EPA
United States
Environmental Protection
Agency

High-throughput transcriptomic (HTTr) screening with targeted RNA-Seq: applications for *in vitro* point-of-departure estimation and *in vitro* to *in vivo* extrapolation

Joshua A. Harrill, USEPA National Center for Computational Toxicology (NCCT)

Novel genetic-based tools for evaluating toxicity potential, mechanism of action, and population dynamics
SOT Annual Meeting, Baltimore, MD
March 11th, 2019

EPA United States Environmental Protection Agency

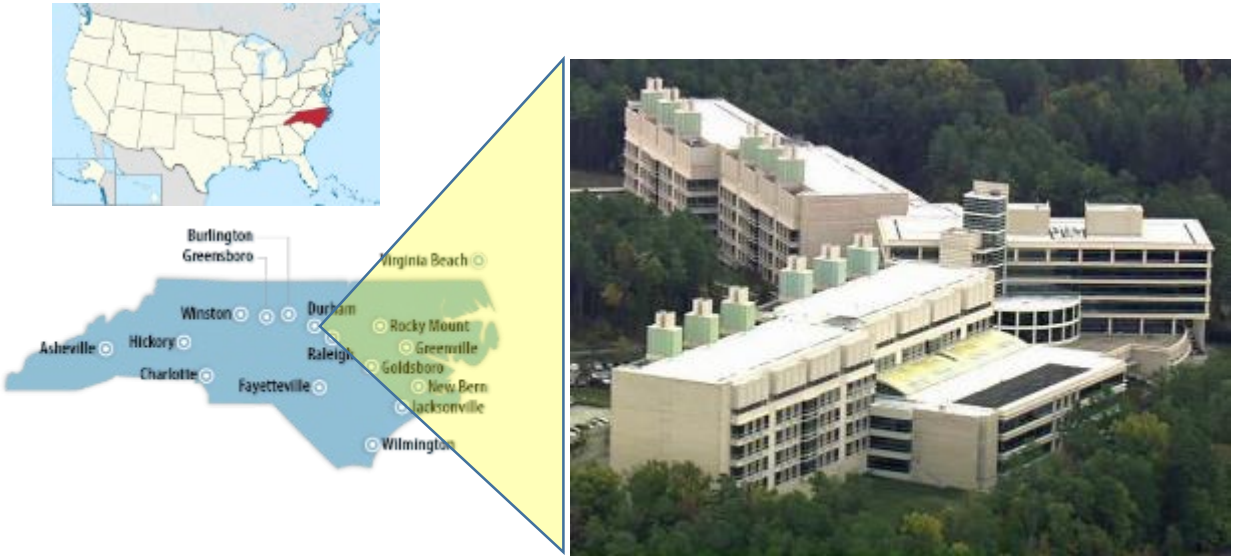
Disclaimer

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.

Outline

- **Background**
  - Who is NCCT?
  - What does NCCT Do?
  - USEPA Computational Toxicology Blueprint
- **High-Throughput Transcriptomics (HTTr)**
  - Technology Overview
  - High-Throughput Screening Workflows
  - Concentration-Response Modeling
- **Potential Applications for Regulatory Decision Making**
  - Bioactivity to Exposure Ratio (BER) Analysis
    - DMSO Soluble Chemicals
    - Volatile Chemicals.
  - Putative Mechanism of Action Exploration & Prediction
- **Concluding Remarks**

Who is NCCT?



Research Triangle Park Campus

National Center for Computational Toxicology

Mission Statement:

A research organization tasked with advancing the science of toxicity testing through the **development and/or application of novel experimental and computational approaches** for rapidly characterizing the biological activity, exposure potential and potential human health risks associated with chemicals.

The Next Generation of Computational Toxicology at USEPA

NCCT research programs focus on developing the **tools, approaches and data** needed to accelerate the pace of chemical risk assessment and foster incorporation of non-traditional toxicity testing data into regulatory decision-making processes.

- **ToxCast:** Use of 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 → protein Y*)
- Incomplete coverage of biological space.

- **New Approach for Hazard Evaluation:** Employ broad-based (i.e. non-targeted) screening assays that cast the broadest net possible for capturing potential hazard associated with chemical treatment and may be used to group chemicals based on similarity in response profiles.

The NexGen Blueprint of CompTox as USEPA
Tox Sci (2019), *in press*

Tiered Hazard Evaluation Approach

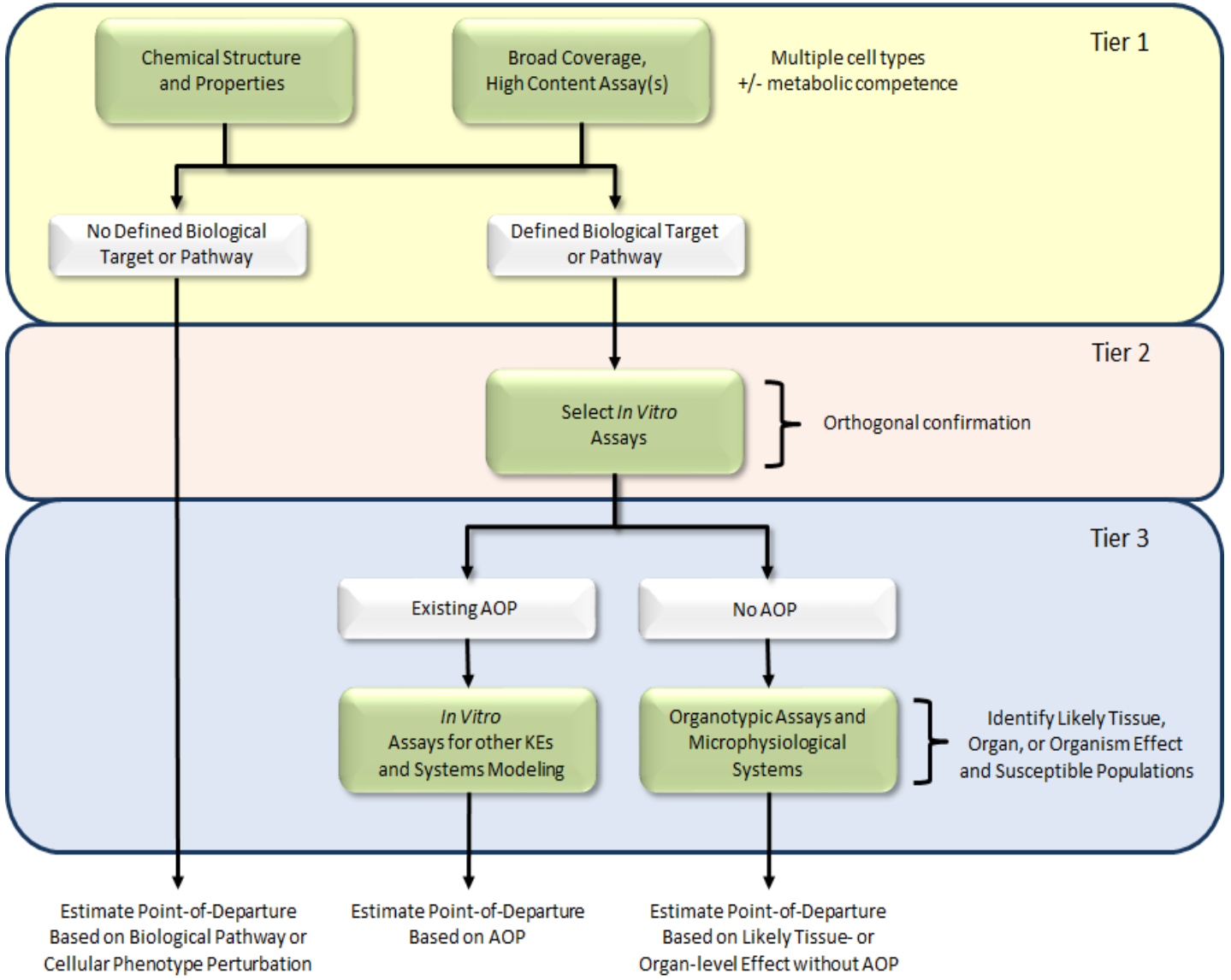
- **Tier 1 assays:**
  - Broad coverage
  - High throughput
  - Conc.-response mode
  - High content outputs
  - Tractable across many cell types / assay formats

- Increasing efficiency and declining cost has made **high-throughput transcriptomics (HTTr)** a practical option for broad coverage *in vitro* chemical screening.

- Bioactivity-based **potency estimates** can be used to identify *in vitro* **bioactivity thresholds**.

- Gene expression **profiles** can potentially be used for **mechanistic prediction** and evaluation of chemical similarity.

High-Throughput Phenotypic Profiling
(Abstract 2089/P481) Tuesday Afternoon
Willis et al., *BMD Modeling of Image-Based Phenotypic Profiling Data Yields More Potent Estimates of Chemical Bioactivity Compared to Cell Viability and Apoptosis Assays.*

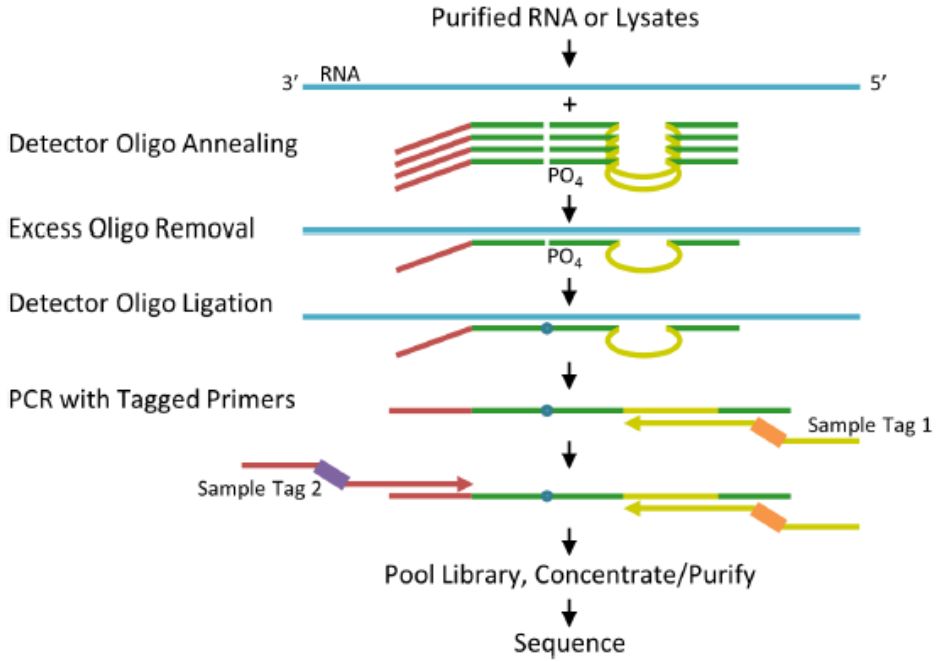


Templated Oligo with Sequencing Readout (TempO-Seq)

Technology

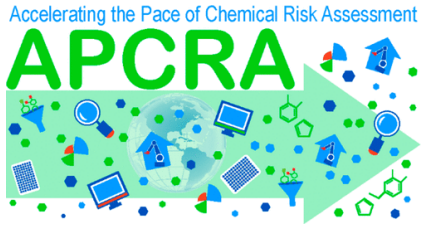
- The **TempO-Seq** human whole transcriptome assay measures the expression of greater than 20,000 transcripts.
- Requires only picogram amounts of total RNA per sample.
- Compatible with purified RNA samples or **cell lysates**.
- Transcripts in cell lysates generated in 384-well format are barcoded according to well position 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
- Per sample fastq files are generated and aligned to BioSpyder sequence manifest to generate integer count tables.

TempO-Seq Assay Illustration



HTTr MCF-7 Screen: Experimental Design

Parameter	Multiplier	Notes
Cell Type(s)	1	MCF-7
Culture Condition	1	DMEM + 10% HI-FBS [a]
Chemicals	2,112 (420)	ToxCast ph1, ph2, e1k / ph3 (APCRA)
Time Points:	1	6 hours
Assay Formats:	2	TempO-Seq HCI Cell Viability & Apoptosis
Concentrations:	8	3.5 log_{10} units; semi log_{10} spacing
Biological Replicates:	3	--



Kavlok et al. (2018)

- International collaboration of regulatory scientists focused on developing case studies for evaluating the use of New Approach Methodologies (NAMs) in chemical risk assessment.
- ECHA Workshop (2017) case study focuses on **deriving quantitative estimates of risk based on NAM-derived potency information and computational exposure estimates**

Experimental Workflow

Cryopreserved Cell Stocks

Cell Expansion

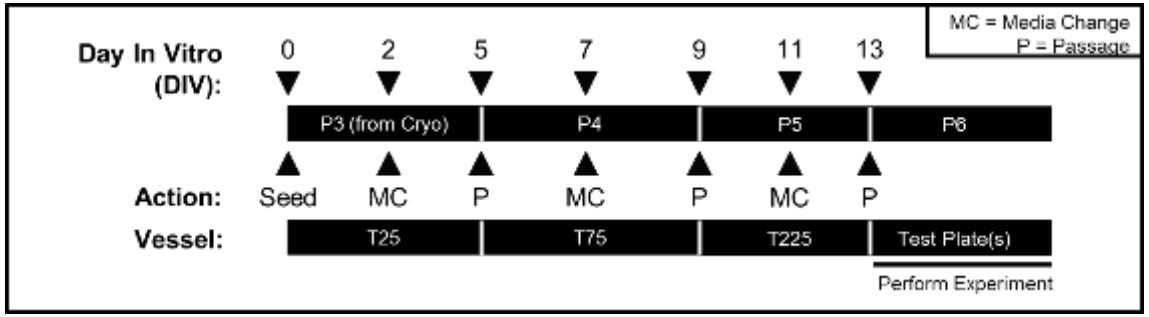
Cell Plating

BioTek MultiFlo ™ FX

Dispensing Test Chemicals

LabCyte Echo® 550 Liquid Handler

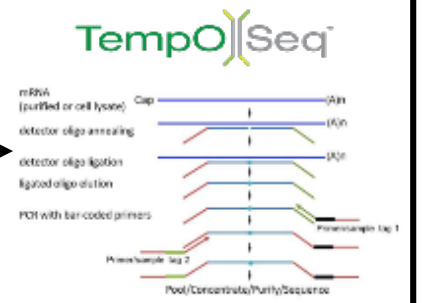
Standardized Expansion Protocol



Track 1: Targeted RNA-Seq

Generating Cell Lysates

TempO-Seq WT



Track 2: Apoptosis / Cell Viability

Reagent Dispensing

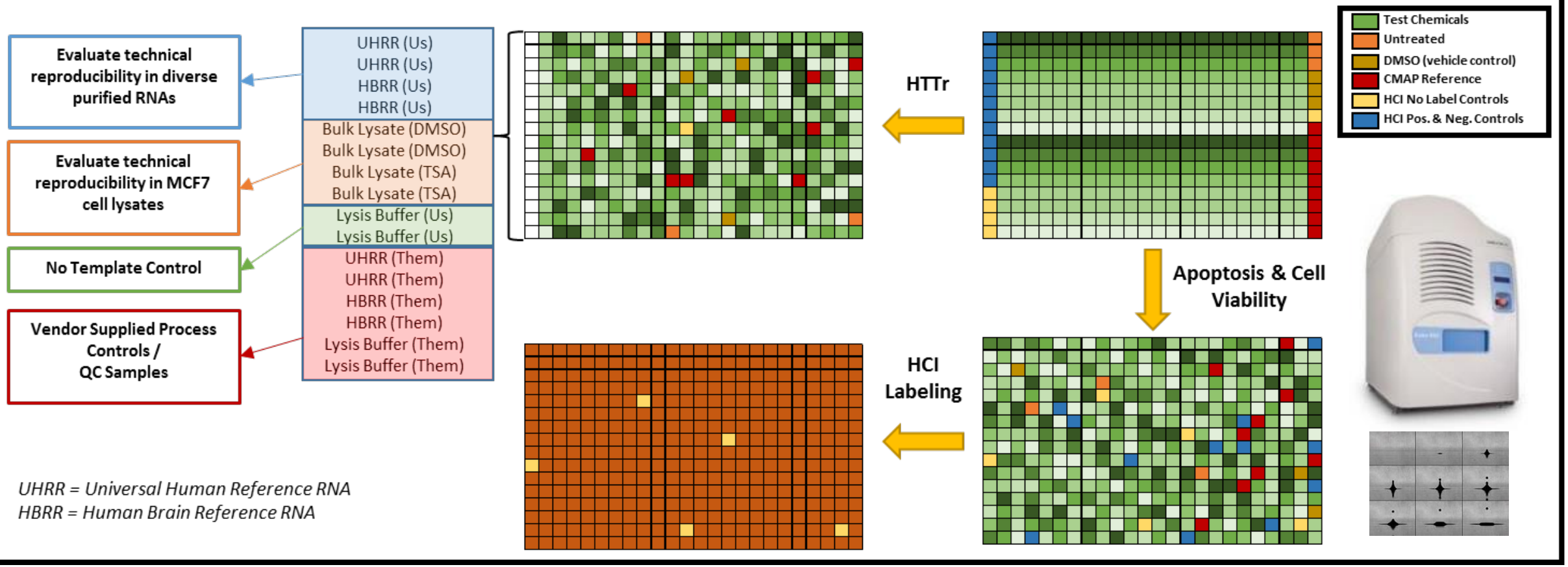
High Content Imaging

Perkin Elmer Opera Phenix™ High Content Screening System

Treatment Randomization & Quality Control Samples

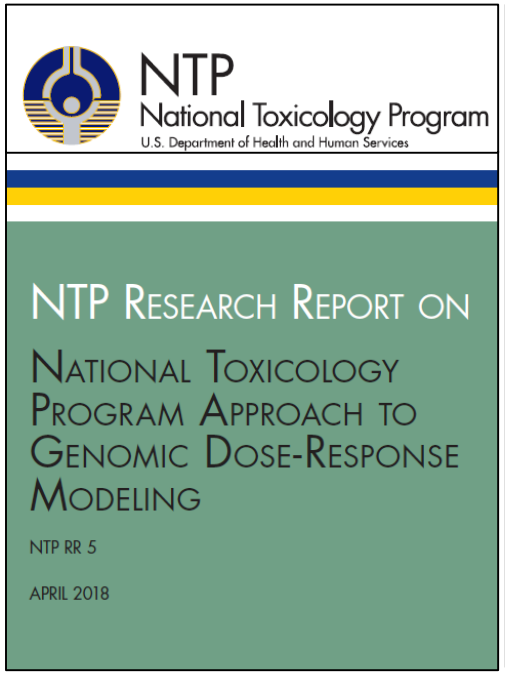
Treatment Randomization: *Each test plate uniquely randomized with respect to treatment.*

QC Samples: *Quality Control samples included on each plate*



UHRR = Universal Human Reference RNA
HBRR = Human Brain Reference RNA

Concentration Response Modeling

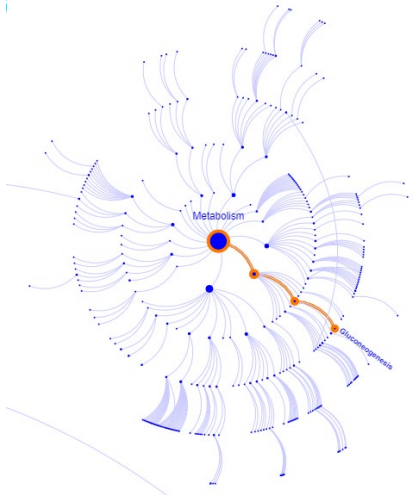


Parameter	Criteria [a]
Pre-filter:	William's Trend Test ($p_{raw} < 0.05$ & $\|FC\| \geq 1.5$)
Models	Hill, Power, Linear, Poly2, Exp2|3|4|5
BMR Factor:	1.349 (10 %)
Best Model Selection:	Nested χ^2 within poly models → Lowest AIC
Hill Model Flagging:	'k' < 1/3 Lowest Positive Dose Retain Flagged Models
Pathway Analysis:	Genes with BMD <= Highest Dose ≥ 3 ≥ 5% Gene Set Coverage
Gene Set Collections:	Reactome, *MSigDB, GO-BP, BioPlanet*

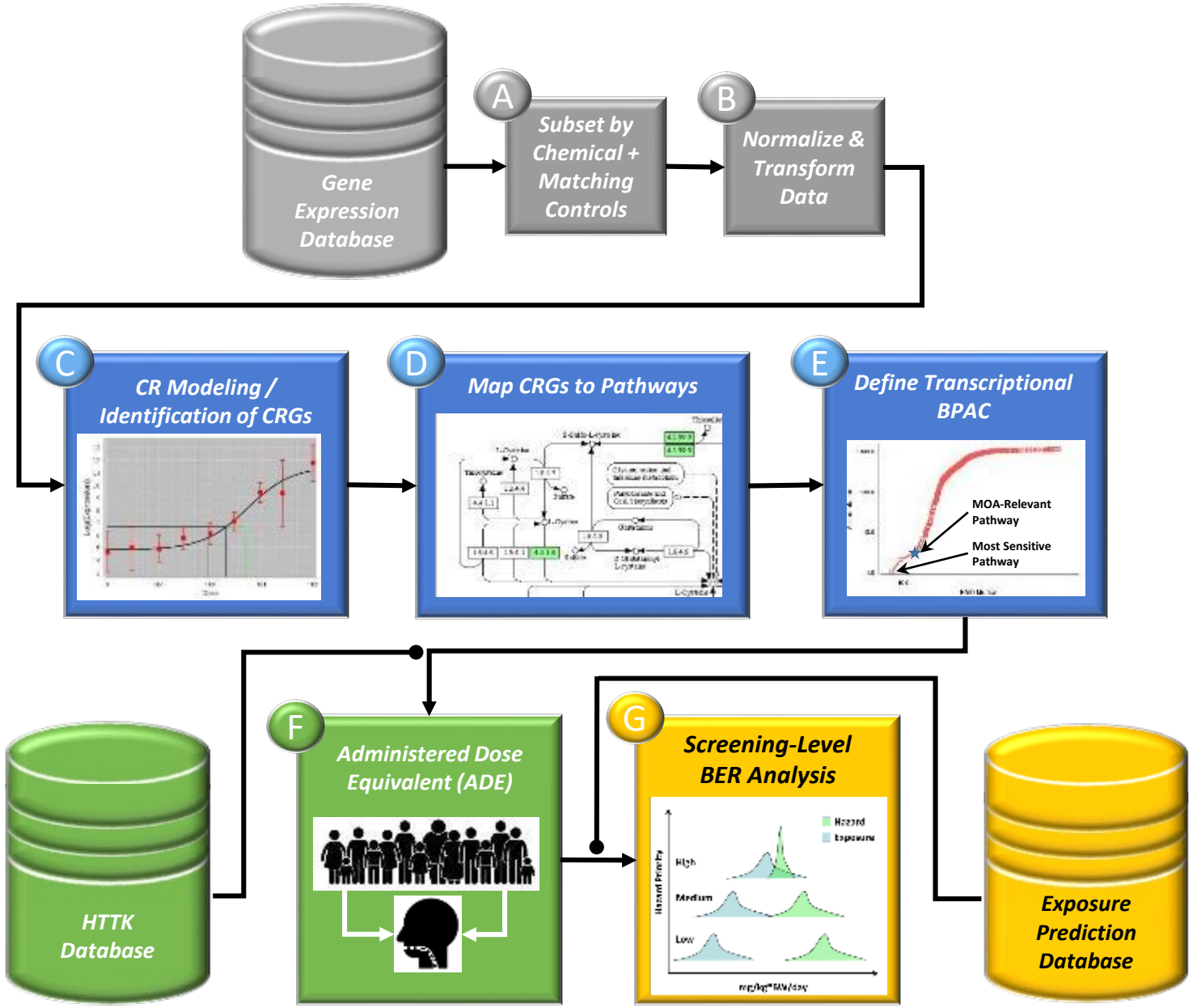
[a] *Exploratory analysis – modeling criteria not finalized*

REACTOME
A CURATED PATHWAY DATABASE

- Free, open-source, curated and peer-reviewed pathway database.
- Biological entities (i.e. proteins, complexes, etc.) participate in reactions that form a network of biological interactions and are organized into a hierarchical pathway structure.
- Bioinformatics tools for visualization, interpretation and analysis of biological pathway knowledge.
- www.reactome.org



Bioactivity Exposure Ratio (BER) Analysis Using HTTr



Bioactivity to Exposure Ratio Comparisons Using Reverse Dosimetry

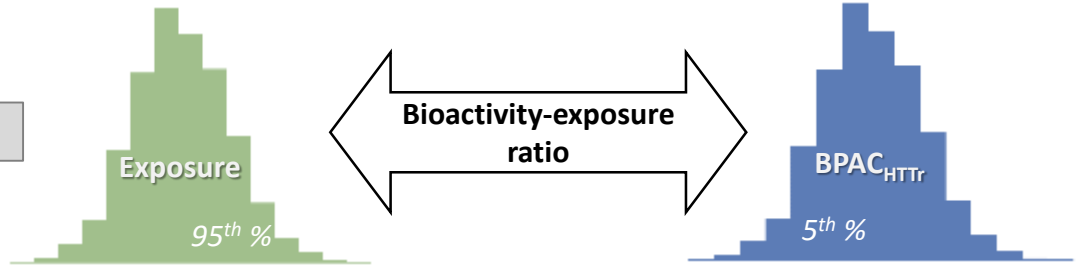
- *Median BMD for most sensitive pathway*

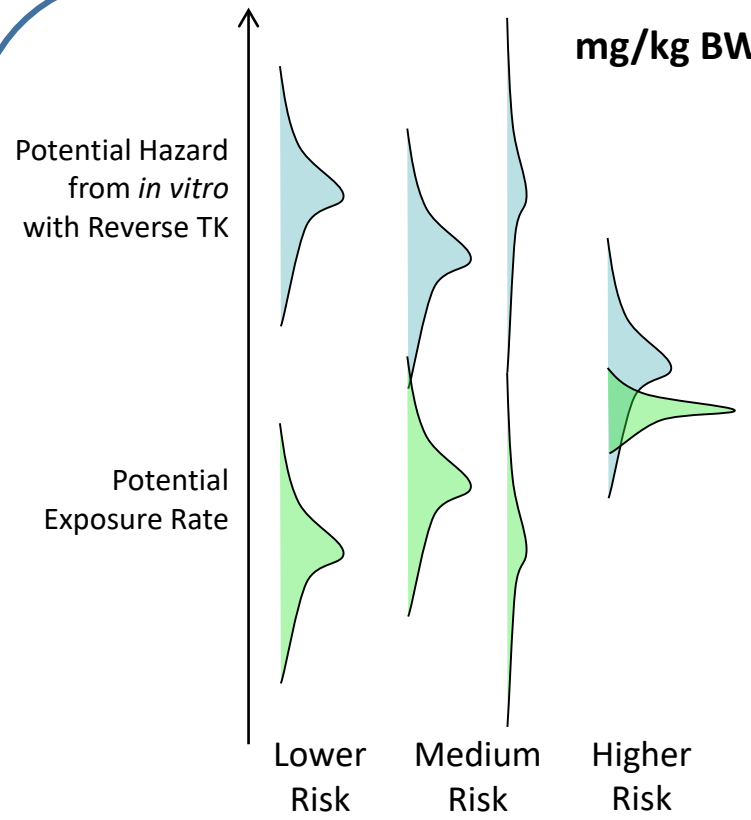
HTTr BPAC (µM)

Apply high-throughput toxicokinetics (httk) to get mg/kg/day

- *Using httk v1.9 values for humans*
- *Default to a simple model with no partition coefficients and use of C_{ss}.*
- *Assume 100% bioavailability and restrictive clearance.*
- *Monte Carlo simulations to obtain distribution for healthy human population*

EPA - ExpoCast



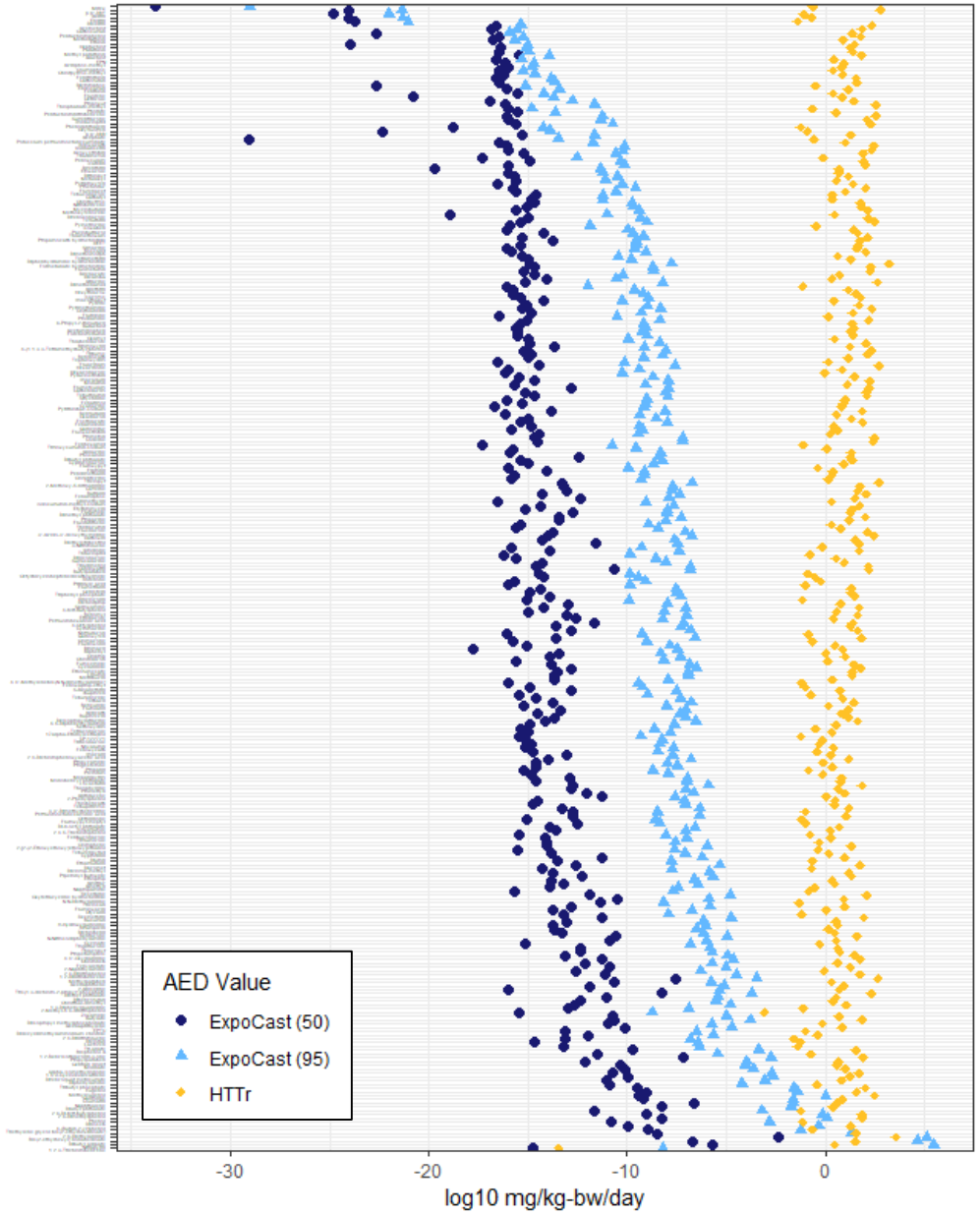


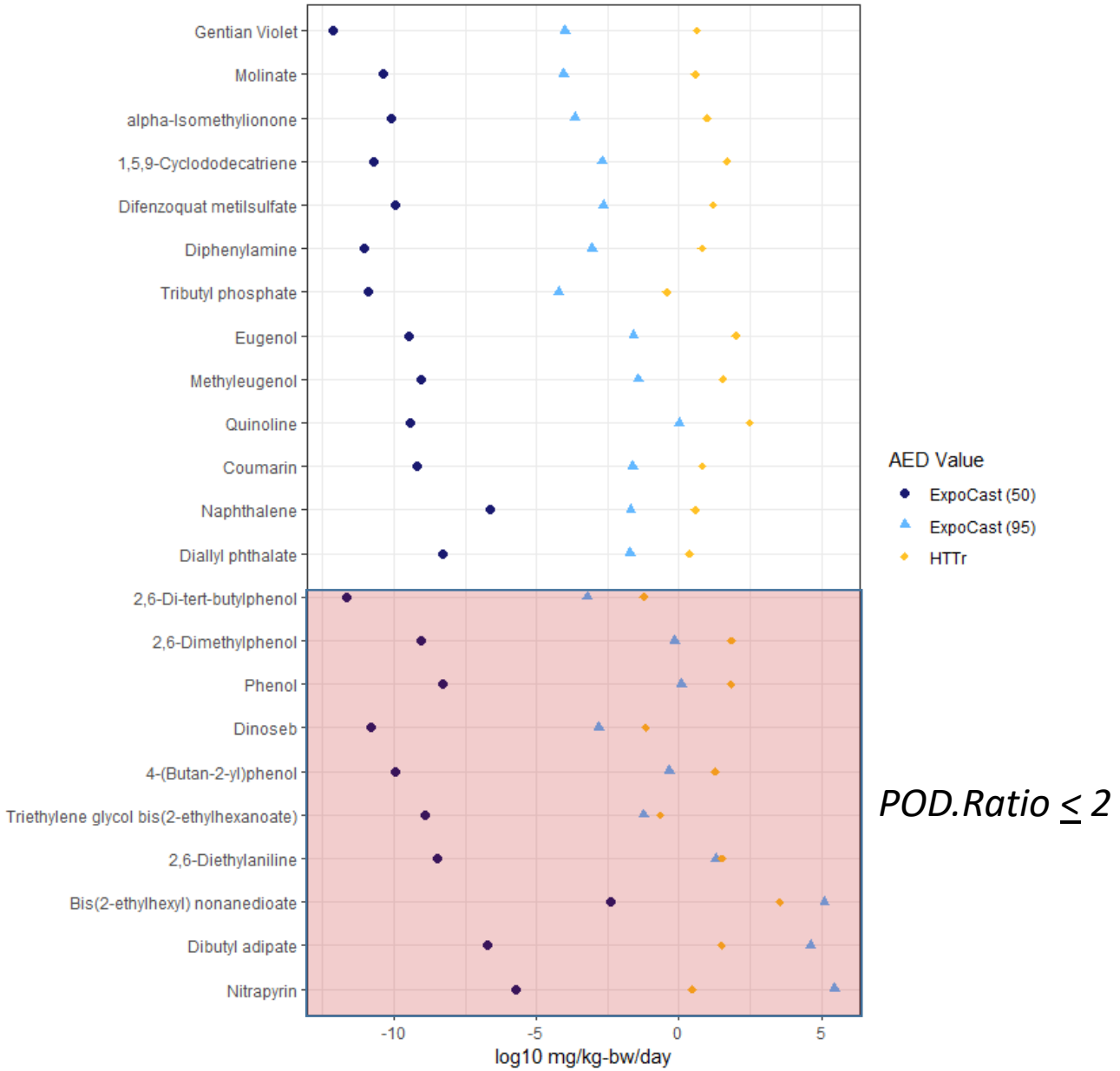
High-throughput toxicokinetic (httk) modeling: Conversion of *in vitro* bioactivity to *in vivo* steady state concentration (C_{ss})

Reverse dosimetry: Conversion of predicted C_{ss} to an administered equivalent dose (AED)

Figure Courtesy of Katie Paul-Friedman & John Wambaugh

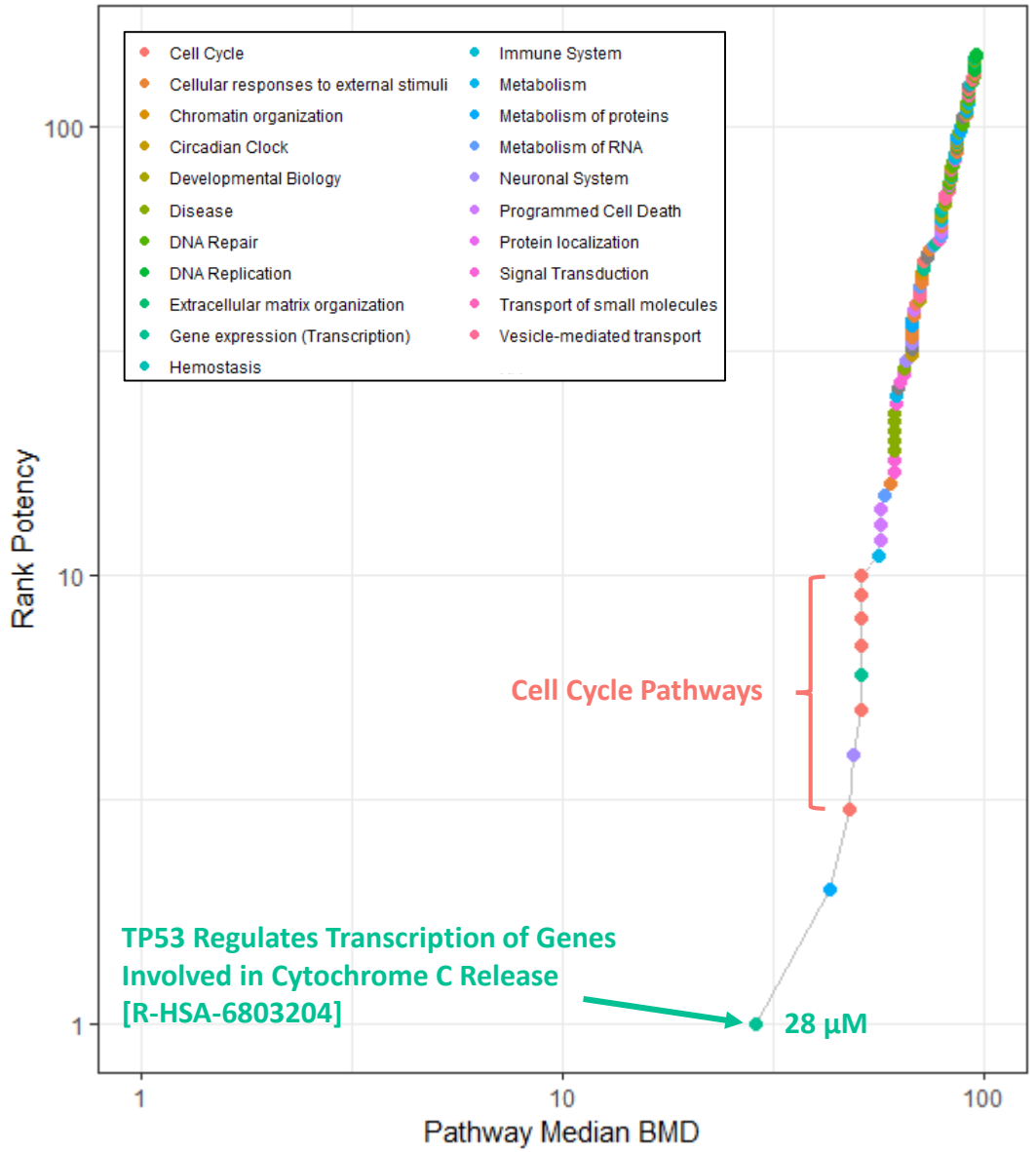
HTTr Bioactivity-to-Exposure Ratio Results



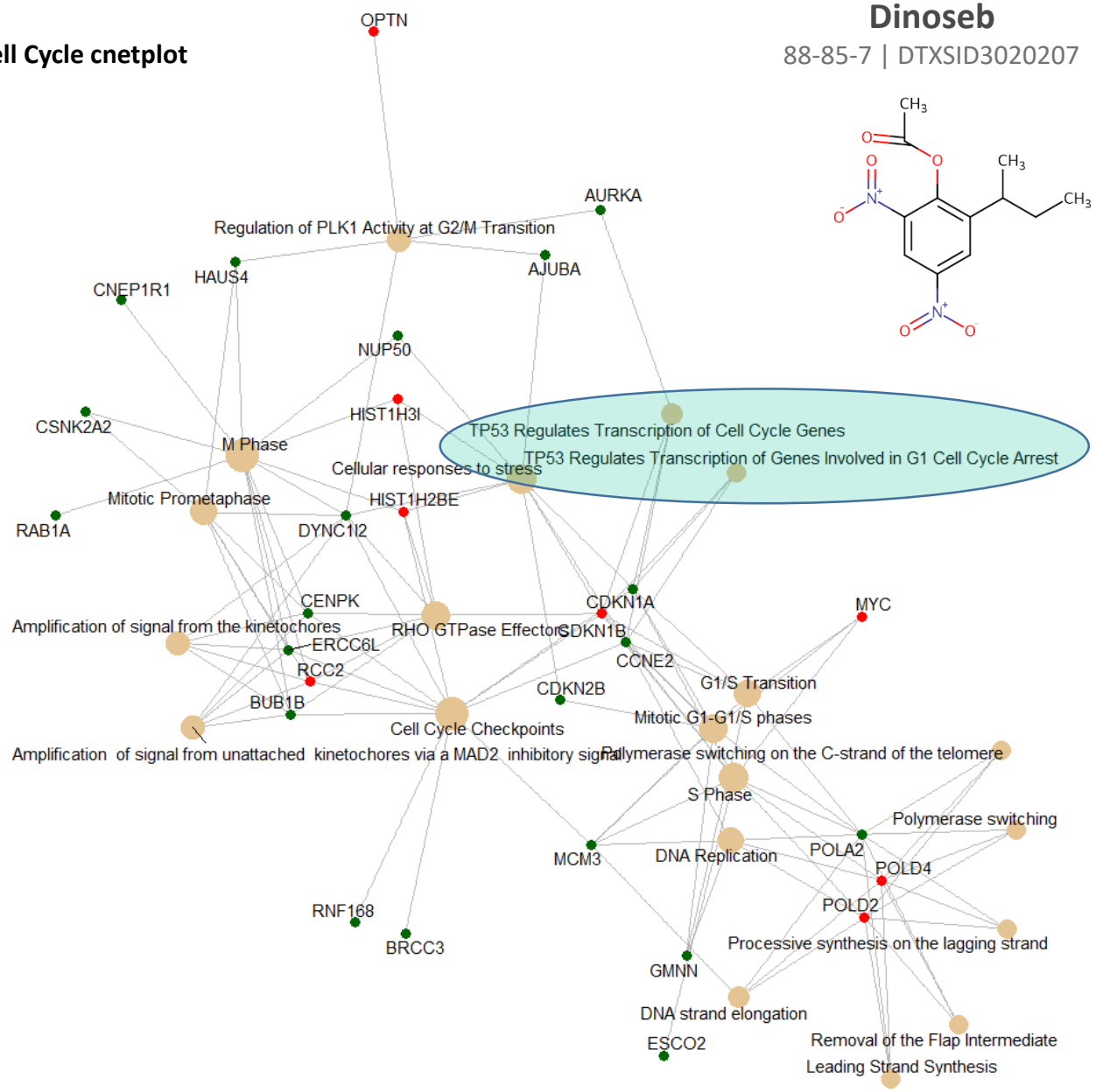


POD.Ratio ≤ 2

BPAC Potency and Putative Mechanism



Cell Cycle cnetplot

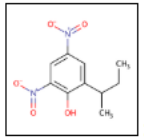


Dinoseb

88-85-7 | DTXSID3020207

Reactome PA: *Guangchuang Yu, Qing-Yu He. Molecular BioSystems 2016, 12(2):477-479*

Cell Cycle: Comparison to ToxCast



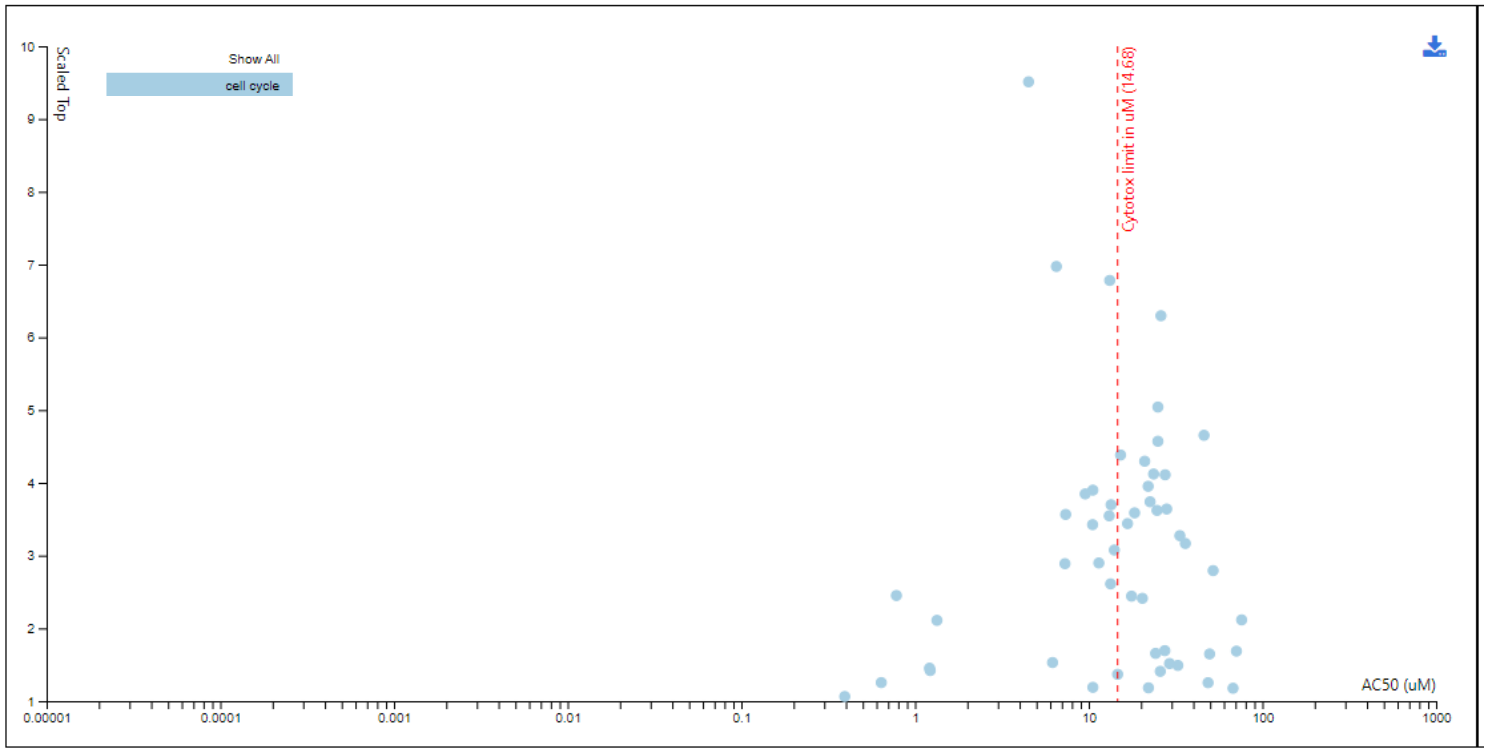
Dinoseb

88-85-7 | DTXSID3020207

Searched by DSSTox Substance Id.

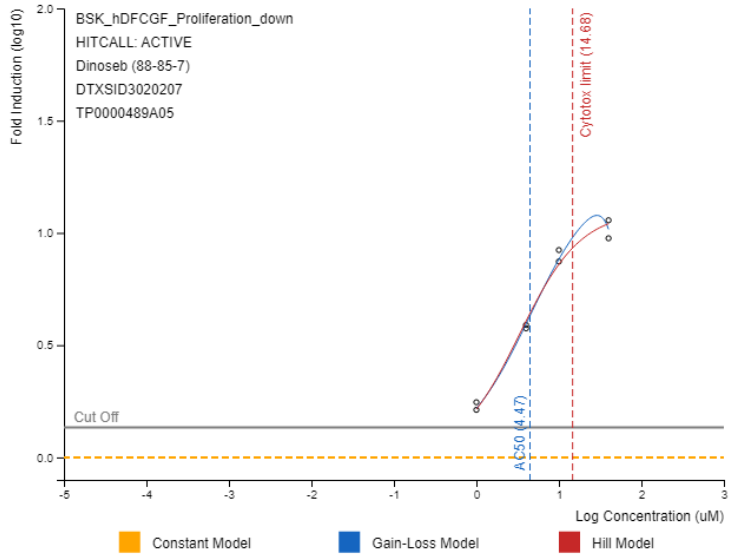
Chemical Activity Summary

TOXCAST DATA



http://comptox-prod.epa.gov/dashboard/dsstoxdb/results?search=DTXSID3020207#invitrodb

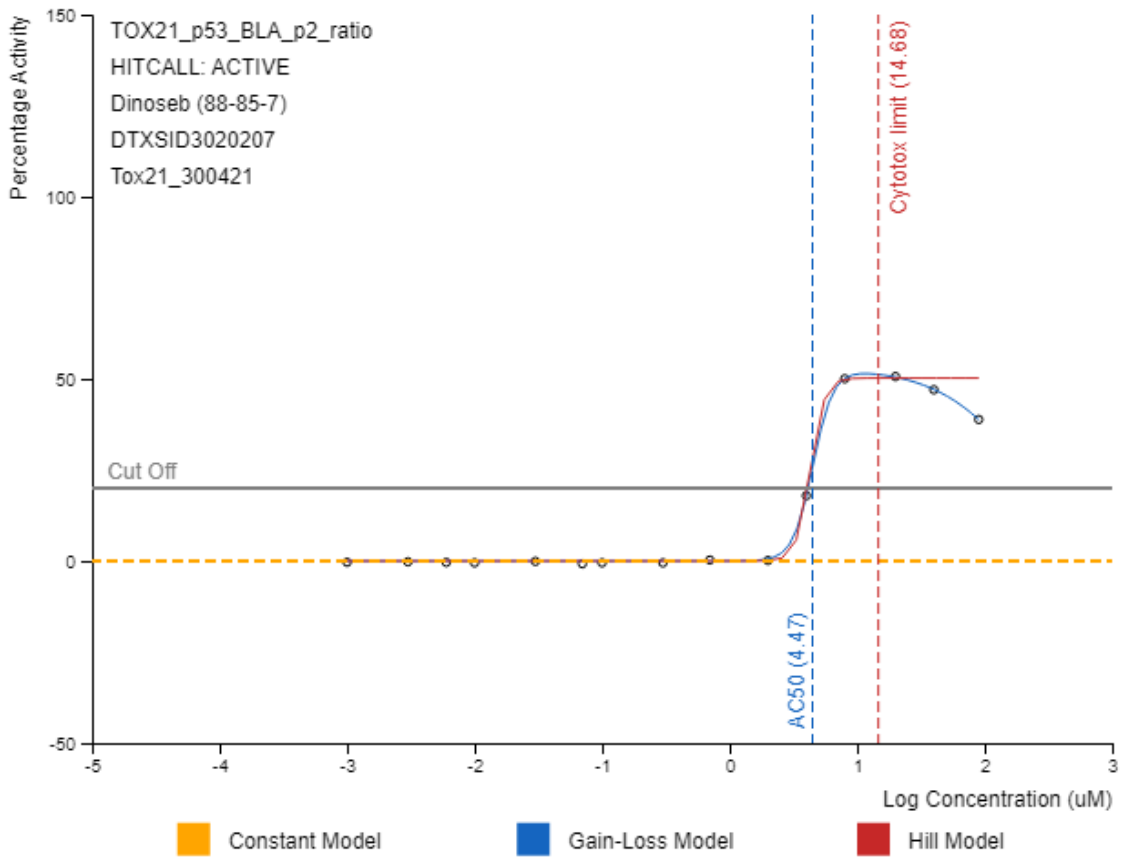
Assay NAME	HIT_CALL	AC50
BSK_SAg_Proliferation_down	ACTIVE	13.22
BSK_hDFCGF_Proliferation_down	ACTIVE	4.51
BSK_CASM3C_Proliferation_down	ACTIVE	26.06
BSK_3C_Proliferation_down	ACTIVE	6.52



Winning Model	Model	AIC	RMSE	Top	AC50	Slope
	Constant	21.53	0.75	-	-	-
✓	Gain-Loss	-21.45	0.03	1.27	4.51	1.04
	Hill	-21.44	0.04	1.1	3.34	1.17

P53 Signaling: Comparison to ToxCast

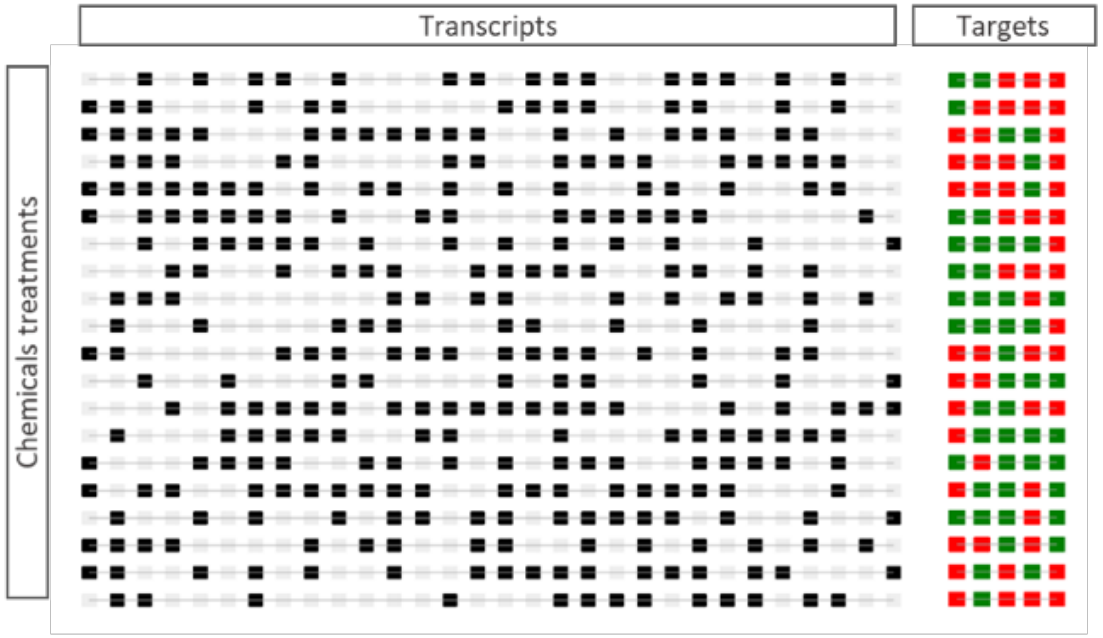
Assay NAME	HIT_CALL	AC50	SCALED_TOP
TOX21_p53_BLA_p1_ratio	INACTIVE	-	-
TOX21_p53_BLA_p1_viability	INACTIVE	-	-
TOX21_p53_BLA_p2_ratio	ACTIVE	4.51	2.65
TOX21_p53_BLA_p2_viability	INACTIVE	-	-
TOX21_p53_BLA_p3_ratio	ACTIVE	0.94	1.97
TOX21_p53_BLA_p3_viability	INACTIVE	-	-
TOX21_p53_BLA_p4_ratio	ACTIVE	10.7	1.14
TOX21_p53_BLA_p4_viability	INACTIVE	-	
TOX21_p53_BLA_p5_ratio	ACTIVE	4.1	1.37
TOX21_p53_BLA_p5_viability	INACTIVE	-	-
APR_HepG2_p53Act_24h_dn	INACTIVE	-	-
APR_HepG2_p53Act_24h_up	INACTIVE	69.8	0.57
APR_HepG2_p53Act_72h_up	INACTIVE	-	
APR_HepG2_p53Act_72h_dn	ACTIVE	78.6	1.21

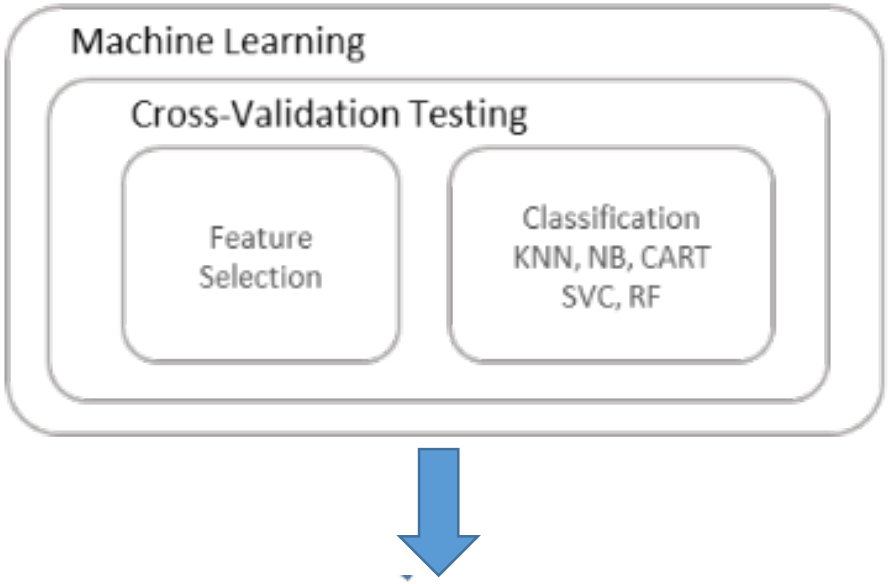


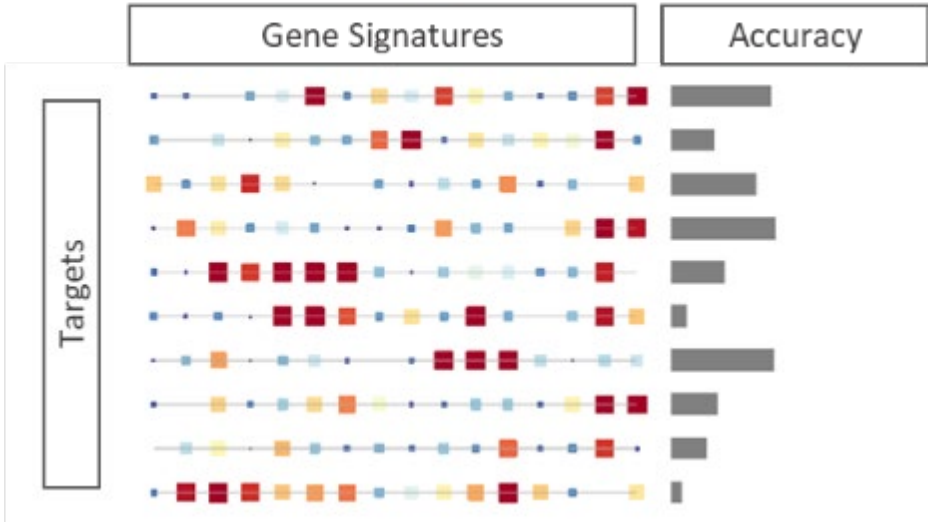
Winning Model	Model	AIC	RMSE	Top	AC50	Slope
	Constant	150.85	24.64	-	-	-
✓	Gain-Loss	26.33	0.37	52.99	4.51	5.54
	Hill	61.48	3.07	50.22	4.31	8

http://comptox-prod.epa.gov/dashboard/dsstoxdb/results?search=dinoseb#details

Signatures/Classifiers For Putative Target Prediction







CONNECTIVITY MAP

BROAD INSTITUTE

- Manually curated a sub-set of the Connectivity Map (v2) MCF7 database with target associations.
- Created a series of target-centric signatures.
- Queried against both CMAP and TempO-Seq HTTr database

Slide by Imran Shah

ER Model (any Mode) Derived from CMAP

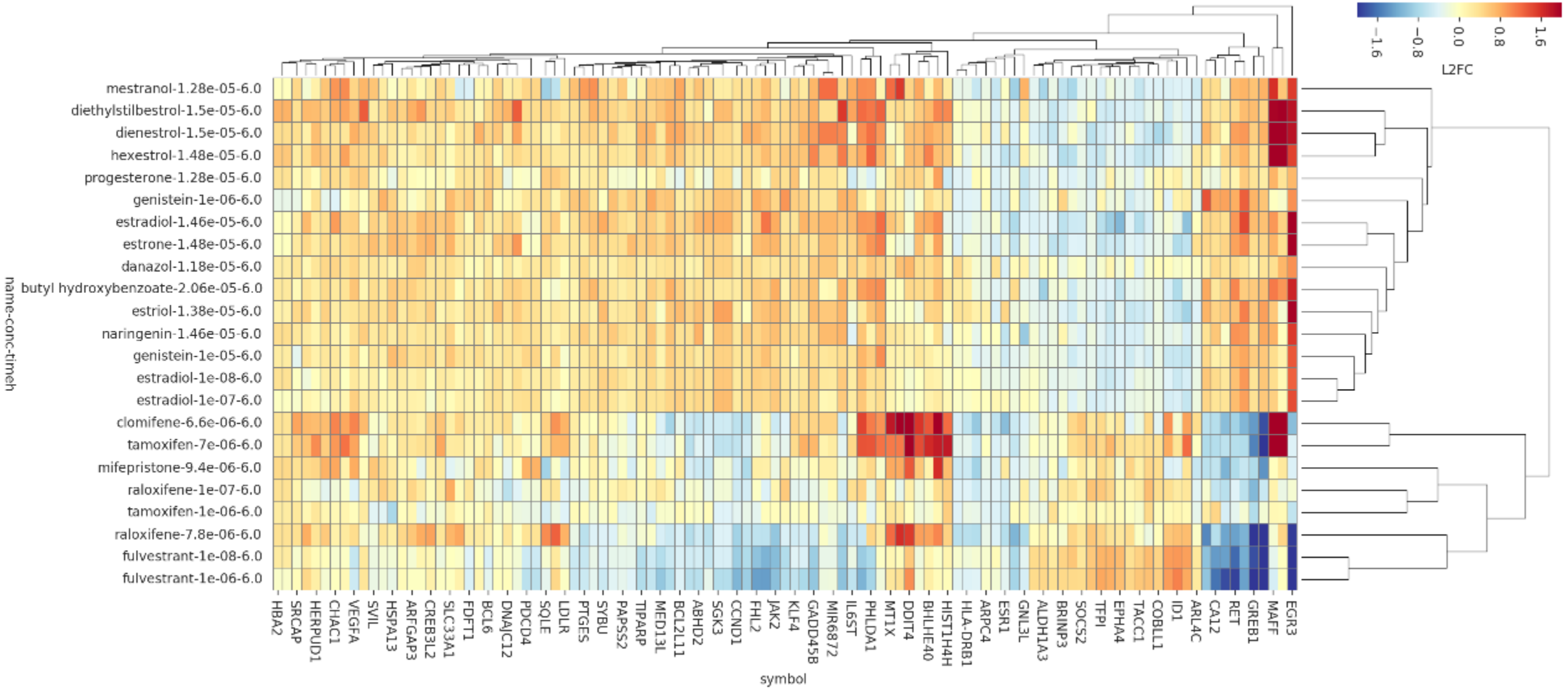


Figure Courtesy of Imran Shah

EPA United States Environmental Protection Agency

Performance of Signatures for Putative Target Prediction in HTTr Data

Putative Target	CMap v2 / Affymetrix	BioSpyder HTTr-Phase I (n = 352)			
	Signature size	PPV	Positives	Positive Chemicals found (Curated)	Top 5 Prediction (Uncurated)
CYP2C9	131	1	1	Fluconazole	Emodin, Phenazopyridine hydrochloride, Lactofen, Hexachlorophene, 2-Amino-5-azotoluene
ESR1	257	1	11	o,p'-DDT, Genistein, 4-Nonylphenol, 4-Hydroxytamoxifen, Diethylstilbestrol, Raloxifene hydrochloride, Bisphenol A, 17beta-Estradiol, 5alpha-Dihydrotestosterone, Mifepristone, 4-(1,1,3,3-Tetramethylbutyl)phenol	dl-Norgestrel, SSR504734, Haloperidol, Cyclosporin A, Astemizole
HDAC1	124	1	2	Trichostatin A, Valproic acid	2-(Thiocyanomethylthio)benzothiazole, Azinphos-methyl, Sodium (2-pyridylthio)-N-oxide, 3,3'-Dichlorobenzidine dihydrochloride
DHFR	215	1	2	Pyrimethamine, Methotrexate	Adriamycin hydrochloride, PharmaGSID_48505, Etoposide, Resveratrol, Nisoldipine
NR1I2	139	1	2	17beta-Estradiol, Bisphenol A	dl-Norgestrel, Endosulfan, Isodrin, Genistein, 17alpha-Estradiol
PGR	115	1	1	Mifepristone	Flurandrenolide, Fluorometholone, Dexamethasone, Melengestrol acetate, Betamethasone
HMGCR	236	1	1	Lovastatin	Resveratrol, dl-Norgestrel, o,p'-DDT, Tamoxifen, Chlorhexidine
ABCC2	357	1	1	Methotrexate	4-Nitrosodiphenylamine, Resveratrol, Adriamycin hydrochloride, Nisoldipine, 8-Hydroxyquinoline sulfate
TYMS	329	1	1	Methotrexate	Etoposide, Resveratrol, 4-Nitrosodiphenylamine, Cytarabine hydrochloride, PharmaGSID_48505
ESR2	281	0.86	7	Genistein, Diethylstilbestrol, 4-Nonylphenol, Bisphenol A, 4-Hydroxytamoxifen, 17beta-Estradiol	dl-Norgestrel, 17alpha-Estradiol, Haloperidol, Cyclosporin A, Isodrin
AR	261	0.78	9	o,p'-DDT, 17beta-Estradiol, 5alpha-Dihydrotestosterone, Flutamide, Bisphenol A, Mifepristone, 17-Methyltestosterone	dl-Norgestrel, Melengestrol acetate, Dehydroepiandrosterone, 8-Hydroxyquinoline, Genistein
NR3C2	352	0.5	2	Mifepristone	Fluocinolone acetonide, Bexarotene, 1-Naphthol, Dexamethasone, dl-Norgestrel
ABCB1	117	0.5	2	Reserpine	Fabesetron hydrochloride, Abamectin, SAR115740, SSR69071, Chlorobenzilate
NR3C1	148	0.5	4	Triamcinolone, Mifepristone	Medroxyprogesterone acetate, Fluorometholone, Melengestrol acetate, Dexamethasone, Prednisolone
CA1	176	0.5	4	Phenol, Sodium nitrite	Triclopyr, Triclopyr butotyl, p-Bromodiphenyl ether, 2-Fluoroacetamide, 1-Ethyl-2-methylbenzene
CA2	341	0.5	4	Celecoxib, Phenol	PharmaGSID_48509, Acenaphthylene, CP-105696, Aloe-emodin, 2-Fluoroacetamide
PTGS1	307	0.25	4	Indomethacin	SSR69071, 17alpha-Estradiol, Chlordane, Cetylpyridinium bromide, Zoxamide

Slide by Imran Shah

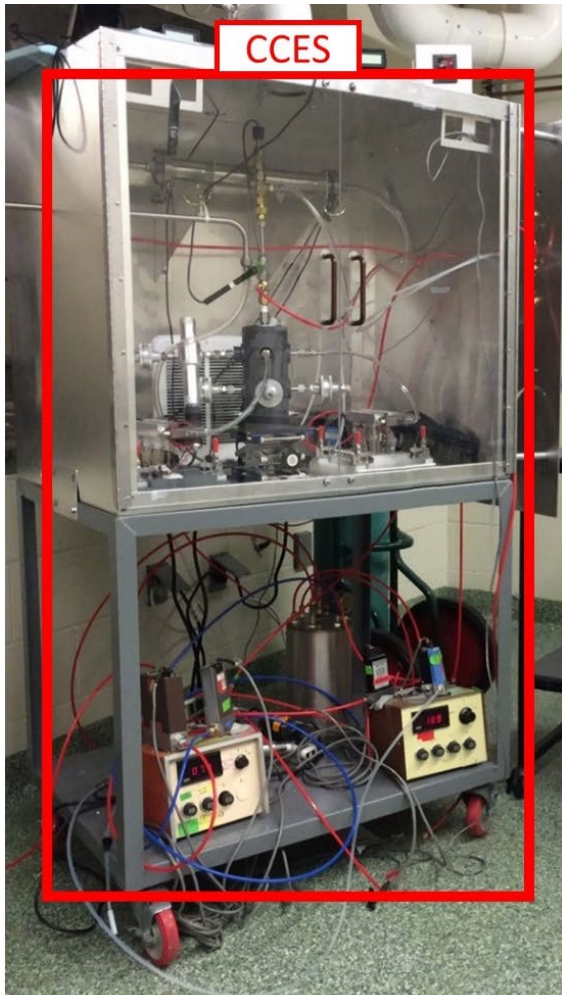
Table Courtesy of Imran Shah

Volatile Chemical Screening with HTTr

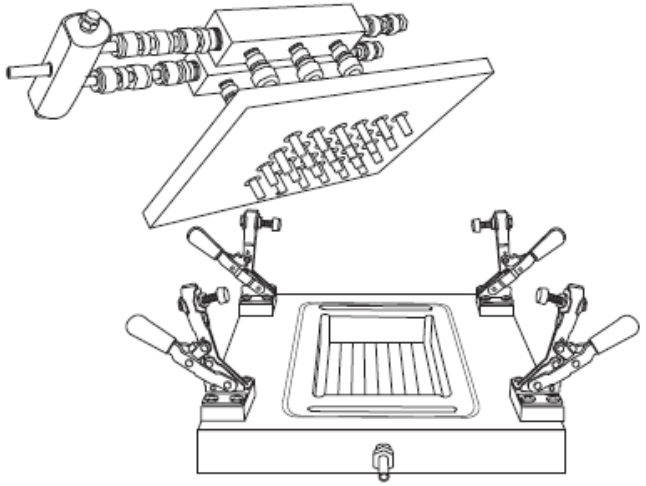
Cell Types	Primary Human Bronchial Epithelial Cells BEAS-2B cells
Test Chemical	1,3-Butadiene Acrolein Formaldehyde Acetaldehyde Trichloroethylene 1-Bromopropane Carbon Tetrachloride Dichloromethane
Test Concentrations	• n = 6, plus control
Exposure Duration	• 2 hours
Technical Replicates	• n = 2
Biological Replicates	• n = 3
Assay Formats	• TempO-Seq • Cytotoxicity [LDH Release]

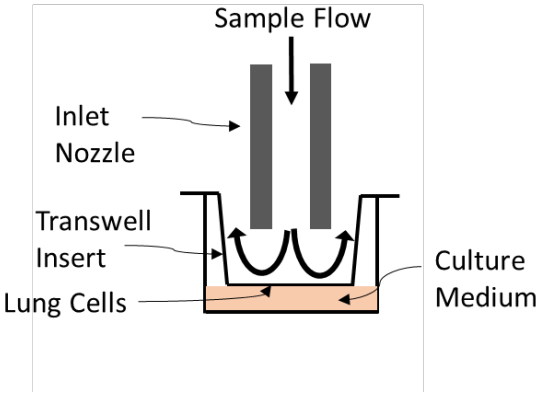
- Cells were cultured in 24 well format net wells at an air liquid interface.
- HPBE cells required ~3-4 weeks to differentiate in culture prior to testing.
- Cells were exposed to volatile chemicals using a custom designed exposure manifold developed by ***Mark Higuchi, Todd Krantz*** and ***Jose Zavala-Mendez***.

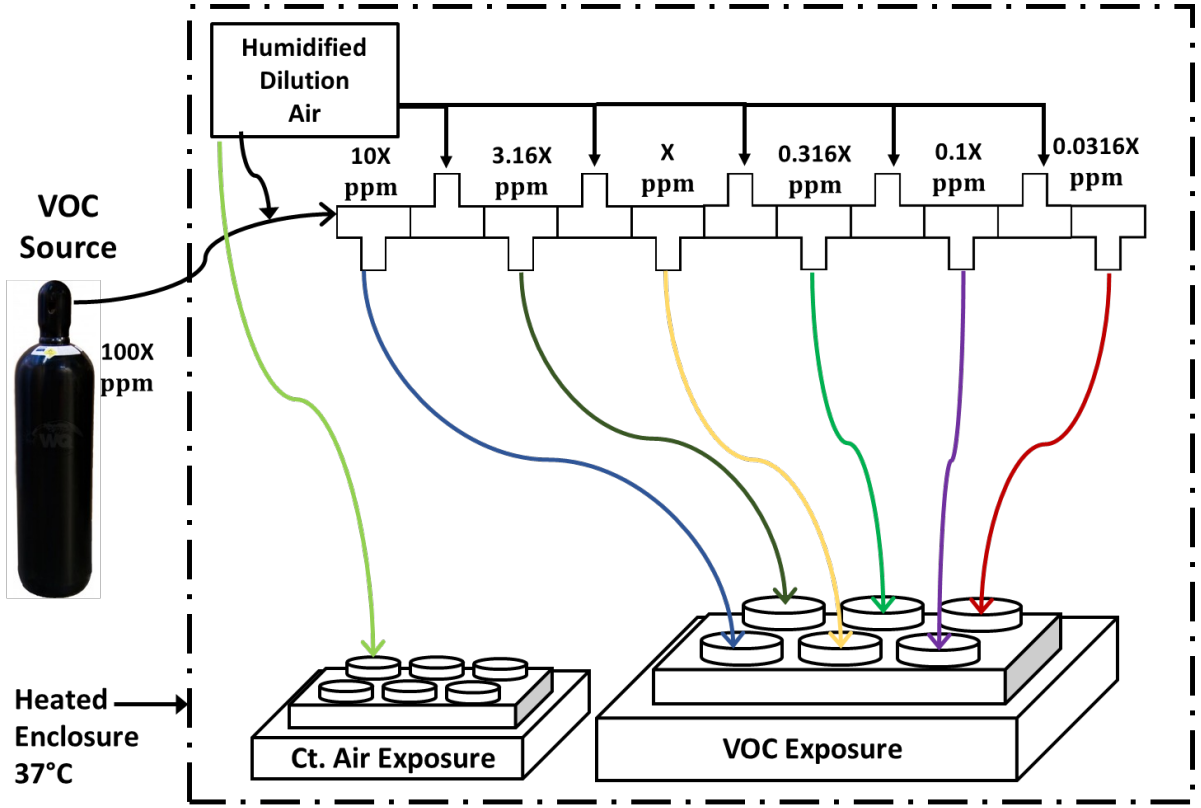
Cell Culture Exposure System (CCES)



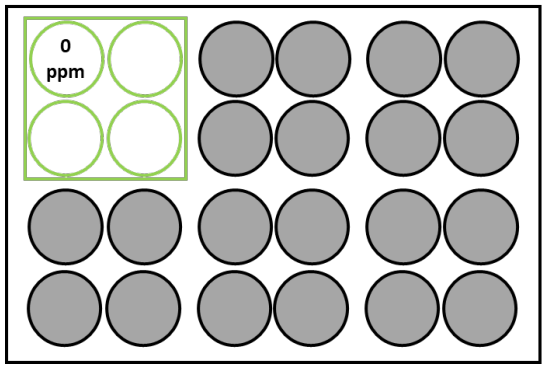
24-Well Format



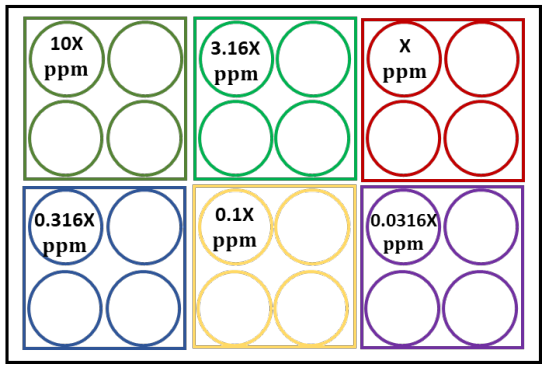




Control Air Exposure

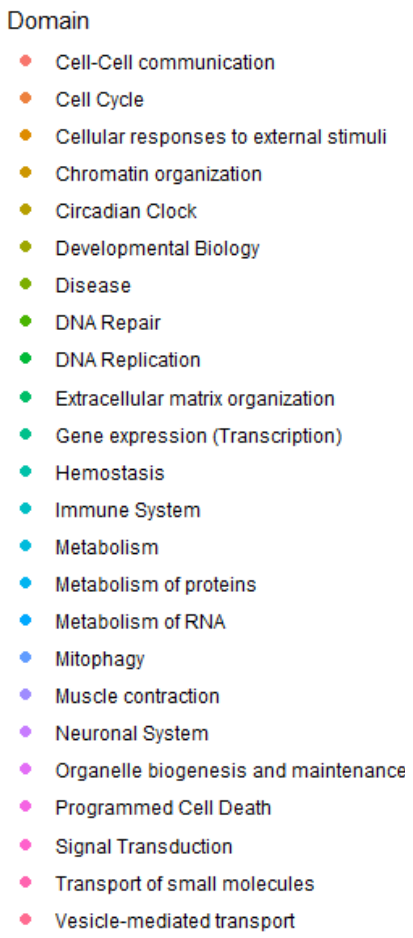


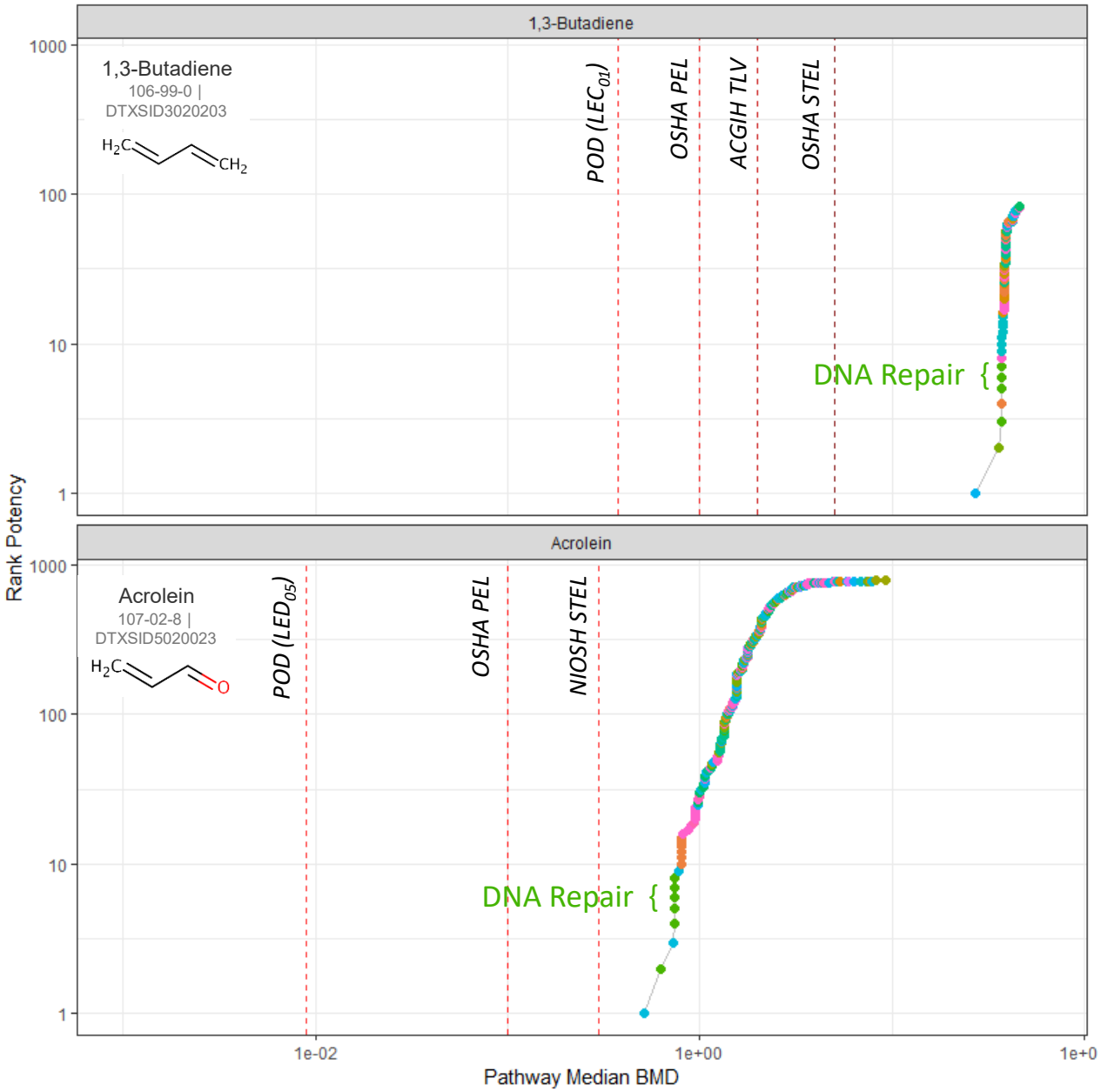
VOC Exposure

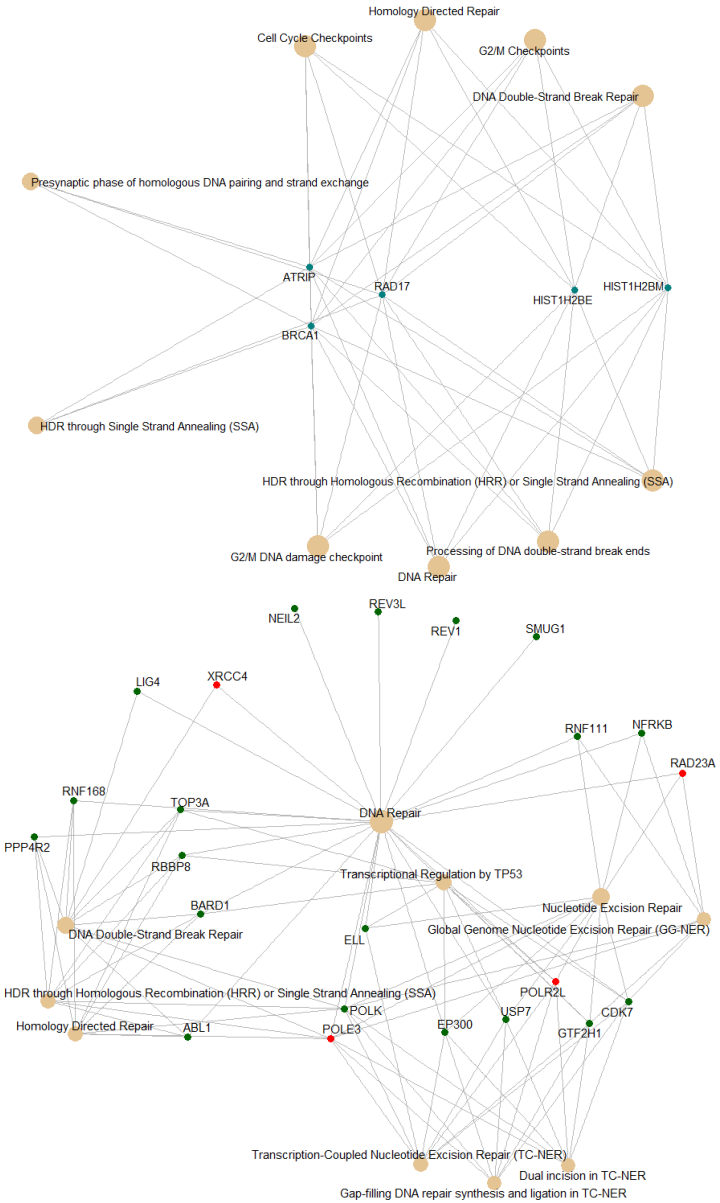


Diagrams courtesy of Mark Higuichi, Todd Krantz and Jose Zavala-Mendez

HTTr Volatile Screening Results







HTTr Summary Slide

- **Technology:** Targeted RNA-Seq based HTTr is a promising platform for comprehensive and cost-effective evaluation of chemically-induced disruption of biological processes/pathways.
- **Workflow:** We have developed a standardized, scalable and portable workflow to generate large-scale HTTr data for thousands of chemicals.
- **Concentration-Response Analysis:** Incorporation of concentration-response modeling into the analysis pipeline enables identification of transcriptional BPACs at the biological pathway/process level.
- **Bioactivity Exposure Ratio:** HTTr data may be used in combination with httk and ExpoCast estimates to identify chemicals with bioactivity thresholds in human relevant exposure ranges
- **MIE/MOA Identification:** Multiple analysis approaches are being investigated for identification of MIE/MOA. Target-centric signatures derived from annotated reference chemicals and machine learning techniques show promise for identification of putative MIE/MOAs.
- **Volatiles:** The TempO-Seq technology can be used to characterize transcription changes in cells grown in air-liquid interface and exposed to volatile compounds. Future work will attempt to use transcriptomic signatures from air-liquid interface cultures to distinguish systemic toxicants from local irritants.

Acknowledgments

NCCT:
Imran Shah
Richard Judson
Woody Setzer
Clinton Willis
Derik Haggard
Katie Paul-Friedman
Rusty Thomas

NHEERL:
Jose Zavala-Mendez
Todd Krantz
Jerrin Gause
Adam Speen
Mark Higuichi

National Toxicology Program:
Scott Auerbach

Sciome:
Jason Phillips
Ruchir Shah

BioSpyder
Joel McComb
Bruce Seligmann
Jo Yeakley
Milos Babic
Pete Shepard
Kyle LeBlanc

Unilever:
Paul Carmichael
Andy White
Paul Russell
Sharon Scott
Sophie Malcomber

National Center for Computational Toxicology

EPA United States Environmental Protection Agency

BONUS SLIDES

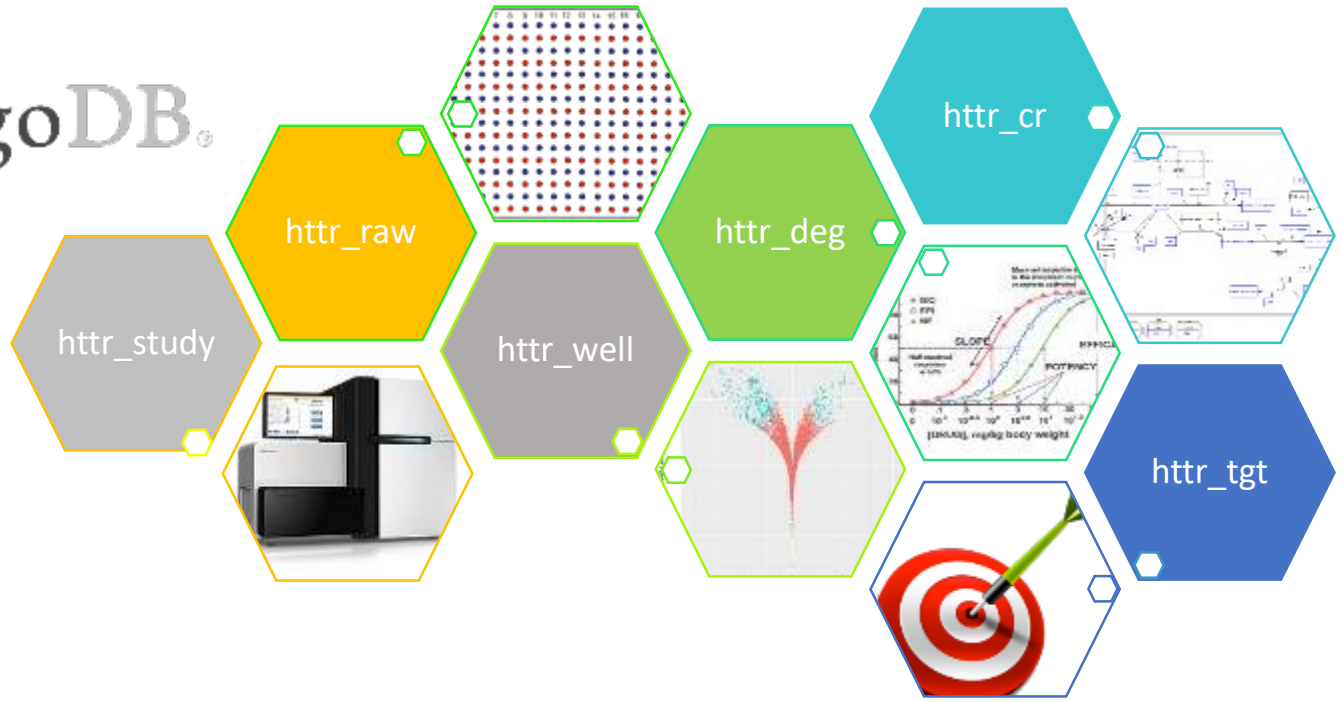
HTTr Computational Framework and Infrastructure

Python & R analysis pipeline

jupyter

TempO-Seq data processing → Differential expression analysis → Concentration response analysis → Target/ Pathway Prediction

mongoDB.



REST API

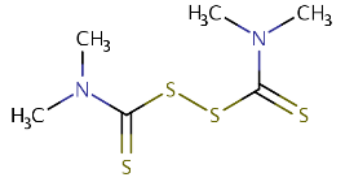
Flask

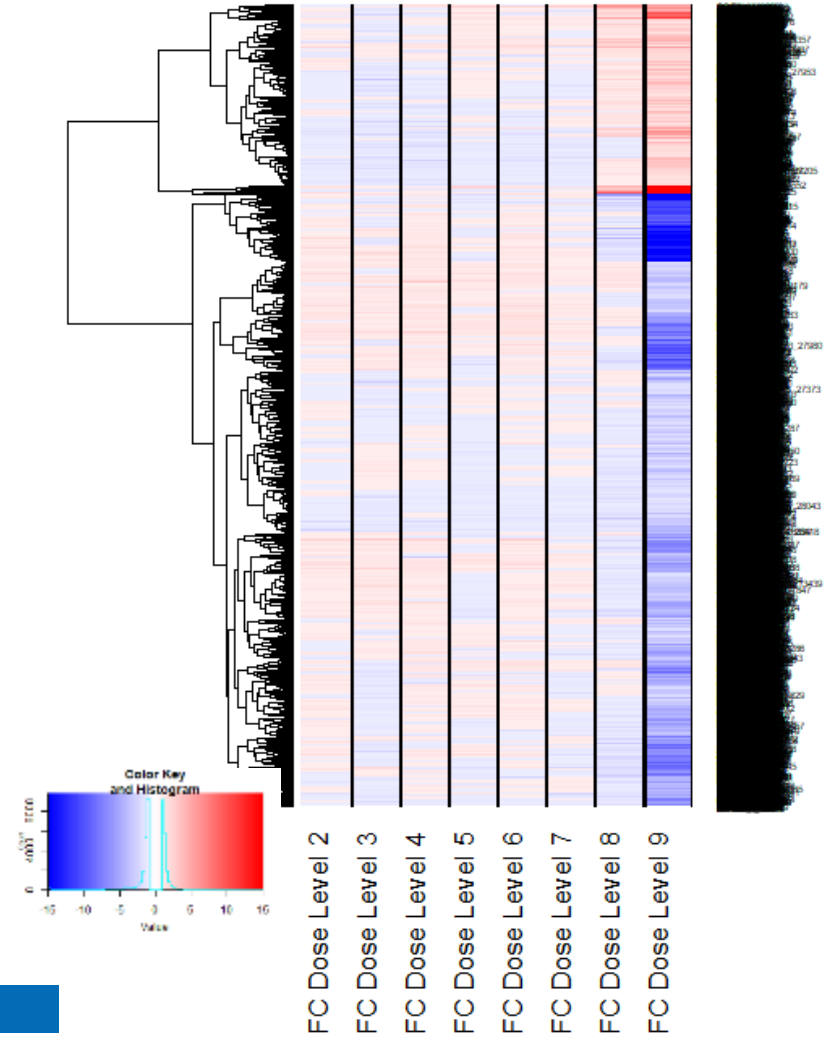
```
http://httr-dev.epa.gov/api/httr/v1/

searchChem
getChemPlates
getPlateInfo
getPlateGroups
getChemProbeCounts
getChemDEG

getChemCRG
getChemTargets
```

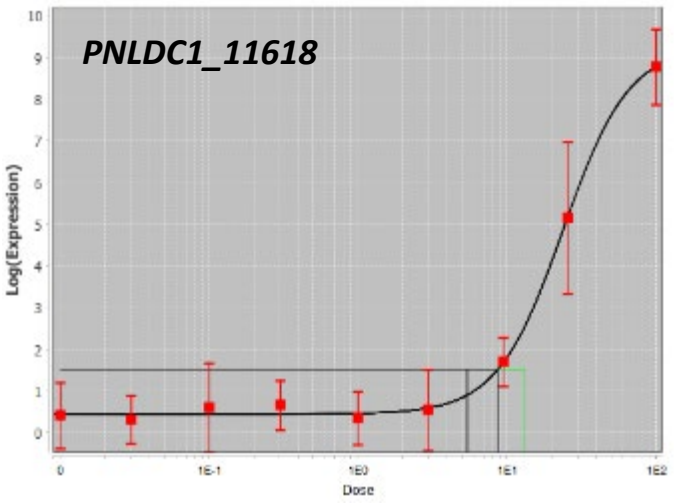
http://bitbucket.zn.epa.gov/projects/HTTR

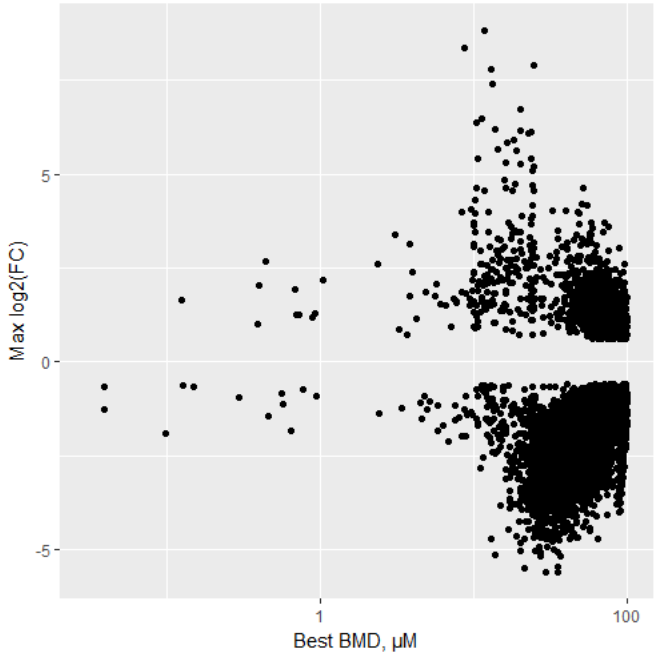
Concentration Response Modeling Example

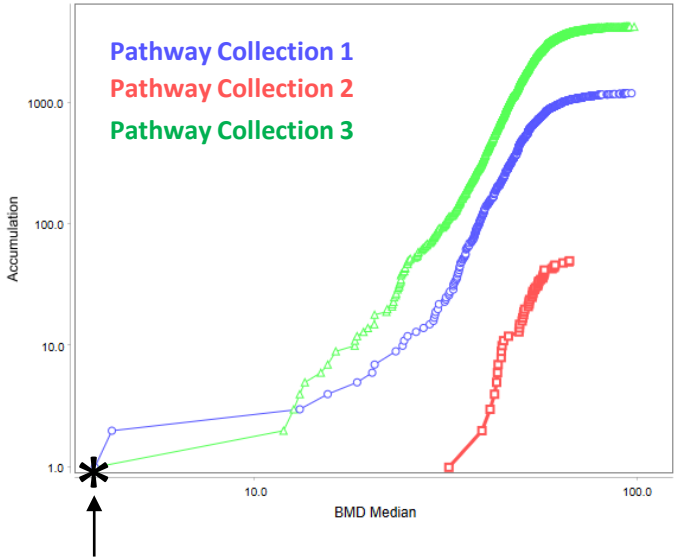
Thiram
137-26-8
DTXSID5021332







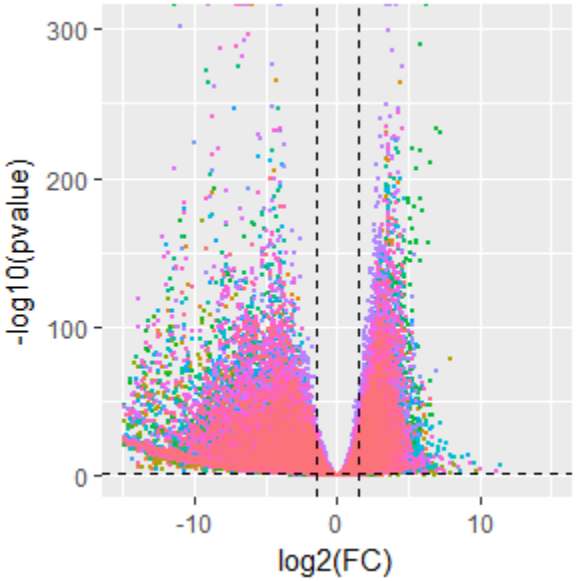




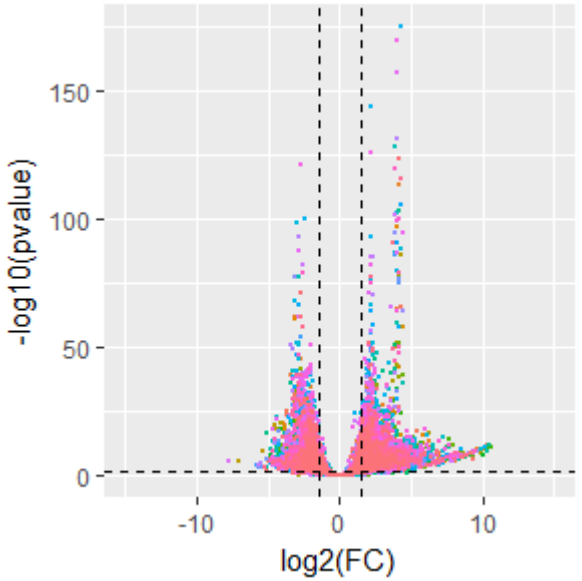
Biological Pathway Altering Concentration (BPAC)

Reproducibility of Log_2(FC) Estimates

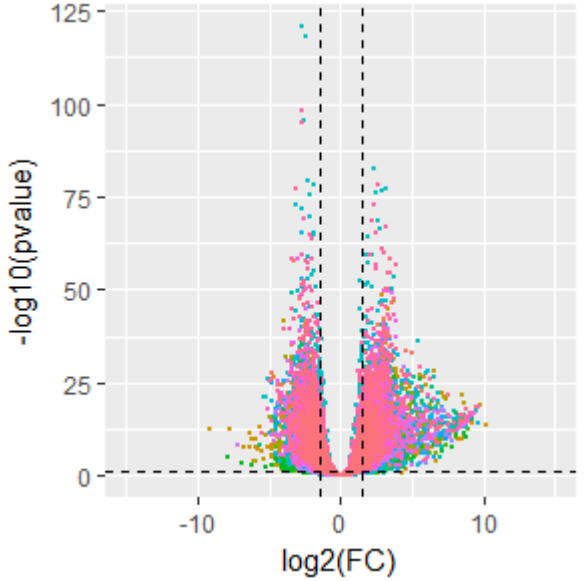
UHRR v HBRR

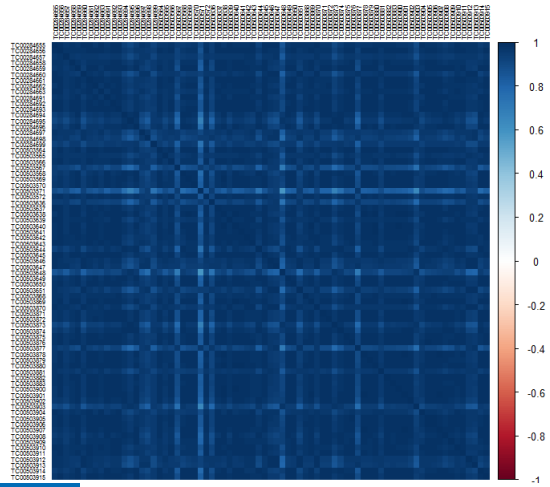


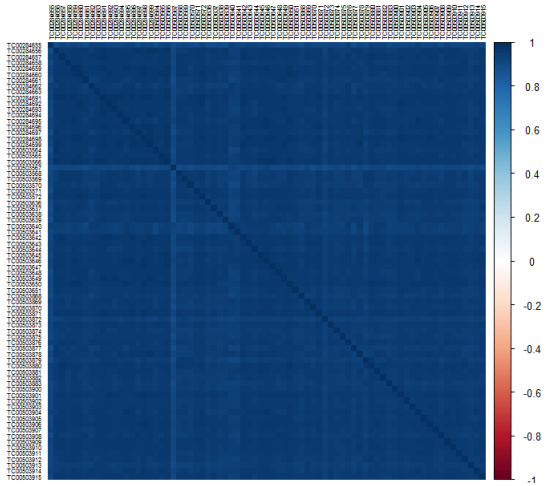
DMSO vs TSA (Bulk Lysate)

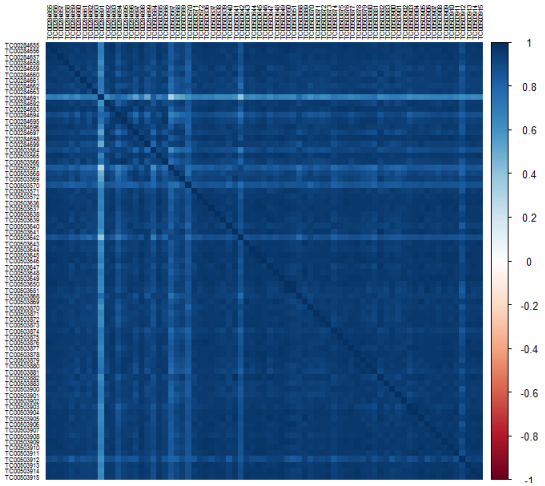


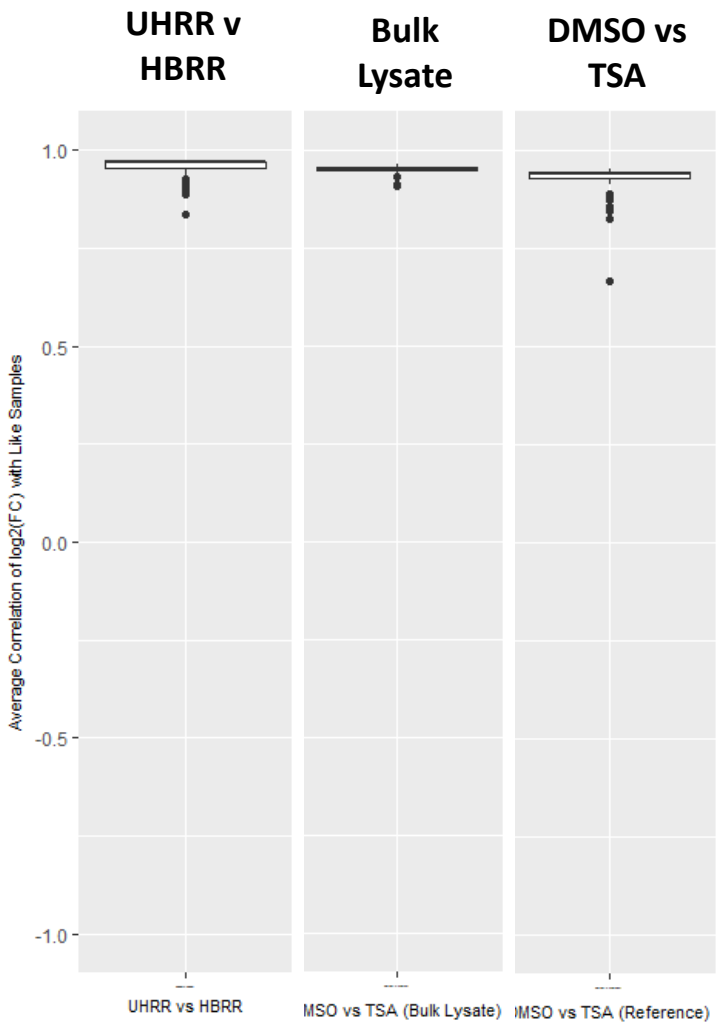
DMSO vs TSA (Treatment)











Comparisons across 144 assay plates
Blue Indicates positive correlation in FC estimates

United States
Environmental Protection
Agency