Tox21 Federal Collaboration Cross Partner Project # 5: Project Update

Development of a Common Reference Chemical Dataset for Interpretation of High-Throughput Transcriptomics Screening Data

Office of Research and Development
Full Name of Lab, Center, Office, Division or Staff goes here.

Tox21 Face-to-Face Meeting, Winter 2020, Bethesda, MD
**2020-01-14**

Disclaimer

Project Background

– CPP5 Partner organizations (USEPA, NTP) have both incorporated high-throughput transcriptomics (HTTr) screening into their research portfolios.

– To date, the cell models, gene panel and chemical sets evaluated by the respective organizations do not completely overlap.

Organization	Cell Models	Gene Panel	Chemical Sets
NTP	2-D HepaRG (Prolif & Diff), 3-D HepaRG	TempO-Seq S1500+	Liver Toxicants (mostly)
EPA	MCF-7, U-2 OS, 2-D HepaRG (Diff)	TempO-Seq whole transcriptome	ToxCast Chemicals

– It is anticipated that the diversity of cell models and chemical sets evaluated will increase over time for both organizations.

– There is a need for datasets and tools that can facilitate interpretation of HTTr data & identify putative molecular targets / mechanisms, especially for poorly characterized or unknown chemicals.

Project Hypothesis

Development of a common set of transcriptional profiles from reference chemicals will allow more robust interpretation of high-throughput transcriptomic screens to link chemicals to biological-response pathways and molecular initiating events.
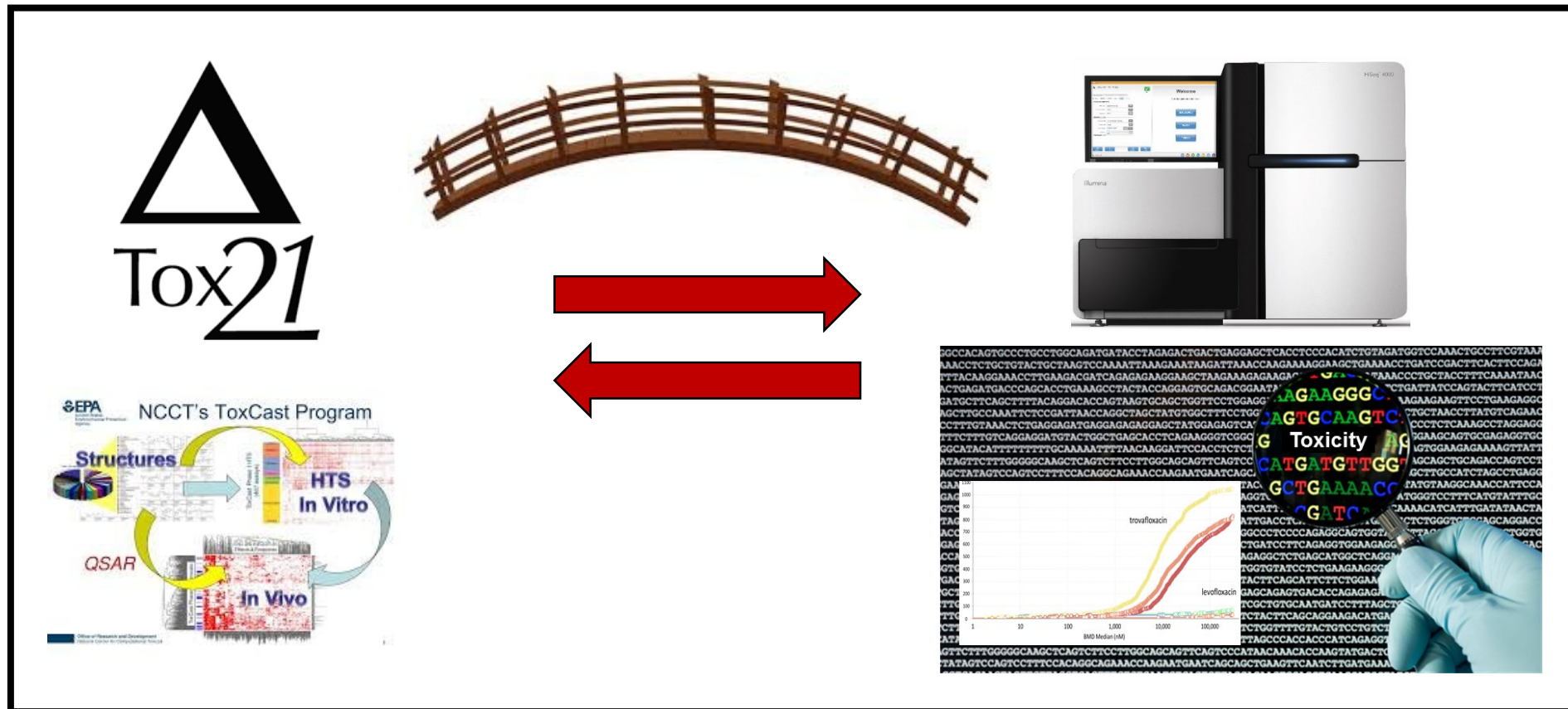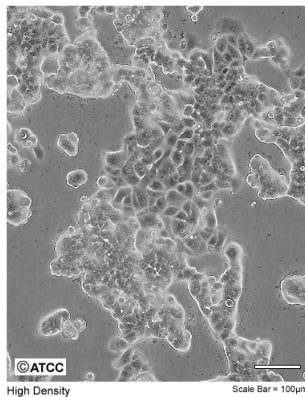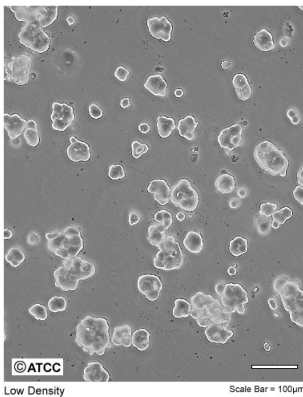
Project Goals

- Identify a diverse set of ~200-500 chemicals from the Tox21 chemical library that overlap with multiple annotated sources of reference chemicals (to biological pathways)

- Build a robust transcriptomics dataset of concentration-response biological perturbations to these reference chemicals using MCF-7 and HepaRG cultures employed by Tox21 partners in high-throughput screens

- Develop transcriptomic signatures that accurately identify specific molecular targets/biological pathways perturbed by the reference chemicals, and their respective 'firing orders' across the range of concentrations examined.

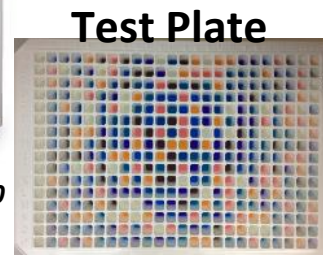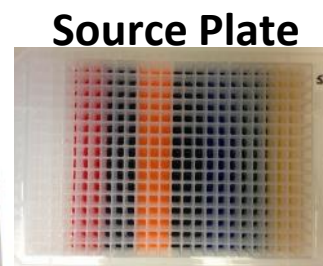# Step 1: Fit-for-purpose approach compatible with Tox21 partner capabilities

- 4 independent experimental runs, 3 averaged for HTT
- 1 technical replicate per treatment group
- 6 chemical concentrations (half-log spacing)
- concentrations span reference target potencies from reporter assays (e.g., IC_{50}, EC_{50})

Assays summary

- **Cell viability/Cytotoxicity**

  - Cell viability will be evaluated based on each lab's respective established methods

    - NTP: Liver enzyme leakage assays (LDH-Glo) in spent culture media with photomicrographs for each well for 3D HepaRG spheroids

    - EPA: Employed a high content imaging (HCI)-based cell painting approach that assesses apoptosis and cell viability assay for use with MCF7 cells grown in monolayer culture.

- **High throughput transcriptomics**

  - Cell cultures will be lysed with BioSpyder lysis buffer for 15m, frozen at -80°C, and shipped to BioSpyder for TempO-Seq analysis using the Human Whole Transcriptome Gene Set to be analyzed at a minimum average read depth of ~300-500 reads/transcript.

- Freezing leftover culture media for future contextual follow-up (e.g., chemical exposures, metabolomics) with 3D HepaRG spheroids

- **Goal:**

  - Develop a list of ~300 chemicals covering ~75-100 biological-response pathways (i.e., 3 or more for a given pathways) with well-annotated associations to specific molecular targets or biological-response pathways and 10-20 "negatives"

Identifying a diverse set of reference chemicals for specific biological targets (2)

RefChemDB

A database of *chemical_target_mode_activity* associations created by Judson et al. from information contained in the public domain.
- > 2900 biological targets
- > 37,000 unique chemicals

Intended to be used in a semi-automated workflow for development *candidate reference chemical lists* for a molecular target

Candidate lists are then refined using expert knowledge.

For a given chemical, more than one literature source may support a given chemical_target_mode association (i.e. *level of support*).

Source	Chemicals	Targets	Chemical-target-mode-activity combinations	Mean multiplicity	PMIDs
ChEMBL	28,832	2,238	310,984	1.16	11,520
ChEMBL Drug	1,187	738	4,099	1	0
CTD	2,317	7,904	25,606	1.22	5,280
DrugBank	1,630	1,169	3,623	3.41	6,274
Eurofins Biochemical	206	570	925	1	0
Eurofins Functional	211	239	706	1	0
Iuphar BPS	1,860	941	5,081	1	0
KEGG Drug	661	263	1,201	1	0
KIDB	535	450	6,532	1	0
KInaseDB	133	168	676	1	1
LitDB	2,654	88	8,348	4.94	27,909
Open Targets	1,031	820	3,973	1	0
Prodrug	41	33	41	1	1
Repurposing Hub	2,279	2,172	10,209	1	0
ToxCast	9,136	343	852,470	1.03	0
TTD	3,916	1,575	11,557	1.00	0
Web Curation	3,940	1,059	5,617	1.01	0
Total	**37,301**	**11,055**	**123,4580**	**1.02**	**49,883**

Identifying a diverse set of reference chemicals for specific biological targets (3)

Assay Performance (A)

Chemical Performance (B)

Potency vs. Support
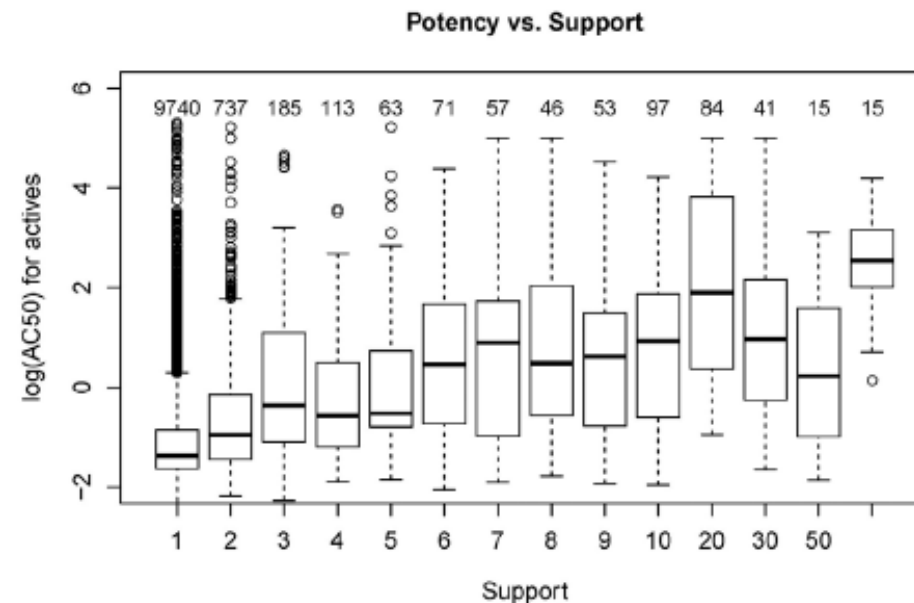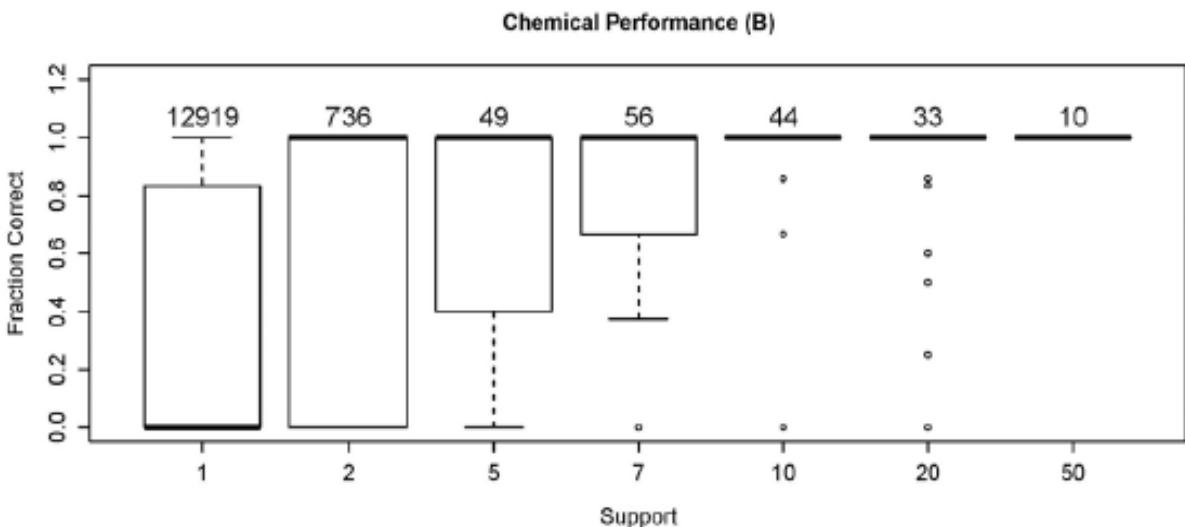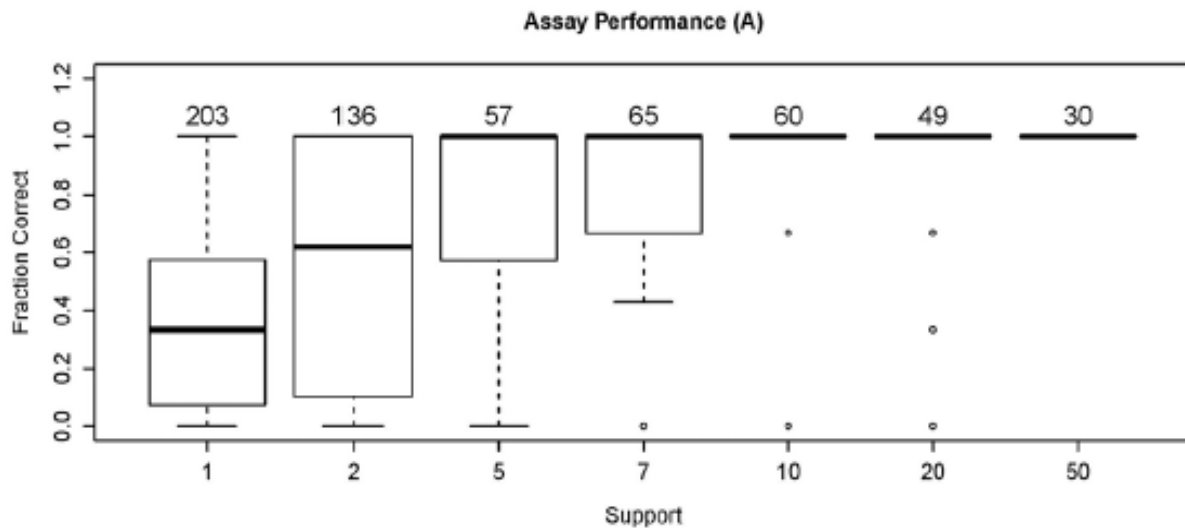
Higher level-of-support is associated with greater reliability of chemical_target_model predictions using the ToxCast/Tox21 assay results as a benchmark.

Chemical_target_mode predictions with level of support <= 5 are suspect.

Higher level-of-support is also associated with greater potency at the predicted molecular target.

Identifying a diverse set of reference chemicals targeting specific biological targets (4)

- **Approach:**

  - RefChemDB filtered for chemicals with > 5 chemical_target_mode associations for a given pathway (n = 3263 chemicals)

  - An additional filter was applied to remove "promiscuous" chemicals that map to > 3 targets.

  - Two agonists / antagonists were then selected (if available) for each remaining target with preference toward more selective chemicals (n = 330 chemical candidate list).

  - The list was then manually curated to confirm target / mode associations, remove spurious target/mode associations and identify targets / pathways of interest that were not selected using the automated methods.

  - The candidate list includes chemicals that target classical signal transduction pathways (e.g., nuclear receptors) and upstream / downstream proteins (e.g., GPCRs, enzymes, transporters) as "pathways".

Targets by Target_Class

315 chemicals → 143 Targets Across → 13 Target Classes
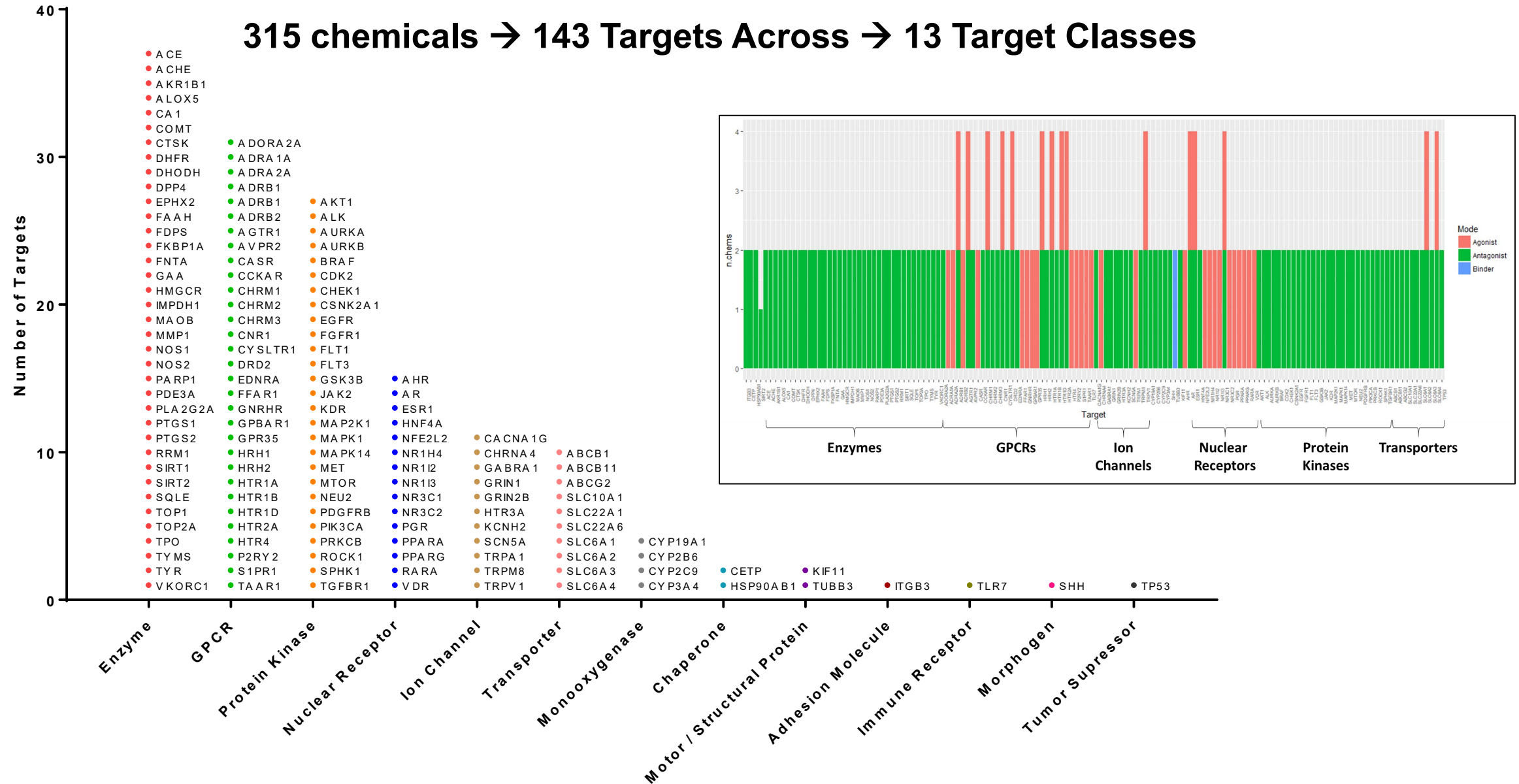
Procurement Plan:

- 58 chemicals available in existing ToxCast library

- 168 chemicals obtained by EPA for CPP5 and other projects.

- 89 compounds designated for procurement by NTP.

  - 8 of 89 chemicals were present in the MRI chemical inventory

  - Initial sourcing of the remaining 81 chemicals targeted 5 g procurements of each to provide enough for preparation of 20 mM solutionsin DMSO plus 1 g of neat material.

  - Initial sourcing cost was prohibitively high (**~$1.4M**) 🙁

Procurement Progress:

- NTP Chemistry worked with Steve Ferguson and Josh Harrill to identify substitutes for the most expensive chemicals on the sourcing list.

- Replacement chemicals:
  - Selected to target the same biomolecular pathways(s) and mode as the original chemicals.
  - Sourced and CAS number/chemical name/chemical form issues were resolved.
    - New sourcing cost: **$181,400** ☺

- Chemicals were ordered through MRI Global.

- **2019-12-23:** All non-DEA schedule CPP5 compounds are in house at MRI and EPA-required paperwork has been completed. Project all non-DEA solubilized chemicals will be received by Evotec in Feb-March 2020.

Examples of manually curated chemical substitutions

Preferred_Name	Original_Substance_Requested	Target_Mode	Reason for substitution
Citronellol	Substitute for JYL-1511 (623166-14-3)	TRPV1 _ Agonist,TRPV1 _ Antagonist	Already in NTP inventory
Proguanil hydrochloride	Substitute for Piritrexim (72732-56-0)	DHFR _ Antagonist	Already in NTP inventory
Anisotropine methylbromide	Substitute for Biperiden (514-65-8)	CHRM1 _ Antagonist	Already in NTP inventory
Fludrocortisone acetate	Substitute for Fludrocortisone (127-31-1)	NR3C2 _ Agonist	Already in NTP inventory
Solifenacin	Substitute for 4-Diphenylacetoxy-1,1-dimethylpiperidinium (81405-11-0)	CHRM3 _ Antagonist	Already in NTP inventory
SB202190	Substitute for SB202190 (152121-30-7)	MAPK1 _ Antagonist,MAPK14 _ Antagonist	Already in NTP inventory
Montelukast sodium	Substitute for Montelukast (158966-92-8)	CYSLTR1 _ Antagonist,CYSLTR1 _ Binder	Already in NTP inventory
Rabusertib	Substitute for granulatimide (219828-99-6)	CHEK1 _ Antagonist	Already in NTP inventory
SU9516	Substitute for 3-[(1H-Imidazol-5-yl)methylidene]-5-methoxy-1,3-dihydro-2H-indol-2-one (377090-84-1)	CDK2 _ Antagonist	Already in NTP inventory
Apatinib myselate	Substitute for CP 547632 (252003-65-9)	KDR _ Antagonist	Already in NTP inventory
Ifenprodil (+)-tartrate salt	Substitute for Ifenprodil (23210-56-2)	GRIN2B _ Antagonist	Already in NTP inventory
Scopolamine hydrochloride	Substitute for Methoctramine (104807-46-7)	CHRM2 _ Antagonist	Already in ToxCast
Caffeine	Substitute for ST-1535 (496955-42-1)	ADORA2A _ Antagonist	Already in ToxCast
ZM-447439	Substitute for Barasertib (722543-31-9)	AURKB _ Antagonist	Cost savings
Leflunomide	Substitute for Brequinar (96187-53-0)	DHODH _ Antagonist	Cost savings
Doxycycline hydrochloride	Substitute for Prinomastat (192329-42-3)	MMP1 _ Antagonist,MMP2 _ Antagonist	Cost savings
Picrotoxin	Substitute for Mephobarbital (115-38-8)	GABRA1 _ Antagonist	Cost savings
Flt-3 inhibitor (TCS 359)	Substitute for Lestaurtinib (111358-88-4).	FLT3 _ Antagonist	Cost savings
Mebendazole	Substitute for MMAD (Monomethylauristatin D) (203849-91-6).	tubulin	Cost savings
AZD8055	Substitute for Temsirolimus (162635-04-3).	MTOR _ Antagonist	Cost savings
SCH-900776	Substitute for isogranulatimide (219829-00-2)	CHEK1 _ Antagonist	Cost savings
Ellagic acid	Substitute for 4,5,6,7-tetrabromobenzimidazole (577779-57-8)	CSNK2A1 _ Antagonist,CSNK2A2 _ Antagonist	Cost savings
Cetrorelix acetate	Substitute for Cetrorelix (120287-85-6)	GNRHR _ Agonist,GNRHR _ Antagonist	Cost savings
Tyramine	Substitute for 3-iodothyronamine (712349-95-6).	TAAR1 _ Agonist	Cost savings
Uridine 5'-triphosphate	Subsitute for Diquafosol (59985-21-6)	P2RY2 _ Agonist	Cost savings
Zopolrestat	Substitute for Sorbinil (68367-52-2)	AKR1B1 _ Antagonist	Cost savings
Cetirizine	Substitue for Levocabastine hydrochloride (79547-78-7)	HRH1 _ Agonist	Cost savings
Zanamivir	Substitute for 2-Deoxy-2,3-dehydro-N-acetylneuraminic acid (24967-27-9)	NEU2 _ Antagonist	Cost savings
HU 210	Substitute for Tetrahydrocannabinol (1972-08-3). Class I DEA Compound	CNR1 _ Agonist,CNR1 _ Binder,CNR2 _ Binder	DEA Schedule I substance
Ergotamine tartrate	Substitute for Ergotamine (113-15-5). Class I DEA compound	HTR1D _ Agonist	DEA Schedule I substance
Milnacipran	Substitute for MDMA (42542-10-9). Class I DEA Compound	SLC6A4 _ Agonist	DEA Schedule I substance

CompTox Chemistry Dashboard Registration

HTTr Reference chemical list registered in EPA Chemicals CompTox Dashboard. Not public (…yet).

Hypothesis & Specific Aims

- Project Goals (Initiated January 2018)

☑ 1. Design a fit-for-purpose approach to link reference chemicals for specific molecular targets (e.g., nuclear receptors, kinases) and transcriptomic signatures that is compatible with Tox21 partner capabilities

☑ 2. Identify a diverse set of reference chemicals targeting relevant biological-response pathways

☑ 3. Procure reference chemicals and prepare stock solutions

June 2020 4. Build a robust reference transcriptomics dataset with selected reference chemicals in 2 cell culture models (MCF-7 and 3D HepaRG) currently employed by Tox21 partners (384-well)

Oct 2020 5. Develop transcriptomic signatures that characterize cellular responses to reference chemicals & the utility of a 3-concentration approach to screen larger numbers of Tox21 compounds & untested chemicals

Dec 2020 6. Relate transcriptomic responses to reference chemicals (e.g., DEGs, BMCs) with established responses for reference biological targets (e.g., IC_{50}, K_i, EC_{50})

2021 7. Extend approach to thousands of compounds in a data-driven manner