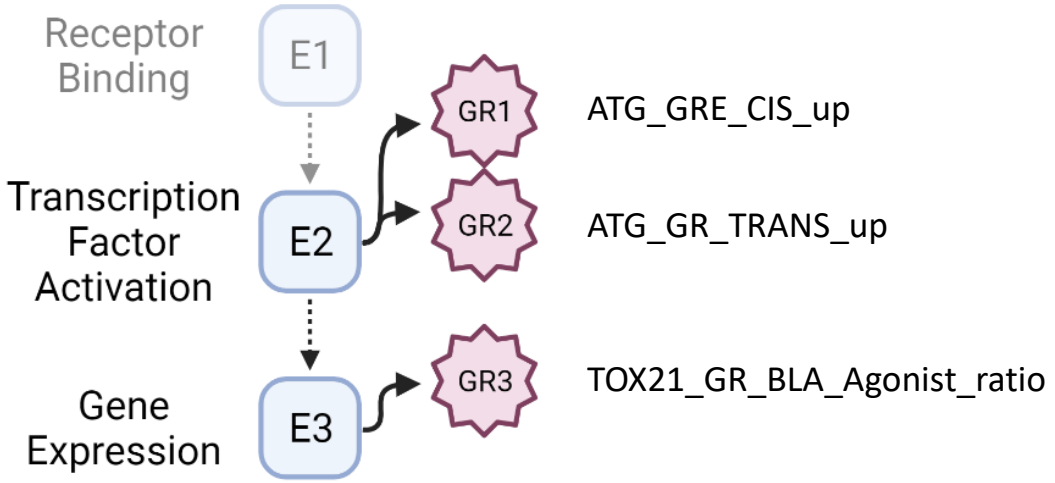
vs. overall Tier 1/2 PODs

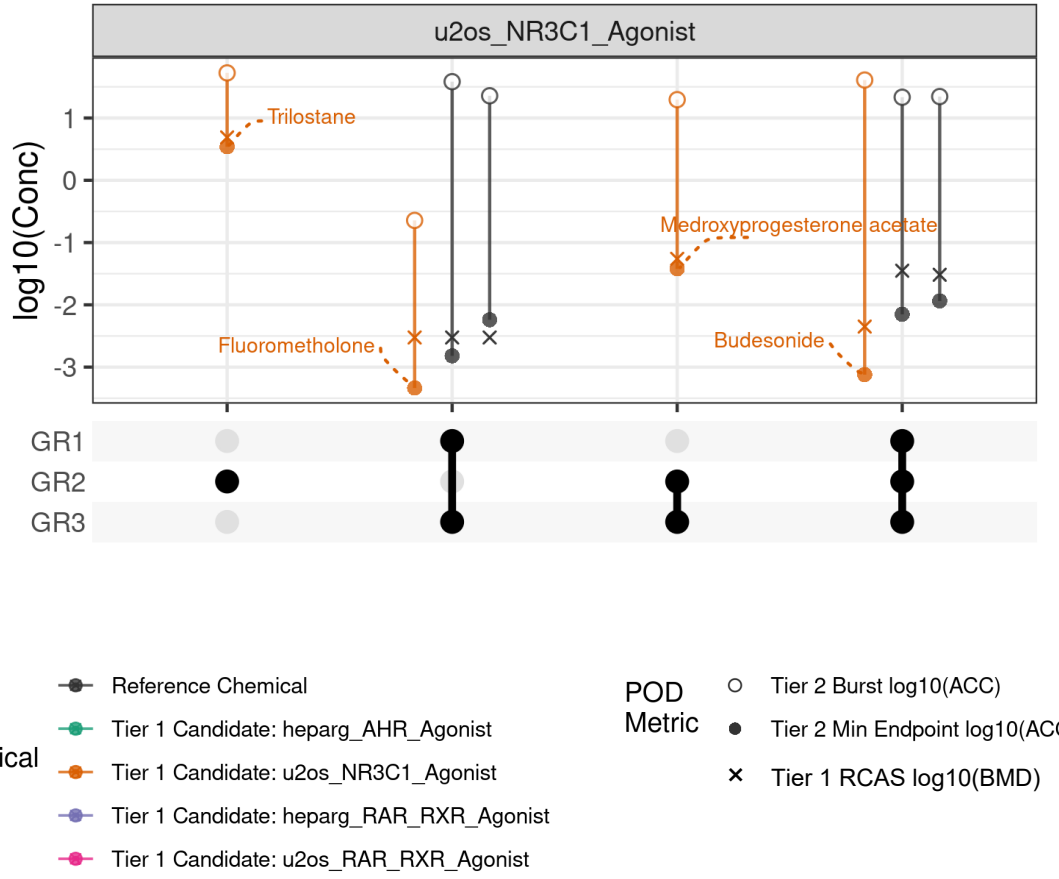


- **80±2%** of Tier 1-bioactive chemicals demonstrate MoA-specific BMD within 0.5-log units of overall HTTr POD or below
- **20±14%** of chemicals within 0.5-log units of overall ToxCast POD or below

Candidate NR3C1 Agonists Reflect Synthetic and Minor Glucocorticoids

Tier 2-Selective candidates demonstrate selective bioactivity in one or more orthogonal ToxCast endpoints:

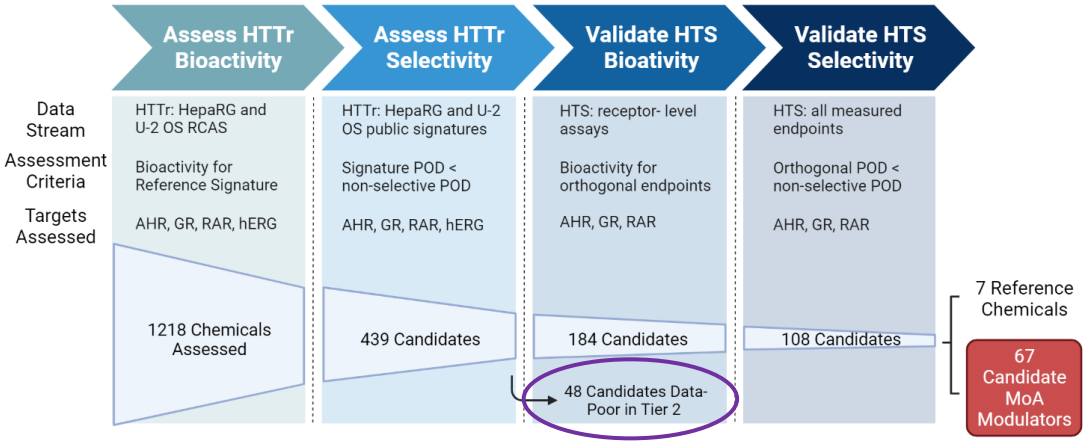




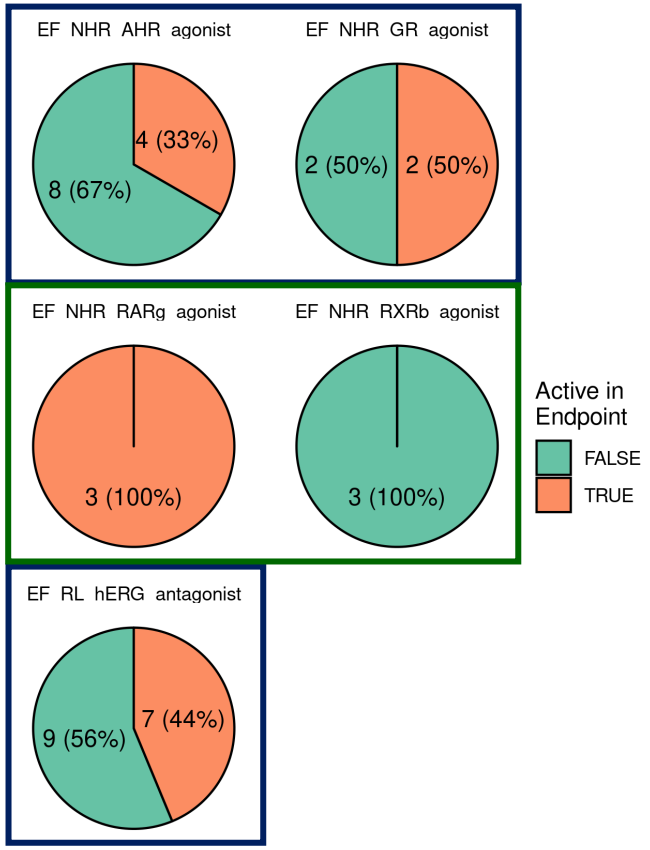
- *Fluorometholone*: active ingredient for treatment of eye inflammation
- *Medroxyprogesterone Acetate*: repression of interleukin secretion in normal human lymphocytes and amnion mesenchymal cells via minor GRE induction (Bamberger *et al. J Clin Endocrinol Metab* 1999, Marinello *et al. Front Physiol* 2020)

External Assessment of Data-Poor Chemicals Demonstrates Necessity of Multiple NAMs

Candidates with limited existing Tier 2 data profiled in orthogonal receptor-level assays:



Target	Vendor	Assay Type	No. Test Chemicals	Doses (uM)
AHR	Eurofins DiscoverX	Protein-Protein Interaction	12	0.3–30
GR	Eurofins DiscoverX	Protein-Protein Interaction	4	0.3–30
hERG	Eurofins Panlabs	Radioligand Binding	16	0.3–30
RARg	Eurofins Panlabs	Functional Coactivator	4	0.3–30
RXRb	Eurofins Panlabs	Functional Coactivator	4	0.3–30

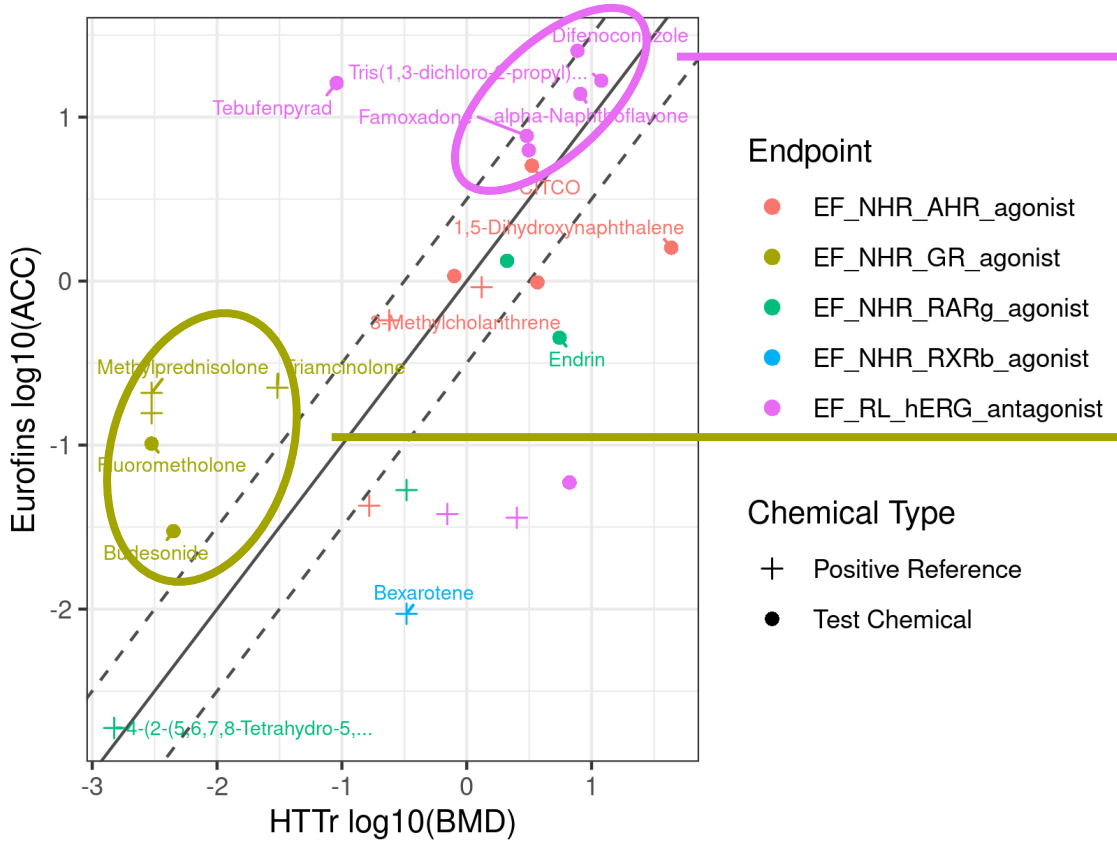


AHR/GR/hERG Candidates: Tier 2 endpoints can further support priority chemicals and de-prioritize others

RAR/RXR Candidates: Tier 2 endpoints distinguish between targets with similar transcriptomic profiles

External Assessment of Data-Poor Chemicals Demonstrates Necessity of Multiple NAMs

Estimated potency values from orthogonal endpoints compared to target-specific Tier 1 PODs:



hERG Inhibitors: Tier 1 PODs within 0.5-log units of orthogonal endpoint POD or below for 6/9 reference and test chemicals

GR Agonists: Tier 1 PODs at least 0.5-log units below orthogonal endpoint POD for 5/5 reference and test chemicals

- Sensitivity of HTTr towards individual mechanisms may be dependent on biological context
- Inclusion of multiple NAMs in pathway-specific PODs may be necessary to ensure confidence

Conclusions

- Univariate gene identification strategy paired with signature-level concentration response analysis allows for assessment of putative MoAs for transcriptomic-based toxicity testing

- Confirmation of transcriptional bioactivity via targeted Tier 2 assays identifies selectively-acting environmental chemicals and pharmaceuticals

- *Next Steps*: Inclusion of additional data streams to further support tiered testing (e.g. high throughput phenotypic profiling)

Acknowledgements

Project Mentors

- Katie Paul-Friedman
- Logan Everett

Management

- John Cowden
- Sid Hunter

BCTD/CTBB Team Members

- Joshua Harrill
- Richard Judson
- Imran Shah
- Joseph Bundy
- Derik Haggard
- Beena Vallanat
- Bryant Chambers
- Laura Taylor
- Sarah Davidson-Fitz

Email: Rogers.jesse@epa.gov

0000-0001-6667-2342

jessedrogers

Literature Mining Links Chemicals to Putative Targets

- *RefChemDB*: automated mining of multiple literature databases for chemical-molecular target interactions (Judson et. al. *ALTEX* 2019)

- Chemical assignment to molecular targets based on *support*, i.e. number of sources containing evidence of interaction
  - Hierarchical clustering of molecular target annotations based on Jaccard distance
  - Assignment of chemicals to clusters based on support of constituent molecular targets

- **13 clusters represent unique mechanisms-of-action (MoAs)** after cross-referencing with current high-throughput transcriptomics screening data

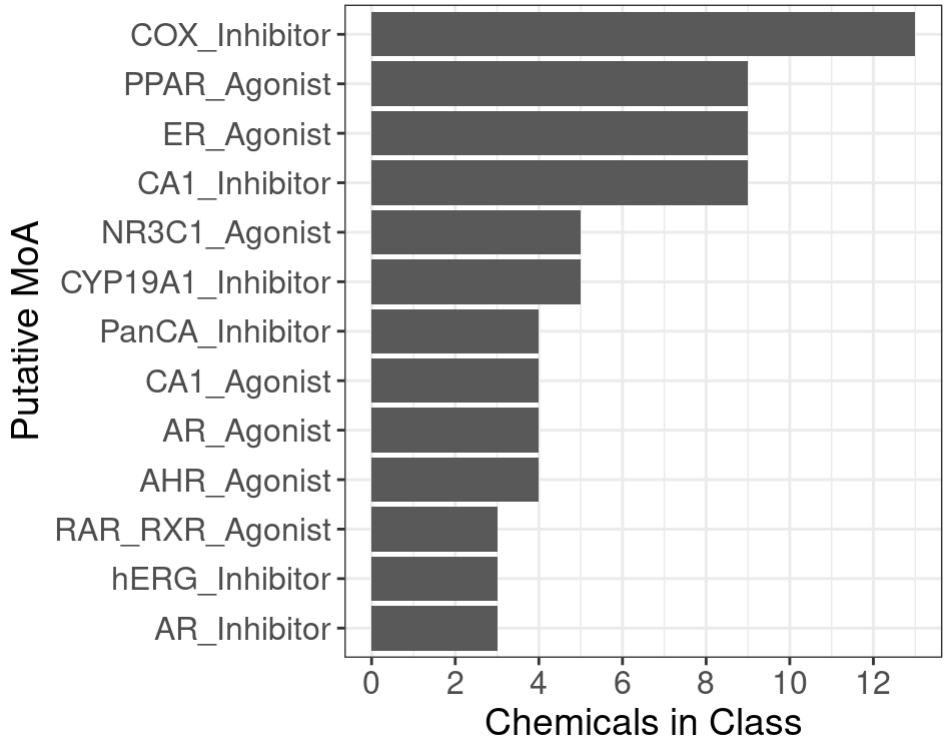
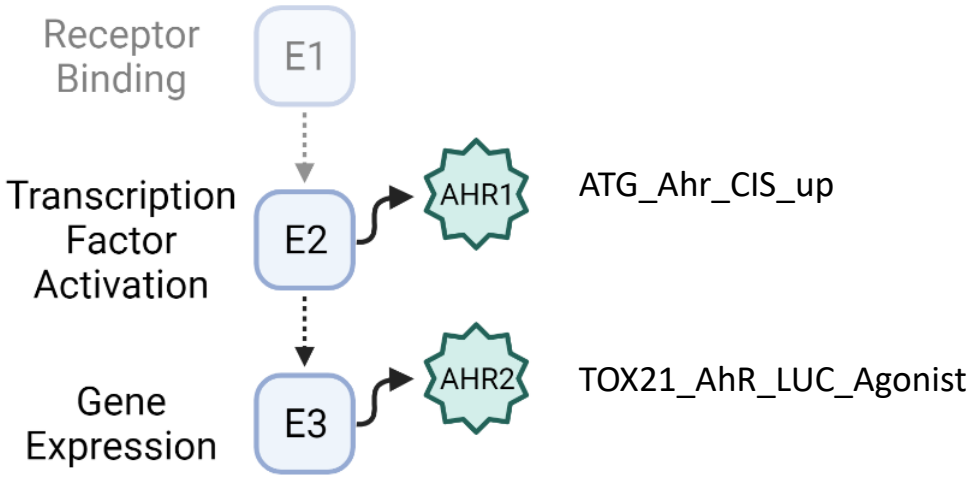
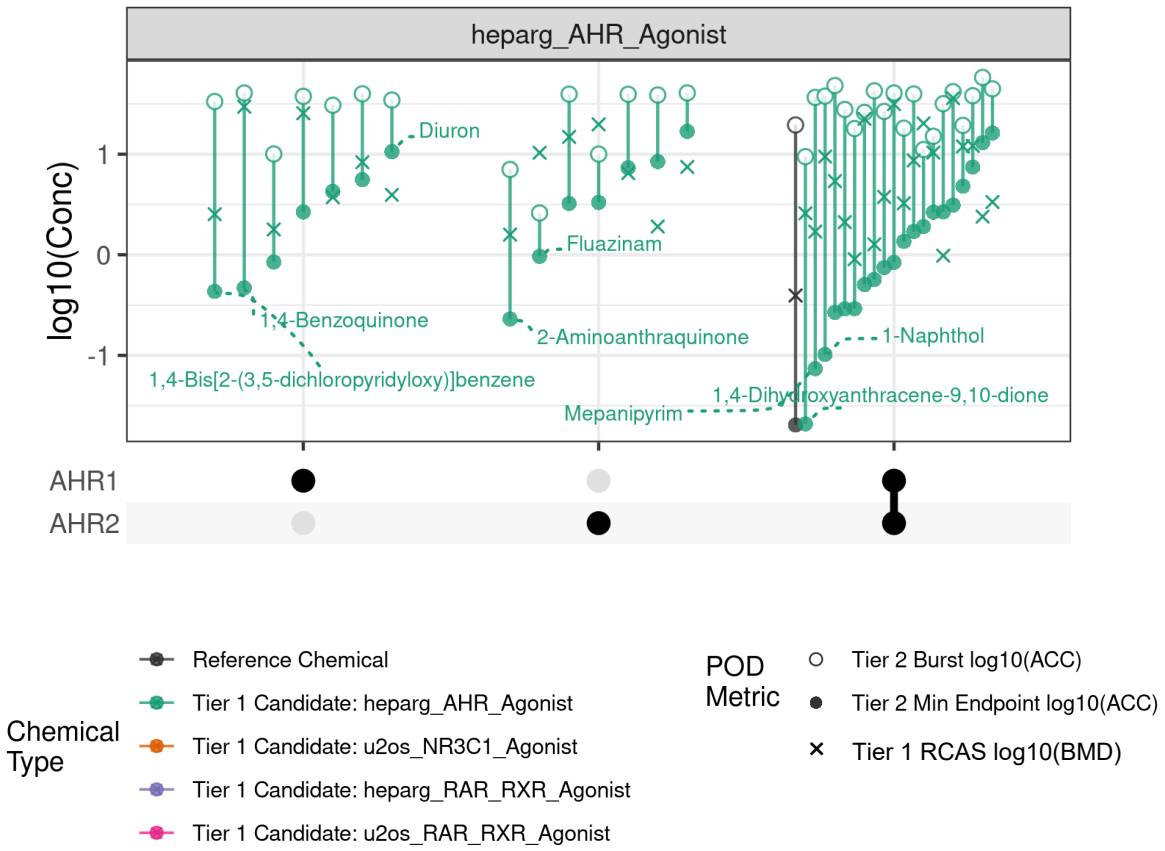


Figure indicates chemicals (selective and non-selective) associated with each signature (out of 1218 screened chemicals)

Candidate AHR Agonists Relate to Known Carcinogens

Tier 2-Selective candidates demonstrate selective bioactivity in one or more orthogonal ToxCast endpoints:





- *Anthraquinone Derivatives*: chronic oral exposure in Fisher rats increased rates of carcinogenesis, primarily in liver (Doi *et al J Environ Health B* 2006)

Candidate Retinoid Agonists Relate to …

Target	Cell Type	Tier1+2-Selective Chemicals / Tier 1-Selective Chemicals
NR3C1	U-2 OS	8/8 (100%)
RAR/RXR	**U-2 OS**	**12/35 (34.3%)**
AHR	HepaRG	35/115 (30.4%)
RAR/RXR	**HepaRG**	**24/52 (46.2%)**