

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

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

<sup>1</sup> Center for Computational Toxicology and Exposure

<sup>2</sup> ORAU/ORISE Participant

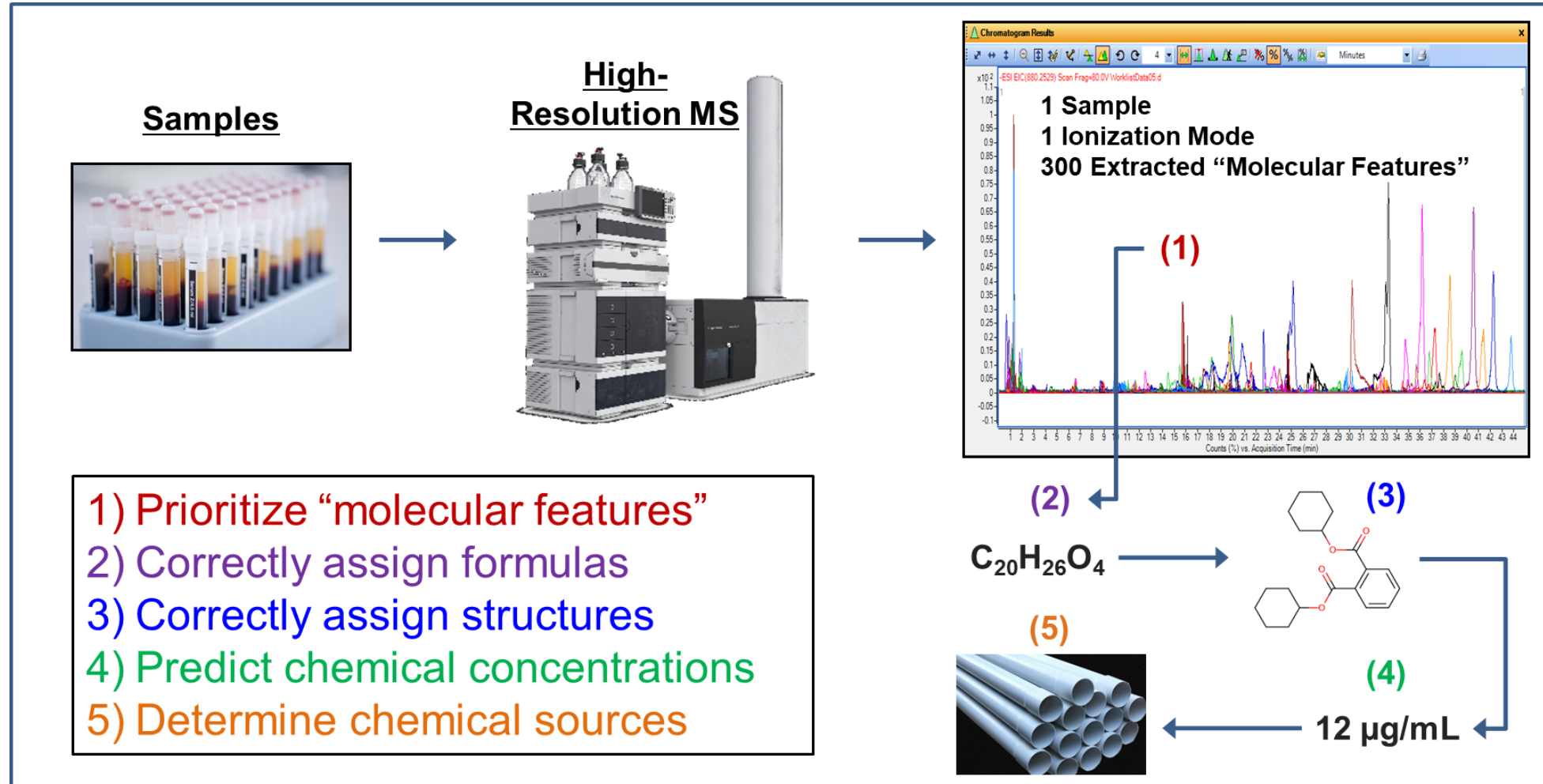
# What's So Great About NTA?

Rapidly screen  
for “knowns”

Discover  
“unknowns”

Uncover historical  
exposures

Generate source  
fingerprints...



# NTA State-of-the-Science

## Environmental Science & Technology

Viewpoint

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

[pubs.acs.org/est](https://pubs.acs.org/est)

### Is Nontargeted Screening Reproducible?

Ronald A. Hites<sup>\*,</sup>

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

Karl J. Jobst<sup>\*</sup>

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



Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: [www.elsevier.com/locate/scitotenv](http://www.elsevier.com/locate/scitotenv)



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

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

<sup>a</sup> Colorado School of Mines, Golden, CO, USA

<sup>b</sup> St. Cloud State University, St. Cloud, MN, USA

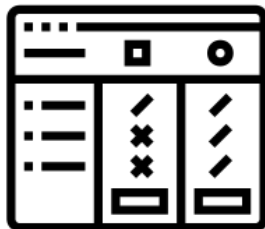
<sup>c</sup> Sciex, Boston, MA, USA

<sup>d</sup> Sciex, Toronto, Canada

<sup>e</sup> 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.”*

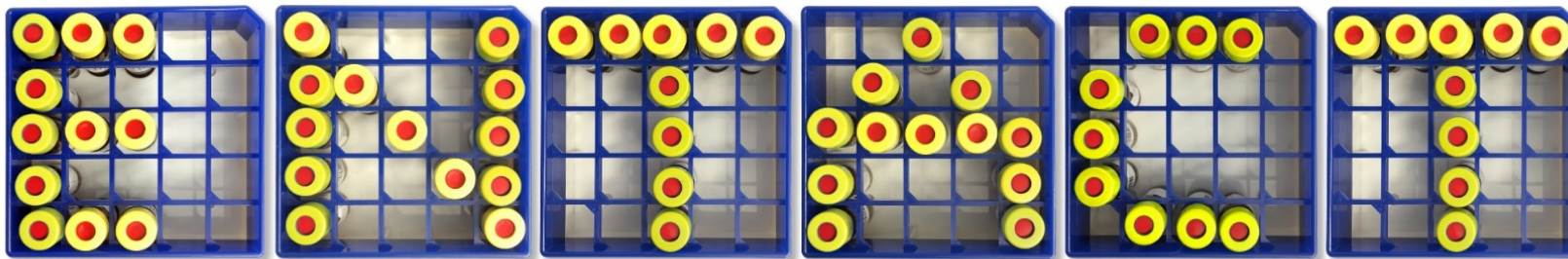
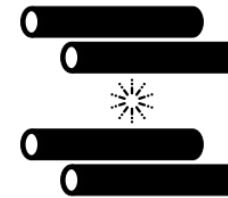
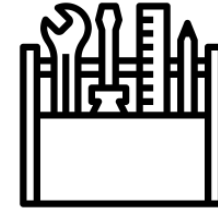


*“The novelty of nontarget analysis, particularly its current lack of implementation by regulatory agencies, has prevented the establishment of streamlined quality assurance and quality control (QA/QC) procedures.”*



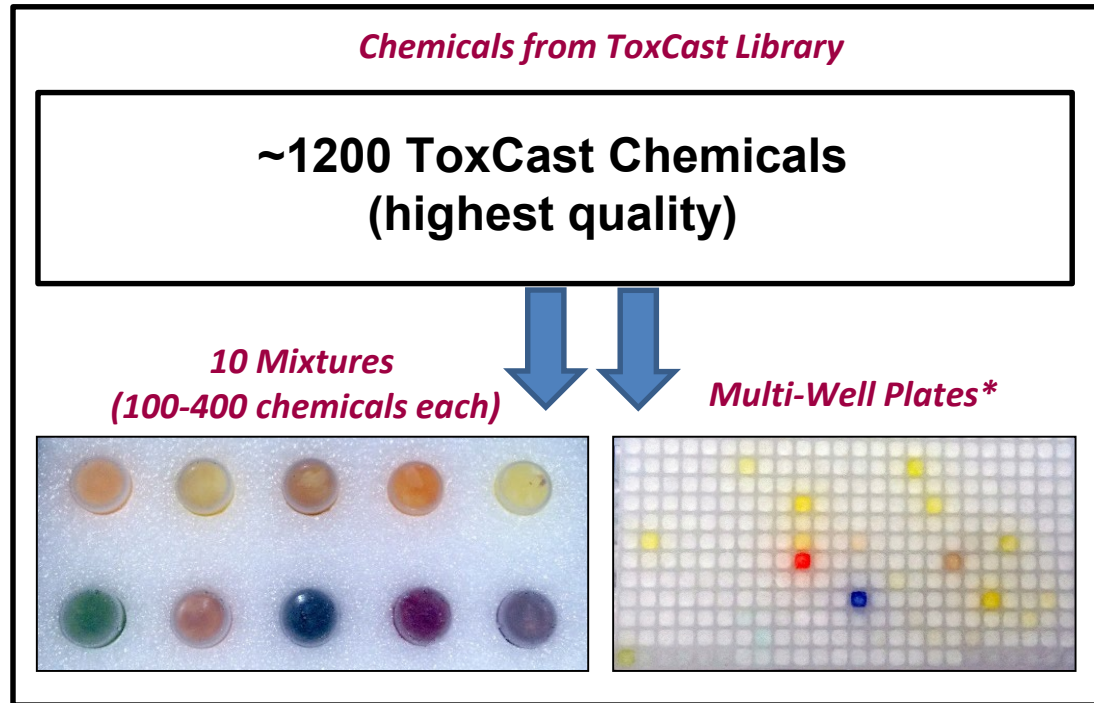
# Science Questions for Research Community

- 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



~25 Collaborators & 5 Contractors\*:

1<sup>st</sup>: Blinded analysis

2<sup>nd</sup>: Unveiling of chemicals

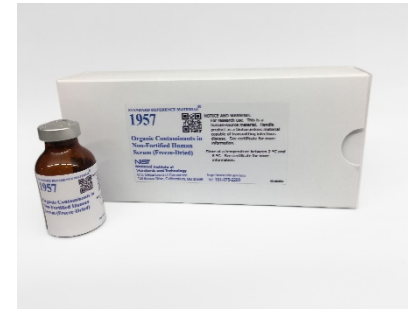
3<sup>rd</sup>: Unblinded evaluation

## ENTACT Part 2

*Reference & Fortified House Dust*



*Reference & Fortified Human Serum*

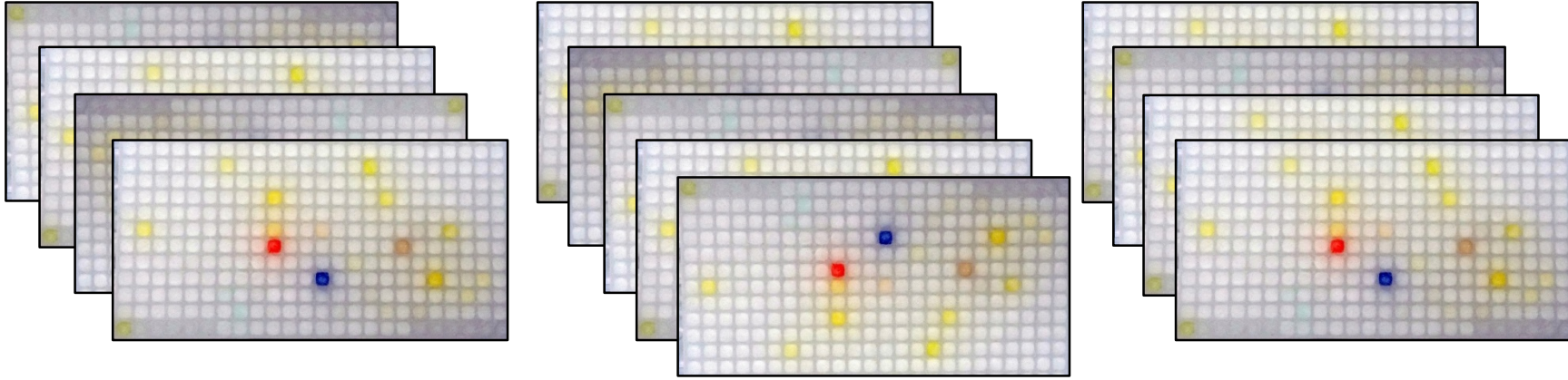


*Reference & Fortified Silicone Wristbands*





# ENTACT Part 3



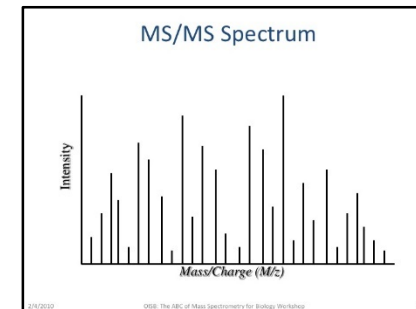
~4600 ToxCast substances



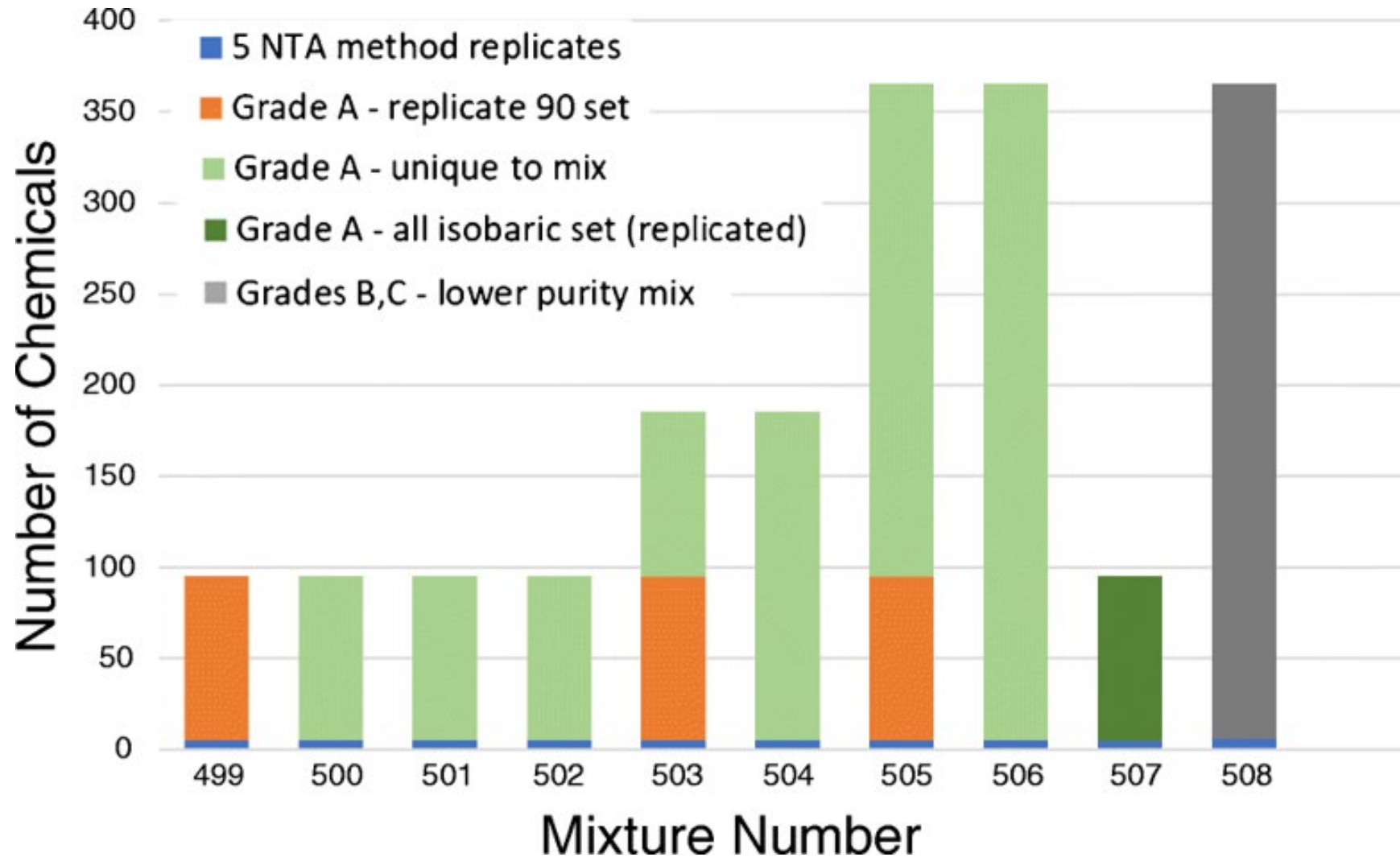
Instrument/software vendors & select labs



Reference libraries for the public



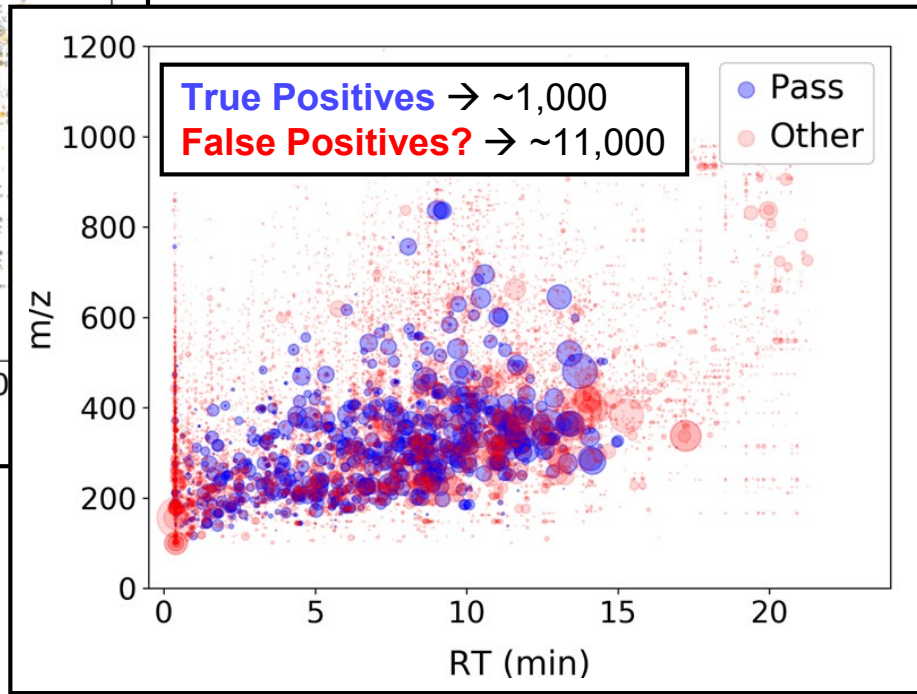
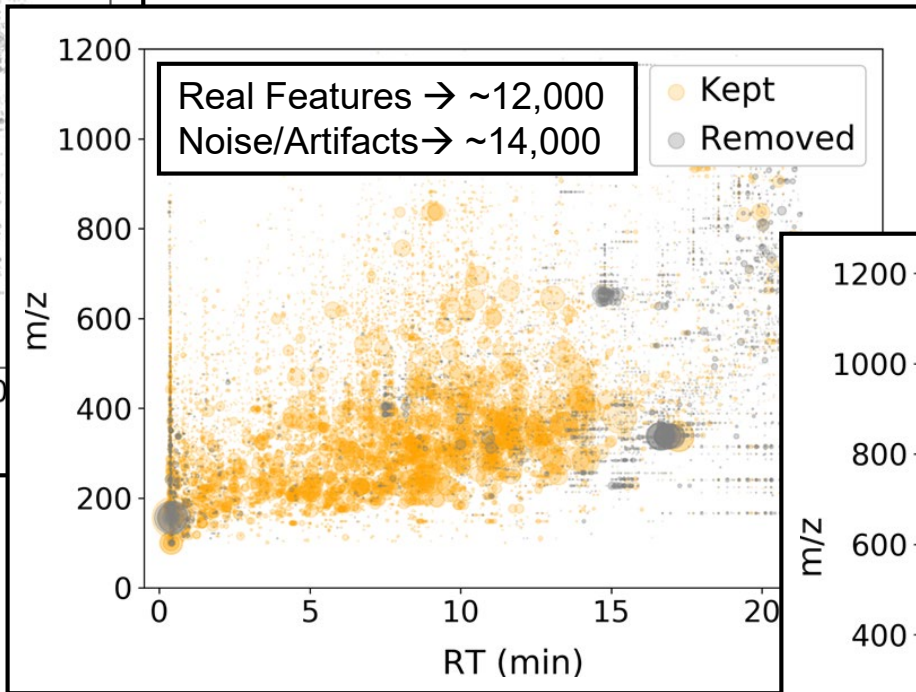
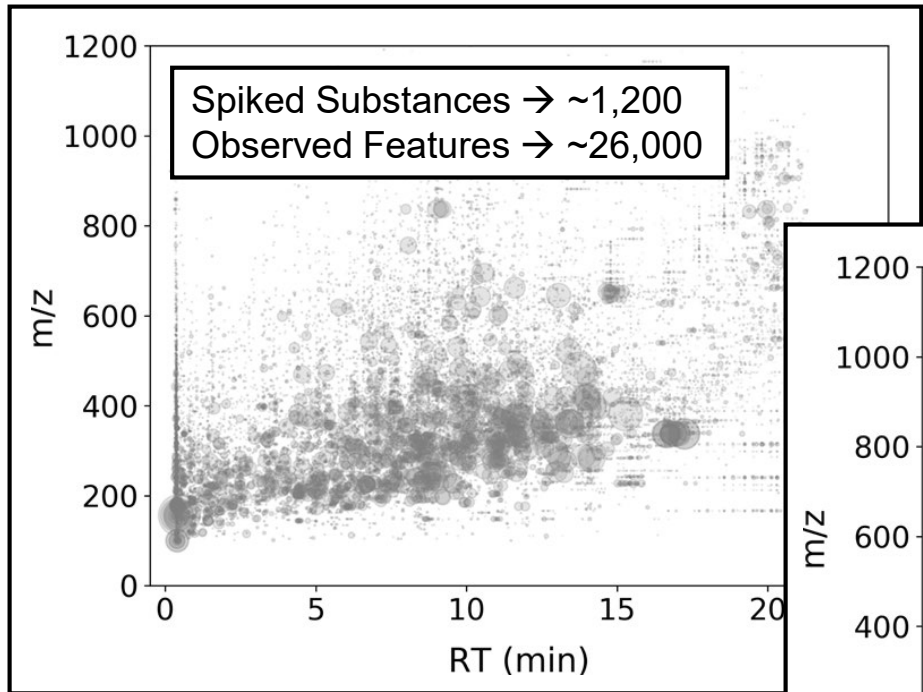
# Design of ENTACT Mixtures



# EPA Lab Results



**LC-QTOF HRMS  
(ESI+ and ESI-)**



## Substance Spiked?

Yes

No

**Substance  
Identified?**

Yes

No

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



# Who Else is Working on ENTACT?

## Contractors:



**19 Blind  
submissions**

**15 Unblinded  
submissions**

## Vendors:



## General Participants:



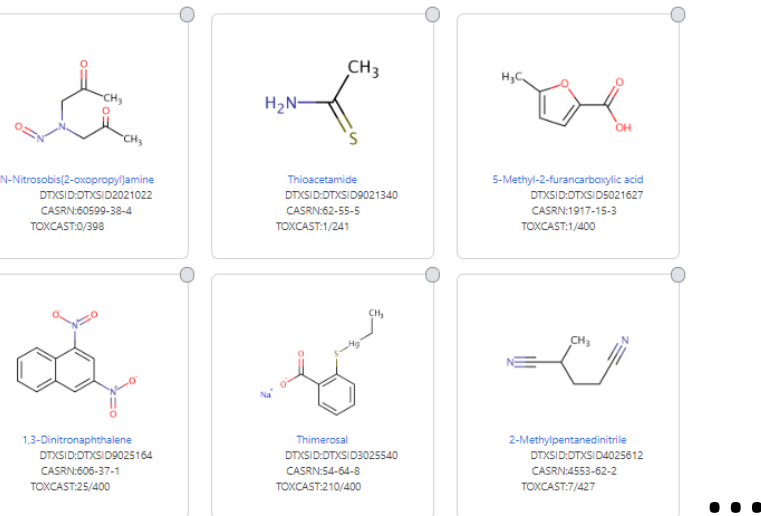
# Processing ENTACT Data Submissions

- 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  $\geq 1$  time

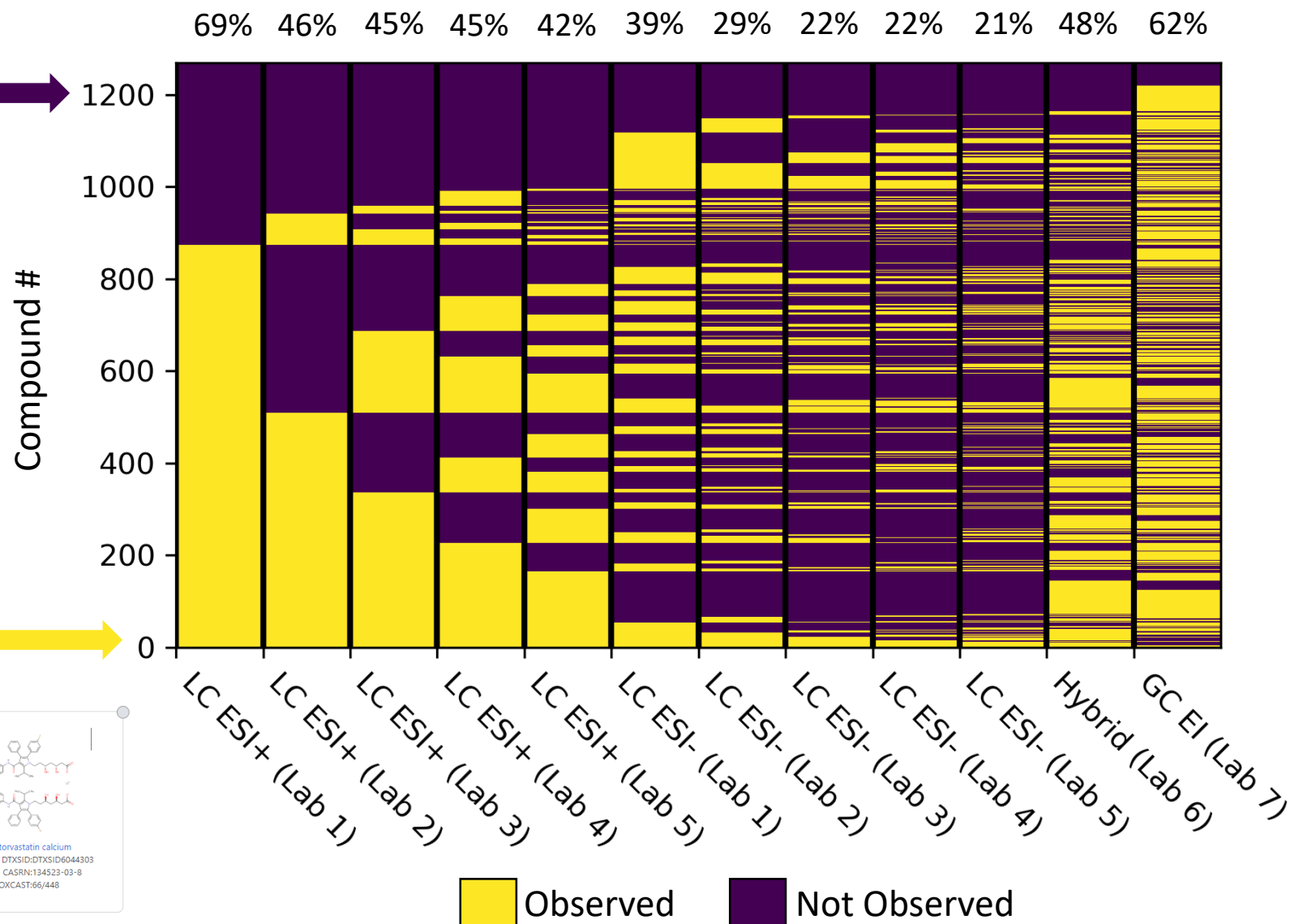
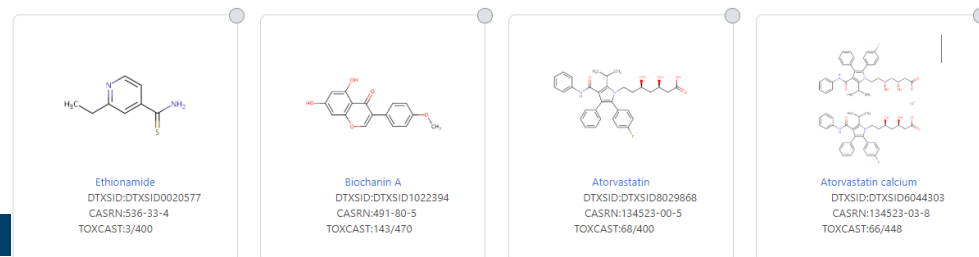
# Lab Comparison: “Observed” Compounds

## 7 Labs, 12 Methods

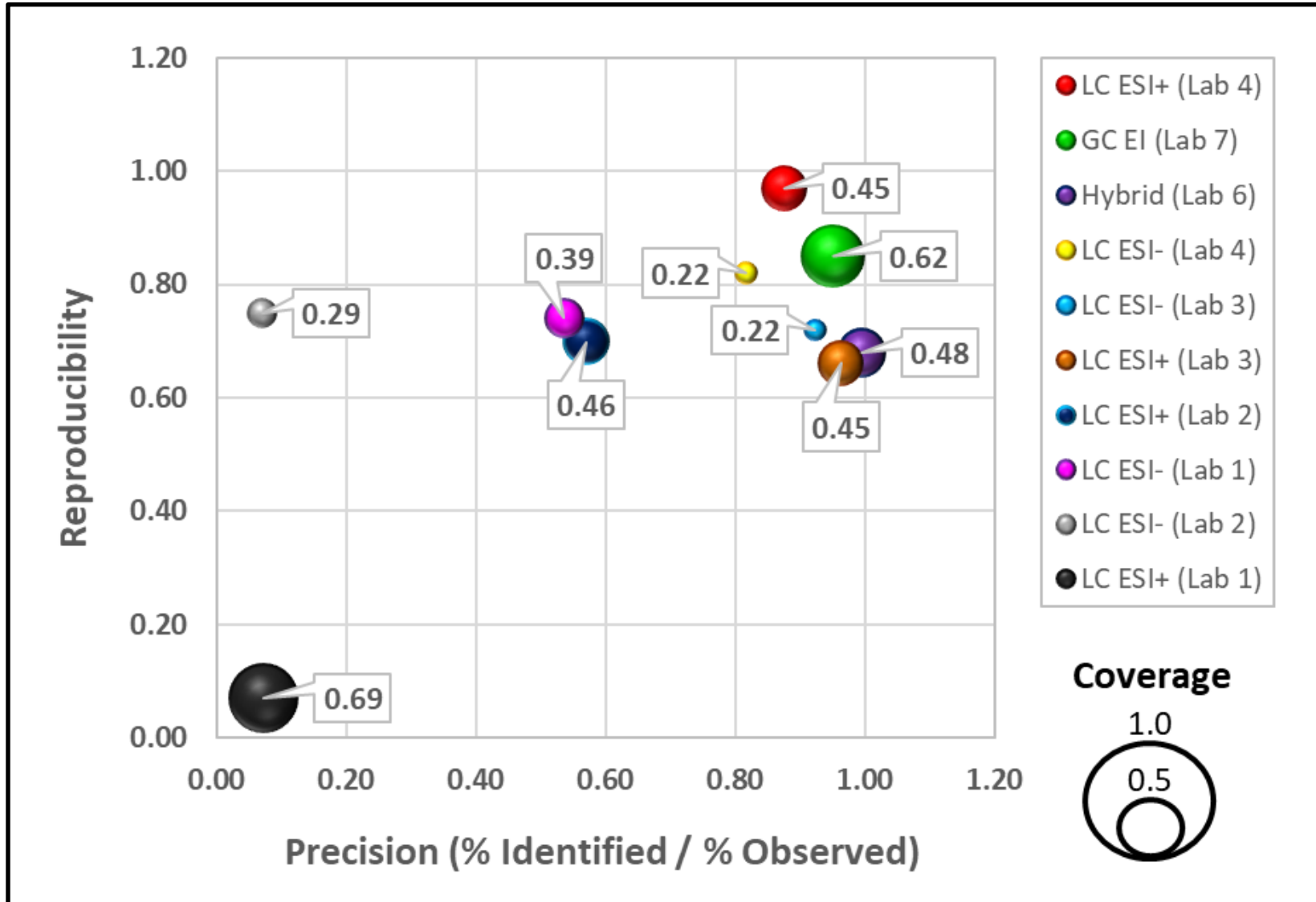
~5% Not Observed by Any Method



<1% Observed by All 12 Methods



# Lab Comparison: Total Performance



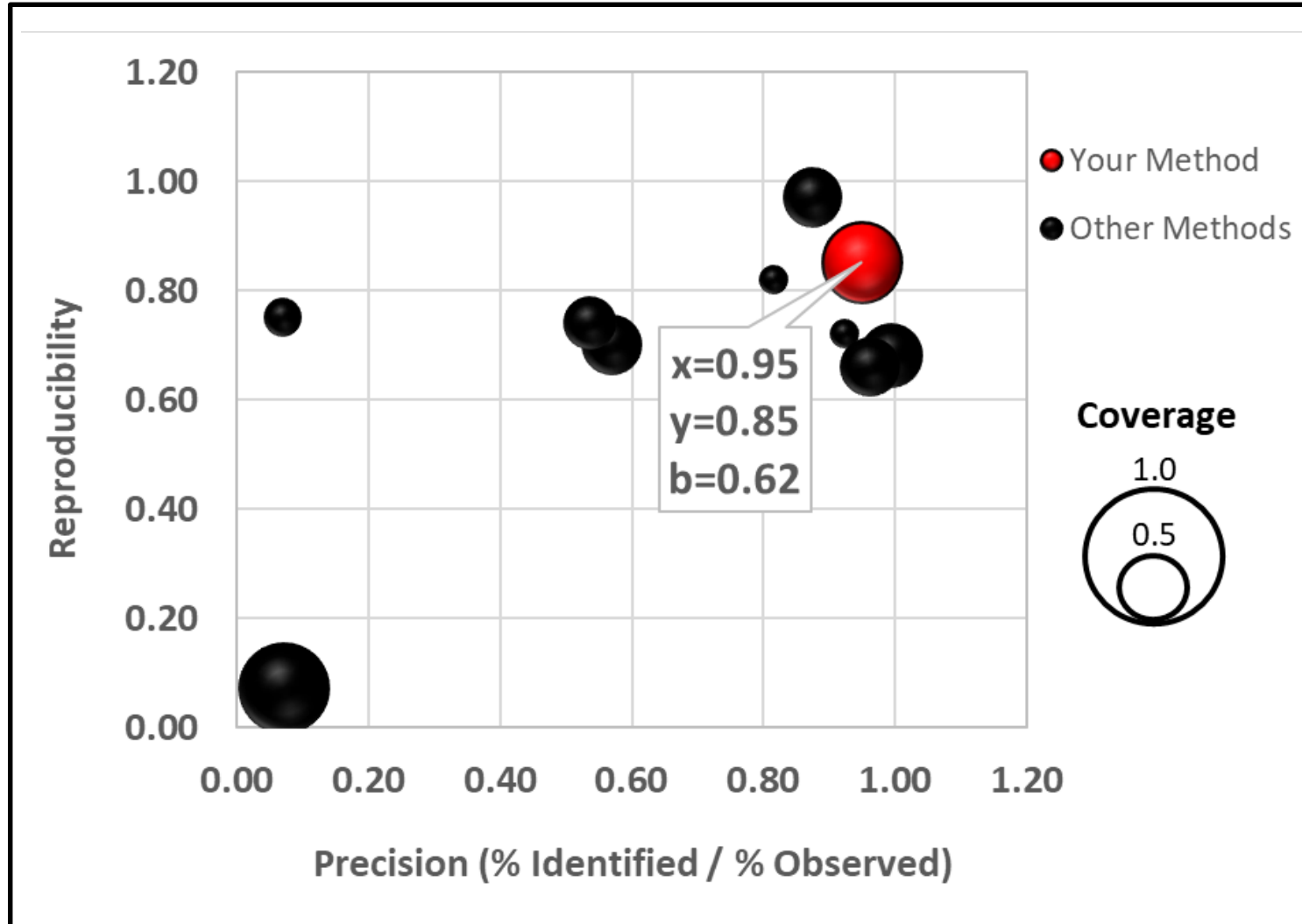
**Metrics (all %):**

X-Axis →  
How often correct?

Y-Axis →  
How consistent?

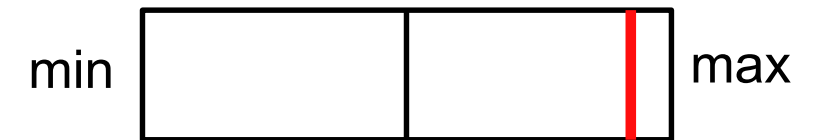
Bubble Size →  
How much coverage?

# Example Performance Report

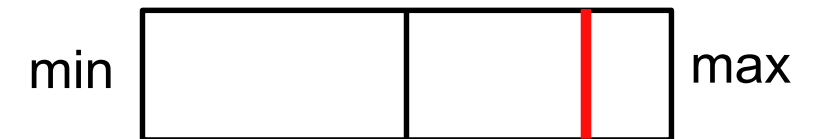


## Performance Scores: (% of max score)

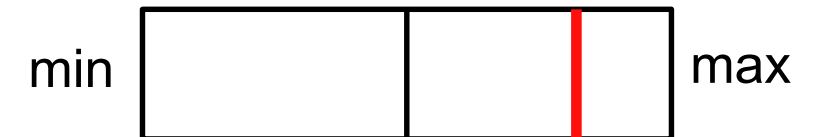
Precision: **95%**



Reproducibility: **87%**

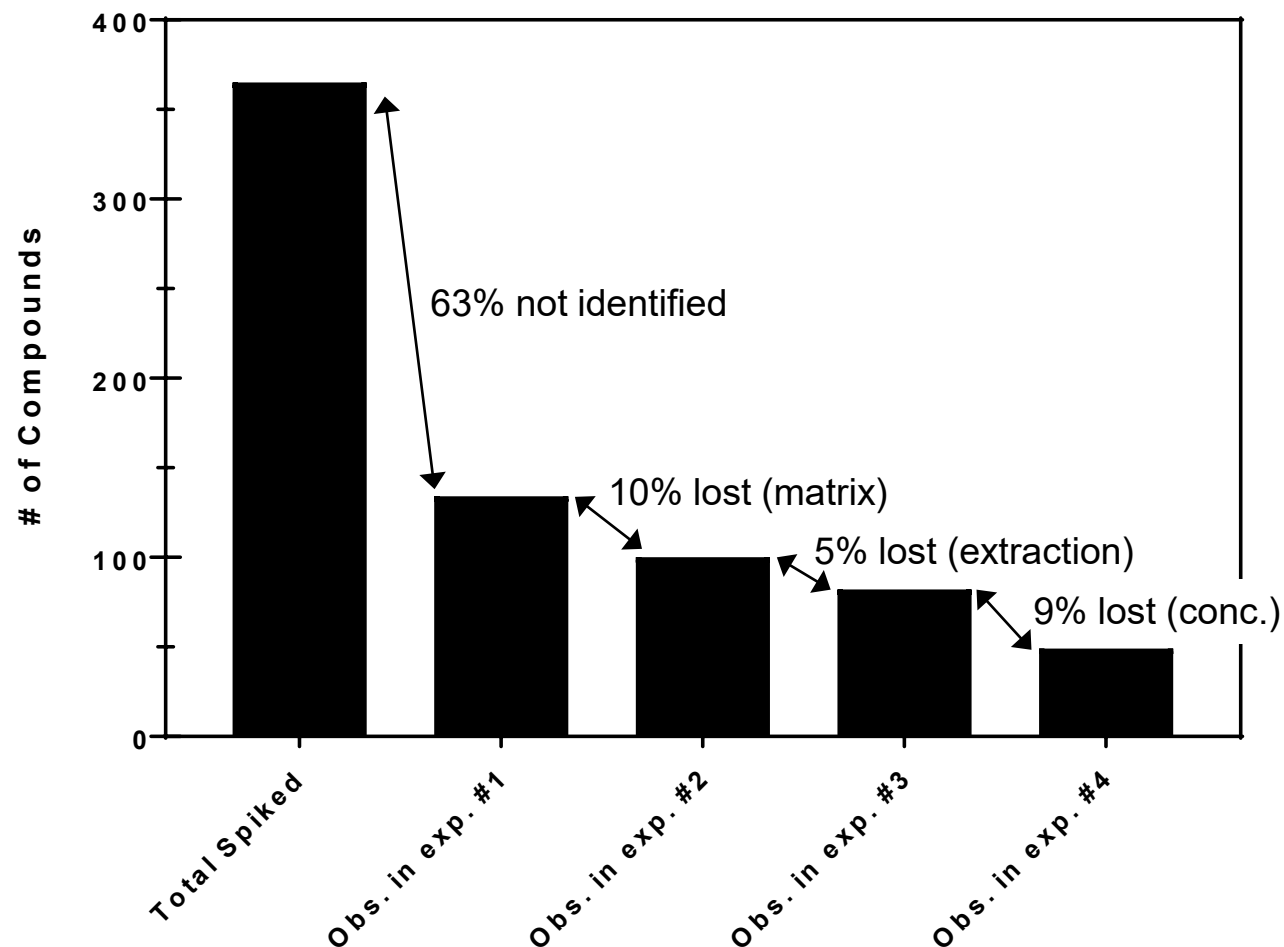
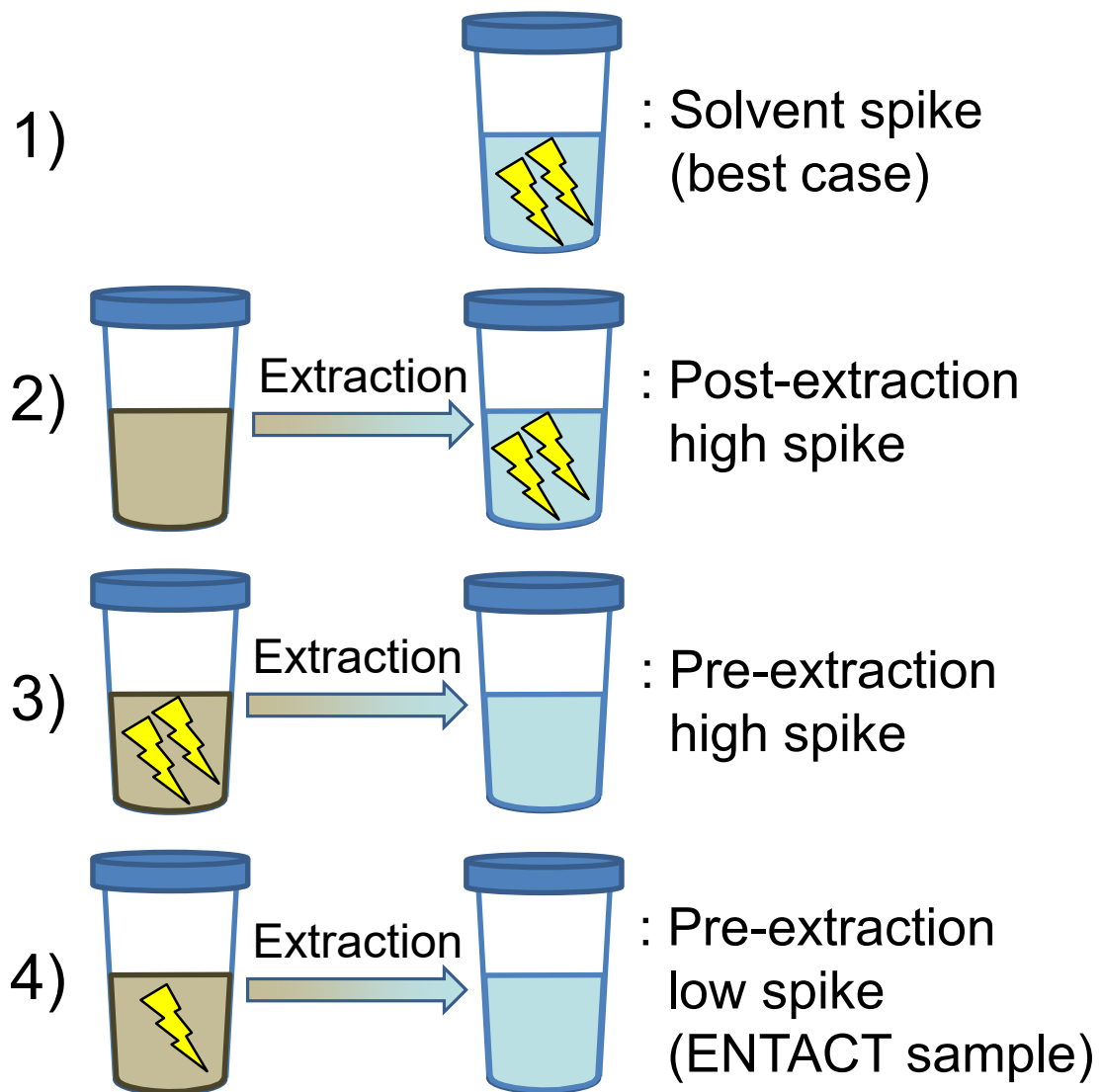


Coverage: **86%**





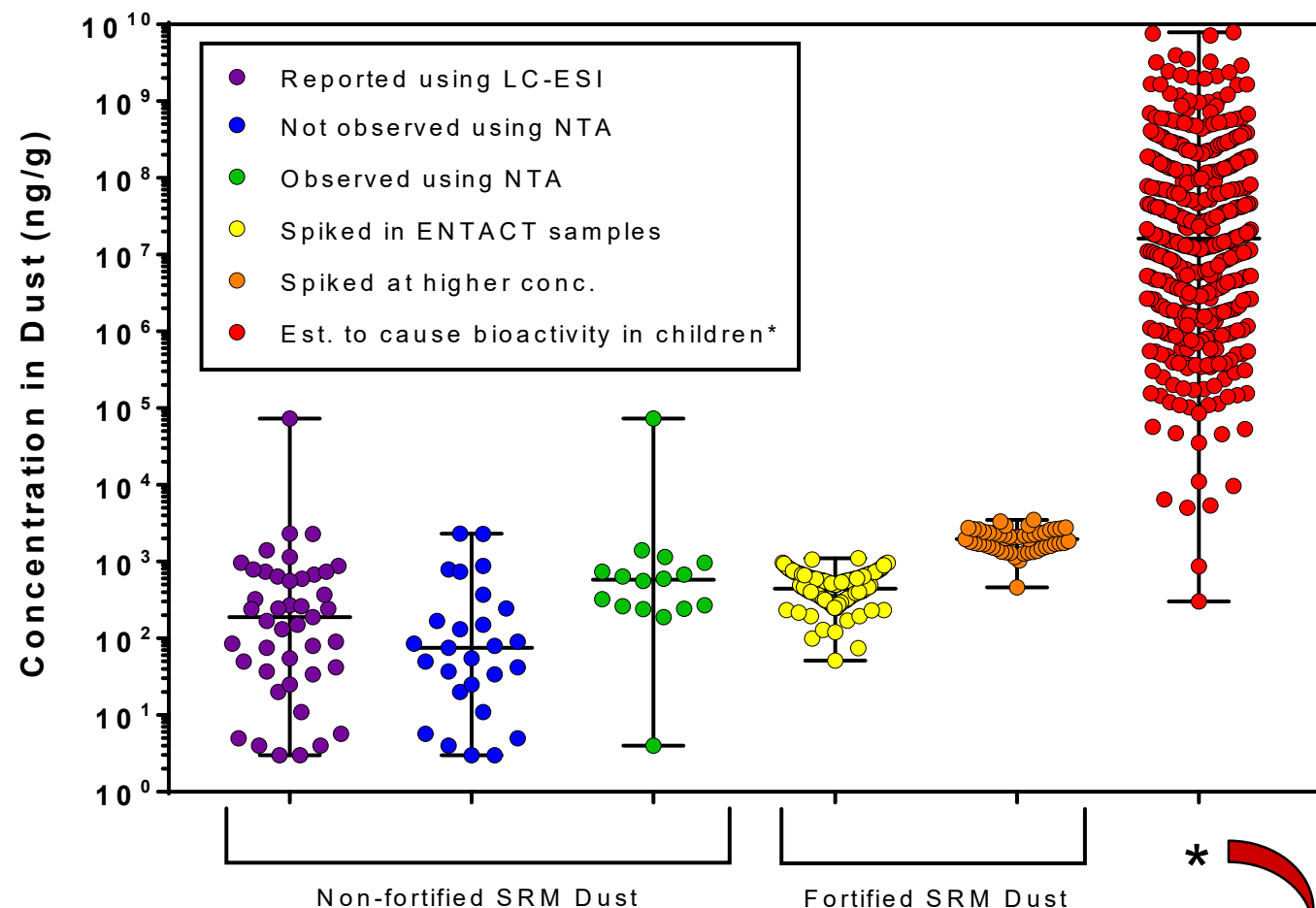
# Experiments with SRM Dust



# 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 dose that would be needed in the most-sensitive 5% of the population to produce a steady-state plasma concentration equal to [the 10<sup>th</sup>] percentile of the ToxCast AC50 distribution across assays for the given chemical."

# Evaluation of *in silico* Spectra

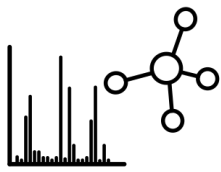
Metabolomics (2015) 11:98–110  
DOI 10.1007/s11306-014-0676-4

## ORIGINAL ARTICLE

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

Felicity Allen · Russ Greiner · David Wishart

Training Set

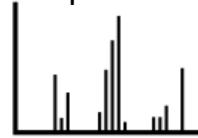


Fragmentation  
Prediction Model

DSSTox structures



DSSTox MS2  
spectra



SCIENTIFIC DATA

OPEN

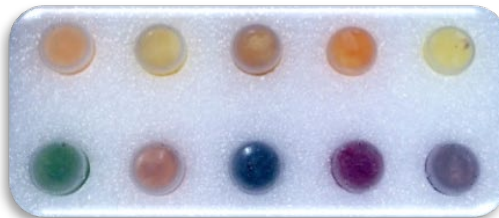
DATA DESCRIPTOR

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

Andrew D. McEachran<sup>1,2</sup>, Ilya Balabin<sup>1</sup>, Tommy Cathey<sup>4</sup>, Thomas R. Transue<sup>4</sup>,  
Hussein Al-Ghoul<sup>1</sup>, Chris Grukke<sup>1</sup>, Jon R. Sobus<sup>5</sup> & Antony J. Williams<sup>2</sup>

10 Synthetic Mixtures:

1,269 Unique ToxCast Substances



LC-QTOF HRMS:  
Data Dependent Acquisition

MS2  
Reference  
Library

Probable  
Structures

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

Tentative  
Structures

No Library  
Matches

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

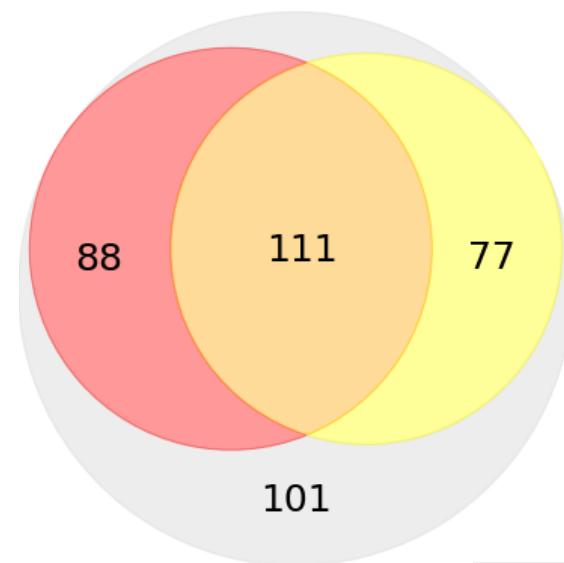
Alex Chao<sup>1,2</sup> · Hussein Al-Ghoul<sup>1,2</sup> · Andrew D. McEachran<sup>1,3</sup> · Ilya Balabin<sup>4</sup> · Tom Transue<sup>4</sup> · Tommy Cathey<sup>4</sup> · Jarod N. Grossman<sup>2,3</sup> · Randolph Singh<sup>1,5</sup> · Elin M. Ulrich<sup>2</sup> · Antony J. Williams<sup>6</sup> ·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

377 ENTACT Compounds  
with MS2 Spectra

Top  
Reference  
Library  
Match



Top *in silico*  
Library  
Match



Not Top Match



# Summary of ENTACT Findings

- NTA methods are suitable for many ToxCast chemicals
  - ~5% of ENTACT compounds not observed across all methods
- Performance determined across 3 categories:
  - **Coverage** = Ability to Observe → (Range = 21% to 69%)
  - **Precision** = Ability to Identify those Observed → (Range = 7% to 99%)
  - **Reproducibility** = Ability to Consistently Identify → (Range = 7% to 97%)
- Multiple methods required for broad characterization
  - No “one size fits all” method
  - <1% of ENTACT compounds observed using all methods
- 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



Credit: the Research Triangle Foundation

## **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.*