INTRODUCTION TO GROUPING AND READ-ACROSS: NATIONAL ACADEMY OF SCIENCES REPORT

Gina M. Solomon, M.D., M.P.H.
David Reif, Ph.D.

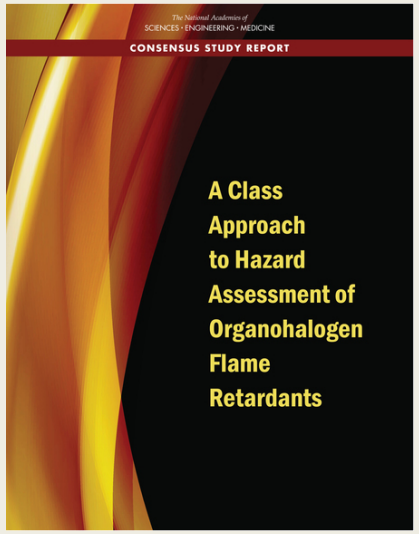
The Problem of PFAS Contamination
Brower Center, Berkeley, CA
December 13, 2019

Report: A Class Approach to Hazard Assessment of Organohalogen Flame Retardants

June 2019

Requested by the Consumer Product Safety Commission

Triggered by a 2015 petition



DAVID C. DORMAN *(Chair)*, NC State University, Raleigh, NC

HUGH A. BARTON *(retired)*, Pfizer, Groton, CT

KAREN BLACKBURN, Procter & Gamble Company, Cincinnati, OH

JOHN BUCHER, NIEHS, Durham, NC

JULIE L. DANIELS, University of North Carolina, Chapel Hill, NC

JENNIFER L. FREEMAN, Purdue University, West Lafayette, IN

KAMEL MANSOURI, Integrated Laboratory Systems, Raleigh, NC

CARMEN MESSERLIAN, Harvard School of Public Health, Boston

DAVID M. REIF, North Carolina State University, Raleigh, NC

GINA M. SOLOMON, Public Health Institute, Oakland, CA

CHIHAE YANG, Molecular Networks GMbH, Columbus, OH

Statement of Task

- Survey available data for organohalogen flame retardants (OFRs) and identify data needs.
- Identify at least one approach for scientifically assessing OFRs as a class for hazard assessment.
- Provide a plan on how to most efficiently and effectively conduct research needed to evaluate OFRs.

> "CPSC needs the hazard assessment plan…when executed, to be readily integrated with a separate quantitative exposure assessment to complete a human health risk assessment."

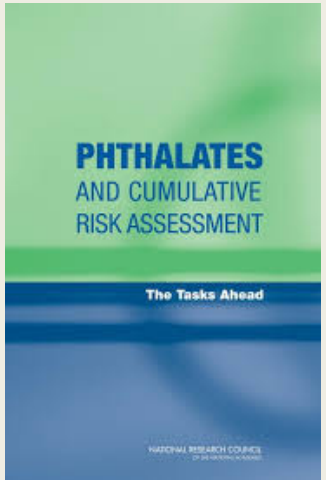
Committee Findings

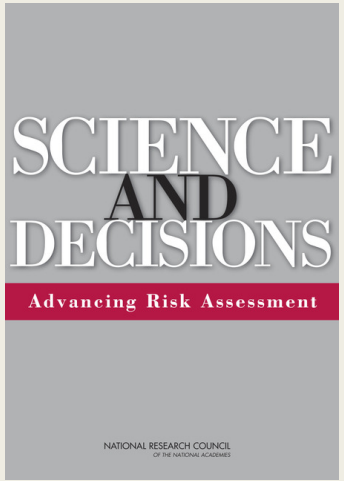
Chemical-by-chemical risk assessment has three main problems:

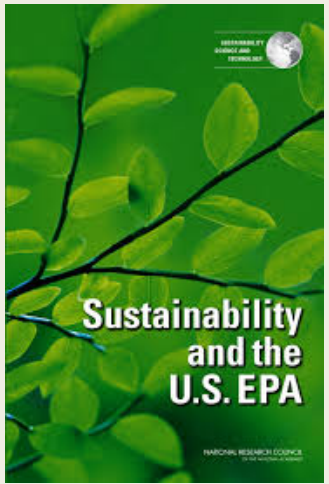
1. Chemicals on which data are insufficient are often deemed not hazardous.
2. Untested chemicals are often substituted for known hazardous chemicals.
3. Cumulative exposure and risk are often ignored.

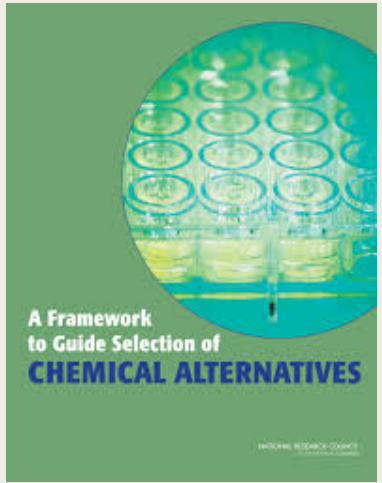
“Ultimately, the sheer number of chemicals in use today demands a new approach to risk assessment.”

The report emphasizes that these findings are consistent with multiple prior NAS reports:

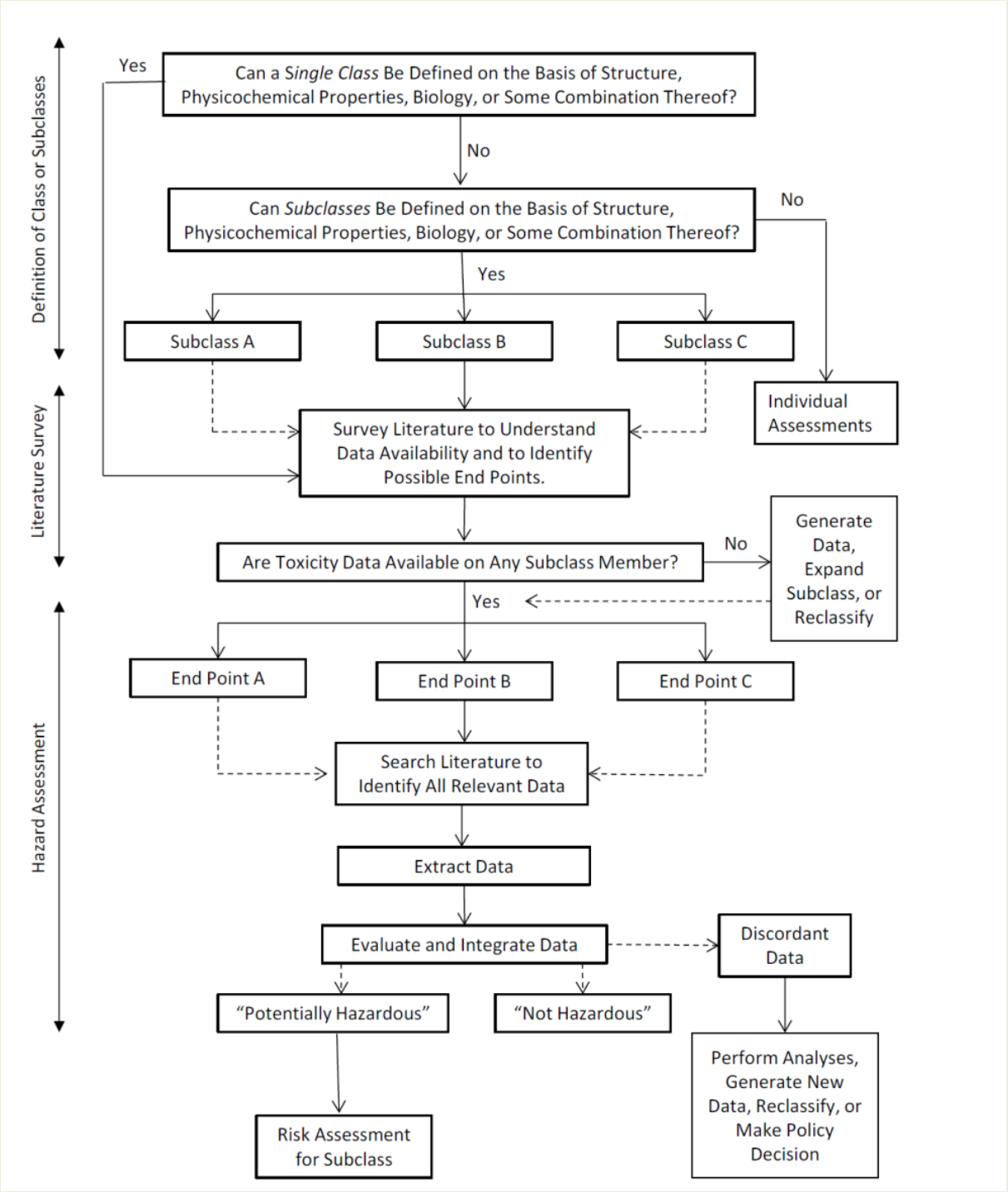






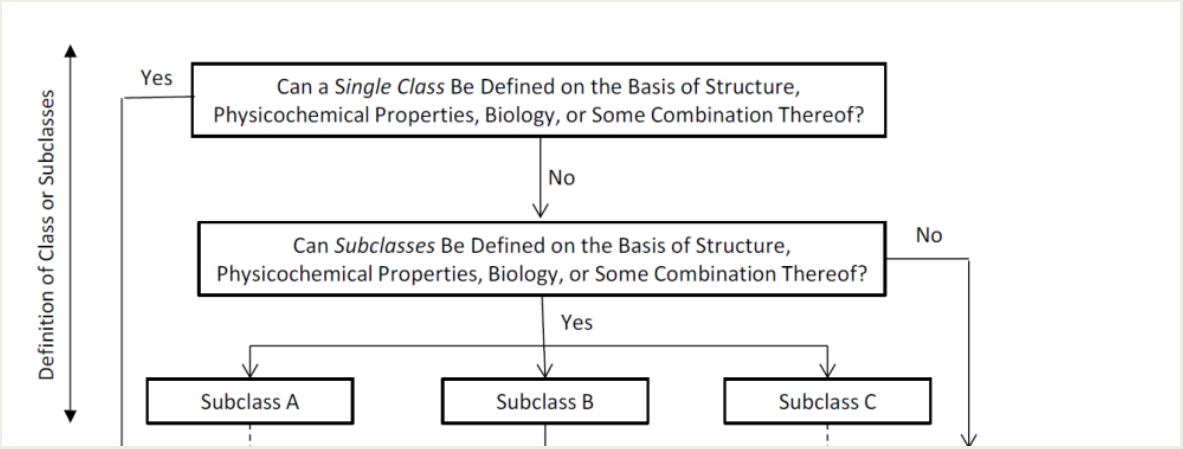


Approach to Defining a Chemical Class



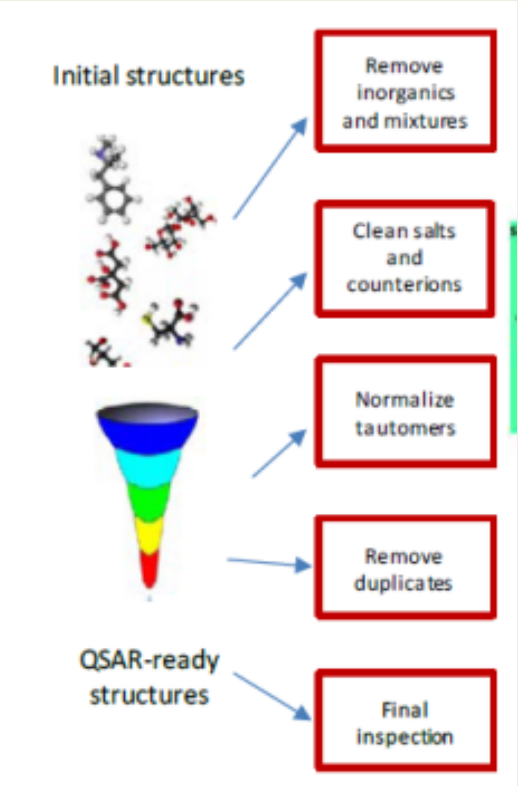
First Question: Can a Class be Defined?

1. Identify and characterize a “Seed” set of chemicals as a working inventory of the class.
2. Generate an “Expanded” set of chemical analogues of the Seed set based on structural similarities.
3. Evaluate the similarity of the Seed set to the analogues to evaluate whether the OFRs are distinguishable as a single class.
4. If not a single class, attempt to define subclasses for hazard evaluation.



Developing the “Seed set”

- Searched 7 data sources and identified 161 OFRs.
- CAS numbers and structures verified on the EPA Dashboard, and chemical structures normalized to be QSAR-ready.
- Several duplicates and four mixtures were eliminated to create an inventory of 148 unique chemical structures.
- Chemical space of the seed set was characterized using OPEn structure-activity/property Relationship App (OPERA) (open access). Predicted phys-chem properties, fate, toxicity, etc.



Data Sources: (1) Eastmond (2015); (2) Danish Environmental Protection Agency (Danish EPA 2016); (3) Environment Agency of the United Kingdom (2003); (4) the WHO International Programme on Chemical Safety (IPCS 1997); (5) European Food Safety Authority (EFSA 2010, 2011a,b,c, 2012a,b); (6) Consumer Product Safety Commission (TERA 2016); and (7) US Environmental Protection Agency (EPA 2015).

Developing the “Expanded set”

1. Used Konstanz Information Miner (KNIME) (open source). Developed automated workflow to identify all organohalogens in the U.S. EPA DSSTox database (~200,000).
2. Used chemistry-development kit (CDK) fingerprints with a Tanimoto Similarity Index of 80% to identify the most similar OFR analogues.
3. Identified 1,073 similar analogues.
4. Normalized structures, used OPERA to generate QSAR predictions on the expanded set.

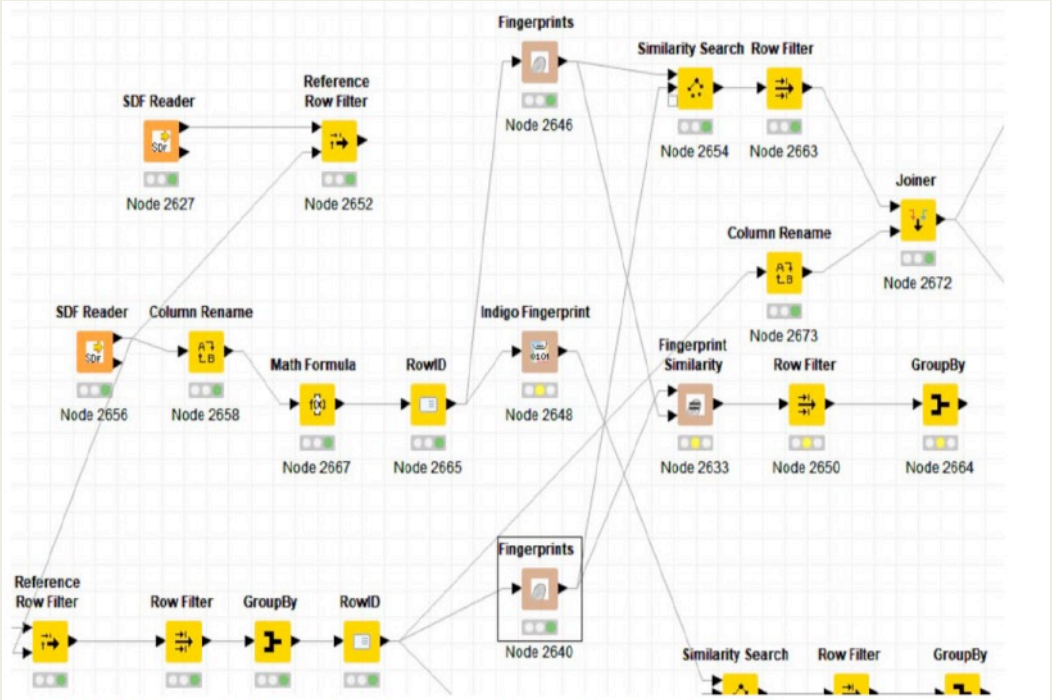


FIGURE B-2 KNIME workflow used to identify organohalogens from DSSTox and determine analogues of OFR seed chemicals.

Comparing the “Seed set” and the Analogues

Evaluated the similarity of the seed set to the analogues using both supervised and unsupervised methods:

1. A principal components analysis (PCA) of OPERA physicochemical properties
2. ToxPrint Chemotype Enrichment
3. Machine Learning Classification using k-nearest-neighbors (kNN) coupled with genetic algorithms (GAs)
4. Supervised PCA

Separation with 80% balanced accuracy using vapor pressure, soil adsorption, and water solubility.

Conclusion: Some differences, but other organohalogens share the same properties as OFRs.

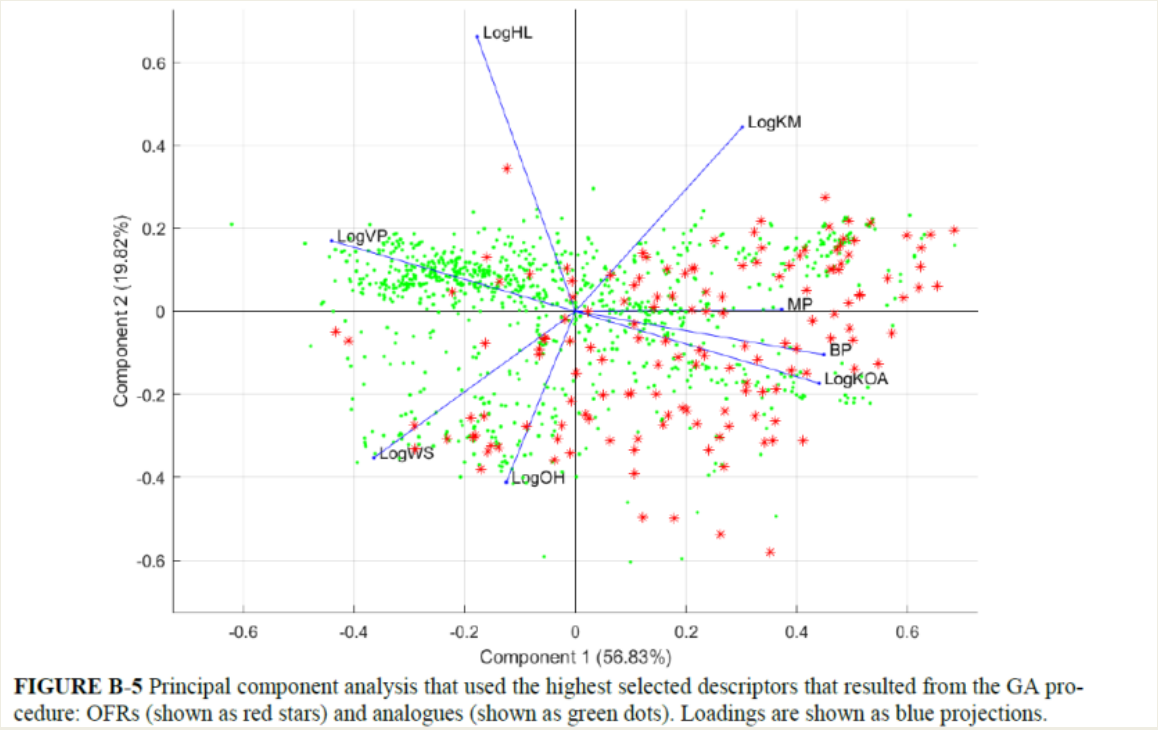
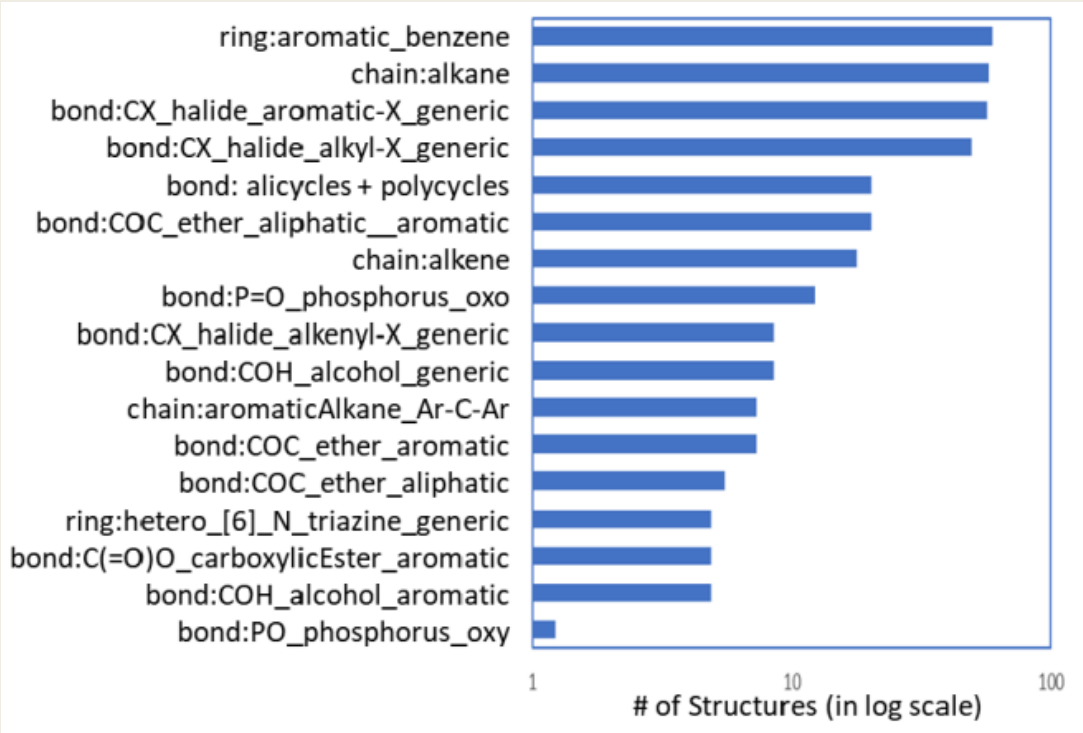


FIGURE B-5 Principal component analysis that used the highest selected descriptors that resulted from the GA procedure: OFRs (shown as red stars) and analogues (shown as green dots). Loadings are shown as blue projections.

Define Sub-Classes

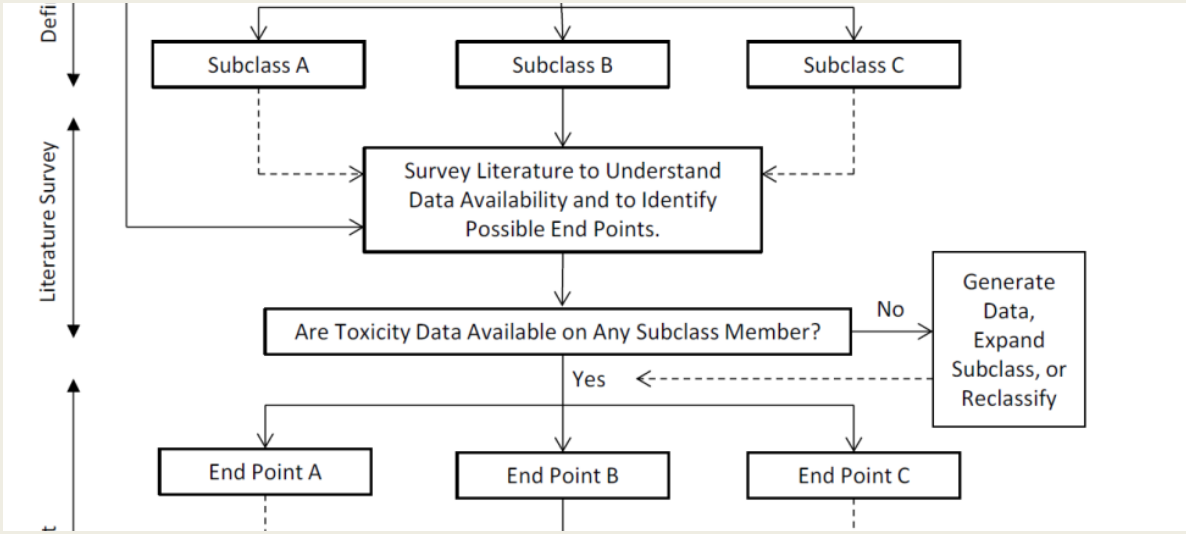
- Used ToxPrint Chemotypes (ChemoTyper) to characterize the 148 structures in the “Seed set”
- Used expert judgment to group the OFR inventory on the basis of predicted biologic activity (such as GABA receptors, aromatase activity, and ER/AR modulators)
  - *8 biology-informed categories identified.*
- Merged structural and biological information.
  - *14 biological/structural subclasses, containing 4-22 members.*
- Some chemicals in more than one class.



Major top-level chemotypes present in the OFR seed set

Evaluation of Subclasses

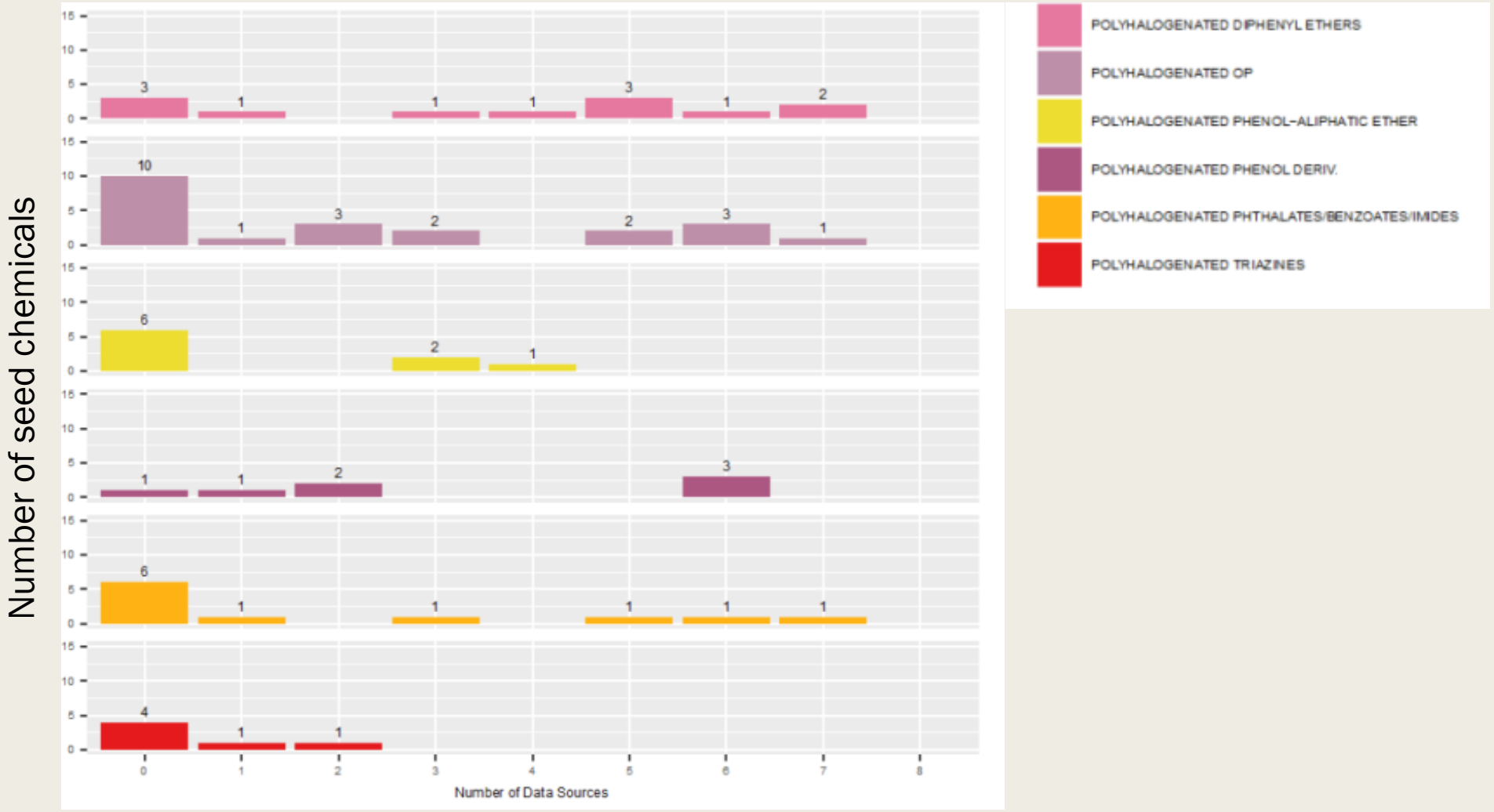
- Broadly survey the data for both the “Seed set” and the “Expanded set”.
- Create an evidence table or map to identify data gaps, and data-rich areas.
- Create analysis plan describing endpoints and relevant data streams to consider.



Data Sources:

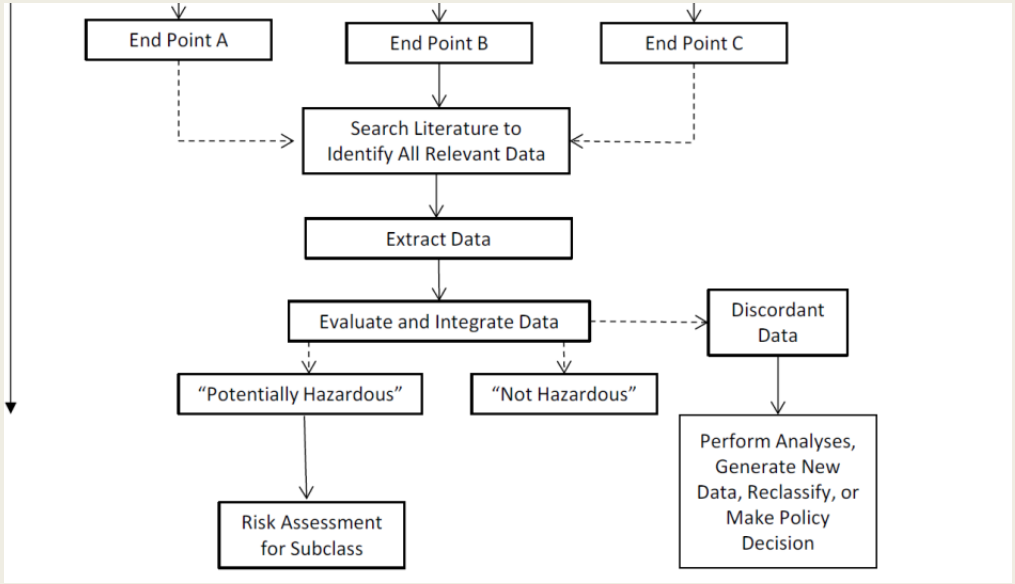
1. Comparative Toxicogenomics Database (CTD)
2. EPA Chemical Dashboard
3. Hazardous Substances Data Bank (HSDB)
4. Integrated Risk Information System (IRIS)
5. ToxCast/Tox21
6. Toxicity Reference Database (ToxRefDB)
7. PubChem
8. ChEMBL

Survey of Available Data



Perform a Hazard Assessment on the Subclass

- Refine the analysis plan based on the survey.
- Review the literature (systematic review or other)
- Include New Approach Methodology (NAM) studies.
- Extract, evaluate, and integrate data.



Ideal Situation:

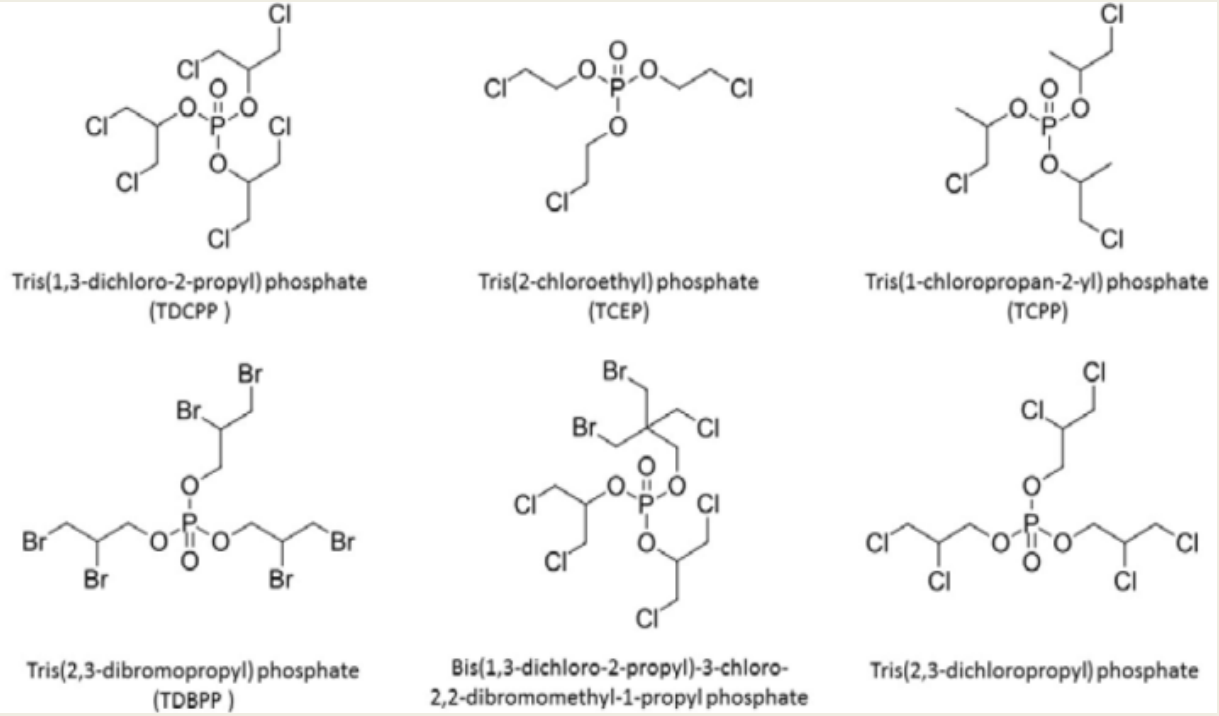
- Data-rich subclass; data are concordant
- Example: PBDEs – 12 member class
  - 7 have been studied to some degree.
  - BDE-47 has strong data on DNT, others generally concordant.
  - *“The present committee concludes that because the data are concordant for the well-studied members of the subclass, a designation of “potentially hazardous” can be applied to the entire subclass.”*

Other Possible Scenarios

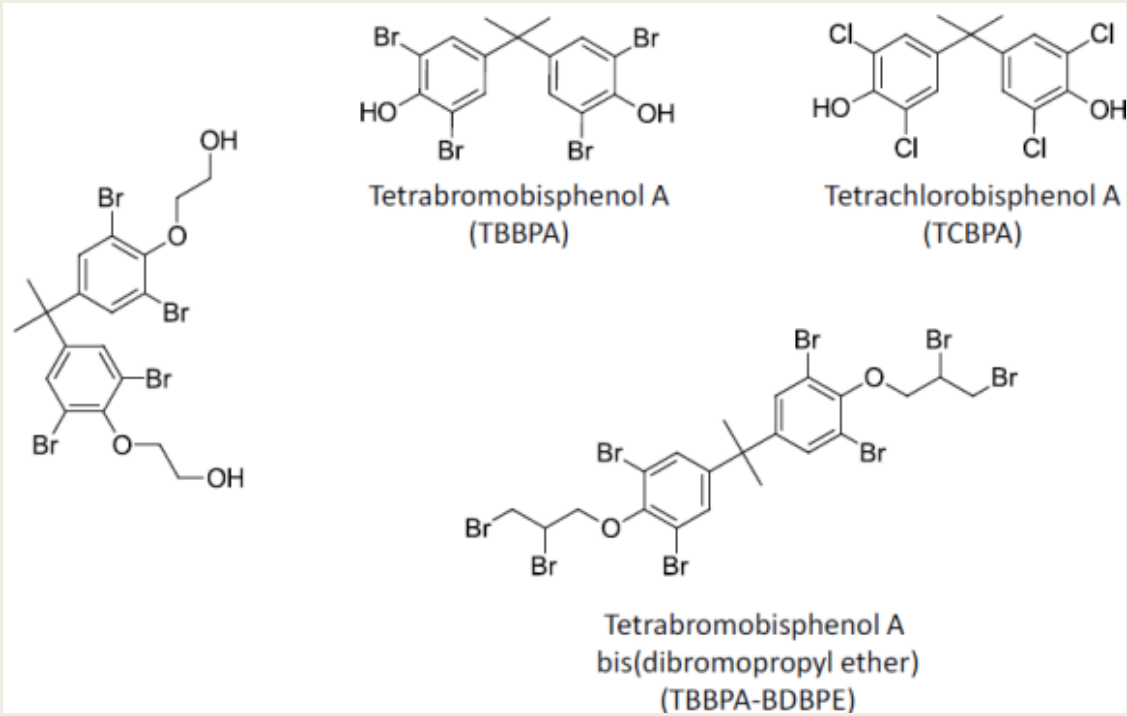
- No data on any member of the subclass. Three options:
  1. Generate new data
  2. Broaden the subclass (e.g., Include “Expanded set”)
  3. Reclassify.
- Data on 1-2 chemicals; no data on the rest. Three options:
  1. Science-based policy decision to treat them all like the ones with data
  2. Extrapolate/interpolate
  3. Generate some data.

Two Case Studies

Polyhalogenated organophosphates (OPs)



Polyhalogenated bisphenol aliphatics



Case Study Results

- Performed literature search and data extraction, including traditional toxicology, Zebrafish data; ToxCast/Tox21 data.
- Focused on developmental toxicity and thyroid homeostasis.
- Evaluated and integrated data.
- Result: “The available data are too heterogeneous or inconsistent on biologic activity.”
- Conclusion: **Discordant data**.

Polyhalogenated organophosphates (OPs)

- 22 Chemicals in the subclass
- 5 with chronic mammalian bioassays
- 4 with mammalian and zebrafish DNT (TDCPP, TCEP, TCPP, TDBPP)
- 5 with some human epidemiology
- Focused on developmental toxicity.

TABLE 3-9 Summary of Experimental Evidence of Developmental Effects in Mammals and Zebrafish Associated with Polyhalogenated Organophosphates

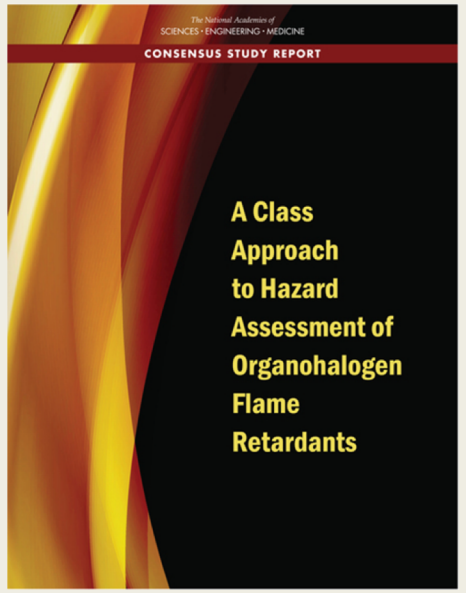
	Mammalian		Zebrafish	
Chemical	Teratogenic	Developmental Neurotoxicity	Teratogenic	Developmental Neurotoxicity or Altered Locomotor Activity
TDCPP	–	–	+	+
TCEP	–	–	+	+
TCPP	–	Not determined	–	+
TDBPP	–	Not determined	–	+

What to Do if Discordant?

- Option 1: Make a policy decision, for example, to extend the most conservative conclusion regarding hazard to the subclass.
- Option 2: Reclassify members to improve their biologic similarity; generate data to increase confidence that reclassification has resulted in biologically similar members.
- Option 3: Perform analyses to explain the discordance and allow the assessment to move forward.
- Option 4: Generate new data that could increase clarity and the scientific basis for a decision.

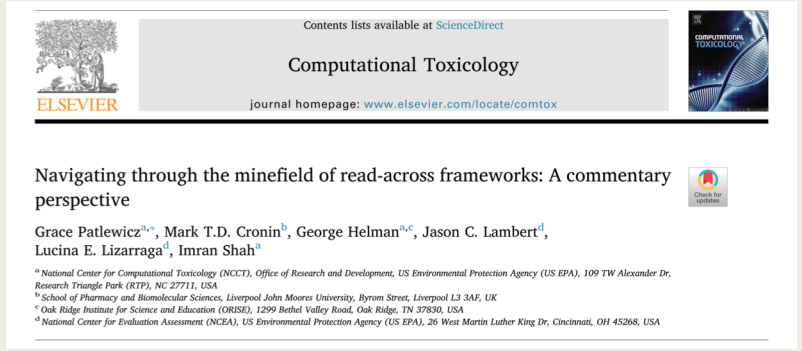
Conclusions

1. A class approach to chemicals is scientifically justifiable in all decision contexts, but the approach to forming classes may differ.
2. In a risk assessment context, classes should be based on a combination of chemistry and predicted biology.
3. If the available data are relatively concordant, it is scientifically justifiable to extrapolate to class members that do not have data.
4. NAMs can be useful for establishing classes and supporting extrapolations across classes.
5. Discordant data within classes is a challenge that will require additional investigation.



What about grouping and read-across for PFAS?

1. The NAS report is an example of a class approach, but the approach to forming classes may differ.
2. Lessons learned can be applied to the PFAS “class”
3. How can we begin to address the key first question of “Can a class be defined?”


Contents lists available at ScienceDirect

Computational Toxicology

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

Navigating through the minefield of read-across frameworks: A commentary perspective

Grace Patlewicz[a,*], Mark T.D. Cronin[b], George Helman[a,c], Jason C. Lambert[d], Lucina E. Lizarraga[d], Imran Shah[a]

[a] National Center for Computational Toxicology (NCCT), Office of Research and Development, US Environmental Protection Agency (US EPA), 109 TW Alexander Dr, Research Triangle Park (RTP), NC 27711, USA
[b] School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK
[c] Oak Ridge Institute for Science and Education (ORISE), 1299 Bethel Valley Road, Oak Ridge, TN 37830, USA
[d] National Center for Evaluation Assessment (NCEA), US Environmental Protection Agency (US EPA), 26 West Martin Luther King Dr, Cincinnati, OH 45268, USA

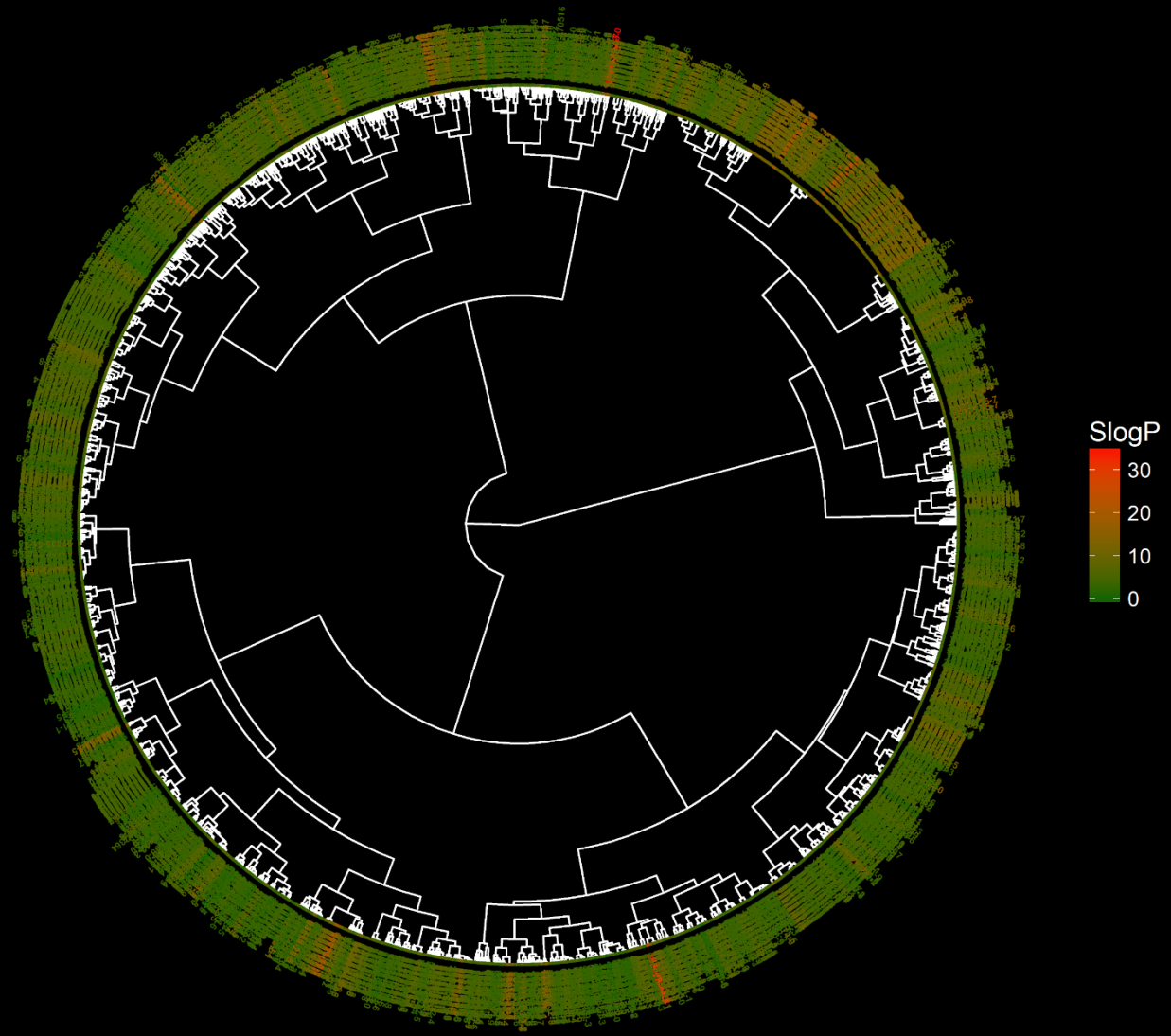
Denis Fourches,
Assistant Professor,
Department of Chemistry,
Bioinformatics Research Center,
North Carolina State University

Automatic Structural Clustering of PFAS Analogues

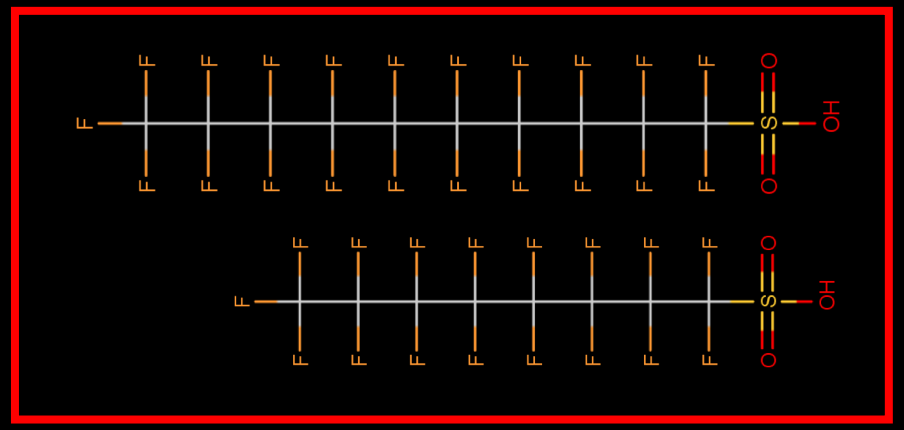
3,430 unique PFAS compounds compiled and integrated from several EPA subsets
(accessed via CompTox Dashboard)

Circular dendrogram obtained with 2D RDKit descriptors
(molecular fingerprints)

→ Objective mapping of the PFAS space based upon chemical structure



Can a class (or subclasses) of PFAS be defined?



PFDS

PFOS

Compounds with similar structures and chemical features are clustered together

Well-known PFAS derivatives are spread over the structural map

→ Rational selection of subsets must be made within a global context

What is relevant biological knowledge (predicted or experimental) that we need to add?