Abstract Book

NonTarget 2022

Nontarget Analysis for Environmental Risk Assessment

22–26 May 2022 | Durham, North Carolina, USA + Online

SETAC®

SOCIETY OF ENVIRONMENTAL TOXICOLOGY AND CHEMISTRY FOCUSED TOPIC MEETING

Abstract Book

Nontarget Analysis for Environmental Risk Assessment
SETAC Focused Topic Meeting

Table of Contents

This book comprises the abstracts of the presentations for the platform and poster sessions of the Society of Environmental Toxicology and Chemistry (SETAC) Focused Topic Meeting on "Nontarget Analysis for Environmental Risk Assessment," conducted in-person and online from 22–26 May 2022. The abstracts are reproduced as accepted by the Scientific Program Committee.

In each abstract, the presenting author's name is listed in bold letters. The author index cross-references the corresponding page numbers.

International Standard Serial Number 1087-8939

About SETAC

The Society of Environmental Toxicology and Chemistry (SETAC), with offices in North America and Europe, is a nonprofit, professional society established to provide a forum for individuals and institutions engaged in the study, analysis and solution of environmental problems, the management and regulation of natural resources, environmental education, and research and development.

Specific goals of the society are:

- Promote research, education and training in the environmental sciences
- Promote the systematic application of all relevant scientific disciplines to the evaluation of chemical hazards
- Participate in the scientific interpretation of issues concerned with hazard assessment and risk analysis
- Support the development of ecologically acceptable practices and principles
- Provide a forum (meetings and publications) for communication among professionals in government, business, academia and other segments of society involved in the use, protection and management of our environment

These goals are pursued through the conduct of numerous activities, which include:

- Conduct meetings with study and workshop sessions, platform and poster presentations, and achievement and merit awards
- Publish scientific journals, a newsletter and special technical publications
- Provide funds for education and training through the SETAC Scholarship and Fellowship Program
- Organize and sponsor chapters and branches to provide a forum for the presentation of scientific data and for the interchange and study of information about local and regional concerns
- Provide advice and counsel to technical and nontechnical persons through a number of standing and ad hoc committees

SETAC membership currently comprises about 4,500 individuals from government, academia, business and nongovernmental organizations with backgrounds in chemistry, toxicology, biology, ecology, atmospheric sciences, health sciences, earth sciences, environmental engineering, hazard and risk assessment, and life cycle assessment.

If you have training in these or related disciplines and are engaged in the study, use or management of environmental resources, SETAC can fulfill your professional affiliation needs.

All members receive the SETAC Globe newsletter highlighting environmental topics and SETAC activities, reduced fees for meetings and discounts on SETAC books. All members receive online access to *Environmental Toxicology and Chemistry* (ET&C) and *Integrated Environmental Assessment and Management* (IEAM), the peer-reviewed journals of the society. Members may hold office and, with the Emeritus Members, constitute the voting membership.

For further information, contact one of our offices:

SETAC Africa
SETAC Europe
Avenue des Arts, 53
B-1000 Brussels, Belgium
T +32 2 772 72 81
E setaceu@setac.org

SETAC Latin America
SETAC North America
PO Box 12804
Pensacola, Florida, 32591-2804 USA
T +1 850 469 1500
E setac@setac.org

SETAC Asia-Pacific
Avenue des Arts, 53
B-1000 Brussels, Belgium
T +32 2 772 72 81
E setaceu@setac.org

www.setac.org

Environmental Quality Through Science®

Advances in Instrumentation and Approaches

Keynote: Utilizing Ion Mobility Spectrometry, Mass Defect Analysis and Machine Learning to Uncovering Xenobiotics in Nontargeted Analyses (Invited Keynote)

***Erin Baker**, North Carolina State University*

The identification of xenobiotics in nontargeted analyses is a vital step in understanding human exposure. Xenobiotic metabolism, excretion, and co-existence with other endogenous molecules however greatly complicates nontargeted studies. While mass spectrometry (MS)-based platforms are commonly used in these measurements, deconvoluting endogenous metabolites and xenobiotics is often challenged by the lack of xenobiotic parent and metabolite standards as well as the numerous isomers possible for each m/z feature. This presentation will demonstrate how combining mass defect analysis, machine learning and liquid chromatography, ion mobility spectrometry and mass spectrometry (LC-IMS-MS) separations provides an in-depth understanding of molecules present and molecular responses occurring due to chemical exposures.

Improved Compound Identification via Simultaneous GC-CI&EI-TOFMS Analysis

***Sonja Klee** and Steffen Braekling, TOFWERK, Switzerland*

Non-target analysis (NTA) approaches are already well-implemented using liquid chromatography (LC) coupled to mass spectrometry (MS) instrumentation in applications as e.g., water and food analysis. In comparison using gas chromatography (GC) coupled to MS for NTA is still an emerging field. Most commonly, GC is used in combination with 70 eV electron ionization (EI) installed to a low-resolution quadrupole mass spectrometer (QMS). In this combination reliable compound identification strongly depends on existing mass spectral libraries and therefore is highly limited for NTA approaches. Although EI based libraries are very comprehensive, non-targeted high-level compound identification can become difficult when the generated fragment mass spectra are not characteristic enough or the compound is not listed in the library. Another issue using EI as ionization method in NTA is the frequent absence of molecular ion information. Later is offering valuable information about the compound and is often lost during the high energetic EI process. Even when obtained, in a GC-QMS experiment the molecular ion signal is just available without high mass resolution (HR) and accurate mass information. Chemical ionization can increase the chance of molecular ion preservation and offers, in combination with HRMS, the possibility of sum formula annotation for the compound in many cases. The combination of EI and CI information for NTA approaches using GC-HRMS appears to be the ideal setup for the identification process of volatile substances. This way structural, molecular and physico-chemical information of a compound can be used to increase the level of certainty in the identification. A GC coupled high resolution time-of-flight (HRTOF) instrument operating separate EI and CI sources simultaneously is presented. After the GC separation the effluent is split, and each sample part is lead to one of the ion sources. After the ionization process the sample is alternately lead via fast switchable ion optics (up to 100 Hz) into the TOF mass analyzer. Undisturbed EI and CI mass spectra, respectively, are constantly generated quasi-simultaneously during a single chromatographic separation step. The principle of operation is shown using different GC standard mixtures. Increased identification capabilities for NTA are shown using samples from different application e.g., material emission. The advantages of chemical ionization are presented via usage of various reactant gases.

Ion Mobility Separations Incorporated in a Non-Target High Resolution Mass Spectrometry Screening Workflow

***Alberto Celma Tirado**, Swedish University of Agricultural Sciences*

The analysis of environmental samples, particularly when dealing with high matrix complexity, low concentrations, and thousands of potential contaminants, is a challenging task. High resolution mass spectrometry (HRMS) is a powerful technique to identify emerging organic contaminants. However, comprehensive HRMS screening strategies typically result in large and complex data sets of features and the burden of work when revising this data can be overwhelming. Despite that, huge steps forward and ongoing

efforts are made to further develop/improve non-target screening workflows. Ion mobility separation (IMS) combined with HRMS analyzers introduces a new dimension to this workflow, which allows to further adjust and improve data processing and compound identification, making use of the drift time of an ion i.e., the time an ion takes to travel through the mobility cell. The collision cross section (CCS) of an ion is a value unique to IMS, which can be derived from the drift time and provides information about the shape and size of an ionized molecule. The use of CCS values to identify compounds, alongside retention time (RT) and accurate mass, is of special interest to analytical chemists since it is known to be unaffected by the matrix. The creation of empirical CCS databases is, however, pivotal for an enhanced and robust screening strategy. In this work, a CCS library developed for hundreds of organic contaminants is presented for its application to monitor aquatic samples using IMS-HRMS instruments. Attention has been paid to the much cleaner and higher-quality spectra obtained after drift time alignment of (de)protonated molecules and their fragment ions. Moreover, the use of prediction tools to prioritize features and/or improve the identification capabilities in high-complex samples will be discussed. This will be supported by some illustrative examples of contaminants found. Finally, the information and expertise gained during the study is used to propose the inclusion of CCS values as additional parameter in widely adopted confidence levels for the identification of compounds. Hence, ion mobility separations incorporated non-target HRMS screening workflow provide an additional criterion for reliable identification and can considerably alleviate the burden on revising large data sets. We expect IMS-HRMS to be more and more applied in different areas of analytical research, in the near future.

On-Site Screening with a Transportable High-Resolution Mass Spectrometer Platform -Insights into Water-Quality Dynamics in Real Time

***Heinz Singer**, Michael Stravs, Christoph Ort, Juliane Hollender, Julian Bosshard, Johannes Schorr and Philipp Longree, Eawag - Swiss Federal Institute of Aquatic Science and Technology, Switzerland*

High-resolution mass spectrometry coupled to liquid chromatography (LC-HRMS) has become an established technology for the comprehensive analysis of target, suspect and non-target compounds in water samples. However, LC-HRMS measurements still rely on conventional workflows consisting of sample collection in the field, transportation to the lab, and storage with subsequent preparation and measurement. This procedure is time- and cost-intensive and limits therefore the number of samples analyzed. Furthermore, the risk of degradation during sampling, transport and storage is high for labile compounds. To overcome those limitations, a transportable high-resolution mass spectrometer platform was developed that enables on-site monitoring of water quality with high time resolution. The technical setup of the field deployable unit, MS2field, consists of a water bypass loop, a self-cleaning filtration device, an online solid-phase extraction-LC system and a Q-Exactive HF mass spectrometer. The Q-Exactive acquires full scan mass spectra with either data-dependent or data-independent MS/MS in polarity switching mode over a 20min LC run. The analysis is fully automated and allows autonomous operation in the field. A programmable logic controller connected to multiple sensors, a surveillance camera, and remote internet access enables monitoring of the system and failsafe operation. Measured LC-HRMS files are automatically processed and the results transferred immediately to an online platform with a live view for selected compounds. To demonstrate the performance of the developed platform under field conditions, measurements in a wastewater treatment plant, in a small creek, in a larger river, and in a karst spring were performed over several weeks. In total more than 7,000 measurements were acquired at the investigated field sites. The comprehensive dataset was processed with an in-house software workflow in a targeted, suspected and non-targeted fashion. To illustrate the potential of the MS2field platform, several examples like time series of pharmaceuticals in industrial wastewater as well as pesticide dynamics in surface waters will be presented. In addition, the analysis of time series measurements in untreated wastewater using frequency domain transformation and unsupervised clustering will be demonstrated. With this analysis, the pronounced temporal dynamics of identified non-target compounds originating from different sources (industry, households, buildings) could be demonstrated.

Applying Nontargeted Analyses to Identify Per- and Polyfluoroalkyl Substances, Pesticides, and Pesticide Transformation Products in North Carolina Water

***Nancy Lee McLean Alexander**, James Dodds, Noelle DeStefano, Catherine LePrevost, Erin Baker and Detlef Knappe, North Carolina State University*

North Carolina (NC) is a national leader in agriculture and agribusiness with $92.7 billion, or over 16% of the gross state product, attributed to food, fiber, and forestry. However, to meet increasing demands and obtain optimal production, many NC farmers rely on the use of pesticides and other agricultural chemicals. As a result, NC water sources are vulnerable to impacts from emerging organic contaminants, including pesticides, pesticide transformation products, and per- and polyfluoroalkyl substances (PFAS), which are directly associated with pesticide products, pesticide packaging, and agricultural biosolids application. As the list of pesticide active ingredients, inactive ingredients, and packaging byproducts continues to evolve, it is imperative that we develop nontargeted analytical approaches to assess the impact of these chemicals on drinking water quality. To address this challenge, we sampled surface water near 9 biosolids application sites and 80 private wells in NC and performed analysis with liquid chromatography and quadrupole time-of-flight mass spectrometry (LC-MS). The LC-MS data were then assessed using various identification algorithms, an online WebApp, and a Python script developed by the US EPA to generate a list of tentative identifications. An important observation was that pesticides such as atrazine and imidacloprid were among the list of organic contaminants detected using this workflow. To provide additional identification capabilities and confidence, a nontargeted analytical workflow coupling LC, ion mobility spectrometry and MS (LC-IMS-MS) was employed. Since pesticide and PFAS standards have illustrated unique relationships for the IMS collision cross section (CCS) and exact mass values, the multidimensional LC-IMS-MS information for each detected feature provided additional confidence in our identifications. Currently, the multidimensional data is being processed to improve our understanding of drinking water quality and provide faster remediation response. These results will be showcased in this presentation.

Advances in Instrumentation and Approaches (Poster)

A Non-Target Approach to Identify Complexing Agents in Atmospheric Aerosol Particles and Rainwater

***Jan Beck**, Leibniz Institute for Tropospheric Research, Germany*

Atmospheric aerosol samples from the gas and liquid phases consist of numerous organic substances. For example, secondary organic aerosol (SOA), which is formed by oxidation of biogenic and anthropogenic volatile organic compounds (VOCs), can contain up to 100 000 different organic substances. Some of these can form complexes in aerosol particles together with metal cations of iron, copper, or manganese. Photochemical conversion reactions can be triggered by such complexes and they contribute to the enhanced solubility of metal ions into the atmospheric aqueous phase. Complexing agents are therefore important organic components of aerosol particles, but rather difficult to identify. In this study, we developed a liquid chromatography coupled with high resolution mass spectrometry (LC/HRMS) non-target screening (NTS) approach, which allows the selective detection of complexing agents in aerosol particle extracts and rainwater. Therefore, a T-junction is installed between the LC outlet and the ion source to implement a post-column complexation with a 155 ¬µM FeCl3 solution. Subsequently, the obtained mass spectra are analyzed for the three characteristic iron complexes [M-H+FeCl3]-, [M-2H+FeCl2]- and [M-3H+FeCl]- with mass differences (‚àÜm/z) between the signal of the respective complexing agent and the iron complex of 160.8416, 124.8648, and 89.8959. In aerosol particle samples from Germany and China, up to 29 different di- and tricarboxylic acids could be detected as complexing agents, even at low extract concentration of only 50 nM. Thirteen complexing agents are detected even in measurements without post-column iron addition from complexation with background Fe3+ traces from the analytical system. At least for the most highly concentrated complexing agents, the proposed screening approach is thus applicable without instrument modification, making it particularly attractive for use in routine analysis. Besides carboxylic acids, 4-Nitrophenol and 4-Nitrocatechol were identified as further complexing

agents in rainwater samples, which shows the applicability of the method to other matrices and to a range of different complexing agents.

A Comparison of Solid Phase Extraction Methods for Non-Target Water Analysis
Alyssa Mianecki *and Gregory LeFevre, University of Iowa*
Non-target, high-resolution mass spectra are influenced by the initial sample preparation method. Compounds in water samples are highly selective to different solid phase extraction (SPE) methods. Selectivity is a desired trait for targeted compound analysis in water to lower detection limits and to reduce background, however, selective SPE methods used for non-target analysis of water can significantly alter the overall spectral signatures. Extraction methods may introduce unintended bias into non-target environmental samples. The objective of this study was to compare spectral signatures of wastewater impacted stream water using multiple SPE extraction methods. For this study, water samples (1L) were collected in parallel from Muddy Creek in Coralville, Iowa. Parallel water samples were then filtered with no extraction, extracted with an Oasis HLB cartridge, or extracted with a Strata X-C cartridge to allow for direct spectral comparison. A UPLC-Q-Exactive was used for high-resolution, non-target screening and all samples were subjected to identical instrumentation. Compounds including pharmaceuticals, personal care products, and pesticides are identified in the samples. Due to different chemical affinities to the SPE sorbents, we anticipate that identified compounds in the water samples will be significantly different in each of the extracted samples.

A Practical Lock-Mass Calibrant Introduction Method for Improved Mass Accuracy and Reduced False Positive Identifications in Non-Targeted Analysis
Christine Fisher (O'Donnell), ***Brian Ng****, Shannon Murphy and Ann Knolhoff, U.S. Food and Drug Administration*
In addition to isotopic distribution, accurate mass obtained with liquid chromatography/high resolution-mass spectrometry (LC/HR-MS) is vital for unknown identification in complex samples using Non-Targeted and Suspect Screening Analysis (NTA and SSA, respectively). Minimizing mass accuracy error improves molecular formula generation and reduces false positive identifications. Lower mass accuracy errors can be achieved with an internal/lock-mass recalibration. Background ions can be used as lock-masses; however, they require identification and can change over time. We have developed a reliable method for introducing lock-mass calibrant ions to a Thermo Q-Exactive mass spectrometer using a calibrant reservoir kit (Bruker) placed ~2 cm inside the window of the HESI source. Briefly, the kit consists of a stainless-steel reservoir that holds 2 sponges with absorbed calibrant solution and a perforated cap. The calibrants used included methyl stearate ($[M+H]^+$, *m/z* 299), hexakis (2,2-difluoroethoxy)phosphazene ($[M+H]^+$, *m/z* 622; $[M+formate]^-$, *m/z* 666) and hexakis (1H,1H,3H-tetrafluoropropoxy)phosphazene ($[M+H]^+$, *m/z* 922; $[M+formate]^-$, *m/z* 966). Co-dosing the lower *m/z* calibrants on the same sponge under a 5-hole cap and the highest *m/z* calibrant on a separate sponge under a 12-hole cap generated stable signal for all calibrants for at least 36 hrs. The lock-masses used in the method can be set to "best" (using the best calibrant for recalibration) or "if all present" ("all"; requiring all calibrants to be present for recalibration). Both settings corrected mass accuracy errors of 3-5 ppm to < 1 ppm for >92% of the ~80 compounds in our standard mixture. In comparison, using the background ions diisooctyl phthalate ($[M+H]^+$, *m/z* 391) and sodium formate ($[2M-2H+Na^+]^-$, *m/z* 112), reduced mass accuracy errors to < 1 ppm for 73% of the standard compounds. Using Compound Discoverer, top-ranked, correct molecular formula were obtained for 30/67 ("best") and 35/67 ("all") standard compounds compared to 23/67 without recalibration. MS/MS match scores improved for 43/52 compounds with lock-mass calibration, with a slight preference for the "all" condition. However, the "all" condition resulted in fewer recalibrated spectra throughout the chromatogram compared to the "best" condition. Therefore, the "best" condition is considered optimal for NTA. This reliable method for lock-mass calibrant introduction should also be amenable to other instruments using the HESI source.

Non-Target-Screening in Sediment and Suspended Particulate Matter by Liquid-Chromatography-High Resolution Mass Spectrometry: A Sample Preparation Protocol

***André Macherius**, Uwe Kunkel and Manfred Sengl, Bavarian Environment Agency (LfU), Germany*

Suspect- and Non-Target-Screening studies on water samples using liquid chromatography-high resolution mass spectrometry are frequently applied to assess the surface water quality and to identify so far not considered anthropogenic compounds. However, for sediment and suspended particulate matter (SSPM) respective studies are less frequent, even though they act as a sink for hydrophobic and ionic compounds. Therefore, more comprehensive data on the past and present status of water bodies might be derived from SSPM analysis. One reason for the comparatively small number of NTS studies on SSPM is the lack of appropriate sample preparation protocols including sample extraction and clean-up. Considering structurally and physicochemically diverse suspects and non-targets and the high content of matrix compounds in SSPM pose challenges for the method development. We present a sample preparation method based on an extraction by pressurized liquid extraction (PLE). The protocol was optimized based on the recoveries and matrix effects of a set of 247 target compounds covering a broad range of applications as well as structural and physicochemical properties. The three-step extraction of a spiked test sediment using PLE provides a median recovery rate of 0.9 and 80% of the test compounds show recovery rates between 0.5 and 1.1. However, 66% of the investigated compounds are subject to moderate to strong matrix effects. To reduce these effects, the PLE extracts are treated with two parallel clean-up steps by 1) gel permeation chromatography (GPC) and 2) normal phase adsorption chromatography with aluminum oxide. This reduces the matrix content so that only 27% of the test compounds are affected by matrix effects. However, the median recovery also decreases to 0.7 and 20% of the compounds show recovery rates below 0.3. The protocol was applied to SSPS samples from two German rivers, in which 87 and 90 compounds were tentatively identified by suspect-screening. In both samples, organophosphate flame retardants and quaternary ammonium compounds were among the most intense signals, but also PFAS, pharmaceuticals, pesticides and 6PPD were identified. Next steps will include the further optimization of the clean-up to increase recoveries of the so far poorly retrieved compounds. Furthermore, quality control measures such as a reference SSPM matrix as well as appropriate internal standards will be implemented to establish a protocol applicable for monitoring campaigns.

Applications to Ecological Exposure and Effects Assessment

Keynote: Effect-Directed Analysis: Integration of Effect-Based Tools into the Suspect/Nontarget Screening Workflow for Identification of Toxic Chemicals

***Marja Lamoree**, Environment & Health, VU University, Department Environment & Health, Netherlands*

The quality of our living environment is of great importance for our health and well-being. Unfortunately, environmental quality is under significant pressure due to many reasons, of which pollution by man-made chemicals is a relevant one. The classical approach for exposure assessment using environmental but also human monitoring based on only a few pre-selected chemicals needs revision in order to keep track of the mixtures of chemicals that are present in our environment. The rapid developments in the field of suspect and nontarget screening are promising: ideally, this type of comprehensive data, carefully archived, is expected to permit a more holistic assessment of the presence of certain chemicals in environmental but also human samples. The scientific community that focuses on data processing workflows and annotation is very lively, and major advancements are reported regularly, showing a steady movement of the field into the realm of big data.

In this presentation, I will highlight the benefits of the inclusion of effect-based, in vitro tools to enhance the identification success of suspect and nontarget screening. In Effect-Directed Analysis, the detection of bioactive chemicals of emerging concern (CECs) is pursued by combining toxicity testing and high-resolution mass spectrometry (HRMS). In recent years, we have achieved higher throughput at the effect-based side of our approach, using high resolution fractionation of samples and extracts, and progress was also made by the development of the CECscreen database.

Although our initial efforts were focused on sediments and biota, the further development of EDA towards a more high throughput approach was done with surface and wastewater as matrices of choice. For these EDA studies, we focused on the identification of bioactive chemicals interfering with steroid hormone systems but also with mutagenic or antibiotic activity. Based on the notion that environmental and human health are closely interlinked, the application domain of our high throughput EDA platform has shifted towards alternative approaches for human exposure assessment, firstly by studying human samples, but possibly also towards the identification of human metabolites of chemicals of emerging concern in community wide exposure studies using wastewater.

Effect-Directed Analysis on Water Samples to Identify Potent Antimicrobial Compounds With Suspect- and Nontarget Screening

Tim Jonkers[1]*, Peter Keizers*[2]*, Frederic Magnus Been*[3]*, Jeroen Meijer*[1]*, Corine Houtman*[4]*, Timo Hamers*[5] *and Marja Lamoree*[5]*, (1) Vrije Universiteit, Netherlands, (2)Dutch National Institute for Public Health and the Environment (RIVM), Netherlands, (3) KWR Watercycle Research Institue, Netherlands, (4) The Water Laboratory, Netherlands, (5) VU University Amsterdam, Netherlands*

Antibiotics in the environment may cause adverse biological effects on biota and promote bacterial resistance. Also, the metabolites and transformation products of antibiotics that are often not included in analytical monitoring programs, may still be biologically active. An effect-directed analysis (EDA) study was performed to identify contaminants of emerging concern with antimicrobial properties in the water cycle. SPE-extracts were prepared from influent, effluent, and surface water (up and downstream of the discharge location) of two wastewater treatment plants. High-resolution fractionation of the extracts was performed with liquid chromatography and a spotter device (FractioMate^{TM}), collecting 80 fractions per sample. Sample extracts were also subjected to high-resolution mass spectrometry analysis on a quadrupole-time-of-flight (QTOF) mass spectrometer for suspect and nontarget screening. An antibiotics bioassay that is sensitive to a broad range of antibiotics was applied to identify bioactive fractions. Chemical features were annotated with spectral libraries (EU Massbank, MoNA), suspect screening libraries, and analytical standards. Suspect libraries included CECscreen, the S6 ITN Antibiotics NORMAN list, and *in silico* predicted phase-I metabolites of the S6 ITN list (predictions performed with BioTransformer 3.0). Annotations were assigned an identification confidence level. Further, the Retention Time Indices platform (University of Athens) was applied to predict retention times based on the annotated structure. Five bioactive fractions overlapped between the fractionated effluent extracts of the two locations. The macrolide antibiotics azithromycin and clarithromycin were identified at a high identification confidence level, explaining the bioactivity of two fractions. In a third fraction, an *in silico* predicted hydroxylated clarithromycin metabolite was tentatively identified. Possible positions of hydroxylation in the structure were suggested based on the MS2 spectra, although chemical confirmation is still required. Other tentatively identified metabolites were not expected to cause bioactivity based on literature. The annotations of the remaining two bioactive fractions were prioritized further on signal intensity, retention time predictions, and MetFrag to match predicted and measured MS2 fragments. In this presentation, we will discuss the biological and chemical confirmation of the identified compounds with antimicrobial activity.

A Holistic Approach for a Comprehensive Water Quality Evaluation by Integrating Chemical and Biological Analyses

Alberto Celma Tirado[1]*, Geeta Mandava*[1]*, Frank Menger*[1]*, Foon Yin Lai*[1]*, Oksana Golovko*[2]*, Juan V Sancho*[3]*, Felix Hernandez*[3]*, Lutz Ahrens*[1]*, Karin Wiberg*[1]*, Agneta Oskarsson*[1]*, Johan Lundqvist*[1] *and Lubertus Bijlsma*[3]*, (1) Swedish University of Agricultural Sciences, (2) University of South Bohemia, Sweden, (3) Jaume I University, Spain*

High resolution mass spectrometry (HRMS) is gaining popularity in environmental research due to its powerful capability for wide-scope (non)target screening analyses, making it possible to obtain large data set for identifying relevant chemicals present in a sample. With the recent development of ion mobility separation (IMS) coupled to HRMS, the capabilities of the instruments have been enhanced. However, dealing with such

large complex data sets to be mined for the detection and identification of (un)known chemicals is time-consuming. Thus, feature prioritization strategies are pivotal for this task. From a different perspective, *in vitro* biological analysis enables the establishment of the toxicological fingerprint of a sample, and therefore measures the combined effects of all chemicals present in the sample. Therefore, biological analysis provides additional information for water quality evaluation alongside the application of IMS-HRMS screening strategies. The Spanish Mediterranean basin is particularly susceptible to climate change and human activities, making it vulnerable to the influence of anthropogenic contaminants. Hence, conducting comprehensive water quality assessment in relevant water bodies of this basin is of high interest. In this work, eight *in vitro* bioassays (aryl hydrocarbon receptor (AhR), estrogen receptor (ER) activation and antagonism, nuclear factor erythroid 2-related factor 2 (Nrf2), androgen receptor (AR) activation and antagonism, and vitamin D receptor (VDR) activation and antagonism) were applied for establishing the toxicological fingerprint of 11 surface water samples from coastal lagoons along the Mediterranean coast of Spain. Additionally, a screening for more than 1,500 organic micropollutants (OMPs), together with the application of *in silico* prediction tools for retention time and collision cross section values, revealed the presence of 96 OMPs. Pharmaceuticals and pesticides were the most prevalent groups of compounds. However, chemical analysis has limitations in revealing the complete water quality. Thus, in this work, measured biological activities were integrated with chemical observations by means of *in silico* toxicity prediction tools to gather an improved overview of the quality of the studied water bodies. Our work will present a stepwise approach for the complete untargeted workflow from *in vitro* bioassays to feature detection, *in silico* prediction and final compound identification.

Suspect and Non-Target Screening of Emerging Contaminants in the Arctic Environment

Linyan Zhu[1]*, Rossana Bossi*[1]*, Pedro Carvalho*[1]*, Frank Riget*[1]*, Jan Christensen*[2]*, Pal Weihe*[3]*, Eva Bonefeld-Jørgensen*[1] *and Katrin Vorkamp*[1]*, (1) Aarhus University, Denmark, (2) University of Copenhagen, Denmark, (3) The Faroese Hospital System, Faroe Islands*

Chemicals emitted at lower latitudes can reach the Arctic if they are persistent and can be transported over long-distance. These characteristics, together with bioaccumulation and toxicity, are criteria of the United Nations Stockholm Convention regulating Persistent Organic Pollutants (POPs). POPs and mercury have been monitored in the Arctic for several decades. More recently, interest has increased in "chemicals of emerging Arctic concern", also including contaminants associated with consumer products and potentially locally emitted in the Arctic. These have commonly been addressed in target analysis of a limited number of substances. Given the high number of chemicals in commerce, it seems likely that additional chemicals with large production and use volumes could occur in the Arctic, potentially causing exposure to animals and humans. We have conducted suspect and non-target screening (NTS) of environmental samples (biota and air) from Greenland using high resolution mass spectrometry in combination with gas and liquid chromatography (GC- and LC-HRMS). Based on a list derived from *in silico* screening of chemicals with potential POP characteristics, an in-house library was built. Suspect screening was limited to GC-HRMS and based on matches with the in-house library. NTS was performed to identify unknown compounds by both GC- and LC-HRMS. Air samples originated from high volume air sampling at Villum Research Station (Northeast Greenland). Filters and polyurethane foam were used for GC-HRMS analysis while only filters were used for LC-HRMS analysis. Biota samples included ringed seals (*Pusa hispida*) and pilot whales (*Globicephala metas*), both from East Greenland. Blubber and liver samples were analysed by GC- and LC-HRMS, respectively. Sample preparation was optimized towards balanced non-selectivity vs. matrix removal. Based on preliminary results from suspect screening, tetrabromophthalic anhydride was found in blubber samples. Multiple regulated POPs were found using NTS. Other compounds preliminarily detected in the NTS approach included chlorinated cyclic aliphatic and aromatic hydrocarbons, chlorinated pyrroles, methoxylated polybrominated diphenyl ethers and phthalates. Following further study to increase identification confidence, these chemicals could be considered for more systematic studies in the Arctic, to explore their occurrence and potential risks.

Data Mining for Linking Environmental Exposures to Toxic Effects via a Hazard Driven Prioritisation

Katarzyna Arturi[1]*, Alan J. Bergmann*[2] *and Juliane Hollender*[1]*, (1) Eawag - Swiss Federal Institute of Aquatic Science and Technology, Switzerland, (2) Oekotoxzentrum Eawag-EPFL, Switzerland*

High-resolution mass spectrometry (HRMS) followed by nontarget screening (NTS) workflows for converting thousands of molecular 'features' into quantified chemical structures have started to become cornerstone tools for comprehensive mapping of global pollution. However, since a complete elucidation of a large number of HRMS signals is not yet feasible, prioritization of the "relevant" chemicals is a crucial step. While the traditional HRMS prioritization approaches are based exclusively on signals' abundance or frequency outlying environmental exposures, we propose a methodology taking into account the pollutants' toxicological relevance assessing overall risks (risk = hazard [toxicity] x exposure [concentration]). For that purpose, we have developed MLinvitroTox, a machine learning (ML) tool for the prediction of toxicity fingerprints (ToxFps) for thousands of unknown HRMS features based on their MS and MSMS characteristics. MLinvitroTox was trained on ToxCast and Tox21 (high-throughput screening toxicity databases from EPA), and it joins under one roof a plethora of optimized supervised classification algorithms (naive Bayes, logistic regression, generalized linear models, decision trees, random forests, extreme boosted trees, as well as neural networks) to predict over 500 toxic molecular endpoints used to compile ToxFps. Since each endpoint represents a cellular-level perturbation caused by a chemical/target interaction, the methodology is suitable for ML-derived linking of structural moieties with molecular-level disturbances in pathways that can be subsequently linked through adverse outcome pathways (AOPs) to specific toxic outcomes in living organisms. MLinvitroTox coupled to the existing sophisticated NTS tools results in a novel data mining pipeline (EXPECTmine), incorporating, in addition to a hazard-driven prioritization, a putative identification of the prioritized signals via various molecular structure annotation packages (MetFrag, CSI:FingerID) and tentative quantification via ML-derived ionization efficiency prediction (Quantem). The pipeline is under continued development and has so far shown excellent accuracy rates (more than 90 % of correctly predicted single toxic outcomes/endpoints) for predicting ToxFps of genotoxic model compounds such as 4-nitroquinoline-1-oxide, nalidixic acid, and nitrofurantoin, as well as genotoxicants such as 5-chloro-2-methyl-4-isothiazolin-3-one (CMIT) suspected to be present in food packaging materials.

Applications to Ecological Exposure and Effects Assessment (Poster)

Application of Target, Non-Target and Suspect Screening to Explore the Occurrence of Anthropogenic Chemicals in the Open Baltic Sea

Jana Johansson*, Stockholm University, Sweden*

The global oceans are thought to be recipients of many anthropogenic substances but they can also act as a source of pollutants back to the atmosphere via sea spray aerosol (SSA). However, the number of anthropogenic substances regularly monitored in seawater is low and few studies have attempted to measure emerging substances. One reason for this can be anticipated low levels, as a consequence of dilution, which would make detection and identification challenging. The Baltic Sea is a semi-enclosed water body with an industrialized and densely populated catchment area. It is known as one of the most polluted seas globally. The aim of this study is to use this polluted sea to i) assess the occurrence of potential seawater pollutants rarely monitored in the marine environment, and ii) identify pollutants which may be transferred to the atmosphere via SSA, using enrichment in the sea surface microlayer (SML) as a proxy for enrichment in SSA. Paired seawater and SML samples were collected in the Baltic Sea in June 2018. Filtered samples were solid phase extracted on HLB sorbent. Analysis was performed by liquid chromatography high-resolution mass spectrometry on a Quadrupole-Orbitrap mass spectrometer. A targeted method was applied to determine the concentrations of a set of known micropollutants, predominantly pharmaceuticals and perfluoroalkyl acids. Suspect screening was performed using a curated suspect list of ca. 1000 features in combination with matching against three mass spectral libraries: mzCloud, MassBank and an in-house mzVault library which covers ca. 250 compounds. Additionally, non-targeted screening was applied to identify 'unknown' chemicals which enrich in the SML:

peak areas were used to calculate SML enrichment factors for each feature in each sample pair. For the substances which had enrichment factors >2, tentative identification was achieved using the above-mentioned mass spectral databases and the in silico fragmentation model SIRIUS. Overall, a wide variety of compounds were identified at levels 1-3. The results of this study can inform policy as well as future monitoring programs.

A Non-Target Screening-Directed Long-Term Time-Trend Study of Organic Contaminants in the Baltic Top Consumers

Andriy Rebryk *and Peter Haglund, Umea University, Sweden*

The key species of the Baltic ecosystem suffer from impaired health. It is believed that anthropogenic hazardous substances along with some halogenated natural products are one of the reasons for that, as they accumulate in organisms. Since international restrictions on hazardous chemicals went into force in the 1970-80s, time trends of many environmental persistent organic pollutants (POPs), including PCBs, DDT and its metabolites, and others, have been decreasing constantly for decades. However, the amounts of legacy POPs in the Baltic Sea region have leveled off and stopped declining recently and new contaminants are frequently discovered. There is, hence, a need for comprehensive approaches to study the occurrence and time trends of the legacy POPs and contaminants of emerging concern (CECs), amounts of which are increasing, in the ecosystem. Here, we report the first time-trend study which utilizes a non-target screening (NTS) approach for this purpose. The studied samples included common guillemot egg extract (1986-2019, 12 samples), white-tailed sea eagle muscle (1965-2017, 8 samples), and harbor porpoise blubber (1988-2019, 9 samples). A major challenge in NTS of biota is the removal of matrix components, like lipids that may interfere with the detection and identification of analytes. A sufficient level of lipid removal was achieved by a combination of high-resolution gel permeation chromatography and Florisil fractionation. The data were acquired using GC-QTOF/MS in both electron ionization (EI) and electron capture negative ionization (ECNI) mode to maximize contaminant coverage. Acquired EI data sets were processed using a new highly-automated adjustable workflow. The ECNI data were manually processed and reviewed. The Mann-Kendall test was used to discover time trends and the Theil-Sen regression was used to estimate annual change for the features exhibiting significant time trends. Altogether, more than 300 tentatively identified contaminants were found to have significant time trends in samples studied. Significant decreases were found for many regulated chemicals, as could be expected, e.g., PCBs, DDT, and other organochlorine pesticides. Significant increases were observed for, e.g., small PAHs, heptaBDEs, and CECs. The CECs include, among others, a novel plasticizer tributyl acetylcitrate and two compounds used in polymer production trimethyl isocyanurate and 2-mercaptobenzothiazole, which have not previously been reported in biota.

Biodegradation of the Albendazole in the Water: Detection and Toxicity Assessment of the Novel Degradation Products

Josipa Papac[1]*, Sanja Radman*[2]*, Dijana Grgas*[1]*, Marija Vuković Domanovac*[1]*, Tibela Landeka Dragičević*[1] *and Danijela Ašperger*[1]*, (1)University of Zagreb, Croatia, (2)Mediterranean Institute for Life Sciences, Croatia*

For many years pharmaceuticals have been used for medical therapeutic purposes and therefore have been uncontrollably entered in the environment, through excretion, inappropriate disposal or because of inadequate waste water treatment plants. Recently they have been recognized as a potentially harmful pollutants, so accordingly that the term “emerging contaminants” have been formed. There is still no sufficient data to predict their behavior and the impact on the environment. That is reason for increasingly investing in the development of instruments with sufficiently low limits of detection and quantification, to determine the threshold limit value of pharmaceuticals in the environment. Many studies have pointed to a possible biodegradation of these compounds, so that can be set in systems for wastewater treatment. Moreover, degradation products may pose additional risk to the environment, therefore their detection and identification is necessary. Accordingly, it is important to investigate the toxicity of such compounds, to be sure what are the possible negative effects they have regarding the environment. The aim of this study was to optimize the conditions of biodegradation veterinary pharmaceutical anthelmintic albendazole in a batch reactor by monitoring the physical and chemical

parameters in water media. During such process, degradation products that appears are detected by Ultra Performance Liquid Chromatography coupled with the detector tandem Mass Spectrometry (UPLC-MS/MS), based on a quadrupole–time–of–flight technology (QqTOF). For the toxicity investigation of the samples, methodology of the zone inhibition was applied. For that puprose three isolated species characteristic for soil and water were chosen: *Aspergillus niger, Streptomyces rimosus* and *Bacillus sp.* In this study the biodegradation of albendazole in the water with activated sludge have been achieved whereby the impact of hydrolysis on that process could not be ignored. During the process three degradation products of the albendazole were detected and also their structures were proposed. Also, toxicity of this compounds and associated samples were assessed as low or non-toxic which makes this process non-harmful to the environment.

Chemical Characterization of Oil Spills in Aquatic Environments by Solid-Phase Microextraction Coupled With Gas Chromatography

Dr. Nasima Chorfa[1]*, Tahereh Boloori*[2]*, Benjamin de Jourdan*[2] *and Feiyue Wang*[1]*, (1)University of Manitoba, Canada, (2)Huntsman Marine Science Centre, Canada*

The risk assessment of a petroleum oil spill in aquatic environments is mainly driven by the chemical exposure of living organisms. This is typically done by monitoring chemical profiles in the weathered oil, and assessing its toxicity in the laboratory and under field conditions. Due to the complex chemical composition of oil, an efficient and easy-to-use analytical tool is needed. This work presents solid-phase microextraction (SPME) as a solvent-free, non-depletive technique to integrate sampling and concentration. The technique employs a polymer impregnated on a metal fiber, which is immersed in the headspace (HS) of or directly in an aqueous sample for the extraction of organic compounds. The SPME fiber is then injected in a gas chromatograph where the extract is thermally desorbed, identified and quantified. Here we show two applications of the SPME technique in oil spill research. The first application is on the determination of BTEX (i.e., benzene, toluene, ethylbenzene and xylenes) in the water accommodated fraction (WAF) and chemically enhanced WAF (CEWAF, by the addition of a Corexit dispersant) of oil-contaminated seawater. The co-polymer Divinylbenzene/Carbon Wide Range/Polydimethylsiloxane exhibits excellent BTEX selectivity and sub-ppb detection limit. Our HS-SPME analysis showed disparity in BTEX concentrations of CEWAF/WAF as the presence of the dispersant creates a microemulsion that reduces the interfacial tension between oil and seawater and increases the apparent BTEX solubility. The other application is the biomimetic extraction (BE-SPME) of bioavailable acid-extractable organics in oil contaminated waters. Our results show a 20-fold increase in bioavailable concentrations after acidification of oil sands process-affected water. It emphasizes the importance of acidification to convert naphthenate salts into naphthenic acids when assessing the bioavailability of polar compounds. The BE-SPME method has the potential to be a simple and effective technique to predict the acute aquatic toxicity of petroleum mixtures, and the resulting bioavailable concentrations could be used as a water quality guideline for oil spill response.

Contaminants of Emerging Concern Identified by Suspect and Non-target Screening in Marine Environment: A Scoping Review

Mr. Kelsey Ng[1]*, Nikiforos Alygizakis*[1]*, Nikolaos Thomaidis*[2] *and Jaroslav Slobodnik*[1]*, (1)Environmental Institute, Slovakia, (2)National and Kapodistrian University of Athens, Greece*

Numerous contaminants of emerging concern (CECs) are present in the marine environment and there is a knowledge gap to establish their occurrence profiles. Many CECs and their transformation products (TPs) are overlooked using the conventional target screening approach as reference standards are not always available. Suspect and non-target screening of CECs and their TPs has been made possible with the advances in gas and liquid chromatography coupled to high resolution mass spectrometry techniques. In this study a literature review has been conducted searching for “marine”, “non-target” and “high-resolution” in the title, abstract and keywords of published papers as of October 2021. In total, 175 articles were retrieved, of which 25 were considered relevant. References of these articles were also searched to avoid missing any important additional

study. The search revealed documented the presence of 1,586 CECs of anthropogenic origin detected in the marine environment samples in 23 countries from five continents. The study indicated that a continuous development of suspect and non-target screening allows for investigation of ever wider scope of CECs in the marine environment with a potential for use of the results in the regulatory framework. It has been found out that advances in data sharing platform and harmonization of analytical protocols are required for future regulatory applications. The use of available computational tools and data exchange platforms, such as the NORMAN Suspect List Exchange, NORMAN Substance Database and MassBank Europe was endorsed. An overview on the progress in harmonization and automation of NTS workflows has been provided.
Acknowledgements: Kelsey Ng was financially supported through Marie Skłodowska-Curie grant agreement No. 859891 (MSCA-ETN) for PhD fellowship.

Epidemiology-Based Wastewater Monitoring for Ecological Risk Assessment of Anti-Tuberculosis Drugs Mixture Effects

Wei-Yu Chen[1]*, Yi-Fang Chen*[1] *and Hsin-Mei Huang*[2]*, (1) Kaohsiung Medical University, Taiwan, (2) National University of Tainan, Taiwan*

Real-time monitor environmentally relevant anti-infective drugs and precise predictions for the mixture toxicity for anti-infective drugs are still limited. This study seeks to provide a practical methodology, epidemiology-based wastewater monitoring (EBWM), as early warning signals for assessing the mixture toxicity risks of anti-tuberculosis (TB) drugs in water environments. We incorporated a TB transmission dynamic model and per capita water uses to predict the environmentally relevant concentrations of anti-TB drugs, isoniazid (INH), rifampicin (RMP), and ethambutol (EMB) in Taiwan's high TB incidence area. Using our model, the predicted TB cases and environmentally relevant concentration of anti-TB drugs were decreasing to 2025. The EMB has the highest residues in water environments (INH: 0.010 μg L^{-1}; RMP: 0.024 μg L^{-1}; EMB: 0.136 μg L^{-1}). For mixture toxicity, we find the antagonistic effect among anti-TB drugs from daphnia (*Daphnia magna*) immobilization tests. There are no ecological risk concerns on the mixture toxicity of anti-TB drugs in water environments. The EBWM predicts environmentally anti-TB drugs in real-time that will be validated with the measured environmental residues of anti-TB drugs and removal efficiency of wastewater treatment plants from fieldwork data to enhance the predictive power in future work.

Fingerprinting Water Sources in Southeast Florida: Using Nontarget Analysis to Identify Water Sources and Contamination

Kassidy Troxell*, Brian Ng, Ingrid Ley and Piero Gardinali, Florida International University*

Water contamination is a growing source of concern in Southeast Florida due to diminished onsite treatment of wastewater as a consequence of sea-level rise and a limestone dominated geology providing a relatively unobstructed connection pathway to groundwater, canals, and Biscayne Bay. This study was designed to investigate an environmental system potentially influenced by compromised septic tanks and their effluents. Using a high-resolution mass spectrometry (HRMS) Q Exactive orbitrap Nontarget Analysis (NTA) method and a newly developed Compound Discoverer (CD) 3.1 prioritization workflow a fingerprint of the chemical space of each sample was created and assessed. Water samples were collected from areas in and around the Charles Deering Estate property in the Village of Palmetto Bay including samples from the freshwater environment (Everglades-based), urban impacted areas (septic tank driven), and coastal (Biscayne Bay) endmembers. It is imperative to assess how the fresh water from the Everglades changes traveling through the urban systems around Deering Estate, and in what condition the water is being discharged into the at-risk Biscayne Bay endmember. Different water sources (Everglades based, urban influenced, and Biscayne Bay) contained a unique combination of organic compounds tentatively identified by the data processing workflow providing value based on frequency of detection and environmental relevance. NTA was utilized in this study to tentatively identify source discriminating compounds throughout this system giving rise to new potential tracers and contaminates from various categories such as dyes, personal care products (PCPs), and pharmaceuticals. The proposed tracers associated with a septic tank source in this area are azelaic acid, decanophenone, methyl

violet, monoolein, metoprolol, and 1-stearoylglycerol. CD HRMS data presented as Kendrick mass defect plots and Principal Component Analysis (PCA) showed clear differences between the water sources throughout this system. To compliment the NTA, sucralose was targeted and utilized as a conservative tracer of septic tanks with 3,594 ± 94 ng/L of sucralose at the most septic influenced site with a declining gradient reaching the lowest concentrations in Biscayne Bay of 122 ± 94 ng/L. This approach has been applied to the Miami River in efforts to characterize the different water and contamination sources within the system that, like Deering Estate, discharge into Biscayne Bay.

Nontargeted Analysis of Polymer Additives in Five Microplastic Fiber Types

Anna Michelle Lewis[1]*, Lee Ferguson*[1]*, Melissa Chernick*[2] *and Abigail Joyce*[1]*, (1) Duke University, (2) Duke University Nicholas School of the Environment*

Microplastic fibers are the dominant synthetic particle type observed in common environmental samples, including ambient waters and gut content recovered from aquatic organisms. Most synthetic polymer-based textiles use a wide variety of chemical additives to enhance the functionality, water-resistance, and coloring of the material —such as plasticizers, per- and poly fluoroalkyl substances (PFAS), and disperse azobenzene dyes (azo dyes)— many of which are suspected endocrine disrupters and toxicants. We hypothesize that microplastic fibers may act as a delivery mechanism for chemical additives to aquatic environments and organisms that ingest or are otherwise exposed to the synthetic particles. To test this hypothesis and investigate what portion of extractable polymer additives are also leachable from the particles, clear polypropylene (PP) fibers and both clear and dyed polyester (PES) fibers were subject to ethyl acetate solvent extractions and 3-day leaching experiments in freshwater. Black and blue PES fibers were isolated from a water-repellant fleece and couch sleeve, respectively. The PP and green PES fibers were sourced from a Chinese fabric store. Clear PES was purchased from Sigma Aldrich. Polymer additives were analyzed by HPLC with high resolution Orbitrap MS/MS. Compound annotations were based on matches to spectral libraries and a custom, curated database of known organic polymer additives and dyes. Results of analysis showed extractable and readily leachable polymer additive complements unique for each fiber type. Based on library matching and/or computational mass spectrometry evidence, several known polymer additives were annotated, including multiple plasticizers (Dibutyl phthalate), PFAS (Perfluoro-1-butane sulfonic acid), antioxidants (Irganox 1010), and surfactants (PEGn). Annotated azo dyes include Disperse Blue 79:1 in blue PES extracts, Disperse Blue 373, and the chlorinated analog of Disperse Violet 93 in black PES fiber extracts, but none in the corresponding freshwater leachates. Results will be discussed in the context of the applied nontargeted analysis workflow and potential for microplastic additive release to the aquatic environment.

Quantifying Chemical Habitat of Aquatic Species Assemblages in the Willamette River, Oregon

Daphne Guo[1]*, Anna Beran*[1]*, Brooke Penaluna*[2]*, Francisco Pickens*[1] *and Gerrad Jones*[1]*, (1) Oregon State University, (2) USDA Forest Service, Pacific Northwest Research Station*

Fish biodiversity is imperiled worldwide, with many species known to be threatened, extirpated, or extinct. Restoration efforts for improving survival and fitness are often targeted at improving physical in-stream habitats; however, the chemical “habitats” of aquatic ecosystems, which affect physiological condition of fish, are virtually ignored during restoration efforts. Salmonids are sensitive to chemicals present in aquatic environments, and the vast majority of ecotoxicological studies conducted on salmonids focus on the toxic effects of relatively small numbers (n = 1-20) of specific target chemicals (e.g., pesticides). With 10’s of thousands of chemical features present in aquatic environments, targeted analyses ignore >99% of the chemical data present in a water sample. First, many toxicants are novel and are yet to be described by science (e.g., degradation products); therefore, scientists are not currently screening for relevant compounds. Second, the cumulative effects of multiple chemicals can elicit adverse effects in fish even if the concentration of any individual toxicant is present a low/trace levels; therefore, scientists may inadvertently overlook important drivers of fish declines. We used non-target chemicals to identify the chemical features that are diagnostic of the presence of different fishes. We electrofished >40 sections of the Willamette River, Oregon. Previously, using a

previously developed chemical fingerprinting workflow to predict the presence/absence of different pollution sources, we selected the chemical features that were most diagnostic of the presence/absence of different fish species. While the workflow's accuracy was near perfect when predicting the presence/absence of pollution sources, the accuracy was considerably lower when predicting the presence/absence of different fish species. This is likely due to higher false negative rates (i.e., failing to detect a true fish presence) when sampling organisms. While true negatives were correctly classified with near perfect accuracy, false negatives in the model were high. Approaches that utilize 95% convex hulls could be used to help identify false negatives and improve the performance of the model. While methodological limitations need to be addressed, our initial results suggest that chemical fingerprinting could be useful for quantifying the chemical conditions that are unsuitable for supporting aquatic community assemblages in addition to source apportionment and pollution source tracking.

Uplc-Tims-Hrms Combined with a Sophisticated Annotation Workflow for Comprehensive Xeno- and Endo-Metabolome Coverage of Zebrafish Exposed to Xenobiotics

Dimitrios E. Damalas[1]*, Eleni Panagopoulou*[1]*, Adamantia Agalou*[2]*, Dimitris Beis*[2] *and Nikolaos Thomaidis*[1]*, (1) National and Kapodistrian University of Athens, Greece, (2) Biomedical Research Foundation Academy of Athens, Greece*

Given the large number of xenobiotics, there is an important gap in the literature concerning their adverse effects on aquatic organisms. The impact of xenobiotics in the aquatic ecosystem is evaluated in more depth when the whole xeno-metabolome (xenobiotics and their biotransformation products (bio-TPs)) and endo-metabolome (metabolites and lipids) of aquatic organisms is studied. HRMS-based workflows appear to be a powerful tool for the identification of bio-TPs and metabolites/lipids. Beyond the cases that identification fails because of reduced ionization efficiency, false negatives arise from the inadequate separation. Despite HRMS's high applicability and accuracy, separation of isomers/isobars is not always possible, as they may pose identical chromatographic and spectral profiles. Thus, unique analytical challenges still exist concerning the identification task. Additional "dimensions" of separation are required to provide additional evidence and reliable identifications. Combining complementary chromatographic modes with HRMS seems a powerful alternative tool for the comprehensive and high-throughput identification of unknowns. Lately, Trapped Ion Mobility Spectrometry (TIMS) has been proved an up-and-coming technology, separating molecules based on their 3-D size and charge in the gas phase. Developing comprehensive annotation workflows that utilize the analytical evidence to their maximum is also imperative. The objective of this study is to highlight a high-end analytical platform (LC-TIMS-HRMS) that combines multiple dimensions of separation with HRMS, in order to provide extensive experimental evidence for the identification of bio-TPs and endogenous metabolites/lipids. In addition, a data treatment workflow was developed, consisting of suspect and non-target screening approaches combining different annotation tools. Overall, the aim was to highlight a holistic approach for comprehensive xeno- and endo-metabolome coverage to facilitate toxicity assessment of aquatic organisms exposed to xenobiotics. ZFE exposed to xenobiotics were used as case studies. The developed workflow encompasses a wide range of tools to identify bio-TPs of xenobiotics, endogenous metabolites and lipids in the same dataset. Examples highlighting the added value of TIMS in the identification of bio-TPs will be presented. Finally, the annotation of several lipids will be showcased, leading to the observation of lipid accumulation as a toxic response to triclosan exposure.

Using Nontarget LCMS and In Silico Toxicological Approaches for the Detection and Prioritization of Contaminant Mixtures in a Regional Irrigation Canal

Chloe Fender*, Stephen Good and Manuel Garcia-Jaramillo, Oregon State University*

Water quality and quantity is impacted by land management practices, industrial development, and climate change. Water scarcity, combined with an increase in water contamination, has diminished our access to safe water for food production, manufacturing, and ecosystems health. Contaminants of concern that are usually found include pesticides, pharmaceuticals, personal care products, per-and polyfluoroalkyl substances, micro

and nanoplastics, cyanotoxins associated with algal blooms, and infectious agents such as *Escherichia coli*and norovirus. The potential toxic effects derived from the exposure to these contaminant mixtures on ecosystems and human health, as well as their effects on crops production, require further investigation. This study sought to monitor changes in water composition and contaminants along an irrigation canal, and their association with evaporation trends, in an arid region of central Oregon. Water collection sites were evenly spaced along the 110 km stretch. Water stable isotopes ($^{2}H/^{1}H$ and $^{18}O/^{16}O$) were measured using cavity ring-down spectroscopy. The evaporation over inflow ratio (E/I) was estimated for each site. Chemical analyses were performed by ultra-high performance liquid chromatography coupled to high resolution mass spectrometry (UPLC-HRMS). Acquired data was processed using MS Dial, an open-source software. In addition, collected water was tested for toxicity endpoints in the zebrafish model. The E/I was linear ($R^2 = 0.81$) along the irrigation canal and increased after a reservoir used as a water storage facility. Significant behavioral changes were observed in larval responses to light stimulus after the exposure to irrigation canal water. Approximately 300 chemicals were annotated using high-resolution MS libraries. Overall, the abundance of the annotated compounds decreased from the beginning to the end of the canal, as expected due to degradation processes. Annotated compounds were investigated for potential toxicity using the EPA CompTox chemical database. Based on their bioactivity and changes in chemical abundance over the length of the canal, a mixture of 5 contaminants was prioritized for further *in vitro* and *in vivo* bioassays. Investigating chemical distribution trends in irrigation canals, and their correlation with the location and characteristics of water supply networks, have the potential to help us to predict the behavior of relevant contaminants, and their persistence in the environment.

Applications to Human Exposure and Effects Assessment

Keynote: Establishing an Analytical Base to Support Large-Scale Exposome Epidemiology

***Douglas Walker**, Icahn School of Medicine at Mount Sinai*

Over a lifetime, humans experience thousands of chemical exposures from multiple sources. A more complete estimate of environmental exposures across the lifespan would be a transformative research initiative. The use of high-resolution, mass spectrometry (HRMS) provides a key platform for assessing the exposome and provides measures of thousands of chemical signals in a single human sample; however, application of internal exposome profiling in large populations has been limited due to challenges in sample throughput, instrument robustness, and data handling methods. To support large-scale population research for exposome epidemiology, we have established a high-throughput untargeted HRMS analytical framework combining parallel analysis by liquid (LC) and gas (GC) chromatography to enable robust analysis of up to 20,000 blood samples per year. Sample preparation is achieved using low-cost and open-source automated liquid handlers allowing parallel processing of extracts for LC and GC analysis from a single blood aliquot of 150µL, eliminating operator effects and allowing daily preparation of 96 samples in under 3.5 hours. Exposome profiles are measured using five different HRMS analytical configurations that include reverse phase and HILIC chromatography with both positive and negative ionization, and GC-HRMS to provide detection of 60,000-100,000 chemical signals. To support large-scale data extraction and alignment, we implement batchwise data extraction algorithms that allow alignment of batches, rather than samples, to improve peak detection and robustness. Resources for annotation include a standard library of over 7,000 compounds covering a wide range of environmental, drug and endogenous metabolites. The resulting exposome profiles are being assembled within a framework built upon SQL databases, providing a cumulative resource for assembling exposome-disease atlases and laying the foundation for the Human Exposome Project. Development of a robust and scalable analytical framework for large-scale exposome epidemiology is a critical first step to provide a robust foundation for exposome research and to facilitate development of a knowledge base of environmental chemicals, their products, distributions, and associated effects.

Integrative Exposomic, Transcriptomic, Epigenomic Analyses of Human Placental Tissues Links Understudied Chemicals to Preeclampsia

Alex Chao[1], *Jarod Grossman*[2], *Celeste Carberry*[3], *Yunjia Lai*[4], *Antony Williams*[1], *Jeffrey Minucci*[1], *Steven Purucker*[1], *John Szilagyi*[3], *Kun Lu*[3], *Kim Boggess*[3], *Rebecca Fry*[3], *Jon Sobus*[1] *and Julia Rager*[3], *(1) U.S. Environmental Protection Agency, (2) Agilent Technologies, Inc., (3) University of North Carolina at Chapel Hill, (4) Columbia University*

Pregnancy is a time of heightened susceptibility to environmental exposures. Biological perturbations induced by the presence of xenobiotic chemicals can result in adverse health effects for both the mother and fetus. Traditional exposure monitoring for fetal exposure has been performed with targeted methods on a limited set of chemicals. More recently, non-targeted analysis (NTA) has been applied to maternal and cord blood as surrogate matrices for fetal exposure. With limited knowledge on the fetal exposome, we are constrained in our understanding of potential drivers of developmental and prenatal diseases. Preeclampsia represents one such disease that affects up to 8% of pregnancies worldwide and can result in complications including death for both the mother and fetus. While evidence suggests that environmental exposures are related to this disease, the exact causes remain unknown. We sought to better characterize and understand potential sources of preeclampsia through analyses of placental tissue. Specifically, we focused on the following analyses of placental tissue (from both preeclamptic and normotensive patients): (1) high-resolution mass spectrometry-based NTA to identify xenobiotic chemicals present, (2) transcriptomic and epigenomic analyses to characterize biological signatures associated with preeclampsia, and (3) integration of both exposomic and biological -omic results, to elucidate chemical-biological interactions. NTA results revealed 183 molecular features with significantly higher or lower abundance in placental samples of preeclamptic patients, with 53 features identified (corresponding to 40 unique chemicals, categorized as “exogenous drugs”, “exogenous non-drugs”, or “endogenous metabolites”) through chemical standards, fragmentation spectra matching, and chemical metadata. Unsupervised clustering grouped molecular features into eight groups. Of the identified features, endogenous metabolites (such as acyl carnitines and amino acids) were found at the highest abundance and had strongest relation to biological -omic signatures. This integrative analysis of exposomic, transcriptomic and epigenomic results is a novel approach to not only identify potential chemical risk factors for preeclampsia, but to lay the groundwork for understanding underlying linkages between gene signatures, human metabolites, and physiological effects of preeclampsia. *Abstract does not necessarily reflect Agency policy.*

Effect-Directed Analysis of Thyroid Hormone System Disrupting Compounds in Human Cord Blood Samples

Jeroen Meijer[1], *Jelle Vlaanderen*[2], *Roel Vermeulen*[2], *Timo Hamers*[3] *and Marja Lamoree*[3], *(1) Vrije Universiteit, Netherlands, (2) Utrecht University, Netherlands, (3) VU University Amsterdam, Institute for Environmental Studies (IVM), Netherlands*

Thyroid hormone system disruption is associated with adverse neurodevelopmental effects in humans, which might result from the exposure to xenobiotic substances that compete with the thyroid hormone T_4 for the binding site on the transporter protein transthyretin (TTR). There are many known compounds that show affinity for the binding site of TTR including PCBs, PFASs and PBDEs. However, it is possible that humans are also exposed to unknown compounds with TTR-binding potential. Therefore, it is important to identify and subsequently monitor these currently unknown TTR-binding compounds. An Effect-Directed Analysis (EDA) workflow was developed that incorporates toxicological testing and chemical identification of fractionated serum extracts, which resulted in the successful identification of two unknown TTR-binding compounds in fetal calf serum (fipronil sulfone and bisphenol S) in previous research. The aim of the present study, as part of the HBM4EU project, was to apply this newly developed EDA workflow and identify unknown TTR-binding compounds in human samples. Six human cord blood samples were extracted and tested in the T_4-FITC TTR-binding assay, which shows a perturbed fluorescent signal in the presence of TTR-binding compounds. Subsequently, the extracts were chemically analyzed using LC-QTOF MS(/MS) and fractionated into 80 fractions using the FractioMate, each encompassing 13.5 seconds of the chromatographic run. Each fraction was then tested in the TTR-binding assay. The chemical features present in the chromatographic window of the active fractions were prioritized for identification, reducing the complexity of the samples. All the

unfractionated extracts showed TTR-binding activity higher than what would be expected if the activity was caused by native T_4 alone. However, after fractionation, activity was only found in two separate fractions from two samples. Chemical analysis (in both positive and negative mode) resulted in the detection of over 14,000 chemical features, of which 317 features were present in the retention time window of the most active fraction. Suspect screening using the CECscreen database (includes >70,000 chemicals of emerging concern) resulted in the annotation of 142 compounds that might be responsible for the measured activity. Future work will focus on further prioritization and elucidation of those suspected candidates in order to find the chemical(s) responsible for the TTR-binding activity.

Characterization of Micro- and Nanoplastic Exposure in Human Tissues Using an Exposomic Framework

Anna Robuck[1]*, Zoe Coates Fuentes*[1]*, Brooklynn McNeil*[1]*, Mariona Bustamante*[2]*, Virissa Lenters*[3]*, Martine Vrijheid*[2]*, Tim Nawrot*[4]*, Roel Vermeulen*[5] *and Douglas Walker*[1]*, (1) Icahn School of Medicine at Mount Sinai, (2) Barcelona Institute for Global Health, Spain, (3) Universitair Medisch Centrum Utrecht, Netherlands, (4) Hasselt University, Belgium, (5) University of Utrecht, Netherlands*

Despite rapid advances in our understanding of micro-and nanoplastics (MNPs) in environmental matrices, significant data gaps hinder our understanding of MNP exposure in humans. Studies involving small sample sets investigating specific tissues have thus far established the presence of MNPs in human feces and placenta. Further work is required to more fully understand MNP exposure across a range of human tissues to deduce associated health outcomes. However, this goal is hindered by analytical challenges associated with the heterogenous and complex nature of MNPs. Most environmental studies of MNPs focus on a few candidate polymers or chemicals that capture only a small fraction of MNP complexity. To overcome these and other limitations we have developed an untargeted HRMS analytical framework to characterize known and unknown MNPs, their monomers and additives, and related exposure biomarkers and biological response. This analytical framework combines 1) pyrolysis (GC) gas-chromatography with HRMS (Pyr-HRMS) to characterize MNP levels, 2) alkaline-assisted hydrolysis with liquid chromatography HRMS (AAH-HRMS) to measure particle monomers and additives, and 3) untargeted liquid chromatography with HRMS (LC-HRMS) to characterize the metabolome for the presence of MNP constituents and additives and biological response profiles. This framework was applied to placenta samples, virgin plastics, and weathered plastics to validate our approach. Pyr-HRMS detected a range of common MNP types in placenta tissue, including PS, PE and PMMA which were detected at a range of concentrations from 10-2650 ng/g. LC-HRMS revealed that free and particle-bound small molecule profiles were significantly different between virgin and weathered particles. Results to date validate the application of this framework to contemporaneously characterize polymers, free monomers, additives, and biomarkers of disease in human tissues; this framework provides an important toolset to expand our understanding of MNP exposure in humans and other biological receptors.

Evaluation of the Human Exposure to a Broad Spectrum of Organic Chemicals and the Potential Use of Sewage Sludge to Prioritize Hazardous Substances

Rubén Gil-Solsona[1]*, Maria-Christina Nika*[2]*, Nikiforos Alygizakis*[3]*, Mariona Bustamante*[4]*, Cristina Villanueva*[5]*, Esteban Restrepo*[1]*, Maria Foraster*[5]*, Maria Dolores Gómez-Roig*[6]*, Elisa Llurba-Olive*[7]*, Jordi Sunyer*[8]*, Payam Dadvand*[5]*, Nikolaos Thomaidis*[2] *and Pablo Gago Ferrero*[1]*, (1) Institute of Environmental Assessment and Water Research (IDAEA), Spanish Research Council (CSIC), Spain, (2) National and Kapodistrian University of Athens, Greece, (3) Environmental Institute, Greece, (4) Barcelona Institute for Global Health, Spain, (5)ISGlobal-CREAL, Spain, (6) Hospital Sant Joan de Déu, Spain, (7) Universitat Autònoma de Barcelona, Spain, (8) Centre for Research in Environmental Epidemiology, Spain*

Chemicals are part of our daily lives and we are exposed to numerous chemicals through multiple pathways. Relevant scientific evidence contributing to the regulation of hazardous chemicals require a holistic approach to assess simultaneous exposure to multiple compounds. Currently, the main way to obtain data on the exposure to organic chemicals is through human biomonitoring, that requires very complex and costly sampling campaigns.

Finding efficient proxies to predict the risk of chemical exposure in humans is an urgent need to cover large areas and populations at a reasonable cost. We conducted an exploratory study to characterize the human chemical exposome in maternal blood and placenta samples of a population-based birth cohort in Barcelona (2018-2021). Ultimate HRMS-based approaches were applied including wide-scope target, suspect (for >2300 and >10000 chemicals, respectively) and non-target screening. Forty-two chemicals were identified including pesticides, personal care products or industrial compounds, among others, in the range of ng/mL and ng/g. In parallel, sewage sludge from the wastewater treatment plants serving the residence areas of the studied population were also screened, showing correlations with the type and concentrations of chemicals found in humans. Our findings were suggestive for the potential use of sewage sludge as a proxy of the human exposure and its application in early-warning systems to prevent chemical threats.

Applications to Human Exposure and Effects Assessment (Poster)

Chemical Exposure Identification in Relation to Gestational Diabetes Diagnosis Using Non-Targeted Analysis

***Garret Bland**, University of California*

Gestational diabetes mellitus (GDM) is a significant health issue because of its increasing prevalence as a risk factor for several complications in pregnant populations and their offspring. Chemical exposure can increase the risk of adverse maternal health outcomes, but there is a lack of research covering the wide range of chemical exposures among pregnant populations and relationships to GDM. The objective of this study is to use non-targeted analysis to identify and evaluate the range of chemical exposures during pregnancy and their relationship to GDM. Our hypothesis is that specific endocrine disrupting chemicals could influence metabolite levels that regulate glucose levels which may lead to increased GDM risk. We used non-targeted analysis to match and identify exogenous chemicals and endogenous metabolites in 233 pairs (n = 466) of matched maternal and cord blood samples using High-Resolution Mass Spectrometry (HRMS) data measured by LC-QTOF/MS. We used a comprehensive data treatment workflow to match the MS1 data with chemicals from EPA's DSSTox Database (~860,000 reported chemicals) and match MS2 data to the MS-Dial database (~26,000 unique compounds). We evaluated the relationship between molecular chemical and metabolite features and blood glucose levels, and GDM diagnosis, adjusting for covariates (body mass index and maternal age). Fifteen participants had a definitive GDM diagnosis, and 16 participants were marginal with higher glucose levels (n =31). We identified 685 chemical features. Eighty-three chemical features correlated with glucose measurements ($p < 0.05$) and had distinct clustering that separate participants based on glucose levels. Twenty-four of these had a chemical confidence level of 3 or better with 14 being exogenous chemicals and the rest endogenous metabolites. These select exogenous chemicals include perfluorinated and organosilicon compounds. We found multiple exogenous chemicals in this sample of pregnant people (and in their newborn offspring cord blood), some of which are known to have toxic effects including endocrine-disruption. This study is one of the first to match chemical exposure with GDM risk in the prenatal exposome and advances the application of NTA analysis to better understand relationships between chemical exposure and adverse pregnancy-related health outcomes.

Addressing Public Health Concerns of Yurok Tribe Using Non-Target Analysis

***Brittany Saleeby**, University of California*

The objective of current research activities is to apply non-targeted workflows for the analysis of high-resolution mass spectrometry on affected environmental samples, with prioritizing compounds identified by analysis via in-vitro assay for biological activity (bioassay). In collaboration with UC-Davis, the Yurok Tribe Environmental Program (YTEP) collected sediment suspected of contamination from previous logging and agricultural industry activity for non-target analysis of compounds linked to the community's adverse health issues. Scientist from YTEP collected sediment from affected regions between 2018-2020. At UC-Davis, the sediments were extracted, analyzed on multiple MS platforms, and vetted for QA/QC information derived from

target analytes and internal standards. All sample extracts and blanks that were prepared for analysis were also submitted to collaborators for bioassay results. Complete sample datasets have been compiled and optimized on MS-Dial application, and the resulting alignment compound features have been exported for statistical analysis. From the exported MS-Dial alignment data and bioassay data from collaborators, Spearman's rank correlation coefficients were calculated for ordinal correlation between bioassay activity and ion intensity. Some of these features are automatically identified by spectral libraries, but others will require identification via non-target molecular formula generating programs like MS-Finder and SIRIUS. Unlike traditional targeted and suspect analytical workflows where compounds of concern are known, the non-targeted analytical workflow supports compound discovery by broader and systematic methods of detecting unknown analytes. A benefit but also a challenge to this methodology is that non-target workflows result in a list of molecular features typically in the order of thousands even after filtering and post processing steps. Due to the immense number of compounds that could be detected in any given chemical space, these biological indicators are necessary to guide action for regulation of these contaminants. Ultimately, by using Spearman's rank correlation coefficients to highlight molecular features of interest, the list can be reduced to hundreds, rather than thousands, of compounds.

Characterizing Azobenzene Disperse Dyes in House Dust via Non-Targeted Mass Spectrometry Analysis
Kirsten Overdahl[1]*, Chris Kassotis*[2]*, Gordon Getzinger*[1]*, Allison Phillips*[3]*, Kate Hoffman*[1]*, Heather M. Stapleton*[1] *and Lee Ferguson*[1]*, (1) Duke University, (2) Wayne State University, (3) U.S. Environmental Protection Agency*
Disperse dyes are a class of substituted anthraquinone- or azobenzene- based dyes used to color synthetic fabrics such as polyester, nylon, and acrylic. Many of these dyes are chlorinated or brominated, and account for roughly 70% of the 9.9 million tons of industrial dye colorant used annually. Azobenzene dyes are well-characterized as electrophilic mutagens and contact allergens in clothing, but little is known about occurrences of these dyes in the indoor environment, nor their health implications following exposure. Furthermore, no comprehensive library of disperse dyes exists to date. Here, we report on identifications of twenty azobenzene compounds in house dust samples collected from 190 homes in the Toddlers Exposure to SVOCs in Indoor Environments (TESIE) study in central North Carolina. House dust samples were collected in 2014-2016 by research investigators from homes with small children. We used a data-dependent, suspect-screening analytical strategy with HPLC-ESI-HRMS/MS analysis to identify azobenzene-class disperse dyes in organic extracts of house dust. A substructure search in CAS was used to assemble a suspect list of compounds containing p-aminoazobenzene as the core structure of disperse dyes. Extracted house dust samples were then analyzed using an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific) in positive and negative modes with data-dependent MS/MS selection. Tentative identifications were assigned using a weight-of-evidence non-targeted analysis approach with multiple open-source, *in silico* fragmentation algorithms combined with curated data source, scientific reference, and patent information for each candidate compound. We identified 20 azobenzene compounds in dust that are present at high relative abundances; of these, we quantified 12 azo disperse dyes for which we had reference standards and quantified at least one dye in every house dust sample. Detection frequencies ranged from 11% to 89%; of the dyes that were detected in at least 50% of the samples, geometric mean levels ranged from 32.4 to 360 ng/g. This study indicates that azo disperse dyes and related compounds are ubiquitous in the indoor environment. Future studies are needed to quantify additional dyes in dust, particularly those identified here via suspect screening, and to examine exposure pathways of dyes in the indoor environment where children are concerned.

Chemical Characterization of Recycled Consumer Products Using Suspect Screening Analysis
Charles Lowe[1]*, Katherine Phillips*[1]*, Kristin Favela*[2]*, Alice Yau*[2]*, John Wambaugh*[1]*, Jon Sobus*[1]*, Antony Williams*[1]*, Ashley Pfirrman*[3] *and Kristin Isaacs*[1]*, (1) U.S. Environmental Protection Agency, (2) Southwest Research Institute, (3) Oak Ridge Associated Universities*
Recycled materials are found in many consumer products as part of a circular economy; however, the chemical content of recycled products is generally uncharacterized. A suspect screening analysis using two-dimensional

gas chromatography time-of-flight mass spectrometry (GC × GC-TOFMS) was applied to 210 products (154 recycled, 56 virgin) across seven categories. Chemicals in products were tentatively identified using a standard spectral library or confirmed using chemical standards. A total of 918 probable chemical structures identified (112 of which were confirmed) in recycled materials versus 587 (110 confirmed) in virgin materials. Identified chemicals were characterized in terms of their functional use and structural class. Recycled paper products and construction materials contained greater numbers of chemicals than virgin products; 733 identified chemicals had greater occurrence in recycled compared to virgin materials. Products made from recycled materials contained greater numbers of fragrances, flame retardants, solvents, biocides, and dyes. The results were clustered to identify groups of chemicals potentially associated with unique chemical sources, and identified chemicals were prioritized for further study using high-throughput hazard and exposure information. While occurrence is not necessarily indicative of risk, these results can be used to inform the expansion of existing models or identify exposure pathways currently neglected in exposure assessments. *This abstract does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.*

Describing the Chemical Component of the Exposome in Urban Environments of Chile Using Two-Dimensional Gas Chromatography

Carlos Manzano[1]*, Carolina Molina*[1]*, Laila Hamzai*[2]*, William Richardot*[2]*, Penelope Quintana*[2]*, Nathan Dodder*[2] *and Eunha Hoh*[2]*, (1) Universidad de Chile, (2) San Diego State University*

The exposome is defined as the composite of every contaminant exposure to which an individual is subjected from conception to death. Thus, it contains an individual description of the environmental pollutants that surround us and could be thought of as the equivalent of the human genome or microbiome. This project aimed to explore part of the exposome of people living in urban environments of Chile, focusing on environmental organic contaminants. Samples were collected using personal passive samplers (silicone wristbands) at different urban centers in Chile (in the northern, central, and southern regions). The passive samplers were worn by a group of adults in each city for a full work week (Monday to Friday) in 2 seasons of the same year in each urban area and later extracted using minimal sample cleanup to preserve the original sample complexity. The resulting extracts were concentrated and analyzed directly by GC×GC/ToF-MS. The results were processed using the Statistical Compare tool in ChromaTOF, with a minimal S/N ratio of 50, and compounds identities were assigned for similarity scores greater than 750 against the NIST library. More than 800 chromatographic features were detected in each silicone wristband including personal care products, polycyclic aromatic hydrocarbons, pesticides, and insecticides, among others. The resulting compounds were classified and grouped based on their uses and sources. We present groupings of these contaminants by urban area and hypothesize these groups can reflect a local signature for organic environmental contaminant exposure.

Development of a Semi-Automated Quadruplex Mass Spectrometry Analytical Framework for Measuring the Human Exposome

Brismar Pinto-Pacheco*, Kathryn Dunn, Xiangping Lin, Brooklynn McNeil, Anna Robuck, Zoe Coates Fuentes and Douglas Walker, Icahn School of Medicine at Mount Sinai*

In the past decade, application of genome-wide association studies suggest that heritability alone is a poor predictor for most common diseases, and 80−85% of disease risk is linked to environment and its interaction with the genome. Most of these environmental exposures are largely uncharacterized and have minimal or no evaluation for health effects or risk assessment. The use of untargeted high-resolution mass spectrometry combined with advanced data extraction algorithms provide a key measurement of the spectrum of chemicals in a variety of biologic samples; however, most of these assays in human studies are optimized for detection of endogenous metabolite and lack the sensitivity for low level exposure biomarkers. To improve detection of environmental chemicals and biological response measures in blood, we have developed an Orbitrap-based untargeted analytical workflow that combines dual chromatography and ionization polarity for enhancing chemical coverage of the exposome. This quadruplex method includes analysis using two reverse-phase (RPLC) columns and two hydrophilic interaction liquid chromatography (HILIC) columns in which the positive and

negative mode are run in parallel. Sample preparation has been automated using low-cost, open-source liquid handlers, enabling rapid scale-up and translation of sample preparation methods across multiple laboratories. The combination of both analyses (30-minute runtime per sample) enables the systematic analysis of both lipophilic and hydrophilic compounds. Untargeted profiles for over 3500 samples have been analyzed to date and show detection of over 40,000 unique features; annotation suggests a wide range of environmental chemicals and endogenous metabolites are detected. Including internal standards for common exposures, including nicotine-related metabolites and perfluoroalkyl substances (PFAS) show quantitative measurements are possible. Comparison of untargeted PFAS measurements to concentrations determined using gold-standard targeted methods showed high agreement (correlation >0.95), and participation in international proficiency testing demonstrates accuracy is within specifications for targeted biomonitoring programs. Continued application of these methods to human populations will provide new insight into exposome chemical burdens across populations and provides a strategy for exposome epidemiology.

Development of Non-Targeted Analysis Methods on Gc-Hrms for Silicone Wristbands Used as Passive Samplers

Danielle Blum[1]*, Scott Clifton*[2]*, Seth Newton*[2] *and Elin M. Ulrich*[2]*, (1) Brown University, (2) U.S. Environmental Protection Agency*

Rising industrial rates has caused concern regarding personal exposure and negative impacts on human health and the environment. Past atmospheric monitoring efforts have utilized stationary sampling systems which can be limited in number/location and could over or underestimate exposure to the population. Silicone wristbands have potential for capturing a wide range of chemicals to which humans are exposed. Targeted methods are limited by unique analytical methods and may be unable to identify unknowns. Non-target analysis (NTA) methods are beneficial for identifying emerging contaminants and more rapidly identifying known pollutants. NTA methods have been used to broaden the understanding of personal exposure by qualitatively identifying emerging contaminants and estimating concentrations. Gas chromatography (GC) methods can investigate broad swaths of chemical space and are somewhat orthogonal to more commonly used liquid chromatography (LC). A standardized and tested procedure analyzing silicone wristbands on GC-HRMS has not been established at EPA. In order to expand NTA in-house capabilities, method development was undertaken, with the intent to enable application of the method to the EPA's Non-Targeted Analysis Collaborative Trial wristbands and real-world samples. Multiple methods for cleaning and extraction of the wristbands were tested for capture of the most analytes with the least amount of interference. Two cleaning methods were tested in conjunction with four extraction techniques. Isotopically labeled standards were used to evaluate recovery and instrument performance, and blanks were included to evaluate ubiquitous signals. Sample data was analyzed using Xcalibur and Compound Discover. The optimal cleaning method was chosen by lowest number of peaks produced, while the extraction method was chosen by highest number of peaks and best recovery rate of isotopically labeled standards that vary in mass and boiling point. This optimized procedure was applied to silicone wristbands deployed in various environments. Unknowns will be identified in accordance with Schymanski et al. Identification by NTA techniques using GC-HRMS data will allow greater confidence and breadth for identification of emerging contaminants and unknown species. This abstract does not necessarily reflect agency policy.

Discovering Novel Chemicals in Drinking Water: Point-of-Use Filters with Non-Targeted Analysis

Seth Newton[1]*, Hannah Liberatore*[1] *and John Sloop*[2]*, (1) U.S. Environmental Protection Agency, (2) Oak Ridge Institute for Science and Education*

A direct link has been established between contaminated drinking water and negative public health outcomes such as increased rates of gastrointestinal illnesses, nervous system or reproductive effects, and chronic diseases such as cancer. However, hundreds to thousands of unidentified, trace contaminants can be observed in typical drinking water in the United States, which has relatively stringent standards for known drinking water contaminants. The consequences of widespread, chronic exposure to trace contaminants on public health cannot

be evaluated until the contaminants are identified. Non-targeted analysis (NTA) has become one of the most powerful tools available to identify unknown contaminants in environmental matrices. However, higher concentrations are often needed for correct identifications compared to targeted methods. Previous work has shown that carbon-based point-of-use filters can capture a wide array of contaminants from large volumes of drinking water (~100 gallon or ~380 liters) and can serve as excellent preconcentrating sampling devices for NTA studies. To advance previous work, a method has been developed to include both gas chromatography (GC) and liquid chromatography (LC) NTA methods. Spiking experiments were performed to evaluate extraction recovery of spiked compounds from the filters that cover a wide range of vapor pressures and octanol-water partition coefficients (K_{ow}). Finally, a limited number of samples have been collected via mail from large drinking water systems around the country such as in New York City and Dekalb County, Georgia. MS1 LC data was matched to MS-Ready formulas and metadata from the CompTox Chemicals Dashboard (~883,000 chemicals) while MS2 LC data was matched to spectral libraries and *in silico* predicted spectra. GC-EI data were matched against NIST and Wiley spectral libraries with CI data serving as further confirmation of their identities. Results show the presence of a large number of compounds not typically monitored in drinking water studies, including current-use pesticides, fungicides used in consumer products, and industrial chemicals. The method developed in this study is ready for deployment on a larger scale, and the results of the pilot samples demonstrate a national drinking water study aimed at discovering emerging contaminants is warranted. Abstract does not necessarily reflect Agency policy.

Evaluating Lipidomic Alterations in Exosomes to Explore Human Health Differences

Rebecca Beres[1]*, Sandip Patel*[2]*, Jacob Rose*[2]*, Judith Campisi*[2]*, Birgit Schilling*[2] *and Erin Baker*[1]*, (1) North Carolina State University, (2) Buck Institute for Age Research*

Exosomes are small, lipid-bilayer enclosed, cell-derived nanoparticles that play a vital role in cell-to-cell signaling. To date, many molecular changes occurring in exosomes remain largely unexplored, however, due to factors such as isolation difficulties even though they could lead to potential biomarkers with less molecular interferences than blood or tissue samples. In this study, exosomes from both plasma and specific tissues were isolated with 2 different methods to understand the optimal approach for lipid evaluations. Specifically, in an initial study, exosomes from plasma samples for a cohort having different age ranges were isolated with two different extraction schemes; size-exclusion chromatography (SEC) which is commonly employed in proteomics, and then an exosome isolation kit that contains CD9,CD63, and CD81 antibodies. Following the isolations, the lipids were extracted from the exosomes and the samples were analyzed utilizing Liquid Chromatography, Ion Mobility Spectrometry, and Mass Spectrometry (LC-IMS-MS) separations. Using a Skyline library containing 765 lipids, 400 lipids were identified. Upon using statistical measures, such as t-tests to compare exosomes from adults ranging from 22-26 years old and 65-74 years old, triglycerides were found to be highly downregulated in the samples where SEC was used, while sphingomyelins and phosphatidylcholines were downregulated in the samples where the isolation incorporating the antibody method. Interestingly, more statistically significant lipids such as sphingomyelins were found in samples where the exosomes were isolated using the antibody kit. We think that larger lipid vesicles are co-isolated and interfere with exosome lipid content in the first isolation method (SEC). Upon knowing about that interference, the second isolation method (antibodies) was utilized in specific comparison that will be illustrated.

Non-Target Screening of Chemicals Migrating from Plastic Food Packaging by Gas and Liquid Chromatography with High Resolution Mass Spectrometry

Raegyn Taylor *and Yelena Sapozhnikova, U.S. Department of Agriculture*

Plastic food packaging is used to protect foodstuff from spoiling, damage, and contamination, but components of the packaging can transfer onto food in a process termed migration. While migration studies of known food packaging components (intentionally added substances; IAS) are required and relatively straightforward, identification of non-intentionally added substances (NIAS; unknowns) is imperative to better characterizing food safety. To this aim, migration was investigated across 24 unique plastic food packaging products including

storage bags, vacuum bags, plastic wrap, and meat trays. Migration studies were carried out over ten days with a fatty food simulant according to the US FDA guidelines. Gas and liquid chromatography separation systems coupled with Orbitrap mass analyzers were used for comprehensive non-target screening of migrants. Compound Discoverer software was used for data processing where tentative identifications of features were made by searching commercial databases (e.g., NIST, Extractables and Leachables), retention time indices, and fragmentation pattern matching. Retention time prediction models created using isotopically labeled standards and log K_{ow} values were used to further confirm (or not) compound annotations. Most migrants were classified as IAS and included plasticizers, anti-statics, azodyes, monomers, and slip agents. The NIAS tentatively identified with Chemspider and fragment ion searching showed structural similarities to the IAS and included potential breakdown products and impurities. Ten of the identified compounds were listed in the Food Contact Chemicals database as substances of potential concern or priority hazardous substances, warranting additional studies on their migration into various foods. Overall, these results provide information on the types of chemicals migrating from different food contact materials over time and expand our currently limited understanding of NIAS migrating into food.

Precision Environmental Health Monitoring by Longitudinal Exposome and Multi-Omics Profiling
Peng Gao[1]*, Xiaotao Shen*[1]*, Xinyue Zhang*[1]*, Chao Jiang*[2] *and Michael Snyder*[1]*, (1) Stanford University, (2) Zhejiang University, China*
Conventional environmental health studies primarily focused on limited environmental stressors at the population level, which lacks the power to dissect the complexity and heterogeneity of individualized environmental exposures. Here we integrated deep-profiled longitudinal personal exposome and internal multi-omics to systematically investigate how the exposome shapes a single individual's phenome as a pilot case study. We annotated thousands of chemical and biological components in the personal exposome cloud and found they were significantly correlated with thousands of internal biomolecules, which was further cross validated using corresponding clinical data. Our results showed that agrochemicals and fungi predominated in the highly diverse and dynamic personal exposome, and the biomolecules and pathways related to the individual's immune system, kidneys, and liver were highly associated with the personal external exposome. Overall, this data-driven longitudinal monitoring study demonstrates the potential dynamic interactions between the personal exposome and internal multi-omics, and the impact of the exposome on precision health by producing abundant testable hypotheses. Preprint link:
https://www.biorxiv.org/content/10.1101/2021.05.05.442855

The Chemical Exposome in Brain Cancer: An Exploratory Study
Rubén Gil-Solsona[1]*, Albert Pons-Escoda*[2]*, Noemí Vidal-Sarro*[2]*, Marta Turull-Lopez*[1]*, Daniel Gutierrez-Martin*[1]*, Jordi Bruna*[2]*, Merce Gari*[1]*, Sergi Díez*[1]*, Payam Dadvand*[3]*, Carles Majos*[2] *and* ***Pablo Gago Ferrero***[1]*, (1) Institute of Environmental Assessment and Water Research (IDAEA), Spanish Research Council (CSIC), Spain, (2) Hospital Universitari de Bellvitge, Spain, (3) ISGlobal-CREAL, Spain*
Diffuse gliomas are a highly heterogeneous and aggressive brain tumours with poor prognosis and survival and few established risk factors. Environmental exposures are suspected in the pathogenesis of these tumours; however, results of existing studies are limited and inconsistent, particularly for exogenous organic chemicals, with no available characterization of the chemical exposome of these tumours. Also, better understanding of phenotypic differences in tumour types is needed in order to improve clinical decision making and provision of personalised treatment recommendations. In this proof of concept study we analysed 33 glioblastoma samples (Bellvitge Glioma Cohort (BGC), Spain, 2005-present), including 16 methylated and 17 non-methylated tumours combining HRMS-based wide-scope target and suspect strategies. Forty-six exogenous chemicals were identified in the tumour tissue samples (31 confirmed with standard) including a variety of industrial chemicals (e.g. plastic additives or perfluorinated compounds), personal care products and pharmaceuticals. Our findings provide novel evidence on the presence of these chemicals in brain tissue, highlighting the need for comprehensive evaluations of their potential effects in the tumour pathogenesis. Finally, after applying

metabolomics methods we observed clear differences in the profiles of endogenous chemicals among the studied glioma subtypes, and identified possible biomarkers. These chemicals have potential to be determined in a non-invasive manner, either by LC-HRMS-based blood analysis or using complementary techniques (proton magnetic resonance (1H-MRS)). These are inspiring results since methylation is a strong independent predictor of survival as well as tumour response to chemotherapy for glioblastoma. Indeed, its non-invasive and pre-surgical determination would have a major impact on patient management. Our preliminary data is suggestive for the potential of nontargeted exposome methods to find new valuable biomarkers for diffuse gliomas diagnostic and prognostic stratification.

The Potential Link Between CKDu and Organic Pollutants: A Non-Targeted Assessment of Organic Contaminants in Sri Lankan Waters

Jake Ulrich[1]*, Mangala De Silva*[2]*, Nishad Jayasundara*[1] *and Lee Ferguson*[1]*, (1) Duke University, (2) University of Ruhuna, Sri Lanka*

Chronic kidney disease (CKD) affects ~15% of the worldwide population. The role of environmental contaminants in kidney dysfunction is poorly understood. An increase in the prevalence of CKD of unknown origins (CKDu) has become a significant public health concern in Sri Lanka, with 5-20% of the adult population in endemic regions diagnosed with the disease. The goal of our study was to use non-targeted screening methods based on high-resolution mass spectrometry to explore the prevalence of organic pollutants in Sri Lankan drinking water sources in CKDu-endemic areas. The analytical strategy used was based on UHPLC coupled to high-resolution, accurate mass MS/MS (LC-HRMS), using an Orbitrap Fusion Lumos mass spectrometer (ThermoFisher). High resolution, data-dependent MS/MS with HCD fragmentation and Orbitrap mass analysis was performed. High-resolution product ion spectra were used to interrogate tandem MS spectral libraries for compound identification, together with an *in-silico* mass spectrometry workflow using chemoinformatic methods integrated with a commercial HRMS data processing system. Using our workflow, 1400+ unique chemical features were identified in a preliminary study with 18 wells. A total of 28 compounds were identified using authentic standards. Using library matching with a spectral library containing >30,000 unique compounds, 127 compounds were annotated with a library match score >75%. This includes compounds such as Atrazine, Bisphenol S, Diuron, Genistein, and Verapamil. Of the tentatively identified chemicals, only 58 were found to be tested in the EPA ToxCast database. These chemicals displayed activity in over 650 assays that target a range of toxicological endpoints using various cell tissues. Also, 79 assays had 10 or more of the 58 compounds display activity, suggesting there are multiple compounds present in these waters that can amplify common targets as toxicological endpoints. Cell proliferation and regulation of transcription factor activity were the most abundant active assays of these assays in this preliminary study, suggesting that these may contribute to biological processes linked to nephrotoxicity in Sri Lanka well water. However, a larger field study examining 100 wells in rural Sri Lanka is currently in progress and will help to further elucidate chemical suspects in relation to CKDu.

Untargeted Profiling of the Adipose Tissue Exposome

Brooklynn McNeil[1]*, Anna Robuck*[1]*, Thomas Ziegler*[2]*, Rob McConnell*[3]*, Brittney Baumert*[3]*, Michele La Merrill*[4]*, Lida Chatzi*[3] *and Douglas Walker*[1]*, (1) Icahn School of Medicine at Mount Sinai, (2) Emory University, (3) University of Southern California, (4) University of California*

The global prevalence of obesity and type 2 diabetes (T2DM) has been rising over the past four decades. While changing diet and physical activity patterns are well recognized contributors, exposure to persistent organic pollutants and their interaction with adipose tissue biology may increase risk of obesity and related complications. These chemicals, which accumulate in adipose tissue, have the potential to interact with metabolic pathways and influence fat cell metabolism and inflammation. To better characterize exposome profiles in adipose tissue, we have developed a robust untargeted analytical workflow based upon gas chromatography with high resolution mass spectrometry (GC-HRMS) allowing hybrid targeted and untargeted analysis of human adipose tissues. The method was optimized using a series of human and whale tissue

samples; greatest sensitivity, robustness and reproducibility was achieved using a modified QUEChERS method that requires only 250 mg of tissue and provides limits of detection < 100 pg/g. To evaluate the adipose tissue exposome in human samples, we analyzed adipose tissue biopsies from 107 adults undergoing bariatric surgery in either the US or Norway, and 69 adolescent bariatric surgery patients from the US. Targeted analysis detected multiple persistent organic pollutants, with the most frequently detected including DDE, multiple PCB congeners, and PBDE-47. Untargeted data extraction detected over 30,000 features, the majority of which did not match known environmental exposures. Comparison of exposures across the tissue collection locations suggest potential differences in exposome profiles; however, greater changes were detected when comparing adolescent and adult exposure profiles. This was expected since lipophilic compounds show a time-dependent accumulation across the lifespan. These results provide one of the first exposomic analyses of human adipose tissue using an untargeted GC-HRMS workflow. Future work within these studies will evaluate how the adipose tissue exposome burden influences weight loss trajectories following bariatric surgery and redistribution of pollutant burden using plasma samples collected multiple time points after surgery.

Using Estrogen-Responsive Reporter Gene Assays to Identify Nontarget Mammary Gland Carcinogens in Drinking Water

***Gabrielle Black**, Thomas Young, Michael Denison and Guochun He, University of California*

Four regions in California have been identified by the California Breast Cancer Research Program and Public Health Institute as areas with disproportionally high rates of breast cancer compared to the rest of the state. This study aims to identify chemicals in household drinking water correlated to measured estrogenicity collected from these breast cancer "hot spot" zones. Tap water was collected from 120 homes in California, half of which encompass breast cancer "hot spots" and analyzed using High Resolution LC- and GC-QTOF-MS. Estrogenicity was measured using reporter-gene assays and statistically correlated with nontarget feature abundances. Annotation and identification were carried out using a suite of chemometric tools to identify potentially new mammary gland carcinogens. Suspect screening and targeted analysis were performed synchronously to support this work and has included quantitative analysis for per- and poly-fluorinated compounds, disinfection byproducts, and trace inorganics.

Applications to Regulatory Frameworks and Monitoring Programs

Lessons Learned During Environmental Forensic Investigation of Chemical Manufacturing and Use of PFAS by Nontargeted Analysis

***Mark Strynar** and James McCord, U.S. Environmental Protection Agency*

The intersection between high resolution mass spectrometry (HRMS), non-targeted analysis (NTA) and per- and polyfluoroalkyl substances (PFAS) brings three complex topics together. Over the past several years we have obtained numerous samples of industrial effluent and contaminated surface, ground, and wastewater from collaborations with several state and regional EPA offices investigating PFAS production sites. NTA using HRMS (e.g. QTOF or Orbitrap) was used to examine the chemical identities of PFAS within these samples and investigate the abundances of legacy and emerging chemical contaminants. Through this approach, we have been able to identify legacy PFAS, as well as elucidate alternative PFAS used by a variety of chemical manufacturers. Alternative PFAS include a previously unidentified family of chlorinated perfluoropolyethers (Cl-PFECAs), polyfluorinated side products of polyfluorovinylidine (PVDF), as well as several other novel PFAS fluoroether species exhibiting ether linkages and acid head groups. This presentation will discuss lessons learned while conducting PFAS site investigation case studies. Topics will include decisions concerning sample preparation and analysis conditions, isolation of PFAS from large data files, prioritization of detected chemicals, data reduction strategies, alternative data mining techniques and determination of data analysis completion. In many instances there is no "one size fits all" approach for the above decisions being made. However, shared experiences with the greater scientific community may benefit veteran and novice investigators alike.

LIFE APEX - Improving the Systematic Use of Contaminant Data from Apex Predators and Their Prey in Chemicals Management

Jaroslav Slobodnik[1], *Nikiforos Alygizakis*[1], *Jürgen Arning*[2], *Alexander Badry*[2], *Kevin Bauer*[2], *Ľuboš Čirka*[1], *Alessandra Cincinelli*[3], *Daniela Claßen*[2], *Rene Dekker*[4], *Guy Duke*[5], *Wiebke Drost*[2], *Natalia Glowacka*[1], *Georgios Gkotsis*[6], *Burkhard Knopf*[7], *Tania Martellini*[3], *Paola Movalli*[4], *Maria-Christina Nika*[6], *Peter Oswald*[1], *Heinz Ruedel*[7], *Richard Shore*[8], *Nikolaos Thomaidis*[6], *Lee Walker*[8], *Gabriele Treu*[2] *and Jan Koschorreck*[2], *(1) Environmental Institute, Slovakia, (2) German Environment Agency (Umweltbundesamt), (3) University of Florence, Italy, (4) Naturalis Biodiversity Center, Netherlands, (5) University of Oxford, United Kingdom, (6) National and Kapodistrian University of Athens, Greece, (7) Fraunhofer IME - Institute for Molecular Biology and Applied Ecology, Germany, (8) UK Centre for Ecology & Hydrology, (12)UK Centre for Ecology & Hydrology*

The hypothesis of whether new methods of chemical analysis, together with collections of wildlife samples, can effectively support chemical management to control substances with persistent and bioaccumulative properties has been tested. The EU-funded LIFE APEX project aims at demonstrating feasibility of wide-scope target and suspect chemical screening of top predators and their prey for regulatory purposes at the European scale, in support of the European Chemicals Strategy for Sustainability, REACH and Biocidal Products Regulation. A well-structured network of 95 European environmental specimen banks, natural history museums, and research collections has been identified that systematically collects and archives wildlife samples. Altogether 198 liver samples from 20 countries were successfully taken from top predators, *i.e.* birds of prey (buzzard) as well as otters and marine mammals (dolphins, porpoises, seals) along with fish filets from their same regions. Subsequent wide-scope target screening of each sample for more than 2,400 substances and a suspect screening of more than 65,000 substances by gas and liquid chromatography coupled to high resolution mass spectrometry techniques provided information on presence/absence of the compounds in the samples and a quantitative or semi-quantitative estimate of their concentrations. The raw data obtained were digitally archived in the NORMAN Digital Sample Freezing Platform for future retrospective screening of additional substances. Additionally, 116 legacy persistent organic pollutants were analysed to benchmark the screening results. Guidelines for quality assurance in the collection, archiving and shipping of samples for analyses have been developed to support future standardization and networking efforts. All data and supporting information have been stored in four modules of the LIFE APEX Database System which will be made accessible for regulators, researchers, and the public. The detected substances were ranked based on their exposure (Frequency of Appearance) and hazard assessment (model-predicted persistency, bioaccumulation, toxicity (PBT) properties). All detected substances were subjected to the detailed PBT assessment by the regulatory partner - German Environment Agency, and the frequently appearing chemicals with P and B properties were prioritised.

Acknowledgement This research has received funding from the European Union through the program LIFE17 ENV/SK/000355 “Systematic use of contaminant data from apex predators and their prey in chemicals management”.

The NORMAN NTS Prioritization Workflow in Support of Environmental Risk Assessment and Chemicals Management

Valeria Dulio[1], *Nikiforos Alygizakis*[2], *Peter Von der Ohe*[3], *Emma Schymanski*[4], *Juliane Hollender*[5], *Reza Aalizadeh*[6], *Nikolaos Thomaidis*[6], *Jan Koschorreck*[3] *and Jaroslav Slobodnik*[2], *(1) INERIS, France, (2) Environmental Institute, Slovakia, (3) German Environment Agency (Umweltbundesamt), (4) University of Luxembourg, (5) Eawag - Swiss Federal Institute of Aquatic Science and Technology, (6) National and Kapodistrian University of Athens, Greece*

It is increasingly recognized that the number of compounds measured via target chemical analysis is not sufficient to provide an overview of the status of the environment. The application of non-target screening (NTS) techniques is a suitable approach to detect overlooked substances and their mixtures. The NORMAN Association (www.norman-network.net) has built a comprehensive infrastructure to support laboratories using NTS techniques and promote their application in regulation. Screening results from target, suspect and non-

target analysis by LC- and GC-HRMS are stored in different modules of the NORMAN Database System (NDS; https://www.norman-network.com/nds/), enabling retrospective screening. Lists of substances of interest are contributed by numerous research groups and regulators worldwide via the Suspect List Exchange initiative (https://www.norman-network.com/nds/SLE/). A merged, curated list of all substances is available as NORMAN SusDat (https://www.norman-network.com/nds/susdat/), which contains 109,619 substances as of January 2022 and is constantly growing. Key screening parameters, physico-chemical, hazard and (eco)toxicological properties such as PNECs are available in the interlinked modules of the NDS. Any changes are traceable, thus allowing for an automated and transparent prioritisation process in a dedicated module. Thanks to this comprehensive infrastructure, NORMAN promotes the application of a prioritization workflow where NTS is applied as a screening tool in a step-wise approach to support decision-makers for further investigation and follow-up regulatory action. Using this workflow data are retrospectively screened, detected suspect compounds are assigned a reproducible and transparent identification score based on MS, MS/MS, retention time information, and finally semiquantified using the structurally most similar target compound. The workflow enables allocation of substances to different categories for further regulatory actions, depending on a series of indicators and metadata, such as the frequency of appearance, frequency of exceedance of PNEC and type of PNEC (experimental / predicted). An overview of the NORMAN prioritization workflow will be shown, together with results of selected case studies. The practicality, risk assessment potential and regulatory implications will be demonstrated on datasets from large-scale projects covering European marine and riverine environment, including biota and wastewater effluents.

Building Stakeholder Relationships to Facilitate Expanded Application of Non-Targeted Analysis Methods

Sara Nason[1]*, Yong-Lai Feng*[2]*, Jon Sobus*[3]*, Christine Fisher (O'Donnell)*[4]*, Ruth Marfil-Vega*[5]*, Krystal Pollitt*[6] *and Matthew Newmeyer*[7]*, (1) Connecticut Agricultural Experiment Station, (2) Health Canada, (3) U.S. Environmental Protection Agency, (4) U.S. Food and Drug Administration, (5) Shimadzu Scientific Instruments, (6) Yale University, (7) Johns Hopkins University*

By nature, working groups focused on sharing and improving analytical methods are primarily comprised of those who actively work on method development research. As part of the Benchmarking and Publications for Non-Targeted Analysis working group (BP4NTA), we recognize that the potential audience for our products extends far beyond our current membership. Non-targeted analysis (NTA) using high resolution mass spectrometry is valuable for a wide range of applications including environmental monitoring, food safety evaluation, human exposure and biomonitoring, forensic science, epidemiology, chemical threat detection, and more. Stakeholders who can benefit from implementing current NTA workflows and developments in NTA science include scientists who run these analyses in the lab, as well as those who may use NTA data for developing new environmental, food safety and health policies. The BP4NTA Stakeholder Subcommittee is working to engage with stakeholders from various fields through a series of meetings, workshops, and surveys where the main goal will be to listen to the attendees, learn about the strongest drivers for NTA adoption, and clarify primary limitations preventing these potential stakeholders from collecting and utilizing NTA data. The subcommittee is further planning to develop instructional materials such as fact sheets and videos to introduce NTA and fill in common stakeholder knowledge gaps. These resources will be made publicly available through the BP4NTA website (https://nontargetedanalysis.org/), enabling easy access for users across diverse fields. Outcomes from these initial efforts will establish a framework for enabling broader implementation of NTA workflows. Our long-term goals include collaborating with stakeholders to develop guidance materials and creating standardized NTA methods. In this presentation, we will share findings from the stakeholder meetings and workshops we have hosted so far and show content from the materials we have developed. In addition, we hope this presentation will prompt additional feedback so we can further improve these materials and facilitate additional collaborations with stakeholders across the globe. This abstract does not necessarily reflect agency policy.

A Decade of Nontargeted Analysis in the Great Lakes Fish Monitoring and Surveillance Program

Junda Ren[1], Aikebaier Renaguli[1], Sadjad Fakouri Baygi[1], Sujan Fernando[1], Michael Milligan[2], James Pagano[3], Philip Hopke[1], Thomas Holsen[1] and ***Bernard Crimmins****[1], (1) Clarkson University, (2) SUNY Fredonia, (3) State University of New York, Oswego*

The Great Lakes Fish Monitoring and Surveillance Program (GLFMSP) has measured the legacy contaminant burdens in the Great Lakes basin for over 5 decades. In 2010, GLFMSP expanded the scope of contaminant monitoring to include emerging contaminants and novel contaminant discovery using the legacy top predator fish monitoring framework. An array of instrumentation was deployed based on the compound classes (un)targeted. For example, 2-dimensional gas chromatography coupled to a high-resolution time of flight mass spectrometer (GC×GC - HRT) is used for halogenated semi-volatiles and liquid chromatography coupled to a quadrupole time of flight (LC-QToF) mass spectrometer for more water-soluble species such as polyfluoroalkyl substances (PFAS). Data processing workflows were developed to handle the large volume of complex data generated by these platforms. The current presentation will highlight a decade of nontargeted analysis (NTA) and novel chemical discovery in GLFMSP. The discussion will include lessons learned by our group in developing approaches for data interpretation and adapting to the analytical/computation limitations. We will show methods we have used to maximize our understanding of the fate and distribution of novel chemical signatures and future work needed to adapt our workflows to other monitoring programs.

Applications to Regulatory Frameworks and Monitoring Programs (Poster)

Combatting Illegal Logging and Trafficking of Endangered Wood Species by Using High-Resolution LC/MS and GC/MS Data with Machine Learning Statistics

Pamela Brunswick[1], ***Daniel Cuthbertson****[2], Jeffrey Yan[1], Candice Chua[1], Isabelle Duchesne[3], Peter Gasson[4], Geoffrey Kite[4], Joy Bruno[1], Graham van Aggelen[1] and Dayue Shang[1], (1) Environment and Climate Change Canada, (2) Agilent Technologies, Inc., (3) Natural Resources Canada - Canadian Forest Service, (4) Royal Botanic Gardens, United Kingdom*

Illegal logging and trafficking of endangered timber products is a profitable form of income for major organized crime syndicates, with associated deforestation of protected habitats and other ecological and social damages. Enforcement by CBSA, Interpol, and other law enforcement agencies has been hampered by a lack of practical assessment tools, including the inability of commonly used DNA fingerprinting methods to reliably identify wood species. In this study, unbiased data acquisition methods allowed the development of machine learning statistical tools for the accurate identification of commonly trafficked *Dalbergia* rosewood species. *Dalbergia* wood reference specimens were extracted in acidified methanol and adjusted for appropriate concentration based on extract coloration. Aliquots were analyzed by either gas chromatography (GC) quadrupole time-of-flight (QTOF), or liquid chromatography (LC) QTOF, mass spectrometry method. Both systems were optimized to produce characteristic chromatographic profiles. A variety of traditional univariate statistics were assessed, in addition to machine learning approaches including Random Forest decision trees and Partial Least Squares Discriminant Analysis (PLS-DA). The selected machine learning models were then validated with naïve specimens. Initial comparisons of *Dalbergia* species against look-alike genera *Milcia and Patymiscium* yielded clear differentiation using Principle Components Analysis (PCA). A second set of Random Forest -models within closely similar *Dalbergia* species only, were able to reliably predict the species with low out-of-bag (OOB) error rates. An exception in this species set was apparent for *D.spruceana,* for which a sub-model was required for more accurate prediction.Here we demonstrated the feasibility of a complimentary chemotyping method for endangered *Dalbergia* spp. using both LC and GC coupled to high-resolution mass spectrometers, together with machine learning statistics, as methods to successfully classify unknown wood samples. These models and techniques may be further developed to serve as tools for customs investigations of imported woods.

Nontarget Analysis of Polluted Surface Waters: A Case Study in Bangladesh

Bénilde Bonnefille[1]*, Oskar Karlsson*[1]*, Rubhana Raqib*[2]*, M. Faruque Parvez*[3]*, M. Sirajul Islam*[2] *and Jonathan Martin*[4]*, (1) Stockholm University, Sweden, (2) International Centre for Diarrhoeal Disease Research, Bangladesh, (3) Columbia University Mailman School of Public Health, (4) University of Alberta, Sweden*

In Bangladesh, the high density of population and industry, as well intensity of agriculture practices, is leads to major surface water contamination, including by organic micropollutants. Given that Bangladesh's population also depends on local fish as a staple dietary component, this contamination poses unknown risk to both human and environmental health. Thus, in this study we undertook a surface water monitoring survey with comprehensive analysis by nontarget mass spectrometry. Surface waters from 12 sites located on 4 different rivers around the capital city, Dhaka, were were spiked with internal standards and filtered before analysis. Analyses were performed using liquid chromatography high-resolution mass spectrometry in 4 modes: positive and negative electrospray (± ESI), and atmospheric pressure chemical ionization (± APCI). Among 39,030 total detected molecular features, there was relatively little redundancy between ionization modes, and over 26,000 unique features were detected. Given that 51.9-77.5% of the features in each mode were unique, the combination of these 4 ionization modes is necessary to cover the relevant chemical space in polluted surface waters. In a principal component analysis, components PC1 and PC2 explained 54.8% and 6.4% of the overall data variability, respectively, and the scores plot revealed two groups of samples separated by geographical location on PC1, and an upstream-downstream distribution on PC2. This suggested strong chemical profile differences between highly urbanized and more remote sampling sites, and a pollution gradient along the flowpath of rivers moving towards the Bay of Bengal. A nontarget workflow using open science tools allowed us to confirm or annotate many anthropogenic contaminants, including pharmaceuticals (e.g. diclofenac), pesticides (e.g. diuron), and industrial chemicals (e.g. anionic surfactants). An example annotation that may be related to local textile industry was indigo dye (level 2a), which had a high MS^2 similarity score (796/1000) compared to the public MS^2 database available on MS-DIAL website (v2021/04/13, http://prime.psc.riken.jp/compms/msdial/main.html#MSP). Most compounds identified had a spatial distribution with higher response intensity closer to Dhaka. Work is ongoing to identify contaminants at higher confidence levels, their likely sources, and bioaccumulation in fish sold at local markets.

Non-Targeted Screening of Pesticides in Honey From Canada Using LC-QTOF-MS

Lei Tian[1]*, Caren Akiki*[1]*, Shaghig Bilamjian*[1]*, Tarun Anumol*[2]*, Daniel Cuthbertson*[2] *and* ***Stéphane Bayen***[1]*, (1) McGill University, Canada, (2) Agilent Technologies, Inc.*

The global honey market, and notably the demand for organic and high quality honeys, is currently growing, driven by the booming interest of consumers for natural sweeteners and healthy food products. Honey bees can be exposed to environmental contaminants, including pesticides, via contact and the consumption of contaminated pollen and nectar. Traces of these environmental contaminant residues may in turn accumulate in honey and other bee products. In our previous study, a method based on direct injection in LC-QTOF-MS was developed for the suspected-target screening of plastic-related chemicals in honey. This novel method relies on the analysis of small amount of honey (0.2 mg), requires little sample preparation (< 20 min.), and can be automated. In this study, a non-targeted workflow was adapted from this earlier approach and assessed for its capacity to screen for pesticide residues. Twelve model pesticides were used to validate the non-targeted workflow and another six mass-labeled standards were used for method validation. Good instrumental linearity (1-100 ng/L, r^2>0.99) and recoveries (98-105%) were achieved. The workflow was applied to honey samples collected from the Canadian market, and several pesticides were successfully detected and tentatively identified. Unique patterns based on non-targeted analysis were found in some specific types of honey with different floral origin. These results further confirm that the present method is quick, simple and effective, which shows the potential to screen many classes of contaminants to ensure the quality, authenticity and safety of honey.

Time for Change - a Modern Concept for Water Monitoring Using LC-HRMS

Susanne Brüggen*, North Rhine Westphalian State Agency for Nature, Environment and Consumer Protection*

(LANUV NRW), Germany
Water monitoring in North Rhine-Westphalia, Germany is carried out by the North Rhine-Westphalia Office of Nature, Environment and Consumer Protection (LANUV) at over 2.000 monitoring sites using quantitative results. This monitoring requires the application and maintenance of several different target LC-MS methods, each with complex quantification and quality assurance. However, information on new substances are completely missing. With the establishment of LC-HRMS in recent years, future water monitoring may consist of the application of an LC-HRMS screening method to ensure water monitoring and additionally fill the knowledge gap. The advantage of LC-HRMS screening is the different possibilities of evaluation: 1) Target - quantitative results A simplified quantification concept with a calibration point of 0.1 µg/L was developed in this work. A comparison of 211 analytes over three matrices, which were evaluated with the classical quantification concept using a multi-point calibration and the simplified quantification concept, showed no significant difference between the results, the concepts are comparable. 2) Suspected Target - qualitative results The *Gewässersteckbriefe (water characteristics)* developed in this work, containing qualitative information on the occurrence and frequency of e.g. organic trace substances, which have not yet been analysed in a LANUV routine method so far. By comparing different surface waters with each other, it is possible to identify substances that are ubiquitously distributed or occur only in specific water bodies. 3) Non Target - possibility to identify unknown substances and retrospective The mass 315.0544 Da in negative mode became conspicuous due to an increasing time course in the winter months and could be clearly identified after synthesising the standard. Currently, attempts are being made to identify the source of the emission. For each strategy, examples are explained in detail on the poster. This concept could change water monitoring and assessment, and make it much more efficient without losing information. There is a chance to measure less but learn more about the water bodies.

Interactive Session: Harmonization and Assessment of Approaches

A Roadmap of Harmonization Efforts From Benchmarking and Publications for Non-Targeted Analysis (BP4NTA): Successes, Challenges, and Future Developments

Ruth Marfil-Vega[1]*, Christine Fisher (O'Donnell)*[2]*, Benjamin Place*[3] *and Elin M. Ulrich*[4]*, (1) Shimadzu Scientific Instruments, (2) U.S. Food and Drug Administration, (3) National Institute of Standards and Technology, (4) U.S. Environmental Protection Agency*
Non-Targeted Analysis (NTA) using high resolution mass spectrometry has been implemented to explore complex samples and identify novel contaminants by many research communities, from the environmental science and food chemistry fields to ‘omics and toxicology. Recent technological developments in hardware and software have contributed to the exponential growth of this field within the research communities (the “NTA practitioners”) as well as the increasing interest in the outcomes from NTA by multiple stakeholders developing new environmental, food safety and health policies (the “NTA data end-users”). Despite global efforts to identify, disseminate, and adopt community-wide definitions, methods, quality control, and more, there are still foundational knowledge gaps (e.g., semi-quantitative analysis, procedures for method performance assessment) and additional work is needed to harmonize these efforts. As a result, the broader adoption of NTA studies and the understanding of its outcomes by stakeholders can be hindered. The Benchmarking and Publications for Non-Targeted Analysis (BP4NTA) working group was established in 2018 to address these challenges. Since its creation, the group has created two peer-reviewed publications and launched a website hosting a consensus-based glossary of NTA terms, a Study Reporting Tool, and reference content, among other resources, to guide scientists in planning, reporting, and reviewing NTA studies, results, and manuscripts. The website can be found at https://nontargetedanalysis.org. Current projects include upcoming manuscripts focused on performance metrics for evaluating NTA results and understanding the chemical space covered by NTA studies, accompanied by tools to help the growing community use NTA in their fields. A recently kickstarted project is aimed at engaging with stakeholders to identify and understand their needs, fulfill their knowledge gaps regarding NTA science and data, and co-create tools and resources to

enable broader implementation of NTA. In this presentation, we will provide a detailed overview of the activities and outcomes from the group since its inception and our future activities. Additionally, we will share the challenges we have encountered along the way and our emerging needs as a new consensus group, with the aim of opening a discussion and strengthening collaborations with other like-minded scientists and organizations across the globe. Abstract does not necessarily reflect agency policy.

Non-Target Screening Techniques to Support Environmental Risk Assessment and Chemicals Management: The NORMAN Approach

Jaroslav Slobodnik[1]*, Nikiforos Alygizakis*[1]*, Reza Aalizadeh*[2]*, Nikolaos Thomaidis*[2]*, Emma Schymanski*[3]*, Peter Von der Ohe*[4]*, Valeria Dulio*[5]*, Stellan Fischer*[6]*, Jan Koschorreck*[4] *and Juliane Hollender*[7]*, (1) Environmental Institute, Slovakia, (2) National and Kapodistrian University of Athens, Greece, (3) University of Luxembourg, (4) German Environment Agency (Umweltbundesamt), (5) INERIS, France, (6) Swedish Chemicals Inspectorate, (8) Eawag - Swiss Federal Institute of Aquatic Science and Technology, Switzerland*

It is increasingly recognised that the number of compounds measured today via target chemical analysis is not sufficient to provide an exhaustive overview of the status of the environment and that the application of non-target screening (NTS) techniques is needed in order to detect the presence of harmful substances potentially overlooked with target analysis. The NORMAN Association (www.norman-network.net) is an independent network of more than 80 research institutions, reference laboratories, industry, universities and regulatory bodies to support policy-makers in all decisions for the identification, monitoring, assessment and prioritisation of relevant emerging pollutants and their transformation products in the environment. Results from wide-scope target, suspect and non-target screening, as well as raw LC- and GC-HRMS data, allowing for their retrospective screening are stored in different modules of the NORMAN Database System (NDS; https://www.norman-network.com/nds/). NORMAN substances are contributed by numerous research groups and regulators worldwide and stored in the Suspect List Exchange (https://www.norman-network.com/nds/SLE/). A merged, curated list of all substances with structures (*e.g.* SMILES, InChI and InChIKey), the exact mass (molecular ion, adducts etc.) and chromatography (Retention Time Index) characteristics, is in NORMAN SusDat (https://www.norman-network.com/nds/susdat/), which contained 65,697 substances as of 22 January 2020 and is constantly growing. NORMAN puts significant effort in systematic collection of additional information on exposure (via use), hazard (PBT, vPvB, ED, CMR, PMT properties) and risk (exceedance of (eco)toxicological threshold values) of each of these substances to complement data on their occurrence in the environment and prioritise them for the follow up regulatory actions. All above data are closely interlinked in the NDS. Any changes/updates are traceable, and thus allowing the prioritisation process to be automated and transparent. An overview of the NORMAN NTS data collection, archiving and prioritisation tools will be provided together with results of selected case studies carried out in support of the European Commission, European Chemicals Agency, river basin managers, sea conventions and national authorities. Their practicality, risk assessment potential and regulatory implications will be demonstrated on datasets from large-scale projects covering European marine and riverine environment, wastewater effluents and biota. **Acknowledgements** Authors would like to acknowledge all NORMAN network partners for continuous support of NTS activities and European Union for its funding through the LIFE17 ENV/SK/000355 project “Systematic use of contaminant data from apex predators and their prey in chemicals management”.

Mapping Chemical Space Coverage in Non-Targeted Analysis

Gabrielle Black[1]*, Charles Lowe*[2]*, Tarun Anumol*[3]*, Jessica Bade*[4]*, Kristin Favela*[5]*, Christine Fisher (O'Donnell)*[6]*, Yong-Lai Feng*[7]*, Alan Hood*[6]*, Ann Knolhoff*[6]*, Andrew D. McEachran*[3]*, Jamie Nunez*[4]*, Katherine Peter*[8]*, Natalia Soares Quinete*[9]*, Jon Sobus*[2]*, Eric Sussmann*[6]*, William Watson*[5]*, Antony Williams*[2] *and Samanthi Wickramasekara*[10]*, (1) University of California, (2) U.S. Environmental Protection Agency, (3) Agilent Technologies, Inc., (4) Pacific Northwest National Laboratory, (5) Southwest Research Institute, (6) U.S. Food and Drug Administration, (7) Health Canada, (8) University of Washington, (9) Florida International*

University, (10) Arizona Laboratory of Emerging Contaminants
Non-targeted analysis (NTA) using high-resolution mass spectrometry allows scientists to detect and identify a broad range of compounds for further analysis, future monitoring, and potential prioritization for toxicological and exposure assessment in diverse matrices without a priori chemical knowledge. Relative to targeted methods, these NTA methods present the opportunity to describe the constituents of a sample across a multidimensional swatch of chemical properties, referred to as chemical space. However, understanding and communicating which "corner" of chemical space is detectable by an NTA workflow remains challenging and non-standardized. For example, sample collection, preparation, and NTA data acquisition, processing, and filtering steps will all influence the types of chemicals that are detected and identified. Accordingly, it is challenging to assess whether analyte non-detection in an NTA study indicates true absence in a sample (above a detection limit) or is a false negative result driven by limitations of the workflow. Such information gaps can reduce the utility and reliability of chemical screening data, particularly in the context of exposure and risk assessment. Here, we describe the need for accessible approaches that enable chemical space mapping in NTA studies. We identify a suite of existing predictive and analytical tools that can be used in combination to generate scores that relate the likelihood that each identified compound was plausibly detected according to the predicted chemical space of the NTA workflow. High-scoring compounds, therefore, correlate to a high likelihood that the compound could be detected and identified in the given workflow, whereas low scoring compounds have a low probability of being detected (even if truly present in the samples of interest) and could be used in decoy libraries to help researchers determine appropriate minimum scoring thresholds throughout the analytical workflow. This work also highlights the chemometric tools needed to make such a tool robust and usable across a wide range of NTA disciplines. Ultimately, development of a chemical space mapping tool strives to enable further standardization of NTA by improving method transparency and communication around false detection rates, thus allowing for more direct method comparisons between studies and improved reproducibility. This, in turn, is expected to promote further widespread application of NTA beyond research-oriented settings. This abstract does not necessarily represent the policies of the USEPA.

Quantitative Non-Targeted Analysis: From Data to Decisions

***Jon Sobus**, Louis C. Groff and James McCord, U.S. Environmental Protection Agency*

Over the past decade, non-targeted analysis (NTA) methods have fundamentally transformed the manner in which emerging contaminants are identified in environmental and human health studies. A wide variety of qualitative methods for rapid and accurate chemical identification have been developed and implemented with great speed and success. Yet, the full potential of NTA methods cannot be realized until techniques are expanded to facilitate comprehensive quantitative evaluation. In many NTA studies, chemical standards are procured after the initial analysis to confirm and quantify compounds of interest. Many tentatively identified chemicals, however, cannot be readily purchased, obtained, or synthesized, providing a considerable barrier to further quantitative examination. To overcome this challenge, a number of studies have recently proposed methods for estimating point concentrations for tentatively identified chemicals using analytical surrogates and/or model-based predictions. To date, limited consideration has been given to estimating uncertainty stemming from experimental error (random and systematic), surrogate selection strategies, model prediction error, and sample preparation procedures. As part of this interactive session, we will briefly review existing qNTA methods, covering an assortment of strategies utilized for compound quantitation in environmental studies. We will then describe newly proposed methods to estimate and communicate prediction uncertainty via the reporting of concentration confidence intervals for individual analytes in individual samples. We will further initiate group discussion on procedures for relating estimated concentrations in prepared sample extracts to estimated concentrations in original sample matrices. Finally, we will discuss strategies to relate qNTA concentration estimates to hazard-based levels-of-interest to set research priorities for more in-depth risk-based examinations. *This abstract does not necessarily reflect agency policy.*

Interactive Session: Nontargeted Analysis of Per- and Polyfluoroalkyl Substances

Comprehensive Annotation and Validation of PFAS and PFAS Transformation Products Using FluoroMatch Software and Interactive LC-HRMS/MS Visualizations

Jeremy Koelmel[1], *Paul Stelben*[1], *Carrie McDonough*[2], *David Godri*[3], *Jiarong Qi*[1], *Elizabeth Lin*[1], *Antony Williams*[4], *John Bowden*[5], *Emma Rennie*[6] *and Krystal Pollitt*[1], *(1) Yale University, (2) Stony Brook University, (3) 3rd Floor Solutions, (4) U.S. Environmental Protection Agency, (5) University of Florida, (6) Agilent Technologies, Inc.*

Per- and polyfluoroalkyl substances (PFAS) are pervasive, persistent, can bioaccumulate, and have been found to have a wide range of deleterious health effects. Thousands of legacy and alternative PFAS have been identified in consumer products, the environment, and human body. Liquid chromatography high-resolution tandem mass spectrometry (LC-HRMS/MS) provides orthogonal evidence to assign structures to tens to thousands of PFAS molecules. Combining retention time, accurate mass, and fragmentation to assign structure is challenging. In addition to peak picking and blank filtering, FluoroMatch 2 automates assignment of features to homologous series based on retention time order and accurate mass, accurate mass matching to EPA compiled PFAS databases, MS/MS screening across 777 fluorine containing fragments, predicting, and matching against in-silico fragmentation, and MS/MS matching using class-based fragmentation rules established using molecular standards. Many PFAS are biologically transformed in the environment and body. To capture these transformation products, we introduce a new in-silico fragmentation database in FluoroMatch 2 for annotating PFAS biological transformation products. Our original PFAS list used for library development consisted of 8,951 compounds based on the PFASSTRUCTv4 Comptox dashboard list, after transformation of SMILES to MS-READY format, and removal of duplicates. In contrast, our new in-silico fragmentation database includes tens of thousands of possible biological transformation products and precursors (name, enzymatic reaction, formula, and predicted MS/MS spectra) generated using Biotransformer, Chemical Transformation Simulator (CTS) and FluoroMatch Generator. Automated non-targeted workflows can often have high false positive rates. We have further expanded FluoroMatch 2 to include FluoroMatch Visualizer, which can be used for manual validation by visualizing mass spectral data using retention time vs mass plots, normalized mass defect plots (e.g., to CF_2), the annotated and scored FluoroMatch table, and MS/MS plots with fragment assignments. Using FluoroMatch 2 with the new PFAS biologically transformation in-silico fragmentation database and FluoroMatch Visualizer, we show decreased false positives and discovery of unknown previously undetected PFAS in enzymatically treated AFFF. This abstract does not necessarily reflect environmental protection agency policy.

PFAS Detection in Non-Target LC-HRMS Data Using Machine Learning Algorithms

Mathieu Feraud[1], *Saer Samanipour*[2], *Jake O'Brien*[1], *Kevin Thomas*[1], *Sarit Kaserzon*[1] *and Pradeep Dewapriya*[1], *(1)The University of Queensland, Australia, (2)University of Amsterdam, Nederland*

Per- and Polyfluoroalkyl Substances (PFAS) are a diverse class of chemicals that are ubiquitous in the environment. Because of their adverse effects on the environment and human health, there is an increasing interest in detecting and monitoring them. Conventional target analysis only focuses on a fraction of PFAS in the environment. It is suspected that there still remains many unknown PFAS in nature to identify. Non-target analysis (NTA) employing high-resolution mass spectrometry (HRMS) has been developed as a key approach to tackling this problem. Typically, identification PFAS from non-target data analysis requires manual curation of data from expert knowledge. However, this process can be prone to errors, can be less comprehensive, and requires a significant amount of time. Many of the decisions and analyses an expert makes when trying to identify PFAS compounds can be automated into a single algorithm. Because often the values which are used to make decisions are often ambiguous as they are close to the ranges, and the correlation between variables used to make a decision are still not well set, and there could be still unknown PFAS patterns that would help the detection of them, a Machine Learning ML algorithm was used to create an algorithm to detect PFAS. Most ML algorithms have the ability and advantage of detecting known and unknown correlations in the data, without

overfitting to noise or false correlation. Furthermore, the process also automates the detection of PFAS from the data. By compiling HRMS data from MassBank Project(Masbank.jp), our algorithm was developed to detect fluorinated compounds from complex matrices, and the performance of the algorithm was tested on various LC-HRMS datasets generated from a range of matrices with success.

Quantitative Structure Retention Relationships to Identify Per-/Poly-Fluoroalkyl Substances and Other Contaminants of Emerging Concern

Scott Simpson[1]*, Antle Jonathan*[2]*, Lahiruni Halwatura*[2]*, Michael Larock*[1]*, Rebecca Dickman*[2] *and Diana Aga*[3]*, (1) St. Bonaventure University, (2) University at Buffalo, (3) State University of New York, Buffalo*

Nontarget analysis using liquid chromatography–high resolution mass spectrometry (LC–HRMS) is a valuable approach in characterizing for contaminants of emerging concern (CECs) in the environment. However, identification of these analytes can be quite costly or taxing without proper analytical standards. To circumvent this problem we utilize Quantitative structure-retention relationships (QSRR) models to predict elution order and retention times. Properties calculated from density functional theory (DFT) and the conductor-like screening model for real solvents (COSMO-RS) theory are used to produce our QSRR models, which can be calculated for virtually any analyte. We show that this methodology has been successful in identification of per-/poly-fluoroalkyl substances (PFAS) and other contaminants. Nontarget analysis using liquid chromatography–high resolution mass spectrometry (LC–HRMS) is a valuable approach in characterizing for contaminants of emerging concern (CECs) in the environment. However, identification of these analytes can be quite costly or taxing without proper analytical standards. To circumvent this problem we utilize Quantitative structure-retention relationships (QSRR) models to predict elution order and retention times. Properties calculated from density functional theory (DFT) and the conductor-like screening model for real solvents (COSMO-RS) theory are used to produce our QSRR models, which can be calculated for virtually any analyte. We show that this methodology has been successful in identification of per-/poly-fluoroalkyl substances (PFAS) and other contaminants.

Use of Electron Activated Dissociation (EAD) on the SCIEX ZenoTOF 7600 System to Elucidate PFAS Structures

Karl Oetjen*, Craig Butt, Simon Roberts, Diana Tran, Megumi Shimizu and Matt Noestheden, Sciex*

Poly- and perfluoroalkyl substances (PFAS) are well-known environmental contaminants and are widely detected in humans and wildlife, as well as water, soil, and air. PFAS are primarily used for their stain repellency properties as well as their surfactant characteristics (i.e. in foams to combat petroleum fires). Even though there are an estimated 5000 unique PFAS industrially manufactured, most monitoring efforts are focused on only 20-30 compounds. Non-target acquisition using high-resolution accurate mass spectrometry are beneficial for elucidating unknown compound structures, such as PFAS. However, traditional fragmentation methods using collision-induced dissociation (CID) are occasionally too aggressive to form diagnostic MS/MS spectra. Electron Activated Dissociation (EAD) – has shown potential as an alternative form of fragmentation. Standard solutions of PFAS compounds were infused on the SCIEX ZenoTOF 7600 System using both CID and EAD fragmentation modes. In separate EAD experiments, the kinetic energy (KE) was ramped from -10 to 25 V and the electron beam current ramped from 0-8000 V. Further, the reaction time was varied (10, 35, 100 ms). EAD generated a more comprehensive MS/MS fragmentation spectrum as compared to CID which resulted in additional structural information for improved compound elucidation. Preliminary tests were focused on the 5:1:2 fluorotelomer betaine in ESI+ ionization mode. Using CID fragmentation the only fragment formed was the $[C_3H_8N]^+$ ion at m/z 58.065 Da. The SCIEX Fluorochemical HR-MS/MS Spectral Library 2.0 contains the fragmentation spectra for several fluorotelomer betaines and shows that the m/z 58.065 Da ion is the only significant fragment formed with CID. In contrast, the EAD fragmentation spectrum showed many fragment ions which corresponded to the unzipping of the carbon backbone of the molecule. For example, sequential losses of CF2 were observed. This is important because the more comprehensive MS/MS fragmentation

spectrum with EAD is more diagnostic of the unique fluorotelomer betaine compounds and can provide more structural information.

Expanding Per- and Polyfluoroalkyl Substances (PFAS) Nontargeted Analysis Capabilities with Ion Mobility Spectrometry

Kaylie Kirkwood *and Erin Baker, North Carolina State University*

Per- and polyfluoroalkyl substances (PFAS) are a class of manmade compounds comprised of highly fluorinated aliphatic substances used in various household and industrial materials. Production of PFAS began in the 1940s, however traditional legacy PFAS have mostly been phased out of use due to their toxicity and environmental persistence. Manufacturers have therefore shifted to structurally modified alternatives which has led to the production of over 6000 known PFAS. Moreover, researchers are now tasked with both the quantitative monitoring and discovery of novel PFAS. Nontargeted approaches are often applied to this problem but can be subject to quantitation, separation, and data annotation challenges. Here we demonstrate the ability of liquid chromatography-ion mobility spectrometry-mass spectrometry (LC-IMS-MS) platforms to overcome many of these nontargeted analysis limitations via simultaneous hydrophobicity, size, and mass evaluations. Additionally, a discovery and targeted monitoring (DTM) approach, originally developed for proteomic studies, was optimized for PFAS assessment and applied to complex matrices including serum, liver tissue, and plant material to both quantify specific targets with matched internal standards and discover novel PFAS in a single sample analysis. A data independent acquisition scheme with ramped collision-induced dissociation was then developed to appropriately fragment all PFAS ions rather than a set of targets or the most abundant ions. Furthermore, the addition of IMS aided in size-based isomer and matrix separations, giving less noise and enhanced quantitative accuracy following filtering of matrix biomolecule signals. Challenges in unknown identification were also addressed via drift alignment of precursors and their corresponding fragments as well as the unique collision cross section (CCS) trends of PFAS subclasses.

Interactive Session: Workflows A

Predicting Compound Amenability with Liquid Chromatography Mass Spectrometry to Improve Non-Targeted Analysis

Charles Lowe[1]*, Kristin Isaacs*[1]*, Andrew D McEachrans*[2]*, Chris Grulke*[1]*, Jon Sobus*[1]*, Elin M. Ulrich*[1]*, Ann Richard*[1]*, Alex Chao*[1]*, John Wambaugh*[1] *and Antony Williams*[1]*, (1) U.S. Environmental Protection Agency, (2) Agilent Technologies, Inc.*

With the increasing availability of high-resolution mass spectrometers, suspect screening and non-targeted analysis are becoming popular compound identification tools for environmental researchers. Samples of interest often contain a large (unknown) number of chemicals spanning the detectable mass range of the instrument. In an effort to separate these chemicals prior to injection into the mass spectrometer, a chromatography method is often utilized. There are numerous types of gas and liquid chromatographs that can be coupled to commercially available mass spectrometers. Depending on the type of instrument used for analysis, the researcher is likely to observe a different subset of compounds based on the amenability of those chemicals to the selected experimental techniques and equipment. It would be advantageous if this subset of chemicals could be predicted prior to conducting the experiment, in order to minimize potential false positive and false negative identifications. In this work, we utilize experimental datasets to predict the amenability of chemical compounds to detection with liquid chromatography mass spectrometry (LC-MS). The assembled dataset totals 5,517 unique chemicals either explicitly detected or not detected with LC-MS. The resulting detected/not-detected matrix has been modeled using specific molecular descriptors to predict which chemicals are amenable to LC-MS, and to which form(s) of ionization. Random forest models, including a measure of the applicability domain of the model for both positive and negative modes of the electrospray ionization source, were successfully developed. The outcome of this work will help to inform future suspect screening and non-targeted analyses of

chemicals by better defining the potential LC-MS detectable chemical landscape of interest. *This abstract does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.*

Quantum Chemical Annotated Graphical Neural Networks for Mass Spectral Prediction
***Richard Overstreet**, Danielle Ciesielski and Ethan King, Pacific Northwest National Laboratory*
Prediction of the tandem mass spectrum for a given molecule is often accomplished via three generalized approaches: rules-based methods for bond breaking, deep learning, or quantum mechanical modeling. Rules-based approaches are often limited by diverse bonding environments and perform poorly for chemistries with few defined rules. Deep learning models often require a molecular fingerprint which generalizes chemical complexity resulting in poor prediction accuracy. Quantum chemical modeling is theoretically robust but requires significant amounts of computational time to produce a spectrum for a given molecule. In the current work we construct a neural network that encodes a graphical representation of molecular structure eliminating the reliance on molecular fingerprints. In addition, this approach allows for the integration of additional chemical features from quantum modeling. Target molecules are annotated using extended tight-binding (xTB) quantum chemistry methods for rapid calculation of graph features. Current results for peptides with [M+H] precursors show an average cosine similarity score of 0.77 across 10-30eV collision energy ranges. Current performance is better than a null model with an average score 0.27 and is an improvement over existing architectures.

Identification of Xenobiotic Metabolites Using In Silico Tools and Non-Targeted Analysis
***Matthew Boyce**[1], Kristin Favela[2], Alex Chao[1], Grace Patlewicz[1], Jon Sobus[1], Antony Williams[1] and John Wambaugh[1], (1) U.S. Environmental Protection Agency, (2) Southwest Research Institute*
Biomarkers of xenobiotic substance are often used to infer exposure and understand metabolic processes. Characterization of metabolic profiles associated with data poor or novel substances increasingly relies on non-targeted analyses (NTA), which aim to profile unknown compounds within a sample using high-resolution mass spectrometry (HRMS). A lack of spectral databases for metabolites makes structural identification of compounds difficult beyond formula-level assignments. In silico tools can be used to fill this data gap through the prediction of metabolite structures and fragmentation spectra. In this work, we used a suite of metabolite prediction tools (TIMES, Nexus Meteor, BioTransformer, and QSAR Toolbox) to prepare a suspect screening list to aid non-targeted analysis of 33 substances and their expected metabolites. A spectral database was also prepared for metabolites within the suspect screening list using Competitive Fragmentation Modeling (CFM-ID). The 33 starting substances were selected from the Environmental Protection Agency (EPA) ToxCast inventory and include pharmaceutical, agrochemical, and industrial chemicals. Metabolites were generated using a high-throughput assay with primary human hepatocytes, and cell lysates were analyzed via HRMS. Resulting NTA data were processed using a combination of vendor software and EPA's NTA web application. Processed data were further analyzed using custom python-based scripts to select clusters of features corresponding to metabolites of dosed substances. The analysis workflow was applied to six of the starting substances, and eight metabolites were identified across these substances with varying levels of confidence. The method outlined in this study provide a framework for using in silico tools to guide NTA of xenobiotic substances and their metabolites. This abstract does not necessarily reflect agency policy.

Interactive Session: Workflows B

Mass-Suite a Novel Open-Source Python Package Designed for HRMS Data Analysis
***Ximin Hu**, Derek Mar, Nozomi Suzuki, David Beck and Edward Kolodziej, University of Washington*
High-Resolution Mass Spectrometry (HRMS) has been widely used in chemical and biological analysis. Expanding from enhanced targeted analysis with precise results, non-targeted analysis utilized the data collection capacity of HRMS becomes more and more popular in recent years, especially in omics study and environmental analysis. However, comparing to the overwhelming data collection efficiency from the

instrument, the analysis pipeline of such HRMS data for water quality assessment is still in its infancy, with many basic aspects of data reduction, analysis, and interpretation requires optimization. Therefore, our team developed Mass-suite, a Python based open-source package that designed to better utilize HRMS data, especially for water quality assessment studies. The package provides flexible and various options to process the HRMS data: from basic functions, such as peak picking and data alignment; to advanced data analysis including statistical analysis, chemical fingerprint extraction or source quantification. Multiple algorithms were deployed within the package to enhance the data analysis performance including supervised machine learning and unsupervised clustering using Sci-kit learn Python package. The package also provides fully interactive access to the raw data, including visualization tools, formula prediction tools and MS2 online searching tool. Furthermore, the package developed in a modularized concept that different combinations of functions are available to accomplish desired tasks. The package has undergone tests on the feature detection coverage results in 99% detections out of a reference list of 400 known peaks from known environmental contaminants standard samples and proved to be work well with real world experimental dataset for non-target screening and prioritization purposes. Building up using the python core enables the utilization of cloud computational resource such as supercomputers, AWS and google colab, which frees up the space and cost of lab computers. By providing this package, we hope to open a new space for HRMS data analysis in environmental science field, resulting in more rapid and detailed studies in this area.

DEIMoS: An Open-Source Tool for Processing High-Dimensional Mass Spectrometry Data

Sean Colby*, Christine Chang, Jessica Bade, Jamie Nunez, Madison Blumer, Daniel Orton, Kent Bloodsworth, Ernesto Nakayasu, Richard Smith, Yehia Ibrahim, Ryan Renslow and Thomas Metz, Pacific Northwest National Laboratory*

We present DEIMoS: Data Extraction for Integrated Multidimensional Spectrometry, a Python application programming interface and command-line tool for high-dimensional mass spectrometry (MS) data analysis workflows that offers ease of development and access to efficient algorithmic implementations. Functionality includes feature detection, feature alignment, collision cross section (CCS) calibration, isotope detection, and MS/MS spectral deconvolution, with the output comprising detected features aligned across study samples and characterized by mass, CCS, tandem mass spectra, and isotopic signature. Notably, DEIMoS operates on N-dimensional data, largely agnostic to acquisition instrumentation; algorithm implementations simultaneously utilize all dimensions to (i) offer greater separation between features, thus improving detection sensitivity, (ii) increase alignment/feature matching confidence among datasets, and (iii) mitigate convolution artifacts in tandem mass spectra. We demonstrate DEIMoS with liquid chromatography–ion mobility spectrometry–tandem mass spectrometry (LC-IMS-MS/MS) data to illustrate the advantages of a multidimensional approach in each data processing step.

A Comprehensive Workflow for Compound Consolidation and Structure Annotation of Organic Pollutants in Non-Targeted Mass Spectrometry Analysis

Lee Ferguson *and Gordon Getzinger, Duke University*

High-resolution accurate-mass mass spectrometry (HR/AM MS) is capable of comprehensive structural elucidation of numerous trace-level organic contaminants in the environment. However, comprehensive sample characterization is made difficult by inefficiencies in the consolidation, annotation, and prioritization of detected features and postulated structures based on tandem MS analysis. We have developed a holistic, non-targeted analysis workflow to address these challenges, which combines ultra-high resolution LC-tandem mass spectrometry, cheminformatics, and computational mass spectrometry. This workflow utilizes custom and open-source algorithms to reduce redundancy and artifacts due to adducts, isotopes, and multimers in electrospray ionization through a feature consolidation step utilizing RAMClustR, followed by annotation of in-source fragment ions through automated comparison of data-dependent MS/MS spectra with retention time-grouped precursor spectra. These consolidation steps allow down-selection of molecular features to approach an ideal one-compound-one-feature paradigm for further annotation. Selected features are then passed to a

comprehensive spectral library assembled using custom scripts from MassBank, NIST, and mzCloud MS/MS libraries, providing spectra representing > 30,000 unique compounds. Custom-annotation of mass accuracy for each peak in each library spectrum allowed dynamic mass error thresholding during library matching. Further cheminformatic approaches are applied to validate spectral library matches among candidates, assess unexpected ion adduct formation, and report best confidence annotations. We have observed structure annotation rates of 10 – 20% of consolidated features in environmental samples using this approach. Finally, for compounds giving no spectral library matches, structure annotation from the PubChem database is accomplished via automated tandem mass spectral annotation and scoring (MetFrag CL, MAGMa, CSI:FingerID, and CFM-ID). A novel machine learning-based approach was developed and deployed to provide probability-based compound identification assessment from *in silico* MS/MS predictions, compound metadata (patents, references, and sources). This workflow is integrated into a distributable R package. We illustrate our approach using both formulated chemical mixtures (e.g. the EPA ENTACT collaborative trial samples) and extracts of environmental water and wastewaters.

NORMAN Digital Sample Freezing Platform - a European Virtual Platform to Exchange and Screen High Resolution-Mass Spectrometry Data

Nikiforos Alygizakis[1], *Nikolaos Thomaidis*[2], *Emma Schymanski*[3], *Reza Aalizadeh*[2], *Tobias Schulze*[4], *Juliane Hollender*[5] *and Jaroslav Slobodnik*[1], *(1) Environmental Institute, Slovakia, (2) National and Kapodistrian University of Athens, Greece, (3) University of Luxembourg, Luxembourg, (4) Helmholtz Centre for Environmental Research - UFZ, Germany, (5) Eawag - Swiss Federal Institute of Aquatic Science and Technology, Switzerland*

Environmental samples are nowadays routinely analyzed by liquid and gas chromatography coupled to high resolution-mass spectrometry (HRMS) detectors from various vendors (e.g. Thermo Fisher Scientific, Agilent Technologies, Bruker, Waters, AB Sciex). A novel platform for archiving and processing raw HRMS data was established by the NORMAN network (www.norman-network.net) to enable inter-comparison of results and increase the level of exploitation of information from HRMS chromatograms. The platform was termed 'NORMAN Digital Sample Freezing Platform' (DSFP) as it allows for retrospective screening of the 'digitally stored' samples. The DSFP enables fast and effective searching of thousands of substances known or suspected to be present in the environment registered in NORMAN Substance Database (SusDat) and even unknowns across many samples and different matrices. In DSFP, a standardized workflow allows the evaluation of raw mass chromatograms from any HRMS instrumentation vendor after they are converted into the standardized open mzML format. Mass spectral and chromatographic (normalized retention time) information on hundreds up to thousands of components typically contained in each environmental sample is then extracted into spreadsheet-based standardized Data Collection Templates (DCTs). Such 'digitalized' samples can then be retrospectively screened for a presence or absence of virtually any compound detectable by a given LC- or GC-HRMS technique. Concentrations of identified substances can be estimated via a semi-quantification algorithm based on their structure similarity with known chemicals analyzed under the same conditions. The DSFP was thoroughly tested with data obtained from various European and international projects. Examples of its successful application will be shown on several datasets, including the Joint Black Sea Surveys (JBSS; EU/UNDP EMBLAS-II; http://www.emblasproject.org/), wastewater effluents (SOLUTIONS; http://www.solutions-project.eu/), data from top predators and their prey (LIFE APEX; https://lifeapex.eu/) and Joint Danube Survey 4 (JDS4; http://www.danubesurvey.org/jds4/about). The latest updates, underlying informatic technologies and key non-target screening functionalities will be presented. The potential of the platform to become a global tool for storing and processing environmental HRMS data will be discussed.

Acknowledgements Authors would like to acknowledge NORMAN network for funding development of DSFP through the NORMAN Joint Programme of Activities and European Union for its support through the project LIFE17 ENV/SK/000355 "Systematic use of contaminant data from apex predators and their prey in chemicals management" (APEX).

Method and Performance Evaluation

Capabilities and Performance of Nontarget and Suspect Screening Methods: Lessons Learned From (Collaborative) Comparison Studies

***Martin Krauss**, Helmholtz Centre for Environmental Research - UFZ, Germany*

Within the last 15 years, the use of nontarget and suspect screening has gained momentum in the research community and a considerable number of different analytical methods and data processing workflows have been developed. A range of studies compared analytical and data processing approaches, either looking at individual steps (e.g., different peak detection algorithms, ionization techniques), or a more general picture using large-scale collaborative trials such as those organized by NORMAN (indoor dust, water), USEPA (ENTACT) or the German Water Chemistry Society. This presentation summarizes the main outcomes and implications of these studies on current method capabilities and performance: The compound domain covered is an essential characteristic of an analytical method, depending on the compound domains covered by each individual procedural step. While first approaches to characterize this domain are only recently emerging, it is feasible to assess it using a set of target compounds with a wide range of physical-chemical properties and diversity of functional groups. A reliable and comparable suspect screening is feasible, as almost all laboratories in interlaboratory studies are able to detect and confirm a large fraction of the candidate compounds if a common suspect list is used. For a true non-target screening, the variability of results (numbers of detected compounds, levels of identification confidence) among laboratories and workflows is huge. This depends not only on the choice of analytical settings discriminating against specific compound domains, in rare cases poor instrument performance, but to a large extent on data processing workflows (peak detection and alignment of different samples). Data processing requires a thorough development and validation along with the analytical method, as the appropriate choice of settings has a huge impact on the results. A strict QA/QC and performance evaluation and quality control scheme is essential, including all steps of the workflow using spiked and unspiked quality control samples and internal standards, as these can used to monitor sample processing performance (as for classical target analysis), instrumental performance (retention time stability, mass accuracy, matrix effects) and performance of peak detection algorithms. A recent large-scale pesticide suspect screening study within the HBM4EU project, in which five laboratories analyzed each 400 urine samples, shows that based on harmonized analytical methodologies and well-established quality control procedures consistent and reliable data sets covering 2000 samples can be generated among different laboratories.

Approaches for Assessing Performance of High-Resolution Mass Spectrometry-Based Non-Targeted Analysis Methods

***Katherine Peter**[1], Christine Fisher (O'Donnell)[2], Seth Newton[3], Andrew Schaub[4] and Jon Sobus[3], (1) University of Washington, (2) U.S. Food and Drug Administration, (3) U.S. Environmental Protection Agency, (4) Southwest Research Institute*

Non-targeted analysis (NTA) using high-resolution mass spectrometry has enabled the detection and identification of unknown and unexpected compounds of interest in a wide range of sample matrices. Despite the benefits of NTA methods, standardized procedures do not yet exist for assessing performance, limiting stakeholders' abilities to suitably interpret and utilize NTA results. This is particularly challenging for regulatory stakeholders that are accustomed to interpreting the results of targeted analyses. Accordingly, members of the Benchmarking and Publications for Non-Targeted Analysis (BP4NTA) working group have undertaken an effort to understand current approaches for performance assessment and lay the groundwork for further development of standardized procedures. As a reference point, existing performance assessment metrics for targeted analyses will be summarized to provide context and clarify terminology that may be shared between targeted and NTA performance assessment methods (e.g., accuracy, precision, sensitivity, and selectivity). To structure the discussion of NTA performance assessment methods, three types of NTA study objectives are defined: sample classification, chemical identification, and chemical quantitation. For each of these study objectives, promising approaches for assessing performance will be presented, including strengths, key caveats,

and areas in need of further development. For qualitative study outputs (i.e., focusing on sample classification and/or chemical identification) the traditional confusion matrix is employed, with some challenges and limitations, particularly related to bounding the confusion matrix. For quantitative study outputs, performance can be assessed using estimation procedures developed for targeted methods, although consideration of additional sources of uncontrolled experimental error is necessary. By laying out these performance assessment approaches, we intend to stimulate discussion and further efforts to develop and improve procedures for assessing NTA method performance. Ultimately, improved performance assessments will enable accurate communication of NTA results and actionable utilization of NTA data by stakeholders. This abstract does not necessarily reflect agency policy.

Evaluation of Semi-Quantification Methods in Suspect Screening of Water Samples Using Liquid Chromatography Electrospray Ionisation High Resolution Mass Spectrometry

Louise Maria Eleonora Malm *and Anneli Kruve, Stockholm University, Sweden*

The surface waters of the world are continuously polluted by an abundance of chemicals, which may degrade to more toxic or persistent transformation products. Most commonly, liquid chromatography electrospray ionisation high resolution mass spectrometry (LC/ESI/HRMS) is used for analysis, with increasing use of non-targeted or suspect screening workflows. Due to the vast variation in response between compounds in ESI, the need for robust semi-quantification approaches is of importance.Here, five semi-quantification methods are tested and evaluated in a NORMAN network organised collaborative trial with 45 participating laboratories.Three methods use the response factor of the (1) structurally most similar calibrant; (2) parent compound (only applicable for parent compound – transformation product pairs); and (3) closest eluting calibrant to estimate the concentration of the suspect compound. Two methods predict the ionisation efficiency of the suspect compound using (4) a random forest model based on molecular and eluent descriptors; and (5) quantitative structure-property relationship modelling with molecular descriptors. The predicted ionisation efficiency is related to a predicted response factor, which is used to semi-quantify the suspects.Briefly, water samples spiked with 45 suspect compounds are analysed on different LC/HRMS instruments together with 41 calibrants at six known concentrations. The semi-quantification methods are applied, and the estimated concentrations are compared with the real concentrations to obtain the prediction error.Preliminary results from one instrument show that methods using predicted ionisation efficiencies yield the lowest errors, with a median error of 2.72× and 2.42× for approach (4) and (5), respectively. Corresponding mean errors are 6.68× and 30.75×. For approach (1), (2) and (3), 75%, 72% and 63% of the compounds are predicted with less than 10× error, while for approach (4) and (5) these percentages are 86% and 88%, respectively. Comparison of results between laboratories and instruments are underway, with preliminary results expected by the time of the conference.The results from this study are expected to provide insights in the performance of existing semi-quantification methods across instruments and laboratories, and to identify limitations with them, as a first step towards standardisation of semi-quantification methods.

The Use of Non-Targeted Analysis for Rapid and Emergency Response: Demonstration Through Mock Scenarios

John Sloop*, Alex Chao, Jennifer Gundersen, Allison Phillips, Jon Sobus, Elin M. Ulrich, Antony Williams and Seth Newton, U.S. Environmental Protection Agency*

The U.S. Environmental Protection Agency (EPA) is mandated by the Executive Branch of the federal government to respond to intentional and unintentional chemical releases. Several thousand releases occur annually in the U.S., with the contents of many releases being of unknown composition. Existing rapid response workflows typically rely on low-resolution screening methods and targeted analytical approaches that can only measure a limited number of compounds (100s – 1000s). When targeted methods are unable to identify the chemicals present, alternate approaches such as non-targeted analysis (NTA) can be used in conjunction with expert analysis of the available data to reach conclusive determinations; however, the analysis required to reach those determinations at a confident level is a very time-consuming process. With recent advances in high-

resolution mass spectrometry (HRMS) instrumentation and cheminformatics tools comes the potential for a more expedited process of chemical identification. To demonstrate this potential, we have designed multiple "mock scenarios" that mimic real-world rapid response situations involving one or more unknown chemicals, including various examples, such as the release of a chemical warfare agent and an accidental industrial spill. We have developed NTA methods and workflows that utilize advanced HRMS instrumentation and cheminformatics tools to identify the unknown chemical(s) in a timeframe that is relevant and useful to responders (typically 24-48 hours post-release). In each of these scenarios, one analyst is charged with scenario design and generation of the real-world sample. A second analyst (representing a scientist responsible for identification of the unknown chemicals) receives details of the mock scenario (i.e., observations that would be made in the field during an actual scenario), and is responsible for carrying out sample preparation, data acquisition, and data analysis, to identify the unknown chemical(s) in the sample. Using the developed NTA workflows, we have been able to identify the most important chemicals of interest in each of these designed mock scenarios in a rapid manner. The results strongly support the use of HRMS NTA workflows in rapid response scenarios, especially when unknown stressors need timely and confident identification. This abstract does not reflect Agency policy.

Effects of Different Detection Algorithms on Non-Target Analysis Results

Bastian Schulze[1]*, Amy Heffernan*[2]*, Cameron Veal*[3]*, Kevin Thomas*[1]*, Saer Samanipour*[5] *and Sarit Kaserzon*[1]*, (1) University of Queensland, Australia*, *(2)Eurofins Environment Testing Australia & New Zealand, Australia, (3) Seqwater, Afghanistan, (4) University of Amsterdam, Nederland*

Non-target analysis (NTA), used as a comprehensive approach to characterize unknown chemicals, including chemicals of emerging concern has seen a steady increase in application recently. Given the relative novelty of this type of analysis, new approaches get developed regularly. However, given the complexity and multi-step approach required for NTA, even minor changes to a non-target workflow can substantially influence the result reliability and reproducibility. This does not only include obvious changes in methodology like the type of sampling, chromatographic or spectrometric settings, but also how the data is processed afterwards. For all of these steps it is important to know how exactly they influence the final results. One way to compare different data processing approaches is to process the same sample using different workflows and compare the similarities and discrepancies between the results. However, this ignores the different sub steps that are usually combined within the data processing, like peak picking/detection, componentization, filtering, prioritization and identification. Usually, every software has different algorithms for these steps, making it difficult to know where differences in the results originate from. Therefore, in this study we investigate how the peak picking in particular influences the overall non-target results and their interpretation, using a real-world dataset. Samples were acquired on a bi-annual basis from 2014 to 2019 at several sites across South East Queensland, Australia and can therefore be used to investigate temporal and spatial trends. Samples were run on a generic LC-HRMS method. Peak detection was done using 4 different algorithms, two vendor supplied (SciexOS and MarkerView) and 2 open source algorithms (In House and MSDial) with parameters set as comparable as possible. The obtained feature lists were then filtered in Excel, before submitting it to statistical analysis. Using these results, we examine and discuss how the overall conclusion regarding spatial and temporal trends, together with results for individual samples changes.

A Quality Control Standard Mixture for Assessing Non-Targeted Method Performance: Development, Implementation, and Efforts Toward Commercialization

Christine Fisher (O'Donnell), Jacob Premo and ***Ann Knolhoff****, U.S. Food and Drug Administration*

Measuring the quality of non-targeted analysis (NTA) and suspect screening analysis (SSA) data and results is critical for ensuring accurate and reliable interpretations. However, evaluating NTA/SSA method performance is challenging given the vast chemical space to consider. A limiting factor is the lack of commercially available standard mixtures covering the breadth/diversity of physicochemical properties encountered in NTA/SSA. Thus, we developed a non-targeted standard quality control (NTS/QC) mixture to assess liquid

chromatography/high resolution-mass spectrometry (LC/HR-MS) method performance. The mixture contains 89 compounds covering a broad range of molecular weights (126-1111 Da), $\log K_{ow}$ values (-8.08 to 8.51), diverse elemental compositions (C, H, N, O, S, P, Cl, Br, F), and compound classes (e.g., pesticides, pharmaceuticals, toxins, etc.) that ionize in positive and/or negative ion polarities by electrospray ionization. We used this mixture to evaluate our NTA/SSA workflow including sample preparation, data quality, and data processing. We also determined appropriate thresholds for reducing potential false positives based on mass accuracy errors, isotopic fit, variability, generated molecular formula ranking, and mzCloud match scores. This work also highlighted areas of improvement, including retention of polar analytes and molecular formula assignment for fluorinated compounds. Although incredibly useful, preparing such a large standard mixture is time-consuming, presenting a significant barrier to broader implementation. To work toward a commercially available standard for the assessment of LC/HR-MS NTA method performance, we have conducted an electronic poll to gage interest in the community and to further develop the standard to meet the needs of different NTA/SSA applications. Preliminary feedback indicates a significant interest/need for a standard like this. Although, some preferences differ among researchers (e.g., number of compounds, inclusion of labeled compounds), the current range of physicochemical properties included in our NTS/QC mixture would fit the needs of a majority of those using reversed-phase LC/ESI-HR-MS. We hope to obtain additional feedback from NTA practitioners at the SETAC NTA conference and continue to evaluate potential modifications to the mixture to enable broader adoption in the NTA community. We are also communicating with companies to determine routes for beta-testing and commercialization.

Method and Performance Evaluation (Poster)

A Novel Approach to Evaluating Disease Risk Through the Incorporation of Lipidomic Data

***Jessie Chappel**, Erin Baker and David Reif, North Carolina State University*

A key component in the early diagnosis of diseases or possible prevention is identifying the risk an individual carries for a certain condition or begins to molecularly express throughout life. To accomplish this, many biomedical researchers have taken to developing risk scores, which use data such as genetic information, geographical location, and previous/current health-related conditions to estimate an individual's level of risk for other conditions and diseases. Optimizing the predictive power of risk scores has however proven to be quite cumbersome since many diseases exhibit various molecular changes, and thus mandate the integration of many diverse datatypes. Risk score development has primarily focused on utilizing genetic data, discounting the early molecular changes that are often observed by lipidomic and metabolomic alterations. As a result, it is essential to further explore the predictive power and determine how to best calculate risk based on these datatypes alone before combining them with other data sources. To this end, a novel method for deriving risk scores based on lipidomics was explored by evaluating various panels of significant lipids in different case versus control comparisons. Numerous weighting functions and approaches were therefore evaluated and optimized to explore how well the magnitude of the scores correlated with each condition studied. Here, we illustrate the strengths and limitations of each scoring scheme and the potential of how integrating lipidomics with other datasets can produce a more comprehensive risk score.

A Novel Floatation-Based Sample Processing Method for Raw Wastewater Testing

***Casey Wegner**[1], Douglas Sieglaff[2], Carolina Livi[1], Jon Roussey[1] and Stephan Baumann[2], (1) Akadeum Life Sciences, (2) Agilent Technologies, Inc.*

From an analytical perspective, wastewater is an atypical matrix to process and analyze. Starting from EPA method 1615, many wastewater testing labs have protocols that perform reasonably well but are very labor intensive and require large volumes of wastewater to achieve target sensitivity levels. Filtration based workflows challenges include throughput and requirement of time-consuming pasteurization and centrifugation processes for wastewater samples. Ultrafiltration also has the downside of removing RNA-rich solids to avoid clogging filters. With this novel microbubble approach, we look to simultaneously improve ease-of-use,

sensitivity, and method robustness starting with processing smaller volumes of wastewater using functionalized microbubbles to extract RNA for qPCR analysis. Specifically, using nucleic acid binding, buoyant, functionalized microbubbles to extract nucleic acids enables the analysis of large volumes of wastewater without the need for pre-processing such as the removal of solids from sludge. The efficiency of this approach allows over 600ng of RNA and over 900ng of DNA to be extracted from as little as 3mls of wastewater. When combined with RNA binding columns and qPCR this workflow assesses requires significantly less sample volume and receives equivalent to increased in sensitivity compared to commonly practiced wastewater processing methods. This approach also removed the need for paste pasteurization as the lysis buffer is added directly to the raw wastewater. Moreover, microbubbles float on top of the extraction tube which makes removal of the subnatant cumbersome when using a traditional tube with Pasteur pipettes. To address this inconvenience a custom 50ml microbubble separation tube containing an integrated stopper was developed for wastewater workflows. This custom microbubble tube eliminates the need for Pasteur pipette fluid transfer steps, which in turn enhanced the sensitivity of Sars-CoV2 N1 and N2 gene detection by 2.9 and 4.0 Cts respectively. A microbubble functionalized to bind virus particles or bacteria directly could also be incorporated into the workflow enabling a multi-modal testing method when combined with culture or molecular assays.

Automated Near Real Time High Resolution Mass Spectrometry Quality Assurance
***Tobias Bader**, Wolfgang Schulz and Rudi Winzenbacher, Landeswasserversorgung, Germany*
Nontargeted approaches based on liquid chromatography high-resolution mass spectrometry (LC-HRMS) are widely used in different fields of analytical chemistry. Despite the widespread application of HRMS, sufficiently harmonized and recognized quality assurance (QA) practices are scarce. While QA is a key aspect to control and maintain data quality, it is often perceived a burdensome and time-consuming process. The objective of this work was thus to implement an automated near real time QA procedure to control the HRMS instruments in our lab. For all instruments, a recalibration by infusion of a reference solution is performed after every forth LC-HRMS run. After acquisition, these calibration files were automatically processed and used to extract time series data on i) mass accuracy, ii) resolving power and iii) signal intensity for both MS1 and MS2. Based on historical data, we set upper and lower thresholds for each of these parameters. Moreover, the standard deviation within the current batch was found to be a valuable problem indicator. If any tolerance limits were exceeded, the user is directly notified via email, which allows a timely initiation of measures. This approach leads to a reduction of the manual workload, while still maintaining a comprehensive QA. We could show that the automated QA allows the sensitive recognition of deviations, variations and creeping trends whereas the thorough compilation of QA data strongly supports the identification of problematic areas and is a great help for service engineers.

Comparing and Harmonizing Results of Non-Target Workflows for Environmental Matrices
***Brittany Saleeby**, University of California*
There are a variety of software available to analyze non-target datasets from privatized companies and open-source developers. In general, the software transforms raw high-resolution mass spectrometry (HRMS) data into lists of non-target molecular features and compounds. Majority of these software are developed for "-omic" applications (ex. metabolomics), so the abilities of these software to perform their tasks for environmental matrices are largely untested. As a collaboration between labs at UC-Davis and Duke University, this project objective aims for harmonizing and comparing non-targeted workflows for environmental matrices by the analysis of two sets of HRMS data through different open-source software. Raw data files were transferred between the two data repositories using file transfer protocol software. These files were converted from vendor-specific formats into open-source data formats required for the different workflows. Both groups applied procedures to the datasets that remove adducts, isotopes, and contamination or other artifacts. To validate the successful conversion of data files, compounds used for analytical QA/QC from the originating lab were qualitatively identified within datasets using R studio. This first step also probes the ability of spectral libraries to correctly match to the standard's spectra. To further probe the spectral library's ability, R studio was also

used to create list of matching and non-matching features between both datasets. When a feature is matched to a known compound in the spectral library, commonly known as suspect screening, both software will assign the appropriate InChiKey to the feature. The list of merged InChiKey highlights the similarities and differences between each software's ability to suspect screen and assign identity to a molecular feature. The remaining features that have not been identified are considered the non-target features. The features are considered matched under the criteria that the m/z of the two features are within 3 ppm and the retention time is within 0.9 minutes. Using R-Studio, thousands of features from the individual datasets are matched to the thousands of features that were generated by the other lab's software. The resulting dataset are matching non-target features, which are vetted further for spectral similarity. The ability of each workflow to produce molecular formulas will also be considered and analyzed in the same manner as the suspect and non-target features.

A First Guideline Towards Standardization of Nontarget Screening

*Wolfgang Schulz[1], Thomas Lucke[1] and **Christian Zwiener**[2], (1) Landeswasserversorgung, Germany, (2) University of Tübingen, Germany*

The application of nontarget screening (NTS) with liquid chromatography high resolution mass spectrometry (LC-HRMS) in water analysis has spread widely within the last few years. Against this background, an expert committee (EC) was established to summarize the experiences with NTS in a guideline, which is freely available at https://www.wasserchemische-gesellschaft.de/images/HAIII/NTS-Guidline_EN_s.pdf. The guideline's structure predominantly reflects that of analytical methods. Nontarget analysis has always to start with the definition of the problem or task which also has an impact on the sampling strategy. Attention should be paid to suitable blanks since any contact of materials and chemicals with the water sample may result in sample contamination. Thus, adequate blanks should cover all aspects of the analytical procedure from sampling, transport, sample preparation in the lab to instrumental measurement. The reproducibility of LC-HRMS measurements should be continuously monitored and documented with a suitable quality assurance and is crucial for the long-term comparability of LC-HRMS data. For data interpretation and validation of results additional information (metadata) on all aspects of the sample may be helpful. The comparability of analysis results across different instrumental platforms and data evaluation software is a prerequisite for the establishment and successful application of NTS. For this purpose, the EC organized two comparative tests in the form of collaborative trials. Different Time-Of-Flight and Orbitrap systems were used by 18 participating laboratories. In the first trial sensitivity and mass accuracy of defined reference standards in ultrapure water were considered. In the second trial, the complete laboratory specific NTS workflow was applied to spiked river water samples (25 ng/L and 500 ng/L) and aimed at assessing the challenges and limits of detection and identification of unknowns in real matrices. The results are summarized in the guideline.

Exploring Novel Relationships Between Persistent Organic Pollutants in Environmental Analyses Using Ion Mobility Spectrometry-Mass Spectrometry

***James Dodds**, Anna Boatman, Detlef Knappe, Astrid Schnetzer and Erin Baker, North Carolina State University*

Persistent organic pollutants (POPs) are an ensemble of man-made chemical compounds that are resistant to degradation and known to cause adverse outcomes in human health, wildlife, and the environment. POPs are comprised of several biochemical classes including pesticides, PCBs, and PFAS which are produced for a variety of industrial, agricultural and household applications. The chemical persistence and adverse effects of POPs have resulted in environmental regulation efforts and decreased production, yet new, unregulated replacement compounds pose a continual challenge to environmental monitoring. Here we evaluate the efficacy of a low-cost passive sampling device (Solid Phase Adsorption Toxin Tracking, or SPATT) to simultaneously monitor pesticides, PFAS, and cyanotoxins coupled with appropriate analytical detection methods, such as ion mobility spectrometry-mass spectrometry (IMS-MS). The IMS-MS platform utilized in this work (Agilent 6560) has been extensively characterized in previous publications evaluating PFAS, pesticides, and other pollutant classes. These studies highlight the unique advantages afforded by LC-IMS-MS pertaining to spatial

resolution of halogenated species from traditional biomolecular classes as a consequence of their relatively large mass compared to overall size (*e.g.* fluorine vs. hydrogen atoms in PFAS and chlorine in pesticides). While quantitative methods for PFAS, pesticides, and cyanotoxins are well established, the unique challenges addressed in this study focus on two primary aims: 1) evaluate the overall utility of SPATT samplers to comprehensively assess multiple POP classes and 2) establish a generalized extraction method for SPATTs to characterize multiple POP classes and cyanotoxins concurrently. While previous studies have correlated high pesticide concentrations in agriculture with proliferation of harmful algal blooms, the relationship between these molecules and PFAS contamination is unknown. Provided the methods established herein are applicable to comprehensively assess POPs from a single passive sampler, evaluation of these field samplers endeavor to uncover the unique intricacies of pollutant mixtures in the environment and their biological consequences on wildlife.

GC/MS Target Screening Method for 176 Agricultural Chemicals in Drinking Water Samples
Norihiro Kobayashi[1]*, Yuko Tsuchiya*[1]*, Sokichi Takagi*[2] *and Yoshiaki Ikarashi*[1]*, (1) National Institute of Health Sciences, Japan, (2) Osaka Institute of Public Health, Japan*
Water suppliers in Japan ensure the safety of drinking water by monitoring the levels of agricultural chemicals in raw and tap water in Japan. However, some of the official Japanese analytical methods regarding agricultural chemicals in drinking water by the Ministry of Health, Labour and Welfare (MHLW) has complicated analytical procedures, and a more practical alternative method is required. This study aimed to develop a GC/MS target screening method for the measurement of agricultural chemicals in drinking water samples. The target compounds included 176 agricultural chemicals listed in the Japanese Waterworks Act. Standards solutions (0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, and 2 mg/L) of all the agricultural chemicals were measured using GC/MS, and. retention times, mass spectra, and calibration curves of each agricultural chemical were collected in a database. Then, a total of 108 river water samples and 5 ground water samples were collected from 21 prefectures in Japan between May and September 2018. The sample preparation method is stipulated in the official Japanese analytical methods was closely followed, which involved concentrating a 500 mL of water samples to 1 mL using solid phase extraction. The GC/MS analysis detected 52 agricultural chemicals in the river water samples and no agricultural chemicals in the ground water samples. The concentrations of agricultural chemicals in the river water samples determined using the GC/MS target screening method and the official MHLW analytical methods were compared. The levels of agricultural chemicals measured using the GC/MS target screening method varied between half and twice the level detected using the official MHLW analytical methods. The GC/MS target screening method was sufficiently accurate for use as an alternative to the official MHLW analytical methods for agricultural chemicals in drinking water. We will conduct the validity test of the GC/MS target screening method with several examination organizations to promote its acceptance as a new standard analytical method in Japan.

Hydrophilic Lipophilic Balanced Retainability in the Context of Chemical Space
Anna Feerick*, Gabrielle Black and Thomas Young, University of California*
Hydrophilic Lipophilic balanced (HLB) cartridges are widely used in solid-phase extraction for sample cleanup and concentration for nontarget analysis. Their ability to retain a diverse set of compounds while simultaneously removing interferences makes them a preferred step in many workflows. Despite the diverse set of compounds that HLB is capable of recovering, there are inevitably many that are not successfully retained. The boundaries of chemical space, i.e., the set of known and possible compounds, covered by HLB extraction remains undefined, limiting the identification confidence of suspect and unknown contaminants annotated during nontarget analytical workflows. Defining the "*HLB detectability domain*", an area of chemical space where HLB is capable of extracting versus not based on a multitude of molecular descriptors, is crucial for improving the confidence of feature annotation. We propose the use of computational machine learning-based models to predict the detectability domain of nontarget methods by examining the extraction potential of HLB sorbents. For this purpose, a dataset from a previous study consisting of 414 priority and emerging pollutants in water

was used. One- and two-dimensional molecular descriptors for each compound were obtained through PaDeL. Both classification and regression algorithms were evaluated for their potential in developing a quantitative structure-property relationship (QSPR) model that will describe which molecular descriptors most affect HLB retainability. The classification methods include classification and regression trees (CART) and genetic algorithm (GA)-support vector machines. Both linear (GA-multiple linear regression) and non-linear (GA-support vector regression) regression methods were employed. Knowing the range of compounds that can be readily recovered by HLB resins, confines nontarget annotations to this area of chemical space. Understanding the bounds of HLB extraction (nontarget space) is an important step in nontarget method standardization and reproducibility, both of which are necessary for improving regulatory acceptance of the methods for environmental monitoring and assessment.

Non-Targeted Analysis Study Reporting Tool: A New Framework to Improve Reproducibility and Transparency

*Katherine Peter[1], Allison Phillips[2], Piero Gardinali[3], Ann Knolhoff[4], Carlos Manzano[5], Kelsey Miller[2], Manuel Pristner[6], Lyne Sabourin[7], Mark Sumarah[7], Benedikt Warth[6] and **Jon Sobus**[2], (1) University of Washington, (2) U.S. Environmental Protection Agency, (3) Florida International University, (4) U.S. Food and Drug Administration, (5) San Diego State University, (6) University of Vienna, Austria, (7) Agriculture and Agri-Food Canada (AAFC), Canada*

High resolution mass spectrometry (HRMS) and non-targeted analysis (NTA) methods have broadened the chemical lens through which researchers examine complex samples. Despite increasing refinement of instrumentation and workflows, universally accepted reporting standards for NTA studies have yet to be realized. Currently, proposed benchmarks address only specific elements of NTA reporting – most notably, confidence in compound identification. While critically important, such guidance is limited in scope (relative to an entire NTA workflow) and therefore insufficient to ensure scientific transparency and reproducibility. To address the need for standardized reporting criteria, the Benchmarking and Publications for Non-Targeted Analysis (BP4NTA) working group developed the NTA Study Reporting Tool (SRT), the first easy-to-use approach for rigorous evaluation of NTA reporting practices. The SRT is organized by NTA study chronology and contains 13 sub-categories for scoring that cover all aspects of study design; data acquisition, analysis methods, and outputs; and quality assurance/quality control metrics. To test the SRT, eleven NTA practitioners applied it to evaluate the quality of reporting in eight published manuscripts covering environmental, food, and health-based exposomic applications. Results highlighted NTA areas where current reporting practices need significant improvement and demonstrated that the SRT provides a valid framework for evaluating NTA reporting quality. In fact, 70% of scores self-assigned by authors of the evaluated manuscripts fell within the range of peer-assigned scores, indicating that future use of the SRT will further strengthen reporting practices. Reviewer feedback directed an evolution of the SRT to resolve ambiguity and coverage gaps. It further helped refine a final scoring system that provides an accessible and objective appraisal of overall study reporting quality. Widespread implementation of the SRT is anticipated to improve the efficiency and rigor of NTA study design and review, and ultimately, lend credence to NTA applications within both research and regulatory arenas. *This abstract does not necessarily reflect agency policy.*

Nontargeted Analysis of Per- and Polyfluoroalkyl Substances

Novel Per- and Polyfluoroalkyl Substances in a Current-Use C6-based Aqueous Film Forming Foam Formulation

***Jennifer Guelfo**[1], Marzieh Shojaei[1], Abigail Joyce[2] and Lee Ferguson[2], (1) Texas Tech University, (2) Duke University*

Studies have identified hundreds of per- and polyfluoroalkyl substances (PFAS) in aqueous film forming foam (AFFF) using high resolution mass spectrometry (HRMS), and there is increasing reliance on HRMS with suspect screening research in areas such as PFAS occurrence and fate and transport. Characterization of

contemporary AFFF formulations is crucial for maintaining representative lists of candidate PFAS for such efforts. The objective of this study was to characterize an AFFF currently certified for use by the U.S. military. Targeted analysis, total oxidizable precursor assay (TOP), HRMS with suspect screening, and non-targeted analysis (NTA) were used to determine PFAS composition and concentration. The sum of PFAS identified during targeted analysis and suspect screening was compared to TOP, demonstrating that >90% (20.2 mM) of the estimated total PFAS concentration (22.4 mM) was comprised of "unknown" PFAS. Tandem and multi-stage tandem mass spectrometry spectra were used to annotate 10 PFAS within 9 classes, 8 of which have not been previously reported. A subset of these compounds are isomers of legacy PFAS of different manufacturing origin. Thus, suspect screening efforts that rely solely on accurate mass matching may mis-annotate PFAS presented here as isomers that will have key differences in properties such as biotransformation pathways. The total estimated concentration of the 10 PFAS was ~20 mM, consistent with the "unknown" fraction identified by TOP.

Tracking Biological Transformation of Polyfluoroalkyl Substances in Commercial Mixtures Using In Vitro Assays and High-Resolution Mass Spectrometry

*David Dukes[1], Jeremy Koelmel[2], Paul Stelben[2], Joseph Okeme[2], Emily Parry[3], Tarun Anumol[3], Krystal Pollitt[2] and **Carrie McDonough**[1], (1) Stony Brook University, (2) Yale University, (3) Agilent Technologies, Inc.*

Per- and polyfluoroalkyl substances (PFAS) are in the blood of virtually all (>99%) of Americans. This is a serious public health threat, as exposure to PFAS has been linked with myriad health problems, including kidney and testicular cancer, high cholesterol, liver toxicity, immunosuppression. Well-characterized PFAS like perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA) are a major focus of current research due to their widespread presence in human blood. However, the majority of PFAS currently used in consumer products are pre-PFAAs, a sub-class of PFAS that can be transformed via environmental and biological processes to ultimately form persistent perfluoroalkyl acid (PFAA) “forever chemicals.” The vast majority of diverse pre-PFAA structures are not available as neat standards and have undergone no toxicological evaluation. To understand the implications of direct human exposure to pre-PFAAs, information is urgently needed on the occurrence of pre-PFAAs in consumer products and the extent to which pre-PFAAs are transformed *in vivo*, acting as an internal source of known toxic PFAAs as well as novel transformation products. Here we will demonstrate a novel high resolution mass spectrometry (HRMS) workflow incorporating FluoroMatch 2 for screening complex PFAS mixtures before and after digestion with *in vitro* metabolism assays to describe biologically relevant pre-PFAA transformation. Novel transformation products identified via this workflow are entered into an HRMS spectral library to improve suspect screening coverage for human serum and urine. In our preliminary experiments using mouse liver S9 sub-cellular fractions and an historical aqueous film-forming foam (AFFF) mixture, we observe substantial formation (log fold-change > 2) of branched and linear PFHxS and 6:2 FTS after a 60-minute incubation, as well as potential formation of several novel substituted PFAAs, some of which have been identified as bioaccumulative in our previous human biomonitoring and animal dosing studies. Here we will present compositional changes in complex PFAS mixtures occurring during *in vitro* biological transformation and use these *in vitro* findings to inform screening of biological tissues and fluids.

Non-Target Screening Approaches for Poly- and Perfluorinated Alkyl Substances (PFAS) in Contaminated Soil Samples

***Boris Bugsel**, Jonathan Zweigle, Markus Schmitt and Christian Zwiener, University of Tübingen, Germany*

The class of poly- and perfluorinated alkyl substances (PFAS) comprises more than 3000 compounds which have a broad application area, from industrial processes to consumer products. For example, PFAS are used to produce paper and cardboard with water and grease repelling properties. Paper sludge from impregnated paper products was presumably applied on agricultural soils during a period of ten years and caused contamination of several millions of square meters on a site in southwest Germany. Since no information is available on the identity of the contaminants, analytical screening approaches have to be used to characterize the contamination.

Original PFAS chemicals and their transformation products have been identified in soil samples by liquid chromatography-high-resolution mass spectrometry (LC-HRMS) screening approaches. Data analysis comprised Kendrick mass defect analysis and systematic retention time shifts to assign homologous series and matching with a suspect list. LC-HRMS analysis of contaminated soil samples identified a total of 65 individual substances from 13 different substance classes. In particular, the substance class of disubstituted perfluoroalkyl phosphates (diPAPs) as well as their transformation products could be identified as major contaminants. PFAS distribution patterns in contaminated soil samples have further been compared with PFAS distribution patterns in impregnated papers collected at about the same time as the contamination occurred. The results confirm the hypothesis that impregnated papers are major sources of the contamination. Environmental processes have been mimicked by photochemical and electrochemical experiments which demonstrated that commercially used PFAS are slowly converted to transformation products (TPs, e.g., perfluoroheptanoic acid from 6:2 diPAP) which are affecting the groundwater quality in the long term.

Hunting the Missing Fluorine: Nontarget Identification of PFAS in Aqueous Film Forming Foams

Dunping Cao[1]*, Jennifer Field*[1]*, Thierry Fouquet*[2] *and Ivan Titaley*[1]*, (1) Oregon State University, (2) National Institute of Advanced Industrial Science and Technology, Japan*

Aqueous film-forming foams (AFFF) are complex proprietary mixtures used by the military sites, airport, and municipal fire stations to fight hydrocarbon fuel fires. AFFF is a recognized source of PFAS in watersheds that now impact drinking water sources. The PFAS composition of AFFF is proprietary; thus non-target approaches are needed to identify the PFAS components in AFFF. While some detailed information on PFAS in AFFFs is known, few studies approach the question using a mass balance (total fluorine) approach. A combination of analytical approaches is needed to identify all the of the PFAS in AFFF formulations, including techniques that capture non-volatile and volatile PFAS as well as any PFAS that are 'silent' under electrospray ionization (LC) or electron impact or chemical ionization (GC) conditions. Using a mass balance approach on total fluorine, we quantitatively analyzed 15 AFFFs manufactured over a 30-year time period by four manufacturers who supplied the military with MilSpec AFFF. ^{19}F NMR was used to quantify total fluorine and liquid chromatography (electrospray) with quadrupole time-of-flight mass spectrometry (LC-QToF) was used to detect ionic PFAS in negative and positive mode and gas chromatography mass spectrometry (GC-MS) with electron impact/chemical ionization was used to quantify volatile PFAS. In addition to target PFAS, suspects found by suspect screening with the new NIST list were quantified using a new quantification approach (PFAS curve). Next, non-target screening and structural elucidation were employed to identify remaining PFAS detected by LC-QToF. Other advanced techniques, including Atmospheric Solids Analysis Probe Mass Spectrometry (ASAP-MS) combined with kendo software for data visualization was used to expand the chemical space to find additional fluorine mass Target, suspect, newly identified non-target PFAS, as well as volatile PFAS accounted for 62-122% for two 3M AFFFs from the 1970s, but only accounted for 20% for three additional 3M AFFFs from the 1980s. For nine out of ten fluorotelomer-based AFFFs, < 20% of the total fluorine was quantified. Work is underway to identify the missing mass using ASAP-MS to move into a chemical space not covered by the more conventional techniques. Closing the mass balance on the PFAS chemical composition of the AFFFs will aid in the development of environmental forensics for PFAS and to more accurately attribute PFAS observed in the environment to sources and to differentiate the effects of biodegradation and transport on PFAS composition.

Occurrence and Bioaccumulation of Novel PFAS in the Delaware River Estuary, USA

Anna Robuck[1]*, James McCord*[2]*, Mark Strynar*[2]*, Mark Cantwell*[2]*, Douglas Walker*[1]*, Michaela Cashman*[2] *and Rainer Lohmann*[3]*, (1) Icahn School of Medicine at Mount Sinai, (2) U.S. Environmental Protection Agency, (3) University of Rhode Island, Narragansett*

With the phase-out of some legacy per- and polyfluoroalkyl substances (PFAS), new PFAS have been introduced in consumer and industrial applications. Little is known about the occurrence, biological accumulation, or potential impacts of novel PFAS. Here, we used high resolution mass spectrometry to evaluate

legacy and novel PFAS in surface water, passive samplers, fish muscle, and fish liver from the Delaware River, USA. This region is home to historical and current PFAS producers and users. Samples were collected within the mainstem River and tributaries at sites located between Bristol, PA downstream to Elsinboro, NJ, spanning over 50 river miles. Targeted, suspect, and non-targeted analysis were conducted; features identified using suspect lists and non-targeted analysis were semi-quantified using surrogate normalization, or presented as raw abundances. We identified a series of ether-based novel PFAS in surface water and fish tissue from across the sampled area, with the highest abundances adjacent to and downstream from a suspected point source in southwest New Jersey. Some of the chlorinated and hydrogenated species within this series were previously identified in soil, groundwater, and local surface water; here, we present the first information establishing the presence of these and additional chlorinated, hydrogenated, and fluorinated congeners in the wider ambient aquatic ecosystem. Some novel compounds were orders of magnitude more abundant than legacy PFAS in each sampled matrix. In white perch liver, the novel Cl-PFPECAs were an order of magnitude more bioaccumulative compared to PFOS based on bioaccumulation factor calculations using surface water and fish liver tissue from the same location. We also describe previously unidentified PFAS associated with recent changes in fluoropolymer production process; spectral evidence suggests these unique structures possess variable degree of fluorination, incorporation of ethers, and multiple acid moieties. Legacy PFAS including PFNA, PFOS, and PFOA were also identified in surface water and fish tissue, with some fish muscle concentrations exceeding human consumption advisories issued by the state of NJ. This work highlights the continuing evolution of PFAS occurrence in the environment, and underscores the importance of non-targeted methods to screen for PFAS beyond limited targeted lists.

Nontargeted Analysis of Per- and Polyfluoroalkyl Substances (Poster)

A Novel 'PFAS' Curve for Estimating Suspect and Nontarget Per- and Polyfluoroalkyl Substance (PFAS) Concentrations

Trever Schwichtenberg *and Jennifer Field, Oregon State University*

There are 1,000s of PFAS curated on suspect lists (e.g. NIST's "Suspect List of Possible Per- and Polyfluoroalkyl Substances" and the NORMAN database) plus more being discovered by nontargeted analysis, yet analytical standards exist for < 100 of these PFAS. Estimates of PFAS concentrations without analytical standards are needed to improve our understanding of PFAS and prioritize the future synthesis of PFAS standards. Current methods for estimating suspect and nontarget PFAS concentrations involve matching the suspect/nontarget PFAS to a structurally similar PFAS with an analytical standard, assuming an equal molar response, borrowing the calibration curve, and then adjusting for molar mass differences. However, the number of suspect PFAS is growing rapidly, and decisions for borrowing are complex and likely to lead to inconsistencies between studies over time. Studies may also use different standards, limiting the ability to use published matching schemes. For nontargeted studies structures are not always known, limiting the ability to accurately borrow a standard. An alternative method was developed to simplify, unify, and make robust over time the estimation of suspect/nontarget PFAS concentrations. The approach uses a single 'PFAS' curve composed of all target analytes. Response factors are generated by ratioing the average target area count to the average surrogate area count for each point in the calibration curve. The suspect/nontarget concentration is estimated by a response factor consisting of the average surrogate area count of a sample in the denominator and the suspect/nontarget area count in the numerator. Results include closing the fluorine mass balance on an aqueous film-forming foam (AFFF) characterized by ^{19}F-NMR, and assigning suspect concentrations to a wide variety of environmental matrices (e.g., groundwater, landfill leachates, wastewater treatment plants, and biosolids-amended soil leachates) and using those concentrations for a machine-learning analysis. This quantitation method involves simple mathematics with only small deviations depending on the analytical standards used. Thus, it is an effective solution to estimating suspect/nontarget PFAS concentrations for government, academic, and contract research organizations.

A Quantitative Comparison of Acquisition Methods for Non-Target Analysis of Per- and Polyfluorinated Alkyl Substances by Liquid Chromatography HRMS

Tarun Anumol[1]*, Ralph Hindle*[2]*, Kathy Hunt*[2]*, James Pyke*[1] *and Chris Klein*[1]*, (1) Agilent Technologies, Inc., (2) Vogon Labs, Canada*

Per- and polyfluorinated alkyl substances (PFAS) are ubiquitous in the environment, with thousands of compounds identified in this chemical classification. When the use of perfluorooctanoic acid (PFOA) and perfluorooctane sulfonate (PFOS) were replaced, industry developed many other compounds that were unregulated, increasing the complexity of identification and quantitation. Liquid chromatography with triple quadrupole (LC/TQ) mass spectrometry (MS) is the gold standard for sensitivity and highly selective quantitative analysis of PFOS, PFOA, and numerous other PFAS where commercial analytical reference materials are available, but the data cannot be used for additional compounds that were unknown at the time of analysis. Quadrupole time-of-flight (QTOF) MS provides high resolution accurate mass (HRAM) data that provides similar sensitivity to TQ, wide linearity ranges, accurate mass specificity, and can be retrospectively searched for additional compounds not included in an initial targeted list. QTOF data can be collected in a choice of acquisition modes to provide not only precursor ion detection, but also successively more specific modes that aid in determining product ions that can be used for structural elucidation of previously unknown PFAS. This study systematically compared quantitative performance for 25 PFAS in a number of data acquisition modes including targeted MS/MS (data dependant acquisition) and All Ions and Q-RAI (data independent acquisitions) on an LC-Q/TOF and MRM mode on an LC/TQ with a view of identifying the impact of different data acquisition modes on quantitative performance and data accuracy. Comparison metrics included limits of detection (LOD), linear dynamic range (LDR), accuracy and precision of replicates along with other 'soft' factors like data processing speed, file size, ability to perform suspect and unknown screening and others. This work is the first known to make such a comparison for PFAS quantification using same chromatography and multiple acquisition modes using a TQ and QTOF-HRMS.

An Integrated Approach for Determination of Total Per- and Polyfluoroalkyl Substances (PFAS)

Marzieh Shojaei*, Naveen Kumar and Jennifer Guelfo, Texas Tech University*

Per- and polyfluoroalkyl substances (PFAS) are difficult to analyze in environmental media due challenges such as extraction recovery and lack of analytical standards. The total oxidizable precursor (TOP) assay and suspect screening analysis coupled with semiquantitative concentration estimates (SQ) are two approaches to assess total PFAS in environmental media, but studies are needed to optimize workstreams for total PFAS analysis. This study applied two soil extraction methods, TOP assay, and SQ analysis to 3 aqueous film-forming foams (AFFF) and 3 AFFF-impacted soils. In soils, total PFAS estimated using a recently published extraction method utilizing sequential acidic and basic solvents led to a 35% increase in precursors during TOP assay in one of 3 soils tested, but concentrations did not increase significantly in remaining soils. Sample-specific dilution schemes were required to overcome matrix effects caused by the acidic extraction step that influenced estimates of total PFAS by SQ. In 3 AFFFs, suspect screening of post-TOP samples identified 8 classes of PFAS present after oxidation. Concentrations of 3 classes increased, suggesting they are new TOP endpoints. The remaining 5 classes decreased after TOP but their concentrations either remained the same or decreased slightly. The post-TOP signatures may be unique to the PFAS manufacturing origin. As a result, combined TOP and SQ workstreams may yield the most representative assessment of total PFAS composition and concentration and/or information useful in forensics applications. New TOP endpoints and PFAS recalcitrant to TOP identified in this study did not degrade in harsh conditions and are structurally similar to PFCAs and PFSAs, suggesting a similar degree of persistence and a need for routine monitoring in the environment.

Environmental Forensic Investigation of Chemical Manufacturing and Use of PFAS by Nontargeted Analysis

James McCord *and Mark Strynar, U.S. Environmental Protection Agency*

Industrial producers and users of per- and polyfluorinated alkyl substances (PFAS) have been a major source of

chemical contamination to nearby communities due to historical releases of legacy PFAS. Following the general phase-out of legacy PFAS such as PFOA/PFOS, industrial usage has shifted to alternative PFAS chemicals and it has been the domain of non-targeted analysis (NTA) to identify replacement species. Knowledge of emerging chemical contaminants is of significant environmental and human health concern due to the potential for widespread release, persistence, and bioaccumulation of PFAS species, along with their concomitant toxicological effects, which are frequently understudied. Over the past several years, we have obtained samples of industrial effluent and contaminated ground and wastewater from numerous state and regional EPA offices for investigation into the presence of legacy and emerging PFAS chemistries. NTA using high-resolution mass spectrometry was used to examine the chemical identities of PFAS within these samples and compare the abundances of legacy and emerging chemical contaminants. NTA has been able to identify legacy contaminants of use, as well as elucidate replacement chemistries used by a variety of chemical manufacturers. These include a previously identified family of chlorinated perfluoropolyethers (ClPFPECAs), polyfluorinated side products of polyfluorovinylidine (PVDF), as well as several other novel PFAS fluoroether species exhibiting ether linkages and acid head groups. Several of these families have been reported as components of fluoropolymer manufacturing in multiple locations in the United States, while some are novel to specific active sites. Measurements indicate the presence of effluent-derived contaminants in localized groundwater and recirculation through processes that are "non-contact" is common, and that both intentionally added processing aids and polymeric production byproducts are frequent PFAS sources. This presentation will discuss case studies of recent work by EPA ORD in multiple PFAS site investigations.

Increasing Confidence in Resolving Isomers of Emerging Per-And Polyfluoroalkyl Substances (PFASs) by Using DFT and Cosmo-Rs to Predict Chromatographic Retention Factors

Jonathan Antle[1]*, Mary Grace Guardian*[1]*, Diana Aga*[2] *and Scott Simpson*[3]*, (1) University at Buffalo, (2) State University of New York, Buffalo, (3) St. Bonaventure University*

Per-and polyfluoroalkyl substances (PFAS) are a class of contaminants of emerging concern frequently used in products like aqueous firefighting foams and non-stick coatings due to their stability and surfactant-like qualities. The lack of analytical standards for many emerging PFAS have severely limited our ability to comprehensively identify unknown PFAS contaminants in the environment, especially those that occur as isomers. Annotation of small molecules and identification of unknowns based only on elemental composition and mass fragmentation patterns remain major challenges in nontarget analysis employing liquid chromatography with high-resolution mass spectrometry (LC-HRMS). In this study, chromatographic retention factors (k) and mass spectral fragmentation patterns of 32 known PFAS were determined using our optimized parameters in LC-HRMS. The same method was then used to analyze previously unidentified PFAS in actual environmental samples. Using characteristic ions observed in the MS fragmentation of PFAS, the most probable isomeric structures of the detected PFAS were predicted. To increase confidence in the predicted molecular structure, Density Functional Theory and Conductor-like Screening Model for Realistic Solvents (COSMO-RS) calculations were used to predict physicochemical properties of different constitutional isomers. The DFT calculations facilitated geometric optimization, determination of polarizability, and calculation of the chemical potential the isomers. COSMO-RS uses the chemical potential to predict thermodynamic properties of molecules such as pKa, solubility, and K_{ow}. These properties were then used to make a multi-variable linear regression to predict k values. The model was trained using 32 known PFAS. The properties used were log K_{ow} of the neutral and anion species of the PFAS, and their polarizability. The model was specific enough to predict significantly different k values of unknown compounds with similar structures, which facilitated assignment of isomeric structures of PFAS.

Integrating Non-Targeted Analysis into High-Throughput Monitoring of PFAS Contamination in Drinking Water: When Is NTA Appropriate?

Abigail Joyce[1]*, Noelle DeStefano*[2]*, Gordon Getzinger*[1]*, Patrick Faught*[1]*, Detlef Knappe*[2] *and Lee Ferguson*[1]*, (1) Duke University, (2) North Carolina State University*

Per- and Polyfluorinated alkyl substances (PFASs) contaminate drinking water sources throughout the United States and have been measured at elevated concentrations at several locations in the state of North Carolina (NC). North Carolina has several documented point PFAS sources, e.g. fluorochemical and textile manufacturing discharge to rivers, firefighting activities at military installations, and application of biosolids to agricultural fields leading to runoff into reservoirs. The diverse and widespread commercial use of ever-changing PFAS chemicals and resulting complex mixtures compounded with environmental transformations make characterizing PFAS contamination difficult. Prompted by the known PFAS contamination and mandated by state legislative action, a holistic monitoring study was initiated to survey PFAS occurrence throughout drinking water sources in the state of North Carolina. A total of 376 samples were collected, including surface water intakes and groundwater. These samples were characterized using a quantitative HPLC-MS/MS method targeting 47 individual PFAS and by non-target analysis (NTA) via an online-solid phase extraction (SPE), HPLC-HRMS method for comprehensive and trace-level suspect screening of PFAS chemicals. Suspect screening included a list of 7,267 known PFAS molecules and 41,629 predicted chemical and biological transformation products of these PFAS molecules. Targeted analysis showed 20 sources had total PFAS measurements greater than 70 ppt, all but two of these sources were from the Haw River and Cape Fear River basins-both of known PFAS contamination. Nine of the ten most concentrated sources were from the Cape Fear River basin, downstream from a PFAS manufacturing plant. Approximately 50 tentative "novel" PFAS compound annotations were obtained from NTA analysis, primarily from the Cape Fear River Basin and well contaminated with AFFF. NTA showed consistent detections of compounds targeted by LC-MS/MS analysis and allowed for detection of Nafion-byproduct-4 in the Cape Fear River basin that had previously been undetected by targeted analysis. Novel compound annotations were detected at overall lower abundances than were targeted PFAS compounds. Overall, our results indicated that targeted analysis accounted for the majority of detectable PFAS in drinking water, but that non-targeted analysis allowed for detection of several unanticipated PFAS compounds present at low relative abundance.

Methodology for Identifying Sources of PFAS Precursors, Transformation Products, and Alternatives Within a Wastewater Treatment Plant

Rebecca Dickman[1]*, Alysson Bermudez*[1] *and Diana Aga*[2]*, (1) University at Buffalo, (2) State University of New York, Buffalo*

Occurrence of per and polyfluoroalkyl substances (PFAS) in effluents of wastewater treatment plants (WWTPs) has been documented over the last decade due to the ubiquity and persistence of this chemical class. Previous studies have mainly focused on the analysis of common carboxylate and sulfonate PFAS, and the majority of compounds within the class (9,000+ PFAS) remain unidentified. By focusing on only a small fraction of PFAS, targeted analyses of PFAS within WWTPs can yield misleading results since many precursors and transformation products might not be properly identified. Degradation pathways resulting in the production of perfluorooctanoic acid (PFOA) and perfluorooctane sulfonic acid (PFOS) have been observed in the environment, suggesting that these two ubiquitous PFAS can form from unidentified PFAS in influent sources. Treated wastewater from WWTP effluents are released directly into surface waters, and subsequently, drinking water sources. Additionally, land application of biosolids is the most common route of disposal of sewage sludge, and can potentially introduce the contaminants to food sources when used for fertilization in agriculture. Therefore, understanding the extent of PFAS within WWTP is crucial to drive efforts for minimizing human exposure to the contaminant. This study aimed to pair targeted quantitative analysis with suspect screening techniques on five aqueous and two solid samples from a WWTP to achieve a more complete understanding of mass flows and partitioning of PFAS within the treatment process. The analysis required rigorous sample clean-up and thorough data inspection due to the samples' complexity. The workflow successfully identified 17 PFAS within the WWTP. It was observed that 6 detections were unique to the aqueous samples, and 4 detections were unique to the solid samples. Each detection required exact mass and fragmentation matches, a mass defect within -0.25 to 0.1 amu (a range supporting a fluorinated molecule), and isotope trend matching to predict the chemical structure and formula. When the resources were available, detections were confirmed with reference

standards or spectral databases. The identified compounds ranged from PFAS precursors, possible transformation products, and PFAS alternatives that are either unsaturated or polyfluorinated. The concentration of newly identified PFAS were estimated using the area of the most structurally similar isotopically labelled internal standard.

Nontarget Analysis of PFAS in Waterbodies in Europe, Australia and the South Pacific
Bastian Schulze[1]*, Sara Ghorbani Gorji*[1]*, Carly Beggs*[1]*, Rachel Mackie*[1]*, María José Gómez Ramos*[2]*, Kevin Thomas*[1] *and Sarit Kaserzon*[1]*, (1) The University of Queensland, Australia, (2) University of Almería, Spain, (4)*
Per- and polyfluorinated alkylic substances (PFAS) are a group of over 3000 chemical compounds that are considered persistent and potentially cancerous substances. PFAS have been used in numerous products and applications, for example in firefighting foams and non-stick coatings. Awareness of this group of compounds is growing and regulations are being put in place to limit their impact on the environment. Presently, only a handful of standards exist for quantification and identification of about 2% of known PFAS compounds. Nontarget analyses using high resolution mass spectrometry (HRMS) is an important tool that is helping to elucidate the identification of a wider range of PFAS in environmental and biological samples. Information obtained through non target analysis (NTA) combined with harmonized HRMS reporting criteria could help provide a better representation of the presence and risk from PFAS in the environment without focusing on only a selected few or help discover sources of PFAS. In this study, aquatic passive sampling devices designed for the monitoring of PFASs were deployed in several water bodies in Europe, Australia and the South Pacific Islands and analyzed using HRMS NTA method. For each dataset separate feature lists were produced. Obtained feature lists were then filtered using blanks, Kendrick mass defect and relative standard deviations of pooled quality control samples. Filtering reduced the number of features 90-99 % giving lists of 88-1111 features. Using different statistical approaches, similarities and differences in PFAS composition between the different datasets were determined and compared to potential source zones at the sampling sites. The approach highlighted described a sample processing method that could help characterize PFAS in the environment and help determine the source of water contamination.

Nontargeted Analysis for Identification of PFAS Biotransformation Products in Rat Plasma
Denise MacMillan[1]*, Aero Renyer*[2]*, Michael DeVito*[1]*, Michael Hughes*[1] *and Leah Wehmas*[1]*, (1) U.S. Environmental Protection Agency, (2) Oak Ridge Institute for Science and Education*
Per- and polyfluorinated substances (PFAS) are frequently used in industrial processes and commercial products, leading to environmental contamination, exposure, and the potential for adverse health effects. With toxicological data available for only a few of the more than 4700 known PFAS, the United States Environmental Protection Agency is using short term exposure studies to establish interim benchmark dose levels for emergent PFAS and promote understanding of the biological responses to PFAS exposure. Introduced exogenous chemicals are often transformed *in vivo* to increase water solubility and promote elimination. Little is known about biotransformations of PFAS, however. Here we present results of nontargeted analysis of plasma to identify potential biotransformation products from rats exposed over five days to multiple concentrations of perfluoro(2,5,8-trimethyl-3,6,9-trioxadodecanoic) acid (PF-TODoA) or 2,2,3,3-tetrafluoro-3-(trifluoromethoxy) propionic acid (PF-MOPA). Extracts were analyzed on a Sciex X500R QTOF using independent data analysis (IDA) and sequential window acquisition of all theoretical mass spectra (SWATH) scanning. Data were processed with Sciex OS 2.1 and MarkerView 1.3.1. BioTransformer 3.0 was used to predict PFAS metabolites. Several biotransformations were predicted for PF-TODoA, a twelve carbon perfluoroether carboxylic acid. At least two ions with masses consistent with predicted biotransformations were observed from PF-TODoA-exposed rat plasma. For both ions, formulas with less than 5 ppm mass error compared to the predicted products were generated with Sciex OS. One ion was tentatively identified as the molecular ion of the product formed by *O*-dealkylation of the PF-TODoA carboxylate moiety to give an alcohol. The MS/MS spectrum included ions that are similarly observed for PF-TODoA. All acquired spectra were searched also for the presence of fragments consistent with structural components of a PF-TODoA

precursor. Numerous instances of characteristic fragments were detected, suggesting the presence of additional species of interest. The only biotransformation predicted for PF-MOPA, a small four carbon carboxylic acid, was *O*-glucuronidation, which was not observed in initial experiments. These data improve our understanding of *in vivo* disposition of PFAS and contribute to an important aspect of their risk assessment. Disclaimer: This abstract does not necessarily reflect EPA policy.

Non-Targeted Analysis for the Screening of Per- and Polyfluoroalkyl Substances in Drinking and Surface Water Samples from South Florida Environments

***Xuerong Li**, Danni Cui, Brian Ng, Piero Gardinali and Natalia Quinete, Florida International University*

Per- and polyfluoroalkyl substances (PFAS) are a group of anthropogenic pollutants that are found ubiquitous present in surface and drinking water supply. It poses big concerns on human and environmental exposure, considering their persistent feature, bioaccumulative potential, and significant adverse health effects associated at low concentrations. With the effort of sensitive target analysis, identification and quantitation of some legacy and emerging PFAS compounds can be achieved. However, the information on thousands of unknown PFAS products from precursors, degradants, and metabolites is limited by the availability of analytical standards and scientific knowledge on fate and transformation of these contaminants. With the advances in high-resolution mass spectrometry (HRMS) methods, there has been an increasing number of non-targeted analysis (NTA) approaches that allow for a more comprehensive characterization of total PFAS present in environmental samples. A full assessment with both target and non-targeted approaches will play a crucial role in the understanding of the toxicological and environmental impact of PFAS, which are largely underestimated due to the lack of such information for many PFAS. In this study we have developed a NTA workflow based on an online solid phase extraction (SPE, Water WAX)- liquid chromatography (LC)- HRMS method using a Q-Exactive Orbitrap system for the screening of PFAS species in drinking waters from populated counties in South Florida, as well as in surface waters from Biscayne Bay, Key west, and Everglades, complementing our previous target analysis of 30 PFAS congeners. Water samples were run full scan negative mode with a scan range from 100.0 to 800.0 m/z at a resolution of 140,000, followed by data dependent MS/MS with a normalized collision energy of 30 and at a resolution of 35,000. Data post-processing was conducted using the small molecules identification software Compound Discoverer 3.3 and FluoroMatch, a software specific for PFAS NTA, with the goal to compare the variability in processing NTA data for PFAS analysis. A Semi-quantitation method for PFAS identified by NTA was also achieved using existing native and internal standards (IS), which concentration estimates were determined by a regression-based model and IS response factors.

Non-Targeted High-Resolution Mass Spectrometry Workflow

***Ying Long** and Lam H Leung, The Chemours Company*

Background: Non-targeted high-resolution mass spectrometry (HRMS) methodology was used to characterize potential unknown per- and polyfluoroalkyl substances (PFAS) analytes in selected process and non-process wastewater and stormwater samples at the Chemours Fayetteville Works. **Challenge and problem:** High mass accuracy of HRMS allows predication of molecular formula and identification of unknown structures. Most per- and poly- fluorinated compounds have negative mass defects due to the high degree of fluorination of the compounds. Some may have slightly positive mass defects when multiple hydrogens are present in molecules of polyfluoroalkyl substances. This mass defect feature is combined with other molecular features (adducts, isotopic distribution) to screen preliminarily as many PFAS as possible. There are challenges in this task for unknown screening and identification of PFAS: 1) the trace-level detection limit (parts-per-billion) requires both sensitivity and accuracy of screening method; 2) complex mass spectra resulting from various ion adduction and in-source fragmentation cause difficulty to determine true molecular weight and empirical formula of analytes; 3) mass spectral data analysis is time-consuming for large data sets generated in workflow; 4) the screening approach of unknown PFAS based on mass defect of fluorine may cause incorrect assignment of empirical formula resulting in potential false positives. **Methods:** To address these challenges, several approaches have been attempted: 1) the enrichment of analytes was achieved through either large injection

volume or optimized solid phase extraction (SPE); 2) verification of findings from data analysis by multiple experimental designs, especially for potential false positives; 3) Identification of unknown structures based on a combination of accurate mass, empirical formula, tandem mass spectrometry (MS/MS), and discussion with process and synthetic chemists. **Results:** Four interim reports have been provided to NC-DEQ since 2020 on efforts to identify previously unknown PFAS in collected water samples at the Chemours Fayetteville Works. **Conclusion:** Improvement in method sensitivity and potential "false positives" identification will be demonstrated and how best to address this.

Optimizing a Method for Per- and Polyfluoroalkyl Substances (PFAS) Extraction in Fish Fillet to Explore Spatial Distributions Using Sunfish

***Anna Boatman**, James Dodds, Scott Belcher and Erin Baker, North Carolina State University*

Per- and polyfluoroalkyl substances (PFAS) are a class of synthetic molecules gaining global attention as contaminants of concern due to their persistence, bioaccumulation, and toxicity. Biological effects of exposure to complex mixtures of these compounds is a growing area of research interest, requiring robust analytical techniques to monitor the accumulation of the thousands of different PFAS congeners in tissue. Fish muscle tissue provides a unique avenue to assess human exposure, both through fish consumption and through drinking or recreational use of contaminated waters. While multiple PFAS extraction methods currently exist in literature for fish tissue, their respective advantages and disadvantages for processing a large number of biological samples have not been compared. For example, a commonly used clean up step in PFAS extractions is solid-phase extraction (SPE) using a weak anion exchange cartridge to remove molecules that could interfere with PFAS detection. SPE is also time- and resource-intensive, and the increase in sensitivity provided by the addition of this step may come at a cost of congener bias and limited sample throughput. Each step in the extraction process also adds time and has the potential to introduce contamination. Thus, we explored clean up steps and other parameters to identify an optimal extraction method that supports simple, fast fish tissue extractions while retaining our ability to detect PFAS congeners of interest. In this study, we used a standard reference material from the National Institutes of Standards and Technology (NIST SRM 1947), consisting of fish tissue and known quantities of five different PFAS. PFAS standards were also spiked prior to extraction to determine efficiency of the extractions and to support quantitation of both legacy and emerging PFAS by liquid chromatography-ion mobility spectrometry-mass spectrometry (LC-IMS-MS). After optimization and evaluation of the different methods, our procedure will be applied to sunfish caught in North Carolina lakes to assess PFAS presence and spatial distributions.

The Use of Non-Targeted Analysis to Determine the Efficacy of Per- and Polyfluoroalkyl Substance (PFAS) Destruction Technologies

***Ariel Wallace**, Erin Shields and William Linak, U.S. Environmental Protection Agency*

Per- and polyfluoroalkyl substances (PFAS) are large family of man-made chemicals with strong carbon-fluorine bonds and low surface energies. These characteristics impart PFAS with unique properties allowing their useful applications in many consumer products and industrial processes, such as stain and water-resistant coatings, firefighting foams (aqueous film forming foam (AFFF)), polymer dispersions, and many others. Due to their widespread use and chemical and thermal stability, PFAS are becoming increasingly ubiquitous within the environment and can be found in water, soil, biosolids, food, dust, and air. Their thermal stability suggests that new techniques are required to better define the destruction efficiencies of PFAS and to characterize any products of incomplete destruction (PIDs) that may be produced during thermal treatment processes. In this study, three PFAS of interest, including 8:2 fluorotelomer alcohol (FTOH), perfluorooctane sulfonamide (PFOSA), and N-ethyl-N-(2-hydroxyethyl) perfluorooctane sulfonamide, were thermally treated in the gas-phase from 200 to 800 °C in an externally heated quartz tube furnace and 0.5-1 L samples were actively collected using multi-sorbent bed passivated stainless steel universal thermal desorption (TD) tubes. The same three PFAS were also treated in the same tube furnace with a commercially available solid alumina catalyst. Samples were analyzed using TD-gas chromatography-QExactive-mass spectrometry (GC-MS) using multiple

ionization modes to monitor the destruction of the original PFAS compounds and to detect non-targeted PIDs. As the parent compound's peak area decreased as the furnace temperature increased, one or more PIDs started to appear for most of the compounds tested. The electron ionization (EI) spectra of the PIDs were searched using mass spectral libraries to determine tentative identifications and positive chemical ionization (PCI) spectra were used for molecular mass confirmation. This study shows the importance of non-targeted analysis in evaluating new and existing PFAS destruction technologies to ensure that both the parent PFAS and their PIDs are fully destroyed (mineralized). This research was funded by U.S. Environmental Protection Agency (EPA). The views expressed in this abstract are those of the authors and do not represent the views and policies of the U.S. EPA.

Use of Non-Target Analysis for Matrix Mitigation and Accurate Quantification of Legacy PFAS
***Charles Neslund**, Eurofins Lancaster Laboratories Environmental, LLC*
For the majority of analytical applications for the analysis of PFAS in environmental samples, LC/MS/MS has become the analytical workhorse. As we explore the prevalence of PFAS, the analytical techniques have evolved to include "non-traditional" matrices that are substantially different from the analysis of drinking water, ground water, soils/sediment and biosolids. While these matrices have their own set of analytical challenges, the expansion of these analytical techniques and the use of LC/MS/MS to matrices like consumer products, cosmetics, non-stick coatings, AFFF (aqueous film forming foam) and oils, in some cases has stretched the bounds of reliability of the triple quad platform. We will report on projects where the concentrations detected of legacy PFAS in these unique matrices was higher than anticipated based on estimations from the source materials used for manufacture. The "high bias" that was observed by the triple quad platform was investigated using qTOF (quadrupole Time of Flight) Mass spectrometry. The investigation uncovered transitions uncharacteristic of the legacy PFAS, but not resolvable, initially, by the LC/MS/MS. Using the exact mass transitions determined by qTOF, we were able to construct a set of acquisition conditions to differentiate between the legacy PFAS and the other co-extracted material.

Process-oriented Studies in Natural and Engineered Systems

Probing Environmental Processes Using High Resolution Mass Spectrometry
***Thomas Young**, University of California*
Chromatographic separation followed by high resolution accurate mass ion abundance scans can produce aligned datasets that can be used to obtain critical insight into chemical transport and transformation across time and space. Example applications include comprehensive investigation of compounds degraded and formed in treatment systems, comparison of alternative treatment system performance, analysis of process stability, and the identification of marker compounds unique to particular sources or processes. This talk explores example applications of nontarget analysis in wastewater and biogas treatment with an emphasis on the data processing required to derive valid inferences. One challenge in the statistical analysis of nontarget feature abundances over time are the existence of batch effects. To assess and correct for the influence of batch effects on the analysis of wastewater samples collected from various monitoring points within a sewer system over the span of seven months, principal variance component analysis and the empirical Bayes method, ComBat, were applied. This novel application of ComBat appears to significantly reduce the influence of batch effects. Samples from particular commercial facility types (pest control operators, laundromats, and pet grooming facilities) are compared to average wastewater influent samples to characterize the specific chemical signatures of wastewater from these operations. Degradation of parent compounds and formation of byproducts is also assessed. A second application area examined is the carbonyl and semi volatile organic compound (SVOC) composition of biogas produced from the digestion of wastewater sludge, solid waste, food waste and dairy manure and the changes in composition occurring as the gas is combusted in appliances, vehicles or on-site power generators and subsequently transformed during atmospheric aging.

Evaluating the Removal of Roadway-Derived Contaminants in Compost Amended Bioswales Using Non-Targeted Analysis

Zhenyu Tian[1], *Katherine Peter*[2], *Benjamin Leonard*[3], *Christopher Wu*[2], *Bowen Du*[4], *Jill Wetzel*[3], *Jenifer McIntyre*[3] *and Edward Kolodziej*[2], *(1) Northeastern University, (2) University of Washington, (3) Washington State University, (4) Southern California Coastal Water Research Project*

Urban stormwater is a significant contributor to impaired water quality in receiving waters. Contaminated roadway runoff may cause various adverse effects to aquatic species, ranging from sublethal impacts to acute toxicity. Stormwater treatment technologies are often installed to address such water quality challenges. However, performance of many treatment technologies is only evaluated or optimized with regulated parameters, and little is known with respect to performance for removal of broader suites of contaminants, long-term performance, or effectiveness for protection of aquatic organisms. In this study, we evaluated the roadway runoff treatment performance of compost-amended bioswales (CABS) in both laboratory and field conditions. The field CABS was a 100 ft full-scale system on State Route 518 (SR518) in SeaTac, WA, and the lab CABS was a 48 ft triplicate laboratory system constructed at the Washington Stormwater Center (WSU-Puyallup, Puyallup, WA), treating the runoff from SR520 of Seattle. Over 3 years, we collected 9 sets of samples from the field CABS during storms and 8 sets of samples from the lab CABS, and compared the contamination profile in untreated and treated runoff samples to assess the performance. We collected samples at different lengths and flowrates in the lab system to analyze the removal mechanisms. We measured regulated parameters and target contaminants, including TSS, Cu, Zn, PAHs, and phthalates, and we used high-resolution mass spectrometry (HRMS) to conduct a non-targeted analysis, characterizing the removal of a broad range of organic contaminants. The results showed that longer swale length and longer hydraulic retention time (HRT) improved contaminant removal. The removal mechanisms of contaminants varied according to the properties, and hydrophobic sorption and ion exchange were the primary removal mechanisms. In general, the removal rates of non-polar organics were satisfactory, but in consistent removal of dissolved metals were observed. 6PPD-quinone, a recently discovered toxicant that kills coho salmon, was efficiently (>80%) removed in both laboratory and field CABS system.

Non-Target Analysis to Discover the Seasonal Dynamics in the Chemical Composition of a Small Stream

Cheng Shi[1], *Damian Helbling*[2] *and Gerrad Jones*[1], *(1) Oregon State University, (2) Cornell University*

Understanding the temporal patterns of non-target chemical composition is critical to develop contaminant monitoring and source apportionment strategies in aquatic ecosystems. Temporal variability in source composition and the effects of temporal changes in ecosystem processes lead to the temporal variation of chemical composition in river waters. However, one question that remains unanswered is if there is a predictable pattern in the non-target chemical composition of river waters and what ecosystem factors determine the pattern. We obtained high-resolution mass spectrometry acquisition files from weekly composite samples collected from a river monitoring station over 18 months. We used XCMS to extract non-target chemical features from each of the acquisition files and performed Fourier transformation to convert the time series of the chemical abundances to frequencies. The extracted frequencies will be used to represent seasonalities in the chemical composition and determine if there are multiple parallel seasonal patterns (frequencies) mixed in the time series of non-target chemicals. We anticipate that the difference in seasonality among the non-target chemicals will reveal their different sources and/or different ecosystem processes driving the gradients within this system. The environmental covariates with the same seasonal patterns will indicate their contribution to the specific non-target chemicals. The findings from this study will help develop new workflows to predict future changes in non-target chemical composition and determine their potential sources and ecosystem processes.

Tracking Pharmaceutical Emissions in River Water Used for Drinking Water Production

Tobias Bader *and Rudi Winzenbacher, Landeswasserversorgung, Germany*

The Landeswasserversorgung (LW) is one of the largest remote water suppliers in Germany and responsible for the drinking water supply of more than three million people. The LW directly extracts river water from the

Danube River and uses a multi-stage treatment process comprising ozonation and activated carbon filtration for drinking water treatment. In contrast to groundwater, the river water is especially subjected to anthropogenic emissions and thus poses a potential threat to the drinking water quality. Since 2014, liquid chromatography high resolution mass spectrometry (LC-HRMS) is used for the monitoring of raw and drinking water samples at least twice per week. After processing, the data comprise tens of thousands of features and hundreds of samples. In this study, time series data were used to prioritize compounds, which could be attributed to the pharmaceutical industry. The source of these emissions was traced back to a municipal wastewater treatment plant (WWTP) located at a small tributary. Around 100 pharmaceuticals and transformation products could be linked to a single company. The quantification of a subset of pharmaceuticals revealed concentrations up three digit µg/L in the WWTP effluent and in the tributary. Further investigations are conducted to correlate the production data with the emission data from the WWTP. In terms of drinking water quality, the pharmaceutical emissions were tracked across the treatment process. Furthermore, time series correlation of known pharmaceutical precursors in the river water with unknown features in the drinking water revealed some ozonation transformation products in the final drinking water. The findings were discussed with the competent authority, which imposed mitigation measures to reduce the emissions into the environment. Further monitoring will help to assess the success of these measures in the near future.

The Fate of Polypropylene and Polyurethane Microplastic-Associated Chemicals in a Freshwater Laboratory and Mesocosm Study

Imari Walker[1]*, Jeanne Hankett*[2]*, Cara Peters*[1]*, Nicholas Geitner*[1]*, Mark Wiesner*[1]*, Wendel Wohlleben*[3]*, and Lee Ferguson*[1]*, (1) Duke University, (2) BASF Corporation, (3) BASF SE, Germany*

Microplastics have become ubiquitous in freshwater environments due to the pervasiveness of plastic use and disposal globally. The risk of microplastics in the environment can be attributed to the ability for these small particles to be transported in water whilst releasing potentially hazardous polymer associated chemicals (PACs). In this study, we utilized non-targeted analysis to examine the release and fate of PACs from polypropylene (PP) and polyurethane (PU) microplastics during 12-month freshwater wetland mesocosm experiments and 3-month freshwater laboratory photolysis and leaching studies. There were significant differences in the number of features detected by polymer type (128 for PP and 34 for PU) likely associated with the physical/chemical properties of the polymer and the types of PACs incorporated. Laboratory controlled leaching studies determined important PACs that were readily leachable and able to undergo transformation including PP polymer additive Tinuvin 770 (Bis(2,2,6,6-tetramethyl-4-piperidyl) sebacate) and PU monomer 4-(4-formamidobenzyl)phenyl) formamide. Three PACs (ricinine, Tinuvin 770, and 2,2,6,6-Tetramethyl-4-piperidinol (TMPO)) were chosen for quantitation and were modeled for release behavior in the mesocosm experiments based on laboratory experimental findings. These compounds demonstrated opposite leaching behavior between the two studies (mesocosm: concentration decreased over time; laboratory: concentration increased over time). Decreased leaching behavior in the complex mesocosm may have been due to microbial degradation and/or sorption to organic particulates and subsequent sedimentation in the system. The role of UV exposure in PAC leaching and transformation was also investigated in the laboratory leaching studies. Release of ricinine, a potentially toxic impurity leachable from PU microplastics, was shown to be inhibited by UV light, indicating possible photodegradation of the compound. Monomeric derivatives in PU, 4-(4-formamidobenzyl)phenyl)formamide and 2-formylamido benzoic acid, were predominantly formed through combined photooxidation and hydrolysis reactions. Release behavior of Tinuvin 770 hydrolysis derivatives were not significantly impacted by UV pretreatment to PP. These results show that polymer additives, monomers, unintentionally added substances, and their degradation products are important leachable chemicals in PP and PU microplastics.

Process-oriented Studies in Natural and Engineered Systems (Poster)

The Occurrence of Water Soluble Organic Pollutants in PM2.5 Collected During Wildfire Events in the Western United States

Phillip Berger[1]*, Sophie Brenner*[1]*, Nicole Manalis*[1] *and Imari Walker*[2]*, (1) RTI, (2) Duke University*

Increasing trends in dry conditions, particularly in the western region of the U.S, have led to an increase in wildfires. Wildfires release a variety of Volatile Organic Chemicals (VOCs) during burning and smoldering. Volatile organic compounds (VOCs) undergo gas and particle phase photochemical reactions to form fine particles $PM_{2.5}$. $PM_{2.5}$ not only degrades air quality but can also have detrimental effects on human health. Fine particles collected for the Interagency Monitoring of Protected Visual Environments (IMPROVE) network are routinely analyzed for inorganic species, sulfate, nitrate and total carbon. The characterization of polar organic chemicals in $PM_{2.5}$ has not been routinely defined. In this study, we utilized non-targeted analysis (NTA) to characterize polar organic chemicals associated with fire impacted $PM_{2.5}$. We investigated fire impacted samples from the IMPROVE network from locations in California, Oregon, and Colorado collected during September-October 2020. $PM_{2.5}$ from multiple time points for each site were extracted with deionized water and then composited, dried and concentrated. Composites were analyzed on the Thermo Fisher Orbitrap IQ-X Tribrid Mass Spectrometer (ESI +/-). Differences in the abundance and types of features were noted by site location and collection dates. Out of the 384 features observed in positive and 380 features observed in negative mode, a significant portion of the compounds were nitrogen containing organic chemicals (268 in POS and 138 in NEG) and a small quantity were sulfur containing organic compounds (24 in POS and 68 in NEG). Structural annotation for 54 features was completed via spectral library matching (MzCloud) and included several chemicals previously measured in wildfires (mannitol and xylose), agrochemicals (4-nitrophenol), and chemicals associated with fire-retardants (melamine). Analytical standards confirmed the identity of multiple common organic acids associated with $PM_{2.5}$ (succinic acid, methyl succinic acid, glutaric acid, and fumaric acid). Next steps include further method development and analysis of samples collected from additional sites and timepoints throughout the U.S.

Suspect and Nontarget Screening of Organic Micropollutants and Their Transformation Products in Groundwater Impacted by Onsite Wastewater Treatment Systems

Shiru Wang*, Syracuse University*

Onsite wastewater treatment systems (OWTS) are increasingly recognized as an alternative to treat and dispose of sewage in suburban and rural areas. Over past decades, numerous studies have highlighted OWTS as a significant source of nutrients, suspended particles, and pathogens to the groundwater. However, little is known about the occurrence of organic micropollutants (OMPs) and their transformation products (TPs) in OWTS and the underlying groundwater. In this study, we worked with the USGS New York Water Science Center to characterize the occurrence pattern of OMPs and their TPs in groundwater and public water supply systems impacted by OWTS in Long Island, NY. We collected 37 grab samples from groundwater monitoring sites and 10 composite samples from 5 OWTS between 2017 and 2021. We applied a suspect screening method using liquid chromatography-high resolution mass spectrometry (LC-HRMS) to screen for a broad suite of OMPs and their TPs. Overall, suspect screening confirmed the presence of 76 pharmaceuticals and their TPs, 9 pesticides and their TPs, and 14 personal care products and their TPs in 47 water samples. The non-target screening approach was then applied to identify unknown OMPs and TPs with regard to their source, based on their occurrence at sites that are OWTS impacted. In this regard, hierarchical clustering was applied to prioritize nontargets with predominating occurrence in groundwater samples collected downgradient of the OWTS. Moreover, the ratio of a nontarget's intensity in groundwater samples collected upgradient of the OWTS to its intensity in groundwater samples collected downgradient of the OWTS was calculated to prioritize the nontargets that are likely originated from OWTS. Non-target screening tentatively identified a total of 1,182 unique features in the downgradient groundwater samples, which still need further verification. Our ongoing work focuses on using non-target screening to identify and quantify OMPs and their TPs, and investigating the relationship between the occurrence pattern of OMPs and their TPs and the land use data.

Uncovering Use Patterns and Overlooked Contaminants of Concern in Residential Wastewater and Onsite Treatment Systems Using High-Resolution Mass Spectrometry

***Rachel Smolinski**, Stony Brook University*

Onsite wastewater treatment systems (OWTSs) often do not remove organic contaminants, releasing them to groundwater and local surface waters and polluting critical water resources. On Long Island, the majority of residences are served by OWTSs, resulting in widespread groundwater contamination and nutrient pollution. Nitrogen removing biofilters (NRBs) are promising alternative OWTSs, though many knowledge gaps remain concerning organic contaminant occurrence in residential wastewater and interactions with NRBs. The objective of this work is to use an HRMS nontarget analysis workflow to identify unknowns in residential wastewater across eleven pilot NRB installation sites. The approach focuses on emphasizing features with high detection frequency and abundance, then assessing seasonal usage patterns and contaminant fate, including transformation product screening. Wastewater influent and effluent from all NRB sites is analyzed by liquid chromatography coupled to quadrupole time of flight mass spectrometry (LC-QTOF-MS). Analysis of grab and passive samples has resulted in the tentative identification of > 120 compounds including pharmaceuticals, plasticizers, and per- and polyfluorinated substances. Most of these compounds have yet to be considered in NRB monitoring studies, demonstrating that many organics are overlooked in targeted approaches. Fold change analysis of features detected in influent and effluent brings attention to significant compound removal and likely transformation product generation in NRBs. Additional data visualization through heat map analysis highlights feature clusters that are unique to season and site, which can aid in future site characterization. This work will inform optimization of next-generation treatment systems, improve identification and prioritization of contaminants in residential wastewater, and improve our understanding of the occurrence of transformation products in treated effluent.

Characterization of Micropollutants in Urban Stormwater Runoff Using Non-Target Screening

***Daeho Kang**, Changwon National University, South Korea*

The intestive daily use of chemcials in urban life have led to the extensitve occurrences of micropollutants in adjacent watershed. Stormwater event would convey massive amount of the pollutants to surrounding watersheds, impacting the water quality. Most of stormwater runoff studies have focused the transport dynamics of pollutants mainy for legacy pollutants whereas nontargeted unexpected substances were remained unknown for the dynamics. In this study, we identified micropollutants in urban receiving water during stormwater event using non-target screening (NTS) method based LC-HRMS. Intensive composite sampling was conducted to delineate the occurrence trends during stormwater events (~1h interval). In order to detect as many pollutants as possible, the SPE method using a multi-layer cartridge was applied. The configured non-target interset was grouped by using hierarchical statistical cluster. We classified non-target peaks into two groups: 'first-flush' and 'stormwater event'. The peaks were evaluated by the concentration trend of highthoroughput target screening (n=179). Total 30,782 non-target peaks configured 18,795 profiles. Among them, highly ranked 10,000 profiles were prioritized for data analysis. After criteria cutoff (peak intensity > 1E+05, detection frequency >70%) 2,348 profiles were remained. Finally, 456 profiles was clustered into 'first-flush' group including 18 targets (e.g., metformin, valsartan, HMMM, TPPO, TBEP). Among non-target peaks in the group, 62 profiles was part in homologous series of C2H4O for polyethylene glycols. Sulfonated polyethylene glycol (e.g., C10H20O5SO4, C12H24O6SO4, C14H28O7SO4) was tentatively identified as the major stormwater relevant pollutants. Among identified substances categorized as first flush influenced, the mass first flush curve was drawn to evaluate their potential role as non-point pollutants in urban water. This result suggested that the occurrence trend of stormwater pollutants is highly affected by meteorological factors such as precipitation intensity. Lots of glycol homologue series was also concerned as potential non-point source pollutants in urban watershed during stormwater event.

Biodegradation of Bisphenol F and S with Activated Sludge: Kinetics and Biotransformation

***Ana Kovacic**[1], Celine Gys[2], Martin Gulin[1], Tjaša Gornik[1], Tina Kosjek[1], David Heath[1], Adrian Covaci[2] and*

Ester Heath[1], (1) Jozef Stefan Institute, Slovenia, (2) University of Antwerp, Belgium
Bisphenol F (BPF) and bisphenol S (BPS) are becoming widespread in the environment, with both compounds exhibiting a similar endocrine-disrupting activity similar to bisphenol A; however, information regarding their fate during wastewater treatment and in the environment is scarce. This study assessed the biodegradation kinetics of BPF and BPS during biological wastewater treatment with activated sludge using GC- MS/MS and accurate mass high-resolution mass spectrometry LC-QTOF-MS for the identification and characterization of their biotransformation products. The studied conditions better reflect the natural environment or conditions in a wastewater treatment plant, where microbial consortia degrade organic pollutants rather than a single bacterial strain. The results showed that BPF and BPS degrade readily and are unlikely to accumulate in biosolids or wastewater effluent (c_i = 0.1 mg L^{-1}, half-lives < 4.3 days). The first-order kinetic model revealed that BPF (k_t = 0.20–0.38) degrades faster than BPS (k_t = 0.04–0.16) and that degradation rate decreases with an increasing initial concentration of BPS (half-lives 17.3 days). The absence of any additional organic carbon source significantly slowed down degradation, particularly that of BPS (lag phase on day 18 instead of day 7). The suspect and non-target based screening, using a machine-learning algorithm, identified one novel and confirmed three previously reported biotransformation of BPF and ten new and one existing biotransformation products of BPS. Based on these data, possible novel biotransformation pathways were postulated, namely sulphation, methylation, cleavage, and the coupling of smaller bisphenol moieties. These results provide new insights into the fate of bisphenols in the environment and their removal during wastewater treatment.

De Facto Water Reuse I: Development of Non-Targeted and Suspect Screening Workflows for Chemical Prioritization in Environmental Water Samples

***Angela Batt**[1], Laura Brunelle[2], Susan T. Glassmeyer[1], Dana Kolpin[3], Edward Furlong[3], Marc Mills[1], David Alvarez[3] and Diana Aga[4], (1) U.S. Environmental Protection Agency, (2) University at Buffalo, (3) U.S. Geological Survey, (4) State University of New York, Buffalo*
The analysis of contaminants of emerging concern (CECs) has been an ongoing priority due to their potential for adverse effects on human health and wildlife. With many tens of thousands of chemicals in use today, prioritizing specific chemicals or mixtures of chemicals that should be targeted for analysis has become increasingly difficult. We have developed non-targeted chemical screening tools to look for a broad range of possible CECs present in environmental waters using liquid chromatography and quadrupole/time of flight mass spectrometry with electrospray ionization (LC-QTOF-EI). The applications of this method will serve as a first step in prioritizing CECs for future targeted methods development and determining potential risk to human health and aquatic ecosystems. Two chromatographic techniques were evaluated, including reverse phase C18 and hydrophilic interaction liquid chromatography (HILIC) to capture the more polar CECs that may be present in water. Optimal method properties needed for data analysis and assessment workflow, including blank subtraction, retention time monitoring, replicate injections, peak alignment, and sequence set up with sample randomization were investigated. An extensive suspect screening library of over 1,400 possible CECs, including pesticides, pharmaceuticals and personal care products, illicit drugs and drugs of abuse, and various anthropogenic markers was developed to compliment the non-targeted workflow. The suspect screening library was created using analytical reference standards to produce a searchable database with retention times and experimental fragmentation data for each compound. The non-targeted methods were developed using environmental samples collected with polar organic chemical integrative samplers (POCIS), and these samples included surface water, wastewater treatment plant effluent, untreated drinking water, and treated drinking water. The corresponding analysis of these samples with the suspect screening standards were used to optimize method conditions and assess performance.

De Facto Water Reuse II: Application of Non-Targeted and Suspect Screening for Studying the Fate and Transport of Contaminants of Emerging Concern in Environmental Waters

***Laura Brunelle**[1], Angela Batt[2], Susan T. Glassmeyer[2], Dana Kolpin[3], Edward Furlong[3], Marc Mills[2] and Diana Aga[4], (1) University at Buffalo, (2) U.S. Environmental Protection Agency, (3) U.S. Geological Survey, (4) State*

University of New York, Buffalo
Wastewater is a point source for many contaminants of emerging concern (CECs) and surface waters receiving their discharge often serve as source water for downstream drinking water treatment plants (DWTPs). Non-targeted chemical analysis methods and suspect screening tools have been developed to capture a range of CECs in environmental water samples using liquid chromatography and quadrupole/time-of-flight mass spectrometry with electrospray ionization. Our workflow has been applied to residence-time weighted samples collected on three separate seasonal sampling events throughout a watershed in Concord, Massachusetts. This sampling design included grab samples collected upstream, at a wastewater treatment plant, the receiving surface water, untreated drinking water, and treated drinking water from the downstream DWTP. The main goal of this work was to examine the occurrence, fate, and transport of CECs throughout the watershed, and prioritize chemicals that occur frequently, or survive drinking water treatment for identification and future targeted method development. Positive and negative ionization modes were both assessed with 1,052 and 888 compounds detected across the three sets of samples collected, respectively. Comparisons show 86 compounds detected by positive ionization in all three sampling seasons/flow conditions, with 94 compounds being detected by negative ionization in all seasons/flow conditions. On average, 84% of positive mode detections and 81% of negative mode detections were highest in abundance in the effluent samples, with 12% and 25% (respectively) detected in only the effluent samples. Fewer than 100 compounds were found to persist through drinking water treatment in at least two seasons/flow conditions, with 18 of those compounds being present in all three sampling periods. Suspect screening tools were applied as a first step to identify compounds detected in the watershed. An in-house database was created for pesticides, pharmaceuticals and personal care products, illicit drugs, drugs of abuse, and anthropogenic markers using over 1,400 analytical standards. The suspect screening database was searched against the detected compounds to confirm identities based on retention time, exact mass, and the collected experimental fragmentation patterns. Utilization of the in-house database provided identification for over 80 of the detected compounds, with further compounds tentatively identified using additional available databases.

Ecological Risk Assessment for Chemicals of Emerging Concern in Reclaimed Wastewater Infiltrated into Groundwater

Kate McPeek[1]*, Berit Bergquist*[1]*, Brian Church*[1]*, John Toll*[1] *and Wendy Steffensen*[2]*, (1) Windward Environmental LLC, (2) LOTT Clean Water Alliance*

The LOTT Clean Water Alliance (LOTT) provides wastewater treatment and management for the urban areas of Lacey, Olympia, and Tumwater in Thurston County, WA (at the southern end of Puget Sound). LOTT's long-range plan relies on the production and beneficial use of reclaimed water, including the infiltration of unused reclaimed water into groundwater. A multiyear reclaimed water infiltration study was recently conducted to address questions about chemicals of emerging concern (CECs) that might remain in reclaimed water (e.g., pharmaceuticals, per- and polyfluoroalkyl substances [PFAS], and other chemicals found in household and personal care products). One of the study tasks was to conduct an ecological risk assessment (ERA) to evaluate the potential risks to aquatic receptors that utilize nearby water bodies with groundwater inputs containing reclaimed water. The general lack of ecological toxicity data for many CECs poses a challenge to risk assessors. In the present study, CECs were screened—based on guidance from the US Environmental Protection Agency (EPA) and the Toxic Substances Control Act—to identify chemicals of potential ecological concern (COPECs) for detailed risk evaluation. The screening procedure included a two-part evaluation to 1) compare concentrations of chemicals detected in reclaimed water to chronic screening-level benchmarks derived from a variety of sources, including EPA's Ecological Structure Activity Relationships model and ECOTOX database; and 2) evaluate the persistence and bioaccumulative potential of each chemical using half-lives and bioaccumulation factors from EPA's Estimation Program Interface Suite software. Chemicals with benchmark exceedances and those considered highly persistent or bioaccumulative were classified as COPECs. The list of COPECs was then refined using a groundwater fate and transport model; the refined list of COPECs, which included four PFAS, was evaluated in detail in the ERA exposure analysis

and effects and risk characterizations. The conclusion was that LOTT's proposed use of reclaimed water for groundwater recharge does not pose unacceptable risk to aquatic-dependent organisms using nearby streams where groundwater discharges. The methods used in this ERA can be applied to other studies evaluating the potential ecological risks of CECs.

Identification and Environmental Occurrence of the Ozonation Transformation Products of Tire Rubber Antioxidant 6PPD

Haoqi Zhao[1]*, Ximin Hu*[1]*, Zhenyu Tian*[2]*, Katherine Peter*[1]*, Melissa Gonzalez*[1]*, Craig Rideout*[1]*, Michael Dodd*[1] *and Edward Kolodziej*[1]*, (1) University of Washington, (2) Northeastern University*

Contaminants in urban stormwaters are problematic due to their potential to adversely impact ecosystem health in receiving waters. It is a clear research priority to identify unknown contaminants of emerging concern in urban stormwater. Recently, our group identified 6PPD-quinone from tire rubbers that is highly toxic towards coho salmon (*Oncorhynchus kisutch*) in the U.S. Pacific Northwest. 6PPD (*N*-(1,3-dimethylbutyl)-*N*′-phenyl-*p*-phenylenediamine) is ubiquitously used as an antioxidant in tire rubbers at mass fractions of 0.4-2%, where it reacts with ground-level ozone to prevent oxidation of rubber elastomers. Beyond 6PPD-quinone, 6PPD is transformed into a suite of ozonation transformation products (TPs) during the product life that are currently uncharacterized in the environment. In this study, we conducted gas phase ozonation of 6PPD, identified 6PPD ozonation TPs, and quantified (for level 1 TPs) or detected (for major TPs without analytical standards) their occurrence in multiple environmental media. 24-hour heterogeneous gas-phase ozonation was conducted in a column reactor under both dark and room light conditions. TP structures were proposed based on high resolution mass spectrometry (HRMS) evidence and confirmed with authentic standards when available. Tandem mass spectrometry (MS/MS) methods were developed to quantify/detected TPs in representative environmental media, including tire tread wear particles (TWPs), roadway runoff, and creek stormwater samples.

Non-Target (and Target) Screening of Priority Organic Chemicals in Recycled Plastics in Canada

Leah Chibwe[1]*, Derek C.G. Muir*[2]*, Amila De Silva*[2]*, Christine Spencer*[2]*, Camilla Teixeira*[2]*, Mary Williamson*[2] *and Xiaowa Wang*[2]*, (1) Oregon State University, (2) Environment and Climate Change Canada*

The increasing global plastic consumption and waste generated has encouraged sustainability initiatives and programs, including demand for recycled plastic. However, there are growing concerns that hazardous chemicals could be introduced into recycling streams, and ultimately into repurposed plastic products. Our objective was to characterize contaminants in various recycled products including both flakes (cleaned and chipped original recycled materials) and pellets (produced by heating and extrusion of flakes). We obtained samples from several Canadian recyclers and companies, including low- and high-density polyethylene pellets and flakes, polypropylene flakes, high impact polystyrene pellets, and polyethylene terephthalate pellets. We conducted non-target analyses (NTA) using comprehensive two-dimensional gas chromatography coupled to time-of-flight mass spectrometry (GC×GC/TOF-MS) and liquid chromatography Orbitrap high-resolution mass spectrometry (LC-HRMS). Additionally, we used target analyses to screen for halogenated flame retardants (e.g., polybromodiphenyl ethers), perfluoroalkyl acids, organophosphate esters and metal(loids) using GC-negative ion MS, LC/MS/MS and inductively coupled plasma-mass spectrometry. We identified over 100 chemicals, including polycyclic aromatic compounds, benzoic acid esters and benzoates, using mass spectral library matching to the NIST EI library (>80%) in GC×GC/TOF-MS. Furthermore, we detected many known commercial and industrial antioxidants, stabilizers or plasticizers, such as bisphenol A, octabenzone and several phthalates using LC-HRMS (accurate mass/fragmentation patterns and mzCloud database). In total, we tentatively identified 233 unique features using NTA as level 2 (probable structure) compounds, with on-going work to confirm the detected structures (i.e., level 1). Overall, target analyses showed that the concentrations of detected chemicals were generally 10^3 to 10^6 -fold lower than guideline limits (e.g., 0.1% or 1 mg/g) for Substances of Very High Concern in products in Europe, i.e., would be considered residual trace contaminants and impurities. However, given the wide range of chemicals identified/detected, albeit at low levels, we

recommend further evaluation of a wider range of recycled plastics and consideration of additive effects. The results of this study are of use to the broad scientific community, including regulators and the recycling industry.

Simultaneous Targeted and Non-Targeted Analysis of Chemical Contaminants in Electronic Waste-Impacted Soil

*Jingyun Zheng[1], Nil Basu[1], Julius Fobil[2] and **Stéphane Bayen**[1], (1) McGill University, Canada, (2) University of Ghana*

Rapid technological advancement and the growing position of electronic products in modern society have resulted in increasing amounts of electronic waste (e-waste) globally. Both the improper disposal of e-waste and e-waste dismantling activities have been identified as sources of environmental pollution by toxic metals or by well-known persistent organic pollutants such as polybrominated diphenyl ethers. However, little is known about the myriad of other plastic-related chemicals (unreacted monomers, additives, degradation products) which could virtually leach out of e-waste and their occurrence around e-waste dismantling sites. There is a need to expand the chemical coverage of surveillance tools to fully assess the environmental impact of e-waste dismantling activities. In this study, an analytical method was developed and validated for the simultaneous targeted analysis and suspect screening of plastic-related contaminants in soil. Freeze-dried soil samples were extracted with acetonitrile in a sonicator. Extracts were filtered, reconstituted in methanol, and analyzed with liquid chromatography coupled to quadrupole time-of-flight mass spectrometry (LC-QTOF-MS) in full scan mode. A suite of phthalates and bisphenols was selected as targets to assess the overall performances of the method (e.g. recovery, matrix effect, limit of detection, precision, etc.). The validated method was then applied to 53 actual soil samples from an e-waste site in Agbogbloshie (Ghana) and from neighboring agricultural fields. Relatively high levels of bisphenol A (up to ~ 20,000 ng/g dry weight of soil) and of various phthalates (up to ~ 26,000 ng/g dry weight for Bis(2-ethylhexyl) phthalate) were recorded at the e-waste site. In addition, relatively higher levels of these same pollutants were measured in the nearby agricultural sites compared to control agricultural sites. A non-targeted workflow was then developed to investigate contamination patterns across e-waste and agricultural sites, in order to identify a tentative list of pollutants released by e-waste into the nearby environment. A new list of plastic-related contaminants, previously unreported in e-waste surveillance studies, was established using this non-targeted approach.

Using Nontargeted Screening High Resolution Mass Spectrometry to Assess Coastal Groundwater Vulnerability to Flood Events

Hayden Rudd[1]*, Elizabeth Nichols[1], Damian Shea[2], David Genereux[1], Andy Neal[3] and Summer Xia[2], (1) North Carolina State University, (2) Statera Environmental, Inc., (3) NC Department of Environmental Quality*

Communities in the North Carolina Coastal Plain (NCCP) depend on safe and reliable groundwater for private domestic wells, agriculture, and industry. As rising global temperatures result in more frequent and intense storms, we need to understand the impacts of these events on the water quality of wells in coastal aquifers. Nontargeted screening high resolution mass spectrometry (HRMS) was employed to develop the first comprehensive organic chemical profiles of NCCP monitoring wells and to investigate groundwater vulnerability to flood events. Pre- and post-flood groundwater samples from the NCCP were collected around Hurricane Florence (in 2018) and subsequent extreme flood events over the next two years. These groundwater samples were analyzed for regulated semi-volatile organics using standard gas chromatography-mass spectrometry (GC-MS) following EPA Method 625/8270D. The same sample extracts were then analyzed for organics by GC-HRMS using a nontargeted and suspect screening workflow. Additionally, tritium analysis was conducted for a subset of these groundwater samples to determine the presence of young water (recharged after 1953) in the wells. The paired MS analyses were completed for 150 groundwater samples, and tritium analysis was completed for 67 samples. While targeted GC-MS analyses frequently resulted in no detections, the higher sensitivity GC-HRMS method detected an average of ~5,200 chemical features per sample. Across all groundwater samples, a total of 396 unique chemicals were tentatively identified using the NIST 20 mass

spectral database (M^1). The identified chemicals included aromatic hydrocarbons, pesticides, and phthalates, among other classes. Regardless of well depth, the number of chemicals of emerging concern detected in post-flood samples increased more frequently for well sites that experienced flooding. Chemical profiles of post-flood samples included several regulated organic compounds that were not detected prior to flooding. Tritium was not detected in deep wells (>100 m) that did not experience flooding, but ~73% of deep wells that experienced flooding had detectable tritium concentrations. The HRMS and tritium analyses suggest the intrusion of young water to deep wells in confined aquifers and the vulnerability of NCCP wells to contamination from flood events. Further, this study demonstrates the utility of nontargeted screening HRMS as a tool for advancing knowledge of groundwater dynamics and vulnerability.

Workflow and Software Advances

Keynote: Open, Dynamic Databases, Workflows and Transformations to Support Non-Target Analysis
***Emma Schymanski**, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Luxembourg*
The multitude of chemicals in our environment is ever increasing, with over 111 million chemicals in the largest open chemical databases, over 350,000 in global use inventories, and over 70,000 estimated to be in household use alone. Detectable molecules can be captured using non-target high resolution mass spectrometry (HRMS), which provides a 'snapshot' of all chemicals present in a sample and allows for retrospective data analysis through digital archiving. However, there is no 'one size fits all' data or analytical method, and scientists cannot yet identify most of the tens of thousands of features in each sample, leading to critical bottlenecks in identification and data interpretation. It is, however, getting much easier to find, identify and interpret 'known unknowns' quickly. Despite the size of the open chemicals databases, critical knowledge gaps still exist in them, with environmental transformation products and metabolites forming a key part of these missing puzzle pieces. Defining the chemical space to search and prioritizing efforts to find significant environmental chemicals, metabolites or transformation products as well as filling the knowledge gaps are the key to transition non-target HRMS into routine applications. This involves reconciling complex samples with expert knowledge and careful validation in an automatable and reliable manner. Thus, this talk will cover European and worldwide community initiatives and resources to help connect environmental expert knowledge and observations towards generating real, interpretable outcomes from non-target data. Various open cheminformatics and computational mass spectrometry approaches including patRoon, MassBank, MetFrag and PubChemLite for Exposomics will be covered, including how FAIRifying transformation information in the NORMAN Suspect List Exchange and PubChem helps close database gaps and create new possibilities for dynamic suspect screening efforts via patRoon. The performance improvements associated with PubChemLite and use of related metadata for interpretation of sample results will also be presented. This talk will show how interdisciplinary efforts and data sharing can facilitate research in environmental sciences, metabolomics, exposomics and beyond, towards a not-to-distant future of routine non-target HRMS. The very many contributors to these collaborative efforts will be acknowledged throughout the talk.

MSNovelist: De Novo Structure Generation from Mass Spectra
***Michael Stravs**[1], Kai Dührkop[2], Sebastian Boecker[2] and Nicola Zamboni[3], (1) Eawag - Swiss Federal Institute of Aquatic Science and Technology, Switzerland, (2) Friedrich-Schiller-University Jena, Germany, (3) ETH Zurich, Switzerland*
Structure elucidation remains a key challenge in mass spectrometry and metabolomics. Searches with spectral libraries and structure databases cannot detect truly novel, unexpected structures. Instead, strategies for unknown identification rely on *in silico* database expansion, substructure identification, or chemical class prediction. Recent deep learning methods for directed generation of chemical structures make it possible to query unknown chemical space without extensive candidate enumeration. We propose a two-step method for *de novo* structure elucidation: Molecular fingerprints are predicted from MS^2 spectra using CSI:FingerID and used as an input to a recurrent neural network (RNN) to predict chemical structures as SMILES codes. Using beam

search, up to 128 sequences are generated per fingerprint and ranked using the modified Platt score. Importantly, the neural network is trained with molecular fingerprint data, bypassing the bottleneck of limited spectral training data.*MSNovelist* was evaluated in a structure-disjoint cross-validation setup. For 3863 MS^2 reference spectra from the GNPS library, *MSNovelist* retrieved the correct structure for 45% of instances and identified it as the top candidate in 25% of instances (compared to 39% for database search). For structures identified correctly by database search, *MSNovelist* identified 61% correctly and retrieved 68%. In contrast, a naïve model generating structures without fingerprint input retrieved only 31% (GNPS) or 37% (DB top1) of structures. In structure similarity metrics, *MSNovelist* was able to consistently outperform training set data and match database search results. As a proof of concept application, the method was applied to discover novel metabolites in a public bryophyte dataset. For seven instances in 576 spectra, the *de novo* generated structure strongly outscored the best database candidates, indicating potential novel metabolites. A feature with m/z 381.1020 was tentatively annotated as a novel polyphenol with a flavonoid core ($C_{21}H_{16}O_7$). The proposed structure explains all observed fragments, whereas the database search result is inconsistent with two peaks. For four of seven cases, the *de novo* structure can explain the spectrum better than the database candidates and for one case equally well, thus providing candidates for further validation. The model is available open-source under https://github.com/meowcat/MSNovelist and as a Docker container.

Introducing Molecular Hypernetworks for Exploration of Multi-Dimensional Metabolomics Data
*Emilie Purvine[1], Aivett Bilbao[1], Corey Broeckling[2], Sean Colby[1], **Cliff Joslyn**[1], Andy Lin[1], Thomas Metz[1] and Madelyn Shapiro[1], (1) Pacific Northwest National Laboratory, (2) Colorado State University*
High resolution mass spectrometry (MS) coupled to liquid or gas chromatography (LC/GC) and ion mobility spectrometry (IMS) separations provide insight into sample composition to address the challenge of complete annotation of molecules in untargeted metabolomics. But computational tools are necessary to exploit these extremely complex data for metabolite annotation. Molecular network (MN) approaches, such as Global Natural Products Social Molecular Networking (GNPS), are an increasingly popular computational strategy for visualizing and interpreting complex tandem mass spectrometry (MS/MS) spectra. While powerful, this approach misses potential multi-way similarities among groups of MS/MS spectra. Moreover, it does not consider similarities in separation dimensions (e.g., retention time for LC or GC, drift time or collision cross section for IMS). Our approach greatly extends MN by (1) considering drift time, retention time, and precursor m/z in addition to MS/MS spectrum to analyze multi-dimensional similarity of ions, and (2) extending beyond graph edges to consider multi-way relationships among ions. We produce a hypergraph structure, a molecular hypernetwork (MH). Our MHs generalize MNs to represent similarities of arbitrary numbers of ions. Two ions are no longer either connected (similar) or not, as in an MN, but they can participate in multiple multi-way similarities in an MH. Annotations can also be propagated using the MH structure, but the existence of multiple multi-way similarities among pairs of ions provides added confidence of annotation propagation. We are evaluating our MH approach using reference mixtures of known molecular composition (e.g., Cambridge Isotope Laboratories QReSS™ Kit) and exploring its utility to characterize complex samples. In this talk we will describe our MH approach for characterizing complex samples and demonstrate an interactive MH visualization tool for exploring sample composition.

Estimating Uncertainty of Predicted Chemical Concentrations via Quantitative Non-Targeted Analysis
***Louis C Groff**[1], Jarod Grossman[2], Anneli Kruve[3], Jeffrey Minucci[1], Charles Lowe[1], James McCord[1], Dustin Kapraun[1], Katherine Phillips[1], Steven Purucker[1], Alex Chao[1], Caroline Ring[1], Antony Williams[1] and Jon Sobus[1], (1) U.S. Environmental Protection Agency, (2) Agilent Technologies, Inc., (3) Stockholm University, Sweden*
For decades, environmental studies have looked to targeted analysis methods for accurate quantification of chemical exposures. The applicability of targeted analysis methods is limited to chemicals for which there are readily obtainable analytical standards, thus inhibiting characterization of chemicals of emerging concern for which standards may not be readily available. To complement targeted methods, non-targeted analysis (NTA)

methods are now commonly used to identify tens to hundreds of chemicals in any sample of interest without prior knowledge of chemical presence within the sample. While NTA chemical identification methods are rapidly maturing, quantitative NTA (qNTA) methods have been somewhat slower to progress. Several qNTA methods are now able to generate point concentration estimates for identified chemicals, but none of these methods fully characterize the uncertainty of their estimates (*e.g.*, with confidence intervals). A clear need therefore exists for methods to bound chemical concentration estimates if NTA is to be utilized within regulatory contexts to support provisional risk characterizations. Here, the mathematical basis of traditional calibration methods used in targeted analyses is detailed, as well as procedures to extend these methods to support quantitative estimates for all chemicals observed in an NTA experiment. A non-parametric univariate method is detailed for bounding concentration estimates based on bootstrapping a distribution of instrument response factors for observed chemicals. A more complex method is then described, which utilizes predicted ionization efficiencies (calculated from machine-learning models using physicochemical descriptors) and linear mixed-effects models. The prediction uncertainty associated with both qNTA methods is compared to the uncertainty observed when performing semi-automated NTA experiments with traditional calibration curve-based approaches. We show that, in lieu of analytical standards, qNTA confidence limit estimation is feasible using either method given a base set of empirical response factors. *The views expressed are those of the author(s) and do not necessarily reflect the views or policies of the US EPA.*

Nontarget Mass Spectrometry and In Silico Molecular Characterization of Air Pollution from the Indian Subcontinent

Stefano Papazian[1]*, Lisa D'Agostino*[2]*, Ioannis Sadiktsis*[1]*, Jean Froment*[1]*, Bénilde Bonnefille*[1]*, Kalliroi Sdougkou*[1]*, Hongyu Xie*[1]*, Ioannis Athanassiadis*[1]*, Krishnakant Budhavant*[1]*, Sanjeev Dasari*[1]*, August Andersson*[1]*, Örjan Gustafsson*[1] *and Jonathan Martin*[1]*, (1) Stockholm University, Sweden, (2) Geosyntec Consultants, Sweden*

Airborne fine particulate-matter (PM2.5) is a major component of air pollution that contains unresolved complex mixtures of organic compounds. A deep molecular characterization of PM2.5 would contribute to improved understanding of air pollution impacts on global climate and human health. Here, we collected PM2.5 for 3 months at a remote receptor site in the Maldives that regularly intercepts polluted air outflows of the Indian subcontinent, and occasional pristine air from the Southern Indian Ocean. To achieve a broad molecular characterization of organic compounds, each PM2.5 sample (n = 40) was extracted with a range of solvents and analyzed by combining gas- and liquid chromatography high resolution mass spectrometry, i.e. (GC)- and (LC)-HRMS. A wide range of known anthropogenic contaminants were confirmed, including legacy persistent organic pollutants, polycyclic aromatic hydrocarbons (PAHs), plasticizers, flame retardants, pharmaceuticals, pesticides and associated transformation intermediates, some of which were detected in air for the first time (e.g. 2-hydroxy-atrazine). Unprecedented complexity was observed by nontarget analysis (60,030 features). However, only a minor proportion of features (0.5%) could be annotated with spectral databases (318 compounds) or identified with authentic standards (53 compounds at Level 1, and 36 compounds as closely related isomers) - thus leaving vast *'molecular dark matter'* structurally uncharacterized. An integrated workflow that leveraged on molecular networking and in-silico fragmentation predictions allowed us to reach a structural-level characterization for over 10,000 molecules, of which over 1,000 were hallmarks distinguishing polluted air of the Indian subcontinent from pristine air of the Indian Ocean. Clean air from the Indian Ocean primarily contained simple organooxygen molecules (Ox class; x=1-3) while polluted air was dominated by organodinitrogen (i.e. N2, N2Ox class), organosulfur and highly-oxygenated molecules (O4-O7 class). Overall, molecules associated with polluted regions occupied a broader and highly oxidized chemical space with physicochemical properties of high relevance to human toxicity and regional climate. Altogether, these results highlight how nontarget analyses and in-silico structure predictions can be implemented as advanced tools to explore deeper molecular-level insights and hypotheses on the health and climate impacts of complex organic compound mixtures in airborne PM.

Workflow and Software Advances (Poster)

Automated Low-Level Signal Capture in Two Dimensional Gas Chromatography High Resolution Mass Spectrometry

***Kristin Favela**, Michael Hartnett, William Watson, Jake Janssen, Keith Pickens and David Vickers, Southwest Research Institute*

Two Dimensional Gas Chromatography (GCxGC) leverages high resolution separation and is coupled with mass spectrometry (MS) for chemical identification. Hardware advances facilitating the availability of high resolution MS for GCxGC (e.g. GCxGC-HRT) have brought the promise of increased confidence in identification but also new challenges for efficient data management. The challenges associated with leveraging GCxGC-HRT for high throughput non-targeted analysis are associated with the deconvolution process necessary for assigning fragmentation ions to a given peak. This is because, unlike most LC workflows, ions are fragmented as they elute without isolation of a precursor ion. Complex samples present a wide dynamic range in chemical concentration. Many low-level signatures and individual fragment ions are incorrectly removed by the deconvolution process, partly because of mass drift that occurs with low signals. The result is poor library scores and incorrect identifications due to spectral distortion. In a worst-case scenario key diagnostic ions, such as unfragmented precursor ion, are incorrectly removed requiring time-consuming manual intervention to correctly extract the mass spectrum. Our team overcomes this complexity by combining iterative processing with machine learning (ML) to allow detection of low-level compounds otherwise missed by traditional peak finding algorithms. We leverage the information emergent from the batch to overcome the challenge of relying on peak finding and deconvolution. We first apply ML to automatically rank spectral signals into a quality hierarchy by first-pass peak finding. Next, a representative signal is selected according to quality at each chromatographic retention time in each sample. The high-resolution data is exploited by identifying the mass spectral fingerprint of each high-quality molecular feature. This fingerprint is leveraged in a second iteration of processing to extract quantitative information across the batch of samples by searching for specific ion signatures. The technique is evaluated with an EPA 8270 cocktail assayed in small concentration steps demonstrating the failure point for the commercial software routines. By contrast, our technique results in recovery of the targets well below the commercial failure point. The technique is also demonstrated in by spiking the 8270 targets into complex dust extracts, resulting in a higher true positive rate for our iterative process.

EPA's Non-Targeted Analysis WebApp: Combining Tentative Structure Identification with Risk Prioritization

***Matthew Boyce**, Alex Chao, Jeffrey Minucci, Steven Purucker, Antony Williams and Jon Sobus, U.S. Environmental Protection Agency*

Identifying chemical stressors present in various environments is necessary to evaluate potential exposures and their subsequent health risks. Targeted analysis has been the traditional method for chemical monitoring; however, emergent chemicals and their transformation products are often untenable for targeted analysis due to a lack of available reference standards. Technological and computational advancements have enabled increased focus on non-targeted analysis (NTA) methods, which utilize high-resolution mass spectrometry to identify chemicals without prior knowledge of their presence. While NTA provides greater chemical coverage, these methods require more rigorous data preparation and interpretation to reach confident identifications. To facilitate the adoption of NTA methods, the EPA has developed a web application, the “NTA WebApp”, that provides a transparent and user-friendly interface for processing both MS1 and MS2 NTA data. After input of NTA data, a number of processing steps are performed by the NTA WebApp: (i) features are filtered based on quality control parameters, (ii) multiple databases are queried to provide chemical identifications and associated meta data, and (iii) an output file is provided to the user which summarizes the processed feature and chemical results. The NTA WebApp is a continually evolving application, with recent developments including the incorporation of toxicity data drawn from EPA’s Hazard Comparison Dashboard. The Hazard Comparison

Dashboard serves as a comprehensive resource of available toxicity data reported for human health effects, ecotoxicity, and fate. As part of this demonstration, we will provide an example of how the toxicity data can be used to guide risk-based prioritization when interpreting chemical results. This abstract does not necessarily reflect agency policy.

Exploring the Unexplored: Characterizing Samples and Potential Exposure

***Abdullah Shouaib**, Jon Wahl and Jamie Nunez, Pacific Northwest National Laboratory*

Comparison of samples across different conditions traditionally requires some level of compound characterization upfront, which in turn relies on an often-incomplete reference library. This type of characterization can be problematic for novel compounds comprised within the sample that would otherwise aid in such an investigation. Moreover, this becomes extremely difficult when examining various complex samples across studies. We present a non-targeted workflow for the analysis of mass spectral (MS) data by extracting features from raw data with the intent to compare and/or characterize samples from unknown sources. Such a capability has the potential to help in many use cases such as (1) quickly comparing different conditions and which features to investigate further, (2) screen samples for signatures associated with illicit drugs and toxins, and (3) characterize new signatures for conditions of interest. Our workflow compares a variety of established machine learning tools, as previous studies have demonstrated some techniques are better at highlighting different inferences of MS data. Using this workflow, we investigate the utility of different models and when a specific model (or ensemble) may be preferred, using human tissue samples as an example.

In Silico Methods for the Identification of Wastewater-Based Emerging Contaminants via Non-Targeted Screening Approaches

***Lahiruni Halwatura**[1], Jonathan Antle[2], Scott Simpson[2] and Diana Aga[3], (1) University at Buffalo, (2) St. Bonaventure University, (3) State University of New York, Buffalo*

A large number of micropollutant types have attracted the attention of scientific communities in recent years to be listed as contaminants of emerging concern (CECs). Liquid chromatography coupled high-resolution mass spectrometry (HRMS) under a non-targeted approach plays a major role in identification of CECs in aqueous matrices. To avoid the bias between reporting environmental contaminants, a classification of confidence criteria has been introduced. Accordingly, the highest level of confidence (level 1) requires a match in the chromatographic retention time and mass spectra with a reference standard. However, it can be challenging to report all the detected compounds up to the highest identification level if the reference standards are not available or costly. Herein we introduce an in silico-based approach to increase the confidence of compounds that fail to be recognized up to the highest confidence. In the current work, quantitative-structure retention relationship (QSRR) methods have been implemented for the prediction of chromatographic retention. In order to construct the regression models, different molecular descriptors such as solubility of molecules at varying states of the chromatographic gradient program and octanol-water partition coefficient values were taken into consideration. The utilized method involved an in silico software package, COSMOtherm, which is based on "conductor-like screening model for real solvents" (COSMO-RS) theory to calculate the molecular descriptors. Analysis of compounds was performed using a QExactive Focus LC-Orbitrap^{TM} operated under positive electrospray ionization mode with the use of Cortecs $C18^{+}$ 2.1 × 150 mm (2.7 μm particle size) analytical column in reversed-phase. The training set and test set consisted of compounds with diverse functional groups in their chemical structures. A total of 83 level 2 assigned compounds including natural products, pharmaceuticals, human metabolites, personal care products, and industrial chemicals were used to test the regression models. COSMO-RS to provide further confidence in level 2 identified compounds was successfully demonstrated in this work. Thus, depending on the collected mass spectral data the COSMOtherm model introduced in this study can be useful to elevate the identity of compounds that lack the requirements to be confirmed up to the highest identification level.

Non-Target Workflow for Detecting the Occurrence of Xenobiotic Fingerprints in Surface Waters
Mulatu Yohannes Nanusha, Emil Frøkjær and ***Martin Metabolomics Hansen****, Aarhus University, Denmark*
The release of chemicals to the aquatic environment is increasing from day to day due to their elevated application in everyday life to meet the needs of humankind. On the other hand, the characterization of chemical entities of the aquatic environment remains challenging. Advancements in computational data processing and integrated software platforms have enabled the increasing use of liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS) for the identification of chemical entities in environmental samples. Even though HRMS provides spectral data which helps for the structural elucidation of unknown contaminants, the interpretation of spectra without any prior information is complicated and requires extensive data analysis and mining processes. Moreover, the selection or deselection of candidate chromatographic peaks depends on the workflow used. Thus, in this study, we developed a non-target workflow comprising multistep peak prioritization in combination with various software tools (e.g. mzCloud, MetFrag, MassBank, and SIRIUS4) and compound database (e.g. PubChem and ChemSpider) for the structural determination of unknown contaminants in surface water. The workflow was applied in water samples collected from three locations in Denmark; namely, two creeks near Faxe and Ringsted with urban and agricultural input, and various surface waters from Jutland. Analytes were enriched using solid-phase extraction (HLB) and eluted with methanol, which was later reconstituted in 5 % methanol, prior to chemical analysis. The nanoLC-HRMS raw data were subjected to peak detection followed by peak prioritization by employing a lower cut-off intensity and blank correction to remove the noise and background contaminants. Further clearing of false positives and peak filtering using isotopic pattern and peak shape, allowed us to isolate several thousands of chemical entities of interest, and of these, the chemical structure for around 250 peaks was identified to level 1 (confirmed structure) and level 2 (probable structure) of identification confidence. Compounds from different environmental applications such as pesticides (fluopyram, mecoprop, and metazachlor), herbicide (2-methyl-4-chlorophenoxyacetic acid), and pharmaceuticals (citalopram, furosemide and fexofenadine) were among the identified (to level 1) compound classes. Current activities focus on further identification, quantification, and in-vitro androgen receptor toxicological assay.

Nuclear Magnetic Resonance Spectral Similarity in Structural Clusters of Pesticide Chemical Space
Jessica Bade*, Jamie Nunez, Kabrena Rodda and John Cort, Pacific Northwest National Laboratory*
Chemical space, in brief, refers to the expansive and diverse molecular universe or a subset thereof. The ability to effectively define and fully harness such a resource has many potential applications (e.g., decoy molecule generation, novel compound discovery, and property prediction) that will aid in nontargeted identification and sample characterization approaches. A recently reported black-box approach applies hierarchical clustering to 2D descriptors of new psychoactive substances (NPS) and validates with nuclear magnetic resonance (NMR) spectra of structurally and spectrally representative molecules. This approach is a significant step forward, as having representative molecules available for reference enables faster traversal of theoretical chemical space. Here, we supply a comparable open-source, open-access method, which is essential to greater collaboration, development, and transparency. The proposed workflow utilizes Tanimoto similarity scoring, hierarchical clustering according to maximum common substructure, and selecting the medoid of a cluster as the representative molecule. We apply this workflow to a chemical space of obsolete pesticides and evaluate NMR spectral similarities between representative molecules and their respective clusters. Future applications include investigation and validation with fragmentation spectra.

Target-Decoy Strategy for Controlling False Discovery Rates in Structure Annotation of Small Organic Molecules by Computational Mass Spectrometry
Gordon Getzinger *and Lee Ferguson, Duke University*
Structure annotation of small molecules by high-resolution mass spectrometry (HRMS) provides insights into the occurrence and fate of unknown organic contaminants. Various algorithms rank molecule spectrum matches (MSMs) by predicting molecular fragmentation in tandem mass spectrometry, but such approaches generally

have unknown error rates. Therefore, approaches are needed for determining false discovery rates (FDR) in MSMs. Here we describe an approach for controlling error rates in MSMs using a target-decoy strategy. Target-decoy strategies compare score distributions of candidate molecules (i.e., targets) with those derived from molecules known to be incorrect (i.e., decoys). The target-decoy strategy is common to HRMS-based proteomics, in which reversed peptide sequences comprise decoy databases. However, few approaches are available for generating small molecule decoy databases. Here, we describe an approach to generating small molecule structure decoys based on the RDKit cheminformatics software and a custom Python library (SMdecoy). Our approach defines the target database as the set of known molecules for a given molecular formula present in the PubChem database. Next, decoys are generated by mutating and rearranging molecules. Mutation generates new molecules by swapping substructures in a molecule with isomeric substructures from the set of all target and decoy molecules. Rearrangement swaps moieties within a molecule to generate new structures. After mutating and rearranging the non-redundant set of target and decoy molecules are combined and the algorithm recurses until a pre-defined number of decoys are generated. Finally, the distribution of MSM scores for target and decoy structures are evaluated based on a pre-defined FDR. In this presentation we will present evaluation of the described target-decoy strategy using MSMs for tandem mass spectra from >400 unique structures using the computational MS tools MetFrag, MAGMa, and CFM-ID. Results demonstrate how FDR control can increase confidence in the high-throughput structure annotation of small molecules by HRMS.

Targeted Lists of Chemicals to Support Non-Targeted Analysis via the US-EPA CompTox Chemicals Dashboard

Antony Williams *and Charles Lowe, U.S. Environmental Protection Agency*

Non-targeted analysis (NTA) uses high-resolution mass spectrometry to better understand the identity of a wide variety of chemicals present in environmental samples (and other matrices). However, data processing remains challenging due to the vast number of chemicals detected in samples, software and computational requirements of data processing, and inherent uncertainty in confidently identifying chemicals from candidate lists. Analysis of the resultant mass spectrometry information relies on cheminformatics to identify and rank chemicals and the US EPA has developed functionality within the US-EPA CompTox Chemicals Dashboard (https://comptox.epa.gov/dashboard) to address challenges related to this analysis. These tools include the generation of "MS-Ready" structures to optimize database searching, retention time prediction for candidate reduction, consensus ranking using chemical metadata, and *in silico* MS/MS fragmentation prediction for spectral matching. A number of chemical lists (https://comptox.epa.gov/dashboard/chemical_lists) have also been developed to support specific applications including the identification of chemicals in human breath (i.e., the volatilome), chemicals in human media (e.g., blood, saliva), and in water. Other lists include PFAS (per- and polyfluoroalkyl substances) chemicals and toxins (e.g.. microcystins and cyanotoxins). Combining the synergies of a database containing ~900,000 chemicals with over 200 segregated chemical lists with dashboard search functionality into a comprehensive workflow provides a powerful tool to support NTA. This poster will review how the CompTox Chemicals Dashboard provides a freely available web-based application to support structure identification and NTA and how expansion of the types of data hosted in the database can support different targeted and non-targeted projects. *This abstract does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.*

The US-EPA CompTox Chemicals Dashboard - an Online Data Integration Hub Supporting Non-Targeted Analysis

Antony Williams *and Charles Lowe, U.S. Environmental Protection Agency*

High resolution mass spectrometry (HRMS) and non-targeted analysis (NTA) are advancing the identification of emerging contaminants in environmental matrices, improving the means by which exposure analyses can be conducted. However, confidence in structure identification of unknowns in NTA presents challenges to analytical chemists. Structure identification requires integration of complementary data types such as reference databases (either commercial or open databases), fragmentation prediction tools, and retention time prediction

models. One goal of our research is to optimize and implement structure identification functionality within the US EPA's CompTox Chemicals Dashboard, an open chemistry resource and web application containing data for ~900,000 substances. Database searching using mass or formula-based inputs has been optimized for structure identification using "MS-Ready Structures": de-salted, stripped of stereochemistry, and mixture separated to replicate the form of a chemical observed via HRMS. Functionality to conduct batch searching of molecular formulae and monoisotopic masses has also been implemented. While the increasing number of free online databases are of value to support chemical structure verification and elucidation there are known issues regarding data quality and careful data curation is a very necessary part of the development of these resources. This presentation will provide an overview of our latest enhancements to the dashboard to support mass spectrometry, incorporation of specific datasets (e.g., to support breath research (the volatilome) and household dust analysis), and the value of metadata and predicted fragmentation spectral matching to support structure identification. *This abstract does not necessarily represent the views or policies of the U.S. Environmental Protection Agency.*

Tools for the Measurement of the Uncertainty of Fragmentation Mass Spectra and Incorporation of Mass Spectra with Affiliated Metadata into an Open-Source Database

Benjamin Place *and Jared Ragland, National Institute of Standards and Technology*

The collection and analysis of fragmentation mass spectra for unknown compounds is an essential part of most nontargeted analysis (NTA) workflows; as it is related to the identification of the unknown compounds. Conventionally, fragmentation mass spectra that are collected from single samples are averaged over a chromatographic peak to create a single representative mass spectrum for each compound. The average mass spectrum is not representative of the measurement uncertainty of the mass spectrometer, and confidence of identification using the mass spectrum cannot be fully communicated without an understanding of this uncertainty. The presented study is an in-depth investigation of the measurement uncertainty of a Quadrupole-Orbitrap mass spectrometer as it is related to the generation of consensus mass spectra. Guidance regarding the estimation of measurement uncertainty and tools to compare mass spectra with associated uncertainties will be presented. In addition, the use of consensus mass spectra for compound identification depends on the data (and affiliated metadata) to be available and usable. The development of an open-source database infrastructure for the storage and management of chemical information, analytical method metadata, and consensus mass spectra will be presented. The novel database infrastructure will enable users to search known PFAS for potential identities of their unknown mass spectra, as well as provide context for the clear communication of the confidence of those identifications. Furthermore, by making the database scheme publicly accessible, researchers will be able to develop computational tools for their own novel investigations of the data. Per- and polyfluoroalkyl substances (PFAS) will be the model chemical class used for the demonstration of the uncertainty measurements and for the database. The current interest in PFAS provides a unique opportunity to explore the generation of consensus mass spectra without the availability of analytical standards, which is a significant issue within PFAS research, and the database provides a data tool for researchers share their own novel PFAS mass spectra with other laboratories.

Winning the Chemical Space Race: Introducing an Automated Application-Informed Platform for Optimized Compound Representation

Kate Schultz[1]*, Danielle Ciesielski*[1]*, Madison Blumer*[1]*, Javier Flores*[1]*, Gabrielle Black*[2]*, Karen Wahl*[1] *and Jamie Nunez*[1]*, (1) Pacific Northwest National Laboratory, (2) University of California*

Robust representations of chemical space—broadly defined as a region in which molecules are placed according to their descriptors—are highly desired in a diverse range of applications. Such models of chemical space have the potential to aid in the identification of new compounds and chemical moieties with activity at desired targets; predict toxicity and other undesirable characteristics; locate new and/or relevant reaction pathways; aid in the development of decoy libraries for false discovery rate assessment; and assist in the generation of novel compounds with desired properties. Currently, many challenges hinder the creation of robust chemical space

models. For example, model development is highly dependent on the application space and descriptors (e.g., chemical fingerprints, properties, and class) used. The descriptors chosen to define a chemical space have ramifications for training, visualization, and projection of new compounds onto a predefined model. A chemical space model that performs well for a particular application or target will likely not achieve comparable performance for other applications or targets. Given these challenges and considerations, it is feasible that an all-encompassing, application-independent chemical space representation is not attainable. As such, understanding how design considerations affect specific use case becomes critical. Here, we present an automated pipeline for building application-tailored chemical spaces. We provide and discuss a variety of iterations of chemical space design using different combinations of models, descriptors, and applications. Furthermore, we outline best-practice considerations for developing chemical space models, including those associated with (1) data limitations; (2) big data; (3) curse of dimensionality; and (4) model selection.

Society of Environmental Toxicology and Chemistry
Environmental Quality Through Science®

www.setac.org