

# Interactive software for visualization of nontargeted mass spectrometry data—FluoroMatch visualizer

Jeremy P. Koelme<sup>1,†</sup>, Paul Stelben<sup>1,†</sup>, David Godri<sup>2</sup>, Jiarong Qi<sup>1</sup>, Carrie A. McDonough<sup>3</sup>, David A. Dukes<sup>3</sup>, Juan J. Aristizabal-Henao<sup>4</sup>, John A. Bowden<sup>4,5</sup>, Sandi Sternberg<sup>6</sup>, Emma E. Rennie<sup>7</sup> and Krystal J. Godri Pollitt<sup>1,\*</sup>

<sup>1</sup>Department of Environmental Health Sciences, Yale School of Public Health, New Haven, CT, USA

<sup>2</sup>3rd Floor Solutions, Toronto, ON, Canada

<sup>3</sup>Department of Civil Engineering, Stony Brook University, Stony Brook, NY, USA

<sup>4</sup>Department of Chemistry, University of Florida, Gainesville, FL, USA

<sup>5</sup>Center for Environmental and Human Toxicology & Department of Physiological Sciences, University of Florida, Gainesville, FL, USA

<sup>6</sup>Innovative Omics Inc, Sarasota, FL, USA

<sup>7</sup>Agilent Technologies Inc, Santa Clara, CA, USA

\*To whom correspondence should be addressed: Email: [krystal.pollitt@yale.edu](mailto:krystal.pollitt@yale.edu)

†These authors contributed equally to this work.

## Abstract

There are thousands of different per- and polyfluoroalkyl substances (PFAS) in everyday products and in the environment. Discerning the abundance and diversity of PFAS is essential for understanding sources, fate, exposure routes, and the associated health impacts of PFAS. While comprehensive detection of PFAS requires the use of nontargeted mass spectrometry, data processing is time intensive and prone to error. While automated approaches can compile all mass spectrometric evidence (e.g., retention time, isotopic pattern, fragmentation, and accurate mass) and provide ranking or scoring metrics for annotations, confident assignment of structure often still requires extensive manual review of the data. To aid this process, we present FluoroMatch Visualizer, an open-source free software which was developed to provide interactive visualizations which include normalized mass defect plots, retention time versus accurate mass plots, MS/MS fragmentation spectra, and tables of annotations and metadata. All graphs and tables are interactive and have cross-filtering such that when a user selects a feature, all other visuals highlight the feature of interest. Several filtering options have been integrated into this novel data visualization tool, specifically with the capability to filter by PFAS chemical series, fragment(s), assignment confidence, and MS/MS file(s). FluoroMatch Visualizer is part of FluoroMatch Suite, which consists of FluoroMatch Modular, FluoroMatch Flow, and FluoroMatch Generator. FluoroMatch Visualizer enables annotations to be extensively validated, increasing annotation confidence. The resulting visualizations and datasets can be shared online in an interactive format for community-based PFAS discovery. FluoroMatch visualizer holds potential to promote harmonization of nontargeted data processing and interpretation throughout the PFAS scientific community.

**Keywords:** mass spectrometry; PFAS; software; nontargeted analysis; aqueous film forming foam; liquid chromatography

## Introduction

The diversity in synthesized chemicals and biproducts released into the environment continues to increase in part due to technological improvements reducing the cost and time needed to design and synthesize new chemicals, along with the use of anthropogenic chemicals in virtually all areas of life.<sup>1–3</sup> Alongside the benefits of novel chemicals are the unknown health impacts of exposure.<sup>2,4</sup> Databases of chemicals that may be found in the environment range from 360 000<sup>5</sup> to over 800 000 compounds,<sup>6</sup> with additional compounds added annually. Determining which chemicals pose a health risk is challenging without an understanding of the persistence, bioaccumulation, fate and transport, transformation products, and toxicity of the compound. One major chemical class that is receiving scientific and regulatory attention is the expanding list of per- and polyfluoroalkyl substances (PFAS).

In this manuscript, we define PFAS as any molecule containing the set of substructures defined by the EPA in the CompTox

Dashboard, which are a carbon–fluorine chain (CF<sub>2</sub>) of two units, or two units of CF<sub>2</sub> separated by oxygen, CH<sub>2</sub> group, CHF group, or one unit of CF<sub>2</sub> with a CH<sub>2</sub> and CHF.<sup>6</sup> There are thousands of known PFAS, which combined with the fact that they are pervasive,<sup>7–10</sup> persistent, and are often noted as toxic, makes these chemicals an important target for comprehensive nontargeted analysis. The health burden associated with PFAS can depend on structure, and most studies on PFAS health effects primarily focus on perfluorooctanoic acid (PFOA) and perfluorooctanesulfonic acid (PFOS). Most PFAS bioaccumulate<sup>11–16</sup> and certain PFAS have been linked to high cholesterol<sup>17,18</sup> and triglycerides,<sup>19,20</sup> thyroid disease,<sup>18</sup> pregnancy-induced hypertension,<sup>21,22</sup> ulcerative colitis,<sup>23</sup> cancers,<sup>24,25</sup> and a weakened immune system.<sup>26–29</sup> While there has been increased regulatory,<sup>30,31</sup> academic, and public awareness, scrutiny, and action to limit PFAS exposures, novel PFAS often are outside the scope of these efforts.<sup>32–34</sup> To accompany the growing need for regulatory changes, tools, such as high-resolution mass spectrometry,<sup>35–37</sup> are needed to provide

Received: May 02, 2022. Revised: July 18, 2022. Accepted: July 28, 2022

© The Author(s) 2022. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

the capability to rapidly assess the diversity, abundance, and health impacts of environmental exposures in a comprehensive fashion to better determine and understand emerging and legacy chemicals of concern.

Liquid chromatography high-resolution mass spectrometry (LC-HRMS/MS) utilizes universal chemical principles to determine chemical structure including chemical polarity, size, or charge (retention time), bond connectivity (fragmentation), molecular formula (mass and isotopic pattern), and chemical moieties/chemical class (fragmentation and spectral similarity).<sup>35-37</sup> The specificity and universality of mass spectral evidence allow for nontargeted analysis: the analysis of thousands to hundreds of thousands of signals representing individual chemicals or structurally similar isobars (features within 1Da) or isomers (chemicals with the exact same formula but different structure). However, annotating chemical structure from mass spectral signal is often the bottleneck in nontargeted LC-HRMS/MS and even with sophisticated software, manual review of data is generally required for high confidence.<sup>38-41</sup>

For nontargeted PFAS data processing, we have previously released FluoroMatch Flow, FluoroMatch Modular, and FluoroMatch Generator which cover all steps of the PFAS data processing workflow.<sup>42,43</sup> FluoroMatch Flow performs all steps in automatic fashion, FluoroMatch Modular can be used to annotate feature tables acquired using the users' own peak picking and filtering algorithms, and FluoroMatch Generator can be used to develop PFAS MS/MS libraries from a few class-representative standards. While a low false positive rate (< 5%) has been shown for FluoroMatch Flow, using class-based fragmentation rules from standards,<sup>42,45</sup> annotations using fragment screening, homologous series detection, and other more nontargeted approaches have higher false positive rates.<sup>43</sup> Manual validation is often required for the annotation of the majority of PFAS suspected in nontargeted high-resolution mass spectrometry datasets.

To aid in manual validation of PFAS annotations, we introduce FluoroMatch Visualizer. This tool provides interactive, cross-linked, and cross-filtered graphics and tables to facilitate evaluation of annotations in complex data. Here we describe a suggested step-by-step workflow employing the tool, exemplifying this workflow using PFAS detected in Aqueous Film Forming Foam (AFFF).

## Methods

### FluoroMatch visualizer integration with FluoroMatch Flow/Modular

FluoroMatch Visualizer was developed to work with outputs from FluoroMatch Flow or Modular.<sup>42,43</sup> FluoroMatch v2.6 currently outputs a csv file, which contains over 35 columns of information for each feature and includes confidence scores, SMILES structures, identifiers, names, formulas, mass-to-charge ratios, retention times, peak area(s), mass defects, fragments, and annotations, MS/MS files, and adducts. The code was modified to also output annotated MS/MS spectra combined across all MS/MS-containing files in a format readable by Microsoft Power BI. These are the only two inputs needed for FluoroMatch Visualizer.

### Data acquisition and data processing

An AFFF mixture was collected from a holding tank at a field site in 1999; this holding tank contained a mixture of legacy AFFF products. Therefore, characterization of this AFFF mixture can serve as a proxy for legacy AFFF-contaminated sites.<sup>46</sup> The sample was diluted 100 000 fold in 70:30 water: methanol (Fisher

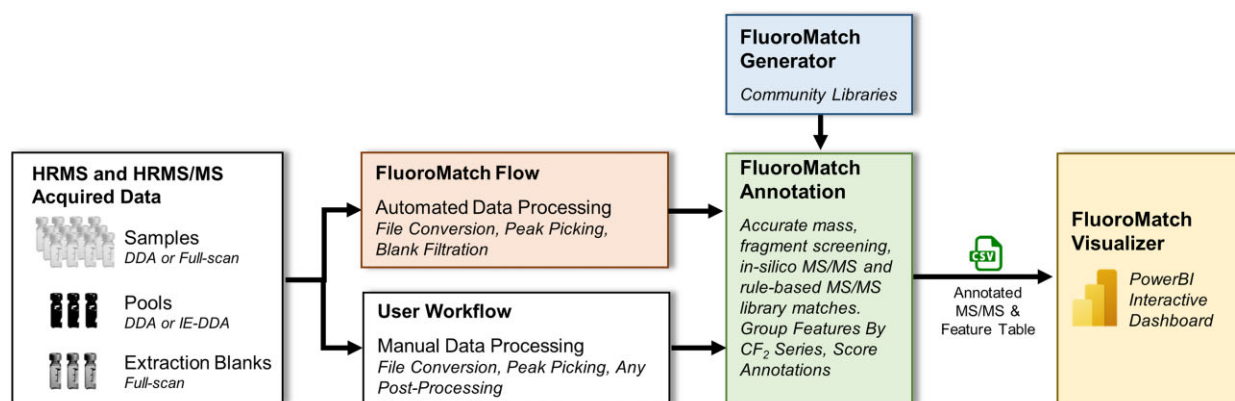
Scientific Optima® LCMS-grade). The diluted sample was injected 4 times for iterative exclusion information-data dependent analysis (iterative MS/MS), with a 50 µL injection volume onto an Agilent 1290 Infinity II ultra-high-performance liquid chromatography (UHPLC) system connected to an Agilent 6545 quadrupole time-of-flight mass spectrometer (Q-TOF MS). The LC-QTOF method used has been published previously in PFAS suspect screening studies.<sup>16,47</sup> The column used was an Agilent Poroshell EC-C18, 3 × 100 mm, 4 µm. Blanks were acquired every other injection for blank filtering. PFAS were detected in negative electrospray ionization mode. Data were acquired from *m/z* 100-1100, with MS/MS collision energy set to 0, 25, and 40 eV. Source parameters and further acquisition parameters for this dataset have been previously described.<sup>16,43</sup>

## Results and discussion

### FluoroMatch workflow

FluoroMatch and FluoroMatch Visualizer were developed to be comprehensive, modifiable, user-friendly, and open source. The software is flexible, but a few sample types are recommended. During sample preparation, the user should incorporate method blanks which are carried throughout all field collection, extraction, storage, and sample handling procedures. These blanks can be analyzed throughout the acquisition queue to also account for carryover, instrument contamination, and other acquisition-related background and artifacts. This also enables the software to remove any peaks which are not related to the samples. Furthermore, MS/MS is required for annotation, and we recommend acquiring MS/MS on pooled samples or representative samples in an iterative MS/MS or intelligent acquisition mode to increase coverage.<sup>42,48</sup> FluoroMatch can use data-dependent and targeted MS/MS files as well, but currently does not support data-independent approaches, as these approaches are challenging when most precursors share the same fragments (e.g., [C<sub>2</sub>F<sub>5</sub>]<sup>-</sup>). Finally, full-scan data is required for all samples for peak integration and statistics across samples; MS/MS files can be used for full-scan data if at least 10-15 scans are acquired across the chromatographic peak.

The data processing workflow starts with FluoroMatch Flow (or Modular), which can be used to directly process vendor files (Figure 1); the software is freely available at [innovativeomics.com/software](https://innovativeomics.com/software). FluoroMatch Modular uses the same annotation algorithms as FluoroMatch Flow but can be used to annotate feature tables generated using the users' own peak picking and filtering algorithms/software (vendor or open source including Profinder, Compound Discoverer, MZMine, and XCMS). FluoroMatch Flow, which covers the entire PFAS nontargeted data processing workflow (Figure 1), uses MZMine,<sup>49</sup> MSConvert,<sup>50</sup> and in-house code using R and C# in the background to process the files. FluoroMatch Flow works for most vendor files including Waters, Agilent, Thermo, and Bruker. Together FluoroMatch Flow, FluoroMatch Modular, FluoroMatch Generator, and FluoroMatch Visualizer are termed FluoroMatch Suite, and this suite of software offers several workflows (Figure 1). After following the proper naming conventions outlined for this software program, users simply drag files onto the interface, choose blank filtering parameters and their preferred output directory, and click run. The software performs file conversion, peak processing (including peak picking, alignment, and gap-filling), blank filtering, compilation of annotation evidence, annotation, and assignment of confidence (Figure 1). Blank filtering is performed by removing features that have sample quartiles that are not above a certain fold-change of the average of the blanks plus 3 times the standard



**Figure 1.** The FluoroMatch Suite covers the entire PFAS non-targeted data processing and manual validation workflow. User acquisition and data processing workflow using FluoroMatch Modular, FluoroMatch Flow, and FluoroMatch Visualizer are shown. DDA, data-dependent analysis; MS/MS, fragmentation spectra.

deviation of the blanks (threshold parameters, such as which quartile to use of the samples is experiment dependent and user-modifiable). Confidence assignments are based on an in-house scoring framework<sup>42,43</sup> (Figure 2) and the Schymanski scheme.<sup>51</sup> Recently, a specific schema has been developed by the community for assigning confidence to PFAS annotations,<sup>52</sup> and this schema may be integrated into future releases of FluoroMatch. Comparison of FluoroMatch Flow with other software, including Compound Discoverer and EnviMass, shows unique advantages and disadvantages of various software.<sup>44,45</sup> In previous studies, FluoroMatch Flow has been shown to have similar or lower false positive rates among high-confidence annotations and has similar or more comprehensive annotation when the same peak list is used,<sup>44,45</sup> but vendor or open-source peak picking algorithms designed specifically for the methods/instruments deployed often have better peak picking coverage than built-in algorithms in FluoroMatch Flow. To increase coverage, users can use their own peak picking workflow optimized for their instrument and gradient, using FluoroMatch Modular for annotation, compiling evidence, and scoring; this workflow is currently recommended for expert users. FluoroMatch Modular works with any peak picking/processing software (both vendor and open source), which outputs a txt file with  $m/z$  and retention time information.

FluoroMatch Visualizer requires the desktop version of Power BI which can be freely downloaded at <https://powerbi.microsoft.com/en-us/downloads/>. An online version is being developed for future use. Using the FluoroMatch Visualizer Power BI file, users upload their data processing output file from FluoroMatch Modular or FluoroMatch Flow.

### FluoroMatch visualizer description and possible workflow


























In this section, different methods for validating and discovering likely PFAS series are covered. The software automatically determines series by grouping features which have the same  $CF_2$  normalized mass defect (e.g., within 5 ppm, or user-selected precursor mass accuracy), and are spaced out by 50 Da ( $CF_2$  nominal mass). Other series (e.g.,  $CF_2CF_2O$ ), as well as multiple series simultaneously, can be searched for as well (see tutorial videos). By default, retention time order is also used to include members of a series, and if features do not increase in retention time with increasing nominal mass, they are flagged as potential false positives. This series detection does not include retention time trends other than order, nor MS/MS spectra (other than in scoring

members of a series), nor isotopic pattern, and therefore manual review using FluoroMatch Visualizer and other tools is helpful to validate series. Because exact mass alone cannot distinguish isomers or even provide an exact formula in most cases, these other lines of evidence must be reviewed before confidently assigning a series.



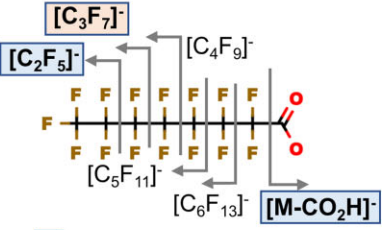
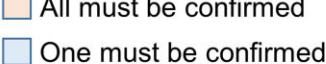



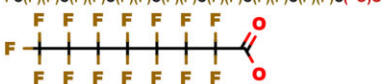

FluoroMatch Visualizer includes three interactive graphs:  $m/z$  vs retention time (Figure 3B), normalized mass defects (Figure 3C), and MS/MS spectra (Figure 3E). These graphs can be filtered by score and chemical series (Figure 3A). Data is further summarized in two tables: annotated and scored features (Figure 3F) and annotated fragments for fragment screening (Figure 3D). For all tables and graphs, users can drag over entries to visualize PFAS chemical structures, or substructures related to fragmentation. For this, we developed a novel tool for interpreting SMILES within Power BI.

All graphs and tables can be filtered by user-selected feature(s). For example, if a row in the feature table is selected then only that feature is displayed in the charts and vice versa. Cross-filtering is important so that all evidence for a feature, PFAS series, or other group of features, can be analyzed simultaneously. FluoroMatch Visualizer additionally includes tool tips, which provide further information when the user hovers over a feature or fragment in the MS/MS spectra. When hovering over a feature in graphs showing  $m/z$  vs retention time (Figure 3B) or normalized mass defects (Figure 3C), the chemical series, formula, SMILES structure, name/class, score, and  $x/y$ -axis values are shown. In the graph of MS/MS spectra (Figure 3E), any fragment annotation, potential SMILES substructure of fragment, fragment  $m/z$ , fragment intensity, and ppm error of any fragment annotation are shown. In addition, features aligned with presented fragments are presented, which is important when MS/MS spectra from multiple features are overlaid.

User workflows employing FluoroMatch Visualizer can be diverse, especially considering that using Power BI Desktop new graphs, variables, and tables can be designed and added by users familiar with the platform. 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 splitters and filters can be developed. Here we describe a simple workflow for determining false positives and true positives, in “Application of FluoroMatch visualizer to identify PFAS in a legacy AFFF mixture” section. This workflow will be exemplified using an AFFF mixture.

Score	Description	Score	Description
A+	Confirmed by standards (RT, MS/MS, MS)	C+	 $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> In series w/ confident ID
A	 $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> Confident ID (MS/MS, MS) and series	C	 $\longleftrightarrow$  $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> (CF <sub>2</sub> ) <sub>n</sub> In series (3+) w/ any tentative ID
A-	 Confident ID (MS/MS, MS)	C-	 $\longleftrightarrow$  or  $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> (CF <sub>2</sub> ) <sub>n</sub> In series w/ any tentative ID
B+	 $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> Tentative ID (MS/MS, MS) and series	D+	 $\longleftrightarrow$  $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> (CF <sub>2</sub> ) <sub>n</sub> MD or MS and in series (3+)
B	 $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> Tentative ID (MS/MS, MS) and series	D	 $\longleftrightarrow$  $\longleftrightarrow$  (CF <sub>2</sub> ) <sub>n</sub> (CF <sub>2</sub> ) <sub>n</sub> In series (3+)
B- *B--	 or  Any tentative ID (MS/MS, MS)	D-	 MD or MS

  Derived MS/MS rules from class-based standards  	 Confident Annotation (MS/MS rules derived from class-based standards) Example: PFCA MS: [M-H] <sup>-</sup> MS/MS: [C <sub>3</sub> F <sub>7</sub> ] <sup>-</sup> AND ([C <sub>2</sub> F <sub>5</sub> ] <sup>-</sup> OR [M-COOH] <sup>-</sup> )   Tentative Annotation (1+ fragment from class-based standards) Example: PFCA MS: [M-H] <sup>-</sup> MS/MS: [M-H] <sup>-</sup> , [M-CO <sub>2</sub> H] <sup>-</sup> , [C <sub>2</sub> F <sub>5</sub> ] <sup>-</sup> , [C <sub>3</sub> F <sub>7</sub> ] <sup>-</sup> , [C <sub>4</sub> F <sub>9</sub> ] <sup>-</sup> , [C <sub>5</sub> F <sub>11</sub> ] <sup>-</sup> OR [C <sub>6</sub> F <sub>13</sub> ] <sup>-</sup>
 In-silico fragments from substructures  c(=O)O: [CO <sub>2</sub> F] <sup>-</sup> , [M-CO <sub>2</sub> H] <sup>-</sup> c(F)(F): [CF <sub>2</sub> ] <sup>-</sup> , [CF <sub>3</sub> ] <sup>-</sup> c(F)(F)c(F)(F): [C <sub>2</sub> F <sub>5</sub> ] <sup>-</sup> c(F)(F)c(F)(F)c(F)(F): [C <sub>3</sub> F <sub>5</sub> ] <sup>-</sup> , [C <sub>3</sub> F <sub>7</sub> ] <sup>-</sup> c(F)(F)c(F)(F)c(F)(F)c(F)(F): [C <sub>4</sub> F <sub>7</sub> ] <sup>-</sup> , [C <sub>4</sub> F <sub>9</sub> ] <sup>-</sup>	 Tentative Annotation (1+ F-containing fragment from <i>in-silico</i> MS/MS)  Example: PFCA MS: [M-H] <sup>-</sup> MS/MS: [M-HF] <sup>-</sup> , [CO <sub>2</sub> F] <sup>-</sup> , [M-F] <sup>-</sup> , [CF <sub>2</sub> ] <sup>-</sup> , [CF <sub>3</sub> ] <sup>-</sup> , [C <sub>2</sub> F <sub>5</sub> ] <sup>-</sup> , [C <sub>3</sub> F <sub>5</sub> ] <sup>-</sup> , [C <sub>3</sub> F <sub>7</sub> ] <sup>-</sup> , [C <sub>4</sub> F <sub>7</sub> ] <sup>-</sup> , [C <sub>4</sub> F <sub>9</sub> ] <sup>-</sup> , [C <sub>5</sub> F <sub>9</sub> ] <sup>-</sup> , [C <sub>5</sub> F <sub>11</sub> ] <sup>-</sup> , [C <sub>6</sub> F <sub>11</sub> ] <sup>-</sup> , [C <sub>6</sub> F <sub>13</sub> ] <sup>-</sup> , [C <sub>7</sub> F <sub>13</sub> ] <sup>-</sup> , [C <sub>7</sub> F <sub>15</sub> ] <sup>-</sup> , OR [C <sub>2</sub> F <sub>3</sub> ] <sup>-</sup>

**Figure 2.** FluoroMatch Suite includes a systematic scoring framework to communicate confidence for every single feature, alongside also reporting confidence levels via the Schymanski et al. schema. All scores are shown alongside visual and descriptive definitions. Circles refer to features and are depicted within or not within mass defect series. The type of fills of the circles represents evidence used for PFAS assignment. B- are features with fluorine-containing fragments using a list of 777 fluorine-containing fragments for screening derived from standards, but the fragments observed are not common fragments with only one possible predicted formula.

### Step 1: Determine PFAS series with high-confidence annotations

By filtering features to only those having score “A” using the Score Filter (Figure 3B), the most confident annotations are retained (~5% false positive rate).<sup>42,43,45</sup> The remaining members of series containing one or more confident annotation can then be selected using the series type identifier filter (Figure 3A). Then in the Score Filter, all scores can be added back (no filtering) and now the user can view only those series with confident annotations. Confident series should be documented separately for future use.

### Step 2: Determine false positives using retention time vs m/z plots

Opening the annotated feature table (.csv) in excel from FluoroMatch Suite (the file with FIN in the name), a new column can be added for comments and for flagging potential false positives or true positives (note these columns will automatically be integrated in later releases). The file should be saved with a different name in excel format (.xlsx). Following generation of a new file to track new assignments of confidence and structures via manual review of the visualizations, the user can then select the first



defined, or series can be selected and investigated based on containing these moieties.

### Step 6: Discovery of PFAS series without MS/MS evidence

Depending on the application, further validation of series which do not have any structural information but may be PFAS based on accurate mass can be performed. In this case, using methods in Step 1 through Step 2 series with D+ and D can be investigated for retention time patterns (no MS/MS evidence exists for these). In addition, these series can further be narrowed down to only those series within the normalized mass defect where high-confidence annotations were observed, as these are more likely to be PFAS (Figure 3C). One way to navigate these series is to sort the table (Figure 3F) by number of compounds in the series and look at series with many members, these series are likely to be more confident. Finally, as in Step 4, isotopic pattern matching, and formula prediction can be used to validate formula.

### Application of FluoroMatch Visualizer to identify PFAS in a legacy AFFF mixture

Nontargeted data provide a wealth of information and in most datasets the majority of chemical features are not PFAS related. By looking at all features simultaneously, it can be difficult to determine patterns and discover potential PFAS series. Often thousands of features are observed (3686 for the AFFF mixture after blank filtering), and no user can look at all the data and make sense of it

simultaneously; for example, when all features are displayed the resulting plots show few clear patterns (Figures 4 and 5).

Because of the complexity and richness of nontargeted data, users need to look at a subsection of PFAS features and multiple filters are provided in FluoroMatch Visualizer to prioritize which group of features to investigate. Ideally, the user would want to know which series have confident annotations and which series have no annotations, in order to evaluate the most confident annotations first. Filtering by score (Figure 6) allows users to determine which features to focus on based on the evidence provided, by selecting the series which contains the highest scoring features. A description of the scores can be found in Figure 2 and is described previously.<sup>43</sup> Briefly, features assigned as A (A, A-) contain class-based MS/MS fragments required for confident assignments, features assigned B (B+, B, and B-) have an *in-silico* or fragment screening match, features assigned C (C+, C, and C-) are in homologous series with B's or A's, and D's form homologous series or have accurate mass evidence consistent with PFAS but no MS/MS information; E's are the remaining features which are likely not PFAS and therefore can be generally ignored, drastically reducing the number of features needing manual validation.<sup>43</sup>

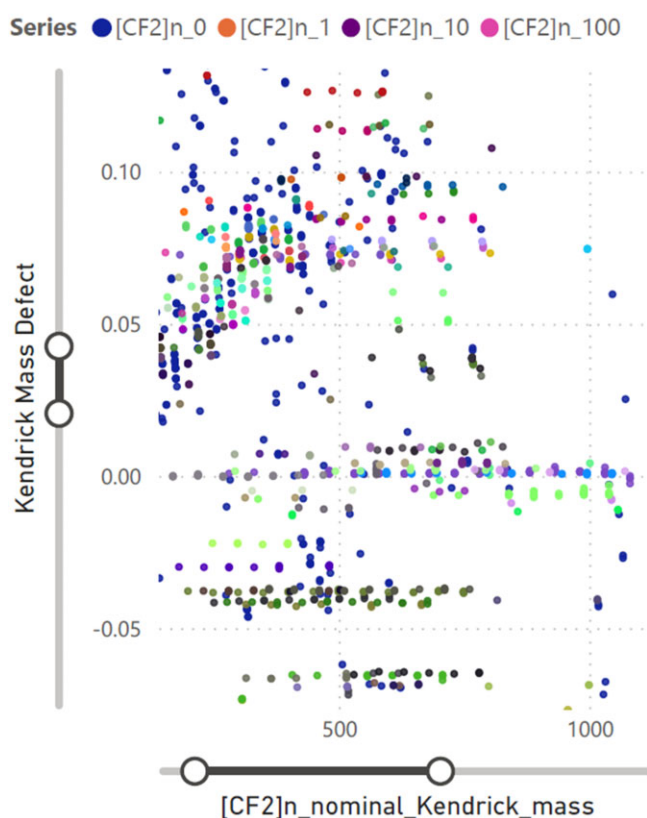
By filtering to retain only features with high-confidence scores (e.g., scores of A; Figure 6A), known series (about a 5% false positive rate)<sup>42,43</sup> can be readily viewed. In the case of the AFFF mixture, 18 PFAS series were determined, covering 62 individual species with a high-confidence score (Figure 6A). These series can then be expanded to all features which fall within the series (including lower confidence annotations or no annotations), in which case the coverage of species with likely annotations expanded to 257 individual species for the AFFF mixture (Figure 6C). After establishing these series containing high-confidence annotations it is important to distinguish which members of the series are false positives via manual review.

Organizing features by series is essential for determining false positives, as clear patterns are observed for each series. If a member of a series falls outside of this pattern, it may be considered a false positive. Therefore, using the drop-down menu (or by selecting series in the legend), users can select one or more series to start determining true and false positives, as well as false negatives. Each member of a series should follow a clear diagonal pattern in *m/z* vs retention time plots as shown for only high-confidence annotations of perfluorosulfonic acids (PFSAs) (Figure 6B) and any outliers can be considered false positives (Figure 6D). Furthermore, members of the series should also follow a clear trend in the normalized mass defect plots. Because the tool automatically assigns series by mass defect and nominal mass, the PFAS series will follow trends in these plots automatically and no manual removal of false positives is needed (Figure 6A and C).

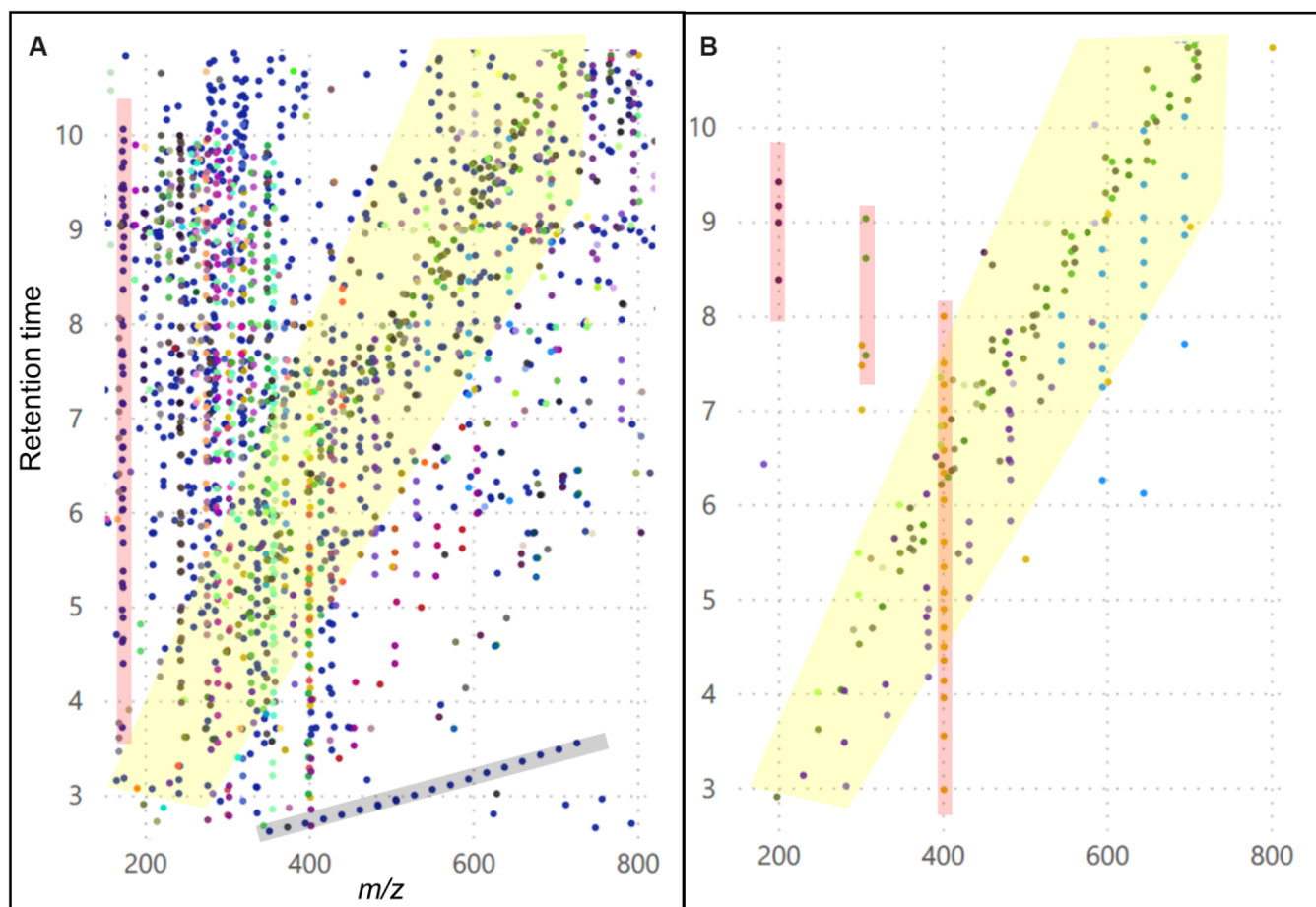
Cross-filtering is especially important when looking at MS/MS spectra, as multiple MS/MS spectra can be overlaid. If certain members of a series MS/MS spectra are overlaid, for example by sorting and selecting them in the plots (e.g., as selected in Figure 6A for the PFSA series) or feature table, the trends for neutral losses and fragment ions become clear (Figure 7). Any MS/MS spectra which do not have fragments consistently observed in other features of the same series can then be labeled as potential false positives. Furthermore, assignments with low confidence may attain higher confidence by the user after surveying the MS/MS spectra overlaid with other members of the series.

Fragment screening can also be achieved by annotated fragments from MS/MS spectra, which can aid in finding unknowns and identifying chemical series/classes. Common fragments which were annotated can be selected and all MS/MS spectra and features across all charts will

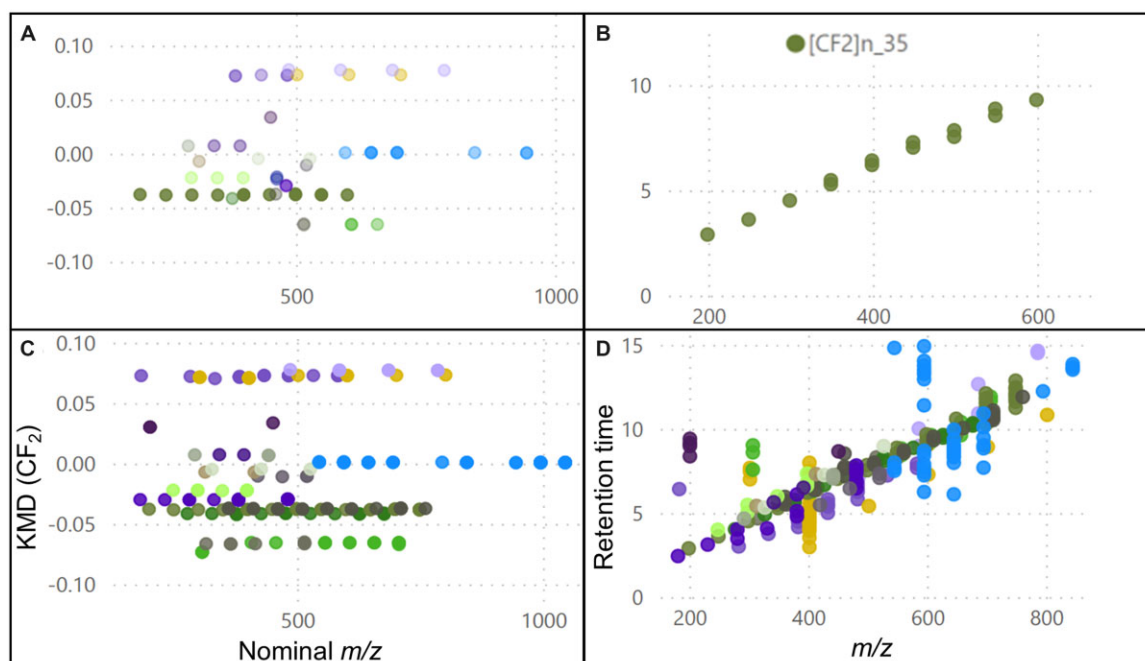
### KMD vs NKM



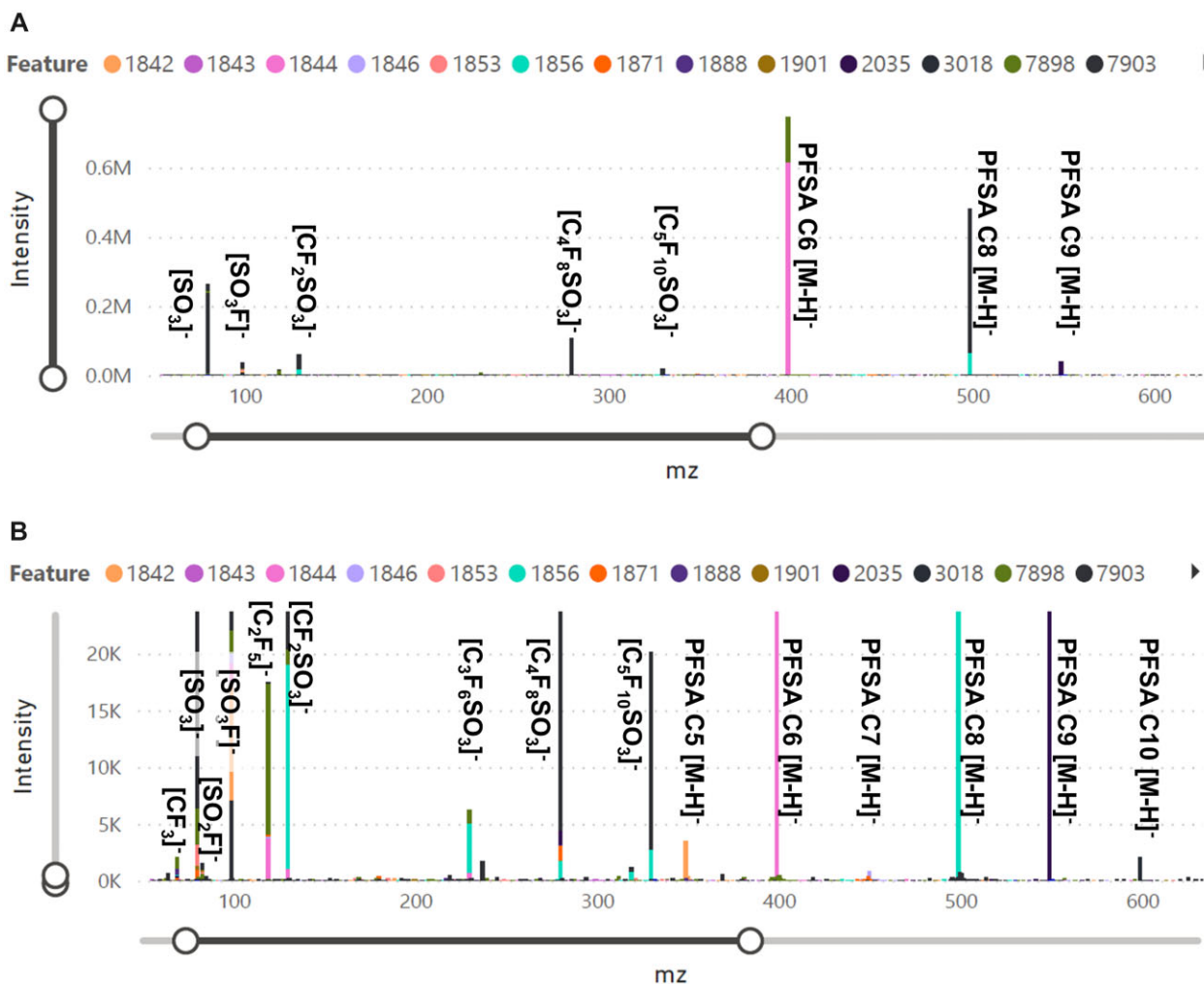
**Figure 4.** Several PFAS series can be discovered by looking at certain regions of mass defect plots normalized to CF<sub>2</sub>. Without filtering data by series or score the visual is overly complex and challenging to make sense of. Mass defect plot (normalized to CF<sub>2</sub>) showing a PFAS-specific region and all features. Horizontal evenly spaced (intervals of 50 Da) points of the same color represent potential PFAS homologous series.



**Figure 5.** Retention time vs  $m/z$  plots are overly complex without filtering; certain PFAS-rich regions, blank signal, and other non-PFAS series can readily be discovered using these plots. Retention time vs  $m/z$  for all features (A) and for features that belong to high-confidence series with one or more PFAS with a score of A (B). A specific region in the plot is where most PFAS can be found, and certain outliers of high confident series can easily be found (B). Red highlighted areas show background noise, yellow highlighted area shows PFAS region of plot, and gray highlighting shows another potential non-PFAS series.



**Figure 6.** By filtering to only include those series with the most confident annotations, true positives and false negatives can manually be determined using clear trends. Shown above are Kendrick mass defect plots normalized to  $CF_2$  (KMD [ $CF_2$ ]) for annotations with high confidence (A) and series containing one or more high confident annotations (C). As an example, for high-confidence annotation, the series for perfluoroalkyl sulfonic acids (PFSA) is selected (A) and the retention time vs  $m/z$  plot is shown for this series (B). Retention time vs  $m/z$  plots is also shown for all high-confidence series' (D). Note other series beyond  $CF_2$  can also be automatically classified by the software, and these series repeating units can be generated by the user.



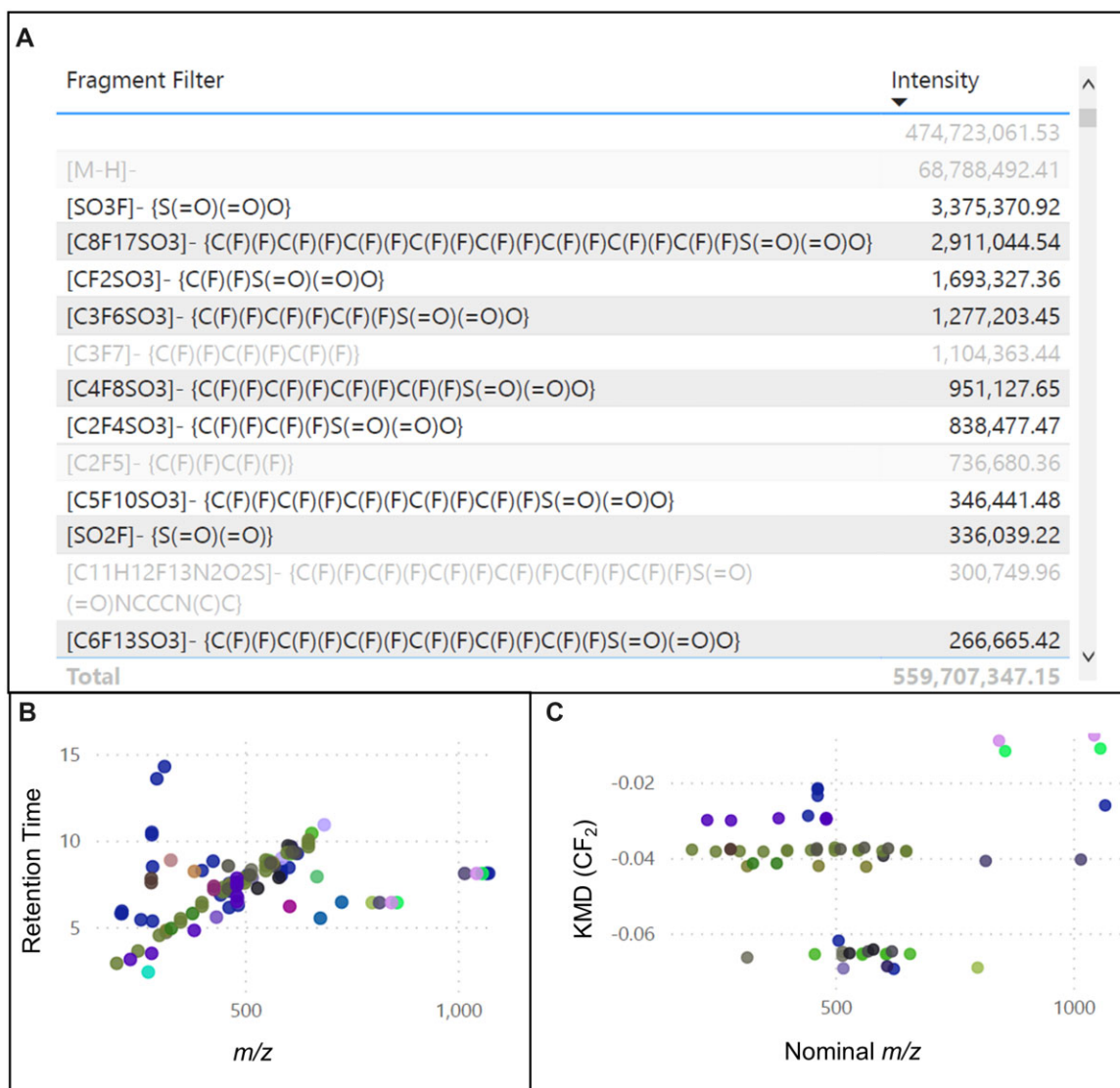
**Figure 7.** Averaging spectra across series can be readily done in the visualizer platform and used to determine common fragmentation patterns and false positives. Combined mass spectra after highlighting the perfluoroalkyl sulfonic acids (PFSA) series (Figure 5A), zoomed out (A) and zoomed in (B) with annotations (annotations are provided by the software upon dragging over a fragment peak).

be shown, which contain these annotated fragments. This can aid in grouping compounds that contain certain moieties. For example, [SO<sub>3</sub>F]<sup>-</sup> and [SO<sub>2</sub>F]<sup>-</sup> may be selected to find PFAS with sulfonic acid groups, and [CF<sub>3</sub>]<sup>-</sup> and [C<sub>2</sub>F<sub>5</sub>]<sup>-</sup> may be selected to identify carbon-fluorine chains. PFAS-containing sulfonic acid groups (Figure 8) and PFAS-containing pentafluorosulfide (Figure 9) were screened in this manner. One strategy to determine which fragments to use for fragment screening is to sort the annotated fragments by fragment abundance and then select those representing a moiety of interest. Fragments related to sulfonic acid and the PFAS carbon-fluorine chain were the most abundant fragments when all fragments across all features were summed (automatically provided by FluoroMatch Visualizer) (Figure 8A). Many potential false positives were determined using retention time vs *m/z* plots (Figure 8B), showing the importance of this visualization.

Fragment screening using the [SF<sub>5</sub>]<sup>-</sup> fragment for pentafluorosulfide-containing PFAS showed that while the fragment was of low abundance across all features (Figure 9B) 11 features were determined (Figure 9A). When looking at all members of those series containing at least one feature with a [SF<sub>5</sub>]<sup>-</sup>

fragment, 83 features were determined (Figure 9C and D). For one series, many false positives (teal series containing *m/z* 257.0472, 307.0419, and 357.0414) were observed (Figure 9D) and the entire series (blue) was also likely a false positive (no retention time order, data not shown).

Beyond fragment filtering and overlapping series MS/MS spectra, MS/MS spectra and features can be filtered by MS/MS file, in case certain files represent very different sample types, which may have different PFAS isomers. Furthermore, techniques using different MS/MS methods or parameters can be cross-compared readily using filter by MS/MS file. Here we compared different acquisitions of iterative exclusion (IE), where repeated injections of a sample using a rolling exclusion list of ions previously selected are deployed (ensuring that each new injection has MS/MS acquired on new ions not previously selected). As shown in mass defect plots in the PFAS region filtered by each MS/MS file, new series and new members are discovered for each additional round of IE (Figure 10A).<sup>42,48</sup> Note these plots are cumulative, showing all past acquisitions and new acquisitions. Furthermore, by summing and plotting the number of features and filtering by



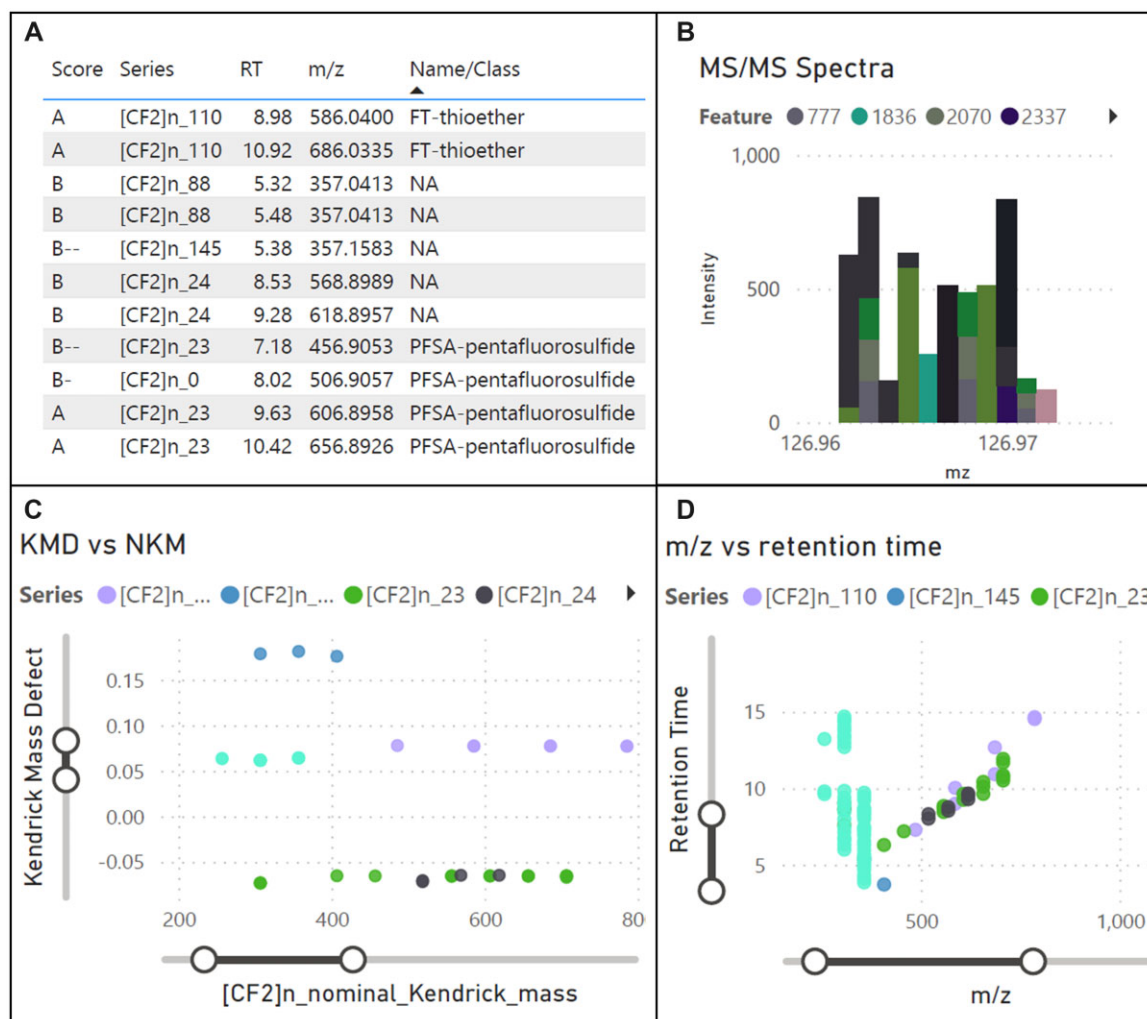
**Figure 8.** Fragment screening can be deployed to determine PFAS series with specific moieties and by sorting annotated fragments by abundance, the most common PFAS fragments can be determined out of a database containing over 777 potential PFAS fragments. Topmost abundant annotated fragments for an entire dataset containing fluorine atoms (A), sulfonic acid-related fragments are selected (A) to identify all PFAS series (C) with sulfonic acids. Numerous outliers are shown in retention time vs  $m/z$  plots which may not belong to PFAS series, as well as likely true positives (B).

score, it's shown that the advantage dissipates for high-confidence annotations after about the fourth injection, meaning fewer injections are needed if high-confidence annotations are the goal (Figure 10B–D).

## Conclusion and future developments

Nontargeted liquid chromatography high-resolution mass spectrometry (LC-HRMS/MS) provides evidence using universal chemical properties (polarity, functional groups, molecular formula, bond linkages, and strength) for comprehensive characterization of molecules. Sifting through this rich set of evidence provided by LC-HRMS/MS, for structural characterization of the molecules, is a major bottleneck in nontargeted workflows. Our automated

approach previously released for PFAS (FluoroMatch) incorporates mass defect, retention time order, exact mass matching, homologous series grouping, and fragmentation pattern searching (class-based fragment rules, common PFAS fragment screening, and structure to fragmentation *in-silico* approaches) to annotate features. While scores are provided to communicate confidence in the results, for most annotations, manual review is required. FluoroMatch Visualizer was developed to meet this need with interactive tables and charts, allowing the user to examine visually several lines of evidence for annotating a specific feature. Visual evidence includes Kendrick mass defect type plots normalize to  $CF_2$  or other repeating unit type,  $m/z$  versus retention time plots, and annotated tandem mass spectra, while tables provide a wealth of other evidence for each selected feature.



**Figure 9.** Certain fragments are very specific to chemical moieties and hence can be advantageous to use in fragment screening, for example, the  $[\text{SF}_5]^-$  fragment is specific to pentafluorosulfide-containing species found in AFFFs. Features (A) and their scores, homologous series, retention time, mass to charge ratio ( $m/z$ ), and level A annotations, filtered by performing fragment screening on the  $[\text{SF}_5]^-$  fragment is shown.  $[\text{SF}_5]^-$  fragment intensities from different fragment scans/features are shown within a 5-ppm window (B), and these fragments were used to screen for pentafluorosulfide-containing series (Figure 9C). Most species form clear trends in retention time vs  $m/z$  plots, with some obvious outliers (Figure 9D).

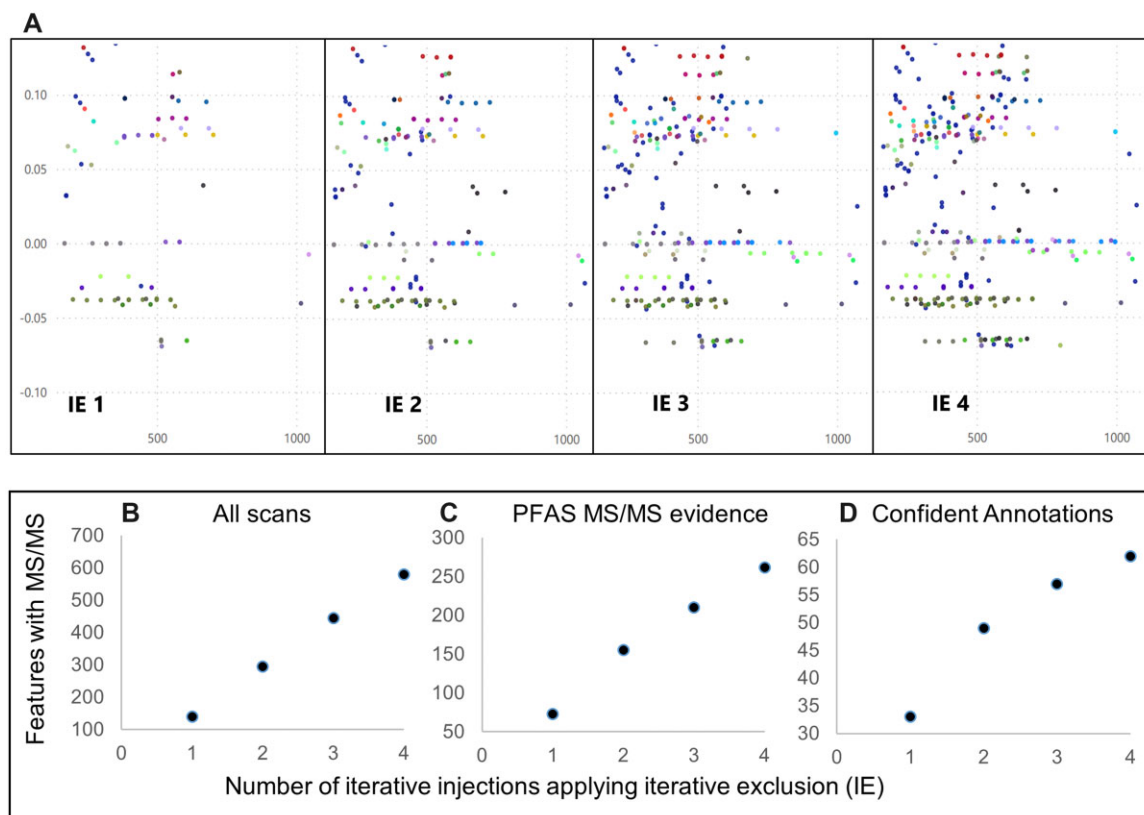
While this software currently covers PFAS, the framework can readily be adapted to other chemical series, and work is being undertaken to expand capabilities to polymers and lipids.

To investigate trends across PFAS, the user can narrow down the number of features to examine using multiple methods. One of the most useful methods is to select individual homologous series, automatically determined using nominal mass and normalized mass defect. When series are selected, all visuals, including MS/MS spectra, are updated to show all members of a series overlaid. The patterns can easily be observed, and outliers determined. Tens to hundreds of series often exist, and therefore series can further be narrowed down easily by those containing high scores or certain characteristic PFAS fragments. Beyond series automatically assigned, PFAS can be determined in another nontargeted fashion, by selecting all features containing a specific fragment. In an application to a mixture of aqueous fire-fighting foams (AFFF), certain PFAS functional groups and structural characteristics could be determined, for example, sulfonic acid, alcohol, pentafluorosulfide, double bond, carbon-

fluorine chain, and carboxylic acid containing species in this application.

Future developments will further improve the utility of this software. For example, statistics could also be performed in Power BI including boxplots, tests of significance, ANOVA, and volcano plots. The advantage of embedding statistics here is that the statistical graphs and tables would update depending on the features or series the user selects. Furthermore, this visual platform can be expanded to other applications being developed, including lipids and polymers.

By using FluoroMatch Visualizer as part of FluoroMatch Suite, researchers can quickly cross-compare results in a visual manner and share reports and data online in an interactive fashion. This will ideally aid in transparency and community-based validation of results in nontargeted PFAS analysis. The nontargeted community uses various methodologies and lines of evidence and this harmonization of reporting and data sharing can improve trust and comprehension in the field of LC-HRMS/MS PFAS analysis.



**Figure 10.** By filtering by MS/MS file, the benefit of different MS/MS acquisition techniques can readily be explored. Features with MS/MS after four iterations of IE, showing increased coverage in the PFAS region of Kendrick mass defect plots with additional injections (A). The increase in MS/MS coverage is more representative of features with low abundance after iterative injections; whereas the number of new features with MS/MS increases linearly throughout all injections (B), after the third and fourth injections new ions selected are of too low abundance for MS/MS to have high enough spectral quality for most confident assignments (C and D). All scans represent any score A through E (B), scans with PFAS MS/MS evidence represent scores of A through B (except B-) (C), and confident annotations refer to those with scores of A and A- (D).

## Acknowledgments

A large array of PFAS standards used to develop libraries were donated by SynQuest Labs, Inc (<http://www.synquestlabs.com>) and Oakwood Products, Inc. (<https://www.oakwoodchemical.com>).

## Funding

J.P.K. and K.J.G.P. received support from Agilent Technologies ACT-UR grant mechanism. We acknowledge support from Innovative Omics Inc. for funding the development of training videos and software. J.P.K. is supported by the Yale Cancer Prevention and Control Training Program (NIH T32 CA250803). K.J.G.P. acknowledges funding from NIH R01ES032196. J.A.B. received support from the U.S. Environmental Protection Agency under the Science to Achieve Results (STAR) grant programs: EPA-G2018-STAR-B1—Grant#: 83962001-0 and EPA-G2019-STAR-E1—Grant#: 84004501-0.

## Conflict of interest statement

J. P. K. is CEO of Innovative Omics that provides trainings on the PFAS data-processing workflow, which include trainings on software including the software presented in this manuscript. All other authors declares no competing financial interest.

## Data availability

The raw Agilent “.d” files can be downloaded at: [ftp://massive.ucsd.edu/MSV000086811/updates/2021-02-05\\_jeremykoelmel\\_e5b21166/raw/McDonough\\_AFFF\\_3M\\_ddMS2\\_Neg.zip](ftp://massive.ucsd.edu/MSV000086811/updates/2021-02-05_jeremykoelmel_e5b21166/raw/McDonough_AFFF_3M_ddMS2_Neg.zip) (Note use Google Chrome or Firefox, Microsoft Edge and certain other browsers are unable to download from an ftp link). The Power BI interactive AFFF results can be viewed here: <https://pollittlab.weebly.com/pfas.html>

## Code availability

The FluoroMatch Software Suite and written and video tutorials are available at: <http://innovativeomics.com/software> The Power BI file can be found at <http://innovativeomics.com/software/fluoromatch-flow-covers-entire-pfas-workflow/>. It can be found in the following directory after unzipping the file: FluoroMatch-2.6/FluoroMatch\_Visualizer/2022\_05\_01\_Visualizer\_v9\_FIN.pbix. YouTube Tutorials can be found here: [https://youtube.com/playlist?list=PLZtU6nmcTb5kv9gvYuikgDXc\\_1\\_mhxTqO](https://youtube.com/playlist?list=PLZtU6nmcTb5kv9gvYuikgDXc_1_mhxTqO)

## References

1. Easterlin RA. Industrial revolution and mortality revolution: two of a kind? *J Evol Econ.* 1995;5(4):393–408. <https://doi.org/10.1007/BF01194368>

2. Whitesides GM. Assumptions: taking chemistry in new directions. *Angew Chem Int Ed Engl.* 2004;43(28):3632–3641. <https://doi.org/10.1002/anie.200330076>
3. Aaron H, Schwartz WB. *Coping with Methuselah: The Impact of Molecular Biology on Medicine and Society.* Brookings Institution Press; 2004.
4. Naidu R, Biswas B, Willett IR, et al. Chemical pollution: a growing peril and potential catastrophic risk to humanity. *Environ Int.* 2021;156:106616. <https://doi.org/10.1016/j.envint.2021.106616>
5. Schymanski EL, Kondić T, Neumann S, Thiessen PA, Zhang J, Bolton EE. Empowering large chemical knowledge bases for exposomics: PubChemLite meets MetFrag. *J Cheminform.* 2021;13(1):19. <https://doi.org/10.1186/s13321-021-00489-0>
6. Williams AJ, Grulke CM, Edwards J, et al. The CompTox chemistry dashboard: a community data resource for environmental chemistry. *J Cheminform.* 2017;9(1):61. <https://doi.org/10.1186/s13321-017-0247-6>
7. Olsen GW, Mair DC, Lange CC, et al. Per- and polyfluoroalkyl substances (PFAS) in American red cross adult blood donors, 2000–2015. *Environ Res.* 2017;157:87–95. <https://doi.org/10.1016/j.envres.2017.05.013>
8. Göckener B, Weber T, Rüdell H, Bücking M, Kolossa-Gehring M. Human biomonitoring of per- and polyfluoroalkyl substances in German blood plasma samples from 1982 to 2019. *Environ Int.* 2020;145:106123. <https://doi.org/10.1016/j.envint.2020.106123>
9. Smithwick M, Mabury SA, Solomon KR, et al. Circumpolar study of perfluoroalkyl contaminants in polar bears (*Ursus Maritimus*). *Environ Sci Technol.* 2005;39(15):5517–5523. <https://doi.org/10.1021/es048309w>
10. Hoppin J, Kotlarz N, de Kort T, Ng-A-Tham J, Starling A, Adgate J, Jakobsson K. An overview of emerging PFAS in drinking water worldwide. *Environ Epidemiol.* 2019;3:162–163. <https://doi.org/10.1097/01.EE9.0000607564.20698.d8>
11. Langberg HA, Breedveld GD, Grønning HM, Kvennås M, Jenssen BM, Hale SE. Bioaccumulation of fluorotelomer sulfonates and perfluoroalkyl acids in marine organisms living in aqueous film-forming foam impacted waters. *Environ Sci Technol.* 2019;53(18):10951–10960. <https://doi.org/10.1021/acs.est.9b00927>
12. Zhao S, Zhu L, Liu L, Liu Z, Zhang Y. Bioaccumulation of perfluoroalkyl carboxylates (PFCAs) and perfluoroalkane sulfonates (PFASs) by earthworms (*Eisenia Fetida*) in soil. *Environ Pollut.* 2013;179:45–52. <https://doi.org/10.1016/j.envpol.2013.04.002>
13. Martín J, Hidalgo F, García-Corcoles MT, et al. Bioaccumulation of perfluoroalkyl substances in marine echinoderms: results of laboratory-scale experiments with *Holothuria Tubulosa* Gmelin, 1791. *Chemosphere* 2019;215:261–271. <https://doi.org/10.1016/j.chemosphere.2018.10.037>
14. Haukås M, Berger U, Hop H, Gulliksen B, Gabrielsen GW. Bioaccumulation of per- and polyfluorinated alkyl substances (PFAS) in selected species from the Barents sea food web. *Environ Pollut.* 2007;148(1):360–371. <https://doi.org/10.1016/j.envpol.2006.09.021>
15. Munoz G, Desrosiers M, Vetter L, et al. Bioaccumulation of. *Environ Sci Technol.* 2020;54(3):1687–1697. <https://doi.org/10.1021/acs.est.9b05102>
16. McDonough CA, Choyke S, Ferguson PL, DeWitt JC, Higgins CP. Bioaccumulation of novel per- and polyfluoroalkyl substances in mice dosed with an aqueous film-forming foam. *Environ Sci Technol.* 2020;54(9):5700–5709. <https://doi.org/10.1021/acs.est.0c00234>
17. Nelson JW, Hatch EE, Webster TF. Exposure to polyfluoroalkyl chemicals and cholesterol, body weight, and insulin resistance in the general U.S. population. *Environ Health Perspect.* 2010;118(2):197–202. <https://doi.org/10.1289/ehp.0901165>
18. Lopez-Espinosa M-J, Mondal D, Armstrong B, Bloom MS, Fletcher T. Thyroid function and perfluoroalkyl acids in children living near a chemical plant. *Environ Health Perspect.* 2012;120(7):1036–1041. <https://doi.org/10.1289/ehp.1104370>
19. Steenland K, Tinker S, Frisbee S, Ducatman A, Vaccarino V. Association of perfluorooctanoic acid and perfluorooctane sulfonate with serum lipids among adults living near a chemical plant. *Am J Epidemiol.* 2009;170(10):1268–1278. <https://doi.org/10.1093/aje/kwp279>
20. Olsen GW, Burris JM, Burlew MM, Mandel JH. Epidemiologic assessment of worker serum perfluorooctanesulfonate (PFOS) and perfluorooctanoate (PFOA) concentrations and medical surveillance examinations. *J Occup Environ Med.* 2003;45(3):260–270. <https://doi.org/10.1097/01.jom.0000052958.59271.10>
21. Darrow LA, Stein CR, Steenland K. Serum perfluorooctanoic acid and perfluorooctane sulfonate concentrations in relation to birth outcomes in the Mid-Ohio Valley, 2005–2010. *Environ Health Perspect.* 2013;121(10):1207–1213. <https://doi.org/10.1289/ehp.1206372>
22. Szilagyi JT, Avula V, Fry RC. Perfluoroalkyl substances (PFAS) and their effects on the placenta, pregnancy, and child development: a potential mechanistic role for placental peroxisome proliferator-activated receptors (PPARs). *Curr Environ Health Rep.* 2020;7(3):222–230. <https://doi.org/10.1007/s40572-020-00279-0>
23. Steenland K, Zhao L, Winquist A, Parks C. Ulcerative colitis and perfluorooctanoic acid (PFOA) in a highly exposed population of community residents and workers in the Mid-Ohio Valley. *Environ Health Perspect.* 2013;121(8):900–905. <https://doi.org/10.1289/ehp.1206449>
24. Barry V, Winquist A, Steenland K. Perfluorooctanoic acid (PFOA) exposures and incident cancers among adults living near a chemical plant. *Environ Health Perspect.* 2013;121(11–12):1313–1318. <https://doi.org/10.1289/ehp.1306615>
25. Shearer JJ, Callahan CL, Calafat AM, et al. Serum concentrations of per- and polyfluoroalkyl substances and risk of renal cell carcinoma. *J Natl Cancer Inst.* <https://doi.org/10.1093/jnci/djaa143>
26. Sunderland EM, Hu XC, Dassuncao C, Tokranov AK, Wagner CC, Allen JG. A review of the pathways of human exposure to poly- and perfluoroalkyl substances (PFASs) and present understanding of health effects. *J Expo Sci Environ Epidemiol.* 2019;29(2):131–147. <https://doi.org/10.1038/s41370-018-0094-1>
27. Grandjean P, Heilmann C, Weihe P, et al. Estimated exposures to perfluorinated compounds in infancy predict attenuated vaccine antibody concentrations at age 5-years. *J Immunotoxicol.* 2017;14(1):188–195. <https://doi.org/10.1080/1547691X.2017.1360968>
28. DeWitt JC, Peden-Adams MM, Keller JM, Germolec DR. Immunotoxicity of perfluorinated compounds: recent developments. *Toxicol Pathol.* 2012;40(2):300–311. <https://doi.org/10.1177/0192623311428473>
29. McDonough CA, Ward C, Hu Q, Vance S, Higgins CP, DeWitt JC. Immunotoxicity of an electrochemically fluorinated aqueous film-forming foam. *Toxicol Sci.* 2020;178(1):104–114. <https://doi.org/10.1093/toxsci/kfaa138>
30. Bălan SA, Mathrani VC, Guo DF, Algazi AM. Regulating PFAS as a chemical class under the California Safer Consumer Products program. *Environ Health Perspect.* 2021;129(2):25001. <https://doi.org/10.1289/EHP7431>
31. Brennan NM, Evans AT, Fritz MK, Peak SA, von Holst HE. Trends in the regulation of per- and polyfluoroalkyl substances (PFAS):

- a scoping review. *IJERPH*. 2021;18(20):10900. <https://doi.org/10.3390/ijerph182010900>
32. McDonald FA. Omnipresent chemicals: TSCA preemption in the wake of PFAS contamination. *Environ Claims J*. 2020;32(4):265–288. <https://doi.org/10.1080/10406026.2020.1754631>
  33. Denison RA. Ten essential elements in TSCA reform. *Envtl L Rep News Analysis*. 2009;39:10020.
  34. Ekey K. Tick toxic: the failure to clean up TSCA poisons public health and threatens chemical innovation. *Wm Mary Envtl L Pol Rev*. 2013;38:169.
  35. Aurich D, Miles O, Schymanski EL. Historical exposomics and high resolution mass spectrometry. *Exposome*. 2021;1(1):osab007. <https://doi.org/10.1093/exposome/osab007>
  36. Getzinger GJ, Ferguson PL. Illuminating the exposome with high-resolution accurate-mass mass spectrometry and nontargeted analysis. *Curr Opin Environ Sci Health*. 2020;15:49–56. <https://doi.org/10.1016/j.coesh.2020.05.005>
  37. Andra SS, Austin C, Patel D, Dolios G, Awawda M, Arora M. Trends in the application of high-resolution mass spectrometry for human biomonitoring: an analytical primer to studying the environmental chemical space of the human exposome. *Environ Int*. 2017;100:32–61. <https://doi.org/10.1016/j.envint.2016.11.026>
  38. Koelmel JP, Ulmer CZ, Jones CM, Yost RA, Bowden JA. Common cases of improper lipid annotation using high-resolution tandem mass spectrometry data and corresponding limitations in biological interpretation. *Biochim Biophys Acta Mol Cell Biol Lipids*. 2017;1862(8):766–770. <https://doi.org/10.1016/j.bbalip.2017.02.016>
  39. Köfeler HC, Eichmann TO, Ahrends R, et al. Quality control requirements for the correct annotation of lipidomics data. *Nat Commun*. 2021;12(1):4771. <https://doi.org/10.1038/s41467-021-24984-y>
  40. Chaleckis R, Meister I, Zhang P, Wheelock CE. Challenges, progress and promises of metabolite annotation for LC-MS-based metabolomics. *Curr Opin Biotechnol*. 2019;55:44–50. <https://doi.org/10.1016/j.copbio.2018.07.010>
  41. Matsuda F. Technical challenges in mass spectrometry-based metabolomics. *Mass Spectrom (Tokyo)*. 2016;5(2):S0052. <https://doi.org/10.5702/massspectrometry.S0052>
  42. Koelmel JP, Paige MK, Aristizabal-Henao JJ, et al. Toward comprehensive per- and polyfluoroalkyl substances annotation using fluoromatch software and intelligent high-resolution tandem mass spectrometry acquisition. *Anal Chem*. 2020;92(16):11186–11194. <https://doi.org/10.1021/acs.analchem.0c01591>
  43. Koelmel JP, Stelben P, McDonough CA, et al. FluoroMatch 2.0—making automated and comprehensive non-targeted PFAS annotation a reality. *Anal Bioanal Chem*. 2022;414(3):1201–1215. <https://doi.org/10.1007/s00216-021-03392-7>
  44. Jacob P, Wang R, Ching C, Helbling DE. Evaluation, optimization, and application of three independent suspect screening workflows for the characterization of PFASs in water. *Environ Sci Process Impacts*. 2021;23(10):1554–1565. <https://doi.org/10.1039/D1EM00286D>
  45. Nason SL, Koelmel J, Zuverza-Mena N, et al. Software comparison for nontargeted analysis of PFAS in AFFF-contaminated soil. *J Am Soc Mass Spectrom*. 2021;32(4):840–846. <https://doi.org/10.1021/jasms.0c00261>
  46. Anderson RH, Thompson T, Stroo HF, Leeson A. US Department of Defense-funded fate and transport research on per- and polyfluoroalkyl substances at aqueous film-forming foam-impacted sites. *Environ Toxicol Chem*. 2021;40(1):37–43. <https://doi.org/10.1002/etc.4694>
  47. McDonough CA, Choyke S, Barton KE, et al. Unsaturated PFOS and other PFASs in human serum and drinking water from an AFFF-impacted community. *Environ Sci Technol*. 2021;55(12):8139–8148. <https://doi.org/10.1021/acs.est.1c00522>
  48. Koelmel JP, Kroeger NM, Gill EL, et al. 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>
  49. 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 Bioinform*. 2010;11(1):395. <https://doi.org/10.1186/1471-2105-11-395>
  50. Chambers MC, Maclean B, Burke R, et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol*. 2012;30(10):918–920. <https://doi.org/10.1038/nbt.2377>
  51. Schymanski EL, Jeon J, Gulde R, et al. Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environ Sci Technol*. 2014;48(4):2097–2098. <https://doi.org/10.1021/es5002105>
  52. Charbonnet JA, McDonough CA, Xiao F, et al. Communicating confidence of per- and polyfluoroalkyl substance identification via high-resolution mass spectrometry. *Environ Sci Technol Lett*. 2022;9(6):473–481. <https://doi.org/10.1021/acs.estlett.2c00206>