























An international laboratory comparison of dissolved organic matter composition by high resolution mass spectrometry: Are we getting the same answer?^a

Jeffrey A. Hawkes ^{1,*a} Juliana D'Andrilli ^{2a} Jeffrey N. Agar^{3,4} Mark P. Barrow ⁵
Stephanie M. Berg ⁶ Núria Catalán ^{7,8} Hongmei Chen⁹ Rosalie K. Chu ¹⁰ Richard B. Cole ¹¹
Thorsten Dittmar^{12,13} Rémy Gavard ^{5,14} Gerd Gleixner ¹⁵ Patrick G. Hatcher⁹ Chen He ¹⁶
Nancy J. Hess¹⁰ Ryan H. S. Hutchins ¹⁷ Amna Ijaz¹⁸ Hugh E. Jones ^{5,14} William Kew¹⁰
Maryam Khaksari^{18,19} Diana Catalina Palacio Lozano ⁵ Jitao Lv²⁰ Lynn R Mazzoleni^{18,19}
Beatriz E. Noriega-Ortega ²¹ Helena Osterholz^{12,22} Nikola Radoman²³ Christina K. Remucal ^{6,24}
Nicholas D. Schmitt^{3,4} Simeon K Schum^{18,19} Quan Shi ¹⁶ Carsten Simon ¹⁵ Gabriel Singer ²⁵
Rachel L. Sleighter⁹ Aron Stubbins ^{3,26,27} Mary J. Thomas ^{5,14} Nikola Tolic ¹⁰ Shuzhen Zhang²⁰
Phoebe Zito²⁸ David C. Podgorski ²⁸

¹Department of Chemistry, BMC, Uppsala University, Sweden

²Louisiana Universities Marine Consortium, Chauvin, Louisiana

³Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts

⁴Barnett Institute of Chemical and Biological Analysis, Northeastern University, Boston, Massachusetts

⁵Department of Chemistry, University of Warwick, Coventry, United Kingdom

⁶Environmental Chemistry and Technology Program, University of Wisconsin-Madison, Madison, Wisconsin

⁷Catalan Institute for Water Research (ICRA), Girona, Spain

⁸Universitat de Girona, Girona, Spain

⁹Department of Chemistry and Biochemistry, Old Dominion University, Norfolk, Virginia

¹⁰Environmental Molecular Science Laboratory, Pacific Northwest National Laboratory, Richland, Washington

¹¹Sorbonne Université, Institut Parisien de Chimie Moléculaire, UMR 8232, Paris, France

¹²ICBM-MPI Bridging Group for Marine Geochemistry, Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University, Oldenburg, Germany

¹³Helmholtz Institute for Functional Marine Biodiversity, Carl von Ossietzky University, Oldenburg, Germany

¹⁴Molecular Analytical Sciences Centre for Doctoral Training, University of Warwick, Coventry, United Kingdom

¹⁵Molecular Biogeochemistry, Department Biogeochemical Processes, Max Planck Institute for Biogeochemistry, Jena, Germany

¹⁶State Key Laboratory of Heavy Oil Processing, China University of Petroleum, Beijing, China

¹⁷Department of Biological Sciences, University of Alberta, Edmonton, Canada

¹⁸Department of Chemistry, Michigan Technological University, Houghton, Michigan

¹⁹Chemical Advanced Resolution Methods Laboratory, Michigan Technological University, Houghton, Michigan

²⁰State Key Laboratory of Environmental Chemistry and Ecotoxicology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China

²¹Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

*Correspondence: jeffrey.hawkes@kemi.uu.se

Additional Supporting Information may be found in the online version of this article.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

^aAuthor Contribution Statement: J.A.H. and J.D. contributed equally to this work by organizing and managing the project: formulated research questions, objectives, experimental design soliciting collaborators and project management. J.A.H., J.D., G.G., and D.C.P. contributed in experimental sample preparation or IHSS sample procurement. J.A.H., R.L.S., H.M.C., P.G.H., A.I., M.K., L.M., R.K.C., N.J.H., L.V., S.Z., C.H., Q.S., R.H., D.C.P.L., M.B., H.O., T.D., C.S., G.G., S.M.B., C.K.R., N.C., R.B.C., B.N.-O., G.S., N.R., N.D.S., A.S., J.N.A., P.Z., and D.C.P. collected, facilitated, or supported the acquisition of mass spectra. J.A.H., S.S., N.T., R.G., H.E.J., and H.O. wrote the code for molecular formula assignments and metrics or performed statistical analysis. J.A.H., J.D., and D.C.P. composed the original draft of manuscript and figures. All authors revised the manuscript for publication

²²Marine Chemistry, Leibniz-Institute for Baltic Sea Research Warnemuende, Rostock, Germany

²³Department of Environmental Science, Stockholm University, Stockholm, Sweden

²⁴Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, Wisconsin

²⁵Department of Ecology, University of Innsbruck, Innsbruck, Austria

²⁶Department of Civil and Environmental Engineering, Northeastern University, Boston, Massachusetts

²⁷Department of Marine and Environmental Science, Northeastern University, Boston, Massachusetts

²⁸Pontchartrain Institute for Environmental Sciences, Department of Chemistry, Chemical Analysis & Mass Spectrometry Facility, University of New Orleans, New Orleans, Louisiana

Abstract

High-resolution mass spectrometry (HRMS) has become a vital tool for dissolved organic matter (DOM) characterization. The upward trend in HRMS analysis of DOM presents challenges in data comparison and interpretation among laboratories operating instruments with differing performance and user operating conditions. It is therefore essential that the community establishes metric ranges and compositional trends for data comparison with reference samples so that data can be robustly compared among research groups. To this end, four identically prepared DOM samples were each measured by 16 laboratories, using 17 commercially purchased instruments, using positive-ion and negative-ion mode electrospray ionization (ESI) HRMS analyses. The instruments identified ~1000 common ions in both negative- and positive-ion modes over a wide range of m/z values and chemical space, as determined by van Krevelen diagrams. Calculated metrics of abundance-weighted average indices (H/C, O/C, aromaticity, and m/z) of the commonly detected ions showed that hydrogen saturation and aromaticity were consistent for each reference sample across the instruments, while average mass and oxygenation were more affected by differences in instrument type and settings. In this paper we present 32 metric values for future benchmarking. The metric values were obtained for the four different parameters from four samples in two ionization modes and can be used in future work to evaluate the performance of HRMS instruments.

High-resolution mass spectrometry (HRMS) has become a central tool in the analysis of dissolved organic matter (DOM), due to seminal work reported over the last three decades (Fievre et al. 1997; Kujawinski et al. 2002; Stenson et al. 2002; Stenson et al. 2003; Dittmar and Koch 2006; Sleighter and Hatcher 2007; Reemtsma et al. 2008; Gonsior et al. 2009) and the detailed biogeochemical insight afforded by molecular compositional patterns in large and diverse sample sets (Flerus et al. 2012; Jaffé et al. 2012; Kellerman et al. 2014; Lechtenfeld et al. 2014; Hertkorn et al. 2016; Drake et al. 2019). Since Fievre et al. collected the first HRMS spectrum of DOM in 1997 (Fievre et al. 1997), there have been large increases in both the number of researchers using HRMS for DOM analysis and the variety of instrument types being employed. This upward trend has been facilitated by the propagation of commercial Fourier transform ion cyclotron resonance (FT-ICR), Orbitrap, and high-resolution quadrupole time of flight (q-TOF) mass spectrometers (Hawkes et al. 2016; Lu et al. 2018; Pan et al. 2020). While the application of HRMS for the molecular-level assessment of DOM continues to grow, leading to improved methods that are adopted by the research community, so do the challenges with data comparison and interpretation that arise from the use of instruments with different resolving powers (mass/[full width at half maximum] at m/z centroid; $m/\Delta m$), source conditions, ion optics, and users.

The wide range of potential conditions for analysis, resulting interpretation of the spectra, and the conclusions based on those interpretations poses potential problems with common interpretation across the community using different instruments, laboratories, and practices. As both climate and

ecosystems change (Williamson et al. 2015; Drake et al. 2019), it is critical for the community to understand the degree to which archived data can be compared to newer measurements, and if data generated among various research groups can be robustly compared and related to emerging global trends.

In the vast majority of recent DOM research, HRMS has been coupled to electrospray ionization (ESI), as this 'soft' ionization technique allows intact molecular ions to enter the mass spectrometer without fragmentation (Fenn et al. 1989; Henry et al. 1989; Novotny et al. 2014). With ESI-HRMS, signal response is not necessarily linearly related to the concentration of analyte (Kujawinski et al. 2002), or the linear response range may be very narrow. Signal magnitude is rather a function of concentration *and* ionization efficiency, but ionization efficiency can fluctuate as a function of the sample matrix (for example, the ionic strength, pH, and analytical complexity of the sample) (Tang and Kebarle 1993; Brown and Rice 2000; Oss et al. 2010) and the instrument tune settings. Owing especially to matrix signal suppression, direct infusion ESI-HRMS is considered to be a "qualitative," or at best (under controlled conditions), a "semi-quantitative" technique for untargeted analysis of complex mixtures (Dittmar and Koch 2006; Mopper et al. 2007; Dubinenkov et al. 2015; Liu and Kujawinski 2015; Luek et al. 2018). This implies that, when constructing response vs. concentration relationships, standard reference solutions need to be highly matrix-matched, but absolute analyte concentration cannot be quantified without standards. Even within the data acquisition, the signal response is not uniform across the m/z range. The combination of the user-defined ionization source conditions and HRMS instrument settings for ion optics and detection can

often favor low m/z values vs. high m/z values or vice versa—this translates to biases in detection and the potential for variability in relative ion abundance results among instruments (Cao et al. 2016; Kew et al. 2018a). Such biases may confound molecular-level assessments of DOM chemical assignments and interpretations and limit the ability to utilize datasets generated by different research groups, for example, for meta-analyses of ecosystem trends.

A further concern is the lack of “standards” for complex DOM mixtures that limits the possibilities for standardizing reproducible DOM HRMS results and enabling uniform assessment of instrument performance. Currently, the research community relies on ‘standard’ reference materials provided in large quantities by the International Humic Substances Society (IHSS), like Suwannee River fulvic acid (SRFA). These reference materials, described as a collection of humic acids, fulvic acids, and other isolates, are routinely collected and homogenized, and have been used throughout the history of HRMS, setting the foundation to understand complex DOM mixtures (Fiebre et al. 1997; Stenson 2008; Witt et al. 2009; Gaspar et al. 2010; Herzsprung et al. 2015; Kew et al. 2017; Simon et al. 2018). These reference materials are often used as measures of instrument performance in studies that analyze newly collected samples, either as a DOM comparison or to control HRMS settings (Koch et al. 2005; Mangal et al. 2016; Li et al. 2017; Hawkes et al. 2018b; Solihat et al. 2019). Even though the reference samples are highly processed to ensure an appropriate degree of batch-to-batch uniformity, they still contain a high level of chemical complexity, and are considerably more complex than other commercially available laboratory chemicals.

Previous HRMS studies have defined important metrics with respect to instrument performance and capabilities, such as signal to noise ratio, mass measurement accuracy, and resolving power (Marshall et al. 1998; Hawkes et al. 2016; Simon et al. 2018; Smith et al. 2018), have extended optimizations based on IHSS reference materials such as Suwannee River and/or Pony Lake Fulvic Acids (SRFA and PLFA) (Stenson et al. 2003; Koch et al. 2005; D’Andrilli et al. 2013, 2015; Mangal et al. 2016). Currently, however, laboratories have little access to raw data from other research groups, limiting detailed comparison of instrument performance and results.

The objective of this study was to investigate the differences in DOM assessments from the same samples across multiple instruments due to variations in the employed HRMS instrumentation and the optimized settings developed by different laboratories. This study is a first step toward understanding the interpretational variability caused by different laboratories’ established protocols, routines, and instrumental differences. The two goals of this study were (1) to evaluate whether different HRMS instruments can identify the same molecular-level trends in chemical composition from a range of samples processed by an identical formula assignment protocol and (2) to provide metrics (identified variables and calculated data ranges) for the HRMS DOM community that can be used for reference when developing new methods

or validating new datasets. For this work, a set of IHSS reference materials was used that represents a compositional gradient of DOM commonly found along freshwater-associated environmental continua. The current lack of a seawater reference material prohibited the inclusion of a sample that represented marine systems. Nonetheless, altogether, these IHSS reference materials represent important DOM endmembers that may relate to similar complexities found in marine environments. For example, PLFA and marine environments both represent microbially sourced end-members and therefore may share some DOM production pathways and measurable compositional similarities (Kellerman et al. 2018; Zark and Dittmar 2018).

Four identically prepared DOM samples were sent to 16 laboratories operating 17 commercially available HRMS instruments, including various magnetic field strengths of FT-ICR mass spectrometers and various models of Orbitrap mass spectrometers (Table 1). The samples were SRFA, PLFA, Elliot soil fulvic acid (ESFA), and Suwannee River Natural Organic Matter (SRNOM). Laboratories were instructed to tune their instrument to maximize the total signal for SRFA and then measure the samples in positive- and negative-ion mode ESI using their independently developed optimized settings. The data interpretation was conducted by one centralized method to exclude all molecular formula assignment variability that might otherwise originate from different data processing algorithms (detailed descriptions in Section 1.2).

After calibration and molecular formula assignment, the data (sum-normalized ion abundances) for all instruments were aligned based on formula assignments. The instruments were compared based on the number of ions with an assigned formula and by comparing ion abundance patterns among the four samples using statistical techniques. Further, the average values for the metrics were calculated for each sample in each ESI mode using common ions that were found by the majority of instruments (i.e., relative abundance weighted averages of oxygenation, O/C_{wa} ; hydrogen saturation, H/C_{wa} ; mass to charge ratio, m/z_{wa} and modified aromaticity index; $AI_{mod\ wa}$ [Koch and Dittmar 2016]). The calculated ranges of each metric are presented, along with a discussion of outliers relative to average results found in the study. The results can be used by the community to facilitate comparison of HRMS DOM data from both past and future data sets that use similar instruments and any of the same reference materials.

Materials and procedures

Materials

Sample selection

Four DOM reference materials were obtained from the IHSS (<http://humic-substances.org/>). Suwannee River fulvic acid (SRFA; 2S101F) and Suwannee River Natural Organic Matter (SRNOM; 2E101N) were collected from the blackwater Suwannee River, Georgia, which originates from the Okefenokee peat swamp. The fulvic acid sample was prepared by isolating DOM onto XAD-8 resin from filtered water that was acidified

Table 1 List of participating institutions, country of institution, and mass spectrometer instrument type including magnetic strength or model. All FT-ICRs are manufactured by Bruker Daltonics and use an Apollo II ESI source (Bremen, Germany), all Orbitraps are manufactured by Thermo Fisher and use an Ion Max API ESI source (Bremen, Germany).

Institution	Country	Instrument
Department of Chemistry, University of Alberta	Canada	9.4 T Apex Qe FT-ICR
Chinese Academy of Sciences, Beijing State Key Laboratory of Heavy Oil Processing China	China	15 T SolariX FT-ICR
University of Petroleum, Beijing	China	9.4 T Apex Ultra FT-ICR
Sorbonne Université, CNRS, Institut Parisien de Chimie Moléculaire, Paris	France	7 T SolariX FT-ICR
Institute for Chemistry and Biology of the Marine Environment	Germany	15 T SolariX FT-ICR
Carl-von-Ossietzky University, Oldenburg		
Max-Planck-Institute for Biogeochemistry, Jena	Germany	Orbitrap Elite
Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin	Germany	Orbitrap Velos
Catalan Institute for Water Research (ICRA), Girona	Spain	Orbitrap Velos
Department of Environmental Science and Analytical Chemistry (ACES)	Sweden	Orbitrap QE
Stockholm University		
Department of Chemistry—BMC Uppsala University	Sweden	Orbitrap Velos, Orbitrap QE
Department of Chemistry University of Warwick	United Kingdom	12 T SolariX FT-ICR
Department of Civil and Environmental Engineering University of Wisconsin-Madison	United States of America	12 T SolariX XR FT-ICR
College of Sciences Major Instrumentation Cluster (COSMIC), Old Dominion University	United States of America	12 T Apex Qe FT-ICR
EMSL for the USDOE	United States of America	12 T SolariX FT-ICR
Pacific Northwest National Laboratory		
Barnett Institute of Chemical and Biological Analysis	United States of America	9.4 T SolariX XR FT-ICR
Northeastern University		
Department of Chemistry Michigan Technological University	United States of America	Orbitrap Elite

to pH < 2, followed by elution with NaOH, and precipitation of “humic acids” at pH = 1.0. The fulvic acids remain in solution at this low pH. The SRNOM sample was prepared by reverse osmosis from the same site (Green et al. 2015). Pony Lake Fulvic Acid (PLFA; 1R109F) was prepared by the same procedure as SRFA at pH = 2.0 (± 0.1) and originates from Pony Lake, Antarctica, which is characterized by the absence of higher order plants. The site is dominated by autochthonous, microbially derived carbon (Brown et al. 2004). Elliot soil fulvic acid (ESFA; 5S102F) was prepared by the same protocol

as SRFA after soil homogenization of a fertile terrestrial soil commonly found in Illinois, Indiana, and Iowa. Detailed information regarding IHSS standard (SRFA, SRNOM, and ESFA) and reference (PLFA) material sample locations, preparation, and chemical data can be found on the IHSS website: <http://humic-substances.org/>.

Sample preparation

Each powdered sample (2 mg) was added to separate 2 mL combusted amber glass vials and dissolved in 2 mL of HPLC

grade methanol by vigorous shaking, creating stock solutions of individual reference materials. Using sterile pipette tips, 40 μL of each “standard” was pipetted into separate 2 mL combusted amber glass vials, dried under nitrogen gas, and capped (64 samples total). Sample volumes were calculated to achieve $\sim 20 \text{ mg L}^{-1}$ of carbon (C) once redissolved in 1 mL solvent (assuming 50% C). One sample set, containing SRFA, ESFA, PLFA, SRNOM, and an empty, sealed 2 mL combusted amber glass vial, was shipped to each of the 16 participating laboratories (80 vials total). Upon sample arrival, each participating laboratory prepared 50–100 mL ultraclean 50% methanol (LCMS grade methanol in 18.2 M Ω ultraclean water, sourced in each laboratory individually) in combusted glassware. One milliliter of clean 50% methanol was added to each sample vial using methanol-washed glass syringes. One milliliter of clean 50% methanol was also added to the empty vial and served as each laboratory’s methodological blank for the experiment. DOM samples were sonicated for 5 min to ensure complete dissolution prior to analyses. One Laboratory, M, diluted their samples to 5 mg L^{-1} C before analysis, while the others analyzed at 20 mg L^{-1} C. Samples were stored in the dark at 4°C pending MS analysis.

Procedures

Participating laboratory instructions

This study included data from 16 HRMS DOM community members in laboratories spanning eight countries of the Northern Hemisphere (Table 1). Each laboratory received a sample set prepared at Montana State University (Bozeman, Montana) and an identical set of instructions (established by Hawkes and D’Andrilli) specific to sample storage and reconstitution, method blank preparation, acquisition in positive- and negative-ion mode, and data exportation.

Each instrument was assigned a code letter from A-G (Orbitraps) and H-Q (FT-ICRs) for laboratory and instrument anonymity. This minimized focus on each instrument’s perceived performance, and instead placed emphasis upon the variability or similarity of the data.

Electrospray ionization and HRMS tune settings

The ESI sources were thoroughly cleaned with LCMS grade methanol or ultraclean 1 : 1 methanol : water and tuned to maximize the signal in negative- and positive-ion modes using the SRFA sample and according to each laboratory’s own optimization routines. All five samples were analyzed in the m/z range of 150–1000, and the resulting data were exported as mass lists (signal to noise ratio $[S/N] \geq 4$ for FT-ICR instruments and all data on Orbitraps) in comma delimited (.csv) format with three to four columns: “ m/z ,” “Intensity,” and “Resolving Power,” plus “Noise” for the Orbitrap analyzers. Five further data acquisitions of SRFA were carried out on

instrument A (Orbitrap) and L (FT-ICR) in order to assess intra-instrument variability.

Data processing

All data processing (mass spectral calibration and molecular formula assignments) was conducted in MATLAB version R2017b with the Statistics and Machine Learning Toolbox (v11.2), Bioinformatics Toolbox (v.4.9) and R (R v.3.6.1) with vegan package at Uppsala University, Sweden. The raw mass lists and data processing code are available in Data S1. Each sample was processed separately to recalibrate the mass axis, remove noise, and then assign molecular formulas to individual peaks (Fig. 1). This post-acquisition processing was executed on all data in an identical manner to minimize differences arising from differing software programs and algorithms used by various laboratories, and also eliminated potential user error, thereby ensuring consistency among the data sets.

Noise removal

Noise was assessed using the concepts from the “KMD slice” method from the R package MFAssignR (Schum et al. 2019), which was modified from Riedel and Dittmar 2014. Briefly, noise level is calculated based on signals with highly improbable mass defects that are likely to arise from electronic noise. In this case, the mass defects selected were calculated as a window of Kendrick mass defects (KMDs). Masses were converted to Kendrick mass (KM) based on CH_2 as $\text{KM} = \text{mass} \times 14/14.01565$, and the KMD was then computed by subtraction of KM from nominal mass. The KMD window for noise was taken as $0.0011232(\text{KM}) + 0.05$ to $0.0011232(\text{KM}) + 0.2$. All peaks in this noise window were collected, and the 99th percentile of their abundances was taken as the upper limit of noise, in order to allow the most intense 1% of these peaks to be considered as potential analytes. Mass spectra from samples that exhibited fewer than 100 peaks in the KMD window were not subjected to noise reduction treatment. In the rest of the samples, peaks throughout the spectrum below the upper limit of noise were removed (Fig. 1).

Theoretical formula list

To assign formulas, a theoretical neutral molecule formula list was generated based on the following constraints: C_{4-50} , H_{4-100} , O_{2-40} , N_{0-2} , S_{0-1} , $^{13}\text{C}_{0-1}$, $150 < m/z < 1000$, $0.3 \leq \text{H}/\text{C} \leq 2.2$, $0 < \text{O}/\text{C} \leq 1.2$, $\text{KMD} \leq 0.4$ or ≥ 0.9 , valence neutral (nitrogen rule), and double bond equivalents minus oxygen ($\text{DBE}-\text{O} \leq 10$) (Herzprung et al. 2014). Beyond $\text{C}_c\text{H}_h\text{O}_o$ containing molecular formulas, heteroatomic or isotopic formulas were allowed to contain one of the following: N_{1-2} , S_1 , or $^{13}\text{C}_1$. Formulas above m/z 500 were restricted from N_2 assignments. In positive-ion mode, the S-containing formulas were removed from consideration to reduce mis-assignments due to variable resolving power among instruments. In positive mode all formula masses were calculated as sodium cation adducts ($\text{C}_c\text{H}_h\text{O}_o\text{N}_n\text{Na}_1$). In negative-ion mode, theoretical formula

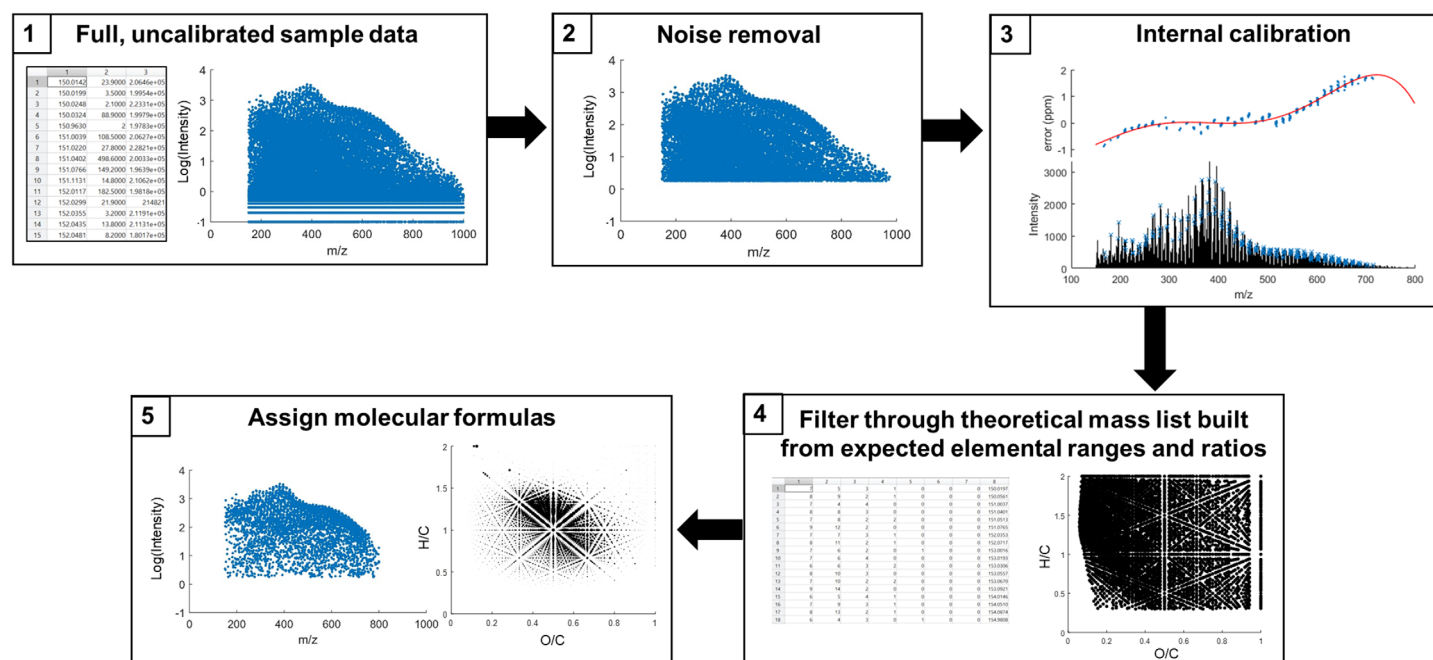


Fig 1 Schematic summarizing data processing steps from raw data acquisition to molecular formula assignment.

masses were calculated as deprotonated analytes ($M - H$)⁻. This created theoretical mass lists with 75,059 entries in negative-ion mode and 54,847 entries in positive-ion mode. These formula lists were chosen after close inspection of the data—Na adducts dominated over protonated species in positive-ion mode, and due to the narrow peak split between NaH and C₂ (2.4 mDa), could not be fully resolved over the full mass range on the lower resolving power instruments (i.e., Orbitraps), as this requires $m/\Delta m > 3 \times 10^5$ at m/z 400 (Δm is mass split between the two peaks to be resolved). This often led to a single centroid peak in the individual sample peak lists, rather than two, and for this reason, only the more abundant Na ions were considered for molecular formula assignment. The almost doubled number of peaks in each spectrum due to Na adduct formation in positive-ion mode also obscured the S-containing ions at higher masses and lower resolution. Thus, S assignments could not be made by the lower resolving power instruments in positive-ion mode. The objective of this study was comparison of ion abundances for confident assignments and not full sample coverage; therefore, this conservative approach was appropriate. Clearly, more complete sample coverage can be achieved with higher resolving power instruments and more complete formula assignment routines.

Formula assignment continues to be debated in the ESI-HRMS community, and here the conservative approach was taken to more severely constrain assigned formulas as compared to many other studies, bearing in mind that the samples chosen are not among the most complex examples,

such as petroleum or mixtures containing fresh metabolites (Gonsior et al. 2019; Palacio Lozano et al. 2019). Rates of formula assignment heavily depend on the resolving power of the chosen instrument. The limitations of our study reflect the fact that multiple instruments with very different resolving powers were included, while only one formula assignment routine was applied throughout. Furthermore, we focused our data analysis to the subset of signals that were reliably detected by most instruments to provide benchmarking data for other users. This approach obviously improves the ability to compare the resulting data, but concomitantly means that a major part of information obtained by high-resolution instruments were left unassigned. Since our goal was to evaluate common peaks that allow comparison of the instruments and data in both this study and future applications, this conservative approach was applied to minimize false positive assignments and provide the most robust metrics and data possible.

Internal calibration

The spectra were first internally calibrated using molecular formulas with DBE-O = -1, 0, or 1. These formulas have previously been shown to be among the most abundant in DOM samples (Herzprung et al. 2014) and are therefore extremely likely assignments for some high abundance ions. These highly probable formula assignments were thus used to perform a fine internal mass calibration (5th order polynomial) over the full mass range, using a similar approach to Sleighter et al. (2008) (Fig. 1). Three instruments required an initial

mass adjustment before the internal calibration due to pre-acquisition calibrations that were beyond acceptable tolerance (common for initial instrument use) – these were Lab M positive mode (42.67 ppm), Lab G positive mode (101.5 ppm), and Lab G negative mode (−98 ppm).

Formula assignment

After calibration with the reduced formula list, detected peaks in the full noise-filtered and calibrated peak lists were assigned to the closest molecular formulas from the full theoretical mass list within a mass tolerance of ± 1 ppm (thereby allowing evaluation of the best and worst performance characteristics among the instruments). The large majority of assignments had a mass error < 0.5 ppm across the full dataset (85% negative mode, 87% positive mode), with better mass accuracy being exhibited by higher resolution instruments. In all cases, the next nearest formula from the theoretical peak list was also noted in a separate matrix of potential interferences. These generally had mass errors > 2 ppm. Each assigned peak was thereby attributed to only a single formula. The ion abundance, mass error, and closest alternative formula's mass distance were recorded for subsequent analyses and error monitoring. The formula assignments at a single nominal mass were compared and confirmed with previously published values for SRFA and PLFA in negative-ion mode at m/z 311 (D'Andrilli et al. 2013) and positive-ion mode at m/z 411 (Podgorski et al. 2012). The elemental combinations used in this study were sufficiently diverse to assign formulas to the majority of detected ions in the datasets (usually $> 70\%$ of total signal abundance; Fig. S1) but may require modification when analyzing different samples or when using different ionization methods. Some low abundance peaks were not assigned in most datasets, either due to the elemental and isotopic constraints or due to the requirement for single charge. This conservative approach reduced false positive assignments.

Contaminant and rare peak removal

Contaminant ions for each instrument were defined as those detected in the blank sample at $> 20\%$ of the maximum abundance (base peak) ion of the blank mass spectrum, and such contaminant ion abundances were adjusted to zero in all datasets from the respective instrument. This high threshold allows low abundance carryover to exist in the blank and prevents erroneous “blanket subtraction” of low abundance ions that may represent real DOM signals in the samples. Finally, ion formulas from all samples and instruments were aligned separately in negative- and positive-ion mode, and this list was cropped to only include formulas that were found in > 5 analyses throughout the entire data set and to not include ^{13}C isotopologues. The value of 5 was chosen so that each ion had to be found in at least two HRMS instruments. The resulting matrices contained 9291 rows (distinct monoisotopic formulas) in negative-ion mode and 7945 rows in positive-ion mode.

Labs G, H, and N did not provide positive-ion mode data, so positive-ion mode data was generated from three fewer instruments.

Classification of commonly assigned ions, metrics, composition dissimilarity, and compound classes

Assigned ions that were common for a given sample across all instruments in each ionization mode or in all-but-one instrument were labeled as “common,” and these ions were used for the metric calculations (separated for positive- and negative-ion modes). Ions detected in fewer than $n - 1$ instruments were categorized as “50%— $n - 1$ ” (at least half of instruments, but not common), “3–50%” (at least three instruments but fewer than half), and “1–2” (fewer than three instruments). Ions that were confidently detected and assigned by at least three instruments (i.e., the sum of the first three categories) were categorized as “detected ≥ 3 ” ions.

Each metric value was calculated using the common ions for each sample for all instruments. The common ion signals were normalized so that abundances summed to 10,000, and the weighted average (wa) metrics (O/C, H/C, m/z , and AI_{mod}) were calculated from the common ion formulas $k = 1 : n$ as in Eq. 1, where X is the metric, n is the number of common ions, and I is the relative ion abundance.

$$X_{\text{wa}} = \frac{\sum_{k=1}^{k=n} X_k \cdot I_k}{\sum I} \quad (1)$$

To assess compositional dissimilarity between paired samples, the percent Bray Curtis dissimilarity (%BC dissimilarity) was calculated for the common formulas $k = 1 : n$ between samples p and q as in Eq. 2. Because the common ion list was different for each reference material, the list of ions included for consideration included all ions that were “common” in at least one of the reference samples.

$$\% \text{BC dissimilarity} = 100 \frac{\sum_{k=1}^{k=n} |I_{p,k} - I_{q,k}|}{\sum_{k=1}^{k=n} (I_{p,k} + I_{q,k})} \quad (2)$$

%BC dissimilarity is a useful tool for HRMS datasets because it allows zeros and takes into account abundance information. %BC dissimilarity approaches 100% when samples have no assigned formulas in common, thus signifying completely different DOM molecular composition. The full dissimilarity matrix was used for principal coordinate analyses (PCoA) to obtain Eigenvector scores using classical multidimensional scaling with the *cmdscale* function of MATLAB. Cluster center positions were calculated for each sample in the multidimensional principal coordinate space, and distances to cluster centers were calculated

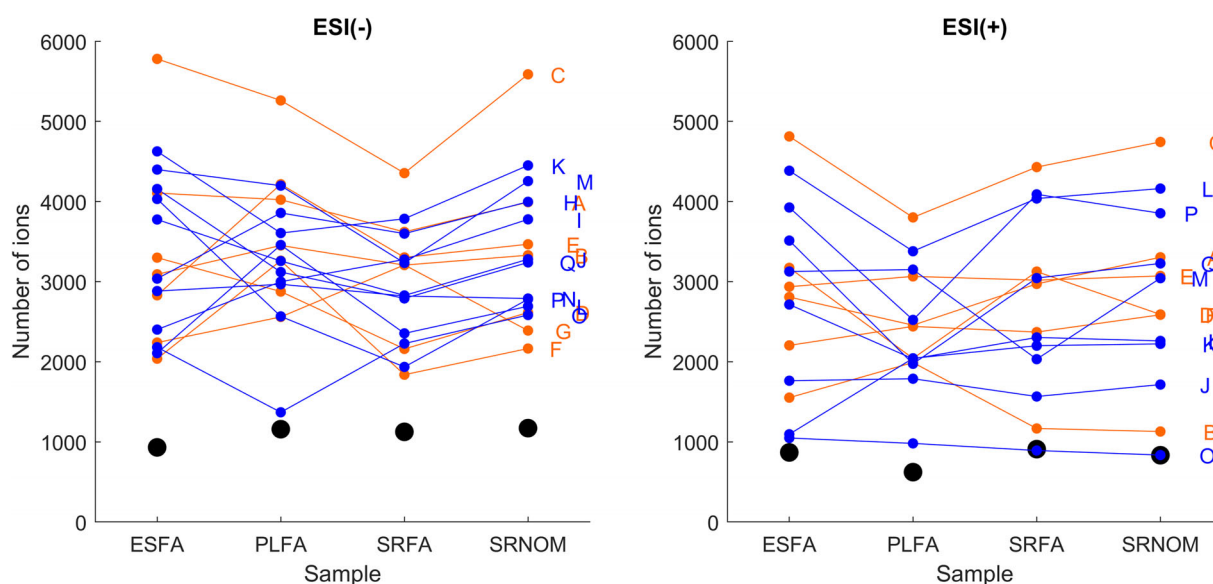


Fig 2 Number of detected ≥ 3 ions assigned in negative-ion and positive-ion Electrospray Ionization (ESI) modes for the four samples by each instrument. Instruments are identified by their code letters. Orbitraps are shown in orange and FT-ICRs are shown in blue. The number of ‘common’ ions (≥ 1) for each sample is shown as larger black circles. Samples included Elliot Soil, Pony Lake, and Suwannee River Fulvic Acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM).

for every dataset with the *betadisper* function in R (vegan package). This was done to calculate dispersion of the instrumental data from the average value for each sample.

Molecular formulas were assigned to compound classes based on atomic ratios. Relying upon the characteristic clustering of DOM data on van Krevelen diagrams, diverse chemical class divisions have been suggested (Sleighter and Hatcher 2008; Kellerman et al. 2014; Rivas-Ubach et al. 2018). For simplicity, in this study, we used five broad chemical compound classes, based on O/C, H/C, and AI_{mod} (Koch and Dittmar 2016). The classes were “aliphatic” ($H/C \geq 1.5$), “low O unsaturated” ($H/C < 1.5$, $AI_{mod} < 0.5$, $O/C < 0.5$), “high O unsaturated” ($H/C < 1.5$, $AI_{mod} < 0.5$, $O/C \geq 0.5$), “aromatics” ($0.5 < AI_{mod} < 0.67$), and “condensed aromatics” ($AI_{mod} \geq 0.67$). These classes represent broad compositional groups from HRMS data projected on van Krevelen diagrams and they are presented here to highlight key chemical differences between the samples and to demonstrate the inter-instrument variability in interpreting these results. At the same time, the broad grouping prevents us from over-interpreting molecular formula data, which allows only limited structural insight.

Assessment

Common, detected ≥ 3 , and total assigned ions

In total, data for the samples and blanks was compared from 17 and 14 different instruments in negative- and positive-ion modes, respectively. The number of detected ≥ 3 ions detected per sample varied widely across instruments from ~ 1500 to ~ 6000 in negative-ion mode and from ~ 1000 to ~ 5000 in positive-ion mode (Fig. 2). The total number of

assigned ions did not vary significantly among samples but did vary across instruments (ANOVA: F_{sample} 1.86, $F_{instrument}$ 11.2; Table 2). This variability in peak assignment number was not solely due to the differing resolving powers of the instruments—in fact, resolving power had only a small influence on the number of assigned peaks (Fig. S2).

Instruments with fewer assigned formulas generally had fewer total ions (assigned and unassigned) in each mass spectrum, indicating poorer detection limits rather than lower mass accuracy and assignment ratio (Figs. S1, S3, S4). Unassigned ions, i.e., resolved peaks that did not fit within the elemental constraints applied, were often more prominent at lower m/z where the resolution was better. Orbitraps had consistently greater numbers of unassigned ions, which may be a result of variables such as instrument sensitivity, noise reduction, and variability in mass accuracy. Regardless of the exact cause(s), this finding highlights the need for careful evaluation of formula assignment routines.

The variance in the number of assigned formulas was therefore partially attributable to instrument sensitivity as well as resolving power, ESI settings, the dynamic range and mass range of the instrument, and the calculated level of noise. The number of common ions varied between 622 and 1171 per sample among the four samples in positive- and negative-ion modes (Table 3). These ions were found in the mass range (m/z 223–661 negative, 221–713 positive) where all instruments’ analytical windows overlapped (Fig. 3) and are therefore not a complete representation of the full chemical diversity of ionizable organic species contained within the samples. The common peaks made up the majority of abundance (Fig. 4, S5), specifically accounting for the following

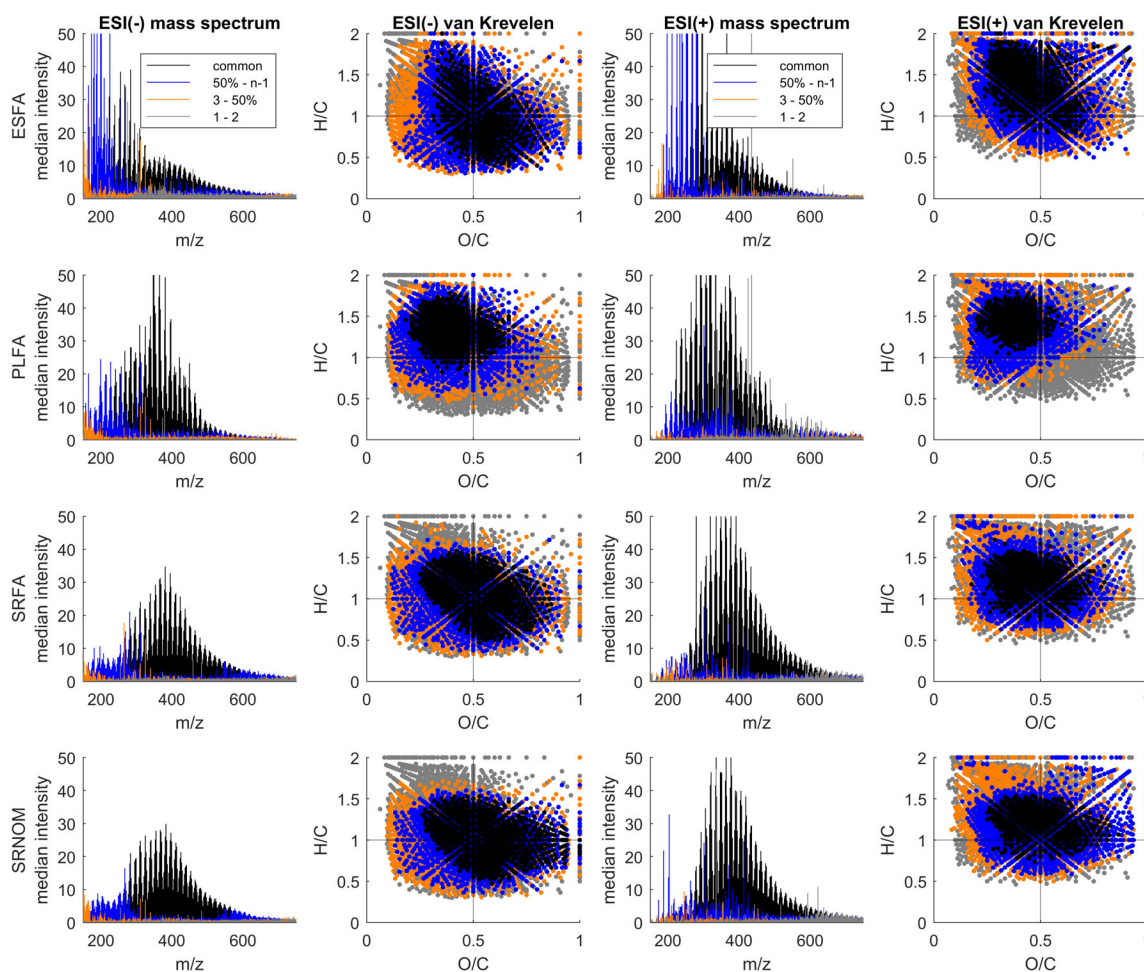


Fig 3 Ions detected in the four samples, shown as mass spectra and dissolved organic matter compositional ratios in van Krevelen space (H/C vs. O/C for each formula assigned) in negative-ion (left) and positive-ion (right) modes. The median relative abundance (summed to 10,000 for each sample) is shown as line height (cropped at 50) in the mass spectra. Detection rate categories from Table 2 are shown as different colors, with the more common categories shown in the background for mass spectra and in the foreground for van Krevelen diagrams. The overlap in the assigned molecular ions occurred in the central part of the mass spectral distribution (between m/z 300–550). Samples included Elliot Soil, Pony Lake, and Suwannee River Fulvic Acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM).

Table 2 ANOVA F values of variance explained by sample and instrument differences. The critical F value at 95% confidence is shown at the bottom of the table, and significant F values are shown in italics. The ratio of $F_{\text{sample}} : F_{\text{instrument}}$ is also displayed, to show the relative importance of the two factors.

Dataset	Parameter	Negative-ion mode			Positive-ion mode		
		F_{sample}	$F_{\text{instrument}}$	Ratio	F_{sample}	$F_{\text{instrument}}$	Ratio
Common ions	O/C _{wa}	541	11.1	48.6	435	9.53	45.6
	H/C _{wa}	1034	3.91	264.4	933	7.61	122.6
	m/z_{wa}	109	21.3	5.1	228	95.2	2.4
	Al _{mod,wa}	546	2.49	219.2	716	7.51	95.4
	High O unsaturated	136	6.09	22.3	249	7.69	32.4
	Low O unsaturated	379	11.1	34.2	232	6.73	34.5
	Aromatics	185	4.00	46.2	16.7	8.85	1.9
	Condensed aromatics	41.4	2.78	14.9	4.67	3.65	1.3
	Aliphatic	365	1.89	193.6	289	3.42	84.6
	All ions	# Assigned peaks	1.86	11.2	0.2	2.44	24.3
Critical F , 95% confidence		2.80	1.86		2.84	1.98	

assigned abundance percentages across the laboratories in negative-ion mode: ESFA ($46 \pm 11\%$), PLFA ($67 \pm 8\%$), SRFA ($72 \pm 9\%$), and SRNOM ($71 \pm 10\%$), each given as mean and standard deviations across the instruments. In positive-ion mode, the common ions represented assigned abundance percentages as follows: ESFA ($56 \pm 14\%$), PLFA ($66 \pm 12\%$), SRFA ($77 \pm 13\%$), and SRNOM ($72 \pm 13\%$).

The ions that were detected in at least three instruments (detected ≥ 3 ; Table 3) are probably more accurate measures of the “true” diversity of species that can be ionized and resolved by direct infusion ESI employing “broadband” (full mass range) acquisitions on instruments with resolving powers on the order of 10^5 – 10^6 (at m/z 401). Notably, broadband HRMS does not represent the full chemical diversity of the reference samples, just the portion that is ionizable by ESI. Furthermore, isomeric diversity will not be revealed by this approach (Hertkorn et al. 2008; Zark et al. 2017; Hawkes et al. 2018a). Formula lists (common and detected ≥ 3) for each sample in each ionization mode are available in Data S1 and online at <https://go.warwick.ac.uk/InterLabStudy>.

HRMS metrics for DOM comparison

Calculated weight-averaged values for each metric using the common ions are shown in Table 4. Across the HRMS instruments, for each sample, H/C_{wa} and $AI_{mod,wa}$ values had lower variability in negative- and positive-ion modes than O/C_{wa} and m/z_{wa} values (Fig. 4; see $F_{instrument}$ values from ANOVA analysis in Table 2). In this diverse IHSS sample set, the DOM sources had a larger effect on composition variability for all metric values than the different instruments ($F_{sample} > F_{instrument}$, Table 2). The only evaluated variable for which instruments led to higher variability than the DOM source was the total number of assigned peaks (Table 2). The Orbitrap instruments produced data with lower overall m/z_{wa} values than the FT-ICR instruments in both ESI modes (Fig. 4). In negative-ion mode, the two instrument types were significantly different (Student's t -test, 95% confidence level) for m/z_{wa} for all samples, significantly different for O/C_{wa} for ESFA and PLFA, and significantly different for H/C_{wa} and $AI_{mod,wa}$ only for ESFA. In positive-ion mode, the two instrument types were significantly different for m/z_{wa} for all

samples and significantly different for O/C_{wa} only for PLFA. Generally, the two instrument types (Orbitrap and FT-ICR) significantly overlapped for the metrics chosen (except m/z). However, in some cases it may be worth comparing newly obtained metric values in the context of instrument type.

The deviation of each instrument from the median instrument result with regard to O/C_{wa} and m/z_{wa} metric values is shown for negative-ion mode data in Figs. 5 and S6. The trends in ranking order of instruments are similar for each sample in each ionization mode. Clearly, some instruments have consistent metric value offsets across the samples (e.g., instrument B for O/C_{wa}), while others had highly outlying values, such as instrument I for O/C_{wa} in ESFA. As stated previously, the Orbitrap instruments gave lower metric values for m/z and higher values for O/C , which indicated consistent differences in sensitivity among instrument types.

Reference sample and instrument dissimilarity

Bray Curtis dissimilarities (%BC dissimilarity) were computed among all pairwise combinations of samples for each instrument (Fig. 6). Assessment of pairwise dissimilarities allows evaluation of how DOM composition differences between each sample pair were perceived by the various instruments. Dissimilarities between PLFA, SRFA, and SRNOM were consistent within $\sim 20\%$, while pairwise dissimilarities to ESFA were more variable (40%), but had similar ranges ($\sim 20\%$) when considering only one instrument type (Orbitrap or FT-ICR). These results were consistent between both ionization modes, and the larger dissimilarity with ESFA is likely related to the significant difference in obtained average m/z and O/C between these two instrument types (Fig. 5). For reference, the dissimilarity for five replicate analyses of SRFA on instruments A and L were 5.7% and 5.9%, respectively, using the common ion abundances.

The dispersion homogeneity analysis determined that average distances to cluster centers in principal coordinate space was significantly higher for ESFA in negative-ion mode (Tukey's HSD, $p < 0.05$), while dispersion was not significantly different for the other reference samples and all four samples in positive-ion mode. This method allowed the classification of outliers, as indicated in Fig. 7 (i.e., instrument I for ESFA in negative mode, instrument P for three samples in positive mode). Note that

Table 3 Number of peaks detected as $\geq n - 1$ (common), $> 50\%$, $< n - 1$, ≥ 3 – 50% , and < 3 instruments for the Elliot Soil, Pony Lake, and Suwannee River fulvic acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM) samples after the molecular formula assignments and blank subtraction.

Category	No. peaks detected by ESI (–)				No. peaks detected by ESI (+)			
	ESFA	PLFA	SRFA	SRNOM	ESFA	PLFA	SRFA	SRNOM
$\geq n-1$	931	1158	1125	1171	868	622	909	833
$\geq 50\%$, $< n-1$	2110	2186	1746	2024	1794	1886	1695	1854
≥ 3 , $< 50\%$	2337	944	1054	1544	1595	990	1270	1524
< 3	1973	1472	2108	1872	2268	1625	2432	2281
≥ 3	5378	4288	3925	4739	4257	3498	3874	4211

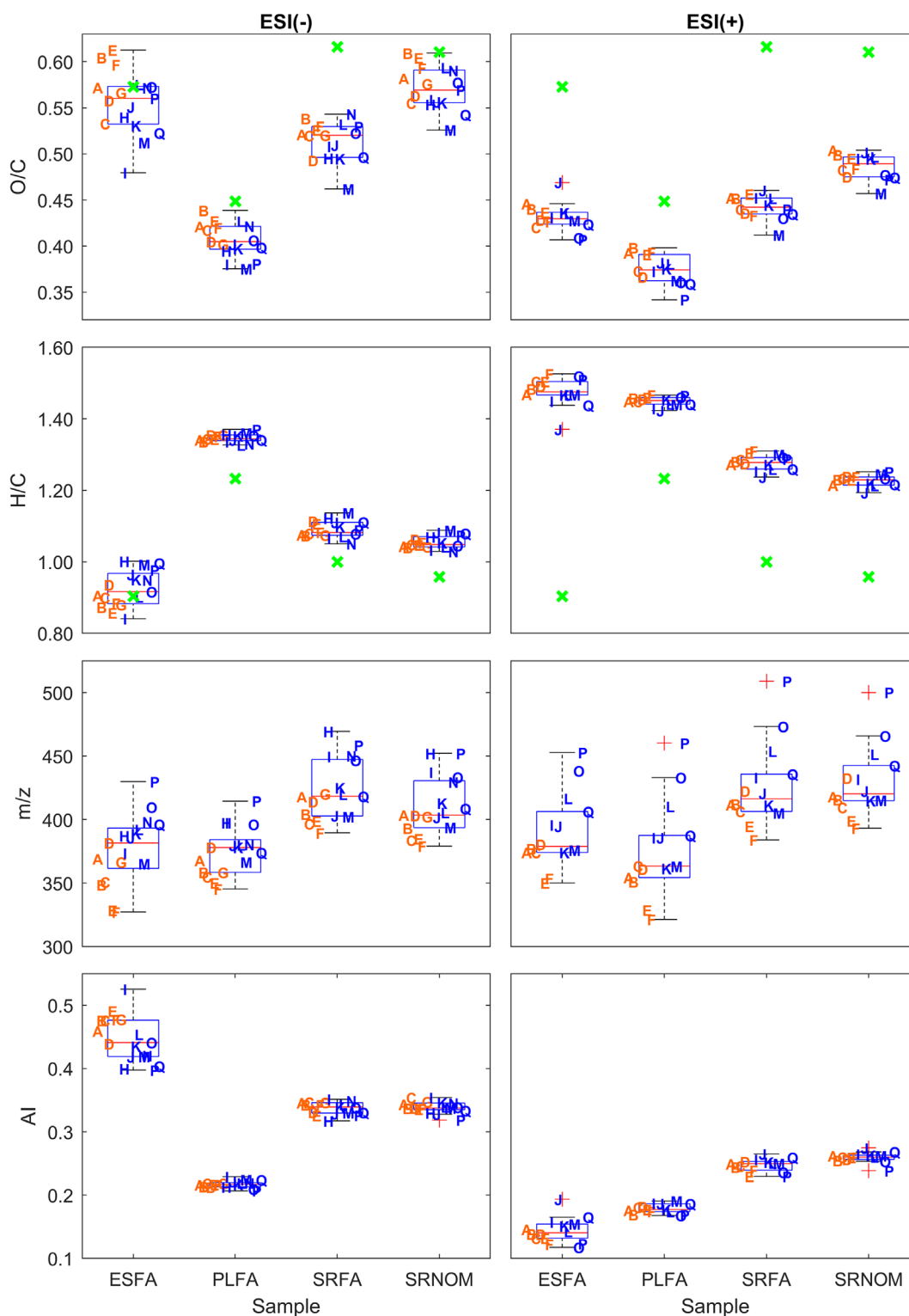


Fig 4 Box plots showing median (red line), interquartile range (blue lines), non-outlier range (dotted line and black bar) and outliers (red crosses). The data shown are the measured values of the four calculated metrics (weighted averages for oxygenation, hydrogen saturation, mass to charge ratio, and double bond equivalents minus oxygen: O/C_{war} , H/C_{war} , m/z_{war} and $AI_{mod,wa}$) for the four samples across 17 instruments (negative-ion mode, left panels) and 14 instruments (positive-ion mode, right panels) using common ions. FT-ICR instruments are displayed in blue and Orbitraps in orange. Where applicable, the IHSS bulk elemental ratios (IHSS website, 2019) are indicated with a green 'x'. Samples included Elliot Soil, Pony Lake, and Suwannee River Fulvic Acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM).

Table 4 Mean and standard deviation (SD) for four metrics (oxygenation, hydrogen saturation, mass to charge ratio, and modified aromaticity index: O/C_{wa} , H/C_{wa} , m/z_{wa} , and $AI_{mod,wa}$) for the four IHSS samples: Elliot Soil, Pony Lake, and Suwannee River fulvic acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM) in negative-ion and positive-ion mode. These values were calculated for the common ions without outliers, as in Fig. 4, which were defined as points greater than 1.5X the IQR above or below the 75th and 25th percentile values (Q3 and Q1), mathematically: $> Q3 + 1.5(Q3 - Q1)$ or $< Q1 - 1.5(Q3 - Q1)$.

	O/C		H/C		m/z		AI _{mod}	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Negative								
ESFA	0.56	0.033	0.93	0.049	376	26	0.45	0.035
PLFA	0.41	0.018	1.35	0.011	375	18	0.22	0.005
SRFA	0.52	0.016	1.09	0.022	423	23	0.34	0.01
SRNOM	0.57	0.022	1.05	0.016	405	17	0.34	0.009
Positive								
ESFA	0.43	0.011	1.48	0.039	386	24	0.14	0.019
PLFA	0.37	0.015	1.45	0.012	376	37	0.18	0.007
SRFA	0.44	0.012	1.28	0.02	420	23	0.25	0.01
SRNOM	0.49	0.013	1.23	0.015	424	20	0.26	0.005

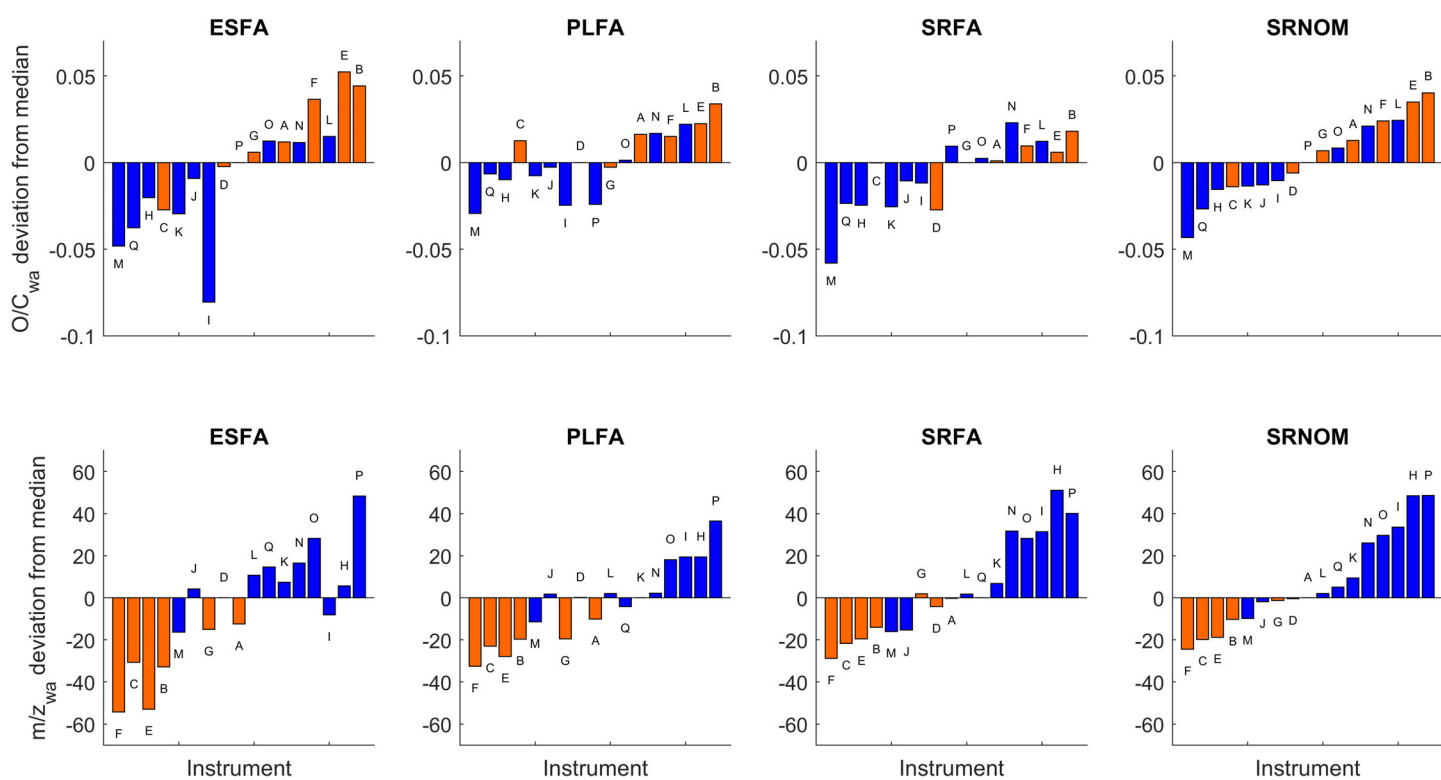


Fig 5 The deviation of the weighted averaged O/C and m/z from the median for the four different samples using all common ions in negative-ion mode. Orbitraps are shown in orange and FT-ICRs are shown in blue. Samples included Elliot Soil, Pony Lake, and Suwannee River Fulvic Acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM). The instruments are ordered based on their deviation from the median for SRNOM.

ESFA for instrument I was separated from the other instruments on the third PCoA dimension (not shown).

The %BC dissimilarity measured for each sample across instruments was also assessed in individual PCoAs. The resulting PCoA diagrams (first two coordinates) are shown in

Fig. 8. In each diagram, the correlation coefficients (Pearson) for the four calculated metrics and total number of assigned peaks with PCoA 1 and 2 are overlaid (significant ones, $p < 0.05$, are shown in bold purple font). The %BC dissimilarity among instruments generally averaged ~24%, except for the more

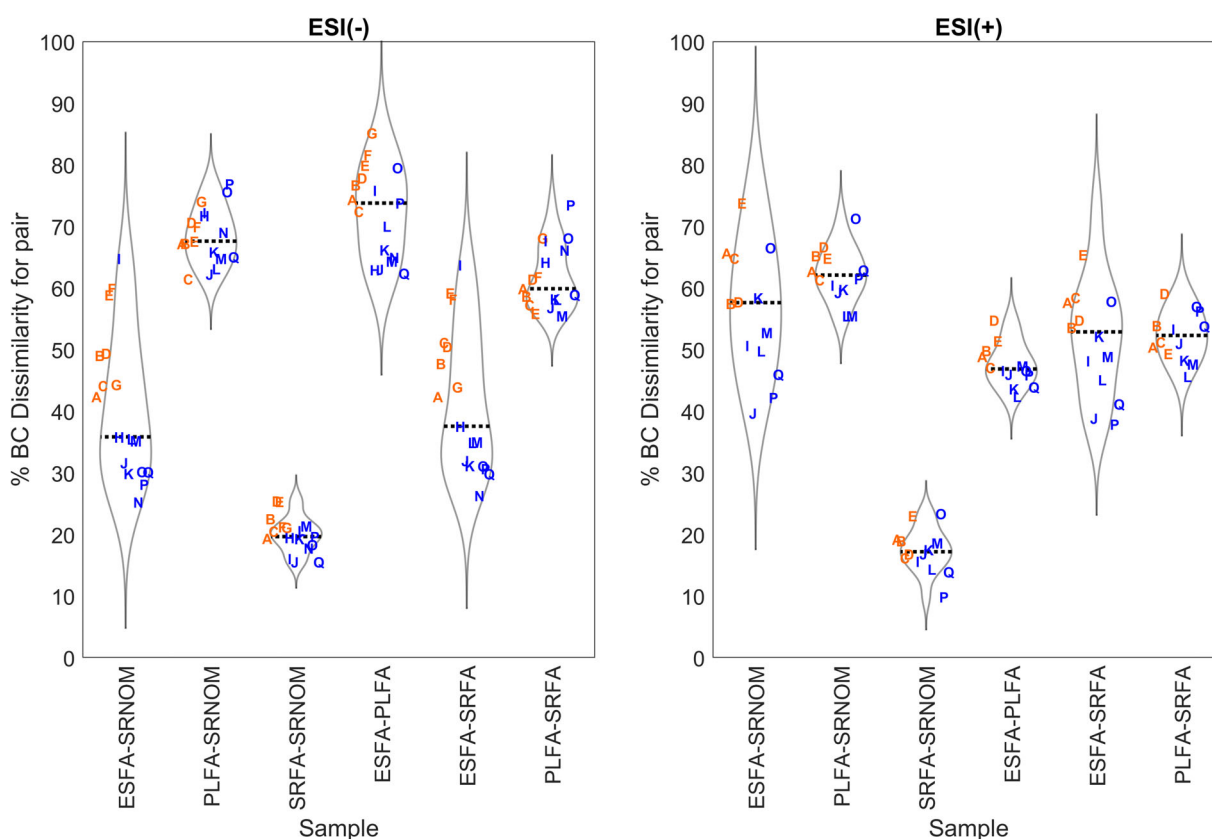


Fig 6 Percent Bray-Curtis Dissimilarities (%BC Dissimilarity) between Suwannee River Natural Organic Matter (SRNOM) and Elliot Soil, Pony Lake, and Suwannee River Fulvic Acids (ESFA, PLFA and SRFA) samples, shown as violin plots with an estimated kernel density function (Hoffmann, 2020) (negative-ion mode, left and positive-ion mode, right) with each instrument as one labeled data point using common ions. The median value is shown as a dotted black line. FT-ICR MS and Orbitrap instruments are represented by blue and orange, respectively.

disperse ESFA, which averaged 34–35%. Inter-instrument %BC Dissimilarity values were observed to be as low as 7%, close to intra-instrument variability (5–6%). The correlations with obtained metrics indicate that mass tuning was the principal cause of %BC dissimilarity, as m/z_{wa} correlated significantly with the first principal coordinate in every case. In negative-ion mode, O/C_{wa} usually correlated significantly with the second coordinate, and the total number of peak assignments and H/C_{wa} also frequently correlated significantly. In positive-ion mode, O/C_{wa} was typically not significantly correlated, while the number of assignments and saturation (H/C_{wa} or $AI_{mod,wa}$) were.

Compound class proportions

Using the full set of assigned formulas, the relative proportions of five compound classes were calculated (Figs. 9, 10, Table 5). In negative-ion mode (Fig. 9), the two most abundant classes were “high O unsaturated” and “low O unsaturated,” which signify formulas with $H/C < 1.5$ and $AI_{mod} < 0.5$. The ESFA sample had the highest proportion of “aromatic” or “condensed aromatic” formulas ($AI_{mod} \geq 0.5$ and ≥ 0.67 , respectively), and PLFA had the highest proportion of “aliphatic” formulas ($H/C \geq 1.5$). DOM compositional differences found within each reference sample

were far greater than differences across instruments (see Table 2 for ANOVA statistics). In line with differences in O/C_{wa} , highest instrumental variability was linked to the molecular groups of “high O unsaturated” and “low O unsaturated” formulas. In the case of ESFA, the proportion of “aromatics” and “condensed aromatics” was particularly variable among instruments (Simon et al. 2018). In positive-ion mode (Fig. 10) the trends were similar, but ESFA had a far greater proportion of ‘aliphatic’ formulas in this ionization mode ($59 \pm 11\%$ in positive-ion, $8 \pm 3\%$ negative-ion mode; mean \pm SD). “Aromatics” and “condensed aromatics,” conversely, were poorly ionized in positive-ion mode.

Discussion

Number of assigned peaks, mass ranges, and ion distributions across mass spectra

The number of assigned molecular formulas varied greatly among instruments (Fig. 2, Table 2) and did not solely correlate with instrument resolving power (Fig. S2). This is likely a consequence of variable quality of instrumental tuning related to the ESI source, the ICR or Orbitrap cell, and the ion transfer optics. Optimized instrument tuning will improve the detection limit and widen the m/z range, thereby generating a

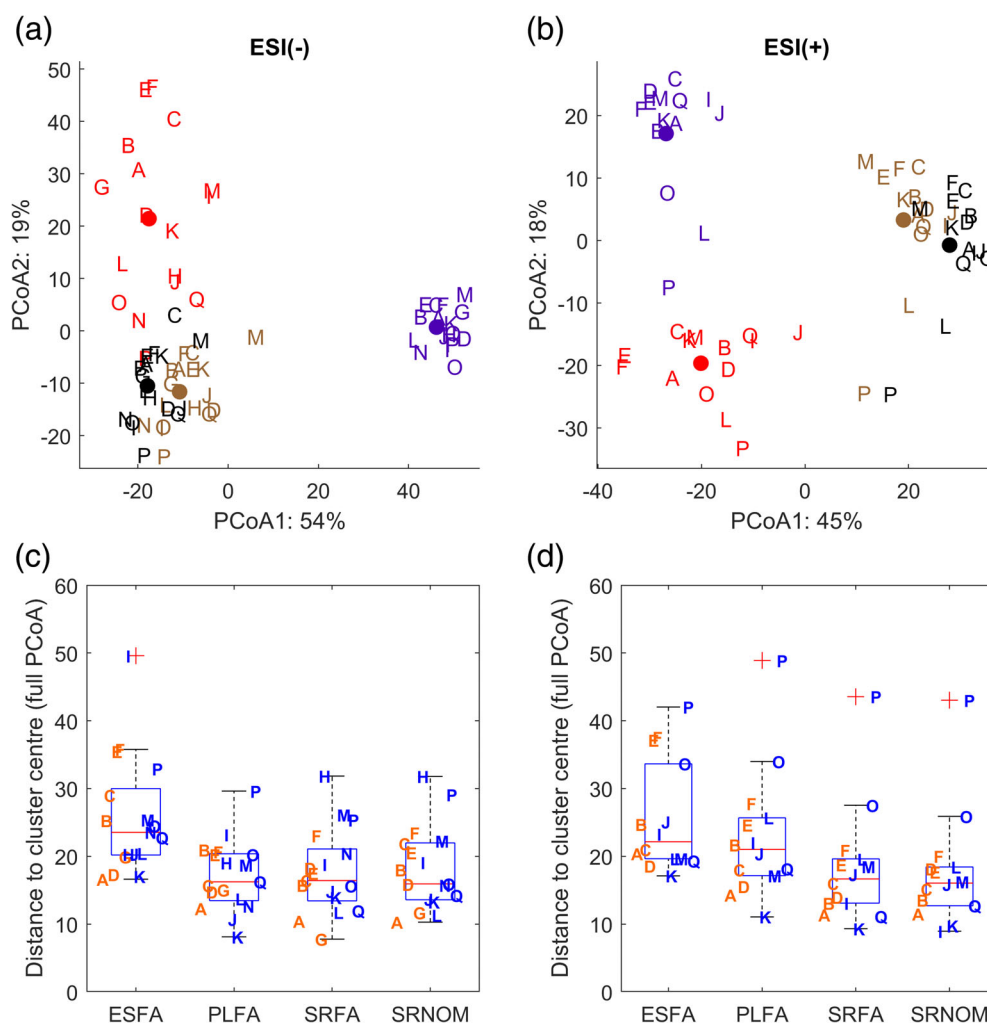


Fig 7 A-B: Principal coordinate Analysis (PCoA) diagrams (top) constructed from pairwise %BC Dissimilarities in negative-ion (left) and positive-ion (right) modes. Instruments are indicated by their designated letters and cluster centers are shown with filled circles. ESFA: Elliot Soil Fulvic Acid, red, PLFA: Pony Lake Fulvic Acid, purple, SRFA: Suwannee River Fulvic Acid, brown, SRNOM: Suwannee River Natural Organic Matter, black. **C-D:** Box plots (bottom) showing distance to cluster centers in **A-B** for the four samples in negative-ion (left) and positive-ion (right) mode. The box plots in panels **C-D** show median (red line), interquartile range (blue lines), non-outlier range (dotted line and black bar) and outliers (red crosses).

larger list of assigned formulas. Some of the instruments with higher resolving powers would be capable of assigning a wider range of compound classes than were allowed in the formula assignment routine, and so the total assigned peak number was constrained by the conservative approach taken. The minimum instrument resolving power required for any study depends on both the research question and the chemical diversity of the sample. Research that focusses on broad geochemical trends can often be achieved with lower resolving power, e.g., when determining presence and absence of specific chemical compositions or easily resolved ions, and when determining the abundance changes of the most prominent ions is sufficient (Hawkes et al. 2016; Simon et al. 2018). However, some research applications may benefit from maximizing resolving power and extending dynamic range to enable a more thorough assignment of compound classes

(Pohlabein and Dittmar 2015; Smith et al. 2018; Palacio Lozano et al. 2019).

The common ions for a given reference sample covered a broad mass range, and their assignments tended to occupy the central region of the van Krevelen diagram (Fig. 3). This indicates that these chemical species were highly abundant and/or easily ionizable in the samples, and that all instruments obtained a large overlap in the molecular composition of the samples. While the common ions represented the majority of abundance in three reference materials (PLFA, SRFA, and SRNOM), ESFA showed stronger variability due to the major differences obtained between Orbitrap and FT-ICR instruments. Even so, the common ions of a sample were detected consistently across all instruments, so can be considered as a reasonable and robust benchmark for evaluation of instrument performance in DOM-related applications.

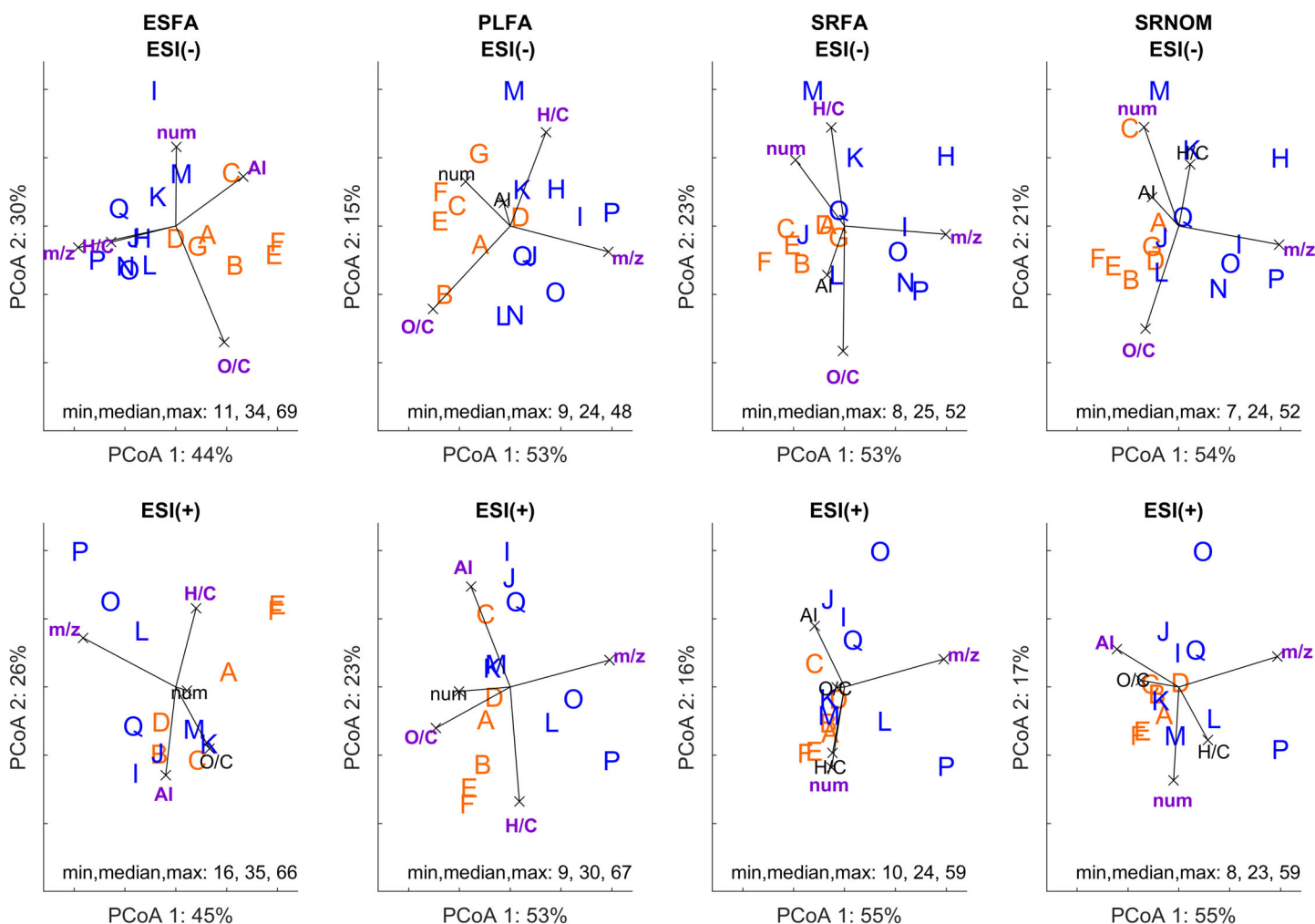


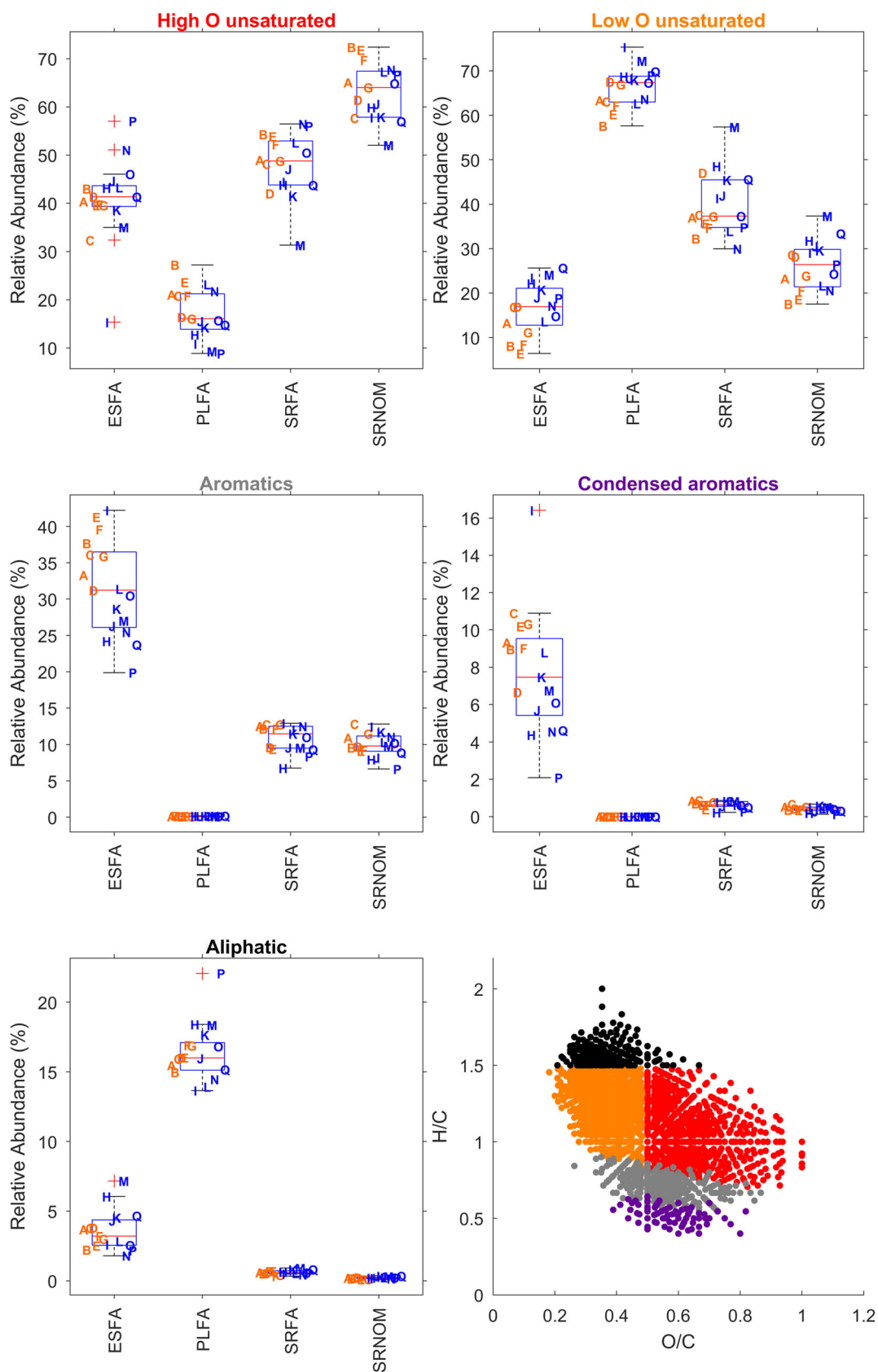
Fig 8 Principal coordinate diagrams based on %BC dissimilarity of normalized sample data among instruments for each sample in negative-ion mode (top) and positive-ion mode (bottom), using common ions. Each instrument is indicated by its designated letter and Orbitraps are orange and FT-ICRs are blue. The PCoA scores are normalized in both dimensions so that the highest value is 1 (scale not shown). Overlaid are Pearson's correlation scores showing covariance of the metric values and the PCoA score. A higher correlation is therefore manifested as a longer arrow, and the direction of the arrow indicates which axis the metric correlates with. Metric values with a significant correlation at $p < 0.05$ are indicated in bold and purple. The minimum, median, and maximum %BC dissimilarity among instruments is annotated at the bottom of each plot. Samples included Elliot Soil, Pony Lake, and Suwannee River fulvic acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM).

Variability in the detection of ions and metric values

Metric values were calculated using commonly detected ions (see definition in Section 1.2.3). The largest instrument variability in these metrics was found in average *m/z* and *O/C* ratios (Table 2). The detection of these common ions may be considered as an anchor point to evaluate instrument performance when assessing a new instrument or when troubleshooting with an established instrument. We suggest that a detection rate of $> 95\%$ of common ions is a sensible level for reasonable performance. There were instruments in our dataset that did not achieve this level (e.g., instrument I for ESFA in negative-ion mode), indicating that some tuning of instrument I may be required to analyze ESFA and other soil DOM samples at a level similar to that of other laboratories

(Fig. S7). Improvements in tuning or calibration may be made to instrument O in positive-ion mode and instrument D in negative-ion mode. Due to the overall higher variability across instruments for sample ESFA, we suggest not to choose ESFA as a routine standard material for instrument evaluation.

The weight-averaged metric values of commonly detected ions can be used to assess instrument tuning bias, particularly with regards to average mass and oxygenation. Saturation levels (*H/C* or Al_{mod}) were more consistent across instruments and can be used as effective guides for instrument comparability, or may be used as benchmarks to gauge tuning acceptability for new instruments to give results in a context comparable to those of the international community (Pan et al. 2020).



Bias due to ionization

The choice of ionization technique and polarity has a well-known and large effect on the ions produced and detected from any complex mixture (Hertkorn et al. 2008; Hockaday et al. 2009; Barrow et al. 2010; Hertzog et al. 2017; Kew et al. 2018b). ESI is a popular choice, and negative-ion mode is the most commonly selected mode in aquatic biogeochemistry because DOM is acidic by nature and bears many O-containing functional groups (Perdue and Ritchie 2013). However, these technical choices lead to specific perspectives on analytical mixtures (D'Andrilli et al. 2010; Gross 2010). The selectivity of the ESI mode can be shown in our data simply by comparing negative-ion and positive-ion mode data. The O/C_{wa} and H/C_{wa} metric values in negative-ion mode were similar to the published bulk elemental ratios of these mixtures (Fig. 4). The positive-ion mode O/C_{wa} and H/C_{wa} metric values were considerably different from the published bulk elemental ratios. It is unknown how much of each sample mixture is ionized in representative abundances, but these findings suggest that negative-ion mode ESI recovers a more representative portion of the mixture, albeit at lower oxygenation. Indications are that positive-ion mode is better suited to sensitively investigate aliphatic compounds, with higher average H/C ratios and lower aromaticity (Hertkorn et al. 2008; Hockaday et al. 2009).

Each HRMS instrument was tuned to analyze a specific ionizable portion of the material present in the samples (Fig. S8). It is unlikely that any instrument at any particular tune setting can capture the full representative distribution of ionizable compounds in these complex DOM mixtures in a single broadband analysis (Southam et al. 2007; Hawkes et al. 2019; Palacio Lozano et al. 2019). Indeed, the various instruments produced a relatively broad range in O/C_{wa} in both ionization modes (Fig. 4). Although some values were similar to the bulk elemental ratios, we acknowledge that fractions of the DOM material remained non-ionized and thus undetected. Obtaining the published elemental ratio using HRMS should thus not necessarily be the goal. Although outside the scope of this paper, aspects of differential ionization will need to be further assessed by the community (Hertkorn et al. 2008).

Inter-sample DOM composition and inter-instrument differences

The reference material samples that were expected to exert the largest differences in DOM molecular composition (e.g., PLFA vs. SRNOM) indeed showed large and consistent %

BC dissimilarity values across instruments. In contrast, samples from similar sources (e.g., SRFA vs. SRNOM), yielded consistently lower %BC Dissimilarity values for every instrument (Fig. 6). The ESFA to SRNOM dissimilarity had the widest range, depending strongly on the HRMS instrument type, and particularly m/z and O/C biases that greatly influenced ESFA data for Orbitrap instruments. With the exception of the number of assigned peaks, all metric values and proportions of compound classes had higher variability arising from inter-sample compositional differences than from instrumental differences (Table 2). However, our study covers a more diverse sample set compared to studies that focus on DOM temporal or spatial trends (Kellerman et al. 2014; Hertkorn et al. 2016; Hawkes et al. 2018b; Drake et al. 2019; Roth et al. 2019). For sample sets with higher DOM compositional similarity, instrument bias may become more important and trends in features such as compound class proportion may not be reproducibly determined among research groups. For this reason, inclusion of a known reference sample such as SRFA, PLFA, ESFA, or SRNOM in future HRMS DOM research is of high importance in order to give technical/instrumental context to the results. As mentioned above, ESFA is less likely to be a good reference sample for this purpose, since it exhibits greater variability of results among instruments.

As an example comparison with data collected outside of the present study, we evaluated previously published data (assigned mass lists for $S/N > 6$) for PLFA and SRFA (D'Andrilli et al. 2013) from a 9.4 T custom built FT-ICR MS instrument. Eighty-six percent of our "common PLFA peaks" (negative-ion mode) were found in the previous PLFA dataset, and the metric values were different from our published means by the following number of standard deviations: $O/C_{wa} + 1.17$, $H/C_{wa} - 2.25$, $m/z_{wa} + 0.29$, and $AI_{wa,mod} - 0.6$. Seventy-five percent of the common SRFA peaks were found in the previous SRFA dataset, and metric values were different from our published means by the following number of standard deviations: $O/C + 1.44$, $H/C + 0.32$, $m/z - 1.9$, $AI_{mod} - 1.0$. This suggests that D'Andrilli et al. (2013) had fewer peak assignments in common with our study (possibly due to their more conservative detection limit) and a bias in O/C that led to higher values than the averages in our study. However, most of these metric values are within 2 SDs of these average results (with the exception of H/C for PLFA), meaning that they are not statistically different to the sample of instruments studied here at the 95% confidence level. The results from that study should be considered in the context that the detected ions

Fig 9 Box plots showing summed abundances of molecular formulas for common ions in each of five compound classes for the four samples in negative-ion mode. The box plots show median (red line), interquartile range (blue lines), non-outlier range (dotted line and black bar) and outliers (red crosses). Orbitraps are shown in orange and FT-ICRs are shown in blue. Samples included Elliot Soil, Pony Lake, and Suwannee River fulvic acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM). The compound classes are represented by different colors in the example van Krevelen diagram (bottom right), which shows formulas that were common across instruments in at least one reference sample according to their oxygenation and hydrogen saturation (O/C and H/C).

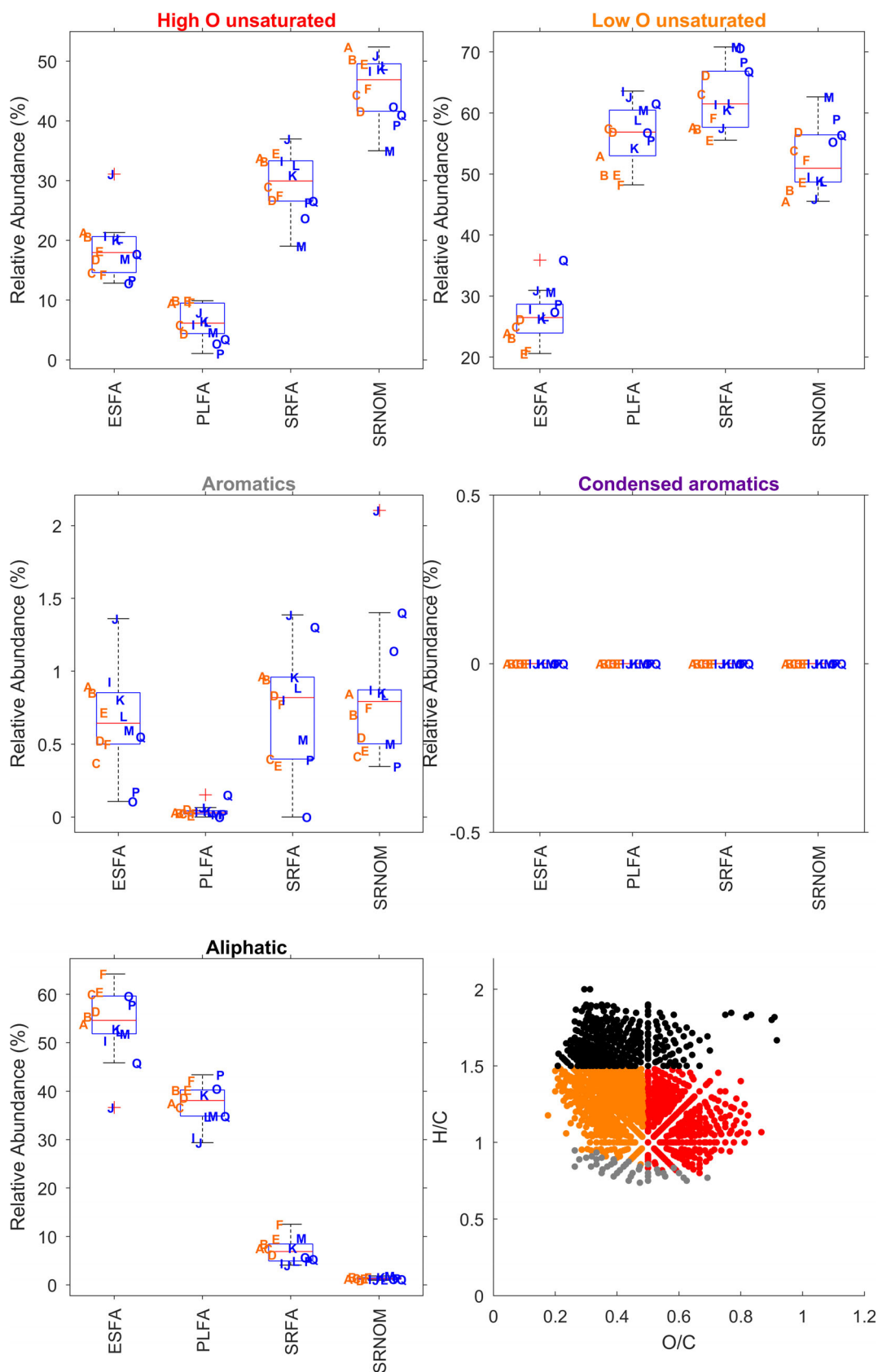


Table 5 Mean and SD for the five compound classes of the four IHSS samples: Elliot Soil, Pony Lake, and Suwannee River fulvic acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM) in negative-ion and positive-ion modes. These values were calculated for the common ions without outliers, as in Figs. 9, 10, which were defined as points greater than 1.5X the IQR above or below the 75th and 25th percentile values (Q3 and Q1), mathematically: $>Q3 + 1.5(Q3 - Q1)$ or $< Q1 - 1.5(Q3 - Q1)$.

	ESFA		PLFA		SRFA		SRNOM	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Negative								
High O unsaturated	41	2.9	17	5.3	48	6.4	63	5.8
Low O unsaturated	17	5.8	66	4.5	39	5.6	26	5.5
Aromatics	31	6.6	0.14	0.035	11	1.9	10	1.6
Condensed aromatics	7.2	2.5	0	0	0.64	0.2	0.41	0.14
Aliphatic	3.4	1.1	16	1.4	0.61	0.15	0.21	0.078
Positive								
High O unsaturated	18	3	6.3	2.8	30	4.9	46	5.1
Low O unsaturated	27	4.1	56	4.9	63	5.1	52	5.2
Aromatics	0.65	0.32	0.03	0.017	0.75	0.38	0.74	0.3
Condensed aromatics	0	0	0	0	0	0	0	0
Aliphatic	55	5	37	4.1	7	2.4	1.3	0.25

were more oxygenated and less saturated (in the case of PLFA) than the average of the international community.

DOM chemical characterization by each instrument

HRMS data of DOM samples is often used to characterize biomolecular compound classes as well as to quantify the relative abundances of each, as a way to reduce complex data into a smaller number of variables. Consistent with the %BC dissimilarity results (Fig. 6), the PLFA sample was the most compositionally distinct in negative-ion mode, being rich in aliphatic and low oxygen molecules and having a considerably lower abundance of unsaturated and aromatic molecules (Figs. 4, 9). As expected, the SRFA and SRNOM samples were the most similar to each other, although SRNOM had higher heteroatom content (i.e., a greater number of formulas containing N or S; Fig. S9). ESFA had the highest abundance of unequivocally aromatic and condensed aromatic compounds, as defined by AI_{mod} (Figs. 4, 9), but calculated aromaticity was highly variable among instruments according to the instrument sensitivity to low mass defects (i.e., high oxygen and/or low hydrogen) species. Generally, the Orbitraps detected more “aromatic” and fewer “low O unsaturated” ions in ESFA. Instrument I, which was clearly different for ESFA for nearly every calculated metric value, also exhibited a high abundance of “aromatic” and “condensed aromatic” ions. In positive-ion mode, ESFA was the most compositionally distinct, with the

majority of ions ($59 \pm 11\%$) being assigned to the “aliphatic” region (Fig. 10). Notably, “aromatics” and “condensed aromatics” were proportionally lower for every sample, including ESFA, in positive-ion mode.

The results for condensed aromatics ($AI_{mod} \geq 0.67$) for ESFA in negative-ion mode highlight a key problem for comparison of data between studies, with the large range in its percent compound class being attributable to the differences among instruments (Fig. 9, Table 2). This finding emphasizes the need for one or more standard mixture validations of instrument performance for future studies. Since it cannot be stated in an absolute sense that a sample has $x\%$ condensed aromatics from HRMS data alone, we suggest that only relative comparisons are used when applying a metric value comparison, e.g., one sample has more condensed aromatics than a known reference mixture.

It should also be emphasized that no ESI-HRMS analysis of a highly complex mixture is thoroughly comprehensive, especially in direct infusion mode, where analytes in the mixture compete for available charge during ionization. Very few aliphatic compounds were detected in the ESFA sample in negative-ion mode, but these compounds were highly ionized in positive-ion mode. In negative-ion mode, it is likely that ionization of aliphatics was suppressed by the large abundance of oxygen-rich compounds that can readily accommodate a negative charge.

Fig 10 Box plots showing summed abundances of molecular formulas for common ions in each of five compound classes for the four samples in positive-ion mode. The box plots show median (red line), interquartile range (blue lines), non-outlier range (dotted line and black bar) and outliers (red crosses). Orbitraps are shown in orange and FT-ICRs are shown in blue. Samples included Elliot Soil, Pony Lake, and Suwannee River fulvic acids (ESFA, PLFA, and SRFA) and Suwannee River Natural Organic Matter (SRNOM). The compound classes are represented by different colors in the example van Krevelen diagram (bottom right), which shows formulas that were common across instruments in at least one reference sample according to their oxygenation and hydrogen saturation (O/C and H/C).

Comments and recommendations

It is clear that a great effort would be required to obtain virtually identical mass spectra for complex DOM mixtures among laboratories. Instruments can be initially tuned to modify the “analytical window” that is covered (Kew et al. 2018a; Pan et al. 2020), but it is conceptually difficult to decide a priori what the desired spectrum should look like. Here we begin this effort by establishing the averages and ranges obtained at the *status quo*, including the presence of commonly detected ions and weight-averaged metric values of these ions. The use of reference samples to evaluate HRMS instrument bias will be an important feature of future studies in order to generate robust and comparable data, and will thereby help to minimize irreproducibility. The results presented here enhance our ability and awareness to reveal differential instrument responses in HRMS studies focusing on DOM. This will also enable more focus in discussions about HRMS analytical challenges. In a follow-up paper, HRMS parameter tuning effects will be investigated to gain better understanding of observed data variability, so that other laboratories may set up their instruments in a more uniform manner. It is our hope that this will help the DOM community to disentangle instrumental from interpretational variability in sample sets from identical (or comparable) environmental contexts.

We recommend that at least one of the four IHSS reference mixtures used in this study be included in future datasets, and that the % common ions detected and their metrics be reported, so that the instrument bias and extent of analytical window can be understood. For freshwater studies, SRFA is the preferred option to include (if only one is to be selected) because more studies have used this reference sample in the past. While marine studies would benefit from the provision of a marine reference material (Green et al. 2014), the current preference may be to select PLFA. However, in all cases SRFA or SRNOM would also be suitable to provide overall context to the results. Large deviations from the detection of common ions (< 95%) or metric values (> ± 2 SD offset from mean) should be investigated and explained. To facilitate this, we have set up an online data comparison tool – see <https://go.warwick.ac.uk/InterLabStudy>. We also recommend that when analyzing a large number of samples for comparison that 10–20% of the samples be analyzed by replicate injections, to ensure that variability observed between samples is larger than that between replicates.

Care should be taken to describe results in a relative sense, because the values obtained (abundances and averaged calculated metric values) are not absolute. Even though more work is needed to understand ion abundance pattern reproducibility, its determinants and what fraction of DOM is ionized by ESI, this work initiates a unification among ESI-HRMS users that can continue to advance DOM research at the molecular level.

As open communication among HRMS laboratories for optimal data collection continues, more reliable comparisons across the community will be made. Including a commercially

available (and continuously collected) reference material in a data set, either made in-house or ideally with internationally certified mixtures, adds substantial confidence that the data has been collected and assigned in a reproducible manner. Moreover, it allows the new data to be compared with previous findings from the community in a less isolated way, enabling a more direct integration of new results in the broader context of existing data. It remains to be investigated whether and how a complex reference material could be used to standardize a dataset in order to increase comparability of results among instruments. Furthermore, the community should aim to work together such that raw and processed peak lists along with data processing methods are available in a central, accessible repository as part of routine data publishing and archiving. In this manner, future work in environmental sciences can reproduce and reliably build upon already published efforts.

References

- Barrow, M. P., M. Witt, J. V. Headley, and K. M. Peru. 2010. Athabasca oil sands process water: Characterization by atmospheric pressure photoionization and electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **82**(9): 3727–3735. doi.org/10.1021/ac100103y.
- Brown, T. L., and J. A. Rice. 2000. Effect of experimental parameters on the ESI FT-ICR mass spectrum of fulvic acid. *Anal. Chem.* **72**(2): 384–390. <https://doi.org/10.1021/ac9902087>.
- Brown, A., D. M. McKnight, Y. P. Chin, E. C. Roberts, and M. Uhle. 2004. Chemical characterization of dissolved organic material in Pony Lake, a saline coastal pond in Antarctica. *Mar. Chem.* **89**(1–4): 327–337. <https://doi.org/10.1016/j.marchem.2004.02.016>.
- Cao, D., J. Lv, F. Geng, Z. Rao, H. Niu, Y. Shi, Y. Cai, and Y. Kang. 2016. Ion accumulation time dependent molecular characterization of natural organic matter using electrospray ionization-Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **88**(24): 12210–12218. <https://doi.org/10.1021/acs.analchem.6b03198>
- D’Andrilli, J., T. Dittmar, B. P. Koch, J. M. Purcell, A. G. Marshall, and W. T. Cooper. 2010. Comprehensive characterization of marine dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry with electrospray and atmospheric pressure photoionization. *Rapid Commun. Mass Spectrom.* **24**: 643–650. <https://doi.org/10.1002/rcm.4421>.
- D’Andrilli, J., C. M. Foreman, A. G. Marshall, and D. M. McKnight. 2013. Characterization of IHSS pony lake fulvic acid dissolved organic matter by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry

- and fluorescence spectroscopy. *Org. Geochem.* **65**: 19–28. <https://doi.org/10.1016/j.orggeochem.2013.09.013>
- D'Andrilli, J., W. T. Cooper, C. M. Foreman, and A. G. Marshall. 2015. An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability. *Rapid Commun. Mass Spectrom.* **29**(4): 2385–2401. <https://doi.org/10.1002/rcm.7400>.
- Dittmar, T., and B. P. Koch. 2006. Thermogenic organic matter dissolved in the abyssal ocean. *Mar. Chem.* **102**(3–4): 208–217. <https://doi.org/10.1016/j.marchem.2006.04.003>
- Drake, T. W., K. Van Oost, M. Barthel, M. Bauters, A. M. Hoyt, D. C. Podgorski, J. Six, P. Boeckx, S. E. Trumbore, N. L. Cizungu, and R. G. M. Spencer. 2019. Mobilization of aged and biolabile soil carbon by tropical deforestation. *Nat. Geosci.* **12**(July) 541–546. <https://doi.org/10.1038/s41561-019-0384-9>.
- Dubinenkov, I., R. Flerus, P. Schmitt-Kopplin, G. Kattner, and B. P. Koch. 2015. Origin-specific molecular signatures of dissolved organic matter in the Lena Delta. *Biogeochemistry* **123**(1–2): 1–14. <https://doi.org/10.1007/s10533-014-0049-0>
- Fenn, J. B., M. Mann, C. K. A. I. Meng, S. F. Wong, and C. M. Whitehouse. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**(4926): 64–71. <https://doi.org/10.1126/science.2675315>.
- Fievre, A., T. Solouki, A. G. Marshall, and W. T. Cooper. 1997. High-resolution Fourier transform ion cyclotron resonance mass spectrometry of humic and fulvic acids by laser desorption/ionization and electrospray ionization. *Energy Fuel* **11**(3): 554–560. <https://doi.org/10.1021/ef970005q>.
- Flerus, R., O. J. Lechtenfeld, B. P. Koch, S. L. McCallister, P. Schmitt-Kopplin, R. Benner, K. Kaiser, and G. Kattner. 2012. A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* **9**(6): 1935–1955. <https://doi.org/10.5194/bg-9-1935-2012>
- Gaspar, A., M. Harir, N. Hertkorn, and P. Schmitt-Kopplin. 2010. Preparative free-flow electrophoretic offline ESI-Fourier transform ion cyclotron resonance/MS analysis of Suwannee River fulvic acid. *Electrophoresis* **31**(12): 2070–2079. <https://doi.org/10.1002/elps.200900726>
- Gonsior, M., B. M. Peake, W. T. Cooper, D. Podgorski, J. D'Andrilli, and W. J. Cooper. 2009. Photochemically induced changes in dissolved organic matter identified by ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry. *Environ. Sci. Technol.* **43**(3): 698–703. <https://doi.org/10.1021/es8022804>
- Gonsior, M., L. C. Powers, E. Williams, A. Place, F. Chen, A. Ruf, N. Hertkorn, and P. Schmitt-kopplin. 2019. The chemodiversity of algal dissolved organic matter from lysed *Microcystis aeruginosa* cells and its ability to form disinfection by-products during chlorination. *Water Res.* **155**: 300–309. <https://doi.org/10.1016/j.watres.2019.02.030>
- Green, N. W., E. M. Perdue, G. R. Aiken, K. D. Butler, H. Chen, T. Dittmar, J. Niggemann, and A. Stubbins. 2014. An intercomparison of three methods for the large-scale isolation of oceanic dissolved organic matter. *Mar. Chem.* **161**: 14–19. doi:10.1016/j.marchem.2014.01.012
- Green, N. W., D. McInnis, N. Hertkorn, P. A. Maurice, and E. M. Perdue. 2015. Suwannee River natural organic matter: Isolation of the 2R101N reference sample by reverse osmosis. *Environ. Eng. Sci.* **32**: 38–44. doi:10.1089/ees.2014.0284
- Gross, J. 2010. *Mass spectrometry: A textbook*, 2nd Edition. Springer Heidelberg.
- Hawkes, J. A., T. Dittmar, C. Patriarca, L. Tranvik, and J. Bergquist. 2016. Evaluation of the Orbitrap mass spectrometer for the molecular fingerprinting analysis of natural dissolved organic matter. *Anal. Chem.* **88**: 7698–7704. doi:10.1021/acs.analchem.6b01624
- Hawkes, J. A., C. Patriarca, P. J. R. Sjöberg, L. J. Tranvik, and J. Bergquist. 2018a. Extreme isomeric complexity of dissolved organic matter found across aquatic environments. *Limnol. Oceanogr. Lett.* **3**: 21–30. doi:10.1002/lol2.10064
- Hawkes, J. A., N. Radoman, J. Bergquist, M. B. Wallin, L. J. Tranvik, and S. Löfgren. 2018b. Regional diversity of complex dissolved organic matter across forested hemiboreal headwater streams. *Sci. Rep.* **8**: 16060. doi:10.1038/s41598-018-34272-3
- Hawkes, J. A., P. J. R. Sjöberg, J. Bergquist, and L. J. Tranvik. 2019. Complexity of dissolved organic matter in the molecular size dimension: Insights from coupled size exclusion chromatography electrospray ionisation mass spectrometry. *Faraday Discuss.* **218**: 52–71. doi:10.1039/c8fd00222c
- Henry, K. D., E. R. Williams, B. H. Wang, F. W. McLafferty, J. Shabanowitz, and D. F. Hunt. 1989. Fourier-transform mass spectrometry of large molecules by electrospray ionization. *Proc. Natl. Acad. Sci. USA* **86**: 9075–9078. doi:10.1073/pnas.86.23.9075
- Hertkorn, N., M. Frommberger, M. Witt, B. P. Koch, P. Schmitt-kopplin, and E. M. Perdue. 2008. Natural organic matter and the event horizon of mass spectrometry. *Anal. Chem.* **80**: 8908–8919. doi:10.1021/ac800464g
- Hertkorn, N., M. Harir, K. M. Cawley, P. Schmitt-Kopplin, and R. Jaffé. 2016. Molecular characterization of dissolved organic matter from subtropical wetlands: A comparative study through the analysis of optical properties, NMR and FTICR/MS. *Biogeosciences* **13**: 2257–2277. doi:10.5194/bg-13-2257-2016
- Hertzog, J., V. Carré, Y. Le Brech, C. L. Mackay, A. Dufour, O. Mašek, and F. Aubriet. 2017. Combination of electrospray ionization, atmospheric pressure photoionization and laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry for the investigation of complex mixtures – Application to the petroleomic analysis. *Anal. Chim. Acta* **969**: 26–34. doi:10.1016/j.aca.2017.03.022
- Herzprung, P., N. Hertkorn, W. Von Tumpling, M. Harir, K. Friese, and P. Schmitt-Kopplin. 2014. Understanding molecular formula assignment of Fourier transform ion cyclotron

- resonance mass spectrometry data of natural organic matter from a chemical point of view. *Anal. Bioanal. Chem.* **406**: 7977–7987. doi:[10.1007/s00216-014-8249-y](https://doi.org/10.1007/s00216-014-8249-y)
- Herzsprung, P. V., W. Tümpling, N. Hertkorn, M. Harir, K. Friese, and P. Schmitt-Kopplin. 2015. High-field FTICR-MS data evaluation of natural organic matter: Are CHONSS2 molecular class formulas assigned to ^{13}C isotopic m/z and in reality CHO components? *Anal. Chem.* **87**: 9563–9566. doi:[10.1021/acs.analchem.5b02549](https://doi.org/10.1021/acs.analchem.5b02549)
- Hockaday, W. C., J. M. Purcell, A. G. Marshall, J. A. Baldock, and P. G. Hatcher. 2009. Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: A qualitative assessment. *Limnol. Oceanogr. Methods* **7**: 81–95. doi:[10.4319/lom.2009.7.81](https://doi.org/10.4319/lom.2009.7.81)
- Hoffmann, H. 2020. Violin plot for MATLAB.
- IHSS Website. 2019. Elemental analysis of IHSS products. Wheat Ridge, Colorado: Huffman Laboratories, Available from: <http://humic-substances.org/elemental-compositions-and-stable-isotopic-ratios-of-ihss-samples/>
- Jaffé, R., Y. Yamashita, N. Maie, W. T. Cooper, T. Dittmar, W. K. Dodds, J. B. Jones, T. Myoshi, J. R. Ortiz-Zayas, D. C. Podgorski, and A. Watanabe. 2012. Dissolved organic matter in headwater streams: Compositional variability across climatic regions of North America. *Geochim. Cosmochim. Acta* **94**: 95–108. doi:[10.1016/j.gca.2012.06.031](https://doi.org/10.1016/j.gca.2012.06.031)
- Kellerman, A. M., T. Dittmar, D. N. Kothawala, and L. J. Tranvik. 2014. Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. *Nat. Commun.* **5**: 1–8. doi:[10.1038/ncomms4804](https://doi.org/10.1038/ncomms4804)
- Kellerman, A. M., F. Guillemette, D. C. Podgorski, G. R. Aiken, K. D. Butler, and R. G. M. Spencer. 2018. Unifying concepts linking dissolved organic matter composition to persistence in aquatic ecosystems. *Environ. Sci. Technol.* **52**: 2538–2548. doi:[10.1021/acs.est.7b05513](https://doi.org/10.1021/acs.est.7b05513)
- Kew, W., J. W. T. Blackburn, D. J. Clarke, and D. Uhrin. 2017. Interactive van Krevelen diagrams - Advanced visualisation of mass spectrometry data of complex mixtures. *Rapid Commun. Mass Spectrom.* **31**: 658–662. doi:[10.1002/rcm.7823](https://doi.org/10.1002/rcm.7823)
- Kew, W., J. W. T. Blackburn, and D. Uhrin. 2018a. Response to comment on “laser desorption/ionization coupled to FTICR mass spectrometry for studies of natural organic matter”. *Anal. Chem.* **90**: 5968–5971. doi:[10.1021/acs.analchem.8b00332](https://doi.org/10.1021/acs.analchem.8b00332)
- Kew, W., C. L. Mackay, I. Goodall, D. J. Clarke, and D. Uhrin. 2018b. Complementary ionization techniques for the analysis of scotch whisky by high resolution mass spectrometry. *Anal. Chem.* **90**: 11265–11272. doi:[10.1021/acs.analchem.8b01446](https://doi.org/10.1021/acs.analchem.8b01446)
- Koch, B., M. Witt, R. Engbrodt, T. Dittmar, and G. Kattner. 2005. Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim. Cosmochim. Acta* **69**: 3299–3308. doi:[10.1016/j.gca.2005.02.027](https://doi.org/10.1016/j.gca.2005.02.027)
- Koch, B. P., and T. Dittmar. 2016. Erratum: From mass to structure: An aromaticity index for high-resolution mass data of natural organic matter (*Rapid Communications in Mass Spectrometry* (2006) 20 (926-932) DOI: 10.1002/rcm.2386). *Rapid Commun. Mass Spectrom.* **30**: 250. doi:[10.1002/rcm.7433](https://doi.org/10.1002/rcm.7433)
- Kujawinski, E. B., M. A. Freitas, X. Zang, P. G. Hatcher, K. B. Green-Church, and R. B. Jones. 2002. The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Org. Geochem.* **33**: 171–180. doi:[10.1016/S0146-6380\(01\)00149-8](https://doi.org/10.1016/S0146-6380(01)00149-8)
- Lechtenfeld, O. J., G. Kattner, R. Flerus, S. L. McCallister, P. Schmitt-Kopplin, and B. P. Koch. 2014. Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim. Cosmochim. Acta* **126**: 321–337. doi:[10.1016/j.gca.2013.11.009](https://doi.org/10.1016/j.gca.2013.11.009)
- Li, Y., M. Harir, J. Uhl, B. Kanawati, M. Lucio, K. S. Smirnov, B. P. Koch, P. Schmitt-Kopplin, and N. Hertkorn. 2017. How representative are dissolved organic matter (DOM) extracts? A comprehensive study of sorbent selectivity for DOM isolation. *Water Res.* **116**: 316–323. doi:[10.1016/j.watres.2017.03.038](https://doi.org/10.1016/j.watres.2017.03.038)
- Liu, Y., and E. B. Kujawinski. 2015. Chemical composition and potential environmental impacts of water-soluble polar crude oil components inferred from ESI FT-ICR MS. *PLoS One* **10**: 1–18. doi:[10.1371/journal.pone.0136376](https://doi.org/10.1371/journal.pone.0136376)
- Lu, K., W. S. Gardner, and Z. Liu. 2018. Molecular structure characterization of riverine and coastal dissolved organic matter with ion mobility quadrupole time-of-flight LCMS (IM Q-TOF LCMS). *Environ. Sci. Technol.* **52**: 7182–7191. doi:[10.1021/acs.est.8b00999](https://doi.org/10.1021/acs.est.8b00999)
- Luek, J. L., M. Harir, P. Schmitt-Kopplin, P. J. Mouser, and M. Gonsior. 2018. Temporal dynamics of halogenated organic compounds in Marcellus Shale flowback. *Water Res.* **136**: 200–206. doi:[10.1016/j.watres.2018.02.055](https://doi.org/10.1016/j.watres.2018.02.055)
- Mangal, V., N. L. Stock, and C. Guéguen. 2016. Molecular characterization of phytoplankton dissolved organic matter (DOM) and sulfur components using high resolution Orbitrap mass spectrometry. *Anal. Bioanal. Chem.* **408**: 1891–1900. doi:[10.1007/s00216-015-9295-9](https://doi.org/10.1007/s00216-015-9295-9)
- Marshall, A. G., C. L. Hendrickson, and G. S. Jackson. 1998. Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **17**: 1–35. doi:[10.1002/\(SICI\)1098-2787\(1998\)17:1<1::AID-MAS1>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K)
- Mopper, K., A. Stubbins, J. D. Ritchie, H. M. Bialk, and P. G. Hatcher. 2007. Advanced instrumental approaches for characterization of marine dissolved organic matter: Extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Chem. Rev.* **107**: 419–442. doi:[10.1021/cr050359b](https://doi.org/10.1021/cr050359b)
- Novotny, N. R., E. N. Capley, and A. C. Stenson. 2014. Fact or artifact: The representativeness of ESI-MS for complex

- natural organic mixtures. *J. Mass Spectrom.* **49**: 316–326. doi:10.1002/jms.3345
- Oss, M., A. Kruve, K. Herodes, and I. Leito. 2010. Electrospray ionization efficiency scale of organic compound. *Anal. Chem.* **82**: 2865–2872. doi:10.1021/ac902856t
- Palacio Lozano, D. C., R. Gavard, J. P. Arenas-Diaz, M. J. Thomas, D. D. Stranz, E. Mejía-Ospino, A. Guzman, S. E. F. Spencer, D. Rossell, and M. P. Barrow. 2019. Pushing the analytical limits: New insights into complex mixtures using mass spectra segments of constant ultrahigh resolving power. *Chem. Sci.* **10**: 6966–6978. doi:10.1039/C9SC02903F
- Pan, Q., X. Zhuo, C. He, Y. Zhang, and Q. Shi. 2020. Validation and evaluation of high-resolution orbitrap mass spectrometry on molecular characterization of dissolved organic matter. *ACS Omega.* **5**: 5372–5379. doi:10.1021/acsomega.9b04411
- Perdue, E. M., and J. D. Ritchie. 2013. Dissolved organic matter in freshwaters. In *Treatise on geochemistry*, 2nd ed. Elsevier. doi:10.1016/B978-0-08-095975-7.00509-X
- Podgorski, D. C., A. M. McKenna, R. P. Rodgers, A. G. Marshall, and W. T. Cooper. 2012. Selective ionization of dissolved organic nitrogen by positive ion atmospheric pressure photoionization coupled with Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **84**: 5085–5090. doi:10.1021/ac300800w
- Pohlbeln, A. M., and T. Dittmar. 2015. Novel insights into the molecular structure of non-volatile marine dissolved organic sulfur. *Mar. Chem.* **168**: 86–94. doi:10.1016/j.marchem.2014.10.018
- Reemtsma, T., A. These, M. Linscheid, J. Leenheer, and A. Spitzy. 2008. Molecular and structural characterization of dissolved organic matter from the deep ocean by FTICR-MS, including hydrophilic nitrogenous organic molecules. *Environ. Sci. Technol.* **42**: 1430–1437. doi:10.1021/es7021413
- Riedel, T., and T. Dittmar. 2014. A method detection limit for the analysis of natural organic matter via Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **86**: 8376–8382. doi:10.1021/ac501946m
- Rivas-Ubach, A., Y. Liu, T. S. Bianchi, N. Tolić, C. Jansson, and L. Paša-Tolić. 2018. Moving beyond the van Krevelen diagram: A new stoichiometric approach for compound classification in organisms. *Anal. Chem.* **90**: 6152–6160. doi:10.1021/acs.analchem.8b00529
- Roth, V.-N., M. Lange, C. Simon, N. Hertkorn, S. Bucher, T. Goodall, R. I. Griffiths, P. G. Mellado-Vázquez, L. Mommer, N. J. Oram, A. Weigelt, T. Dittmar, and G. Gleixner. 2019. Persistence of dissolved organic matter explained by molecular changes during its passage through soil. *Nat. Geosci.* **12**: 755–761. doi:10.1038/s41561-019-0417-4
- Schum S. K., L. R. Mazzoleni and L. E. Brown. 2019. MFAssignR: An R package for data preparation and molecular formula assignment, R package version 0.0.4. Available at: <http://github.com/skschum/MFAssignR>.
- Simon, C., V.-N. Roth, T. Dittmar, and G. Gleixner. 2018. Molecular signals of heterogeneous terrestrial environments identified in dissolved organic matter: A comparative analysis of Orbitrap and ion cyclotron resonance mass spectrometers. *Front. Earth Sci.* **6**: 1–16. doi:10.3389/feart.2018.00138
- Sleighter, R. L., and P. G. Hatcher. 2007. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J. Mass Spectrom.* **42**: 559–574. doi:10.1002/jms.1221
- Sleighter, R. L., and P. G. Hatcher. 2008. Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mar. Chem.* **110**: 140–152. doi:10.1016/j.marchem.2008.04.008
- Sleighter, R. L., G. A. Mckee, Z. Liu, and P. G. Hatcher. 2008. Naturally present fatty acids as internal calibrants for fourier transform mass spectra of dissolved organic matter. *Limnol. Oceanogr. Methods* **6**: 246–253. doi:10.4319/lom.2008.6.246
- Smith, D. F., D. C. Podgorski, R. P. Rodgers, G. T. Blakney, and C. L. Hendrickson. 2018. 21 tesla FT-ICR mass spectrometer for ultrahigh-resolution analysis of complex organic mixtures. *Anal. Chem.* **90**: 2041–2047. doi:10.1021/acs.analchem.7b04159
- Solihat, N. N., T. Acter, D. Kim, A. F. Plante, and S. Kim. 2019. Analyzing solid-phase natural organic matter using laser desorption ionization ultrahigh resolution mass spectrometry. *Anal. Chem.* **91**: 951–957. doi:10.1021/acs.analchem.8b04032
- Southam, A. D., T. G. Payne, H. J. Cooper, T. N. Arvanitis, and M. R. Viant. 2007. Dynamic range and mass accuracy of wide-scan direct infusion nano-electrospray Fourier transform ion cyclotron resonance mass spectrometry-based metabolomics increased by the spectral stitching method. *Anal. Chem.* **79**: 4595–4602. doi:10.1021/ac062446p
- Stenson, A. C. 2008. Reversed-phase chromatography fractionation tailored to mass spectral characterization of humic substances. *Environ. Sci. Technol.* **42**: 2060–2065. doi:10.1021/es7022412
- Stenson, A. C., W. M. Landing, A. G. Marshall, and W. T. Cooper. 2002. Ionization and fragmentation of humic substances in electrospray ionization Fourier transform-ion cyclotron resonance mass spectrometry. *Anal. Chem.* **74**: 4397–4409. doi:10.1021/ac020019f
- Stenson, A. C., A. G. Marshall, and W. T. Cooper. 2003. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* **75**: 1275–1284. doi:10.1021/ac026106p
- Tang, L., and P. Kebarle. 1993. Dependence of ion intensity in electrospray mass spectrometry on the concentration of the

- analytes in the electrosprayed solution. *Anal. Chem.* **65**: 3654–3668. doi:10.1021/ac00072a020
- Williamson, C. E., E. P. Overholt, R. M. Pilla, T. H. Leach, J. A. Brentrup, L. B. Knoll, E. M. Mette, and R. E. Moeller. 2015. Ecological consequences of long-term browning in lakes. *Sci. Rep.* **5**: 1–10. doi:10.1038/srep18666
- Witt, M., J. Fuchser, and B. P. Koch. 2009. Fragmentation studies of fulvic acids using collision induced dissociation fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **81**: 2688–2694. doi:10.1021/ac802624s
- Zark, M., J. Christoffers, and T. Dittmar. 2017. Molecular properties of deep-sea dissolved organic matter are predictable by the central limit theorem: Evidence from tandem FT-ICR-MS. *Mar. Chem.* **191**: 9–15. doi:10.1016/j.marchem.2017.02.005
- Zark, M., and T. Dittmar. 2018. Universal molecular structures in natural dissolved organic matter. *Nat. Commun.* **9**: 1–8. doi:10.1038/s41467-018-05665-9

Acknowledgments

We thank two anonymous reviewers for their comments and suggestions, which improved the manuscript. We thank R.A. Payn and the Environmental Analytical Laboratory at Montana State University for laboratory space, instruments, and materials during sample preparation. J.H. held a grant from the Swedish Research Council VR (2018-04618). N.C. held a Beatriu de Pinos (2016-00215) grant and a Marie Skłodowska-Curie Action of the European Commission (CHROME-839709). C.R. and S.B. acknowledge the UW-Madison Human Proteomics Program Mass Spectrometry Facility (initially funded by the Wisconsin partnership funds) for support in obtaining mass spectrometry data, NIH S10OD018475 for the acquisition of ultrahigh resolution mass spectrometer. This work was also supported by the National Science Foundation (CBET 1802388 and DEB 1824723). P.Z. and D.C.P. were supported by Shell Global Solutions

and the Award for Building Research or Artistic Development (ABROAD) provided by the University of New Orleans Office of Research and Sponsored Programs. Financial support from the French National FT-ICR network (FR 3624 CNRS) for conducting this research is also gratefully acknowledged. A.I., M.K., S.S., and L.M. acknowledge NSF MRI grant (AGS-1531454) and the Chemical Advanced Resolution Methods Laboratory for ultrahigh resolution mass spectrometry research support. N.R.'s contribution was funded through his PhD scholarship provided by the Department of Environmental Science and Analytical Chemistry (ACES), Stockholm University, Sweden. C.S. and G.G. acknowledge financial support from the Max-Planck-Gesellschaft (MPG) and the German Research Foundation (DFG, Deutsche Forschungs-Gemeinschaft) as part of the CRC 1076 "Aqua Diva." CS received a Ph.D. stipend from the International Max Planck Research School for Global Biogeochemical Cycles (IMPRS-gBGC). D.C.P.L. and M.P.B. would like to thank the Newton Fund award (reference number 275910721), Research Agreement No. 5211770 UIS-ICP, and COLCIENCIAS (project No. FP44842-039-2015). R.G., H.E.J., M.J.T. thank EPSRC for a PhD studentship through the EPSRC Centre for Doctoral Training in Molecular Analytical Science, grant number EP/L015307/1. A portion of the research was performed using EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research. R.L.S., H.M.C., and P.G.H. were supported by the Old Dominion University Frank Batten Endowment. B.N.O. and G.S. were supported by the European Research Council (ERC-STG 716196). R.H. was supported by the Campus Alberta Innovates Program.

Submitted 01 September 2019

Revised 22 April 2020

Accepted 03 May 2020

Associate editor: Hayley Schiebel