EPA
United States
Environmental Protection
Agency

Characteristics of the ToxCast In Vitro Datasets from Biochemical and Cellular Assays

Keith Houck
TDAS
May 14-15, 2009

This work was reviewed by EPA and approved for presentation but does not necessarily reflect official Agency policy. Mention of trade names or commercial products does not constitute endorsement or recommendation by EPA for use.

Office of Research and Development
National Center for Computational Toxicology

Outline: ToxCast Data Sets Overview

- Data Sets: Bio-centric View
  - Assay overview
  - How were the data calculated?
  - How did the replicates perform?
  - Example results
  - What did the assay measure?
  - Potential assay artifacts

ToxCast Phase I Data Sets

- Biochemical
  - Novascreen
- Cell-Based
  - ACEA
  - Attagene
  - BioSeek
  - Cellumen
  - CellzDirect
  - Gentronix
  - NCGC
  - Solidus

Novascreen: 239 Biochemical Assays (Abstract 59)

United States Environmental Protection Agency

- Protein super-families
  - GPCR
  - Kinase
  - Phosphatase
  - Protease
  - Ion channel
  - Nuclear receptor
  - Other enzyme
  - CYP P450 inhibition

- Various formats:
  - Radioligand receptor binding
  - Fluorescent receptor binding
  - Fluorescent enzyme substrate-intensity quench
  - Fluorescent enzyme substrate-mobility shift
- Initial screening:
  - 25 μM in duplicate
  - 10 μM in duplicate (CYPs)
- Normalize data to assay window
  - % of control activity (central reference – scalar reference)

Novascreen: What do the assays measure?

- Mainly direct effects of chemical on target protein
  - Enzyme activity
  - Ligand binding
- False positives:
  - Fluorescent compounds—fluorescing and quenching
  - Reactive compounds/covalent modification of target
  - Physical effects—colloid aggregation of target
  - Operational
- False negatives:
  - Solubility
  - Inappropriate enzymatic conditions
  - Operational
  - Target protein not physiological
  - Lack of biotransformation

Data Correction Algorithm Examples (Additive)

normalized

AChE

	1	2	3	4	5	6	7	8	9	10	11	12
A	-3	3	-29	-100	-98	-100	-68	-68	-89	-89	-98	-98
B	-4	7	-29	-12	-10	-12	-27	-22	-30	-30	-25	-16
C	8	2	-23	-23	-19	-26	-25	-26	-26	-28	-25	-13
D	-14	1	-17	-21	-23	-25	-18	-26	-26	-27	-25	-17
E	-11	-5	-23	-25	-16	-20	-26	-22	-27	-30	-26	-13
F	-16	0	-7	-7	-17	-21	-22	-24	-29	-33	-19	-17
G	-5	4	-32	-18	-25	-17	-25	-28	-25	-32	-20	-13
H	-85	-78	-22	-22	2	-14	-17	-17	NaN	NaN	NaN	NaN

Caspase 10

	1	2	3	4	5	6	7	8	9	10	11	12
A	25	0	-3	-97	-100	-101	-35	-33	-76	-80	-100	-99
B	9	-17	-16	-23	-28	-17	-13	-16	-22	-15	-22	-3
C	3	-22	-10	-27	-24	-20	-39	-31	-24	-28	-26	-6
D	15	-22	-15	-27	-23	-27	-13	-15	-17	-18	-20	-3
E	3	-16	-16	-17	-18	-7	-6	-13	2	5	-17	12
F	-11	-34	-18	-26	-29	-21	-22	-19	-18	-19	-25	0
G	-23	-30	-27	-24	-17	-6	-8	-11	-22	-12	-9	10
H	-15	-28	-17	-20	-14	-28	-5	4	-19	-4	-18	5

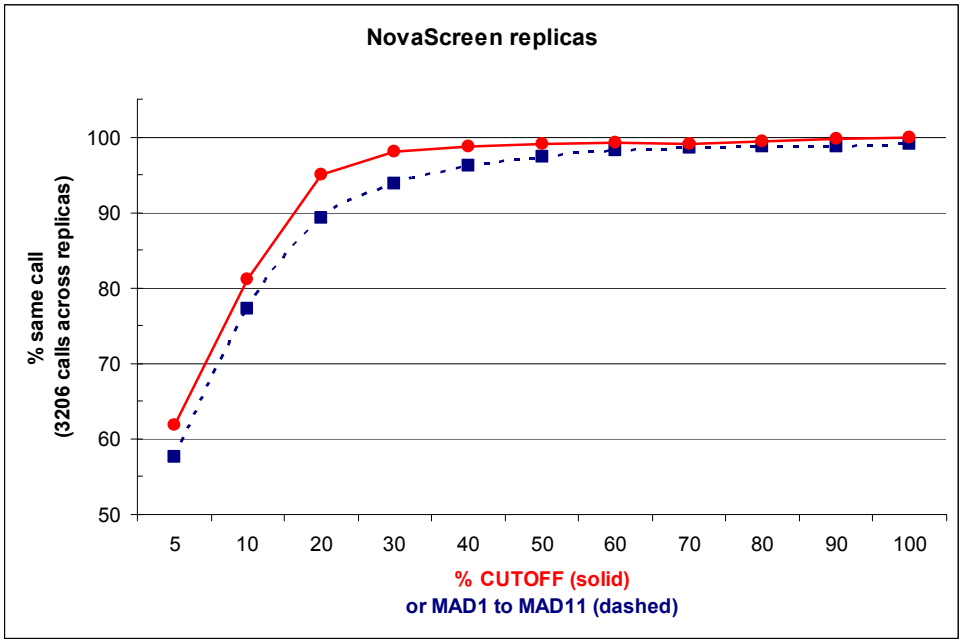
Normalized & corrected

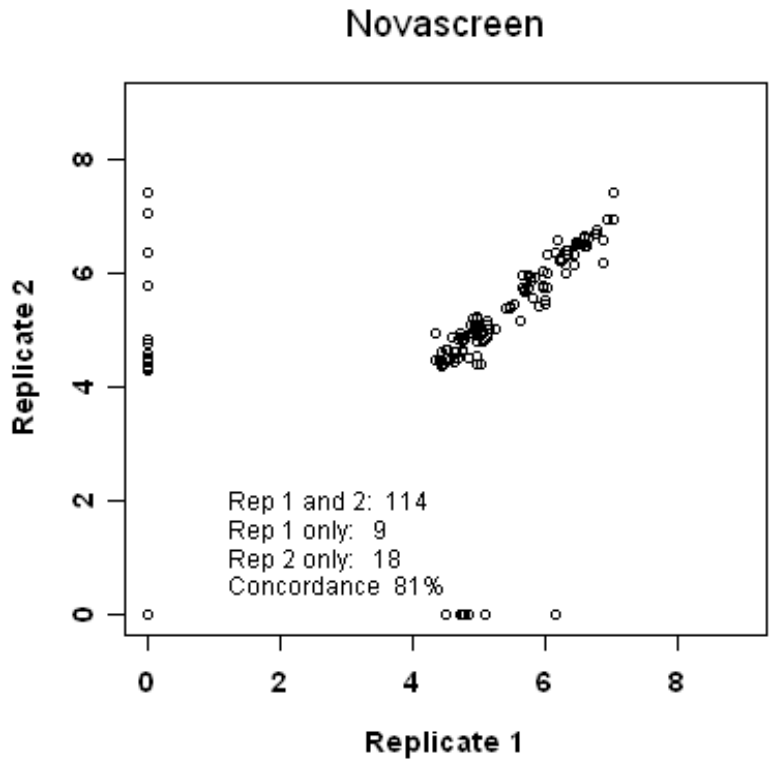
	1	2	3	4	5	6	7	8	9	10	11	12
A	-1	1	-19	-100	-98	-100	-67	-67	-89	-89	-98	-98
B	0	7	-20	-10	-2	0	2	9	-15	-16	-9	-4
C	10	10	-3	-3	0	-6	-2	-5	-16	-18	-11	-1
D	4	14	7	2	2	-5	6	0	-13	-11	-9	-3
E	7	13	4	0	7	4	-2	2	-15	-18	-10	-1
F	5	22	16	16	9	6	3	3	-15	-17	-3	-2
G	12	23	-4	6	1	5	3	0	-11	-16	-1	4
H	-52	-44	7	9	20	8	4	5	NaN	NaN	NaN	NaN

	1	2	3	4	5	6	7	8	9	10	11	12
A	23	0	-3	-97	-100	-101	-36	-34	-76	-81	-100	-99
B	13	6	7	0	-6	-3	3	2	-6	1	-5	0
C	6	-1	10	-7	-4	-5	-21	-13	-5	-10	-7	-5
D	18	0	2	-7	0	-7	6	4	0	-1	-4	1
E	1	1	2	-3	-5	-2	7	0	14	17	-6	9
F	-9	-12	4	-3	-11	-9	-4	-1	-7	-4	-11	-1
G	-17	-5	-5	-4	-1	-1	13	6	-2	3	5	8
H	-9	-11	2	0	4	-14	9	15	-4	10	-7	8

Novascreen Concentration-Response Testing

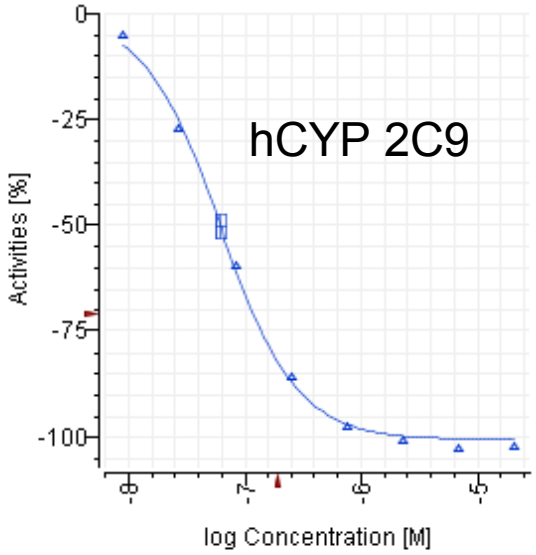
- Retest actives:
  - Median absolute deviation (MAD) median | *x-xmed* | two MADs or 30% activity
  - 8 conc/3-fold serial dilutions
    - 50 μM high conc
    - 25 μM high conc for CYPs
- Normalize to assay window
- Fit % Activity data to 3- or 4-parameter Hill function
  - Sometimes had to fix top or bottom of curve
  - Did not extrapolate beyond testing range
  - Manual or automated removal of obvious outliers

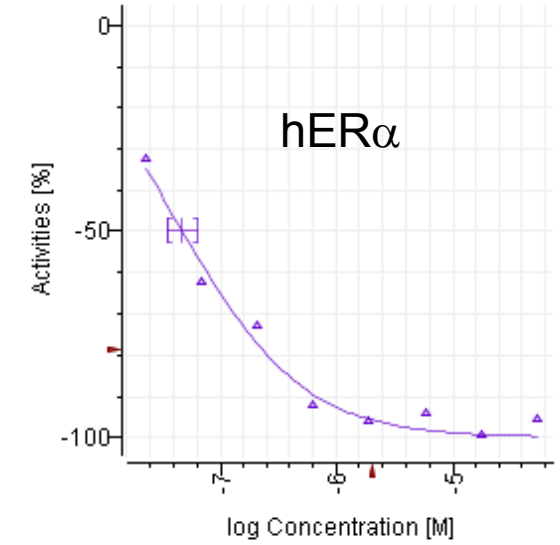


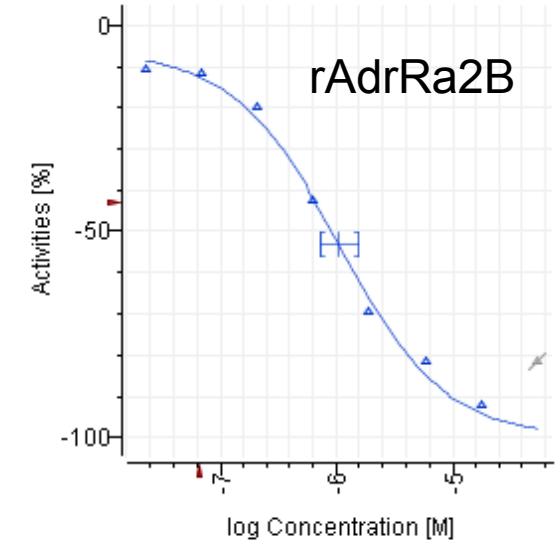


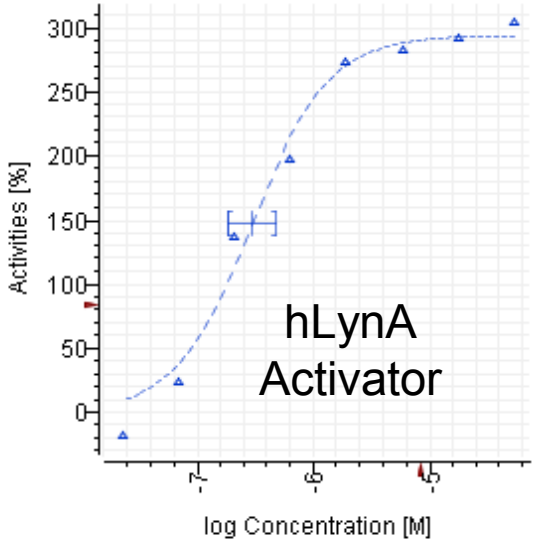
Novascreen: Example Curve Fits

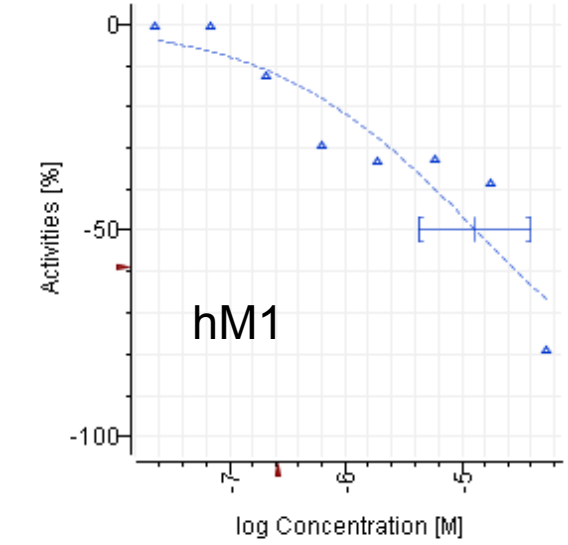
United States Environmental Protection Agency

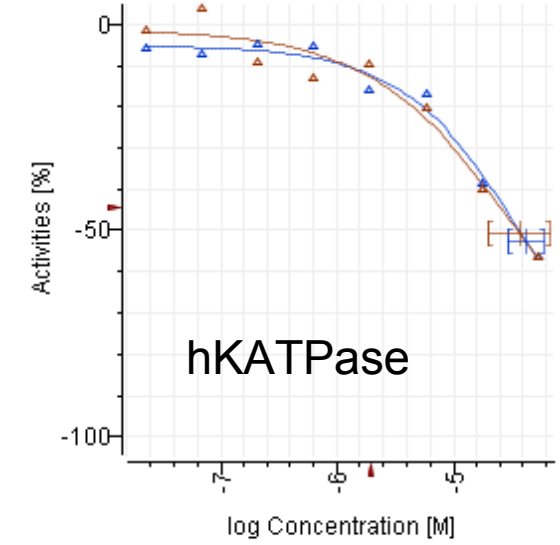








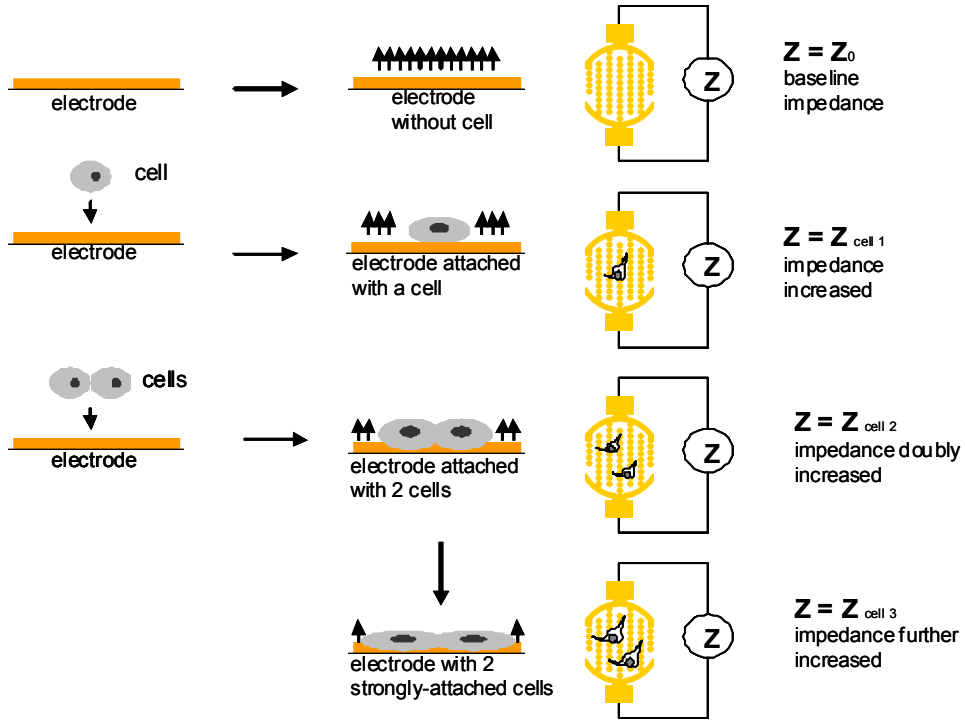




ACEA: Real Time Cell Growth Kinetics

United States Environmental Protection Agency

- Cytotoxicity with potential mechanistic interpretation
- Human A549 lung carcinoma cell line
  - ACEA experience with line
  - Reference compound effects
- Concentration-response testing
  - 8 conc/3-fold serial dilutions
  - Duplicate wells
- Real-time measurements during exposure (0-72 hr)
- IC50 and LECs calculated



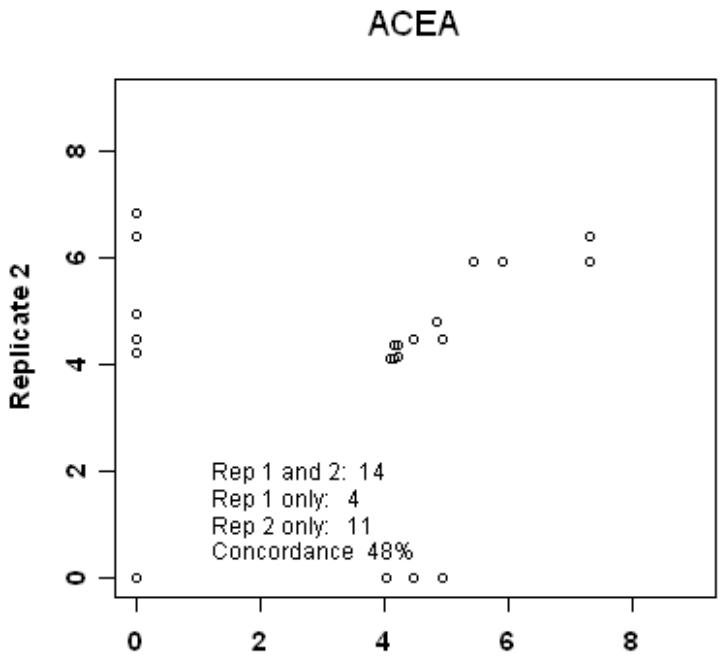
ACEA: What is measured?

- General cytotoxicity in transformed cell line
- False positives
  - None obvious
  - Operational
- False negatives
  - Operational error
  - Solubility
  - Lack of appropriate toxicity targets (irrelevant cell line)
  - Lack of biotransformation

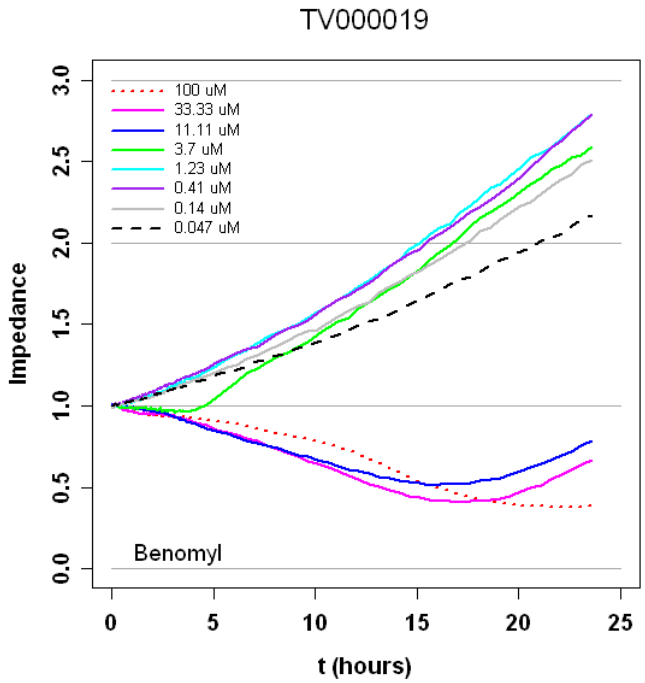
ACEA: Data Examples

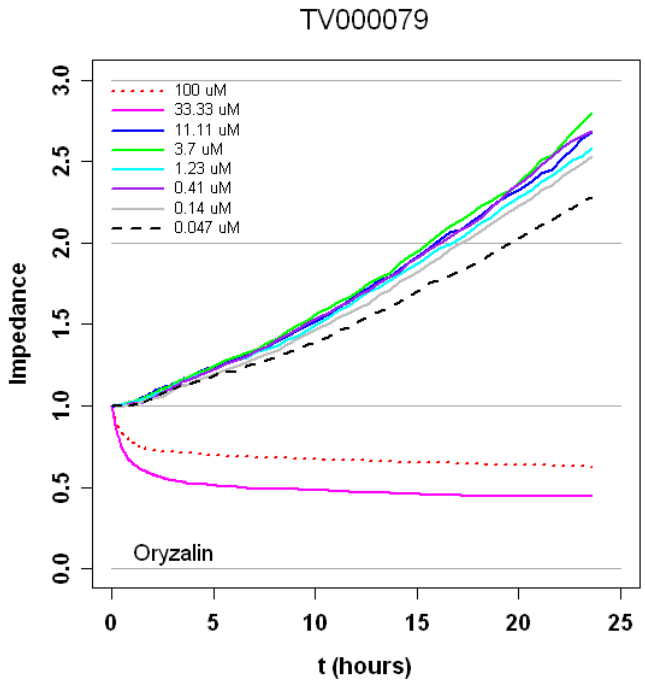
United States Environmental Protection Agency

Replicate Analysis:



Example Plots:





Attagene Multiplexed Transcription Reporter Assays

- Modulation of TF activity in human hepatoma HepG2 cells
- Multiplexed reporter gene assay
  - *cis* 52 assays (response element driving reporter)
  - *trans* 29 assays (GAL4-NR_LBD driving reporter) “ligand detection”
- IC50 for cytotoxicity measured first in HepG2
- High concentration either 100 μM or 1/3 calculated IC50 for cytotoxicity
- Seven concentrations, 3-fold serial dilutions, 24 hr exposure
- Cells harvested, RNA isolated, processed for reporter gene quantitation
- LEC provided in data set

Attagene Technology

United States Environmental Protection Agency

Multiplexed Reporter Gene Assay

Library of RTUs

Cell Transfection

Transcription

RNA Isolation

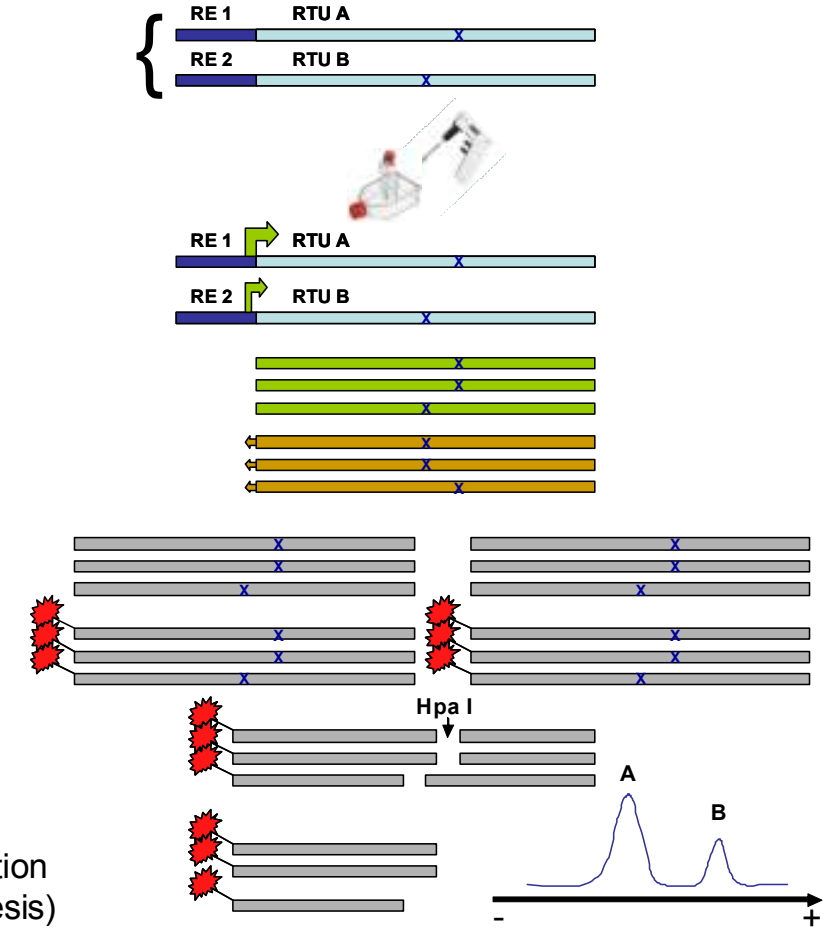
Reverse transcription

PCR amplification

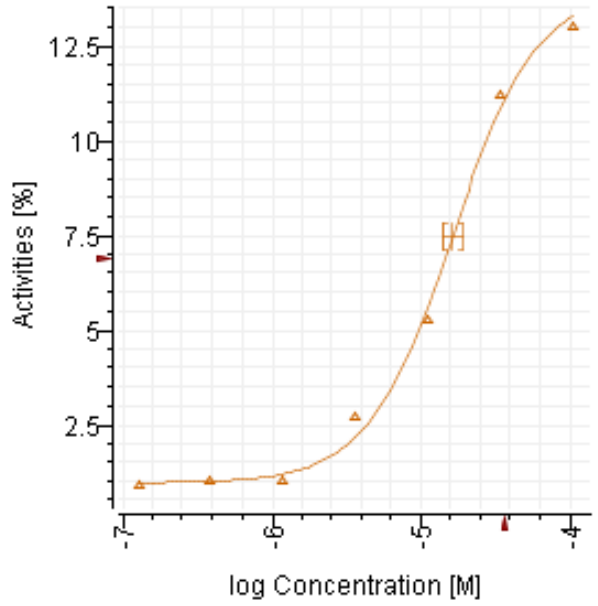
Labeling

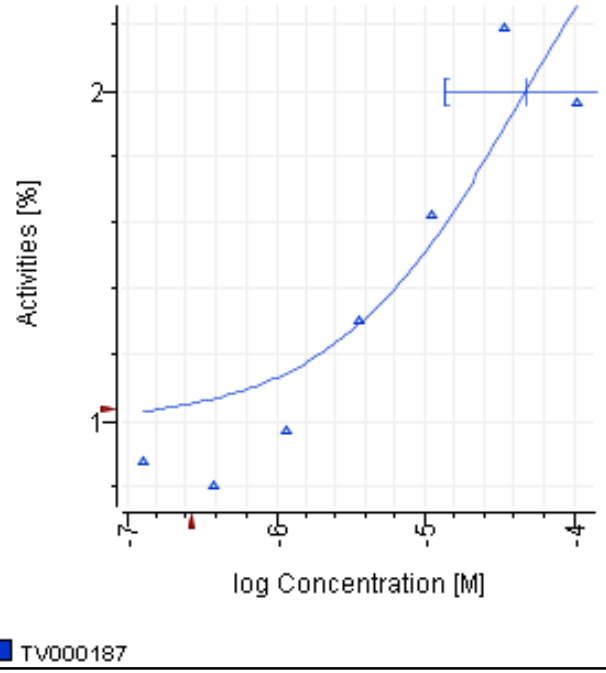
Processing (Hpa I)

Separation and detection (capillary electrophoresis)



Cis: AhR





Attagene: What Is Being Measured?

- cis Assays
  - Up/down regulation of endogenous transcription factor activity in transformed cell line
  - False positives
    - General cytotoxic response resulting in non-specific transcriptional activity
    - Promiscuity of response elements
    - Statistical, not biologically, significant response
    - Operational
  - False negatives
    - Solubility
    - Cytotoxicity
    - Operational
    - Lack of endogenous machinery
    - Lack of biotransformation
- trans Assays
  - NR agonist activity
  - False positives
    - General cytotoxic response resulting in non-specific transcriptional activity
    - Statistical, not biologically, significant response
    - Operational
  - False negatives
    - Solubility
    - Cytotoxicity
    - Operational
    - Lack of endogenous machinery
    - Lack of biotransformation

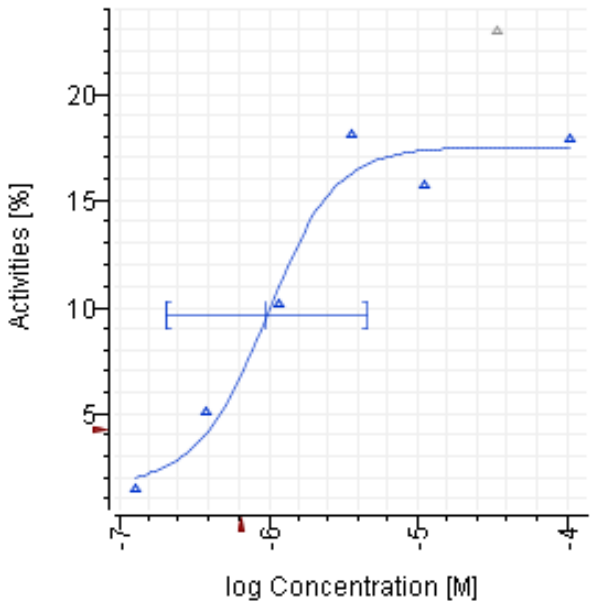
Attagene: Corresponding *cis* and *trans* Assays

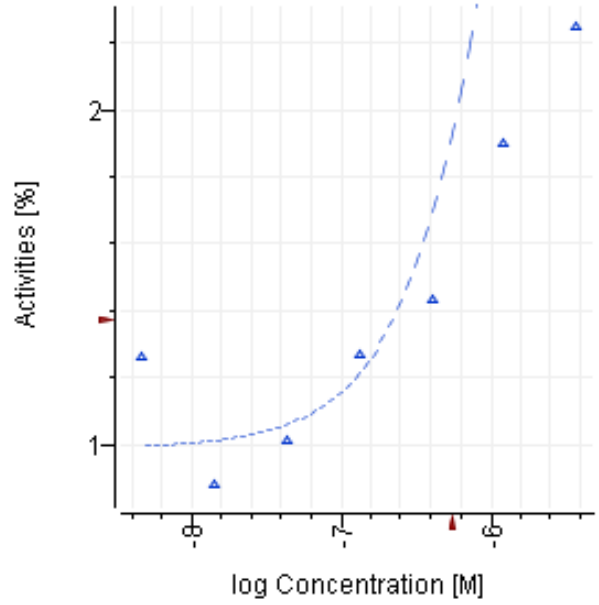
United States Environmental Protection Agency

Bisphenol A

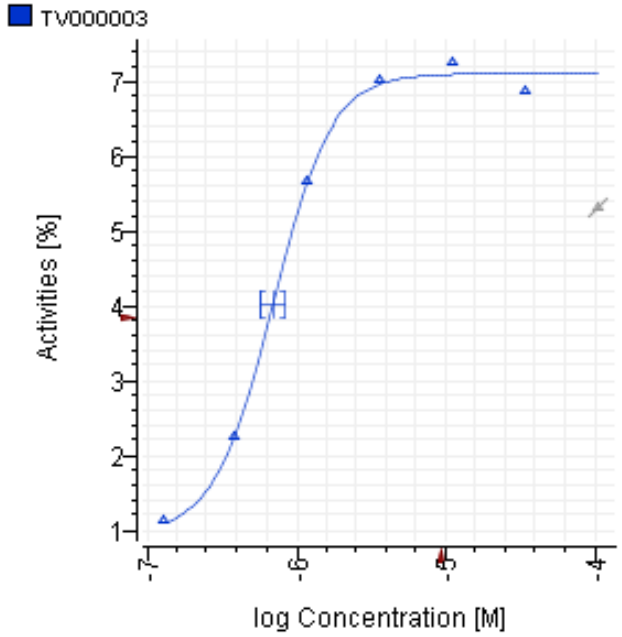
HPTE

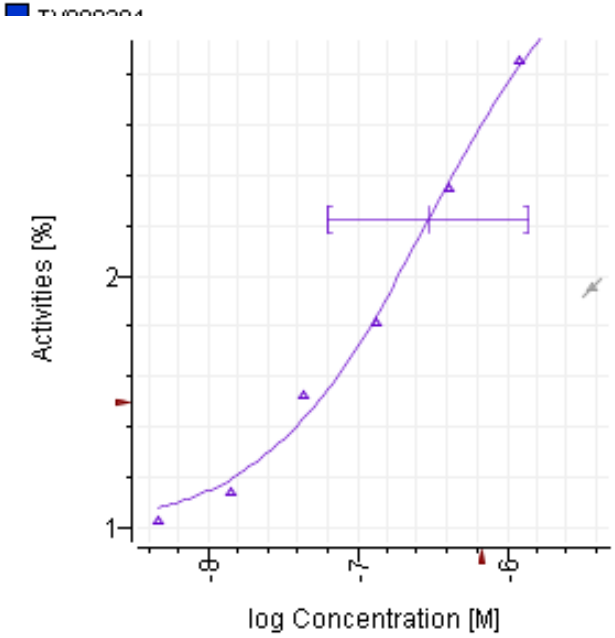
***trans*: ERa**





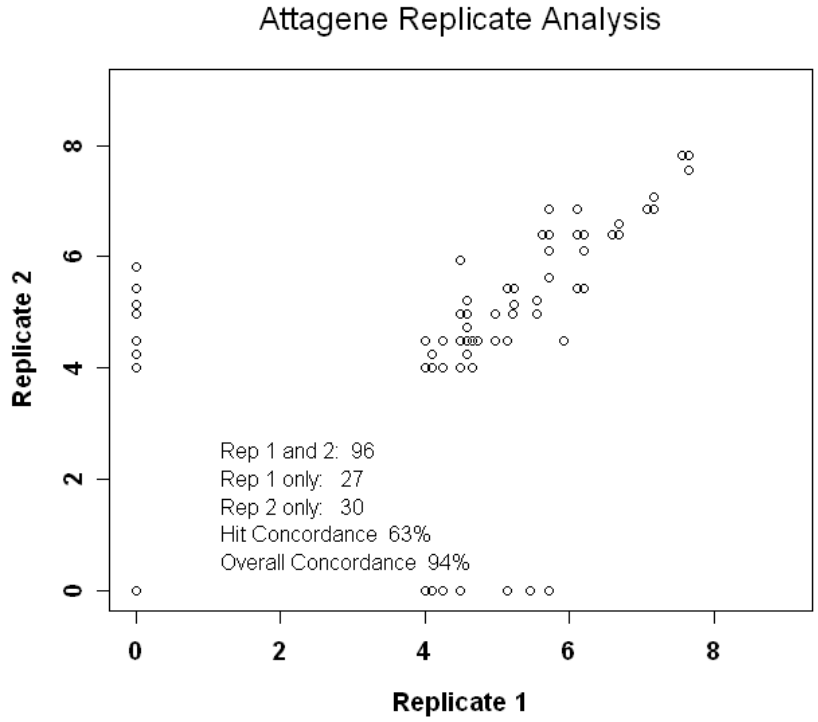
***cis*: ERE**

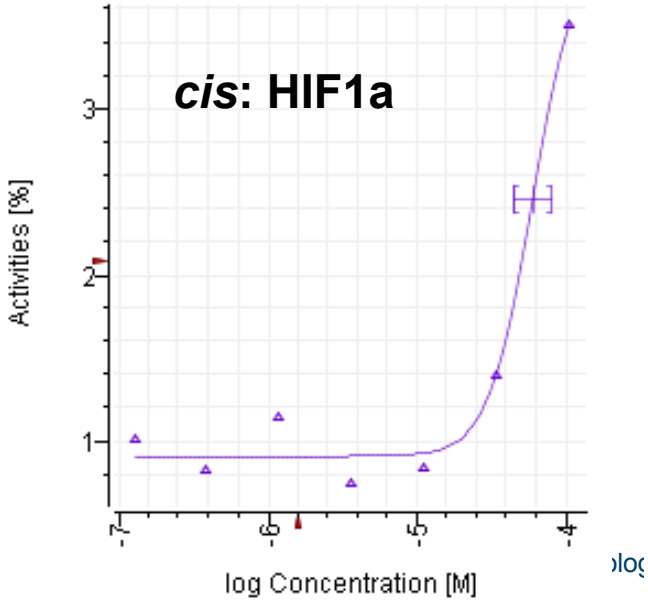


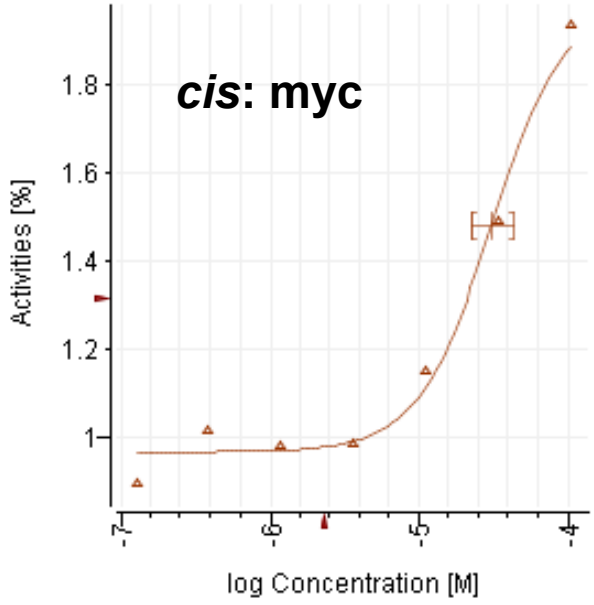


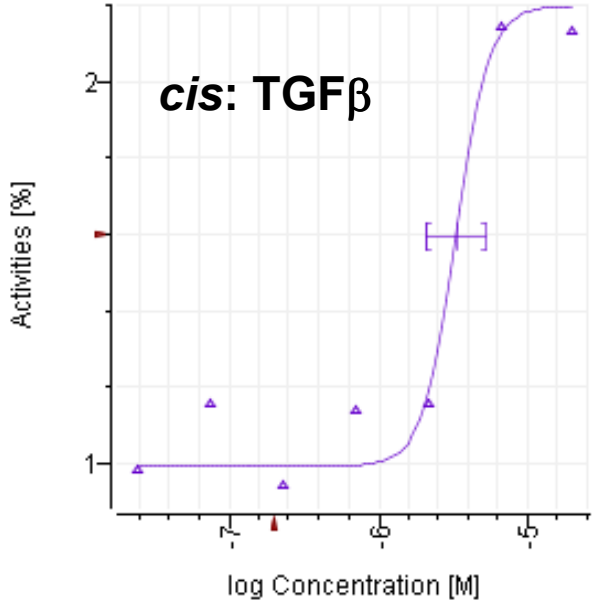
Attagene: Data Calculation Challenges

- No positive reference compound for each endpoint
- Responses, especially for *cis* assays, tended to be monotonic so potency value difficult to derive
- Biological vs statistical significance for LEC







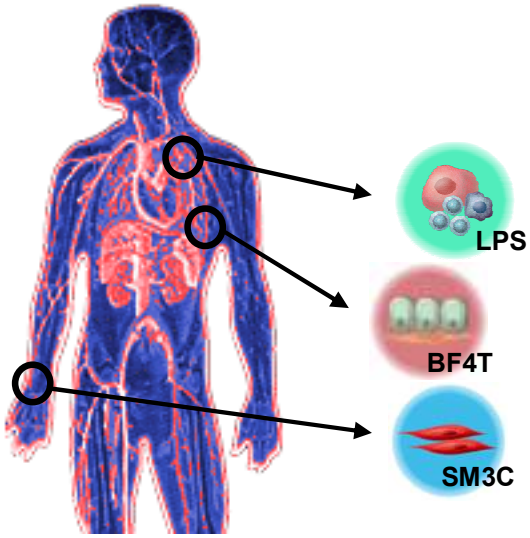


TV000278 TV000273 TV000142

BioSeek: BioMAP® Technology Platform (Abstract 24)

United States Environmental Protection Agency

Assays

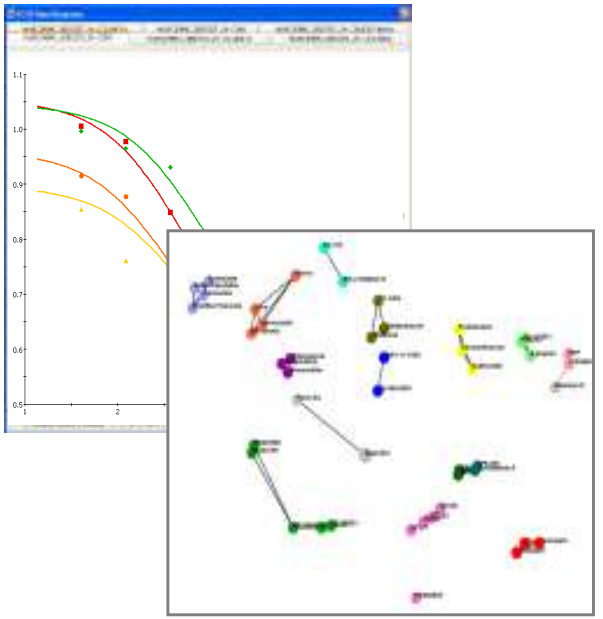


Human primary cells
Disease-like culture conditions

Profile Database

Biological responses to drugs and stored in the database

Informatics



Specialized informatics tools are used to mine and analyze biological data

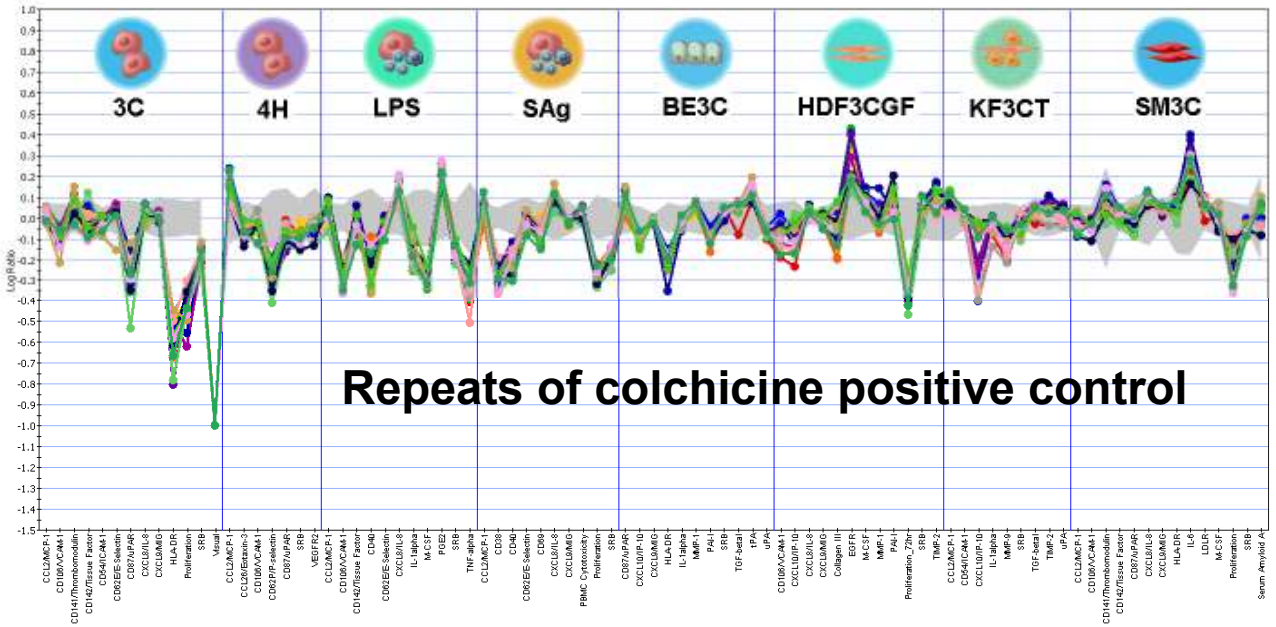
Primary Human Cell-Based Assay Platform for Human Pharmacology

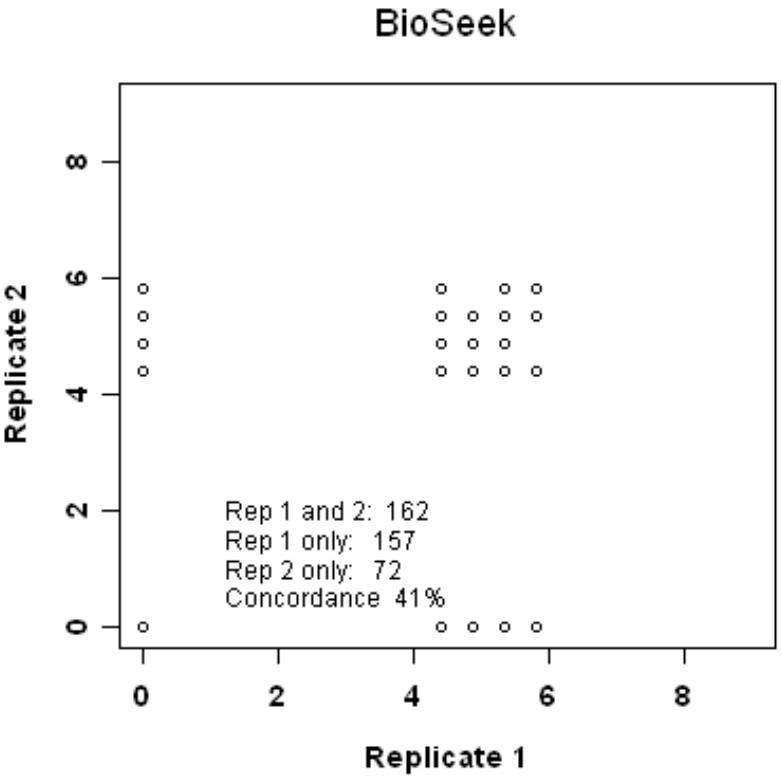
BioSeek Assays Tested Against ToxCast_320

System		Cell Types	Environment	Readouts
3C		Endothelial cells	IL-1β+TNF-α+IFN-γ	MCP-1, VCAM-1, ICAM-1, Thrombomodulin, Tissue Factor, E-selectin, uPAR, IL-8, MIG, HLA-DR, Prolif., Vis., SRB (13)
4H		Endothelial cells	IL-4+histamine	VEGFRII, P-selectin, VCAM-1, uPAR, Eotaxin-3, MCP-1, SRB (7)
LPS		Peripheral Blood Mononuclear Cells + Endothelial cells	TLR4	CD40, VCAM-1,Tissue Factor, MCP-1, E-selectin, IL-1a, IL-8, M-CSF, TNF-a, PGE2, SRB (11)
SAg		Peripheral Blood Mononuclear Cells + Endothelial cells	TCR	MCP-1, CD38, CD40, CD69, E-selectin, IL-8, MIG, PBMC Cytotox., SRB, Proliferation (10)
BE3C		Bronchial epithelial cells	IL-1β+TNF-α+IFN-γ	uPAR, IP-10, MIG, HLA-DR, IL-1a, MMP-1, PAI-1, SRB, TGF-b1, tPA, uPA (11)
HDF3CGF		Fibroblasts	IL-1β+TNF-α+IFN-γ +bFGF+EGF+PDGF-BB	VCAM-1, IP-10, IL-8, MIG, Collagen III, M-CSF, MMP-1, PAI-1, Proliferation, TIMP-1, EGFR, SRB (12)
KF3CT		Keratinocytes + Fibroblasts	IL-1β+TNF-α+IFN-γ +TGF-β	MCP-1, ICAM-1, IP-10, IL-1a, MMP-9, TGF-b1, TIMP-2, uPA, SRB (9)
SM3C		Vascular smooth muscle cells	IL-1β+TNF-α+IFN-γ	MCP-1, VCAM-1, Thrombomodulin, Tissue Factor, IL-6, LDLR, SAA, uPAR, IL-8, MIG, HLA-DR, M-CSF, Prolif., SRB (14)

BioSeek Assays

- Chemicals tested at 4 concentrations: 40, 13.3, 4.4, 1.5 μM, single well
- Exposure started 1 hr before stimulation of cell signaling pathways
- Following 24 hr exposure, endpoints measured by ELISA (also Alamar blue, SRB staining, and microscopy)
- Data calculated vs 6 DMSO controls as fold-change
- Up and down regulation distinguished
- LECs determined





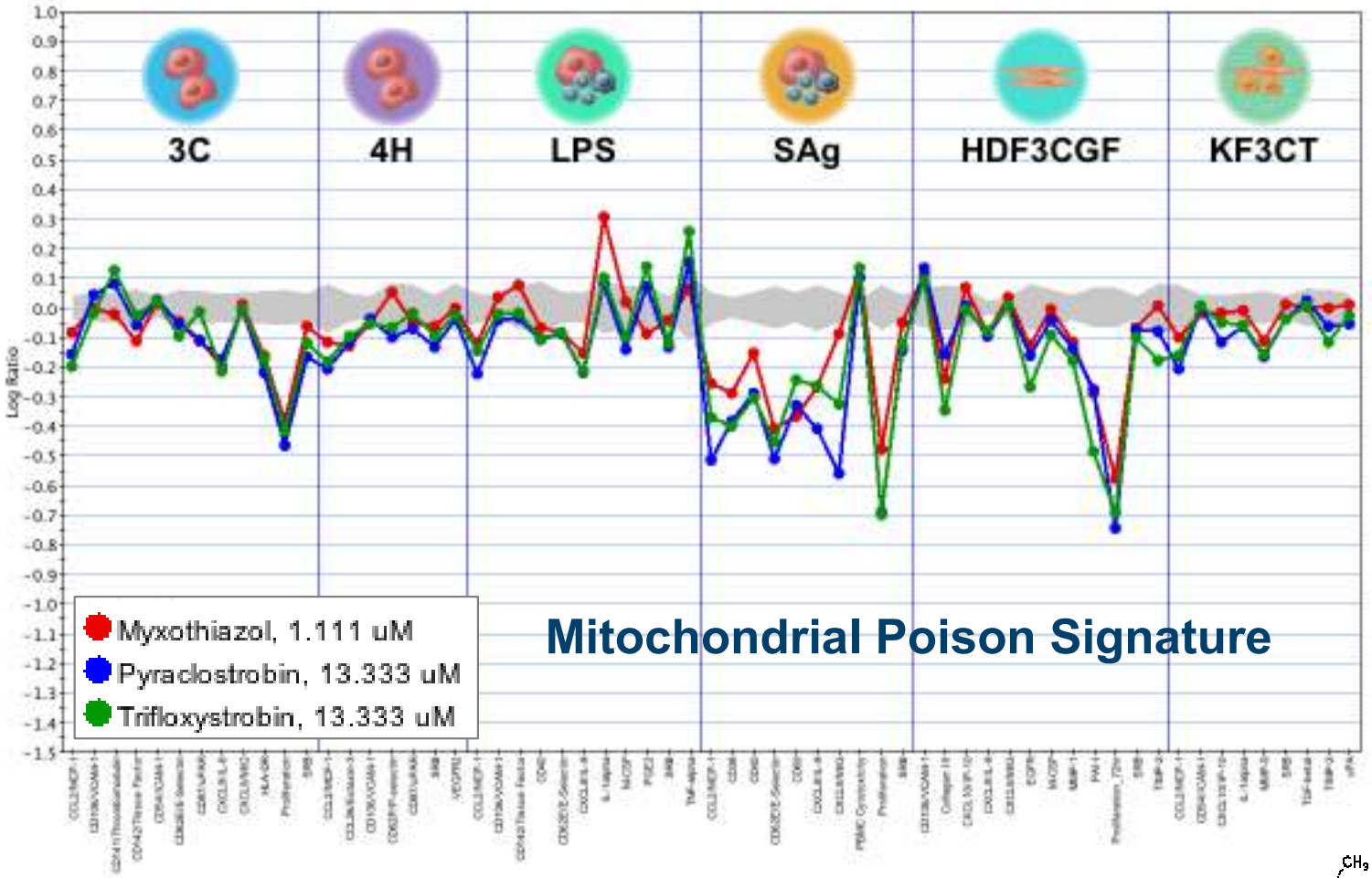
BioSeek Assays: What is Being Measured?

- Effects of chemicals on signaling pathways in primary human cells
- False positives
  - Cytotoxicity (down-regulation in particular)
  - Statistical vs biological significance
  - Operational
- False negatives
  - Solubility
  - Lack of biotransformation
  - Cytotoxicity
  - Operational

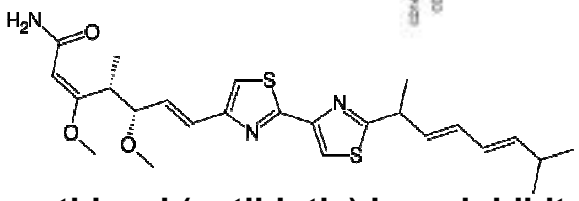
BioSeek MOA Signatures

Defined by reference cmpd behavior in assays used

United States Environmental Protection Agency

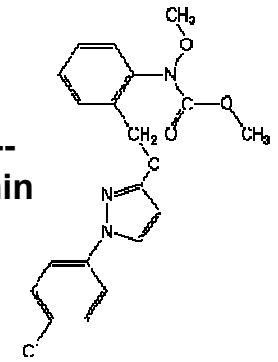


Mitochondrial Poison Signature

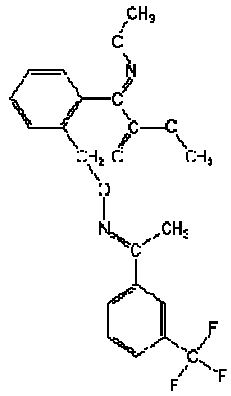


Myxothiazol (antibiotic) is an inhibitor of coenzyme Q - cytochrome c reductase (aka complex III). It prevents electron transfer from Fe-S center to ubiquinone in the electron transport chain.

Strobilurin fungicides: MOA-- blocks electron transfer within respiratory chain



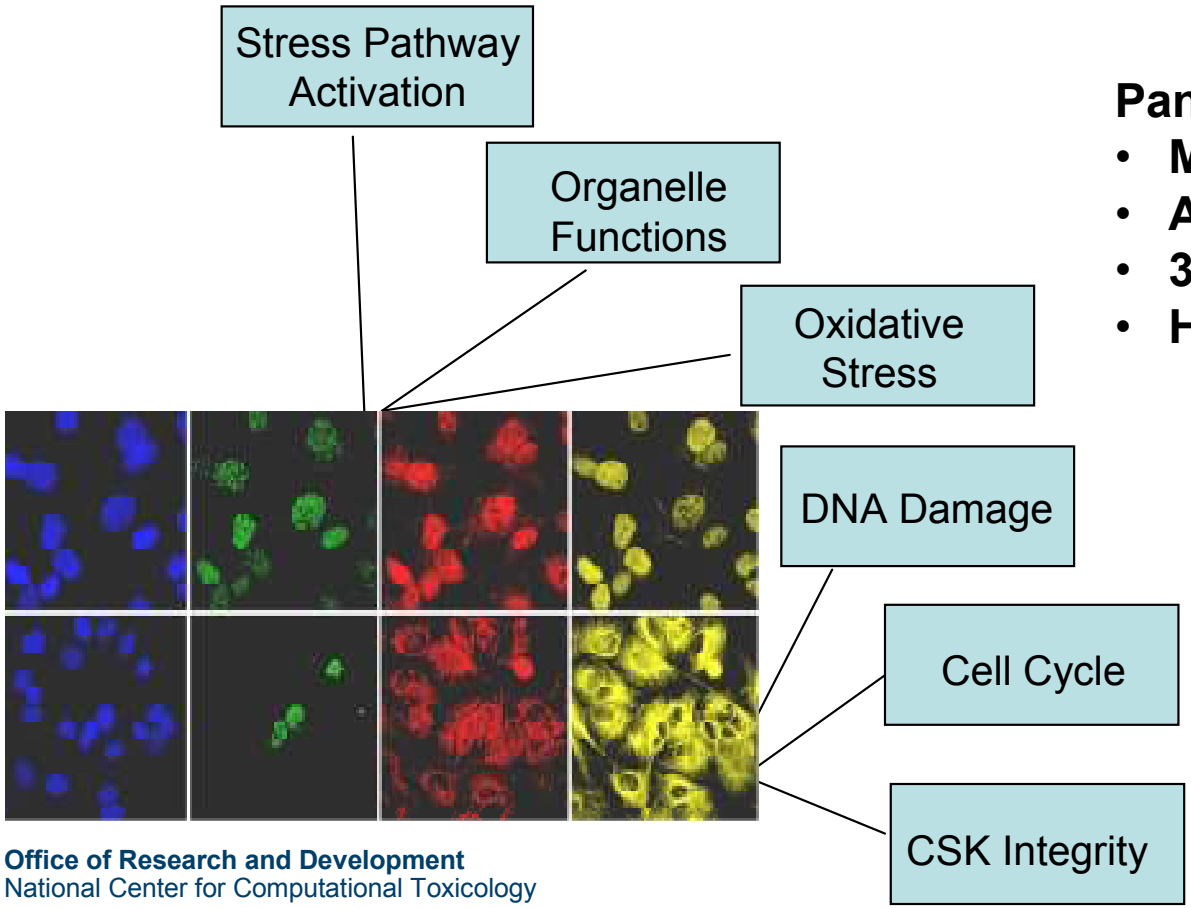
Pyraclostrobin



Trifloxystrobin

Cellumen: High-Content Screening of Cellular Phenotypic Toxicity Parameters (Abstract 38)

- Technology: automated fluorescent microscopy
- Objective: Determine effects of chemicals on toxicity biomarkers in a cell culture of HepG2 and primary rat hepatocytes



Panel 1 design*:

- **Multiple mechanisms of toxicity**
- **Acute, early & chronic exposure**
- **384-well capacity**
- **HepG2**

Cellumen: CellCiphr™ Cytotoxicity Panel

- 10-point conc-response (200 μM-39 nM) in duplicate
- Three time points (1 hr, 24 hr, 72 hr)
- 11 endpoints per assay

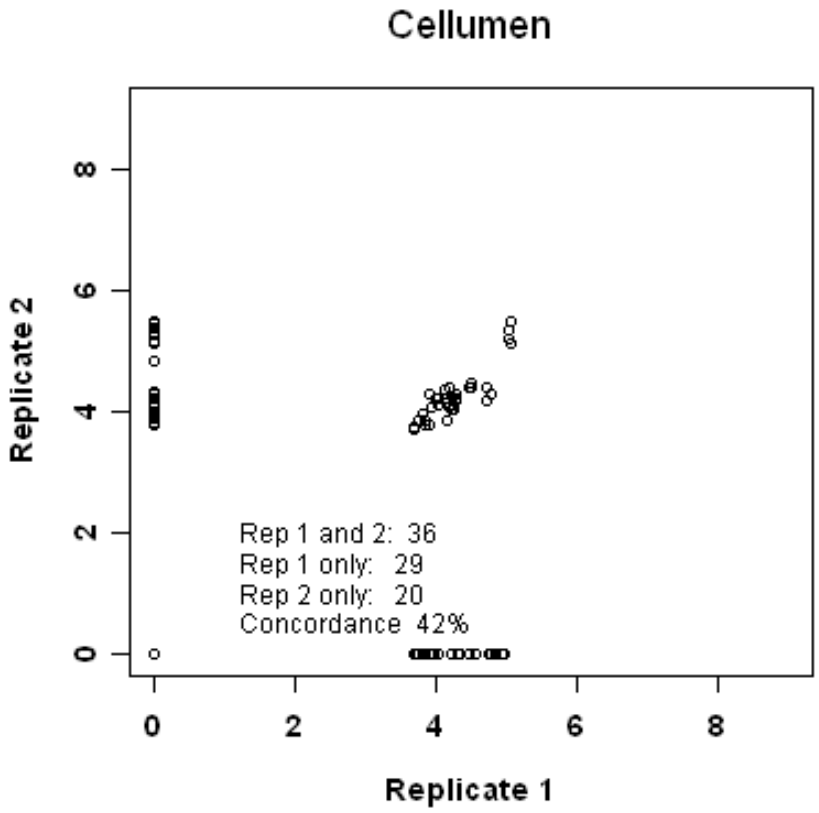
Biomarker	Measurement	Positive Control	Z'
Stress Pathway	Phospho-c-jun	Anisomycin	0.63
Oxidative Stress	Phospho-Histone H2A.X	Camptothecin	0.7
Mitochondrial Function	Mitochondrial membrane potential	CCCP	0.55
Mitochondrial Mass	Mitochondrial mass	CCCP	0.35
Cell Loss	Cell number	Camptothecin	0.56
Cell Cycle	DNA content	Paclitaxel	0.54
DNA Degradation	DNA structure	Paclitaxel	0.6
Nuclear Size	Area of nuclear region	Paclitaxel	0.63
DNA Damage	Detection of p53	Camptothecin	0.43
Mitotic Arrest	Phospho-Histone-H3	Paclitaxel	0.63
Cytoskeletal Integrity	Detection of -tubulin	Paclitaxel	0.3

Cellumen: What is Being Measured?

- Cellular toxicity phenotypes in a transformed cell line
- False positives
  - Imaging artifacts
  - Fluorescent compounds
  - Statistical vs biological significance
  - Operational
- False negatives
  - Solubility
  - Cytotoxicity
  - Lack of biotransformation
  - Operational

Cellumen: Data Calculation

- Data not normalized to controls (with a few exceptions)
- Fit to 3- or 4-parameter Hill equation
- AC50 reported with these rules:
  - For Cell Loss, AC50 is reported as calculated
  - For other endpoints, if AC50 for endpoint is > AC50 for Cell Loss at the corresponding exposure time, AC50 for endpoint is set to 100 μM (to account for imaging artifacts of cytotoxicity
- Issues:
  - Lack of positive controls for all endpoint/time combinations
  - Large differences in maximal response
  - noisy curves due in part to effects of cytotoxicity



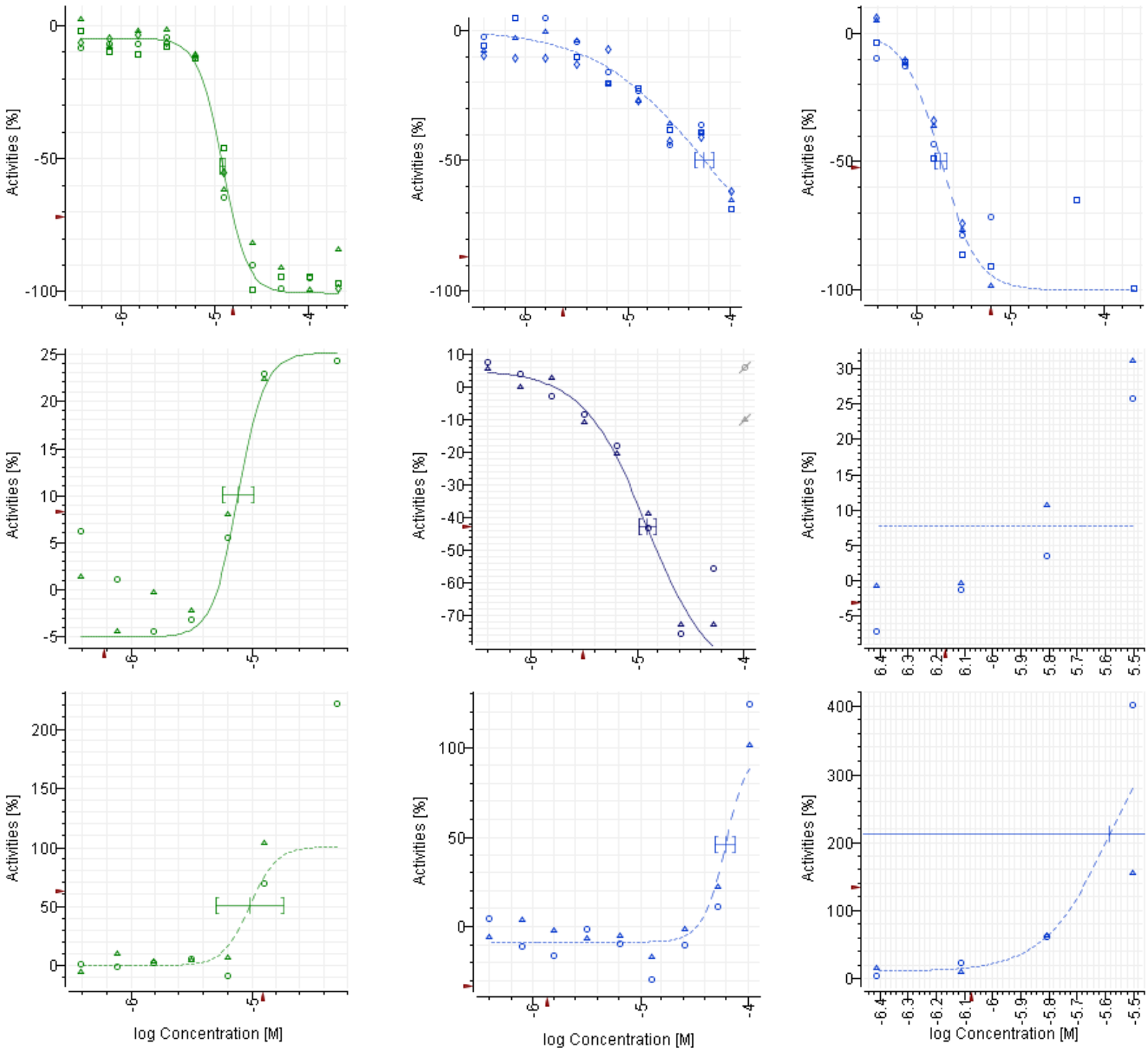
Cellumen: Data Examples

EPA United States Environmental Protectic Agency

Cell Loss

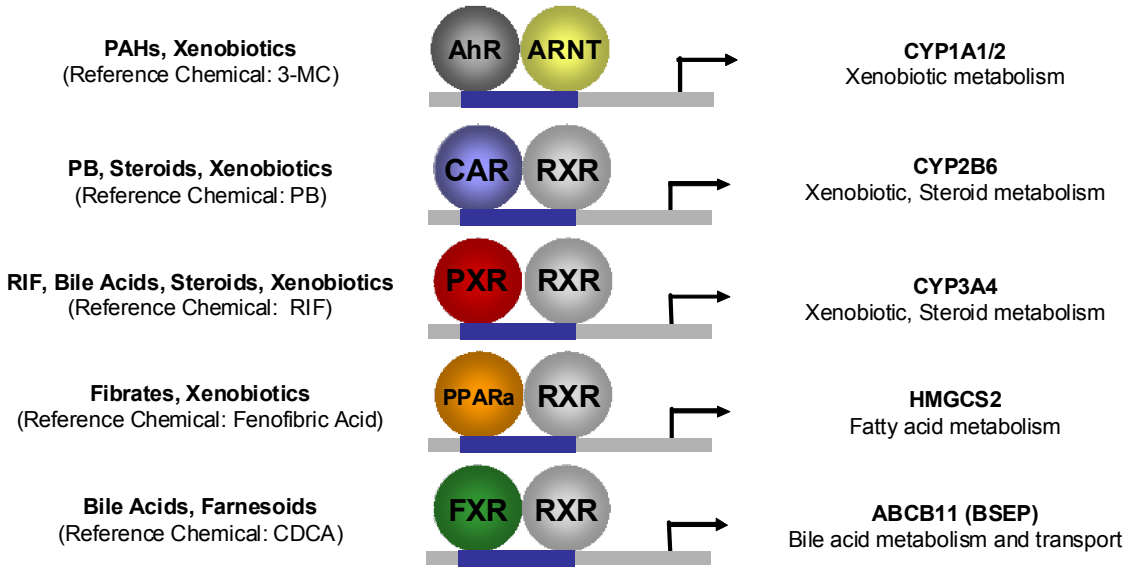
Mitochondrial Membrane Potential

DNA Damage



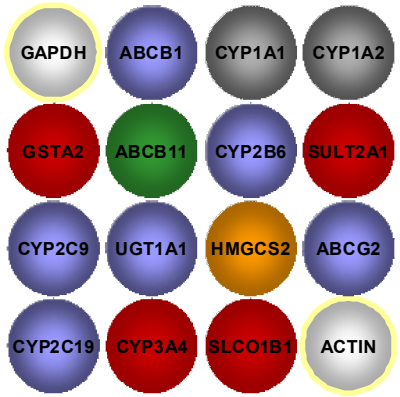
CellzDirect: XME Gene Expression in Primary Human Hepatocytes (Abstract 22)

- Primary human hepatocytes from two donors used
- Cells exposed for 6, 24, and 48 hr; medium/chemical refreshed daily
- Concentrations tested: 40, 4, 0.4, 0.04, and 0.004 µM
- 16 Genes measured in multiplexed RNAse protection assay (qNPA)
- Genes targeted XME and transporters



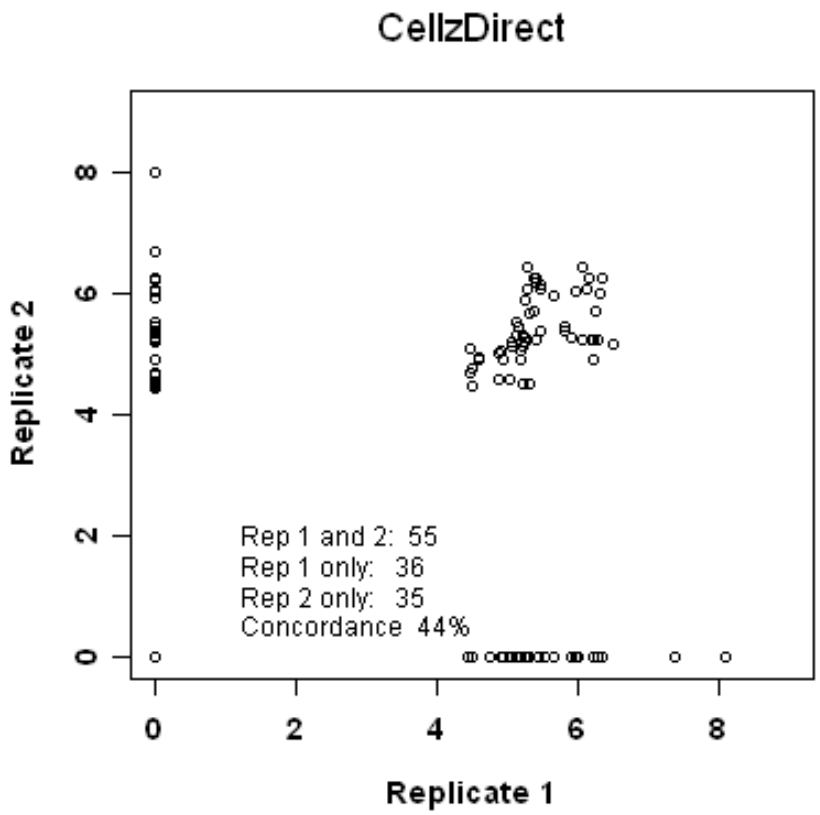
Target Gene Categories

CYP450 (6)
Transporter (4)
Phase II Metabolism (3)
Cholesterol Synthesis (1)
Endogenous control (2)



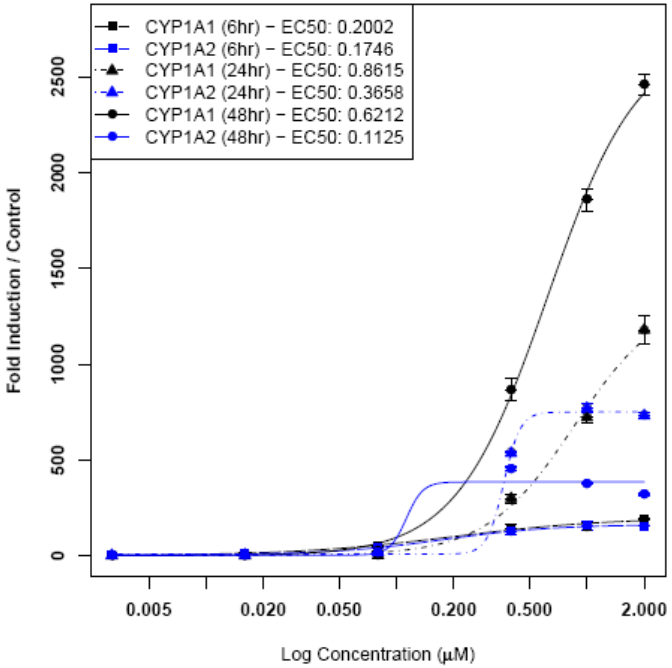
CellzDirect: Data Calculation

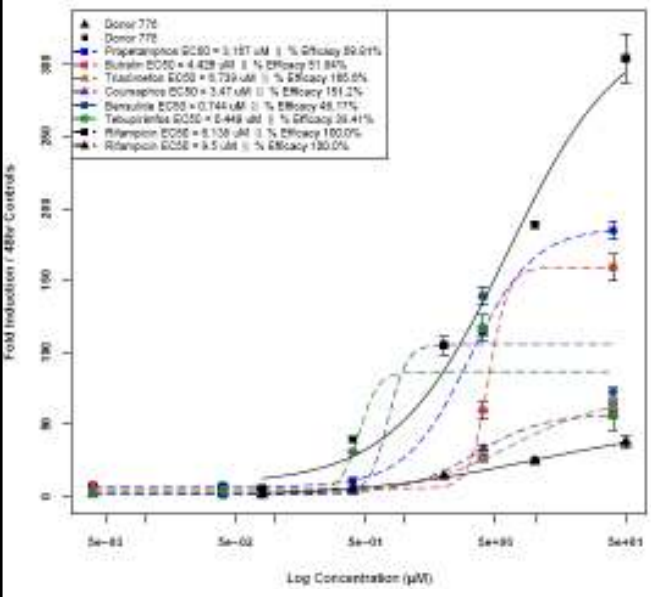
- Normalized to solvent control; expressed as fold-change
- Curves fit to Hill equation
  - Upper and lower limits defined by minimum and maximum responses observed over dataset of a particular gene/donor/time
- LECs determined
- Large variations in maximal responses
  - Biological vs statistical significance
- Measuring endogenous promoter activity reflects complex, multifactorial regulation of gene expression
- 6 hr exposure data not provided due to high variability associated with limited time for gene induction



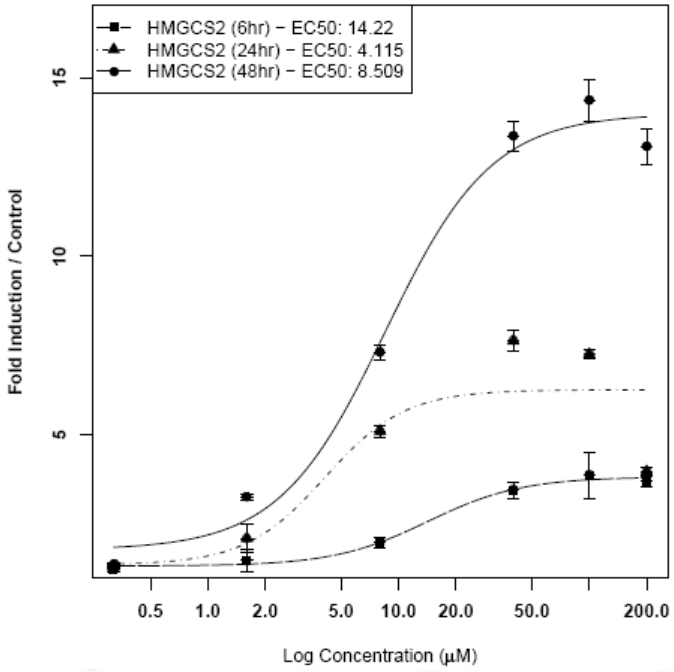
CellzDirect: Data Examples

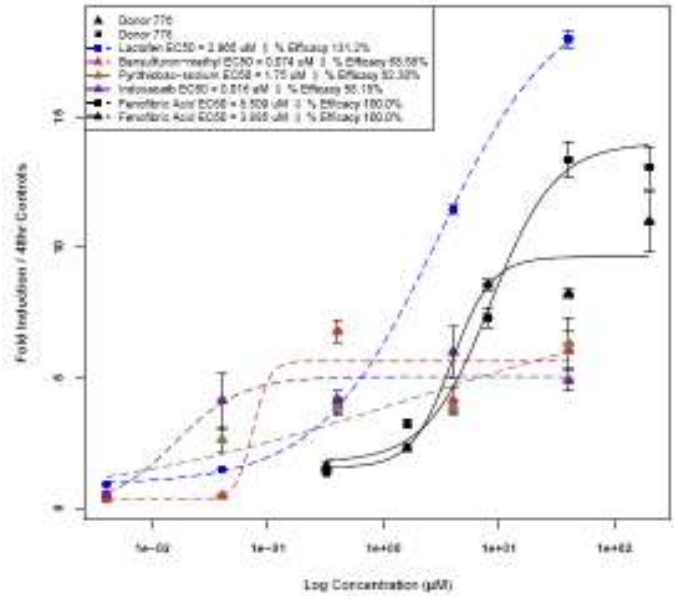
CYP1A1-AhR



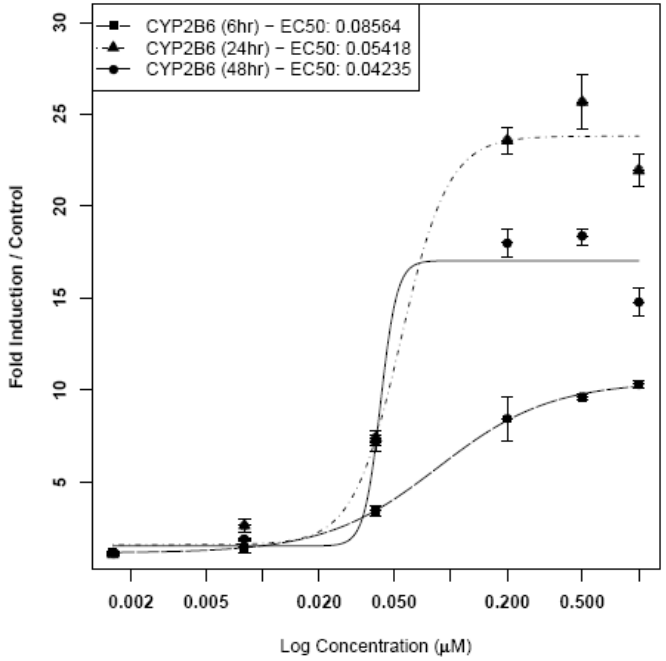


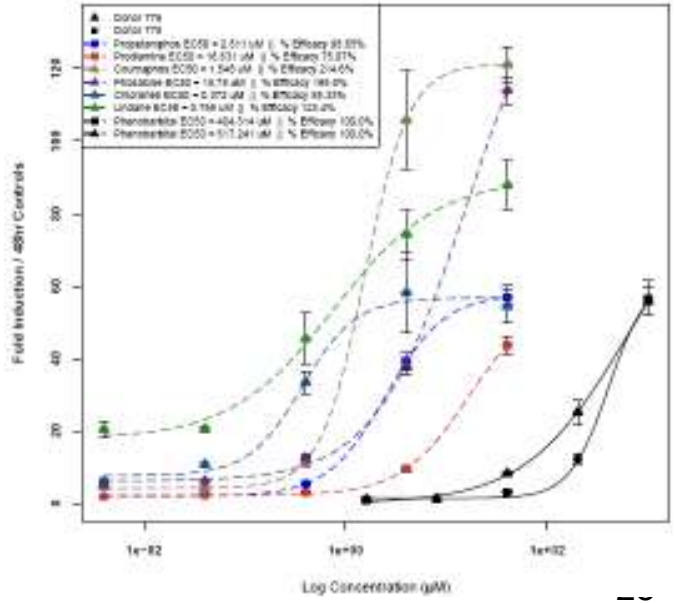
HMGCS2-PPAR α





CYP2B6-CAR



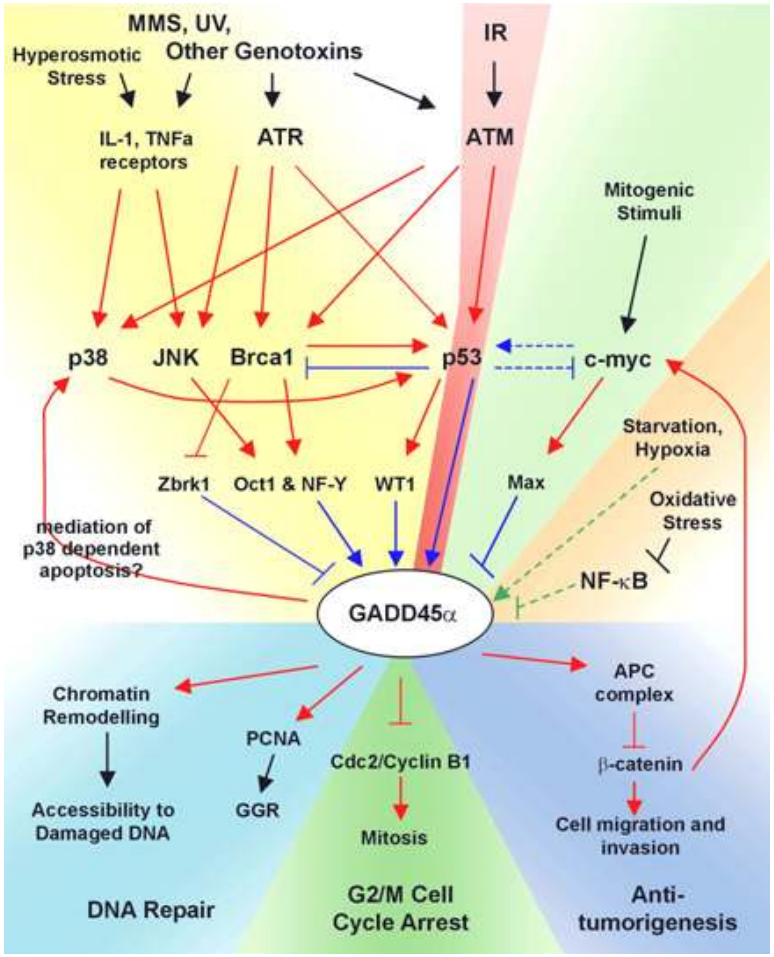


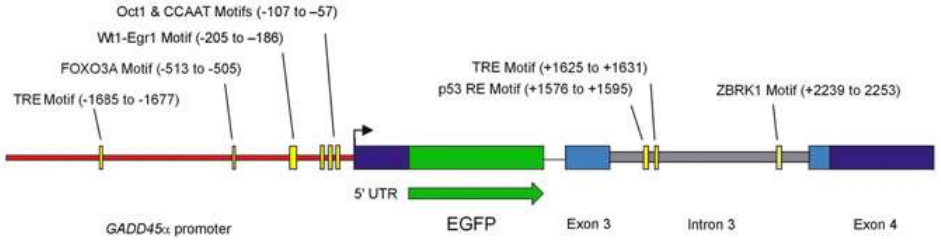
CellzDirect: What is Being Measured?

- Up/down regulation of mRNA for XME and transporters in primary human hepatocytes
- False positives:
  - General effect of cytotoxicity on transcriptional activity
  - Statistical vs biological significance
  - Operational
- False negatives:
  - Solubility
  - Cytotoxicity
  - Operational
  - Lack of biotransformation
  - Inter-individual donor variation

Gentronix: GADD45a Reporter Gene Assay for DNA Damage (Abstract 41)

- TK6 cell line expressing GFP under control of GADD45a promoter
- Cells exposed at 200, 100, 50 μM for 24 and 48 hr
- Cytotox assay to discount artifacts
- Retested at lower conc if cytotoxic





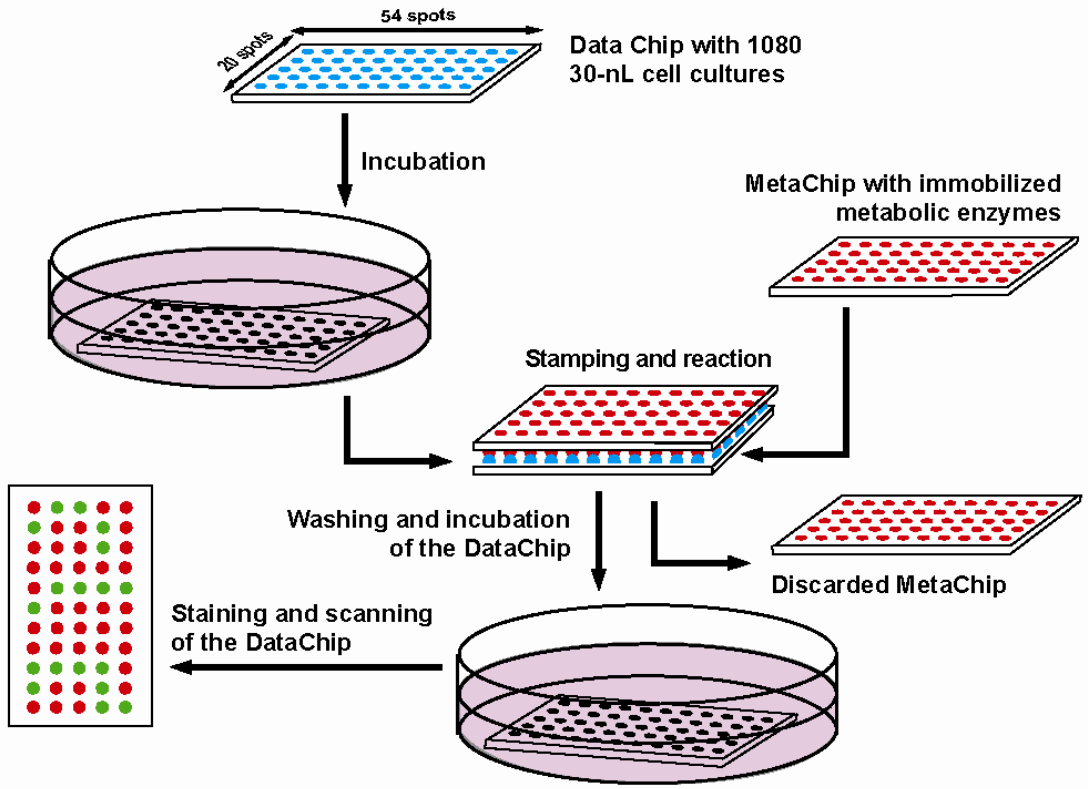
Gentronix: Data Calculations

- Induction of GFP fluorescence >50% = genotoxic
- If over 2 or 3 concentrations, strongly genotoxic
- LELs calculated
- Replicate analysis: no actives among replicates

- False positives:
  - Cytotoxicity resulting in general transcriptional activity
  - Cellular stress other than DNA damage
  - Statistical vs biological validation
  - Operational
- False negatives:
  - Solubility
  - Cytotoxicity
  - Lack of biotransformation
  - Operational

Solidus: Biotransformation Chip and Effect of Cytotoxicity (Abstract 30)

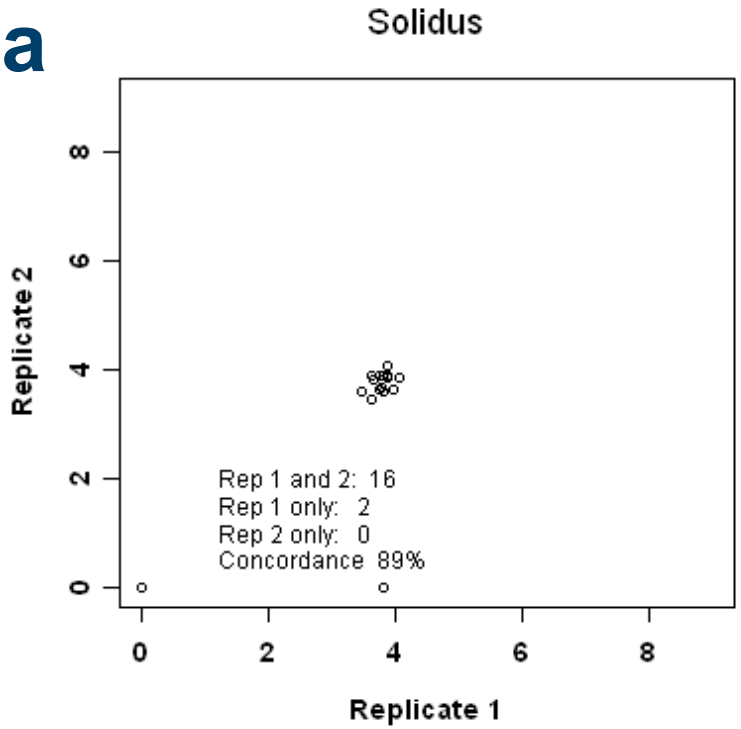
- Alginate-immobilized Phase I and Phase II enzymes
- ToxCast_320 exposed 6 hr to:
  - Control
  - Ph I
  - Ph II
  - Ph I and Ph II
- 960 μM high conc/4-fold serial dilutions/9 concentrations/5 replicates
- Cytotoxicity in Hep3B measured 48 hr later

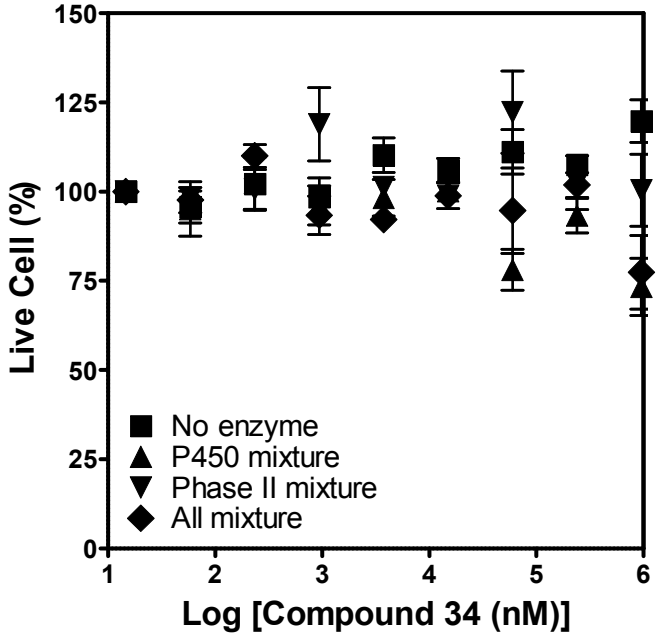


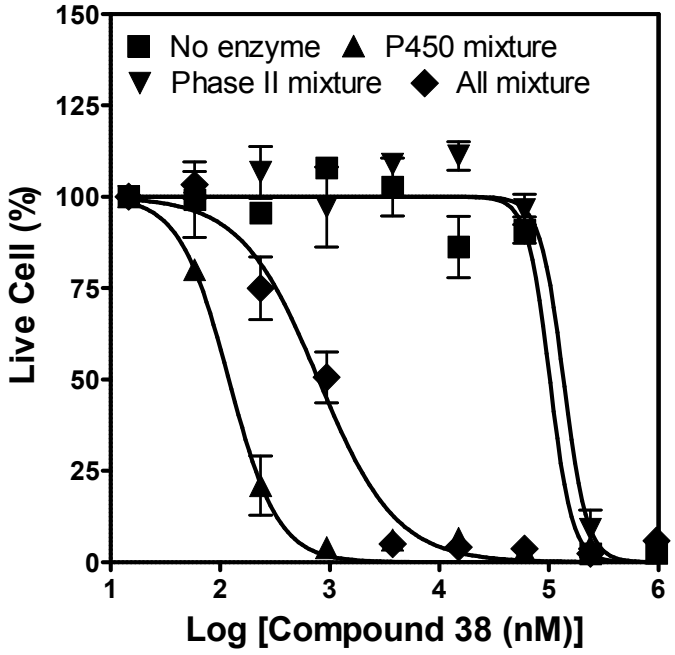
Solidus: Example Data

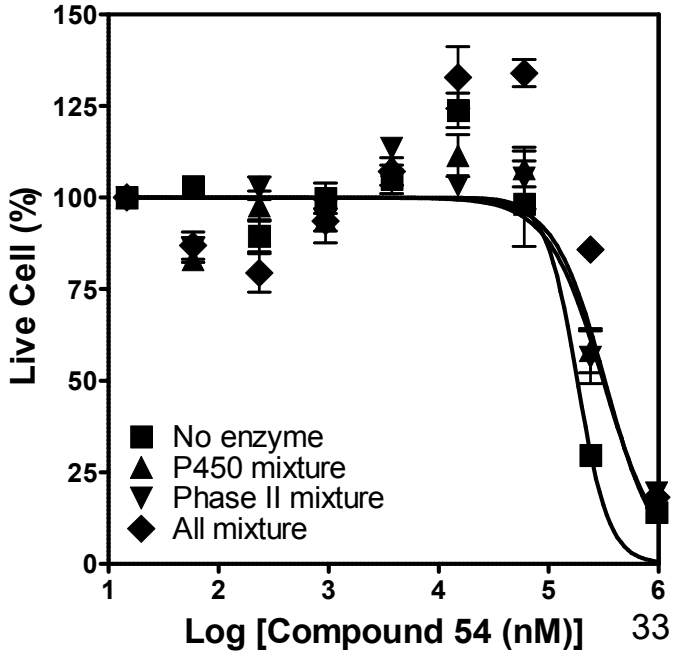
United States Environmental Protection Agency

- Data normalized to control values
- Concentration-response data fit to Hill equation
- LC50 determined for each assay condition









Solidus: What is Being Measured?

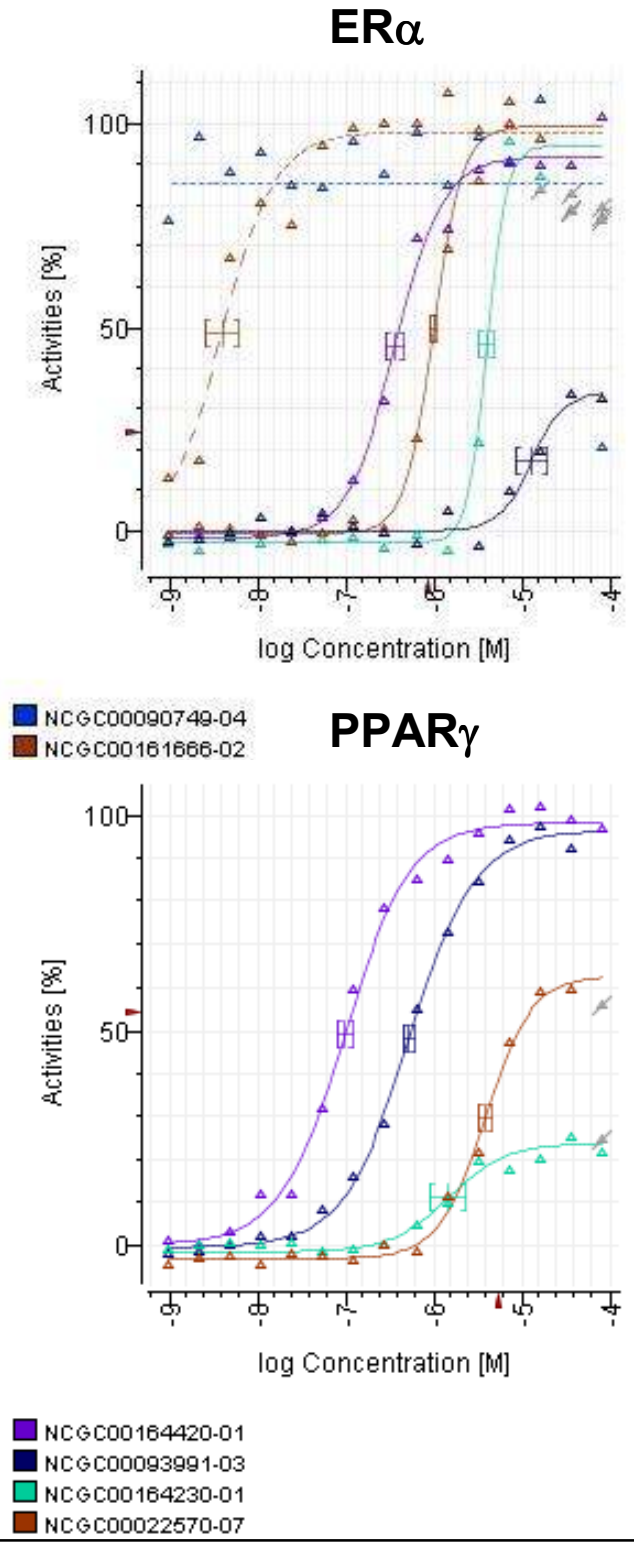
- Effect of Phase I and Phase II enzymes on cytotoxicity activity of chemicals against a transformed cell line
- False positives:
  - Possibly not optimized Phase I and/or Phase II mix
  - Operational
- False negatives
  - Solubility
  - Possibly not optimized Phase I and/or Phase II mix
  - Availability of compound from alginate-immobilized enzyme matrix
  - Operational

NCGC Reporter Gene Assays

- Nuclear Receptors
  - GAL4 System (ligand detection assay)
  - 11 human receptors
  - 1 rat (PXR)
  - β-lactamase reporter gene assays except:
  - PXR assays are luciferase reporter gene assays
- p53 Reporter Gene assay
  - β-lactamase reporter gene assay
- Parental cell lines mostly HEK293 (also HeLa and DPX-2)
- 12-15 point concentration-response curves (single replicate)

NCGC: Data Calculations

- Data normalized to reference compound effect
- Curves fit to 3- or 4-parameter Hill equation
- Artifacts removed where obvious fluorescence or cytotoxicity detected
- Required at least 25% efficacy of control compound to calculate AC50
- AC50 values provided
- Antagonist format assays challenging due to effects of cytotoxicity
- LXR assay problematic—contaminated with GR reporter line?



NCGC Assays: What is Being Measured?

- NR assays are ligand-detection assays
- False positives
  - Fluorescent compounds
  - Statistical vs biological significance
  - Gal4_NR-LBD not physiological
  - Cytotoxicity (antagonist format)
  - Operational
- False negatives
  - Fluorescent compounds
  - Cytotoxicity
  - Gal4_NR-LBD not physiological
  - Operational
  - Lack of biotransformation

Additional Data Sets To Be Added Soon:

- NHEERL
  - Zebrafish developmental toxicity (Padilla) Poster
  - Neurite outgrowth and neuronal proliferation (Mundy and Shafer)
  - Stress gene reporter assays (Simmons)
  - ES cell differentiation (Hunter)
- Plasma protein binding and hepatocyte clearance (Thomas)
- PPARα and AhR (NCGC)
- Primary rat hepatocyte HCS Cell Health (Cellumen)