

Online and Offline Prioritization of Chemicals of Interest in Suspect Screening and Non-targeted Screening with High-Resolution Mass Spectrometry

Drew Szabo, Travis M. Falconer, Christine M. Fisher, Ted Heise, Allison L. Phillips, Gyorgy Vas, Antony J. Williams, and Anneli Kruve*




Cite This: *Anal. Chem.* 2024, 96, 3707–3716



Read Online

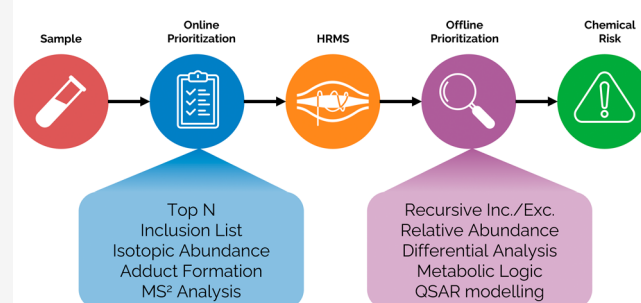
ACCESS |

 Metrics & More

 Article Recommendations

ABSTRACT: Recent advances in high-resolution mass spectrometry (HRMS) have enabled the detection of thousands of chemicals from a single sample, while computational methods have improved the identification and quantification of these chemicals in the absence of reference standards typically required in targeted analysis. However, to determine the presence of chemicals of interest that may pose an overall impact on ecological and human health, prioritization strategies must be used to effectively and efficiently highlight chemicals for further investigation. Prioritization can be based on a chemical's physicochemical properties, structure, exposure, and toxicity, in addition to its regulatory status. This Perspective aims to provide a framework for the strategies used for chemical prioritization that can be implemented to facilitate high-quality research and communication of results. These strategies are categorized as either “online” or “offline” prioritization techniques. Online prioritization techniques trigger the isolation and fragmentation of ions from the low-energy mass spectra in real time, with user-defined parameters. Offline prioritization techniques, in contrast, highlight chemicals of interest after the data has been acquired; detected features can be filtered and ranked based on the relative abundance or the predicted structure, toxicity, and concentration imputed from the tandem mass spectrum (MS^2). Here we provide an overview of these prioritization techniques and how they have been successfully implemented and reported in the literature to find chemicals of elevated risk to human and ecological environments. A complete list of software and tools is available from <https://nontargetedanalysis.org/>.

HRMS Prioritization



INTRODUCTION

The scale and impact of novel entities, such as persistent organic pollutants (POPs), toward the sustainability of the planet's ecosystem are poorly understood.¹ Domains contributing to the impact on this planetary boundary include industrial manufacturing,² the supply and quality of foods,³ and finished pharmaceutical products.⁴ The Chemical Abstracts Service (CAS, American Chemical Society) currently reports >200 million organic substances, alloys, coordination compounds, minerals, mixtures, polymers, and salts.⁵ Global production of many chemicals is increasing each year, exceeding our ability to assess the risk posed by these chemicals, which may lead to adverse impacts on the environment.^{2,6} Methods utilizing analytical chemistry tools, such as high-resolution mass spectrometry (HRMS), are constantly being improved upon to detect and measure the mass of a wide range of chemicals with extreme precision. However, the amenability of a chemical to a specific analytical method depends on many factors, including the extraction

methodology, chromatographic conditions, instrument response factor, and the concentration of the chemicals present in the sample (*vide infra*).⁷

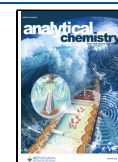
A variety of HRMS instruments are used to detect chemicals in a sample where the mass-to-charge ratio (m/z) of each chemical is measured with a mass resolution >10,000.⁸ The instruments most often used are time-of-flight (TOF) and orbital ion trap mass analyzers coupled with quadrupole devices to enable ion isolation (m/z window widths between 0.4 to 2 Da) and/or dissociation for multidimensional data acquisition.⁹ With most data-dependent acquisition (DDA)

Received: December 14, 2023

Revised: January 30, 2024

Accepted: February 7, 2024

Published: February 21, 2024



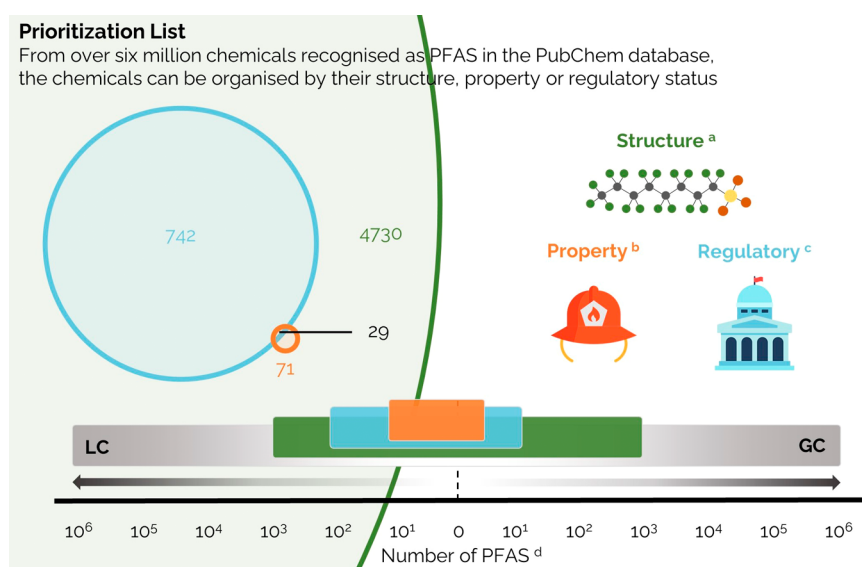


Figure 1. An example of the structural, property, and regulation-based characterization for PFAS in the LC- and GC-amenable chemical spaces. From over 6 million chemicals defined as PFAS,³³ 4730 were identified by the OECD,³⁴ at least 71 chemicals have been identified in aqueous-film-forming-foam (AFFF),³⁵ and 742 are listed as REACH chemicals (EU-1272/2008). Source: NORMAN Suspect List Exchange.³⁶ ^dChromatography predictions were calculated by Alygizakis et al.³⁷ and extrapolated to the entire PubChem list.

methods, precursor ions are isolated and fragmented using molecular dissociation, where the resulting fragments are analyzed with a high-resolution mass analyzer. Unfortunately, the number of spectra per unit time is limited such that DDA does not enable MS² analysis of all precursor ions; therefore, the MS² spectra must be triggered based on certain criteria. Alternatively, data-independent acquisition (DIA) methods set the quadrupole to pass through all ions or ions in a wide m/z range (typically a few tens of Da), to the collision cell for dissociation. In DIA, the resulting MS² spectrum contains a mixture of fragment ions from all precursors in the MS¹ spectrum. There are two strategies used to deconvolute the precursor and fragment mass spectrum: post hoc analysis of the chromatographic peak profile¹⁰ and multiway curve resolution. The former has been thoroughly tested and used, but it can produce lower-quality MS² spectra when presented with numerous coeluting precursors.¹¹ The latter is an emerging technique that has been demonstrated to improve deconvolution by adding the analysis of the relative peak intensity to the data cube.¹² Ion mobility separation is a complementary technique that can be coupled to HRMS instruments to separate coeluting chemicals and deconvolute MS² spectra based on their collisional cross sections (CCSs).¹³

The two main approaches for high-throughput detection and identification of unknown chemicals using HRMS are known as suspect screening and non-targeted screening¹⁴ (also non-targeted analysis). Generally, suspect screening matches the acquisition data with a list of chemicals of interest in the sample, thereby ignoring any other chemicals which may be in the sample. On the other hand, non-targeted screening does not require *a priori* assumptions for the presence or absence of chemicals in a sample, which often results in a wealth of information-rich data that can be more challenging to process efficiently. Careful design of suspect and non-targeted screening workflows is crucial for the successful detection and identification of chemicals present in samples. The large number of features that can be detected, even in seemingly “clean” matrices, illustrates the complexity of non-targeted

data. For example, in previous studies of both tap water and tertiary-treated wastewater, tens of thousands of features were detected across multiple samples.^{15,16} Here, “feature” refers to a data tuple consisting of the retention time of the chromatographic peak and m/z of the precursor adducts and isotopes.¹⁷ The confident structural annotation of unknown features has been a topic of great interest, with molecular networking,^{18,19} forward prediction,^{20,21} and inverse prediction^{22,23} technologies recently becoming available. The confidence of annotation performance can be quantified for communication of results.^{24,25} Regardless of the selected workflow, a considerable amount of computational and research time is currently required to annotate each feature confidently. Therefore, there is a need to develop strategies that will reduce analysis time without sacrificing confidence.

Prioritization strategies address this need by highlighting the most relevant features to the study goal, which can significantly reduce the vast number of features detected by HRMS that are selected for further investigation. The field is working to standardize the evaluation and performance of suspect and non-targeted screening workflows,^{26–28} and prioritization strategies are playing an important role in these workflows. Here, we provide a framework for the categorization of online and offline prioritization techniques and methodologies for the harmonization of prioritization strategies by demonstrating concepts and tools available to researchers that can facilitate both high-quality HRMS analysis and communication of the results, including confidence in those results.

■ PRIORITIZATION LISTS

Lists of various chemicals can be employed in both online and offline prioritization strategies to narrow down potential suspects. Both chemical databases and mass spectral libraries can be used to curate a “prioritization list”. While the former contains useful information for each chemical, such as its structure, names, physicochemical properties, functional uses, and toxicity,²⁹ mass spectral libraries contain empirical MS¹ and/or MSⁿ spectra for each chemical that can be used as a

direct comparison with experimental data for identification³⁰ (see **Suspect Screening**). Studies have used hundreds,³¹ thousands,¹⁴ or tens of thousands of chemicals³² relevant to the scope of the analysis; however, in most cases, it would be antithetical to the aim of prioritization to report hundreds or thousands of chemicals based on matching exact masses with a list or database without further context. Therefore, building a list of suspect chemicals for online or offline prioritization is of utmost importance and should be clearly described and reported. Depending on the scope of a particular study, chemical lists based on their structural, property, and regulatory information may be used to more effectively prioritize chemicals of interest for further investigation and provide clear context in the reporting.

Structure-based lists incorporate sets of chemicals of similar functional groups, moieties, or the presence of specific elements. For example, the detection of organic pollutants in the environment can include common structure-based lists including polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (PBDE), and per- and polyfluoroalkyl substances (PFAS).^{38,39} Incidentally, each of these three groups of chemicals may be assembled into a larger organohalogen compounds (OHCs) list. While PCBs and PBDEs are commonly separated by gas chromatography (GC), PFAS from the Organisation for Economic Co-operation and Development (OECD) priority list are more often LC-amenable (Figure 1). Due to the influence of physicochemical properties and amenability to specific techniques, the selection of chemicals for specific structure-based lists will depend on instrument selection and sample preparation.⁷ Additionally, the toxicity or hazard of chemicals with similar structural and molecular descriptors can be predetermined and used to rank chemicals based on their predicted risk to human and ecological health, as demonstrated in a proof-of-concept software application developed by the United States Environmental Protection Agency (US-EPA).⁴⁰

Property-based lists may incorporate chemicals that share similar physicochemical parameters or are used to elicit the same desired outcome. For example, herbicides, fungicides, insecticides, or rodenticides, collectively known as pesticides, are property-based lists because they all have the same aim: to cause adverse effects in their target organism.⁴¹ Similarly, pharmaceuticals and personal care products (PPCPs) are another property-based group of chemicals that are all designed to improve the quality of life. Property-based lists may also include chemicals with similar physicochemical properties or environmental fate, such as transformation products from pesticides or pharmaceuticals, which can be included to monitor drinking water quality before and after treatment.⁴² Such lists are readily available from resources such as the NORMAN Suspect List Exchange³⁶ and the US-EPA's CompTox Chemicals Dashboard.⁴³

Regulatory-based lists of chemicals are prepared by local, state, or federal governments to limit the exposure of chemicals with known adverse impacts to human, societal, or ecological health. For example, the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH; EU/1907/2006) and the US-EPA's Toxic Substance Control Act (TSCA; 15 U.S.C. §2601 1976) each contain a list of substances that have been used to screen for restricted chemicals in textiles and surface waters.^{44,45} Similarly, lists of controlled or illicit substances can be used to extend the capability of forensic laboratories,⁴⁶ for example, the Scientific Working Group for

the Analysis of Seized Drugs (SWGDRUG), which publishes structure-based lists and mass spectral libraries.⁴⁷ Typically, it takes many years of scientific reporting for a chemical to appear in regulatory-based lists and is subject to political processes.^{48,49} Many of these chemicals are still manufactured and used within regulatory frameworks, often to balance the adverse effects on humans and the environment with the economic value of the chemical. For these reasons, regulatory-based lists are a popular starting point for many studies, as the detection and quantification of these chemicals provide key monitoring information to enforcement agencies.⁵⁰ It should be noted that transformation products of regulated chemicals, resulting from biological metabolism and environmental degradation, should also be taken into account (**Structural and Molecular Analysis**).

ONLINE PRIORITIZATION

Online prioritization includes the variety of parameters that are used to perform DDA, predefined by the user for automatic selection by an instrument in real-time during each duty cycle. The five online prioritization techniques that are used to overcome the duty cycle limitations in DDA are (1) Top *N*, (2) Inclusion/Exclusion List, (3) Isotopic Abundance, (4) Adduct Formation, and (5) Low Res MS² Trigger. The simplest and most widely available strategy is the "Top *N*" selection (Figure 2a), where *N* most abundant MS¹ peaks from a single scan, above a set intensity threshold, are selected for MS² analysis.^{42,51–53} Most HRMS instruments also allow triggering MS² experiments via matching with an inclusion list of predefined *m/z* values (Figure 2b),^{44,54,55} which may be built from a prioritization list as described in **Prioritization Lists**.

A useful approach for the curation of the inclusion list is a post hoc analysis of MS¹ data previously acquired. Here, precursor *m/z* values can be matched with a suspect screening prioritization list to identify potential candidates. Due to the lack of relative confidence in matching only MS¹ data and limitations of duty cycle, these candidates can be fed back into an inclusion list, where each of the candidates can be isolated for fragmentation and analyzed to obtain a higher degree of annotation confidence and to target more ions for MS/MS. Conversely, features previously triggered for MS² acquisition can be added to an exclusion list, allowing other peaks to be interrogated in subsequent runs.⁵⁶ These approaches require multiple injections and data analysis that will increase the total user time for acquisition. However, pooled quality control samples can be used to generate an inclusion list without requiring multiple injections of every sample.⁵⁷

Triggering strategies can also involve the real-time determination of the isotopic abundance pattern; for example, chemicals containing Cl and Br can be triggered for MS² based on the increased intensity of the [A + 2]⁺ isotope caused by their naturally occurring stable isotopes (Figure 2c). Similarly, the presence of adduct formations ([M + H]⁺, [M + Na]⁺, [M + K]⁺, etc.) can be determined and selected for fragmentation accordingly (Figure 2d).⁵⁸ Finally, in instruments with two or more mass analyzers, the *m/z* or $\Delta m/z$ from a rapid (40 Hz) and low-resolution (>0.1 Da) MS² spectrum can be monitored for peaks of interest.⁵⁹ This can, in turn, trigger a high-resolution MS² experiment in the next cycle (Figure 2e). Triggering MS² spectra based on the measured isotopic ratio, or the presence of adducts in the MS¹ scan is, to the authors' knowledge, available only for the instruments from one vendor

Data-dependent acquisition (DDA) triggers

Five currently available strategies for selecting and isolating precursor m/z for high-resolution MS² analysis

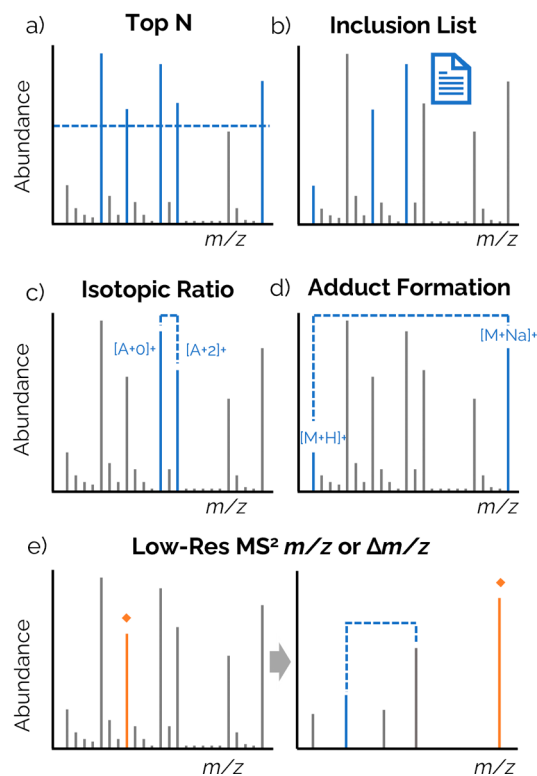


Figure 2. Examples of current online prioritization strategies for the selection and isolation of precursor m/z for fragmentation: (a) Top N, (b) inclusion list. Then, (c) isotopic ratio, (d) adduct formation, and (e) low-resolution MS² signal processing are currently limited to a few instruments.

(Thermo Fisher Scientific) and has not been widely employed for the analysis of environmental contaminants. However, these strategies were recently used to prioritize the presence of potentially harmful chemicals in water by triggering MS² acquisition based on the presence of halogen isotopic ratios in the MS¹ and by neutral losses associated with structural alerts.⁵⁸ These approaches appear to be suitable for the detection of small molecules though they have not yet been extensively tested, optimized, and validated for applications to suspect and non-targeted screening approaches.

Each of these approaches aims to select chemicals of interest for rapid and efficient fragmentation but have certain limitations. First, chemicals of equal concentrations can have differences in response factors by several orders of magnitude.⁶⁰ Particularly in complex matrices like foods and biosolids, chemicals of interest may yield low signal intensity due to low ionization efficiency or matrix effects; therefore, the MS² spectra may not be acquired for these ions with “Top N”. Finally, as mentioned earlier, only a limited number of m/z values matched to an inclusion list may be isolated depending on the size of the inclusion list and the extent of coelution of these compounds. Certain strategies can be used to address some of these limitations, that help optimize the selection criteria of MS² acquisitions.

The chromatographic method also plays an important role in determining the extent of coelution and can limit the

number of precursor ions that can be triggered for MS² acquisition.⁶¹ Shorter gradients typically yield sharper, more intense peaks, while longer gradients often increase peak widths and have reduced intensities. Longer gradient elutions also enable the detection and fragmentation of more precursor ions in highly complex samples by improving the overall chromatographic peak resolution.⁶² These effects can contribute to the efficiency of the instrument to select and isolate peaks within the duty cycle time of the instrument for reliable analysis. As a result, 20–30 min gradient lengths have been found to provide a good balance between peak resolution and separation, cycle time, and total run time.^{63–65} Additionally, “active” or “dynamic” exclusion windows can avoid over-triggering multiple MS² spectra for the same precursor so those scans can be used to generate MS² spectra for more unique precursors. The duration of these exclusion windows should be set based on general peak widths of the chromatographic method to ensure that MS² spectra are collected at or near the apex of a peak (i.e., for a 10 s peak, the exclusion window should be <5 s).

In DDA which is applied by using an inclusion list, instruments from many vendors also allow for the selection of chemicals within a specified retention time window to increase the number of chemicals monitored in a given duty cycle. If the retention time is unknown or not measured with analytical standards, the retention time for any chemical may be predicted using advanced machine learning models.^{66–68} Here, the calculated or predicted retention time index (RTI) for each chemical in the inclusion list can be used to assign an appropriate retention time window. Additionally, employing machine learning algorithms to predict the polarity and LC amenability can be used to filter inclusion lists.^{37,69} To the authors’ knowledge, this process has yet to be applied to explicitly filter chemicals in inclusion lists for DDA but is expected to improve prioritization given sufficiently accurate models for predicting retention time.⁷⁰

Improvements in hardware continually improve signal quality at lower cycle times, and developments in real-time signal processing will allow more useful and dynamic background exclusion and inclusion. Given the current limitations of DDA on chemical selection for fragmentation, DIA provides an alternative instrumental technique to improve MS² coverage,⁷¹ although because it lacks the ability to isolate precursors, it is not amenable to online prioritization techniques. Deconvolution of DIA data requires complex post hoc signal processing to link MS² spectra with corresponding MS¹ peaks. Retention time correlation models can provide adequate deconvolution results,⁷² and developing multivariate curve resolution analysis can refine the signal processing even further.⁷³

■ OFFLINE PRIORITIZATION

Suspect Screening. Offline prioritization strategies employ post hoc analysis of MS¹ and MS² (both DIA and DDA) data. Regardless of the acquisition methodology, offline prioritization ranges in complexity from suspect screening to non-targeted screening, each of which are performed independently of each other. Data processing in suspect screening begins by mining the data directly for acquired MS¹ features with the exact mass and isotopic distribution in a prioritization list (see [Prioritization Lists](#)). Measured and/or predicted retention times help to filter the candidates if multiple matches are observed.⁶⁶ Substances with unique mass

defects or repeating moieties, such as PFAS, can be extracted and prioritized in suspect screening workflows.^{74,75} Finally, matching the MS² spectra with a publicly available spectral library or an in-house library increases the confidence of a positive match.²⁵ Inventories that detail the volume or concentration of the discharge of chemicals to the environment have been used to prioritize substances emitted from a wastewater treatment plant.⁷⁶ The advantage of this strategy is the ability to assess the risk of exposure to downstream environments, based on the relative discharge rates declared by licensed companies.

Prioritization lists containing thousands of chemicals may be used in offline prioritization; however, the propensity for false positives increases with the length of the prioritization list. Hence, it is beneficial to reduce these lists where possible depending on study questions and goals. Specifically, caution is urged when using complete PubChem and ChemSpider databases containing over 100 million live chemicals from 799 and 277 sources, respectively.^{77,78} The probability of most of these chemicals being detected in their respective matrix is extremely small, so an effort has been made to limit the chemical list used to support mass spectrometry analysis to a much smaller fraction of the database: the so-called PubChemLite for Exposomics list that is updated regularly.⁷⁹

Intensity and Concentration Exclusion. A strategy common for the non-targeted screening of impurities in food, pharmaceuticals, and medical devices involves filtering features based on the instrument response relative to an internal standard of known concentration (Figure 3a).^{80,81} Threshold of toxicological concern (TTC), also known as safety concern threshold (SCT), analytical evaluation threshold (AET), or dose-based threshold (DBT), are offline prioritization strategies developed to highlight chemicals by excluding those expected to have low concentrations or risk of adverse effects.⁸² This approach has been used to prioritize the presence of chemical migrants in food that originated from food contact material^{83,84} and potentially harmful chemicals detected in natural and drinking water.⁸⁵ This strategy was also demonstrated for the analysis of extractables and leachables (E&L) from a coronary implant, where polymer-based chemicals and transformation products were positively identified.⁸⁶

There are notable exceptions to the TTC approach, where chemicals containing known structural alerts should be evaluated regardless of their intensity or concentration, commonly known as the “cohort of concern”.^{82,87} However, in an NTS approach, the absence of *a priori* structural information can obfuscate confident annotation of chemicals with these structural alerts and may be omitted from the analysis. Furthermore, poor quantification accuracy often results when using surrogate internal standards⁸⁸ due to variability in the ionization efficiency between chemicals.⁶⁰ For example, bifenthrin (CASRN: 82657-04-3), a potent insecticide and aquatic toxicant, has a relatively low ionization efficiency ($\log_{10}IE: -1.19$ to 0.79),⁸⁹ where its concentration could be seriously underestimated using standard practices.⁸²

Differential Analysis. Differential analysis relates to a number of methodologies that aim to help isolate features of interest, without a predefined list of chemicals, based on the analysis of multidimensional data within a sample, and the statistical differences between and within samples (Figure 3b). These types of analysis determine differences in feature presence and intensity between any two (or more) groups

Offline Prioritization

Examples of four strategies for feature analysis by abundance and structural evaluation

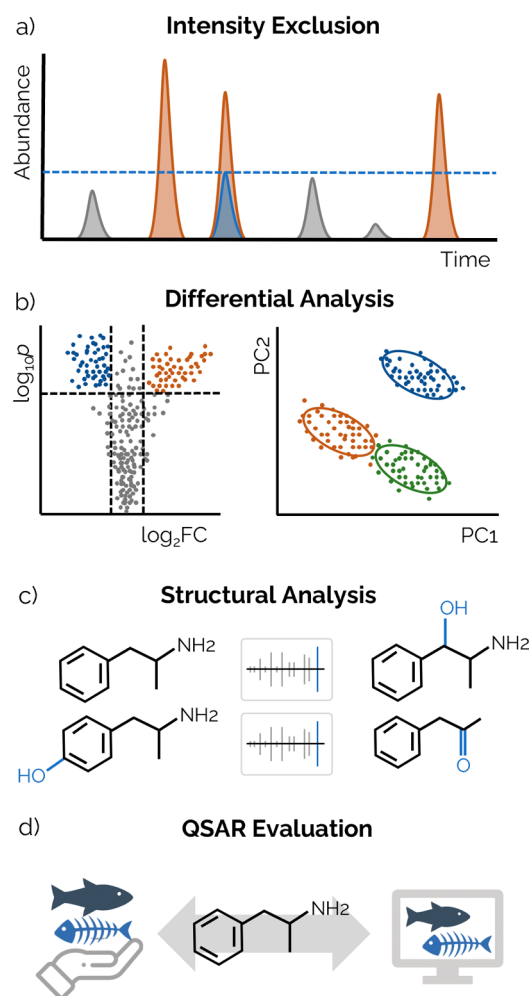


Figure 3. Examples of offline prioritization techniques used to highlight or rank features detected by HRMS. (a) Peaks with lower response than a spiked chemical are excluded from analysis. (b) Fold-change analysis and principal component analysis reveal grouped features. (c) Metabolic logic and analysis of MS² spectra group features of similar structure. (d) Empirical or predicted QSAR rank features based on biological endpoints.

(temporal, spatial, biological, etc.) and are a valuable approach to identify chemicals of interest in a particular sample group. For example, the chemical diversity from different brands and collection dates of oats were compared using principal component analysis (PCA), where even this relatively simple food-type resulted in >3000 unique features. This work concluded that differential analysis was an efficient prioritization approach for highlighting QCs in a spiked vs unspiked sample, but that the best approach depends on the goal and may require a combination of multiple approaches.⁹⁰ This type of trend analysis has also been useful in the analysis of human blood,⁵⁵ drinking water,⁹¹ food,⁹² biological matrices,⁹³ contaminated beverages,⁹⁴ and human biomonitoring.⁹⁵

As data acquired with HRMS is highly dimensional and can contain many more variables and observations, univariate or multivariate statistical tests can be scaled up by applying supervised and unsupervised machine learning algorithms to

build predictive models that resolve complex relationships in the data. For example, advanced trend analysis has been applied to characterize and prioritize features that were detected in wastewater-impacted surface waters, where 25 contaminants were successfully identified.⁹⁶ These features could also be categorized into periodic, spill, increasing, or decreasing trends that can enable more efficient filtering and chemical risk assessment. The accuracy and precision of these predictive models continue to be developed and evaluated toward the high-throughput and multiresidue analysis of chemicals from samples.⁹⁷ In any case, the selection and purpose of the statistical methodologies should be reported. Furthermore, since the number of features generated by HRMS is often much greater than the number of samples or observations, the application and interpretation of statistical analysis should be verified and carefully considered to avoid overfitting.⁷⁵

Structural and Molecular Analysis. Prioritization of features found in samples can be conducted by grouping potential transformation products (metabolic logic).⁹⁸ This analysis is performed first by identifying features with a mass difference equal to known biotransformation reactions, including basic modifications, conjugation, and deconjugation (Figure 3c).⁹⁹ As many basic modifications are more polar than the parent compound, a simple filter can be applied to features with longer retention times in reverse-phase LC. Then, the similarity of MS² of the parent and potential transformation products can be evaluated, where the fragments of transformation products should generally align with the parent chemical. This strategy was partially successful in the prioritization of features detected in treated wastewater¹⁰⁰ and later more successful in prioritizing pharmaceutical transformation products in sludge.¹⁰¹ Tools commonly used for the prioritization and identification of transformation products include BioTransformer,¹⁰² enviPath,¹⁰³ and the US-EPA's Chemical Transformation Simulator (CTS).¹⁰⁴

Molecular networking can be used to prioritize features based on expected structural similarities determined by comparing high-quality MS² data between features.¹⁰⁵ Similar to metabolic logic, this analysis can reveal chemicals of analogous structure or chemicals with similar bioactive chemicals and metabolites. Molecular networking is a well-suited application for the dereplication of novel pharmaceuticals, especially from existing natural products.¹⁰⁶ Frameworks are available to freely communicate these types of analysis and often include powerful visualization tools.^{18,107} Molecular networking was used to help identify toxic alkaloids from a case of poisoning from an unidentified plant root, where chemical profiling helped to identify the genus and species of the plant that produced the toxins.¹⁰⁸

The negative mass defect of fluorine and identification of homologous series can be leveraged to prioritize legacy and emerging PFAS. Specific tools have been developed to detect differences in mass defect,¹⁰⁹ with the ability to discern PFAS by evaluating the relationship between mass defect (MD) and mass (*m*) relative to the number of carbons (*C*) (MD/*C* – *m*/*C*).¹¹⁰ These analyses has been applied to the discovery of novel PFAS in historical pine-needle samples by the combination of the negative mass defect and characteristically low CCS values.¹¹¹

Quantitative Structure–Activity Relationship Evaluation. Hazard profile listings for chemicals contain various types of empirical human and ecological toxicity endpoints, including

acute and chronic, reproductive, and behavioral, from oral, inhalation, and dermal exposure pathways.¹¹² Chemicals from these listings can be matched with tentative structural annotation from suspect and non-targeted screening workflows to evaluate the potential chemical risk. When these hazard profiles are paired with cheminformatic data, quantitative structure–activity relationship (QSAR) models can be developed to provide hazard estimates from chemicals with no empirical evidence (Figure 3d). QSAR-based strategies for offline prioritization aim to rank a list of putative identifications according to their associated risk to receiving organisms and environments.¹¹³ QSAR models are typically determined by the relationship between a chemical's structural and/or molecular descriptors and its human or ecological-based toxicity endpoint. These models can range in complexity; for example, bioaccumulation in aquatic organisms can simply be related to a chemical's octanol–water coefficient ($\log K_{ow}$),¹¹⁴ and toxicity can be predicted by analyzing key substructures that are known to negatively impact survivability and reproduction.¹¹⁵

Human and ecological toxicity predictions can also be used to prioritize hazardous substances from a list of features. Analysis of the relative abundance of chemicals detected in household dust (via ToxPi¹¹⁶) revealed 15 compounds that were reported in dust samples for the first time.¹¹⁷ However, toxicity predictions that are derived from structural information¹¹⁸ or the presence of structural alerts⁵⁸ contain two sources of uncertainty that can compound the error of the estimate: the putative structural annotation and the toxicity prediction model.²⁵ To address this issue, Peets et al.¹¹⁹ modeled acute aquatic toxicity in fish-derived only from the MS² spectrum (MS2Tox), precluding confident annotation. Additionally, there have been recent advances to quantify chemicals identified by non-targeted screening through machine learning algorithms, enabling more accurate evaluation of the risk posed to organisms exposed to contaminated environments.^{88,113,120}

CONCLUSION

For suspect or non-targeted screening methodologies to be most useful for the evaluation of chemical risk to human or ecological health, the chemical space and prioritization strategies must be clearly defined. Careful consideration of the data acquisition and data analysis parameters presented in this study can vastly improve the confidence and efficiency for the prioritization of chemicals of interest in environmental, biological, food, medical device, and other matrices. We encourage HRMS users employing suspect and non-targeted screening workflows to conceptualise prioritization strategies in the framework presented here. Online prioritization techniques require minimal input from the user and can enable the generation of high-quality MS² spectra for a greater percentage of relevant features depending on the study goals. Through the use of one or more offline prioritization techniques, both DIA and DDA information can be used to highlight chemicals based on their abundance, chemometrics, or predicted impacts to a selection of toxicological endpoints. Notable implementations of each of these strategies have been presented in this study, and we support the continued development of each of these strategies as the field matures.

AUTHOR INFORMATION

Corresponding Author

Anneli Kruve – Department of Materials and Environmental Chemistry, Stockholm University, Stockholm 106 91, Sweden; Department of Environmental Science, Stockholm University, Stockholm 106 91, Sweden; orcid.org/0000-0001-9725-3351; Email: anneli.kruve@su.se

Authors

Drew Szabo – Department of Materials and Environmental Chemistry, Stockholm University, Stockholm 106 91, Sweden; orcid.org/0000-0002-0089-9218

Travis M. Falconer – Forensic Chemistry Center, Office of Regulatory Science, Office of Regulatory Affairs, US Food and Drug Administration, Cincinnati, Ohio 45237, United States

Christine M. Fisher – Center for Food Safety and Applied Nutrition, US Food and Drug Administration, College Park, Maryland 20740, United States

Ted Heise – MED Institute Inc, West Lafayette, Indiana 47906, United States

Allison L. Phillips – Center for Public Health and Environmental Assessment, US Environmental Protection Agency, Corvallis, Oregon 97333, United States

Gyorgy Vas – VasAnalytical, Flemington, New Jersey 08822, United States; Intertek Pharmaceutical Services, Whitehouse, New Jersey 08888, United States

Antony J. Williams – Center for Computational Toxicology and Exposure, Office of Research and Development, US Environmental Protection Agency, Durham, North Carolina 27711, United States; orcid.org/0000-0002-2668-4821

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.analchem.3c05705>

Author Contributions

Drew Szabo: Writing - Original Draft, Conceptualization, Visualization. **Travis M. Falconer**: Writing - Review & Editing. **Christine M. Fisher**: Writing - Review & Editing. **Ted Heise**: Writing - Review & Editing. **Allison L. Phillips**: Writing - Review & Editing. **Gyorgy Vas**: Writing - Review & Editing. **Antony J. Williams**: Writing - Review & Editing. **Anneli Kruve**: Writing - Review & Editing, Supervision, Funding acquisition.

Funding

Open-access funding is provided by Stockholm University. This project is funded by Formas (Grant No. 2022-00440).

Notes

The authors whose names are listed certify that they have no affiliations with or involvement in any organization or entity with any financial interest, or nonfinancial interest in the subject matter or materials discussed in this manuscript. The opinions expressed in this manuscript are those of the authors and do not reflect current or future policy of the FDA, the EPA, or any other US government agency. The mention of manufacturers or specific products is for clarity and does not constitute endorsement. The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the leadership and members of the Best Practices for Non-Targeted Analysis (BP4NTA, <https://nontargetedanalysis.org/>) expert working group for their support in this project. Particularly, we thank Keaton Nahan

for his contributions to the manuscript. We thank the members of the analytical chemistry unit at Stockholm University for their support, particularly the students and staff in the KruveLab (<https://kruvelab.com/>). The artificial intelligence model (<https://scite.ai>) was used to help evaluate references and provided the authors with reading suggestions.

REFERENCES

- (1) Steffen, W.; Richardson, K.; Rockström, J.; Cornell, E.; Fetzer, I.; Bennett, E. M.; Biggs, R.; Carpenter, S. R.; de Vries, W.; de Wit, C. A.; Folke, C.; Gerten, D.; Heinke, J.; Mace, G. M.; Persson, L. M.; Ramanathan, V.; Reyers, B.; Sörlin, S. *Science* **2015**, *347* (6223), 1259855.
- (2) Persson, L.; Carney Almroth, B. M.; Collins, C. D.; Cornell, S.; de Wit, C. A.; Diamond, M. L.; Fantke, P.; Hassellöv, M.; MacLeod, M.; Ryberg, M. W.; Sogaard Jørgensen, P.; Villarrubia-Gómez, P.; Wang, Z.; Hauschild, M. Z. *Environ. Sci. Technol.* **2022**, *56* (3), 1510–1521.
- (3) Campbell, B. M.; Beare, D. J.; Bennett, E. M.; Hall-Spencer, J. M.; Ingram, J. S. I.; Jaramillo, F.; Ortiz, R.; Ramankutty, N.; Sayer, J. A.; Shindell, D. *Ecol. Soc.* **2017**, *22* (4), 8.
- (4) Thornber, K.; Adshead, F.; Balayannis, A.; Brazier, R.; Brown, R.; Comber, S.; Court, C.; Davidson, I.; Depledge, M.; Farmer, C.; Gibb, S.; Hixson, R.; Kirchhelle, C.; Moore, K.; Motta, M.; Niemi, L.; Owen, S.; Pencheon, D.; Pflieger, S.; Pitchforth, E.; Powell, N.; Schmidt, W.; Smith, R.; Sowman, G.; Tyler-Batt, W.; Wilkinson, H.; Wilson, E. C. F.; Fleming, L.; Gaze, W.; Tyler, C. *Lancet Planet. Health* **2022**, *6* (12), e935–e937.
- (5) Chemical Abstract Service (CAS) CAS Registry. <https://www.cas.org/cas-data/cas-registry> [Date Accessed: September 2023].
- (6) Cousins, I. T.; Johansson, J. H.; Salter, M. E.; Sha, B.; Scheringer, M. *Environ. Sci. Technol.* **2022**, *56*, 11172.
- (7) Black, G.; Lowe, C.; Anumol, T.; Bade, J.; Favela, K.; Feng, Y.-L.; Knolhoff, A.; McEachran, A.; Nuñez, J.; Fisher, C.; Peter, K.; Quinete, N. S.; Sobus, J.; Sussman, E.; Watson, W.; Wickramasekara, S.; Williams, A.; Young, T. *Anal. Bioanal. Chem.* **2023**, *415*, 35.
- (8) 2002/657/EC: Commission Decision of 12 August 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (Text with EEA relevance) (notified under document number C(2002) 3044). **2002**; pp 8–36.
- (9) Kaufmann, A.; Bromirski, M., Selecting the best Q Exactive Orbitrap mass spectrometer scan mode for your application. **2018**, *White Paper* 65147.
- (10) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* **2015**, *12* (6), 523–526.
- (11) Tokiyoshi, K.; Matsuzawa, Y.; Takahashi, M.; Takeda, H.; Hasegawa, M.; Miyamoto, J.; Tsugawa, H. *Anal. Chem.* **2024**, *96* (3), 991–996.
- (12) Kronik, O. M.; Liang, X.; Nielsen, N. J.; Christensen, J. H.; Tomasi, G. *J. Chromatogr. A* **2022**, *1682*, 463501.
- (13) Marsden-Edwards, E.; Tomczyk, N.; Wildgoose, J. Making Ion Mobility Mass Spectrometry Routine. *Chromatography Online*. https://www.waters.com/webassets/cms/library/docs/lcgc1215_waters_11-18_wp_pr4f.pdf [Date Accessed: May 2023].
- (14) Schymanski, E. L.; Singer, H. P.; Slobodnik, J.; Ipolyi, I. M.; Oswald, P.; Krauss, M.; Schulze, T.; Haglund, P.; Letzel, T.; Grosse, S.; Thomaidis, N. S.; Bletsou, A.; Zwiener, C.; Ibáñez, M.; Portolés, T.; de Boer, R.; Reid, M. J.; Onghena, M.; Kunkel, U.; Schulz, W.; Guillon, A.; Noyon, N.; Leroy, G.; Bados, P.; Bogianni, S.; Stipanicev, D.; Rostkowski, P.; Hollender, J. *Anal. Bioanal. Chem.* **2015**, *407* (21), 6237–6255.
- (15) Newton, S. R.; McMahan, R. L.; Sobus, J. R.; Mansouri, K.; Williams, A. J.; McEachran, A. D.; Strynar, M. *J. Environ. Pollut.* **2018**, *234*, 297–306.
- (16) Schollée, J. E.; Hollender, J.; McArdeall, C. S. *Water Res.* **2021**, *200*, 117209.

- (17) Samanipour, S.; Martin, J. W.; Lamoree, M. H.; Reid, M. J.; Thomas, K. V. *Environ. Sci. Technol.* **2019**, *53* (10), 5529–5530.
- (18) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsubawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kamenik, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin, H. C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. *Nat. Methods* **2020**, *17* (9), 905–908.
- (19) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; van der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Bandeira, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109* (26), No. E1743-E1752.
- (20) Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S. *J. Cheminformatics* **2016**, *8* (1), 3.
- (21) Wang, F.; Lüigand, J.; Tian, S.; Arndt, D.; Greiner, R.; Wishart, D. S. *Anal. Chem.* **2021**, *93* (34), 11692–11700.
- (22) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. *Nat. Methods* **2019**, *16* (4), 299–302.
- (23) Goldman, S.; Wohlwend, J.; Stražar, M.; Haroush, G.; Xavier, R. J.; Coley, C. W. *Nat. Mach. Intell.* **2023**, *5*, 965.
- (24) Alygizakis, N.; Lestremou, F.; Gago-Ferrero, P.; Gil-Solsona, R.; Arturi, K.; Hollender, J.; Schymanski, E. L.; Dulio, V.; Slobodnik, J.; Thomaidis, N. S. *Trends Anal. Chem.* **2023**, *159*, 116944.
- (25) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. *Environ. Sci. Technol.* **2014**, *48* (4), 2097–2098.
- (26) Fisher, C. M.; Peter, K. T.; Newton, S. R.; Schaub, A. J.; Sobus, J. R. *Anal. Bioanal. Chem.* **2022**, *414* (22), 6455–6471.
- (27) Peter, K. T.; Phillips, A. L.; Knolhoff, A. M.; Gardinali, P. R.; Manzano, C. A.; Miller, K. E.; Pristner, M.; Sabourin, L.; Sumarah, M. W.; Warth, B.; Sobus, J. R. *Anal. Chem.* **2021**, *93* (41), 13870–13879.
- (28) Knolhoff, A. M.; Premo, J. H.; Fisher, C. M. *Anal. Chem.* **2021**, *93* (3), 1596–1603.
- (29) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. *Nucleic Acids Res.* **2016**, *44* (D1), D1202–D1213.
- (30) Dettmer, K.; Aronov, P. A.; Hammock, B. D. *Mass Spectrom. Rev.* **2007**, *26* (1), 51–78.
- (31) Moschet, C.; Piazzoli, A.; Singer, H.; Hollender, J. *Anal. Chem.* **2013**, *85* (21), 10312–10320.
- (32) Dürig, W.; Tröger, R.; Andersson, P. L.; Rybacka, A.; Fischer, S.; Wiberg, K.; Ahrens, L. *Chemosphere* **2019**, *222*, 904–912.
- (33) Schymanski, E. L.; Chirsir, P.; Kondic, T.; Thiessen, P. A.; Zhang, J.; Bolton, E. E. PFAS and Fluorinated Compounds in PubChem Tree. <https://pubchem.ncbi.nlm.nih.gov/classification/#hid=120> [Date Accessed: July 2022].
- (34) Organisation for Economic Co-operation and Development (OECD). *Toward a new comprehensive global database of per- and polyfluoroalkyl substances (PFAS)*; Paris, May, 2018.
- (35) Glüge, J.; Scheringer, M.; Cousins, I. T.; DeWitt, J. C.; Goldenman, G.; Herzke, D.; Lohmann, R.; Ng, C. A.; Trier, X.; Wang, Z. *Environ. Sci. Process. Impacts* **2020**, *22*, 2345.
- (36) Mohammed Taha, H.; Aalizadeh, R.; Alygizakis, N.; Antignac, J.-P.; Arp, H. P. H.; Bade, R.; Baker, N.; Belova, L.; Bijlsma, L.; Bolton, E. E.; Brack, W.; Celma, A.; Chen, W.-L.; Cheng, T.; Chirsir, P.; Ćirka, L.; D'Agostino, L. A.; Djoumbou Feunang, Y.; Dulio, V.; Fischer, S.; Gago-Ferrero, P.; Galani, A.; Geueke, B.; Glowacka, N.; Glüge, J.; Groh, K.; Grosse, S.; Haglund, P.; Hakkinen, P. J.; Hale, S. E.; Hernandez, F.; Janssen, E. M. L.; Jonkers, T.; Kiefer, K.; Kirchner, M.; Koschorreck, J.; Krauss, M.; Krier, J.; Lamoree, M. H.; Letzel, M.; Letzel, T.; Li, Q.; Little, J.; Liu, Y.; Lunderberg, D. M.; Martin, J. W.; McEachran, A. D.; McLean, J. A.; Meier, C.; Meijer, J.; Menger, F.; Merino, C.; Muncke, J.; Muschket, M.; Neumann, M.; Neveu, V.; Ng, K.; Oberacher, H.; O'Brien, J.; Oswald, P.; Oswaldova, M.; Picache, J. A.; Postigo, C.; Ramirez, N.; Reemtsma, T.; Renaud, J.; Rostkowski, P.; Rüdell, H.; Salek, R. M.; Samanipour, S.; Scheringer, M.; Schliebner, I.; Schulz, W.; Schulze, T.; Sengl, M.; Shoemaker, B. A.; Sims, K.; Singer, H.; Singh, R. R.; Sumarah, M.; Thiessen, P. A.; Thomas, K. V.; Torres, S.; Trier, X.; van Wezel, A. P.; Vermeulen, R. C. H.; Vlaanderen, J. J.; von der Ohe, P. C.; Wang, Z.; Williams, A. J.; Willighagen, E. L.; Wishart, D. S.; Zhang, J.; Thomaidis, N. S.; Hollender, J.; Slobodnik, J.; Schymanski, E. L. *Environ. Sci. Eur.* **2022**, *34* (1), 104.
- (37) Alygizakis, N.; Konstantakos, V.; Bouziotopoulos, G.; Kormentzas, E.; Slobodnik, J.; Thomaidis, N. S. *Metabolites* **2022**, *12* (3), 199.
- (38) Batt, A. L.; Wathen, J. B.; Lazorchak, J. M.; Olsen, A. R.; Kincaid, T. M. *Environ. Sci. Technol.* **2017**, *51* (5), 3021–3031.
- (39) Liu, Y.; D'Agostino, L. A.; Qu, G.; Jiang, G.; Martin, J. W. *Trends Anal. Chem.* **2019**, *121*, 115420.
- (40) Black, G. P.; Lowe, C.; Anumol, T.; Bade, J. L.; Favela, K.; Fisher, C.; Feng, Y.-L.; Hood, A.; Knolhoff, A. M.; McEachran, A.; Nuñez, J. R.; Peter, K.; Quinete, N.; Sobus, J.; Sussman, E. M.; Watson, W.; Williams, A.; Wickramesekara, S. In *Mapping Chemical Space Coverage in Non-Targeted Analysis*; SETAC - Non-Targeted Analysis Meeting, Durham, NC, USA, 2022; DOI: 10.23645/epacomptox.20126375.v1.
- (41) Kiefer, K.; Müller, A.; Singer, H.; Hollender, J. *Water Res.* **2019**, *165*, 114972.
- (42) Tröger, R.; Ren, H.; Yin, D.; Postigo, C.; Nguyen, P. D.; Baduel, C.; Golovko, O.; Been, F.; Joers, H.; Boleda, M. R.; Polesello, S.; Roncoroni, M.; Taniyasu, S.; Menger, F.; Ahrens, L.; Yin Lai, F.; Wiberg, K. *Water Res.* **2021**, *198*, 117099.
- (43) 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. *J. Cheminformatics* **2017**, *9* (1), 61.
- (44) Carlsson, J.; Iadaresta, F.; Eklund, J.; Avagyan, R.; Östman, C.; Nilsson, U. *Anal. Bioanal. Chem.* **2022**, *414* (3), 1403–1413.
- (45) Zahn, D.; Mucha, P.; Zilles, V.; Touffet, A.; Gallard, H.; Knepper, T. P.; Frömel, T. *Water Res.* **2019**, *150*, 86–96.
- (46) Colby, J. M.; Thoren, K. L.; Lynch, K. L. *J. Anal. Toxicol.* **2018**, *42* (4), 207–213.
- (47) Feeney, W.; Moorthy, A. S.; Sisco, E. *Forensic Chem.* **2022**, *31*, 100459.
- (48) Kudsk, P.; Mathiassen, S. K. *Weed Sci.* **2020**, *68* (3), 214–222.
- (49) Krinsky, S. *PLOS Biol.* **2017**, *15* (12), No. e2002404.
- (50) Schymanski, E. L.; Kondić, T.; Neumann, S.; Thiessen, P. A.; Zhang, J.; Bolton, E. E. *J. Cheminformatics* **2021**, *13* (1), 19.
- (51) Awad, E.; Zhang, X.; Bhavsar, S. P.; Petro, S.; Crozier, P. W.; Reiner, E. J.; Fletcher, R.; Tittlemier, S. A.; Braekelvel, E. *Environ. Sci. Technol.* **2011**, *45* (19), 8081–8089.
- (52) Blum, K. M.; Andersson, P. L.; Renman, G.; Ahrens, L.; Gros, M.; Wiberg, K.; Haglund, P. *Sci. Total Environ.* **2017**, *575*, 265–275.
- (53) Stincone, P.; Shah, A. K. P.; Schmid, R.; Graves, L.; Lambidis, S. P.; Torres, R.; Xia, S.-N.; Minda, V.; Aron, A.; Wang, M.; Hughes, C. C.; Petras, D. Evaluation of Data Dependent MS/MS Acquisition Parameters for Non-targeted Metabolomics and Molecular Networking of Environmental Samples - Focus on the Q Exactive Platform. *ChemRxiv* **2023**, This content is a preprint and has not been peer-reviewed. 1.
- (54) Menger, F.; Gago-Ferrero, P.; Wiberg, K.; Ahrens, L. *Trends Environ. Anal. Chem.* **2020**, *28*, No. e00102.
- (55) Plassmann, M. M.; Fischer, S.; Benskin, J. P. *Environ. Sci. Technol. Lett.* **2018**, *5* (6), 335–340.
- (56) Broeckling, C. D.; Hoyes, E.; Richardson, K.; Brown, J. M.; Prenni, J. E. *Anal. Chem.* **2018**, *90* (13), 8020–8027.

- (57) Hollender, J.; Schymanski, E. L.; Ahrens, L.; Alygizakis, N.; Béen, F.; Bijlsma, L.; Brunner, A. M.; Celma, A.; Fildier, A.; Fu, Q.; Gago-Ferrero, P.; Gil-Solsona, R.; Haglund, P.; Hansen, M.; Kaserzon, S.; Kruve, A.; Lamoree, M.; Margoum, C.; Meijer, J.; Merel, S.; Rauert, C.; Rostkowski, P.; Samanipour, S.; Schulze, B.; Schulze, T.; Singh, R. R.; Slobodnik, J.; Steininger-Mairinger, T.; Thomaidis, N. S.; Togola, A.; Vorkamp, K.; Vulliet, E.; Zhu, L.; Krauss, M. *Environ. Sci. Eur.* **2023**, *35* (1), 75.
- (58) Meekel, N.; Vughs, D.; Béen, F.; Brunner, A. M. *Anal. Chem.* **2021**, *93* (12), 5071–5080.
- (59) Zabrouskov, V.; Damoc, E.; Arrey, T. N.; Pashkova, A.; Zeller, M.; Christian, Hock; Stewart, H.; Hermanson, D. Rethink what is possible with the Orbitrap Astral mass spectrometer. 2023, *White paper 001800*.
- (60) Kruve, A. *Anal. Chem.* **2020**, *92* (7), 4691–4699.
- (61) Krauss, M.; Singer, H.; Hollender, J. *Anal. Bioanal. Chem.* **2010**, *397* (3), 943–951.
- (62) Anderson, B. G.; Raskind, A.; Habra, H.; Kennedy, R. T.; Evans, C. R. *Anal. Chem.* **2021**, *93* (48), 15840–15849.
- (63) Bade, R.; Causanilles, A.; Emke, E.; Bijlsma, L.; Sancho, J. V.; Hernandez, F.; de Voogt, P. *Sci. Total Environ.* **2016**, *569*–570, 434–441.
- (64) Du, B.; Lofton, J. M.; Peter, K. T.; Gipe, A. D.; James, C. A.; McIntyre, J. K.; Scholz, N. L.; Baker, J. E.; Kolodziej, E. P. *Environ. Sci.: Process. Impacts* **2017**, *19* (9), 1185–1196.
- (65) Knolhoff, A. M.; Kneapler, C. N.; Croley, T. R. *Anal. Chim. Acta* **2019**, *1066*, 93–101.
- (66) Aalizadeh, R.; Alygizakis, N. A.; Schymanski, E. L.; Krauss, M.; Schulze, T.; Ibáñez, M.; McEachran, A. D.; Chao, A.; Williams, A. J.; Gago-Ferrero, P.; Covaci, A.; Moschet, C.; Young, T. M.; Hollender, J.; Slobodnik, J.; Thomaidis, N. S. *Anal. Chem.* **2021**, *93* (33), 11601–11611.
- (67) Celma, A.; Bade, R.; Sancho, J. V.; Hernandez, F.; Humphries, M.; Bijlsma, L. *J. Chem. Inf. Model.* **2022**, *62*, 5425.
- (68) Souihi, A.; Mohai, M. P.; Palm, E.; Malm, L.; Kruve, A. *J. Chromatogr. A* **2022**, *1666*, 462867.
- (69) Lowe, C. N.; Isaacs, K. K.; McEachran, A.; Grulke, C. M.; Sobus, J. R.; Ulrich, E. M.; Richard, A.; Chao, A.; Wambaugh, J.; Williams, A. *J. Anal. Bioanal. Chem.* **2021**, *413* (30), 7495–7508.
- (70) Nika, M.-C.; Alygizakis, N.; Arvaniti, O. S.; Thomaidis, N. S. *Curr. Opin. Environ. Sci. Health* **2023**, *32*, 100430.
- (71) Guo, J.; Huan, T. *Anal. Chem.* **2020**, *92* (12), 8072–8080.
- (72) Kind, T.; Fiehn, O. *Bioanal. Rev.* **2010**, *2* (1), 23–60.
- (73) de Juan, A.; Tauler, R. *Anal. Chim. Acta* **2021**, *1145*, 59–78.
- (74) Zweigle, J.; Buggel, B.; Zwiener, C. *Anal. Bioanal. Chem.* **2023**, *415*, 1791.
- (75) Fisher, C. M.; Croley, T. R.; Knolhoff, A. M. *Trends Anal. Chem.* **2021**, *136*, 116188.
- (76) Gago-Ferrero, P.; Krettek, A.; Fischer, S.; Wiberg, K.; Ahrens, L. *Environ. Sci. Technol.* **2018**, *52* (12), 6881–6894.
- (77) Royal Society of Chemistry ChemSpider: Data Sources. <http://www.chemspider.com/DataSources.aspx> [Date Accessed: September 2023].
- (78) National Institutes of Health PubChem: Data Sources. <https://pubchem.ncbi.nlm.nih.gov/sources#sort=Live-Substance-Count> [Date Accessed: September 2023].
- (79) Bolton, E.; Schymanski, E.; Kondic, T.; Thiessen, P.; Zhang, J. *J. Zenodo* **2023**, DOI: 10.5281/zenodo.5995885.
- (80) Patlewicz, G.; Wambaugh, J. F.; Felter, S. P.; Simon, T. W.; Becker, R. A. *Comput. Toxicol.* **2018**, *7*, 58–67.
- (81) Sussman, E. M.; Oktem, B.; Isayeva, I. S.; Liu, J.; Wickramasekara, S.; Chandrasekar, V.; Nahan, K.; Shin, H. Y.; Zheng, J. *ACS Biomater. Sci. Eng.* **2022**, *8* (3), 939–963.
- (82) Kroes, R.; Kleiner, J.; Renwick, A. *Toxicol. Sci.* **2005**, *86* (2), 226–230.
- (83) Hu, Y.; Du, Z.; Sun, X.; Ma, X.; Song, J.; Sui, H.; Debrah, A. A. *Food Chem.* **2021**, *345*, 128625.
- (84) Pinter, E.; Rainer, B.; Czerny, T.; Riegel, E.; Schilter, B.; Marin-Kuan, M.; Tacker, M. *Foods* **2020**, *9*, 237.
- (85) Sjerps, R. M. A.; Vughs, D.; van Leerdam, J. A.; ter Laak, T. L.; van Wezel, A. P. *Water Res.* **2016**, *93*, 254–264.
- (86) Yu, H.; Kiley, K.; Kullar, S.; Fu, K.; Tran, T. N.; Wang, H.; Hu, J.; Kamberi, M. *ACS Biomater. Sci. Eng.* **2023**, *9* (5), 2277–2291.
- (87) Cramer, G. M.; Ford, R. A.; Hall, R. L. *Food Cosmet. Toxicol.* **1976**, *16* (3), 255–276.
- (88) Palm, E.; Kruve, A. *Molecules* **2022**, *27* (3), 1013.
- (89) Sepman, H.; Malm, L.; Peets, P.; MacLeod, M.; Martin, J.; Breitholtz, M.; Kruve, A. *Anal. Chem.* **2023**, *95* (33), 12329–12338.
- (90) Knolhoff, A. M.; Fisher, C. M. *Food Chem.* **2021**, *350*, 128540.
- (91) Rosén, J.; Westerberg, E.; Pekar, H.; Cappelli, P.; Karki, A. J.; Mören, L.; Åstot, C.; Hellenäs, K.-E. *Water* **2022**, *14*, 2586.
- (92) Knolhoff, A. M.; Zweigenbaum, J. A.; Croley, T. R. *Anal. Chem.* **2016**, *88* (7), 3617–3623.
- (93) Rebryk, A.; Gallampos, C.; Haglund, P. *Sci. Total Environ.* **2022**, *829*, 154620.
- (94) Falconer, T. M.; Kern, S. E.; Brzezinski, J. L.; Turner, J. A.; Boyd, B. L.; Litzau, J. J. *Forensic Sci. Int.* **2017**, *270*, e5–e11.
- (95) Tkalec, Ž.; Codling, G.; Tratnik, J. S.; Mazej, D.; Klánová, J.; Horvat, M.; Kosjek, T. *Environ. Pollut.* **2022**, *313*, 120091.
- (96) Nikolopoulou, V.; Aalizadeh, R.; Nika, M.-C.; Thomaidis, N. S. *J. Haz. Mater.* **2022**, *428*, 128194.
- (97) Buckley, T. J.; Egeghy, P. P.; Isaacs, K.; Richard, A. M.; Ring, C.; Sayre, R. R.; Sobus, J. R.; Thomas, R. S.; Ulrich, E. M.; Wambaugh, J. F.; Williams, A. J. *Environ. Int.* **2023**, *178*, 108097.
- (98) Helbling, D. E.; Hollender, J.; Kohler, H.-P. E.; Fenner, K. *Environ. Sci. Technol.* **2010**, *44* (17), 6628–6635.
- (99) Ellis, L. B. M.; Gao, J.; Fenner, K.; Wackett, L. P. *Nucleic Acids Res.* **2008**, *36*, W427–W432.
- (100) Schollée, J. E.; Schymanski, E. L.; Avak, S. E.; Loos, M.; Hollender, J. *Anal. Chem.* **2015**, *87* (24), 12121–12129.
- (101) Spaan, K. M.; Seilitz, F.; Plassmann, M. M.; de Wit, C. A.; Benskin, J. P. *Environ. Sci. Technol. Lett.* **2023**, *10* (4), 328–336.
- (102) Djoumbou-Feunang, Y.; Fiamoncini, J.; Gil-de-la-Fuente, A.; Greiner, R.; Manach, C.; Wishart, D. S. *J. Cheminformatics* **2019**, *11* (1), 2.
- (103) Wicker, J.; Lorschach, T.; Gütlein, M.; Schmid, E.; Latino, D.; Kramer, S.; Fenner, K. *Nucleic Acids Res.* **2016**, *44* (D1), D502–D508.
- (104) Yuan, C.; Tebes-Stevens, C.; Weber, E. J. *Environ. Sci. Technol.* **2021**, *55* (9), 5950–5958.
- (105) Petras, D.; Koester, I.; Da Silva, R.; Stephens, B. M.; Haas, A. F.; Nelson, C. E.; Kelly, L. W.; Aluwihare, L. I.; Dorrestein, P. C. *Front. Mar. Sci.* **2017**, *4*, 405 DOI: 10.3389/fmars.2017.00405.
- (106) Zhao, Y.; Gericke, O.; Li, T.; Kjaerulf, L.; Kongstad, K. T.; Heskes, A. M.; Möller, B. L.; Jørgensen, F. S.; Venter, H.; Coriani, S.; Semple, S. J.; Staerk, D. *Anal. Chem.* **2023**, *95* (9), 4381–4389.
- (107) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapon, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya, P.; C, A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodriguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters,

K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Lington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. *Nat. Biotechnol.* **2016**, *34* (8), 828–837.

(108) Allard, S.; Le Daré, B.; Allard, P.-M.; Morel, I.; Gicquel, T. *Forensic Toxicol.* **2020**, *38* (2), 447–454.

(109) 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. *Anal. Bioanal. Chem.* **2022**, *414*, 1201.

(110) Zweigle, J.; Bugsel, B.; Fabregat-Palau, J.; Zwiener, C. PF Δ Screen - An open-source tool for automated PFAS feature prioritization in non-target HRMS data. *ChemRxiv* **2023**, *This content is a preprint and has not been peer-reviewed.* 1.

(111) Kirkwood, K. I.; Fleming, J.; Nguyen, H.; Reif, D. M.; Baker, E. S.; Belcher, S. M. *Environ. Sci. Technol.* **2022**, *56* (6), 3441–3451.

(112) Vegosen, L.; Martin, T. M. *Clean Technol. Environ. Policy.* **2020**, *22* (2), 441–458.

(113) Been, F.; Krueve, A.; Vughs, D.; Meekel, N.; Reus, A.; Zwartsen, A.; Wessel, A.; Fischer, A.; ter Laak, T.; Brunner, A. M. *Water Res.* **2021**, *204*, 117612.

(114) Lunghini, F.; Marcou, G.; Azam, P.; Patoux, R.; Enrici, M. H.; Bonachera, F.; Horvath, D.; Varnek, A. *SAR QSAR Environ. Res.* **2019**, *30* (7), 507–524.

(115) Sushko, I.; Salmina, E.; Potemkin, V. A.; Poda, G.; Tetko, I. V. *J. Chem. Inf. Model.* **2012**, *52* (8), 2310–2316.

(116) Reif, D. M.; Martin, M. T.; Tan, S. W.; Houck, K. A.; Judson, R. S.; Richard, A. M.; Knudsen, T. B.; Dix, D. J.; Kavlock, R. J. *Environ. Health Perspect.* **2010**, *118* (12), 1714–1720.

(117) Rager, J. E.; Strynar, M. J.; Liang, S.; McMahan, R. L.; Richard, A. M.; Grulke, C. M.; Wambaugh, J. F.; Isaacs, K. K.; Judson, R.; Williams, A. J.; Sobus, J. R. *Environ. Int.* **2016**, *88*, 269–280.

(118) Yang, J.; Zhao, F.; Zheng, J.; Wang, Y.; Fei, X.; Xiao, Y.; Fang, M. *J. Haz. Mater.* **2023**, *448*, 130893.

(119) Peets, P.; Wang, W.-C.; MacLeod, M.; Breitholtz, M.; Martin, J. W.; Krueve, A. *Environ. Sci. Technol.* **2022**, *56*, 15508.

(120) Groff, L. C.; Grossman, J. N.; Krueve, A.; Minucci, J. M.; Lowe, C. N.; McCord, J. P.; Kapraun, D. F.; Phillips, K. A.; Purucker, S. T.; Chao, A.; Ring, C. L.; Williams, A. J.; Sobus, J. R. *Anal. Bioanal. Chem.* **2022**, *414* (17), 4919–4933.