United States Environmental Protection Agency

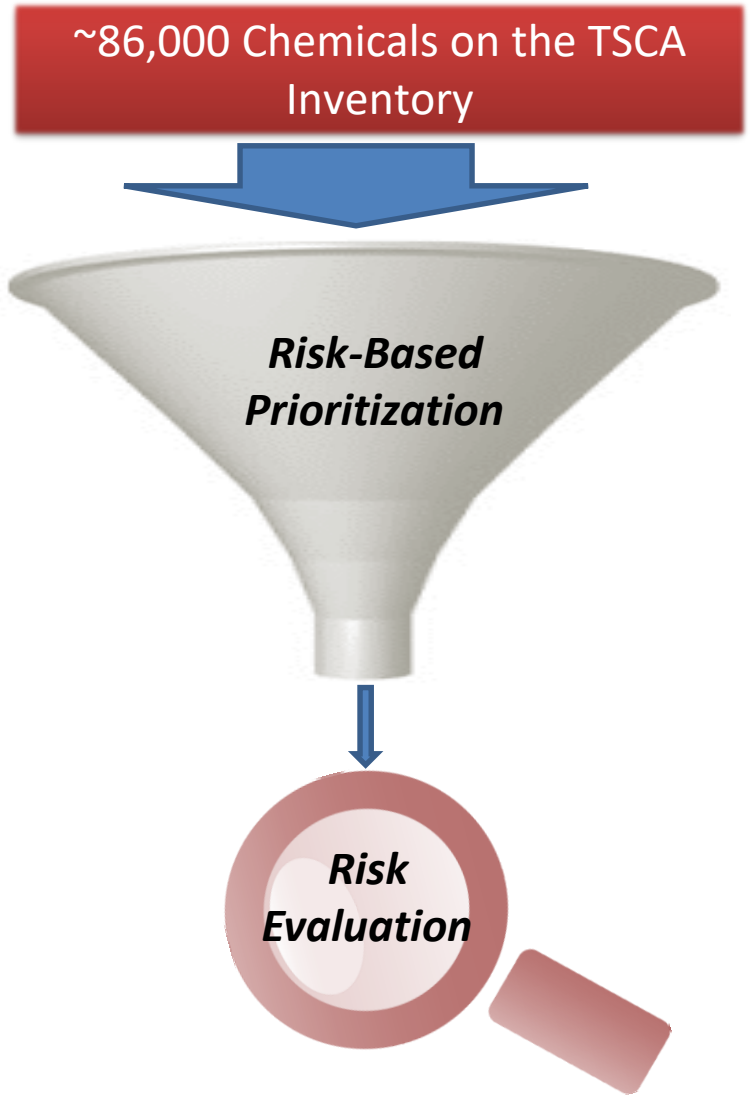
Findings from EPA's Non-Targeted Analysis Collaborative Trial (ENTACT)

Jon Sobus[1], Elin Ulrich[1], Jarod Grossman[2], Alex Chao[2], Seth Newton[1], Antony Williams[1], Ann Richard[1], Chris Grulke[1], Andrew McEachran[2], Randolph Singh[2], Hussein Al-Ghoul[2], Louis Groff[2]

[1] Center for Computational Toxicology and Exposure
[2] ORAU/ORISE Participant

Drivers for EPA Research Initiatives

- Many industrial & commercial chemicals are covered by the Toxic Substances Control Act (TSCA), which is administered by EPA.
- TSCA updated in June 2016 to allow *risk-based* evaluation of existing and new chemicals.
- Characterization of risk requires exposure and hazard data.
- EPA's Office of Research and Development (ORD) is developing new approach methodologies (NAMs) for rapid risk characterization.



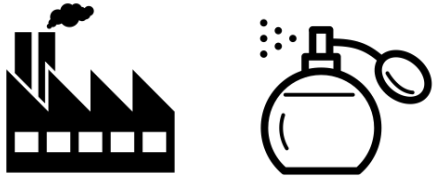
The Era of High-Throughput Assessments

United States Environmental Protection Agency

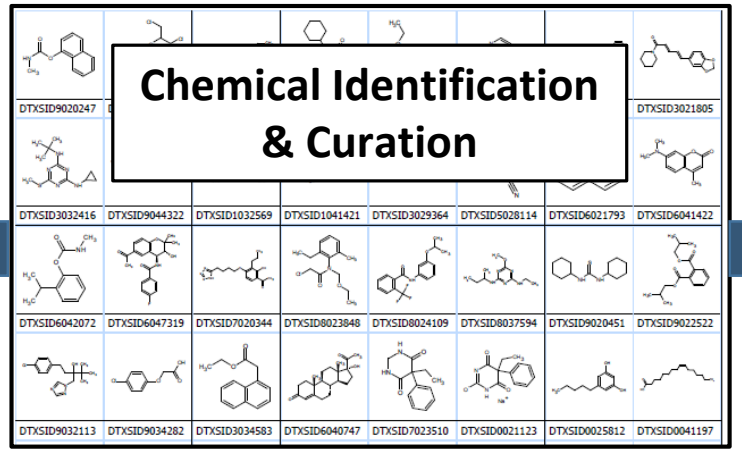
Exposure Forecasting (ExpoCast)

Toxicity Forecasting (ToxCast)

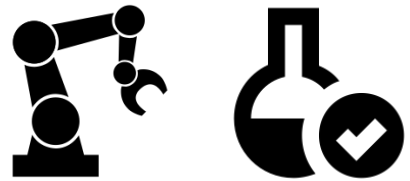
Production, Use & Scenario Data



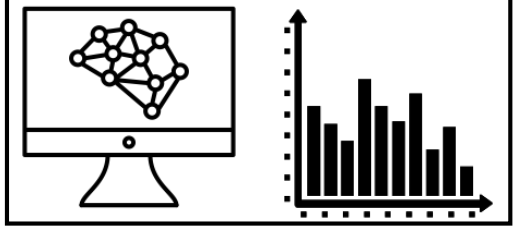
Chemical Identification & Curation



Procurement or Synthesis

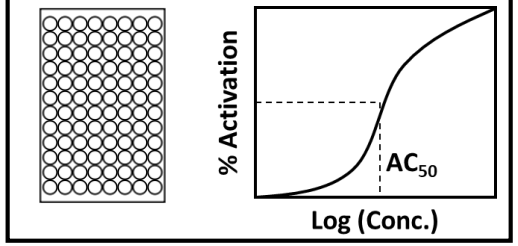


HT Exposure Prediction

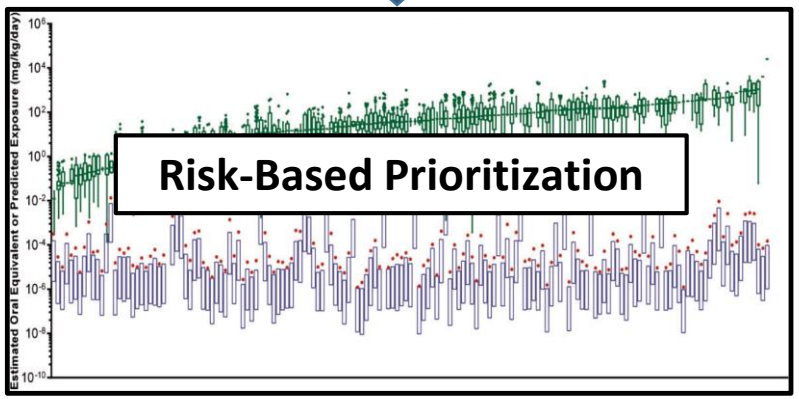


Extrapolation

HT Tox Assays



Risk-Based Prioritization



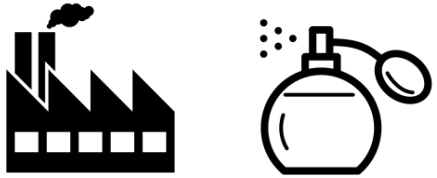
The Era of High-Throughput Assessments

United States Environmental Protection Agency

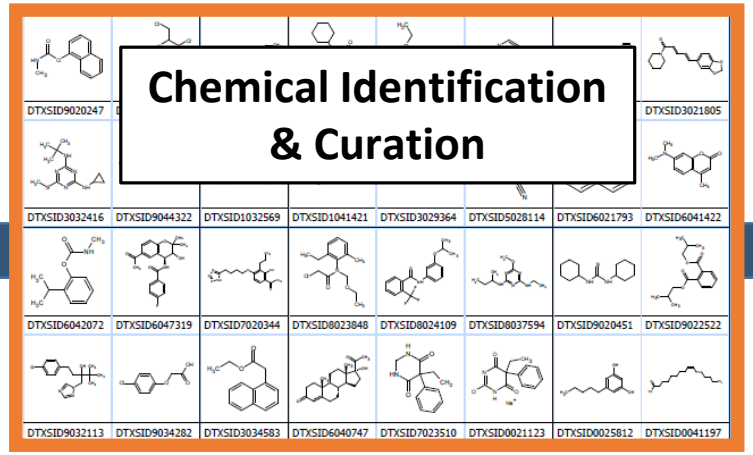
Exposure Forecasting (ExpoCast)

Toxicity Forecasting (ToxCast)

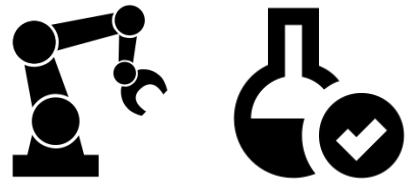
Production, Use & Scenario Data



Chemical Identification & Curation



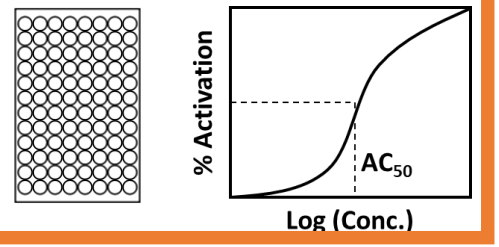
Procurement or Synthesis



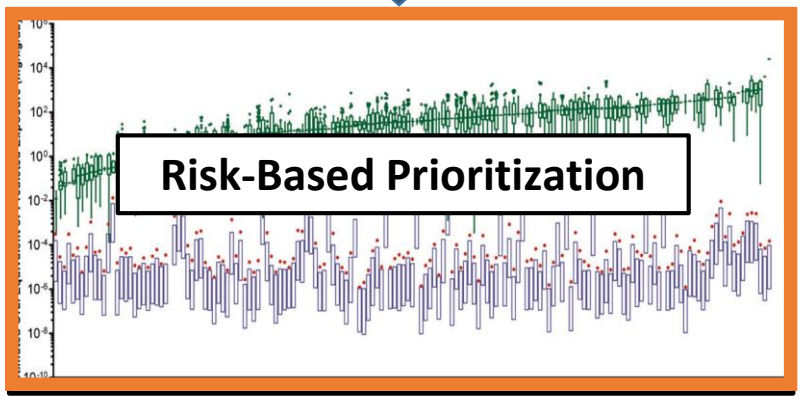
HT Exposure Prediction

Extrapolation

HT Tox Assays



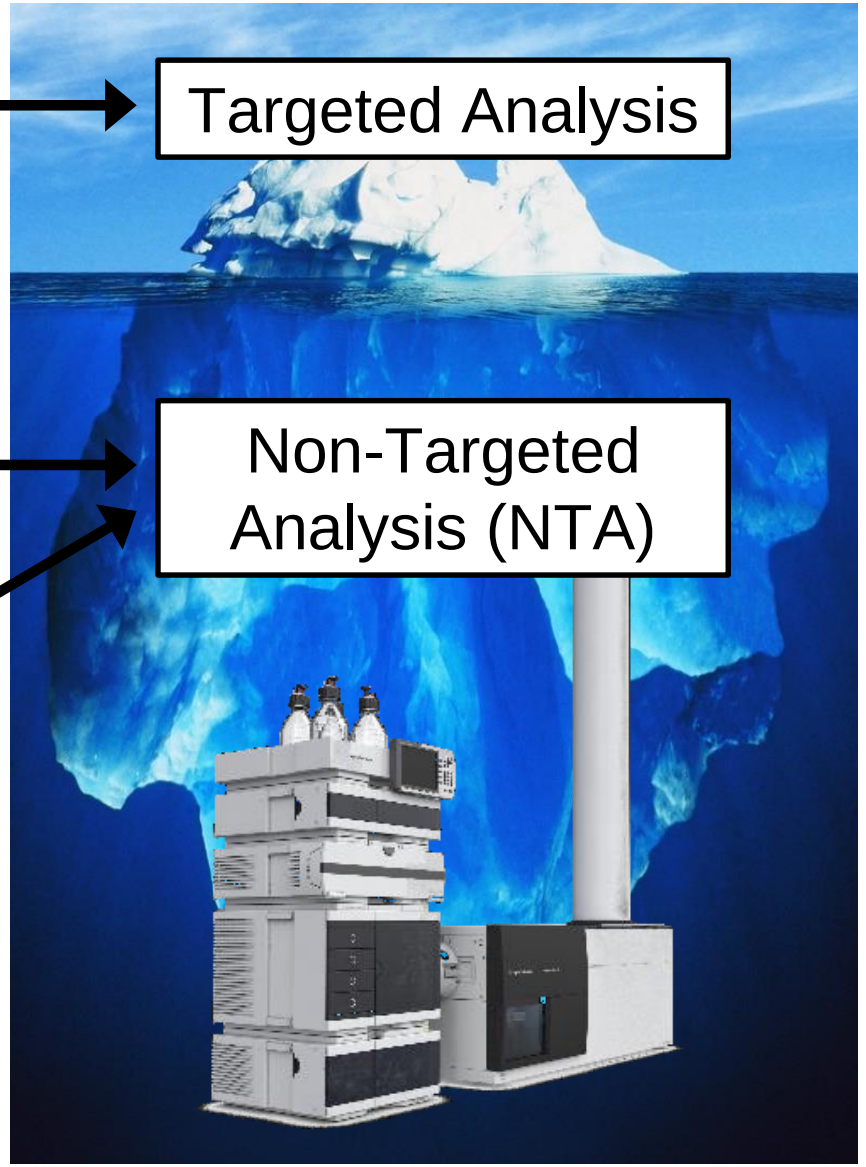
Risk-Based Prioritization



Guided and/or enhanced by Measurement Data!!

The Need for Chemical Measurement Data

- **Well-known chemicals**
  - 100s - 1,000s (e.g., NHANES)
  - Quality exposure data
- **Known but data-poor chemicals**
  - 1,000s - 1,000,000s (e.g., TSCA)
  - Limited exposure data
- **Chemicals not yet known to exist**
  - Unknown #
  - No exposure data



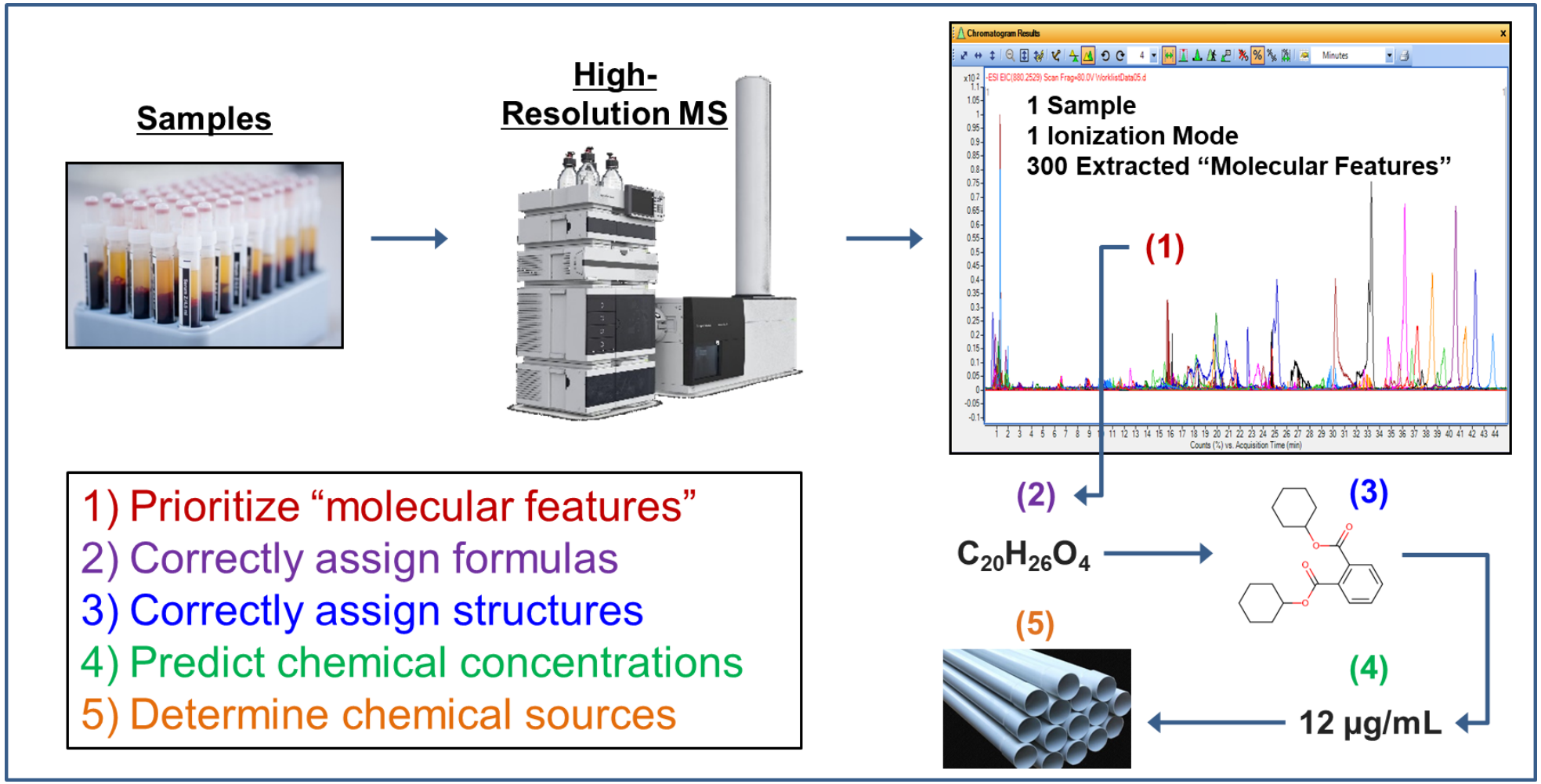
What's So Great About NTA?

Rapidly screen for "knowns"

Discover "unknowns"

Uncover historical exposures

Generate source fingerprints…



Example Uses and Requirements

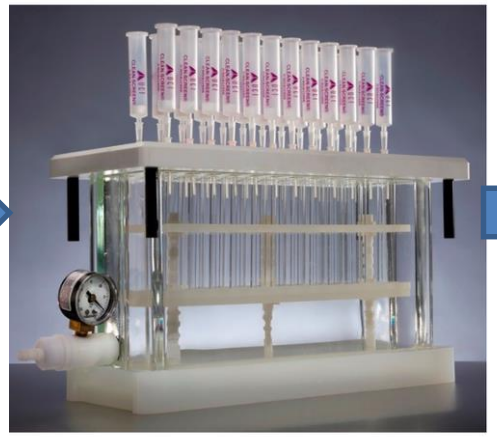
Study Level	Decision Context			Example Uses of NTA Data	Example Stakeholders
	Sample Classification	*Chemical Annotation*	*Semi-Quantitation*		
1	Required	Optional	Optional	- Classify locations impacted by point-source emitters - Classify locations impacted by inadvertent environmental releases - Classify exposure status for active or former military personnel - Classify food items not meeting criteria for product certification	- EPA, USGS - FEMA, EPA - DoD, VA - FDA, NIST
2	Required	Required	Optional	- Identify natural or synthetic chemical nerve agents - Identify chemicals associated with product-related illness - Identify chemicals released in emergency response scenarios - Identify designer drugs used for athletic performance enhancement	- DHS, CDC - CPSC, FDA - FEMA, EPA - DEA, FDA
3	Required	Required	Required	- Assess occupational health risks from exposure to fire-fighting foams - Assess consumer health risks from exposure to household products - Assess ecological health risks from exposure to urban wastewater - Assess maternal and infant health risk from exposure during pregnancy	- NIOSH, DoD - CPSC, EPA - USGS, EPA - NIEHS, EPA

Semi-Quant NTA is a Multi-Step Process

United States Environmental Protection Agency

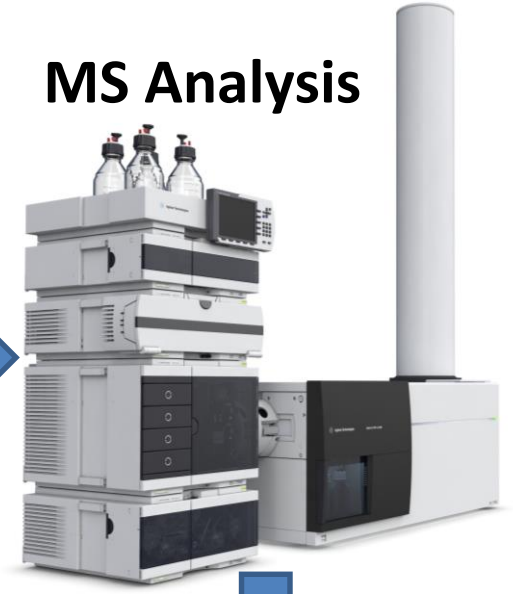
Media Sample

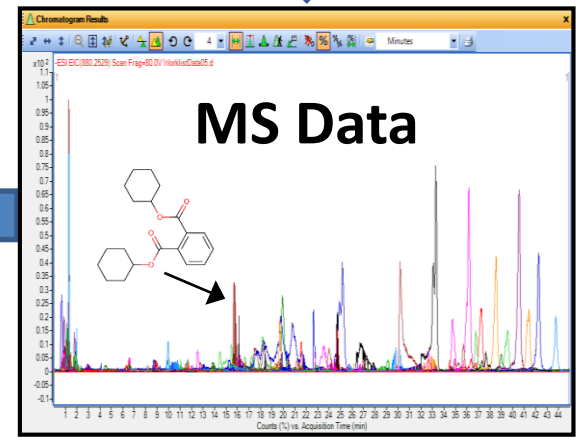
Extraction & Cleanup



Prepared Sample

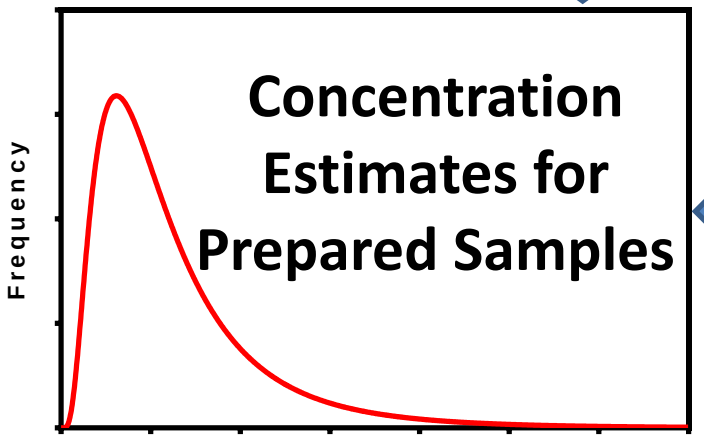
MS Analysis





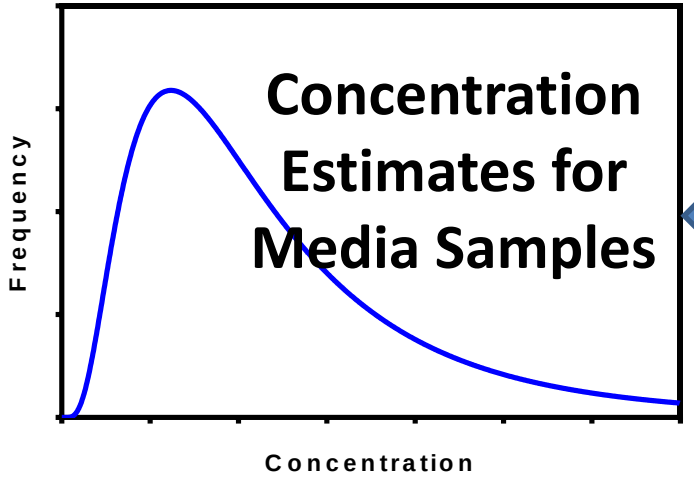
?

Concentration Estimates for Prepared Samples



?

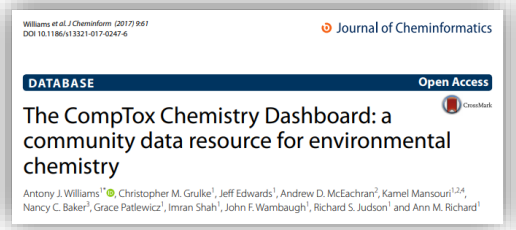
Concentration Estimates for Media Samples



Building an NTA Research Program

United States Environmental Protection Agency

Data, Tools & Informatics


Williams et al. J Cheminform (2017) 9:61
DOI 10.1186/s13321-017-0247-6

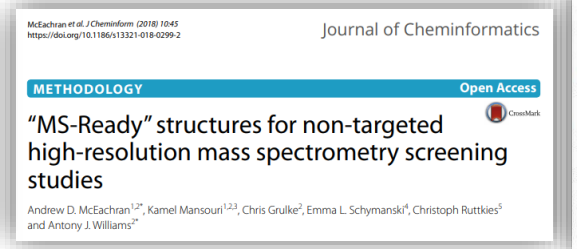
Journal of Cheminformatics

DATABASE Open Access

The CompTox Chemistry Dashboard: a community data resource for environmental chemistry

Antony J. Williams[1*], Christopher M. Grulke[1], Jeff Edwards[1], Andrew D. McEachran[2], Kamel Mansouri[1,2,4], Nancy C. Baker[3], Grace Patlewicz[1], Imran Shah[1], John F. Wambaugh[1], Richard S. Judson[1] and Ann M. Richard[1]

Chemicals Dashboard


McEachran et al. J Cheminform (2018) 10:45
https://doi.org/10.1186/s13321-018-0299-2

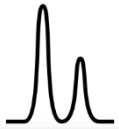
Journal of Cheminformatics

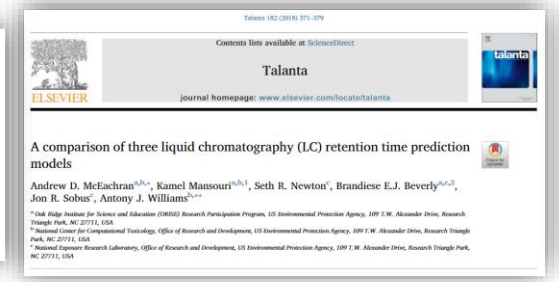
METHODOLOGY Open Access

"MS-Ready" structures for non-targeted high-resolution mass spectrometry screening studies

Andrew D. McEachran[1,2*], Kamel Mansouri[1,2,3], Chris Grulke[2], Emma L. Schymanski[4], Christoph Ruttkies[5] and Antony J. Williams[2*]

"MS Ready" Structures




Talanta 182 (2018) 371–379

Contents lists available at ScienceDirect

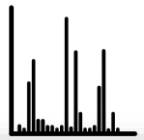
Talanta

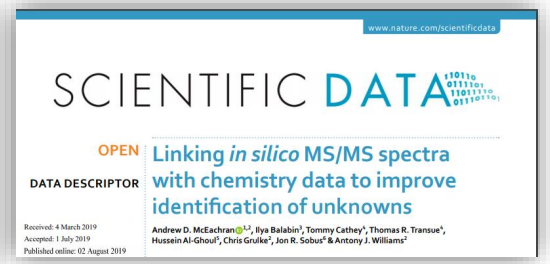
journal homepage: www.elsevier.com/locate/talanta

A comparison of three liquid chromatography (LC) retention time prediction models

Andrew D. McEachran[a,b,*], Kamel Mansouri[a,b,1], Seth R. Newton[c], Brandiese E.J. Beverly[a,c,2], Jon R. Sobus[a], Antony J. Williams[a,*]

RT Models




www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Linking *in silico* MS/MS spectra with chemistry data to improve identification of unknowns

Received: 4 March 2019
Accepted: 1 July 2019
Published online: 02 August 2019

Andrew D. McEachran[1,2], Ilya Balabin[3], Tommy Cathey[4], Thomas R. Transue[4], Hussein Al-Ghoul[3], Chris Grulke[2], Jon R. Sobus[2] & Antony J. Williams[2]

In Silico Spectra

Semi-Quant. Methods

EPA NTA WebApp

2016… 2017… 2019… 2020…

NTA Applications

Brita Filters

Environmental Pollution 234 (2018) 297–306

Contents lists available at ScienceDirect

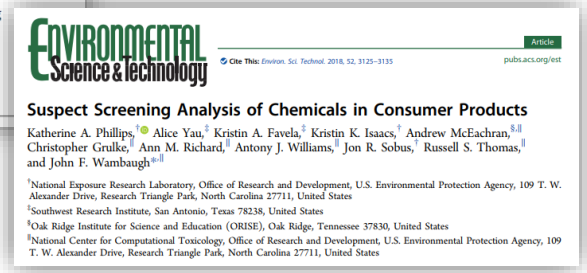
Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Suspect screening and non-targeted analysis of drinking water using point-of-use filters☆

Seth R. Newton[a,*], Rebecca L. McMahen[a,b], Jon R. Sobus[a], Kamel Mansouri[b,c,1], Antony J. Williams[c], Andrew D. McEachran[b,c], Mark J. Strynar[a]

Consumer Products


Environmental Science & Technology

Article

Cite This: Environ. Sci. Technol. 2018, 52, 3125−3135

pubs.acs.org/est

Suspect Screening Analysis of Chemicals in Consumer Products

Katherine A. Phillips,[†] Alice Yau,[‡] Kristin A. Favela,[‡] Kristin K. Isaacs,[†] Andrew McEachran,[§,∥] Christopher Grulke,[∥] Ann M. Richard,[∥] Antony J. Williams,[∥] Jon R. Sobus,[†] Russell S. Thomas,[∥] and John F. Wambaugh[*,∥]

[†]National Exposure Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, 109 T. W. Alexander Drive, Research Triangle Park, North Carolina 27711, United States
[‡]Southwest Research Institute, San Antonio, Texas 78238, United States
[§]Oak Ridge Institute for Science and Education (ORISE), Oak Ridge, Tennessee 37830, United States
[∥]National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, 109 T. W. Alexander Drive, Research Triangle Park, North Carolina 27711, United States

Model Organisms

OXFORD SOT Society of Toxicology academic.oup.com/toxsci Tox Spotlight

TOXICOLOGICAL SCIENCES, 172(1), 2019, 109–122

Advance Access Publication Date: September 5, 2019

Research Article

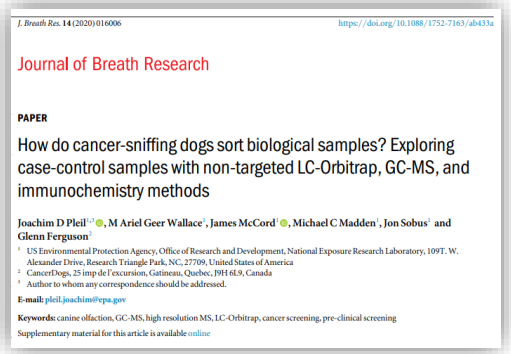
Triclosan-Selected Host-Associated Microbiota Perform Xenobiotic Biotransformations in Larval Zebrafish

Chelsea A. Weitekamp,[*,1] Drake Phelps,[*,1] Adam Swank,[†] James McCord,[*] Jon R. Sobus,[‡] Tara Catron,[*] Scott Keely,[§] Nichole Brinkman,[§] Todd Zurlinden,[¶] Emily Wheaton,[§] Mark Strynar,[‡] Charlene McQueen,[‖] Charles E. Wood,[‖,2] and Tamara Tal[‖,3]

[*]Oak Ridge Institute for Science and Education/U.S. EPA/ORD/NHEERL/ISTD, Research Triangle Park, North Carolina, 27711; [†]U.S. EPA/ORD/NHEERL/RCU, Research Triangle Park, North Carolina, 27711; [‡]U.S. EPA/ORD/NERL/EMMD, Research Triangle Park, North Carolina, 27711; [§]U.S. EPA/ORD/NERL/SED, Cincinnati, Ohio, 45220; [¶]U.S. EPA/ORD/NCCT, Research Triangle Park, North Carolina, 27711; [‖]University of Arizona, Department of Pharmacology and Toxicology, Tucson, Arizona, 85721; and [⊥]U.S. EPA/ORD/NHEERL/ISTD, Research Triangle Park, North Carolina, 27711

[1]These authors contributed equally to this study.

Hospital Masks


J. Breath Res. 14 (2020) 016006

https://doi.org/10.1088/1752-7163/ab433a

Journal of Breath Research

PAPER

How do cancer-sniffing dogs sort biological samples? Exploring case-control samples with non-targeted LC-Orbitrap, GC-MS, and immunochemistry methods

Joachim D Pleil[1,3], M Ariel Geer Wallace[1], James McCord[1], Michael C Madden[1], Jon Sobus[1] and Glenn Ferguson[2]

[1] US Environmental Protection Agency, Office of Research and Development, National Exposure Research Laboratory, 109T. W. Alexander Drive, Research Triangle Park, NC, 27709, United States of America
[2] CancerDogs, 25 imp de l'excursion, Gatineau, Quebec, J9H 6L9, Canada
[3] Author to whom any correspondence should be addressed.

E-mail: pleil.joachim@epa.gov

Keywords: canine olfaction, GC-MS, high resolution MS, LC-Orbitrap, cancer screening, pre-clinical screening

Supplementary material for this article is available online

Recycled Products

Human Placentas

NTA State-of-the-Science

Environmental Science & Technology

Viewpoint

Cite This: *Environ. Sci. Technol.* 2018, 52, 11975−11976

pubs.acs.org/est

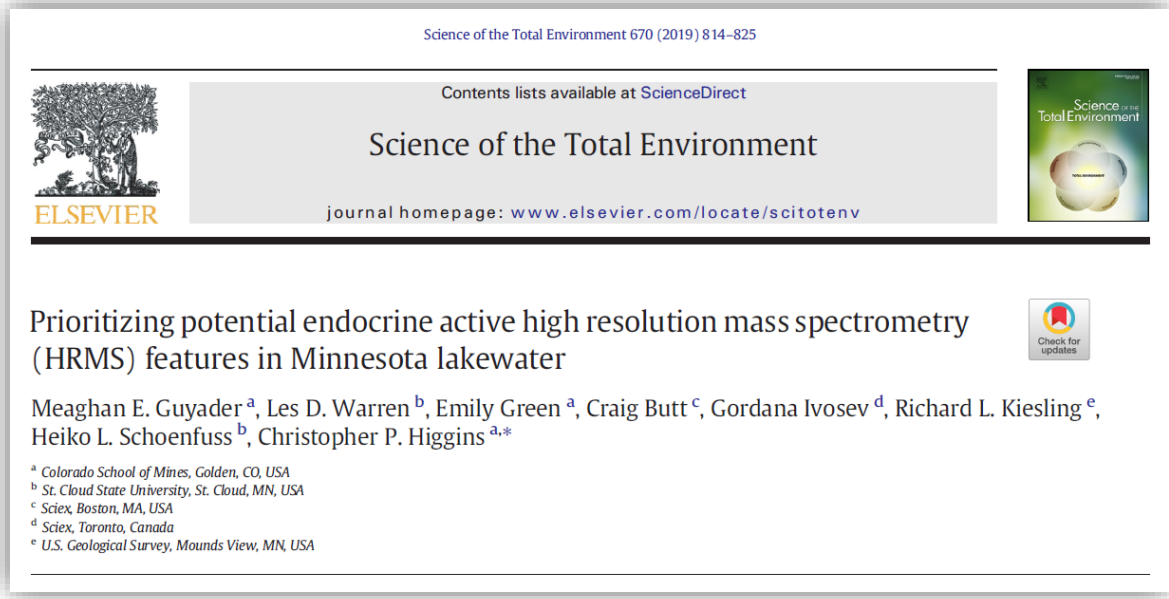
Is Nontargeted Screening Reproducible?

Ronald A. Hites*

School of Public and Environmental Affairs, Indiana University, Bloomington, Indiana 47405, United States

Karl J. Jobst*

Department of Chemistry and Chemical Biology, McMaster University, Hamilton, Ontario L8S 4M1, Canada



Science of the Total Environment 670 (2019) 814–825

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

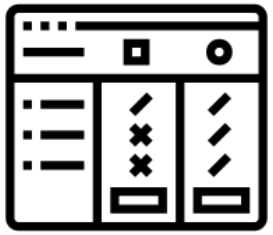
Prioritizing potential endocrine active high resolution mass spectrometry (HRMS) features in Minnesota lakewater

Meaghan E. Guyader [a], Les D. Warren [b], Emily Green [a], Craig Butt [c], Gordana Ivosev [d], Richard L. Kiesling [e], Heiko L. Schoenfuss [b], Christopher P. Higgins [a,*]

[a] Colorado School of Mines, Golden, CO, USA
[b] St. Cloud State University, St. Cloud, MN, USA
[c] Sciex, Boston, MA, USA
[d] Sciex, Toronto, Canada
[e] U.S. Geological Survey, Mounds View, MN, USA

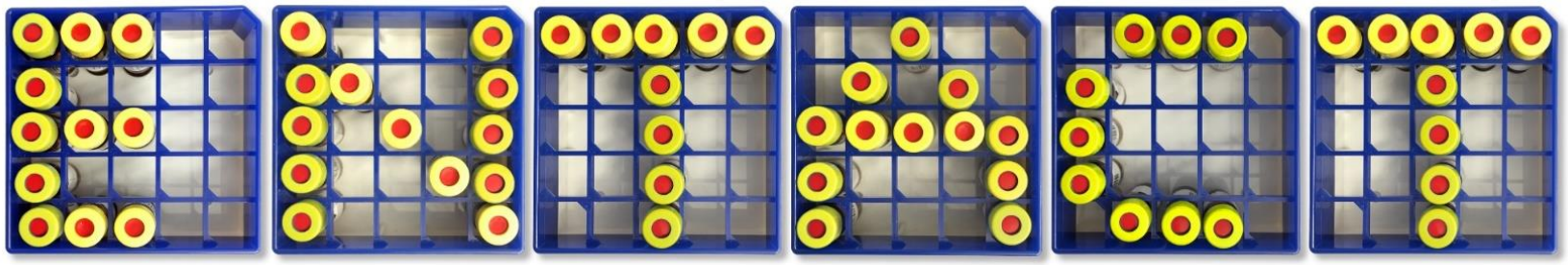
"*No single analytical technique is suitable for the analysis of all compounds, and successful nontargeted screening will require the development of multiplatform approaches, facilitated and validated through interlaboratory collaborations*."

"*The novelty of nontarget analysis, particularly its current lack of implementation by regulatory agencies, has prevented the establishment of streamlined quality assurance and quality control (QA/QC) procedures.*"



Science Questions for Research Community

- How variable are tools and results from lab to lab?
- Are some methods/workflows better than others?
- How does sample complexity affect performance?
- What chemical space does a given method cover?
- How sensitive are specific instruments/methods?



EPA's Non-Targeted Analysis Collaborative Trial

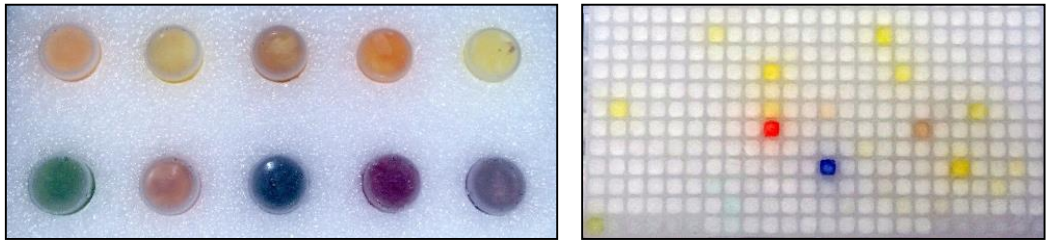
ENTACT Part 1

Chemicals from ToxCast Library

~1200 ToxCast Chemicals (highest quality)

10 Mixtures (100-400 chemicals each)

*Multi-Well Plates**



~25 Collaborators & 6 Contractors*:

1st: Blinded analysis

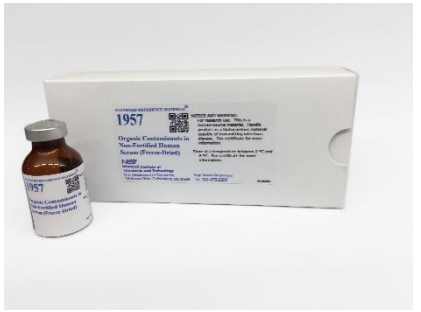
2nd: Unveiling of chemicals

3rd: Unblinded evaluation

ENTACT Part 2

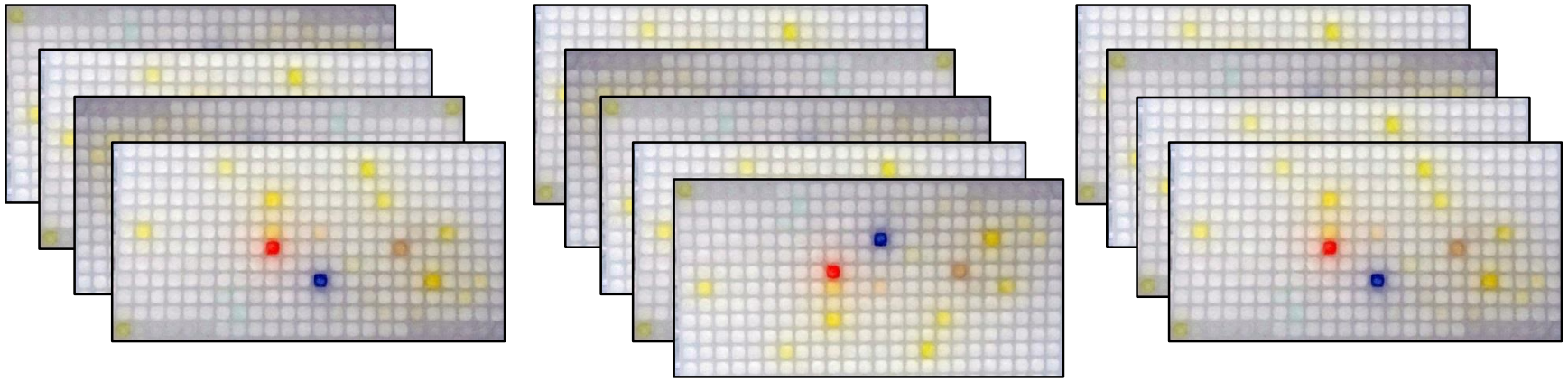
Reference & Fortified House Dust

Reference & Fortified Human Serum



Reference & Fortified Silicone Wristbands

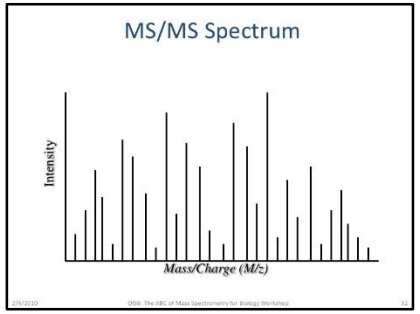
ENTACT Part 3



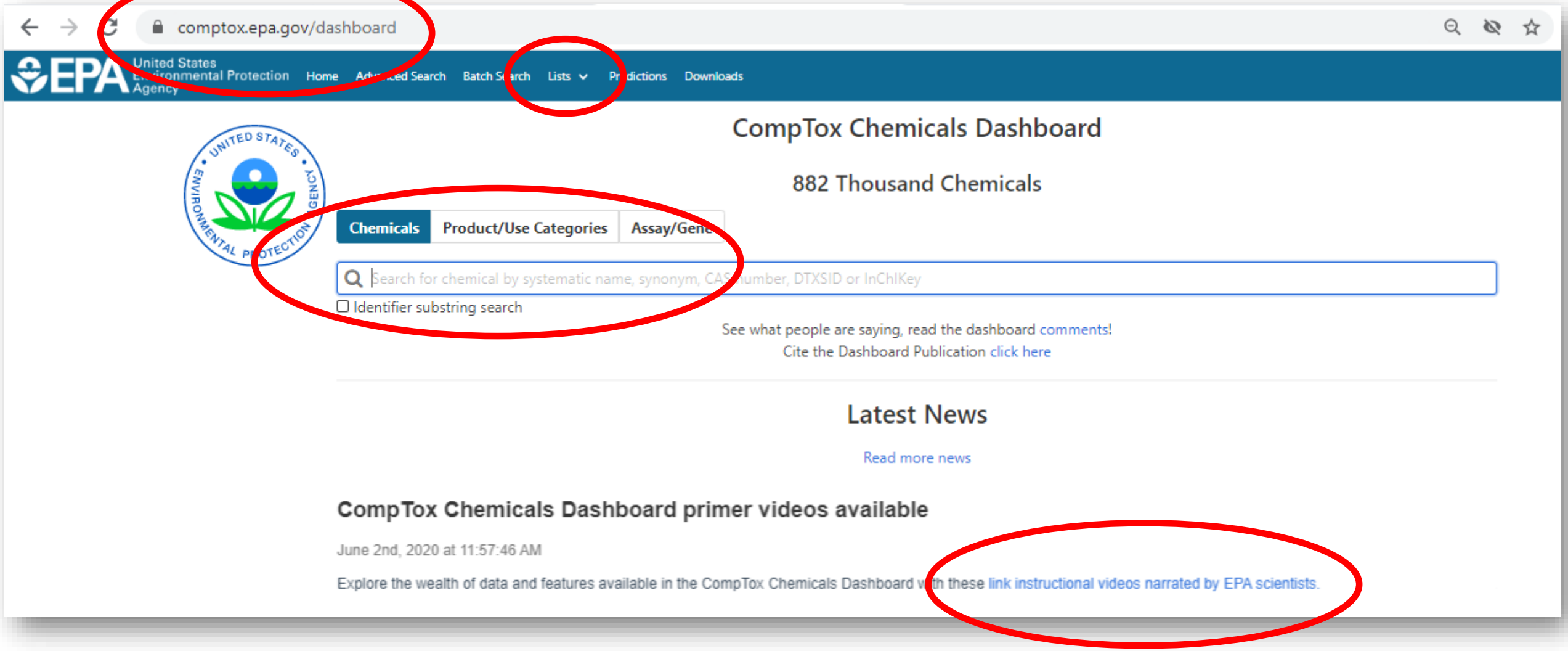
~4600 ToxCast substances

Instrument/software vendors & select labs

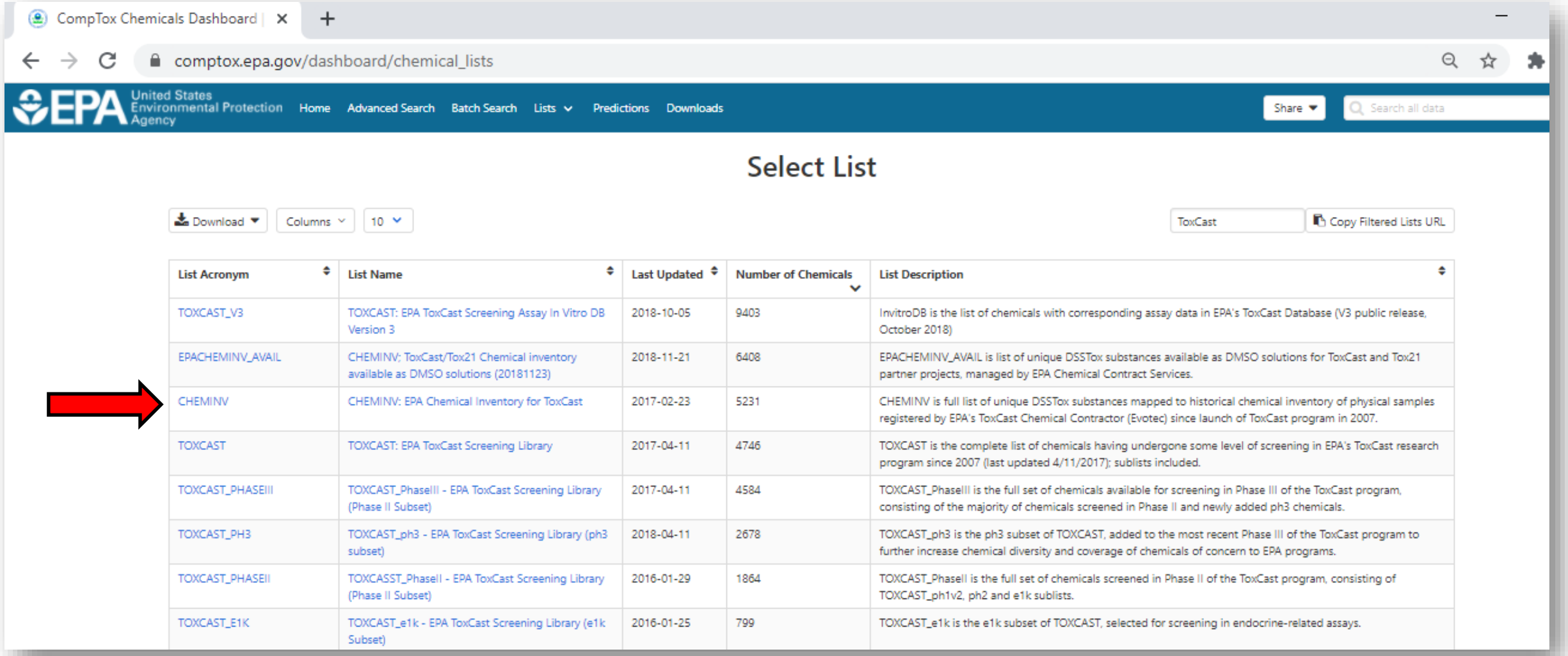
Reference libraries for the public



Accessing ENTACT Chemistry Data

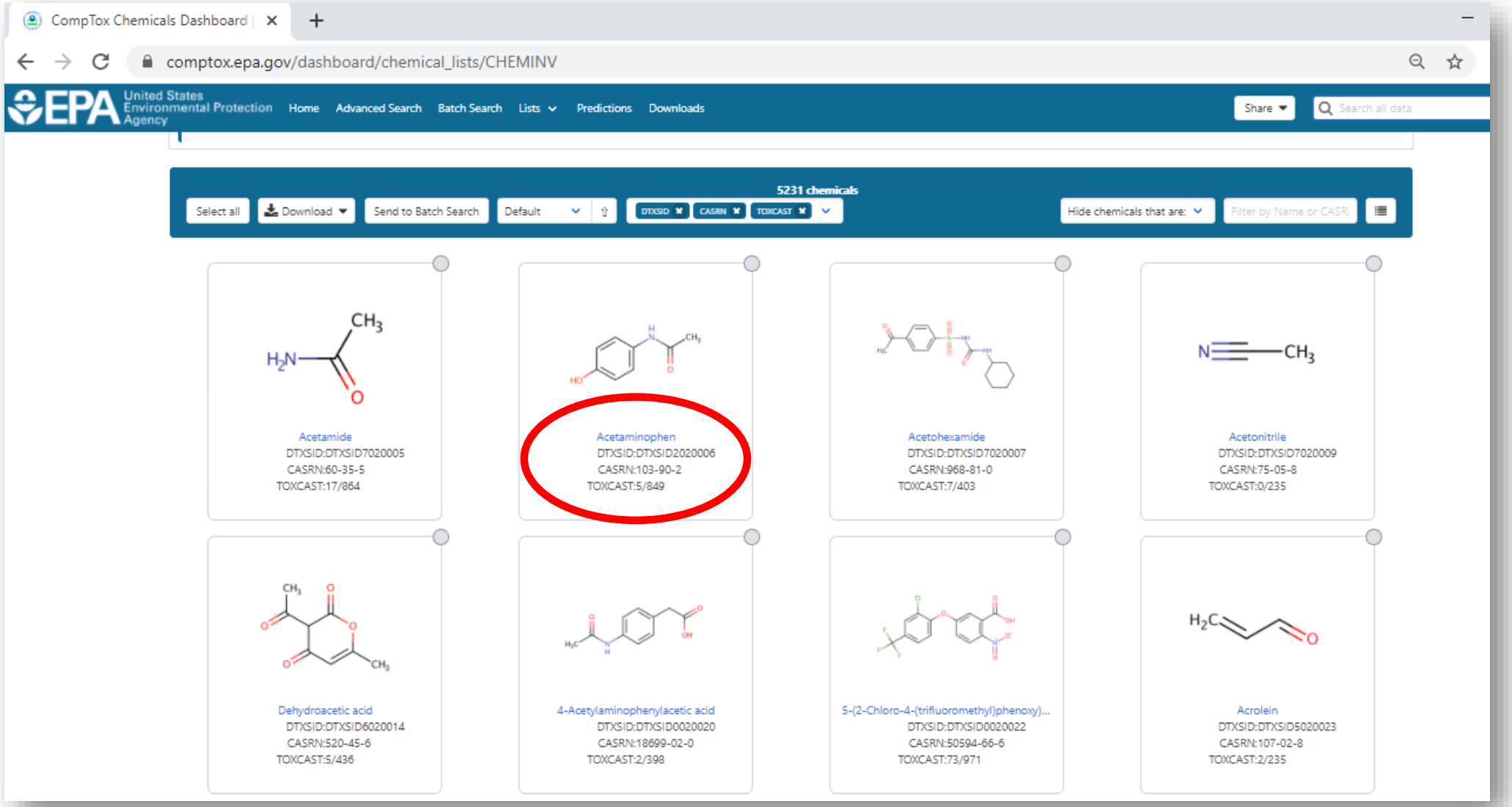


Chemicals in the ToxCast Physical Library

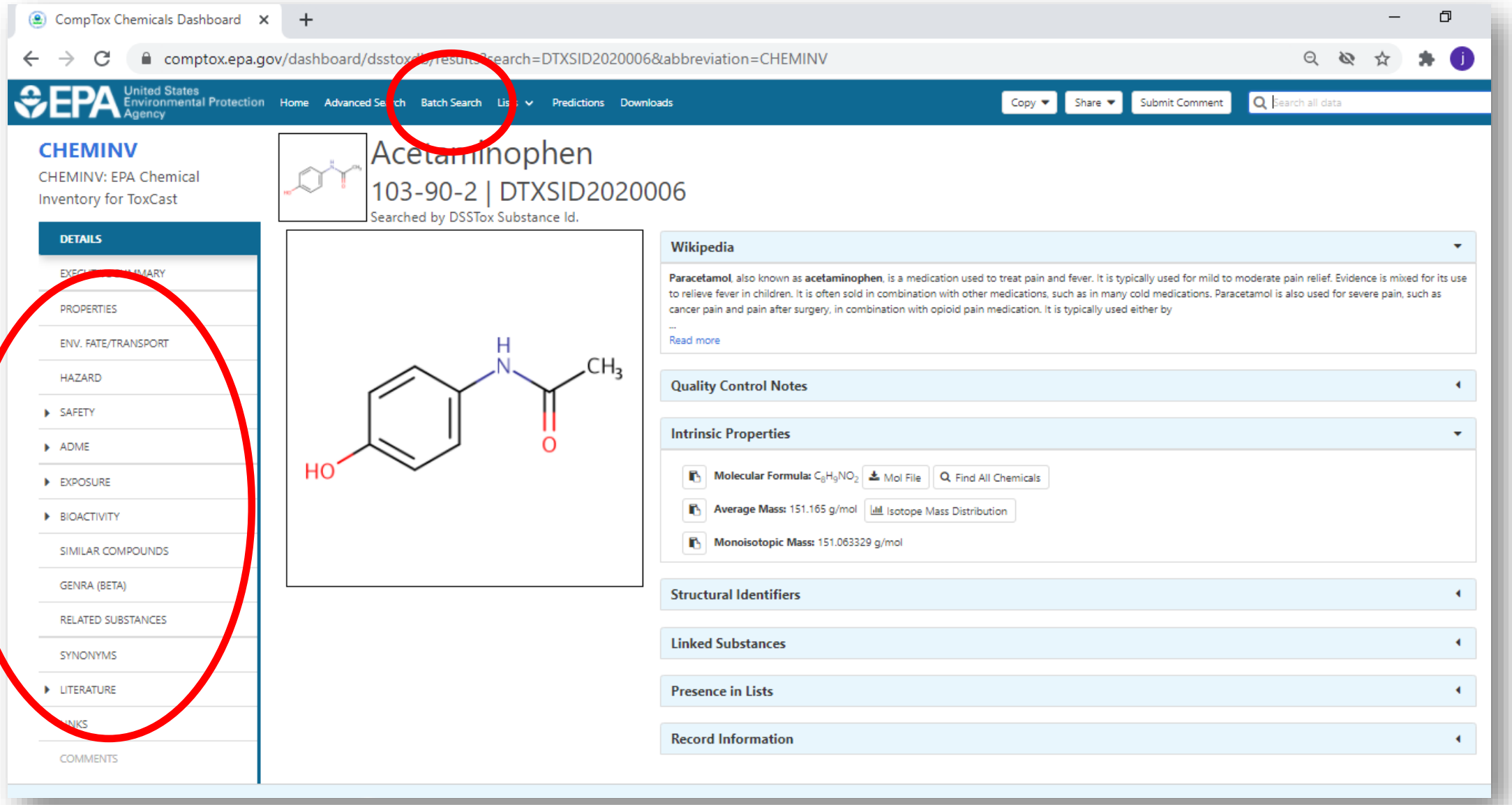


List Acronym	List Name	Last Updated	Number of Chemicals	List Description
TOXCAST_V3	TOXCAST: EPA ToxCast Screening Assay In Vitro DB Version 3	2018-10-05	9403	InvitroDB is the list of chemicals with corresponding assay data in EPA's ToxCast Database (V3 public release, October 2018)
EPACHEMINV_AVAIL	CHEMINV: ToxCast/Tox21 Chemical inventory available as DMSO solutions (20181123)	2018-11-21	6408	EPACHEMINV_AVAIL is list of unique DSSTox substances available as DMSO solutions for ToxCast and Tox21 partner projects, managed by EPA Chemical Contract Services.
CHEMINV	CHEMINV: EPA Chemical Inventory for ToxCast	2017-02-23	5231	CHEMINV is full list of unique DSSTox substances mapped to historical chemical inventory of physical samples registered by EPA's ToxCast Chemical Contractor (Evotec) since launch of ToxCast program in 2007.
TOXCAST	TOXCAST: EPA ToxCast Screening Library	2017-04-11	4746	TOXCAST is the complete list of chemicals having undergone some level of screening in EPA's ToxCast research program since 2007 (last updated 4/11/2017); sublists included.
TOXCAST_PHASEIII	TOXCAST_PhaseIII - EPA ToxCast Screening Library (Phase II Subset)	2017-04-11	4584	TOXCAST_PhaseIII is the full set of chemicals available for screening in Phase III of the ToxCast program, consisting of the majority of chemicals screened in Phase II and newly added ph3 chemicals.
TOXCAST_PH3	TOXCAST_ph3 - EPA ToxCast Screening Library (ph3 subset)	2018-04-11	2678	TOXCAST_ph3 is the ph3 subset of TOXCAST, added to the most recent Phase III of the ToxCast program to further increase chemical diversity and coverage of chemicals of concern to EPA programs.
TOXCAST_PHASEII	TOXCASST_PhaseII - EPA ToxCast Screening Library (Phase II Subset)	2016-01-29	1864	TOXCAST_PhaseII is the full set of chemicals screened in Phase II of the ToxCast program, consisting of TOXCAST_ph1v2, ph2 and e1k sublists.
TOXCAST_E1K	TOXCAST_e1k - EPA ToxCast Screening Library (e1k Subset)	2016-01-25	799	TOXCAST_e1k is the e1k subset of TOXCAST, selected for screening in endocrine-related assays.

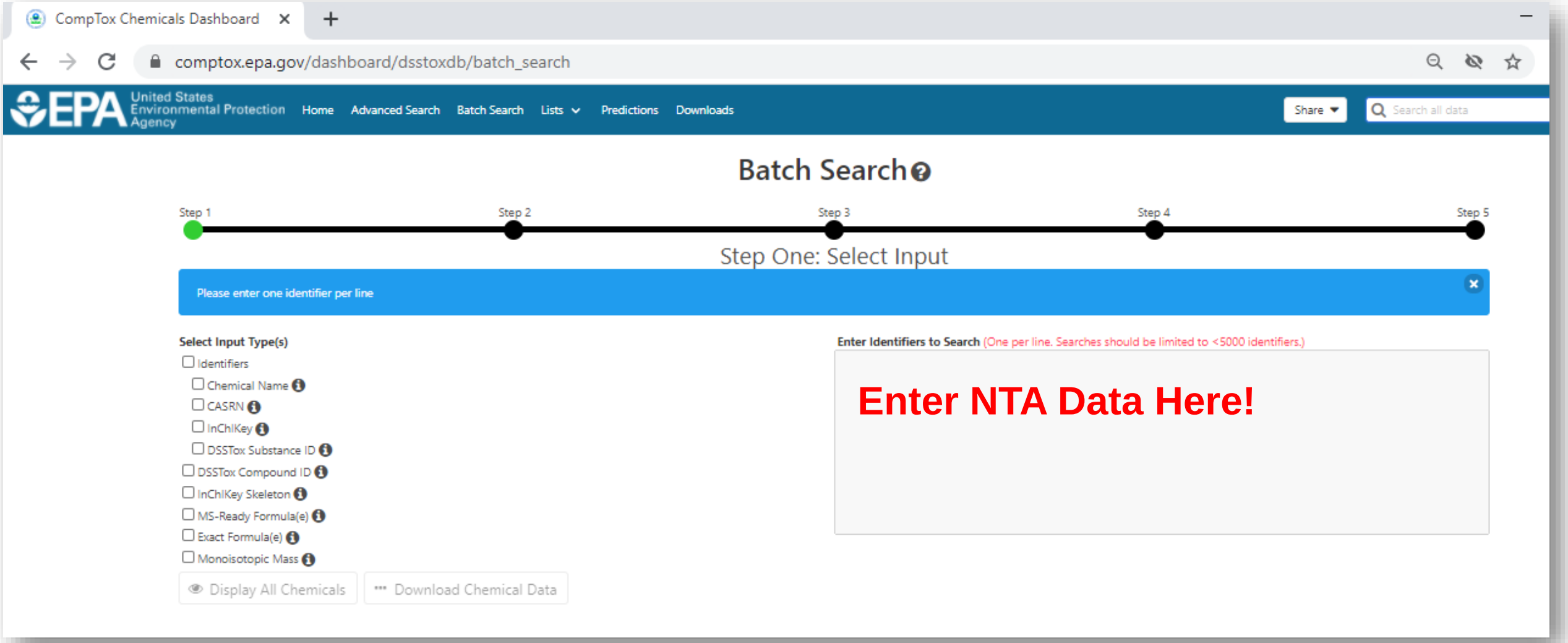
Accessing Specific Chemicals



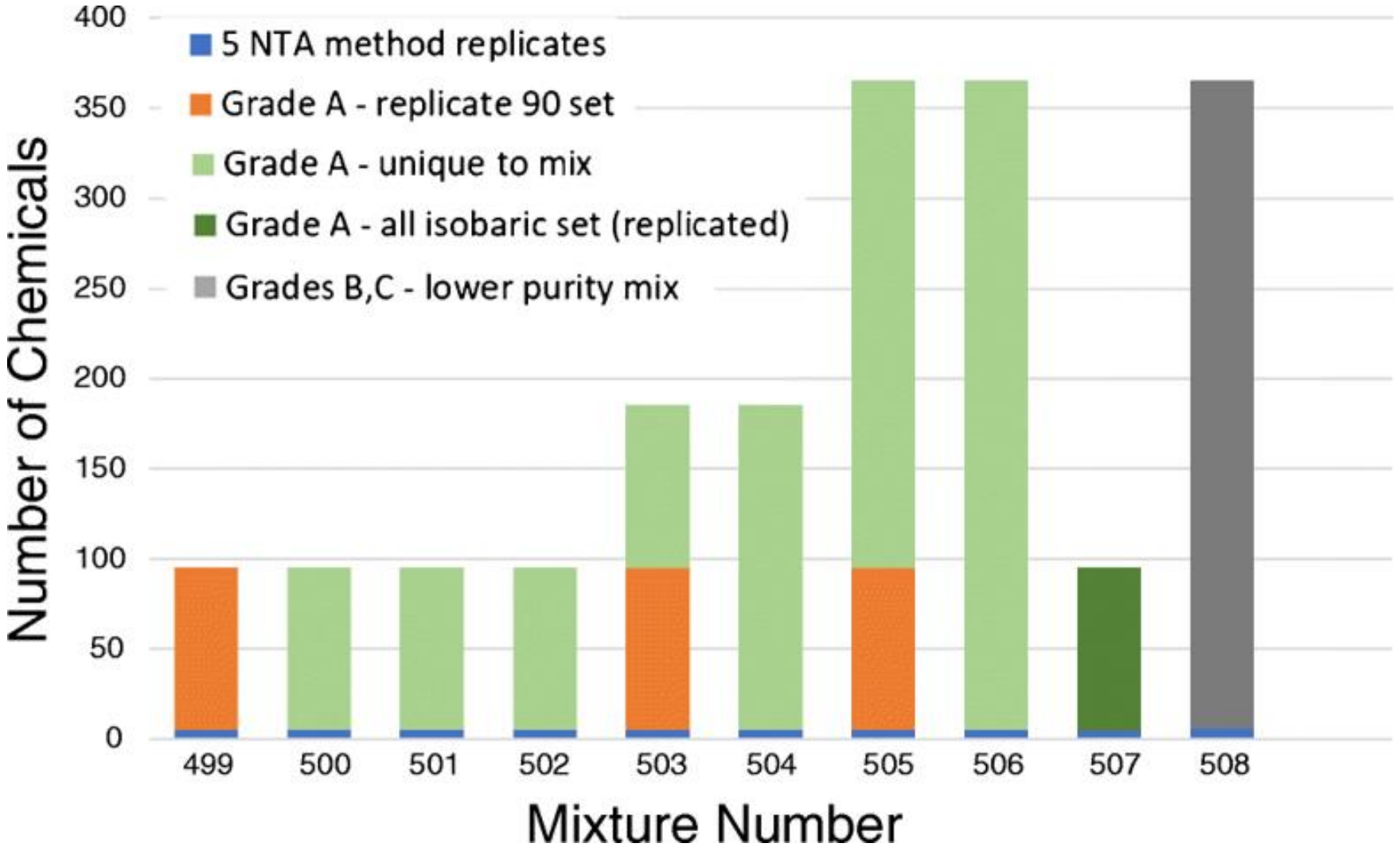
Accessing Chemical-Specific Info



Accessing Chemical Info via Batch Search



Design of ENTACT Mixtures

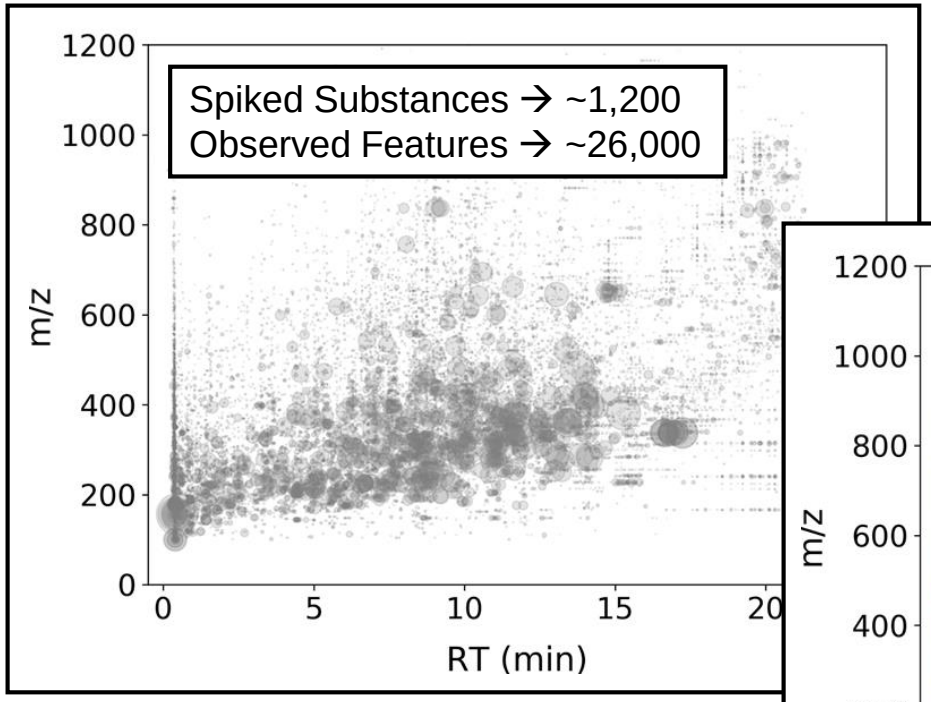


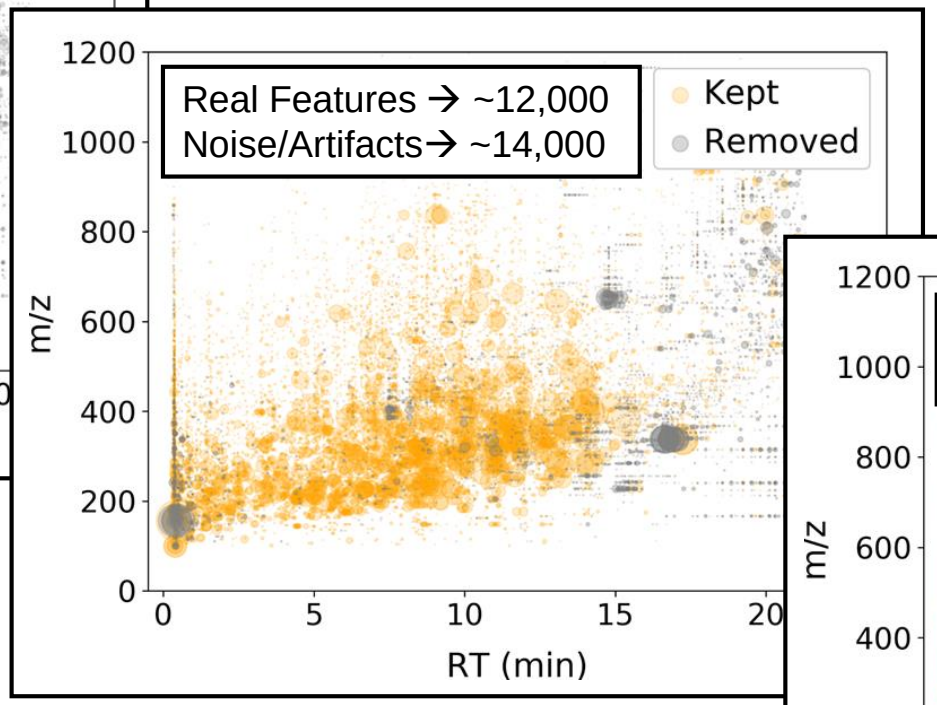
Replication in substance spikes offers a unique means to assess NTA method reproducibility!

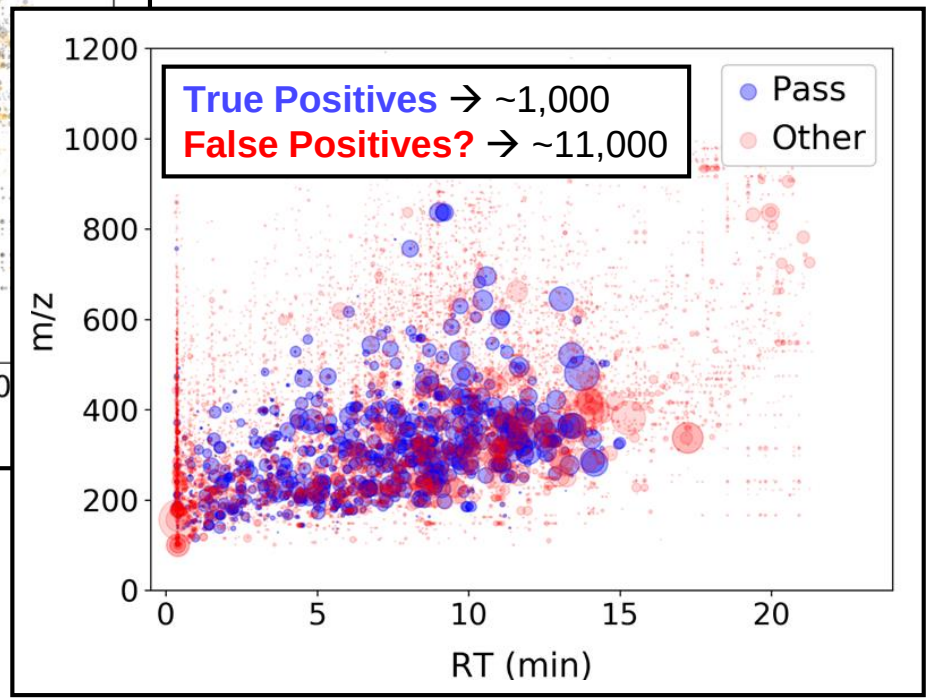
Ulrich et al. 2019. doi: 10.1007/s00216-018-1435-6

EPA Lab Results for ENTACT Mixtures

LC-QTOF HRMS (ESI+ and ESI-)







Substance Identified?	Substance Spiked? Yes	Substance Spiked? No
Yes	True Positives (≤ 65%)	False Positives?
No	False Negatives (≥ 35%)	True Negatives?

Sobus et al. 2019. doi: 10.1007/s00216-018-1526-4

Who Else is Working on ENTACT?

Contractors:

University of Alberta
Emory University
Duke University
Oregon State University
San Diego State University
Pacific Northwest National Laboratory

19 Blind submissions

15 Unblinded submissions

Vendors:

The Scripps Research Institute
Agilent Technologies
Thermo Scientific
Leco
AB SCIEX
Waters

General Participants:

Université du Luxembourg
University of Florida
Wisconsin State Laboratory of Hygiene, University of Wisconsin-Madison
NIST
NC State University
Recetox
Cornell University
UW Tacoma
Mount Sinai
University of Birmingham
DTSC Department of Toxic Substances Control
eawag aquatic research
United States Environmental Protection Agency
FIU

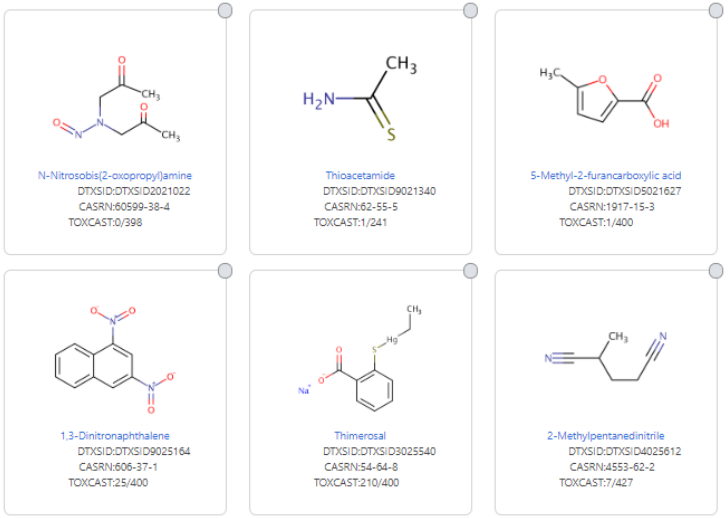
Processing ENTACT Data Submissions

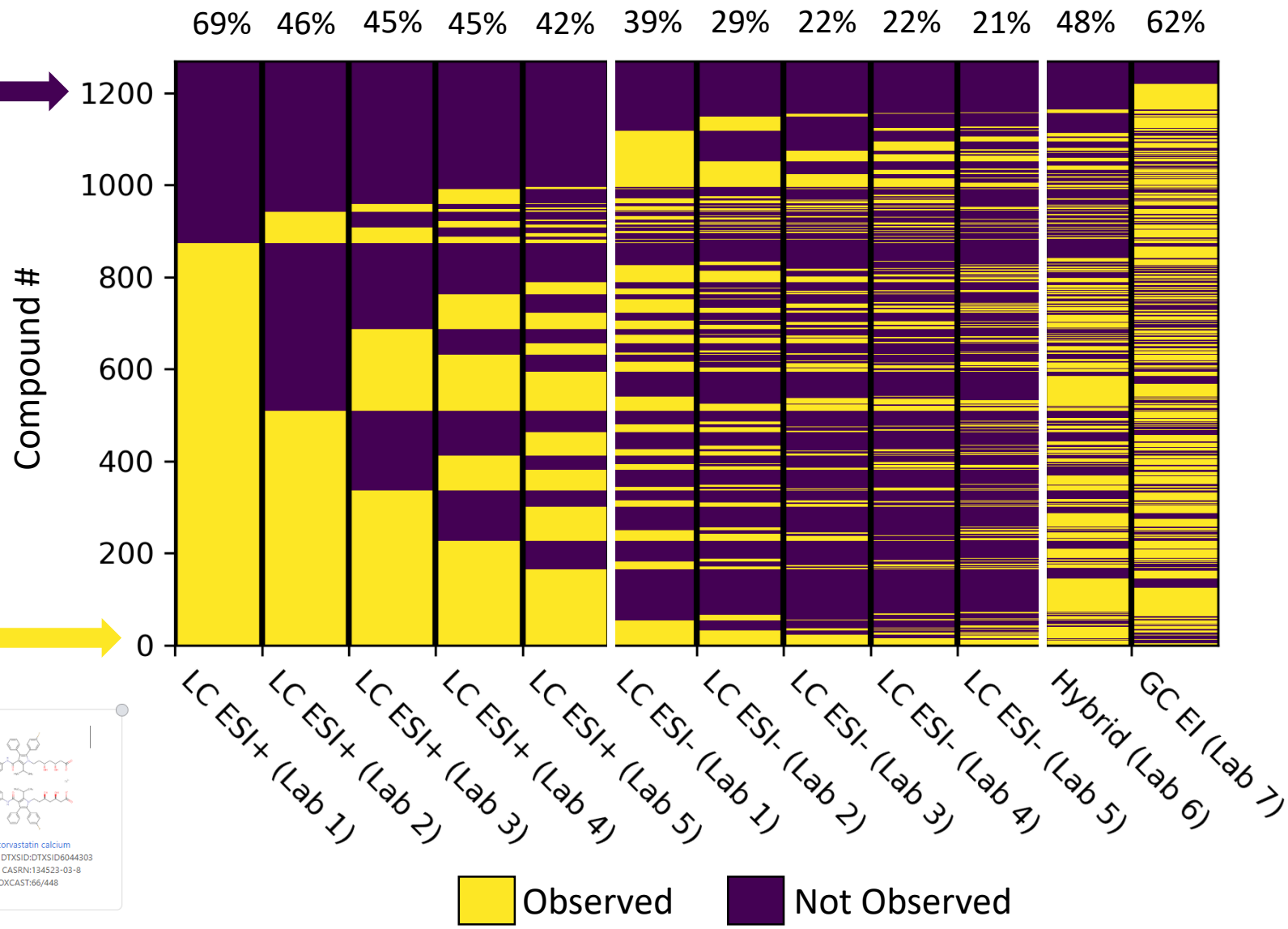
- Individual methods treated separately (if appropriate)
- One candidate mass/formula/compound per feature
- Confidence level revised as needed (with consensus)
- Matching to spiked substances by mass, formula & structure
- "**Observed**" if structure or formula (no spiked isomers) match
- "**Identified**" if structure match
- "**Reproducible**" if correctly ID'd >50% of the time
  - For compounds spiked >1 time and identified ≥1 time

Method Comparison: “Observed” Compounds

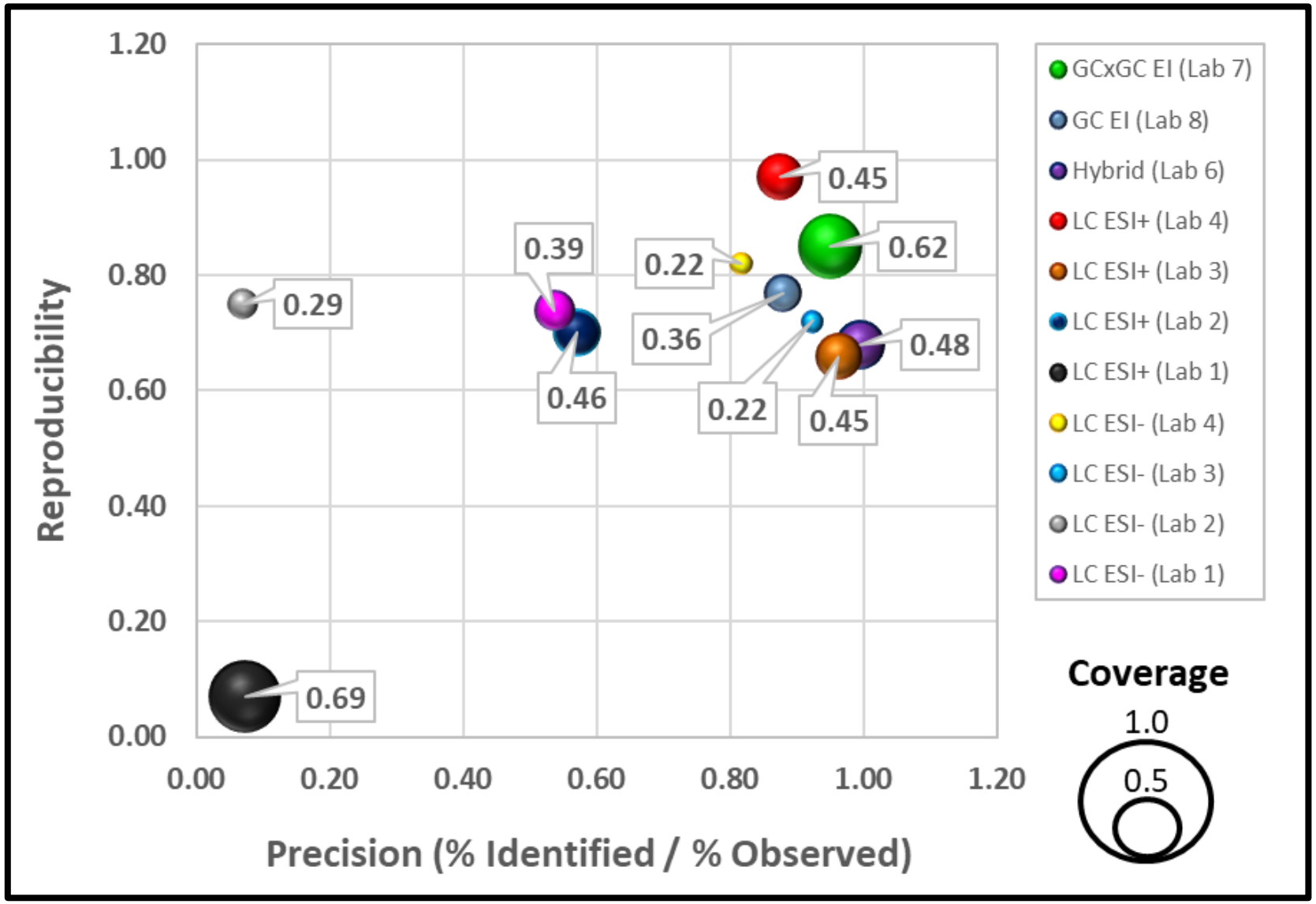
7 Labs, 12 Methods

~5% Not Observed by Any Method





Method Comparison: Total Performance



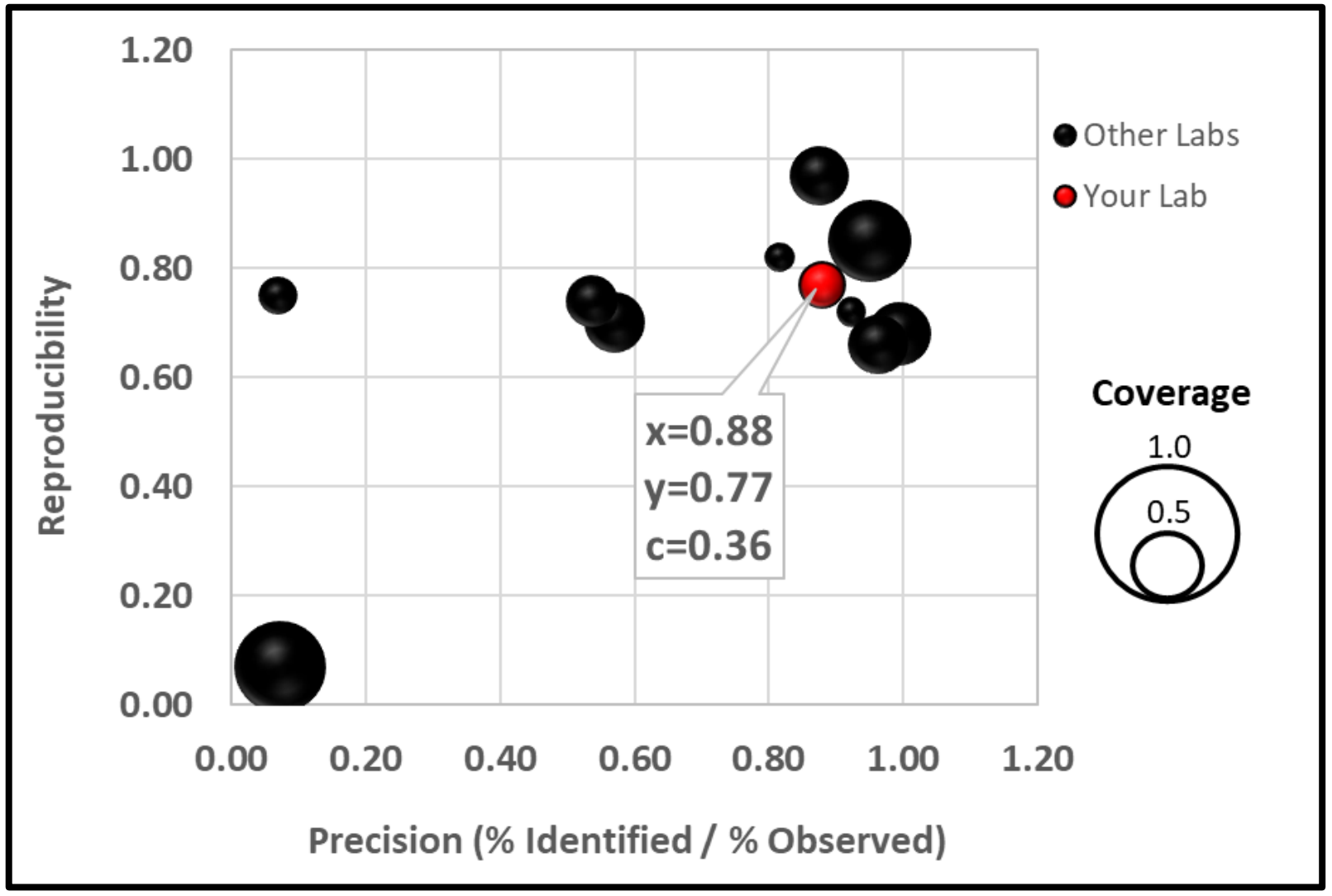
Metrics (all %):

X-Axis →
How often correct?

Y-Axis →
How consistent?

Bubble Size →
How much coverage?

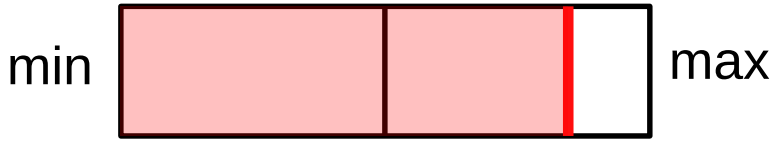
Example Performance Report



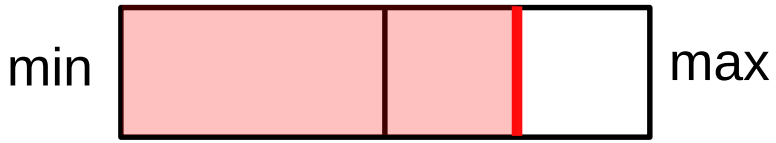
Performance Scores:

(% of max score)

Precision: **88%**



Reproducibility: **78%**



Coverage: **30%**

Additional Results for Collaborators

- Simple performance summary file (n=1 per method):
  - # and % correct identifications per sample
- Individual results files (n=10 per method):
  - Mass match (yes/no), formula match (yes/no), compound match (yes/no)
  - Highest confidence level (as reported or after consensus revision)
- Composite results file (n=1 per method):
  - For each spiked substance (n=1,269)
    - # of spikes (1-10), # of isomer spikes (1-5)
    - # mass hits, # formula hits, # compound hits
    - Observed (yes/no/undetermined), Correct ID (yes/no), Reproducible (yes/no)

Some Challenges (to date)

- Multiple chemical candidate submissions per feature
- Inconsistent & inaccurate use of scoring metrics
- Inconsistent & inaccurate chemical naming procedures
- Inconsistent and unclear feature filtering protocols
- Limited engagement regarding collaborator follow-up
- Determining false positives vs. unanticipated true positives
- Determining true negatives and dependent metrics
- Slow evaluation process vs. rapid method development processes

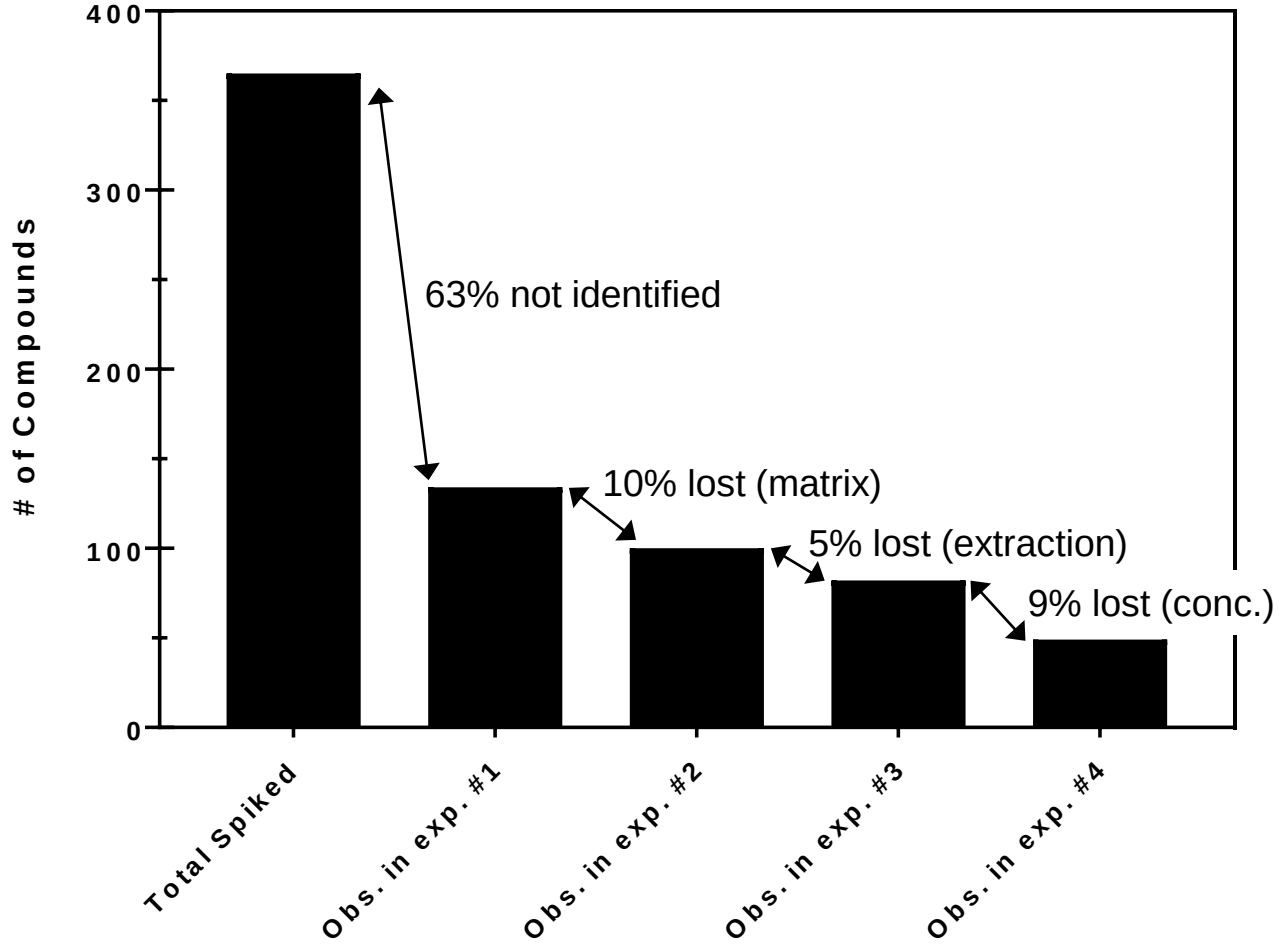
EPA Experiments with SRM Dust

1) : Solvent spike (best case)

2) Extraction : Post-extraction high spike

3) Extraction : Pre-extraction high spike

4) Extraction : Pre-extraction low spike (ENTACT sample)

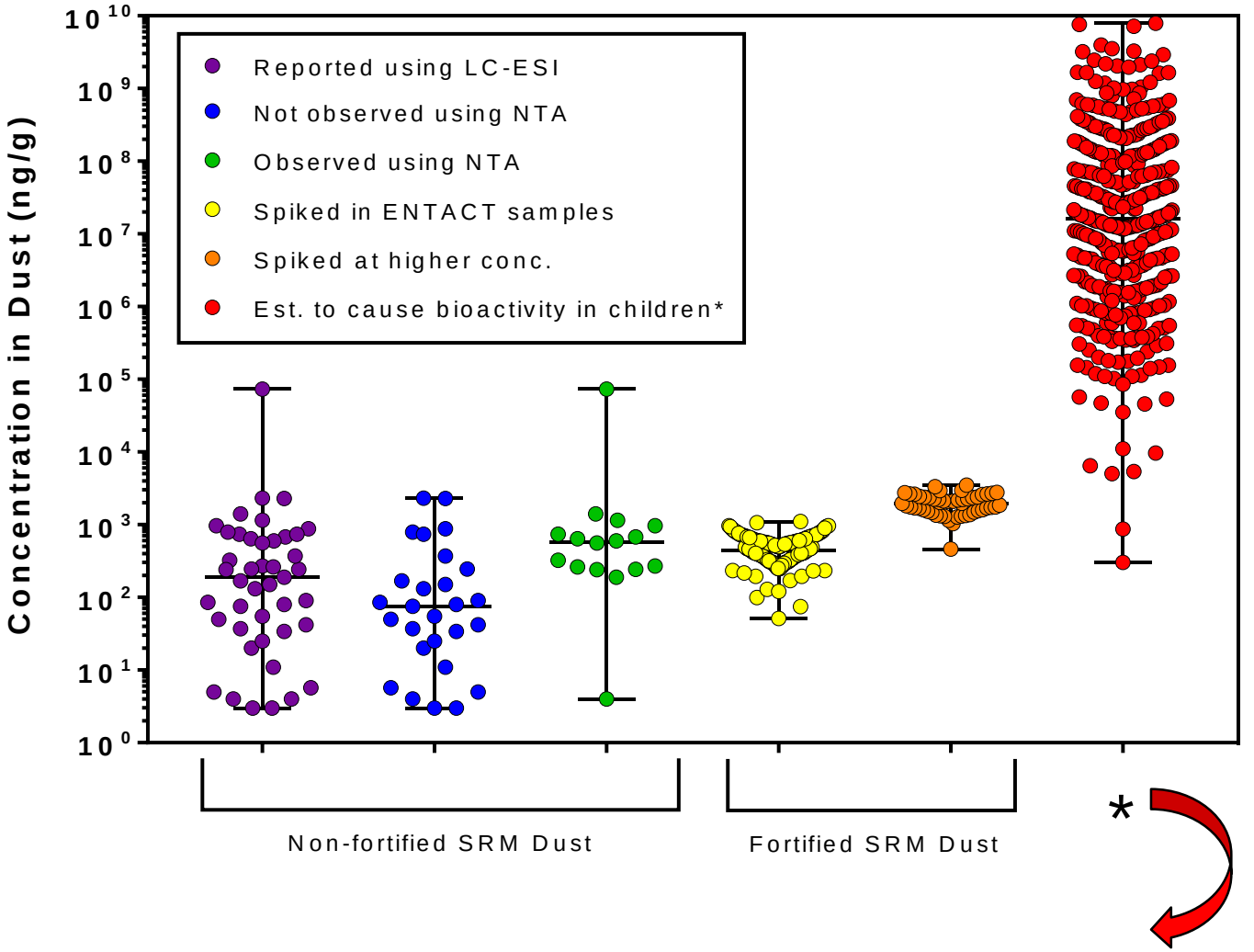


Newton et al. 2020. doi: 10.1007/s00216-020-02658-w

EPA Experiments with SRM Dust

Results for Unfortified SRM Dust

Chemical Class	All Reported Compounds	Reported Using LC-ESI	Observed Using NTA
PAHs	69	0	0
PCBs	44	0	0
PFAS	31	31	12
BFRs	30	3	0
OCPs	15	0	0
OPEs	12	9	4
Phthalates	7	0	2
Total	208	**43**	**18**



***** the concentration that would be needed (for the most-sensitive 5% of the population) to produce a steady-state plasma concentration equal to the 10th percentile of the ToxCast AC50 distribution across assays for the given chemical.

Newton et al. 2020. doi: 10.1007/s00216-020-02658-w

EPA Evaluation of *in silico* Spectra

Metabolomics (2015) 11:98–110
DOI 10.1007/s11306-014-0676-4

ORIGINAL ARTICLE

Competitive fragmentation modeling of ESI-MS/MS spectra for putative metabolite identification

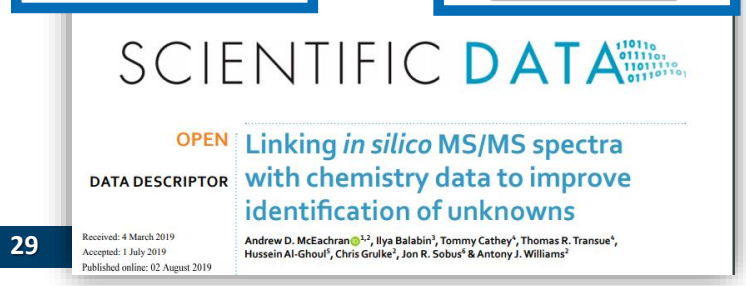
Felicity Allen · Russ Greiner · David Wishart

Training Set

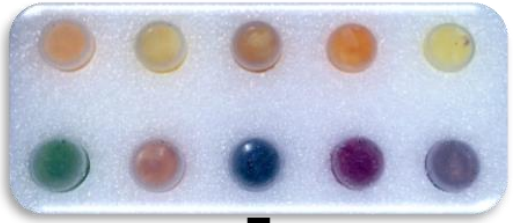
Fragmentation Prediction Model

DSSTox structures

DSSTox MS2 spectra



10 Synthetic Mixtures:
1,269 Unique ToxCast Substances



LC-QTOF HRMS:
Data Dependent Acquisition

MS2 Reference Library

No Library Matches

Probable Structures

MS2 *in silico* Library (~765,000 DSSTox Substances)

Tentative Structures

Analytical and Bioanalytical Chemistry
https://doi.org/10.1007/s00216-019-02351-7

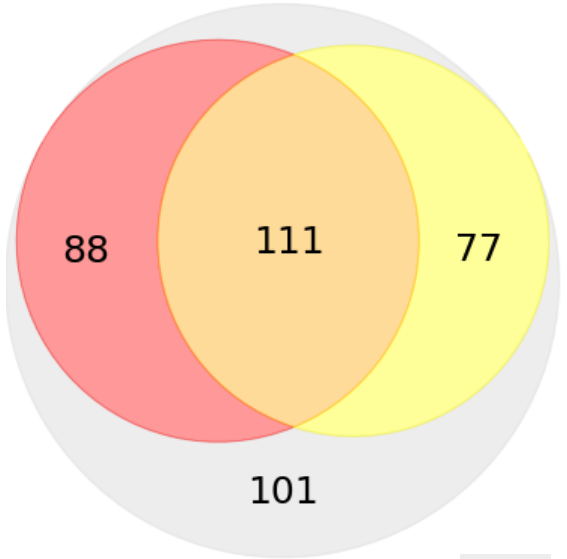
RESEARCH PAPER

In silico MS/MS spectra for identifying unknowns: a critical examination using CFM-ID algorithms and ENTACT mixture samples

Alex Chao[1,2] · Hussein Al-Ghoul[1,2] · Andrew D. McEachran[1,3] · Ilya Balabin[4] · Tom Transue[4] · Tommy Cathey[4] · Jarod N. Grossman[2,3] · Randolph Singh[1,5] · Elin M. Ulrich[2] · Antony J. Williams[6] · Jon R. Sobus[2]

377 ENTACT Compounds with MS2 Spectra

Top Reference Library Match



Top *in silico* Library Match

Not Top Match

Overall Summary

- Regulatory drivers necessitate NAMs for rapid risk characterization
- Measurement data are needed to inform and evaluate NAMs
- Targeted measurement methods can’t keep pace with needs of NAMs
- NTA methods may meet needs, but require development and validation
- EPA/ORD is working to:
  - Develop tools to support NTA studies
  - Apply NTA methods to identify and prioritize chemicals based on anticipated risk
  - Evaluate NTA state-of-the-science via ENTACT

Summary of ENTACT Findings

- NTA methods are suitable for many ToxCast chemicals
  - ~5% of ENTACT compounds not observed by any method
- Multiple methods required for broad characterization
  - No “one size fits all” method
  - <1% of ENTACT compounds observed using all methods
- Performance determined across 3 categories:
  - **Coverage** = Ability to Observe → (Range = 22% to 69%)
  - **Precision** = Ability to Identify those Observed → (Range = 7% to 99%)
  - **Reproducibility** = Ability to Consistently Identify → (Range = 7% to 97%)
- Concentration, media, and extraction techniques will affect performance
- Mixtures/data are highly valuable for NTA method development/evaluation

Contributing Researchers

This work was supported, in part, by ORD's Pathfinder Innovation Program (PIP) and an ORD EMVL award

EPA ORD

Hussein Al-Ghoul*
Alex Chao*
Louis Groff*
Jarod Grossman*
Chris Grulke
Kristin Isaacs
Sarah Laughlin*
Charles Lowe
Kamel Mansouri*
James McCord
Andrew McEachran*
Jeff Minucci
Seth Newton
Katherine Phillips

EPA ORD (cont.)

Tom Purucker
Ann Richard
Randolph Singh*
Mark Strynar
Elin Ulrich
John Wambaugh
Antony Williams

GDIT

Ilya Balabin
Tom Transue
Tommy Cathey

* = ORISE/ORAU

Questions?

sobus.jon@epa.gov

The views expressed in this presentation are those of the author and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.