Developing a list of per- and polyfluoroalkyl (PFAS) chemicals from public domain databases

Antony J. Williams[1], Charles Lowe[1], Chris Grulke[1], Grace Patlewicz[1] and Ann Richard[1]
[1]Center for Computational Toxicology & Exposure, ORD, U.S. EPA
ORCID: 0000-0002-2668-4821

OBJECTIVES

- To build and maintain a list of PFAS chemical substances and associated structures, iterating and versioned over time
- The list of PFAS chemicals to include those chemicals represented as structures and also UVCB chemicals (Unknown or Variable Composition, Complex Reaction Products and Biological Materials)
- Associate the PFAS with experimental and predicted data sourced from literature articles and public databases

APPROACH

- Use a set of selected substructures to search all chemical structures on the CompTox Chemicals Dashboard [1] (https://comptox.epa.gov/dashboard) and combine with text substrings to identify "UVCB chemicals and polymers"
- Deliver lists of PFAS chemicals via the Dashboard that represent different foci – e.g. PFAS in water, PFAS in various international reports, PFAS in food contact materials etc https://comptox.epa.gov/dashboard/chemical_lists/?search=PFAS

MAIN RESULTS

- A list of 9252 PFAS substances (version 2) is available https://comptox.epa.gov/dashboard/chemical_lists/PFASMASTER
- The list is subsetted into both structures and UVCB chemicals. This includes delivery of a set of PFAS chemical category representations, represented as "Markush structures", that expand to sets of chemical structures

IMPACT

- The PFAS lists represent highly curated aggregations [2] of chemical structures and associated information that can be easily accessed by the community
- Lists continue to expand and are updated with each release of the Dashboard
- **For more information, contact:** Antony Williams,

Developing a list of per- and polyfluoroalkyl (PFAS) chemicals from public domain databases

APPROACH
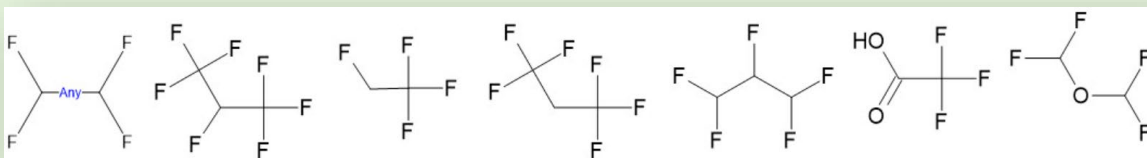
There is no community accepted standard definition for what defines a PFAS chemical. We are using, for now, the substructures above to search all database content.

Resulting structure set of 8163 chemicals is available: https://comptox.epa.gov/dashboard/chemical_lists/PFASSTRUCT. Downloadable as Excel/SDF file with associated metadata

List of PFAS chemicals without explicit structures https://comptox.epa.gov/dashboard/chemical_lists/PFASDEV1 List was assembled by searching on the following substring list: Perfluoro, Polyfluoro, Fluoroethylene, Fluoropropylene, Fluorobutene, Fluoropolymer, "Ethene, 1,1,2,2-tetrafluoro" (the PTFE monomer unit), Chlorotrifluoroethylene, Difluoromethylene, Vinyl fluoride, Tetrafluoro, Pentafluoro, Hexafluoro, Heptafluoro, Octafluoro, Nonafluoro, Decafluoro, Dodecafluoro and filtering out chemical

Developing a list of per- and polyfluoroalkyl (PFAS) chemicals from public domain databases

Many PFAS chemicals fit into "categories" that aggregate sets of chemicals. The PFAS Categories collection contains 112 PFAS categories but is under constant curation as new categories of chemicals are added to the database.

The category "Fluorotelomer (linear) sulfonic acids" maps to 8 chemicals with enumeration of the Markush structure https://comptox.epa.gov/dashboard/dsstoxdb/results?search=DTXSID50892558#related-substances.
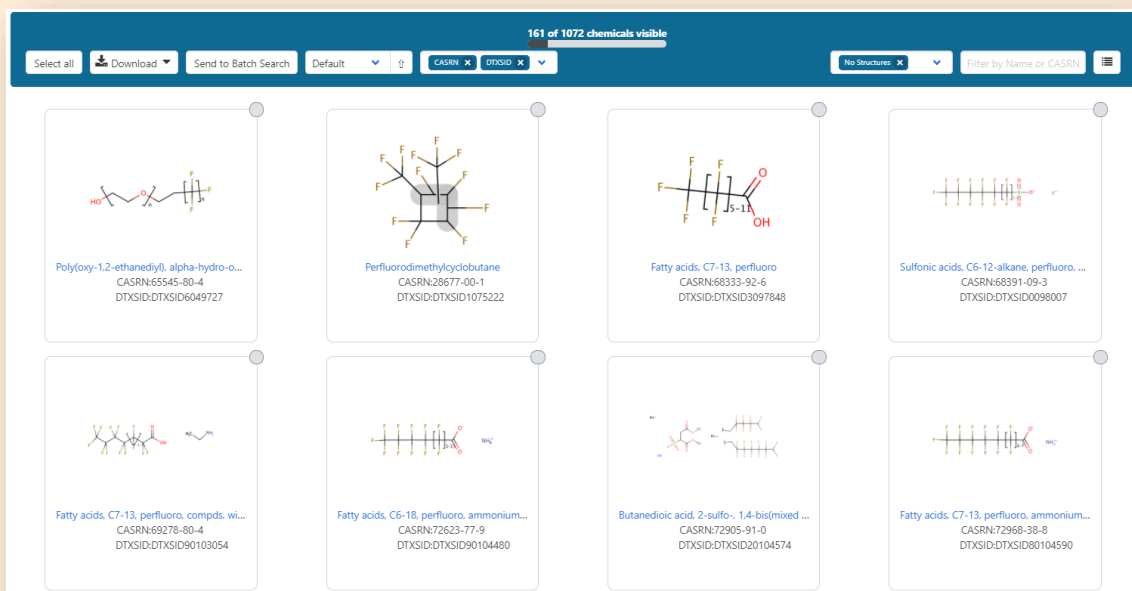
Categories are represented as "Markush structures" that can be enumerated to the set of chemicals that map to these category structure representations. These are
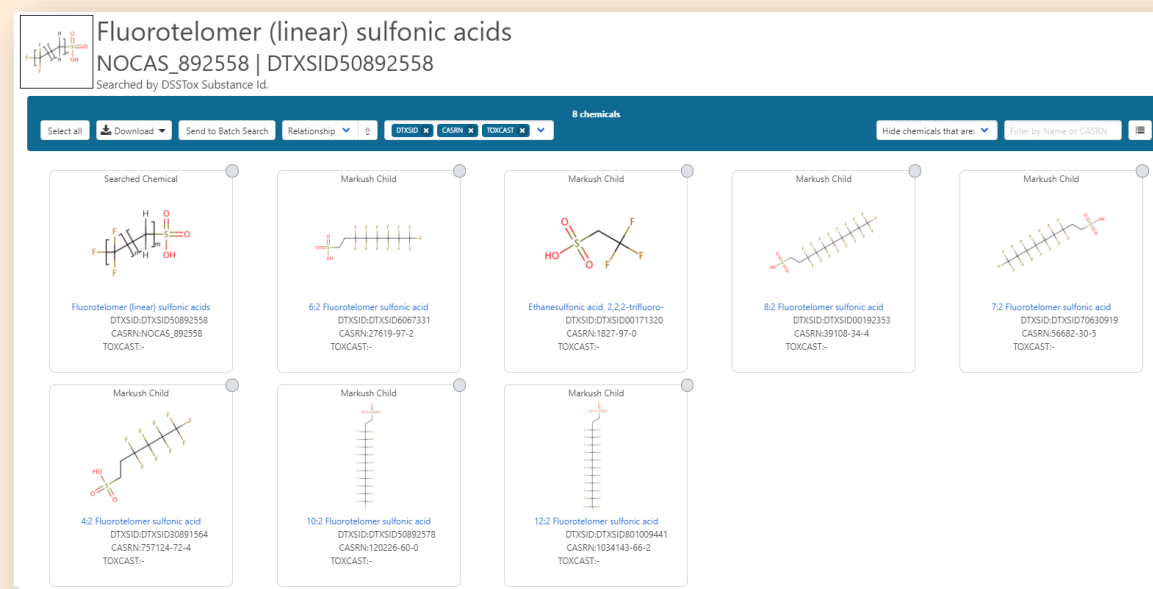
The individual structures can be mapped to experimental and predicted properties, toxicity and bioactivity data and other data on the CompTox Chemicals Dashboard.

Developing a list of per- and polyfluoroalkyl (PFAS) chemicals from public domain databases

Summary

- ~30 lists of PFAS related chemicals available from the CompTox Chemicals Dashboard
- ~9250 PFAS substances presently available based on defined substructures and substrings
- Substances are mapped to available toxicity data, information regarding associated consumer products containing the chemicals, experimental/predicted physicochemical and environmental fate and transport properties

Future Plans

- List(s) of PFAS substances will continue to expand with each versioned release of the Dashboard
- Community feedback will help us continue to iterate on relevant substructures and text substrings for the list
- Aggregate experimental physicochemical and fate and transport data and build new QSAR prediction models
- Publication regarding how the (existing) list was assembled is in progress. All data will be shared as a FAIR dataset for the community to repurpose.

References

1. Williams, A. J. et al. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. Journal of Cheminformatics 9, 61 (2017)
2. Grulke, C. M. et al. EPA's DSSTox database: History of development of a curated chemistry resource supporting computational toxicology research. Computational Toxicology 12, 100096 (2019)
3. G Patlewicz et al, A chemical category-based prioritization approach for selecting 75 per-and polyfluoroalkyl substances (PFAS) for tiered toxicity and toxicokinetic testing, Environmental health perspectives, 127(1), 2019