Towards comprehensive PFAS annotation using FluoroMatch Software and Intelligent LC-HRMS/MS Acquisition Methods

Authors:

Jeremy P. Koelmel[1], Matthew Paige[1], Juan J. Aristizabal-Henao[2], Nicole M. Robey[3], Sara Nason[4], Paul Stelben[1], Yang Li[5], Nicholas M. Kroeger[5], Michael P. Napolitano[6], Tina Savvaides[1], Vasilis Vasiliou[1], Pawel Rostkowski[7], Timothy Garrett[5,8], Elizabeth Lin[1], Chris Deigl[9], Karl Jobst[10], Timothy G. Townsend[3], Krystal J. Godri Pollitt[1]*, John A. Bowden[2,8]*

[1]School of Public Health, Yale University, New Haven, CT, United States
[2]Center for Environmental and Human Toxicology & Department of Physiological Sciences, University of Florida, Gainesville, FL, United States
[3]Department of Environmental Engineering Sciences, University of Florida, Gainesville, FL, United States
[4]Departments of Environmental Sciences and Analytical Chemistry, The Connecticut Agricultural Experiment Station, New Haven, CT, United States
[5]Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL, United States
[6]Independant researcher and consultant, Charleston, SC, United States
[7]Environmental Chemistry Department, Norwegian Institute for Air Research, Kjeller, Norway
[8]Department of Chemistry, University of Florida, Gainesville, FL, United States
[9]SynQuest Laboratories, Alachua, FL, Unites States
[10]Department of Chemistry, Memorial University of Newfoundland, St. John's, NL, Canada

Contact Information:
Co-corresponding authors
Krystal J. Godri Pollitt
krystal.pollitt@yale.edu
John A. Bowden
john.bowden@ufl.edu

Abstract:

Thousands of per- and polyfluoroalkyl substances (PFAS) exist in the environment and pose a potential health hazard. Suspect and non-target screening with liquid chromatography (LC) high-resolution tandem mass spectrometry (HRMS/MS) can be used for comprehensive characterization of PFAS. To date no automated open source PFAS data analysis software exists to mine these extensive datasets. We introduce FluoroMatch, which automates file conversion, chromatographic peak picking, blank feature filtering, PFAS annotation based on precursor and fragment masses, and annotation ranking. The software library currently contains ~7,000 PFAS fragmentation patterns based on rules derived from standards and literature and the software

automates a process for users to add additional compounds. The use of intelligent data-acquisition methods (iterative exclusion) nearly doubled the number of annotations. The software application is demonstrated by characterizing PFAS in landfill leachate as well as in leachate foam generated to concentrate the compounds for remediation purposes. FluoroMatch had wide coverage, returning 27 PFAS annotations for landfill leachate samples, explaining 71% of the all-ion fragmentation $(CF_2)_n$ related fragments. By improving the throughput and coverage of PFAS annotation, FluoroMatch will accelerate the discovery of PFAS posing significant human risk.

Introduction

Over 7,500 per- and polyfluoroalkyl substances (PFAS) are compiled in the US EPA's CompTox Chemistry Dashboard ("The PFAS Master List"),[1,2] yet most targeted PFAS analyses cover fewer than 30 compounds[3–5]. Rarely screened and replacement PFAS may prove to have similar or higher toxicity compared with traditionally-measured PFAS.[2,6–8] Therefore, user-friendly and high-throughput workflows and algorithms for characterizing these PFAS are needed. One major bottleneck is the comprehensive characterization of both known and unknown PFAS in high-resolution tandem mass spectrometry data. Current non-targeted approaches take advantage of the higher mass of fluorine-containing compounds and negative mass defect of fluorine compared to hydrogen to differentiate fluorinated from non-fluorinated compounds.[9–12] In addition, many PFAS compounds can be grouped into classes and sub-classes consisting of per- and polyfluorinated alkyl and alkoxy-based polymers attached to unique chemical moieties. Common repeating structural units include CF_2, CH_2CF_2, and CF_2O. Mass defects calculated using fluorine-containing repeating units (e.g., CF_2) instead of ^{12}C, can be plotted against nominal mass to determine homologous series of PFAS sub-classes.[13–17] The success of this approach as annotation based solely on exact mass hinges on the assumption that most polyfluorinated compounds occupy a region of compositional space[18] that is devoid of other non-halogenated chemicals. An analysis of the PubChem library as well as inventories of high production volume chemicals (e.g., the US Toxic Substances Control Act's Chemical Substance Inventory)[19] show that this assumption is not true. Additionally, the possible formulas for an exact mass containing fluorine atoms are numerous; for example, the exact masses of 142 formulae fall within 1 ppm of that of PFOA¥1. Furthermore, PFAS compounds have a large number of isomers, with 25% of compounds in the EPA "PFAS Master List" having at least one matching PFAS isomer, thereby being indistinguishable by exact mass alone. This underscores the need to use complementary data to confidently filter suspected PFAS from complex mass spectrometric data, such as isotopic ratios which can narrow the list of potential formulae matches[21,22] and (predicted) fragmentation which can provide further structural information.

MS^2 fragments shared across many PFAS compounds (e.g., $[C_2F_5]^-$) can be used to distinguish fluorine-containing compounds from other exact mass matches.[9,14,23]

¥Atomic constraints: C0-100 H0-100 N0-3 O0-10 F0-200 Cl0-17 Br0-17 I0-17 S0-3 P0-3
Algorithm used for formula prediction: MS-FINDER[20]

Fragmentation patterns can also be used to differentiate PFAS with the same or similar exact mass-to-charge ratio (*m/z*) (e.g., $[M\text{-}H_5FCO]^-$ for fluorotelomer alcohols).

We developed FluoroMatch to address the described challenges. As part of the FluoroMatch software package, we introduce FluoroMatch Generator, which generates in-silico fragmentation libraries of perfluorinated and polyfluorinated molecules from a set of fragmentation rules and molecular structures indicating repeating units in order to continue to expand FluoroMatch *in-silico* libraries with the help of the PFAS research community. FluoroMatch can be used to annotate PFAS from data-dependent acquisition (DDA) and data-independent acquisition (DIA) experiments. In the case of DIA, spectral deconvolution is used to reconstruct precursor-fragment relationships. In addition, we employ iterative exclusion (IE) acquisition[24], which generates exclusion lists from ions selected for fragmentation in previous injections, and fragments the next most abundant ions in iterative sample injections. This technique improves fragmentation coverage, especially of low abundance ions, as compared to DDA.[24] We demonstrated the utility of FluoroMatch to characterize the breadth of PFAS molecules in liquid effluent from solid waste landfills (landfill leachate), as well as in foam generated to concentrate the compounds for remediation purposes.

Experimental Section

Library Generation – FluoroMatch Generator

In-silico fragmentation can be generated via: extracting fragment rules for specific classes of compounds where all members of the class share common fragments/neutral losses,[25–27] predicting spectra based on fragmentation mechanisms and/or machine learning approaches,[28–30] or using quantum mechanics to predict fragmentation[31]. In this case, because PFAS have similar fragmentation across a given sub-class, class-based neutral losses and fragment ions were assigned, as this technique requires minimal assumptions and margin for error as fragments are hand-annotated from standards across multiple chain lengths and patterns are easily discerned. FluoroMatch consists of *in-silico* libraries that include over 7,000 species with predicted fragmentation (fragment and precursor formulae and *m/z* values, without intensity) across 72 PFAS sub-classes. See the supplementary files FluoroMatch_Libraries.xlsx and FlouroMatch_Supplemental.pdf for standards and literature used to generate libraries, compound sub-classes and fragmentation rules, and further details on library development. FluoroMatch Generator was developed in R[32] to allow users to expand in-silico libraries using their own standards and annotated PFAS spectra. The workflow for FluoroMatch Generator and examples are both shown in Figure 1 and describe in detail in FluoroMatch_Supplemental.pdf. Figure 1 shows that as PFAS differ in chain lengths, while the intensities of fragments fluctuate, certain main fragments are characteristic of all species. In the case where this assumption was not met (e.g. short-chain PFAS with less than 4 CF_2 units), a separate library was generated.

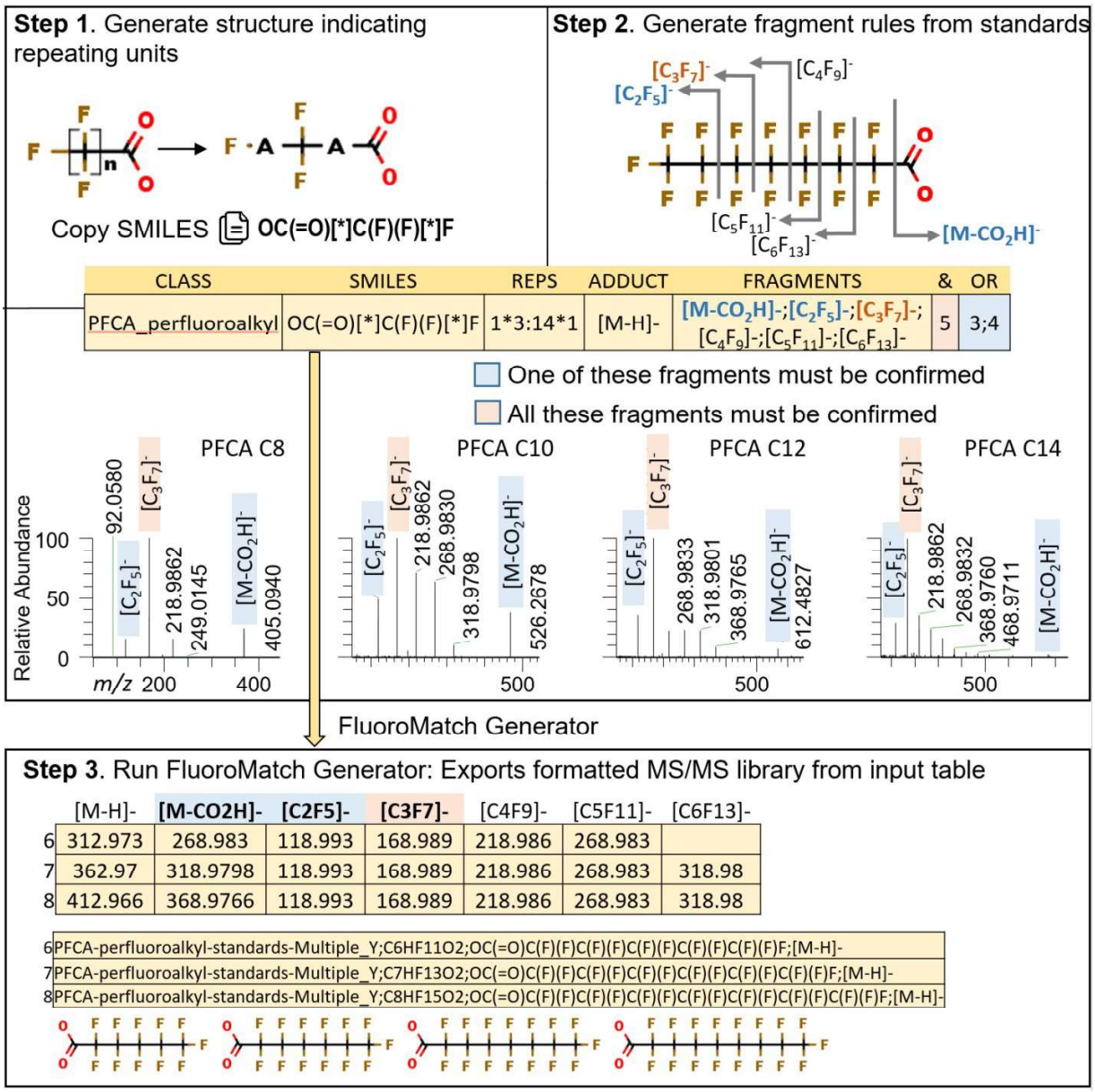


Figure 1: Workflow to generate PFAS in-silico libraries using FluoroMatch Generator. The first step (**Step 1**) is to draw the PFAS structure in any chemical drawing software, indicating repeating units via the letter "A" on each side of the repeating unit. Then, the user right clicks and copies the SMILES structure from the drawing, pastes it into the correctly formatted input table and indicates the number of possible repeating units in the adjacent column. In **Step 2**, the user generates fragmentation rules from accurately annotated spectra, ideally with three or more species spectra per sub-class spanning the allowable number of repeating units to generalize fragment rules (column 4, 5, 6, and 7 in the topmost table). The table is imported in FluoroMatch Generator in **step 3** and the software outputs molecular information and fragment masses to be used as libraries for FluoroMatch. This process employing FluoroMatch Generator was used to generate the current libraries included with FluoroMatch using both standards and literature.

Non-targeted and suspect screening workflow – Automated and Modular Modules

FluoroMatch can be applied to several acquisition workflows. Acquisition requires, at minimum, files containing MS/MS data and at least 10 full scans across chromatographic peaks for good integration. We recommend applying iterative exclusion data-dependent analysis (IE-DDA) to obtain the most coverage, at least for pooled samples (as this requires multiple injections of a single sample). Traditional data-dependent workflows and targeted MS/MS acquisition to increase coverage of known PFAS are also supported. All-ion fragmentation (AIF; also referred to as MS^E) can also be acquired and deconvoluted using the software (currently only supported for Thermo) using algorithms previously developed for lipidomics.[25]

FluoroMatch uses libraries generated according to the section above and modified software previously developed for lipidomics.[25] Two versions of FluoroMatch have been developed: a standalone R script that can be used in a modular workflow (FluoroMatch Modular) and FluoroMatch Flow, which covers many aspects of the non-targeted and suspect screening PFAS workflow including file conversion using msConvert,[33] a unique untargeted chromatographic peak picking strategy implementing MZmine 2.26,[34] and blank feature filtering (BFF).[35] Both versions output annotations using exact mass and fragment masses, rankings of multiple annotations for features, and compilations of meta-data on fragmentation information and peaks used to annotate features (Figure 2 and FluoroMatch_FormattedOutput.xlsx). It is important to note that blank filtering is an important step employed to remove features which are from sample collection, processing, and acquisition, and not inherent to the sample itself. This is an important step in non-targeted PFAS analysis,[36] especially given the various sources of PFAS cross-contamination that can occur.[37] Note that for BFF to work best, blanks must be field blanks: they must have gone through the same process of sample collection, transport, extraction, and acquisition, as samples. The BFF method employed uses a stringent filter (Equation 1), and therefore samples should be chosen with the highest levels of all PFAS as the reference samples. More details and algorithms behind FluoroMatch are provided in the Supplementary Information and in-depth manuals are provided with the software.

To use FluoroMatch Flow, users drag vendor files (no conversion necessary) onto the software interface and click run after choosing an export directory. Current vendor formats supported by FluoroMatch Flow include .d (Agilent), .raw (Thermo), and .wiff/.wiff2 (SCIEX). FluoroMatch Modular can be used to annotate feature tables generated by any approach (for example, XCMS or vendor software such as Compound Discoverer), and hence, annotations can be appended after prior non-targeted steps including selecting homologue features using mass defect plots.[13–17] The modular version supports any vendor; specific methods and conversion parameters are provided in the manual for Waters and Bruker's files, along with the previous three vendors mentioned.

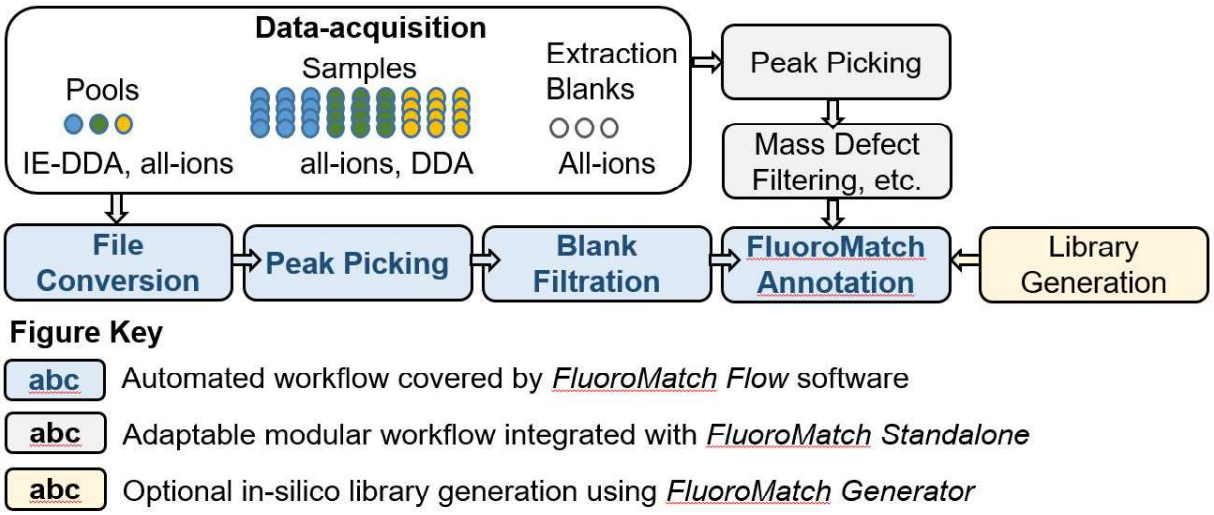


Figure 2: Workflow to use FluoroMatch software. FluoroMatch consists of FluoroMatch Flow (covering the entire workflow from file conversion to annotation), FluoroMatch Modular (covering only annotation, which can be integrated with upstream data-processing workflows), and FluoroMatch Generator (for developing new PFAS in-silico libraries). In blue is a workflow using FluoroMatch Flow; the user simply specifies which files are samples for chromatographic peak picking, which are MS/MS files for identification, and which samples are blanks, and then clicks run (a single sample can be used for multiple categories). Chromatographic peak picking is performed in an untargeted fashion on a subset of samples (e.g., pools), and the resulting peak list is used to target peaks across all samples. This two-step process improves throughput and the ability to handle large sample sizes by reducing the amount of samples which undergo the computationally expensive untargeted peak picking step. In grey is a modular workflow where the user can use their own peak picking algorithms and processing steps and use FluoroMatch Modular only for annotation.

Leachate Collection and Data-Acquisition

The air-water interfacial partitioning tendencies of some PFAS species suggests that increasing the surface area of a PFAS-contaminated liquid may be an effective way to concentrate PFAS in a volume-reduced waste stream. In a laboratory experiment, sintered glass aquarium air stones were used to bubble air and produce substantial foaming (i.e., high air-water interfacial area) in samples of landfill leachate collected from an active MSW landfill in central Florida, US. The foam was collected using a stainless-steel mesh skimmer and allowed to coalesce back to a liquid form. Unlike destructive technologies, such as sonication[38] which also relies upon the surfactant qualities of certain PFAS, aeration foaming is a low-energy sequestration technique that is shown to be effective across most of the PFAS which were analyzed using a targeted approach. Therefore, the highly concentrated coalesced foam samples as well as untreated leachate for comparison were of interest for FluoroMatch application.

Leachate, foam, and over 100 standards were acquired on a Thermo Vanquish UHPLC system coupled to a Thermo Q-Exactive Orbitrap mass spectrometer using a Phenomenex Gemini column (C18 with TMS endcapping, 110 Å, 100 mm × 2.1 mm × 3.0 µm) with the mobile phase consisting of A: 100% water and B: 100% methanol, both with 5 mM ammonium acetate. For validation of FluoroMatch annotations, 25 labeled standards (^{13}C or ^{2}H labeled) were spiked into leachate samples acquired using the Q-

Exactive LC-HRMS approach (Table S2). For further validation and comparison to targeted approaches, a targeted QqQ approach was used across 51 standards (Table S1) on a Thermo Quantis triple-quadrupole mass spectrometer using the same UHPLC settings. Further acquisition details (scanning events, gradient, etc.) are in the Supplemental Information.

Results and Discussion

Feature Filtering and Annotation

We tested feature filtering and annotation in FluoroMatch by analyzing LC/HRMS data acquired from six leachate samples (three raw leachate samples and three samples of coalesced foam generated from the leachate via bubble aeration) from a municipal solid waste landfill in Florida. Of the 73,718 features determined using MZmine, 21,551 of the integrated peak areas were above the background threshold using blank feature filtering (BFF). The BFF filter was set as:

Equation 1: $S_{Q1} > 5 \times (\overline{B} + (3 \times B_{\sigma}))$

Where S_{Q1} = first quartile of samples, $\overline{B}$ = blank average, and B_{σ} = blank standard deviation

Therefore, 71% of features were likely background contamination, emphasizing the importance of blank feature filtering.

Further filtering was manually performed using mass defect filtering to determine the fraction of chemicals likely containing fluorines (negative mass defect). With a mass defect from −0.25 to 0.1, 17,911 features were found, thereby removing 76% of features in combination with the BFF approach. FluoroMatch Flow performs an exact mass search against the EPA PFAS Master List (with over 7,500 PFAS compounds). Using FluoroMatch Flow 1,286 features had exact mass hits, or 7% of the filtered features. Of these, FluoroMatch returned 27 matches based on MS/MS matching of the experimental data and PFAS *in-silico* MS/MS generated using FluoroMatch Generator (0.2% of filtered features). These annotations have sufficient evidence to be level 2a identifications (probable structure) using the schema from Schymanski et al.[39–41], having exact mass and literature/standard based MS/MS matching, except for the fact that the fragmentation cannot always be used to discern between certain PFAS isomers (e.g., branched versus linear chain). Hence, identifications are Level 3 (tentative); for example, both branched and linear chain perfluorocarboxylic acids are reported for the same feature by FluoroMatch.

Validation of FluoroMatch Annotations

Validation showed no false positives for compounds where standards were available. The annotations were validated against labeled standards spiked into samples using the Q-Exactive non-targeted approach. Labeled standards were observed for 16 of the 26 unique chemical structures annotated by FluoroMatch (supplemental excel). All 16 chemicals were confirmed (level 1 annotation – matching retention time, MS/MS, and exact mass)[41] using the labeled standards. A wider breadth of chemicals were targeted for comparison to a triple quadrupole (QqQ) approach screening 52 PFAS compounds

using the same chromatography (supplemental Excel). Of the 32 compounds confirmed above the limit of detection using the QqQ approach (level 1 annotation)[41], 23 of these had chromatograms above the noise threshold using the Q-Exactive (the remaining showed no signal, or less than 10 scans across the peak). This suggests that selected reaction monitoring (SRM) scans using the QqQ are more sensitive for the detection of low abundance compounds than non-targeted approaches employing the Q-Exactive. Note that targeted t-SIM or similar experiments employing the Q-Exactive may be as or more sensitive. Of the 23 detected using the targeted approach, 20 were annotated using FluoroMatch, 15 by MS/MS (validating 15 of the 26 FluoroMatch annotations) and five only by exact mass. Of the remaining three, one was removed using BFF, while the remaining two were not picked up by the MZmine algorithms embedded in FluoroMatch Flow. Hence, no false positives were observed using this second confirmatory technique.

Twelve compounds were annotated by FluoroMatch but not contained in the extensive targeted approach (52 compounds), showing the potential of FluoroMatch to enhance PFAS coverage compared to extensive targeted approaches. Unique annotations included pentafluoropropionic acid (PFPrA); ethyl trifluoromethanesulfonate; N-methyl perfluoro sulfonamido acetic acid: n=3 (MeFPrSAA), n=4 (MeFBSAA), n=5 (MeFPeSAA), and n=6 (MeFHxSAA); n-ethyl-N-tridecafluorohexyl sulfonyl glycine; or related isomers. While exact mass and multiple PFAS specific fragments were assigned by FlouroMatch for these species (see supplemental excel for fragments annotated), further validation with standards would be necessary for confirmation of the exact molecular structure.

While targeted approaches and labeled standards spike into samples can be used to validate known PFAS, other non-targeted approaches are needed to validate PFAS for which standards do not exist and to increase PFAS coverage. Mass defect plots can be used to determine homologous series. In this case, the masses were normalized to CF_2 (Figure 3). Most PFAS with polar head groups and CF_2 chains had a normalized (CF_2) mass defect between −0.03 and 0.05. The normalized mass defect plot for this range can be seen in Figure 3. Forty-eight PFAS compounds were found across 11 homologous series based on the same normalized mass defect within 0.003 Da for a series and mass differences divisible by 49.9968 (CF_2). These non-targeted approaches can be used as a benchmark for the performance of FluoroMatch. *Using current in-silico libraries, FluoroMatch was able to annotate the structure of 56% of the features that followed homologous series (Figure 3)*. Additionally, it is important to note that MS/MS is required for high-confidence annotation: for example, MS/MS was manually examined for 669.9611, and while the species followed the homologues series, fragmentation did not confer evidence that the compound belonged to the N-methylperfluorooctane sulfonamidoacetic acid class, nor PFAS with CF_2 repeating units in general. Hence, coverage maybe be higher than 56%, with FluoroMatch removing false positives by necessitating fragmentation information for annotation.

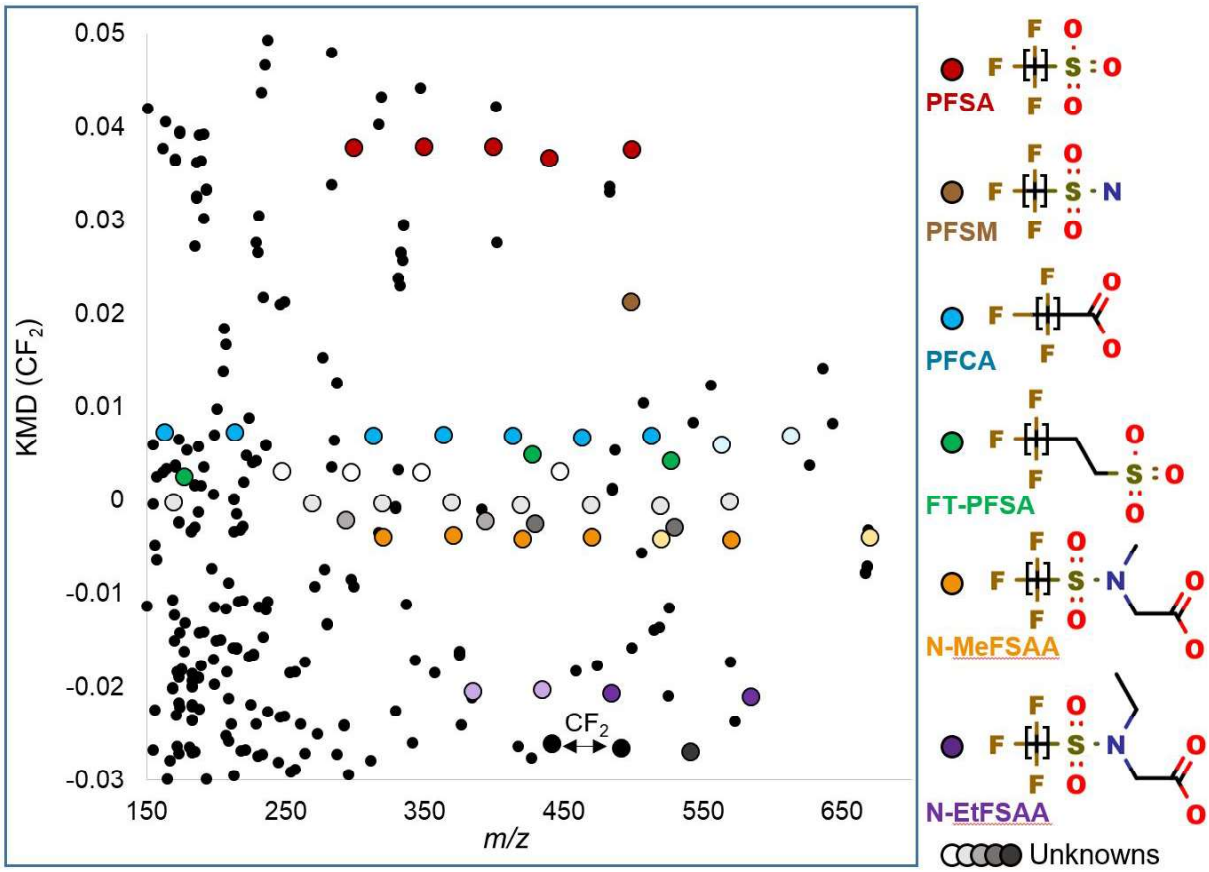


Figure 3: Mass defect plot with masses (MS1) normalized to CF_2 (manually generated) showing PFAS homologue series and associated FluoroMatch coverage across these series. Homologous series are identified by shifts in the x-axis divisible by 49.9968 (CF_2) and the same CF_2 normalized mass defect (within 0.003 normalized mass units in this case). Colored markers are those that were identified by FluoroMatch by MS/MS (either AIF or IE-ddMS^2-topN). Lighter shaded colors for a given series are those that follow the homologues series but were not identified by FluoroMatch. Greyscaled larger markers are those that follow homologues series but had no identifications from FluoroMatch. In the plot, 283 features are shown, 48 of which followed homologues series, and 23 of which were identified by MS/MS using FluoroMatch. The names of the compounds from top to bottom are as follows: perfluorosulfonic acid (PFSA), perfluorosulfonamide (PFSM), perfluorocarboxylic acid (PFCA), fluorotelomer perfluorosulfonic acid (FT-PFSA), N-methylperfluorooctane sulfonamidoacetic acid (N-MeFSAA), and N-ethylperfluorooctane sulfonamidoacetic acid (N-EtFSAA). Note that for PFSM only one species existed, but as with all darker colored series, MS/MS evidence was used to annotate the compound.

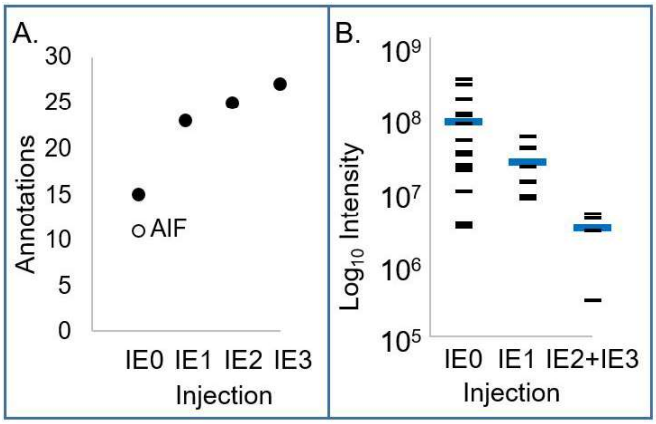
Iterative-Exclusion, Data-Dependent Analysis, and Data-Independent Analysis

In order to improve confidence in PFAS annotation, it is essential to have information orthogonal to the exact mass of the precursor. The accurate mass of a precursor ion can provide a restricted number of potential molecular formulas depending on the mass accuracy of the MS detector,[42] whereas tandem mass spectrometry can provide evidence of certain structural components within a molecule. Because of the low abundance and high diversity of PFAS species in some samples, it is both difficult and essential to improve fragmentation coverage in comparison to traditional MS/MS

techniques. For this purpose, samples were analyzed using three different acquisition methods: data-dependent top N analysis (ddMS^2-topN), iterative exclusion top N analysis (IE-ddMS^2-topN), and data-independent analysis *viz.* all-ion fragmentation (AIF). DDA and IE use ion selection (1-Da isolation window in this work) prior to fragmentation, thereby providing explicit precursor fragment relationships assuming no other ions within 1 Da co-elute. DDA and IE have limited coverage, even with IE-ddMS^2-topN, where ions selected in a previous injection are excluded and the next top-most abundant ions are fragmented iteratively, fragmentation for all ions are not generally obtained. In addition, if isomers co-elute, the limited MS/MS scans can lead to mixed spectra (contributed from both isomers) that cannot be deconvoluted. Data-independent approaches, in this case AIF, obtains fragmentation across all species and has enough scans to deconvolute fragmentation from closely eluting isomers, but fragment-precursor relationships are lost. Therefore, the techniques are complimentary.

When the three acquisition methods were used separately, 11 PFAS from the leachate samples (both unadulterated and foam) were annotated using AIF, 15 using ddMS^2-topN, and 27 using IE. Figure 4A shows the increase in annotations using IE, showing an 80% improvement in coverage by applying IE to the traditional data-dependent approach. This advantage is due to the increase in lower abundance features with MS/MS (Figure 4B) and is similar to previous findings in lipidomics[24]. The low number of annotations produced using AIF is due to the need for correlation between precursor ion and fragment ion intensities for reconstructing precursor-fragment relationships: because common PFAS fragments (eg., $[C_2F_5]^-$) may co-elute, the R^2 between some precursors and fragments will be below the FluoroMatch threshold of 0.6, and the compound will not be identified. While lower in number, the AIF annotations increased confidence in compound determination; all AIF annotations were replicated in the IE analysis validating the deconvolution algorithm.

Figure 4: Annotation coverage using iterative exclusion data-dependent analysis (IE-ddMS^2-topN).



Iterative injections, shown with black filled markers (**A**), exhibit an increase in the cumulative number of PFAS annotations as a function of injection number applying IE sequentially, which levels off at the fourth injection (IE3). The number of annotations for AIF (white marker) is shown for comparison to the first data-dependent scan (**A**). In the rigt panel (**B**) IE both increased the number of PFAS annotations, and specifically increased annotation of lower abundance compounds. The blue rectangle represents the mean; all differences were significant based on multiple t-tests (p-value < 0.05).

While AIF using FluoroMatch deconvolution algorithms underperformed in terms of total annotations, AIF data can be used to estimate the number of PFAS compounds consisting of CF_2 repeating units and assess the coverage of FluoroMatch annotations. The most common fragments observed for CF_2 repeating units in LC/MS are $[C_2F_5]^-$, $[C_3F_7]^-$, $[C_4F_9]^-$, $[C_5F_{11}]^-$, and $[C_6F_{13}]^-$. Fragment chromatographic traces from AIF

acquisition of leachate samples are shown in Figure 5A. Twenty chromatographic peaks can be estimated based on fragment traces (Figure 5A), suggesting 20 or more PFAS compounds. It is important to note that when overlaying precursor ion traces with fragment traces, different precursor ions often co-eluted, leading to each fragment trace representing one *or more* PFAS compounds (Figure 5B). After overlaying precursors identified in FluoroMatch annotations with fragment traces, 71% (15/21) of fragment traces were accounted for and the remaining traces were generally of low abundance (Figure 5A, Figure 5B, and Figure S2). This suggests that FluoroMatch was able to annotate a large portion of the CF_2-containing PFAS compounds, especially the dominant species.

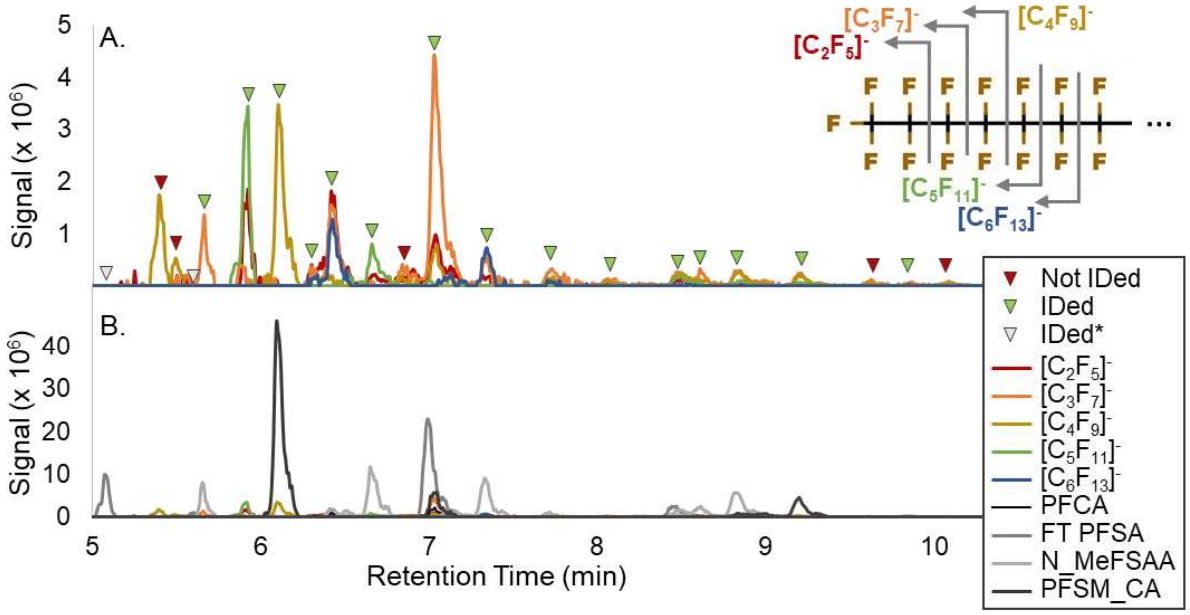


Figure 5: Fragment and precursor traces provide an estimate of PFAS coverage by FluoroMatch. All-ion fragmentation data of fragment ions of the carbon-fluorine chain (**A**) and full-scan data of precursor ions (colored by class) identified by FluoroMatch (**B**). Triangles in (**A**) indicate whether a fragment trace had one or more precursors that matched based on the elution profiles. MS/MS identifications using FluoroMatch Flow were able to explain 15/21 (71%) of fragment traces. Fragment traces for repeating CF_2 units with ether linkages are shown in Figure S2.
*Grey triangles represent peaks that were identified but did not have carbon-fluorine chain fragments in the top panel (**A**).

While 27 species were annotated by MS/MS and FluoroMatch, 23 had unique structures, meaning that the remaining four were likely different isomers with subtle structural differences. An example is shown for an N-methylperfluorooctane sulfonamidoacetic acid (N-MeFSAA) species with a CF_2 chain length of eight in Figure 6. In this case, three resolved peaks were determined, with the second and third species in terms of elution order identified by both AIF and IE analysis (Figure 6B and Figure 6C), showing high confidence that annotation were not false positives from trace fragments from co-eluting species. The first species had AIF fragmentation shifted to slightly earlier retention times (Figure 6B) and had a fragmentation profile slightly different than the second and third peak (Figure 6C), although the fragmentation

provides ample evidence of similar structural motifs. In this case, the major plausible possibility is that species could differ in terms of chain branching, with branched isomers eluting prior to straight-chain PFAS[43], indeed, the latest eluting species and dominant peak (Figure 6) was confirmed as the non-branched form using a deuterated standard (supplemental validation excel file). Future studies might distinguish branched isomers using MS^3 approaches.[44]

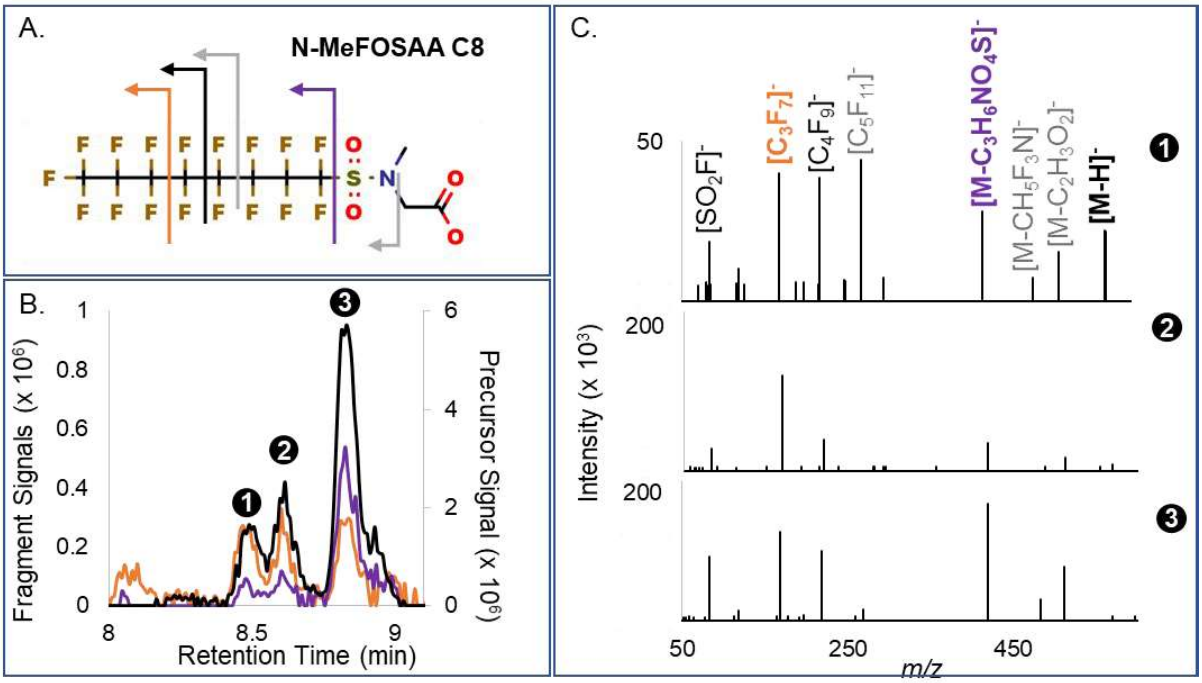


Figure 6: Three PFAS isomers all showing MS/MS evidence as N-methylperfluorooctane sulfonamidoacetic acid (N-MeFSAA). **A**: The structure for N-MeFOSAA with eight carbons and some common fragmentations. Necessary fragments for annotation are colored, whereas other library fragments are in black, and fragments which were not contained in FluoroMatch libraries in grey. **B**: All-ion fragmentation (AIF) data and precursor full-scan data, with fragment profiles highly correlated with the precursor ion trace for the second and third peak, but less correlated for the first peak (possibly due to fragment overlap). **C**: Data-dependent fragment scans at retention time 8.51, 8.64, and 8.82. Fragmentation patterns give evidence that all three isomers are highly similar in structure, and only subtle structural differences (e.g., branching) are responsible for the differences in retention time.

Effect of Mass Accuracy on Confidence in PFAS Annotations

High-resolution mass spectrometers across vendors differ in mass accuracy (e.g. Q-TOF, orbitrap, and FTICR mass spectrometers), with mass accuracy arguably one of the most important determinants in MS/MS confidence, especially given that FluoroMatch does not use intensity profiles. To test the effect of mass accuracy on false positive rate in annotations, the mass accuracy of the software was toggled across a tolerance of ± 5ppm, 10ppm, 20ppm, and 30ppm. Any additional annotations with lower mass accuracy are false positives, as the exact mass at 5ppm was incorrect.

Toggling the mass accuracy parameter of the FluoroMatch software showed a slight increase in false positives for confident annotations (Figure S3). There was no

difference in confident annotations (rule based annotation) between ± 5ppm and ± 10ppm mass tolerance. For ± 20ppm, only 1 false positive was obtained (4% false positive rate), whereas for ± 30ppm 2 false positives were obtained (8% false positive rate). For lower confidence annotations a higher percentage of false positives were obtained with decreasing mass accuracy, due to the less stringent criteria (any fragment and exact mass needed for confirmation). While only 1 additional false positive was obtained at ± 10ppm (3%), at ± 30ppm 23% (9/39) false positives were observed. This highlights the importance of mass accuracy for accurate annotation of PFAS, noting that this will become even more essential as libraries expand and the chances of false annotations increases for this reason.

Evaluations of FluoroMatch Identified PFAS for Municipal Solid Waste Leachate Characterization and Remediation

FluoroMatch was applied to unadulterated leachate and leachate foam generated through bubble aeration. Further in-depth discussion of findings is provided in the supplemental materials. Based on a one-sided t-test (p-value < 0.05), 19 of the 27 PFAS compounds were significantly higher in the foam than unadulterated leachate by an average factor of 5. Partitioning of PFAS between the foam and the unadulterated leachate was modeled against predicted water solubility (mol/L) of the PFAS molecules[45] using an exponential trend (R^2 = 0.70) (Figure S1A). While short-chain PFAS species were found to have a lower concentration in the foam fraction, these species were characterized by a decreased potential for bioaccumulation and ecological[46] and mammalian toxicity[47] (Figure S1B).

The detection of PFSM, FT-PFSA, N-MeFSAA, and N-EtFSAA species in landfill leachate suggest that these species should be further studied as they are recycled back into the environment without full conversion to PFCA and PFSA from waste water treatment plants. This is especially urgent given that certain species may be significantly more toxic than PFCA and PFSA. For example, evidence shows that PFSM is a more potent neurotoxin compared to other PFAS[48,49].

Conclusion

FluoroMatch automates PFAS annotation using in-silico PFAS fragmentation libraries and rule-based annotation. We introduce in-silico fragmentation libraries containing over 7000 PFAS across 72 PFAS sub-classes, built using spectra from literature and authentic standards. To further expand libraries in the future, we include FluoroMatch Generator, for users to generate in-silico PFAS libraries using fragment annotations and SMILES structures for representative compounds. This software workflow falls between suspect screening and non-targeted approaches: a wide range of PFAS are screened, but evidence for each PFAS is stronger than in traditional non-targeted approaches as each library is based on experimental fragmentation.

Validating the percent coverage and accuracy of annotations in real-world samples is challenging due to the case of known-unknowns and unknown-unknowns. Here, we use 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 (remaining being false negatives). Here, we use all-ion fragmentation to estimate that FluoroMatch covered 71% of CF2 containing PFAS compounds with fragmentation, and CF2 normalized mass defect plots to estimate 56% coverage of compounds (remaining being false negatives). Furthermore, using both a targeted method (LC-QqQ) and spiked in internal standards (LC-HRMS), 100% of FluoroMatch annotated features with corresponding standards were confirmed suggesting a low false positive rate. Based on application to MSW leachate, the most abundant species were annotated using FluoroMatch, and the use of intelligent data-acquisition, specifically iterative exclusion, nearly doubled the number of PFAS annotated, especially for those of low abundance.

Further work on isotopic pattern scoring, fragment prediction where no standards or experimental spectra exists, and automating the use of CF_2 normalized mass defect plots are being implemented to continually improve the coverage and accuracy of our FluoroMatch software platform. In addition, the research community stands to benefit from retention time index models for PFAS, fully automated fragmentation generation (for example improvement of the CFM-ID algorithm[50] by increasing coverage of PFAS in the training set), and improved deconvolution algorithms for AIF. As users continue to validate and expand the open source software, especially in terms of modifiable databases, we expect FluoroMatch to continue to improve as a rapid, automated, and comprehensive tool for researchers interested in the environmental and clinical consequences, as well as industrial applications, of PFAS molecules.

Acknowledgements

This work was in part supported by the EPA STAR grant (P0134825, AWD06571). The researchers (NMR, JAB, TGT, JJAH) have received support from the U.S. Environmental Protection Agency under the Science To Achieve Results (STAR) grant program (EPA-G2018-STAR-B1; Grant#: 83962001-0) to investigate the fate and transport of PFAS in municipal solid waste landfills. Researcher SLN received support from USDA NIFA Hatch funds (CONH00789). A large array of standards used to develop libraries were donated by SynQuest Labs, Inc. JAB would like to acknowledge Startup funds from the UF College of Veterinary Medicine which were used to support this work. We would also like to thank Alina Timshina, Atiye Ahmadireskety, Bianca Ferreira Da Silva, Brian Martinez, Emily Griffin, Gina Del Pozo, Heather Ditz, and Maggie Zawoy for annotating PFAS MS/MS spectra for library development.

Supporting Information:

Data files (leachate and blanks):
ftp://massive.ucsd.edu/MSV000085244/raw/Leachate/Leachate/

FluoroMatch Software (FluoroMatch Flow, Modular, and Generator):
http://innovativeomics.com/software/fluoromatch-flow-covers-entire-pfas-workflow/
Note: download the zip. The manuals for FluoroMatch software are contained within the zip file.

More in-depth discussions of algorithms and comparison of leachate foam and unadulterated leachate. The following are included: Table S1, Figure S1 and Figure S2.
FluoroMatch_Supplemental.docx

Output from FluoroMatch Software – formatted and non-formatted version with description:
FluoroMatch_FormattedOutput.xlsx

Current list of Names, SMILES, and fragment rules used to develop FluoroMatch libraries:
FluoroMatch_Libraries.xlsx

A comparison of results using a targeted QqQ approach versus the non-targeted/suspect screeing approach with FluoroMatch annotation:
Validation_Targeted_vs_FluoroMatch.xlsx

References

(1) Chemistry Dashboard | PFASMASTER Chemicals https://comptox.epa.gov/dashboard/chemical_lists/pfasmaster (accessed Feb 2, 2020).

(2) Patlewicz Grace; Richard Ann M.; Williams Antony J.; Grulke Christopher M.; Sams Reeder; Lambert Jason; Noyes Pamela D.; DeVito Michael J.; Hines Ronald N.; Strynar Mark; Guiseppi-Elie Annette; Thomas Russell S. A Chemical Category-Based Prioritization Approach for Selecting 75 Per- and Polyfluoroalkyl Substances (PFAS) for Tiered Toxicity and Toxicokinetic Testing. *Environmental Health Perspectives* **2019**, *127* (1), 014501. https://doi.org/10.1289/EHP4555.

(3) Yoo, H.; Washington, J. W.; Jenkins, T. M.; Ellington, J. J. Quantitative Determination of Perfluorochemicals and Fluorotelomer Alcohols in Plants from Biosolid-Amended Fields Using LC/MS/MS and GC/MS. *Environ. Sci. Technol.* **2011**, *45* (19), 7985–7990. https://doi.org/10.1021/es102972m.

(4) Mulabagal, V.; Liu, L.; Qi, J.; Wilson, C.; Hayworth, J. S. A Rapid UHPLC-MS/MS Method for Simultaneous Quantitation of 23 Perfluoroalkyl Substances (PFAS) in Estuarine Water. *Talanta* **2018**, *190*, 95–102. https://doi.org/10.1016/j.talanta.2018.07.053.

(5) Munoz, G.; Vo Duy, S.; Budzinski, H.; Labadie, P.; Liu, J.; Sauvé, S. Quantitative Analysis of Poly- and Perfluoroalkyl Compounds in Water Matrices Using High Resolution Mass Spectrometry: Optimization for a Laser Diode Thermal Desorption Method. *Analytica Chimica Acta* **2015**, *881*, 98–106. https://doi.org/10.1016/j.aca.2015.04.015.

(6) Gomis, M. I.; Vestergren, R.; Borg, D.; Cousins, I. T. Comparing the Toxic Potency in Vivo of Long-Chain Perfluoroalkyl Acids and Fluorinated Alternatives. *Environment International* **2018**, *113*, 1–9. https://doi.org/10.1016/j.envint.2018.01.011.

(7) Conley Justin M.; Lambright Christy S.; Evans Nicola; Strynar Mark J.; McCord James; McIntyre Barry S.; Travlos Gregory S.; Cardon Mary C.; Medlock-Kakaley Elizabeth; Hartig Phillip C.; Wilson Vickie S.; Gray L. Earl. Adverse Maternal, Fetal, and Postnatal Effects of Hexafluoropropylene Oxide Dimer Acid (GenX) from Oral Gestational Exposure in Sprague-Dawley Rats. *Environmental Health Perspectives 127* (3), 037008. https://doi.org/10.1289/EHP4372.

(8) Sheng, N.; Cui, R.; Wang, J.; Guo, Y.; Wang, J.; Dai, J. Cytotoxicity of Novel Fluorinated Alternatives to Long-Chain Perfluoroalkyl Substances to Human Liver Cell Line and Their Binding Capacity to Human Liver Fatty Acid Binding Protein. *Arch Toxicol* **2018**, *92* (1), 359–369. https://doi.org/10.1007/s00204-017-2055-1.

(9) Baduel, C.; Mueller, J. F.; Rotander, A.; Corfield, J.; Gomez-Ramos, M.-J. Discovery of Novel Per- and Polyfluoroalkyl Substances (PFASs) at a Fire Fighting Training Ground and Preliminary Investigation of Their Fate and Mobility. *Chemosphere* **2017**, *185*, 1030–1038. https://doi.org/10.1016/j.chemosphere.2017.06.096.

(10) Barzen-Hanson, K. A.; Roberts, S. C.; Choyke, S.; Oetjen, K.; McAlees, A.; Riddell, N.; McCrindle, R.; Ferguson, P. L.; Higgins, C. P.; Field, J. A. Discovery of 40 Classes of Per- and Polyfluoroalkyl Substances in Historical Aqueous Film-Forming Foams (AFFFs) and AFFF-Impacted Groundwater. *Environ. Sci. Technol.* **2017**, *51* (4), 2047–2057. https://doi.org/10.1021/acs.est.6b05843.

(11) D'Agostino, L. A.; Mabury, S. A. Identification of Novel Fluorinated Surfactants in Aqueous Film Forming Foams and Commercial Surfactant Concentrates. *Environ. Sci. Technol.* **2014**, *48* (1), 121–129. https://doi.org/10.1021/es403729e.

(12) Place, B. J.; Field, J. A. Identification of Novel Fluorochemicals in Aqueous Film-Forming Foams Used by the US Military. *Environ. Sci. Technol.* **2012**, *46* (13), 7120–7127. https://doi.org/10.1021/es301465n.
(13) Myers, A. L.; Jobst, K. J.; Mabury, S. A.; Reiner, E. J. Using Mass Defect Plots as a Discovery Tool to Identify Novel Fluoropolymer Thermal Decomposition Products. *Journal of Mass Spectrometry* **2014**, *49* (4), 291–296. https://doi.org/10.1002/jms.3340.
(14) Yu, N.; Guo, H.; Yang, J.; Jin, L.; Wang, X.; Shi, W.; Zhang, X.; Yu, H.; Wei, S. Non-Target and Suspect Screening of Per- and Polyfluoroalkyl Substances in Airborne Particulate Matter in China. *Environ. Sci. Technol.* **2018**, *52* (15), 8205–8214. https://doi.org/10.1021/acs.est.8b02492.
(15) Wang, Y.; Yu, N.; Zhu, X.; Guo, H.; Jiang, J.; Wang, X.; Shi, W.; Wu, J.; Yu, H.; Wei, S. Suspect and Nontarget Screening of Per- and Polyfluoroalkyl Substances in Wastewater from a Fluorochemical Manufacturing Park. *Environ. Sci. Technol.* **2018**, *52* (19), 11007–11016. https://doi.org/10.1021/acs.est.8b03030.
(16) Newton, S.; McMahen, R.; Stoeckel, J. A.; Chislock, M.; Lindstrom, A.; Strynar, M. Novel Polyfluorinated Compounds Identified Using High Resolution Mass Spectrometry Downstream of Manufacturing Facilities near Decatur, Alabama. *Environ. Sci. Technol.* **2017**, *51* (3), 1544–1552. https://doi.org/10.1021/acs.est.6b05330.
(17) Liu, Y.; Pereira, A. D. S.; Martin, J. W. Discovery of C5–C17 Poly- and Perfluoroalkyl Substances in Water by In-Line SPE-HPLC-Orbitrap with In-Source Fragmentation Flagging. *Anal. Chem.* **2015**, *87* (8), 4260–4268. https://doi.org/10.1021/acs.analchem.5b00039.
(18) Zhang, X.; Di Lorenzo, R. A.; Helm, P. A.; Reiner, E. J.; Howard, P. H.; Muir, D. C. G.; Sled, J. G.; Jobst, K. J. Compositional Space: A Guide for Environmental Chemists on the Identification of Persistent and Bioaccumulative Organics Using Mass Spectrometry. *Environment International* **2019**, *132*, 104808. https://doi.org/10.1016/j.envint.2019.05.002.
(19) US EPA, O. TSCA Chemical Substance Inventory https://www.epa.gov/tsca-inventory (accessed Apr 8, 2020).
(20) Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihira, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn, O.; Arita, M. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. *Anal. Chem.* **2016**, *88* (16), 7946–7958. https://doi.org/10.1021/acs.analchem.6b00770.
(21) Meusel, M.; Hufsky, F.; Panter, F.; Krug, D.; Müller, R.; Böcker, S. Predicting the Presence of Uncommon Elements in Unknown Biomolecules from Isotope Patterns. *Anal. Chem.* **2016**, *88* (15), 7556–7566. https://doi.org/10.1021/acs.analchem.6b01015.
(22) Treutler, H.; Neumann, S. Prediction, Detection, and Validation of Isotope Clusters in Mass Spectrometry Data. *Metabolites* **2016**, *6* (4). https://doi.org/10.3390/metabo6040037.
(23) Rotander, A.; Kärrman, A.; Toms, L.-M. L.; Kay, M.; Mueller, J. F.; Gómez Ramos, M. J. Novel Fluorinated Surfactants Tentatively Identified in Firefighters Using Liquid Chromatography Quadrupole Time-of-Flight Tandem Mass Spectrometry

and a Case-Control Approach. *Environ. Sci. Technol.* **2015**, *49* (4), 2434–2442. https://doi.org/10.1021/es503653n.

(24) Koelmel, J. P.; Kroeger, N. M.; Gill, E. L.; Ulmer, C. Z.; Bowden, J. A.; Patterson, R. E.; Yost, R. A.; Garrett, T. J. Expanding Lipidome Coverage Using LC-MS/MS Data-Dependent Acquisition with Automated Exclusion List Generation. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (5), 908–917. https://doi.org/10.1007/s13361-017-1608-0.

(25) Koelmel, J. P.; Kroeger, N. M.; Ulmer, C. Z.; Bowden, J. A.; Patterson, R. E.; Cochran, J. A.; Beecher, C. W. W.; Garrett, T. J.; Yost, R. A. LipidMatch: An Automated Workflow for Rule-Based Lipid Identification Using Untargeted High-Resolution Tandem Mass Spectrometry Data. *BMC Bioinformatics* **2017**, *18* (1), 331. https://doi.org/10.1186/s12859-017-1744-3.

(26) Hutchins, P. D.; Russell, J. D.; Coon, J. J. Mapping Lipid Fragmentation for Tailored Mass Spectral Libraries. *J Am Soc Mass Spectrom* **2019**, *30* (4), 659–668. https://doi.org/10.1007/s13361-018-02125-y.

(27) Kind, T.; Liu, K.-H.; Yup Lee, D.; DeFelice, B.; Meissen, J. K.; Fiehn, O. LipidBlast - in-Silico Tandem Mass Spectrometry Database for Lipid Identification. *Nat Methods* **2013**, *10* (8), 755–758. https://doi.org/10.1038/nmeth.2551.

(28) Djoumbou-Feunang, Y.; Pon, A.; Karu, N.; Zheng, J.; Li, C.; Arndt, D.; Gautam, M.; Allen, F.; Wishart, D. S. CFM-ID 3.0: Significantly Improved ESI-MS/MS Prediction and Compound Identification. *Metabolites* **2019**, *9* (4). https://doi.org/10.3390/metabo9040072.

(29) Ruttkies, C.; Neumann, S.; Posch, S. Improving MetFrag with Statistical Learning of Fragment Annotations. *BMC Bioinformatics* **2019**, *20* (1), 376. https://doi.org/10.1186/s12859-019-2954-7.

(30) Tsugawa, H.; Kind, T.; Nakabayashi, R.; Yukihira, D.; Tanaka, W.; Cajka, T.; Saito, K.; Fiehn, O.; Arita, M. Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. *Anal. Chem.* **2016**, *88* (16), 7946–7958. https://doi.org/10.1021/acs.analchem.6b00770.

(31) Janesko, B. G.; Li, L.; Mensing, R. Quantum Chemical Fragment Precursor Tests: Accelerating de Novo Annotation of Tandem Mass Spectra. *Analytica Chimica Acta* **2017**, *995*, 52–64. https://doi.org/10.1016/j.aca.2017.09.034.

(32) R Development Core Team. *R: A Language and Environment for Statistical Computing. Vienna, Austria : The R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available Online at Http://Www.R-Project.Org/.*; 2019.

(33) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. *Bioinformatics* **2008**, *24* (21), 2534–2536. https://doi.org/10.1093/bioinformatics/btn323.

(34) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular Framework for Processing, Visualizing, and Analyzing Mass Spectrometry-Based Molecular Profile Data. *BMC Bioinformatics* **2010**, *11* (1), 395. https://doi.org/10.1186/1471-2105-11-395.

(35) Patterson, R. E.; Kirpich, A. S.; Koelmel, J. P.; Kalavalapalli, S.; Morse, A. M.; Cusi, K.; Sunny, N. E.; McIntyre, L. M.; Garrett, T. J.; Yost, R. A. Improved Experimental Data Processing for UHPLC–HRMS/MS Lipidomics Applied to

Nonalcoholic Fatty Liver Disease. *Metabolomics* **2017**, *13* (11), 142. https://doi.org/10.1007/s11306-017-1280-1.

(36) Liu, Y.; D'Agostino, L.; Qu, G.; Jiang, G.; Martin, J. High-Resolution Mass Spectrometry (HRMS) Methods for Nontarget Discovery and Characterization of Poly- and Per-Fluoroalkyl Substances (PFASs) in Environmental and Human Samples. *TrAC Trends in Analytical Chemistry* **2019**. https://doi.org/10.1016/j.trac.2019.02.021.

(37) Bartlett, S. A.; Davis, K. L. Evaluating PFAS Cross Contamination Issues. *Remediation Journal* **2018**, *28* (2), 53–57. https://doi.org/10.1002/rem.21549.

(38) Moriwaki, H.; Takagi, Y.; Tanaka, M.; Tsuruho, K.; Okitsu, K.; Maeda, Y. Sonochemical Decomposition of Perfluorooctane Sulfonate and Perfluorooctanoic Acid. *Environ. Sci. Technol.* **2005**, *39* (9), 3388–3392. https://doi.org/10.1021/es040342v.

(39) Fiehn, O.; Robertson, D.; Griffin, J.; van der Werf, M.; Nikolau, B.; Morrison, N.; Sumner, L. W.; Goodacre, R.; Hardy, N. W.; Taylor, C.; Fostel, J.; Kristal, B.; Kaddurah-Daouk, R.; Mendes, P.; van Ommen, B.; Lindon, J. C.; Sansone, S.-A. The Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3* (3), 175–178. https://doi.org/10.1007/s11306-007-0070-6.

(40) Spicer, R. A.; Salek, R.; Steinbeck, C. A Decade after the Metabolomics Standards Initiative It's Time for a Revision. *Sci Data* **2017**, *4*. https://doi.org/10.1038/sdata.2017.138.

(41) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098. https://doi.org/10.1021/es5002105.

(42) Kind, T.; Fiehn, O. Metabolomic Database Annotations via Query of Elemental Compositions: Mass Accuracy Is Insufficient Even at Less than 1 Ppm. *BMC Bioinformatics* **2006**, *7* (1), 234. https://doi.org/10.1186/1471-2105-7-234.

(43) Pellizzaro, A.; Zaggia, A.; Fant, M.; Conte, L.; Falletti, L. Identification and Quantification of Linear and Branched Isomers of Perfluorooctanoic and Perfluorooctane Sulfonic Acids in Contaminated Groundwater in the Veneto Region. *Journal of Chromatography A* **2018**, *1533*, 143–154. https://doi.org/10.1016/j.chroma.2017.12.036.

(44) U, B.; I, L.; M, O.; R, K. Comparison of Three Types of Mass Spectrometers for HPLC/MS Analysis of Perfluoroalkylated Substances and Fluorotelomer Alcohols. *Eur J Mass Spectrom (Chichester)* **2004**, *10* (5), 579–588. https://doi.org/10.1255/ejms.679.

(45) US EPA, O. Toxicity Estimation Software Tool (TEST) https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test (accessed Mar 11, 2020).

(46) Martin, T. M.; Young, D. M. Prediction of the Acute Toxicity (96-h LC50) of Organic Compounds to the Fathead Minnow (Pimephales Promelas) Using a Group Contribution Method. *Chem. Res. Toxicol.* **2001**, *14* (10), 1378–1385. https://doi.org/10.1021/tx0155045.

(47) Raevsky, O. A.; Grigorev, V. Y.; Yarkov, A. V.; Tinkov, O. V. QSAR Modeling of Mammal Acute Toxicity by Oral Exposure. *Biomedical Chemistry: Research and Methods* **2018**, *1* (3), e00066–e00066. https://doi.org/10.18097/BMCRM00066.
(48) Slotkin, T. A.; MacKillop, E. A.; Melnick, R. L.; Thayer, K. A.; Seidler, F. J. Developmental Neurotoxicity of Perfluorinated Chemicals Modeled in Vitro. *Environ Health Perspect* **2008**, *116* (6), 716–722. https://doi.org/10.1289/ehp.11253.
(49) Fields, S. Brain Effects Stick Around: PFCs and Neural Cell Development. *Environ Health Perspect* **2008**, *116* (6), A258.
(50) Djoumbou-Feunang, Y.; Pon, A.; Karu, N.; Zheng, J.; Li, C.; Arndt, D.; Gautam, M.; Allen, F.; Wishart, D. S. CFM-ID 3.0: Significantly Improved ESI-MS/MS Prediction and Compound Identification. *Metabolites* **2019**, *9* (4). https://doi.org/10.3390/metabo9040072.

Table of Contents Graphic