Streamlined PFAS Annotation and Visualization with FluoroMatch Flow and Visualizer

Stephan Baumann[1], Jeremy Koelmel[2], Ziyun Xu[3], LAN Liu[3], Stéphane Bayen[3]

[1]Agilent Technologies, Santa Clara, CA, 95051, USA

[2]Yale School of Public Health, 60 College St., New Haven, CT, 06510, USA

[3]McGill Department of Food Science, 21,111 Lakeshore Ste Anne de Bellevue, Quebec H9X 3V9, Canada

RAFA 2022
Migrants from food contact materials

Agilent
Trusted Answers

Overview

Unidentified organofluorine chemicals account for a significant percentage of organofluorine content in food contact material samples.

Introduction

Greaseproof packaging often contains per and polyfluoroalkyl substances, often abbreviated as PFAS. They are composed of a head group and a fluoroalkyl chain that can vary in length, creating homologous series. In fact, there are more than 12,000 PFAS compounds that have been identified.[1] Despite the pervasiveness of PFAS compounds, there are a little over a hundred commercially available standards, and those tend to be expensive.

Tentatively identifying PFAS in food contact material (FCM) requires a screening instrument and annotation software. Typically, screening for these compounds in food contact materials (FCM) is done by liquid chromatography based high-resolution tandem mass spectrometry. To date, no automated open-source PFAS data analysis software exists to mine fluorine-containing organic compounds.

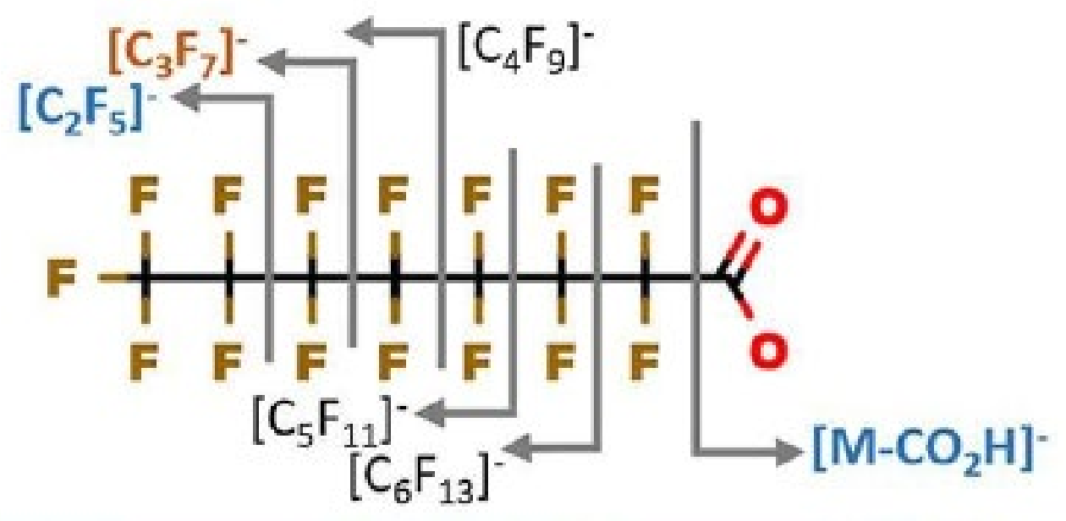


Figure 1: Around 7000 PFAS fragmentation patterns have been evaluated from standards and literature to develop PFAS-specific annotation rules.

We introduced FluoroMatch, which automates file conversion, chromatographic peak picking, blank feature filtering, PFAS annotation based on retention time, precursor masses and fragment masses, annotation ranking, and confidence assignment.[2]

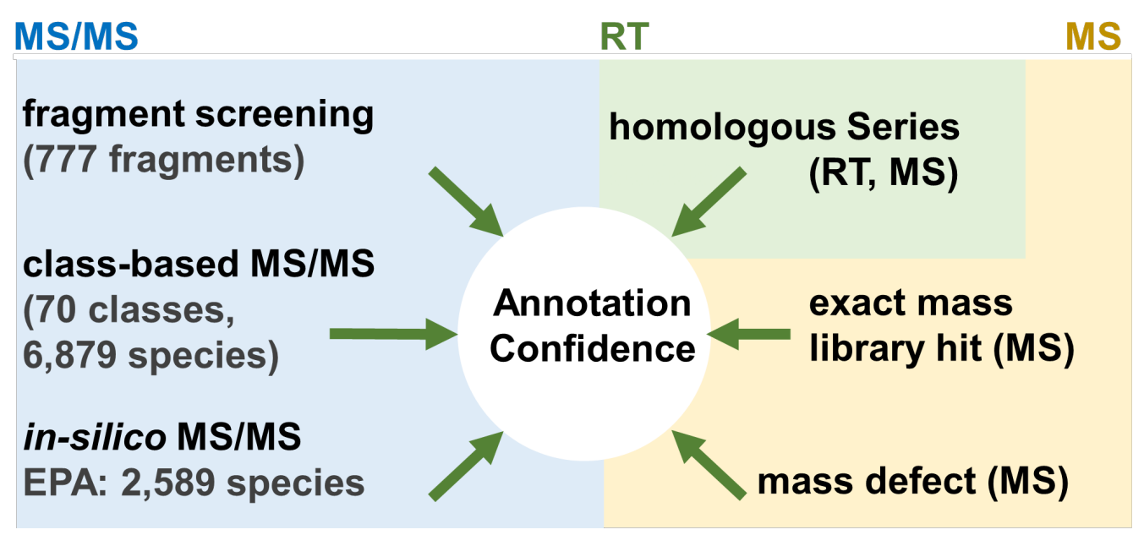


Figure 2: Annotation ranking and confidence assignment, require many lines of evidence. FluoroMatch uses the Schymanski Scale to rank annotations.

To aid interpretation by making homologous series more identifiable, we have added a Visualizer tool to the FluoroMatch suite of software utilizing Microsoft PowerBI. It provides interactive mass defect plots, accurate mass vs. retention time plots, MS/MS fragmentation plots, annotation tables, and fragment screening. Selecting a feature in one graph will adjust what is displayed in other views. This interactive cross-filtering allows simplified evaluation of a feature, PFAS series, or other groups of features.

This work is the first application of FluoroMatch automated PFAS annotation using in-silico PFAS fragmentation to food packaging.

Methods

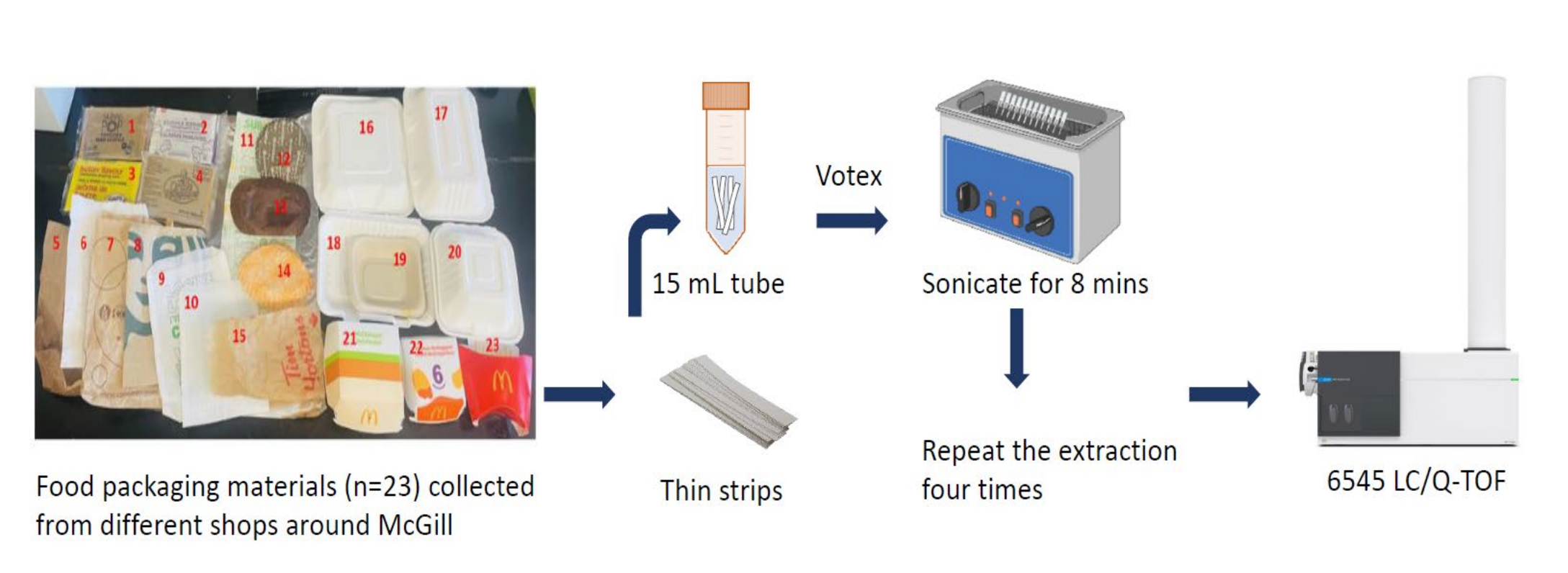
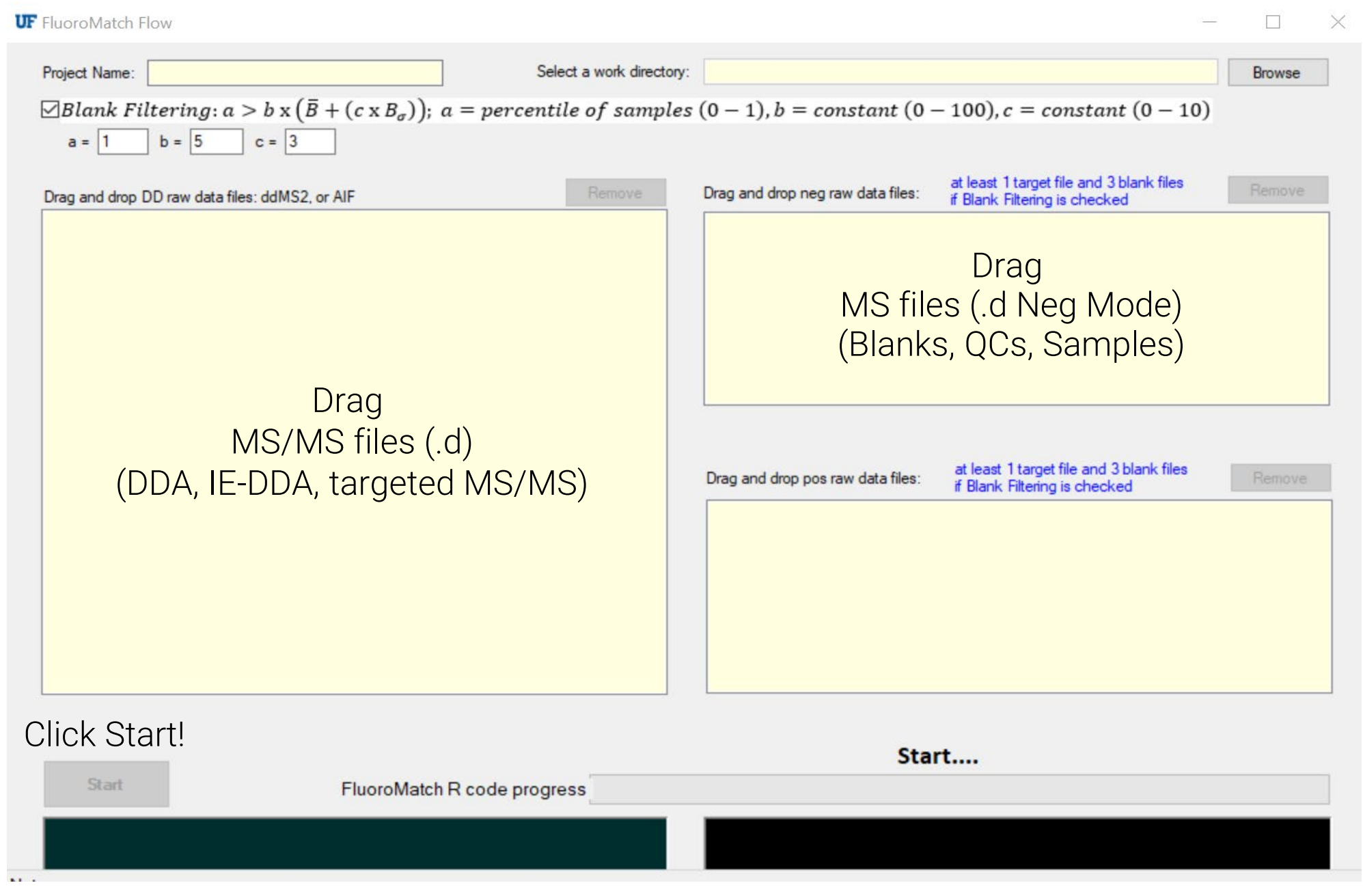


Figure 3: General workflow showing the selected food contact materials commonly found around McGill University, the extraction, and the 6545 LC/Q-TOF used for analysis.

Approximately 0.2 g samples were taken from 23 packing samples suspected of containing PFAS. Suspect matrices include (takeaway) containers, deli paper, paper plates, microwave popcorn bags, and baking supplies. These samples were cut into thin strips and placed into 15 mL centrifuge tubes. To each sample, 4 mL of fresh EPA 1633 diluent was added. To aid extraction, the samples were vortexed for 1 min, sonicated for 8 min, and stored (extracted) at room temperature for an hour. The vortex, sonication, and storage steps were repeated four additional times. After approximately 5 hr of extraction, a 200 µL aliquot was removed for analysis. Next, a 500 µL aliquot was taken. To these aliquots, internal standards (ISTDs) and extracted internal standards (EIS) were added for potential quantification. Five process banks were also prepared by adding 4 mL of fresh EPA 1633 diluent to 15 mL centrifuge tubes and extracting this neat solution. There were also three LC/MS grade methanol blanks. Six PFAS standards prepared in LC/MS grade methanol [range from 0.1; 0.5; 1; 5; 10; 50 ng/mL] with the 10 ng/mL standard being run five times in total during the batch. There are also three matrix matched calibration standards prepared at the same concentrations as the methanol standards. One pooled sample containing extracts from all 23 samples was also prepared for normalization. The diluted samples were injected four times for iterative exclusion data-dependent analysis (iterative MS/MS), with a 10 µL injection volume onto an Agilent 1290 Infinity II ultra-high-performance liquid chromatography (UHPLC) system connected to an Agilent 6545 quadrupole time-of-flight mass spectrometer (LC/Q-TOF). Blanks were acquired every other injection for blank filtering. PFAS were detected in negative electrospray ionization mode. Data was acquired from m/z 100-1100, with MS/MS collision energy set to 0, 25, and 40 eV.

FluoroMatch Flow and Visualizer were designed as an open-source solution for PFAS Annotation that is freely available from innovativeomics.com/software.



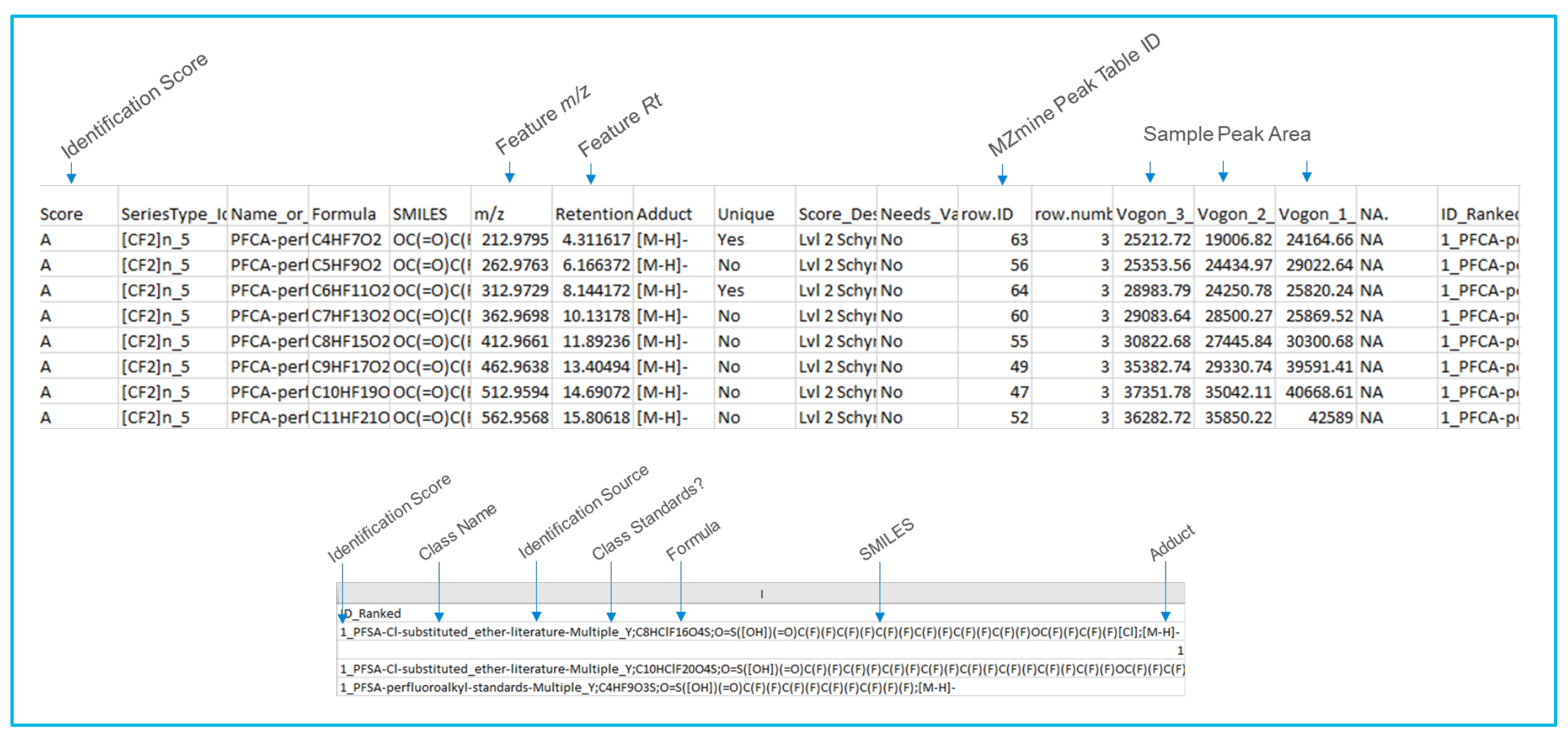
☑ $Blank\ Filtering: a > b \times (\bar{B} + (c \times B_\sigma));\ a = percentile\ of\ samples\ (0-1), b = constant\ (0-100), c = constant\ (0-10)$

Figure 4: FluoroMatch Flow is designed for ease-of-use with drag-and-drop capability. FluoroMatch Flow directly processes vendor files and includes a systematic scoring framework to communicate confidence for every feature.

FluoroMatch Visualizer is built on the Microsoft Power BI Desktop software. The Visualizer PBIX report file was created to provide researchers with a PFAS-specific template. User workflows can be diverse. With the Power BI Desktop, new graphs, variables, and tables can be designed and added. For example, new columns can be added to tables containing information of interest, new plots, for example mass defect versus retention time can be added, and new splicers and filters can be developed.

Results

Many structure elucidation algorithms are focused on the six main chemical elements that are necessary for life, namely carbon (C), hydrogen (H), nitrogen (N), oxygen (O), phosphorus (P), and sulfur (S). This means that they tend to perform poorly when confronted with anthropogenic compounds such as PFAS. FluoroMatch Flow automates the PFAS nontargeted and suspect screening workflow using multiple lines of evidence. The integrated steps include file conversion using msConvert, a unique untargeted chromatographic peak picking strategy implementing MZmine 2.26 (users own peak picking workflow can be integrate into FluoroMatch Modular), and blank feature filtering (BFF). It outputs annotations using exact mass and fragment masses, rankings of multiple annotations for features, and compilations of metadata on fragmentation information and peaks used to annotate features.



Score	SeriesType_Id	Name_or_	Formula	SMILES	m/z	Retention	Adduct	Unique	Score_Des	Needs_Va	row.ID	row.numb	Vogon_3_	Vogon_2_	Vogon_1_	NA.	ID_Ranked
A	[CF2]n_5	PFCA-per	C4HF7O2	OC(=O)C(	212.9795	4.311617	[M-H]-	Yes	Lvl 2 Schy	No	63	3	25212.72	19006.82	24164.66	NA	1_PFCA-p
A	[CF2]n_5	PFCA-per	C5HF9O2	OC(=O)C(	262.9763	6.166372	[M-H]-	No	Lvl 2 Schy	No	56	3	25353.56	24434.97	29022.64	NA	1_PFCA-p
A	[CF2]n_5	PFCA-per	C6HF11O2	OC(=O)C(	312.9729	8.144172	[M-H]-	Yes	Lvl 2 Schy	No	64	3	28983.79	24250.78	25820.24	NA	1_PFCA-p
A	[CF2]n_5	PFCA-per	C7HF13O2	OC(=O)C(	362.9698	10.13178	[M-H]-	No	Lvl 2 Schy	No	60	3	29083.64	28500.27	25869.52	NA	1_PFCA-p
A	[CF2]n_5	PFCA-per	C8HF15O2	OC(=O)C(	412.9661	11.89236	[M-H]-	No	Lvl 2 Schy	No	55	3	30822.68	27445.84	30300.68	NA	1_PFCA-p
A	[CF2]n_5	PFCA-per	C9HF17O2	OC(=O)C(	462.9638	13.40494	[M-H]-	No	Lvl 2 Schy	No	49	3	35382.74	29330.74	39591.41	NA	1_PFCA-p
A	[CF2]n_5	PFCA-per	C10HF19O	OC(=O)C(	512.9594	14.69072	[M-H]-	No	Lvl 2 Schy	No	47	3	37351.78	35042.11	40668.61	NA	1_PFCA-p
A	[CF2]n_5	PFCA-per	C11HF21O	OC(=O)C(	562.9568	15.80618	[M-H]-	No	Lvl 2 Schy	No	52	3	36282.72	35850.22	42589	NA	1_PFCA-p

```
ID_Ranked
1_PFSA-Cl-substituted_ether-literature-Multiple_Y;C8HClF16O4S;O=S([OH])(=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)OC(F)(F)C(F)(F)[Cl];[M-H]-
                                                                                                                                        1
1_PFSA-Cl-substituted_ether-literature-Multiple_Y;C10HClF20O4S;O=S([OH])(=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)OC(F)(F)C(F)
1_PFSA-perfluoroalkyl-standards-Multiple_Y;C4HF9O3S;O=S([OH])(=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F;[M-H]-
```

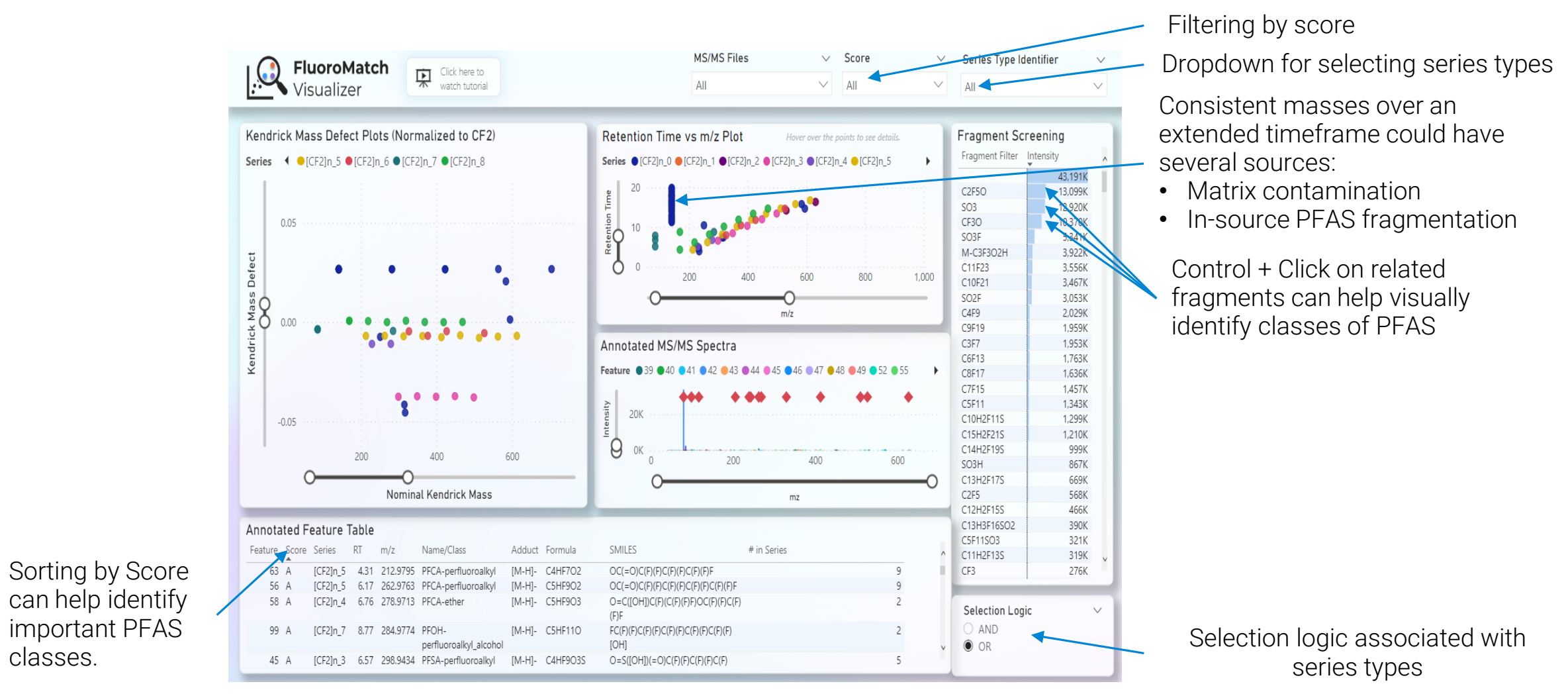
Figure 5: A primary FluoroMatch output file is in the format of a CSV file. It also generates a PBIX file for FluoroMatch Visualizer.

	PFAS chemicals			Concentration (ng/g per sample)						
	Name	m/z	RT (PFC system)	S2	S6	S16	S17	S18	S19	S20
1	HFPO-DA	284.9733	7.388							
2	PFBS	298.9430	5.600							
3	PFHxA	312.9728	6.891		6.5	13	7.4	12	8.6	8.2
4	PFHpA	362.9696	8.695			6		7.5		
5	ADONA	376.9689	8.937							
6	PFHxS	398.9366	8.892							
7	PFOA	412.9664	10.402			35		33		
8	PFNA	462.9632	11.909			6.8				
9	PFOS	498.9302	11.953	6						
10	PFDA	512.9600	13.215			25		19		
11	9Cl-PF3ONS	530.8956	12.774							
12	PFUnDA	562.9568	14.348							
13	MeFOSAA	569.9568	13.884							
14	EtFOSAA	583.9830	14.472							
15	PFDoDA	612.9537	15.326			17		13		
16	11Cl-PF3OUdS	630.8892	14.925							
17	PFTrDA	662.9505	16.189							
18	PFTeDA	712.9473	16.941							

Table 1: Thirty-two calibration standards were run. One set was made in pure solvent and three sets were made in sample matrices at eight levels: 0.1, 0.5, 1, 2.5, 5, 7.5, 10, 50 ng/mL. Along with the samples shown here, there were procedural blanks, pooled samples, and quality control samples. Samples 16 and 18, identified as biodegradable, grease-proof packaging in Figure 3, have the highest levels of PFAS.

To aid in interpretation, FluoroMatch Visualizer interface was designed so that all relevant information could be observed simultaneously. The interface consists of three filters: by MS/MS file, score, and chemical series. It has three visuals: m/z vs retention time, normalized mass defect plot, and MS/MS spectra. It also contains a table of fragments, and table of annotated features. EICs, isotopic pattern, and statistical visualizations.[3]

Because of the complexity and richness of nontargeted data, users need to prioritize which group of features to investigate. Filtering by score and/or fragments allows FluoroMatch Visualizer users to determine which PFAS features to focus on based on features, including annotation quality.



Feature	Score	Series	RT	m/z	Name/Class	Adduct	Formula	SMILES	# in Series
63	A	[CF2]n_5	4.31	212.9795	PFCA-perfluoroalkyl	[M-H]-	C4HF7O2	OC(=O)C(F)(F)C(F)(F)C(F)(F)F	9
56	A	[CF2]n_5	6.17	262.9763	PFCA-perfluoroalkyl	[M-H]-	C5HF9O2	OC(=O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F	9
58	A	[CF2]n_4	6.76	278.9713	PFCA-ether	[M-H]-	C5HF9O3	O=C([OH])C(F)(C(F)(F)F)OC(F)(F)C(F)(F)F	2
99	A	[CF2]n_7	8.77	284.9774	PFOH-perfluoroalkyl_alcohol	[M-H]-	C5HF11O	FC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)[OH]	2
45	A	[CF2]n_3	6.57	298.9434	PFSA-perfluoroalkyl	[M-H]-	C4HF9O3S	O=S([OH])(=O)C(F)(F)C(F)(F)C(F)(F)C(F)	5

Filtering by score

Dropdown for selecting series types

Consistent masses over an extended timeframe could have several sources:

- Matrix contamination
- In-source PFAS fragmentation

Control + Click on related fragments can help visually identify classes of PFAS

Sorting by Score can help identify important PFAS classes.

Selection logic associated with series types

Figure 6: FluoroMatch Visualizer provides interactive mass defect plots, accurate mass vs. retention time plots, MS/MS fragmentation plots, annotation tables, and fragment screening. Selecting and sorting on fragments and features can aid in confirming annotations in homologous series.

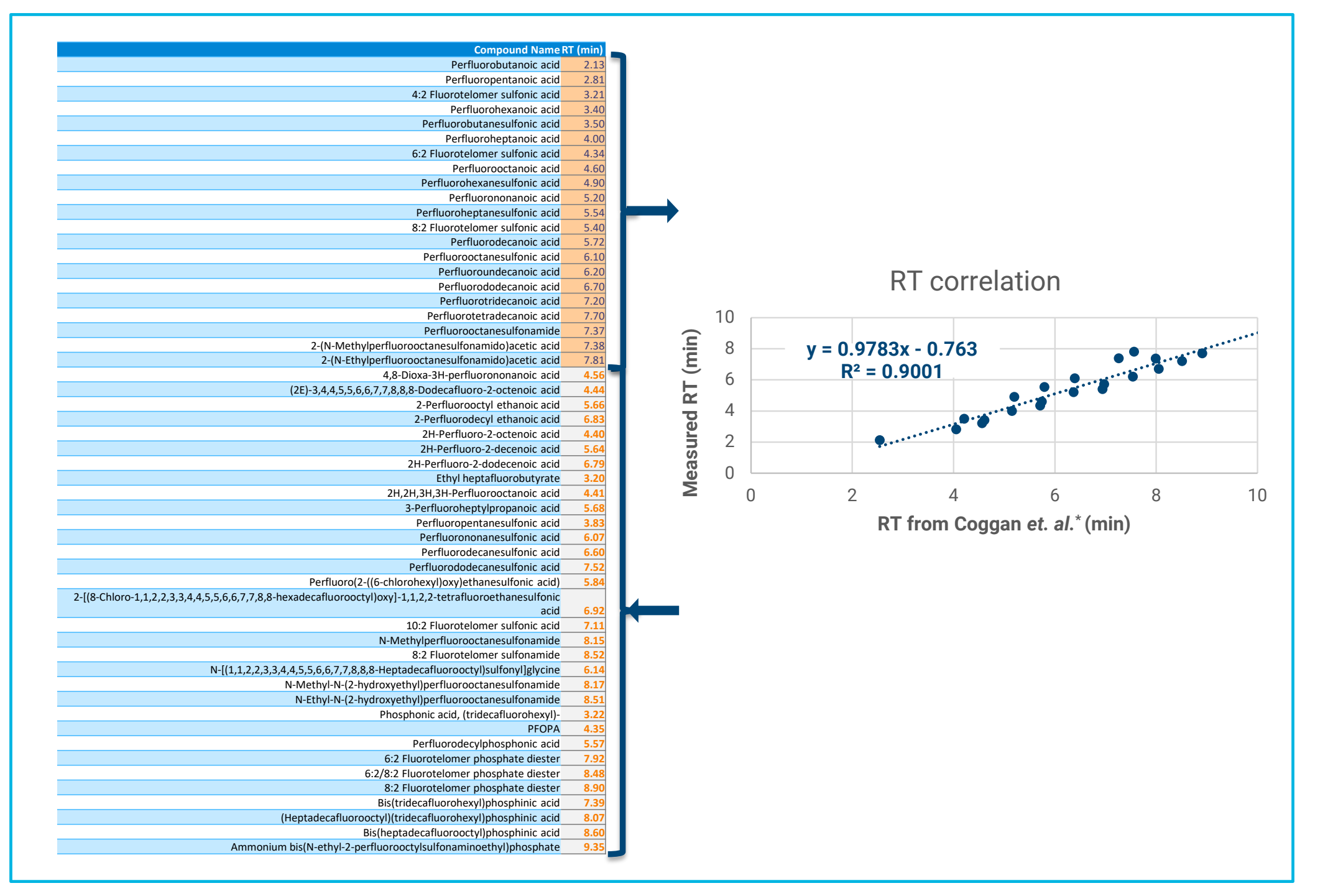
Conclusions

FluoroMatch automated PFAS annotation using in-silico PFAS fragmentation libraries and rule-based annotation. We introduce in-silico fragmentation libraries containing over 7,000 PFAS across 72 PFAS subclasses, built using spectra from literature and authentic standards.

Validating the percent coverage and accuracy of annotations in real-world samples was challenging due to the case of known unknowns and unknown-unknowns. Here, we used all-ion fragmentation to estimate that FluoroMatch covered 71% of CF_2 containing PFAS compounds with fragmentation and CF_2 normalized mass defect plots to estimate 56% coverage of compounds with the remaining being false negatives.

FluoroMatch Visualizer allowed the investigation of trends across PFAS by narrowing down the number of features. One of the most useful approaches was to select individual homologous series, automatically determined using nominal mass and normalized mass defect. When series were selected, all visuals, including MS/MS spectra, were updated to show all members of a series overlaid. Then patterns could easily be observed, and outliers determined. Tens to hundreds of series often exist, and these series can be reduced by those containing high scores or certain characteristic PFAS fragments.

Future Developments



Compound Name	RT (min)
Perfluorobutanoic acid	2.13
Perfluoropentanoic acid	2.81
4:2 Fluorotelomer sulfonic acid	3.21
Perfluorohexanoic acid	3.40
Perfluorobutanesulfonic acid	3.50
Perfluoroheptanoic acid	4.00
6:2 Fluorotelomer sulfonic acid	4.34
Perfluorooctanoic acid	4.60
Perfluorohexanesulfonic acid	4.90
Perfluorononanoic acid	5.20
Perfluoroheptanesulfonic acid	5.54
8:2 Fluorotelomer sulfonic acid	5.40
Perfluorodecanoic acid	5.72
Perfluorooctanesulfonic acid	6.10
Perfluoroundecanoic acid	6.20
Perfluorododecanoic acid	6.70
Perfluorotridecanoic acid	7.20
Perfluorotetradecanoic acid	7.70
Perfluorooctanesulfonamide	7.37
2-(N-Methylperfluorooctanesulfonamido)acetic acid	7.38
2-(N-Ethylperfluorooctanesulfonamido)acetic acid	7.81
4,8-Dioxa-3H-perfluorononanoic acid	4.56
(2E)-3,4,4,5,5,6,6,7,7,8,8,8-Dodecafluoro-2-octenoic acid	4.44
2-Perfluorooctyl ethanoic acid	5.66
2-Perfluorodecyl ethanoic acid	6.83
2H-Perfluoro-2-octenoic acid	4.40
2H-Perfluoro-2-decenoic acid	5.64
2H-Perfluoro-2-dodecenoic acid	6.79
Ethyl heptafluorobutyrate	3.20
2H,2H,3H,3H-Perfluorooctanoic acid	4.41
3-Perfluoroheptylpropanoic acid	5.68
Perfluoropentanesulfonic acid	3.83
Perfluorononanesulfonic acid	6.07
Perfluorodecanesulfonic acid	6.60
Perfluorododecanesulfonic acid	7.52
Perfluoro(2-((6-chlorohexyl)oxy)ethanesulfonic acid)	5.84
2-[(8-Chloro-1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8-hexadecafluorooctyl)oxy]-1,1,2,2-tetrafluoroethanesulfonic acid	6.92
10:2 Fluorotelomer sulfonic acid	7.11
N-Methylperfluorooctanesulfonamide	8.15
8:2 Fluorotelomer sulfonamide	8.52
N-[(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,8-Heptadecafluorooctyl)sulfonyl]glycine	6.14
N-Methyl-N-(2-hydroxyethyl)perfluorooctanesulfonamide	8.17
N-Ethyl-N-(2-hydroxyethyl)perfluorooctanesulfonamide	8.51
Phosphonic acid, (tridecafluorohexyl)-	3.22
PFOPA	4.35
Perfluorodecylphosphonic acid	5.57
6:2 Fluorotelomer phosphate diester	7.92
6:2/8:2 Fluorotelomer phosphate diester	8.48
8:2 Fluorotelomer phosphate diester	8.90
Bis(tridecafluorohexyl)phosphinic acid	7.39
(Heptadecafluorooctyl)(tridecafluorohexyl)phosphinic acid	8.07
Bis(heptadecafluorooctyl)phosphinic acid	8.60
Ammonium bis(N-ethyl-2-perfluorooctylsulfonaminoethyl)phosphate	9.35

Figure 7: Direct experimental mapping of known PFAS compounds allows for the long-term goal of building in silico libraries. The precision of this approach is approximately ± 30 s.[4]

References

[1]U.S. Environmental Protection Agency. Comptox Chemicals Dashboard: Master List of PFAS Substances (Version2). Accessed March 11, 2021. https://comptox.epa.gov/dashboard/chemical_lists/pfasmaster

[2]Koelmel, J.P. et al., Toward Comprehensive Per- and Polyfluoroalkyl Substances Annotation Using FluoroMatch Software and Intelligent High-Resolution Tandem Mass Spectrometry Acquisition, *Analytical Chemistry 2020 92 (16), 11186-11194.*

[3]Koelmel, J.P. et al., Interactive Software for Visualization of Non-Targeted Mass Spectrometry Data – FluoroMatch Visualizer, ChemRxiv. Cambridge: Cambridge Open Engage; 2022; This article has been accepted for publication by Exposomics.

[4]Coggan, T.L., et al., Single analytical method for the determination of 50 legacy and emerging per- and polyfluoroalkyl substances (PFAS) in aqueous matrices, Analytical and Bioanalytical Chemistry (2019) 411:3507-3520.

Acknowledgments

The development of the FluoroMatch suite of software tools is in part funded by an Agilent ACT-UR grant. Also, in 2021, Stéphane Bayen has been recognized as an Agilent Thought Leader for his contribution to the understanding of both human and environmental food safety risk assessment. This work is a continuation of that quest. Jeremy Koelmel would like to specifically call out the Bowden group at the University of Florida and the Pollitt group at Yale. Jeremy also thanks SynQuest Labs for providing him PFAS standards. We would finally like to thank the unattributed Agilent, McGill, and Yale team members who made this work possible.

DE08849523

5994-6579EN
July 21, 2023