The True Positives in EPA's Non-Targeted Analysis Collaborative Trial (ENTACT)

Christopher Grulke[1]
Elin Ulrich[2]
Mark Strynar[2]
Jon Sobus[2]
Antony Williams[1]
Ann Richard[1]

1. National Center of Computational Toxicology, U.S. EPA
2. National Exposure Research Laboratory, U.S. EPA

American Chemical Society Meeting, Spring 2019
3 April 2019, Orlando FL

Who am I?

- Name: Christopher M. Grulke
- Undergraduate Education:
  - BSE Chemical Engineering
- Job 1: Research Informatics at Pfizer
- Graduate Education
  - PhD, Pharmaceutical Sciences
  - Focus: QSAR modeling
- Current Position at EPA:
  - Title: Computational Chemist
  - Duties: Research (Chem)informatics so one
- Analytical Chemistry Background: 0

NTA Workshop 2015

Non-Targeted Analysis Workshop

Home | Agenda | Registration | Abstract Submission | Logistics

The U.S. Environmental Protection Agency (EPA) will host the Non-Targeted Analysis Workshop August 18-19, 2015 at EPA's Research Triangle Park Campus.

The EPA is responsible for ensuring the safety of tens-of-thousands of chemicals registered for use in the United States. Quantitative exposure data are available for only a small fraction of registered chemicals, hindering assessments of potential health risks to humans, wildlife, and ecosystems. New methods are therefore required to efficiently generate measurement data for a growing number of chemicals that remain largely unexamined. Non-targeted measurement methods, including those based on high-resolution mass spectrometry platforms, offer unique means to screen for xenobiotic chemicals in a variety of environmental and biological media.

This workshop will bring together experts in non-targeted screening to discuss innovative methods and best practices for collecting, analyzing, interpreting, storing, and exchanging measurement data related to xenobiotic chemical exposures. Participants of this workshop will discuss non-targeted analysis "success stories" and help identify:

- Best practices for using non-targeted measurement methods to identify and characterize xenobiotic chemicals in environmental and biological samples.
- Innovative and practical approaches for the analysis, interpretation, and storage of measurement data obtained via non-targeted methods.
- Challenges that may impede widespread use of non-targeted measurement techniques for exposure science.
- Collaborative opportunities to overcome challenges related to existing methodologies and/or information infrastructures.

The EPA has developed a draft agenda and is soliciting abstracts for poster presentations. Travel awards ($500 each) will be given to two selected abstracts to support the registrant to orally present their research on the second day. Please visit the Abstract Submission tab for more details and to submit your abstract.

Workshop Details:
August 18-19, 2015

Workshop Location:
Auditorium C111
U.S. EPA
Research Triangle Park
109 TW Alexander Dr.
Durham, NC 27709

Check out related ORISE research opportunities here.

Contact:
Jon Sobus
919-541-2232
sobus.jon@epa.gov

Elin Ulrich
919-541-3717
ulrich.elin@epa.gov

Mark Strynar
919-541-3706
strynar.mark@epa.gov

I don't think NTA is that great!

Oh yeah, we find everything!

I bet if we make some mixtures of known composition, you won't find half of the chemicals

Bring it on!!!!

https://sites.google.com/site/nontargetedanalysisworkshop/home

Goals for Mixture Design

Question(s):

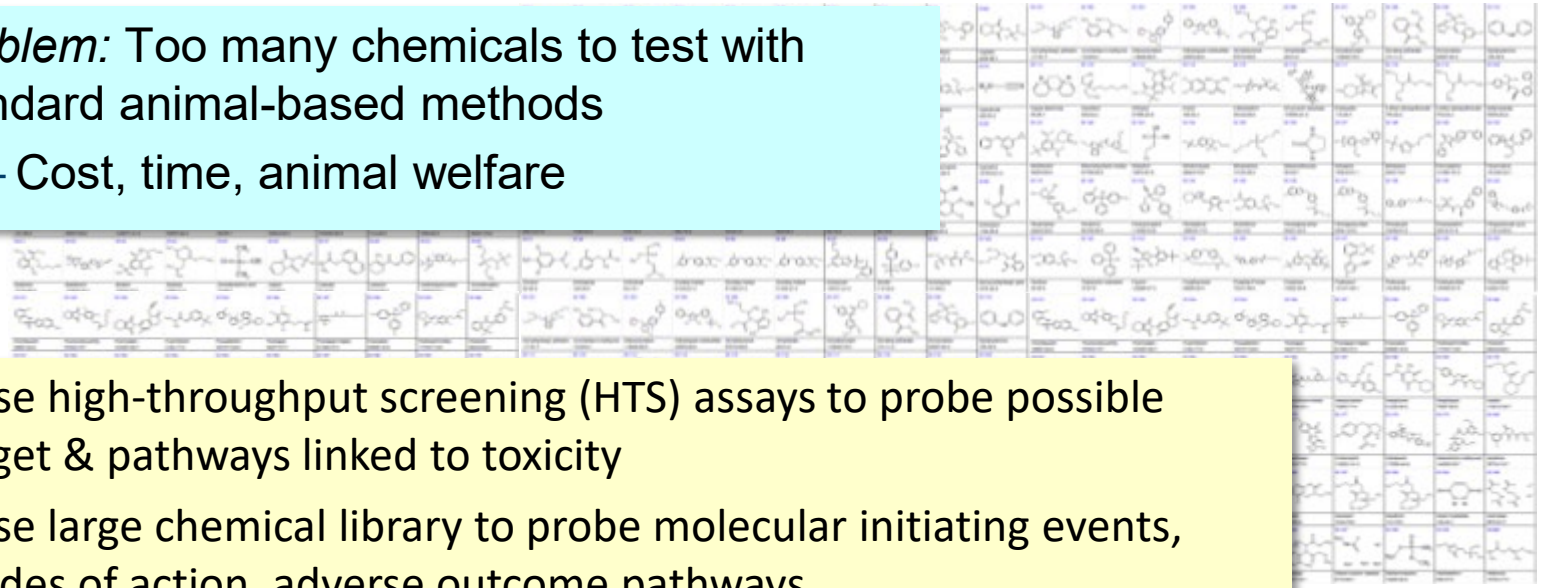
- **What is the relationship between NTA identification (methods) and chemical space detected?**
- Are there chemicals that cannot be detected
- If mixtures are bigger, is detection frequency less?
- If I put the same chemical in a different mixture, can they find it?
- If I put the same chemical in a bigger mixture, can they find it?
- If I put the same chemical in a mixture with an isobaric chemical, can they find it?
- If I throw in some bad QC chemicals, can they find them?
- What if I toss an ill-defined chemical in the mixture?
- Hmmm, polymers are fun…
- Maybe I should vary concentration…
- …

But Where Will We get Samples?

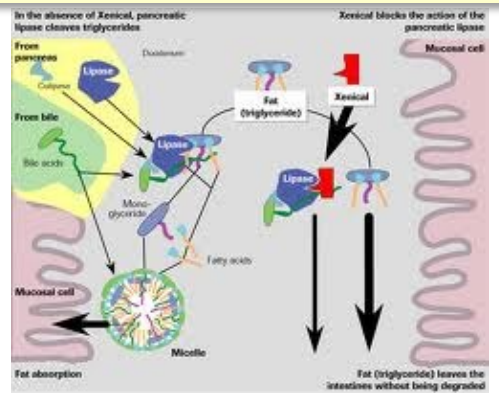
ToxCast Goals

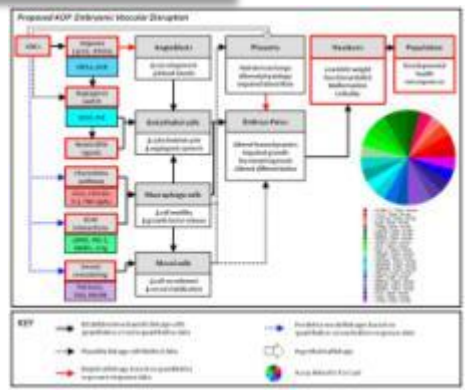
Problem: Too many chemicals to test with standard animal-based methods

- Cost, time, animal welfare



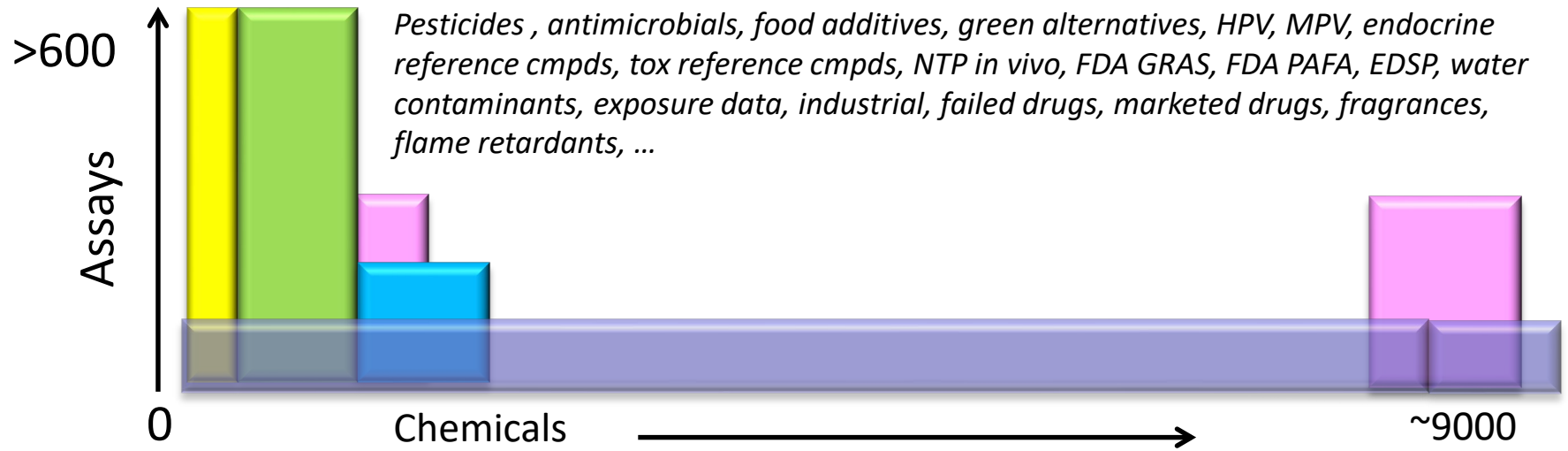
- Use high-throughput screening (HTS) assays to probe possible target & pathways linked to toxicity
- Use large chemical library to probe molecular initiating events, modes of action, adverse outcome pathways
- Develop models to predict in vivo toxicity outcomes





ToxCast & Tox21 Inventories: Chemicals, Data & Timelines

Set	Chemicals	Assays	Endpoints	Completion	Available
ToxCast Phase I	293	~600	~700	2011	Now
ToxCast Phase II	767	~600	~700	03/2013	Now
ToxCast E1K	800	~50	~120	03/2013	Now
Tox21	~9000	~80	~150	Ongoing	Ongoing
ToxCast Phase III	~900	~300	~300	Ongoing	Ongoing



Organizing Our Chemical Library Processes

EPA

1: EPA Orders New Chemical

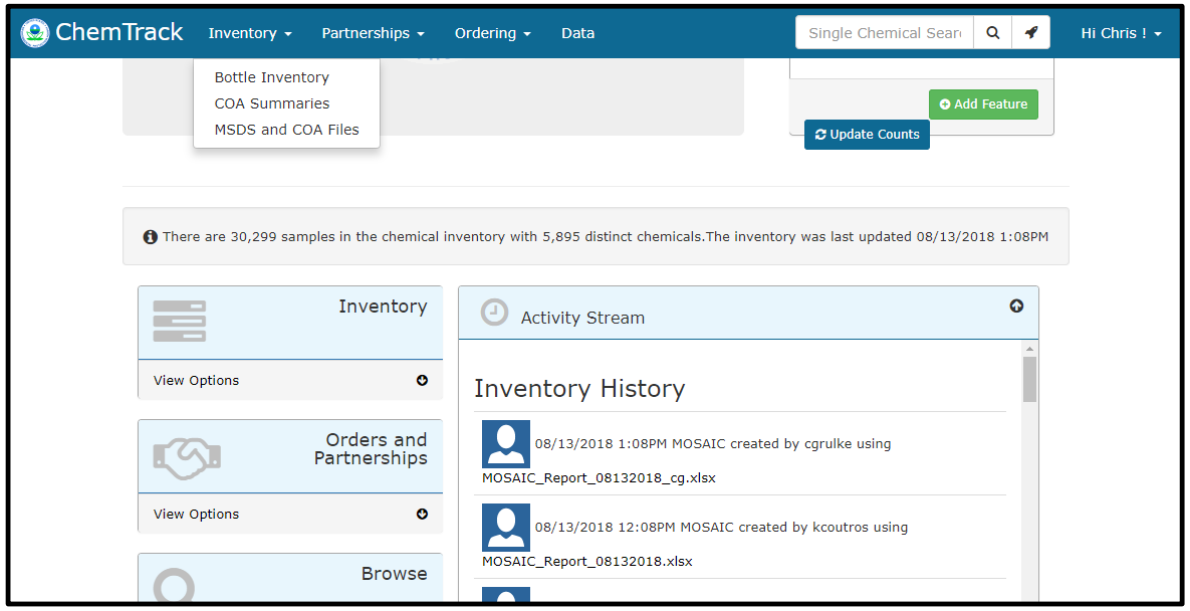
2: CC Procures Chemicals

3: CC Sends EPA
a. COA/MSDS
b. Inventory Report

4: EPA Submits Order

5a: EVOTEC Sends Plate Map

Chemical Contractor (CC)



5b: CC Ships Plates

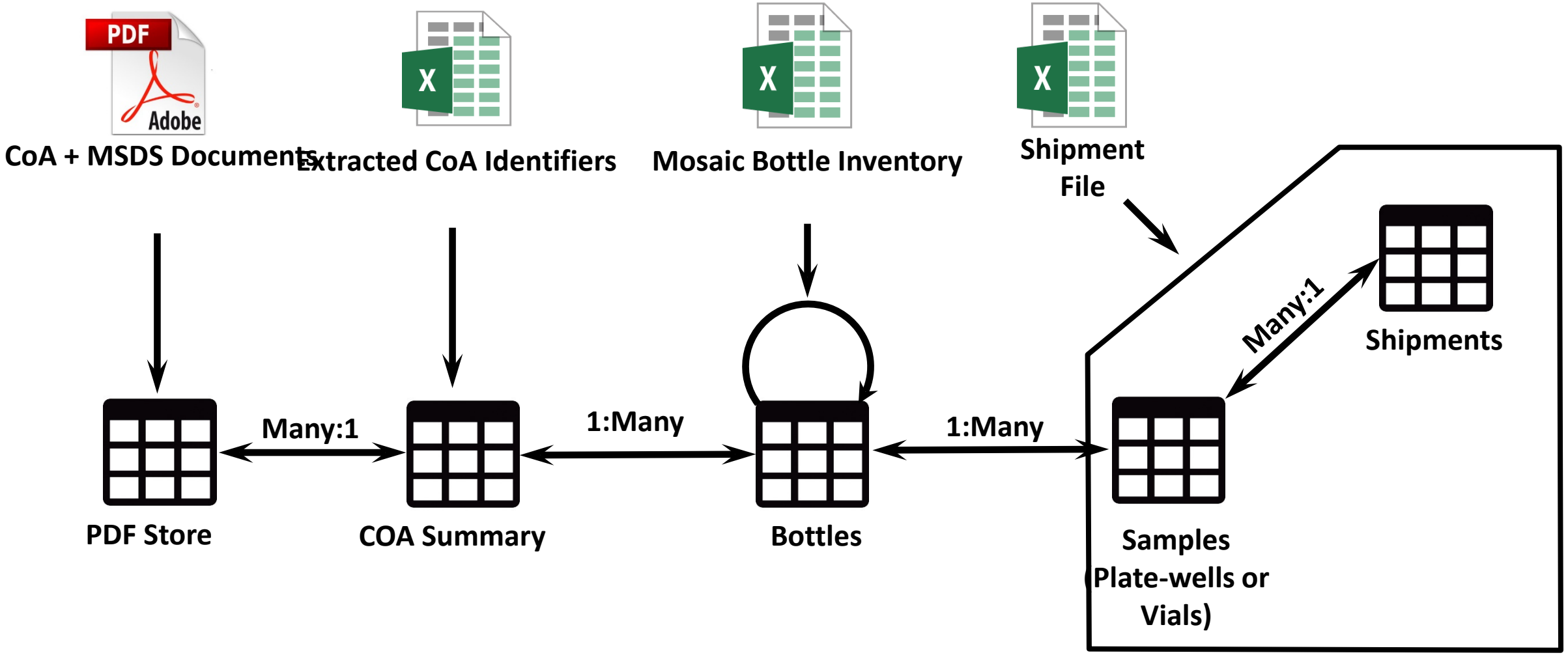
6: EPA Sends Blinded Plate Map

7: Partner Sends Data

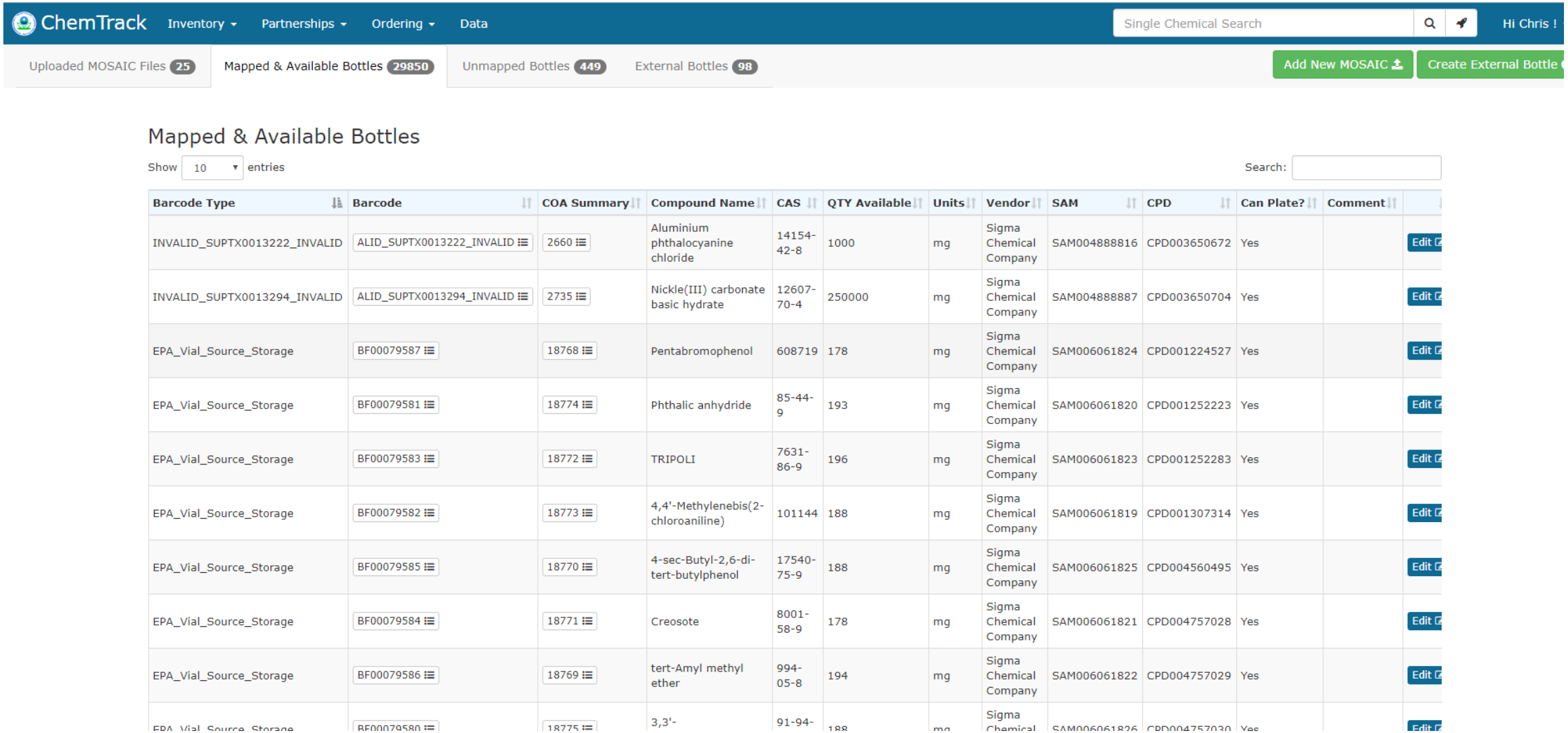
8: EPA Sends Unblinded Plate Map

Screening Partner

ChemTrack Simple Data Model



ChemTrack Bottles



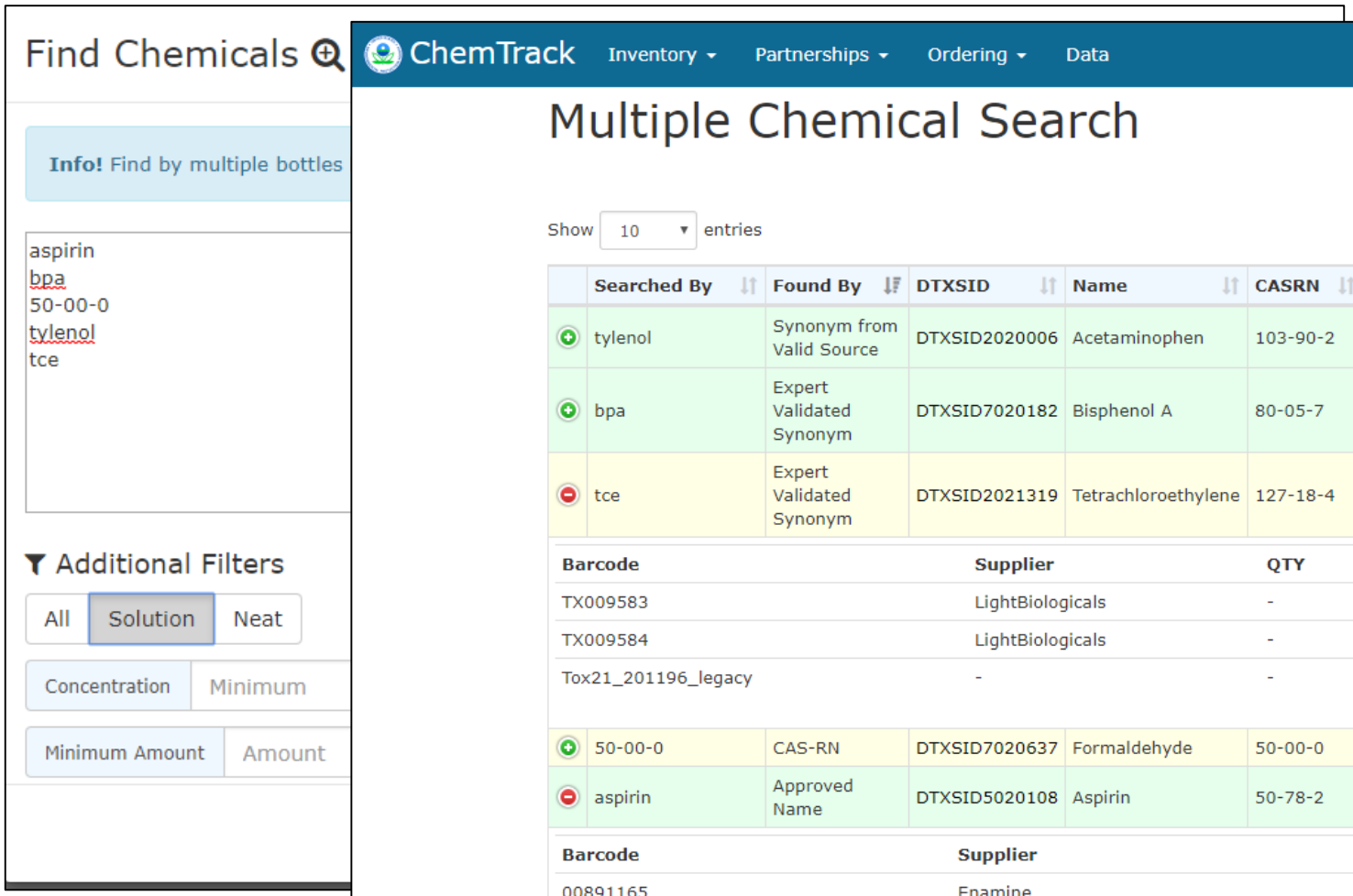
Mapped & Available Bottles

Show 10 entries

Search:

Barcode Type	Barcode	COA Summary	Compound Name	CAS	QTY Available	Units	Vendor	SAM	CPD	Can Plate?	Comment	
INVALID_SUPTX0013222_INVALID	ALID_SUPTX0013222_INVALID	2660	Aluminium phthalocyanine chloride	14154-42-8	1000	mg	Sigma Chemical Company	SAM004888816	CPD003650672	Yes		Edit
INVALID_SUPTX0013294_INVALID	ALID_SUPTX0013294_INVALID	2735	Nickle(III) carbonate basic hydrate	12607-70-4	250000	mg	Sigma Chemical Company	SAM004888887	CPD003650704	Yes		Edit
EPA_Vial_Source_Storage	BF00079587	18768	Pentabromophenol	608719	178	mg	Sigma Chemical Company	SAM006061824	CPD001224527	Yes		Edit
EPA_Vial_Source_Storage	BF00079581	18774	Phthalic anhydride	85-44-9	193	mg	Sigma Chemical Company	SAM006061820	CPD001252223	Yes		Edit
EPA_Vial_Source_Storage	BF00079583	18772	TRIPOLI	7631-86-9	196	mg	Sigma Chemical Company	SAM006061823	CPD001252283	Yes		Edit
EPA_Vial_Source_Storage	BF00079582	18773	4,4'-Methylenebis(2-chloroaniline)	101144	188	mg	Sigma Chemical Company	SAM006061819	CPD001307314	Yes		Edit
EPA_Vial_Source_Storage	BF00079585	18770	4-sec-Butyl-2,6-di-tert-butylphenol	17540-75-9	188	mg	Sigma Chemical Company	SAM006061825	CPD004560495	Yes		Edit
EPA_Vial_Source_Storage	BF00079584	18771	Creosote	8001-58-9	178	mg	Sigma Chemical Company	SAM006061821	CPD004757028	Yes		Edit
EPA_Vial_Source_Storage	BF00079586	18769	tert-Amyl methyl ether	994-05-8	194	mg	Sigma Chemical Company	SAM006061822	CPD004757029	Yes		Edit
EPA_Vial_Source_Storage	BF00079580	18775	3,3'-	91-94-	188	mg	Sigma Chemical	SAM006061826	CPD004757030	Yes		Edit

ChemTrack Search



ChemTrack Inventory Partnerships Ordering Data

Single Chemical Search

Multiple Chemical Search

Show 10 entries

Search:

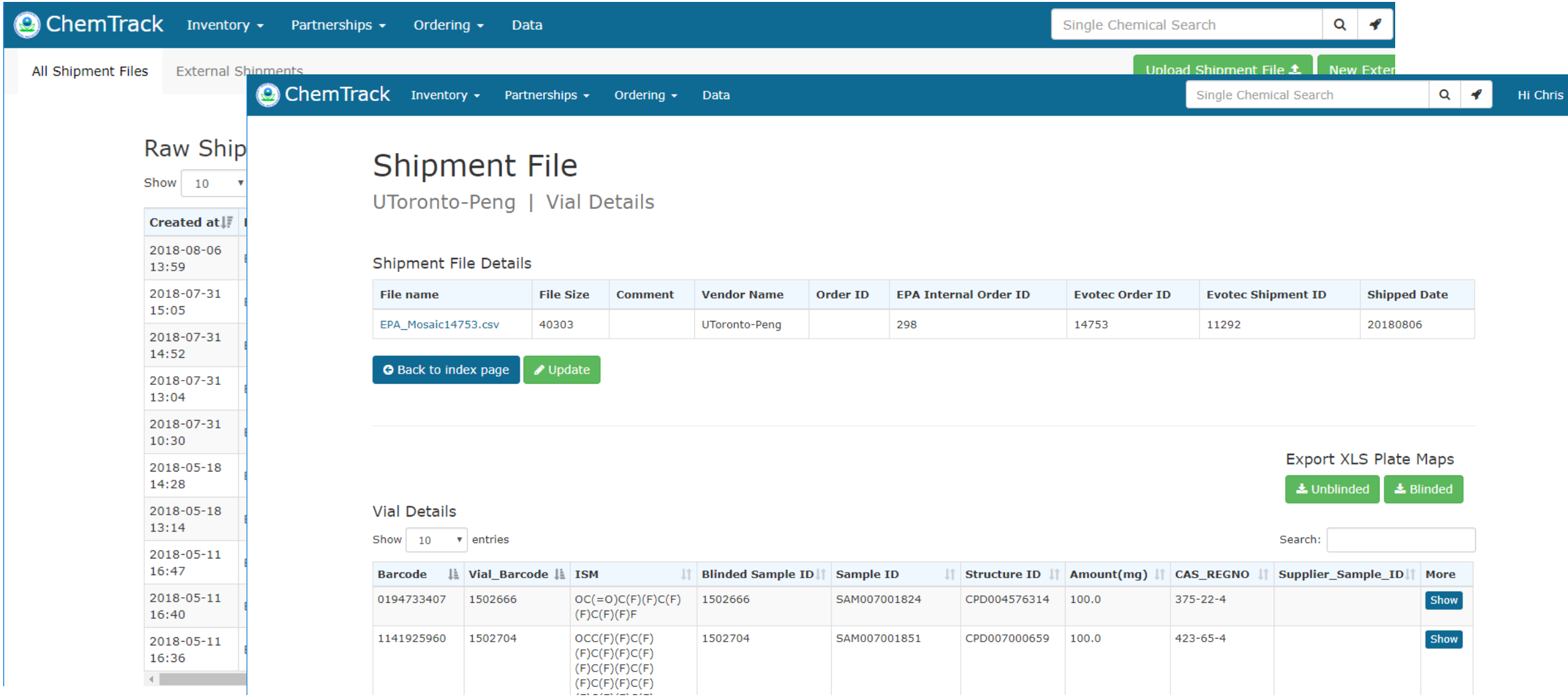
	Searched By	Found By	DTXSID	Name	CASRN	Neat(mg)	0-24mM Stock(ul)	24-100mM Stock(ul)	Number of Bottles
⊕	tylenol	Synonym from Valid Source	DTXSID2020006	Acetaminophen	103-90-2	HIGH	HIGH	HIGH	15
⊕	bpa	Expert Validated Synonym	DTXSID7020182	Bisphenol A	80-05-7	HIGH	HIGH	HIGH	39
⊖	tce	Expert Validated Synonym	DTXSID2021319	Tetrachloroethylene	127-18-4	NONE	NONE	NONE	3

Barcode	Supplier	QTY	Units	Concentration (mM)	Solubility Solvent
TX009583	LightBiologicals	-	mg	-	-
TX009584	LightBiologicals	-	mg	-	-
Tox21_201196_legacy	-	-	-	-	-

⊕	50-00-0	CAS-RN	DTXSID7020637	Formaldehyde	50-00-0	NONE	NONE	NONE	0
⊖	aspirin	Approved Name	DTXSID5020108	Aspirin	50-78-2	HIGH	HIGH	HIGH	14

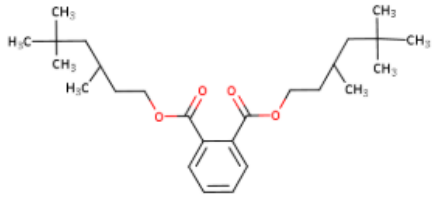
Barcode	Supplier	QTY	Units	Concentration (mM)	Solubility Solvent
00891165	Enamine	-	mg	-	-
TX003515	Sigma Chemical Company	17831	ul	20	DMSO
TX003516	Sigma Chemical Company	19	mg	-	-
TX016586	Sigma Chemical Company	2742	ul	99	DMSO

ChemTrack Shipments

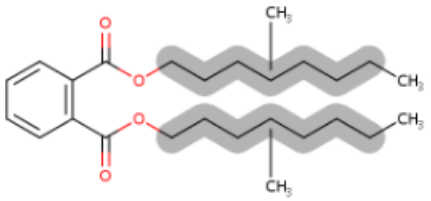


Bottles and Samples Need Chemistry

- Bottle Information
  - Barcode: TX013642
  - CAS-RN: 28553-12-0
  - Name: DIISONONYL PHTHALATE
- CoA Information
  - CAS-RN: 68515-48-0
  - Name: Bis(3,5,5-trimethylhexyl) phthalate
- Which Chemical?????



Bis(3,5,5-trimethylhexyl) phthalate
DTXSID: DTXSID0057889
CASRN: 14103-61-8
TOXCAST: 0



Diisononyl phthalate
DTXSID: DTXSID4022521
CASRN: 28553-12-0
TOXCAST: 10/450

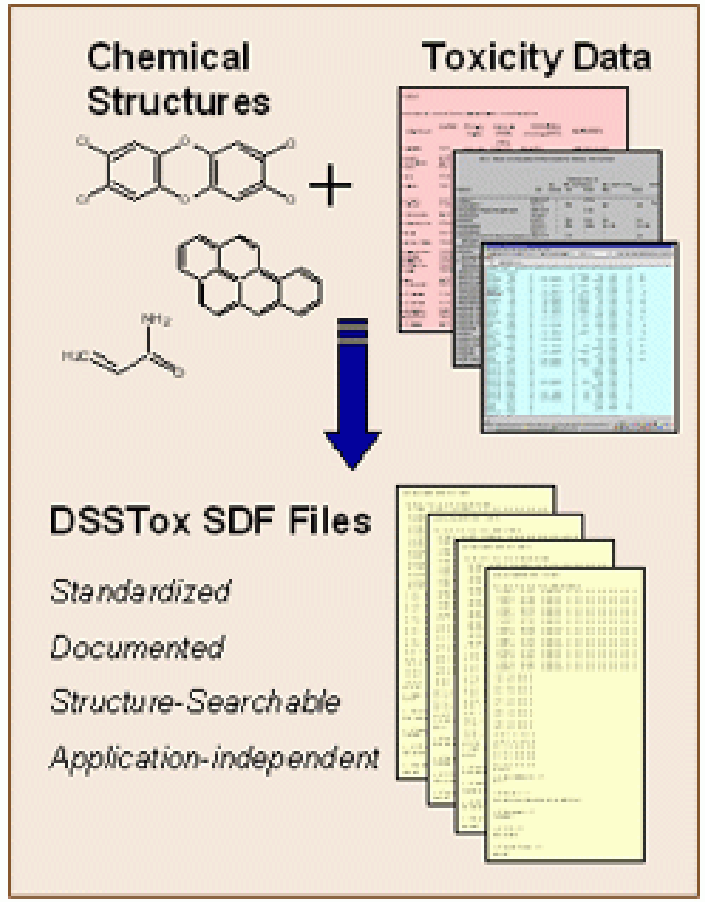
1 related chemical structure with this substance

DINP branched
DTXSID: DTXSID5028665
CASRN: 68515-48-0
TOXCAST: 5/296

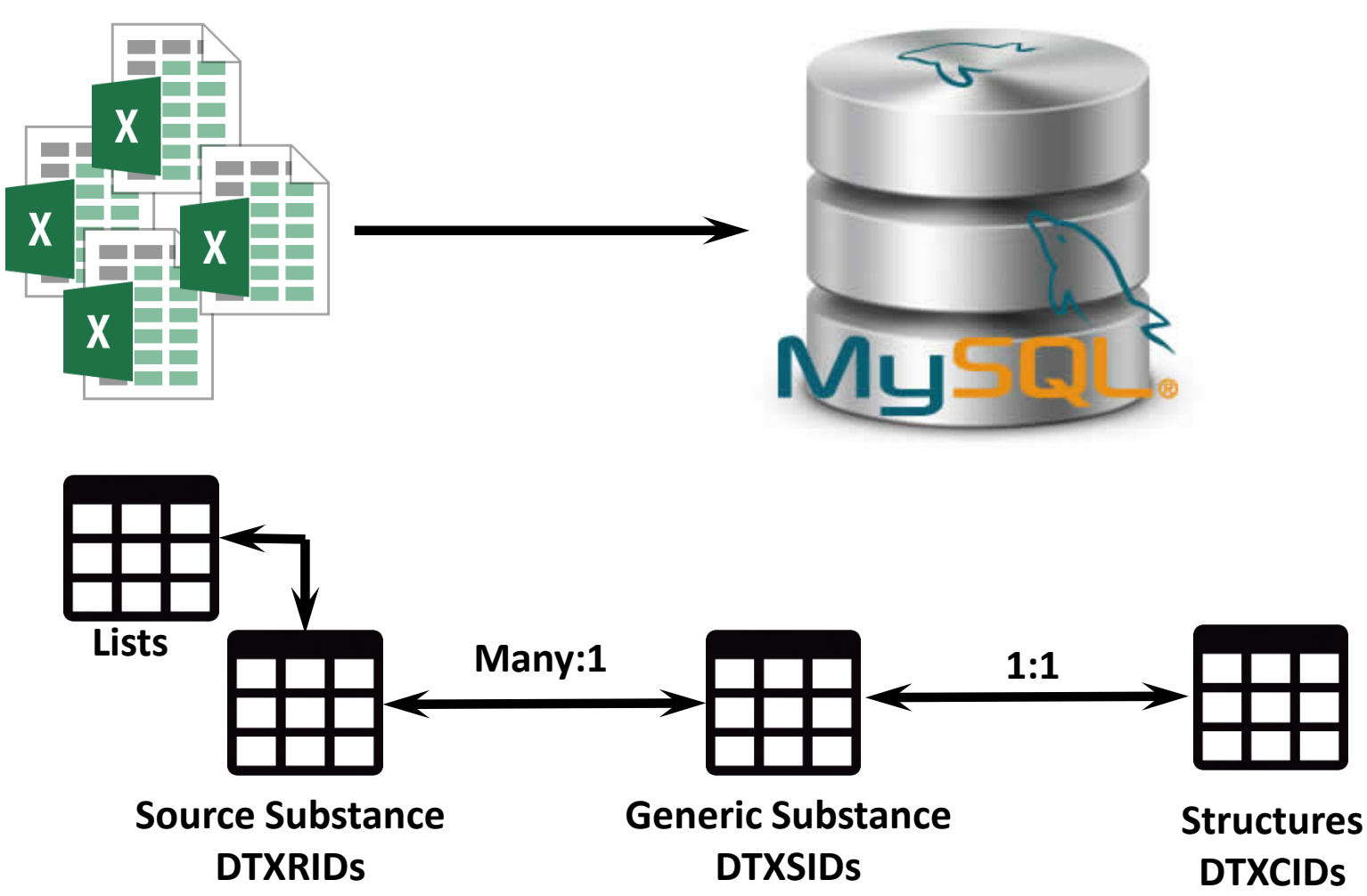
DSSTox Background

Goal: Linking data to chemical structures enabling SAR

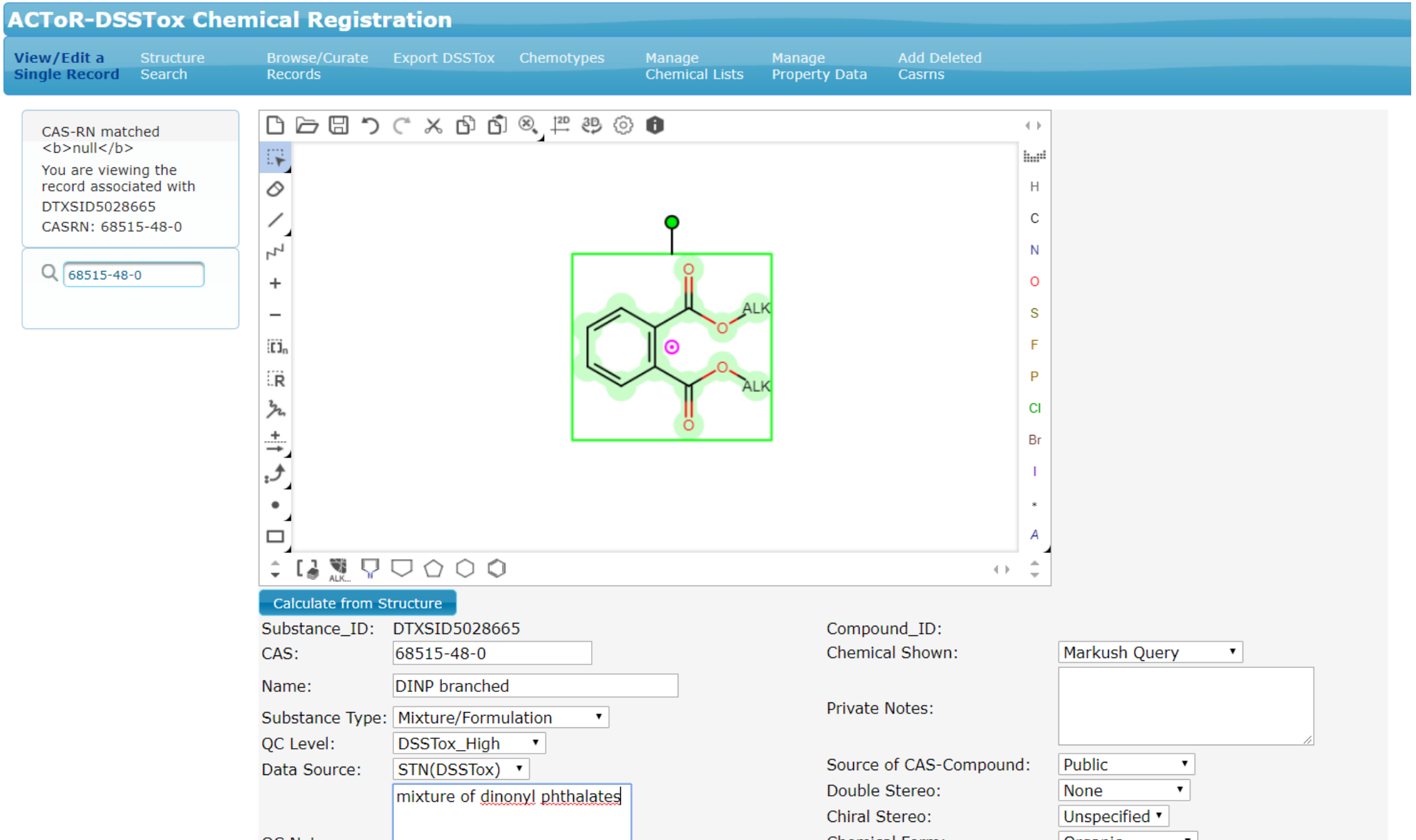
- First release of data files in 2004
- Focused on high impact sets of data
  - Carcinogenic Potency Database
  - Drinking water disinfection by-products
  - EPA's Integrated Risk Information System
  - FDA's Maximum Daily Dose dataset
  - EPA's Fat Head Minnow Toxicity dataset
  - ToxCast and Tox21 chemicals
- Currently contains: 876K records (32K manually curated)
- Check it out: https://comptox.epa.gov/dashboard



Generalized DSSTox Storage Architecture



Chemical Registration



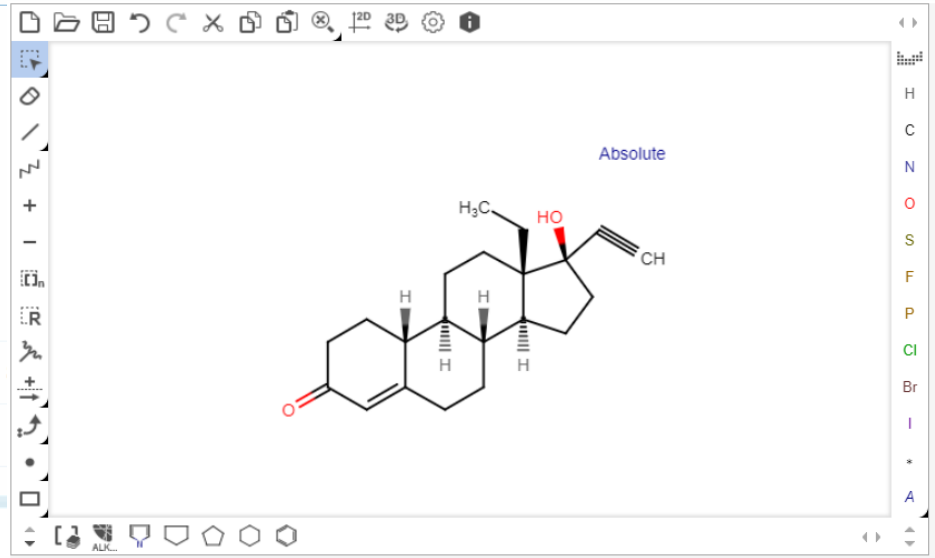
Chemical List Registration

Hits

	ssCAS-RN	ssName	Hit Desc	Hit Substance_ID	Hit Casrn	Hit Name
●		Norgestrel	Structure matched **SMILES**	DTXSID10859541	NOCAS_859541	13-Ethyl-17-ethynyl-17-hydroxy-1,2,6,7,8,9,10,11,12,13,14,15,16,1 tetradecahydro-3H-cyclopenta[a]phenanthren-3-one (non-preferred name)
●		Norgestrel	Mapped Identifier matched **NAME1**	DTXSID3036496	797-63-7	Levonorgestrel
●		Norgestrel	Mapped Identifier matched **NAME1**	DTXSID3047477	6533-00-2	dl-Norgestrel

Map hit | Cancel

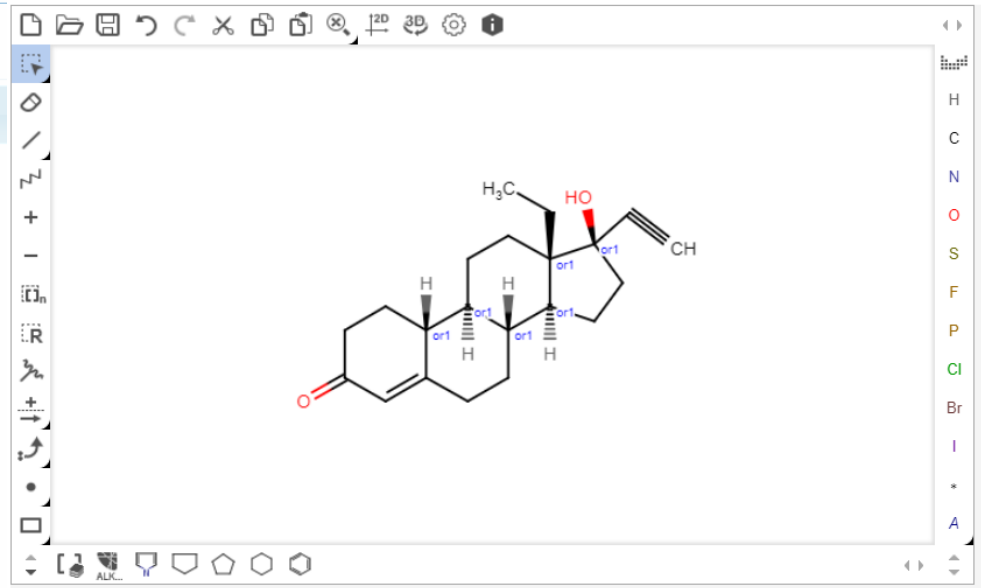
SMILES
Valid Synonym matched **NAME2**
Preferred Name matched **NAME1**
Valid Synonym matched other record: **NAME1**
Unique Synonym matched other record: **NAME2**
Unique Synonym matched other re
Mapped Identifier matched **NAME1**
Mapped Identifier matched **NAME2**
Name2Structure
Export All



Calculate from Structure

Substance_ID: DTXSID3036496
CAS: 797-63-7
Name: Levonorgestrel
Substance Type: Single Compound
QC Level: DSSTox_High
Data Source: STN(DSSTox)

Compound_ID:
Chemical Shown:
Private Notes:
Source of CAS-Compound:

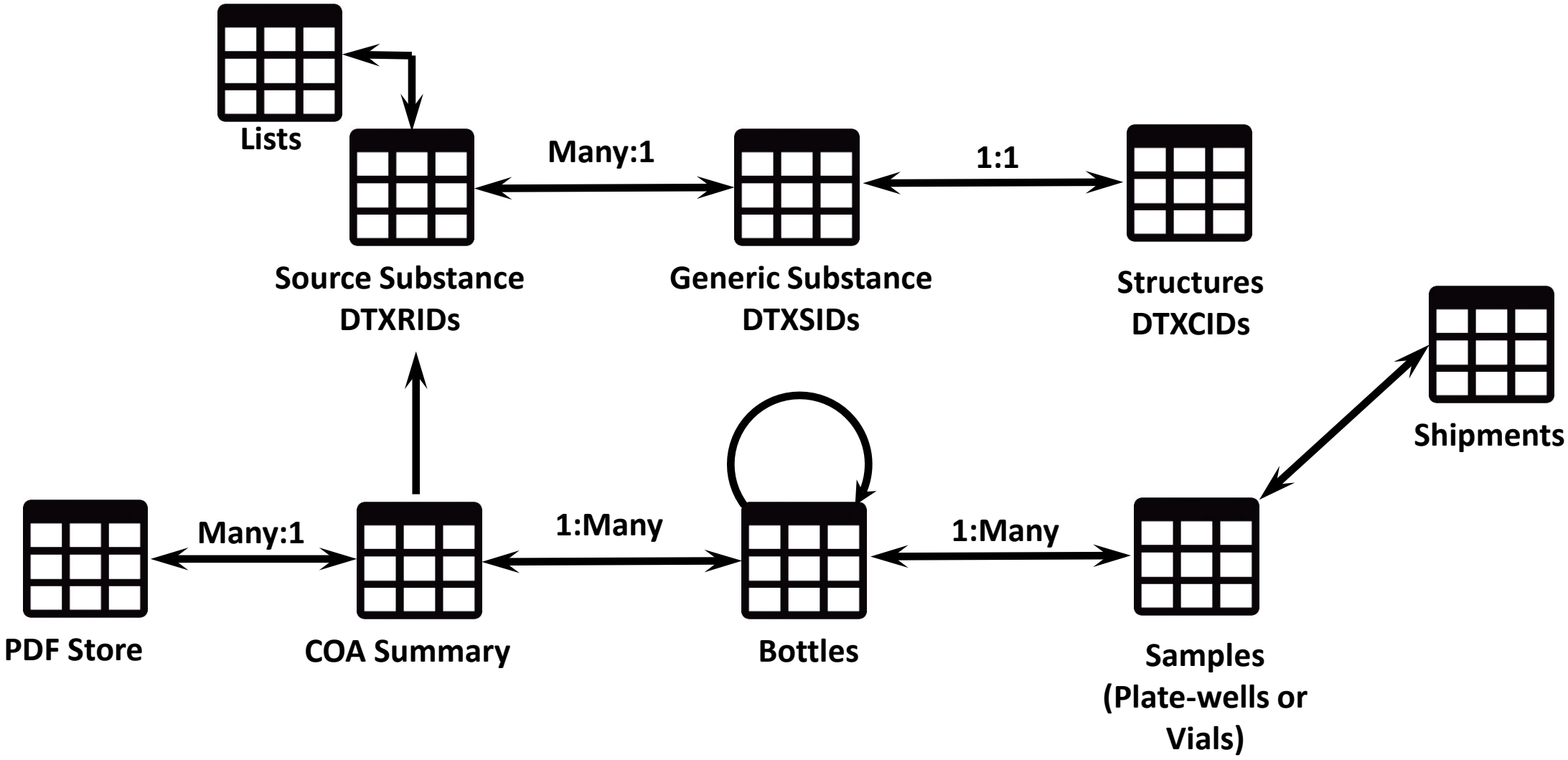


Calculate from Structure

Substance_ID: DTXSID3047477
CAS: 6533-00-2
Name: dl-Norgestrel
Substance Type: Mixture of Stereoisomers
QC Level: DSSTox_High

Compound_ID:
Chemical Shown:
Private Notes:

Link ChemTrack CoA Summaries to DSSTox

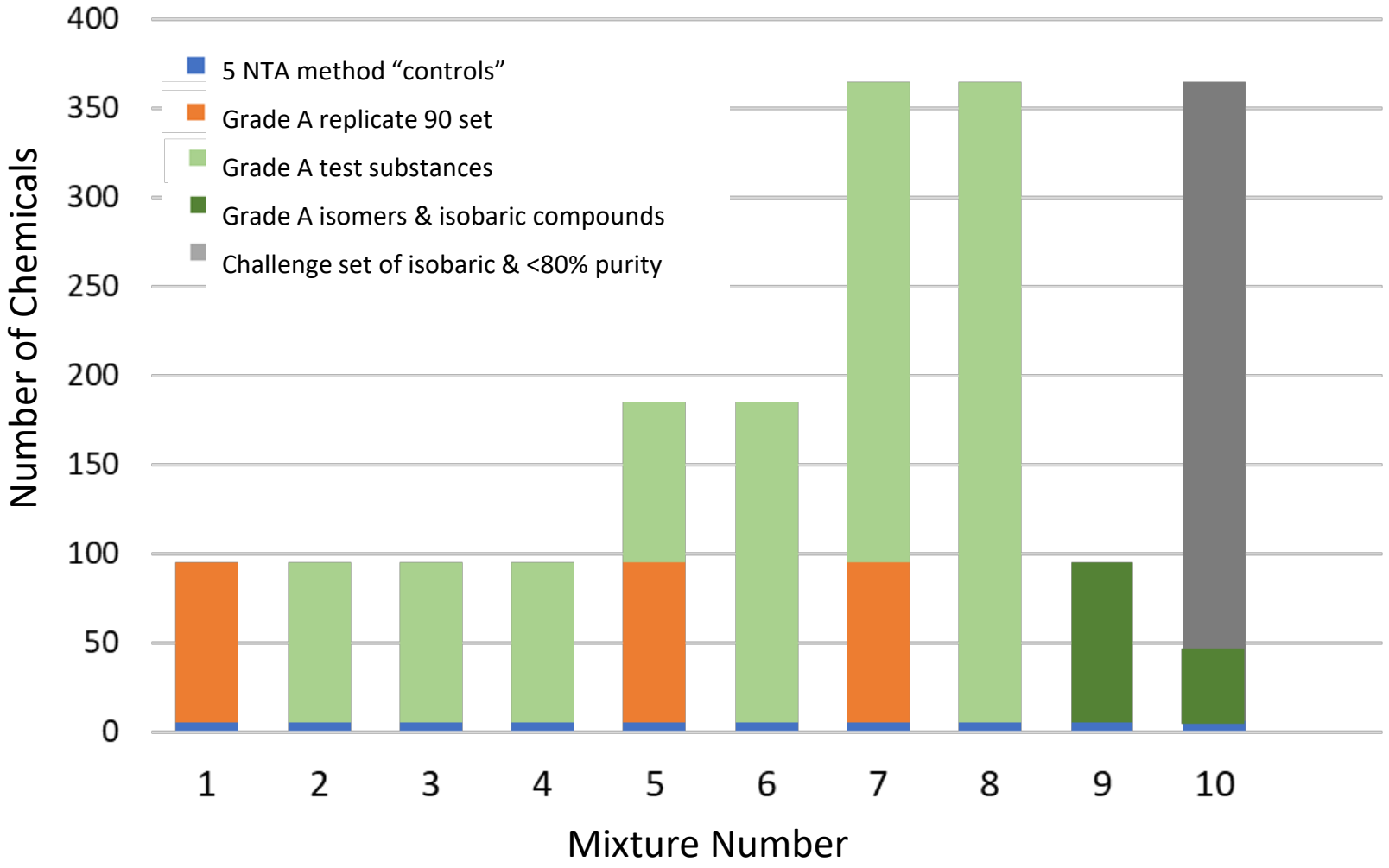


Wait, weren't we making mixtures…

Question(s):

- **What is the relationship between NTA identification (methods) and chemical space detected?**
- Are there chemicals that cannot be detected
- If mixtures are bigger, is detection frequency less?
- If I put the same chemical in a different mixture, can they find it?
- If I put the same chemical in a bigger mixture, can they find it?
- If I put the same chemical in a mixture with an isobaric chemical, can they find it?
- If I throw in some bad QC chemicals, can they find them?
- What if I toss an ill-defined chemical in the mixture?
- Hmmm, polymers are fun…
- Maybe I should vary concentration…

General Mixture Plan…



10 Prepared Mixtures:

1,939 total spiked substances

1,269 unique substances:

1 → spiked 11 times

4 → spiked 10 times

57 → spiked 4 times

33 → spiked 3 times

388 → spiked 2 times

786 → spiked 1 time

Limiting Mixture Difficulty

Comments from Confluence Page on Mixture Creation

September 2015

- *Jon Sobus*: How loose do you want to get with respect to QC levels for the "messy mixtures"? I'm a little worried that a messy mixture with 384 might be too difficult.

- *Chris Grulke*: I think it might be difficult, but not more difficult than dealing with a real media sample…. **Muhuhahahahaha!**

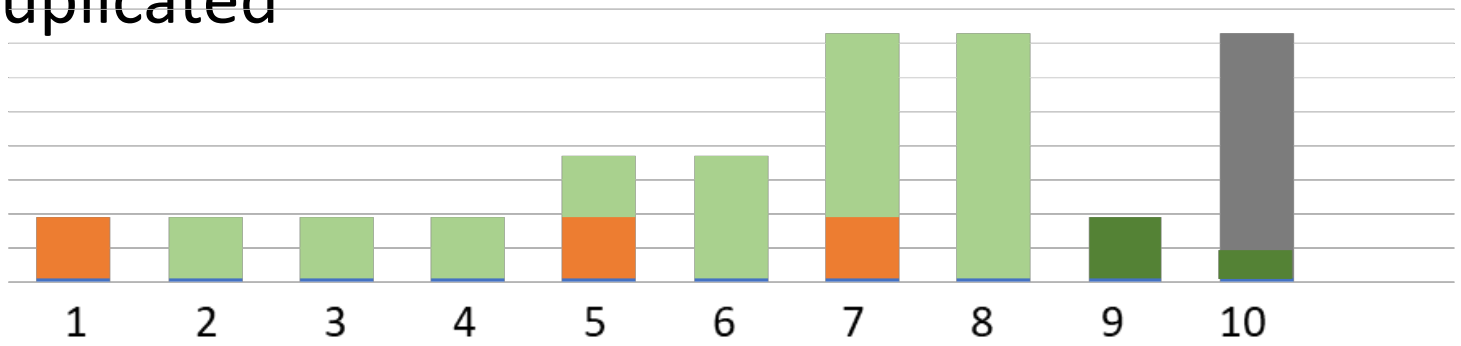
“Easy” Mixtures

General Easy Mixtures (2,3,4,6,8)

1. Sample available in the ToxCast Library
2. Samples Passed Analytical QC
3. No isobaric conflicts (based on 5 ppm resolution)
4. Span the logP and monoisotopic mass range
5. Only controls duplicated

Embedded Set Mixtures (1,5,7)

1. Obey rule 1-4 of General Easy Mixtures
2. Mixtures of 95,185,365 all embedding a common 95 chemicals



Selecting the "Control" Chemicals - Mixture Pilot Successes

- 947 chemicals identifier in dust
- 100 selected and placed in a blinded mixture
- 58 were identified when performing NTA on the mixture
- Jon picked 13
- I picked 5

Rank priority	CAS	DSSTox_name	LC_and_GC	QC_Score	LC_mode	Both concentrations
1	1085-12-7	Heptylparaben	yes	A	neg and pos	yes
2	120-32-1	Clorophene	yes	A	neg and pos	yes
3	13674-87-8	TDCPP	yes	A	neg and pos	yes
4	84-61-7	Dicyclohexyl phthalate	yes	A	neg and pos	yes
5	94-13-3	Propylparaben	yes	A	neg	yes
6	105-99-7	Dibutyl hexanedioate	yes	A	pos	yes
7	63-05-8	4-Androstene-3,17-dione	yes	A	pos	yes
8	63-25-2	Carbaryl	yes	A	pos	yes
9	77-93-0	Triethyl citrate	yes	A	pos	yes
10	78-42-2	Tris(2-ethylhexyl) phosphate	yes	A	pos	yes
11	84-66-2	Diethyl phthalate	yes	A	pos	yes
12	125-33-7	Primidone	yes	A	neg and pos	yes
13	4559-86-8	1,1,3,3-Tetrabutylurea	yes	A	pos	yes

Rager, J.E., et al., 2016. Environ. Int. 88, 269–280

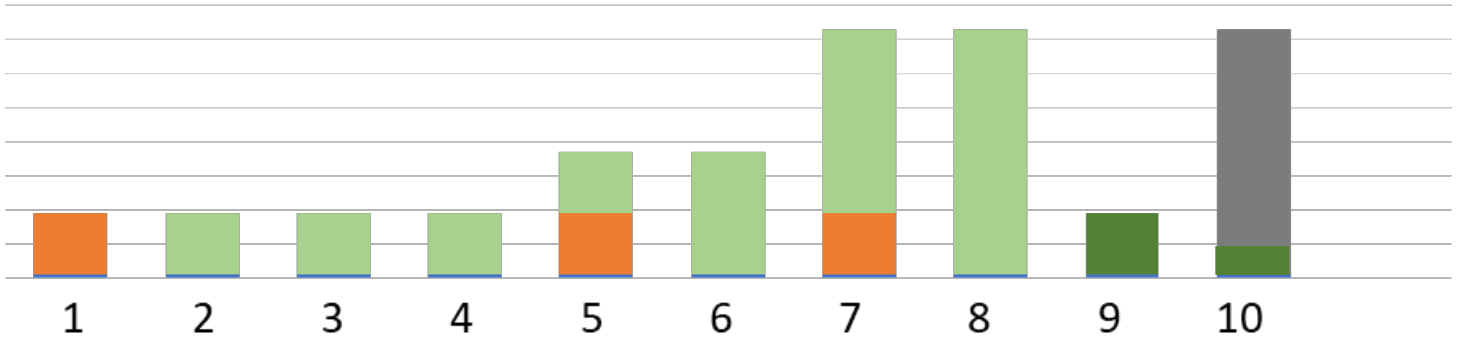
Making Hard Mixtures

Isobar Mixture (9)

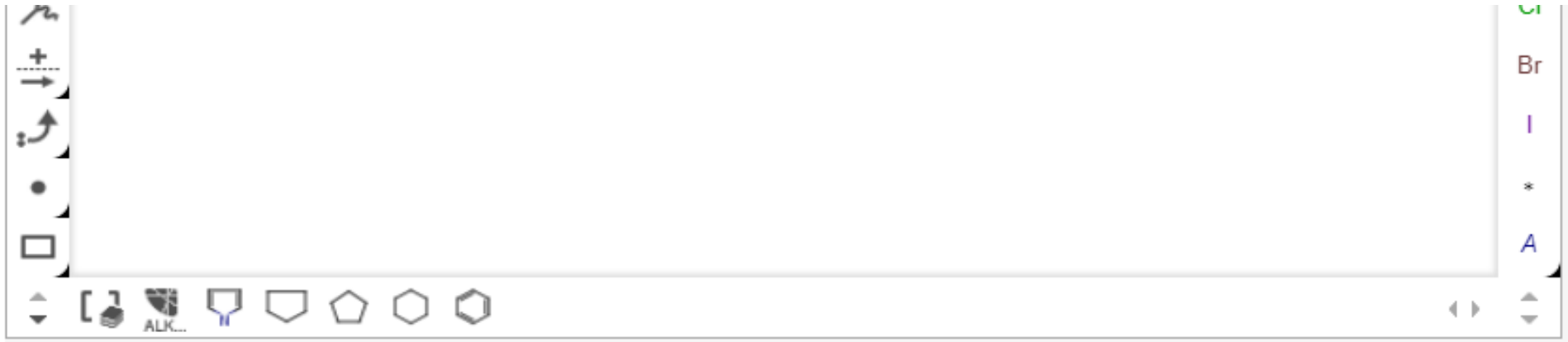
1. Sample available in the ToxCast Library
2. Samples Passed Analytical QC
3. All isobaric conflicts (based on 5 ppm resolution) with conflicted chemicals run individually in Easy Mixtures

Isobar and QC Fail Mixture (10)

1. Dump in the rest of the isobars from the library
2. Add some things which we detected during QC, but with concentration issues.



Mixture Sample Documentation



Calculate from Structure

Substance_ID:	DTXSID30892536	Compound_ID:	
CAS:	NOCAS_892536	Chemical Shown:	No Structure
Name:	NTSMIX_Mix1	Private Notes:	Substance registered to enable loading of EVOTEC shipment files for daughter platings. Additional mapping of successor substances (mixture components) is needed
Substance Type:	Mixture/Formulation		
QC Level:	Incomplete		
Data Source:	Public		
QC Notes:	A mixture of chemicals sent to collaborators as part of a study to evaluate the capabilities and limitations of various NTA methods in identify chemicals	Source of CAS-Compound:	
		Double Stereo:	
		Chiral Stereo:	
		Chemical Form:	

Goals for Mixture Design

Question(s):

- What is the relationship between NTA identification (methods) and chemical space detected?
- Are there chemicals that cannot be detected
- If mixtures are bigger, is detection frequency less?
- If I put the same chemical in a different mixture, can they find it?
- If I put the same chemical in a bigger mixture, can they find it?
- If I put the same chemical in a mixture with an isobaric chemical, can they find it?
- **If I throw in some bad QC chemicals, can they find them?**
- **What if I toss an ill-defined chemical in the mixture?**
- **Hmmm, polymers are fun…**
- **Maybe I should vary concentration…**
- …

Acknowledgements

EPA NCCT ChemTrack
Chris Higgins
Jon Gardner
Kathy Coutros

EPA NCCT IT
Jeff Edwards
Jeremy Dunne
Amar Singh

EPA NCCT Registration
Inthirany Thillainadarajah
Sakuntala Sivasupramaniam

EPA United States Environmental Protection Agency

Questions?