

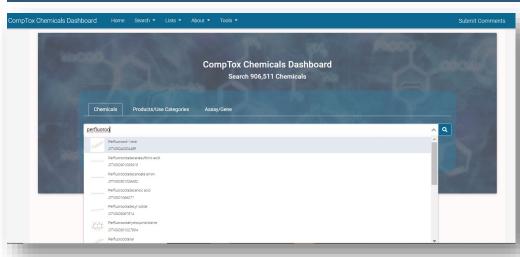
# Targeted Lists of Chemicals to Support Non-Targeted Analysis via the **US-EPA CompTox Chemicals Dashboard**

1. Center for Computational Toxicology & Exposure, Office of Research & Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

### **Overview**

At the US-EPA non-targeted analysis (NTA) uses high-resolution mass spectrometry to better understand the identity of a wide variety of chemicals present in environmental samples. Data processing remains challenging and there is inherent uncertainty in confidently identifying chemicals from candidate lists. We have developed functionality within the CompTox Chemicals Dashboard (available at https://comptox.epa.gov/dashboard) to support this analysis. These tools include the generation of "MS-Ready" structures (1) to optimize database searching, consensus ranking using chemical metadata (2), and in silico MS/MS fragmentation prediction for spectral matching (3). A number of chemical lists have also been developed to support specific applications including the identification of chemicals in human breath (i.e., the volatilome), chemicals in human media (e.g., blood, saliva), and in water. Other lists include PFAS (per- and polyfluoroalkyl substances) chemicals and toxins (e.g., microcystins and cyanotoxins). Combining the synergies of a database containing ~900,000 chemicals with over 320 segregated chemical lists with dashboard search functionality provides a comprehensive workflow to support NTA. The Dashboard provides a freely available web-based application to support structure identification and NTA. Expansion of the types of data hosted in the database can support different targeted and non-targeted projects. This abstract does not reflect EPA policy.

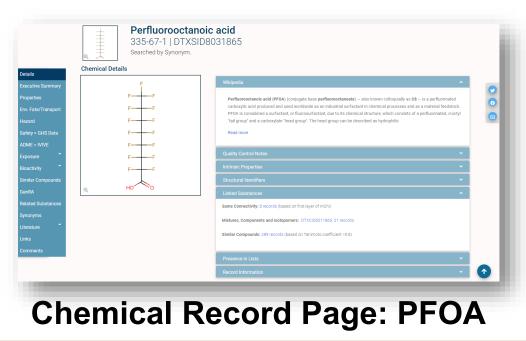
### The CompTox Chemicals Dashboard



### **Dashboard Entry Page**

Where possible, links are provided to related Wikipedia articles. Structure file formats are available for download to the desktop (SMILES and molfile) executive summary and an report regarding chemical toxicity is provided.

landing The page the Of Dashboard is a simple text entry allowing a type-ahead box search for systematic, trade and CAS trivial Registry names, InChI chemical Numbers and identifiers.

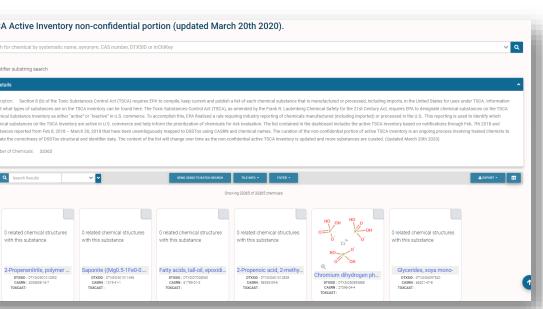


www.epa.gov/research

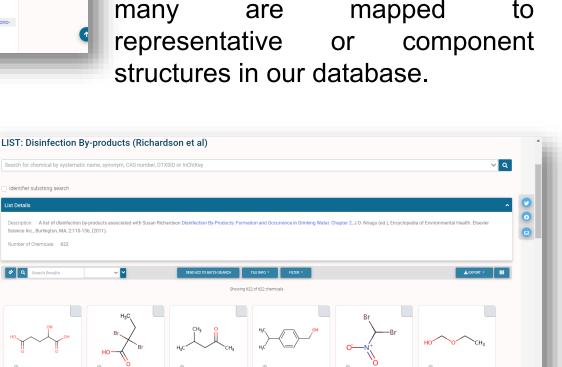
Antony J. Williams<sup>\*</sup> and Charles Lowe

# Accessing Chemical Lists

The dashboard provides access to >320 individual chemical lists. The lists are aggregated from publications, public domain databases, assembled from regulatory lists and assembled to support specific research activities in the agency. Example lists are shown below. Some remain under constant expansion (e.g., the list of ToxCast chemicals) or under versioned release (i.e., the PFAS structure list, presently in its fourth iteration).



The list of Disinfection byproducts is under constant curation and is expanded based on the availability of updated review article from Susan Richardson at University of South the Carolina.



The TSCA (Toxics Substances

Control Act) inventory list is

refreshed with every release, often

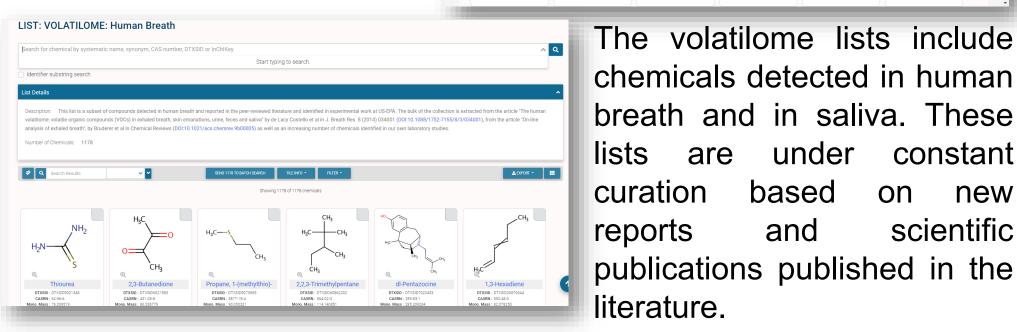
twice a year. The inventory

contains a high proportion of UVCB

(unknown variable composition or

biological substance) chemicals

(without structures) but commonly



Our Center measures the bioactivity of chemicals in hundreds of assays and adds new chemicals and assays every year. These lists are updated with new chemicals for every release of the Dashboard link and to bioactivity data.

rch for chemical by sy	rstematic name, synonym, CAS number, DTXSID or InChIKey
dentifier substring sear	ch
t Details	~
urrent Phase III testing, as vailability), and those disc	onsists of the full list of chemicals having undergone some level of screening in EPA's ToxCast research program from 2007 to the present (last updated 4/11/2017). The list includes all chemicals available for well as discontinued chemicals that underwent limited screening in earlier Phases I and II of the ToxCast program. Discontinued chemicals includes those that were depleted and could not be reprocured (cost, ontinued for other reasons (e.g., limited solubility, instability, volatility). TOXCAST also includes EPA's full, plated contribution of nearly 4000 unique chemicals to the multi-federal agency Tox21 program s/10.1021/acs.chemrestox.6b00135
elated sublists: DXCAST_Phasel: 310 c	hemicals (mostly pesticides) screened in Phase I of the ToxCast program (ph1v1 subset)
OXCAST_Phasell: 1800	chemicals screened in Phase II of the ToxCast program, consisting of TOXCAST_ph1v2, ph2 and e1k sublists
OXCAST_PhaseIII: 4584	chemicals available for screening in the current Phase III of the ToxCast program (as of 4/11/2017)
OXCAST_ph1v2: 293 cf	emicals representing reprocured subset of Phase I (ph1v1) moved into Phase II testing
DXCAST_ph2: 768 cher	nicals added in Phase II of the ToxCast program to increase chemical diversity and coverage of chemicals of concern to EPA programs
DXCAST_e1k: 799 cher	nicals added in Phase II of the ToxCast program that were selected for screening in endocrine-related assays
DXCAST_ph3: 2678 che	micals added in the most recent, ongoing Phase III of the ToxCast program (current as of 4/11/2017)
r more information on E	PA's ToxCast program, see: https://www.epa.gov/chemical-research/toxicity-forecasting
	S data within the EPA ToxCast Dashboard, see: https://www.epa.gov/chemical-research/toxcast-dashboard

are

under

based

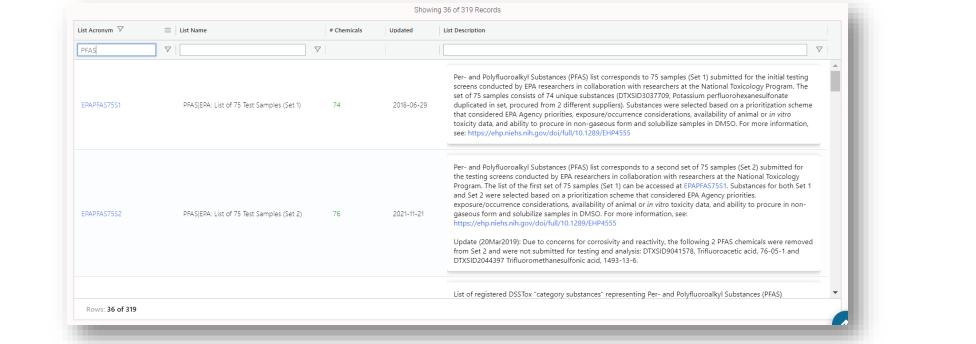
and

constant

scientific

on new

There are >30 individual PFAS lists on the Dashboard. These include a list based on structural filters, UVCB chemicals, PFAS categories represented as Markush structures, and regulatory lists (4).



A growing list of Markush structure representations is available. These structures can be enumerated and distinct mapped to chemical structures as members of the categories. ChemAxon We use enumeration tools.

# References

- 10:45.
- 2) McEachran et al. (2020) Revisiting Five Years of CASMI Contests with EPA Identification Tools. *Metabolites*. 10(6): 260.
- chemistry data to improve identification of unknowns, Scientific Data 6(141)
- 4) Assembly and curation of lists of per- and polyfluoroalkyl substances (PFAS) to support environmental science research, Williams et al. Front. Environ. Sci. 10:850019

# Acknowledgements

The authors thank the CCTE chemical curation team for their rigorous work and the Dashboard software development team for their ongoing development of the dashboard. We acknowledge contributions from the community especially to the NORMAN Network

Innovative Research for a Sustainable Future



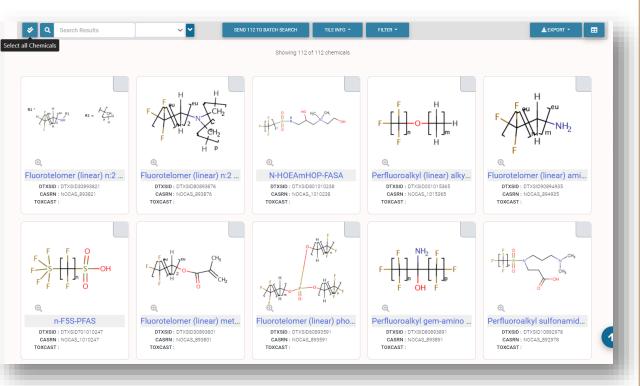
# **SETAC NTA Meeting** Durham, NC May 22-25, 2022



ORCID: 0000-0002-2668-4821 Antony Williams I williams.antony@epa.gov I 919-541-1033

# **PFAS Chemical Lists**

# List of chemical lists of PFAS chemicals: 36 lists and growing



1) McEachran et al. (2018) "MS-Ready" structures for non-targeted high-resolution mass spectrometry screening studies. J Cheminform

3) McEachran et al. (2019) Linking in silico MS/MS spectra with