Interactive Software for Visualization of Non-Targeted Mass Spectrometry Data – FluoroMatch Visualizer

Jeremy P Koelmel[1◊], Paul Stelben[1◊], David Godri[2], Jiarong Qi[1], Carrie A. McDonough[3], David A. Dukes[3], Juan J. Aristizabal-Henao[4], John A. Bowden[4,5], Sandi Sternberg[6], Emma E Rennie[7], Krystal J. Godri Pollitt[1*]

[1]Yale School of Public Health, New Haven, CT, United States
[2]3rd Floor Solutions, Toronto, Canada
[3]Department of Civil Engineering, Stony Brook University, Stony Brook, NY, United States
[4]Department of Chemistry, University of Florida, Gainesville, FL, United States
[5] Center for Environmental and Human Toxicology & Department of Physiological Sciences, University of Florida, Gainesville, FL, United States
[6]Innovative Omics Inc., FL, United States
[7]Agilent Technologies Inc., Santa Clara, CA, United States

◊These authors contributed equally

*Corresponding Author
Email: krystal.pollitt@yale.edu
Address: 60 College Street, Room 523, New Haven, CT, 06520

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 use of non-targeted 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 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 meta-data. 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 non-targeted data-processing and interpretation throughout the PFAS scientific community.

Keywords: mass spectrometry, PFAS, software, non-targeted analysis, aqueous film forming foam, liquid chromatography

Declarations

Conflicts of interest/Competing interests:
All authors confirm no conflict of interest.

Availability of data and material:
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

(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 after navigating to https://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.
YoutTube Tutorials can be found here:
https://youtube.com/playlist?list=PLZtU6nmcTb5kv9gvYuikgDXc_I_mhxTqO

Acknowledgements
A large array of PFAS standards used to develop libraries were donated by SynQuest Labs, Inc and Oakwood Products, Inc.

1. Introduction:

Innovation in the chemical industry often leads to better consumer products,[1–3] longer life expectancy, new scientific discoveries, and higher interconnectivity of society, to name a few benefits. There are revolutionary new prospects of technology[2] enabled by chemistry. These revolutionary benefits are actualized through a diversification of synthesized chemicals. Alongside the benefits of novel chemicals are the unknown health impacts of exposure[2,4]. Databases of chemicals which may be found in the environment range from 360,000[5] to over 800,000 compounds[6], 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).

There are thousands of known PFAS, which combined with the fact that they are pervasive,[7–10] 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 perfluorooctane sulfonic acid (PFOS). Most PFAS bioaccumulate[11–16] and certain PFAS have been linked to high cholesterol[17,18] and triglycerides[19,20], thyroid disease[18], pregnancy-induced hypertension[21,22], ulcerative colitis[23], cancers[24,25], and a weakened immune system[26–29]. While there has been increased regulatory[30,31], academic, and public awareness, scrutiny, and action to limit PFAS exposures, novel PFAS chemicals often are outside the scope of these efforts[32–34]. To accompany the growing need for regulatory changes, tools, such as high-resolution mass spectrometry,[35–37] are needed to provide 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)[35–37]. The specificity and universality of mass spectral evidence allows for non-targeted analysis: the analysis of thousands to hundreds of thousands of signals representing individual chemicals or structurally similar isobars or isomers. 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[38–41].

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. 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 has been shown for FluoroMatch Flow, using class-based fragmentation rules from standards,[42–45] annotations using fragment screening, homologous series detection, and other more non-targeted approaches are more susceptible to very high false positive rates.[45] Most features have modest to low confidence in the noted annotations and hence, 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).

2. Methods:

2.1. FluoroMatch Visualizer integration with FluoroMatch Flow / Modular

FluoroMatch Visualizer was developed to work with outputs from FluoroMatch Flow or Modular[44,45]. 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, formulaes, 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.

2.2. 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[46]. The sample was diluted 100,000 fold in 70:30 water:methanol (Fisher Scientific Optima® LCMS-grade). The diluted sample was injected four 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). Blanks were acquired every other injection for blank filtering. PFAS were detected in negative electrospray ionization mode. Data was acquired from m/z 100-1100, with MS/MS collision energy set to 0, 25, and 40 eV. Source parameters and further acquisition parameters for this dataset have been previously described[45,47].

3. Results and Discussion:

3.1. FluoroMatch Workflow

FluoroMatch and FluoroMatch Visualizer were developed to be comprehensive, modifiable, user-friendly, and open source. The 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. 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,[48] MSConvert,[49] 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**). Confidence assignments are based on an in-house scoring framework[44,45] (**Figure 2**) and the Schymanski scheme[50]. Comparison of FluoroMatch Flow with other software, including Compound Discoverer and EnviMass, shows unique advantages and disadvantages of various software.[42,43] Generally FluoroMatch Flow has similar or lower false positive rates among high confidence annotations and has similar or more comprehensive annotation when the same peak list is used, but vendor or open-source peak picking algorithms designed specifically for the methods / instruments deployed often have better peak picking coverage then 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. 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.
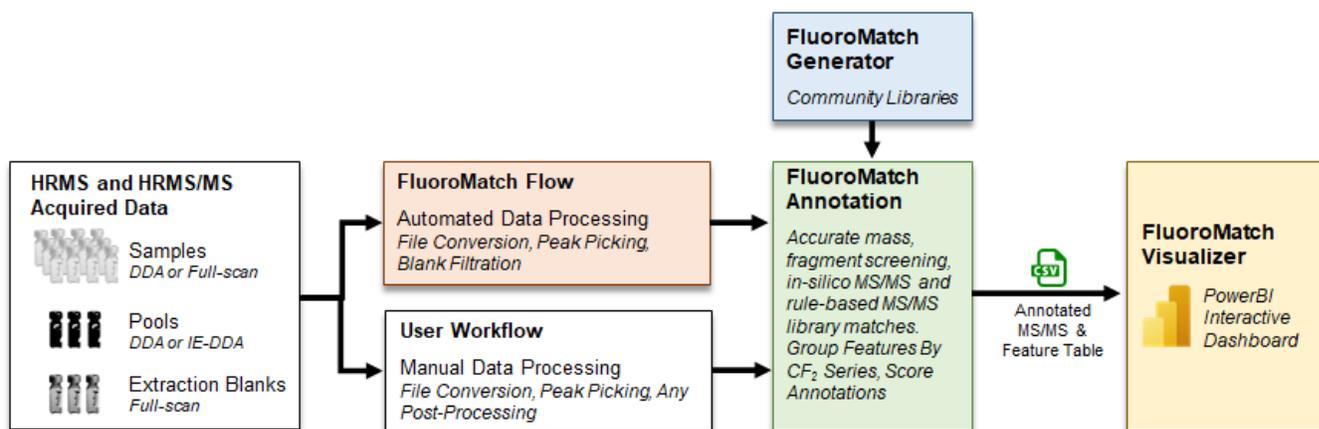
Figure 1: *The FluoroMatch Suite covers the entire PFAS non-targeted data-processing and manual validation workflow.* User acquisition and data-processing workflow using FluoroMatch Modular, FluoroMatch Flow, and FluoroMatch Visualizer are shown. Acronyms: data-dependent analysis (DDA), iterative exclusion (IE), fragmentation spectra (MS/MS).

Score	Description	Score	Description
A+	Confirmed by standards (RT, MS/MS, MS)	C+	●←→○ (CF$_2$)$_n$ In series w/ confident ID
A	●←→○ (CF$_2$)$_n$ Confident ID (MS/MS, MS) and series	C	⊘←→○←→◐ (CF$_2$)$_n$ (CF$_2$)$_n$ In series (3+) w/ any tentative ID
A-	● Confident ID (MS/MS, MS)	C-	⊘←→○ or ◐←→○ (CF$_2$)$_n$ (CF$_2$)$_n$ In series w/ any tentative ID
B+	⊜←→○ (CF$_2$)$_n$ Tentative ID (MS/MS, MS) and series	D+	○←→○←→○ (CF$_2$)$_n$ (CF$_2$)$_n$ MD or MS and in series (3+)
B	○←→○ (CF$_2$)$_n$ Tentative ID (MS/MS, MS) and series	D	○←→○←→○ (CF$_2$)$_n$ (CF$_2$)$_n$ In series (3+)
B- *B--	⊜ or ○ Any tentative ID (MS/MS, MS)	D-	○ MD or MS

●⊜
Derived MS/MS rules from class-based standards

[C$_3$F$_7$]⁻
[C$_2$F$_5$]⁻   [C$_4$F$_9$]⁻

[C$_5$F$_{11}$]⁻
[C$_6$F$_{13}$]⁻ [M-CO$_2$H]⁻
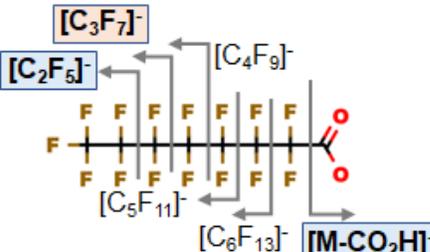
☐ All must be confirmed
☐ One must be confirmed

● Confident Annotation (MS/MS rules derived from class-based standards)
Example: PFCA
MS: [M-H]⁻
MS/MS: [C$_3$F$_7$]⁻ AND ([C$_2$F$_5$]⁻ OR [M-COOH]⁻)

⊜ Tentative Annotation (1+ fragment from class-based standards)
Example: PFCA
MS: [M-H]⁻
MS/MS: [M-H]-, [M-CO2H]-, [C2F5]-, [C3F7]-, [C4F9]-, [C5F11]- OR [C6F13]-

◐
In-silico fragments from substructures
FC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(=O)O
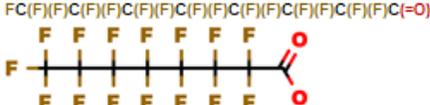
C(=O)O: [CO2F]-, [M-CO2H]-
C(F)(F): [CF2]-, [CF3]-
C(F)(F)C(F)(F): [C2F5]-
C(F)(F)C(F)(F)C(F)(F): [C3F5]-, [C3F7]-
C(F)(F)C(F)(F)C(F)(F)C(F)(F): [C4F7]-, [C4F9]-

◐ Tentative Annotation (1+ F-containing fragment from in-silico MS/MS)

Figure 2: FluoroMatch includes a systematic scoring framework to communicate confidence for every single feature, alongside also reporting confidence levels via the Schymanski schema. All scores are shown alongside visual and descriptive definitions.

*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.

3.2. FluoroMatch Visualizer Description and Possible Workflow

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 splicers and filters can be developed. Here we describe a simple workflow for determining false positives and true positives, in section 3.3. this workflow will be exemplified using an AFFF mixture.

Step 1) Determine series with high confidence series
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). The remaining series containing one or more confident annotation can then be selected in the legend (**Figure 3B** or **Figure 3C**). 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 LipidMatch Flow (the file with FIN in the name), a new column can be added for comments and for flagging potential false positives or true positives. The file should be saved with a different name in excel format (.xlsx). Now, select the first confident series, and look for patterns in retention time vs *m/z* (**Figure 3B**). Members of the same homologous series should follow a diagonal line. Outliers that do not fall along this diagonal can then be noted in the annotated feature table as likely false positives. The remaining members will now have higher confidence, given that they fall withing a clear retention time pattern even for features without MS/MS. This process can be iterated across all remaining high confidence series. Depending on the goals of the study this process can also be done for all series, including series without structural annotations, although this may be too time intensive. Because FluoroMatch is designed to define series using the mass defect plot (**Figure 3C**) no false positives will be observed for series in those plots.

Step 3) Determine false positives using MS/MS spectra
For each series selected the MS/MS will be combined (summed) for all individual PFAS species for that series which have MS/MS acquired (**Figure 3E**). Clear fragmentation patterns can be observed in this manner, with the most abundant and common fragments for this PFAS class readily noted. Any features with MS/MS but without a confident annotation (score of A) can then be elevated if they have the appropriate fragments and fall into the observed fragmentation pattern. Furthermore, false positives which do not have expected fragments can be noted for that class. This information can be noted in a separate column in the final table in excel, and flags for both retention time and MS/MS can be referenced to remove false positives which do not belong to the series.

Step 4) Discovery of PFAS structures without confident assignments
After flagging potential false positives and true positives for confident annotations, then all other annotations with some PFAS MS/MS evidence can be screened. Using the same methodology as in Step 1 through Step 3, series with B+, B, or B- can be determined (and series with A's can be removed from this list) and false positives flagged. For certain features with a score of B+, B, or B-, structures will be assigned using in-silico approaches or fragment screening. MS/MS evidence should be carefully considered and benchmarked against literature or standards for validation of these structures. Furthermore, isotopic pattern matching and formula prediction (not included in the platform) can be used to validate formula.

Step 5) Discovering series and PFAS chemical classes using fragment screening
In addition to investigating by predefined series and scores, series can be determined by filtering based on fragments which are most abundant across the entire dataset, and which represent chemical moieties of interest, using the Fragment Screening Filter **(Figure 3D)**. After these features are discovered, MS/MS evidence (**Figure 3E**) can be investigated to ensure these fragments are of high abundance for each feature, and then new series can be 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.

Figure 3. *The FluoroMatch interface is designed so that all relevant information can be observed simultaneously.* The user view of the FluoroMatch Visualizer interface is shown. The interface consists of three filters (by MS/MS file, score and chemical series; **Figure 3A**), three visuals (*m/z* vs retention time, normalized mass defect plot, and MS/MS spectra; **Figure 3B**, **Figure 3C**, and **Figure 3E**, respectively), and two tables (table of fragments, and table of annotated features; **Figures 3D**, and **Figure 3F**, respectively).

3.3. Application of FluoroMatch Visualizer to identify PFAS in a Legacy AFFF Mixture

Non-targeted 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 (3,686 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 (**Figure 4** and **Figure 5**).
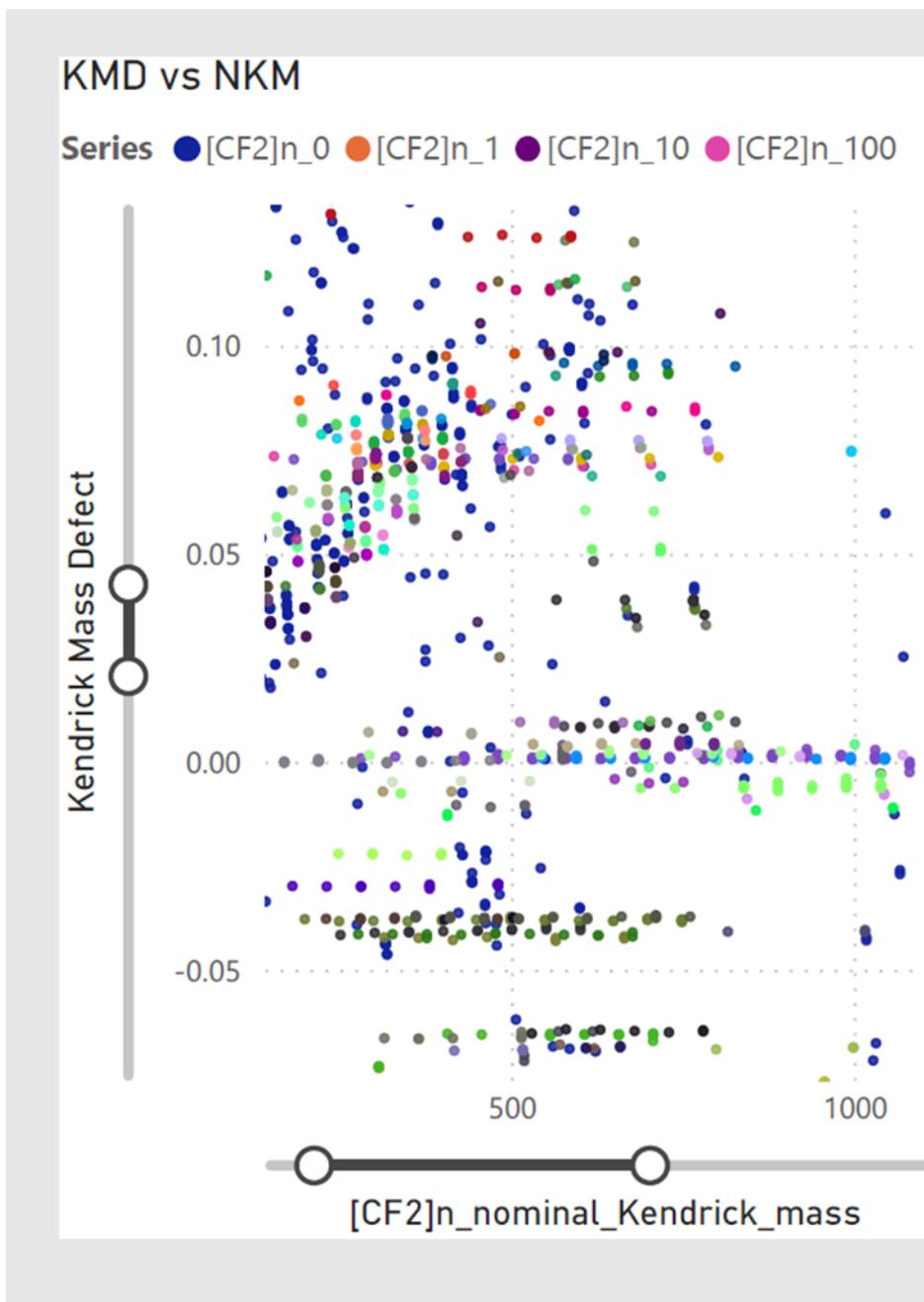
Figure 4. *Several PFAS series can be discovered by looking at certain regions of mass defect plots normalized to CF$_2$. Without filtering data by series or score the visual is*

overly complex and challenging to make sense of. Mass defect plot (normalized to CF2) 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.
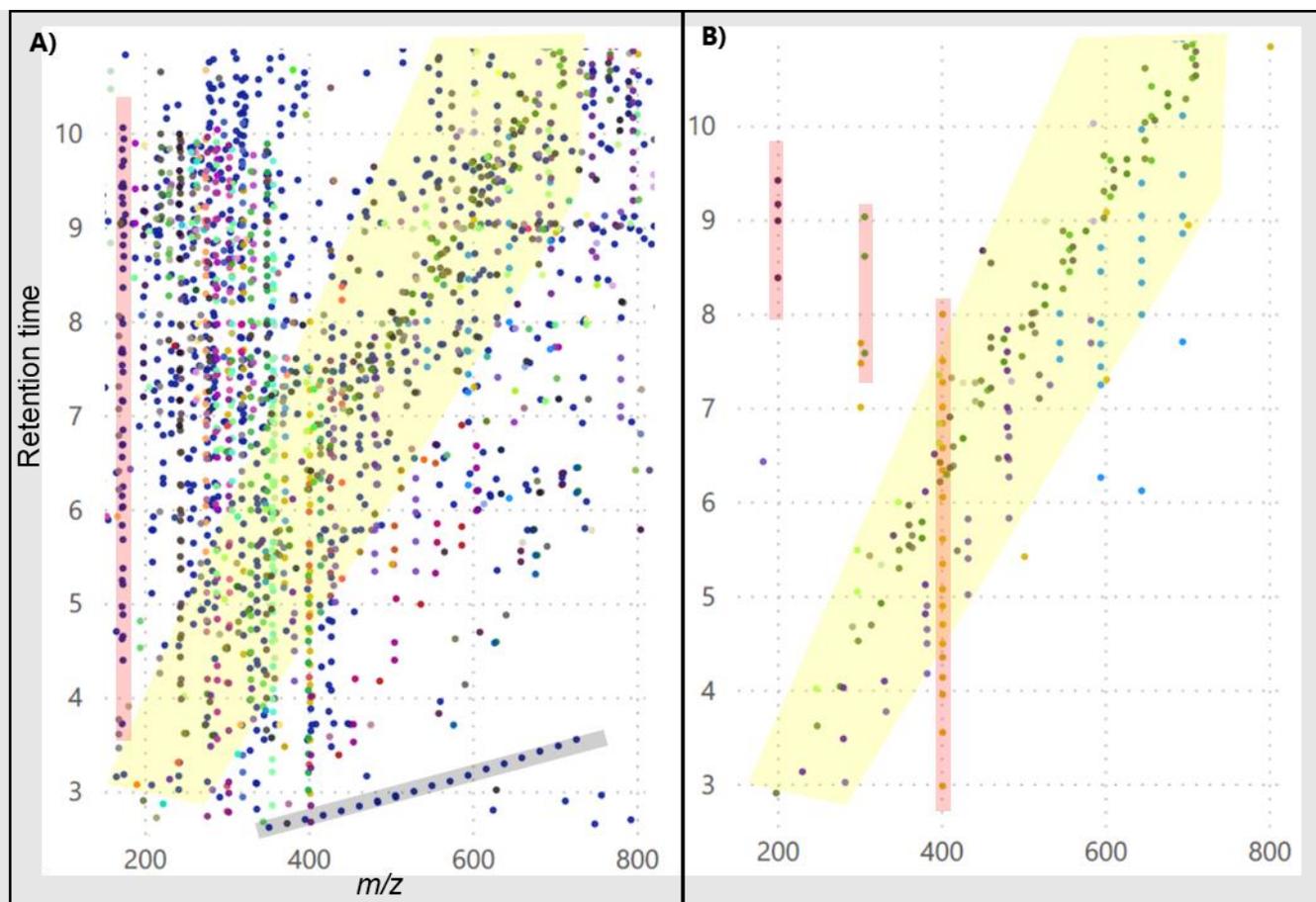

Figure 5. *Retention time vs m/z plots are overly complex without filtering; certain PFAS rich regions, blank signal, and other non-PFAS series can readily be discovered using these plots.* Retention time vs *m/z* for all features (**Figure 5A**) and for features that belong to high confidence series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

Because of the complexity and richness of non-targeted 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 are described previously[45]. 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[45].

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)[44,45] 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 **Figure 6C**).
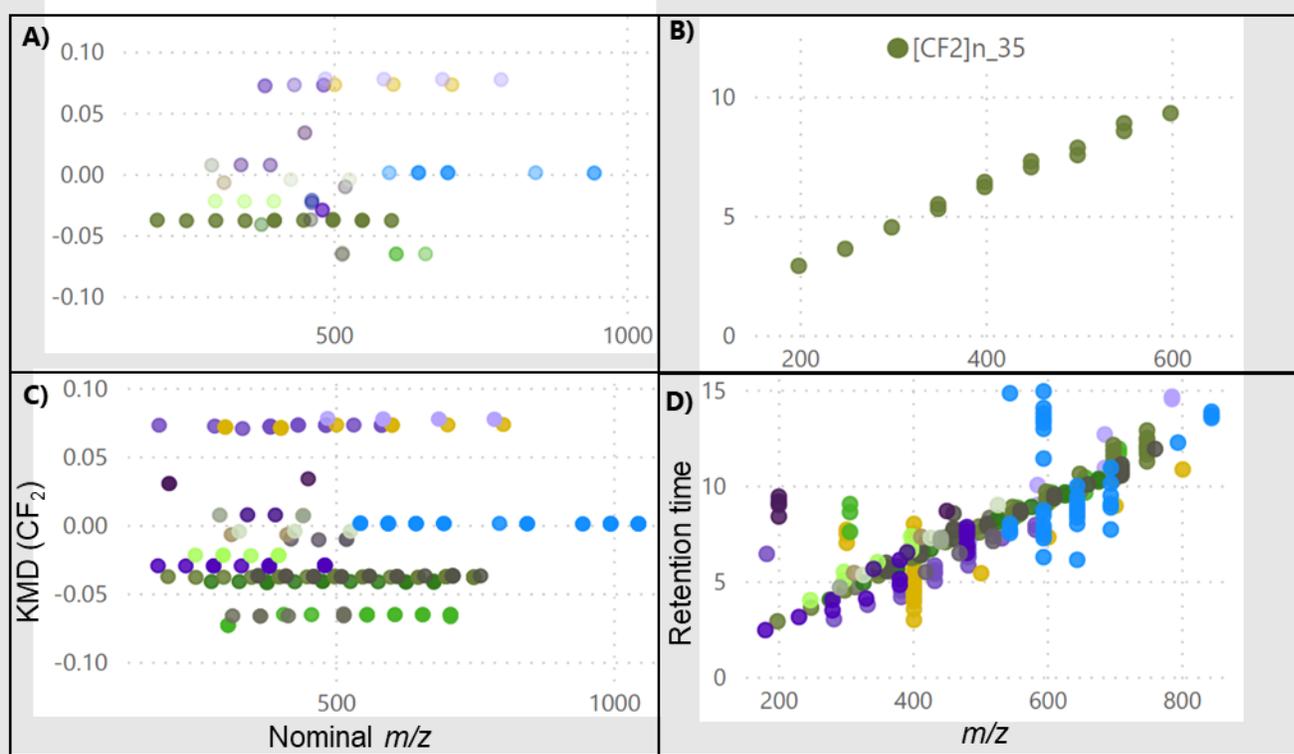
Figure 6. *By filtering to only include those series with the most confident annotations, true positives and false negatives can manually be determined using clear trends.* Shown above are kendrick mass defect plots normalized to CF2 (KMD (CF2)) for annotations with high confidence (**Figure 6A**) and series containing one or more high confident annotations (**Figure 6C**). As an example, for high confidence annotation, the series for perfluoroalkyl sulfonic acids (PFSA) are selected (**Figure 6A**) and the retention time vs *m/z* plot is shown for this series (**Figure 6B**). Retention time vs *m/z* plots are also shown for all high confidence series' (**Figure 6D**).

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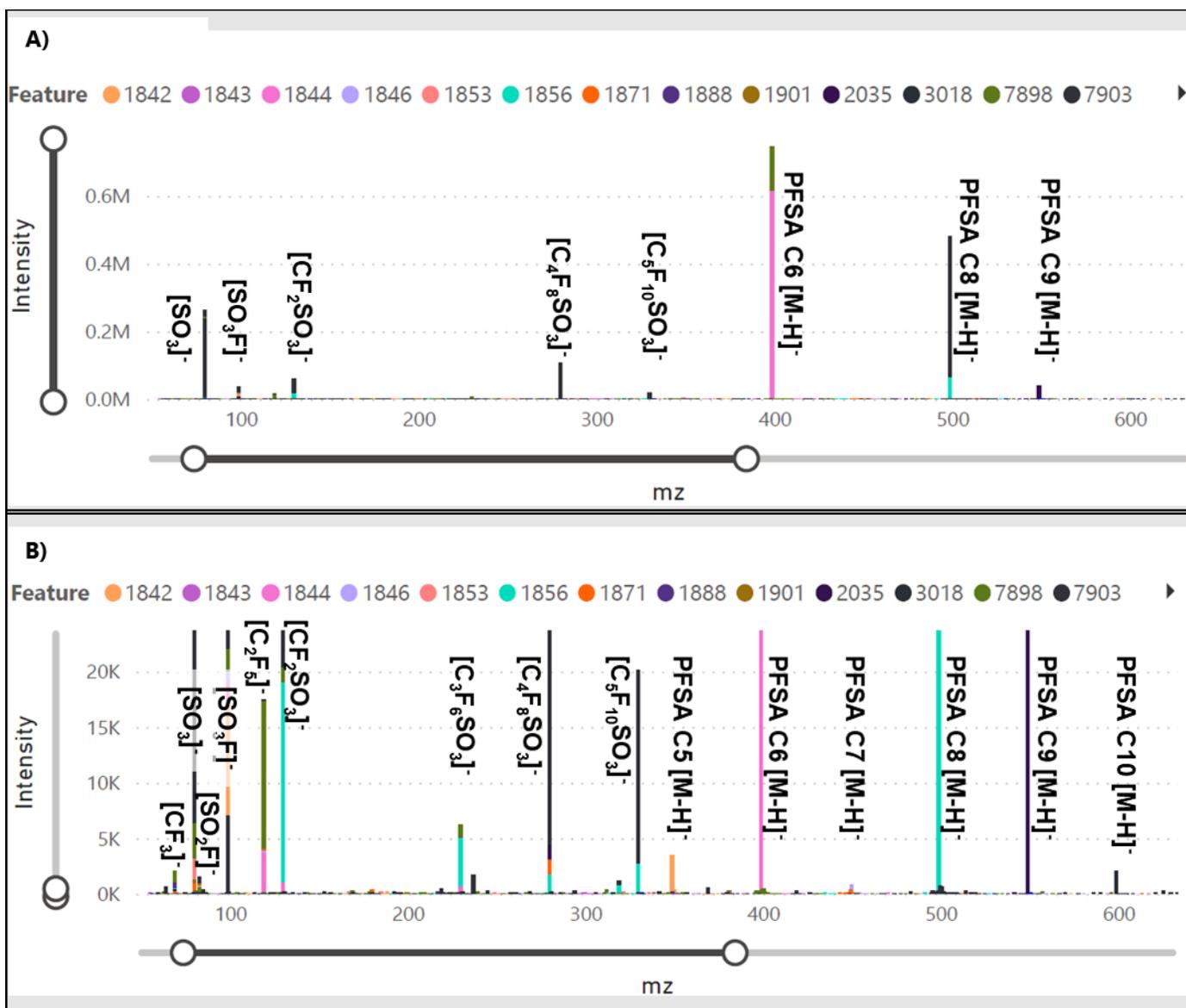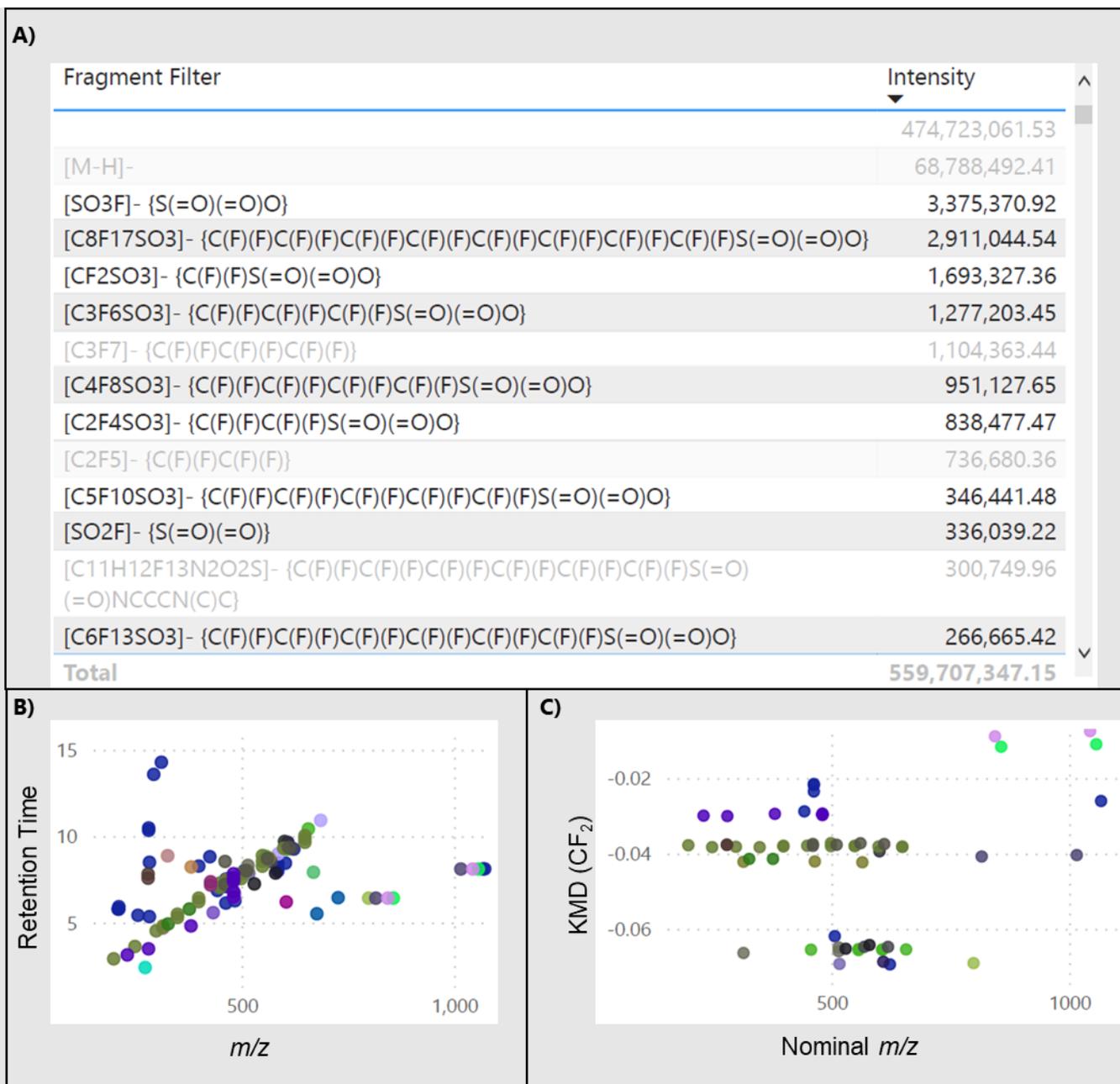
Figure 7. *Averaging spectra across series can be readily done in the visualizer platform and used to determine common fragmentation patterns and false positives.* Combined mass spectra after highlighting the perfluoroalkyl sulfonic acids (PFSA) series (**Figure 5A**), zoomed out (**Figure 7A**) and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

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 be shown, which contain these annotated fragments. This can aid in grouping compounds which contain certain moieties. For example, $[SO_3F]^-$ and $[SO_2F]^-$ may be selected to find PFAS with sulfonic acid groups, and $[CF_3]^-$ and $[C_2F_5]^-$ 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 $[SF_5]^-$ 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 $[SF_5]^-$ fragment, 83 features were determined (Figure 9C and Figure 9D). 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).

A)

Fragment Filter	Intensity
	474,723,061.53
[M-H]-	68,788,492.41
[SO3F]- {S(=O)(=O)O}	3,375,370.92
[C8F17SO3]- {C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)S(=O)(=O)O}	2,911,044.54
[CF2SO3]- {C(F)(F)S(=O)(=O)O}	1,693,327.36
[C3F6SO3]- {C(F)(F)C(F)(F)C(F)(F)S(=O)(=O)O}	1,277,203.45
[C3F7]- {C(F)(F)C(F)(F)C(F)(F)}	1,104,363.44
[C4F8SO3]- {C(F)(F)C(F)(F)C(F)(F)C(F)(F)S(=O)(=O)O}	951,127.65
[C2F4SO3]- {C(F)(F)C(F)(F)S(=O)(=O)O}	838,477.47
[C2F5]- {C(F)(F)C(F)(F)}	736,680.36
[C5F10SO3]- {C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)S(=O)(=O)O}	346,441.48
[SO2F]- {S(=O)(=O)}	336,039.22
[C11H12F13N2O2S]- {C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)S(=O)(=O)NCCCN(C)C}	300,749.96
[C6F13SO3]- {C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)S(=O)(=O)O}	266,665.42
Total	559,707,347.15

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 (**Figure 8A**), sulfonic acid related fragments are selected (**Figure 8A**) to identify all PFAS series (**Figure 8C**) 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 (**Figure 8B**).
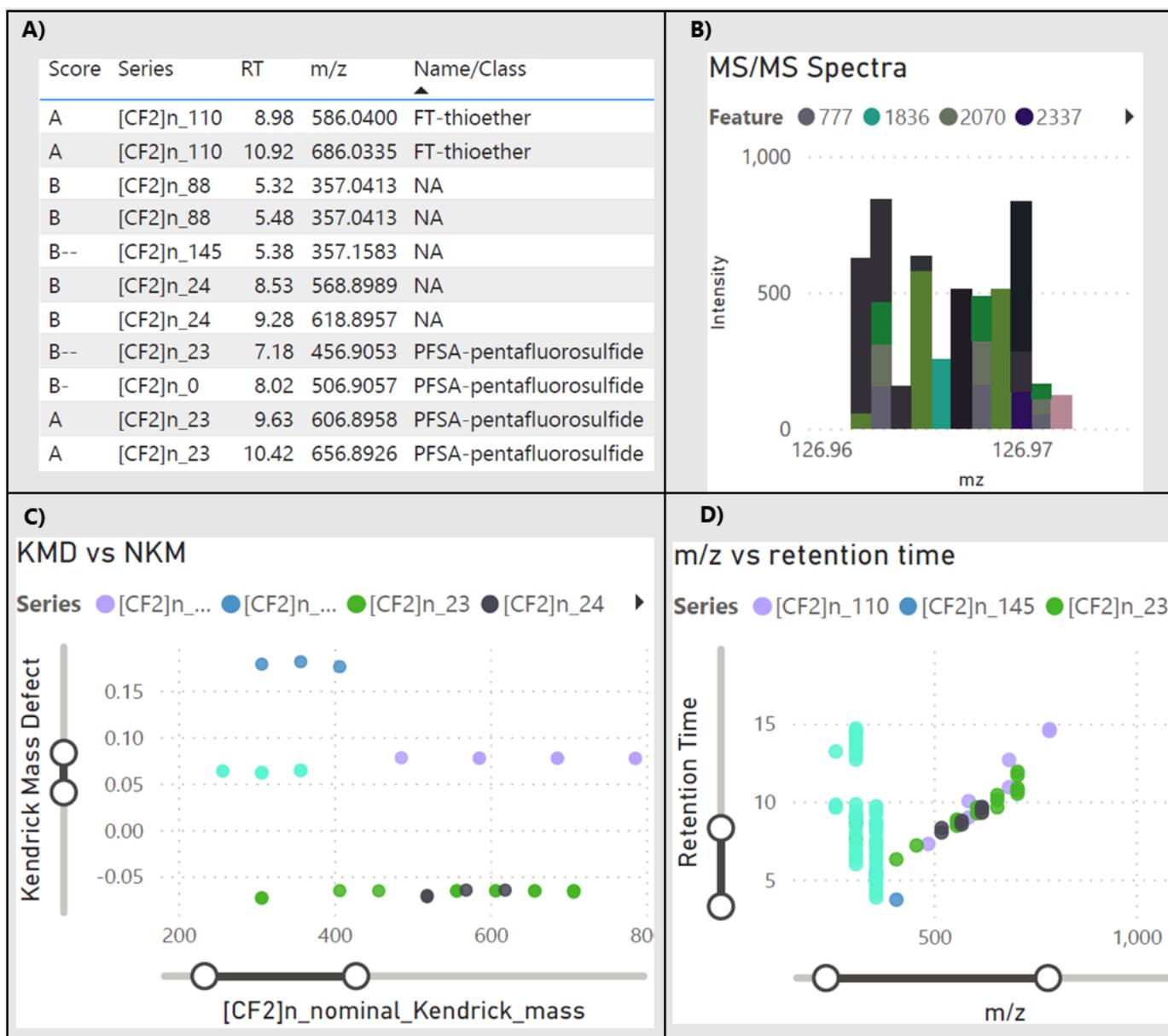
Figure 9. *Certain fragments are very specific to chemical moieties and hence can be advantageous to use in fragment screening, for example the [SF5]- fragment is specific to pentafluosulfide containing species found in AFFFs*. Features (**Figure 9A**) and their scores, homologous series, retention time (RT), mass to charge ratio (*m/z*), and level A annotations, filtered by performing fragment screening on the [SF5]- fragment are shown. [SF5]- fragment intensities from different fragment scans / features are shown within a 5-ppm window (**Figure 9B**), 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**).

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 is 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**).[44,51] Note these plots are cumulative, showing all past acquisitions and new acquisitions. Furthermore, by summing and plotting the number of features and filtering by score, it's shown that the advantage dissipates for high confidence annotations after about the 4th injection, meaning fewer injections are needed if high confidence annotations are the goal (**Figure 10B-D**).
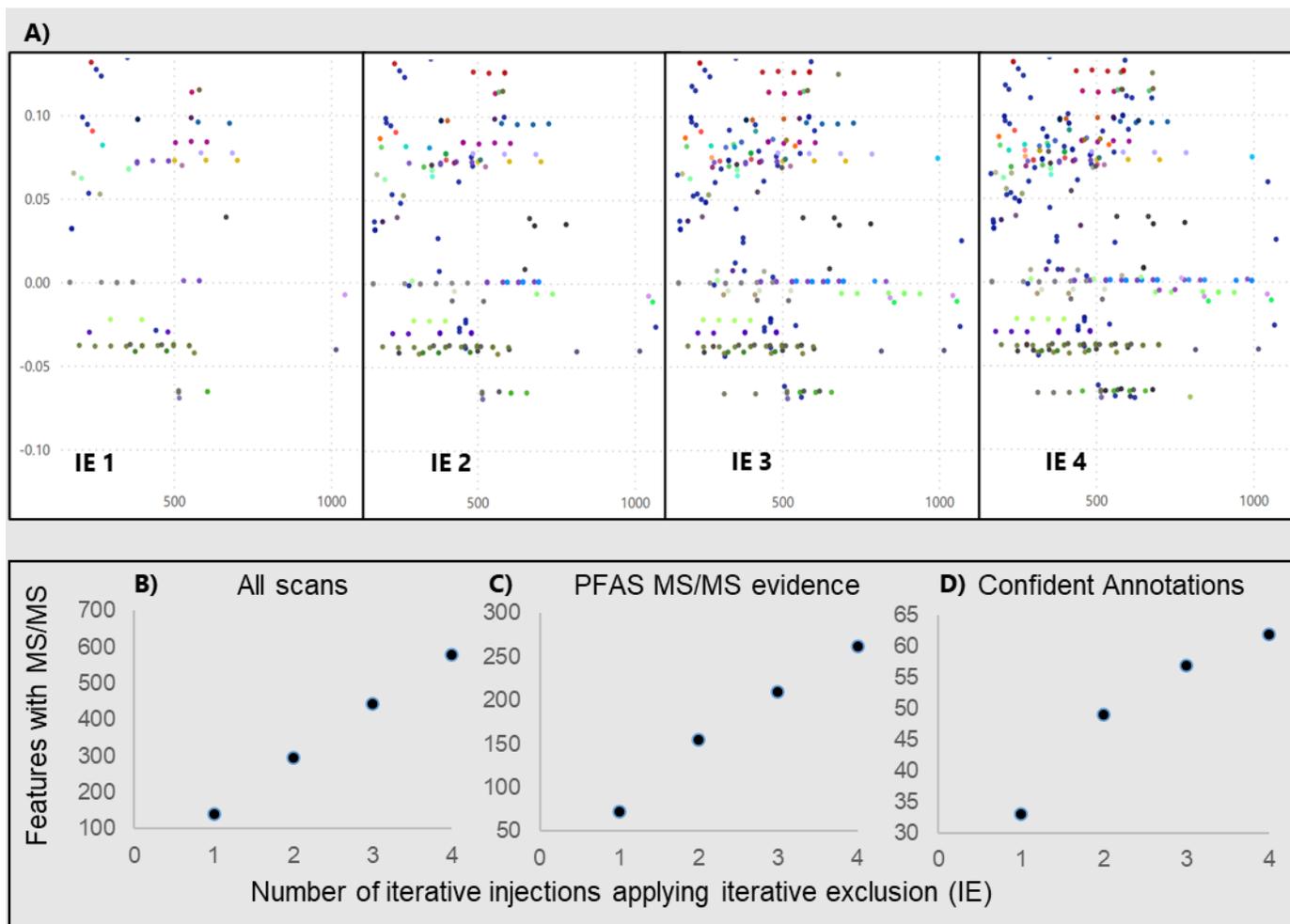
Figure 10. *By filtering by MS/MS file, the benefit of different MS/MS acquisition techniques can readily be explored.* Features with MS/MS after four iterations of iterative exclusion (IE), showing increased coverage in the PFAS region of kendrick mass defect plots with additional injections (**Figure 10A**). 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 (**Figure 10B**), after the 3rd and 4th injection new ions selected are of too low abundance for MS/MS to have high enough spectral quality for most

confident assignments (**Figure 10C** and **Figure 10D**). All scans represent any score A through E (**Figure 10B**), scans with PFAS MS/MS evidence represents scores of A through B (except B--) (**Figure 10C**), and confident annotations refer to those with scores of A and A- (**Figure 10D**).

4. 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. 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. Then 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 firefighting 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 an data sharing can improve trust and comprehension in the field of LC-HRMS/MS PFAS analysis.

References:

(1)   Easterlin, R. A. 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, G. M. Assumptions: Taking Chemistry in New Directions. *Angewandte Chemie International Edition* **2004**, *43* (28), 3632–3641. https://doi.org/10.1002/anie.200330076.
(3)   Aaron, H.; Schwartz, W. B. *Coping with Methuselah: The Impact of Molecular Biology on Medicine and Society*; Brookings Institution Press, 2004.
(4)   Naidu, R.; Biswas, B.; Willett, I. R.; Cribb, J.; Kumar Singh, B.; Paul Nathanail, C.; Coulon, F.; Semple, K. T.; Jones, K. C.; Barclay, A.; Aitken, R. J. Chemical Pollution: A Growing Peril and Potential Catastrophic Risk to Humanity. *Environment International* **2021**, *156*, 106616. https://doi.org/10.1016/j.envint.2021.106616.
(5)   Schymanski, E. L.; Kondić, T.; Neumann, S.; Thiessen, P. A.; Zhang, J.; Bolton, E. E. Empowering Large Chemical Knowledge Bases for Exposomics: PubChemLite Meets MetFrag. *Journal of Cheminformatics* **2021**, *13* (1), 19. https://doi.org/10.1186/s13321-021-00489-0.
(6)   Williams, A. J.; Grulke, C. M.; Edwards, J.; McEachran, A. D.; Mansouri, K.; Baker, N. C.; Patlewicz, G.; Shah, I.; Wambaugh, J. F.; Judson, R. S.; Richard, A. M. The CompTox Chemistry Dashboard: A Community Data Resource for Environmental Chemistry. *Journal of Cheminformatics* **2017**, *9* (1), 61. https://doi.org/10.1186/s13321-017-0247-6.
(7)   Olsen, G. W.; Mair, D. C.; Lange, C. C.; Harrington, L. M.; Church, T. R.; Goldberg, C. L.; Herron, R. M.; Hanna, H.; Nobiletti, J. B.; Rios, J. A.; Reagen, W. K.; Ley, C. A. Per- and Polyfluoroalkyl Substances (PFAS) in American Red Cross Adult Blood Donors, 2000–2015. *Environmental Research* **2017**, *157*, 87–95. https://doi.org/10.1016/j.envres.2017.05.013.
(8)   Göckener, B.; Weber, T.; Rüdel, H.; Bücking, M.; Kolossa-Gehring, M. Human Biomonitoring of Per- and Polyfluoroalkyl Substances in German Blood Plasma Samples from 1982 to 2019. *Environment International* **2020**, *145*, 106123. https://doi.org/10.1016/j.envint.2020.106123.
(9)   Smithwick, M.; Mabury, S. A.; Solomon, K. R.; Sonne, C.; Martin, J. W.; Born, E. W.; Dietz, R.; Derocher, A. E.; Letcher, R. J.; Evans, T. J.; Gabrielsen, G. W.; Nagy, J.; Stirling, I.; Taylor, M. K.; Muir, D. C. G. 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)   J, H.; N, K.; T, de K.; J, N.-A.-T.; A, S.; J, A.; K, J. An Overview of Emerging PFAS in Drinking Water Worldwide. *Environmental Epidemiology* **2019**, *3*, 162–163. https://doi.org/10.1097/01.EE9.0000607564.20698.d8.

(11)   Langberg, H. A.; Breedveld, G. D.; Grønning, H. M.; Kvennås, M.; Jenssen, B. M.; Hale, S. E. 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 (PFSAs) by Earthworms (Eisenia Fetida) in Soil. *Environmental Pollution* **2013**, *179*, 45–52. https://doi.org/10.1016/j.envpol.2013.04.002.

(13)   Martín, J.; Hidalgo, F.; García-Corcoles, M. T.; Ibáñez-Yuste, A. J.; Alonso, E.; Vilchez, J. L.; Zafra-Gómez, A. 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, G. W. Bioaccumulation of Per- and Polyfluorinated Alkyl Substances (PFAS) in Selected Species from the Barents Sea Food Web. *Environmental Pollution* **2007**, *148* (1), 360–371. https://doi.org/10.1016/j.envpol.2006.09.021.

(15)   Munoz, G.; Desrosiers, M.; Vetter, L.; Vo Duy, S.; Jarjour, J.; Liu, J.; Sauvé, S. Bioaccumulation of Zwitterionic Polyfluoroalkyl Substances in Earthworms Exposed to Aqueous Film-Forming Foam Impacted Soils. *Environ. Sci. Technol.* **2020**, *54* (3), 1687–1697. https://doi.org/10.1021/acs.est.9b05102.

(16)   McDonough, C. A.; Choyke, S.; Ferguson, P. L.; DeWitt, J. C.; Higgins, C. P. 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, J. W.; Hatch, E. E.; Webster, T. F. 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, M. S.; 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, G. W.; Burris, J. M.; Burlew, M. M.; Mandel, J. H. Epidemiologic Assessment of Worker Serum Perfluorooctanesulfonate (PFOS) and Perfluorooctanoate (PFOA) Concentrations and Medical Surveillance

Examinations. *Journal of Occupational and Environmental Medicine* **2003**, *45* (3), 260–270. https://doi.org/10.1097/01.jom.0000052958.59271.10.

(21)　Darrow, L. A.; Stein, C. R.; 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, J. T.; Avula, V.; Fry, R. C. 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 Envir Health Rpt* **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, J. J.; Callahan, C. L.; Calafat, A. M.; Huang, W.-Y.; Jones, R. R.; Sabbisetti, V. S.; Freedman, N. D.; Sampson, J. N.; Silverman, D. T.; Purdue, M. P.; Hofmann, J. N. 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, E. M.; Hu, X. C.; Dassuncao, C.; Tokranov, A. K.; Wagner, C. C.; Allen, J. G. A Review of the Pathways of Human Exposure to Poly- and Perfluoroalkyl Substances (PFASs) and Present Understanding of Health Effects. *Journal of Exposure Science & Environmental Epidemiology* **2019**, *29* (2), 131–147. https://doi.org/10.1038/s41370-018-0094-1.

(27)　Grandjean, P.; Heilmann, C.; Weihe, P.; Nielsen, F.; Mogensen, U. B.; Timmermann, A.; Budtz-Jørgensen, E. 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, J. C.; Peden-Adams, M. M.; Keller, J. M.; Germolec, D. R. Immunotoxicity of Perfluorinated Compounds: Recent Developments. *Toxicol Pathol* **2012**, *40* (2), 300–311. https://doi.org/10.1177/0192623311428473.

(29)　McDonough, C. A.; Ward, C.; Hu, Q.; Vance, S.; Higgins, C. P.; DeWitt, J. C. Immunotoxicity of an Electrochemically Fluorinated Aqueous Film-Forming Foam. *Toxicological Sciences* **2020**, *178* (1), 104–114. https://doi.org/10.1093/toxsci/kfaa138.

(30)　B, ălan S. A.; Mathrani, V. C.; Guo, D. F.; Algazi, A. M. Regulating PFAS as a Chemical Class under the California Safer Consumer Products Program. *Environmental Health Perspectives 129* (2), 025001. https://doi.org/10.1289/EHP7431.

(31)　Brennan, N. M.; Evans, A. T.; Fritz, M. K.; Peak, S. A.; von Holst, H. E. Trends in the Regulation of Per- and Polyfluoroalkyl Substances (PFAS): A Scoping

Review. *International Journal of Environmental Research and Public Health* **2021**, *18* (20), 10900. https://doi.org/10.3390/ijerph182010900.

(32) McDonald, F. A. Omnipresent Chemicals: TSCA Preemption in the Wake of PFAS Contamination. *Environmental Claims Journal* **2020**, *32* (4), 265–288. https://doi.org/10.1080/10406026.2020.1754631.

(33) Denison, R. A. 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'y Rev.* **2013**, *38*, 169.

(35) Aurich, D.; Miles, O.; Schymanski, E. L. Historical Exposomics and High Resolution Mass Spectrometry. *Exposome* **2021**, *1* (1), osab007. https://doi.org/10.1093/exposome/osab007.

(36) Getzinger, G. J.; Ferguson, P. L. Illuminating the Exposome with High-Resolution Accurate-Mass Mass Spectrometry and Nontargeted Analysis. *Current Opinion in Environmental Science & Health* **2020**, *15*, 49–56. https://doi.org/10.1016/j.coesh.2020.05.005.

(37) Andra, S. S.; 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. *Environment International* **2017**, *100*, 32–61. https://doi.org/10.1016/j.envint.2016.11.026.

(38) Koelmel, J. P.; Ulmer, C. Z.; Jones, C. M.; Yost, R. A.; Bowden, J. A. Common Cases of Improper Lipid Annotation Using High-Resolution Tandem Mass Spectrometry Data and Corresponding Limitations in Biological Interpretation. *Biochim Biophys Acta* **2017**, *1862* (8), 766–770. https://doi.org/10.1016/j.bbalip.2017.02.016.

(39) Köfeler, H. C.; Eichmann, T. O.; Ahrends, R.; Bowden, J. A.; Danne-Rasche, N.; Dennis, E. A.; Fedorova, M.; Griffiths, W. J.; Han, X.; Hartler, J.; Holčapek, M.; Jirásko, R.; Koelmel, J. P.; Ejsing, C. S.; Liebisch, G.; Ni, Z.; O'Donnell, V. B.; Quehenberger, O.; Schwudke, D.; Shevchenko, A.; Wakelam, M. J. O.; Wenk, M. R.; Wolrab, D.; Ekroos, K. 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, C. E. Challenges, Progress and Promises of Metabolite Annotation for LC–MS-Based Metabolomics. *Current Opinion in Biotechnology* **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 Spectrometry* **2016**, *5* (2), S0052–S0052. https://doi.org/10.5702/massspectrometry.S0052.

(42) Jacob, P.; Wang, R.; Ching, C.; Helbling, D. E. Evaluation, Optimization, and Application of Three Independent Suspect Screening Workflows for the Characterization of PFASs in Water. *Environ. Sci.: Processes Impacts* **2021**, *23* (10), 1554–1565. https://doi.org/10.1039/D1EM00286D.

(43) Nason, S. L.; Koelmel, J.; Zuverza-Mena, N.; Stanley, C.; Tamez, C.; Bowden, J. A.; Godri Pollitt, K. J. 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.

(44) Koelmel, J. P.; Paige, M. K.; Aristizabal-Henao, J. J.; Robey, N. M.; Nason, S. L.; Stelben, P. J.; Li, Y.; Kroeger, N. M.; Napolitano, M. P.; Savvaides, T.; Vasiliou, V.; Rostkowski, P.; Garrett, T. J.; Lin, E.; Deigl, C.; Jobst, K.; Townsend, T. G.; Godri Pollitt, K. J.; Bowden, J. A. 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.

(45) Koelmel, J. P.; Stelben, P.; McDonough, C. A.; Dukes, D. A.; Aristizabal-Henao, J. J.; Nason, S. L.; Li, Y.; Sternberg, S.; Lin, E.; Beckmann, M.; Williams, A. J.; Draper, J.; Finch, J. P.; Munk, J. K.; Deigl, C.; Rennie, E. E.; Bowden, J. A.; Godri Pollitt, K. J. 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.

(46) Anderson, R. H.; Thompson, T.; Stroo, H. F.; Leeson, A. US Department of Defense–Funded Fate and Transport Research on Per- and Polyfluoroalkyl Substances at Aqueous Film–Forming Foam–Impacted Sites. *Environmental Toxicology and Chemistry* **2021**, *40* (1), 37–43. https://doi.org/10.1002/etc.4694.

(47) McDonough, C. A.; Choyke, S.; Ferguson, P. L.; DeWitt, J. C.; Higgins, C. P. 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.

(48) 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.

(49) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat Biotechnol* **2012**, *30* (10), 918–920. https://doi.org/10.1038/nbt.2377.

(50) 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.

(51) 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.