DISSERTATION

UNRAVELING BIOGEOCHEMICAL CYCLING OF CARBON, SULFUR AND NITROGEN WITH HIGH RESOLUTION MASS SPECTROMETRY: FROM CALIFORNIA VINEYARDS TO THAWING PERMAFROST IN SWEDEN

Submitted by

Merritt N. Logan

Department of Chemistry

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2025

Doctoral Committee:

Advisor: Dr. Thomas Borch

Dr. Anthony Rappé
Dr. James Neilson
Dr. Richard Conant

ABSTRACT

UNRAVELING BIOGEOCHEMICAL CYCLING OF CARBON, SULFUR AND NITROGEN WITH HIGH RESOLUTION MASS SPECTROMETRY: FROM CALIFORNIA VINEYARDS TO THAWING PERMAFROST IN SWEDEN

Permafrost—perennially frozen ground covering approximately one-fifth of the Northern Hemisphere's land—plays a crucial role in global biogeochemical cycles. These frozen soils are vast reservoirs, holding roughly twice the organic carbon found in the atmosphere and about half of the global subterranean organic nitrogen. However, rising temperatures, particularly in the Arctic where warming is two to four times faster than the global average, threaten permafrost stability. This thawing exposes previously sequestered organic matter to decomposition, potentially releasing billions of tons of carbon and nitrogen-containing greenhouse gases (GHGs). The unpredictable, long-term behavior of permafrost during and after thaw makes it extremely challenging to forecast its impact on global climate, leading to significant uncertainties and frequent omissions from Earth system models.

While historical permafrost thaw occurred gradually, rising temperatures are now driving rapid thaw events, accelerating permafrost loss within just a few years. These rapidly thawing areas are identified as highly active GHG emitters, expected to contribute nearly half of permafrost-derived GHGs by 2300, despite impacting only about 1% of the permafrost area. Though carbonaceous GHGs often dominate research, emerging studies highlight that released nitrous oxide—formed through the mineralization of nitrogen reserves from deep, previously frozen soils—may also have a significant impact.

The biogeochemical cycling of nitrogen in thawing permafrost is complex, influenced by organic matter composition, thaw rate, soil saturation, and microbial activity. Chapter 2 addresses these complexities by characterizing organic nitrogen across a permafrost thaw gradient at Stordalen Mire, Sweden. Employing Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS) alongside Nuclear Magnetic Resonance (^{1}H NMR) spectroscopy and metatranscriptomic analyses, we found elevated ammonium and dissolved organic nitrogen concentrations in the thaw front, with a reduced proportion of peptide-like or carbohydrate molecules. These findings are critical for understanding how the molecular composition of nitrogen changes during thaw, directly impacting its bioavailability and subsequent GHG emissions.

The analytical techniques utilized for characterizing organic matter in environmental systems are further described in Chapter 3, which provides a critical review of FT-ICR MS applications. This methodological framework addresses analytical challenges and offers recommendations for sample collection, preparation, analysis, and data interpretation, serving as a vital resource for the field. This provides valuable guidelines for the continued and broader application of FT-ICR MS in a variety of environmental systems.

Returning to permafrost, Chapter 4 investigates the interactions between iron minerals and organic carbon across the thaw transition. Reactive iron minerals in permafrost soils sequester organic carbon but can release it upon thaw through reductive dissolution. Using FT-ICR MS, we identified a significant pulse of dissolved organic carbon and $Fe^{2+}_{(aq)}$ at the thaw front, with a higher proportion of aliphatic molecules observed in both dissolved and mineral–adsorbed fractions before thawing. These results suggest that reactive iron minerals at the thaw front play a

crucial role as electron acceptors for anaerobic respiration, directly influencing the fate and mobility of released organic carbon and its potential for GHG production.

Finally, Chapter 5 extends the application of FT−ICR MS to assess the environmental impact of agricultural sulfur use. In an independent study, we examined dissolved organic sulfur (DOS) in California vineyards and their downstream watersheds. Combining FT−ICR MS with sulfur isotope ($\delta^{34}S$) analysis, we detected vineyard-derived DOS in non-agricultural water systems. Our findings highlight that the mobilization of agriculturally derived organic sulfur can influence critical downstream biogeochemical processes, such as mercury methylation, underscoring the broader environmental consequences of land management practices.

In summary, this dissertation demonstrates the power of FT-ICR MS, combined with a comprehensive suite of complementary analytical methodologies, to elucidate the complex molecular transformations of organic carbon, nitrogen, and sulfur in dynamic environmental systems. This molecular-level understanding is crucial for improving predictions of future climate impacts and informing environmental management strategies in a warming world.

ACKNOWLEDGMENTS

I cannot thank my wife Emma and daughter Augusta enough for supporting me, believing in me, and keeping me going through graduate school. You put things into perspective when I was overwhelmed, you cheered me up when I needed encouragement, and you always looked out for me. I could never have finished this process without you, and I am so grateful you were part of this journey. I also need to thank my extended family who provided time, child support, stress relief, encouragement and more. You were all the most wonderful support network, and I am thrilled to share this accomplishment with you.

I am very appreciative of my advisor, Thomas Borch, who mentored me and facilitated my research opportunities, along with introducing me to so many amazing scientists. Thank you for sticking with me though the roller coaster of graduate school and encouraging me at my lowest. I am also extremely grateful to Amy McKenna; you taught me so much about being a scientist and were always empathetic and generous with your time.

I was extremely lucky to collaborate with Monique Patzner, both in planning our field campaign and collaborating on multiple projects. I learned so much about designing and structuring large projects while we were in Abisko. Through our friendship you continued to show me how to manage oneself to succeed in grad school. Thank you so much for your years of support. I was also lucky to work with Anna Hermes, you were a wonderful collaborator and source of solidarity. Thank you for your encouragement and friendship.

To all my group members, you provided invaluable feedback and support through practice talks and research updates and I am grateful to have known all of you through the years. I have learned about so many diverse topics through our interactions and I am honored to be part of such

an inquisitive and thoughtful group of scientists. Finally, I was fortunate to receive invaluable support from numerous distinguished scientists during the development of this work. I am grateful to Andreas Kappler and Casey Bryce, who supported my arctic field work and were always willing to discuss projects and results. I am grateful to Robert Young, who always made time to discuss code, results, or life in general. Finally, I would like to thank all my other collaborators and colleagues who supported these projects and me throughout my time at CSU. This includes Myrna Simpson, Jacob VanderRoest, Rachelle Rehberg, William Bahureksa, Molly McLaughlin, and Lonnie Smith Thomas. Samples from this work were analyzed at the University of Tübingen, the National High Magnetic Field Laboratory, and the University of Toronto.

Portions of this work were supported by NSF Division of Chemistry and Division of Materials Research through DMR-2128556 and the State of Florida. Support was also provided by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy, cluster of Excellence EXC2124, project ID 390838134, the University of Tübingen, the German Academic Scholarship Foundation, Discovery Grant and Tier 1 Canada Research Chair in Integrative Molecular Biogeochemistry through the Natural Sciences and Engineering Research Council of Canada, National Science Foundation Biology Integration Institutes Program, and the Genomic Science Program of the United States Department of Energy Office of Biological and Environmental Research.

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION

1. Climate Change Induced Permafrost Thaw

Permafrost—perennially frozen ground that covers approximately 15% of the northern hemisphere's land surface—plays a crucial role in global cycles, yet it faces increasing risk from progressively warmer temperatures that have risen approximately 0.5 °C per decade since 1980[1,2]. These frozen soils contain about 2.5 to 3 trillion tons of organic carbon—roughly twice the amount in the atmosphere. They also store around 67 petagrams (Pg) of organic nitrogen, approximately half of the global stock[3–7]. Beyond acting as a major carbon and nitrogen sink, permafrost regions are home to tens of millions of people, and encompass more than half of the remaining terrestrial and marine wilderness areas[6,8–12].

The stability of permafrost ecosystems is particularly at risk due to thawing, as air temperatures in these regions have risen two to four times faster than the global average[8,13–16]. As temperatures rise and permafrost ice melts, the sequestered organic matter becomes exposed to ecological processes that are highly sensitive to the availability of oxygen and water[3,17,18]. However, the effects of climate change on permafrost ecosystems are uneven, leading to varying rates of thaw and changes in hydrology, topography, and vegetation. These factors collectively influence overall greenhouse gas emissions and the long-term behavior of permafrost sites or thaw regions[8,19,20]. Until recently, permafrost emissions were frequently overlooked in Earth system models forecasting climate change. Although recent models have begun to incorporate permafrost emissions, they remain unreliable due to the wide range of long-term permafrost soil behaviors[8,21,22]. Consequently, the IPCC 2023 report underscored the necessity for more extensive permafrost research due to the global ramifications of permafrost thaw.

2. Detailed Spatial and Molecular Scale Studies Crucial for Understanding Permafrost Thaw Feedback Mechanisms

Permafrost thaw has historically been measured through large-scale observations and landscape-scale studies using satellites or remote sensing[23]. These studies reveal shifting hydrological features and changing vegetation, typically occurring over timescales ranging from decades to centuries[23,24]. While existing landscape-scale studies provide insights into the extent and rate of permafrost thawing, they often overlook detailed soil processes, such as the mineralization of carbon or nitrogen, that influence net greenhouse gas emissions[7,17,18,25–27]. Moreover, although there are studies of different thaw state ecosystems, they typically examine static thaw states rather than the dynamic thaw transitions that regulate active biogeochemical feedback mechanisms[22,28,29]. These dynamic transitions are particularly important during permafrost thaw collapse, often referred to as thermokarsts or rapid soil slumps, which result in drastic changes in soil moisture and oxygen content[14,30–33]. These highly variable conditions and their influence on biogeochemical cycling and nutrient mineralization during and after thaw determine whether permafrost acts as a source or sink for greenhouse gases[6,19]. This duality manifests in peatland systems where warmer conditions enhance CO_2 sequestration through increased plant productivity while simultaneously converting dry permafrost peatlands to wetter permafrost-free environments that release significantly more CH_4[18]. Although future emissions remain unclear, current emissions from permafrost are significant, as they are equivalent to the annual greenhouse gas emissions of Japan[33,34].

3. Biogeochemical Cycling of Nitrogen in Permafrost Regions is Poorly Understood

The complex dynamics of organic nitrogen in thawing permafrost ecosystems remain a critical area of investigation. Permafrost regions serve as crucial reservoirs for nitrogen storage, with perennially frozen soils containing vast reserves that become available during thaw[20,35]. This nitrogen release creates a biogeochemical paradox: while freshly thawed permafrost releases nitrogen that can stimulate plant activity and microbial respiration, the seasonally thawed active layer often remains nitrogen-limited due to plant uptake[20,35]. The rate, extent, and molecular composition of this released nitrogen, along with its subsequent effects on carbon cycling and greenhouse gas emissions, remain poorly understood.

The biogeochemical fate of nitrogen during permafrost thaw represents a critical knowledge gap in climate research. In permafrost soils, 94-99% of nitrogen exists in organic forms that must undergo mineralization before becoming bioavailable[7,39,40]. The subsequent transformations of this nitrogen through various pathways (mineralization, nitrification, denitrification) can lead to dramatically different ecosystem outcomes, potentially determining whether these systems function as greenhouse gas sources or sinks[36–38].

Chapter 2 addresses these knowledge gaps by employing 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS) to characterize the molecular composition of organic nitrogen across a permafrost thaw gradient at Stordalen Mire, Sweden. This study takes a multi-analytical approach, combining the ultra-high resolution capabilities of FT-ICR MS with complementary nuclear magnetic resonance (NMR) spectroscopy (for molecular structure elucidation) and metatranscriptomic sequencing (to assess microbial activity). This comprehensive integration of techniques allows for an in-depth investigation of nitrogen transformations during thaw transitions, specifically examining how different thaw states affect the composition, concentration, and bioavailability of organic nitrogen compounds,

with particular attention to the dynamic processes occurring at the critical thaw front interface. By analyzing changes in nitrogen-containing molecular classes across the thaw gradient—from frozen permafrost through actively thawing zones to fully thawed bog environments, this research offers new insights into how permafrost thaw alters nitrogen cycling and its broader implications for plant productivity, microbial activity, and greenhouse gas emissions in a warming Arctic.

4. Characterization of Organic Matter Requires Specialized Tools

Organic matter consists of a highly complex and heterogeneous molecular mixture, posing significant analytical challenges[49–51]. FT-ICR MS is often used as a characterization tool for dissolved organic matter (DOM) and soil organic matter (SOM) extracts[52,53]. FT-ICR MS operates by trapping ions in a powerful magnetic field where they cycle at frequencies specific to their mass-to-charge ratios, providing significantly higher resolving power and mass accuracy than Time-of-Flight (TOF) or Orbitrap instruments. This superior analytical capability is critical when characterizing the complex molecular mixtures of DOM and SOM, where thousands of compounds with nearly identical masses must be distinguished at sub-parts-per-million accuracy[53–55]. The development of the 21 Tesla FT-ICR MS has greatly increased resolving power, allowing for 49,000 additional peak assignments over 9.4 Tesla instruments[56]. The parts-per-billion (ppb) resolution provided by FT-ICR MS allows differentiation of the tens of thousands of m/z peaks that make up organic matter[53]. While the high mass accuracy of FT-ICR MS enables precise assignment of monoisotopic molecular formulas to organic matter, achieving a truly representative characterization remains challenging due to biases introduced during ionization and prior sample preparation[53].

Electrospray ionization is the most widely utilized ionization method due to its function as a "soft" ionization, which creates intact molecular ions. In ESI, sample molecules in solution pass through a charged capillary, creating a fine spray of charged droplets that subsequently evaporate, transferring charge to the analyte molecules. The polarity of the applied voltage determines the ionization mode: negative mode applies a negative voltage, preferentially ionizing acidic compounds by deprotonation, while positive mode applies a positive voltage, preferentially ionizing basic compounds through protonation or metal adduction[53,57]. It can be performed in positive or negative mode depending on the target molecules and the sample matrix. Negative ion mode is more widely used for organic matter analysis due to the prevalence of acidic, generally carboxylic acid, functional groups[53]. However, these acidic motifs that improve negative ionization efficiency can suppress the ionization of nitrogen-containing organic species, which more efficiently form positive ions[53,58]. Therefore, one approach to target nitrogen-containing motifs is to utilize positive mode electrospray ionization. However, due to inherent sample and molecular composition differences, positive ESI is not a reliable solution for targeting nitrogen organic matter[53].

While FT-ICR MS excels at detecting individual organic molecules, it does not provide detailed molecular structures. Instead, it uses elemental ratios to infer the most probable molecular classes, which can be compared across samples to assess organic matter composition[59]. However, these classes do not directly measure molecular types or structural abundance. To address this limitation, FT-ICR MS results can be complemented with additional molecular analyses. Nuclear magnetic resonance spectroscopy provides valuable structural information to corroborate FT-ICR MS findings. Although standard ^{1}H or ^{13}C NMR approaches do not specifically target nitrogen compounds, they offer insights into carbon functionality and

hydrogen environments within the organic matter[53,60]. While ^{15}N-NMR could theoretically provide direct information about nitrogen structures, its practical application for environmental DOM/SOM samples is limited by the low natural abundance of ^{15}N and the typically low overall nitrogen content in these samples.

The methodological challenges and considerations for applying FT-ICR MS to environmental organic matter samples are addressed in a comprehensive critical review presented in Chapter 3. This review synthesizes current best practices for soil organic matter characterization using FT-ICR MS, with detailed guidance on sample collection, preparation, instrumental analysis, and data interpretation. The review establishes standardized workflows for processing the complex datasets generated by FT-ICR MS analysis of environmental samples and identifies critical research gaps and biases that must be addressed to advance our understanding of organic matter chemistry in terrestrial ecosystems. By providing this methodological framework, the review serves as a foundation for the analytical approaches applied throughout this dissertation to characterize organic nitrogen, carbon, and sulfur in diverse environmental contexts.

5. Investigating Iron-Bound Organic Carbon Dynamics in Thawing Permafrost

Chapter 4 explores the relationship between reactive iron minerals and organic carbon release during permafrost thaw. At Stordalen Mire, oxy-hydroxy iron minerals are abundant in the soil, with approximately 20% of the present organic matter associated with these minerals[61]. This mineral-organic association raises important questions about carbon stability during thaw transitions. As permafrost thaws, reductive dissolution processes may release previously protected iron-bound organic carbon to porewaters, potentially making this carbon vulnerable to microbial decomposition and subsequent greenhouse gas emissions.

The interactions between mineral iron stability and microbial processes during transitional permafrost thaw remain poorly understood, yet have significant implications for carbon cycling and climate feedback mechanisms. A significant contribution to this research was the implementation of FT-ICR MS to characterize the molecular composition of organic matter in both porewater and mineral-bound fractions. This high-resolution analytical approach enabled insights into the chemical transformations of organic carbon during mineral dissolution processes. Chapter 4 details how FT-ICR MS analysis revealed distinct molecular signatures of released organic matter along thaw gradients, specifically identifying the mobilization of aliphatic organic compounds during iron mineral dissolution. The molecular-level characterization of these organic compounds provided critical evidence connecting mineral-organic interactions to carbon bioavailability and subsequent greenhouse gas production in transitional thaw environments.

6. Molecular Fingerprinting of Agriculturally-Derived Organic Sulfur in Watershed Systems

Agricultural sulfur (S) applications have dramatically increased over the past several decades, raising important questions about environmental impacts beyond their intended agricultural benefits. Excess sulfur in the environment can lead to numerous biogeochemical and ecological consequences, including enhanced methylmercury production, which poses a significant environmental and public health concern. While inorganic sulfur cycling has been well-studied, organic sulfur, the dominant form within soils, has received considerably less attention despite its importance in environmental processes. The transformations and transport of dissolved organic sulfur (DOS) from agricultural lands to watershed systems represent a critical knowledge gap in our understanding of how human activities affect environmental sulfur

cycling. Agricultural inputs can modify not only the quantity but also the molecular composition of DOS, with implications for downstream ecosystems. These modifications may create unique sulfur-containing organic compounds that persist through the watershed, potentially influencing biogeochemical processes far from their original source. Chapter 5 leverages FT-ICR MS to identify and track sulfur-containing organic molecules through the Napa River watershed. The molecular-level precision of FT-ICR MS revealed distinctive $CHOS_2$ compounds unique to vineyard-influenced waters that persisted from field to watershed scales. This molecular fingerprinting demonstrated how agricultural sulfur applications create characteristic organic compounds that can be traced through environmental systems, establishing FT-ICR MS as a powerful tool for connecting land management practices to watershed-scale biogeochemical impacts.

7. Publications and Presentations

This dissertation applies FT-ICR MS as part of a suite of analytical tools to elucidate the composition of organic nitrogen and carbon across a permafrost thaw transition and agricultural systems. Chapter 2 characterizes organic nitrogen across a permafrost thaw gradient using FT-ICR MS supported by ^{1}H NMR and metatranscriptome analysis. This research was published in *Soil and Environmental Health*[62]. Chapter 3 presents a comprehensive review of FT-ICR MS applications to organic matter analysis, addressing analytical challenges and providing recommendations for sample collection, preparation, analysis, and data interpretation. This methodological framework was published in *Environmental Science & Technology*[53]. Chapter 4 investigates both dissolved organic carbon and mineral-associated organic carbon across a permafrost thaw transition, integrating FT-ICR MS with gas analysis, geochemical measurements, microbial community profiling, and Mössbauer spectroscopy. As co-author, I led all FT-ICR MS-

related sampling and analysis, contributed to gas sampling, and co-planned the field campaign. This work was published in *Nature Communications Earth and Environment*[63]. Chapter 5 examines organic sulfur composition in California vineyards, combining FT-ICR MS with $\delta^{34}S$ analysis and XANES spectroscopy. As co-author, I designed the FT-ICR MS sampling approach and led the FT-ICR MS data analysis and interpretation. This research was published in *Environmental Science & Technology*[64]. Research from this dissertation was also presented at the American Chemical Society Spring 2021 conference ("Ultrahigh resolution mass spectrometry analysis of organic nitrogen released upon permafrost thaw") and at the 2024 Goldschmidt conference ("Rapid transformation of peptide-like organic nitrogen upon permafrost thaw").

8. REFERENCES

1 Obu J, How Much of the Earth's Surface is Underlain by Permafrost?, *J Geophys Res Earth Surf* **126**:1–5 (2021).

2 Smith SL, O'Neill HB, Isaksen K, Noetzli J, and Romanovsky VE, The changing thermal state of permafrost, *Nat Rev Earth Environ* **3**:10–23, Springer US (2022).

3 Voigt C, Marushchak ME, Lamprecht RE, Jackowicz-Korczyński M, Lindgren A, Mastepanov M, *et al.*, Increased nitrous oxide emissions from Arctic peatlands after permafrost thaw, *Proc Natl Acad Sci U S A* **114**:6238–6243 (2017).

4 Batjes NH, Total carbon and nitrogen in the soils of the world., *Eur J Soil Sci* **47**:151–163 (1996).

5 Harden JW, Koven CD, Ping CL, Hugelius G, David McGuire A, Camill P, *et al.*, Field information links permafrost carbon to physical vulnerabilities of thawing, *Geophys Res Lett* **39**:1–6 (2012).

6 Schuur EAG, Abbott BW, Commane R, Ernakovich J, Euskirchen E, Hugelius G, *et al.*, Permafrost and Climate Change: Carbon Cycle Feedbacks From the Warming Arctic, *Annu Rev Environ Resour* **47**:343–371 (2022).

7 Voigt C, Marushchak ME, Abbott BW, Biasi C, Elberling B, Siciliano SD, *et al.*, Nitrous oxide emissions from permafrost-affected soils, *Nat Rev Earth Environ* **1**:420–434, Springer US (2020).

8 Abbott BW, Brown M, Carey JC, Ernakovich J, Frederick JM, Guo L, *et al.*, We Must Stop Fossil Fuel Emissions to Protect Permafrost Ecosystems, *Front Environ Sci* **10**:1–13 (2022).

9 Hugelius G, Strauss J, Zubrzycki S, Harden JW, Schuur EAG, Ping CL, *et al.*, Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, *Biogeosciences* **11**:6573–6593 (2014).

10 Abbott BW, Jones JB, Schuur EAG, Chapin FS, Bowden WB, Bret-Harte MS, *et al.*, Biomass offsets little or none of permafrost carbon release from soils, streams, and wildfire: An expert assessment, *Environ Res Lett* **11**, IOP Publishing (2016).

11 Sayedi SS, Abbott BW, Thornton BF, Frederick JM, Vonk JE, Overduin P, *et al.*, Subsea permafrost carbon stocks and climate change sensitivity estimated by expert assessment, *Environ Res Lett* **15** (2020).

12 Mishra U, Hugelius G, Shelef E, Yang Y, Strauss J, Lupachev A, *et al.*, Spatial heterogeneity and environmental predictors of permafrost region soil organic carbon stocks, *Sci Adv* **7**:1–12 (2021).

13 Huang J, Zhang X, Zhang Q, Lin Y, Hao M, Luo Y, *et al.*, Recently amplified arctic warming has contributed to a continual global warming trend, *Nat Clim Chang* **7**:875–879, Springer US (2017).

14 Grosse G, Harden J, Turetsky M, McGuire AD, Camill P, Tarnocai C, *et al.*, Vulnerability of high-latitude soil organic carbon in North America to disturbance, *J Geophys Res Biogeosciences* **116**:1–23 (2011).

15 Mu C, Schuster PF, Abbott BW, Kang S, Guo J, Sun S, *et al.*, Permafrost degradation enhances the risk of mercury release on Qinghai-Tibetan Plateau, *Sci Total Environ* **708**:135127, Elsevier B.V. (2020).

16 Polvani LM, Previdi M, England MR, Chiodo G, and Smith KL, Substantial twentieth-century Arctic warming caused by ozone-depleting substances, *Nat Clim Chang* **10**:130–133, Springer US (2020).

17 Abbott BW and Jones JB, Permafrost collapse alters soil carbon stocks, respiration, CH4, and N2O in upland tundra, *Glob Chang Biol* **21**:4570–4587 (2015).

18 Hugelius G, Loisel J, Chadburn S, Jackson RB, Jones M, MacDonald G, *et al.*, Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw, *Proc Natl Acad Sci U S A* **117**:20438–20446 (2020).

19 Schuur EAG, Bockheim J, Canadell JG, Euskirchen E, Field CB, Goryachkin S V., *et al.*,

Vulnerability of Permafrost Carbon to Climate Change: Implications for the Global Carbon Cycle, *Bioscience* **58**:701–714 (2008).
20 Wegner R, Fiencke C, Knoblauch C, Sauerland L, and Beer C, Rapid Permafrost Thaw Removes Nitrogen Limitation and Rises the Potential for N2O Emissions, *Nitrogen* **3**:608–627 (2022).
21 Fischlin A, Midgley GF, Price J, Leemans R, Gopal B, Turley C, *et al.*, Ecosystems, their properties, goods and services, ed. by Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, and Hanson CE, Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution ofWorking Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, pp. 211–272 (2007).
22 Chang KY, Riley WJ, Crill PM, Grant RF, Rich VI, and Saleska SR, Large carbon cycle sensitivities to climate across a permafrost thaw gradient in subarctic Sweden, *Cryosphere* **13**:647–663 (2019).
23 Grosse G, Goetz S, McGuire AD, Romanovsky VE, and Schuur EAG, Changing permafrost in a warming world and feedbacks to the Earth system, *Environ Res Lett* **11**, IOP Publishing (2016).
24 Torre Jorgenson M, Harden J, Kanevskiy M, O'Donnell J, Wickland K, Ewing S, *et al.*, Reorganization of vegetation, hydrology and soil carbon after permafrost degradation across heterogeneous boreal landscapes, *Environ Res Lett* **8** (2013).
25 Miner KR, Turetsky MR, Malina E, Bartsch A, Tamminen J, McGuire AD, *et al.*, Permafrost carbon emissions in a changing Arctic, *Nat Rev Earth Environ* **3**:55–67, Springer US (2022).
26 Mann BF, Chen H, Herndon EM, Chu RK, Tolic N, Portier EF, *et al.*, Indexing permafrost soil organic matter degradation using high-resolution mass spectrometry, *PLoS One* **10**:1–16 (2015).
27 Fiencke C, Marushchak ME, Sanders T, Wegner R, and Beer C, Microbiogeochemical Traits to Identify Nitrogen Hotspots in Permafrost Regions, *Nitrogen* **3**:458–501 (2022).
28 Wallenstein MD, Hoyt DW, Cotrufo MF, Rithner CD, Covino TP, Boot CM, *et al.*, Dissolved organic matter chemistry and transport along an Arctic tundra hillslope, *Global Biogeochem Cycles*:1–16 (2019).
29 Olefeldt D, Heffernan L, Jones MC, Sannel ABK, Treat CC, and Turetsky MR, Permafrost Thaw in Northern Peatlands: Rapid Changes in Ecosystem and Landscape Functions BT - Ecosystem Collapse and Climate Change, ed. by Canadell JG and Jackson RB, Springer International Publishing, Cham, pp. 27–67 (2021).
30 Bowden WB, Gooseff MN, Balser A, Green A, Peterson BJ, and Bradford J, Sediment and nutrient delivery from thermokarst features in the foothills of the North Slope, Alaska: Potential impacts on headwater stream ecosystems, *J Geophys Res Biogeosciences* **113**:1–12 (2008).
31 Strauss J, Schirrmeister L, Grosse G, Fortier D, Hugelius G, Knoblauch C, *et al.*, Deep Yedoma permafrost: A synthesis of depositional characteristics and carbon vulnerability, *Earth-Science Rev* **172**:75–86, Elsevier (2017).
32 Mao C, Kou D, Wang G, Peng Y, Yang G, Liu F, *et al.*, Trajectory of Topsoil Nitrogen Transformations Along a Thermo-Erosion Gully on the Tibetan Plateau, *J Geophys Res Biogeosciences* **124**:1342–1354 (2019).
33 Turetsky MR, Abbott BW, Jones MC, Anthony KW, Olefeldt D, Schuur EAG, *et al.*, Carbon release through abrupt permafrost thaw, *Nat Geosci* **13**:138–143, Springer US (2020).
34 Cryopshere I, Initiative C, Centre B, Climate W, Network PC, Stocktake G, *et al.*, Global Stocktake Submission : " Country " of Permafrost, 1–5 (2022).
35 Salmon VG, Schädel C, Bracho R, Pegoraro E, Celis G, Mauritz M, *et al.*, Adding Depth to Our Understanding of Nitrogen Dynamics in Permafrost Soils, *J Geophys Res Biogeosciences* **123**:2497–2512 (2018).
36 Hansen HFE and Elberling B, Spatial Distribution of Bioavailable Inorganic Nitrogen From Thawing Permafrost, *Global Biogeochem Cycles* **37**:1–12 (2023).
37 Marushchak ME, Kerttula J, Diáková K, Faguet A, Gil J, Grosse G, *et al.*, Thawing Yedoma permafrost is a neglected nitrous oxide source, *Nat Commun* **12**, Springer US (2021).
38 Ramage J, Kuhn M, Virkkala AM, Voigt C, Marushchak ME, Bastos A, *et al.*, The Net GHG

Balance and Budget of the Permafrost Region (2000–2020) From Ecosystem Flux Upscaling, *Global Biogeochem Cycles* **38**:1–18 (2024).
39 Kicklighter DW, Melillo JM, Monier E, Sokolov AP, and Zhuang Q, Future nitrogen availability and its effect on carbon sequestration in Northern Eurasia, *Nat Commun* **10**, Springer US (2019).
40 Wild B, Schnecker J, Alves RJE, Barsukov P, Bárta J, Čapek P, *et al.*, Input of easily available organic C and N stimulates microbial decomposition of soil organic matter in arctic permafrost soil, *Soil Biol Biochem* (2014).
41 Francis A, Ganeshram RS, Tuerena RE, Spencer RGM, Holmes RM, Rogers JA, *et al.*, Permafrost degradation and nitrogen cycling in Arctic rivers: Insights from stable nitrogen isotope studies, *Biogeosciences* **20**:365–382 (2023).
42 Kuypers MMM, Marchant HK, and Kartal B, The microbial nitrogen-cycling network, *Nat Rev Microbiol* **16**:263–276, Nature Publishing Group (2018).
43 Ramm E, Liu C, Ambus P, Butterbach-Bahl K, Hu B, Martikainen PJ, *et al.*, A review of the importance of mineral nitrogen cycling in the plant-soil-microbe system of permafrost-affected soils—changing the paradigm, *Environ Res Lett* **17**:013004 (2022).
44 AminiTabrizi R, Wilson RM, Fudyma JD, Hodgkins SB, Heyman HM, Rich VI, *et al.*, Controls on Soil Organic Matter Degradation and Subsequent Greenhouse Gas Emissions Across a Permafrost Thaw Gradient in Northern Sweden, *Front Earth Sci* **8** (2020).
45 Hodgkins SB, Tfaily MM, Podgorski DC, McCalley CK, Saleska SR, Crill PM, *et al.*, Elemental composition and optical properties reveal changes in dissolved organic matter along a permafrost thaw chronosequence in a subarctic peatland, *Geochim Cosmochim Acta* **187**:123–140, Elsevier Ltd (2016).
46 Hodgkins SB, Tfaily MM, McCalley CK, Logan TA, Crill PM, Saleska SR, *et al.*, Changes in peat chemistry associated with permafrost thaw increase greenhouse gas production, *Proc Natl Acad Sci* **111**:5819–5824 (2014).
47 Gillespie AW, Sanei H, Diochon A, Ellert BH, Regier TZ, Chevrier D, *et al.*, Perennially and annually frozen soil carbon differ in their susceptibility to decomposition: Analysis of Subarctic earth hummocks by bioassay, XANES and pyrolysis, *Soil Biol Biochem* **68**:106–116, Elsevier Ltd (2014).
48 Gregorich EG, Gillespie AW, Beare MH, Curtin D, Sanei H, and Yanni SF, Evaluating biodegradability of soil organic matter by its thermal stability and chemical composition, *Soil Biol Biochem* **91**:182–191, Elsevier Ltd (2015).
49 Lehmann J, Hansel CM, Kaiser C, Kleber M, Maher K, Manzoni S, *et al.*, Persistence of soil organic carbon caused by functional complexity, *Nat Geosci*:1–8 (2020).
50 Paul EA, The nature and dynamics of soil organic matter: Plant inputs, microbial transformations, and organic matter stabilization, *Soil Biol Biochem* **98**:109–126, Elsevier Ltd (2016).
51 D'Andrilli J, Cooper WT, Foreman CM, and Marshall AG, An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability, *Rapid Commun Mass Spectrom* **29**:2385–2401 (2015).
52 Herzsprung P, Hertkorn N, Von Tümpling W, Harir M, Friese K, and Schmitt-Kopplin P, Molecular formula assignment for dissolved organic matter (DOM) using high-field FT-ICR-MS: Chemical perspective and validation of sulphur-rich organic components (CHOS) in pit lake samples, *Anal Bioanal Chem* **408**:2461–2469 (2016).
53 Bahureksa W, Tfaily MM, Boiteau RM, Young RB, Logan MN, McKenna AM, *et al.*, Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation, *Environ Sci Technol* **55**:9637–9656 (2021).
54 Hawkes JA, D'Andrilli J, Agar JN, Barrow MP, Berg SM, Catalán N, *et al.*, An international laboratory comparison of dissolved organic matter composition by high resolution mass spectrometry: Are we getting the same answer?, *Limnol Oceanogr Methods* **18**:235–258 (2020).
55 Lai YH and Wang YS, Advances in high-resolution mass spectrometry techniques for analysis of

high mass-to-charge ions, *Mass Spectrom Rev* **42**:2426–2445 (2023).
56 Banerjee S and Mazumdar S, Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte, *Int J Anal Chem* **2012**:1–40 (2012).
57 Sleighter RL and Hatcher PG, 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 (2007).
58 Roth HK, Borch T, Young RB, Bahureksa W, Blakney GT, Nelson AR, *et al.*, Enhanced Speciation of Pyrogenic Organic Matter from Wildfires Enabled by 21 T FT-ICR Mass Spectrometry, *Anal Chem* **94**:2973–2980 (2022).
59 Rivas-Ubach A, Liu Y, Bianchi TS, Tolić N, Jansson C, and Paša-Tolić L, Moving beyond the van Krevelen Diagram: A New Stoichiometric Approach for Compound Classification in Organisms, *Anal Chem* **90**:6152–6160 (2018).
60 Simpson AJ, McNally DJ, and Simpson MJ, NMR spectroscopy in environmental research: From molecular interactions to global processes, *Prog Nucl Magn Reson Spectrosc* **58**:97–175 (2011).
61 Patzner MS, Mueller CW, Malusova M, Baur M, Nikeleit V, Scholten T, *et al.*, Iron mineral dissolution releases iron and associated organic carbon during permafrost thaw, *Nat Commun* **11**:1–39 (2020).
62 Logan MN, Patzner MS, VanderRoest JP, McGivern BB, Srikanthan N, Simpson MJ, *et al.*, Role of permafrost thaw transitions in biogeochemical nitrogen cycling, *Soil Environ Heal* **3**:100148, The Author(s) (2025).
63 Patzner M, Logan M, McKenna A, Young R, Zhou Z, Joss H, *et al.*, Microbial iron cycling during palsa hillslope collapse promotes greenhouse gas emissions before complete permafrost thaw, *Commun Earth Environ* **3** (2022).
64 Hermes AL, Logan MN, Poulin BA, McKenna AM, Dawson TE, Borch T, *et al.*, Agricultural Sulfur Applications Alter the Quantity and Composition of Dissolved Organic Matter from Field-to-Watershed Scales, *Environ Sci Technol* (2023).

CHAPTER 2 - ROLE OF PERMAFROST THAW TRANSITIONS IN BIOGEOCHEMICAL NITROGEN CYCLING[1]

1. INTRODUCTION

Approximately 15% of land area in the northern hemisphere is underlain by permafrost, serving as a major sink for both carbon and nitrogen[2–7]. The northern circumpolar permafrost region stores an estimated 1,600 petagrams of carbon and 67 petagrams of nitrogen accumulated over thousands of years in frozen or saturated anoxic soils, representing approximately 33% and 50% of the total global carbon and nitrogen stocks respectively[8–11]. Climate change is expected to release large quantities of these stocks as greenhouse gas (GHG), and between 1971 and 2019 warming in permafrost regions has increased two to four times faster than global averages, accelerating permafrost loss and environmental disturbances[12,13].

Permafrost peatlands, a subset of permafrost, cover ~3.7 million km^2 and store ~415 petagrams of carbon and ~10 petagrams of nitrogen[5,14–16]. Many permafrost peatlands are especially susceptible to thawing as their average annual near-surface temperatures are within 1°C of freezing[17]. Therefore, minimal warming can raise average annual temperatures above 0°C, inducing thaw and exposing the stored organic carbon and nitrogen to biogeochemical mineralization[18,19]. This increased mineralization is expected to release substantial amounts of GHG, contributing to climate change and further accelerating thawing[20]. While many studies focus on carbon cycling and carbonaceous GHG[21–23], research on nitrogen cycling and nitrogenous GHG release is comparatively limited. However, the fate of organic nitrogen is drawing increased

[1] Logan, M. N., Patzner, M. S., VanderRoest, J. P., McGivern, B. B., Srikanthan, N., Simpson, M. J., et al. (2025). Role of permafrost thaw transitions in biogeochemical nitrogen cycling. *Soil and Environmental Health*, *3*(2), 100148. https://doi.org/10.1016/j.seh.2025.100148

interest due in part to its potential transformation to N_2O (a greenhouse gas with a warming potential 298 times greater than CO_2[5,8,24,25]) and the important role that nitrogen plays in nutrient cycling.

Nitrogen availability and utilization in permafrost peatlands are influenced by seasonal thaw depth and thaw conditions[26–29]. The active layer, which thaws each summer, varies in depth depending on seasonal conditions and temperature. Within this layer, the upper 5-15 cm contains lower nitrogen concentrations than deeper, perennially frozen soils. However, nitrogen bioavailability is limited by either restricted root access within the deeper active layer or by frozen conditions in the permafrost below[30,31]. Research has demonstrated that active layer expansion of ~10 cm into permafrost can result in a sevenfold increase in dissolved inorganic nitrogen (DIN) release compared to seasonally thawed regions[30,31]. Nitrogen inputs to peatlands through biological fixation and atmospheric deposition are relatively minor compared to stored nitrogen content[32], despite varying with landscape and vegetation type. In permafrost systems, thaw events provide the primary source of nitrogen for biogeochemical cycling [29,31,33,34].

Nitrogen cycling in terrestrial ecosystems involves complex interactions between organic and inorganic forms, mediated by microbial processes and environmental conditions[35]. In permafrost systems, these cycles are particularly important as they influence ecosystem productivity and greenhouse gas emissions[25,36,37]. Recently, Ramm *et al.* (2022) proposed an updated model for nitrogen cycling within permafrost soils suggesting higher nitrogen cycling activity in seasonally thawed surface soils than previously presumed[32]. This model illustrates the relationships between organic nitrogen pools and the production of bioavailable nitrogen forms (ammonium, nitrate, and N_2). Their work indicates that while protein depolymerization rates are high, they do not directly control nitrogen turnover. Instead, dissolved organic nitrogen (DON)

concentration during seasonal thawing appears to be the primary driver of inorganic nitrogen cycling[32]. Furthermore, the composition of the DON can affect ammonification and production of inorganic nitrogen; however, the controls of ammonification in thawing permafrost are still not well-studied[29,32,38,39]. Soil conditions during thaw also affect inorganic nitrogen concentrations by influencing the activity of nitrogen cycling pathways[29,40]. For example, thaw transition studies have observed increased nitrate concentrations in thawed soil with moderate moisture (40-80% water filled pore space) adjacent to existing permafrost. These studies suggest that increased nitrate concentrations could enhance N_2O emissions via denitrification, provided that an active nitrogen reducing microbial community is present[39,41]. Ultimately, nitrogen demand and fate in permafrost systems are controlled by the interplay of available ammonium, nitrification and denitrification activity, and soil moisture conditions. In permafrost peatlands, thaw creates distinct moisture interfaces that manifest as either thermokarst features or thaw slumps[39,41,42]. When thaw occurs adjacent to bog areas, soil can collapse and become saturated depending on water table depth. In thaw slumps, rather than exhibiting the peat breakage characteristic of thermokarsts, the peat surface undergoes elevation loss while maintaining structural integrity[43]. These thaw transitions represent zones of variable water saturation and rapidly shifting microbial communities and have been identified as potential hotspots of GHG release due to their fluctuating oxic and anoxic conditions[29,39,44]. However, to our knowledge, previous studies examining nitrogen dynamics across thaw transitions have not investigated fully saturated thawed soil conditions[39,41], nor have they characterized the composition of the organic nitrogen pool[45–49]. Understanding biogeochemical nitrogen cycling within these dynamic thaw front environments requires detailed characterization of both the released organic nitrogen composition and the expression of microbial genes involved in nitrogen transformation pathways[8,32,50].

Evaluating organic nitrogen composition within complex organic matter requires molecular-level analytical techniques capable of resolving tens of thousands of polyfunctional molecules. Such characterization demands a mass analyzer with high dynamic range, high sensitivity, and ultrahigh resolving power across a wide molecular weight range (m/z 150-1500)[51,52]. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) uniquely meets these requirements, with the 21 Tesla hybrid linear ion trap/FT-ICR mass analyzer achieving the highest resolving power to date (3,000,000 at *m/z* 200)[51]. The ultrahigh resolution of 21T FT-ICR MS enables accurate molecular formula assignments within complex dissolved organic matter (DOM) matrices[36,53–56]. This capability allows for post-analysis isolation of nitrogen-containing formulae, separating them from the larger, total DOM pool. Such isolation enhances our ability to track patterns of selective degradation or preservation of nitrogenous species that would otherwise be obscured in the total DOM pool. While FT-ICR MS has been successfully employed to study OM composition changes between bogs and fens and in laboratory incubation studies, it has not yet been applied to characterize organic nitrogen within a thaw front[49,57–61].

In this study, we examined the thaw transition from palsa to thaw front to bog, which represents the progressive stages of permafrost thaw. We present a molecular-level characterization of DON and inorganic nitrogen using electrospray ionization FT-ICR MS, NMR spectroscopy, and elemental analysis, allowing us to examine the changes in the composition of both organic and inorganic nitrogen and corroborate shifts in molecular structure. In addition, we complemented our chemical analysis by examining previously collected palsa and bog samples for the expression of genes associated with microbial nitrogen cycling pathways. Although these samples were collected during a separate field campaign than the porewater samples used for molecular and elemental analysis, the site conditions are highly comparable. This allowed us to

evaluate the microbiome response in these distinct systems and provide a possible explanation for the biogeochemical shifts we observed. We hypothesized that (i) the thaw front will have elevated ammonium concentrations but not elevated nitrate concentrations due to wet, anoxic conditions, (ii) across the palsa-to-thaw front transition, peptide-like molecules will be depleted while aromatic molecules will be more abundant due to preferential microbial degradation, and (iii) the predominantly anoxic bog will not show increased microbial gene expression associated with denitrification due to nitrate limitations.

2. Materials and Methods

2.1 Site description

Field sampling was conducted at Stordalen Mire (68 22' N, 19 03'E), a subarctic peatland underlain by discontinuous permafrost in northern Sweden with peat depth between 1 and 3 meters[62]. This site has been studied since the 1970s, with anthropogenic warming driving ecosystem changes through permafrost thaw[15,63,64]. The site comprises three distinct ecosystems (palsa, bog, and fen), whose characteristics and boundaries have remained stable over time, as demonstrated through DOM and gas flux measurements [15,36,43,47,57,65,66]. Although at our sampling site the position of the thaw front has progressed into the palsa, the adjacent palsa and bog have been stable ecosystems since at least 2000[67]. Greenhouse gas emissions at Stordalen are primarily controlled by thaw state and saturation conditions rather than interannual variation, and ecosystem differences remain consistent interannually[43,66]. This temporal stability supports the validity of comparing our samples collected at different time points (cores in 2016; porewater in 2019 and 2022).

We selected Stordalen Mire because its palsa-to-bog thaw transition provides a well-defined system where the shift to waterlogged, anoxic conditions coincides with reductive

dissolution of Fe(III) minerals and substantial DOC release and oxidation[36]. The sampling location within the mire as well as a photo of the sampling site are shown in Figure 1. The three dominant vegetation communities of the mire are (1) ericaceous and woody plants in palsa (intact permafrost); (2) *Sphagnum spp.,* sedges, and shrubs in ombrotrophic peatland or bog (intermediate thaw); (3) sedges, mainly *Eriophorum spp.* in minerotrophic peatland or fen (full permafrost thaw)[68,69].The palsa is oxic, features pH values of 4.5-5.5, and is elevated by underlying intact permafrost, leading to higher elevation than nearby bogs[70,71]. Bogs generally feature pH values of ~3.6, are waterlogged, and contain deep underlying permafrost[48,49,67,72].

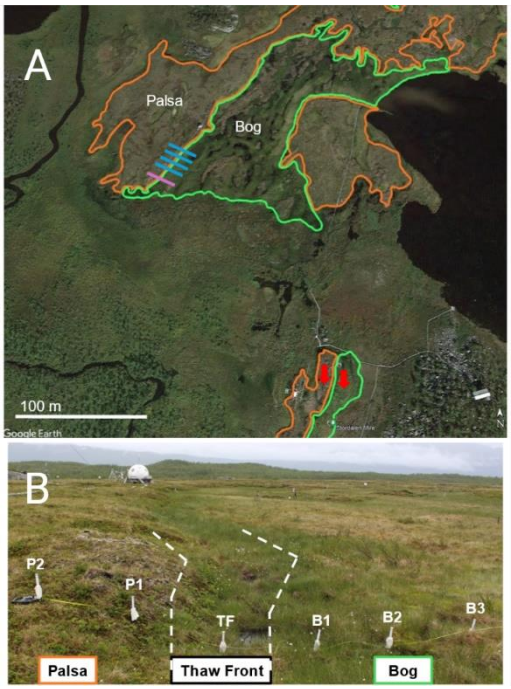


Figure 1. (A) Stordalen mire, Abisko (Sweden) shown with palsa and bog regions. Transects 1-4 are shown as blue lines, transect 5 is shown as a purple line, and the soil cores used for metatranscriptome analysis are shown as red arrows. Palsa and bog regions are based on reported hydrological data[15,68,73] and first hand observation during sampling. Maps data: Google Earth, CNES/Airbus, Lantmateriet/Metria, Maxar Technologies; copyright 2024. (B) Transect one lysimeters shown crossing the thaw transition. Palsa, thaw front, and bog regions in addition to sample points for transect one are shown. Dashed lines show thaw front moving away from the camera. Photo was taken looking north at 68°21'18.70"N, 19° 2'38.00"E.

2.2 Sample Collection

Sampling was conducted along four transects incorporating the thaw transition (Figure 2A). A fifth transect was collected three meters adjacent to the initial sampling grid for inorganic and organic nitrogen content analysis. Transects 1-4 were collected in June/July 2019 and transect 5 was collected in September 2022. Transect locations were based on previously sampled palsa and bog locations[74]. Thaw front location was determined by measuring rapid reductive dissolution of reactive iron(III) (oxyhydr)oxide minerals due to soil saturation, releasing a pulse of aqueous Fe^{2+} [36]. Therefore, the thaw front position was defined as the point in the transect with the highest $Fe^{2+}_{(aq)}$ concentration, correlating with the point of soil saturation. Due to natural variation in the thaw progression, the thaw front occurred at different points in our sampling transects (Figure 2A). Based on thaw front location, the palsa (defined as intact permafrost) and the bog (opposing the palsa on the other side of the thaw front) were both labeled according to their respective distances from the thaw front (Figure 2A). Porewater was collected along each transect at a depth of 30 cm, corresponding to the interface between the seasonally thawed active layer and the underlying perennially frozen permafrost, and consistently located within a transition zone between the organic and mineral horizons[36]. Samples were labeled according to their position in the transect (Figure 2B). The thawed-frozen interface serves as a lateral flow path, directing water from precipitation and active layer deepening from palsa to bog[75]. Porewater was collected using lysimeters that were spaced one meter apart along a given transect. The collected porewater was syringe filtered (0.22 μm) and transferred to dinitrogen-flushed SCHOTT bottles that were wrapped in aluminum foil to prevent photodegradation. The samples were then stored at 4 °C until further analysis, and sample volumes are reported in SI Table S1. Instrument access and sample volume limitations necessitated the use of select transects for FT-ICR MS and NMR, which are shown in Figure 2A. Further description of porewater sampling can be found in previous work[36].

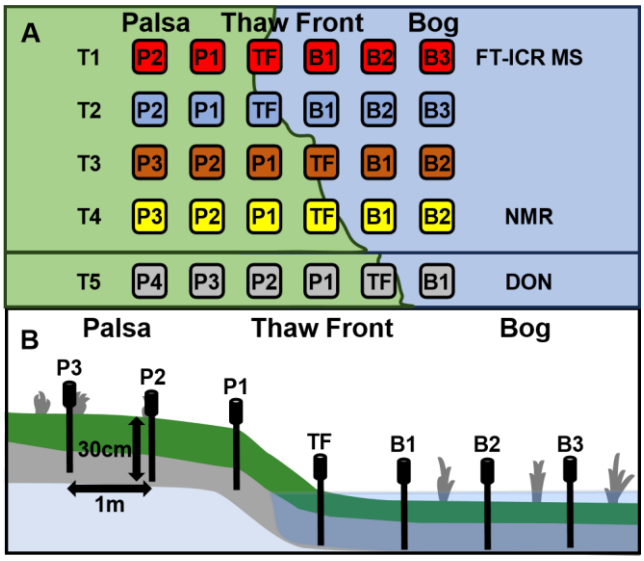


Figure 2. (A) Top-down view of porewater sampling transects with sample points marked by thaw conditions along the transect. Palsa (P) and bog (B) are numbered according to their distance from the thaw front. The thaw front (TF) is identified as the point with the maximum $Fe^{2+}_{(aq)}$ concentration in each transect. Transects are labeled as T1-4 and are spaced 1 meter apart. Transect 5 is located approximately 3 m to the south of T4 and crosses the same thaw transition. Transects 1-4 were used for elemental analysis. FT-ICR MS and NMR analyses were conducted on porewater from T1 and T4, respectively, and dissolved organic nitrogen (DON) was measured for T5. (B) Cross-section of porewater sampling within each transect showing sampling depth at 30 cm, lysimeters spaced every meter, permafrost collapse at the thaw front, and the approximate height of the water table in the bog. Green and gray layers comprise the active layer's primarily organic content and transition between organic and mineral horizons respectively. The light blue region shows permafrost ice.

2.3 Porewater Analysis

Total iron ($Fe_{(s)}$ and $Fe^{2+}_{(aq)}$) was measured by acidifying the porewater using 2M HCl followed by a ferrozine assay and colorimetric analysis measured at 560 nm with further detail described in previous work[36,74,76]. To measure DOC concentration, porewater was acidified with 2M HCl to remove inorganic carbon and analyzed by total carbon analyzer (High TOC II, Elementar, Germany). Dissolved organic nitrogen was determined by measuring total nitrogen alongside DOC and subtracting inorganic nitrogen[36,74,77]. $Ammonium_{(aq)}$, $nitrate_{(aq)}$, and $nitrite_{(aq)}$ concentrations were measured with a flow injection analyzer (Seal Analytical, Germany) equipped with a dialysis membrane to remove $Fe^{2+}_{(aq)}$ and prevent side reactions during

measurement. Due to sample volume limitations, nitrite and nitrate were analyzed in transects 1-4, DOC, $Fe^{2+}_{(aq)}$, and ammonium were analyzed in transects 1-5, and DON was analyzed in transect 5 (Table S2)

The mean analyte concentrations of the palsa samples (n=13), thaw front samples (n=5), and bog samples (n=11) were statistically compared to each other. First, the normality of the analyte concentrations from each thaw state were evaluated using the Shapiro-Wilk test. If the populations of each thaw state were normally distributed, then the average analyte concentration of two thaw states were compared with the Welch's t-test. If one of the populations was not normally distributed, then the Mann-Whitney U test was used.

2.4 Nuclear Magnetic Resonance (NMR) Spectroscopy

Nuclear magnetic resonance spectroscopy was utilized to determine the distribution of ~~structural~~ groups in total organic matter. Although 1H NMR does not select for nitrogen containing compounds, it does provide an overview of the DOM chemistry that can be used for sample comparison with the corroborating molecular class abundances identified using FT-ICR MS53,78,79. Filtered (0.22 μm) porewater was lyophilized and yielded approximately 1 mg of freeze-dried material. The freeze-dried extracts were further dried over phosphorus pentoxide in a vacuum desiccator and then re-dissolved into 60 μL of deuterium oxide (D_2O, 99.9% D) and 5 μL of sodium deuteroxide (NaOD, 99.5% D, 30 wt. % in D_2O). These dissolved samples were then centrifuged (10,000 x g and 4 °C), and the supernatants were transferred into 1.7 mm NMR tubes (Norell). All samples were analyzed using a Bruker BioSpin Avance III 500 MHz NMR spectrometer (Karlsruhe, Germany) equipped with a ^{1}H-^{15}N-^{13}C TXI 1.7 mm microprobe with an actively shielded Z gradient. One-dimensional ^{1}H NMR spectra were collected using water suppression via pre-saturation utilizing relaxation gradients and, the spectra were collected using

2048 scans per sample with a recycle delay of 2 s and 32K time domain points. The spectra were processed using a zero-filling factor of two and were apodized by multiplication with an exponential decay corresponding to 0.3 Hz line broadening[80]. NMR spectra were integrated based on typical DOM chemical shift regions associated with materials derived from linear terpenoids (MDLT; 0.6−1.6 ppm), carboxyl-rich alicyclic molecules (CRAM; 1.6−3.2 ppm), carbohydrates and peptides (3.2−4.5 ppm), and aromatic and phenolic components (6.5−8.4 ppm)[80,81]. The integration was performed using Analysis of Mixtures (AMIX; v. 3.9.15) software and expressed as percentage of the sum of the specific integration regions such that our comparison considered relative changes between these specific regions. It is important to note that some resonances (~1.9 ppm: likely from acetic acid; ~3.3 ppm likely from methanol; and ~8.4 ppm: likely from formic acid)[80–82] do not align with the typical regions of DOM but were not excluded during integration due to overlapping resonances in some samples.

2.5 METATRANSCRIPTOME SEQUENCING

Metatranscriptome sequencing was utilized to evaluate upregulation or downregulation of gene expression of enzymes in nitrogen cycling pathways, including but not limited to ammonification, nitrification, and denitrification. To survey microbial gene expression in Stordalen Mire, we used the genomes recovered in Woodcroft and Singleton et al.[83]. The 1,529 metagenome-assembled genomes (MAGs) were dereplicated at 99% identity with dRep (v2.6.2)[84] into 649 MAGs. The 99% dereplicated MAGs were annotated with DRAM (v1.4.4)[85]. To assess the gene expression of these MAGs in the palsa and bog, we used metatranscriptome samples previously published in Ellenbogen et al.[86]. The metatranscriptomes derive from cores taken in July 2016 from defined palsa and bog environmental conditions at an adjacent thaw transition, and the thaw states have been well defined by previous work[15,43,69,73,87]. We recognize

the possibility of variability between the porewater transects and metatranscriptome peat cores. However, site restrictions prevented the collection of peat cores at the porewater thaw transition location. Furthermore, both sites represent well-defined thaw transitions in Stordalen Mire[15]. Additionally, the conditions of the palsa and bog are consistent at an interannual scale although there are year to year differences, the observed differences between them are consistent[47,57]. Therefore, we are confident that the findings from the metatranscriptome analysis can be applied to the comparable porewater sampled from the palsa to bog transition. Soil cores were sectioned in the field at three depths: "surface" (1-4 cm), "middle" (10-14 cm), and "deep" (20-24 cm) and immediately put into LifeGuard (Qiagen). Metatranscriptome analysis used the deep section of the cores, as the 20-24 cm depth is most comparable to the 30 cm depth of porewater samples. Core sections were stored at -20˚C until analysis. DNA and RNA were co-extracted from 5-10 g peat using the Mobio PowerMax Soil DNA/RNA isolation kit (cat# 12966-10). Details on metatranscriptome library preparation, raw reads, gene annotation, and pathway expression comparison are provided in more detail in the SI (SI Section 1 and SI Data 1).

2.6 21 TESLA FT-ICR MS ANALYSIS

Fourier transform ion cyclotron resonance mass spectrometry assigns nitrogen-containing molecular formulae within the dissolved organic matter pool. The ability to segregate formulae according to elemental classes, like those that include nitrogen, enables us to differentiate compositional variations within each class that might otherwise be indiscernible within the total FT-ICR MS identified OM pool. Porewater DOM was analyzed with 21 Tesla FT-ICR MS to identify and monitor DON compositional changes. Porewater was prepared for FT-ICR MS analysis by solid phase extraction (SPE) under N_2 atmosphere (glove bag) following a modified procedure based on Dittmar *et al.*[88] and Li *et al.*[89], which has been described in previous

work[36,60,74]. The samples were processed under N_2 using degassed solvents to maintain the oxygen concentrations present *in situ*, and Hypersep Retain CX cartridges were selected for SPE due to reported improved nitrogen containing formulae and peptide-like formulae recovery compared to commonly used PPL cartidges[60]. Hypersep Retain CX SPE cartridges (part# 60107-305, Thermo Fisher Scientific, Waltham, MA) were rinsed with 5 mL of HPLC grade methanol (Sigma-Aldrich, Rehovot, Israel) followed by 5 mL of 0.01 M HCl. Each DOM sample was acidified to pH ~2.5 and 0.5 mg C was loaded onto the SPE columns. After sample loading, the SPE cartridges were rinsed with 5 mL of 0.01 M HCl followed by drying with N_2 for 3–5 min. Finally, the samples were eluted with 1 mL of HPLC grade methanol and stored in airtight amber sample vials wrapped in aluminum foil at 4 °C. No additional dilution of the samples was performed prior to FT-ICR MS analysis.

Samples were analyzed as negative ions in a custom-built hybrid linear ion trap FT-ICR mass spectrometer equipped with a 21 T superconducting solenoid magnet[52,90]. Ionization conditions, instrument settings, and molecular formula assignment procedure are described in more detail in the SI (SI Section 2). The molecular formulae were assigned molecular class (condensed aromatic, aromatics/polyphenols, highly unsaturated, unsaturated aliphatic, saturated fatty acids, peptides, and sugars) according to the formula boundaries outlined by Poulin *et al*.[91] (Table S3). For all mass spectra presented herein, between 10,024-11,357 mass peaks were assigned elemental compositions with root-mean-square mass measurement accuracy of 55 to 70 ppb with achieved resolving power of 3,400,000 at m/z 200. The m/z range including all samples was 165.0193 to 1155.211. Data processing post-formula assignment was performed with RStudio utilizing R software (V4.1.2). Table S4 shows the sample name, RMS error, and number of assignments for all spectra discussed herein. PetroOrg files, calibrated peak lists, and peak

assignments are publicly available via the Open Science Framework via DOI 10.17605/OSF.IO/TZKWN.

Table S4. FT-ICR MS total identified peaks
Total identified peaks from FT-ICR MS analysis for samples discussed herein. Total identified peaks include all ^{13}C peaks that were not included in sample comparison as they mirror the higher relative abundance ^{12}C peaks.

Sample Name	Total Identified Peaks	RMS error (ppm)
Palsa 2	18908	0.062219
Thaw Front	17370	0.055763
Bog 2	18782	0.063072
Bog 3	16314	0.070826

3. RESULTS

3.1 THE THAW FRONT FEATURED THE HIGHEST $Fe^{2+}_{(aq)}$, DOC, AMMONIUM, AND DON CONCENTRATIONS

Transects 1-4 exhibited a pulse of Fe^{2+}(aq) and DOC at the saturated thaw front (Figure 3A). Transect 5, which was collected to analyze DON concentrations, showed changes in concentration of Fe^{2+}(aq), ammonium, and DOC consistent with the previously sampled four transects, despite the 3-year gap in sampling time. The palsa, generally dry and oxic[48], had the lowest levels of all measured species (Figure 3 and Table S2). The only exception to this was nitrate, which had no statistically significant change between any of the thaw states (Welch's t-test, $P>0.05$) ranging between 1.27 and 0.07 mg/L. The thaw front had the highest concentrations of all analyzed species, with significantly higher levels (Welch's t-test, $P<0.05$) of DOC and ammonium compared to the palsa and bog. The bog, which the palsa drains into via the thaw front[67], contained concentrations of ammonium that were not significantly different (Welch's t-test, $P>0.05$) from the palsa. Similarly to ammonium, DON concentration decreased by 73%

immediately after the thaw front (Figure 3B). Differences in DON concentrations could not be evaluated with Welch's t-test or the Mann-Whitney U test because DON was only measured along transect 5. Nevertheless, the thaw front DON concentration was substantially higher compared to other sampling points (10-fold increase compared to the average palsa concentration). Since the trends in DOC and Fe^{2+}(aq) are consistent across all transects when aligned to the pulse of Fe^{2+}(aq) (i.e., the thaw front), we can conclude that the geochemical profiles of the transects are comparable. Therefore, we propose that results from more detailed molecular analyses (FT-ICR MS and NMR) that were performed on select transects can be applied to the general thaw transition at Stordalen Mire.

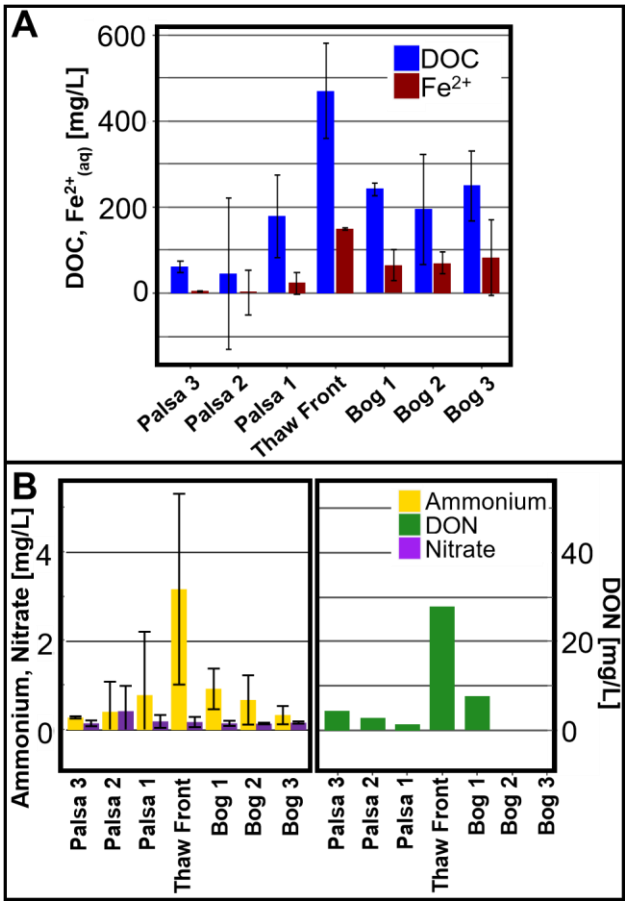


Figure 3. Porewater composition across the thaw transition. (A) DOC and $Fe^{2+}_{(aq)}$ concentrations across the thaw transition. (B) Ammonium, nitrate, and DON concentrations across the thaw transition. Reported values and error bars represent the average and standard deviation of palsa-

to-bog sampling transects. Error bars are not present for DON as only one transect (T5) was analyzed for DON, so n=1 for all DON values. There were no DON values for Bog 2 and Bog 3 since there were no Bog 2 and Bog 3 sampling points for transect 5 due to the position of the thaw front. N-values for the remaining sample positions varied due to different thaw front positions in the transects (see Table S2 for full transect data).

3.2 Thaw front DON has lower H/C and lower relative abundance of aliphatic formula compared to palsa and bog:

Van Krevelen diagram analysis showed clear compositional differences between the palsa and thaw front sites (Figure 4A) The palsa (P2), thaw front, and bog (B3) featured 11272, 10852, and 10024 total monoisotopic formulae, respectively. Higher relative abundance formulae in the thaw front had lower H/C and higher O/C ratios and were more aromatic-like than in the palsa (Figure 4A, 4B). Specifically, the percentage of aromatic/polyphenols-like and condensed aromatic-like formulae increased from 23% in the palsa to 27% in the thaw front followed by a decrease to 23% in the bog. Higher aromatic-like content in the thaw front could be attributed to microbial degradation of polyphenolic compounds[36]. Despite a 7.5-fold increase in DOC and a tenfold increase in the DON concentration at the thaw front, the number of identified DOC and DON formulae underwent minimal change. Assigned DON formulae represented approximately 27% of the total formulae across all conditions. As observed in the total formulae, when only nitrogen-containing formulae (referred from here as CHNO formulae) were compared, the highly abundant CHNO formulae in the thaw front also featured higher O/C and lower H/C ratios and were more aromatic-like than the palsa (Figure 4C).

To further examine compositional differences in the assigned CHNO formulae, we isolated and compared formulae that were unique to each thaw state, hereafter referred to as the "unique pool". When comparing the unique CHNO pools between the palsa and thaw front, there is a clear compositional difference in CHNO formulae, which is distinct from the unique CHO pools.

(Figure 4B, 4C). Comparable to the total formulae pools, the unique CHNO formulae shifted from more aliphatic-like in the palsa to more aromatic-like in the thaw front. In the palsa, most unique CHNO formulae with high relative abundance feature H/C values around 1.5, with only a few formulae in the region of highest unique CHO. Conversely, the thaw front lacks unique formulae with H/C values near 1.5 but features a cluster of formulae in the more aromatic-like region, consistent with its unique CHO pool. The unique CHNO pools reveal a distinct pattern: many formulae found in the palsa are missing from the thaw front, and vice versa (Figure 4C).

To determine which formulae contributed to the differences in the unique CHNO pools, we compared unique CHNO formulae by assigned molecular class (Table S3). The thaw front contained more aromatic-like and less peptide-like unique CHNO formulae when compared to either palsa or bog (Figure 4D, Table S5). Notably, when comparing the palsa and thaw front, the proportion of unique CHNO formulae that were peptide-like decreased from 17% to 0%. Simultaneously, the aromatic/polyphenols-like and condensed aromatic-like formulae increased from 5% to 50%. Although the thaw front drains into the bog, the high proportions of condensed aromatic/polyphenol-like and low proportion of peptide-like unique CHNO formulae of the thaw front were not present in the bog. From the thaw front to the bog, the unique CHNO condensed aromatic and aromatic/polyphenol-like fraction decreased from 43% to 6%. Unique CHNO peptide-like formula increased from the thaw front to the bog, increasing from 1% in the thaw front to 8% in the bog, but did not return to palsa levels.

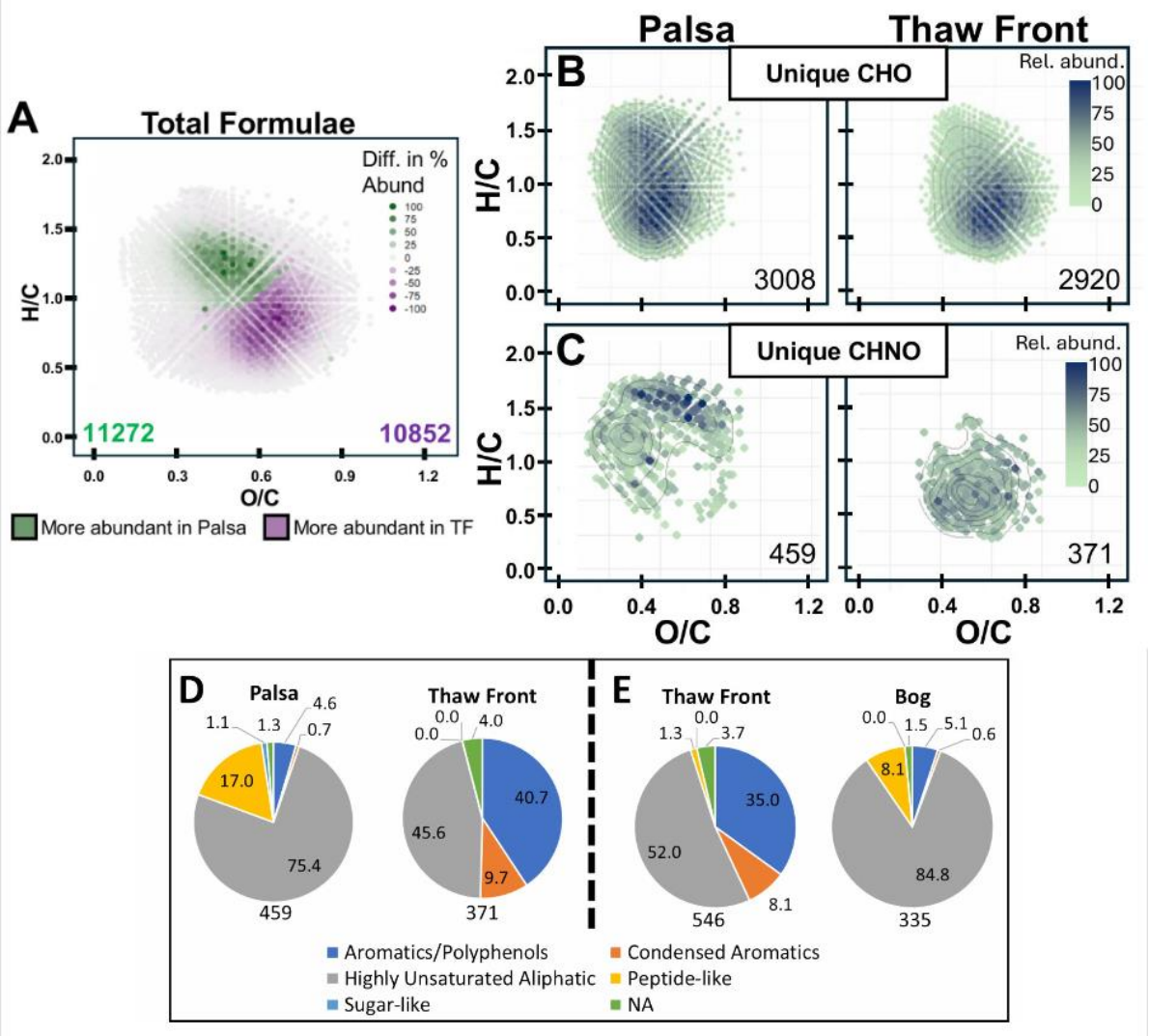


Figure 4. Molecular formula characterization performed by FT-ICR MS of porewater samples. (A) Van Krevelen diagrams displaying comparisons between palsa (P2) and thaw front samples. Distribution of the total formula showing regions of higher abundance in palsa (green) and higher abundance in thaw front (purple). Color is determined by the difference in the relative abundance for a given formula between the palsa and the thaw front, with positive values having higher relative abundance in the palsa and negative values having higher relative abundance in the thaw front. The total count of identified formulae is inserted and color matched to palsa and thaw front. (B) Unique CHO formulae for palsa and thaw front and (C) unique CHNO formulae in palsa and thaw front colored by percent abundance of the formula. Inserted numbers in (B) and (C) are the number of formulae plotted in each plot. (D and E) Formula count of unique CHNO molecules shown as percentages of molecular class assigned by FT-ICR MS between (D) palsa (P2) and thaw front and (E) thaw front and bog (B3). The numbers of identified unique CHNO formulae for each thaw state are shown underneath each chart. Molecular class is defined by the heteroatom content of each molecular formula. Molecular class was defined according to Poulin et al. (2017)[91]. Unsaturated aliphatic and saturated fatty acid classes are not shown as there are no unique CHNO formulae in those classes.

Table S5. FT-ICR MS assigned total and unique monoisotopic formulae

Counts of FT-ICR MS assigned total and unique monoisotopic formula for samples P2, thaw front, and B3 in transect one.

	SAMPLE	TOTAL FORMULAE	AROMATIC/ POLYPHENOL-LIKE	CONDENSED AROMATIC-LIKE	HIGHLY UNSATURATED ALIPHATIC-LIKE	PEPTIDE-LIKE	SUGAR-LIKE	UNSATURATED ALIPHATIC-LIKE	NA	SATURATED FATTY ACID-LIKE
FORMULAE COUNT (% OF TOTAL)	Palsa (P2)	11272	1972(17.5%)	606(5.4%)	7904(70.1%)	110(1.0%)	28(0.2%)	429(3.8%)	223(2.0%)	0(0%)
	Thaw front	10852	2192(20.2%)	694(6.4%)	7239(66.7%)	32(0.3%)	44(0.4%)	411(3.8%)	240(2.2%)	0(0%)
	Bog (B3)	10024	1833(18.3%)	491(4.9%)	6997(69.8%)	52(0.5%)	27(0.3%)	414(4.1%)	210(2.1%)	0(0%)
CHNO FORMULAE COUNT (% OF TOTAL)	Palsa (P2)	3008	551(18.3%)	69(2.3%)	2188(72.7%)	110(3.7%)	5(0.2%)	0(0%)	85(2.8%)	0(0%)
	Thaw front	2920	681(23.3%)	102(3.5%)	2011(68.9%)	32(1.1%)	0(0%)	0(0%)	94(3.2%)	0(0%)
	Bog (B3)	2709	507(18.7%)	60(2.2%)	2011(74.2%)	52(1.9%)	0(0%)	0(0%)	79(2.9%)	0(0%)
PALSA/THAW FRONT COMPARISON UNIQUE CHNO (% OF GROUP)	Palsa (P2)	459	21(4.6%)	3(0.7%)	346(75.4%)	78(17.0%)	5(1.1%)	0(0%)	6(1.3%)	0(0%)
	Thaw front	371	151(40.7%)	36(9.7%)	169(45.6%)	0(0%)	0(0%)	0(0%)	15(4.0%)	0(0%)
THAW FRONT/BOG COMPARISON UNIQUE CHNO (% OF GROUP)	Thaw front	546	191(35.0%)	44(8.1%)	284(52.0%)	7(1.3%)	0(0%)	0(0%)	20(3.7%)	0(0%)
	Bog (B3)	335	17(5.1%)	2(0.6%)	284(84.8%)	27(8.1%)	0(0%)	0(0%)	5(1.5%)	0(0%)

3.3 Carbohydrate- and peptide-like content was lowest at the thaw front:

The FT-ICR MS results were complemented with solution-state ^{1}H NMR spectroscopy data to evaluate the molecular structure of porewater DOM throughout transect 4 (Figure 5A, spectra shown in Figure S1). The percentage of carbohydrates and peptides in the thaw front were approximately half of any other sample point (13% in thaw front and 21-35% in all other samples). The proportional decrease in carbohydrates and peptides correlated with a proportional increase in carboxylic rich alicyclic molecules (CRAM), considered an important product of microbial degradation[92], which increased from 41% in P1 to 60% the thaw front. Like the FT-ICR MS results, the chemical composition in the thaw front did not persist into the bog. Notably the proportional CRAM content of the bog fell within the range of CRAM percentages in the palsa, and carbohydrate and peptide proportional content increased from thaw front to bog, although to only ~75% of the palsa average. The remaining fractions, materials derived from linear terpenoids (MDLT) and aromatic/phenolic, featured nominal proportional changes across the thaw transition.

3.4 Microbial nitrogen cycling gene expression increased in the bog:

To uncover the potential microbial processes contributing to nitrogen cycling in the palsa and bog, we mapped field metatranscriptome reads to a database of Stordalen Mire microbial metagenome assembled genomes[83]. We then aggregated genes into nitrogen cycling pathways and looked for altered gene expression levels between the palsa and bog (Figure 5B, Supplementary Data 1). In the anoxic bog, several gene pathways were significantly upregulated (LIMMA, $p<0.05$) including nitrogenase (*nifDKH*), dissimilatory nitrate reduction to ammonia (DNRA, *nrfA* and *nirBD*), nitrogen transporters, denitrification enzymes (*narGH, nirK, norB*), and peptidase pathways. Additionally, expression of genes in nitrogen assimilation (*nirA, nasA*)

were upregulated in the bog but not to a significant extent (LIMMA, p>0.05). Expression of nitrification (*hao*) was the only pathway to have lower expression in the bog than the palsa, but the decrease was not significant (LIMMA, p>0.05).

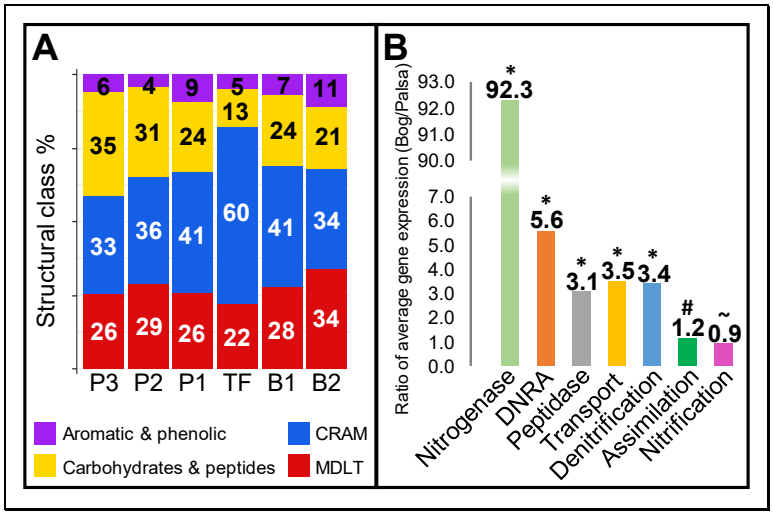


Figure 5. (A) Relative abundance of general structural classes within porewater samples along the thaw transition determined by ^{1}H NMR. Sample position shown on the x-axis as P1-3 (palsa 1-3), TF (thaw front), B1-2 (bog 1-2). CRAM = carboxyl-rich alicyclic molecules, MDLT = materials derived from linear terpenoids. (B) Fold change in average metatranscriptome expression of microbial nitrogen cycling pathways between palsa (n=3) and bog (n=3) cores. The expression of pathways marked with (*) significantly (LIMMA, p<0.05) increased, (#) increased but not significantly (LIMMA, p>0.05), and (~) decreased but not significantly (LIMMA p>0.05) when comparing gene expression in the bog to that in the palsa. Positive y-axis values indicate that the gene featured average greater expression in the bog than the palsa. See SI Data 1 for pathway information and statistics.

4. Discussion

4.1 Thaw-induced conversion of DON to inorganic nitrogen

Several modeling studies suggest that the main factor controlling the production of inorganic nitrogen in thawing permafrost is the release of organic nitrogen[30,32,93]. This release occurs due to the deepening of the active layer and the subsequent transformation of DON into ammonium[32,94,95]. When there is limited ammonium, which is the case for the active layer in

most arctic soils, plant and microbial immobilization of nitrogen is dominant[29,30,39,93,96]. However, two *in situ* studies of permafrost thaw collapse linked DON release to elevated nitrate concentrations due to the improved drainage and oxic conditions after thaw[39,41]. Unlike these studies, the soil during and after the thaw front at our site remained saturated and at or below water level. At the thaw front we observed a large pulse of DON and significantly higher levels of ammonium compared to palsa and bog (Figure 3). Despite the substantial ammonium within the thaw front, there was no significant change in nitrate concentrations across the thaw transition. In the well drained and oxic palsa, the limited nitrate is likely due to the low concentrations of DON and the lack of ammonium, which acts as substrate for nitrification. Conversely, in the thaw front and bog, the high saturation and anoxic conditions were likely the main limiters of nitrification[29,35,39,67]. Therefore, it is likely that the DON released during thaw is primarily mineralized to ammonium and any further nitrification is limited by the anoxic conditions. Our findings align with reported seasonal saturation of thaw slump floors, which observed increased ammonium concentrations and DON without accumulating nitrate[29,39]. We propose that, confirming our first hypothesis, the high levels of ammonium within the thaw front are likely the result of rapid ammonification of DON.

4.2 Shifting composition of DON at the thaw front

Within the thaw front, both NMR and FT-ICR MS results indicated the lowest proportions of peptide-like and carbohydrate molecules throughout the thaw transition (Figures 3 and 4), despite a tenfold increase in DON concentration (Figure 3). FT-ICR MS analysis of nitrogen-containing formulae showed the highest proportions of aromatics/polyphenols and condensed aromatics within the thaw front. NMR analysis of the total DOM pool revealed little change in aromatic and phenolic structures along the thaw transition, but it showed a substantial

increase in CRAM structures at the thaw front. This increase likely reflects the overall oxidation and condensation of the DOM pool observed in previous studies[36], a pattern not captured in the nitrogen-specific FT-ICR MS analysis. The shift toward more condensed, less labile structures in both analyses, coupled with the decrease in peptide-like compounds, suggests active degradation of the nitrogen pool at the thaw front[92,97]. This interpretation aligns with soil amendment and incubation studies showing preferential mineralization of peptides and amino acids over larger nitrogen-containing organic molecules under conditions with abundant labile organic carbon (e.g., glucose)[33,38,98]. Furthermore, arctic plant communities can preferentially uptake amino acids over nitrate, further increasing demand for labile DON[96,99–101]. High nutrient demand in thaw fronts is expected as they have been identified as points of elevated microbial activity and active organic matter decomposition[29,36,40]. While we did not directly measure DON transformation, multiple lines of evidence suggest selective degradation is occurring. We observed a simultaneous decrease in labile peptide-like compounds and enrichment of less labile aromatic or cyclic molecules. This selective transformation of the DON pool parallels patterns documented in anoxic marine sediments, where peptides undergo selective deamination and microorganisms preferentially utilize free amino acids for cellular synthesis[102]. The increased microbial activity previously reported at the thaw front[36,65] would create higher demand for cellular building blocks, driving the observed depletion of available peptides. These findings confirm our second hypothesis and align with previous studies of permafrost nutrient cycling and DOC transformation[32,38,41,47,49,103].

4.3 Organic nitrogen consumption at the thaw front leaves the bog nitrogen-limited

Metatranscriptomic data revealed an upregulation of gene expression for enzymes in nitrogen scavenging pathways in the bog (Figure 5B). These pathways include nitrogen fixation,

DNRA (a process often overlooked in permafrost studies)[32], and mineralization via peptidases. Furthermore, in the water-saturated and anoxic bog, denitrification gene expression was significantly higher than in the palsa. Notably, we saw expression of genes encoding nitrate reductase (*narGH),* nitrite reductase (*nirK*)*,* and nitric oxide reductase (*norB*) but not nitrous oxide reductase (*nosZ*), hinting that incomplete denitrification occurred in the bog. This could represent a pathway of nitrogen loss in the bog and N_2O generation, further contributing to nitrogen limitation[32]. However, saturated wetlands in the arctic have been found to have negligible N_2O emissions, due to the high water table acting as a diffusion barrier[40]. Anaerobic ammonia oxidation and complete ammonia oxidation could be additional pathways of nitrogen loss from the system, yet we did not detect gene expression for either of these processes.

At the thaw front, we observed a substantial increase in total DON concurrent with a reduction in the proportion of peptide-like molecular content. The adjacent bog showed a greater proportional decrease in DON concentration compared to DOC concentration (Figure 3), indicating either preferential removal of nitrogen-containing molecules from solution or subsequent DOC release. This decrease in the proportion of nitrogen-containing organic matter in the bog corresponds with our metatranscriptomic results showing increased expression of nitrogen scavenging genes, suggesting a microbial response to reduced bioavailable DON.

After the pulse of ammonium at the thaw front, the anoxic conditions in the bog would limit nitrification, an aerobic process, which explains the persistence of low nitrate concentrations from thaw front to bog (Figure 3). Furthermore, this limited nitrification would in turn limit denitrification by restricting production of nitrate[29,39,40]. The observed decrease in DON, ammonium, and nitrate concentrations between the thaw front and bog suggests that the most labile and bioavailable nitrogen was likely transformed into ammonium and subsequently

immobilized through biomass incorporation or removed though denitrification to N_2O[8,35]. However, contrary to our third hypothesis, denitrification increased in the bog despite the restricted nitrate substrate availability. This unexpected finding may reflect either increased denitrification activity or, more likely, upregulation of denitrification pathways as a mediating response to limited nitrate availability - similar to the upregulation of nitrogen scavenging pathways observed more broadly in the bog gene expression (Figure 5). Although we did not collect metatranscriptome samples from the thaw front, our previous study documented increased activity of iron reducers and methanogens within it[36]. Therefore, we propose the rapid appearance of high ammonium concentrations in this zone suggests similarly elevated nitrogen cycling activity, particularly ammonification. Although porewater and metatranscriptome samples were collected from different locations within Stordalen Mire, the consistent thaw conditions and ecology across sites[15] allow these complementary datasets to inform our understanding of nitrogen transformations across the thaw transition.

4.4 Nitrogen availability is influenced by peptide content during permafrost thaw.

Our data reveal an inverse relationship between peptide-like DON and ammonium concentrations along the permafrost thaw transition. At the thaw front, we observe depletion of labile peptide-like DON alongside enrichment of aromatic DON and elevated ammonium levels. In contrast, both the bog and palsa show higher peptide content with limited ammonium (Figures 3, 4, 5). The consistent inverse relationship between peptide content and ammonium levels across sites suggests that peptide availability may act as a control point for nitrogen cycling, with peptide degradation regulating the ammonium pool available for subsequent nitrogen transformations. Although we did not directly measure nitrogen transformations at the thaw front, our metatranscriptomic analysis of the bog revealed increased expression of bioavailable

nitrogen scavenging genes. These metatranscriptomic results support the role of microbial communities in driving nitrogen transformations at the thaw front. The relationship between nitrogen and carbon cycling in thawing permafrost remains poorly characterized, though nitrogen availability is expected to influence carbon mineralization and permafrost carbon feedback through both cellular synthesis and DOC mineralization priming[11,40,102]. Based on our findings, we propose that components of Ramm's model[32], particularly the controls on ammonification rates, can be applied to thaw fronts and bogs, with our suggested extensions illustrated in Figure 6. Future studies directly examining metatranscriptomic activity at the thaw front will be crucial for fully understanding the fate of nitrogen-containing organic matter during permafrost thaw.

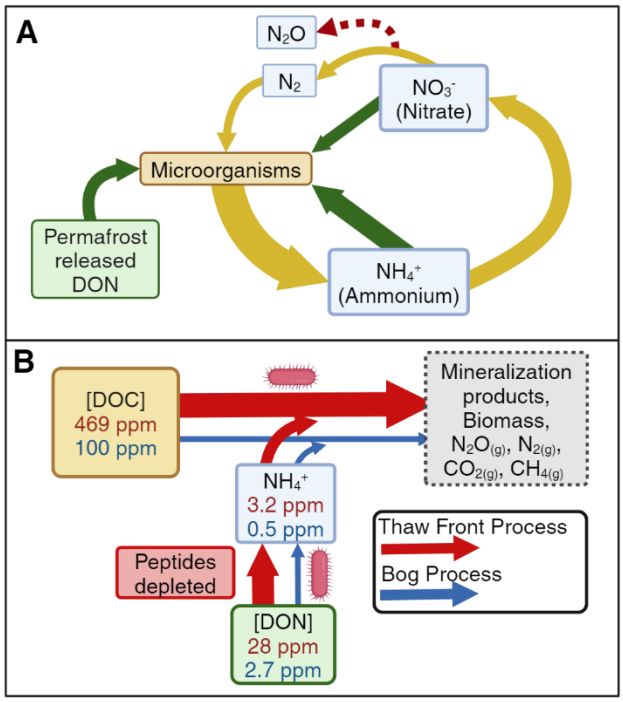


Figure 6. (A) Modified conceptual model for nitrogen cycling within permafrost soils based on Ramm et al. (2022)[32]. Green arrows show uptake processes, yellow arrows show mineral N turnover processes, and red arrows show caseous losses. (B) Adapted model to show palsa vs. thaw front organic matter and nutrient transformation supported by empirical data reported herein. Data shows that high concentrations of organic nitrogen within the thaw front correlate to high ammonium concentrations and reduced peptide content. The production of ammonia can prime the microbial degradation of highly abundant DOC, potentially elevating the production of

GHGs and biomass. Arrows are sized to indicate the expected relative amount of the transformation process based on previous GHG emission and nitrogen priming findings[8,31,94]. Figure 6(A) is based on Figure 1 in "A review of the importance of mineral nitrogen..." by Ramm et al.[32] and is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

5. CONCLUSIONS

Our results provide some of the first detailed characterization of organic nitrogen dynamics within a permafrost thaw transition, revealing that the composition of organic nitrogen released during thaw significantly impacts nutrient cycling and creates biogeochemical hotspots within the thaw front environment. Notably, permafrost collapse and soil saturation resulted in high ammonium concentrations localized to the thaw front. Among nitrogen-containing molecules, the thaw front exhibited lower fractions of peptide and carbohydrate molecules and a higher fraction of aromatic molecules, aligning with previously reported selective degradation patterns of the total organic matter pool[32,36,38]. Nitrogen transformation during and after thaw appeared limited to DON mineralization, with no significant changes in nitrate concentration. However, the upregulation of denitrification under anoxic conditions in the bog suggests microbial metabolism could either contribute to N_2O emissions via incomplete denitrification (nitrate to N_2O) or mitigate N_2O release through complete denitrification (nitrate to N_2)[19,39,104]. The absence of genes encoding N_2O reductase in the bog suggests thaw fronts may serve as previously underrepresented N_2O sources in climate models. Future studies combining metatranscriptome data from the thaw front with N_2O flux measurements would further elucidate nitrogen dynamics and improve our understanding of nitrogen mass balance in these systems. Our findings emphasize that future permafrost peatland studies and GHG models should incorporate thaw transitions due to their rapid and highly localized nitrogen dynamics.

6. Author contributions

The original hypothesis was conceptualized by M.N.L. and T.B. and M.N.L., M.S.P., A.K., C.B., and T.B. designed the project. The field work was conducted by M.N.L. and M.S.P. Data gathered for the main text was collected by M.N.L., M.S.P., B.B.M., N.S., and A.M.M. All FT-ICR MS data was processed and interpreted by M.N.L. NMR data was processed by N.S. and M.J.S. Metatranscriptomic data was processed by B.B.M. and K.C.W. The manuscript was written and all figures were made by M.N.L. J.P.V. contributed to editing the manuscript and designing the figures. All authors contributed to the final preparation of the manuscript and have approved the final version.

7. Copyright

8. REFERENCES

1 Logan MN, Patzner MS, VanderRoest JP, McGivern BB, Srikanthan N, Simpson MJ, *et al.*, Role of permafrost thaw transitions in biogeochemical nitrogen cycling, *Soil Environ Heal* **3**:100148, The Author(s) (2025).

2 Hugelius G, Strauss J, Zubrzycki S, Harden JW, Schuur EAG, Ping CL, *et al.*, Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges and identified data gaps, *Biogeosciences* **11**:6573–6593 (2014).

3 Mishra U, Drewniak B, Jastrow JD, Matamala RM, and Vitharana UWA, Spatial representation of organic carbon and active-layer thickness of high latitude soils in CMIP5 earth system models, *Geoderma* **300**:55–63, The Author(s) (2017).

4 Schuur EAG, Abbott BW, Bowden WB, Brovkin V, Camill P, Canadell JG, *et al.*, Expert assessment of vulnerability of permafrost carbon to climate change, *Clim Change* **119**:359–374 (2013).

5 Hugelius G, Loisel J, Chadburn S, Jackson RB, Jones M, MacDonald G, *et al.*, Large stocks of peatland carbon and nitrogen are vulnerable to permafrost thaw, *Proc Natl Acad Sci U S A* **117**:20438–20446 (2020).

6 Zhang T, Barry RG, Knowles K, Heginbottom JA, and Brown J, Statistics and characteristics of permafros and ground-ice distribution in the Northen Hemisphere, *Polar Geogr* **23**:132–154 (2008).

7 Obu J, How Much of the Earth's Surface is Underlain by Permafrost?, *J Geophys Res Earth Surf* **126**:1–5 (2021).

8 Voigt C, Marushchak ME, Lamprecht RE, Jackowicz-Korczyński M, Lindgren A, Mastepanov M, *et al.*, Increased nitrous oxide emissions from Arctic peatlands after permafrost thaw, *Proc Natl Acad Sci U S A* **114**:6238–6243 (2017).

9 Batjes NH, Total carbon and nitrogen in the soils of the world., *Eur J Soil Sci* **47**:151–163 (1996).

10 Harden JW, Koven CD, Ping CL, Hugelius G, David McGuire A, Camill P, *et al.*, Field information links permafrost carbon to physical vulnerabilities of thawing, *Geophys Res Lett* **39**:1–6 (2012).

11 Schuur EAG, Abbott BW, Commane R, Ernakovich J, Euskirchen E, Hugelius G, *et al.*, Permafrost and Climate Change: Carbon Cycle Feedbacks From the Warming Arctic, *Annu Rev Environ Resour* **47**:343–371 (2022).

12 Abbott BW, Brown M, Carey JC, Ernakovich J, Frederick JM, Guo L, *et al.*, We Must Stop Fossil Fuel Emissions to Protect Permafrost Ecosystems, *Front Environ Sci* **10**:1–13 (2022).

13 Tromsø, Norway: Arctic Monitoring and Assessment Programme (AMAP).

14 Wilson RM, Hopple AM, Tfaily MM, Sebestyen SD, Schadt CW, Pfeifer-Meister L, *et al.*, Stability of peatland carbon to rising temperatures., *Nat Commun* **7**:13723, Nature Publishing Group (2016).

15 Holmes ME, Crill PM, Burnett WC, McCalley CK, Wilson RM, Frolking S, *et al.*, Carbon Accumulation, Flux, and Fate in Stordalen Mire, a Permafrost Peatland in Transition, *Global Biogeochem Cycles* **36** (2022).

16 Harris SA, French HM, Heginbottom JA, Johnston GH, Ladanyi B, Sego DC, *et al.*, Glossary of permafrost and related ground-ice terms Harris, *Tech Memo (National Res Counc Canada Assoc Comm Geotech Res* (1988).

17 Jorgenson MT, Racine CH, Walters JC, and Osterkamp TE, Permafrost degradation and

ecological changes associated with a warming climate in central Alaska, *Clim Change* **48**:551–579 (2001).
18 Schädel C, Schuur EAG, Bracho R, Elberling B, Knoblauch C, Lee H, *et al.*, Circumpolar assessment of permafrost C quality and its vulnerability over time using long-term incubation data, *Glob Chang Biol* (2014).
19 McGuire; AD, Lawrence; DM, Koven; C, Clein; JS, Burke; E, Chen; G, *et al.*, Dependence of the evolution of carbon dynamics in the northern permafrost region on the trajectory of climate change., *Proc Natl Acad Sci 2* **115**:3882–3887 (2018).
20 Schuur EAG, McGuire AD, Schädel C, Grosse G, Harden JW, Hayes DJ, *et al.*, Climate change and the permafrost carbon feedback, *Nature* **520**:171–179 (2015).
21 Schaefer K, Lantuit H, Romanovsky VE, Schuur EAG, and Witt R, The impact of the permafrost carbon feedback on global climate., *Environ Res Lett* **9** (2014).
22 Cooper MDA, Estop-Aragonés C, Fisher JP, Thierry A, Garnett MH, Charman DJ, *et al.*, Limited contribution of permafrost carbon to methane release from thawing peatlands, *Nat Clim Chang* **7**:507–511 (2017).
23 Knoblauch C, Beer C, Liebner S, Grigoriev MN, and Pfeiffer E, Methane production as key to the greenhouse gas budget of thawing permafrost, *Nat Clim Chang* **8**:309–312 (2018).
24 Butterbach-Bahl K, Baggs EM, Dannenmann M, Kiese R, and Zechmeister-Boltenstern S, Nitrous oxide emissions from soils: How well do we understand the processes and their controls?, *Philos Trans R Soc B Biol Sci* **368** (2013).
25 Marushchak ME, Kerttula J, Diáková K, Faguet A, Gil J, Grosse G, *et al.*, Thawing Yedoma permafrost is a neglected nitrous oxide source, *Nat Commun* **12**, Springer US (2021).
26 Harms TK and Jones JB, Thaw depth determines reaction and transport of inorganic nitrogen in valley bottom permafrost soils, *Glob Chang Biol* **18**:2958–2968 (2012).
27 Norby RJ, Slown VL, Iversen CM, and Childs J, Controls on fine-scale spatial and temporal variability of plant-available inorganic nitrogen in a polygonal tundra landscape., *Ecosystems* **22**:528–543 (2019).
28 Walker DA, Raynolds MK, Daniels FJA, Einarsson E, Elvebakk A, Gould WA, *et al.*, The circumpolar Arctic vegetation map., *J Veg Sci* **16**:267–282 (2005).
29 Fiencke C, Marushchak ME, Sanders T, Wegner R, and Beer C, Microbiogeochemical Traits to Identify Nitrogen Hotspots in Permafrost Regions, *Nitrogen* **3**:458–501 (2022).
30 Salmon VG, Schädel C, Bracho R, Pegoraro E, Celis G, Mauritz M, *et al.*, Adding Depth to Our Understanding of Nitrogen Dynamics in Permafrost Soils, *J Geophys Res Biogeosciences* **123**:2497–2512 (2018).
31 Keuper F, van Bodegom PM, Dorrepaal E, Weedon JT, van Hal J, van Logtestijn RSP, *et al.*, A frozen feast: Thawing permafrost increases plant-available nitrogen in subarctic peatlands, *Glob Chang Biol* **18**:1998–2007 (2012).
32 Ramm E, Liu C, Ambus P, Butterbach-Bahl K, Hu B, Martikainen PJ, *et al.*, A review of the importance of mineral nitrogen cycling in the plant-soil-microbe system of permafrost-affected soils—changing the paradigm, *Environ Res Lett* **17**:013004 (2022).
33 Chen H, Yang Z, Chu RK, Tolić N, Liang L, Graham DE, *et al.*, Molecular Insights into Arctic Soil Organic Matter Degradation under Warming, *Environ Sci Technol* **52**:4555–4564 (2018).
34 Burke E, Chadburn S, and Huntingford C, Thawing Permafrost as a Nitrogen Fertiliser:

Implications for Climate Feedbacks, *Nitrogen* **3**:353–375 (2022).
35 Kuypers MMM, Marchant HK, and Kartal B, The microbial nitrogen-cycling network, *Nat Rev Microbiol* **16**:263–276, Nature Publishing Group (2018).
36 Patzner M, Logan M, McKenna A, Young R, Zhou Z, Joss H, *et al.*, Microbial iron cycling during palsa hillslope collapse promotes greenhouse gas emissions before complete permafrost thaw, *Commun Earth Environ* **3** (2022).
37 Abbott BW and Jones JB, Permafrost collapse alters soil carbon stocks, respiration, CH4, and N2O in upland tundra, *Glob Chang Biol* **21**:4570–4587 (2015).
38 Wild B, Schnecker J, Alves RJE, Barsukov P, Bárta J, Čapek P, *et al.*, Input of easily available organic C and N stimulates microbial decomposition of soil organic matter in arctic permafrost soil, *Soil Biol Biochem* (2014).
39 Wegner R, Fiencke C, Knoblauch C, Sauerland L, and Beer C, Rapid Permafrost Thaw Removes Nitrogen Limitation and Rises the Potential for N2O Emissions, *Nitrogen* **3**:608–627 (2022).
40 Voigt C, Marushchak ME, Abbott BW, Biasi C, Elberling B, Siciliano SD, *et al.*, Nitrous oxide emissions from permafrost-affected soils, *Nat Rev Earth Environ* **1**:420–434, Springer US (2020).
41 Mao C, Kou D, Wang G, Peng Y, Yang G, Liu F, *et al.*, Trajectory of Topsoil Nitrogen Transformations Along a Thermo-Erosion Gully on the Tibetan Plateau, *J Geophys Res Biogeosciences* **124**:1342–1354 (2019).
42 Schuur EAG, Crummer KG, Vogel JG, and MacK MC, Plant species composition and productivity following permafrost thaw and thermokarst in Alaskan tundra, *Ecosystems* **10**:280–292 (2007).
43 Varner RK, Crill PM, Frolking S, McCalley CK, Burke SA, Chanton JP, *et al.*, Permafrost thaw driven changes in hydrology and vegetation cover increase trace gas emissions and climate forcing in Stordalen Mire from 1970 to 2014, *Philos Trans R Soc A Math Phys Eng Sci* **380** (2022).
44 Elder CD, Thompson DR, Thorpe AK, Chandanpurkar HA, Hanke PJ, Hasson N, *et al.*, Characterizing Methane Emission Hotspots From Thawing Permafrost, *Global Biogeochem Cycles* **35**:1–22 (2021).
45 Moore MRN, Tank SE, Kurek MR, Taskovic M, McKenna AM, Smith JLJ, *et al.*, Ultrahigh resolution dissolved organic matter characterization reveals distinct permafrost characteristics on the Peel Plateau, Canada, *Biogeochemistry*, Springer International Publishing (2023).
46 Mann BF, Chen H, Herndon EM, Chu RK, Tolic N, Portier EF, *et al.*, Indexing permafrost soil organic matter degradation using high-resolution mass spectrometry, *PLoS One* **10**:1–16 (2015).
47 AminiTabrizi R, Wilson RM, Fudyma JD, Hodgkins SB, Heyman HM, Rich VI, *et al.*, Controls on Soil Organic Matter Degradation and Subsequent Greenhouse Gas Emissions Across a Permafrost Thaw Gradient in Northern Sweden, *Front Earth Sci* **8** (2020).
48 Hodgkins SB, Tfaily MM, McCalley CK, Logan TA, Crill PM, Saleska SR, *et al.*, Changes in peat chemistry associated with permafrost thaw increase greenhouse gas production, *Proc Natl Acad Sci* **111**:5819–5824 (2014).
49 Hodgkins SB, Tfaily MM, Podgorski DC, McCalley CK, Saleska SR, Crill PM, *et al.*, Elemental composition and optical properties reveal changes in dissolved organic matter along a permafrost thaw chronosequence in a subarctic peatland, *Geochim Cosmochim*

Acta **187**:123–140, Elsevier Ltd (2016).
50 Grosse G, Goetz S, McGuire AD, Romanovsky VE, and Schuur EAG, Changing permafrost in a warming world and feedbacks to the Earth system, *Environ Res Lett* **11**, IOP Publishing (2016).
51 Bahureksa W, Borch T, Young RB, Weisbrod C, Blakney GT, and McKenna AM, Improved Dynamic Range, Resolving Power, and Sensitivity Achievable with FT-ICR Mass Spectrometry at 21 T Reveals the Hidden Complexity of Natural Organic Matter, *Anal Chem* **94**:11382–11389 (2022).
52 Smith DF, Podgorski DC, Rodgers RP, Blakney GT, and Hendrickson CL, 21 Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures, *Anal Chem* **90**:2041–2047 (2018).
53 Bahureksa W, Tfaily MM, Boiteau RM, Young RB, Logan MN, McKenna AM, *et al.*, Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation, *Environ Sci Technol* **55**:9637–9656 (2021).
54 Nebbioso A and Piccolo A, Molecular characterization of dissolved organic matter (DOM): A critical review, *Anal Bioanal Chem* **405**:109–124 (2013).
55 Sleighter RL and Hatcher PG, 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 (2007).
56 Hawkes JA and Kew W, High-resolution mass spectrometry strategies for the investigation of dissolved organic matter, *Multidimens Anal Tech Environ Res*, Elsevier Inc. (2020).
57 Wilson RM, Hough MA, Verbeke BA, Hodgkins SB, Chanton JP, Saleska SD, *et al.*, Plant organic matter inputs exert a strong control on soil organic matter decomposition in a thawing permafrost peatland, *Sci Total Environ* **820**:152757, Elsevier B.V. (2022).
58 Zherebker A, Podgorski DC, Kholodov VA, Orlov AA, Yaroslavtseva N V., Kharybin O, *et al.*, The Molecular Composition of Humic Substances Isolated From Yedoma Permafrost and Alas Cores in the Eastern Siberian Arctic as Measured by Ultrahigh Resolution Mass Spectrometry, *J Geophys Res Biogeosciences* (2019).
59 Textor SR, Wickland KP, Podgorski DC, Johnston SE, and Spencer RGM, Dissolved Organic Carbon Turnover in Permafrost-Influenced Watersheds of Interior Alaska: Molecular Insights and the Priming Effect, *Front Earth Sci* **7**:1–17 (2019).
60 Stücheli PE, Niggemann J, and Schubert CJ, Comparison of different solid phase extraction sorbents for the qualitative assessment of dissolved organic nitrogen in freshwater samples using FT-ICR-MS, *J Limnol* **77**:400–411 (2018).
61 MacDonald EN, Tank SE, Kokelj S V., Froese DG, and Hutchins RHS, Permafrost-derived dissolved organic matter composition varies across permafrost end-members in the western Canadian Arctic, *Environ Res Lett* **16** (2021).
62 Malmer N, Johansson T, Olsrud M, and Christensen TR, Vegetation, climatic changes and net carbon sequestration in a North-Scandinavian subarctic mire over 30 years, *Glob Chang Biol* **11**:1895–1909 (2005).
63 Rosswall T and Granhall U, Nitrogen Cycling in a Subarctic Ombrotrophic Mire, *Ecol Bull*:63–95 (1980).
64 ROSSWALL T, FLOWER-ELLIS JGK, JOHANSSON LG, JONSSON S, RYDÉN BE, and SONESSON M, STORDALEN (ABISKO), SWEDEN, *Ecol Bull* **no. 20**:265–294

(1975).
65 Patzner MS, Kainz N, Lundin E, Barczok M, Smith C, Herndon E, *et al.*, Seasonal Fluctuations in Iron Cycling in Thawing Permafrost Peatlands, *Environ Sci Technol* (2022).
66 Lakomiec P, Holst J, Friborg T, Crill P, Rakos N, Kljun N, *et al.*, Field-scale CH4 emission at a subarctic mire with heterogeneous permafrost thaw status, *Biogeosciences* **18**:5811–5830 (2021).
67 Olefeldt D and Roulet NT, Effects of permafrost and hydrology on the composition and transport of dissolved organic carbon in a subarctic peatland complex, *J Geophys Res Biogeosciences* **117**:1–15 (2012).
68 Olefeldt D, Roulet NT, Bergeron O, Crill P, Bäckstrand K, and Christensen TR, Net carbon accumulation of a high-latitude permafrost palsa mire similar to permafrost-free peatlands, *Geophys Res Lett* **39** (2012).
69 Bäckstrand K, Crill PM, Jackowicz-Korczyñski M, Mastepanov M, Christensen TR, and Bastviken D, Annual carbon gas budget for a subarctic peatland, Northern Sweden, *Biogeosciences* **7**:95–108 (2010).
70 Payandi-Rolland D, Shirokova LS, Labonne F, Bénézeth P, and Pokrovsky OS, Impact of freeze-thaw cycles on organic carbon and metals in waters of permafrost peatlands, *Chemosphere* **279** (2021).
71 Bockheim JG and Munroe JS, Organic carbon pools and genesis of alpine soils with permafrost: A review, *Arctic, Antarct Alp Res* **46**:987–1006 (2014).
72 Fofana A, Anderson D, McCalley CK, Hodgkins S, Wilson RM, Cronin D, *et al.*, Mapping substrate use across a permafrost thaw gradient, *Soil Biol Biochem* **175**:108809, Elsevier Ltd (2022).
73 Johansson T, Malmer N, Crill PM, Friborg T, Åkerman JH, Mastepanov M, *et al.*, Decadal vegetation changes in a northern peatland, greenhouse gas fluxes and net radiative forcing, *Glob Chang Biol* **12**:2352–2369 (2006).
74 Patzner MS, Mueller CW, Malusova M, Baur M, Nikeleit V, Scholten T, *et al.*, Iron mineral dissolution releases iron and associated organic carbon during permafrost thaw, *Nat Commun* **11**:1–39 (2020).
75 Koch JC, Ewing SA, Striegl R, and McKnight DM, Rapid runoff via shallow throughflow and deeper preferential flow in a boreal catchment underlain by frozen silt (Alaska, USA), *Hydrogeol J* **21**:93–106 (2013).
76 Stookey LL, Ferrozine-A New Spectrophotometric Reagent for Iron, 779–781 (1970).
77 Treat CC, Wollheim WM, Varner RK, and Bowden WB, Longer thaw seasons increase nitrogen availability for leaching during fall in tundra soils, *Environ Res Lett* **11**, IOP Publishing (2016).
78 Kim S, Kim D, Jung MJ, and Kim S, Analysis of environmental organic matters by Ultrahigh-Resolution mass spectrometry—A review on the development of analytical methods, *Mass Spectrom Rev* **41**:352–369 (2022).
79 Zark M and Dittmar T, Universal molecular structures in natural dissolved organic matter, *Nat Commun* **9**:1–8, Springer US (2018).
80 Mitchell PJ, Simpson AJ, Soong R, and Simpson MJ, Nuclear magnetic resonance analysis of changes in dissolved organic matter composition with successive layering on clay mineral surfaces, *Soil Syst* **2**:1–17 (2018).
81 Woods G and Simpson A, HILIC-NMR: Toward the identification of individual molecular

components in dissolved organic matter, *Environ Sci Technol* **45**:5910 (2011).
82 Tong H, Simpson AJ, Paul EA, and Simpson MJ, Land-Use Change and Environmental Properties Alter the Quantity and Molecular Composition of Soil-Derived Dissolved Organic Matter, *ACS Earth Sp Chem* **5**:1395–1406 (2021).
83 Woodcroft BJ, Singleton CM, Boyd JA, Evans PN, Emerson JB, Zayed AAF, *et al.*, Genome-centric view of carbon processing in thawing permafrost, *Nature* **560**:49–54, Springer US (2018).
84 Olm MR, Brown CT, Brooks B, and Banfield JF, DRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication, *ISME J* **11** (2017).
85 Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL 0003 3527 8101, Solden LM, *et al.*, DRAM for distilling microbial metabolism to automate the curation of microbiome function, *Nucleic Acids Res* **48** (2020).
86 Ellenbogen JB, Borton MA, McGivern BB, Cronin DR, Hoyt DW, Freire-Zapata V, *et al.*, Methylotrophy in the Mire: direct and indirect routes for methane production in thawing permafrost, *mSystems* **9**, American Society for Microbiology (2023).
87 Bäckstrand K, Crill PM, Mastepanov M, Christensen TR, and Bastviken D, Non-methane volatile organic compound flux from a subarctic mire in Northern Sweden, *Tellus, Ser B Chem Phys Meteorol* **60 B**:226–237 (2008).
88 Dittmar T, Koch B, Hertkorn N, and Kattner G, A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater, *Limnol Oceanogr Methods* **6**:230–235 (2008).
89 Li Y, Harir M, Lucio M, Kanawati B, Smirnov K, Flerus R, *et al.*, Proposed Guidelines for Solid Phase Extraction of Suwannee River Dissolved Organic Matter, *Anal Chem* **88**:6680–6688 (2016).
90 Hendrickson CL, Quinn JP, Kaiser NK, Smith DF, Blakney GT, Chen T, *et al.*, 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer: A National Resource for Ultrahigh Resolution Mass Analysis, *J Am Soc Mass Spectrom* **26**:1626–1632 (2015).
91 Poulin BA, Ryan JN, Nagy KL, Stubbins A, Dittmar T, Orem W, *et al.*, Spatial Dependence of Reduced Sulfur in Everglades Dissolved Organic Matter Controlled by Sulfate Enrichment, *Environ Sci Technol* **51**:3630–3639 (2017).
92 Hertkorn N, Benner R, Frommberger M, Schmitt-kopplin P, Witt M, Kaiser K, *et al.*, Characterization of a major refractory component of marine dissolved organic matter, 2990–3010 (2006).
93 Hansen HFE and Elberling B, Spatial Distribution of Bioavailable Inorganic Nitrogen From Thawing Permafrost, *Global Biogeochem Cycles* **37**:1–12 (2023).
94 Mao C, Kou D, Chen L, Qin S, Zhang D, Peng Y, *et al.*, Permafrost nitrogen status and its determinants on the Tibetan Plateau, *Glob Chang Biol* **26**:5290–5302 (2020).
95 Lacroix F, Zaehle S, Caldararu S, Schaller J, Stimmler P, Holl D, *et al.*, Mismatch of N release from the permafrost and vegetative uptake opens pathways of increasing nitrous oxide emissions in the high Arctic, *Glob Chang Biol* **28**:5973–5990 (2022).
96 Jones DL and Kielland K, Soil amino acid turnover dominates the nitrogen flux in permafrost-dominated taiga forest soils, *Soil Biol Biochem* **34**:209–219 (2002).
97 D'Andrilli J, Cooper WT, Foreman CM, and Marshall AG, An ultrahigh-resolution mass spectrometry index to estimate natural organic matter lability, *Rapid Commun Mass Spectrom* **29**:2385–2401 (2015).

98 Wild B, Schnecker J, Capek P, Guggenberger G, Hofhansl F, Kaiser C, *et al.*, Nitrogen dynamics in Turbic Cryosols from Siberia and Greenland c, *Soil Biol Biochem* **67**:85–93 (2013).
99 Kielland K, Amino acid absorption by Arctic plants: Implications for plant nutrition and nitrogen cycling., *Ecology* **75**:2373–2383 (1994).
100 Schimel JP and Stuart Chapin F, Tundra plant uptake of amino acid and NH4+ nitrogen in situ: Plants compete well for amino acid N, *Ecology* **77**:2142–2147 (1996).
101 Wild B, Alves RJE, Bárta J, Čapek P, Gentsch N, Guggenberger G, *et al.*, Amino acid production exceeds plant nitrogen demand in Siberian tundra, *Environ Res Lett* **13** (2018).
102 Abdulla HA, Burdige DJ, and Komada T, Accumulation of deaminated peptides in anoxic sediments of Santa Barbara Basin, *Geochim Cosmochim Acta* **223**:245–258, Elsevier Ltd (2018).
103 Ward CP and Cory RM, Chemical composition of dissolved organic matter draining permafrost soils, *Geochim Cosmochim Acta* **167**:63–79, Elsevier Ltd (2015).
104 Mueller CW, Rethemeyer J, Kao-Kniffin J, Löppmann S, Hinkel KM, and Bockheim JG, Large amounts of labile organic carbon in permafrost soils of northern Alaska, *Glob Chang Biol* **21**:2804–2817 (2015).
105 Bushnell B, BBtools.
106 Langmead B and Salzberg SL, Fast gapped-read alignment with Bowtie 2, *Nat Methods* **9** (2012).
107 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al.*, The Sequence Alignment/Map format and SAMtools, *Bioinformatics* **25** (2009).
108 Liao Y, Smyth GK, and Shi W, FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features, *Bioinformatics* **30** (2014).
109 Smid M, Coebergh van den Braak RRJ, van de Werken HJG, van Riet J, van Galen A, de Weerd V, *et al.*, Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons, *BMC Bioinformatics* **19** (2018).
110 Smyth GK, limma: Linear Models for Microarray Data, Bioinformatics and Computational Biology Solutions Using R and Bioconductor (2005).
111 Emmett MR, White FM, Hendrickson CL, Shi DH, and Marshall AG, Application of micro-electrospray liquid chromatography techniques to FT-ICR MS to enable high-sensitivity biological analysis, *J Am Soc Mass Spectrom* **9**:333–340 (1998).
112 Kaiser NK, McKenna AM, Savory JJ, Hendrickson CL, and Marshall AG, Tailored ion radius distribution for increased dynamic range in FT-ICR mass analysis of complex mixtures, *Anal Chem* **85**:265–272 (2013).
113 Chen T, Beu SC, Kaiser NK, and Hendrickson CL, Note: Optimized circuit for excitation and detection with one pair of electrodes for improved Fourier transform ion cyclotron resonance mass spectrometry, *Rev Sci Instrum* **85**:2012–2015 (2014).
114 Boldin IA and Nikolaev EN, Fourier transform ion cyclotron resonance cell with dynamic harmonization of the electric field in the whole volume by shaping of the excitation and detection electrode assembly, *Rapid Commun Mass Spectrom* **25**:122–126 (2011).
115 Blakney GT, Hendrickson CL, and Marshall AG, Predator data station: A fast data acquisition system for advanced FT-ICR MS experiments, *Int J Mass Spectrom* **306**:246–252, Elsevier B.V. (2011).
116 Xian F, Hendrickson CL, Blakney GT, Beu SC, and Marshall AG, Automated broadband

phase correction of fourier transform ion cyclotron resonance mass spectra, *Anal Chem* **82**:8807–8812 (2010).
117 Savory JJ, Kaiser NK, Mckenna AM, Xian F, Blakney GT, Rodgers RP, *et al.*, Parts-Per-Billion Fourier Transform Ion Cyclotron Resonance Mass Measurement Accuracy with a "Walking" Calibration Equation, *Anal Chem* **83**:1732–1736 (2011).
118 Kendrick E, A mass scale based on CH2 = 14.0000 for high resolution mass spectrometry of organic compounds, *Anal Chem* **35**:2146–2154 (1963).
119 Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, and Qian K, Kendrick Mass Defect Spectroscopy: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra, *Anal Chem* **73**:4676–4681 (2001).
120 van Krevelen DW, "Graphical-statistical method for the study of structure and reaction processes of coal," *Fuel* **29**:269–284 (1950).
121 Kim S, Kramer RW, and Hatcher PG, Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram, *Anal Chem* (2003).
122 Corilo YE, PetroOrg Software, The Florida State University, Tallahassee, Fl, USA 2012.

CHAPTER 3 - SOIL ORGANIC MATTER CHARACTERIZATION BY FOURIER TRANSFORM ION CYCLOTRON RESONANCE MASS SPECTROMETRY (FTICR MS): A CRITICAL REVIEW OF SAMPLE PREPARATION, ANALYSIS, AND DATA INTERPRETATION2

1. INTRODUCTION

Soil organic matter (SOM) is the accumulated, decaying debris of biota in the soil. It is both the largest active reservoir of terrestrial organic carbon,[1] and fuel for microbial activity and plant nutrient production, "giving life" to soil through its decay.[2] SOM affects soil structure, moisture retention, ion exchange capacity, and other properties, and it can be studied from many diverse perspectives, from soil health to contaminant transport. Most recently, urgent attention has been given to its relationship with atmospheric CO_2 concentrations and climate change.[3,4]

The sources of SOM are vast and various, including extracellular compounds like root exudates, extracellular enzymes, and extracellular polymeric substances,[5,6] decomposition products from above-ground plant litter, root litter, and microbial necromass,[7–10] and abiotic reaction byproducts. The succession of biotic and abiotic reactions, and the disparate conditions across both spatial and temporal scales,[11,12] create mixtures with chemical diversity that is far more complex than the original biological source materials.[13,14] Recent observations also show that SOM is heterogenous at high spatial resolution,[15,16] and that physical separations (i.e., water extractable,

[2] Bahureksa, W., Tfaily, M. M., Boiteau, R. M., Young, R. B., Logan, M. N., McKenna, A. M., & Borch, T. (2021). Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation. *Environmental Science and Technology*, *55*(14), 9637–9656. https://doi.org/10.1021/acs.est.1c01135

particulate, and mineral-associated organic matter) reflect the formation, persistence and function of SOM in the environment better than operational fractions produced by solubilities under different pH conditions (i.e., humic acids, fulvic acids, and humins.[17,18] A wide range of complementary analytical techniques are required to study such a complex and dynamic mixture when it is heterogeneously distributed among different physical environments at nearly every measurable scale (i.e., nm to km).[19]

The introduction of Fourier transform ion cyclotron resonance mass spectrometry (FTICR MS) analysis established a means to reliably detect and resolve individual SOM molecules when studying its composition, distribution, or transformation. FTICR MS achieves sufficient mass resolving power over a wide mass range (from a hundred to several thousand Da) to identify an *m/z* difference of a single electron, where *m* is ion mass in daltons (Da) and *z* is the ion charge.[20] This resolution is capable of separating the closely spaced *m/z* peaks that are found in a typical SOM mass spectra (Fig. 1).

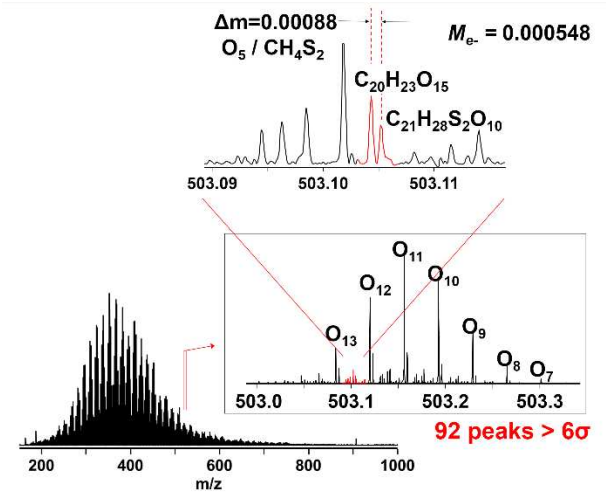


Fig. 1 FTICR mass spectra of a SOM extract displaying the resolution needed to differentiate unique peaks. The peaks ($C_{20}H_{23}O_{15}$ and $C_{21}H_{28}S_2O_{10}$) differ by 0.00088 m/z, requiring a resolution of ~600,000 (resolution = m/Δm) to separate and assign. The mass of an electron (M_e^-) is provided as reference. Unpublished data from the National Magnetic Field Lab.

However, biases that originate from sample preparation and the chosen analysis conditions make full characterization of any SOM sample difficult, whatever the mass resolving power of the FTICR MS. Due to the diversity of molecular components that comprise SOM, any extraction method will only target a specific fraction and must be tailored to a specific research question to collect the desired fraction most effectively. Ionization efficiencies vary greatly between molecules of different chemical composition and sample matrices, which complicates the quantitative comparison of SOM mass spectra. Furthermore, once mass spectra are collected, data analysis requires making assumptions to assign reasonable molecular formula to masses, grouping features by calculated indices, and creating visualizations that reflect chemical differences. Robust interpretation relies on critical assessments of the uncertainties introduced during each of these steps.

Importantly, FTICR MS is not, by itself, able to differentiate between isomers, determine molecular structures, or identify functional groups, meaning that FTICR mass spectra are still simplified representations of SOM.[21] FTICR MS provides a qualitative view into a subset of SOM compositional space. However, these data can be further examined using LC-MS and other complementary approaches (Section 2 of the supporting information) to identify and characterize specific isomers of interest and contextualize relationships between peaks and samples.

There has been significant progress in overcoming the challenges of organic matter analysis using FTICR MS,[22–24] and the user base of FTICR MS analyses is growing rapidly as techniques for soil organic matter analysis become more standardized. Herein, we aim to provide a comprehensive resource to multidisciplinary researchers interested in applying FTICR MS to obtain molecular level insight to soil organic matter chemistry. Workflows for FTICR MS based SOM analysis will be presented from an experimental design perspective, and potential biases

from sample extraction, ionization, and formula assignment methods will be highlighted. Data visualization and exploration methods will also be discussed, with suggested usage to avoid misinterpretations. Finally, techniques complementary to FTICR MS for SOM analysis will be presented with their specific strengths to support more robust conclusions about SOM composition, microbial decomposition pathways and other relevant areas of study (Fig. 2).

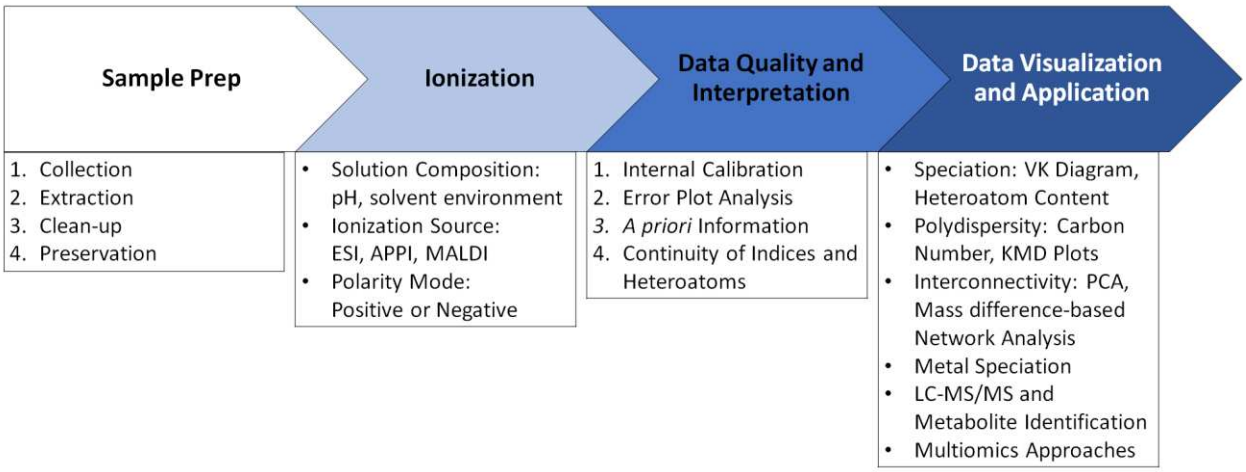


Fig. 2 Workflow for analysis involving FTICR MS to provide meaningful conclusions about a set of samples.

2. Sampling and Extraction of Soil Organic Matter

2.1 Sample collection

SOM sampling methodology depends primarily on the scientific focus of the study and which fraction of SOM will provide the most relevant information. The heterogeneity of SOM, soil surfaces, and microenvironments make a single extraction procedure for all SOM features impossible.[25,26] SOM has been historically isolated into fractions operationally defined by their solubility at different pH.[27] While this method continues to be used,[28,29] other fractions of interest can be collected by targeting specific SOM sorption mechanisms. SOM is bound to soil minerals and other organic matter via different mechanisms, from weak electrostatic, hydrophobic, and

other nonspecific interactions to strong polar covalent bonds produced by ligand exchange with carboxylic and phenolic functional groups.[25,30] (Fig. 3) The most commonly used extraction methods include water, alkaline, mineral dissolution, and organic extractions, and each method has been shown to extract different fractions of SOM. (Table S2) Extracted fractions can be related and even complement each other, however they cannot be assumed to be "complete" or truly representative of total SOM as they, by design, only target specific operational fractions defined by the targeted stabilization mechanism.[31] This review outlines the major extraction methods and the key shortcomings of each that must be considered when collecting samples for FTICR MS.

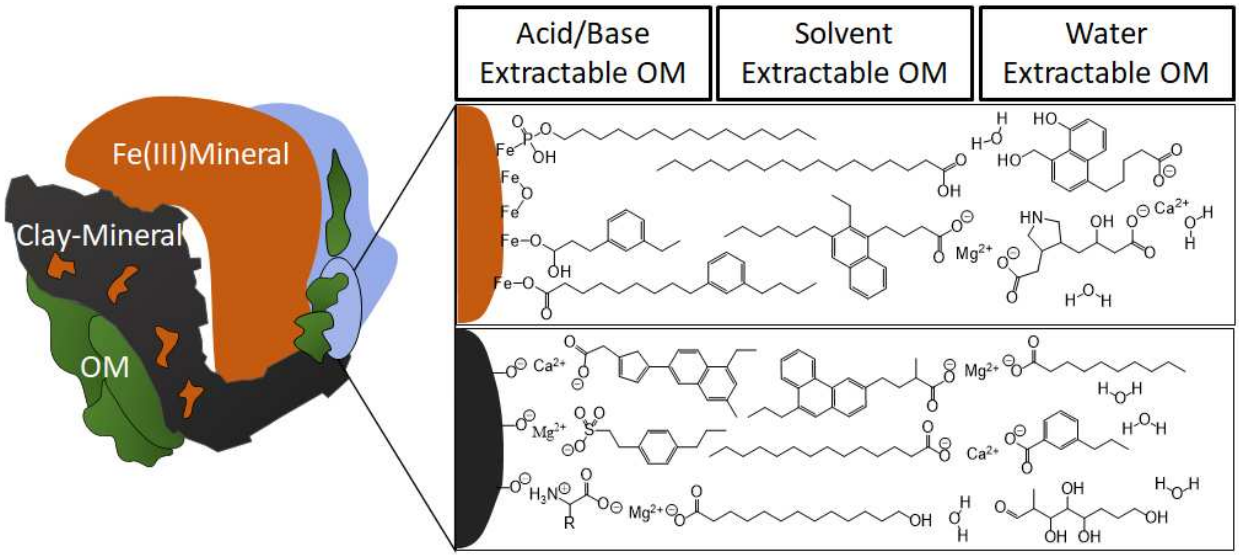


Fig. 3 Model soil particle depicting an Fe(III) mineral (brown) and clay (black) surface exhibiting various possible interactions with model SOM. The variety of possible interactions will create fractions with varying extractability based on the extraction solution (e.g., acid/base, organic solvent, and water), mineral surface, and intermolecular forces.

2.2 Water-based Extraction

For approximately two decades, water extraction has been used as a collection method for SOM that is released to the soil solutions when equilibrium conditions change.[32] The resulting fraction, referred to as water-extractable SOM (WEOM),[33,34] is closely related to dissolved organic matter (DOM), which is defined operationally by filtration (i.e. passing through 0.4 - 0.6 µm pore

sizes)[35] of water-extractable or aquatic organic matter.[36,37] These terms both describe complex natural organic matter and have been shown to be highly related through biomolecular origin and degree of decomposition.[37–40] Therefore, the following discussion on preparation, ionization, and analysis can be similarly applied.

WEOM is considered the most dynamic and bioavailable fraction of SOM and therefore its analysis is highly applicable to a variety of studies.[41–45] The importance and broad application of water extractable SOM is due to its roles in aqueous photochemistry, nutrient cycling, ion distribution, and pH buffering; in addition to acting as an energy and carbon source for microbes.[46–49] Water extraction is performed with ultrapure water or salt solutions such as potassium sulfate (or potassium chloride) which increases ionic strength, generally improving SOM yield.[32,50–53] Additional alteration of extraction conditions, such as increasing temperature, pressure, or time, have failed to produce significant improvement over leaching or batch extractions at room temperature.[45] Overall, water extractions are easily performed but are generally unable to extract SOM adsorbed to mineral surfaces, hydrophobic molecules, or physically protected SOM, which are important when considering carbon sequestration and mineral-organic interactions.[34,50]

2.3 Selective Mineral Dissolution

Study of the role of mineral protection in SOM stability has increased considerably in recent years as mineral-organic associations are thought to increase protection from microbial degradation processes.[38,66] For example, iron oxide minerals are effective sorbents or organic matter through interactions that include co-precipitation, and ligand exchange under more acidic conditions.[67–69] To study these mineral-organic associations directly, mineral dissolving solutions are used to selectively extract SOM associated with those minerals.[67] These extractions function

by either reduction, chelation, or dispersion of a mineral phase; resulting in the release of SOM stabilized by that mineral.[34,44,70–72] Example mineral phases include short range order (SRO) Fe(III) oxides, crystalline Fe(III) oxides, aluminum oxides, and aluminosilicate clays among others. Selection of appropriate extraction solutions has been extensively studied in SOM-mineral or mineral cycling studies.[34,50,67,73–75] Not as well characterized is the potential for SOM alteration by extraction conditions or interactions between the DOM and dissolved minerals.[75–77] After dissolution, new SOM-mineral interactions can be inhibited by inclusion of an appropriate buffering or stabilizing agent, such as citrate for Fe^{2+}, that can complex metal ions and prevent re-adsorption or aggregation of extracted SOM.[76,78,79] All mineral extraction will require a cleanup process to remove the extracting chemicals and stabilizing agents that can affect ionization, discussed below. Finally, mineral extractions have the potential to extract compounds that overlap with water extractions and should be noted when comparing extraction yields and spectra.[51,75,80]

2.4 Organic Solvent Extraction

Soil organic matter is known to contain hydrophobic domains[63] and laboratory studies have indicated that non-selective sorption of DOM contributes substantially to the overall presence of mineral-organic associations.[81] To study these interactions, it is useful to have an extraction method that targets more hydrophobic domains.

Organic solvents can be more efficient extractors than polar solvents depending on the source material and target fraction (e.g., the hydrophobic domain).[82] Tfaily *et al.* showed that extracted H/C and O/C ratios are influenced by solvent polarity by performing a series of parallel extractions, and in a subsequent study show that low/non-polar solvents detect a greater number of low O/C peaks in FTICR-MS analysis.[41,83] Furthermore, Mckee and Hatcher[84] demonstrated

that pyridine extractions of lake sediment isolate a larger range of mass spectral and spectroscopic features compared to other organic and aqueous extractions. Organic extractions can also induce alterations to the SOM, such as esterification following methanol and acetonitrile extraction, that should be considered before usage and when comparing between samples prepared differently.[85] Extraction with organic solvents also dissolves fewer inorganic salts that inhibit SOM ionization compared with aqueous extractions. This is a potentially large benefit as other extractions methods require cleanup that further bias SOM composition.[41]

2.5 Sequential Extraction

Extractants can also be applied sequentially to soils to selectively collect SOM bound at different degrees of stabilization. This can be very beneficial when comparing the quantity and characteristics of SOM stabilized by the different mechanisms mentioned above. Lopez-Sangil and Rovira[50] demonstrated this with seven extracting solutions on four soils of varying pH and organic carbon content. Each extractant collected a fraction of SOM but the majority of SOM was collected with sodium tetraborate or sodium pyrophosphate, indicating that cation bridging or chelating interactions predominated, respectively. Tfaily *et al.*[41] also showed that sequential extractions with water followed by different organic solvents resulted in an overall increase in peak counts by 2- to 4-fold. As above, it is possible that extractants overlap in the SOM collected, therefore each extraction should be performed several times to ensure that as much of the target fraction is extracted as possible before continuing with the next extractant. The additional extracts collected will increase preparation time, however sequential extraction protocol can be tailored to specific research questions based on the mineral-organic associations of interest.

2.6 Sample Preservation and Cleanup (for FT ICR-MS analysis)

Once a soil sample has been collected, it is important to perform an extraction as soon as possible or store the sample to minimize biological degradation of SOM.[86] If extraction soon after sampling is impossible, it is recommended that soil samples be treated as similarly as possible to maintain their comparability, as no storage method can perfectly preserve the *in situ* conditions.[59,78] Before extraction, soil sample storage by freezing or drying should be avoided as both treatments can aggregate the soil or SOM ,[78,87,88] which can alter the structure of the soil matrix, reducing exposure/contact with the extracting chemicals,[78,87,88] and has shown variable carbon yield[78,88] when compared to "wet" or *in situ* samples.

After extraction, long-term storage of SOM extracts is also preferably avoided to minimize post extraction alteration. SOM stability in the extract should be accounted for, but this is easily addressed by sterile filtration, or immediate processing to prepare the samples for FTICR MS. When the extract is aqueous, freezing can provide long-term SOM stability if the solution cannot be sterile filtered, but SOM aggregation can occur so analysis before long-term storage is ideal.[89] The use of acidification or biocides to sterilize sample extracts for storage is also not recommended as there is evidence for alteration of SOM properties.[90,91]

After extraction, SOM often requires concentration and salt removal to make it suitable for FTICR MS analysis. This process is most commonly performed by solid phase extraction (SPE), which utilizes the difference in polarity between small, highly polar inorganic salts and less polar SOM for separation.[92,93] As SPE cartridges use a variety of nonpolar packing material and acidic functional groups are common, retention of the SOM is assisted by adjusting the extract to around pH 2, resulting in protonation of acidic functional groups, decrease in polarity, and improvement of stationary phase binding, however acidification can cause compositional changes in the

SOM.[90,94] Additionally, amino functional groups and N heterocycles can be protonated at low pH, increasing their polarity and preference for cation exchange sorbents.[95]

There are a wide variety of SPE cartridges that retain targeted or broad molecular features, and must be selected based on the design of the experiment.[96] The recovery of SOM from this method is highly variable and depends greatly on the cartridge being used, the origin of the SOM, and the elution procedure,[95,97,98] with recoveries varying between 20 to 90%.[93,99] Addressing the source of variable yields is challenging as there is often no reliable way to identify if the retention or elution were incomplete.[31,93,100]

Alternatives to SPE include ultrafiltration (UF), electrodialysis (ED), and reverse osmosis (RO). Techniques using membranes (i.e., UF and RO) are susceptible to fouling,[101–103] however coupling these techniques with SPE can increase carbon recoveries.[104,105] When coupling SPE with an RO/ED method, Green *et al*.[98] reported near 100% carbon recoveries for marine DOM samples. It is unclear how the higher recovery would affect FTICR MS spectra, and the RO/ED setup costs at least twice as much to perform than SPE.[98] While the combination of methods is still subject to artifacts from procedures in both methods, the improved recovery warrants evaluation for similar performance in extracted SOM.

3. Ionization Techniques and Matrix Effects

Characterization of SOM by any mass spectral technique requires the conversion of nonvolatile analytes from solution to gas phase ions by ionization. All soft ionization techniques are selective, and since OM contains a wide range of polyfunctional species across a wide molecular weight range and ionization potentials, ionization efficiency is determined based on the analyte composition. Different ionization mechanisms preferentially ionize certain components of OM (i.e., functional groups), enhancing their signal and suppressing the signal from other ions,

since ionization efficiency depends on acidity/basicity, hydrophobicity, molecular weight, and degree of conjugation. Thus, experimental design begins with a consideration of the selectivity of each ionization technique including the potential need for using multiple techniques for a better understanding of the formulas and features present.[106–108]

3.1 Electrospray Ionization

Electrospray ionization (ESI) is the most[109] routinely applied ionization techniques for WEOM components. Ionization is achieved by infusing the analyte (e.g., OM) solution through a positively or negatively charged capillary that generates a fine mist of droplets. Solvent removal, often assisted by heating, yields charged molecular ions. This process is one of the softest methods of ionization and during OM analysis is often assumed to produce intact analyte ions, although some analytes may still undergo in-source fragmentation, particularly at high spray voltages.[110–112]

Ion formation by ESI depends largely on the *pKa* of the analyte as well as the ionization polarity (e.g., negative or positive mode). In negative mode, ions are generally formed by deprotonation or formation of adducts with anions such as Cl^-. Negative ionization efficiency generally correlates inversely with *pKa* and the extent of charge delocalization, such as in conjugated and aromatic chemical structures.[113] Common acidic functional groups in SOM such as carboxylic acids and phenolic groups tend to form negative ions. In positive mode, ions form by protonation or adduct formation with cations such as Na^+, K^+, NH_4^+, or metal ions. Many functional groups in SOM such as amines, alcohols, and carbonyls can be protonated to form cations, and their ionization efficiency is largely correlated with basicity (conjugate acid *pKa*).[114] Thus, positive and negative polarity modes may detect distinct components from a SOM mixture.

Analyte hydrophobicity and molecular size influences ESI efficiency. Hydrophobicity often correlates with higher ionization efficiencies, particularly in aqueous solutions, due to enhanced affinity for the surface of droplets where molecules have a greater probability of being desolvated and charged.[115] ESI sensitivity is often also increased when the analyte is already ionized in solution (i.e., acidic conditions for basic analytes and basic conditions for acidic analytes),[116] although specific pH effects vary depending on analyte chemistry.[117,118]

Solvent composition influences the relative solubility and droplet surface affinity for polar and nonpolar components, and therefore ionization efficiency.[115,116] Novotny *et al.*[119] demonstrates that different solvent compositions, including neat solvents, efficiently ionize specific fractions and influence the representativeness of a sample spectra. Solvent mixtures commonly used for ESI include water, methanol, and acetonitrile, and can improve ionization efficiency of target molecules[115] and reproducibility of mass spectra[120] by altering the proportion of organic solvent.

Additives, contaminants, salts, and metals can also enhance or suppress ionization by ESI (i.e., introduce matrix effects) and can also dilute feature signals across multiple adducts, discussed further below. Additives can extend the range of compositional features collected[121,122] and enhance the ionization of specific molecular features.[123] In contrast, there are numerous contaminants that can be introduced accidentally during sample collection and preparation that will suppress SOM ionization. This is especially true for surfactants, which preferentially ionize due to their amphiphilic nature and ionizable head groups (Fig. S2).[81] While the bulk SOM characteristics and most ionizable species will likely still be observable, ionization suppression will limit the number of other peaks that can be detected and assigned near the limit of detection. This can dramatically influence heteroatomic assignments and peak count overall, therefore, it is

crucial that materials used for sampling, filtration, and processing prior to collection and sample analysis are tested.

The use of micro or nano ESI reduces ion suppression from salts and other species, and improves the ionization efficiency of compounds with less surface affinity, because droplet sizes are reduced and desolvation is enhanced compared to electrospray at higher flow rates.[124] However, high proportions of surfactants or other contaminants can still influence the number of SOM peaks that can be detected and assigned because it remains desirable to limit the number of trapped ions during SOM analysis to prevent adverse effects from space charging[125] (discussed below).

3.2 Ionization Techniques for Non-polar and Water-Insoluble Soil Organic Matter

While ESI remains the most common ionization technique used for OM analysis, atmospheric pressure photoionization (APPI) offers different ionization mechanisms and selectivity. Using APPI extends the analytical window of FTICR MS to less polar analytes, evidenced by direct comparisons with ESI that display ionization of formulas with lower oxygen content and higher carbon unsaturation.[106,107,126] This will affect heteroatom abundance and peak assignments (Fig. 4), as less polar nitrogen-containing compounds can be preferentially ionized using positive ion APPI,[127] yet sulfur atoms presumed in polar functional groups are diminished.[106] For positive ion APPI, radical cation ($M^{+\cdot}$) and protonated $[M + H]^{+\bullet}$ species are formed based on ionization energies of the analyte in respect to the dopant. For negative APPI, analyte ions differ based on electron affinity and include $M^{-\cdot}$, $[M - H]^-$, $[M - X + O]^-$, and $[M + A]^-$ where A is an anion (typically a halogen) and X is H, Cl, or NO_2. Developments and mechanistic discussions of APPI have recently been reviewed.[128]

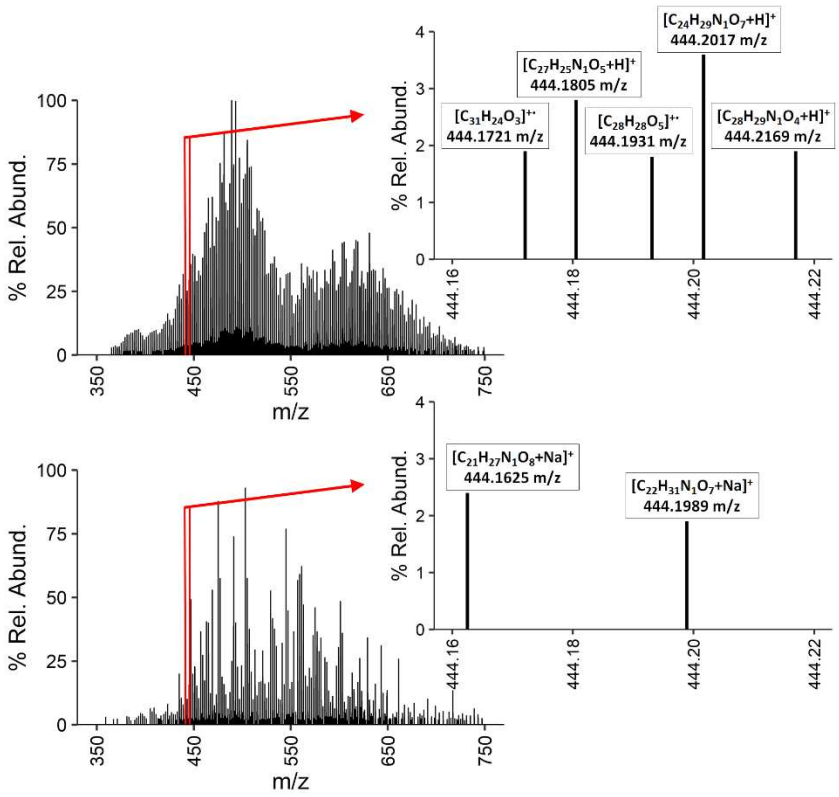


Fig. 4 FTICR mass spectra of DOM from Lake Drummond (Great Dismal Swamp National Forest, VA, USA). Samples were prepared using C_{18} solid phase extraction and spectra were obtained with atmospheric pressure ionization sources: (+)APPI (top) and (+)ESI (bottom). Spectra display peaks assigned with <1 ppm mass error, where insets indicate molecular formulas unique to each ionization source assigned over 444.16 – 444.22 *m/z*. Produced with data obtained from Hockaday *et al*.[107]

The main benefit of APPI over ESI is the simultaneous formation of gas phase ions from both polar and non-polar compounds. Other benefits include less suppression from salt and solvent effects as charge acquisition occurs differently than in ESI.[129] Comparisons of these methods for lipid analysis suggest that APPI is far more sensitive than ESI, and is particularly useful for studies focusing on condensed aromatics and hydrocarbons within natural organic matter.[130] However, the additional ionization mechanisms also increase the number of peaks in APPI compared to ESI,[107] increasing the complexity of the mass spectrum and requiring even higher mass resolving power

to separate the increased number of isobaric species. The amount of sample required for APPI is also generally higher as injection flow rates can be orders of magnitude larger than in ESI.[106,107,126]

Solid phase samples can also be ionized through desorption techniques, including desorption atmospheric pressure photoionization (DAPPI), laser desorption ionization (LDI), and matrix-assisted laser desorption ionization (MALDI). These techniques allow ordinarily insoluble SOM components to be analyzed directly, such as with the application of DAPPI to pyrolyzed carbon.[131,132] Similar to APPI, LDI and MALDI have been shown to compliment ESI-FTICR MS by ionizing more conjugated and aromatic DOM constituents.[133,134] Mechanisms and substrate selection for LDI and MALDI are reviewed by Zenobi *et al*. and Dreisewerd.[135,136]

Importantly, while using multiple techniques will provide a larger window into the OM compositions present, it is not feasible to use every technique nor can the use of every technique unequivocally ionize every organic residue present in injected samples. Technique and polarity mode should therefore be chosen based on *a priori* knowledge of the sample and target fraction, as ionization parameters can greatly influence sample representation. Additionally, the intensity and m/z distribution are sensitive to instrumental parameters, meaning FTICR MS is not well suited for resolving controversies regarding the size distribution of SOM.

4. Analysis and Interpretation of Soil Organic Matter using FTICR MS

When analyzing NOM, FTICR MS typically detects and resolves many thousands of peaks in the *m/z* range from 150 to 1000 *m/z*. Molecular formula assignment involves calibrating the mass spectra, computing formulas that fit each mass within the applicable mass error window, and choosing an appropriate assignment when there are conflicts using all available information (e.g., elemental content, and chemical or structural relationships with other formulas).[137] Uncertainty in assignments rises dramatically with mass and the number of elements[138] as the number of possible

formula increases. While automated assignment procedures often select the formula with lowest mass error or smallest number of non-oxygen heteroatoms (i.e., atoms in the ring of a cyclic compound other than carbon or hydrogen atoms), these approaches can sometimes yield incorrect results.[139] Accordingly, various methods, including *m/z* vs. error[140] and DBE minus O plots[139] can be used to further evaluate dubious or conflicting formula assignments.[141] Following formula assignment, several assumptions will be present in data visualizations and should be carefully considered to avoid misinterpretations. Additional methods, both online and offline to FTICR MS, will be discussed that can provide additional insights into SOM composition, microbial decomposition pathways, and other relevant areas of study.

4.1 Mass Calibration

Molecular formula assignment relies on mass error windows, so mass calibration is a critical part of the formula assignment process.[142] External calibration cannot provide better mass accuracy than a few ppm because the number of ions in the ICR cell varies from sample to sample.[143] As a result, complex organic mixture FTICR mass spectra benefit from internal calibration, wherein the ion cyclotron resonance frequency (*f*) is converted to ion mass-to-charge ratio (*m/z*) by fitting experimental data spanning the entire detected *m/z* range (e.g., multiple homologous series) to the relation $m/z = A/f + B/f^2$, to yield root-mean-square mass error as low as ~10-200 ppb.[144]

Savory *et al*. developed a calibration protocol in which the spectrum is divided into dozens of adjoining segments, with separate calibrations applied to each, to eliminate systematic error with respect to m/z, and introduced a third calibration term to minimize systematic error with respect to ion abundance.[140] In addition, Dittmar *et al.* recently reported improved mass accuracy

by averaging mass spectral data from independent environmental samples, and increased mass precision through peak alignment during post-detection data processing.[145]

The minimum resolution requirements to adequately analyze a sample composition will change depending on signal-to-noise (S/N) thresholds, the dynamic range of the instrument measurements, digital resolution, and isotopic fine structure.[146] Ideally, the instrument and resolving power should be considered based on *a priori* information as additional constraints, such as inclusion or exclusion of specific heteroatoms, can substantially change the minimum resolving power required.[22] Increasing magnetic field strength will increase mass resolution, dynamic range, and peak non-coalescence.[20,147] Increasing the time-domain acquisition period will also increase resolution, however if sample signal magnitude is low (i.e., sample ions are very low in abundance), instrument noise can also erroneously be assigned.[20] Increasing ion accumulation times can increase sensitivity, however high ion densities in the ICR cell can produce space charge effects that coalesce peaks or alter ion frequencies, adversely affecting both formula assignment and mass accuracy.[20,125,148,149] The limited ion capacity during FTICR MS measurement also means that intensity and m/z distributions are sensitive to instrumental parameters. Molecular weight distributions are better determined in low resolution MS systems like linear ion trap and time-of-flight that also capture lower molecular weights (<200 m/z) efficiently.[150,151] Thresholds for S/N will also strongly affect peak detection and reproducibility. Discussions by Sleighter *et al.*[152] indicate how S/N affects minimum common peak assignments and describes methods to evaluate meaningful differences between samples.

4.2 Automated Molecular Formula Assignment

There are a range of tools available for automated molecular formula assignment, including in-house programs (PetroOrg/EnviroOrg at the National High Magnetic Field Laboratory),[153] open-source programs (Formularity, UltraMassExplorer, and ICBM-OCEAN) and publicly-available R or Matlab scripts (MFassignR, CIA, and TRFu).[137,154,155] The Compound Identification Algorithm (CIA) developed by Kujawinski *et al.* computes elemental compositions for low molecular weight compounds (<500 *m/z*), where the potential for conflicting solutions is minimal,[156] and uses a "formula extension" approach to connect the assigned formulas with higher molecular weight compounds via established chemical and functional group relationships (e.g., CH_2, $CH_4 - O$, H_2, CO_2 and NH). The existence of these relationships has also been separately established by mass difference and network analyses, discussed in Section 3 of the supporting information.[157,158] Indeed, some version of formula extension is used in almost every formula assignment algorithm. The CIA algorithm has been incorporated into the freely-available Formularity software platform, which includes an associated database and formula filters based on the "Seven Golden Rules".[159] The CIA component is restricted to C, H, N, O, S and P, but a separate isotopic pattern algorithm (IPA) is available to identify other elements, particularly when they present a strong isotopic signature (e.g., Br, Cl or Hg).[160]

ICBM-OCEAN is a server-based tool that integrates published and novel approaches for standardized processing of ultrahigh resolution mass spectrometry data from complex molecular mixtures.[161] Importantly, ICBM-OCEAN incorporates diagnostic and validation tools for each step in data processing, including noise and systematic error reduction and spectra recalibration and alignment, and it has been tested and applied to both FTICR MS and Orbitrap MS for more than 1000 NOM samples.

After automated formula assignment, equivocal assignments (i.e., peaks above S/N with multiple feasible compositions) will generally remain. Molecular formula assignment for equivocal assignments can be highly subjective, therefore different calculated indices and empirical evidence should be used to increase the reliability of manually distinguished assignments wherever applicable. For instance, the selection of heteroatoms during assignment can dramatically alter the number of equivocal assignments (Fig. S3), however prematurely excluding elements can also result in erroneous assignments.[21] Kendrick mass defect analysis can be used to study patterns in OM assignments, discussed below. Double-bond-equivalents-minus-O (DBE-O) plots are also used that compare equivocal and unequivocal assignments using heteroatoms and oxygen content (Fig. S4).[138]

4.3 Visualization and Data Analysis

Once elemental compositions have been assigned, several parameters can be calculated using the stoichiometry of the assigned molecular formulas. These include rings-plus-double-bonds equivalents (RDBE), aromaticity indices (AI_{mod}),[159,162] and the average nominal carbon oxidation state (NOSC) (Fig. 5B).[163] NOSC has been used to examine chlorinated disinfection byproduct formation,[164] water solubility,[165] and the biogeochemical reactivity of SOM.[166] Indeed, NOSC makes it possible to calculate the Gibbs free energy (ΔG) provided by the oxidation of a particular compound ($\Delta G°_{Cox}$) given the available electron acceptors and environmental conditions.[167–169] Thus, NOSC and $\Delta G°_{Cox}$ have been used to provide a thermodynamically relevant metric for approximating the energy stored in organic matter, or whether microbial oxidation of OM is thermodynamically favorable under a given set of redox conditions.[170]

Kendrick mass defect (KMD) analysis[171] was initially adopted by the organic matter community as the standard means for visualizing complex mass spectra prior to formula assignment.[172–176] Kendrick plots are generated by plotting nominal mass as a function of KMD, most commonly based on a methylene (CH_2) subunit (Fig. 5C). In effect, each *m/z* is normalized by the accurate mass of the subunit (e.g., 14/14.01565 for CH_2 subunits, common to lipid (Fig. S1), sugar, and lignin derivatives[177]). Afterwards, compounds whose elemental compositions only differ by the number of subunits possess the same KMD value and line up horizontally in the associated KMD plot. KMD analysis on SOM is useful for: i) molecular formula assignment by assisting in assigning molecular formulas at higher *m/z* ions when there are more formulas that match an exact *m/z* value, ii) identifying multiply charged ions, and iii) multiply charged polymer ions of plant origin in soils.[178–181]

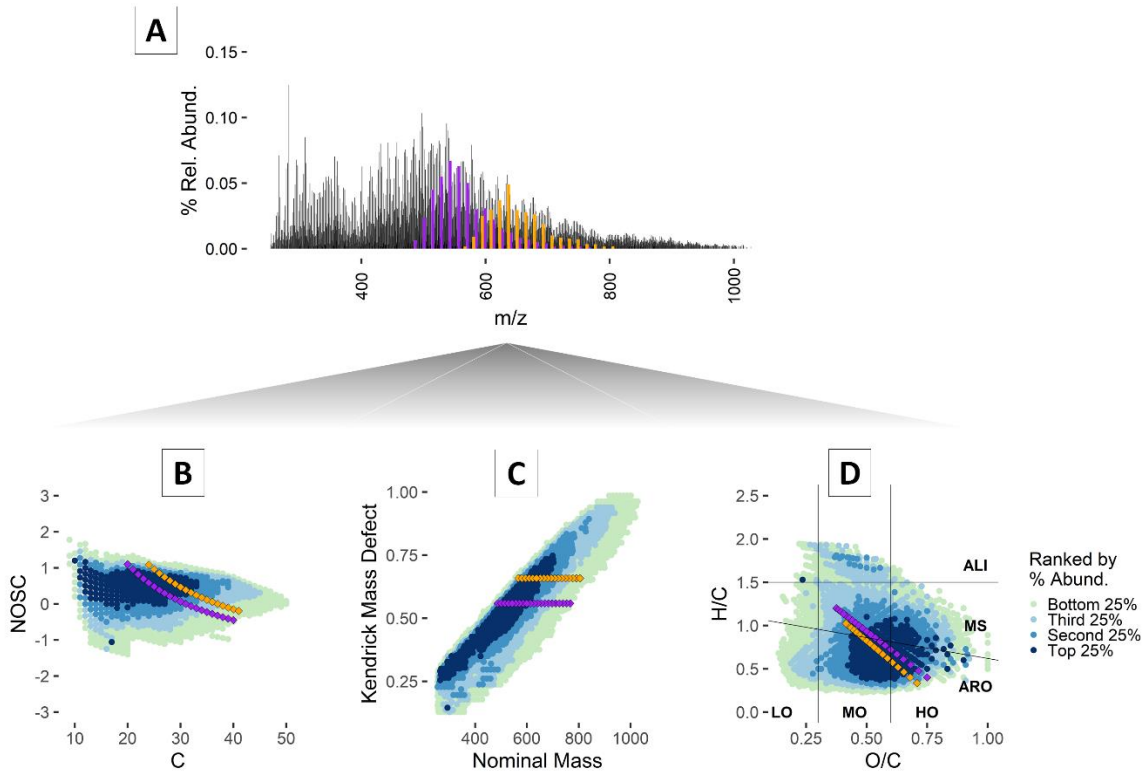


Fig. 5 The 21T FTICR mass spectrum of Pahokee Peat extracted with pyrophosphate with formula assignment conducted using PetroOrg software[153] to assign 15506 assigned formulas. Two series

are highlighted (orange and purple) that represent CH_2 Kendrick mass defect (KMD) series (0.55857 and 0.65806) and are present in each of the plots to illustrate their appearance and patterns through different visualizations: A) mass spectra displaying semi-gaussian abundance distributions, B) nominal oxidation state carbon (NOSC),[163] C) the CH_2 KMD series over the entire mass spectrum, and D) a van Krevelen diagram which sections plot regions indicating oxygen content (LO = low-oxygen, MO = mid-oxygen, HO = high-oxygen content) and degree of saturation (ARO = aromatic,[162] MS = mid-saturated, and ALI = aliphatic[182]). Formulas are colored based on abundance quartiles. Unpublished data from the Borch Lab.

Van Krevelen (VK) diagrams[183] are constructed with assigned formulas using molar ratios of hydrogen-to-carbon (H/C) on the x-axis and oxygen-to-carbon (O/C) on the y-axis (Fig. 5D). Points can be colored, or a z-axis can be added, to provide additional information such as relative abundances, molecular weights, or elemental classes (e.g., CHO vs. CHNO). VK plots are widely used to make inferences about predominant reaction pathways (e.g., methylation, hydrogenation, condensation), and to estimate the abundance of major compound categories based on the major biogeochemical components of the source material (e.g., lignin-like, lipid-like, or carbohydrate-like), which have characteristic H/C versus O/C ratios.[184–186] However, the boundaries of the compound classes vary among published studies, and often overlap. Moreover, "lignin-like" components may be derived from non-lignin source materials,[187] or even reflect synthetic chemicals if anthropogenic impacts are present (Fig. S5), and lignin sources may generate some non-"lignin-like" signatures.[188] Accordingly, while the major compound categories do signify molar ratios similar to well-known biogeochemical classes, they are not reliable indicators of source or structure without additional, complementary information (e.g., on the prevalent reaction pathways).[184] (Section 2 of the supporting information) Recent revisions proposed by Rivas *et al.*[189] improves compound classification in biological samples using C/H/O/N/P stoichiometric ratios, but was only tested on different metabolite compositions and omitted "lignin-like" and "condensed hydrocarbon" compounds that could be a major component of soil organic matter.

Regardless of the VK plot limitation, this tool can still provide useful insights into C cycling and changes in soil organic matter composition with perturbation. Users are advised to use O/C and H/C boundaries that are consistent with the chemistry of their sample and carefully interpret their results while acknowledging that these classifications are only putative. Relative abundance values of different compound classes can also be calculated from count values associated with each observed biomolecule group normalized by the total number of C molecules identified (i.e., number-weighted).[190,191] When comparing multiple samples in side-by-side comparisons or combined analyses, it is important to ensure that systematic biases (e.g., ion suppression or differences in SPE extraction efficiency) are minimized or eliminated, particularly because many SOM constituents are detected at signal intensities close to the detection limit. Molecular weight distribution, heteroatom content, constituent presence or absence, and VKD coverage can all be influenced by these low-intensity signals. Past approaches to eliminate systematic bias include extracting the same number of the most intense signals from every sample,[192] or the formulas common to every sample,[81,193] and then normalizing the extracted signals based on total intensity.

Other visualizations are common in the petroleum and aerosol field for rapid visual comparisons to highlight compositional differences/similarities between samples and can be beneficial to SOM studies.[163,194–196] These include: i) plotting the number of carbons in each formula versus its nominal mass (C versus M), where the molecules are classified into different categories based on their sum of carbon and oxygen atoms (i = C + O), ii) contoured plots of RDBE vs carbon number[197] for members of a single heteroatom class, iii) relative abundance histograms for heteroatom classes, and iv) Kroll diagrams where NOSC is plotted as a function of number of

carbons (Fig. 5B).[193] Kroll diagrams are especially useful when looking at changes in SOM composition under different redox[198] conditions.

When comparing multiple samples, such as DOM from different sources, principle component analysis (PCA)[193,199] and hierarchical clustering analysis (HCA)[200] can be used to group similar samples. PCA takes advantage of the fact that many variables are correlated (e.g., H/C ratios and double bond equivalents), and produces diagrams that group similar samples and show correlations among variables. In one example, PCA was used with intensity-weighted element number, *m/z*, H/C, O/C, AI_{mod}, and DBE values to compare the DOM in ten world rivers, associating the intensity-weighted variables with different levels of anthropogenic influence and other watershed characteristics (Fig. 6).[199] PCA is particularly useful for visualizing the relative similarities and differences of multiple samples, and for identifying the variables that best explain their differences. Importantly, compositional data, such as the %CHO of a DOM sample, are generally not appropriate for PCA.[201] In comparison, HCA uses Bray-Curtis dissimilarity or another distance calculation to perform a series of partitioning calculations that group samples into clusters.[100,202,203] The result is a tree-like dendrogram, where common branches indicate greater similarity.

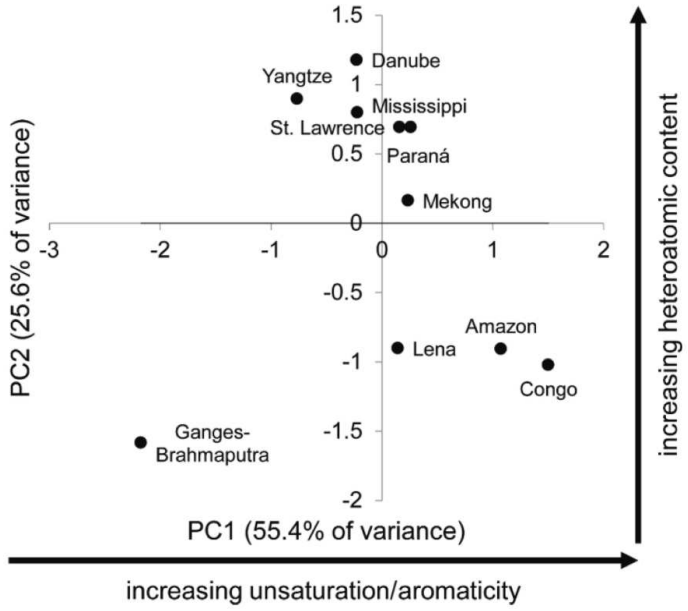


Fig. 6 Ten rivers around the world were evaluated using FTICR MS and compared with PCA to identify sample groupings. The largest variance originated from the unsaturation/aromaticity and heteroatom content. Reproduced from Wagner *et al*.[199]

4.4 Analysis of NOM-Associated Metals

Mineral associated SOM and DOM are known to bind a range of metals and thus control the biogeochemical cycling, mobility, and bioavailability of both toxic metals (e.g., Hg, As, Cd) and essential mineral elements (e.g., Fe, Mn, Zn).[198,204] FTICR MS is particularly well suited for elucidating the speciation of metals and micronutrients that are complexed by or incorporated into organic molecules that influence their solubility and reactivity. In soils, these elements are often binned into operationally defined fractions based on extraction protocols[205,206] that seek to differentiate water soluble, mineral associated, or bioavailable content. However, understanding the origin and dynamics of these elemental pools requires knowledge of their chemical identity. FTICR MS can resolve diagnostic isotopologues and mass defects that are characteristic of many heteroatoms, including trace metals (e.g. Fe,[207–209] Zn, Cu,[210] Ni, Hg[211]) (Fig. 7), as well as halogens (Cl, Br, I[212]), thus providing a means to identify the particular SOM components that play a role in the cycling of these elements.

Direct infusion FTICR MS analysis, however, is not well suited to identifying metal-bound molecules in DOM due to the number of peaks that results in unrelated species that mimic mass differences and relative abundances of metals. Furthermore, molecular formula assignments become more ambiguous as heteroatoms are included in the assignment algorithm (Fig. S3).[21] Boiteau *et al*.[207] developed a method for confident detection of metal-organic complexes by comparing direct infusion FTICR MS spectra of samples spiked with both natural and rare isotopes of metals of interest, comparing the ratio of features across samples. This was compared to another method based on separation using online HPLC-FTICR MS and detection of isotopologues with coherent elution profiles. While the isotope exchange method required significantly shorter analysis time, nearly twice as many metal-organic species were measured by HPLC-FTICR MS. This highlights how ion suppression is one of the key challenges in overcoming the detection of metal-organic species in complex mixtures.

Techniques such as liquid chromatography hyphenated inductively coupled plasma mass spectrometry (LC-ICPMS) have been used to quantify the various metal species present in DOM samples,[210] and FTICR MS and other high-resolution mass spectrometry techniques have been used to determine the molecular formula of these compounds. This approach has been used to identify and quantify chelating agents of biological origin in soil samples and the chemical form of heavy metal or halogenated contaminants in terrestrial environments.[211] Thus, FTICR MS combined with metal quantification provides insight into the processes that govern the transport, fate, and ecological effects of nutrients and contaminants in soils by providing unprecedented information on elemental speciation.

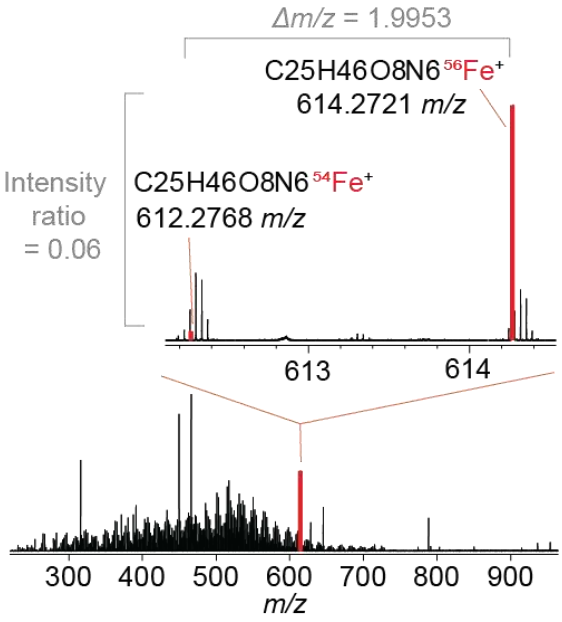


Fig. 7 Mass spectra in a water extract of SOM that depicts the isotope pattern for a siderophore-Fe complex, Ferrioxamine B. Inset displays the mass and abundance difference indicative of iron complexation. Unpublished data from the Boiteau Lab.

4.5 Fragmentation, Separation and Metabolomics

While FTICR MS analysis can attribute chemical formulas to molecules within complex SOM mixtures, identifying specific molecular structures such as metabolites or their transformation products requires pairing FTICR MS with chromatographic and fragmentation analyses. Liquid chromatographic separations prior to offline or online FTICR MS has several benefits. The separation of molecules that must otherwise compete for charge reduces the suppression of poorly ionizing molecules, reducing matrix effects and enabling the detection of a greater number of species.[213–215] In addition, separations can often resolve structural isomers that appear as a single feature with direct infusion FTICR MS (Fig. 9).[216] These structural variants may have distinct sources, functional groups, and fates in soils, and thus distinguishing between isomers by LC-FTICR MS could be important to decipher processes that form, preserve, and degrade SOM. The chromatographic separation modes generally selected for these analyses are aimed at

separating low molecular weight molecules within the *m/z* range of FTICR MS. Reverse phase liquid chromatography is typically applied to organic matter extracts, often mimicking the mode of retention used during solid phase extraction.[207,213,215–219] Hydrophilic interaction chromatography (HILIC), a variant of normal phase chromatography, is also well suited for the separation of polar components of SOM.[220,221] Both separation and ionization are strongly impacted by the choice of mobile phase. Methods commonly employ mass spectrometry grade water, alcohols, and acetonitrile along with volatile pH buffers (e.g., formic acid, acetic acid, ammonium formate, ammonium acetate, or ammonium hydroxide) due to their compatibility with ESI MS. The selectivity of different chromatographic methods for separating DOM has recently been reviewed by Sandon *et al.*[222]

Separation also facilitates the acquisition of tandem MS/MS fragmentation spectra for individual compounds by reducing the number of ions that may appear within a single MS isolation window. Fragment analysis has been used to determine the presence of specific functional groups or structures within DOM and other complex mixtures such as carboxylic acids,[215] sulphates,[223] aromatic archipelagos,[224,225] and metal-binding moieties.[210] MS/MS spectra provide a diagnostic molecular fingerprint that can be compared to the fragmentation spectra of known molecules to facilitate structural elucidation. Libraries of MS/MS fragmentation spectra collected with collision-induced dissociation are ever-growing,[226] and developing computational tools for predicting spectra of molecular structures *in silico* is an active field of research.[227–229] Longnecker *et al.*[230] coupled LC FTICR MS/MS with *in silico* fragmentation computational tools on *Thalassiosira pseudonana* to study novel, intracellular metabolites, finding that they were related through sulfoquinovosyl head groups. Another developing tool is feature-based molecular networking, which links MS/MS spectra by their spectra similarity, permitting the use of precursor

masses and structural information from compounds with library matches to help determine the structures of spectrally similar compounds with no library matches.[231,232] Putative identifications using tools such as *in silico* fragmentation or feature-based molecular networking can be confirmed by comparing retention time and MS/MS spectrum of an authentic compound using the same analytical method used for the soil sample.

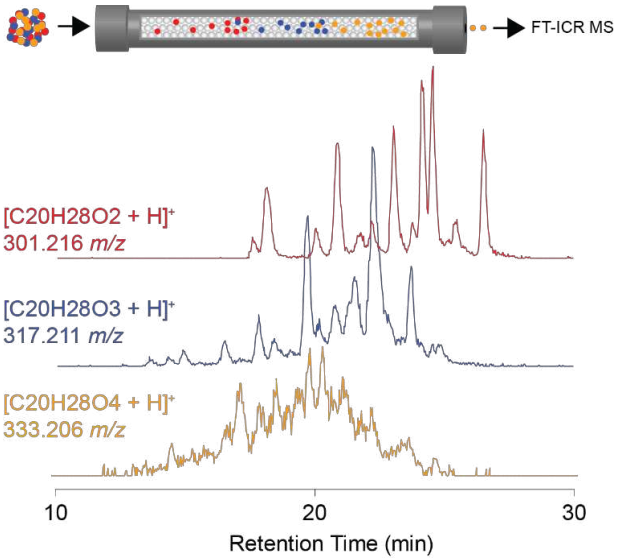


Fig. 8 Extracted ion chromatograms from an LC-FTICR MS (21T) of a soil water extract. Mass ranges about each *m/z* were ±0.005 Da, where each trace represents the protonated monoisotopic form of the molecular formula indicated. Peaks separated for each formula represent different isomers. Unpublished data from the Boiteau Lab.

Finally, while they play a significant role in microbial growth,[233–235] essential metabolites with a molecular weight <200 Da such as acetate, pyruvate, amino acids, glucose, fructose, and succinate are typically either outside the optimum mass range of FTICR MS (i.e., 150-1000 *m/z*) or harder to ionize amongst matrix effects.[96] Gas chromatography mass spectrometry (GC-MS),[236] nuclear magnetic resonance (NMR) spectroscopy,[237,238] and liquid chromatography mass spectrometry (LC-MS),[221] are the preferred instrumentations to identify these critical small metabolites and their structure in SOM. Recently, Swenson *et al.*[236] used GC-MS to estimate microbial metabolite availability in soils while Martins *et al.*[238] used ^{1}H NMR to measure acetate

and methanol metabolite concentrations using samples from the Prairie Pothole Wetlands in North America. Using LC-MS/MS in positive- and negative-ionization modes, Ladd *et al*.[221] were able to identify a wide range of compounds (***e.g***. amino acids, plant/microbial metabolites, sugars, lipids) present in artic soil DOM. Thus, combining multiple analytical techniques can help improve detection and annotation of central and secondary metabolites in SOM important for microbial biogeochemical cycling in soil ecosystems.[169]

Coupling mass spectrometry with microbial techniques (e.g., metaproteomics,[239] community composition profiling[238]) and novel bioinformatics approaches such as correlation networks and metabolite prediction tools[240–243] (Fig. S6) can also be useful for determining microbial decomposition pathways of OM and gaining insight into relationships between bacterial communities and SOM composition.[82,240,242,244–246] Using FTICR MS, ^{1}H NMR, and community composition, Martins *et al*.[238] found that extremely high sulfate reduction rates and methane fluxes in Prairie Pothole Wetlands were driven by abundant carbon substrates. Graham *et al*.[247] also used FTICR MS, gene expression, and community composition to gain mechanistic insight into hyporheic zone organic matter processing. This multi-omics coupled approach revealed a pronounced phenotypic plasticity in the hyporheic zone microbiome with similar microbiome structure, functional potential, and expression across sediments with dissimilar metabolic rates. Diverse nitrogenous metabolites and biochemical transformations as inferred by FTICR MS appeared to be the significant regulatory factor influencing hyporheic zone organic matter processing.

5. Future Directions

The strength of FTICR MS as an analytical tool lies in its unrivaled ability to detect individual organic molecules that make up the tremendously complex mixtures found in soils.[144]

Despite this, challenges associated with selectivity and the numerous equivocal peak assignments ultimately limits the processing and comparison of samples between studies. The utilization of complementary techniques that reinforce and constrain FTICR MS data are often needed to fully interpret the large and complex FTICR MS datasets. The use of spectroscopy provides highly complementary data on molecular origin or structure,[248] and is discussed in Section 2 of the supporting information.

Experimental design and sample preparation are critical steps because data are mostly qualitative and poorly reproducible across laboratories even on similar samples and within the same experiment on the same instrument.[120,152,249,250] Thus, it is important to ensure standardized protocol that minimizes influences on dependent and independent variables, where numerous biases and considerations are summarized (Table 1). Several future developments and needs for the application of FTICR MS to SOM are discussed below.

5.1 Unified Sample Preparation

The high chemical diversity of organic compounds in SOM means there is no universal technique for complete and comprehensive extraction, and the resultant analysis and discussion should reflect the limitations of the selected technique to avoid data misrepresentation. Comparisons between studies using different preparations are therefore limited as our understanding of how each preparation alters and isolates the SOM content is still developing.[95,97,98] Characterizing the effects to SOM during preparation would aid in normalizing SOM content against artifacts from preparation, supporting more effective comparisons between data sets. Therefore, a more thorough understanding of the biases introduced during preparation is needed to develop unified sample preparation techniques specific to a target fraction. Standardized

protocols for specific sample environments are critical in establishing comparable samples between laboratories and studies.

5.2 Automated Internal Calibration

While calibration strategies continue to become more refined in ultrahigh-resolution data, calibration remains dependent on instrument condition and is subject to user input which ultimately contributes to the poor reproducibility across laboratories.[152] Developing automated internal calibration strategies would therefore aid in the consistency and removal of bias during the analysis of FTICR MS samples, such as with the implementation of the ICBM-OCEAN[161] protocol, however procedures remain unique to different laboratories. Standardized SOM processing methods are expected to increase sample reproducibility and provide more opportunity for comparison between datasets.

5.3 Investigation of Ionization Techniques

Negative ESI remains the most commonly applied ionization polarity and technique for SOM, and other techniques, such as APPI, are capable of extending the range of ionizable features to less polar and aromatic functionalities. These techniques can be coupled to MS/MS systems for fragmentation, however collecting MS/MS spectra free from co-isolated precursor fragments has remained a major challenge when fragmenting complex mixtures. Charge acquiring leaving groups such as COO- are abundant during fragmentation of SOM and result in the loss of attributable daughter ions. Structurally informative fragmentation requires the investigation of new solution-based ionization techniques and methods that employ functionality specific ionization. To this end,

the development of metal cationization or solvent pH modification methods for specific precursor compounds may provide much higher specificity in the ionizable fraction.

5.4 Unified Database Repository

Due to the significant effort required to validate mass spectral features and the large number of features in soil that do not match any fragmentation spectra in databases, comprehensive identification of metabolites in soils remains a major challenge. Therefore, the implementation of a unified database repository using standardized protocol with elemental compositions for SOM datasets worldwide is of great interest for comparisons across temporal and spatial scales in environmental samples. Such repositories could be used to rapidly compare and identify peaks as potential metabolites by linking known and unknown compounds in correlation and network analysis, multi-omics techniques, and fragmentation spectra similarity in targeted analysis.[251] These advances would be critical in conducting statistics using larger datasets, identifying stabilization mechanisms and pathways, and in informing multi-omics approaches in SOM analysis.

Table 1. Table of common sources of error/bias during sample preparation and analysis for FTICR MS. Aspects to consider and recommendations are provided.

	Sources of Error/Bias	Considerations and Recommendations
Sample Collection and Preparation	○ Sample Uniformity	○ Ensure samples are well mixed when applicable; analyze pooled samples periodically throughout sample batches to monitor instrument stability and sensitivity.
	○ Contamination	○ Clean sampling equipment (e.g., combusted glass or medical grade polypropylene) prior to sampling
	○ Extraction Methods	○ Using *a priori* knowledge, prepare polar/non-polar/mineral-bound extracts separately to avoid biases in ionization. ○ Extraction and storage conditions (e.g., pH, temperature, volume of extractant, extraction time) should be consistent across sample sets
	○ Extraction Conditions	○ Artifacts (e.g., byproducts formed during extraction or storage) can be introduced from harsh conditions, conditions should be noted in following spectra and visualizations. ○ Solvent of choice should be compatible with FTICR MS.
Data Collection and Mass Spectrometry Acquisition Modes	○ Ionization Method Selection	○ Method and mode should target fraction selected during extraction. For bulk SOM, multiple modes provide the largest window into SOM composition. ○ Quality control strategies and samples (e.g., pooled quality control and standard samples run per analysis session)
	○ Sample matrix/instrument variability	○ S/N should be carefully considered based on the calibration performed (e.g., not set arbitrarily)
Data Pre-processing and Formula Assignment	○ Molecular formula assignment pipeline	○ Heteroatoms included during assignment should be chosen based on *a priori* information. ○ Exercise caution/manually assign and remove known artifacts (e.g., solvent clusters, extraction byproducts) and contaminants (e.g., surfactants, solvent contaminants) in data post-processing.
	○ Formula assignment quality	○ Assignments should be supervised with quality control strategies (e.g., error plots, continuity in calculated indices/heteroatoms)
Data Visualization	○ Data over-interpretation	○ Data visualization approaches will influence the outcome of the results (e.g., grouping by biomolecular class approximations) ○ Methods using assigned formulas (e.g., VK diagrams, NOSC vs. C) only approximate features/origin; approximations need support from *a priori* information, experimental design, and complementary analyses. ○ Formula assignment should not be assumed to represent bulk SOM; only a subset of SOM residues are measured. ○ Not all indices are appropriate for use in statistical technique.

6. AUTHOR CONTRIBUTIONS

William Bahureksa wrote the introduction with Merritt N. Logan, contributed to the "analysis/interpretation" and "future directions", provided data for figures, and assembled the manuscript. Merritt N. Logan wrote the introduction with William Bahureksa and wrote the entirety of the “Sampling and Extraction of Soil Organic Matter” section. René Boiteau wrote the "Ionization Techniques and Matrix Effects” section with Robert Young and contributed to "Future Directions" and provided feedback on "Analysis and Interpretation of Soil Organic Matter using FTICR MS" and contributed data to the “Analysis and Interpretation of Soil Organic Matter using FTICR MS” section. Malak M. Tfaily contributed to "Analysis and Interpretation of Soil Organic Matter using FTICR MS" and provided feedback on "Sampling and Extraction of Soil Organic Matter" and "Future Directions." Robert B. Young wrote the "Analysis and Interpretation of Soil Organic Matter using FTICR MS" section with Rene Boiteau, contributed to "Future Directions", and provided feedback on "Ionization Techniques and Matrix Effects." Amy M. McKenna contributed to "Ionization Techniques and Matrix Effects", " Analysis and Interpretation of Soil Organic Matter using FTICR MS", and "Future Directions " sections. Thomas Borch organized the topic for review and supervised the assembly of the manuscript.

7. ACKNOWLEDGMENTS

A portion of this work was performed at the National High Magnetic Field Laboratory, which is supported by the National Science Foundation Divisions of Chemistry and Materials Research through Cooperative Agreement No. DMR-1644779 and the state of Florida. A portion of the research was also performed using EMSL, a DOE Office of Science User Facility sponsored by the Biological and Environmental Research program. This research was supported, in part, by Grant No. 2018130 from the United States−Israel Binational Science Foundation (BSF), the

National Science Foundation under Grant No. (1512670), Department of Energy, Office of Science Biological and Environmental Research Grant, DE- SC0021349, and by SBR Project DE-SC0020205. We would like to acknowledge Dr. William Hockaday for their correspondence and usage of data for Figure 4. Any use of trade, product, or firm names in the publications is for descriptive purposes only and does not imply endorsement by the U.S. Government.

8. COPYRIGHT

9. REFERENCES

(1) Paustian, K.; Lehmann, J.; Ogle, S.; Reay, D.; Robertson, G. P.; Smith, P. Climate-Smart Soils. *Nature* **2016**, *532* (7597), 49–57. https://doi.org/10.1038/nature17174.

(2) Janzen, H. H. The Soil Carbon Dilemma: Shall We Hoard It or Use It? *Soil Biol. Biochem.* **2006**, *38* (3), 419–424. https://doi.org/10.1016/j.soilbio.2005.10.008.

(3) Schuur, E. A. G.; McGuire, A. D.; Schädel, C.; Grosse, G.; Harden, J. W.; Hayes, D. J.; Hugelius, G.; Koven, C. D.; Kuhry, P.; Lawrence, D. M.; et al. Climate Change and the Permafrost Carbon Feedback. *Nature* **2015**, *520* (7546), 171–179. https://doi.org/10.1038/nature14338.

(4) Stewart, C. E.; Paustian, K.; Conant, R. T.; Plante, A. F.; Six, J. Soil Carbon Saturation: Concept, Evidence and Evaluation. *Biogeochemistry* **2007**, *86* (1), 19–31. https://doi.org/10.1007/s10533-007-9140-0.

(5) Vives-Peris, V.; de Ollas, C.; Gómez-Cadenas, A.; Pérez-Clemente, R. M. Root Exudates: From Plant to Rhizosphere and Beyond. *Plant Cell Rep.* **2020**, *39* (1), 3–17. https://doi.org/10.1007/s00299-019-02447-5.

(6) Kögel-Knabner, I. The Macromolecular Organic Composition of Plant and Microbial Residues as Inputs to Soil Organic Matter: Fourteen Years On. *Soil Biology and Biochemistry*. 2017, pp A3–A8. https://doi.org/10.1016/j.soilbio.2016.08.011.

(7) Poirier, V.; Roumet, C.; Munson, A. D. The Root of the Matter: Linking Root Traits and Soil Organic Matter Stabilization Processes. **2018**. https://doi.org/10.1016/j.soilbio.2018.02.016.

(8) Liang, C.; Amelung, W.; Lehmann, J.; Kästner, M. Quantitative Assessment of Microbial Necromass Contribution to Soil Organic Matter. *Glob. Chang. Biol.* **2019**, *25* (11), 3578–3590. https://doi.org/10.1111/gcb.14781.

(9) Fernandez, C. W.; Kennedy, P. G. Commentary Moving beyond the Black-Box: Fungal Traits, Community Structure, and Carbon. *New Phytol.* **2015**, *205*, 1378–1380. https://doi.org/https://doi.org/10.1111/nph.13289.

(10) Liang, C.; Kästner, M.; Joergensen, R. G. Microbial Necromass on the Rise: The Growing Focus on Its Role in Soil Organic Matter Development. *Soil Biology and Biochemistry*. 2020. https://doi.org/10.1016/j.soilbio.2020.108000.

(11) Piccolo, A. The Supramolecular Structure of Humic Substances: A Novel Understanding of Humus Chemistry and Implications in Soil Science. *Adv. Agron.* **2002**, *75*, 57–134. https://doi.org/10.1016/s0065-2113(02)75003-7.

(12) Lehmann, J.; Hansel, C. M.; Kaiser, C.; Kleber, M.; Maher, K.; Manzoni, S.; Nunan, N.; Reichstein, M.; Schimel, J. P.; Torn, M. S.; et al. Persistence of Soil Organic Carbon Caused by Functional Complexity. *Nat. Geosci.* **2020**, *13* (8), 529–534. https://doi.org/10.1038/s41561-020-0612-3.

(13) Hertkorn, N.; Ruecker, C.; Meringer, M.; Gugisch, R.; Frommberger, M.; Perdue, E. M.; Witt, M.; Schmitt-Kopplin, P. High-Precision Frequency Measurements: Indispensable Tools at the Core of the Molecular-Level Analysis of Complex Systems. *Anal. Bioanal. Chem.* **2007**, *389* (5), 1311–1327. https://doi.org/10.1007/s00216-007-1577-4.

(14) Schmitt-Kopplin, P.; Hemmler, D.; Moritz, F.; Gougeon, R. D.; Lucio, M.; Meringer, M.; Müller, C.; Harir, M.; Hertkorn, N. Systems Chemical Analytics: Introduction to the Challenges of Chemical Complexity Analysis. *Faraday Discuss.* **2019**, *218*, 9–28. https://doi.org/10.1039/c9fd00078j.

(15) Lehmann, J.; Solomon, D.; Kinyangi, J.; Dathe, L.; Wirick, S.; Jacobsen, C. Spatial

Complexity of Soil Organic Matter Forms at Nanometre Scales. *Nat. Geosci.* **2008**, *1* (4), 238–242. https://doi.org/10.1038/ngeo155.
(16) Heister, K.; Höschen, C.; Pronk, G. J.; Mueller, C. W.; Kögel-Knabner, I. NanoSIMS as a Tool for Characterizing Soil Model Compounds and Organomineral Associations in Artificial Soils. *J. Soils Sediments* **2012**, *12* (1), 35–47. https://doi.org/10.1007/s11368-011-0386-8.
(17) Lavallee, J. M.; Soong, J. L.; Cotrufo, M. F. Conceptualizing Soil Organic Matter into Particulate and Mineral-Associated Forms to Address Global Change in the 21st Century. *Glob. Chang. Biol.* **2020**, *26* (1), 261–273. https://doi.org/10.1111/gcb.14859.
(18) Cotrufo, M. F.; Ranalli, M. G.; Haddix, M. L.; Six, J.; Lugato, E. Soil Carbon Storage Informed by Particulate and Mineral-Associated Organic Matter. *Nat. Geosci.* **2019**, *12* (12), 989–994. https://doi.org/10.1038/s41561-019-0484-6.
(19) Young, I. M.; Crawford, J. W. Interactions and Self-Organization in the Soil-Microbe Complex. *Science* **2004**, *304* (5677), 1634–1637. https://doi.org/10.1126/science.1097394.
(20) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17*, 1–35. https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K.
(21) Koch, B. P.; Dittmar, T.; Witt, M.; Kattner, G. Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. *Anal. Chem.* **2007**, *79* (4), 1758–1763. https://doi.org/10.1021/ac061949s.
(22) Reemtsma, T. Determination of Molecular Formulas of Natural Organic Matter Molecules by (Ultra-) High-Resolution Mass Spectrometry. Status and Needs. *J. Chromatogr. A* **2009**, *1216* (18), 3687–3701. https://doi.org/10.1016/j.chroma.2009.02.033.
(23) Zhang, X.; Han, J.; Zhang, X.; Shen, J.; Chen, Z.; Chu, W.; Kang, J.; Zhao, S.; Zhou, Y. Application of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry to Characterize Natural Organic Matter. *Chemosphere* **2020**, *260*, 127458. https://doi.org/10.1016/j.chemosphere.2020.127458.
(24) Qi, Y.; Fu, P.; Volmer, D. A. Analysis of Natural Organic Matter via Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: An Overview of Recent Non-Petroleum Applications. *Mass Spectrom. Rev.* **2020**, 1–15. https://doi.org/10.1002/mas.21634.
(25) Kleber, M.; Sollins, P.; Sutton, R. A Conceptual Model of Organo-Mineral Interactions in Soils: Self-Assembly of Organic Molecular Fragments into Zonal Structures on Mineral Surfaces. *Biogeochemistry* **2007**, *85* (1), 9–24. https://doi.org/10.1007/s10533-007-9103-5.
(26) Bailey, V. L.; Smith, A. P.; Tfaily, M.; Fansler, S. J.; Bond-Lamberty, B. Differences in Soluble Organic Carbon Chemistry in Pore Waters Sampled from Different Pore Size Domains. *Soil Biol. Biochem.* **2017**, *107*, 133–143. https://doi.org/10.1016/j.soilbio.2016.11.025.
(27) Swift, R. S. Organic Matter Characterization. *Methods of Soil Analysis*. January 1, 1996, pp 1011–1069. https://doi.org/https://doi.org/10.2136/sssabookser5.3.c35.
(28) Ikeya, K.; Sleighter, R. L.; Hatcher, P. G.; Watanabe, A. Chemical Compositional Analysis of Soil Fulvic Acids Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2020**, *34* (15), 1–11. https://doi.org/10.1002/rcm.8801.
(29) Ikeya, K.; Sleighter, R. L.; Hatcher, P. G.; Watanabe, A. Characterization of the Chemical Composition of Soil Humic Acids Using Fourier Transform Ion Cyclotron Resonance

Mass Spectrometry. *Geochim. Cosmochim. Acta* **2015**, *153*, 169–182. https://doi.org/10.1016/j.gca.2015.01.002.
(30) Coward, E. K.; Ohno, T.; Sparks, D. L. Direct Evidence for Temporal Molecular Fractionation of Dissolved Organic Matter at the Iron Oxyhydroxide Interface. *Environ. Sci. Technol.* **2019**, *53* (2), 642–650. https://doi.org/10.1021/acs.est.8b04687.
(31) Chen, M.; Kim, S.; Park, J. E.; Jung, H. J.; Hur, J. Structural and Compositional Changes of Dissolved Organic Matter upon Solid-Phase Extraction Tracked by Multiple Analytical Tools. *Anal. Bioanal. Chem.* **2016**, *408* (23), 6249–6258. https://doi.org/10.1007/s00216-016-9728-0.
(32) Kleber, M.; Eusterhues, K.; Keiluweit, M.; Mikutta, C.; Mikutta, R.; Nico, P. S. Mineral-Organic Associations: Formation, Properties, and Relevance in Soil Environments. *Adv. Agron.* **2015**, *130*, 1–140. https://doi.org/10.1016/bs.agron.2014.10.005.
(33) Schwesig, D.; Göttlein, A.; Haumaier, L.; Blasek, R.; Ilgen, G. Soil Organic Matter Extraction Using Water at High Temperature and Elevated Pressure (ASE) as Compared to Conventional Methods. *Int. J. Environ. Anal. Chem.* **1999**, *73* (4), 253–268. https://doi.org/10.1080/03067319908032668.
(34) Fox, P. M.; Nico, P. S.; Tfaily, M. M.; Heckman, K.; Davis, J. A. Characterization of Natural Organic Matter in Low-Carbon Sediments: Extraction and Analytical Approaches. *Org. Geochem.* **2017**, *114* (Supplement C), 12–22. https://doi.org/10.1016/j.orggeochem.2017.08.009.
(35) Chantigny, M. H. Dissolved and Water-Extractable Organic Matter in Soils: A Review on the Influence of Land Use and Management Practices. *Geoderma* **2003**, *113* (3–4), 357–380. https://doi.org/10.1016/S0016-7061(02)00370-1.
(36) Melendez-Perez, J. J.; Martínez-Mejía, M. J.; Awan, A. T.; Fadini, P. S.; Mozeto, A. A.; Eberlin, M. N. Characterization and Comparison of Riverine, Lacustrine, Marine and Estuarine Dissolved Organic Matter by Ultra-High Resolution and Accuracy Fourier Transform Mass Spectrometry. *Org. Geochem.* **2016**, *101*, 99–107. https://doi.org/10.1016/j.orggeochem.2016.08.005.
(37) Koch, B. P.; Witt, M.; Engbrodt, R.; Dittmar, T.; Kattner, G. Molecular Formulae of Marine and Terrigenous Dissolved Organic Matter Detected by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Geochim. Cosmochim. Acta* **2005**, *69* (13), 3299–3308. https://doi.org/10.1016/j.gca.2005.02.027.
(38) Schmidt, M. W. I. I.; Torn, M. S.; Abiven, S.; Dittmar, T.; Guggenberger, G.; Janssens, I. A.; Kleber, M.; Kögel-Knabner, I.; Lehmann, J.; Manning, D. A. C. C.; et al. Persistence of Soil Organic Matter as an Ecosystem Property. *Nature* **2011**, *478* (7367), 49–56. https://doi.org/10.1038/nature10386.
(39) Kellerman, A. M.; Kothawala, D. N.; Dittmar, T.; Tranvik, L. J. Persistence of Dissolved Organic Matter in Lakes Related to Its Molecular Characteristics. *Nat. Geosci.* **2015**, *8* (6), 454–457. https://doi.org/10.1038/NGEO2440.
(40) Kellerman, A. M.; Guillemette, F.; Podgorski, D. C.; Aiken, G. R.; Butler, K. D.; Spencer, R. G. M. Unifying Concepts Linking Dissolved Organic Matter Composition to Persistence in Aquatic Ecosystems. *Environ. Sci. Technol.* **2018**, *52* (5), 2538–2548. https://doi.org/10.1021/acs.est.7b05513.
(41) Tfaily, M. M.; Chu, R. K.; Toyoda, J.; Tolić, N.; Robinson, E. W.; Paša-Tolić, L.; Hess, N. J. Sequential Extraction Protocol for Organic Matter from Soils and Sediments Using High Resolution Mass Spectrometry. *Anal. Chim. Acta* **2017**, *972*, 54–61.

https://doi.org/10.1016/j.aca.2017.03.031.
(42) Casals, P.; Lopez-Sangil, L.; Carrara, A.; Gimeno, C.; Nogués, S. Autotrophic and Heterotrophic Contributions to Short-Term Soil CO 2 Efflux Following Simulated Summer Precipitation Pulses in a Mediterranean Dehesa. *Global Biogeochem. Cycles* **2011**, *25* (3). https://doi.org/10.1029/2010GB003973.
(43) Casals, P.; Gimeno, C.; Carrara, A.; Lopez-Sangil, L.; Sanz, M. J. Soil CO2 Efflux and Extractable Organic Carbon Fractions under Simulated Precipitation Events in a Mediterranean Dehesa. *Soil Biol. Biochem.* **2009**, *41* (9), 1915–1922. https://doi.org/10.1016/j.soilbio.2009.06.015.
(44) Coward, E. K.; Ohno, T.; Plante, A. F. Adsorption and Molecular Fractionation of Dissolved Organic Matter on Iron-Bearing Mineral Matrices of Varying Crystallinity. *Environ. Sci. Technol.* **2018**, *52* (3), 1036–1044. https://doi.org/10.1021/acs.est.7b04953.
(45) Guigue, J.; Mathieu, O.; Lévêque, J.; Mounier, S.; Laffont, R.; Maron, P. A.; Navarro, N.; Chateau, C.; Amiotte-Suchet, P.; Lucas, Y. A Comparison of Extraction Procedures for Water-Extractable Organic Matter in Soils. *Eur. J. Soil Sci.* **2014**, *65* (4), 520–530. https://doi.org/10.1111/ejss.12156.
(46) Nebbioso, A.; Piccolo, A. Molecular Characterization of Dissolved Organic Matter (DOM): A Critical Review. *Analytical and Bioanalytical Chemistry*. Springer January 11, 2013, pp 109–124. https://doi.org/10.1007/s00216-012-6363-2.
(47) Zafiriou, O. C.; Joussot-Dubien, J.; Zepp, R. G.; Zika, R. G. Photochemistry of Natural Waters: Many Compounds and Environments Are Affected by Sunlight-Induced Photochemistry. *Environ. Sci. Technol.* **1984**, *18* (12), 358A-371A. https://doi.org/10.1021/es00130a711.
(48) Fisher, S. G.; Likens, G. E. Energy Flow in Bear Brook, New Hampshire: An Integrative Approach to Stream Ecosystem Metabolism. *Ecol. Monogr.* **1973**, *43* (4), 421–439. https://doi.org/10.2307/1942301.
(49) STEWART, A. J.; WETZEL, R. G. DISSOLVED HUMIC MATERIALS - PHOTODEGRADATION, SEDIMENT EFFECTS, AND REACTIVITY WITH PHOSPHATE AND CALCIUM-CARBONATE PRECIPITATION. *Arch. FUR Hydrobiol.* **1981**, *92* (3), 265–286.
(50) Lopez-Sangil, L.; Rovira, P. Sequential Chemical Extractions of the Mineral-Associated Soil Organic Matter: An Integrated Approach for the Fractionation of Organo-Mineral Complexes. *Soil Biol. Biochem.* **2013**, *62*, 57–67. https://doi.org/10.1016/j.soilbio.2013.03.004.
(51) Jones, D. L.; Willett, V. B. Experimental Evaluation of Methods to Quantify Dissolved Organic Nitrogen (DON) and Dissolved Organic Carbon (DOC) in Soil. *Soil Biol. Biochem.* **2006**, *38* (5), 991–999. https://doi.org/10.1016/j.soilbio.2005.08.012.
(52) Mann, B. F.; Chen, H.; Herndon, E. M.; Chu, R. K.; Tolic, N.; Portier, E. F.; Chowdhury, T. R.; Robinson, E. W.; Callister, S. J.; Wullschleger, S. D.; et al. Indexing Permafrost Soil Organic Matter Degradation Using High-Resolution Mass Spectrometry. *PLoS One* **2015**, *10* (6). https://doi.org/10.1371/journal.pone.0130557.
(53) Ryan, M. G.; Melillo, J. M.; Ricca, A. A Comparison of Methods for Determining Proximate Carbon Fractions of Forest Litter. *Can. J. For. Res.* **1990**, *20* (2), 166–171. https://doi.org/10.1139/x90-023.
(54) Achard, F. K. Chemische Untersuchung Des Torfs. *Crell's Chem. Ann* **1786**, *2*, 391–403.
(55) Hayes, M. H. B. Solvent Systems for the Isolation of Organic Components from Soils.

Soil Sci. Soc. Am. J. **2006**, *70* (3), 986–994. https://doi.org/10.2136/sssaj2005.0107.
(56) Schnitzer, M.; Monreal, C. Chapter Three. Quo Vadis Soil Organic Matter Research? A Biological Link to the Chemistry of Humification. *Adv. Agron.* **2011**, *113*, 143–217. https://doi.org/10.1016/B978-0-12-386473-4.00003-8.
(57) Stevenson, F. J. *Humus Chemistry: Genesis, Composition, Reactions*; John Wiley & Sons, 1994.
(58) Kleber, M.; Lehmann, J. Humic Substances Extracted by Alkali Are Invalid Proxies for the Dynamics and Functions of Organic Matter in Terrestrial and Aquatic Ecosystems. *J. Environ. Qual.* **2019**, *48* (2), 207–216. https://doi.org/10.2134/jeq2019.01.0036.
(59) Olk, D. C.; Bloom, P. R.; Perdue, E. M.; McKnight, D. M.; Chen, Y.; Farenhorst, A.; Senesi, N.; Chin, Y.-P.; Schmitt-Kopplin, P.; Hertkorn, N.; et al. Environmental and Agricultural Relevance of Humic Fractions Extracted by Alkali from Soils and Natural Waters. *J. Environ. Qual.* **2019**, *48* (2), 217–232. https://doi.org/10.2134/jeq2019.02.0041.
(60) DiDonato, N.; Hatcher, P. G. Alicyclic Carboxylic Acids in Soil Humic Acid as Detected with Ultrahigh Resolution Mass Spectrometry and Multi-Dimensional NMR. *Org. Geochem.* **2017**, *112*, 33–46. https://doi.org/10.1016/j.orggeochem.2017.06.010.
(61) DiDonato, N.; Chen, H.; Waggoner, D.; Hatcher, P. G. Potential Origin and Formation for Molecular Components of Humic Acids in Soils. *Geochim. Cosmochim. Acta* **2016**, *178*, 210–222. https://doi.org/10.1016/j.gca.2016.01.013.
(62) Kelleher, B. P.; Simpson, A. J. Humic Substances in Soils: Are They Really Chemically Distinct? *Environ. Sci. Technol.* **2006**, *40* (15), 4605–4611. https://doi.org/10.1021/es0608085.
(63) Masoom, H.; Courtier-Murias, D.; Farooq, H.; Soong, R.; Kelleher, B. P.; Zhang, C.; Maas, W. E.; Fey, M.; Kumar, R.; Monette, M.; et al. Soil Organic Matter in Its Native State: Unravelling the Most Complex Biomaterial on Earth. *Environ. Sci. Technol.* **2016**, *50* (4), 1670–1680. https://doi.org/10.1021/acs.est.5b03410.
(64) Wagner, S.; Jaffé, R.; Stubbins, A. Dissolved Black Carbon in Aquatic Ecosystems. *Limnol. Oceanogr. Lett.* **2018**. https://doi.org/10.1002/lol2.10076.
(65) Lehmann, J.; Kleber, M. The Contentious Nature of Soil Organic Matter. *Nature* **2015**, *528* (7580), 0–8. https://doi.org/10.1038/nature16069.
(66) Kogel-Knabner, I. The Macromolecular Organic Composition of Plant and Microbial Residues as Inputs to Soil Organic Matter. *Soil Biol. Biochem.* **2002**, *105*, A3–A8. https://doi.org/10.1016/j.soilbio.2016.08.011.
(67) Coward, E. K.; Thompson, A. T.; Plante, A. F. Iron-Mediated Mineralogical Control of Organic Matter Accumulation in Tropical Soils. *Geoderma* **2017**, *306*, 206–216. https://doi.org/10.1016/j.geoderma.2017.07.026.
(68) Lalonde, K.; Mucci, A.; Ouellet, A.; Gélinas, Y. Preservation of Organic Matter in Sediments Promoted by Iron. *Nature* **2012**, *483* (7388), 198–200. https://doi.org/10.1038/nature10855.
(69) Gu, B.; Schmitt, J.; Chen, Z.; Liang, L.; McCarthy, J. F. Adsorption and Desorption of Natural Organic Matter on Iron Oxide: Mechanisms and Models. *Environ. Sci. Technol.* **1994**, *28* (1), 38–46. https://doi.org/10.1021/es00050a007.
(70) Mehra, O. P. Iron Oxide Removal from Soils and Clays by a Dithionite-Citrate System Buffered with Sodium Bicarbonate. *Clays Clay Miner.* **1958**, *7* (1), 317–327. https://doi.org/10.1346/ccmn.1958.0070122.
(71) Wagai, R.; Mayer, L. M. Sorptive Stabilization of Organic Matter in Soils by Hydrous

Iron Oxides. *Geochim. Cosmochim. Acta* **2007**, *71* (1), 25–35. https://doi.org/10.1016/j.gca.2006.08.047.
(72) Chao, T. T.; Zhou, L. Extraction Techniques for Selective Dissolution of Amorphous Iron Oxides from Soils and Sediments. *Soil Sci. Soc. Am. J.* **1983**, *47* (2), 225–232. https://doi.org/10.2136/sssaj1983.03615995004700020010x.
(73) von Lützow, M.; Kögel-Knabner, I.; Ekschmitt, K.; Flessa, H.; Guggenberger, G.; Matzner, E.; Marschner, B. SOM Fractionation Methods: Relevance to Functional Pools and to Stabilization Mechanisms. *Soil Biol. Biochem.* **2007**, *39* (9), 2183–2207. https://doi.org/10.1016/j.soilbio.2007.03.007.
(74) Keil, R. G.; Mayer, L. M. Mineral Matrices and Organic Matter. In *Treatise on Geochemistry: Second Edition*; Elsevier Inc., 2013; Vol. 12, pp 337–359. https://doi.org/10.1016/B978-0-08-095975-7.01024-X.
(75) Wagai, R.; Mayer, L. M.; Kitayama, K.; Shirato, Y. Association of Organic Matter with Iron and Aluminum across a Range of Soils Determined via Selective Dissolution Techniques Coupled with Dissolved Nitrogen Analysis. *Biogeochemistry* **2013**, *112* (1–3), 95–109. https://doi.org/10.1007/s10533-011-9652-5.
(76) Patzner, M. S.; Mueller, C. W.; Malusova, M.; Baur, M.; Nikeleit, V.; Scholten, T.; Hoeschen, C.; Byrne, J. M.; Borch, T.; Kappler, A.; et al. Iron Mineral Dissolution Releases Iron and Associated Organic Carbon during Permafrost Thaw. *Nat. Commun.* **2020**, *11* (1), 1–11. https://doi.org/10.1038/s41467-020-20102-6.
(77) Lv, J.; Huang, Z.; Christie, P.; Zhang, S. Reducing Reagents Induce Molecular Artifacts in the Extraction of Soil Organic Matter. *ACS Earth Sp. Chem.* **2020**, acsearthspacechem.0c00194. https://doi.org/10.1021/acsearthspacechem.0c00194.
(78) Fisher, B. J.; Faust, J. C.; Moore, O. W.; Peacock, C. L.; März, C. Technical Note: Uncovering the Influence of Methodological Variations on the Extractability of Iron Bound Organic Carbon. *Biogeosciences Discuss.* **2020**, *2020*, 1–23. https://doi.org/10.5194/bg-2020-399.
(79) Daugherty, E. E.; Gilbert, B.; Nico, P. S.; Borch, T. Complexation and Redox Buffering of Iron(II) by Dissolved Organic Matter. *Environ. Sci. Technol.* **2017**, *51* (19), 11096–11104. https://doi.org/10.1021/acs.est.7b03152.
(80) Chen, C.; Thompson, A. The Influence of Native Soil Organic Matter and Minerals on Ferrous Iron Oxidation. *Geochim. Cosmochim. Acta* **2021**, *292*, 254–270. https://doi.org/10.1016/j.gca.2020.10.002.
(81) Young, R.; Avneri-Katz, S.; McKenna, A.; Chen, H.; Bahureksa, W.; Polubesova, T.; Chefetz, B.; Borch, T. Composition-Dependent Sorptive Fractionation of Anthropogenic Dissolved Organic Matter by Fe(III)-Montmorillonite. *Soil Syst.* **2018**, *2* (1), 14. https://doi.org/10.3390/soilsystems2010014.
(82) Xu, J.; Roley, S. S.; Tfaily, M. M.; Chu, R. K.; Tiedje, J. M. Organic Amendments Change Soil Organic C Structure and Microbial Community but Not Total Organic Matter on Sub-Decadal Scales. *Soil Biol. Biochem.* **2020**, *150* (August), 107986. https://doi.org/10.1016/j.soilbio.2020.107986.
(83) Tfaily, M. M.; Chu, R. K.; Tolić, N.; Roscioli, K. M.; Anderton, C. R.; Paša-Tolić, L.; Robinson, E. W.; Hess, N. J. Advanced Solvent Based Methods for Molecular Characterization of Soil Organic Matter by High-Resolution Mass Spectrometry. *Anal. Chem.* **2015**, *87* (10), 5206–5215. https://doi.org/10.1021/acs.analchem.5b00116.
(84) McKee, G. A.; Hatcher, P. G. A New Approach for Molecular Characterisation of

Sediments with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Extraction Optimisation. *Org. Geochem.* **2015**, *85*, 22–31. https://doi.org/10.1016/j.orggeochem.2015.04.007.
(85) Flerus, R.; Koch, B. P.; Schmitt-Kopplin, P.; Witt, M.; Kattner, G. Molecular Level Investigation of Reactions between Dissolved Organic Matter and Extraction Solvents Using FT-ICR MS. *Mar. Chem.* **2011**, *124* (1–4), 100–107. https://doi.org/10.1016/j.marchem.2010.12.006.
(86) Schwab, K. Effect-Directed Identification of Bioavailable Toxic Organic Compounds in Contaminated Sediments. **2008**.
(87) Raiswell, R.; Canfield, D. E.; Berner, R. A. A Comparison of Iron Extraction Methods for the Determination of Degree of Pyritisation and the Recognition of Iron-Limited Pyrite Formation. *Chem. Geol.* **1994**, *111* (1–4), 101–110. https://doi.org/10.1016/0009-2541(94)90084-1.
(88) Bartlett, R.; James, B. Studying Dried, Stored Soil Samples - Some Pitfalls. *Soil Sci. Soc. Am. J.* **1980**, *44* (4), 721–724. https://doi.org/10.2136/sssaj1980.03615995004400040011x.
(89) Heinz, M.; Zak, D. Storage Effects on Quantity and Composition of Dissolved Organic Carbon and Nitrogen of Lake Water, Leaf Leachate and Peat Soil Water. *Water Res.* **2018**, *130*, 98–104. https://doi.org/10.1016/j.watres.2017.11.053.
(90) Tfaily, M. M.; Podgorski, D. C.; Corbett, J. E.; Chanton, J. P.; Cooper, W. T. Influence of Acidification on the Optical Properties and Molecular Composition of Dissolved Organic Matter. *Anal. Chim. Acta* **2011**, *706* (2), 261–267. https://doi.org/10.1016/j.aca.2011.08.037.
(91) Retelletti Brogi, S.; Derrien, M.; Hur, J. In-Depth Assessment of the Effect of Sodium Azide on the Optical Properties of Dissolved Organic Matter. *J. Fluoresc.* **2019**, *29* (4), 877–885. https://doi.org/10.1007/s10895-019-02398-w.
(92) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A Simple and Efficient Method for the Solid-Phase Extraction of Dissolved Organic Matter (SPE-DOM) from Seawater. *Limnol. Oceanogr. Methods* **2008**, *6* (6), 230–235. https://doi.org/10.4319/lom.2008.6.230.
(93) Li, Y.; Harir, M.; Uhl, J.; Kanawati, B.; Lucio, M.; Smirnov, K. S.; Koch, B. P.; Schmitt-Kopplin, P.; Hertkorn, N. How Representative Are Dissolved Organic Matter (DOM) Extracts? A Comprehensive Study of Sorbent Selectivity for DOM Isolation. *Water Res.* **2017**, *116*, 316–323. https://doi.org/10.1016/j.watres.2017.03.038.
(94) Lam, B.; Baer, A.; Alaee, M.; Lefebvre, B.; Moser, A.; Williams, A.; Simpson, A. J. Major Structural Components in Freshwater Dissolved Organic Matter. *Environ. Sci. Technol.* **2007**, *41* (24), 8240–8247. https://doi.org/10.1021/es0713072.
(95) Stücheli, P. E.; Niggemann, J.; Schubert, C. J. Comparison of Different Solid Phase Extraction Sorbents for the Qualitative Assessment of Dissolved Organic Nitrogen in Freshwater Samples Using FT-ICR-MS. *J. Limnol.* **2018**, *77* (3), 400–411. https://doi.org/10.4081/jlimnol.2018.1791.
(96) Wang, C.; He, L.; Li, D. W.; Bruschweiler-Li, L.; Marshall, A. G.; Brüschweiler, R. Accurate Identification of Unknown and Known Metabolic Mixture Components by Combining 3D NMR with Fourier Transform Ion Cyclotron Resonance Tandem Mass Spectrometry. *J. Proteome Res.* **2017**, *16* (10), 3774–3786. https://doi.org/10.1021/acs.jproteome.7b00457.
(97) McAdams, B. C.; Aiken, G. R.; McKnight, D. M.; Arnold, W. A.; Chin, Y. P. High

Pressure Size Exclusion Chromatography (HPSEC) Determination of Dissolved Organic Matter Molecular Weight Revisited: Accounting for Changes in Stationary Phases, Analytical Standards, and Isolation Methods. *Environ. Sci. Technol.* **2018**, *52* (2), 722–730. https://doi.org/10.1021/acs.est.7b04401.

(98) Green, N. W.; Perdue, E. M.; Aiken, G. R.; Butler, K. D.; Chen, H.; Dittmar, T.; Niggemann, J.; Stubbins, A. An Intercomparison of Three Methods for the Large-Scale Isolation of Oceanic Dissolved Organic Matter. *Mar. Chem.* **2014**, *161*, 14–19. https://doi.org/10.1016/j.marchem.2014.01.012.

(99) Lv, J.; Zhang, S.; Luo, L.; Cao, D. Solid-Phase Extraction-Stepwise Elution (SPE-SE) Procedure for Isolation of Dissolved Organic Matter Prior to ESI-FT-ICR-MS Analysis. *Anal. Chim. Acta* **2016**, *948*, 55–61. https://doi.org/10.1016/j.aca.2016.10.038.

(100) Raeke, J.; Lechtenfeld, O. J.; Wagner, M.; Herzsprung, P.; Reemtsma, T. Selectivity of Solid Phase Extraction of Freshwater Dissolved Organic Matter and Its Effect on Ultrahigh Resolution Mass Spectra. *Environ. Sci. Process. Impacts* **2016**, *18* (7), 918–927. https://doi.org/10.1039/C6EM00200E.

(101) Yamamura, H.; Okimoto, K.; Kimura, K.; Watanabe, Y. Hydrophilic Fraction of Natural Organic Matter Causing Irreversible Fouling of Microfiltration and Ultrafiltration Membranes. *Water Res.* **2014**, *54*, 123–136. https://doi.org/10.1016/j.watres.2014.01.024.

(102) Zularisam, A. W.; Ahmad, A.; Sakinah, M.; Ismail, A. F.; Matsuura, T. Role of Natural Organic Matter (NOM), Colloidal Particles, and Solution Chemistry on Ultrafiltration Performance. *Sep. Purif. Technol.* **2011**, *78* (2), 189–200. https://doi.org/10.1016/j.seppur.2011.02.001.

(103) Sim, L. N.; Chong, T. H.; Taheri, A. H.; Sim, S. T. V.; Lai, L.; Krantz, W. B.; Fane, A. G. A Review of Fouling Indices and Monitoring Techniques for Reverse Osmosis. *Desalination* **2018**, *434* (December 2017), 169–188. https://doi.org/10.1016/j.desal.2017.12.009.

(104) Vetter, T. A.; Perdue, E. M.; Ingall, E.; Koprivnjak, J. F.; Pfromm, P. H. Combining Reverse Osmosis and Electrodialysis for More Complete Recovery of Dissolved Organic Matter from Seawater. *Sep. Purif. Technol.* **2007**, *56* (3), 383–387. https://doi.org/10.1016/j.seppur.2007.04.012.

(105) Koprivnjak, J. F.; Perdue, E. M.; Pfromm, P. H. Coupling Reverse Osmosis with Electrodialysis to Isolate Natural Organic Matter from Fresh Waters. *Water Res.* **2006**, *40* (18), 3385–3392. https://doi.org/10.1016/j.watres.2006.07.019.

(106) D'Andrilli, J.; Dittmar, T.; Koch, B. P.; Purcell, J. M.; Marshall, A. G.; Cooper, W. T.; Andrilli, J. D.; Dittmar, T.; Koch, B. P.; Purcell, J. M.; et al. 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.* **2010**, *24* (24), 3567–3577. https://doi.org/10.1002/rcm.

(107) Hockaday, W. C.; Purcell, J. M.; Marshall, A. G.; Baldock, J. A.; Hatcher, P. G. Electrospray and Photoionization Mass Spectrometry for the Characterization of Organic Matter in Natural Waters: A Qualitative Assessment. *Limnol. Oceanogr. Methods* **2009**, *7* (JAN.), 81–95. https://doi.org/10.4319/lom.2009.7.81.

(108) Ohno, T.; Sleighter, R. L.; Hatcher, P. G. Comparative Study of Organic Matter Chemical Characterization Using Negative and Positive Mode Electrospray Ionization Ultrahigh-Resolution Mass Spectrometry. *Anal. Bioanal. Chem.* **2016**, *408* (10), 2497–2504.

https://doi.org/10.1007/s00216-016-9346-x.
(109) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization-Principles and Practice. *Mass Spectrom. Rev.* **1990**, *9* (1), 37–70. https://doi.org/10.1002/mas.1280090103.
(110) Xu, Y. F.; Lu, W.; Rabinowitz, J. D. Avoiding Misannotation of In-Source Fragmentation Products as Cellular Metabolites in Liquid Chromatography-Mass Spectrometry-Based Metabolomics. *Anal. Chem.* **2015**, *87* (4), 2273–2281. https://doi.org/10.1021/ac504118y.
(111) Gathungu, R. M.; Larrea, P.; Sniatynski, M. J.; Marur, V. R.; Bowden, J. A.; Koelmel, J. P.; Starke-Reed, P.; Hubbard, V. S.; Kristal, B. S. Optimization of Electrospray Ionization Source Parameters for Lipidomics to Reduce Misannotation of In-Source Fragments as Precursor Ions. *Anal. Chem.* **2018**, *90* (22), 13523–13532. https://doi.org/10.1021/acs.analchem.8b03436.
(112) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2–9. https://doi.org/10.1021/ac302789c.
(113) Kruve, A.; Kaupmees, K.; Liigand, J.; Leito, I. Negative Electrospray Ionization via Deprotonation: Predicting the Ionization Efficiency. *Anal. Chem* **2014**, *86*, 57. https://doi.org/10.1021/ac404066v.
(114) Oss, M.; Kruve, A.; Herodes, K.; Leito, I. Electrospray Ionization Efficiency Scale of Organic Compound. *Anal. Chem.* **2010**, *82* (7), 2865–2872. https://doi.org/10.1021/ac902856t.
(115) Liigand, J.; Kruve, A.; Leito, I.; Girod, M.; Antoine, R. Effect of Mobile Phase on Electrospray Ionization Efficiency. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (11), 1853–1861. https://doi.org/10.1007/s13361-014-0969-x.
(116) Kostiainen, R.; Kauppila, T. J. Effect of Eluent on the Ionization Process in Liquid Chromatography-Mass Spectrometry. *J. Chromatogr. A* **2009**, *1216* (4), 685–699. https://doi.org/10.1016/j.chroma.2008.08.095.
(117) Liigand, J.; Laaniste, A.; Kruve, A. PH Effects on Electrospray Ionization Efficiency. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (3), 461–469. https://doi.org/10.1007/s13361-016-1563-1.
(118) Zhou, S.; Cook, K. D. Protonation in Electrospray Mass Spectrometry: Wrong-Way-Round or Right-Way-Round? *J. Am. Soc. Mass Spectrom.* **2000**, *11* (11), 961–966. https://doi.org/10.1016/S1044-0305(00)00174-4.
(119) Novotny, N. R.; Capley, E. N.; Stenson, A. C. Fact or Artifact: The Representativeness of ESI-MS for Complex Natural Organic Mixtures. *J. Mass Spectrom.* **2014**, *49* (4), 316–326. https://doi.org/10.1002/jms.3345.
(120) Kido Soule, M. C.; Longnecker, K.; Giovannoni, S. J.; Kujawinski, E. B. Impact of Instrument and Experiment Parameters on Reproducibility of Ultrahigh Resolution ESI FT-ICR Mass Spectra of Natural Organic Matter. *Org. Geochem.* **2010**, *41* (8), 725–733. https://doi.org/10.1016/j.orggeochem.2010.05.017.
(121) Lobodin, V. V; Juyal, P.; Mckenna, A. M.; Rodgers, R. P.; Marshall, A. G. Tetramethylammonium Hydroxide as a Reagent for Complex Mixture Analysis by Negative Ion Electrospray Ionization Mass Spectrometry. **2013**.
(122) Lobodin, V. V.; Juyal, P.; McKenna, A. M.; Rodgers, R. P.; Marshall, A. G. Lithium Cationization for Petroleum Analysis by Positive Ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Energy & Fuels* **2014**, *28* (11),

6841–6847. https://doi.org/10.1021/ef501683w.
(123) Lobodin, V. V.; Juyal, P.; McKenna, A. M.; Rodgers, R. P.; Marshall, A. G. Silver Cationization for Rapid Speciation of Sulfur-Containing Species in Crude Oils by Positive Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Energy and Fuels* **2014**, *28* (1), 447–452. https://doi.org/10.1021/ef401897p.
(124) El-Faramawy, A.; Siu, K. W. M.; Thomson, B. A. Efficiency of Nano-Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (10), 1702–1707. https://doi.org/10.1016/j.jasms.2005.06.011.
(125) Marshall, A. G.; Chen, T. 40 Years of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Int. J. Mass Spectrom.* **2015**, *377* (1), 410–420. https://doi.org/10.1016/j.ijms.2014.06.034.
(126) Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M.; Hertkorn, N., Frommberger, M., Witt, M., Koch, B.P., Schmitt-Kopplin, P., P.; E.M. Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Anal. Chem.* **2008**, *80* (23), 8908–8919. https://doi.org/10.1021/ac800464g.
(127) Podgorski, D. C.; McKenna, A. M.; Rodgers, R. P.; Marshall, A. G.; Cooper, W. T. Selective Ionization of Dissolved Organic Nitrogen by Positive Ion Atmospheric Pressure Photoionization Coupled with Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2012**, *84* (11), 5085–5090. https://doi.org/10.1021/ac300800w.
(128) Kauppila, T. J.; Syage, J. A.; Benter, T. Recent Developments in Atmospheric Pressure Photoionization-Mass Spectrometry. *Mass Spectrom. Rev.* **2015**, *36* (3), 423–449. https://doi.org/10.1002/mas.21477.
(129) Mopper, K.; Stubbins, A.; Ritchie, J. D.; Bialk, H. M.; Hatcher, P. G. Advanced Instrumental Approaches for Characterization of Marine Dissolved Organic Matter: Extraction Techniques, Mass Spectrometry, and Nuclear Magnetic Resonance Spectroscopy. *Chem. Rev.* **2007**, *107* (2), 419–442. https://doi.org/10.1021/cr050359b.
(130) Cai, S.-S.; Syage, J. a. Atmospheric Pressure Photoionization Mass Spectrometry for Analysis of Fatty Acid and Acylglycerol Lipids. *J. Chromatogr. A* **2006**, *1110*, 15–26. https://doi.org/10.1016/j.chroma.2006.01.050.
(131) Podgorski, D. C.; Hamdan, R.; McKenna, A. M.; Nyadong, L.; Rodgers, R. P.; Marshall, A. G.; Cooper, W. T. Characterization of Pyrogenic Black Carbon by Desorption Atmospheric Pressure Photoionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2011**, *84* (3), 1281–1287. https://doi.org/10.1021/ac202166x.
(132) Hagemann, N.; Joseph, S.; Schmidt, H. P.; Kammann, C. I.; Harter, J.; Borch, T.; Young, R. B.; Varga, K.; Taherymoosavi, S.; Elliott, K. W.; et al. Organic Coating on Biochar Explains Its Nutrient Retention and Stimulation of Soil Fertility. *Nat. Commun.* **2017**, *8* (1), 1–11. https://doi.org/10.1038/s41467-017-01123-0.
(133) Cao, D.; Huang, H.; Hu, M.; Cui, L.; Geng, F.; Rao, Z.; Niu, H.; Cai, Y.; Kang, Y. Comprehensive Characterization of Natural Organic Matter by MALDI- and ESI-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chim. Acta* **2015**, *866*, 48–58. https://doi.org/10.1016/j.aca.2015.01.051.
(134) Solihat, N. N.; Acter, T.; Kim, D.; Plante, A. F.; Kim, S. Analyzing Solid-Phase Natural Organic Matter Using Laser Desorption Ionization Ultrahigh Resolution Mass Spectrometry. *Anal. Chem.* **2019**, *91* (1), 951–957. https://doi.org/10.1021/acs.analchem.8b04032.
(135) Zenobi, R.; Knochenmuss, R. Ion Formation in Maldi Mass Spectrometry. *Mass*

Spectrom. Rev. **1998**, *17* (5), 337–366. https://doi.org/10.1002/(sici)1098-2787(1998)17:5<337::aid-mas2>3.0.co;2-s.
(136) Dreisewerd, K. The Desorption Process in MALDI. *Chemical Reviews*. 2003, pp 395–425. https://doi.org/10.1021/cr010375i.
(137) Kujawinski, E. B.; Behn, M. D. Automated Analysis of Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra of Natural Organic Matter. *Anal. Chem.* **2006**, *78* (13), 4363–4373. https://doi.org/10.1021/ac0600306.
(138) Herzsprung, P.; Hertkorn, N.; von Tümpling, W.; Harir, M.; Friese, K.; Schmitt-Kopplin, P. 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.* **2014**, *406* (30), 7977–7987. https://doi.org/10.1007/s00216-014-8249-y.
(139) Herzsprung, P.; Hertkorn, N.; Von Tümpling, W.; Harir, M.; Friese, K.; Schmitt-Kopplin, P. Molecular Formula Assignment for Dissolved Organic Matter (DOM) Using High-Field FT-ICR-MS: Chemical Perspective and Validation of Sulphur-Rich Organic Components (CHOS) in Pit Lake Samples. *Anal. Bioanal. Chem.* **2016**, *408* (10), 2461–2469. https://doi.org/10.1007/s00216-016-9341-2.
(140) Savory, J. J.; Kaiser, N. K.; McKenna, A. M.; Xian, F.; Blakney, G. T.; Rodgers, R. P.; Hendrickson, C. L.; Marshall, A. G. Parts-per-Billion Fourier Transform Ion Cyclotron Resonance Mass Measurement Accuracy with a "Walking" Calibration Equation. *Anal. Chem.* **2011**, *83* (5), 1732–1736. https://doi.org/10.1021/ac102943z.
(141) Herzsprung, P.; V. Tümpling, W.; Hertkorn, N.; Harir, M.; Friese, K.; Schmitt-Kopplin, P. High-Field FTICR-MS Data Evaluation of Natural Organic Matter: Are CHON5S2 Molecular Class Formulas Assigned to 13C Isotopic m/z and in Reality CHO Components? *Anal. Chem.* **2015**, *87* (19), 9563–9566. https://doi.org/10.1021/acs.analchem.5b02549.
(142) Sleighter, R. L.; Mckee, G. A.; Liu, Z.; Hatcher, P. G. Naturally Present Fatty Acids as Internal Calibrants for Fourier Transform Mass Spectra of Dissolved Organic Matter. *Limnol. Oceanogr. Methods* **2008**, *6* (6), 246–253. https://doi.org/10.4319/lom.2008.6.246.
(143) Qi, Y.; O'Connor, P. B. Data Processing in Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Mass Spectrom. Rev.* **2014**, *33* (5), 333–352. https://doi.org/10.1002/mas.21414.
(144) Smith, D. F.; Podgorski, D. C.; Rodgers, R. P.; Blakney, G. T.; Hendrickson, C. L. 21 Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures. *Anal. Chem.* **2018**, *90* (3), 2041–2047. https://doi.org/10.1021/acs.analchem.7b04159.
(145) Merder, J.; Freund, J. A.; Feudel, U.; Niggemann, J.; Singer, G.; Dittmar, T. Improved Mass Accuracy and Isotope Confirmation through Alignment of Ultrahigh-Resolution Mass Spectra of Complex Natural Mixtures. *Anal. Chem.* **2020**, *92* (3), 2558–2565. https://doi.org/10.1021/acs.analchem.9b04234.
(146) G. Marshall, A.; T. Blakney, G.; Chen, T.; K. Kaiser, N.; M. McKenna, A.; P. Rodgers, R.; M. Ruddy, B.; Xian, F. Mass Resolution and Mass Accuracy: How Much Is Enough? *Mass Spectrom.* **2013**, *2* (Special_Issue), S0009–S0009. https://doi.org/10.5702/massspectrometry.S0009.
(147) Marshall, A. G.; Guan, S. Advantages of High Magnetic Field for Fourier Transform Ion

Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1996**, *10* (14), 1819–1823. https://doi.org/10.1002/(SICI)1097-0231(199611)10:14<1819::AID-RCM686>3.0.CO;2-Z.
(148) Hohenester, U. M.; Barbier Saint-Hilaire, P.; Fenaille, F.; Cole, R. B. Investigation of Space Charge Effects and Ion Trapping Capacity on Direct Introduction Ultra-High-Resolution Mass Spectrometry Workflows for Metabolomics. *J. Mass Spectrom.* **2020**, *55* (10), 1–15. https://doi.org/10.1002/jms.4613.
(149) Kaiser, N. K.; McKenna, A. M.; Savory, J. J.; Hendrickson, C. L.; Marshall, A. G. Tailored Ion Radius Distribution for Increased Dynamic Range in FT-ICR Mass Analysis of Complex Mixtures. *Anal. Chem.* **2013**, *85* (1), 265–272. https://doi.org/10.1021/ac302678v.
(150) McKenna, A. M.; Donald, L. J.; Fitzsimmons, J. E.; Juyal, P.; Spicer, V.; Standing, K. G.; Marshall, A. G.; Rodgers, R. P. Heavy Petroleum Composition. 3. Asphaltene Aggregation. *Energy and Fuels* **2013**, *27* (3), 1246–1256. https://doi.org/10.1021/ef3018578.
(151) McKenna, A. M.; Marshall, A. G.; Rodgers, R. P. Heavy Petroleum Composition. 4. Asphaltene Compositional Space. *Energy and Fuels* **2013**, *27* (3), 1257–1267. https://doi.org/10.1021/ef301747d.
(152) Sleighter, R. L.; Chen, H.; Wozniak, A. S.; Willoughby, A. S.; Caricasole, P.; Hatcher, P. G. Establishing a Measure of Reproducibility of Ultrahigh-Resolution Mass Spectra for Complex Mixtures of Natural Organic Matter. *Anal. Chem.* **2012**, *84* (21), 9184–9191. https://doi.org/10.1021/ac3018026.
(153) Corilo, Y. E. PetroOrg Software; Florida State University, Omics LLC: Tallahassee, Fl,. **2014**. https://doi.org/10.1021/acs.est.7b04445.
(154) Schum, S. K.; Brown, L. E.; Mazzoleni, L. R. MFAssignR: Molecular Formula Assignment Software for Ultrahigh Resolution Mass Spectrometry Analysis of Environmental Complex Mixtures. *Environ. Res.* **2020**, *191*, 110114. https://doi.org/https://doi.org/10.1016/j.envres.2020.110114.
(155) Fu, Q.-L. L.; Fujii, M.; Riedel, T. Development and Comparison of Formula Assignment Algorithms for Ultrahigh-Resolution Mass Spectra of Natural Organic Matter. *Anal. Chim. Acta* **2020**, *1125*, 247–257. https://doi.org/https://doi.org/10.1016/j.aca.2020.05.048.
(156) Kim, S.; Rodgers, R. P.; Marshall, A. G. Truly "Exact" Mass: Elemental Composition Can Be Determined Uniquely from Molecular Mass Measurement at ∼0.1 MDa Accuracy for Molecules up to ∼500 Da. *Int. J. Mass Spectrom.* **2006**, *251* (2-3 SPEC. ISS.), 260–265. https://doi.org/10.1016/j.ijms.2006.02.001.
(157) Kunenkov, E. V.; Kononikhin, A. S.; Perminova, I. V.; Hertkorn, N.; Gaspar, A.; Schmitt-Kopplin, P.; Popov, I. A.; Garmash, A. V.; Nikolaev, E. N. Total Mass Difference Statistics Algorithm: A New Approach to Identification of High-Mass Building Blocks in Electrospray Ionization Fourier Transform Ion Cyclotron Mass Spectrometry Data of Natural Organic Matter. *Anal. Chem.* **2009**, *81* (24), 10106–10115. https://doi.org/10.1021/ac901476u.
(158) Tziotis, D.; Hertkorn, N.; Schmitt-Kopplin, P. Kendrick-Analogous Network Visualisation of Ion Cyclotron Resonance Fourier Transform Mass Spectra: Improved Options for the Assignment of Elemental Compositions and the Classification of Organic Molecular Complexity. *Eur. J. Mass Spectrom.* **2011**, *17* (4), 415–421.

https://doi.org/10.1255/ejms.1135.
(159) Kind, T.; Fiehn, O. Seven Golden Rules for Heuristic Filtering of Molecular Formulas Obtained by Accurate Mass Spectrometry. *BMC Bioinformatics* **2007**, *8*. https://doi.org/10.1186/1471-2105-8-105.
(160) Tolić, N. T.; Liu, Y.; Liyu, A.; Shen, Y.; Tfaily, M. M.; Kujawinski, E. B.; Longnecker, K.; Kuo, L.-J. J.; Robinson, E. W.; Paš, L.; et al. Formularity: Software for Automated Formula Assignment of Natural and Other Organic Matter from Ultrahigh-Resolution Mass Spectra. *Anal. Chem.* **2017**, *89* (23), 12659–12665. https://doi.org/10.1021/acs.analchem.7b03318.
(161) Merder, J.; Freund, J. A.; Feudel, U.; Hansen, C. T.; Hawkes, J. A.; Jacob, B.; Klaproth, K.; Niggemann, J.; Noriega-Ortega, B. E.; Osterholz, H.; et al. ICBM-OCEAN: Processing Ultrahigh-Resolution Mass Spectrometry Data of Complex Molecular Mixtures. *Anal. Chem.* **2020**, *92* (10), 6832–6838. https://doi.org/10.1021/acs.analchem.9b05659.
(162) Koch, B. P.; Dittmar, T. From Mass to Structure: An Aromaticity Index for High-Resolution Mass Data of Natural Organic Matter. *Rapid Commun. Mass Spectrom.* **2006**, *20* (5), 926–932. https://doi.org/10.1002/rcm.2386.
(163) Kroll, J. H.; Donahue, N. M.; Jimenez, J. L.; Kessler, S. H.; Canagaratna, M. R.; Wilson, K. R.; Altieri, K. E.; Mazzoleni, L. R.; Wozniak, A. S.; Bluhm, H.; et al. Carbon Oxidation State as a Metric for Describing the Chemistry of Atmospheric Organic Aerosol. *Nat. Chem.* **2011**, *3* (2), 133–139. https://doi.org/10.1038/nchem.948.
(164) Lavonen, E. E.; Gonsior, M.; Tranvik, L. J.; Schmitt-Kopplin, P.; Köhler, S. J. Selective Chlorination of Natural Organic Matter: Identification of Previously Unknown Disinfection Byproducts. *Environ. Sci. Technol.* **2013**, *47* (5), 2264–2271. https://doi.org/10.1021/es304669p.
(165) Riedel, T.; Biester, H.; Dittmar, T. Molecular Fractionation of Dissolved Organic Matter with Metal Salts. *Environ. Sci. Technol.* **2012**, *46* (8), 4419–4426. https://doi.org/10.1021/es203901u.
(166) LaRowe, D. E.; Van Cappellen, P. Degradation of Natural Organic Matter: A Thermodynamic Analysis. *Geochim. Cosmochim. Acta* **2011**, *75* (8), 2030–2042. https://doi.org/10.1016/j.gca.2011.01.020.
(167) Keiluweit, M.; Nico, P. S.; Kleber, M.; Fendorf, S. Are Oxygen Limitations under Recognized Regulators of Organic Carbon Turnover in Upland Soils? *Biogeochemistry* **2016**, *127* (2–3), 157–171. https://doi.org/10.1007/s10533-015-0180-6.
(168) Graham, E. B.; Tfaily, M. M.; Crump, A. R.; Goldman, A. E.; Bramer, L. M.; Arntzen, E.; Romero, E.; Resch, C. T.; Kennedy, D. W.; Stegen, J. C. Carbon Inputs From Riparian Vegetation Limit Oxidation of Physically Bound Organic Carbon Via Biochemical and Thermodynamic Processes. *J. Geophys. Res. Biogeosciences* **2017**, *122* (12), 3188–3205. https://doi.org/10.1002/2017JG003967.
(169) Wilson, R. M.; Tfaily, M. M. Advanced Molecular Techniques Provide New Rigorous Tools for Characterizing Organic Matter Quality in Complex Systems. *Journal of Geophysical Research: Biogeosciences*. Blackwell Publishing Ltd June 1, 2018, pp 1790–1795. https://doi.org/10.1029/2018JG004525.
(170) Boye, K.; Noël, V.; Tfaily, M. M.; Bone, S. E.; Williams, K. H.; Bargar, J. R. R.; Fendorf, S. Thermodynamically Controlled Preservation of Organic Carbon in Floodplains. *Nat. Geosci.* **2017**, *10* (6), 415–419. https://doi.org/10.1038/ngeo2940.

(171) Kendrick, E. A Mass Scale Based on CH2= 14.0000 for High Resolution Mass Spectrometry of Organic Compounds. *Anal. Chem.* **1963**, *35* (13), 2146–2154. https://doi.org/10.1021/ac60206a048.
(172) Kramer, R. W.; Kujawinski, E. B.; Hatcher, P. G. Identification of Black Carbon Derived Structures in a Volcanic Ash Soil Humic Acid by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environ. Sci. Technol.* **2004**, *38* (12), 3387–3395. https://doi.org/10.1021/es030124m.
(173) Stenson, A. C.; Marshall, A. G.; Cooper, W. T. 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.* **2003**, *75* (6), 1275–1284. https://doi.org/10.1021/ac026106p.
(174) Kujawinski, E. B.; Del Vecchio, R.; Blough, N. V; Klein, G. C.; Marshall, A. G. Probing Molecular-Level Transformations of Dissolved Organic Matter: Insights on Photochemical Degradation and Protozoan Modification of DOM from Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. In *Marine Chemistry*; 2004; Vol. 92, pp 23–37. https://doi.org/10.1016/j.marchem.2004.06.038.
(175) Sleighter, R. L.; Hatcher, P. G. The Application of Electrospray Ionization Coupled to Ultrahigh Resolution Mass Spectrometry for the Molecular Characterization of Natural Organic Matter. *J. MASS Spectrom. J. Mass Spectrom* **2007**, *42*, 559–574. https://doi.org/10.1002/jms.1221.
(176) Ohno, T.; Parr, T. B.; Gruselle, M. C. I.; Fernandez, I. J.; Sleighter, R. L.; Hatcher, P. G. Molecular Composition and Biodegradability of Soil Organic Matter: A Case Study Comparing Two New England Forest Types. *Environ. Sci. Technol.* **2014**, *48* (13), 7229–7236. https://doi.org/10.1021/es405570c.
(177) Dier, T. K. F.; Egele, K.; Fossog, V.; Hempelmann, R.; Volmer, D. A. Enhanced Mass Defect Filtering to Simplify and Classify Complex Mixtures of Lignin Degradation Products. *Anal. Chem.* **2016**, *88* (2), 1328–1335. https://doi.org/10.1021/acs.analchem.5b03790.
(178) Sato, H.; Nakamura, S.; Teramoto, K.; Sato, T. Structural Characterization of Polymers by MALDI Spiral-TOF Mass Spectrometry Combined with Kendrick Mass Defect Analysis. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (8), 1346–1355. https://doi.org/10.1007/s13361-014-0915-y.
(179) Glasser, W. G. About Making Lignin Great Again—Some Lessons From the Past . *Frontiers in Chemistry* . 2019, p 565.
(180) Fouquet, T. N. J.; Cody, R. B.; Ozeki, Y.; Kitagawa, S.; Ohtani, H.; Sato, H. On the Kendrick Mass Defect Plots of Multiply Charged Polymer Ions: Splits, Misalignments, and How to Correct Them. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (8), 1611–1626. https://doi.org/10.1007/s13361-018-1972-4.
(181) Liu, Q.; Luo, L.; Zheng, L. Lignins: Biosynthesis and Biological Functions in Plants. *International Journal of Molecular Sciences*. 2018. https://doi.org/10.3390/ijms19020335.
(182) Lv, J.; Zhang, S.; Wang, S.; Luo, L.; Cao, D.; Christie, P. Molecular-Scale Investigation with ESI-FT-ICR-MS on Fractionation of Dissolved Organic Matter Induced by Adsorption on Iron Oxyhydroxides. *Environ. Sci. Technol.* **2016**, *50* (5), 2328–2336. https://doi.org/10.1021/acs.est.5b04996.
(183) Brockman, S. A.; Roden, E. V.; Hegeman, A. D. Van Krevelen Diagram Visualization of High Resolution-Mass Spectrometry Metabolomics Data with OpenVanKrevelen.

Metabolomics **2018**, *14* (4), 1–5. https://doi.org/10.1007/s11306-018-1343-y.
(184) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **2003**, *75* (20), 5336–5344. https://doi.org/10.1021/ac034415p.
(185) Minor, E. C.; Swenson, M. M.; Mattson, B. M.; Oyler, A. R. Structural Characterization of Dissolved Organic Matter: A Review of Current Techniques for Isolation and Analysis. *Environ. Sci. Process. Impacts* **2014**, *16* (9), 2064–2079. https://doi.org/10.1039/C4EM00062E.
(186) D'Andrilli, J.; Foreman, C. M.; Marshall, A. G.; McKnight, D. M. 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.* **2013**, *65*, 19–28. https://doi.org/10.1016/j.orggeochem.2013.09.013.
(187) Hemmler, D.; Roullier-Gall, C.; Marshall, J. W.; Rychlik, M.; Taylor, A. J.; Schmitt-Kopplin, P. Evolution of Complex Maillard Chemical Reactions, Resolved in Time. *Sci. Rep.* **2017**, *7* (1), 3–8. https://doi.org/10.1038/s41598-017-03691-z.
(188) Waggoner, D. C.; Chen, H.; Willoughby, A. S.; Hatcher, P. G. Formation of Black Carbon-like and Alicyclic Aliphatic Compounds by Hydroxyl Radical Initiated Degradation of Lignin. **2015**. https://doi.org/10.1016/j.orggeochem.2015.02.007.
(189) Rivas-Ubach, A.; Liu, Y.; Bianchi, T. S.; Tolić, N.; Jansson, C.; Paša-Tolić, L. Moving beyond the van Krevelen Diagram: A New Stoichiometric Approach for Compound Classification in Organisms. *Anal. Chem.* **2018**, *90* (10), 6152–6160. https://doi.org/10.1021/acs.analchem.8b00529.
(190) Smith, A. P.; Bond-Lamberty, B.; Benscoter, B. W.; Tfaily, M. M.; Hinkle, C. R.; Liu, C.; Bailey, V. L. Shifts in Pore Connectivity from Precipitation versus Groundwater Rewetting Increases Soil Carbon Loss after Drought. *Nat. Commun.* **2017**, *8* (1), 1–11. https://doi.org/10.1038/s41467-017-01320-x.
(191) Hodgkins, S. B.; Tfaily, M. M.; McCalley, C. K.; Logan, T. A.; Crill, P. M.; Saleska, S. R.; Rich, V. I.; Chanton, J. P. Changes in Peat Chemistry Associated with Permafrost Thaw Increase Greenhouse Gas Production. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (16), 5819–5824. https://doi.org/10.1073/pnas.1314641111.
(192) Riedel, T.; Zark, M.; Vähätalo, A. V.; Niggemann, J.; Spencer, R. G. M.; Hernes, P. J.; Dittmar, T. Molecular Signatures of Biogeochemical Transformations in Dissolved Organic Matter from Ten World Rivers. *Front. Earth Sci.* **2016**, *4* (September), 1–16. https://doi.org/10.3389/feart.2016.00085.
(193) Avneri-Katz, S.; Young, R. B.; McKenna, A. M.; Chen, H.; Corilo, Y. E.; Polubesova, T.; Borch, T.; Chefetz, B. Adsorptive Fractionation of Dissolved Organic Matter (DOM) by Mineral Soil: Macroscale Approach and Molecular Insight. *Org. Geochem.* **2017**, *103*, 113–124. https://doi.org/10.1016/j.orggeochem.2016.11.004.
(194) Mapolelo, M. M.; Rodgers, R. P.; Blakney, G. T.; Yen, A. T.; Asomaning, S.; Marshall, A. G. Characterization of Naphthenic Acids in Crude Oils and Naphthenates by Electrospray Ionization FT-ICR Mass Spectrometry. *Int. J. Mass Spectrom.* **2011**, *300* (2), 149–157. https://doi.org/https://doi.org/10.1016/j.ijms.2010.06.005.
(195) Mazur, D. M.; Harir, M.; Schmitt-Kopplin, P.; Polyakova, O. V; Lebedev, A. T. High Field FT-ICR Mass Spectrometry for Molecular Characterization of Snow Board from Moscow Regions. *Sci. Total Environ.* **2016**, *557–558*, 12–19.

https://doi.org/https://doi.org/10.1016/j.scitotenv.2016.02.178.
(196) Cho, Y.; Ahmed, A.; Islam, A.; Kim, S. Developments in FT-ICR Ms Instrumentation, Ionization Techniques, and Data Interpretation Methods for Petroleomics. *Mass Spectrom. Rev.* **2015**, *34* (2), 248–263. https://doi.org/10.1002/mas.21438.
(197) Jarvis, J. M.; Page-Dumroese, D. S.; Anderson, N. M.; Corilo, Y.; Rodgers, R. P. Characterization of Fast Pyrolysis Products Generated from Several Western USA Woody Species. *Energy and Fuels* **2014**, *28* (10), 6438–6446. https://doi.org/10.1021/ef501714j.
(198) Borch, T.; Kretzschmar, R.; Skappler, A.; Van Cappellen, P.; Ginder-Vogel, M.; Voegelin, A.; Campbell, K. Biogeochemical Redox Processes and Their Impact on Contaminant Dynamics. *Environ. Sci. Technol.* **2010**, *44* (1), 15–23. https://doi.org/10.1021/es9026248.
(199) Wagner, S.; Riedel, T.; Niggemann, J.; Vähätalo, A. V.; Dittmar, T.; Jaffé, R. Linking the Molecular Signature of Heteroatomic Dissolved Organic Matter to Watershed Characteristics in World Rivers. *Environ. Sci. Technol.* **2015**, *49* (23), 13798–13806. https://doi.org/10.1021/acs.est.5b00525.
(200) Brooker, M. R.; Longnecker, K.; Kujawinski, E. B.; Evert, M. H.; Mouser, P. J. Discrete Organic Phosphorus Signatures Are Evident in Pollutant Sources within a Lake Erie Tributary. *Environ. Sci. Technol.* **2018**, *52* (12), 6771–6779. https://doi.org/10.1021/acs.est.7b05703.
(201) Hervé, M. R.; Nicolè, F.; Lê Cao, K. A. Multivariate Analysis of Multiple Datasets: A Practical Guide for Chemical Ecology. *J. Chem. Ecol.* **2018**, *44* (3), 215–234. https://doi.org/10.1007/s10886-018-0932-6.
(202) Hur, M.; Yeo, I.; Park, E.; Kim, Y. H.; Yoo, J.; Kim, E.; No, M.; Koh, J.; Kim, S. Combination of Statistical Methods and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for More Comprehensive, Molecular-Level Interpretations of Petroleum Samples. *Anal. Chem.* **2010**, *82* (1), 211–218. https://doi.org/10.1021/ac901748c.
(203) Kujawinski, E. B.; Longnecker, K.; Blough, N. V.; Vecchio, R. Del; Finlay, L.; Kitner, J. B.; Giovannoni, S. J. Identification of Possible Source Markers in Marine Dissolved Organic Matter Using Ultrahigh Resolution Mass Spectrometry. *Geochim. Cosmochim. Acta* **2009**, *73* (15), 4384–4399. https://doi.org/10.1016/j.gca.2009.04.033.
(204) Temminghoff, E. J. M.; Van Der Zee, S. E. A. T. M.; De Haan, F. A. M. Copper Mobility in a Copper-Contaminated Sandy Soil as Affected by PH and Solid and Dissolved Organic Matter. *Environ. Sci. Technol.* **1997**, *31* (4), 1109–1115. https://doi.org/10.1021/es9606236.
(205) Lindsay, W. L.; Norvell, W. A. Development of a DTPA Soil Test for Zinc, Iron, Manganese, and Copper. *Soil Sci. Soc. Am. J.* **1978**, *42* (3), 421–428. https://doi.org/10.2136/sssaj1978.03615995004200030009x.
(206) Gleyzes, C.; Tellier, S.; Astruc, M. Fractionation Studies of Trace Elements in Contaminated Soils and Sediments: A Review of Sequential Extraction Procedures. *TrAC - Trends Anal. Chem.* **2002**, *21* (6–7), 451–467. https://doi.org/10.1016/S0165-9936(02)00603-9.
(207) Boiteau, R. M.; Fansler, S. J.; Farris, Y.; Shaw, J. B.; Koppenaal, D. W.; Pasa-Tolic, L.; Jansson, J. K. Siderophore Profiling of Co-Habitating Soil Bacteria by Ultra-High Resolution Mass Spectrometry. *Metallomics* **2019**, *11* (1), 166–175. https://doi.org/10.1039/c8mt00252e.

(208) Waska, H.; Brumsack, H. J.; Massmann, G.; Koschinsky, A.; Schnetger, B.; Simon, H.; Dittmar, T. Inorganic and Organic Iron and Copper Species of the Subterranean Estuary: Origins and Fate. *Geochim. Cosmochim. Acta* **2019**, *259*, 211–232. https://doi.org/10.1016/j.gca.2019.06.004.
(209) Waska, H.; Koschinsky, A.; Ruiz Chancho, M. J.; Dittmar, T. Investigating the Potential of Solid-Phase Extraction and Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) for the Isolation and Identification of Dissolved Metal-Organic Complexes from Natural Waters. *Mar. Chem.* **2015**, *173*, 78–92. https://doi.org/10.1016/j.marchem.2014.10.001.
(210) Boiteau, R. M.; Till, C. P.; Ruacho, A.; Bundy, R. M.; Hawco, N. J.; McKenna, A. M.; Barbeau, K. A.; Bruland, K. W.; Saito, M. A.; Repeta, D. J. Structural Characterization of Natural Nickel and Copper Binding Ligands along the US GEOTRACES Eastern Pacific Zonal Transect. *Front. Mar. Sci.* **2016**, *3* (NOV), 243. https://doi.org/10.3389/fmars.2016.00243.
(211) Chen, H.; Johnston, R. C.; Mann, B. F.; Chu, R. K.; Tolic, N.; Parks, J. M.; Gu, B. Identification of Mercury and Dissolved Organic Matter Complexes Using Ultrahigh Resolution Mass Spectrometry. *Environ. Sci. Technol. Lett* **2017**, *4*, 3. https://doi.org/10.1021/acs.estlett.6b00460.
(212) Xu, C.; Chen, H.; Sugiyama, Y.; Zhang, S.; Li, H. P.; Ho, Y. F.; Chuang, C. Y.; Schwehr, K. A.; Kaplan, D. I.; Yeager, C.; et al. Novel Molecular-Level Evidence of Iodine Binding to Natural Organic Matter from Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Sci. Total Environ.* **2013**, *449*, 244–252. https://doi.org/10.1016/j.scitotenv.2013.01.064.
(213) Shen, Y.; Zhao, R.; Tolić, N.; Tfaily, M. M.; Robinson, E. W.; Boiteau, R.; Paša-Tolić, L.; Hess, N. J. Online Supercritical Fluid Extraction Mass Spectrometry (SFE-LC-FTMS) for Sensitive Characterization of Soil Organic Matter. *Faraday Discuss.* **2019**, *218*, 157–171. https://doi.org/10.1039/c9fd00011a.
(214) Rodgers, R. P.; Mapolelo, M. M.; Robbins, W. K.; Chacón-Patiño, M. L.; Putman, J. C.; Niles, S. F.; Rowland, S. M.; Marshall, A. G. Combating Selective Ionization in the High Resolution Mass Spectral Characterization of Complex Mixtures. *Faraday Discuss.* **2019**, *218*, 29–51. https://doi.org/10.1039/c9fd00005d.
(215) Capley, E. N.; Tipton, J. D.; Marshall, A. G.; Stenson, A. C. Chromatographic Reduction of Isobaric and Isomeric Complexity of Fulvic Acids to Enable Multistage Tandem Mass Spectral Characterization. *Anal. Chem.* **2010**, *82* (19), 8194–8202. https://doi.org/10.1021/ac1016216.
(216) Stenson, A. C. Reversed-Phase Chromatography Fractionation Tailored to Mass Spectral Characterization of Humic Substances. *Environ. Sci. Technol.* **2008**, *42* (6), 2060–2065. https://doi.org/10.1021/es7022412.
(217) Fudyma, J. D.; Lyon, J.; AminiTabrizi, R.; Gieschen, H.; Chu, R. K.; Hoyt, D. W.; Kyle, J. E.; Toyoda, J.; Tolic, N.; Heyman, H. M.; et al. Untargeted Metabolomic Profiling of Sphagnum Fallax Reveals Novel Antimicrobial Metabolites. *Plant Direct* **2019**, *3* (11), e00179. https://doi.org/10.1002/pld3.179.
(218) Boiteau, R. M.; Fitzsimmons, J. N.; Repeta, D. J.; Boyle, E. A. Detection of Iron Ligands in Seawater and Marine Cyanobacteria Cultures by High-Performance Liquid Chromatography-Inductively Coupled Plasma-Mass Spectrometry. *Anal. Chem.* **2013**, *85* (9), 4357–4362. https://doi.org/10.1021/ac3034568.

(219) Patriarca, C.; Bergquist, J.; Sjöberg, P. J. R.; Tranvik, L.; Hawkes, J. A. Online HPLC-ESI-HRMS Method for the Analysis and Comparison of Different Dissolved Organic Matter Samples. *Environ. Sci. Technol.* **2018**, *52* (4), 2091–2099. https://doi.org/10.1021/acs.est.7b04508.
(220) Woods, G. C.; Simpson, M. J.; Koerner, P. J.; Napoli, A.; Simpson, A. J. HILIC-NMR: Toward the Identification of Individual Molecular Components in Dissolved Organic Matter. *Environ. Sci. Technol.* **2011**, *45* (9), 3880–3886. https://doi.org/10.1021/es103425s.
(221) Ladd, M. P.; Giannone, R. J.; Abraham, P. E.; Wullschleger, S. D.; Hettich, R. L. Evaluation of an Untargeted Nano-Liquid Chromatography-Mass Spectrometry Approach to Expand Coverage of Low Molecular Weight Dissolved Organic Matter in Arctic Soil. *Sci. Rep.* **2019**, *9* (1), 1–13. https://doi.org/10.1038/s41598-019-42118-9.
(222) Sandron, S.; Rojas, A.; Wilson, R.; Davies, N. W.; Haddad, P. R.; Shellie, R. A.; Nesterenko, P. N.; Kelleher, B. P.; Paull, B. Chromatographic Methods for the Isolation, Separation and Characterisation of Dissolved Organic Matter. *Environ. Sci. Process. Impacts* **2015**, *17* (9), 1531–1567. https://doi.org/10.1039/c5em00223k.
(223) Leclair, J. P.; Collett, J. L.; Mazzoleni, L. R. Fragmentation Analysis of Water-Soluble Atmospheric Organic Matter Using Ultrahigh-Resolution FT-ICR Mass Spectrometry. *Environ. Sci. Technol.* **2012**, *46* (8), 4312–4322. https://doi.org/10.1021/es203509b.
(224) Rodgers, R. P.; Marshall, A. G.; Niles, S. F.; Chacón-Patinõ, M. L.; Smith, D. F. Comprehensive Compositional and Structural Comparison of Coal and Petroleum Asphaltenes Based on Extrography Fractionation Coupled with Fourier Transform Ion Cyclotron Resonance MS and MS/MS Analysis. *Energy and Fuels* **2020**, *34* (2), 1492–1505. https://doi.org/10.1021/acs.energyfuels.9b03527.
(225) Chacón-Patiño, M. L.; Rowland, S. M.; Rodgers, R. P. Advances in Asphaltene Petroleomics. Part 1: Asphaltenes Are Composed of Abundant Island and Archipelago Structural Motifs. *Energy and Fuels* **2017**, *31* (12), 13509–13518. https://doi.org/10.1021/acs.energyfuels.7b02873.
(226) Guijas, C.; Montenegro-Burke, J. R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A. E.; et al. METLIN: A Technology Platform for Identifying Knowns and Unknowns. *Anal. Chem.* **2018**, *90* (5), 3156–3164. https://doi.org/10.1021/acs.analchem.7b04424.
(227) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299–302. https://doi.org/10.1038/s41592-019-0344-8.
(228) Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S. MetFrag Relaunched: Incorporating Strategies beyond in Silico Fragmentation. *J. Cheminform.* **2016**, *8* (1), 3. https://doi.org/10.1186/s13321-016-0115-9.
(229) Djoumbou-Feunang, Y.; Pon, A.; Karu, N.; Zheng, J.; Li, C.; Arndt, D.; Gautam, M.; Allen, F.; Wishart, D. S. CFM-ID 3.0: Significantly Improved ESI-MS/MS Prediction and Compound Identification. *Metabolites* **2019**, *9* (4), 72. https://doi.org/10.3390/metabo9040072.
(230) Longnecker, K.; Kujawinski, E. B. Mining Mass Spectrometry Data: Using New Computational Tools to Find Novel Organic Compounds in Complex Environmental Mixtures. *Org. Geochem.* **2017**, *110*, 92–99.

https://doi.org/10.1016/j.orggeochem.2017.05.008.
(231) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*. Nature Publishing Group September 8, 2016, pp 828–837. https://doi.org/10.1038/nbt.3597.
(232) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; Van Der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; et al. Mass Spectral Molecular Networking of Living Microbial Colonies. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (26), 10150. https://doi.org/10.1073/pnas.1203689109.
(233) Pande, S.; Kost, C. Bacterial Unculturability and the Formation of Intercellular Metabolic Networks. *Trends Microbiol.* **2017**, *25* (5), 349–361. https://doi.org/10.1016/j.tim.2017.02.015.
(234) Pinu, F. R.; Granucci, N.; Daniell, J.; Han, T. L.; Carneiro, S.; Rocha, I.; Nielsen, J.; Villas-Boas, S. G. Metabolite Secretion in Microorganisms: The Theory of Metabolic Overflow Put to the Test. *Metabolomics* **2018**, *14* (4), 1–16. https://doi.org/10.1007/s11306-018-1339-7.
(235) Horak, I.; Engelbrecht, G.; van Rensburg, P. J. J.; Claassens, S. Microbial Metabolomics: Essential Definitions and the Importance of Cultivation Conditions for Utilizing Bacillus Species as Bionematicides. *J. Appl. Microbiol.* **2019**, *127* (2), 326–343. https://doi.org/10.1111/jam.14218.
(236) Swenson, T. L.; Jenkins, S.; Bowen, B. P.; Northen, T. R. Untargeted Soil Metabolomics Methods for Analysis of Extractable Organic Matter. *Soil Biol. Biochem.* **2015**, *80*, 189–198. https://doi.org/10.1016/j.soilbio.2014.10.007.
(237) Bhinderwala, F.; Wase, N.; Dirusso, C.; Powers, R. Combining Mass Spectrometry and NMR Improves Metabolite Detection and Annotation. *J. Proteome Res.* **2018**, *17* (11), 4017–4022. https://doi.org/10.1021/acs.jproteome.8b00567.
(238) Dalcin Martins, P.; Hoyt, D. W.; Bansal, S.; Mills, C. T.; Tfaily, M.; Tangen, B. A.; Finocchiaro, R. G.; Johnston, M. D.; McAdams, B. C.; Solensky, M. J.; et al. Abundant Carbon Substrates Drive Extremely High Sulfate Reduction Rates and Methane Fluxes in Prairie Pothole Wetlands. *Glob. Chang. Biol.* **2017**, *23* (8), 3107–3120. https://doi.org/10.1111/gcb.13633.
(239) Tfaily, M. M.; Wilson, R. M.; Brewer, H. M.; Chu, R. K.; Heyman, H. M.; Hoyt, D. W.; Kyle, J. E.; Purvine, S. O. Single-Throughput Complementary High-Resolution Analytical Techniques for Characterizing Complex Natural Organic Matter Mixtures. *JoVE (Journal Vis. Exp.* **2019**, No. 143, e59035. https://doi.org/10.3791/59035.
(240) Georgii, E.; Jin, M.; Zhao, J.; Kanawati, B.; Schmitt-Kopplin, P.; Albert, A.; Winkler, J. B.; Schäffner, A. R. Relationships between Drought, Heat and Air Humidity Responses Revealed by Transcriptome-Metabolome Co-Analysis. *BMC Plant Biol.* **2017**, *17* (1), 1–24. https://doi.org/10.1186/s12870-017-1062-y.
(241) Osterholz, H.; Kirchman, D. L.; Niggemann, J.; Dittmar, T. Diversity of Bacterial Communities and Dissolved Organic Matter in a Temperate Estuary. *FEMS Microbiol. Ecol.* **2018**, *94* (8), 1–11. https://doi.org/10.1093/femsec/fiy119.
(242) Zhou, L.; Zhou, Y.; Yao, X.; Cai, J.; Liu, X.; Tang, X.; Zhang, Y.; Jang, K. S.; Jeppesen, E. Decreasing Diversity of Rare Bacterial Subcommunities Relates to Dissolved Organic Matter along Permafrost Thawing Gradients. *Environ. Int.* **2020**, *134* (November 2019),

105330. https://doi.org/10.1016/j.envint.2019.105330.
(243) Djoumbou-Feunang, Y.; Fiamoncini, J.; Gil-de-la-Fuente, A.; Greiner, R.; Manach, C.; Wishart, D. S. BioTransformer: A Comprehensive Computational Tool for Small Molecule Metabolism Prediction and Metabolite Identification. *J. Cheminform.* **2019**, *11* (1), 2. https://doi.org/10.1186/s13321-018-0324-5.
(244) Yao, Q.; Li, Z.; Song, Y.; Wright, S. J.; Guo, X.; Tringe, S. G.; Tfaily, M. M.; Paša-Tolić, L.; Hazen, T. C.; Turner, B. L.; et al. Community Proteogenomics Reveals the Systemic Impact of Phosphorus Availability on Microbial Functions in Tropical Soil. *Nat. Ecol. Evol.* **2018**, *2* (3), 499–509. https://doi.org/10.1038/s41559-017-0463-5.
(245) Staley, C.; Ferrieri, A. P.; Tfaily, M. M.; Cui, Y.; Chu, R. K.; Wang, P.; Shaw, J. B.; Ansong, C. K.; Brewer, H.; Norbeck, A. D.; et al. Diurnal Cycling of Rhizosphere Bacterial Communities Is Associated with Shifts in Carbon Metabolism. *Microbiome* **2017**, *5* (1), 65. https://doi.org/10.1186/s40168-017-0287-1.
(246) Li, H. Y.; Wang, H.; Wang, H. T.; Xin, P. Y.; Xu, X. H.; Ma, Y.; Liu, W. P.; Teng, C. Y.; Jiang, C. L.; Lou, L. P.; et al. The Chemodiversity of Paddy Soil Dissolved Organic Matter Correlates with Microbial Community at Continental Scales. *Microbiome* **2018**, *6* (1), 1–16. https://doi.org/10.1186/s40168-018-0561-x.
(247) Graham, E. B.; Crump, A. R.; Kennedy, D. W.; Arntzen, E.; Fansler, S.; Purvine, S. O.; Nicora, C. D.; Nelson, W.; Tfaily, M. M.; Stegen, J. C. Multi 'omics Comparison Reveals Metabolome Biochemistry, Not Microbiome Composition or Gene Expression, Corresponds to Elevated Biogeochemical Function in the Hyporheic Zone. *Sci. Total Environ.* **2018**, *642*, 742–753. https://doi.org/10.1016/j.scitotenv.2018.05.256.
(248) Matilainen, A.; Gjessing, E. T.; Lahtinen, T.; Hed, L.; Bhatnagar, A.; Sillanpää, M. An Overview of the Methods Used in the Characterisation of Natural Organic Matter (NOM) in Relation to Drinking Water Treatment. *Chemosphere* **2011**, *83* (11), 1431–1442. https://doi.org/10.1016/j.chemosphere.2011.01.018.
(249) Brezonik, P. L.; Bloom, P. R.; Sleighter, R. L.; Cory, R. M.; Khwaja, A. R.; Hatcher, P. G. Chemical Differences of Aquatic Humic Substances Extracted by XAD-8 and DEAE-Cellulose. *J. Environ. Chem. Eng.* **2015**, *3* (4, Part B), 2982–2990. https://doi.org/https://doi.org/10.1016/j.jece.2015.03.004.
(250) Hawkes, J. A.; D'Andrilli, J.; Agar, J. N.; Barrow, M. P.; Berg, S. M.; Catalán, N.; Chen, H.; Chu, R. K.; Cole, R. B.; Dittmar, T.; et al. An International Laboratory Comparison of Dissolved Organic Matter Composition by High Resolution Mass Spectrometry: Are We Getting the Same Answer? *Limnol. Oceanogr. Methods* **2020**, *18* (6), 235–258. https://doi.org/10.1002/lom3.10364.
(251) Aron, A. T.; Gentry, E. C.; McPhail, K. L.; Nothias, L. F.; Nothias-Esposito, M.; Bouslimani, A.; Petras, D.; Gauglitz, J. M.; Sikora, N.; Vargas, F.; et al. Reproducible Molecular Networking of Untargeted Mass Spectrometry Data Using GNPS. *Nat. Protoc.* **2020**, *15* (6), 1954–1991. https://doi.org/10.1038/s41596-020-0317-5.

CHAPTER 4 - MICROBIAL IRON(III) REDUCTION DURING PALSA COLLAPSE PROMOTES GREENHOUSE GAS EMISSIONS BEFORE COMPLETE PERMAFROST THAW[3]

1. INTRODUCTION

Climate change has enormous consequences for permafrost environments, causing rapid changes in soil conditions (such as thermal and moisture regime, and aeration) with direct consequences for organic (OC) destabilization[1]. Permafrost soils store ~60% of the world's soil OC in 15% of the global soil area[2,3]. This preserved OC will become increasingly exposed to microbial decomposition and thus can be released from the active layer to the atmosphere as greenhouse gases (GHGs) such as carbon dioxide (CO_2) and methane (CH_4)[4] or discharged by drainage[5]. However, the magnitude of the release of this OC depends strongly on a large variety of factors[6], including the hydrology, soil parent material, organic matter content and the ability of soil minerals to protect OC from degradation, which can regulate long-term preservation of OC[7].

Iron (Fe) minerals are known to stabilize organic carbon by sorption/co-precipitation and protect it from degradation by generating OC-Fe associations that are more persistent in soils[8,9, 10-13]. However, by providing a terminal electron acceptor for anaerobic respiration[14,15], Fe can also enhance decomposition. The fate of Fe and associated OC determines Fe-OC aggregate formation and ultimately accessibility for microbial decomposition[15, 16,17]. Reactive Fe-OC associations (defined as the solid Fe phases that are reductively dissolved by sodium dithionite[11,18,19]) have been

[3] Patzner, M., Logan, M., McKenna, A., Young, R., Zhou, Z., Joss, H., et al. (2022). Microbial iron cycling during palsa hillslope collapse promotes greenhouse gas emissions before complete permafrost thaw. *Communications Earth and Environment*, *3*(76). https://doi.org/10.1038/s43247-022-00407-8

shown to serve as an effective rusty carbon sink and to preserve OC over geological timescales[11]. Previously, it has been shown that reactive Fe-OC associations can mainly be found in intact permafrost soils[20], but cannot preserve OC along a permafrost thaw gradient, following complete permafrost thaw from palsa to bog to fen type wetlands[19]. However, the stability of Fe-OC associations during transitional processes along permafrost thaw gradients remain unstudied. The need to better understand the climate impact of transitional processes in thawing permafrost was stated previously by Shelef *et al.*[21] who emphasize large uncertainty in permafrost carbon stocks (>200%) due to processes at collapsing fronts. Indeed, methane dynamics can also strongly differ between end-members and transitional thaw stages[22].

With permafrost thaw, soils become water-logged and oxygen (O_2) limited, favoring reductive dissolution of reactive Fe(III)[19]. Fe(III)-reducing microorganisms are able to use the reactive Fe(III) as an electron acceptor for anaerobic respiration and, depending on its composition, the associated OC as electron source, resulting in CO_2 and Fe(II) formation[22]. Thus, Fe(III) reduction directly contributes to CO_2 emissions[23]. Fe(III) reduction may also influence CH_4 emissions in thawing permafrost peatlands. On the one hand, Fe(III) reduction is thermodynamically more favorable and thus could outcompete methanogenesis[24]. On the other hand, Fe(III) reduction leads to proton consumption which results in an increasing pH that could favor methanogenesis[25]. The complex balance of these processes that either suppress or promote GHG emissions such as CO_2 and CH_4 highlights the need for a fundamental understanding of microbial Fe metabolisms and their interactions with methanotrophs and methanogens, which is currently lacking.

The release of previously Fe-associated OC into surrounding porewater following reductive dissolution could lead to further microbial decomposition of OC and emission of GHGs

such as CO_2 and CH_4. Mineral-associated OC (MAOC) has been proposed to be comprised of low molecular weight compounds of microbial (e.g. microbial polysaccharides, amino sugars, muramic acid) and plant origin[13,26-30] with low activation energies of MAOC for degradation by microbes. Therefore, the release of MAOC with permafrost thaw is considered an important driver of the composition of arctic surface waters and microbial respiration[31,32]. Recent studies described carboxylic-rich[33] and aliphatic Fe-bound OC in forest soils as more resistant during reductive dissolution[34]. In Siberian permafrost soils, hydrophobic, aromatic DOC was preferentially sorbed by shallower, acidic soil horizons and correlated with an increasing abundance of Fe oxides[35]. The identity of Fe-bound OC in permafrost environments, however, still remains unknown.

To understand the direct impact of the loss of this so-called rusty carbon sink[36] on net GHG emissions in thawing permafrost peatlands, it is essential to further determine (1) the bioavailability of Fe-bound OC and released OC during permafrost thaw and (2) changes in the present and active microbial community, particularly the Fe(III)-reducing bacteria which are key players in reactive Fe mineral dissolution and their interplay with methanotrophs and methanogens.

We followed the dynamic biogeochemical interactions of Fe-OC associations in the active layer along collapsing palsa hillslopes, where palsas underlain by intact permafrost are collapsing into partially-thawed, semi-wet bogs. Fe-OC associations were characterized in the solid phase using selective extractions, scanning electron microscopy (SEM), nanoscale secondary ion mass spectrometry (nanoSIMS), and Mössbauer spectroscopy, and the effect of palsa collapse on porewater geochemistry and CO_2 and CH_4 fluxes was quantified. Reactive Fe-associated OC and DOC in the porewater along the palsa hillslope were investigated at the molecular-level with Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS), and the present and

active microbial community was characterized using DNA- and RNA-based 16S rRNA amplicon (gene) sequencing.

2. Materials and Methods

2.1 Site information

Stordalen Mire (68 22' N, 19 03' E) is a subarctic peatland in northern Sweden underlain by discontinuous permafrost. The mire consists of three distinct sub-habitats: (1) palsa (intact permafrost) with ericaceous and woody plants; (2) ombrotrophic peatland or bog (intermediate thaw) with *Sphagnum spp.*, sedges and shrubs and (3) minerotrophic peatland or fen (fully thawed) with sedges, mainly *Eriophorum spp*[45] (Figure S1). Generally, palsas and bogs are only fed by precipitation and melt water and have more acidic surface waters (pH ~4). Fens are fed by surface water and groundwater, and maintain slightly acidic to alkaline pH[40]. The areal extent of intact palsa across Stordalen mire has declined significantly since 1970 due to progressive warming in the Arctic, while fen habitats have expanded[63]. It is also predicted that the whole mire might be free of permafrost as early as 2050[64,65].

2.2 Gas measurments

To measure CO_2 emissions along the palsa hillslope, two eosense instruments (eosFD Forced Diffusion chamber in conjunction with the eosLink-FD software, EOSENSE INC, Dartmouth, Canada) were installed (Figures S1 and S2): (1) at the top of the palsa hillslope (spot: Palsa A) and (2) at the transition to bog (spot: Front). The collar was situated in a flat location and inserted to near full depth. A centimeter of space was left to aid in installation of the eosFD itself as well as collar retrieval. The collar area was cleared of any rocks or debris, larger vegetation was

removed or avoided. The eosFD was deployed in the installed collar. The collars were deployed at least 24 hours prior to the start of the eosFD measurement collection to avoid disturbance-related fluxes in the early portion of the data collection. The eosFD samples gases from the atmospheric and soil cavities within the device. Briefly, gas is pulled from the atmospheric cavity to the sensor for 20 seconds to purge the sensor cavity, then sampled every 10 minutes for five samples. Gas is then pulled from the soil cavity for 20 seconds, then pulsed every 10 minutes for five samples. Forced diffusion flux is calculated as follows:

$$\frac{V}{A} * \frac{\partial C}{\partial t} = Fs - D\left(\frac{\Delta C}{L}\right)$$

(volume/surface area scaled rate of change in flux rate equal to the flux from the soil surface (Fs) minus the difference in concentration, ΔC (scaled by both the path length L and the diffusivity of the interface (membrane), D)).

The change in the flux rate over the timespan of the concentration measurements (around 60 seconds) is assumed to be zero (steady state):

$$\frac{V}{A}\frac{\partial C}{\partial t} = 0$$

This assumption results in a linear dependence with the path length and interface (membrane) diffusivity being constant and represented by a single coefficient, G:

$$Fs = G\ \Delta C$$

Furthermore, carbon dioxide and methane emissions along the palsa hillslope were measured in triplicate using plastic chambers sealed with a rubber stopper (Figures S1 and S2), as described previously[66]. The metal frames were pushed into the ground at least 24 hours before the measurements to avoid collecting gas emissions from the soil during installation. Again rocks, debris and larger vegetation was avoided. Deionized water was used in the frames to seal off the

chambers from ambient air. Gas chamber samples were collected with a gas-tight syringe (1100TLL 100 mL Gastight, Hamilton, Reno, NV, USA) and directly transferred into evacuated 12 mL exetainer vials[67] until analyzed. The sampling was done every 5 min for a total period of 30 mins in duplicates for palsa and front and in triplicates for bog. All gas samples from the field and standard gases used for calibration were measured with a gas chromatograph (Hewlett Packard, 5890 Series II) equipped with an electron capture detector (^{63}Ni-ECD).

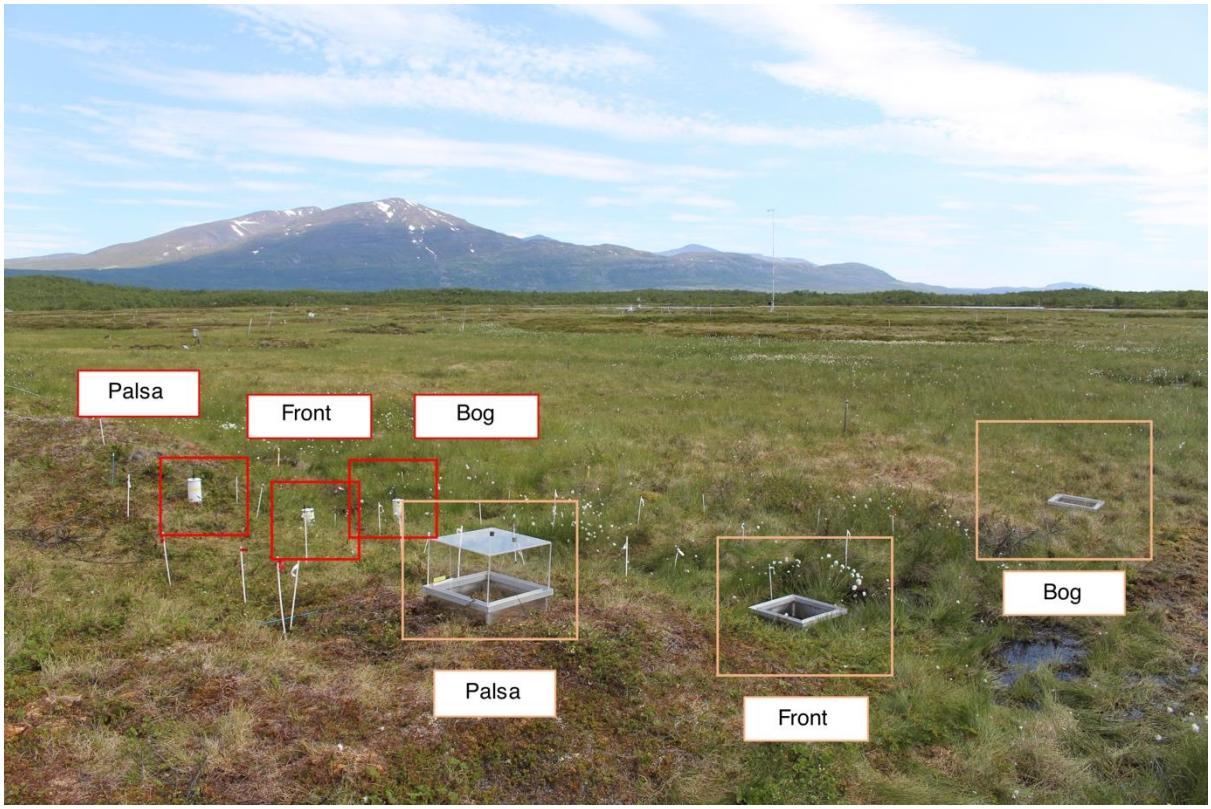


Figure S2. Gas sampling along palsa hillslopes. In transect 1 (see also Figure S1), Eosense gas chambers (eosFD Forced Diffusion chamber in conjunction with the eosLink-FD software, EOSENSE INC, Dartmouth, Canada) (red) were installed to measure carbon dioxide emissions along the palsa hillslope (68°21'18.70"N, 19° 2'38.00"E). Additional gas chambers with plastic chambers, sealed with a rubber stopper, and metal frames (orange) were installed to obtain replicate carbon dioxide and additionally methane emissions along the palsa hillslope into the bog area.

2.3 SAMPLE COLLECTION

In July 2019, cores were taken along three palsa hillslopes (Figure S1 and Figure S7), gently collapsing into bog, following the expected hydrological flow described previously[40]. A Humax corer of 50 cm length and 3-cm diameter with inner liners was used to sample the active layer[19]. The cores for mineral analysis were directly split after sampling under 100% N_2 atmosphere in a glove bag and subsamples stored at -20°C until analysis. The cores for microbial community analysis were split directly in the field, immediately frozen with liquid nitrogen and stored at -80°C until further processing. As previously described[19], the cores were split into three soil horizons based on texture and color changes: (1) A peat or organic horizon, followed by (2) a transition zone between the organic-rich and mineral-rich layer and (3) a mineral horizon.

In July and September 2019, porewater samples were collected from 30 and 60 cm depth below the peat surface along the palsa hillslope (8 transects, Figure S1 and Figure S4) using a luer lock syringe connected to a lysimeter with an effective pore size of 2.5 microns (Simpler Luer-Lock Micro Samplers, Model 1910LL, Soilmoisture Equipment Corp., Santa Barbara, CA). Prior to use, syringes and lysimeters were rinsed 10 times with 50 mL MiliQ water and air dried. Syringe filters (0.22 µm, PES, Merck™ Steritop™, Millipore) were pre-rinsed with 120 mL MilliQ water each to avoid leaching residuals of the filters. The syringes were flushed three times with N_2 and sealed till further use. Syringe filters (0.22 µm) were flushed three times with N_2 and placed into a SCHOTT bottle with N_2 atmosphere till further use. The lysimeters were installed in the soil, pre-flushed by pulling porewater with a syringe and the first 2 mL discharged. Immediately afterwards, the N_2 flushed syringes were unsealed, nitrogen gas pushed out, and then tightly connected to the installed lysimeter. To avoid direct sunlight exposure, syringes were covered with white cotton bags during the time of porewater extraction. After 3-4 hours, the samples were

filtered through a 0.22 µm syringe filter into stoppered, N_2 flushed glass vials, wrapped in aluminum foil to avoid any sunlight exposure. The first 1 mL filtrate from each sample was disposed. The samples were stored at 4°C till further analysis.

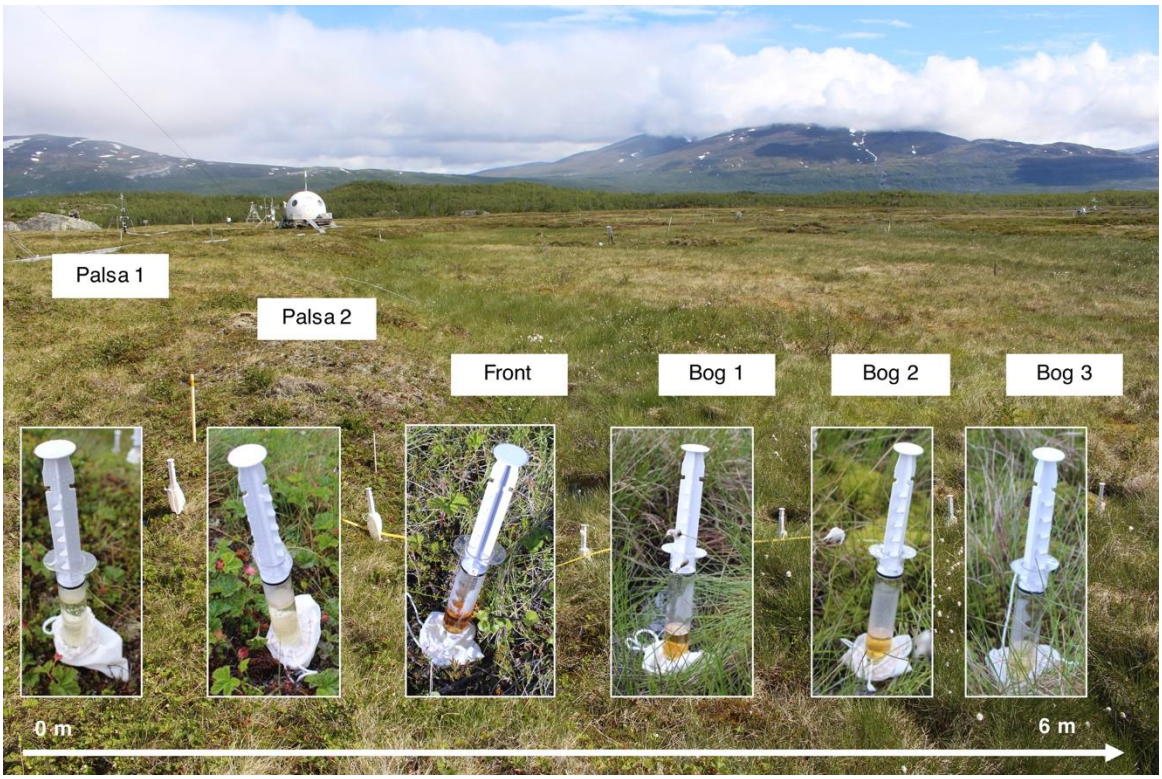


Figure S4. Porewater sampling along palsa hillslopes. Along eight palsa hillslopes, porewater was extracted with lysimeters at six defined sampling points in July 2019. In transect 1 (shown here, see also Figure S1), lysimeters were installed for 3-4 hours along the palsa hillslope (68°21'18.70"N, 19° 2'38.00"E) at a distance of 1 m in between each other from palsa to bog. The porewater with dark brown, reddish color at the collapsing front represents the sample with the highest aqueous iron and dissolved organic carbon along the palsa hillslope into bog.

2.4 Selective iron and carbon extractions

To extract reactive Fe and associated OC, selective dissolutions were conducted as described previously[19]. Briefly, a sodium bicarbonate (0.11 M) sodium dithionite (0.27 M) trisodium citrate (0.27 M) solution was used to reductively dissolve reactive Fe and associated

organic carbon. As described in Lalonde *et al.*[11], a sodium bicarbonate (0.11 M) sodium chloride (1.85 M) solution was used as a control experiment to distinguish between Fe and OC readily desorbed (leachable OC) and organic carbon associated with reactive Fe and only dissolved during reductive dissolution with dithionite. The citrate background in the extract also needs to be subtracted to receive the reactive Fe-associated OC. Thus, only the control corrected values are discussed:

$$Reactive\ Fe = Fe(dithionite\ citrate) - Fe(sodium\ chloride) \quad (1)$$

$$Reactive\ Fe - associated\ OC = DOC(dithionite\ citrate) - DOC(blank\ citrate) - DOC(sodium\ chloride) \quad (2)$$

For each soil horizon (organic horizon, transition zone, mineral horizon), 0.3 g dry soil was weighed into 10 mL glass vials with 6.25 mL extractant and N_2 headspace. After 16 hours at room temperature on a rolling shaker, samples were centrifuged at room temperature for 10 min at 5300 g. The supernatant was decanted and further analyzed for total Fe and DOC.

2.5 GEOCHEMICAL ANALYSIS

To determine total Fe and Fe(II), porewater or extract was acidified in 1 M HCl and quantified spectrophotometrically in triplicates with the ferrozine assay[68]. DOC was measured in triplicates with a total organic carbon analyzer (High TOC II, Elementar, Elementar Analysensysteme GmbH, Germany). Inorganic carbon was removed by acidification with 2 M HCl addition prior to analysis. High performance liquid chromatography (HPLC; class VP with refractive index detector [RID] 10A and photo-diode array detector SPD-M10A VP detectors; Shimadzu, Japan) was used to determine the fatty acid concentrations. To further quantify other elements in the porewater (i.e. phosphorous and sulfur) the samples were acidified in 1% (v/v) HNO_3 and analyzed in triplicates by inductively coupled plasma mass spectrometry (ICP-MS/MS Agilent 8900). A flow injection analysis (FIA) instrument equipped with a dialysis membrane for

removal of Fe to prevent side reactions during measurement (Seal Analytical, Germany) was applied for quantification of NH_4^+, NO_3^- and NO_2^- concentrations.

2.6 Correlative SEM and nanoSIMS analysis

The free particles of the fine fraction of each organic horizon, transition zone and mineral horizon in cores Palsa A (referred to intact palsa), Palsa B (referred to more collapsed palsa) and Front (referred to collapsing front) along the palsa hillslope were analyzed using correlative SEM and nanoSIMS, as described previously[19,69,70]. Briefly, subsamples of each layer (1 mg) were dispersed in anoxic deionized water and gently shaken to obtain the free organo-mineral particles. All larger particles and aggregates were allowed to settle. A drop of 100 µl of the suspension was placed on a silica wafer and dried in an anoxic glovebox (N_2 atmosphere). Finally, samples were sputter-coated with ~30 nm Au/Pd conductive layer using a Bal-Tec SCD005 sputter coater (Baltec GmbH, Germany). To characterize the organo-mineral particles of the fine fraction by size and crystallinity and identify representative particles, a field emission scanning electron microscope (FE-SEM; Jeol JSM-6500F), equipped with secondary electron detector, was used prior to nanoSIMS analysis. The acceleration voltage was set to 5 kV, with a working distance of 10 mm. The nanoSIMS analysis were performed at the Cameca nanoSIMS 50L of the Chair of Soil Science (TU München, Germany). As described previously[19], a primary ion beam (~2 pA) was set at a lateral resolution ~100 nm and scanned over the samples with $^{12}C^-$, $^{16}O^-$, $^{12}C^{14}N^-$, $^{31}P^-$, $^{32}S^-$, $^{27}Al^{16}O^-$ and $^{56}Fe^{16}O^-$ secondary ions collected using electron multipliers.

2.7 Mössbauer spectroscopy

The soil samples for ^{57}Fe Mössbauer spectroscopy were collected under the protection of 100% N_2. Samples from three thaw stages were measured, including Palsa A, Bog and Fen (both wetland cores obtained by a previous campaign see Patzner *et al.*[19]) of transition zone and mineral

horizon (Figure S16 and Table S7). The samples were dried anoxically before loading into a Plexiglass holder. The prepared samples were stored anoxically at -20°C until measurement. Mössbauer spectroscopy was performed in a standard transmission setup (Wissel, Wissenschaftliche Elektronik GmbH), and absorption spectra were collected at 77 and 6 K controlling with a closed-cycle cryostat (SHI-850-I, Janis Research Co). The spectra were calibrated with $\alpha^{57}Fe^{0}$ foil at 295 K, and fitted using the Voigt Based Fitting (VBF)[71] routine in the Recoil software (University of Ottawa, Canada). Results are shown in the Figure S16 and Table S7.

2.8 TOC AND TN ANALYSIS

As described previously[19], total organic carbon (TOC) and total organic nitrogen (TN) were quantified by an Elementar vario El (Elementar Analaysysteme, GmbH, Germany). Soil samples were dried at 60°C until no further weight loss was observed. The dried soils were ground and acidified with 16% HCl to remove the inorganic carbon. After washing with deionized water, followed by drying, the TOC and TN content was analyzed. Results of C/N ratios are shown in the SI (Figure S10).

2.9 MICROBIAL COMMUNITY ANALYSIS

Total RNA and DNA was extracted using the PowerSoil® RNA and DNA isolation kit as described by the manufacturer (MO BIO Laboratories, Carlsbad, CA, USA), with the following modifications: 2-3 g of soil was used from each soil horizon; 10 min bead-beating; centrifugation steps at maximal speed (7000 x g) at 4°C; and longer incubation times at -20°C (1.5 h). RNA and DNA were eluted in 50 µl RNase/DNase-Free water. RNA and DNA concentrations were determined using a Qubit® 2.0 Fluorometer with RNA and DNA HS kits (Life Technologies, Carlsbad, CA, USA). Subsequent DNA digestion and reverse transcription reactions were

performed using a Reverse Transcriptase (Invitrogen, Life Technologies) as described previously by Otte *et al*., 2018[48]. Quantitative PCR (qPCR) specific for the 16S rRNA (gene) of bacteria and archaea was performed as described previously[48]. Microbial 16S rRNA (genes) were amplified using primers 515F and 806R[72]. Quality and quantity of the purified amplicons were determined using agarose gel electrophoresis and Nanodrop (NanoDrop 1000, Thermo Scientific, Waltham, MA, USA). Subsequent library preparation steps (Nextera, Illumina) and sequencing were performed by Microsynth AG (Switzerland) using the 2 × 250 bp MiSeq Reagent Kit v2 on an Illumina MiSeq sequencing system (Illumina, San Diego, CA, USA). From 10,112 to 396,483 (average 113,374) read pairs were generated per sample in three separate sequencing runs on the same MiSeq machine, resulting in total in 8.6 million read pairs. Quality control, reconstruction of 16S rRNA (gene) sequences and taxonomic annotation was performed with nf-core/ampliseq v1.1.2[73,74] with Nextflow v20.10.0[75] using containerized software with singularity v3.4.2[76]. Data from the three sequencing runs were treated initially separately by the pipeline using the option "mulipleSequencingRuns" and ASV tables were merged. Primers were trimmed, and untrimmed sequences were discarded (< 25%, on average 7.7%) with Cutadapt v2.6[77]. Adapter and primer-free sequences were imported into QIIME2 version 2019.10.0[78], processed with DADA2 version 1.10.0[79] to eliminate PhiX contamination, trim reads (before the median quality drops below 35, i.e. position 230 in forward reads and 174 in reverse reads), correct errors, merge read pairs, and remove PCR chimeras; ultimately, in total 9,576 amplicon sequencing variants (ASVs) were obtained across all samples. Alpha rarefaction curves were produced with the QIIME2 diversity alpha-rarefaction plugin, which indicated that the richness of the samples had been fully observed. A Naive Bayes classifier was fitted with 16S rRNA (gene) sequences extracted with the PCR primer sequences from the QIIME compatible, 99%-identity clustered SILVA v132 database[80].

ASVs were classified by taxon using the fitted classifier[81]. 225 ASVs that classified as chloroplasts or mitochondria were removed, totalling to < 7% (average 0.6%) relative abundance per sample, and the remaining 9,351 ASVs had their abundances extracted by feature-table (https://github.com/qiime2/q2-feature-table).

Pathways, i.e. MetaCyc ontology predictions, were inferred with PICRUSt2 version 2.2.0-b (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States)[82] and MinPath (Minimal set of Pathways)[83] using ASVs and their abundance counts. Inferring metabolic pathways from 16S rRNA amplicon sequencing data is certainly not as accurate as measuring genes by shotgun metagenomics, but it yields helpful approximations to support hypotheses driven by additional microbiological and biogeochemical analyses[82].

2.10 FT-ICR-MS ANALYSIS

Soil extracts and DOM in the porewater were analyzed with FT-ICR MS to identify and monitor compositional changes in the mineral-associated organic carbon fraction and the mobile, DOC fraction. All of the samples were prepared for FT-ICR-MS analysis by solid phase extraction (SPE) under N_2 atmosphere (glove bag) following the procedure described by Dittmar *et al.,* 2008[84] and Li *et al.*, 2016[85] with some modifications. In brief, 1 g, 3 mL Bond Elut PPL cartridges (part#12102148, Aglient Technologies, Santa Clara, CA, USA) were conditioned with 5 mL of HPLC grade methanol (Simga-Aldrich, Rehovot, Israel), followed by 5 mL of 0.01 M HCl. Each sample was acidified to pH ~2.5 and then loaded onto the SPE columns, loading volume was adjusted to load a total of 0.5 mg C based on the TOC content. After sample loading, the SPE cartridges were rinsed with 5 mL of 0.01 M HCl followed by drying with N_2 for 3-5 mins. Finally, the samples were eluted with 1 mL of HPLC grade methanol and stored in airtight amber sample

vials wrapped in aluminum foil at 4°C. There was no additional dilution of the samples performed prior to analysis by negative ion electrospray ionization.

The samples were analyzed with a custom-built FT-ICR mass spectrometer, equipped with a 21T superconducting solenoid magnet and a modular software package for data acquisition (Predator)[86]. Sample solution was infused via a microelectrospray source[87] (50 µm i.d. fused silica emitter) at 500 nL/min by a syringe pump. Typical conditions for negative ion formation were: emitter voltage, -3.0 kV; S-lens RF level, 45%; and heated metal capillary temperature, 350 °C. Ions were initially accumulated in an external multipole ion guide (1-5 ms) and released *m/z*-dependently[88]. Ions were excited to *m/z*-dependent radius to maximize the dynamic range and number of observed mass spectral peaks (32-64%)[89], and excitation and detection were performed on the same pair of electrodes[90]. The dynamically harmonized ICR cell in the 21 T FT-ICR is operated with 6 V trapping potential[91,92]. Time-domain transients (100 time-domain acquisitions for all experiments) of 3.1 seconds were acquired with the Predator data station that handled excitation and detection only, initiated by a TTL trigger from the commercial Thermo data station[93]. Mass spectra were phase-corrected[94] and internally calibrated with 10-15 highly abundant homologous series that span the entire molecular weight distribution (~150 to 1300 m/z) based on the "walking" calibration method[95]. Experimentally measured masses were converted from the International Union of Pure and Applied Chemistry (IUPAC) mass scale to the Kendrick mass scale[96] for rapid identification of homologous series for each heteroatom class[97]. Peaks with signal to noise ratios greater than 6 times the noise at the baseline root-mean-square (rms) noise at *m/z* 500 were exported to custom software (PetroOrg©) for additional formula and elemental composition assignment[98]. All assigned formulas were part of a ≥3 peak carbon series and had less than ±0.3 ppm mass error. A LOD of 6 σ was considered sufficient to minimize ionization

difference effects between samples, and therefore biasing by large numbers of low abundance peaks. To further identify macro compositional shifts, analysis of differences between samples was performed only on peaks with ≥20% difference in relative abundance. Additionally, modified aromaticity index (ModAI) was calculated according to Koch&Dittmar[84] and nominal oxidation state of carbon (NOSC) was calculated according to La Rowe&Van Cappellen[99]. Data processing post formula assignment was performed with RStudio utilizing R software (V4.0.3).

2.11 Statistical analysis

The geochemical parameters were checked with the test of homogeneity. Then a one-way analysis of variance (ANOVA) was used to identify differences in the geochemical parameters along the palsa hillslope, combined with a post-hoc test to identify significant differences between the different sampling spots along the palsa hillslope (from palsa to collapsing front to bog). Based on Gloor *et al.*[100] no statistical analysis (such as e.g. one-way ANOVA or unpaired t-test) was chosen for the compositional data obtained by 16S rRNA Amplicon (gene) sequencing.

3. Results and Discussion

3.1 Greenhouse gas emissions promoted by microbial iron cycling in thawing permafrost peatlands.

In the palsa and at the collapsing front (Figures S1 and S2), net CO_2 emissions measured from static flux chambers were similar on average (1.57±0.27 µmol/m^2/s) and slightly decreased in the bog to 1.12±0.51 µmol/m^2/s (Figure 1). Replicate analysis of CO_2 concentrations in automatic Eosense eosFD gas flux chambers showed similar CO_2 emissions along the palsa hillslope (Figure S3). Net CH_4 emissions were very low in the palsa (0.003±0.001 µmol/m^2/sec), significantly increased at the collapsing front to 0.025±0.001 µmol/m^2/s and then slightly decreased in the bog (0.013±0.001 µmol/m^2/s; Figure 1). Emission rates of CO_2 and CH_4 in the

palsa and the bog are in line with previous studies at Stordalen mire [37-39], however, this is the first report of emissions at the collapsing front, where palsa is collapsing into the bog.

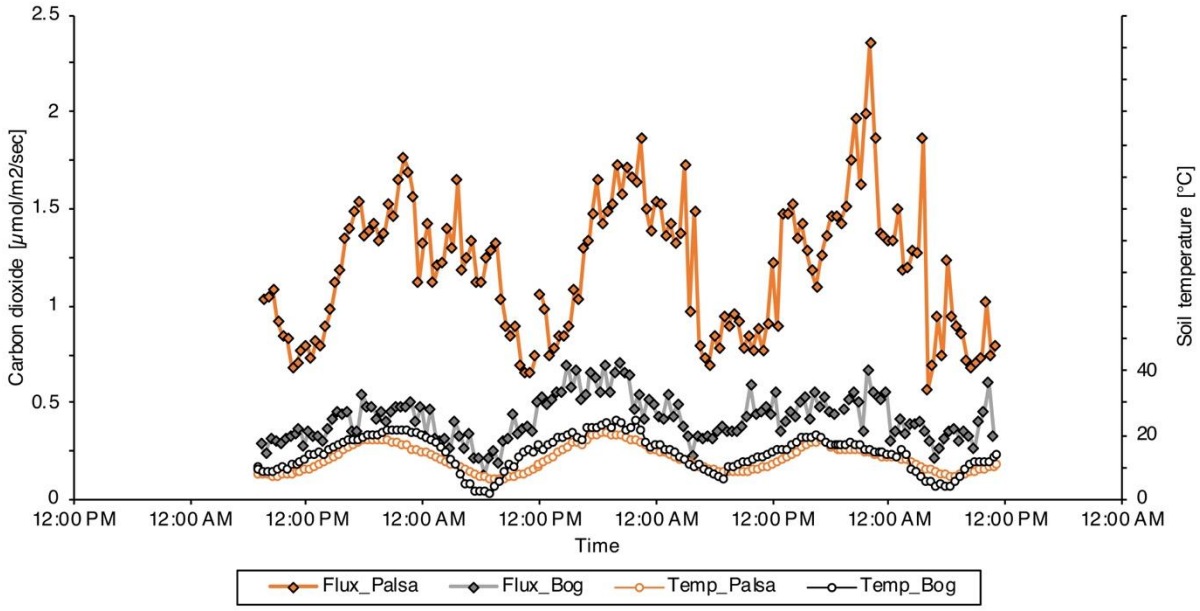


Figure S3. Carbon dioxide emissions along the palsa hillslope (palsa and bog). Eosense gas chambers (eosFD Forced Diffusion chamber in conjunction with the eosLink-FD software, EOSENSE INC, Dartmouth, Canada) were installed along the palsa hillslope and analysis performed from the 8th of July to 10th of July 2019. Unfortunately, the second Eosense instrument at the collapsing front (shown in Figure S2) was broken during shipment and thus excluded in the analysis. The carbon dioxide emissions correlate with the surface soil temperature (measured at 5 cm soil depth at palsa and bog), measured by Integrated Carbon Observation System (ICOS) Sweden Abisko – Stordalen1.

Previous work demonstrated that highest reactive Fe and associated OC contents can be found where the organic and mineral horizons meet, which we have termed the "transition zone"[19]. DOC concentrations in the porewater of the transition zone were low in the intact palsa (Palsa A, 57.97±16.49 mg/L). Porewater DOC significantly increased towards the collapsing front to 207.65±168.16 mg/L in the more collapsed palsa (Palsa B). Highest DOC concentrations were found directly at the collapsing front (535.75±131.45 mg/L) and then significantly decreased in the bog (279.62±113.14 to 206±80.28 mg/L) (Figure 1, Figure S1, Figure S4).

The aqueous Fe^{2+} concentrations show the same trend as the DOC (Figure 1). Aqueous Fe^{2+} concentrations in the palsa were lowest along the palsa hillslope (4.47±3.16 to 22.62±30.14 mg/L; Figure 1) and significantly increased at the collapsing front. Highest aqueous Fe^{2+} concentrations were measured at the collapsing front (153.24±40.14 mg/L) and significantly decreased again at the two measured locations in the bog to 48.86±11.43 and 82.43±47.93 mg/L (Figure 1). Other elements such as dissolved phosphorous (P) also strongly correlated with the aqueous Fe^{2+} pulse at the collapsing front, suggestive of mineral dissolution and release of mineral-associated P (Figures S5- S6).

The release of OC and aqueous Fe^{2+} along the palsa hillslope was accompanied by an increase in the relative 16S rRNA gene sequence abundance (DNA-based) of iron- and methane-cycling microorganisms in the transition zone and mineral horizon from the palsa to the collapsing front (Figure 1; Figure S7). Towards the collapsing front, Fe(III)-reducing bacteria increased from 0.41±0.07 to 2.46±0.34% in the transition zone and from 0.21±0.05 to 2.42±0.27% in the mineral horizon (Figure 1). Fe(II)-oxidizing bacteria also increased from the palsa to the collapsing front from 0.54±0.26 to 2.33±0.33% in the transition zone and from 0.92±0.58 to 1.66±0.44% in the mineral horizon. Methanogens increased along the palsa hillslope from 0.42±0.37 to 2.83±0.26% in the transition zone and from 1.40±1.40 to 11.68±3.12% in the mineral horizon. Methanotrophs increased from the palsa to the collapsing front from 0.90±0.30 to 1.93±0.09% in the transition zone and from 0.58±0.08 to 1.26±0.29% in the mineral horizon (Figure 1). Along the palsa hillslope, the relative 16S rRNA gene sequence abundances of iron- and methane-cycling microorganisms were stable in the organic horizon (Figure 1). The iron- and methane-cycling microorganisms are described in detail in Figure 2. For estimated absolute abundances of bacteria and archaea as well as the manually-compiled database used to identify iron- and methane-cycling

microorganisms and the whole microbial community see Figure 2 and SI (Figures S8 and S9 and Tables S1-S4).

This data reveals that the so-called rusty carbon sink is already destabilized during palsa collapse, even before complete permafrost thaw. Lateral flow by runoff of rain and/or melt water[40,41] in the transition zone between organic and mineral horizon, caused by bulk density shifts (organic horizon: 0.03±0.01 g/cm^3 and mineral horizon: 0.84±0.26 g/cm^3)[19], favors micro-oxic conditions, as also described for other permafrost hillslopes[42]. These redox conditions promote microbial reduction of reactive Fe(III) minerals coupled to carbon oxidation [14,43]. This results in a release of Fe and Fe-associated OC into the surrounding porewater and ultimately contributes to a pulse of aqueous Fe^{2+} and DOC at the collapsing front – where we observed the highest aqueous Fe^{2+} and DOC concentrations ever measured along the whole thaw gradient[19]. The release of OC along the palsa hillslope results from multiple co-occurring processes. These include the release of Fe-associated OC, changes in pH[44], plant community[45] (Figure S10), and in microbial degradation of organic matter[46].

Although Fe(II)-oxidizers are present and active, they cannot prevent the overall loss of reactive Fe and Fe-associated OC along the palsa hillslope. The CO_2 produced from degradation of released carbon, including Fe-associated-OC, further stimulated methanogenic microorganisms at the collapsing front. This CO_2 production was at least partially driven by Fe(III) reduction coupled to carbon oxidation based on the increasing abundance of Fe(III)-reducing bacteria along the palsa hillslope as has also been suggested for subalpine wetland soils[47].

Ultimately, the loss of this so called rusty carbon sink contributes to net GHG emissions of CO_2 and CH_4, directly by Fe(III) reduction coupled to carbon oxidation and indirectly by promoting methanogenesis at the collapsing front.

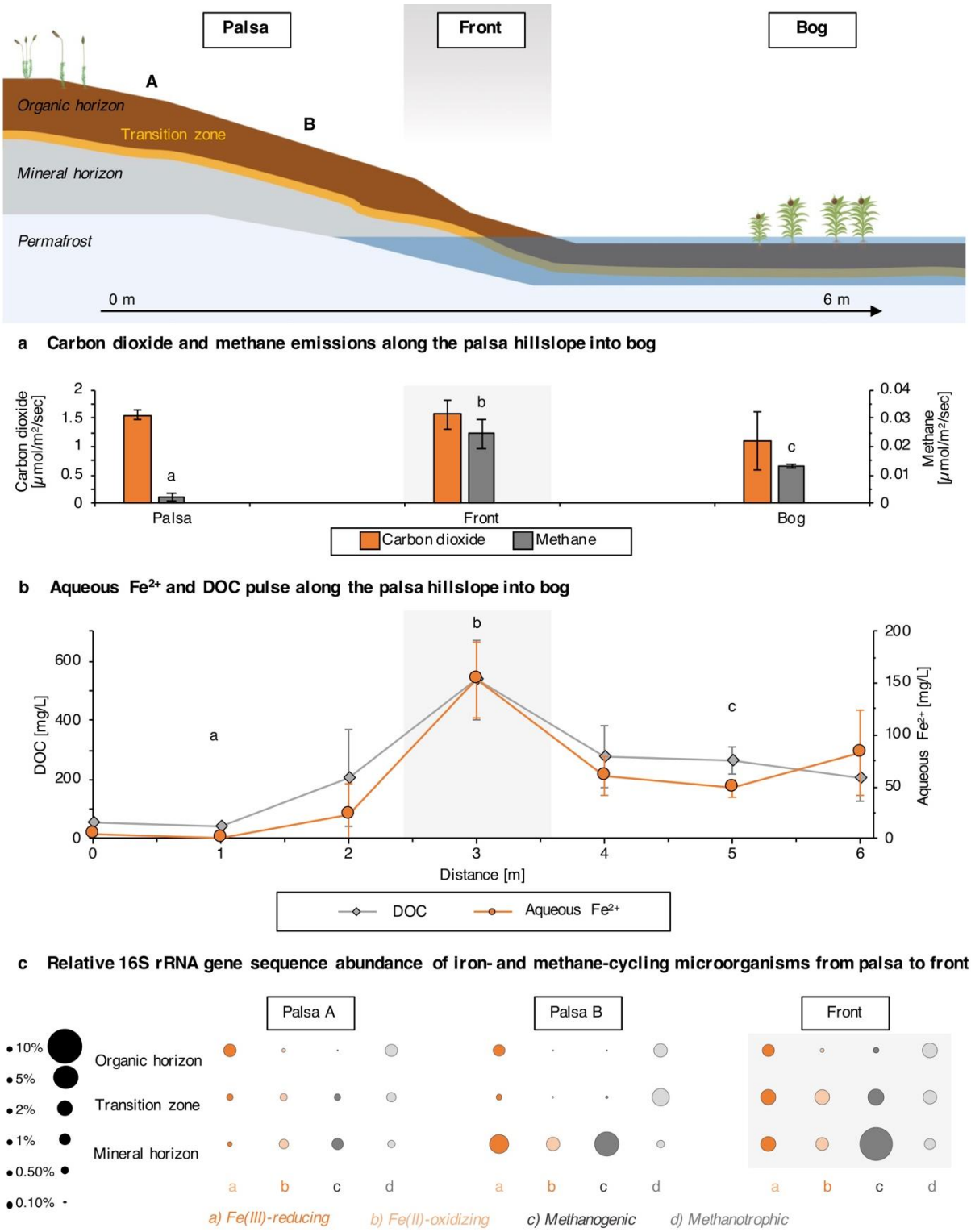


Figure 1. Microbial iron cycling and carbon release as dissolved organic carbon (DOC), carbon dioxide and methane emissions along a palsa hillslope. a, Carbon dioxide and methane emissions along the palsa hillslope with highest emissions at the collapsing front. The reported values and

error bars represent the average and standard deviation of measurements collected on three days at three separate time points. b, Aqueous iron (Fe^{2+}) and DOC pulse along the palsa hillslope at 30 cm depth with highest values at the collapsing front. Reported values and error bars represent the average and standard deviation of eight palsa to bog hillslopes sampled in June/July 2019 c, Relative 16S rRNA gene abundance of iron- and methane-cycling strains along the palsa hillslope with highest abundances at the collapsing front: a) Fe(II)-oxidizing, b) Fe(III)-reducing, c) methanogenic and d) methanotrophic. Small letters above data mean significant differences ($P<0.05$, one-way ANOVA: TukeyHSD test).

3.2 Microbial iron- and methane-cycling communities during palsa collapse.

Along the palsa hillslope, iron- and methane-cycling microorganisms increase in relative abundance, here defined as DNA-based relative 16S rRNA gene abundance, and in potential activity, here defined as RNA-based relative 16S rRNA abundance (Figure 2; for total microbial community and replicate analysis see Figure S9 and S11, Table S1-S4).

Fe(III)-reducers, driving reactive Fe mineral dissolution and associated OC release, are found in high abundance and potential activity along the palsa hillslope. From Palsa A to the collapsing front, *Geobacter* spp., a classical Fe(III)-reducer[24], increased in relative abundance from 0 to 1.55±0.30% in the transition zone and to 1.62±0.18% in the mineral horizon. The potential activity of *Geobacter* spp. rose from 0 to 2.50±0.13% in the transition zone and to 4.75±1.07% in the mineral horizon (Figure 2). *Clostridium* spp., a fermentative Fe(III)-reducer[53], increased in relative abundance from 0 to 0.81±0.02% in the transition zone and 0.76±0.07% in the mineral horizon (Figure 2). Potential activity of *Clostridium* spp. increased from 0 to 2.31±1.15% in the transition zone and to 1.23±0.22% in the mineral horizon (Figure 2). *Rhodoferax* spp., known for dissimilatory Fe(III) reduction[52], only appeared to be present (1.98±1.51%) and potentially active (1.62±0.16%) in the mineral horizon of the more collapsed palsa (Palsa B), close to the collapsing front (Figure 2). *Myxococcales* spp. showed highest relative abundance from 1.67±0.15% in the

intact palsa (Palsa A) to 1.30±0.23% at the collapsing front and potential activity from 9.13±0.08 in the intact palsa to 7.03±2.08% at the collapsing front in the organic horizon (Figure 2).

This microbial community analysis further indicates that the rusty carbon sink is lost by dissimilatory and fermentative Fe(III) reduction. Dissimilatory Fe(III) reduction is conducted along the palsa hillslope by abundant and active Fe(III)-reducing microorganisms such as *Geobacter* spp., *Rhodoferax* spp. and *Myxococcales* spp. (Figure 2; see also absolute abundances in Figure S8 and replicate core analysis in Figure S15)[48,49]. *Myxococcales* spp. are not only capable of Fe(III) reduction, but also e.g. polysaccharide and protein degradation[46]. *Geobacter* spp. and *Rhodoferax* spp. represent classical Fe(III)-reducing microorganisms, that are well studied in different environments[23] with *Rhodoferax* spp. also being described at other permafrost sites[14]. Fermentative Fe(III) reduction is probably performed by *Clostridium* spp. who might use the present DOC as carbon and energy source.

The abundant and active Fe(III)-reducing bacteria are accompanied by less relatively abundant and probably less active Fe(II)-oxidizers. *Gallionella* spp. had a relative abundance of 0.82±1.16% in the present microbial community and 1.42±1.92% in the active community of the mineral horizon of the more collapsed palsa (Palsa B). *Sideroxydans* spp. increased in their relative abundance from below detection to 1.42±0.21% in the transition and to 1.08±0.34% in the mineral horizon. Other *Gallionellaceae*, besides *Gallionella* spp. and *Sideroxydans* spp., were equally distributed in their relative abundance along the palsa hillslope from 0.54±0.26% in the transition zone and 0.86±0.55% in the mineral horizon of the intact palsa (Palsa A) to 0.90±0.12% in the transition zone and 0.58±0.09% in the mineral horizon at the collapsing front. The activity of the other *Gallionellaceae* was probably highest at the collapsing front with 0.53±0.24% in the transition zone and 0.35±0.07% in the mineral horizon. The classical Fe(II)-oxidizing bacteria[48,49]

such as *Gallionella* spp. and *Sideroxydans* spp., observed to be present and potentially active in this system were already described in arctic ponds[50]. In this setting, these cannot sustain or reform the rusty carbon sink during palsa collapse (Figure 2).

The increasing relative 16S rRNA (gene) abundance (DNA- and RNA-based) of classical Fe(III)-reducing bacteria is accompanied by an increase in the relative abundance of methanogenic microorganisms, mainly *Methanobacterium* spp. These significantly increased in their relative abundance in the transition zones from 0.25±0.24% in the intact palsa (Palsa A) to 2.05±0.14% at the collapsing front. In the mineral horizon, they rose in their relative abundance from 1.15±1.22% in the intact palsa (Palsa A) to 10.07±2.84% at the collapsing front (Figure 2). Along the palsa hillslope, only a slight increase in potential activity of *Methanobacterium* spp. was observed in the transition zone from 0 to 0.14±0.05% and in the mineral horizon from 0 to 1.91±0.85% (Figure 2). Other methanogens belonging to *Bathyarchaeia* also increased in relative abundance along the palsa hillslope from 0.17±0.13% to 0.71±0.12% in the transition zone and from 0.25±0.18% to 1.45±0.24% in the mineral horizon. Methanotrophs, such as *Roseiarcus* spp. and other *Beijerinckiacaeae* (i.e. *Methylobacterium* spp. or *Methylocystis* spp.) had an equal relative abundance in the community present along the palsa hillslope (i.e. DNA-based) and had its highest potential activity in the palsa closest to the collapsing front (Palsa B; 12.55±0.30%).

Acetate, formed along the palsa hillslope and accounted for up to 61.70±42.56 mg C/L (10.33% of the total DOC) at the collapsing front (Figure 4). It is expected that this stimulates Fe(III) reduction coupled to acetate oxidation and leads to CO_2 formation by Fe(III)-reducing bacteria such as *Geobacter* spp., known to metabolize acetate[23]. The potential for reductive acetogenesis from CO_2 by *Bathyarchaeia* was previously suggested[51]. Our MetaCyc ontology predictions showed a high potential for acetoclastic methanogenesis (Figure S12), but

contradictory to this, we only saw a high relative abundance of hydrogenotrophic methanogens such as *Methanobacterium* spp. This could be explained by the higher thermodynamic favorability of Fe(III) reduction coupled to acetate oxidation as compared to acetoclastic methanogenesis. H_2 and CO_2, partially produced by fermentation and Fe(III) reduction by e.g. *Clostridium* spp., can be used by hydrogenotrophic methanogens and lead to CH_4 emissions at the collapsing front. The CH_4 is partially oxidized back to CO_2 by methanotrophs as shown by Perryman *et al.*[22] who described highest methane oxidation rates for palsa at the transition between palsa and bog (here referred to as the collapsing front).

Our data clearly shows a co-existence of microbial iron- and methane-cycling microbial communities during palsa collapse, which ultimately cause GHG emissions and effect the balance between CO_2 and CH_4 emissions even before complete permafrost thaw.

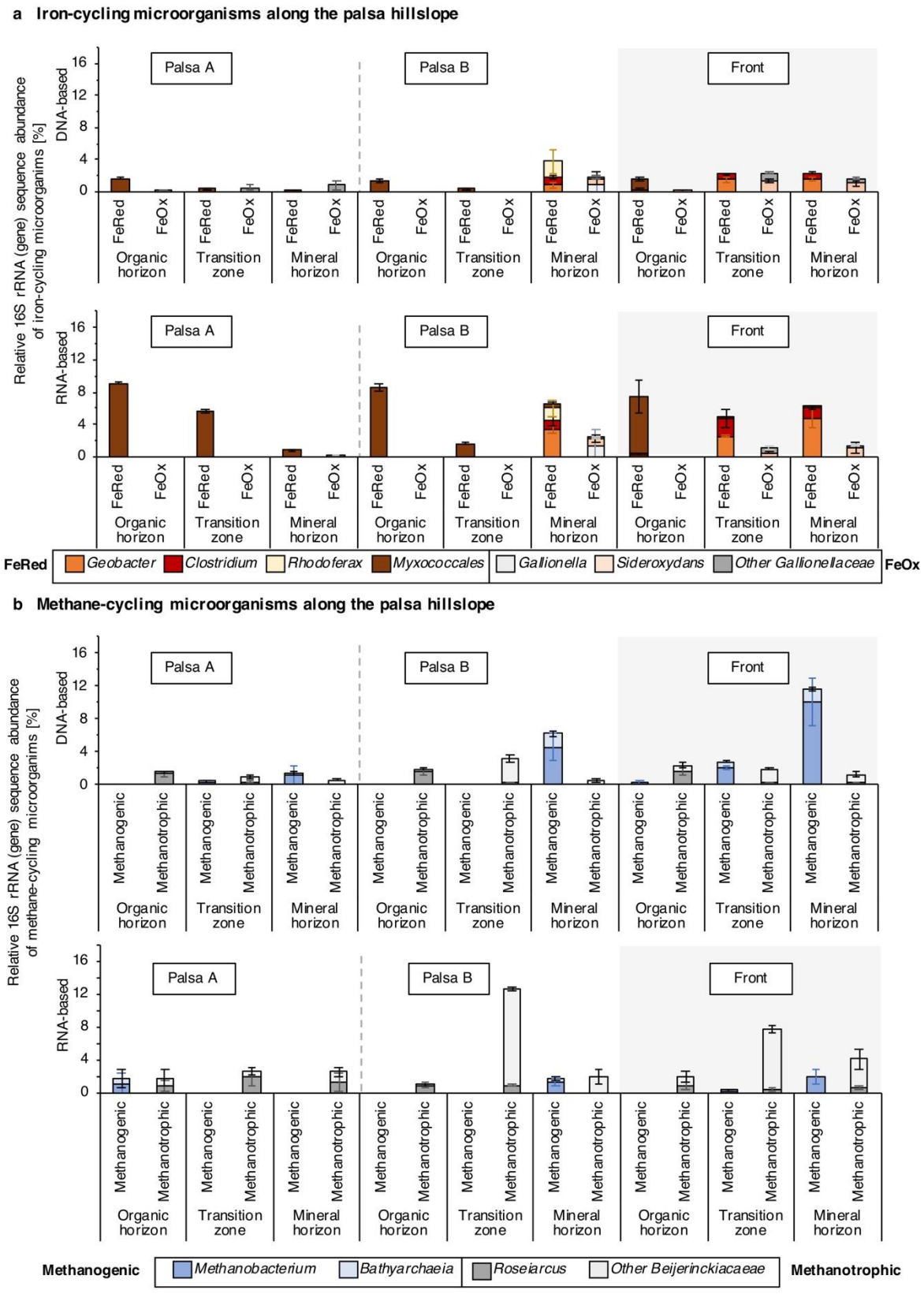

a Iron-cycling microorganisms along the palsa hillslope
Relative 16S rRNA (gene) sequence abundance of iron-cycling microorganisms [%]
DNA-based
RNA-based
Palsa A
Palsa B
Front
FeRed
FeOx
Organic horizon
Transition zone
Mineral horizon
Geobacter
Clostridium
Rhodoferax
Myxococcales
Gallionella
Sideroxydans
Other Gallionellaceae
b Methane-cycling microorganisms along the palsa hillslope
Relative 16S rRNA (gene) sequence abundance of methane-cycling microorganisms [%]
Methanogenic
Methanotrophic
Methanobacterium
Bathyarchaeia
Roseiarcus
Other Beijerinckiacaeae

Figure 2. Relative 16S rRNA (gene) abundance of abundant (DNA-based) and likely active (RNA-based) iron (Fe)- and methane-cycling microbial communities along the palsa hillslope (Palsa A, Palsa B, Front). a, Iron-cycling microorganisms show an increasing relative 16S rRNA (gene) abundance (DNA- and RNA-based) along the palsa hillslope with highest abundances in the transition zone and mineral horizon at the collapsing front. b, Methane-cycling microorganisms are increasing in relative 16S rRNA (gene) abundance along the palsa hillslope. Reported values and error bars represent the average and standard deviation of triplicate analysis of each soil horizon along the palsa hillslope. Replicate cores show similar relative 16S rRNA (gene) abundance of abundant (DNA-based) and potentially active (RNA-based) Fe- and methane cycling microbial communities along the palsa hillslope (Figure S11, Table S1-S4).

3.3 Release of bioavailable iron-associated organic carbon during palsa collapse.

To investigate if the loss of the rusty carbon sink also directly contributes to net GHG emissions by releasing bioavailable, previously Fe-bound organic carbon, into the porewater we determined the quantity and identity of Fe-bound OC in the solid phase (defined as dithionite extractable OC) and of the released OC in the porewater. Dithionite did not affect the identity of extractable OC and did not lead to molecular artifact formation (see SI, Table S5).

Highest reactive Fe concentrations (defined as Fe reductively dissolved by sodium dithionite and control corrected by leachable Fe, see Methods) were found in the transition zone of the most intact palsa (10.04±0.07 mg reactive Fe per g soil; Figure 3). Towards the collapsing front, reactive Fe in the transition zone between the organic and mineral horizons significantly decreased to 3.22±0.06 mg per g soil at the front (Figure 3). Absolute values are listed in Table S6 and replicate core analysis can be seen in Figure S13. The amount of reactive Fe-associated OC (OC dissolved after reductive dissolution of reactive Fe minerals by sodium dithionite and control corrected by leachable OC, see Methods) also decreased from the palsa to the bog in the transition zone (83.69±10.04 and 76.60±16.89 mg Fe-associated OC per g soil in the palsa to 40.88±10.76 mg per g soil in the bog) (Figure 2). In the organic horizons along the palsa hillslope, reactive Fe and Fe-associated OC abundance was the lowest in the soil profile with average values of

0.49±0.25 mg reactive Fe per g soil and 2.08±2.47 mg Fe-associated OC per g soil (Figure 3). In the mineral horizons from the palsa to the collapsing front, reactive Fe was very stable (average 3.81±0.38 reactive Fe per g soil), whereas Fe-associated OC slightly decreased from 47.21±14.30 mg Fe-associated OC per g soil in the palsa to the collapsing front which had only 11.60±8.54 mg Fe-associated OC per g soil (Figure 3). The highest content of Fe-associated OC was found in the most intact palsa along the palsa hillslope. This is supported by the strong spatial associations of OC with Fe minerals in the fine fraction observed by nanoSIMS analysis in the transition zone in this core (Figure 3; see replicate analysis of intact palsa core "Palsa A" in Figures S1, S14-S15). The transition zone and mineral horizons at the collapsing front showed organic-free, co-existing Fe and aluminum (Al), suggestive of Fe-bearing clays (Figure 3). This is also supported by Mössbauer spectroscopy (Figure S16, Table S7) and by previous observations with extended X-ray adsorption fine structure (EXAFS)[19].

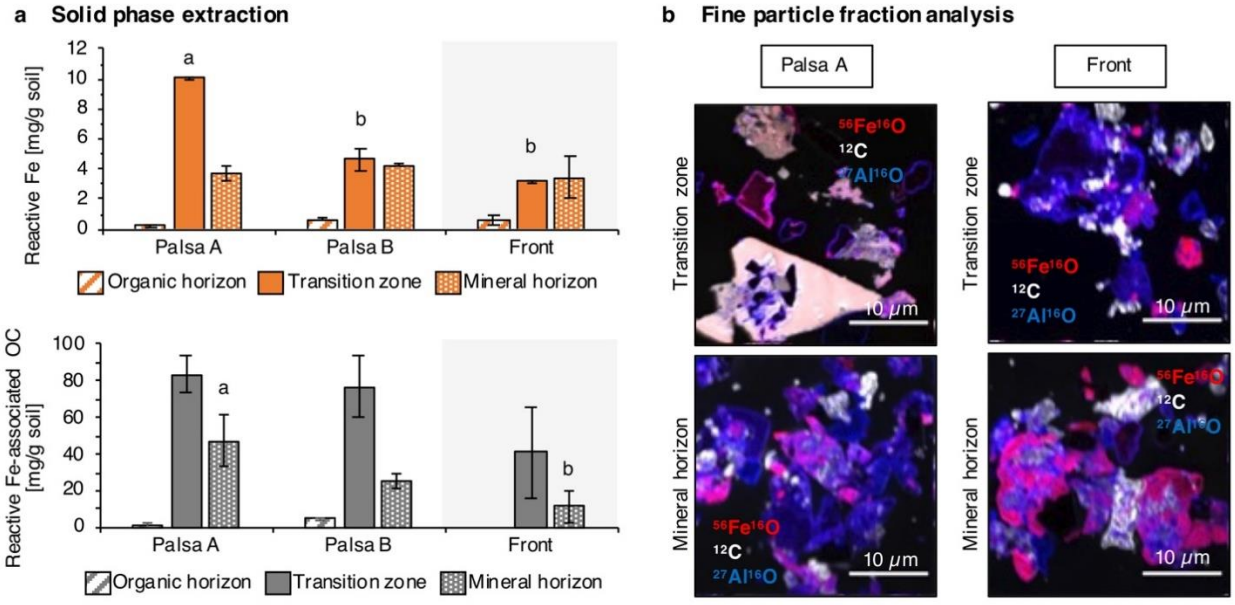


Figure 3. Reactive iron (Fe) and associated organic carbon (OC) from intact palsa to the collapsing front in a, the bulk soil and b, in the fine particle fraction. a, Reactive Fe and Fe-associated organic carbon in the solid phase decrease from the intact palsa (Palsa A) towards the collapsing front. Reactive Fe values are the average of sodium dithionite citrate duplicate extractions, control corrected by sodium chloride bicarbonate extractable Fe (leachable Fe). Fe-associated OC values

are the average of sodium dithionite citrate extractions, control corrected by subtraction of the citrate background and the sodium chloride bicarbonate extractable OC (leachable OC) (see Methods). Error bars of reactive Fe represent a combined standard deviation of sodium chloride bicarbonate extractable Fe and sodium dithionite citrate extractable Fe. Errors of the Fe-associated carbon represent a combined standard deviation of the citrate blank, sodium chloride bicarbonate extractable OC and sodium dithionite citrate extractable OC. Different small letters above bars mean significant differences ($P<0.05$, one-way ANOVA: TukeyHSD test). b, High spatial resolution analysis of Fe-OC associations by nanoSIMS in the fine fraction of the soil, displayed as $^{12}C^-$ (white), $^{56}Fe^{16}O^-$ (red) and $^{27}Al^{16}O^-$ (blue) overlaid in a composite image. For the two end-members, Palsa A and collapsing front, four particles of the fine fractions of each layer were analyzed by correlative SEM and nanoSIMS, all showing the same spatial distribution of Fe, C and Al as shown by the four representatives (Figure S14-S15).

FT-ICR-MS analyses showed that, in the intact palsa, the reactive Fe-associated OC had a higher relative abundance of aliphatic species than the reactive Fe-associated OC at the collapsing front (Figure 4, un-processed van Krevelen diagrams in Figure S17). This more aliphatic-like fraction could represent amino sugar-like, carbohydrate-like and lignin-like compounds (O/C range: 0.3 to 0.6, H/C range: 1.0 to 1.5) [52,53]. A higher fraction of aromatics was associated with reactive Fe phases at the collapsing front compared to the organics bound by reactive Fe in the intact palsa (Figure 4). In general, it should be noted that the amounts (mg/g) of reactive Fe-associated OC are decreasing along the palsa hillslope (Figure 3 and Table S5). The higher relative abundance of the more aliphatic compounds associated with reactive Fe in the intact palsa is lost during reductive dissolution to the surrounding porewater along the palsa hillslope, thus the aliphatic fraction most likely contributes to the aqueous Fe^{2+} and DOC pulse at the collapsing front (Figure 4). Loosely bound OC (salt extractable) appeared in lower quantities and showed less defined but similar identity of organic fractions to the reactive Fe-associated OC (Figure S18, Table S5 and S7).

Porewater extracted from the same soil interface (transition zone), where the rusty carbon sink is lost along the palsa hillslope (Figures 1 and 2), contained a higher relative abundance of

more aliphatic species and more aromatic species compared to porewater extracted at the collapsing front (Figure 4; un-processed van Krevelen diagrams in Figure S17). At the collapsing front, an increased relative abundance of organic molecules, potentially representing tannin-like compounds (O/C range: 0.5 to 0.9, H/C range: 0.5 to 1.4)[52,53], is observed (Figure 4 ; Figure S17). The more aliphatic species had a lower relative abundance in the DOC at the collapsing front, whereas a higher relative abundance of more aromatic species was observed (Figure 4).This could indicate decomposition processes that occur in the palsa hillslope porewater that yield smaller organic compounds, uptake by native microbes, assimilation of organic carbon into biomass and/or further metabolism, and ultimately emissions of GHGs by microbial respiration. Porewater analysis along a replicate palsa hillslope showed the same identity of aliphatic and aromatic species in intact palsa and at the collapsing front (Figure S19).

To provide a more comprehensive molecular understanding of these findings, additional detailed FT-ICR MS data and comparisons are presented in supporting figures. For instance, Figure S17 displays Van Krevelen diagrams for individual solid-phase extracted OC (dithionite extractable) from the transition zone of the intact palsa and collapsing front (transect 1, Figure S1), as well as DOC extracted from 30 cm depth in intact palsa and at the collapsing front. These diagrams illustrate the distinct compositional differences between mineral-associated organic matter (OM) and dissolved organic matter (DOM): specifically, DOM consistently shows higher proportions of formulae with elevated O/C ratios, while the mineral-extracted OM contains higher proportions of more condensed and aromatic formulae, regardless of thaw state.

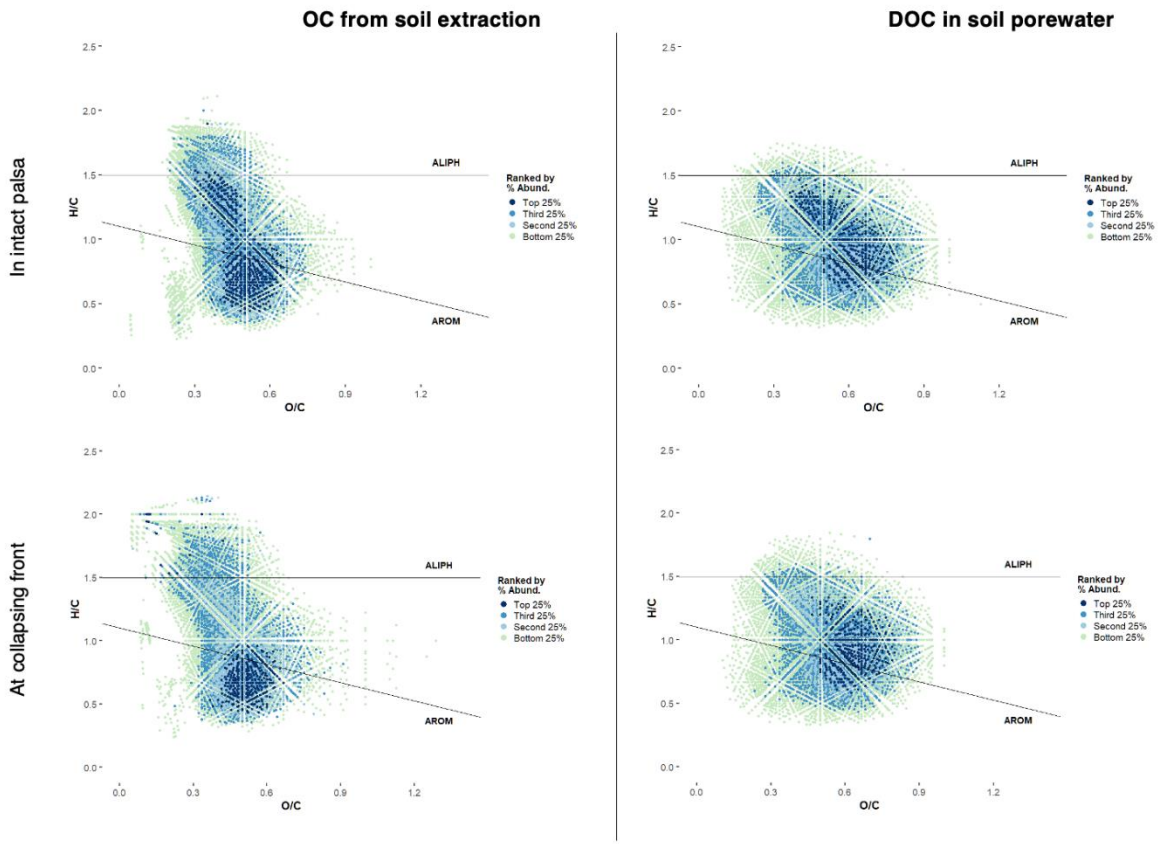


Figure S17. Van Krevelen diagrams for individual samples for solid phase extracted OC (dithionite extractable) from the transition zone of the intact palsa and of the collapsing front (transect 1, Figure S1) and DOC, extracted from 30 cm depth, in intact palsa and at the collapsing front.

To evaluate whether the dithionite extraction introduced artifacts in the DOM analysis and to determine the fraction of DOM released solely by reductive dissolution, the dithionite extraction was compared with a salt solution of comparable ionic strength and pH. Figure S18, which presents Van Krevelen diagrams comparing control extractions directly to the sodium dithionite citrate extractable CHO, demonstrates a distinct fraction extracted by sodium dithionite that generally consists of formulae with lower O/C ratios than those extracted by the salt solution.

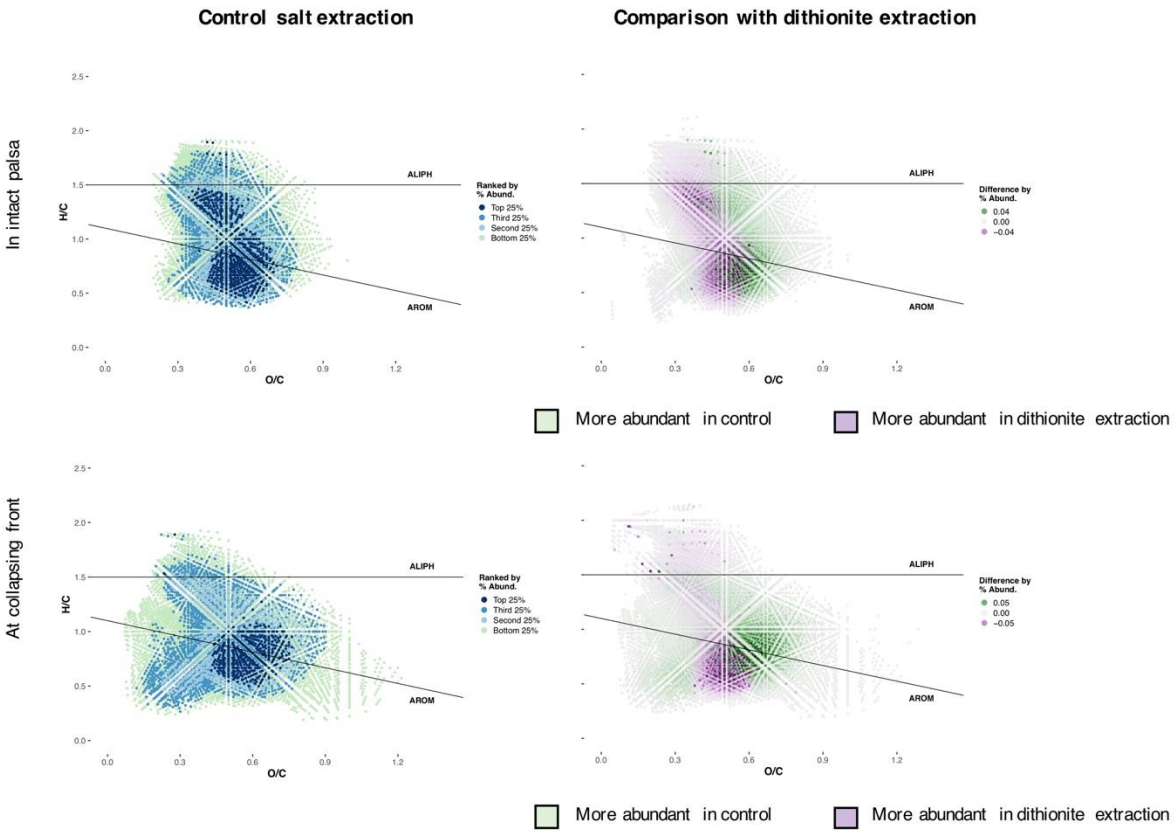


Figure S18. Van Krevelen diagrams of control extractions (sodium chloride bicarbonate with the same ionic strength and pH as the sodium dithionite extraction): left, each individual van Krevelen diagram and right, in direct comparison with the sodium dithionite citrate extractable CHO. Organic carbon (OC) which is more abundant in control (green) represents OC which is leachable of the soil by the same ionic strength and pH as the sodium dithionite extraction. OC which is more abundant in the sodium dithionite extraction (purple) is only dissolved by reductive dissolution.

To show that the observed shift in DOM composition is consistent, replicate porewater samples are compared in Figure S19, revealing the enrichment of more aliphatic compounds in Palsa A and the prevalence of more aromatic molecules towards the Front and Bog. Finally, Figure S20 details the molecular weight (MW) distribution of dissolved organic carbon compounds, highlighting how lower MW compounds are more abundant in intact palsa porewater, while higher MW compounds show higher relative abundance at the collapsing front.

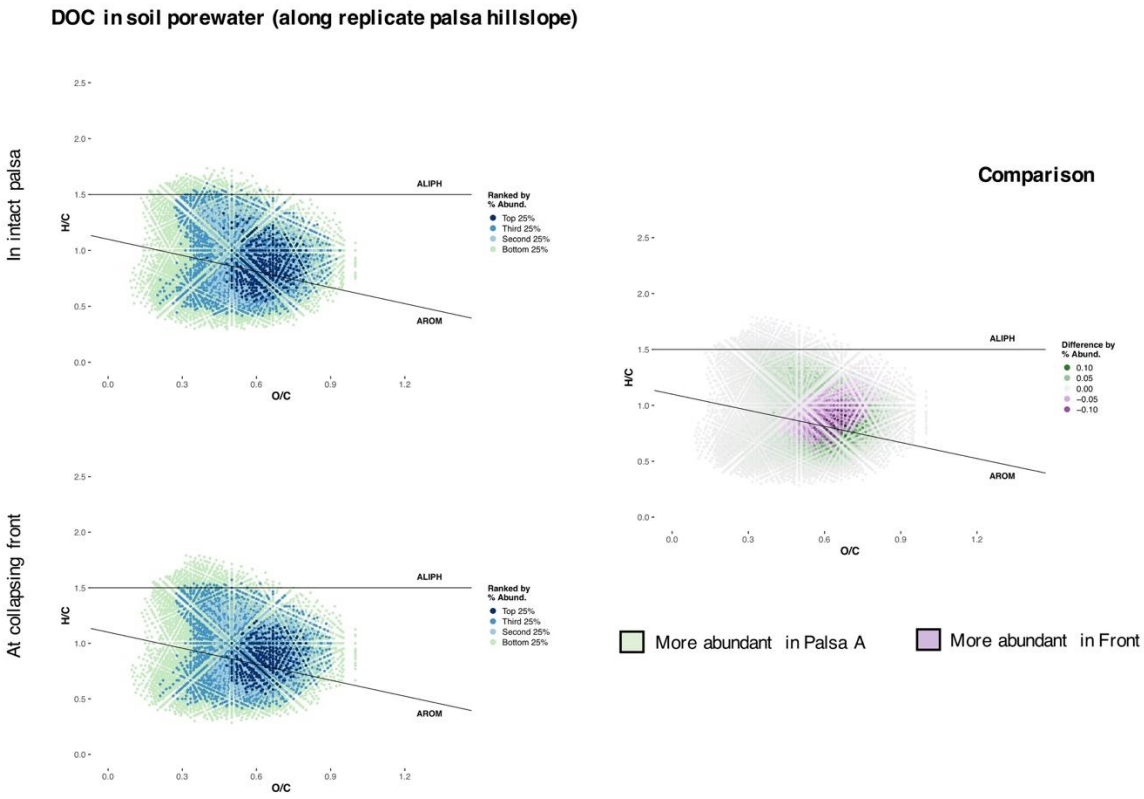


Figure S19. Van Krevelen diagrams for replicate porewater samples (30 cm depth) (left) and in direct comparison to each other (right). Dissolved OC, which is only found in Palsa A, is enriched in more aliphatic compounds (green). Towards Front and Bog, only more aromatic molecules (purple) remain.

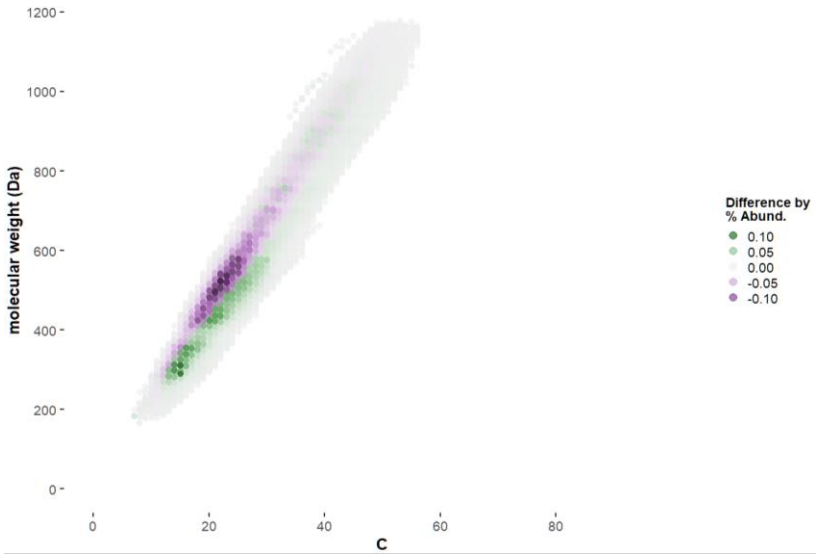


Figure S20. Molecular weight (MW) of dissolved organic carbon compounds in intact palsa versus at the collapsing front. Lower MW compounds have higher relative abundance in

porewater samples from intact palsa at a lower molecular weight for a given carbon number (colored in green). Higher MW compounds have higher relative abundance for a given carbon number at the collapsing front (colored in purple).

Further decomposition of DOC along the palsa hillslope is supported by an increasing nominal oxidation state of carbon (NOSC) of the DOC in the porewater at the transition zone from the palsa towards the collapsing front from 0.12±0.04 to 0.24±0.04 (Figure 4). As the DOC becomes more oxidized, the NOSC increases at the collapsing front. This is in line with an increasing average molecular weight (MW) from 591.24±7.70 in palsa to 614.80± 0.40 at the collapsing front (Figure 4 b and Figure S20). NOSC values slightly decreased in the bog to 0.20±0.02 due to the overall loss of organic carbon mainly as CO_2 and, consequently, enrichment of less decomposed and more reduced DOC in the porewater. The double-bond equivalents (DBE, the number of rings plus double bonds to carbon, calculated from the neutral elemental composition[54]), remained stable along the palsa hillslope (0.39±0.08). The DBE along the palsa hillslope showed lower values than previously reported for bog and fen[44], indicating that bog and fen DOC is overall more unsaturated compared to DOC released along the palsa hillslope.

The further decomposition of released organic carbon contributes to acetate formation (Figure 4) at the collapsing front, probably by pyruvate fermentation, indicated by MetaCyc ontology predictions (Figure S12). Along the palsa hillslope, acetate in the porewater at the transition zone between organic and mineral horizons significantly increased (unpaired *t*-test, N = 8, α = 0.05, p = 0.0024) from 6.24±0.34 mg C/L (3.56% of the total DOC) in the palsa to 61.70±42.56 mg C/L (10.33% of the total DOC) at the collapsing front, the highest acetate concentrations observed across the whole thaw gradient[19]. Further into the bog, the acetate concentrations significantly decreased from 15.13±6.53 to 6.10±1.44 mg C/L.

Previous studies at Stordalen mire focused on the soil organic carbon quantity and identity as well as on dissolved organic matter composition (DOM) and DOC transport along the thaw gradient. These analyses have highlighted shifts towards faster decomposition from partially-thawed bog to fully-thawed fen with an increasing proportion of carbon released as CH_4[40,44,55]. Processes occurring at the transition between palsa and bog had not been studied, thus enhanced production of acetate and its promotion of methanogenesis at this transition has been overlooked.

Our data showed that reactive Fe at the redox boundary between organic and mineral horizons can bind aliphatic organic carbon, probably by downward cycling of DOM (defined as continuous sorption and precipitation of DOM, as well as of microbial processing, desorption and dissolution proportions of more recent plant-derived compounds[56,57]) which is released during reductive dissolution into the surrounding porewater. Lower molecular weight compounds, aliphatic compounds or compounds poor in carboxyl functional groups show lower binding strength to Fe minerals than higher molecular weight compounds, aromatics, or compounds rich in carboxyl functionalities[57]. This is also supported by the leachable OC extractions (same ionic strength and pH as the sodium dithionite extraction; Figure S18). Thus, these compounds are not protected from microbial degradation along the palsa hillslope. The previously Fe-associated aliphatic fraction becomes more bioavailable to microorganisms when it is released from mineral associations[57]. This likely contributes to promotion of microbial growth and respiration of DOM during permafrost thaw[56,58-60]. Relative to aromatic compounds, aliphatic compounds are expected to be even more labile to microorganisms [56,59,60], which is supported by the overall loss of this more aliphatic carbon fraction in the porewater at the collapsing front (Figure 4) with only minor quantities of aromatic organic molecules remaining preserved by reactive Fe minerals after palsa

collapse (Figure 3 and Figure 4). Kawahigashi *et al.* showed that aromatic DOC was preferentially retained by mineral horizons in Siberian tundra containing reactive Fe[35].

Our data clearly suggests that the loss of this rusty carbon sink directly contributes to high DOC concentrations along the palsa hillslope and provides a bioavailable organic carbon source that stimulates microbial respiration and promotes GHG emissions.

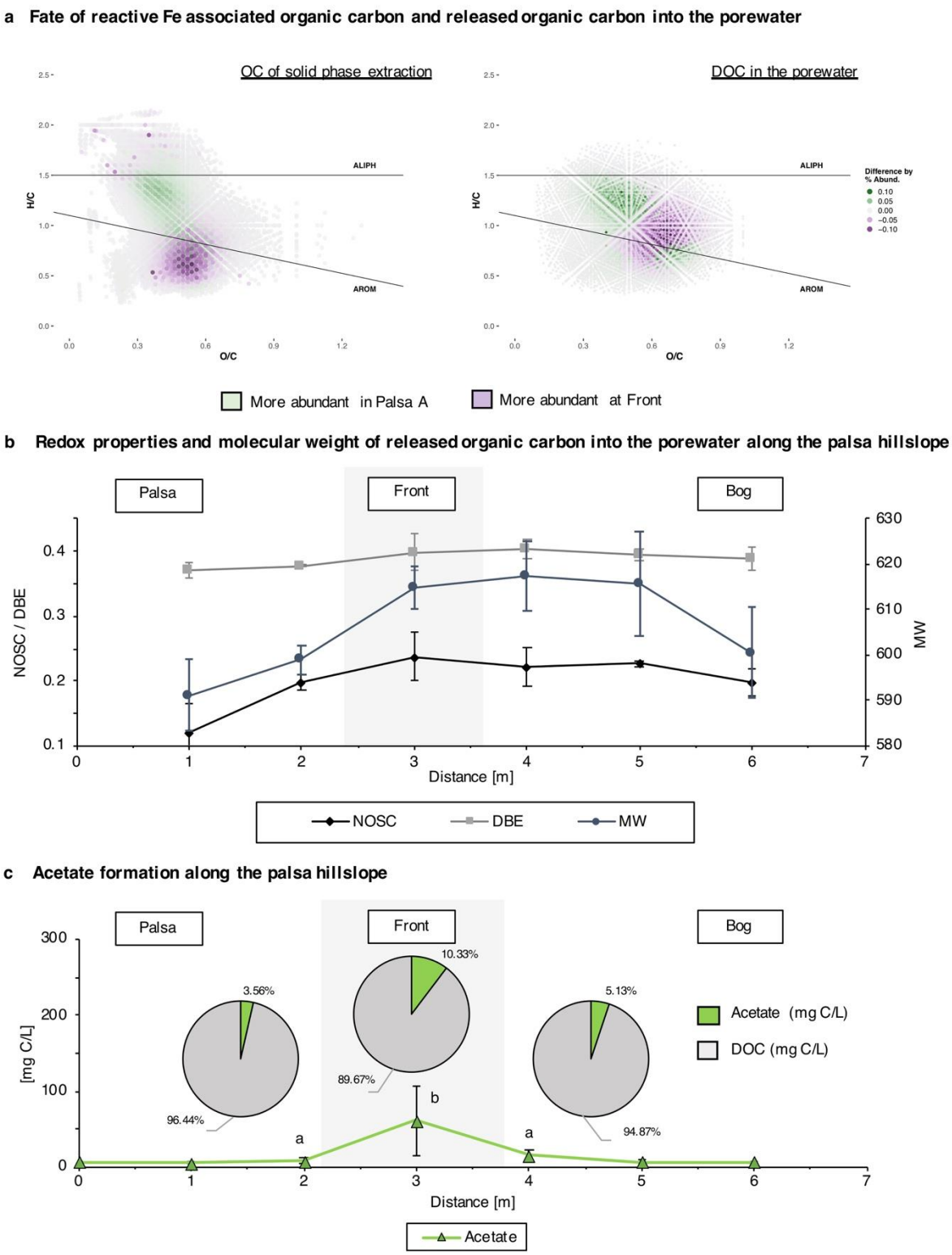


Figure 4. Bioavailability of reactive iron (Fe)-associated organic carbon (OC) released along the palsa hillslope. a, Composition of reactive Fe-associated OC and OC released into the porewater. Fe-bound carbon in palsa soils, underlain by intact permafrost, is comprised of more aliphatic

species (class 1, green) and more aromatic species (class 2, purple). This is lost with reductive dissolution into the porewater. Towards the collapsing front into the bog, the remaining Fe-associated OC fraction (purple) is comprised of less bioavailable organic compounds which are likely associated with clay minerals (common in Palsa A, Front and Bog). Dissolved OC, which is only found in Palsa A, is enriched in more aliphatic compounds (green). Towards Front and Bog, only more aromatic species (purple) remain (Figures S17-S19, Table S5). B, Redox properties and molecular weight of released organic carbon into the porewater along the palsa hillslope (Figure S20). Reported values and error bars represent the average and the range of duplicate porewater analysis along two palsa hillslopes (Figures S1 and Figure S4). a, Acetate formation along the palsa hillslope. Following further decomposition of the dissolved OC, highly bioavailable acetate [mg C/L] is formed which then is again used to further reduce present reactive Fe(III) to Fe^{2+} coupled to acetate oxidation and CO_2 formation. Reported values and error bars represent the average and standard deviation of 8 palsa to bog hillslopes, sampled in June/July 2019 (Figures S1 and Figure S4). Different small letters above data mean significant differences ($P<0.05$, one-way ANOVA: TukeyHSD test).

3.4 Implications for the carbon cycle in thawing permafrost peatlands.

There is a substantial need to piece together carbon sources and sinks in thawing permafrost environments to better understand and quantitatively predict the overall climate impact of permafrost thaw[61]. One such carbon sink or source are Fe-OC associations[36], which sequesters organic carbon in intact permafrost soils[20] but releases it with complete permafrost thaw[19]. Our data now shows that the release of the OC from the rusty carbon sink turns the OC into a source of labile DOC, CO_2 and CH_4 even before permafrost-supported palsas have completely collapsed. With increasing abrupt thaw, occurring in 20% of the permafrost zone, new active hillslope features are formed[62] and thus could speed up the loss of the rusty carbon sink in currently intact permafrost environments. Newest estimates showed that collapsing fronts will occupy 3% of abrupt thaw terrain by 2300, but could emit one-third of abrupt thaw carbon losses[62]. Ultimately, interlinked processes of iron- and carbon cycling in thawing permafrost environments need to be integrated into existing climate models to better understand and predict GHG emissions of thawing permafrost areas and thus better estimate its overall climate impact. For this, it is crucial to further

determine co-occurring Fe(III) reduction rates and CO_2 and CH_4 production rates following Fe mineral dissolution.

4. AUTHOR CONTRIBUTIONS

The original hypothesis was formulated by M.S.P., C.B. and A.K. M.S.P, C.B. and A.K. designed the project, interpreted the data and wrote the manuscript. M.S.P, C.B. and M.L. collected the samples. M.S.P and M.L. gathered the data presented in the main text. M.L. processed, interpreted, and generated all results and figures for the FT-ICR MS analysis. M.S.P. processed, interpreted, and generated all results and figures for the geochemical analysis, SEM and nano SIMS, TOC and TN, and microbial community analysis. A.M. conducted the FT-ICR-MS measurements and contributed to the data interpretation. T.B. and R.Y. contributed to the data analysis and interpretation. Z.Z. performed the Mössbauer spectroscopy and helped interpreting the results. H. J. helped collecting the porewater samples and data interpretation. C.H. and C.W.M., together with M.S.P., collected, analyzed and interpreted the nanoSIMS data. D.S. processed the amplicon sequencing data and, together with S.K., helped with interpretation of the microbial community results. T.S. contributed to project design and data interpretation. All authors contributed to the preparation of the manuscript and have given approval to the final version of the manuscript.

5. ACKNOWLEDGEMENTS

We are grateful for assistance in field work and sampling by Sara Anthony and the Arctic Research station (Abisko Sweden), especially Jennie Wikström, Eric Lundin, Niklas Rakos and Alexander Meire, as well as the Swedish Polar Research Secretariat and SITES for the support of the work done at the Abisko Scientific Research Station. We are thankful to Mette Svenning

(Arctic University of Norway, Tromsø, Norway) for pro- viding the gas chambers with the plastic lids. We thank EOSENSE INC, Dartmouth, Canada for getting the chance to use the EOSENSE gas chambers for our research. We thank Sara Anthony for analysis of gas samples, Miroslava Malusova and Katrin Wunsch for assistance in the lab, Timm Bayer for his help during SEM analysis and Johann Lugmeier for assistance during nanoSIMS analysis. The authors acknowledge infra- structural support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy, cluster of Excellence EXC2124, project ID 390838134. This work was supported by the University of Tuebingen (Pro- gramme for the Promotion of Junior Researchers award to Casey Bryce) and by the German Academic Scholarship Foundation (scholarship to Monique Patzner). Sara Kleindienst is funded by an Emmy-Noether fellowship from the DFG (grant number 326028733). Daniel Straub was funded by the Institutional Strategy of the University of Tuebingen (DFG, ZUK63). The National High Magnetic Field Laboratory is supported by the National Science Foundation Divisions of Chemistry and Materials Research through DMR-1644779, Florida State University, and the State of Florida.

6. COPYRIGHT

7. REFERENCES

1 Ping, C. L., Jastrow, J. D., Jorgenson, M. T., Michaelson, G. J., Shur, Y. L. Permafrost soils and carbon cycling *Soil* **1**, 147-171 (2015).

2 Schuur, E. A. G. *et al.* Climate change and the permafrost carbon feedback. *Nature* **520**, 171-179 (2015).

3 Hugelius, G. *et al.* Estimated stocks of circumpolar permafrost carbon with quantified uncertainty ranges identified data gaps. *Biogeosciences* **11**, 6573-6593 (2014).

4 Nowinski, N., Taneva, L., Trumbore, S., Welker, J. Decomposition of old organic matter as a result of deeper active layers in a snow depth manipulation experiment. *Oecologia* **163**, 785-792 (2010).

5 Vonk, J. E., Mann, P. J., Davydov, S., Davydova, A., Spencer, R. G. M., Schade, J. High biolability of ancient permafrost carbon upon thaw. *Geophys Res Lett* **40**, 2689-2693 (2013).

6 Henkner, J., Scholten, T., Kühn, P. Soil organic carbon stocks in permafrost-affected soils in West Greenland. *Geoderma* **282**, 147-159 (2016).

7 Hemingway, J. D., Rothman, D.H., Grant, K.E., Rosengard, S.Z., Eglinton, T.I., Derry, L.A., Galy, V.V. Mineral protection regulates long-term global preservation of natural organic carbon. *Nature* **570**, 228-231 (2019).

8 Kleber, M., Mikutta, R., Torn, M. S., Jahn, R. Poorly crystalline mineral phases protect organic matter in acid subsoil horizons. *Eur J Soil Sci* **56**, 717-725 (2005).

9 Baldock, J. A., Skjemstad, J.O. Role of soil matrix and minerals in protecting natural organic materials against biological attack. *Org Geochem* **31**, 697-710 (2000).

10 Kleber, M. *et al.* Mineral-Organic Associations: Formation, Properties, and Relevance in Soil Environments. *Adv Agron* **130**, 1-140 (2015).

11 Lalonde, K., Mucci, A., Ouellet, A., Gelinas, Y. Preservation of organic matter in sediments promoted by iron. *Nature* **483**, 198-200 (2012).

12 Wagai, R., Mayer, L. M., Kitayama, K., Shirato, Y. Association of organic matter with iron and aluminum across a range of soils determined via selective dissolution techniques coupled with dissolved nitrogen analysis. *Biogeochemistry* **112**, 95-109 (2013).

13 Kogel-Knabner, I. *et al.* Organo-mineral associations in temperate soils: Integrating biology, mineralogy, and organic matter chemistry. *J Plant Nutr Soil Sc* **171**, 61-82 (2008).

14 Lipson, D. A., Jha, M., Raab, T. K., Oechel, W. C. Reduction of iron (III) and humic substances plays a major role in anaerobic respiration in an Arctic peat soil. *J Geophys Res-Biogeo* **115**, G00I06 (2010).

15 Lovely, D. R., Phillips, E. J. Novel mode of microbial energy metabolism: organic carbon oxidation coupled to dissimiliatory reduction of iron or mangenese. *Appl Environl Microbiol* **57**, 1536-1540 (1988).

16 Asano, M. & Wagai, R. Evidence of aggregate hierarchy at micro- to submicron scales in an allophanic Andisol. *Geoderma* **216**, 62-74005 (2014).

17 Totsche, K. U. *et al.* Microaggregates in soils. *J Plant Nutr Soil Sc* **181**, 104-136 (2018).

18 Coward, E. K., Thompson, A. T. & Plante, A. F. Iron-mediated mineralogical control of organic matter accumulation in tropical soils. *Geoderma* **306**, 206-216 (2017).

19 Patzner, M. S. *et al.* Iron mineral dissolution releases iron and associated organic carbon during permafrost thaw. *Nat Commun* **11,** 6329 (2020).

20 Mu, C. C. *et al.* Soil organic carbon stabilization by iron in permafrost regions of the Qinghai-Tibet Plateau. *Geophys Res Lett* **43**, 10286-10294 (2016).
21 Shelef, E., Rowland, J.C., Wilson, C.J., Hilley, G.E., Mishra, U., Altmann, G.L., Ping, C.L. Large uncertainity in permafrost carbon stocks due to hillslope soil deposits. *Geophys Res Lett* **44**, 6134-6144 (2017).
22 Perryman, C. K. M. *et al.* Thaw Transitions and Redox Conditions Drive Methane Oxidation in a Permafrost Peatland. *J Geophys Res-Biogeo* **125** (2020).
23 Kappler, A., Bryce, C., Mansor, M., Lueder, U., Byrne, J.M., Swanner, E. An evolving view on biogeochemical cycling of iron. *Nat Rev Microbiol* (2021).
24 Van Bodegom, P. M., Scholten, J. C. M.; Stams, A. J. M. Stams. Direct inhibition of methanogenesis by ferric iron. *FEMS Microbiol Ecol* **49**, 261-268 (2004).
25 Wagner, R., Zona, D., Oechel, W., Lipson, D. Microbial community structure and soil pH correspond to methane production in Arctic Alaska soils. *Method Enzymol* **19**, 3398-3410 (2017).
26 Baldock, J. A. & Skjemstad, J. O. Role of the soil matrix and minerals in protecting natural organic materials against biological attack. *Org Geochem* **31**, 697-7102000).
27 Christensen, B. T. Physical fractionation of soil and structural and functional complexity in organic matter turnover. *Eur J Soil Sci* **52**, 345-353 (2001).
28 Sanderman, J., Maddern, T., Baldock, J. Similar composition but differential stability of mineral retained organic matter across four classes of clay minerals. *Biogeochemistry* **121**, 409-424 (2014).
29 Six, J. *et al.* Sources and composition of soil organic matter fractions between and within soil aggregates. *Eur J Soil Sci* **52**, 607-618 (2001).
30 Daugherty, E. E. *et al.* Hydrogeomorphic controls on soil carbon composition in two classes of subalpine wetlands. *Biogeochemistry* **145**, 161-175 (2019).
31 Crump, B., Kling, G., Bahr, M., Hobbie, J. Bacterioplankton community shifts in arctic lake corrleate with seasonal changes in organic matter sources. *Appl Environ Microbiol* **69**, 2253-2268 (2003).
32 Judd, K. E., Crump, B.C., Kling, G.W. Variation in dissolved organic matter controls bacterial production and community composition. *Ecology* **87**, 2068-2079 (2006).
33 Zhao, Q. *et al.* Iron-bound organic carbon in forest soils: quantification and characterization. *Biogeosciences* **13**, 4777-4788 (2016).
34 Adhikari, D. & Yang, Y. Selective stabilization of aliphatic organic carbon by iron oxide. *Sci Rep-Uk* **5**, 11214 (2015).
35 Kawahigashi, M., Kaiser, K., Rodionov, A., Guggenberger, G. Sorption of dissolved organic matter by mineral soils of the Siberian forest tundra. *Global Change Biol* **12**, 1868-1877 (2006).
36 Eglinton, T. I. GEOCHEMISTRY A rusty carbon sink. *Nature* **483**, 165-166 (2012).
37 McCalley, C. K. *et al.* Methane dynamics regulated by microbial community response to permafrost thaw. *Nature* **514**, 478-481 (2014).
38 Mondav, R. *et al.* Discovery of a novel methanogen prevalent in thawing permafrost. *Nat Commun* **5**, 3212 (2014).
39 Singelton, C. M *et al.* Methanotrophy across a natural permafrost thaw environment. *The ISME Journal* **12**, 2544-2558 (2018).

40 Olefeldt, D. & Roulet, N. T. Effects of permafrost and hydrology on the composition and transport of dissolved organic carbon in a subarctic peatland complex. *J Geophys Res-Biogeo* **117**, G01005 (2012).
41 Tang, J. *et al.* Drivers of dissolved organic carbon export in a subarctic catchment: Importance of microbial decomposition, sorption-desorption, peatland and lateral flow *Sci Total Environ* **622-623***,* 260-274 (2018).
42 Lynch, L. M. *et al.* Dissolved organic matter chemistry and transport along an Arctic tundra hillslope. *Global Biogeochem Cy* **33**, 47-62 (2019).
43 Lipson, D. A., Raab, T.K., Goria, D., Zlamal, J. The contribution of Fe(III) and humic acid reduction to ecosystem respiration in drained thaw lake basins of the Arcitc Coastal Plain. *Global Biogeochem Cy* **27**, 399-409 (2013).
44 Hodgkins, S. B. *et al.* Changes in peat chemistry associated with permafrost thaw increase greenhouse gas production. *P Natl Acad Sci USA* **111**, 5819-5824 (2014).
45 Johansson, T. *et al.* Decadal vegetation changes in a northern peatland, greenhouse gas fluxes and net radiative forcing. *Global Change Biol* **12**, 2352-2369 (2006).
46 Woodcroft, B. J. *et al.* Genome-centric view of carbon processing in thawing permafrost. *Nature* **560**, 49-54 (2018).
47 Pallud, C., Rhoades, C. C., Schneider, L., Dwivedi, P., Borch, T. Temperature-induced iron (III) reduction results in decreased dissolved organic carbon export in subalpine wetland soils, Colorado, USA. *Geochim Cosmochim Ac* **280**, 148-160 (2020).
48 Otte, J. M. *et al.* The distribution of active iron-cycling bacteria in marine and freshwater sediments is decoupled from geochemical gradients. *Method Enzymol* **20**, 2483-2499 (2018).
49 Weber, K. A., Achenbach, L. A., Coates, J. D. Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction. *Nat Rev Microbiol* **4**, 752-764 (2006).
50 Emerson, D., Scott, J. J., Benes, J., Bowden, W. B. Microbial iron oxidation in the Arctic tundra and its implications for biogeochemical cycling. *Appl Environ Microbiol* **81**, 8066-8075 (2015).
51 Loh, H. Q., Hervé, V., Brune, A. Metabolic potential for reductive acetogenesis and a novel energy-converting [NiFe] Hydrogenase in Bathyarchaeia from termite guts - a genome-centric analysis. *Front Microbiol* **11,** 635786 (2021).
52 Tfaily, M. M.. Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry. *Anal Chem* **87**, 5026-5215 (2015).
53 Tfaily, M. M. *et al*. Sequential extraction protocol for organic matter from soils and sediments using high resolution mass spectrometry. *Anal. Chim. Acta* **972**, 54-61 (2017).
54 McLafferty, F. W., Choi, J., Tureček, F. & Turecek, F. Interpretation Of Mass Spectra. University Science Books, ISBN 0-935702-25-3 (1993).
55 AminiTabrizi, R. *et al.* Controls on soil organic matter degradation and subsequent greenhouse gas emissions across a permafrost thaw gradient in Northern Sweden. *Frontiers in Earth Science* **8**, 557961 (2020).
56 Ward, P. C., Cory, R. M. Chemical composition of dissolved organic matter draining permafrost soils. *Geochim Cosmochim Ac* **167**, 63-67 (2015).
57 Kaiser, K., Kalbitz, K. Cycling downwards - dissolved organic matter in soils. *Soil Biol Biochem* **52**, 29-32 (2012).
58 Fuchs, G., Boll, M., Heider, J. Microbial degradation of aromatic compounds - From one strategy to four. *Nat Rev Microbiol* **9**, 803-816 (2011).

59 Abbott, B. W., Larouche, J. R., Jones, J. B., Bowden, W. B., Balser, A. W. Elevated dissolved organic carbon biodegradability from thawing and collapsing permafrost. *J Geophys Res* **119**, 2049-2063 (2014).
60 Mann, P. J., Davydova, A., Zimov, N., Spencer, R.G.M., Davydov, S., Bulygina, E., Zimov, S., Holmes, R.M. Controls on the composition and lability of dissolved organic matter in Siberia's Kolyma River basin. *J Geophys Res* **117** (2012).
61 When permafrost thaws. *Nat Geosci* **13**, 765 (2020).
62 Turetsky, M. R. *et al.* Carbon release through abrupt permafrost thaw. *Nat Geosci* **13**, 138-143 (2020).
63 Malmer, N., Johansson, T., Olsrud, M., Christensen, T. R. Vegetation, climatic changes and net carbon sequestration in a North-Scandinavian subarctic mire over 30 years. *Global Change Biol* **11**, 1895-1909 (2005).
64 Fronzek, S., Carter, T. R., Raisanen, J., Ruokolainen, L., Luoto, M. Applying probabilistic projections of climate change with impact models: a case study for sub-arctic palsa mires in Fennoscandia. *Climatic Change* **99**, 515-534 (2010).
65 Parviainen, M., Luoto, M. Climate envelopes of mire complex types in Fennoscandia. *Geogr Ann A* **89a**, 137-151 (2007).
66 Liebner, S. *et al.* Shifts in methanogenic community composition and methane fluxes along the degradation of discontinuous permafrost. *Front Microbiol* **6**, 356 (2015).
67 Glatzel, S., Well, R. Evaluation of septum-capped vials for storage of gas samples during air transport. *Environ Monit Assess* **136**, 307-311 (2008).
68 Stookey, L. L. Ferrozine - a New Spectrophotometric Reagent for Iron. *Anal Chem* **42**, 779-781 (1970).
69 Kopittke, P. M. *et al.* Nitrogen-rich microbial products provide new organo-mineral associations for the stabilization of soil organic matter. *Global Change Biol* **24**, 1762-1770 (2018).
70 Keiluweit, M. *et al.* Nano-scale investigation of the association of microbial nitrogen residues with iron (hydr)oxides in a forest soil O-horizon. *Geochim Cosmochim Ac* **95**, 213-226 (2012).
71 Rancourt, D. G., Ping, J.Y. Voigt-based methods for arbitrary-shape static hyperfine parameter distributions in Mössbauer spectroscopy. *Nucl Instrum Methods Phys Res Sect B* **58**, 85-97 (1991).
72 Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-336 (2010).
73 Ewels, P. A. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* **38**, 276-278 (2020).
74 Straub, D. *et al.* Interpretations of environmental microbial community studies are biased by the selected 16S rRNA (gene) amplicon sequencing pipeline. *Front Microbiol* **11**, 550420 (2020).
75 Di Tommaso, P. *et al.* Nextflow enables reproducible computational workflows. *Nat Biotechnol* **35**, 316-319 (2017).
76 Kurtzer, G. M., Sochat, V., Bauer, M. W. Singularity: Scientific containers for mobility of compute. *PLoS One* **12**, e0177459 (2017).
77 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* **17**, 3 (2011).

78 Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**, 852-857 (2019).
79 Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-583 (2016).
80 Pruesse, E. *et al.* SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**, 7188-7196 (2007).
81 Bokulich, N. A. *et al.* Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 90 (2018).
82 Langille, M. G. I. *et al*. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**, 814-821 (2013).
83 Ye, Y., Doak, T.G. A Parsimony Approach to Biological Pathway Reconstruction/Inference for Genomes and Metagenomes. *PLoS Comput. Biol.* **5**, e1000465 (2009).
84 Dittmar, T., Koch, B., Hertkorn, N., Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol Oceanogr-Meth* **6**, 230-235 (2008).
85 Li, Y. *et al.* Proposed Guidelines for Solid Phase Extraction of Suwannee River Dissolved Organic Matter. *Anal Chem* **88**, 6680-6688 (2016).
86 Blakney, G. T., Hendrickson, C. L., Marshall, A. G. Predator data station: A fast data acquisition system for advanced FT-ICR MS experiments. *Int J Mass Spectrom* **306**, 246-252 (2011).
87 Emmett, M. R., White, F. M., Hendrickson, C. L., Shi, S. D.-H., Marshall, A. G. Application of micro-electrospray liquid chromatography techniques to FT-ICR MS to enable high-sensitivity biological analysis. *J Am Soc Mass Spectrom* **9**, 333-340 (1998).
88 Kaiser, N. K., Savory, J. J., Hendrickson, C. L. Controlled ion ejection from an external trap for extended *m/z* range in FT-ICR mass spectrometry. *J Am Soc Mass Spectrom* **25**, 943-949 (2014).
89 Kaiser, N. K., McKenna, A. M., Savory, J. J., Hendrickson, C. L., Marshall, A. G. Tailored ion radius distribution for increased dynamic range in FT-ICR mass analysis of complex mixtures. *Anal Chem* **85**, 265-272 (2013).
90 Chen, T., Beu, S. C., Kaiser, N. K., Hendrickson, C. L. Note: Optimized circuit for excitation and detection with one pair of electrodes for improved Fourier transform ion cyclotron resonance mass spectrometry. *Rev Sci Instrum* **85**, 066107 (2014).
91 Boldin, I. A. & Nikolaev, E. N. Fourier transform ion cyclotron resonance cell with dynamic harmonization of the electric field in the whole volume by shaping of the excitation and detection electrode assembly. *Rapid Commun Mass Spectrom* **25**, 122-126 (2011).
92 Kaiser, N. K., Quinn, J. P., Blakney, G. T., Hendrickson, C. L., Marshall, A. G. A Novel 9.4 Tesla FT ICR mass spectrometer with improved sensitivity, mass resolution, and mass range. *J Am Soc Mass Spectrom* **22**, 1343-1351 (2011).
93 Blakney, G. T., Hendrickson, C. L., Marshall, A. G. Predator data station: A fast data acquisition system for advanced FT-ICR MS experiments. *Int J Mass Spectrom* **306**, 246-252 (2011).

94 Xian, F., Hendrickson, C. L., Blakney, G. T., Beu, S. C. & Marshall, A. G. Automated Broadband Phase Correction of Fourier Transform Ion Cyclotron Resonance Mass Spectra. *Anal Chem* **82**, 8807-8812 (2010).
95 Savory, J. J. *et al.* Parts-Per-Billion Fourier transform ion cyclotron resonance mass measurement accuracy with a "Walking" calibration equation. *Anal Chem* **83**, 1732-1736 (2011).
96 Kendrick, E. A mass scale based on CH_2 = 14.0000 for high resolution mass spectrometry of organic compounds. *Anal Chem* **35**, 2146-2154 (1963).
97 Hughey, C. A., Hendrickson, C. L., Rodgers, R. P., Marshall, A. G., Qian, K. N. Kendrick mass defect spectrum: A compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal Chem* **73**, 4676-4681 (2001).
98 Corilo, Y. E. The Florida State University: Tallahassee, Fl, USA. *PetroOrg* (2012).
99 LaRowe, D. E. & Van Cappellen, P. Degradation of natural organic matter: A thermodynamic analysis. *Geochim Cosmochim Ac* **75**, 2030-2042 (2011).
100 Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., Egozcue, J. J. Microbiome Datasets are compositional: and this is not optional *Front Microbiol* **8**, 2224 (2017).

CHAPTER 5 - AGRICULTURAL SULFUR APPLICATIONS ALTER THE QUANTITY AND COMPOSITION OF DISSOLVED ORGANIC MATTER FROM FIELD-TO-WATERSHED SCALES[4]

1. INTRODUCTION

Agricultural sulfur (S) applications are increasing worldwide in response to declines in atmospheric S deposition[1,2]. This change in human-driven alteration to the global S cycle mobilizes geologic, relatively non-reactive S into the biosphere. Compelling recent syntheses highlight the potential for excess S to cause significant ecosystem and human health consequences[1,3], which gives new urgency to efforts aimed at determining the forms, amounts, and fates of agriculture-derived S species in the environment. As organic S is the most dominant form of S in both natural and agricultural soils[4–7], investigating agricultural changes to the organic S pool is a critical step towards fully constraining the transport, transformations, and fates of agricultural S in the environment.

Organic S is intricately tied to elements of environmental concern, including mercury. Microbial sulfate reduction (MSR) stimulates methylmercury production, a neurotoxin that bio-accumulates in fish and wildlife[8]. Much of what we understand about the complex interactions between S and Hg stems from decades of research that linked agricultural S runoff to methylmercury production in Florida Everglades wetlands[9,10]. Researchers found that the speciation, atomic S content, and aromaticity of dissolved organic S (DOS) can act either to inhibit

[4] Hermes, A. L., Logan, M. N., Poulin, B. A., McKenna, A. M., Dawson, T. E., Borch, T., & Hinckley, E.-L. S. (2023). Agricultural Sulfur Applications Alter the Quantity and Composition of Dissolved Organic Matter from Field-to-Watershed Scales. *Environmental Science & Technology*. https://doi.org/10.1021/acs.est.3c01347

or stimulate methylmercury production. For example, sulfurization of dissolved organic matter (DOM) in Florida Everglades sediments has been linked to enhanced DOM reactivity[11], as more aromatic DOS or DOS enriched in reduced S functionalities (e.g., thiolates) enhance methylmercury production[12–14]. These studies suggest that characterizing the molecular composition and speciation of DOS derived from agricultural areas is important for understanding agricultural S reactivity in the environment. However, there is a need to broaden investigations of DOS chemistry beyond wetland ecosystems, including upland agricultural areas or regional settings with mixed land use/land cover (LULC).

Within agricultural fields, DOS speciation and molecular composition may affect S storage within field soils versus mobilization to downgradient aquatic ecosystems, where it can interact with Hg and other biogeochemical cycles. Studies in forested ecosystems that received high atmospheric S deposition demonstrated that adsorption of DOS to mineral soils lengthened the residence time of S within soils[15,16]. In terms of mobilization, agricultural areas export primarily low-molecular weight microbially-derived DOM with a higher proportion of S and nitrogen (N) heteroatoms[17–19]. Once in downstream aquatic ecosystems, DOS is subject to additional transformations. Recent studies with DOM and model compounds show that reduced DOS species are selectively photodegraded to inorganic sulfate (SO_4^{2-})[20,21], a process that could enhance agricultural S transport. Finally, complex S cycling and hydrology in both agricultural and non-agricultural areas of a watershed may influence the degree to which agricultural S is mobilized and interacts with other elements downstream. Linking agricultural S applications to the composition of the organic S pool and the fate of agricultural S in broader ecosystems requires (1) analytical approaches that holistically characterize DOS chemistry and (2) an assessment of changes to DOS across spatial scales.

In this study, we employed multiple qualitative and quantitative analytical techniques to advance characterization of agricultural DOS in an upland, mixed LULC watershed. Given higher S inputs to agricultural areas compared to non-agricultural areas, we hypothesized that: (1) the quantity of DOS is higher in agricultural areas compared to non-agricultural areas; (2) agricultural DOS is compositionally distinct from non-agricultural DOS; and (3) agricultural DOS is detectable in surface waters beyond fields. We used negative-ion electrospray ionization ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) at 21 tesla to qualitatively evaluate agricultural and non-agricultural molecular-level compound diversity[22] and to trace agricultural DOS beyond agricultural fields[20]. Sulfur K-edge X-ray absorption near-edge structure (XANES) spectroscopy quantified organic S speciation, or the distribution of S oxidation states[23]. Finally, we measured the ratio of S stable isotopes in organic matter (δ^{34}S-DOS), which has been used to differentiate DOS derived from assimilatory versus dissimilatory SO_4^{2-} reduction[24] and provides insight into S mineralization pathways[25].

We build on prior research of agricultural S in the Napa River Watershed[7,26–29], where a $34 billion wine grape crop industry[30] requires intensive S applications averaging 80 kg S ha^{-1} yr^{-1} as S^{0}[1]. Vineyards are surrounded by forest and shrubland/grassland areas that receive atmospheric and geologic S sources, providing a natural contrast between areas of intensive S application and background S chemistry. To test our hypotheses, we analyzed soil porewater samples collected from vineyard and forest/shrubland/grassland (i.e., "non-agricultural") LULC areas (agricultural field scale, ~0.15-2.5 ha) and sampled surface waters from mixed LULC tributaries and the Napa River (catchment-to-watershed-scales, ~260-47,000 ha). Using our novel combination of advanced DOM characterization techniques, we "fingerprinted" agricultural DOS,

providing a means to holistically investigate the link between agricultural S applications and changes to S chemistry.

2. Materials and Methods

2.1 Site description and sample collection

The Napa River flows into San Pablo Bay, connecting to extensive wetland ecosystems and the greater San Francisco Bay Estuary. The Mediterranean climate results in distinct wet and dry seasons. During the dry, growing season (~May through September), vineyard fields are sprayed weekly to biweekly with S^0 to prevent powdery mildew disease. During the wet season (~October through April), average precipitation ranges from 863 mm in St. Helena to 677 mm in Napa (1991-2020 annual precipitation normal)[31] and intermittent tributaries begin to flow.

Sampling campaigns were carried out during the 2019-2020 California wet season within the Napa River Watershed (Figure 1; Table S1 and Figure S1 in the Supplementary Information (SI)). We collected soil porewater and surface waters between 4-14 December 2019, which was during the first significant rainfall event of the season with 235 mm of precipitation (26 November-13 December 2019) and a maximum of 12.7 m^3s^{-1} streamflow in the Napa River (SI Table S1 and Figure S1). We sampled surface waters again during low flow conditions 29 February-1 March 2020 with 0.7 m^3s^{-1} streamflow in the Napa River (SI Table S1 and Figure S1). At the time of sampling, the 2020 water year was the driest year in a 20-year precipitation record[32], totaling only 424 mm of precipitation in Napa[33].

2.2 SAMPLE PROCESSING AND AQUEOUS CHEMICAL ANALYSES

A complete description of sample processing and analyses can be found in the SI. Briefly, we analyzed soil porewater and stream water samples for dissolved organic carbon (DOC), SO42- and total dissolved sulfur concentrations ([DOC], [SO42-] and [TDS], respectively), DOM specific ultraviolet absorbance at 254 nm, SUVA254, a proxy for DOM aromaticity34, and δ34S-SO42-. Soil porewater samples were processed in an O2-free glove box to retain sample redox conditions prior to analyses.

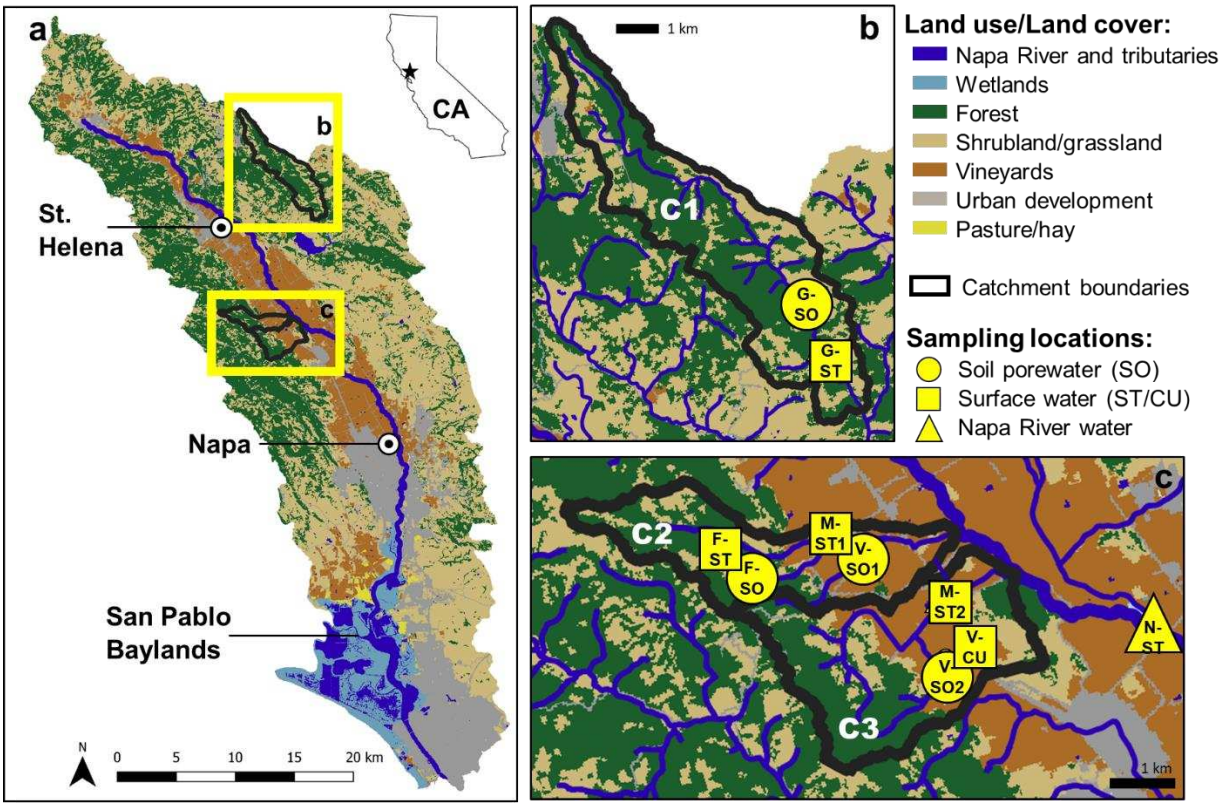


Figure 1. (a) Map and land use/land cover of the Napa River Watershed, which drains into extensive wetlands in the San Pablo Baylands. Sample collection efforts were focused within three sub-catchments of the Napa River: (b) a forest/shrubland/grassland catchment (C1), and (c) two catchments with forested headwaters that then drain through vineyard agriculture (C2 and C3). G = grassland; F = forest; V = vineyard; M = mixed land use/land cover; N = Napa River; PW = soil porewater; RO = surface water runoff.

We used solid-phase extraction (SPE) for parallel DOM characterization using FT-ICR MS, S XANES spectroscopy, DOM carbon and S content ([C]-DOM and [S]-DOM, respectively), and δ^{34}S-DOM analysis (see SI). Salt-free DOM extracts were isolated from soil porewater and stream water samples using modified styrene divinyl benzene polymer resin (Bond Elut-PPL, Varian Inc., USA; modified after Dittmar et al.[35]). We chose to use the PPL sorbent because it has been widely used for FT-ICR MS analysis and performs generally well compared to other sorbents for S extraction[36–38]. We saved aliquots of PPL eluate for [DOC], UV-vis absorption, [SO_4^{2-}], and [TDS] analyses to check for sample breakthrough and calculate C recovery (see SI). We report stable isotope values in conventional δ-notation in parts per thousand (‰), relative to the international standard Vienna-Canyon Diablo Troilite (VCDT). Analytical precision for δ^{34}S-SO_4^{2-} was ± 0.4‰, ± 0.2‰ for δ^{34}S-DOM, and ± 0.35% for [S]-DOM (Table S3).

2.3 FT-ICR MS ANALYSIS

To identify compositional differences between vineyard agriculture and non-agricultural samples, we analyzed a sub-sample of the dried PPL-extracted DOM using FT-ICR MS (see SI). Importantly, S species are lower in abundance and ionization efficiency than O_x compounds and thus identification benefited from improved sensitivity, high dynamic range, and ultrahigh mass resolving power only achievable with 21 tesla FT-ICR MS. Briefly, samples were re-suspended in 100% HPLC-grade methanol (target 200 ppm C)[39] and analyzed in negative ion electrospray ionization mode (ESI-) with a custom-built hybrid linear ion trap FT-ICR MS equipped with a 21 T superconducting solenoid magnet[40,41] at the National High Magnetic Field Laboratory in Tallahassee, Florida. Time-domain transients of 3.1 seconds were conditionally co-added and acquired with the Predator data station that handled excitation and detection only, initiated by a TTL trigger from the commercial Thermo data station, with 100 time-domain acquisitions

averaged for all experiments[42]. Mass spectra were phase-corrected following Xian et al.[43] and internally calibrated with 10-15 highly abundant homologous series that span the entire molecular weight distribution (~150-200 individual calibrants) based on the "walking" calibration method[44]. Molecular formulae were assigned and post-processed using PetroOrg software©[45]. One sample, V-PW1, had poor spectral quality (Figure S11), and was therefore excluded from further FT-ICR MS analyses. All 21 tesla FT-ICR MS mass spectra files are publicly-available via the Open Science Framework (https://osf.io/nh3zc/) at DOI 10.17605/OSF.IO/NH3ZC.

Assigned formulas were summarized into elemental class distributions (normalized to relative intensity), the modified aromaticity index (AI_{mod}), nominal oxidation state of C (NOSC), and compound classes (see SI Table S4). We calculated the percentage of S-containing heteroatom formulas and the formulaic S:C after Poulin et al.[46]:

$$\%\, S\ containing\ formulas = \frac{\sum CHOS, CHNOS}{Total\ formula\ count} * 100\%$$

$$Formulaic\ S{:}C = \frac{\sum Sat}{\sum Cat}$$

where *CHOS* and *CHNOS* are the number of formulas containing S and N plus S; S_{at} and C_{at} are the total number of S and C atoms, respectively, detected in a sample. Finally, we derived a "forest and grassland LULC endmember" by combining all S-containing heteroatom formulas from forest and grassland samples (F-PW, HF F-RO, LF F-RO, G-PW, and G-RO). We then compared the formulas in this general endmember to vineyard samples to determine unique vs. shared S-containing heteroatoms across land use types.

2.4 Sulfur K-edge XANES spectroscopy

We analyzed the PPL-extracted DOM by S K-edge XANES spectroscopy (beamline 4-3, Stanford Synchrotron Radiation Lightsource; complete details in the SI). The atomic fractions of S functionalities were determined by Gaussian curve fitting (GCF)[23], in the software program *Athena*[47]. Spectra were normalized using pre- (2450.0 – 2462.5 eV) and post-edge regions (2515.0 – 2527.5 eV). Precision and accuracy estimates, curve fitting parameters, GCF results, and the atomic fractions of S functionalities are detailed in the SI (Figures S18-S20, Tables S8-S10). We report concentrations of S functionalities (reduced S, sulfoxide, sulfone, sulfonate, and organosulfate) relative to C, calculated by multiplying the fraction of each functionality by the measured atomic S/C.

2.5 Statistical analyses

We analyzed all data using *R* statistical software (v. 4.0.4)[48]. Concentrations and related chemical properties (e.g. atomic S/C-DOM) are reported as median ± interquartile range. To test for differences in molecular composition across samples, we calculated Bray-Curtis dissimilarity matrices of normalized peak intensities[49,50] for all assigned DOM molecular formulae and CHOS + CHNOS formulae using the *vegan* package and "vegdist" function[49–51]. Dissimilarity values were hierarchically clustered into a dendrogram using the "hclust" function with a ward.D2 linkage method[52]. We combined multiple data analyses into a principal component analysis (PCA) to assess overall differences across samples (Figure S21).

3. Results

3.1 Sulfur and carbon concentrations

Comparing across land use/land cover (LULC) types, vineyard sample S concentrations were distinct from non-agricultural samples, but they did not show a trend in [DOC] or DOM

$SUVA_{254}$. Vineyard samples had approximately four-fold higher SO_4^{2-} concentrations relative to non-agricultural samples (14.5 ± 16.7 mg SO_4^{2-}-S L^{-1} vs. 3.22 ± 3.81 mg SO_4^{2-}-S L^{-1}, respectively; Table S2). The highest SO_4^{2-} concentration measured was in catchment 2: vineyard soil porewater (V-PW2), vineyard culvert runoff (V-RO), and mixed LULC stream runoff (M-RO2) had 14.5, 38.5, and 26.7 mg SO_4^{2-}-S L^{-1}, respectively. Like SO_4^{2-}, the influence of S inputs to vineyard areas was apparent in the atomic S/C of PPL-extracted DOM. Vineyard samples had approximately two-fold higher atomic S/C-DOM than non-agricultural samples, with median values of 4.83×10^{-3} and 2.73×10^{-3} mmol S mmol C^{-1}, respectively (Table S3). In contrast to S chemistry, DOC concentration did not differ by LULC type (Table S2) and neither did the DOM $SUVA_{254}$, which ranged from 1.36 L mg C^{-1} m^{-1} in grassland soil porewater (G-PW) to 3.74 L mg C^{-1} m^{-1} in grassland stream water runoff (G-RO; Table S2).

3.2 FT-ICR MS CHARACTERIZATION OF DOS

FT-ICR MS results showed notable differences in S chemistry across LULC types. We observed a significant correlation between atomic S/C-DOM and the percent of S heteroatoms detected in FT-ICR MS spectra (Figure 2a) as well as the formulaic S/C (Figure 2b). The S heteroatom content was also strongly correlated with dissolved SO_4^{2-} concentration ($R^2 = 0.73$; Figure S2). Vineyard samples (V-PW2 and V-RO) had 9% and 13% relative abundance of S-containing heteroatoms, whereas non-agricultural samples had only 3-7% relative abundance of S-containing heteroatoms (Figure 2c). The increase in vineyard S-containing heteroatoms stemmed from an overall increase in the relative abundance of all S-containing heteroatom classes ($CHOS_1$, $CHON_1S_1$, and $CHON_2S_1$; Figure 2c). Mixed LULC runoff M-RO1 and high flow Napa River runoff (HF N-RO) had S heteroatom distributions more similar to non-agricultural samples,

while mixed LULC runoff M-RO2 and low flow Napa River runoff (LF N-RO) had higher proportions of $CHOS_1$ and $CHON_1S_1$ classes, similar to vineyard samples (Figure 2c).

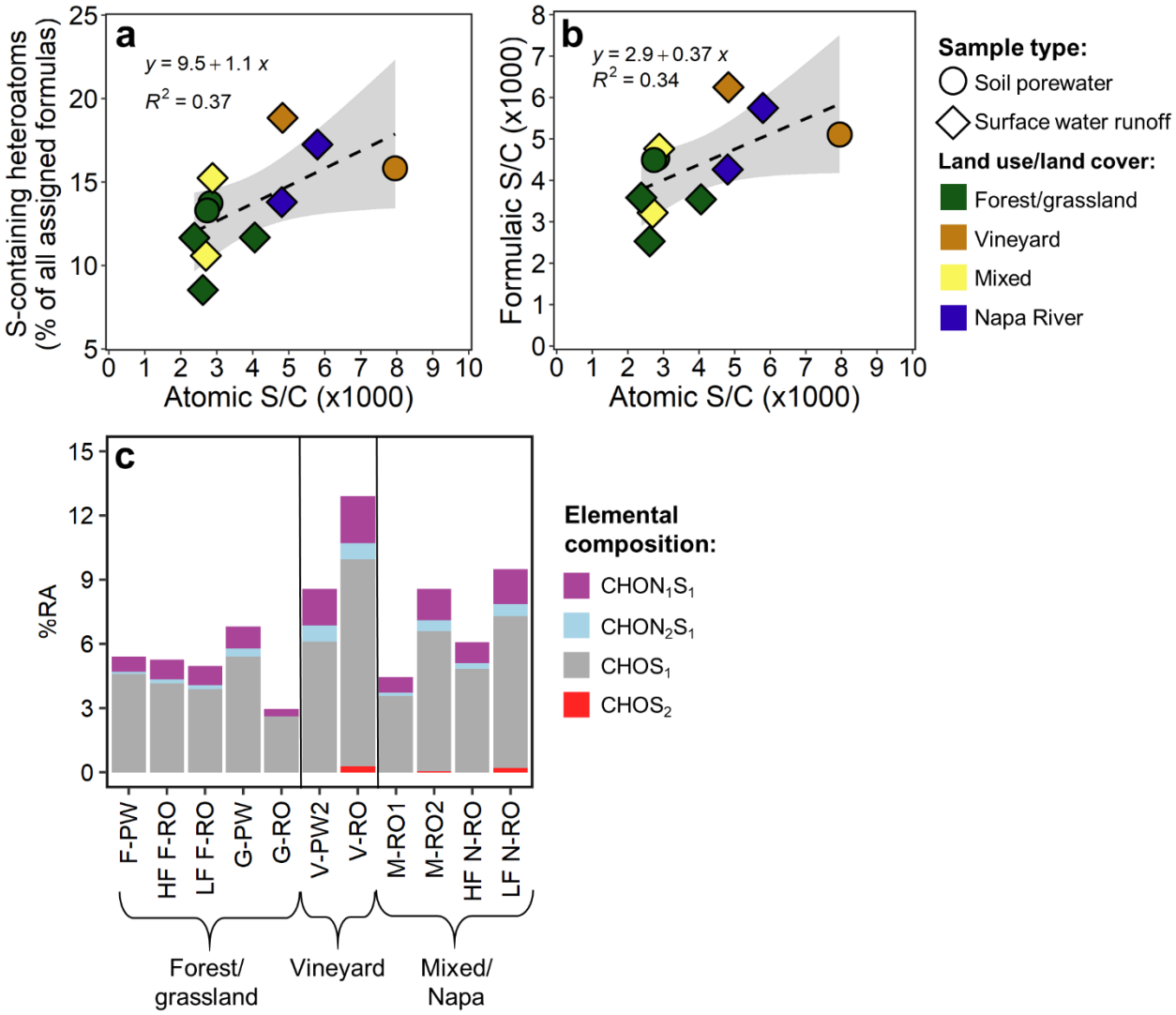


Figure 2. DOM sulfur content and molecular composition for different sample types and land use/land cover. Linear correlations between the atomic S/C content of DOM and (a) S-containing heteroatoms as a percentage of total assigned formulas and (b) FT-ICR MS-derived formulaic S/C ratios show that vineyards have higher DOS content and S-containing formulas compared to forest and grassland samples. In (c), FT-ICR MS-derived S-containing formulas for each sample are normalized to sample relative abundance (RA). Sample abbreviations are as in Figure 1. HF = high streamflow; LF = low streamflow.

We detected low levels of $CHOS_2$ in three samples: vineyard culvert runoff, V-RO (0.27% total relative abundance); mixed LULC stream runoff 2, M-RO2 (0.03% total relative abundance); and the Napa River at low flow, LF N-RO (0.19% total relative abundance; Figure 2c).

Given the higher S-content and proportion of S-containing formula assignments in vineyard samples, we evaluated whether vineyard DOS was compositionally distinct from non-agricultural samples. The Bray-Curtis dissimilarity analysis clustered samples into two groups, which generally split into a "forest/grassland-dominated" group and a "vineyard-dominated" group (Figure 3a). The first group included the majority of non-agricultural samples (G-RO, LF F-RO, HF F-RO, F-PW), as well as mixed stream runoff 1 (M-RO1), and the Napa River at high flow (HF N-RO). Samples within this group had 9-26% dissimilarity from one another. The second group included vineyard (V-PW2 and V-RO) as well as mixed stream runoff 2 (M-RO2), grassland (G-PW), and Napa River at low flow (LF N-RO) samples, and exhibited dissimilarity values slightly higher than the first group from 15-33%. Across groups, the highest dissimilarity was between V-RO and G-RO at 55%, and the lowest dissimilarities were between the forest stream and Napa River at high flow (HF F-RO and HF N-RO) at 9% and between the forest stream at high flow and mixed stream runoff 1 (HF F-RO and M-RO1) at 10%. Soil porewater and culvert runoff samples from the vineyard (V-PW2 and V-RO, respectively) had relatively high dissimilarity at 31% despite coming from the same vineyard.

We further evaluated DOS composition across samples using van Krevelen putative compound categories (defined in Table S4). If Bray-Curtis dissimilarities were driven by differences in the number of formulas only, all samples would be expected to have similar distributions of formulas in van Krevelen space. We observed that highly unsaturated aliphatic formulas dominated all samples at 69-78%, followed by either aromatic (1-12%) or unsaturated

aliphatic formulas (8-17%; Figure 3b). However, several samples exhibited unique compound category distributions. For example, V-PW2 had the highest relative proportions of condensed (1.8%) and aromatic (11.8%) formulas compared to all other samples, and a higher proportion of sugar-like formulas (0.7%) compared to most other samples. V-RO had minimal condensed formulas (< 0.1%), a lower relative proportion of aromatic formulas (8.2%), and no sugar-like formulas. G-PW had the highest proportion of peptide-like (3.7%) and very low aromatic formulas (1.1%). Finally, F-PW was similar to G-PW, except for an increase in sugar-like formulas (1%). Mixed stream and Napa River samples had similar DOS compound category distributions.

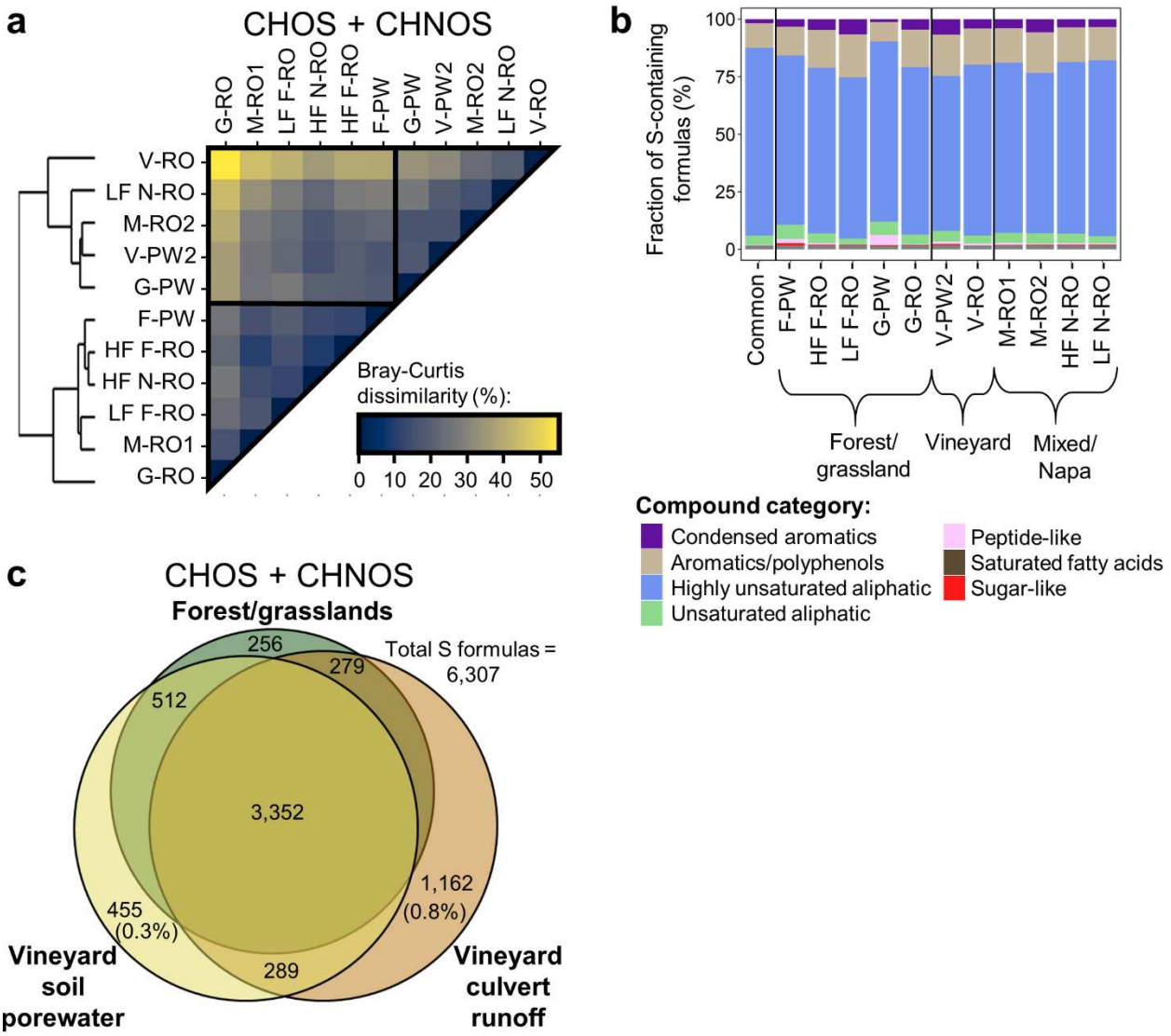


Figure 3. Compositional differences across samples. (a) Bray-Curtis dissimilarity matrix of intensity-weighted CHOS + CHNOS formulas showing that forest and grassland samples and vineyard samples generally cluster separately. Yellow represents higher dissimilarity and blue lower dissimilarity. (b) van Krevelen-derived compound categories shown as percentages of total S-containing formulas. The "common" bar is the composition of CHOS + CHNOS formulas that are shared across all samples. (c) Venn diagram showing the number of unique and common CHOS + CHNOS formulas between V-PW2 (vineyard soil porewater), V-RO (vineyard culvert runoff), and a general forest and grassland LULC endmember, which includes F-PW, HF F-RO, LF F-RO, G-PW, and G-RO. Percentage values are the abundance of unique formulas relative to all assigned formulas.

We examined unique S-containing formulas (CHOS + CHNOS) to isolate drivers of DOS compositional difference, comparing V-PW2, V-RO, and a general "forest/grassland" LULC

endmember (Figure 3c). Roughly 50% of the DOS formulas were common to endmember samples, with 18% unique to vineyard culvert runoff (V-RO), 7% unique to vineyard soil porewater (V-PW2), and only 4% unique to all forest/grassland samples. S-containing formulas unique to vineyard culvert runoff (V-RO) were primarily $CHOS_1$ (31%), $CHON_1S_1$ (24%), and $CHOS_2$ (30%), with the remainder $CHON_2S_1$ (14%; Figure S3). S-containing formulas unique to vineyard soil pore water (V-PW2) were split equally among $CHOS_1$ (35%), $CHON_1S_1$ (33%), and $CHON_2S_1$ (31%). Conversely, S-containing formulas unique to non-agricultural sites were primarily $CHOS_1$ (75%), with smaller contributions from $CHON_1S_1$ (21%) and $CHON_2S_1$ (4%).

The vast majority of mixed stream and Napa River S-containing formulas were shared with LULC endmembers or other mixed LULC samples. Remarkably, M-RO1 and HF N-RO samples shared 98.7% and 99.3% of S-containing formulas with another sample, the majority of which were shared with one or more LULC endmember (F-PW, HF F-RO, FL F-RO, G-PW, G-RO, V-PW2, V-RO; Figure S4). M-RO2 shared 95.9% of its S-containing formulas with another sample, and all of its 54 $CHOS_2$ formulas were shared with V-RO (Figure S5). The LF N-RO sample had the lowest proportion of shared formulas at 88.3% (Figure S4), but the majority of its $CHOS_2$ formulas were also present in the V-RO spectrum (Figure S5T).

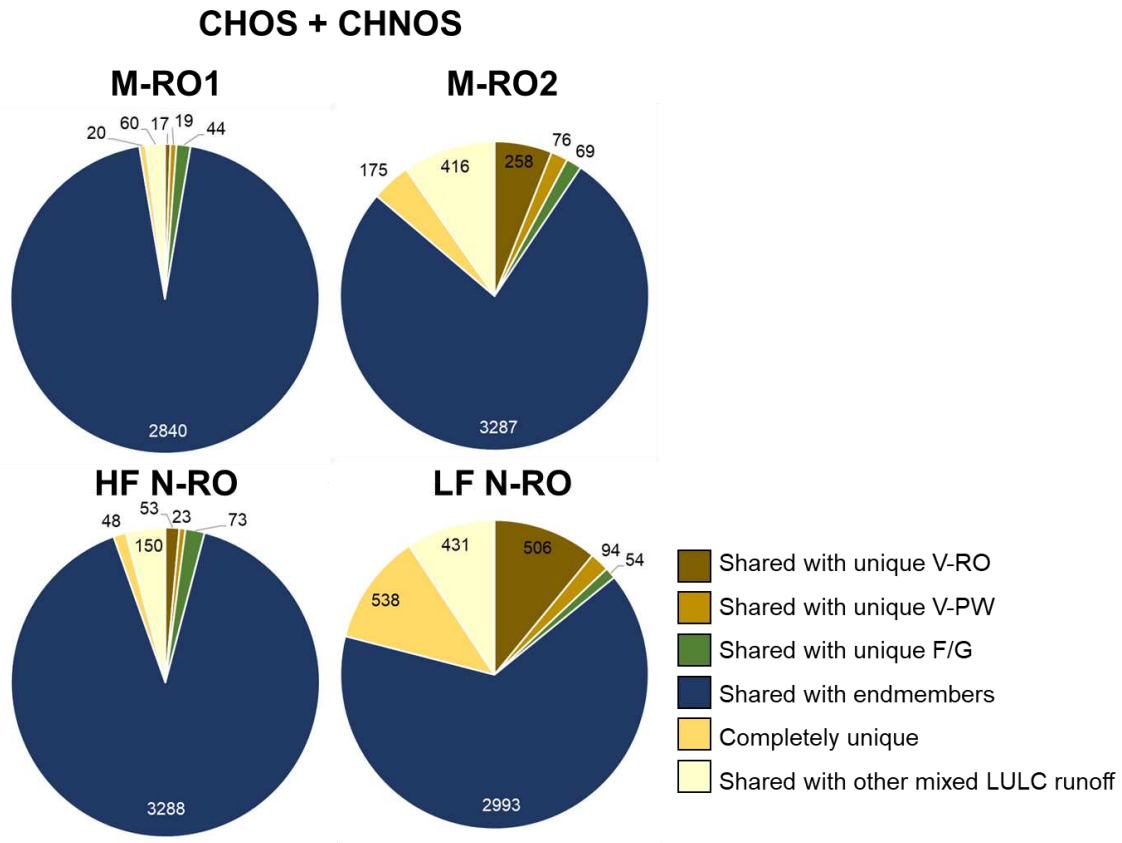


Figure S4. Mixed LULC stream and Napa River CHOS + CHNOS formulas shared with all endmembers (dark blue), shared with unique V-RO formulas (dark brown), shared with unique V-PW2 formulas (light brown), shared with unique forest/grassland formulas ("F/G"; green), shared with other mixed LULC streams or Napa samples (light yellow), or completely unique to the sample itself (gold).

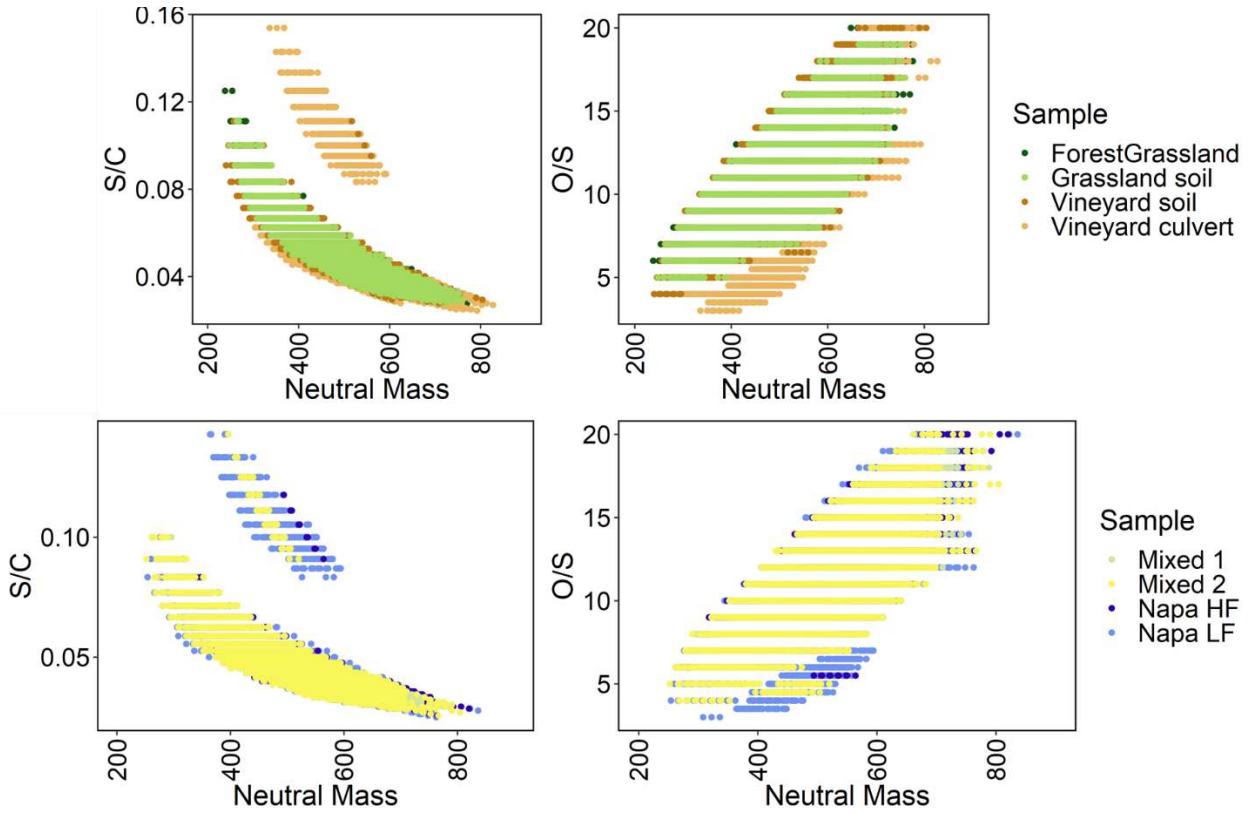


Figure S5. CHOS2 formulas as possible vineyard tracers. Vineyard culvert runoff had 352 unique CHOS2 formulas (top two panels), which re-appeared in Mixed Stream 2 (n = 54) and the Napa River at low flow conditions (n = 291; bottom panels). The CHOS2 formulas in Mixed Stream 2 all matched those in vineyard culvert runoff, and 246/291 CHOS2 formulas in the Napa River at low flow conditions matched vineyard culvert runoff. The CHOS2 formulas have an overall lower average neutral mass relative to other S-containing formulas and have a lower O/S, supported by a lower AImod index.

3.3 Comprehensive FT-ICR MS spectra and van Krevelen diagrams

To provide a complete visual context for the molecular analyses discussed throughout this chapter, Figure S6 presents the full FT-ICR MS spectra, spectra of all assigned formulas, and Van Krevelen diagrams (H/C vs. O/C) for each sample. These diagrams are further separated by heteroatom content (N = 0-3; S = 0-2). Overall, the spectra were of extremely high quality, with the exception of V-PW1, which was excluded from analyses.

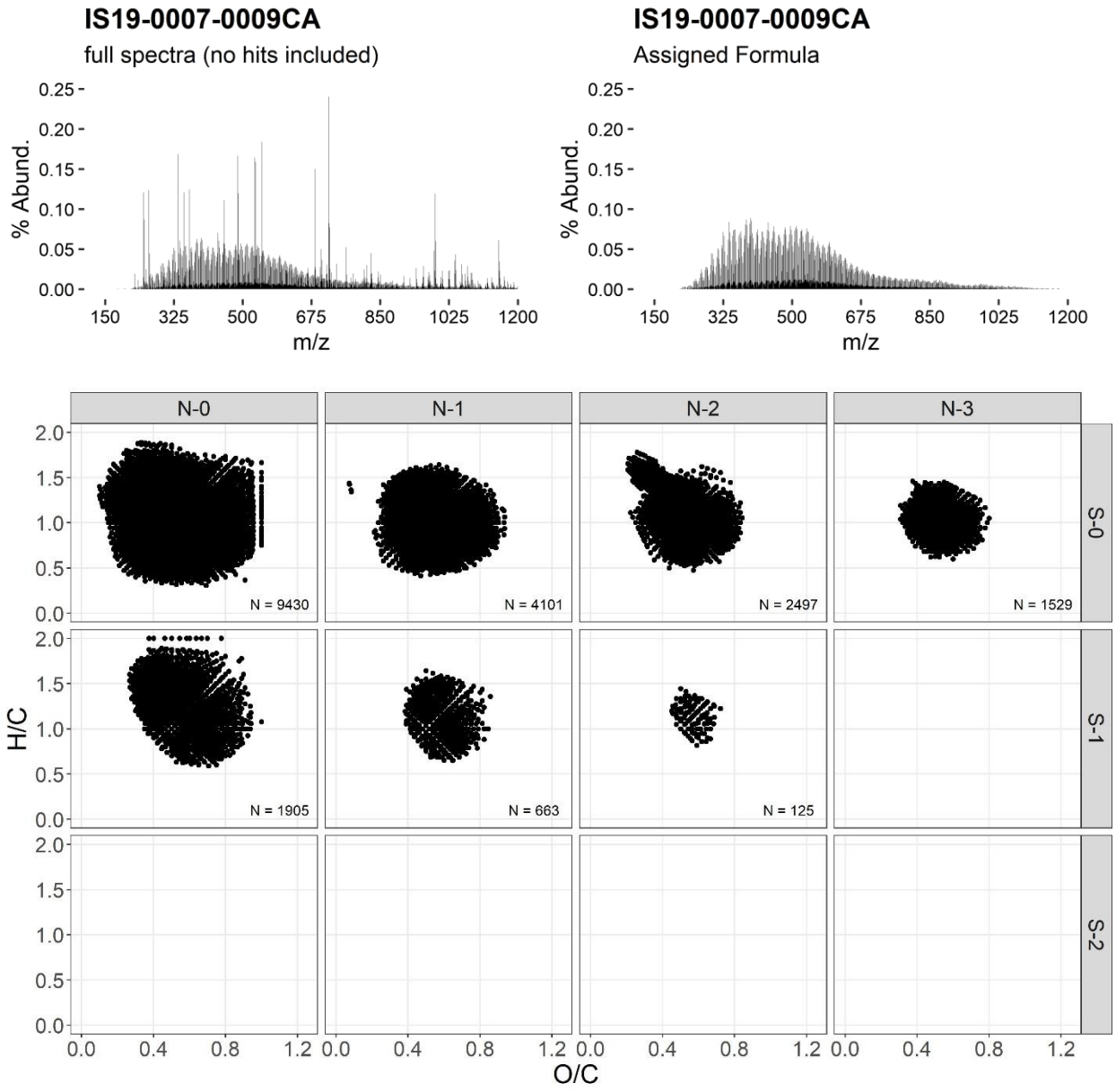


Figure S6. Forest soil porewater (F-PW)

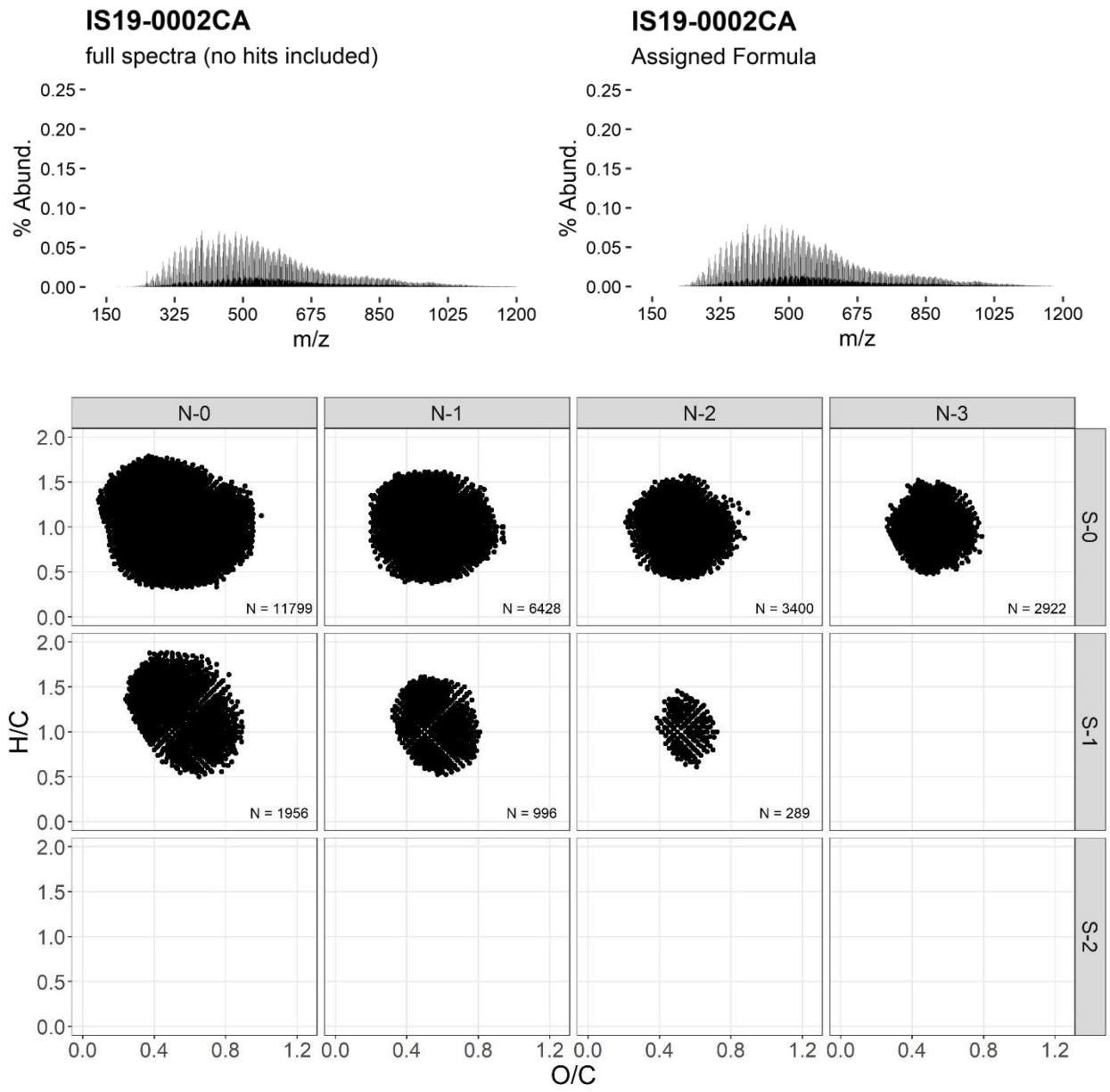


Figure S7. High flow forest stream water runoff (HF F-RO)

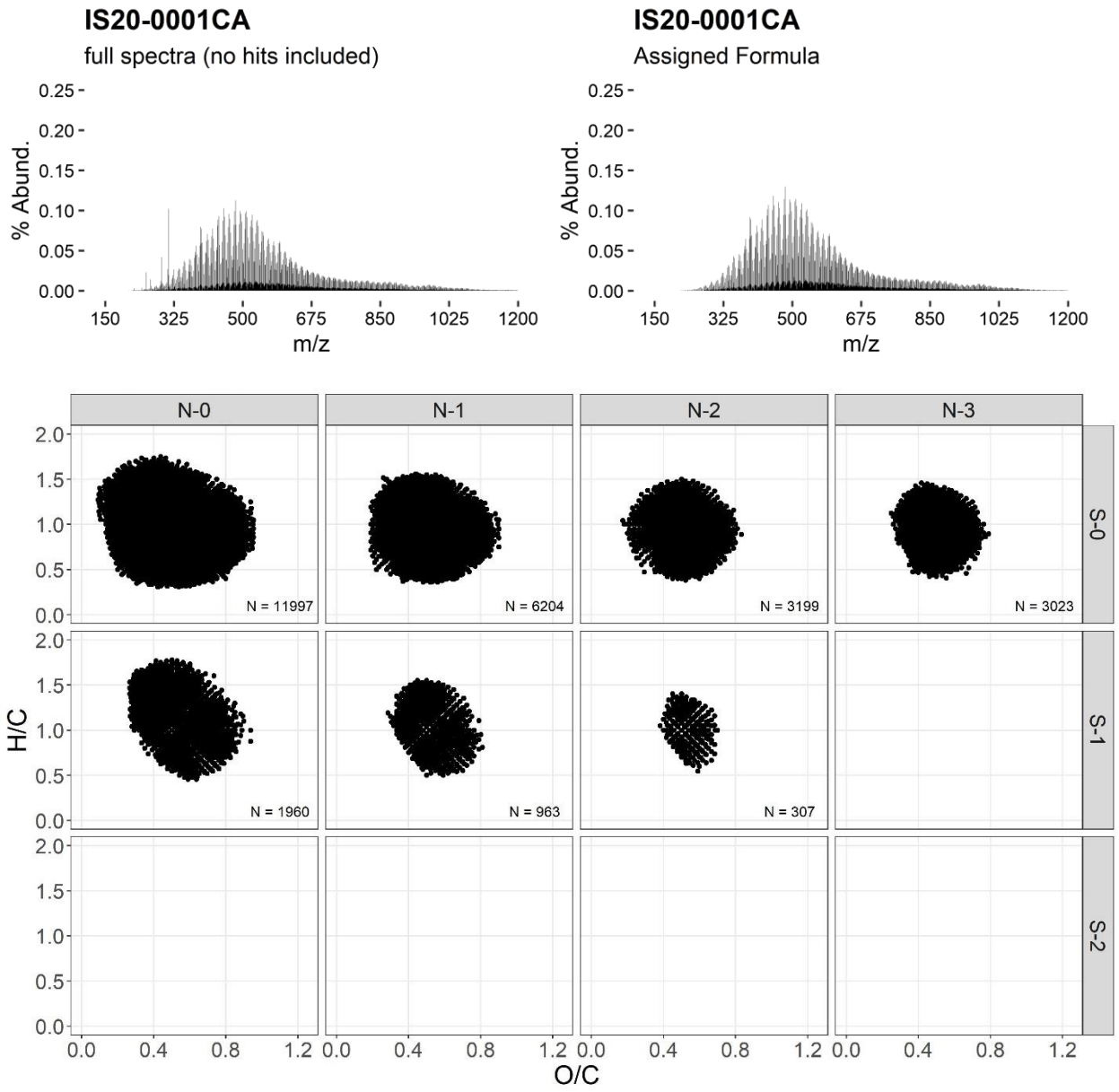


Figure S8. Low flow forest stream water runoff (LF F-RO)

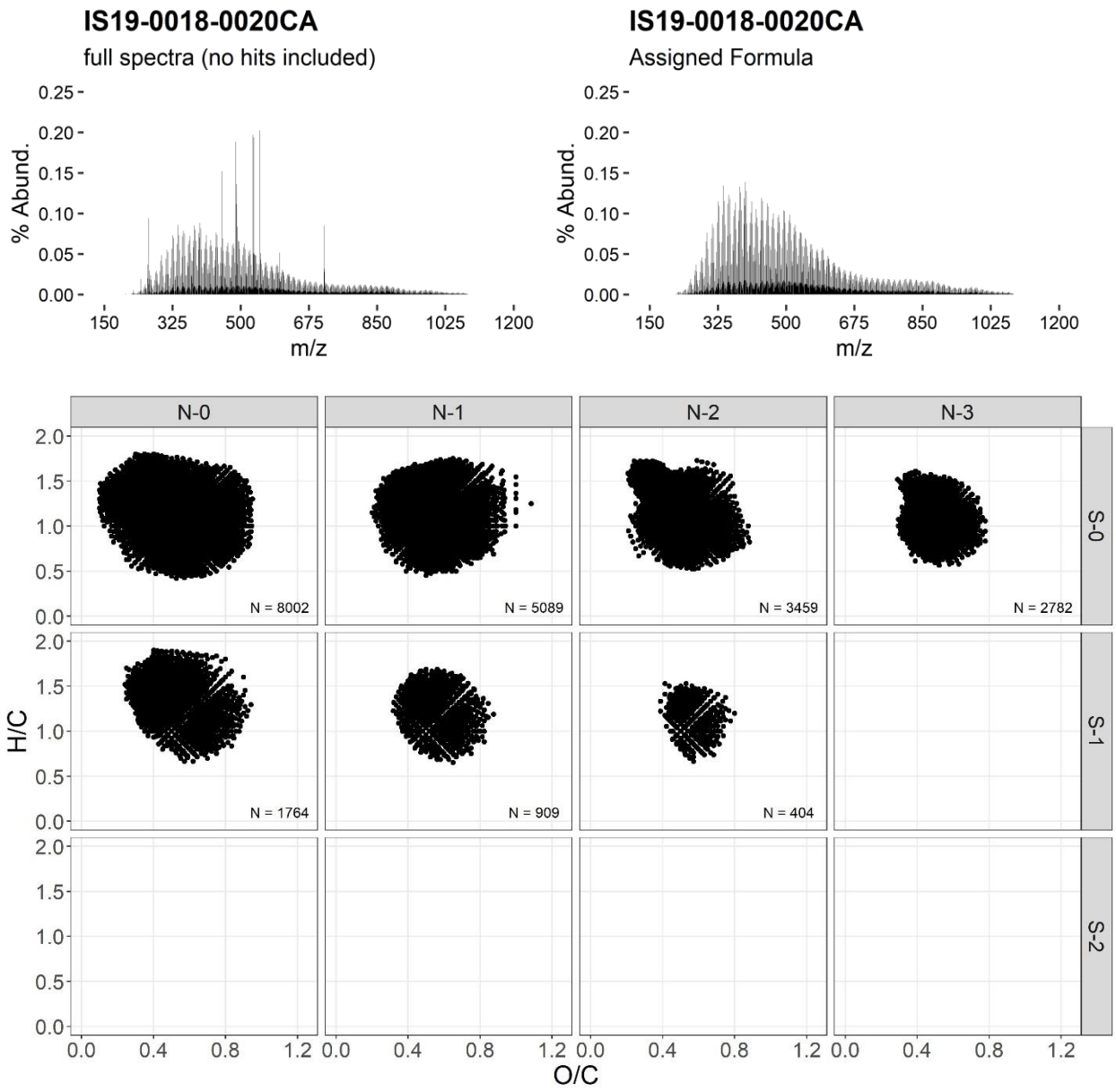


Figure S9. Shrubland/grassland soil porewater (G-PW)

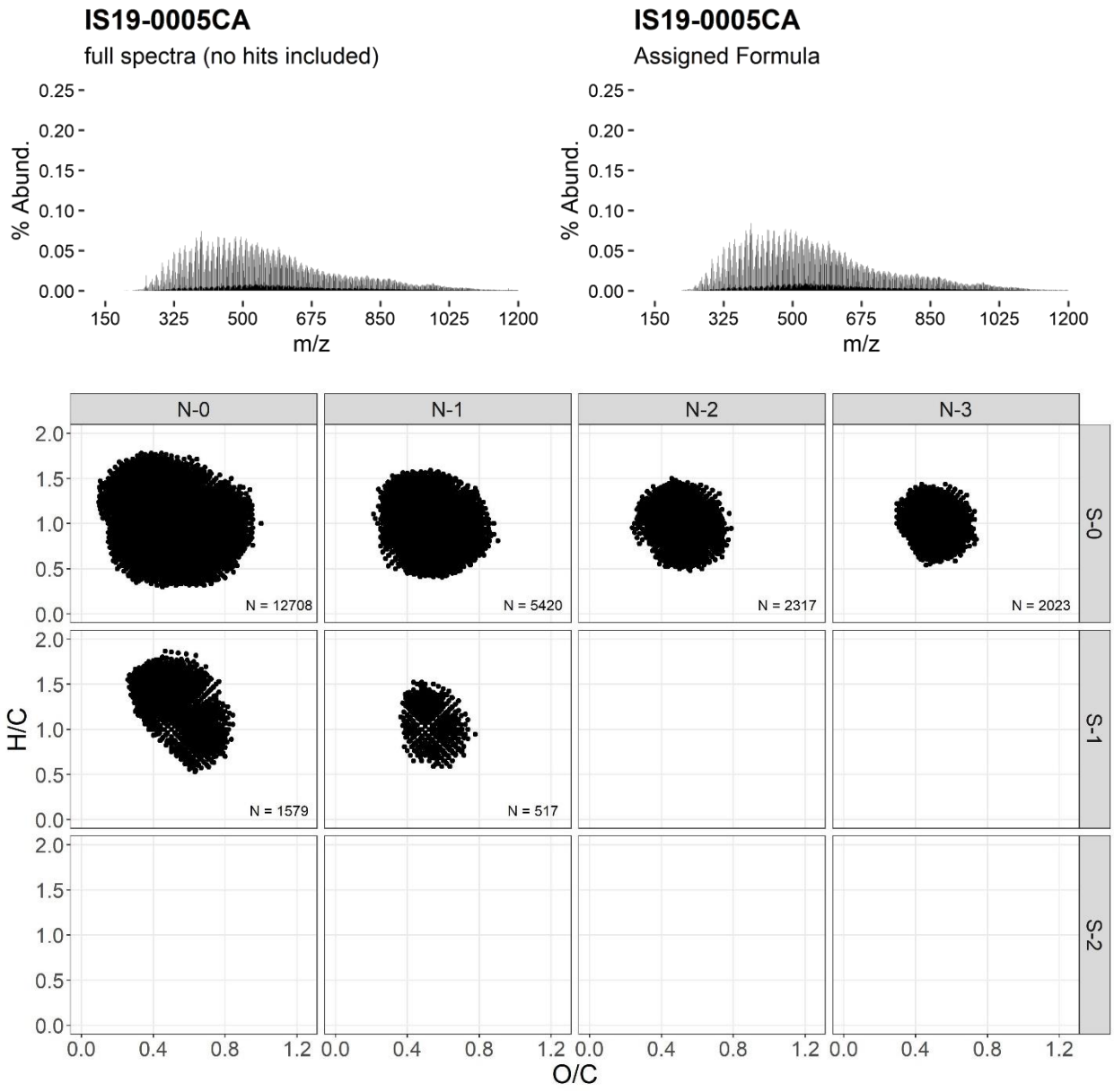


Figure S10. Shrubland/grassland stream water runoff (G-RO)

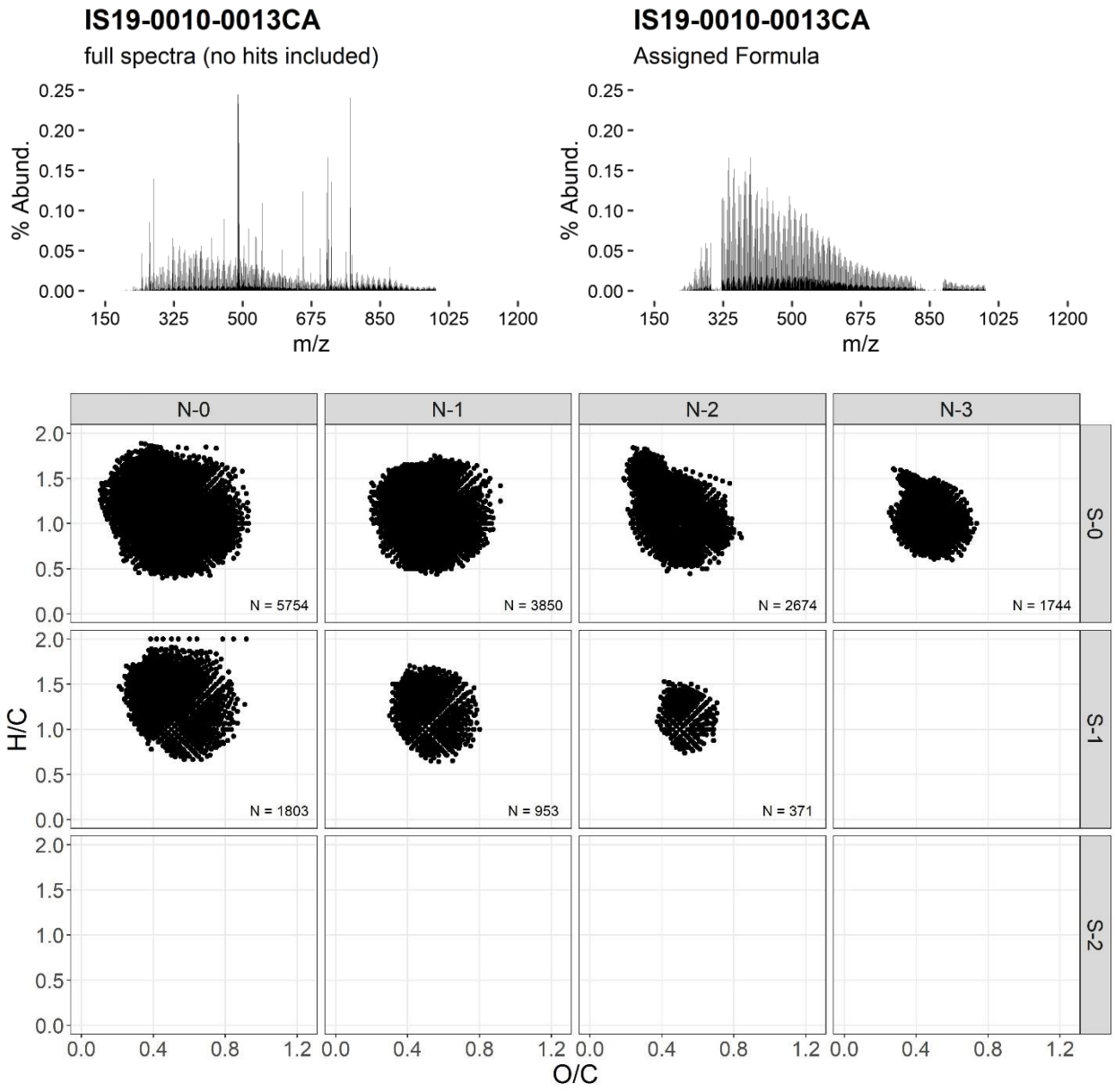


Figure S11. Vineyard soil porewater 1 (V-PW1)

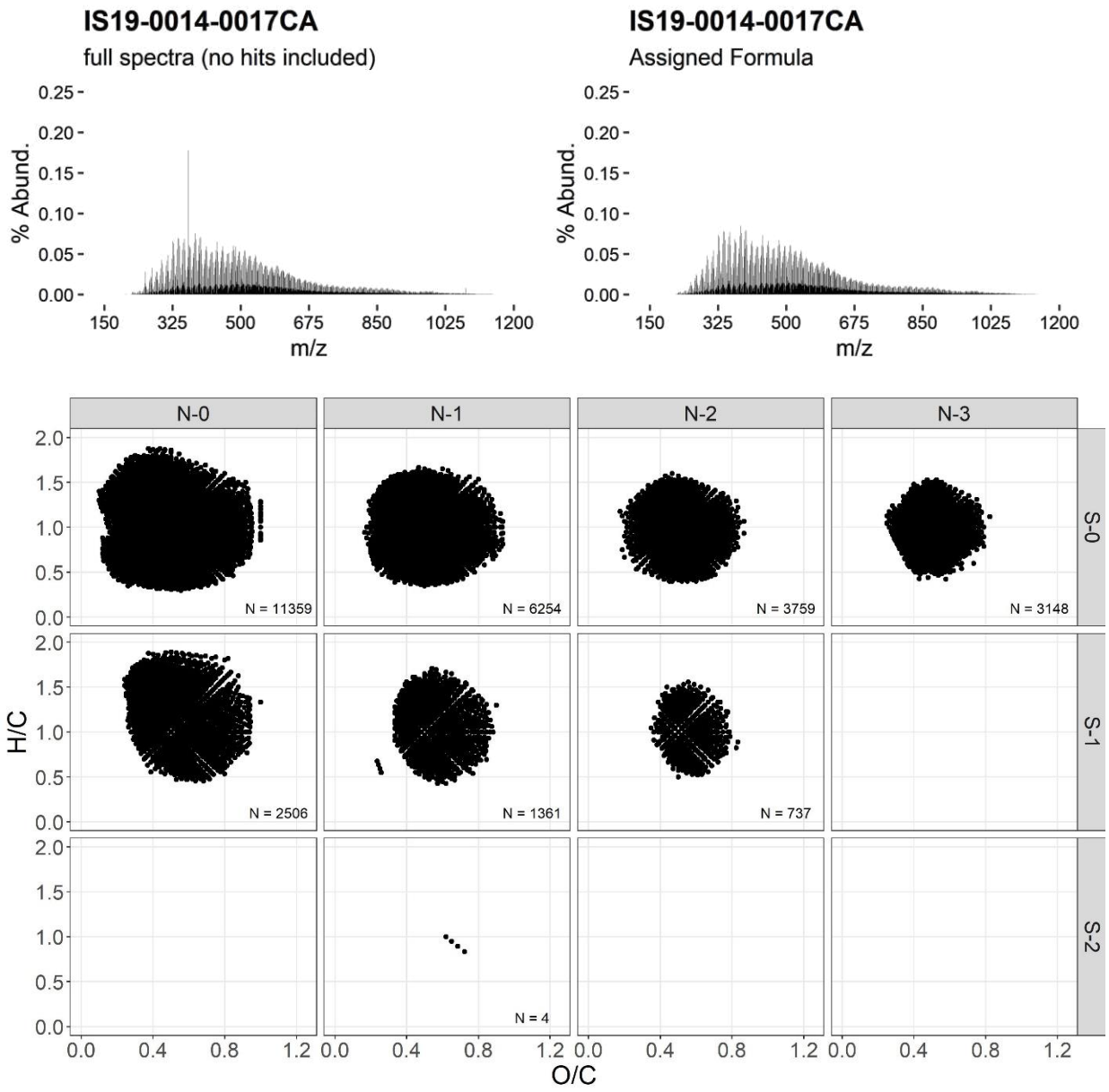


Figure S12. Vineyard soil porewater 2 (V-PW2)

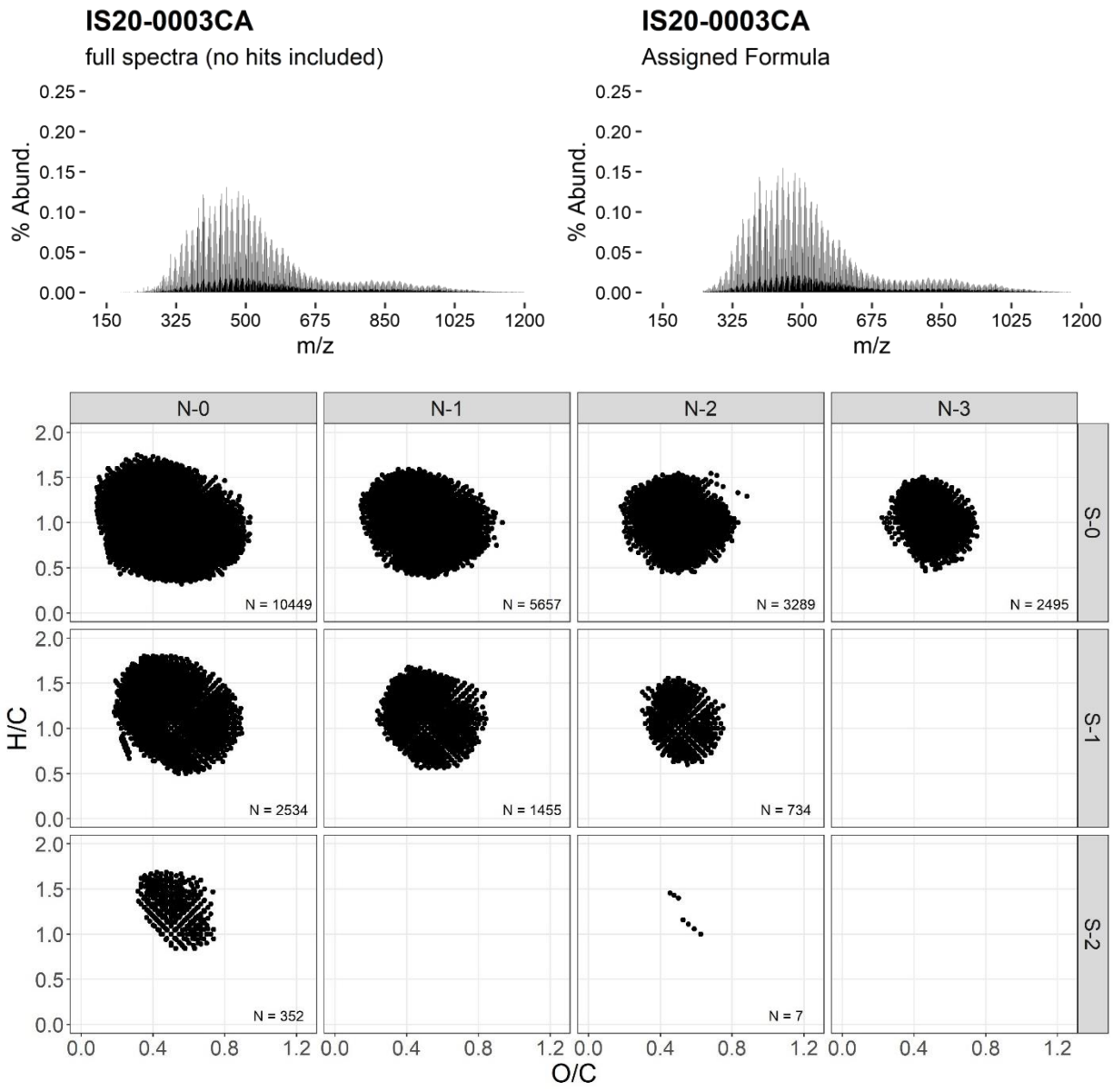


Figure S13. Vineyard culvert runoff (V-RO)

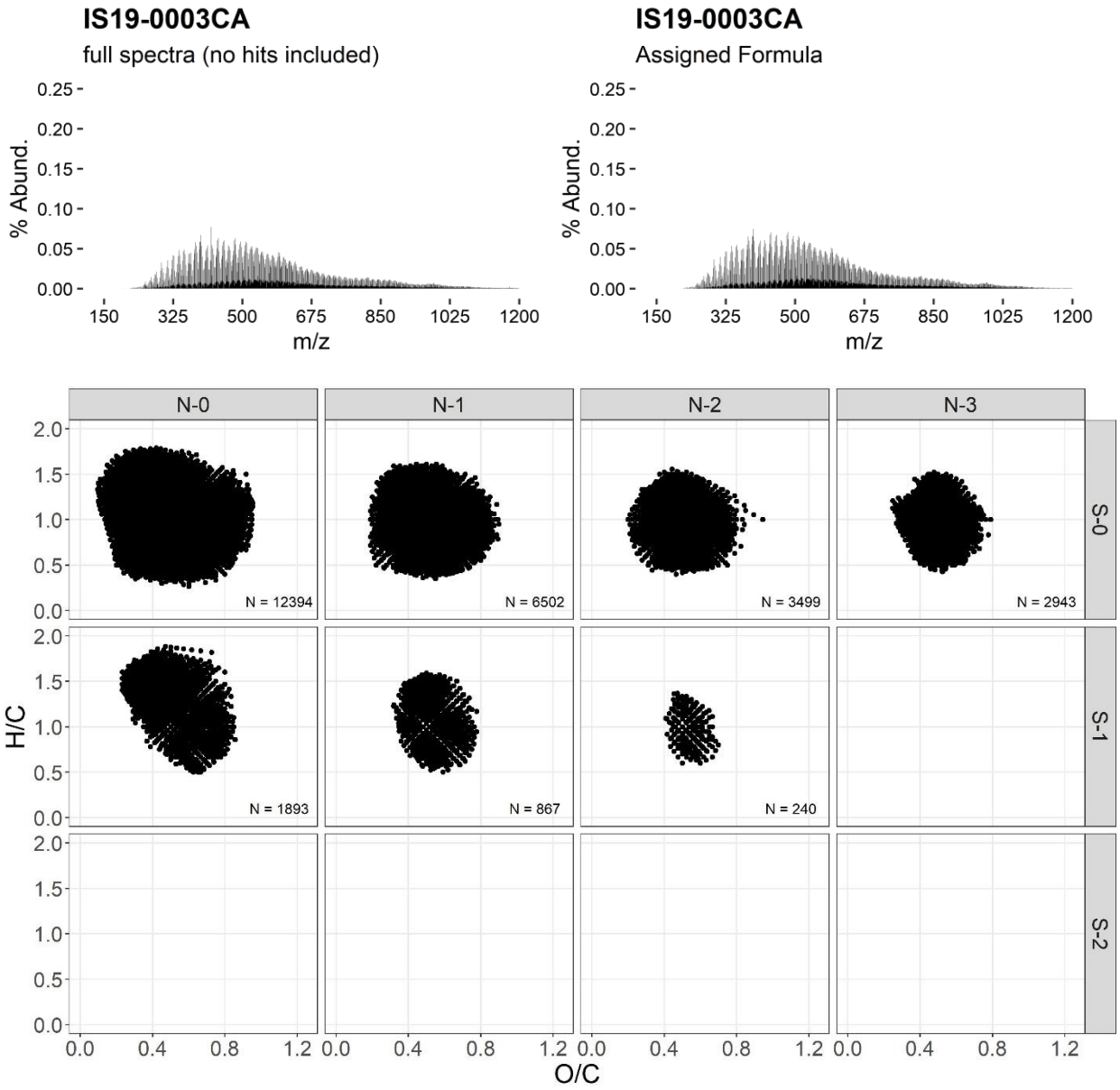
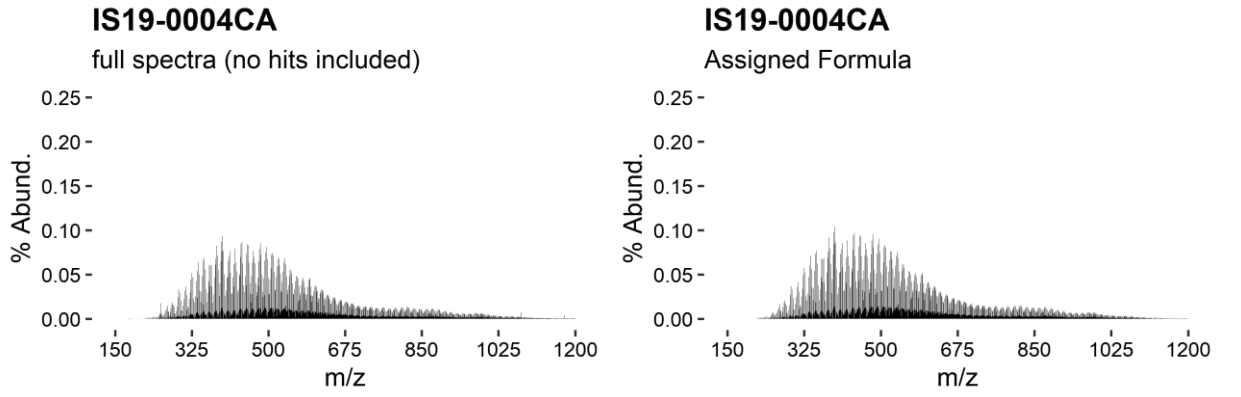


Figure S14. Mixed land use/land cover stream runoff 1 (M-RO1)



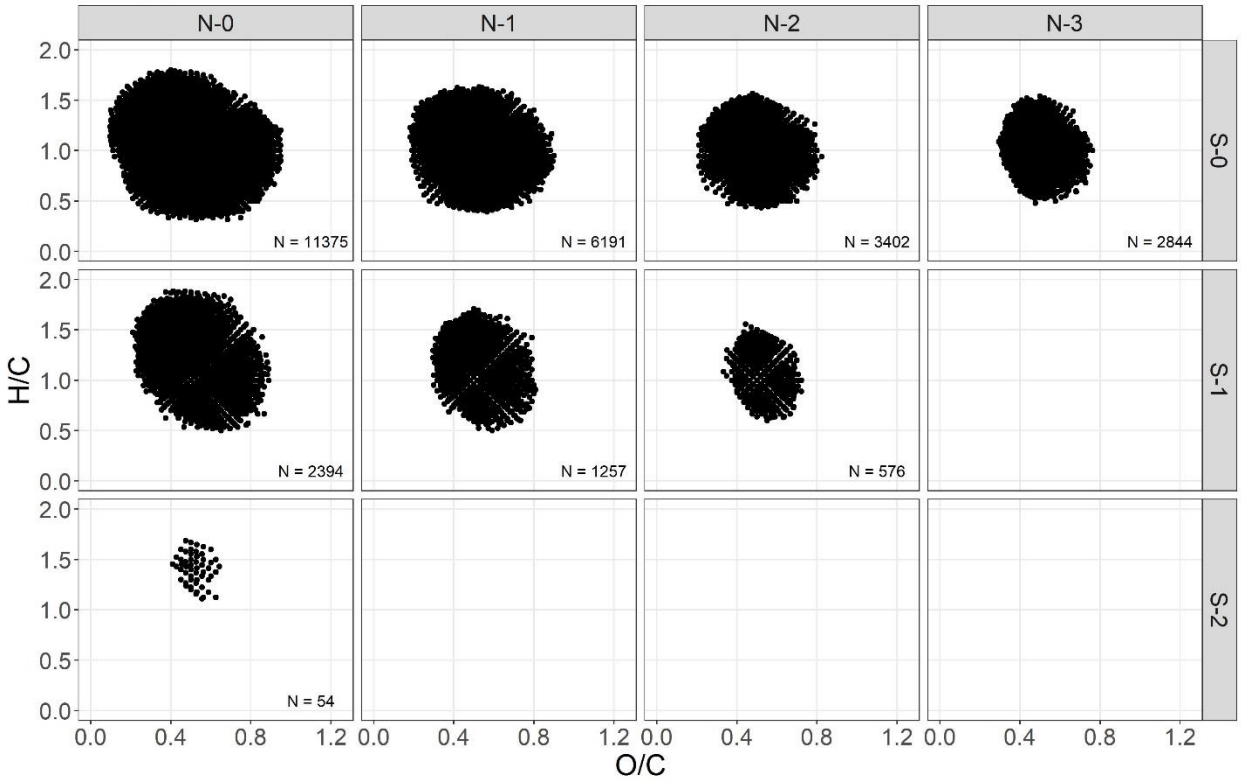
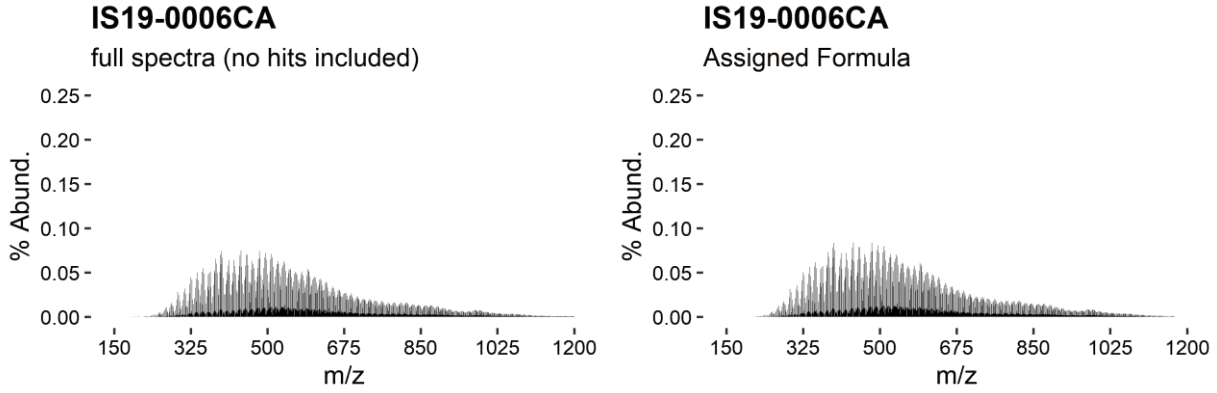


Figure S15. Mixed land use/land cover stream runoff 2 (M-RO2)



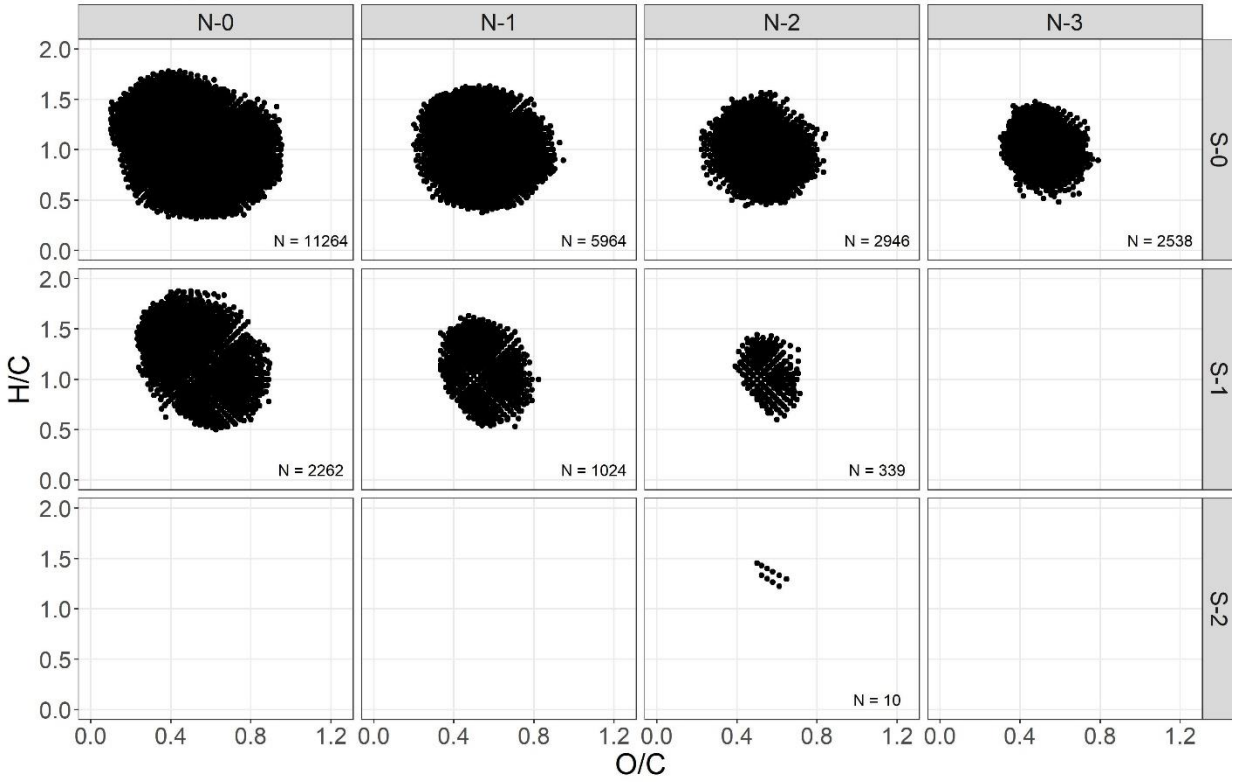
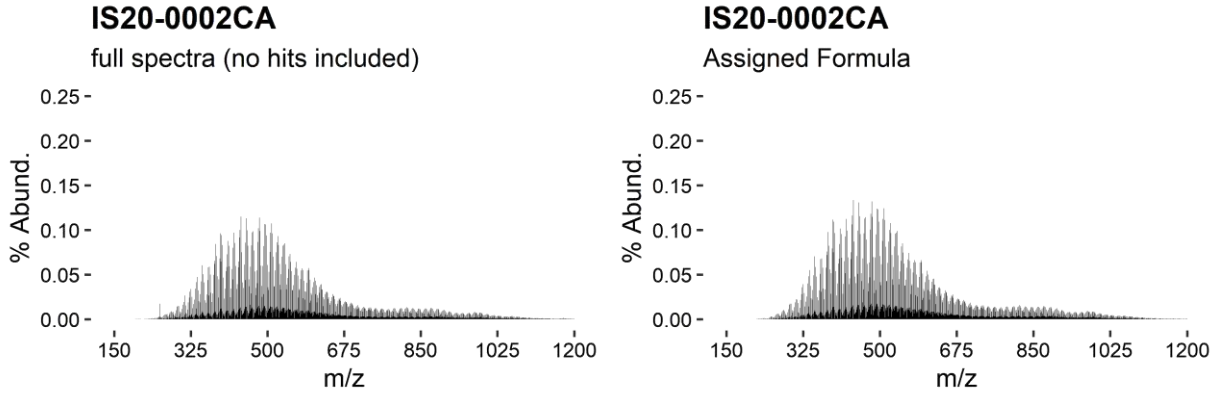


Figure S16. High flow Napa River stream runoff (HF N-RO)



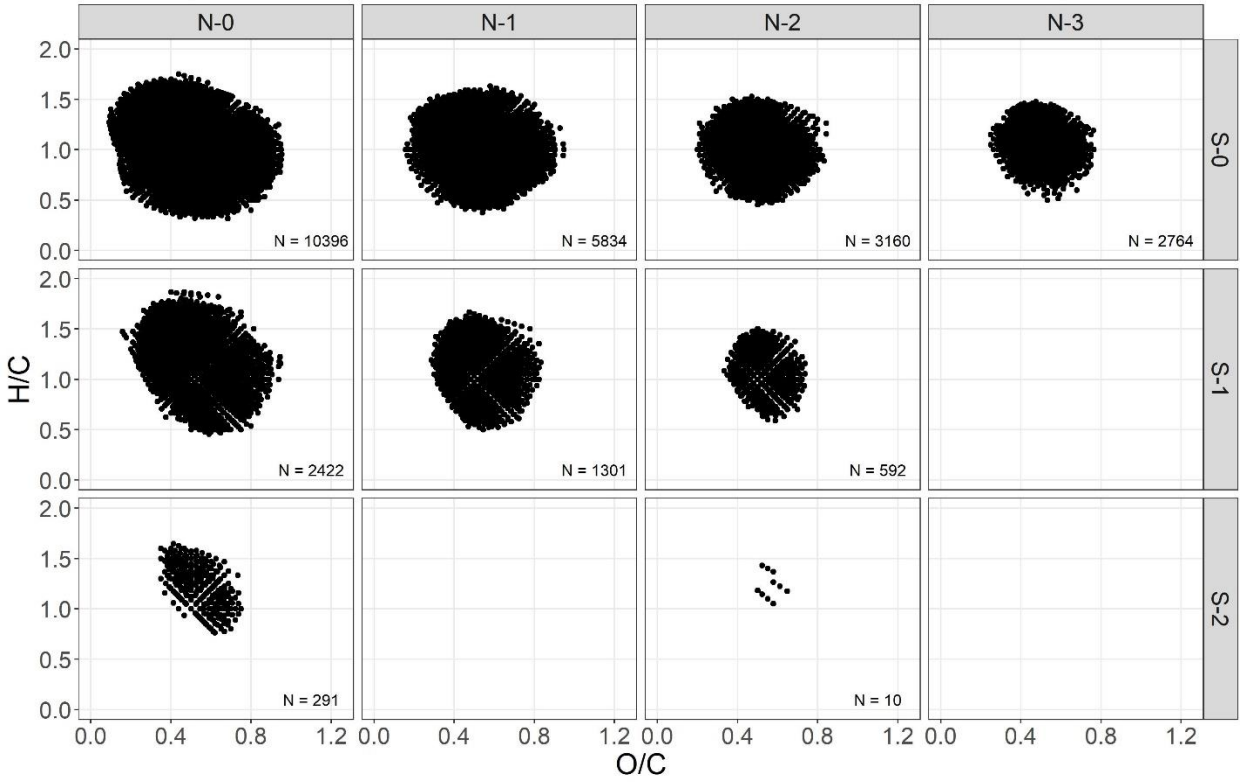


Figure S17. Low flow Napa River stream runoff (LF N-RO)

3.4 DOS OXIDATION STATE

Sulfur XANES results showed variable proportions of organic S oxidation states in soil porewaters and streams (Figure 4). Relative proportions of S functionalities did not differ systematically by LULC or soil/stream sample type. Instead, forest and grassland samples ranged from 42% reduced S (exocyclic + heterocyclic) in G-PW to 72% in F-PW, the highest reduced S proportion observed as primarily exocyclic reduced S (Table S8). Vineyard samples had either roughly equal proportions of reduced and oxidized species (48%, V-PW2) or higher oxidized S at 56-60% (V-PW1 and V-RO, respectively). Mixed LULC streams and the Napa River had 52 – 60% reduced S, similar to HF and LF F-RO, G-RO, and V-PW2. For all samples, sulfonate was

the primary form of oxidized S functionalities. When normalized to atomic S/C, vineyard soil porewater, V-PW2, had the highest reduced S content at 4.2×10^{-3} mmol S mmol C^{-1} (Figure 4b).

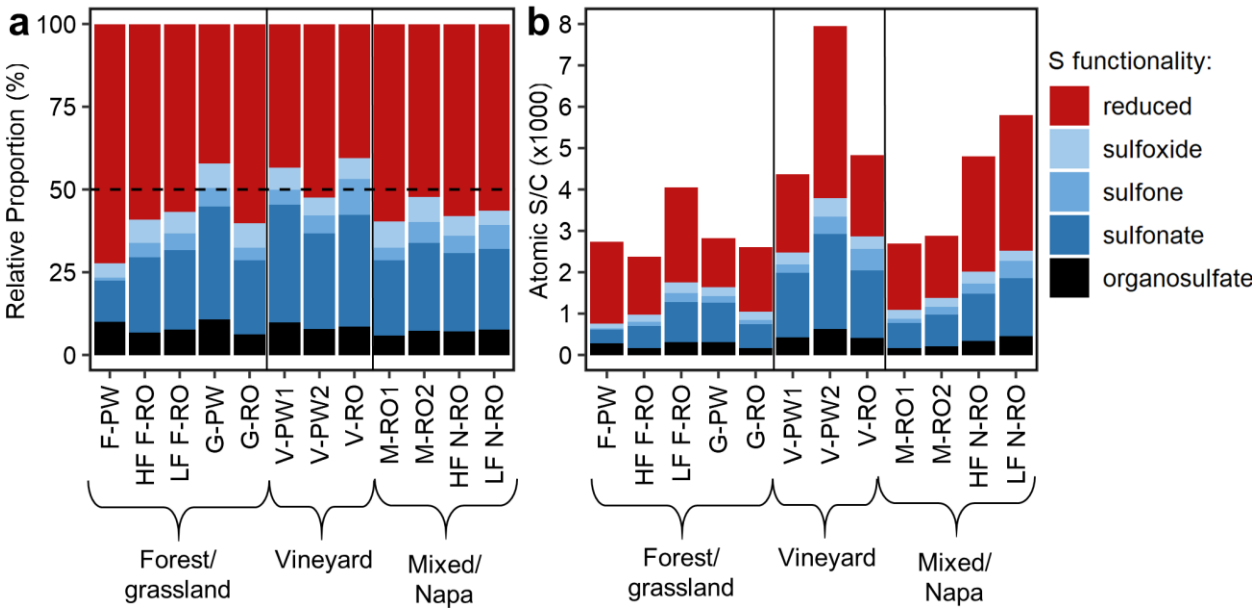


Figure 4. Sulfur K-edge XANES speciation results including the (a) relative proportions of reduced S (summation of exocyclic and heterocyclic reduced S), sulfoxide, sulfone, sulfonate, and organosulfate, and (b) The ratio of sulfur functionalities relative to C in the DOM (as atomic S/C). The dashed line in subplot a represents equal proportions of reduced and oxidized-S species. Sample abbreviations are as in Figure 1.

3.5 Sulfur stable isotope ratios

Sulfur stable isotope ratios were distinct between vineyard and non-agricultural samples, largely driven by differences in $\delta^{34}S\text{-}SO_4^{2-}$ values rather than $\delta^{34}S$-DOS values. Inorganic $\delta^{34}S\text{-}SO_4^{2-}$ values ranged from -1.7 to 12‰ (Figure 5a), and vineyard samples were enriched in ^{34}S by ~12‰ (11.3 ± 2.4‰) compared to forest and grassland samples (-0.81 ± 2.7‰,). Mixed stream runoff 1 (M-RO1) fell within the forest and grassland range of $\delta^{34}S\text{-}SO_4^{2-}$ values at 0.025‰, whereas M-RO2 was more similar to vineyard values at 4.7‰. Napa River samples were within the vineyard $\delta^{34}S\text{-}SO_4^{2-}$ at high and low flows, 6.7‰ and 12.0‰, respectively. Organic S stable isotope values ranged from 0.56 to 8.0‰, spanning a smaller range in $\delta^{34}S$ ratios than inorganic $\delta^{34}SO_4^{2-}$ values (Figure 5b). $\delta^{34}S$-DOS values did not appear to differ by LULC but did differ in

relation to $\delta^{34}S\text{-}SO_4^{2-}$ values. Non-agricultural $\delta^{34}S$-DOS values were 0.44 to 6.2‰ enriched relative $\delta^{34}S\text{-}SO_4^{2-}$ values, whereas vineyard $\delta^{34}S$-DOS values were 2.4 to 6.5‰ depleted relative to $\delta^{34}S\text{-}SO_4^{2-}$ values. The difference between $\delta^{34}S$-DOS and $\delta^{34}S\text{-}SO_4^{2-}$ was negatively correlated with the %S heteroatoms of the DOM ($R^2 = 0.77$). Multivariate PCA analysis reinforced separation in DOS chemistry between vineyard and non-agricultural LULC types (SI Figure S21).

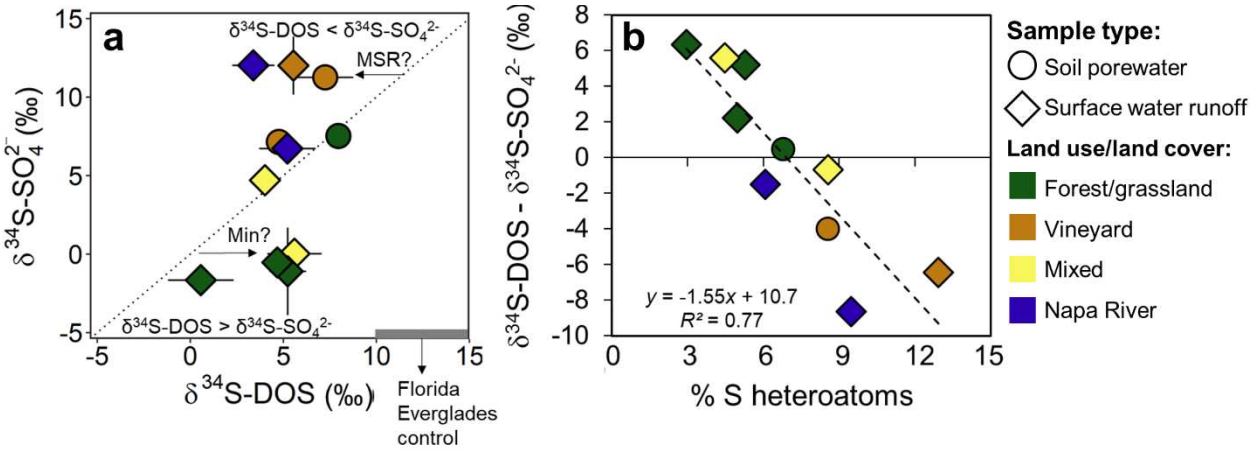


Figure 5. (a) Sulfur stable isotope values for inorganic sulfate ($\delta^{34}S\text{-}SO_4^{2-}$) and organic S ($\delta^{34}S$-DOS). Error bars show ± 1 SD. Dotted line is the 1:1 line of equal isotope values, and the gray bar shows the range of $\delta^{34}S$-DOS measured for control samples from the Florida Everglades (see SI). (b) Relationship between the difference between organic and inorganic sulfur stable isotopes and the percent of S-containing heteroatom formulas as measured by FT-ICR MS.

4. Discussion

4.1 Agricultural S additions increase DOS content

Agricultural S applications increased the S-content of DOM relative to surrounding non-agricultural areas, supporting our first hypothesis. Increased DOS content was apparent in bulk [S]-DOM measurements (Table S3), the FT-ICR MS-derived formulaic S/C and S-containing formula abundances (Figure 2), as well as the strong correlation between SO_4^{2-} concentration and the abundance of S-containing heteroatoms in DOM (Figure S2). The general agreement between bulk measurements and FT-ICR MS results suggests that FT-ICR MS can be used semi-quantitatively to assess DOS content, as previously observed in the Florida Everglades by Poulin

et al.[46]. Our measurements of vineyard atomic S/C (4.4-8.0 × 10^{-3} mmol S mmol C^{-1}) were in between those from agriculturally-non-impacted and impacted surface waters of the Florida Everglades (~4-5 × 10^{-3} mmol S mmol C^{-1} and ~7.5-8.5 × 10^{-3} mmol S mmol C^{-1}, respectively), a highly reducing environment where S is abiotically incorporated into DOM[46]. Higher S-heteroatom content in vineyard samples is also consistent with prior studies showing that agricultural areas have higher heteroatom content in DOM compared to forested ecosystems[17,19,53].

4.2 Agricultural vs. forest and grassland DOS sources and transformations

Dissolved organic S molecular composition and speciation provided insights into DOS sources and transformations in agricultural and non-agricultural soils and streams. Although organic S molecular composition did differ between agricultural and non-agricultural samples, there was also significant formula overlap (Figure 3c), implicating shared S sources and transformations across land use types. This suggests that our second hypothesis that agricultural DOS is compositionally distinct from non-agricultural DOS is partially validated. For example, soil porewater samples from both vineyard agriculture and non-agricultural areas likely reflect microbial S cycling during the early-season sampling campaign, which is supported by high proportions of unsaturated aliphatic, peptide-like, and sugar-like DOM formulas (Figure 3b) and lower DOM $SUVA_{254}$ (Table S2) compared to stream water samples[17,54,55]. One vineyard (V-PW1) and grassland soil porewater (G-PW) had a higher proportion of oxidized than reduced S (Figure 4a), also consistent with a microbial-derived DOS source[56]. Our findings that microbial OM predominated in soil porewaters from both vineyard and non-agricultural areas differs from prior studies showing that cropland streams typically have higher microbial-derived DOM than forests[19,57,58]. In contrast to soil porewater samples, non-agricultural streams (G-RO, HF F-RO, LF

F-RO, and M-RO1) strongly reflect flushing of fresh, plant-derived OM stored in soils[17,19], with lower overall heteroatom content (Table S5), higher $SUVA_{254}$ values, higher AI_{mod}, and higher average formula mass (Figure 5b; Table S2; Table S7).

Table S5. FT-ICR MS assigned formulas. Formula classes show number of formulas and the percentage of CHOS + CHNOS formulas are relative to the total number of formulas (i.e. not intensity-weighted).

Sample Code	Land use/land cover	Sample Type	Total Assigned Formulas	CHO	CHOS	CHNO	CHNOS	% CHOS + CHNOS	$CHOS_2$
F-PW	forest/grassland	soil porewater	20,250	9,430	1,905	8,127	788	13.3	0
HF F-RO	forest/grassland	surface water	27,790	11,799	1,956	12,750	1,285	11.7	0
LF F-RO	forest/grassland	surface water	27,653	11,997	1,960	12,426	1,270	11.7	0
G-PW	forest/grassland	soil porewater	22,409	8,002	1,764	11,330	1,313	13.7	0
G-RO	forest/grassland	surface water	24,564	12,708	1,579	9,760	517	8.5	0
V-PW1[a]	vineyard	soil porewater	17,149	5,754	1,803	8,268	1,324	18.2	0
V-PW2	vineyard	soil porewater	29,128	11,359	2,506	13,161	2,098	15.8	0
V-RO	vineyard	surface water	26,972	10,449	2,886	11,441	2,189	18.8	352
M-RO1	mixed	surface water	28,338	12,394	1,893	12,944	1,107	10.6	0
M-RO2	mixed	surface water	28,093	11,375	2,448	12,437	1,833	15.2	54
HF N-RO	mixed	surface water	26,347	11,264	2,262	11,448	1,363	13.8	0
LF N-RO	mixed	surface water	26,770	10,396	2,713	11,758	1,893	17.2	291

[a] Vineyard soil pore water 1 (V-PW1) had a poor quality spectra, and was therefore excluded from analyses (see "FT-ICR MS spectra and van Krevelen diagrams" section)

Sulfur stable isotope patterns further elucidated differences in DOS transformations across LULC types. Enriched δ^{34}S-DOS values of forest and grassland stream water relative to δ^{34}S-SO_4^{2-} is consistent with measurements in the Black Forest, Germany and spruce forests in the Czech Republic and has been attributed to plant-OM S mineralization, which leads to a slight enrichment in ^{34}S-DOS relative to SO_4^{2-}[59,60]. Samples that had nearly equivalent δ^{34}S-DOS and δ^{34}S-SO_4^{2-} values (Figure 5a) could either reflect DOS produced via pathways with minimal fractionation or mixing of agricultural and non-agricultural S sources. Since FT-ICR MS and S XANES results suggest that soil porewater DOS reflected a microbial source, similar δ^{34}S-DOS and δ^{34}S-SO_4^{2-} values in F-PW and V-PW1 could be further evidence of assimilatory SO_4^{2-} reduction since this process results in minimal S fractionation (~1-3‰)[24,61]. Mixing of sources is a more likely explanation for mixed LULC stream runoff (M-RO2) and the Napa River at high flow (HF N-RO), since stream samples had higher DOM $SUVA_{254}$ and molecular evidence pointed to a flushing of fresh DOS (Table S2).

Vineyard and Napa River samples with δ^{34}S-DOS values lower than δ^{34}S-SO_4^{2-} could either represent simply a difference in organic and inorganic S sources or could be an indicator for SO_4^{2-} reduction followed by DOM sulfurization. Evidence for abiotic incorporation of reduced S species into DOM has grown in recent years across a range of ecosystems including prairie potholes[62], the Florida Everglades[46], groundwater in the Adirondack Mountains, New York[24], mangrove roots[63], Antarctic lakes[64], and oceans[50,65]. In most studies, sulfurization is tied to anaerobic sulfidic conditions, but SO_4^{2-} reduction and subsequent DOM sulfurization could also occur in soil anoxic microsites during periods of intermittent-to-sustained soil saturation[66–68]. We invoke the latter mechanism here because of the strong correlation between the difference in δ^{34}S-DOS and δ^{34}S-SO_4^{2-} values and S heteroatom content, with δ^{34}S-DOS < δ^{34}S-SO_4^{2-} values having

a higher percentage of S-containing formulas (Figure 5). Conversely, if differences in the DOS source were responsible for this pattern, we would have expected S heteroatom content to be uniform across δ^{34}S-DOS and δ^{34}S-SO_4^{2-} values which was not the case. Notably, however, the proposed sulfurization of OM within vineyards is not as extensive as measured in in more reduced ecosystems (e.g., sulfidic wetlands and lakes)[62,64].

4.3 Sub-catchment-to-watershed scale DOS chemistry

Mixed LULC stream and Napa River organic S chemistry provided insights into the traceability of agricultural S through the Napa River watershed. Streams that receive a combination of inputs from different land covers had DOS chemistry that reflected the predominant land cover. For example, the M-RO1 catchment area is 91.6% non-agricultural land cover (Table S1), and M-RO1 DOS chemistry was more similar to non-agricultural samples (F-PW, HF F-RO, and G-RO) than vineyard samples (Figure 3a, Figure S21). In contrast, the M-RO2 catchment is 23.2% vineyard agriculture LULC (Table S1), and DOS had low dissimilarity values and similar S speciation compared to vineyard soil porewater (Figure 3a; Figure 4), Vineyard-derived DOS in M-RO2 was reinforced by the presence, albeit low abundance, of $CHOS_2$ formulas in the vineyard culvert runoff (V-RO; Figure 2c; Figure S4; Figure S5; Table S5), detection of which was enabled by the very high mass resolution of the FT-ICR MS analysis. These results suggest that vineyard-derived DOS is transported beyond fields, is traceable, and influences the DOS chemistry of receiving stream waters, supporting our third hypothesis.

At the watershed scale, Napa River samples showed a shift in DOS composition between high and low flow conditions. At high flow (HF N-RO), the DOS chemistry shared characteristics with non-agricultural DOS, with a lower abundance of S-containing heteroatoms (Figure 2c) and

depleted $\delta^{34}S\text{-}SO_4^{2-}$ values (Figure 5a). At low flow (LF N-RO), the DOS chemistry was strongly influenced by vineyard DOS, with a higher proportion of S-containing heteroatoms (Figure 2c), enriched $\delta^{34}S\text{-}SO_4^{2-}$ values (Figure 5a), and the presence of vineyard-derived $CHOS_2$ formulas (Figure 3c; Figure S4; Figure S5). Several studies have attributed enhanced $CHOS_1$ content in rivers to wastewater inputs[69]. Although there are a number of relatively small wastewater treatment plants within the Napa River watershed, Napa River FT-ICR MS spectra did not have surfactant-like low-molecular weight peaks (typical for domestic wastewater) and had an average m/z of ~ 516, higher than the average of 314 m/z measured from wastewater effluent[69] (Table S6, S7; Figure S16; Figure S17). These results suggest that Napa River DOS is more likely reflecting the vineyard S source. Combining multiple analyses strengthened our ability to differentiate the vineyard DOS signal in this complex and dynamic watershed.

Importantly, only 1% and 12% of Napa River assigned formulas were unique at high and low flow, respectively (Figure S4), suggesting that our endmember sampling captured much of the watershed's molecular diversity. The Napa River results indicate that non-agricultural DOS dominates at watershed scales during high flow, while vineyard-derived DOS dominates at low flows. One explanation for the change in source contributions across flow conditions is that vineyard hydrology is intensively controlled through growing season irrigation and subsurface drainage, potentially extending the duration of vineyard hydrologic transport longer than in the surrounding ecosystems. Nevertheless, our samples represent snapshots of the watershed's hydrology, and future studies are crucial to directly link daily-to-seasonal hydrologic dynamics and DOS composition and transport.

Table S6. FT-ICR MS intensity-weighted summary indices, all formulas

Sample Code	Land use/land cover	Sample Type	Mass	AImod	NOSC	H/C	O/C
F-PW	forest/grassland	soil porewater	566.35	0.302	0.055	1.09	0.55
HF F-RO	forest/grassland	surface water	611.63	0.319	0.033	1.07	0.53
LF F-RO	forest/grassland	surface water	617.61	0.339	0.050	1.04	0.53
G-PW	forest/grassland	soil porewater	564.30	0.255	-0.110	1.19	0.51
G-RO	forest/grassland	surface water	627.27	0.331	0.027	1.06	0.53
V-PW2	vineyard	soil porewater	576.12	0.322	0.049	1.07	0.53
V-RO	vineyard	surface water	582.87	0.307	-0.063	1.11	0.49
M-RO1	mixed	surface water	616.24	0.332	0.026	1.06	0.52
M-RO2	mixed	surface water	598.12	0.306	-0.025	1.10	0.51
HF N-RO	mixed	surface water	616.79	0.318	0.037	1.06	0.53
LF N-RO	mixed	surface water	594.08	0.312	0.010	1.09	0.52

Table S7. FT-ICR MS intensity-weighted summary indices, $CHOS_{1-2}$ formulas

Sample Code	Land use/land cover	Sample Type	Mass	AImod	NOSC	H/C	O/C
F-PW	forest/grassland	soil porewater	497.33	0.137	-0.107	1.31	0.55
HF F-RO	forest/grassland	surface water	505.55	0.144	-0.135	1.30	0.53
LF F-RO	forest/grassland	surface water	528.51	0.191	-0.060	1.21	0.53
G-PW	forest/grassland	soil porewater	487.97	0.096	-0.221	1.38	0.53
G-RO	forest/grassland	surface water	517.40	0.149	-0.145	1.29	0.53
V-PW2	vineyard	soil porewater	501.69	0.171	-0.049	1.25	0.55
V-RO	vineyard	surface water	511.25	0.171	-0.169	1.27	0.50
M-RO1	mixed	surface water	510.13	0.152	-0.168	1.30	0.52
M-RO2	mixed	surface water	509.05	0.154	-0.180	1.30	0.51
HF N-RO	mixed	surface water	517.40	0.149	-0.145	1.29	0.53
LF N-RO	mixed	surface water	516.10	0.188	-0.069	1.22	0.53

4.4 Implications for the fate of agricultural sulfur

When mobilized beyond agricultural fields, DOS remains reactive and is subject to additional transformations such as photodegradation in surface waters. In a laboratory experiment, Ossola et al.[20]found that 10-50% of DOS was mineralized to SO_4^{2-} after irradiation exposure equivalent to one midsummer day, and was enhanced with high DOS-content and microbial-derived OM—akin to the DOS found in vineyard culvert runoff (V-RO). The selective photooxidation of reduced DOS to SO_4^{2-} was confirmed by S XANES[21]. This result suggests that mobilized vineyard DOS likely undergo photodegradation, which could increase SO_4^{2-} concentrations and S mobility. However, if photodegradation significantly altered DOS along flow paths, we might expect that Napa River DOS would have a substantially different molecular composition than source areas. Instead, we found a high degree of formula commonality (Figure S4), suggesting that photodegradation and other S transformations do not strongly alter DOS chemistry at the spatial scale and under the flow conditions of this study.

Our characterization of both agricultural and non-agricultural DOS has implications for the ultimate fate and reactivity of agricultural S applications, especially related to trace elements of Hg and arsenic (As). Mercury methylation is often tied to microbial SO_4^{2-} reduction in wetlands and is enhanced in the presence of aromatic DOM with higher reduced S-content[14,70]. Our measurements show that both non-agricultural and vineyard areas mobilize DOM with these characteristics (Figure 3b; Figure 4b), and vineyard samples had high concentrations of SO_4^{2-} (Table S2). Hence, SO_4^{2-} and DOS mobilized from different LULC areas of a watershed may intersect along flow paths and within wetlands to affect Hg methylation processes or downstream DOM alterations in receiving wetlands. Further, arsenic sequestration and release from soils in wetlands is intimately linked to organic S speciation and sulfur cycling[71]. Our results inform where

process-based studies are needed to better understand the environmental fate of agricultural S. For example, the influence of agricultural S on Hg methylation processes are understudied in the redox gradients associated with stream hyporheic and riparian zones[72,73]. Furthermore, the presence of high Hg methylation rates in the San Pablo Baylands[74–76]—just downgradient of the Napa River Watershed—suggests regional-scale studies are necessary to evaluate the interactions of agricultural S runoff and Hg.

The combined use of δ^{34}S-DOS, S XANES, and FT-ICR MS analyses permitted an atomic- and molecular-level investigation of DOS transformations and identified an agricultural DOS fingerprint in a watershed with complex and dynamic S biogeochemistry. The sampling design enabled the connection between S source applications and DOS chemistry at catchment and watershed scales for the first time, providing an approach to trace this human perturbation to the global S cycle and to enable studies to determine the consequences of excess S in adjacent ecosystems. Together, the findings emphasize that the reactivity of agricultural S, and consequences for diverse biogeochemical processes, are intricately tied to the chemistry of surrounding land use/land cover types. Ultimately, this study reinforces the need to investigate interactions between S and elemental flows from surrounding ecosystems and across scales.

5. AUTHOR CONTRIBUTIONS

Merritt N. Logan designed the FT sampling method, conceived the FT sample processing approach, performed FT-ICR MS sample analysis and interpretation, provided code for processing FT-ICR MS results, developed code to visualize FT-ICR MS results with A. L. Hermes, trained and supervised A. L. Hermes in FT data QA/QC and interpretation, designed the table of contents graphic, and edited the manuscript. Anna L. Hermes conceptualized the project, performed in situ sampling and sample processing, acquired funding, coordinated the field campaign and sampling strategy, analyzed FT-ICR MS results with input from Merritt N. Logan, developed FT-ICR MS

result visualization code with Merritt N. Logan, performed and analyzed aqueous chemical composition, interpreted all results, and wrote and assembled the manuscript. Brett A. Poulin conceptualized the project, supported methods development, supported FT-ICR MS interpretation, conducted XANES analyses, trained A. L. Hermes in XANES data processing, interpreted results, and edited the manuscript. Amy M. McKenna acquired FT-ICR MS results, supported FT-ICR MS QAQC and interpretation, and edited the manuscript. Todd E. Dawson conceptualized the project, supported isotopic analyses, interpreted results, and edited the manuscript. Thomas Borch conceptualized the project, acquired funding, edited the manuscript, and supervised the project. Eve-Lyn S. Hinckley conceptualized the project, acquired funding, conducted field work and sample processing, interpreted results, edited the table of contents graphic, and supervised the project and manuscript assembly.

6. ACKNOWLEDGEMENTS

This research was funded by a National Science Foundation CAREER [EAR-1945388] to E.S. Hinckley and a University of Colorado Center for Water, Earth Science, and Technology.

George R. Aiken Endowed Memorial Research Fellowship to A. Hermes. Sulfur XANES analyses were performed at the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, which is supported by the U.S. Department of Energy, Office of Science, and Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. FT-ICR MS analyses were performed at Ion Cyclotron Resonance User Facility at the National High Magnetic Field Laboratory, which is supported by National Science Foundation Division of Materials Research and Division of Chemistry through Cooperative Agreement No. DMR-1644779 and the State of Florida. T.B. acknowledges support from the National Science Foundation (award # 2114868) and USDA National Institute of Food Agriculture through AFRI grant no. 2021-67019034608. We thank M. Cooper, cooperating vineyard managers and landowners, Napa

County Resource Conservation District, and the Napa County Regional Park and Open Space District for site access and field assistance, as well as S. Mambelli and D. Roth for assistance with sample analyses, and B. Wing for methods development support.

7. COPYRIGHT

8. REFERENCES

(1) Hinckley, E. L. S.; Crawford, J. T.; Fakhraei, H.; Driscoll, C. T. A Shift in Sulfur-Cycle Manipulation from Atmospheric Emissions to Agricultural Additions. *Nat. Geosci.* **2020**, 1–8. https://doi.org/10.1038/s41561-020-0620-3.

(2) Feinberg, A.; Stenke, A.; Peter, T.; Hinckley, E.-L. S.; Driscoll, C. T.; Winkel, L. H. E. Reductions in the Deposition of Sulfur and Selenium to Agricultural Soils Pose Risk of Future Nutrient Deficiencies. *Commun. Earth Environ.* **2021**, *2* (1), 101. https://doi.org/10.1038/s43247-021-00172-0.

(3) Zak, D.; Hupfer, M.; Cabezas, A.; Jurasinski, G.; Audet, J.; Kleeberg, A.; McInnes, R.; Kristiansen, S. M.; Petersen, R. J.; Liu, H.; Goldhammer, T. Sulphate in Freshwater Ecosystems: A Review of Sources, Biogeochemical Cycles, Ecotoxicological Effects and Bioremediation. *Earth-Science Rev.* **2021**, *212*, 103446. https://doi.org/10.1016/j.earscirev.2020.103446.

(4) Tabatabai, M. A. Importance of Sulphur in Crop Production. *Biogeochemistry* **1984**, *1* (1), 45–62.

(5) Germida, J. J.; Janzen, H. H. Factors Affecting the Oxidation of Elemental Sulfur in Soils. *Fertil. Res.* **1993**, *35* (1–2), 101–114. https://doi.org/10.1007/BF00750224.

(6) Likens, G. E.; Driscoll, C. T.; Buso, D. C.; Mitchell, M. J.; Lovett, G. M.; Bailey, S. W.; Siccama, T. G.; Reiners, W. A.; Alewell, C. The Biogeochemistry of Sulfur at Hubbard Brook. *Biogeochemistry* **2002**, *60* (3), 235–316. https://doi.org/10.1023/A:1020972100496.

(7) Hinckley, E.-L. S.; Fendorf, S.; Matson, P. Short-Term Fates of High Sulfur Inputs in Northern California Vineyard Soils. *Nutr. Cycl. Agroecosystems* **2011**, *89* (1), 135–142. https://doi.org/10.1007/s10705-010-9383-3.

(8) Gilmour, C. C.; Henry, E. A.; Mitchell, R. Sulfate Stimulation of Mercury Methylation in Freshwater Sediments. *Environ. Sci. Technol.* **1992**, *26* (20), 480–484.

(9) Bates, A. L.; Orem, W. H.; Harvey, J. W.; Spiker, E. C. Tracing Sources of Sulfur in the Florida Everglades. *J. Environ. Qual.* **2002**, *31*, 287–299. https://doi.org/10.2134/jeq2002.0287.

(10) Orem, W.; Gilmour, C.; Axelrad, D.; Krabbenhoft, D.; Scheidt, D.; Kalla, P.; McCormick, P.; Gabriel, M.; Aiken, G. Sulfur in the South Florida Ecosystem: Distribution, Sources, Biogeochemistry, Impacts, and Management for Restoration. *Crit. Rev. Environ. Sci. Technol.* **2011**, *41* (SUPPL. 1), 249–288. https://doi.org/10.1080/10643389.2010.531201.

(11) Poulin, B.; Gerbig, C. A.; Kim, C. S.; Stegemeier, J. P.; Ryan, J. N.; Aiken, G. R. Effects of Sulfide Concentration and Dissolved Organic Matter Characteristics on the Structure of Nanocolloidal Metacinnabar. *Environ. Sci. Technol.* **2017**. https://doi.org/10.1021/acs.est.7b02687.

(12) Graham, A. M.; Aiken, G. R.; Gilmour, C. C. Effect of Dissolved Organic Matter Source and Character on Microbial Hg Methylation in Hg−S−DOM Solutions. *Environ. Sci. Technol.* **2013**, *47*, 5746–5754. https://doi.org/10.1021/es400414a.

(13) Manceau, A.; Lemouchi, C.; Enescu, M.; Gaillot, A.-C.; Lanson, M.; Magnin, V.; Glatzel, P.; Poulin, B. A.; Ryan, J. N.; Aiken, G. R.; Gautier-Luneau, I.; Nagy, K. L. Formation of Mercury Sulfide from Hg(II)–Thiolate Complexes in Natural Organic Matter. *Environ. Sci. Technol.* **2015**, *49* (16), 9787–9796. https://doi.org/10.1021/ACS.EST.5B02522.

(14) Graham, A. M.; Cameron-Burr, K. T.; Hajic, H. A.; Lee, C.; Msekela, D.; Gilmour, C. C. Sulfurization of Dissolved Organic Matter Increases Hg–Sulfide–Dissolved Organic

Matter Bioavailability to a Hg-Methylating Bacterium. *Environ. Sci. Technol.* **2017**, *51* (16), 9080–9088. https://doi.org/10.1021/acs.est.7b02781.
(15) Mitchell, M. J.; Driscoll, C. T.; Fuller, R. D.; David, M. B.; Likens, G. E. Effect of Whole-Tree Harvesting on the Sulfur Dynamics of a Forest Soil. *Soil Sci. Soc. Am. J.* **1989**, *53* (3), 933–940. https://doi.org/10.2136/SSSAJ1989.03615995005300030050X.
(16) Mitchell, M. J.; Lovett, G.; Bailey, S.; Beall, F.; Burns, D.; Buso, D.; Clair, T. A.; Courchesne, F.; Duchesne, L.; Eimers, C.; Fernandez, I.; Houle, D.; Jeffries, D. S.; Likens, G. E.; Moran, M. D.; Rogers, C.; Schwede, D.; Shanley, J.; Weathers, K. C.; Vet, R. Comparisons of Watershed Sulfur Budgets in Southeast Canada and Northeast US: New Approaches and Implications. *Biogeochemistry* **2011**, *103* (1), 181–207. https://doi.org/10.1007/s10533-010-9455-0.
(17) Wagner, S.; Riedel, T.; Niggemann, J.; Vähätalo, A. V.; Dittmar, T.; Jaffé, R. Linking the Molecular Signature of Heteroatomic Dissolved Organic Matter to Watershed Characteristics in World Rivers. *Environ. Sci. Technol.* **2015**, *49* (23), 13798–13806. https://doi.org/10.1021/acs.est.5b00525.
(18) Seifert, A. G.; Roth, V. N.; Dittmar, T.; Gleixner, G.; Breuer, L.; Houska, T.; Marxsen, J. Comparing Molecular Composition of Dissolved Organic Matter in Soil and Stream Water: Influence of Land Use and Chemical Characteristics. *Sci. Total Environ.* **2016**, *571*, 142–152. https://doi.org/10.1016/J.SCITOTENV.2016.07.033.
(19) Spencer, R. G. M.; Kellerman, A. M.; Podgorski, D. C.; Macedo, M. N.; Jankowski, K.; Nunes, D.; Neill, C. Identifying the Molecular Signatures of Agricultural Expansion in Amazonian Headwater Streams. *J. Geophys. Res. Biogeosciences* **2019**, *124* (6), 1637–1650. https://doi.org/10.1029/2018JG004910.
(20) Ossola, R.; Tolu, J.; Clerc, B.; Erickson, P. R.; Winkel, L. H. E.; McNeill, K. Photochemical Production of Sulfate and Methanesulfonic Acid from Dissolved Organic Sulfur. *Environ. Sci. Technol.* **2019**, *53* (22), 13191–13200. https://doi.org/10.1021/acs.est.9b04721.
(21) Poulin, B. A. Selective Photochemical Oxidation of Reduced Dissolved Organic Sulfur to Inorganic Sulfate. *Environ. Sci. Technol. Lett.* **2023**. https://doi.org/10.1021/ACS.ESTLETT.3C00210.
(22) Bahureksa, W.; Tfaily, M. M.; Boiteau, R. M.; Young, R. B.; Logan, M. N.; McKenna, A. M.; Borch, T. Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation. *Environ. Sci. Technol.* **2021**, acs.est.1c01135. https://doi.org/10.1021/ACS.EST.1C01135.
(23) Manceau, A.; Nagy, K. L. Quantitative Analysis of Sulfur Functional Groups in Natural Organic Matter by XANES Spectroscopy. *Geochim. Cosmochim. Acta* **2012**, *99*, 206–223. https://doi.org/10.1016/j.gca.2012.09.033.
(24) Kang, P. G.; Mitchell, M. J.; Mayer, B.; Campbell, J. L. Isotopic Evidence for Determining the Sources of Dissolved Organic Sulfur in a Forested Catchment. *Environ. Sci. Technol.* **2014**, *48* (19), 11259–11267. https://doi.org/10.1021/es502563n.
(25) Fuller, R. D.; Mitchell, M. J.; Krouse, H. R.; Wyskowski, B. J.; Driscoll, C. T. Stable Sulfur Isotope Ratios as a Tool for Interpreting Ecosystem Sulfur Dynamics. *Water. Air. Soil Pollut.* **1986**, *28* (1–2), 163–171. https://doi.org/10.1007/BF00184078.
(26) Hinckley, E.-L. S.; Kendall, C.; Loague, K. Not All Water Becomes Wine: Sulfur Inputs as an Opportune Tracer of Hydrochemical Losses from Vineyards. *Water Resour. Res.*

2008, *44* (7), 1–14. https://doi.org/10.1029/2007WR006672.
(27) Hinckley, E.-L. S.; Matson, P. A. Transformations, Transport, and Potential Unintended Consequences of High Sulfur Inputs to Napa Valley Vineyards. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (34), 14005–14010. https://doi.org/10.1073/pnas.1110741108.
(28) Hermes, A. L.; Ebel, B. A.; Murphy, S. F.; Hinckley, E.-L. S. Fates and Fingerprints of Sulfur and Carbon Following Wildfire in Economically Important Croplands of California, U.S. *Sci. Total Environ.* **2021**, *750*. https://doi.org/10.1016/j.scitotenv.2020.142179.
(29) Hermes, A. L.; Dawson, T. E.; Hinckley, E.-L. S. Sulfur Isotopes Reveal Agricultural Changes to the Modern Sulfur Cycle. *Environ. Res. Lett.* **2022**, *17* (5). https://doi.org/10.1088/1748-9326/ac6683.
(30) Napa Valley Vintners. Napa valley fast facts https://napavintners.com/press/docs/napa_valley_fast_facts.pdf (accessed Jun 24, 2022).
(31) Arguez, A.; Durre, I.; Applequist, S.; Vose, R. S.; Squires, M. F.; Yin, X.; Heim, R. R.; Owen, T. W. NOAA's 1981–2010 U.S. Climate Normals: An Overview. *Bull. Am. Meteorol. Soc.* **2012**, *93* (11), 1687–1697. https://doi.org/10.1175/BAMS-D-11-00197.1.
(32) Blank, P. October Rainfall Report. *Napa County Resource Conservation District Monthly Newsletter*. November 2021.
(33) Napa County. Napa County Onerain https://napa.onerain.com/.
(34) Weishaar, J. L.; Aiken, G. R.; Bergamaschi, B. A.; Fram, M. S.; Fujii, R.; Mopper, K. Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. *Environ. Sci. Technol.* **2003**, *37* (20), 4702–4708. https://doi.org/10.1021/ES030360X.
(35) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A Simple and Efficient Method for the Solid-Phase Extraction of Dissolved Organic Matter (SPE-DOM) from Seawater. *Limnol. Ocean. Methods* **2008**, *6*, 230–235. https://doi.org/10.4319/lom.2008.6.230.
(36) Pohlabeln, A. M.; Dittmar, T. Novel Insights into the Molecular Structure of Non-Volatile Marine Dissolved Organic Sulfur. *Mar. Chem.* **2015**, *168*, 86–94. https://doi.org/10.1016/J.MARCHEM.2014.10.018.
(37) Li, Y.; Harir, M.; Uhl, J.; Kanawati, B.; Lucio, M.; Smirnov, K. S.; Koch, B. P.; Schmitt-Kopplin, P.; Hertkorn, N. How Representative Are Dissolved Organic Matter (DOM) Extracts? A Comprehensive Study of Sorbent Selectivity for DOM Isolation. *Water Res.* **2017**, *116*, 316–323. https://doi.org/10.1016/J.WATRES.2017.03.038.
(38) Stücheli, P. E.; Niggemann, J.; Schubert, C. J. Comparison of Different Solid Phase Extraction Sorbents for the Qualitative Assessment of Dissolved Organic Nitrogen in Freshwater Samples Using FT-ICR-MS. *J. Limnol.* **2018**, *77* (3), 400–411. https://doi.org/10.4081/jlimnol.2018.1791.
(39) Cao, D.; Rao, Z.; Geng, F.; Niu, H.; Shi, Y.; Cai, Y.; Kang, Y. Advanced Molecular-Fingerprinting Analysis of Dissolved Organic Sulfur by Electrospray Ionization-Fourier Transform Ion Cyclotron Resonance Mass Spectrometry Using Optimal Spray Solvent. *J. Environ. Sci. (China)* **2020**, *97*, 67–74. https://doi.org/10.1016/j.jes.2020.05.008.
(40) Hendrickson, C. L.; Quinn, J. P.; Kaiser, N. K.; Smith, D. F.; Blakney, G. T.; Chen, T.; Marshall, A. G.; Weisbrod, C. R.; Beu, S. C. 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer: A National Resource for Ultrahigh Resolution Mass Analysis. *J. Am. Soc. Mass Spectrom* **2015**, *26*, 1626–1632. https://doi.org/10.1007/s13361-015-1182-2.

(41) Smith, D. F.; Podgorski, D. C.; Rodgers, R. P.; Blakney, G. T.; Hendrickson, C. L. 21 Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures. *Anal. Chem.* **2018**, *90* (3), 2041–2047. https://doi.org/10.1021/ACS.ANALCHEM.7B04159/ASSET/IMAGES/LARGE/AC-2017-04159Q_0007.JPEG.

(42) Blakney, G. T.; Hendrickson, C. L.; Marshall, A. G. Predator Data Station: A Fast Data Acquisition System for Advanced FT-ICR MS Experiments. *Int. J. Mass Spectrom.* **2011**, *306* (2–3), 246–252.

(43) Xian, F.; Hendrickson, C. L.; Blakney, G. T.; Beu, S. C.; Marshall, A. G. Automated Broadband Phase Correction of Fourier Transform Ion Cyclotron Resonance Mass Spectra. *Anal. Chem.* **2010**, *82* (21), 8807–8812. https://doi.org/10.1021/AC101091W/SUPPL_FILE/AC101091W_SI_001.TIF.

(44) Savory, J. J.; Kaiser, N. K.; Mckenna, A. M.; Xian, F.; Blakney, G. T.; Rodgers, R. P.; Hendrickson, C. L.; Marshall, A. G. Parts-Per-Billion Fourier Transform Ion Cyclotron Resonance Mass Measurement Accuracy with a "Walking" Calibration Equation. *Anal. Chem* **2011**, *83*, 1732–1736. https://doi.org/10.1021/ac102943z.

(45) Corilo, Y. E. PetroOrg Software. The Florida Station University: Tallahassee, FL 2012.

(46) Poulin, B. A.; Ryan, J. N.; Nagy, K. L.; Stubbins, A.; Dittmar, T.; Orem, W.; Krabbenhoft, D. P.; Aiken, G. R. Spatial Dependence of Reduced Sulfur in Everglades Dissolved Organic Matter Controlled by Sulfate Enrichment. *Environ. Sci. Technol.* **2017**, *51* (7), 3630–3639. https://doi.org/10.1021/acs.est.6b04142.

(47) Ravel, B.; Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: Data Analysis for X-Ray Absorption Spectroscopy Using IFEFFIT. *J. Synchrotron Radiat.* **2005**, *12* (4), 537–541. https://doi.org/10.1107/S0909049505012719.

(48) R Core Team. R: A Language and Environment for Statistical Computing. R Foundation For Statistical Computing: Vienna 2021.

(49) Seidel, M.; Manecki, M.; Herlemann, D. P. R.; Deutsch, B.; Schulz-Bull, D.; Jürgens, K.; Dittmar, T. Composition and Transformation of Dissolved Organic Matter in the Baltic Sea. *Front. Earth Sci.* **2017**, *0*, 31. https://doi.org/10.3389/FEART.2017.00031.

(50) Gomez-Saez, G. V.; Dittmar, T.; Holtappels, M.; Pohlabeln, A. M.; Lichtschlag, A.; Schnetger, B.; Boetius, A.; Niggemann, J. Sulfurization of Dissolved Organic Matter in the Anoxic Water Column of the Black Sea. *Sci. Adv.* **2021**, *7* (25), 6199–6215. https://doi.org/10.1126/SCIADV.ABF6199.

(51) Oksanen, J.; Guillaume Blanchet, F.; Friendly, M.; Kindt, R.; Legendre, P.; McGlinn, D.; Minchin, P. R.; O'Hara, R. B.; Simpson, G. L.; Solymos, P.; Stevens, M. H. H.; Szoecs, E.; Wagner, H. Vegan: Community Ecology Package. 2020.

(52) Murtagh, F.; Legendre, P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *J. Classif.* **2014**, *31*, 274–295. https://doi.org/10.1007/s00357-014-9161-z.

(53) Roebuck, J. A.; Seidel, M.; Dittmar, T.; Jaffe, R. Controls of Land Use and the River Continuum Concept on Dissolved Organic Matter Composition in an Anthropogenically Disturbed Subtropical Watershed. *Environ. Sci. Technol.* **2020**, *54*, 195–206. https://doi.org/10.1021/acs.est.9b04605.

(54) Brooker, M. R.; Longnecker, K.; Kujawinski, E. B.; Evert, M. H.; Mouser, P. J. Discrete Organic Phosphorus Signatures Are Evident in Pollutant Sources within a Lake Erie Tributary. *Environ. Sci. Technol.* **2018**, *52* (12), 6771–6779.

https://doi.org/10.1021/acs.est.7b05703.
(55) Kellerman, A. M.; Ois Guillemette, F.; Podgorski, D. C.; Aiken, G. R.; Butler, K. D.; Spencer, R. G. M. Unifying Concepts Linking Dissolved Organic Matter Composition to Persistence in Aquatic Ecosystems. *Environ. Sci. Technol.* **2018**, *52*, 2538–2548. https://doi.org/10.1021/acs.est.7b05513.
(56) Edwards, P. J. *Sulfur Cycling, Retention, and Mobility in Soils: A Review*; 1998.
(57) Wilson, H. F.; Xenopoulos, M. A. Effects of Agricultural Land Use on the Composition of Fluvial Dissolved Organic Matter. *Nat. Geosci. 2009 21* **2008**, *2* (1), 37–41. https://doi.org/10.1038/ngeo391.
(58) Chen, S.; Du, Y. X.; Das, P.; Lamore, A. F.; Dimova, N. T.; Elliott, M.; Broadbent, E. N.; Roebuck, J. A.; Jaffé, R.; Lu, Y. H. Agricultural Land Use Changes Stream Dissolved Organic Matter via Altering Soil Inputs to Streams. *Sci. Total Environ.* **2021**, *796*, 148968. https://doi.org/10.1016/J.SCITOTENV.2021.148968.
(59) Mayer, B.; Fritz, P.; Prietzel, J.; Krouse, H. R. The Use of Stable Sulfur and Oxygen Isotope Ratios for Interpreting the Mobility of Sulfate in Aerobic Forest Soils. *Appl. Geochemistry* **1995**, *10* (2), 161–173. https://doi.org/10.1016/0883-2927(94)00054-A.
(60) Novák, M.; Kirchner, J. W.; Fottová, D.; Pr, E.; Acková, I. J.; Krám, P.; Hru, J. Isotopic Evidence for Processes of Sulfur Retention/Release in 13 Forested Catchments Spanning a Strong Pollution Gradient (Czech Republic, Central Europe). *Glob. Biogeochem. Cycles* **2005**, *19*, 4012. https://doi.org/10.1029/2004GB002396.
(61) Kaplan, I. R.; Rittenberg, S. C. Microbiological Fractionation of Sulphur Isotopes. *Microbiology* **1964**, *34* (2), 195–212.
(62) Sleighter, R. L.; Chin, Y.-P.; Arnold, W. A.; Hatcher, P. G.; McCabe, A. J.; McAdams, B. C.; Wallace, G. C. Evidence of Incorporation of Abiotic S and N into Prairie Wetland Dissolved Organic Matter. *Environ. Sci. Technol. Lett.* **2014**, *1* (9), 345–350. https://doi.org/10.1021/ez500229b.
(63) Raven, M. R.; Fike, D. A.; Gomes, M. L.; Webb, S. M. Chemical and Isotopic Evidence for Organic Matter Sulfurization in Redox Gradients Around Mangrove Roots. *Front. Earth Sci.* **2019**, *7* (May), 1–15. https://doi.org/10.3389/feart.2019.00098.
(64) Aiken, G.; Mcknight, D.; Harnish, R.; Wershaw, R. Geochemistry of Aquatic Humic Substances in the Lake Fryxell Basin, Antarctica. *Biogeochemistry* **1996**, *34* (3), 157–188. https://doi.org/10.1007/BF00000900/METRICS.
(65) Abdulla, H. A.; Burdige, D. J.; Komada, T. Abiotic Formation of Dissolved Organic Sulfur in Anoxic Sediments of Santa Barbara Basin. *Org. Geochem.* **2020**, *139*, 103879. https://doi.org/10.1016/j.orggeochem.2019.05.009.
(66) Hansel, C. M.; Fendorf, S.; Jardine, P. M.; Francis, C. A. Changes in Bacterial and Archaeal Community Structure and Functional Diversity along a Geochemically Variable Soil Profile. *Appl. Environ. Microbiol.* **2008**, *74* (5), 1620. https://doi.org/10.1128/AEM.01787-07.
(67) Barton, L. L.; Fauque, G. D. Chapter 2 Biochemistry, Physiology and Biotechnology of Sulfate-Reducing Bacteria. In *Advances in Applied Microbiology*; Laskin, A. I., Sariaslani, S., Gadd, G. M., Eds.; Elsevier Inc., 2009; Vol. 68, pp 41–98. https://doi.org/10.1016/S0065-2164(09)01202-7.
(68) Santana, M. M.; Dias, T.; Gonzalez, J. M.; Cruz, C. Transformation of Organic and Inorganic Sulfur– Adding Perspectives to New Players in Soil and Rhizosphere. *Soil Biol. Biochem.* **2021**, *160*, 108306. https://doi.org/10.1016/J.SOILBIO.2021.108306.

(69) Gonsior, M.; Zwartjes, M.; Cooper, W. J.; Song, W.; Ishida, K. P.; Tseng, L. Y.; Jeung, M. K.; Rosso, D.; Hertkorn, N.; Schmitt-Kopplin, P. Molecular Characterization of Effluent Organic Matter Identified by Ultrahigh Resolution Mass Spectrometry. *Water Res.* **2011**, *45*, 2943–2953. https://doi.org/10.1016/j.watres.2011.03.016.
(70) Graham, A. M.; Aiken, G. R.; Gilmour, C. C. Dissolved Organic Matter Enhances Microbial Mercury Methylation under Sulfidic Conditions. *Environ. Sci. Technol.* **2012**, *46* (2715–2723). https://doi.org/10.1021/es203658f.
(71) Langner, P.; Mikutta, C.; Kretzschmar, R. Arsenic Sequestration by Organic Sulphur in Peat. *Nat. Geosci. 2011 51* **2011**, *5* (1), 66–73. https://doi.org/10.1038/ngeo1329.
(72) Creswell, J. E.; Kerr, S. C.; Meyer, M. H.; Babiarz, C. L.; Shafer, M. M.; Armstrong, D. E.; Roden, E. E. Factors Controlling Temporal and Spatial Distribution of Total Mercury and Methylmercury in Hyporheic Sediments of the Allequash Creek Wetland, Northern Wisconsin. *J. Geophys. Res. Biogeosciences* **2008**, *113* (G2). https://doi.org/10.1029/2008JG000742.
(73) Poulin, B. A.; Aiken, G. R.; Nagy, K. L.; Manceau, A.; Krabbenhoft, D. P.; Ryan, J. N. Mercury Transformation and Release Differs with Depth and Time in a Contaminated Riparian Soil during Simulated Flooding. *Geochim. Cosmochim. Acta* **2016**, *176*, 118–138. https://doi.org/10.1016/J.GCA.2015.12.024.
(74) Marvin-DiPasquale, M. C.; Agee, J. L.; Bouse, R. M.; Jaffe, B. E. Microbial Cycling of Mercury in Contaminated Pelagic and Wetland Sediments of San Pablo Bay, California. *Environ. Geol.* **2003**, *43*, 260–267. https://doi.org/10.1007/s00254-002-0623-y.
(75) Marvin-DiPasquale, M.; Windham-Myers, L.; Agee, J. L.; Kakouros, E.; Kieu, L. H.; Fleck, J. A.; Alpers, C. N.; Stricker, C. A. Methylmercury Production in Sediment from Agricultural and Non-Agricultural Wetlands in the Yolo Bypass, California, USA. *Sci. Total Environ.* **2014**, *484*, 288–299.
(76) Robinson, A.; Richey, A.; Slotton, D.; Collins, J.; Davis, J. *North Bay Mercury Biosentinel Project*; 2018.

CHAPTER 6 - SUMMARY AND FUTURE OUTLOOK

This dissertation uniquely leverages the unparalleled capabilities of Fourier Transform Ion Cyclotron Resonance Mass Spectrometry, in combination with a suite of complementary analytical techniques, to unravel the intricate molecular transformations of organic matter in critically important, dynamic environmental systems. This work has demonstrated the powerful application of FT-ICR MS as a key tool for deciphering organic matter composition and transformation in complex environmental systems, providing novel insights into biogeochemical cycling of carbon, nitrogen, and sulfur across diverse ecosystems. These molecular mechanisms, previously underexplored, are crucial for understanding and predicting environmental responses to global change.

Specifically, this research was applied to two distinct yet equally vital environmental contexts: permafrost thaw transitions and agriculturally impacted watersheds. At the permafrost thaw front, FT-ICR MS characterization of porewater organic nitrogen (Chapter 2) revealed distinct molecular signatures. Our findings indicate that the thaw front is characterized by the lowest proportion of peptide-like compounds, while exhibiting the highest concentrations of ammonium and the greatest proportion of aromatic organic nitrogen. Gene expression associated with ammonium production pathways were observed to be higher in the bog compared to the palsa, suggesting active utilization of the dissolved organic nitrogen pool for ammonification. Despite increased ammonium, nitrate concentration remained unchanged, indicating nitrification limitation due to anaerobic conditions. This detailed molecular understanding of nitrogen dynamics is essential for refining models of nutrient availability and greenhouse gas production in thawing permafrost.

Further exploring permafrost dynamics, FT-ICR MS analysis of both dissolved and mineral-adsorbed organic carbon (Chapter 4) unveiled critical insights into carbon vulnerability during thaw. Our analysis of the porewater revealed a significant pulse of $Fe^{2+}_{(aq)}$ and DOC within the thaw front, attributed to the reductive dissolution of reactive iron(III) (oxyhydr)oxide minerals under soil saturation. We found that this released DOC consists of a higher proportion of more aromatic and less saturated molecules, a composition also observed in organic matter remaining adsorbed to mineral surfaces post-thaw. This demonstrated a profound transformation of the total carbon pool within permafrost during rapid thaw, even before complete permafrost loss. Consequently, this newly mobilized DOC can serve as a potent carbon source, readily converted to CO_2 and CH_4 gases, thereby directly contributing to accelerated warming and further permafrost thaw. Incorporating these molecular-level insights into future climate models is paramount.

Extending the application of this powerful tool beyond permafrost, FT-ICR MS was employed to investigate the composition and migration of organic sulfur (Chapter 5) in California vineyards. This work highlighted the impact of agricultural practices on watershed biogeochemistry, demonstrating that specific sulfur-containing organic molecules derived from vineyards can persist and be traced into downstream non-agricultural water systems. This unique molecular fingerprinting capability reveals critical pathways for contaminant transport and identifies novel compounds whose fate may influence biogeochemical processes such as mercury methylation in wetlands.

The extensive application of FT-ICR MS across these studies unequivocally underscores its utility in dissecting the complex molecular composition of organic matter within dynamic environmental systems. As this dissertation demonstrates, the ultra-high resolution

characterization facilitated by FT-ICR MS, when coupled with complementary analyses, provides indispensable context and validates identified organic matter compositions. The foundational methodology for applying FT-ICR MS to environmental systems was detailed in a critical review (Chapter 3), serving as a primer for robust environmental analysis. This coupled technique approach is especially valuable in systems influenced by multiple variables, enabling the extraction of molecular information not typically accessible.

Looking ahead, the frontiers of FT-ICR MS offer exciting avenues for advancing our understanding of biogeochemical cycling. Integrating techniques such as liquid chromatography (LC-FT-ICR MS) and tandem mass spectrometry (MS/MS) will provide even greater structural detail and enable the study of specific compound transformations with unprecedented precision. The development of new hybrid instruments, like LC-Orbitrap FT-ICR MS, also promise enhanced capabilities. Furthermore, the burgeoning fields of Artificial Intelligence (AI) and Machine Learning (ML) are poised to revolutionize the interpretation of the immense datasets generated by FT-ICR MS, facilitating the discovery of patterns and relationships that would otherwise remain hidden. To maximize the impact of these advancements, there is a critical need for standardization of FT-ICR MS methods and the development of comprehensive molecular databases, which will enable broader comparative studies and facilitate data sharing across the scientific community. These continuous improvements in molecular analysis are crucial for better understanding the subtle, yet profound, molecular changes occurring in permafrost regions and other ecosystems, and their resulting implications for climate change and ecosystem health. In conclusion, this dissertation showcases proficiency in biogeochemistry and unique expertise in approaching complex environmental systems to extract molecular-level information previously unobservable. This work has elucidated the molecular mechanisms driving carbon, nitrogen, and

sulfur cycling in response to environmental change, highlighting that the behavior of these systems, particularly permafrost, is dictated by molecular transformations that are not commonly explored. By demonstrating the power of FT-ICR MS as a non-common tool for environmental investigations, this research not only provides critical insights into current global challenges but also lays a robust foundation for future molecular-scale environmental science, paving the way for more accurate climate predictions and informed management strategies.

APPENDICES

APPENDIX A - SUPPLEMENTARY INFORMATION FOR CHAPTER 2

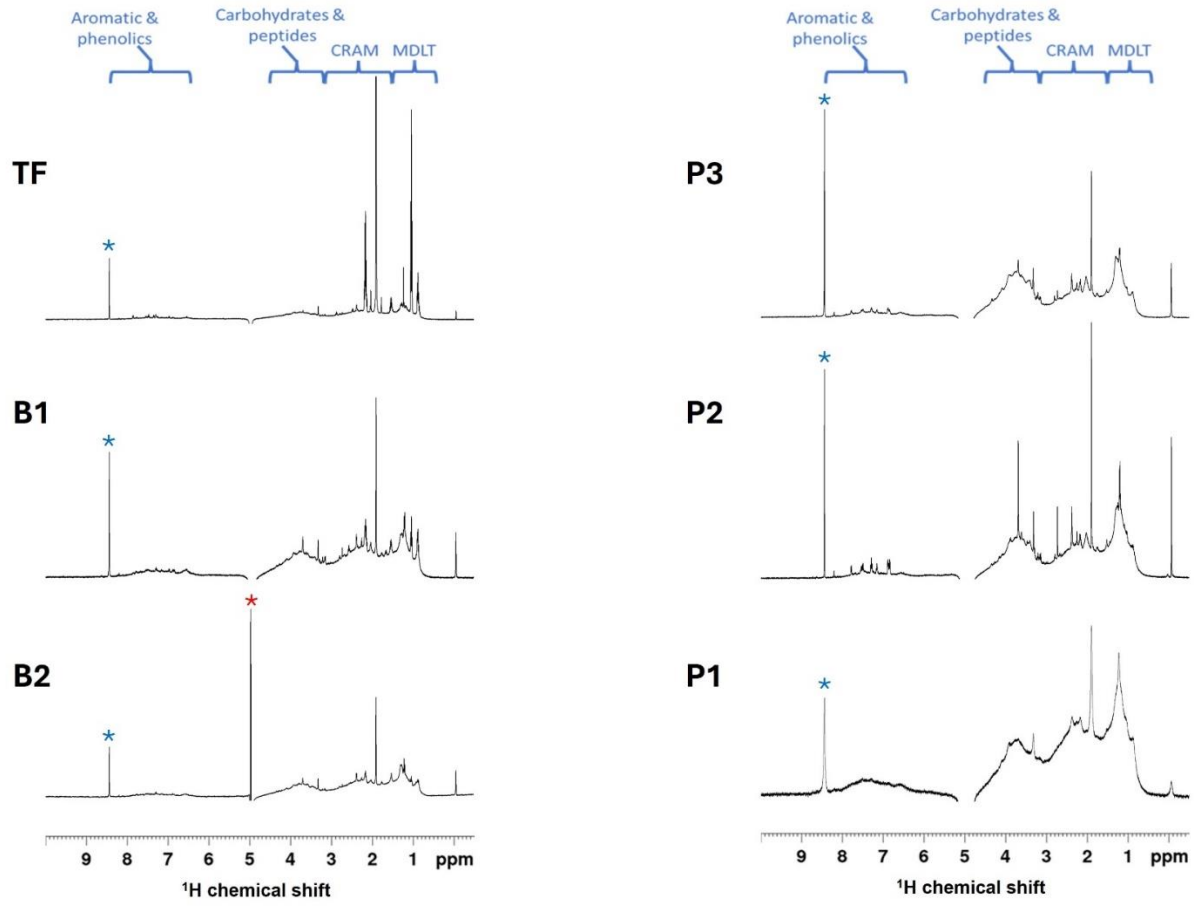


FIGURE S1. SOLUTION STATE ^{1}H NMR SPECTRA OF DISSOLVED ORGANIC MATTER SAMPLES

Samples are labeled as TF (thaw front), B1-2 (bog 1-2), P1-3 (palsa 1-3). All samples are from transect 4. Carboxyl-rich alicyclic molecules (CRAM) region also includes acetic acid (~1.90 ppm) and residual water is marked with a red asterisk and formic acid (~8.44 ppm) is marked with a blue asterisk. The palsa and bog samples exhibited higher proportions of carbohydrates and peptides compared to the thaw front. In contrast, the thaw front was characterized by a lower

proportion of carbohydrates and peptides and a higher proportion of CRAM. Chemical shift integration regions and assignments are based on Woods, 2011[81].

SECTION 1. METATRANSCRIPTOME LIBRARY AND READ ASSIGNMENTS

Metatranscriptome libraries were prepared for the six palsa and bog samples using the QIAseq FastSelect −5S/16S/23S (Qiagen) kit following the kit protocol with the following modifications: addition of probes for plant and yeast and using one-third the probe volumes. Then, the TruSeq Stranded Library Preparation kit (Illumina) was used to prepare the sequencing library. Libraries were sequenced on an Illumina NovaSeq 6000 system at the Genomics Core at the University of Colorado Anschutz Medical Campus.

Raw metatranscriptome reads were quality trimmed and adapters removed using bbduk[105] with the following flags: k=23 mink=11 hdist=1 qtrim=rl trimq=20 minlength=75. Reads were filtered with rqcfilter2[105] using the following flags: jni=t rna=t trimfragadapter=t qtrim=r trimq=0 maxns=1 maq=10 minlen=51 mlf=0.33 phix=t removeribo=t removehuman=t removedog=t removecat=t removemouse=t khist=t removemicrobes=t mtst=t sketch kapa=t clumpify=t tmpdir=null barcodefilter=f trimpolyg=5. Filtered and trimmed reads were mapped against the 99% dereplicated MAGs using bowtie2[106] (v2.4.5) using the following settings: (-D 10 -R 2 -N 1 -L 22 -i S,0,2.50). The output SAM file was converted to BAM using samtools[107] (v1.9), and filtered using the reformat.sh script in the bbtools package using: (idfilter=0.95 pairedonly=t primaryonly=t). Mapped reads were counted using featureCounts[108] (v1.5.3) using the following flags: (-t CDS -g ID -s 2 -p). Read counts were filtered to remove counts <3, and were converted to geTMM[109] in R. To curate nitrogen cycling reactions, geTMM values of genes

annotated as enzymes involved in the nitrogen cycle were summed within samples at a pathway level (Supporting Data 1). Differences in pathway expression between palsa and bog samples were analyzed using LIMMA[110] in R.

SECTION 2. FT-ICR MS INSTRUMENT DETAILS

The sample solution was infused through a microelectrospray source[111] (50 μm i.d. fused silica emitter) at 500 nL / min using a syringe pump. Typical conditions for negative ion formation were as follows: emitter voltage, -2.8-3.2 kV; S-lens RF level: 40%; and heated metal capillary temperature, 350°C. DOM extracts were analyzed with a custom-built hybrid linear ion trap FT-ICR mass spectrometer equipped with a 21 T superconducting solenoid magnet[52,90]. Ions were initially accumulated in an external multipole ion guide (1-5 ms) and released m/z-dependently by decrease of an auxiliary radio frequency potential between the multipole rods and the end-cap electrode[112]. Ions were excited to m/z-dependent radius to maximize the dynamic range and number of observed mass spectral peaks (32-64%)[112], and excitation and detection were performed on the same pair of electrodes[113]. The dynamically harmonized ICR cell in the 21 T FT-ICR is operated with 6 V trapping potential[112,114]. Time-domain transients of 3.1 seconds were conditionally co-added and acquired with the Predator data station that handled excitation and detection only, initiated by a TTL trigger from the commercial Thermo data station with 100 time-domain acquisitions averaged for all experiments[115]. Mass spectra were phase-corrected[116] and internally calibrated with 10-15 highly abundant homologous series that span the entire molecular weight distribution based on the "walking" calibration method[117]. Experimentally measured masses were converted from the International Union of Pure and Applied Chemistry (IUPAC) mass scale to the Kendrick mass scale[118] for rapid identification of homologous series

for each heteroatom class (i.e., species with the same non carbon or hydrogen heteroatom content, differing only in degree of alkylation)[119]. For each elemental composition, the heteroatom class, double bond equivalents (DBE = number of rings plus double bonds to carbon, DBE = C –h/2 + n/2 +1) and carbon number, C, were tabulated for subsequent generation of the relative abundance distributions and the graphical relative-abundance weighted images and van Krevelen diagrams[120]. Peaks with signal magnitude greater than six times the baseline root-mean-square (rms) noise at m/z 400 were exported to peak lists, and molecular formula assignments and data visualization were performed with PetroOrg © software[51,121,122]. Molecular formula assignments with an error >0.3 parts-per-million were discarded, and only chemical classes with a combined relative abundance of ≥0.15% of the total were considered.

TABLE S1. POREWATER SAMPLE VOLUMES

Porewater volumes collected across the thaw transition. While the exact volumes from transect 1 and transect 5 (marked with *) were not documented, they featured comparable volumes to that of transects 2-4 and provided sufficient volume for the required analyses. Cells marked with "-" indicate the transect did not feature that corresponding location.

	Transect 1 (mL)	Transect 2 (mL)	Transect 3 (mL)	Transect 4 (mL)	Transect 5 (mL)
Palsa 3	-	-	22	24	*
Palsa 2	*	26	24	44	*
Palsa 1	*	18	42	48	*
Thaw Front	*	17	21	26	*
Bog 1	*	14	19	16	*
Bog 2	*	29	23	15	-
Bog 3	*	30	-	-	-

TABLE S2. TRANSECT POREWATER CHEMISTRY

Porewater chemistry of each transect and sampling point. Transects (T) are aligned to thaw front position. Sample name (P3, P2, P1, thaw front, B1, B2, B3) refers to the position in the transect relative to the thaw front. The letter refers to the thaw state (P for palsa and B for bog), and the number refers to the number of meters from the thaw front. N-values for each thaw state include all samples across all transects. Cells marked with "-" indicate the transect did not feature that corresponding location.

Sample Name		**P3**	**P2**	**P1**	**Thaw front**	**B1**	**B2**	**B3**
DOC [mg C/L]	T1	-	38.72	40.80	558.25	362.68	292.90	171.23
Palsa n=13	T2	-	52.93	450.89	638.94	324.11	221.78	328.08
Thaw front n=5	T3	69.63	44.72	77.68	322.27	149.36	158.13	-
Bog n=11	T4	46.32	25.38	261.22	386.13	159.89	104.36	-
	T5	67.0	64.0	58.8	440.1	214.1	-	-
Nitrite [mg/L]	T1	-	0.03	0.02	0.03	0.02	0.02	0.02
Palsa n=10	T2	-	0.03	0.02	0.03	0.02	0.02	0.02
Thaw front n=4	T3	0.01	0.01	0.01	0.01	0.01	0.01	-
Bog n=10	T4	0.01	0.01	0.01	0.01	0.01	0.01	-
	T5				N/A			
Nitrate [mg/L]	T1	-	1.27	0.41	0.35	0.23	0.17	0.19
Palsa n=10	T2	-	0.14	0.12	0.12	0.12	0.14	0.15
Thaw front n=4	T3	0.19	0.15	0.12	0.08	0.11	0.12	-
Bog n=10	T4	0.11	0.13	0.14	0.17	0.14	0.15	-
	T5				N/A			
Ammonium [mg/L]	T1	-	1.62	0.37	6.39	1.44	0.55	0.19
Palsa n=13	T2	-	0.15	3.33	4.27	1.04	0.21	0.48
Thaw front n=5	T3	0.25	0.13	0.10	1.07	0.71	0.48	-
Bog n=11	T4	0.31	0.11	0.09	2.01	1.19	1.48	-
	T5	0.29	0.04	0.04	2.12	0.25	-	-
Fe^{2+} [mg/L]	T1	-	0	2.23	136.84	67.02	46.92	18.99
Palsa n=13	T2	-	1.68	76.52	225.09	82.66	62.00	143.54
Thaw front n=5	T3	6.70	2.79	10.61	122.32	27.93	65.91	-
Bog n=11	T4	2.18	2.94	1.08	109.25	80.24	104.25	-
	T5	1.5	0.85	0.39	95.1	67.78	-	-
DON [mg N/L]	T1							
	T2				N/A			
	T3							
	T4							
	T5	4.34	2.59	1.25	27.88	7.5	-	-

512 *TABLE S3. FT-ICR MS MOLECULAR CLASS ASSIGNMENT*

513 FT-ICR MS molecular class assignments based on atomic composition of the assigned formula,

514 adapted from Poulin et al 2017[91].

Condensed Aromatic-like	$Ai_{mod} > 0/67$
Aromatic/Polyphenol-like	$0.5 > Ai_{mod} > 0.67$
Highly Unsaturated Aliphatic-like	$Ai_{mod} < 0.5$; H/C <1.5; O/C <0.9
Unsaturated Aliphatic-like	H/C = 1.5-2.0; O/C < 0.9; N = 0
Saturated Fatty Acid-like	H/C ≥ 2.0; O/C ≤ 0.9
Peptide-like	H/C = 1.5-2.0; O/C < 0.9; N > 0
Sugar-like	O/C > 0.9

515

REFERENCES

(1) Olefeldt, D.; Roulet, N. T.; Bergeron, O.; Crill, P.; Bäckstrand, K.; Christensen, T. R. Net Carbon Accumulation of a High-Latitude Permafrost Palsa Mire Similar to Permafrost-Free Peatlands. *Geophys. Res. Lett.* **2012**, *39* (3). https://doi.org/10.1029/2011GL050355.

(2) Holmes, M. E.; Crill, P. M.; Burnett, W. C.; McCalley, C. K.; Wilson, R. M.; Frolking, S.; Chang, K. -Y.; Riley, W. J.; Varner, R. K.; Hodgkins, S. B.; McNichol, A. P.; Saleska, S. R.; Rich, V. I.; Chanton, J. P. Carbon Accumulation, Flux, and Fate in Stordalen Mire, a Permafrost Peatland in Transition. *Global Biogeochem. Cycles* **2022**, *36* (1). https://doi.org/10.1029/2021gb007113.

(3) Johansson, T.; Malmer, N.; Crill, P. M.; Friborg, T.; Åkerman, J. H.; Mastepanov, M.; Christensen, T. R. Decadal Vegetation Changes in a Northern Peatland, Greenhouse Gas Fluxes and Net Radiative Forcing. *Glob. Chang. Biol.* **2006**, *12* (12), 2352–2369. https://doi.org/10.1111/j.1365-2486.2006.01267.x.

(4) Woods, G.; Simpson, A. HILIC-NMR: Toward the Identification of Individual Molecular Components in Dissolved Organic Matter. *Environ. Sci. Technol.* **2011**, *45* (13), 5910. https://doi.org/10.1021/es201716u.

(5) Bushnell, B. BBtools.

(6) Langmead, B.; Salzberg, S. L. Fast Gapped-Read Alignment with Bowtie 2. *Nat. Methods* **2012**, *9* (4). https://doi.org/10.1038/nmeth.1923.

(7) Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25* (16). https://doi.org/10.1093/bioinformatics/btp352.

(8) Liao, Y.; Smyth, G. K.; Shi, W. FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* **2014**, *30* (7). https://doi.org/10.1093/bioinformatics/btt656.

(9) Smid, M.; Coebergh van den Braak, R. R. J.; van de Werken, H. J. G.; van Riet, J.; van Galen, A.; de Weerd, V.; van der Vlugt-Daane, M.; Bril, S. I.; Lalmahomed, Z. S.; Kloosterman, W. P.; Wilting, S. M.; Foekens, J. A.; IJzermans, J. N. M.; Martens, J. W. M.; Sieuwerts, A. M. Gene Length Corrected Trimmed Mean of M-Values (GeTMM) Processing of RNA-Seq Data Performs Similarly in Intersample Analyses While Improving Intrasample Comparisons. *BMC Bioinformatics* **2018**, *19* (1). https://doi.org/10.1186/s12859-018-2246-7.

(10) Smyth, G. K. Limma: Linear Models for Microarray Data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*; 2005. https://doi.org/10.1007/0-387-29362-0_23.

(11) Emmett, M. R.; White, F. M.; Hendrickson, C. L.; Shi, D. H.; Marshall, A. G. Application of Micro-Electrospray Liquid Chromatography Techniques to FT-ICR MS to Enable High-Sensitivity Biological Analysis. *J. Am. Soc. Mass Spectrom.* **1998**, *9* (4), 333–340. https://doi.org/10.1016/S1044-0305(97)00287-0.

(12) Smith, D. F.; Podgorski, D. C.; Rodgers, R. P.; Blakney, G. T.; Hendrickson, C. L. 21 Tesla FT-ICR Mass Spectrometer for Ultrahigh-Resolution Analysis of Complex Organic Mixtures. *Anal. Chem.* **2018**, *90* (3), 2041–2047. https://doi.org/10.1021/acs.analchem.7b04159.

(13) Hendrickson, C. L.; Quinn, J. P.; Kaiser, N. K.; Smith, D. F.; Blakney, G. T.; Chen, T.; Marshall, A. G.; Weisbrod, C. R.; Beu, S. C. 21 Tesla Fourier Transform Ion Cyclotron

Resonance Mass Spectrometer: A National Resource for Ultrahigh Resolution Mass Analysis. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (9), 1626–1632. https://doi.org/10.1007/s13361-015-1182-2.

(14) Kaiser, N. K.; McKenna, A. M.; Savory, J. J.; Hendrickson, C. L.; Marshall, A. G. Tailored Ion Radius Distribution for Increased Dynamic Range in FT-ICR Mass Analysis of Complex Mixtures. *Anal. Chem.* **2013**, *85* (1), 265–272. https://doi.org/10.1021/ac302678v.

(15) Chen, T.; Beu, S. C.; Kaiser, N. K.; Hendrickson, C. L. Note: Optimized Circuit for Excitation and Detection with One Pair of Electrodes for Improved Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rev. Sci. Instrum.* **2014**, *85* (6), 2012–2015. https://doi.org/10.1063/1.4883179.

(16) Boldin, I. A.; Nikolaev, E. N. Fourier Transform Ion Cyclotron Resonance Cell with Dynamic Harmonization of the Electric Field in the Whole Volume by Shaping of the Excitation and Detection Electrode Assembly. *Rapid Commun. Mass Spectrom.* **2011**, *25* (1), 122–126. https://doi.org/10.1002/rcm.4838.

(17) Blakney, G. T.; Hendrickson, C. L.; Marshall, A. G. Predator Data Station: A Fast Data Acquisition System for Advanced FT-ICR MS Experiments. *Int. J. Mass Spectrom.* **2011**, *306* (2–3), 246–252. https://doi.org/10.1016/j.ijms.2011.03.009.

(18) Xian, F.; Hendrickson, C. L.; Blakney, G. T.; Beu, S. C.; Marshall, A. G. Automated Broadband Phase Correction of Fourier Transform Ion Cyclotron Resonance Mass Spectra. *Anal. Chem.* **2010**, *82* (21), 8807–8812. https://doi.org/10.1021/ac101091w.

(19) Savory, J. J.; Kaiser, N. K.; Mckenna, A. M.; Xian, F.; Blakney, G. T.; Rodgers, R. P.; Hendrickson, C. L.; Marshall, A. G. Parts-Per-Billion Fourier Transform Ion Cyclotron Resonance Mass Measurement Accuracy with a "Walking" Calibration Equation. *Anal. Chem.* **2011**, *83*, 1732–1736.

(20) Kendrick, E. A Mass Scale Based on CH2 = 14.0000 for High Resolution Mass Spectrometry of Organic Compounds. *Anal. Chem.* **1963**, *35* (13), 2146–2154.

(21) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. Kendrick Mass Defect Spectroscopy: A Compact Visual Analysis for Ultrahigh-Resolution Broadband Mass Spectra. *Anal. Chem.* **2001**, *73*, 4676–4681.

(22) van Krevelen, D. W. "Graphical-Statistical Method for the Study of Structure and Reaction Processes of Coal." *Fuel* **1950**, *29*, 269–284.

(23) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **2003**. https://doi.org/10.1021/ac034415p.

(24) Bahureksa, W.; Borch, T.; Young, R. B.; Weisbrod, C.; Blakney, G. T.; McKenna, A. M. Improved Dynamic Range, Resolving Power, and Sensitivity Achievable with FT-ICR Mass Spectrometry at 21 T Reveals the Hidden Complexity of Natural Organic Matter. *Anal. Chem.* **2022**, *94* (32), 11382–11389.

(25) Corilo, Y. E. PetroOrg Software. The Florida State University, Tallahassee, Fl, USA 2012.

(26) Poulin, B. A.; Ryan, J. N.; Nagy, K. L.; Stubbins, A.; Dittmar, T.; Orem, W.; Krabbenhoft, D. P.; Aiken, G. R. Spatial Dependence of Reduced Sulfur in Everglades Dissolved Organic Matter Controlled by Sulfate Enrichment. *Environ. Sci. Technol.* **2017**, *51* (7), 3630–3639. https://doi.org/10.1021/acs.est.6b04142.

APPENDIX B - SUPPLEMENTARY INFORMATION FOR CHAPTER 3

SECTION 1. SUPPORTING TABLES, FIGURES, AND CALCULATIONS

Table. S1 Log *P* and *pKa* values predicted by PubChem[1] and MOSES[2] (via ChemDraw Professional), respectively, for common biomolecular classes associated with SOM. Structures and features relative to each of the examples is shown in Fig. S1. *Intrinsic *pKa* was measured for dioleoyl phosphatidylcholine[3]

Name	Class	Log *P*	*pKa1*	*pKa2*	*pKa3*
glucose	Carbohydrate	-2.6	12.294		
N-Acetyl-Muramic Acid	Amino sugar	-2.1	3.408		
N-Acetyl-D-Glucosamine	Amino sugar	-1.7	11.939		
Tryptophan	Amino Acid	-1.1	2.38	9.39	
Tyrosine	Amino Acid	-2.3	2.2	9.11	10.07
Glycine	Amino Acid	-3.2	2.37	9.78	
Dioleoyl phosphatidylcholine	Phospholipid	13.8	4.053*		
Valoneic acid	Tannin derivative	1.5	8.516		
Hexahydroxydiphenic acid	Tannin derivative	0.5	13.939		
Flavogallonic acid	Tannin derivative	1.6	2.593		
Chlorogenic acid	Lignin derivatives	-0.4	8.099		
Sinapic acid	Lignin derivatives	1.5	4.05		
Vanillic acid	Lignin derivatives	1.4	4.0	9.39	
Gentisic acid	Lignin derivatives	1.3	2.766		
Stearic acid	Fatty acid	7.4	4.75		
Palmitic acid	Fatty acid	6.4	4.75		

Table. S2 Common extraction methods used to extract different fractions of soil organic matter with example conditions used in literature.

Extraction	Example extraction	Target SOM fraction	Reference
Water	Ultrapure (18.2 MΩ), 20+ hours extraction time, and 200g/L (soil/solution)[4]	Water soluble SOM, weakly soil particle bound and most bioavailable SOM	[4–6]
Alkaline	0.5 M NaOH, 20 h, and 200 g/L (soil/solution)[7]	General SOM and SOM compounds with acidic functional groups	[4,7–9]
Mineral dissolution	#1 - 0.1 M dithionite, 16 hours; followed by 0.05 M HCl rinse 1 h; 167 g/L #2 – 0.25 M hydroxylamine hydrochloride, 0.25 M HCl, 50 °C, 1 h, 200 g/L[10]	For extraction #1: short range order Fe(III) oxides and crystalline Fe(III) oxides For extraction #2: short range order Fe(III) oxides	[11–14]

Organic	1 mL acetonitrile (ACN), 100 mg soil, shaken for 2 hours[15]	Low polarity SOM and SOM with less abundant acidic functional groups	5,15–17

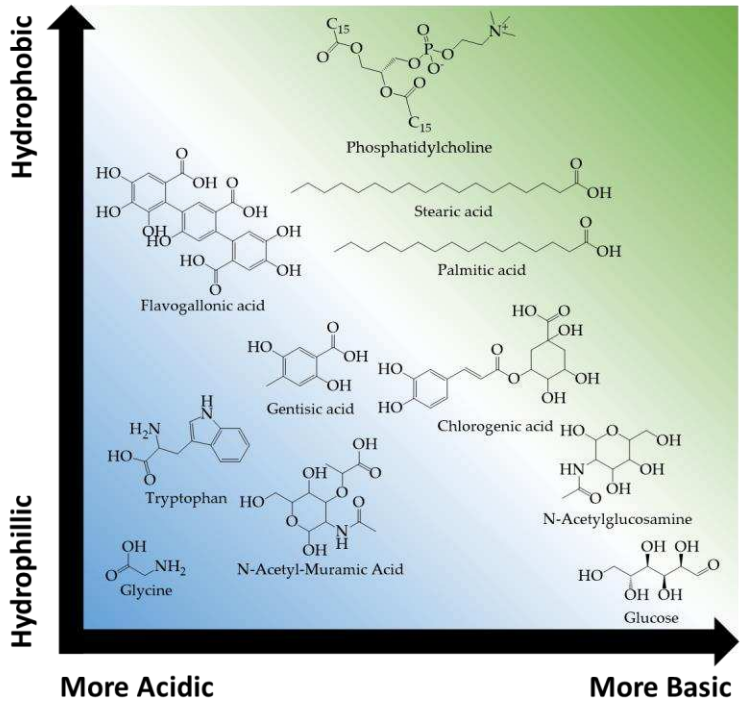


Fig. S1 Examples of SOM compounds with a variety of functional groups that result in a range of acidity and hydrophobicity, impacting their solubility and ionizability. These features determine their optimal extraction and ionization method.

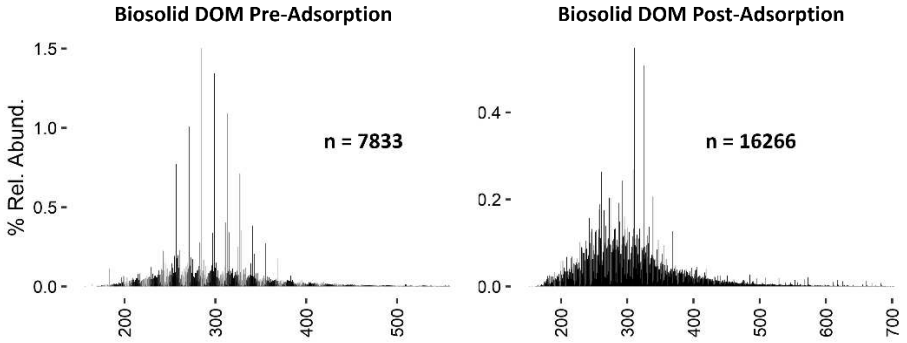


Fig. S2 FTICR mass spectra collected from biosolid-extracted DOM, pre- and post-sorption to Fe(III)-montmorillonite. Surfactant contaminants present in the native biosolid-extracted DOM were absent after sorption to Fe(III)-montmorillonite. Produced with data obtained from Young *et al*.[18]

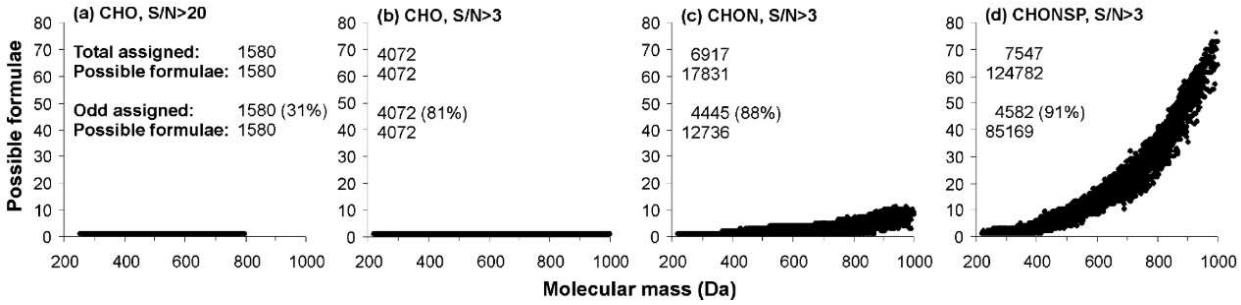


Fig. S3 Possible formula assignments for peaks at a given molecular mass under different heteroatom and S/N conditions. A) and B): the number of assignments increase as S/N thresholds are lowered. B), C), and D): possible compositions for a given peak increase as the number of heteroatoms included and molecular mass increases. Reproduced from Koch *et al*.[19]

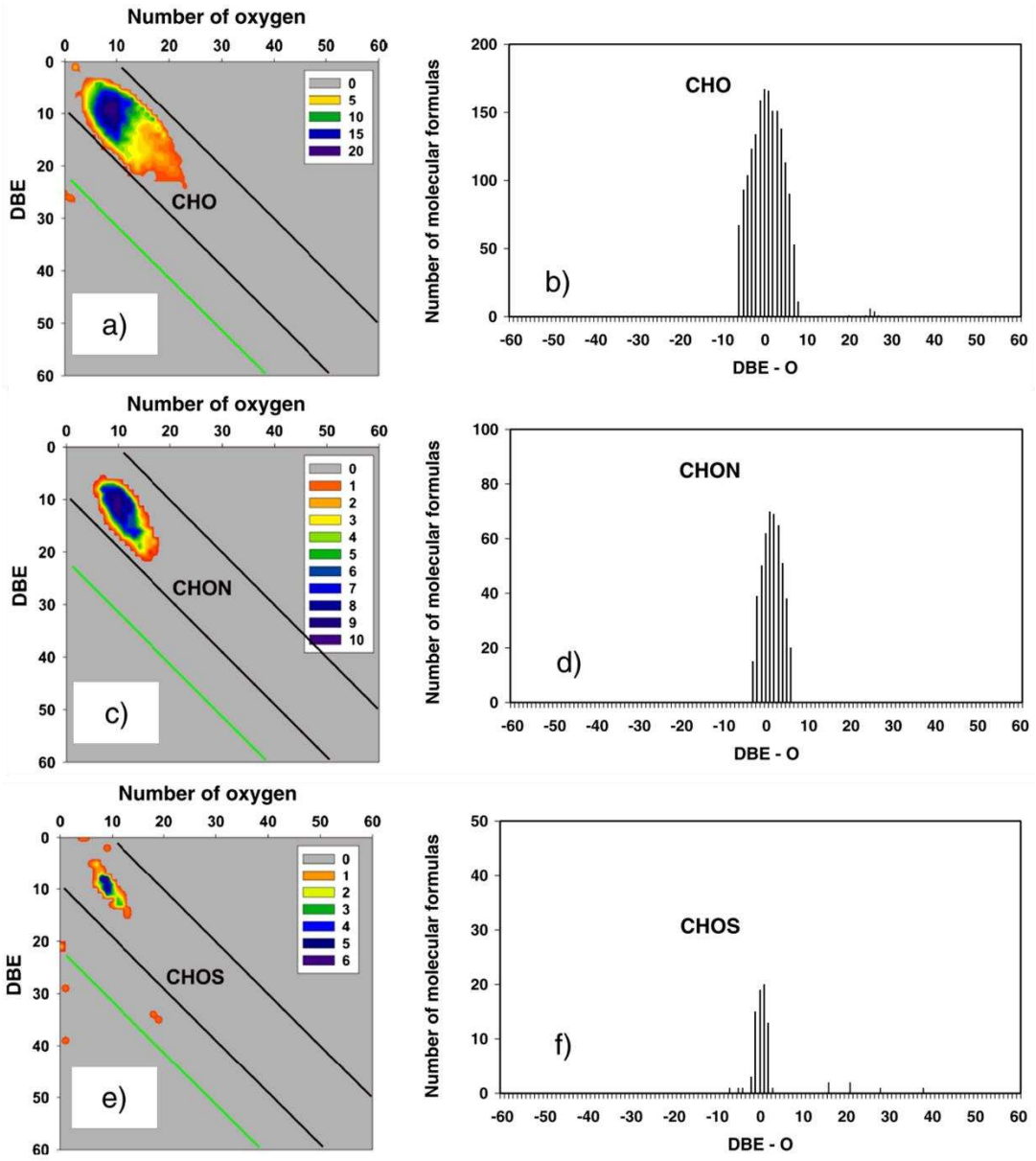


Fig. S4 Contoured plots of DBE vs. O (a, c, e) and bar graphs of DBE – O vs. number of molecular formulas (b, d, f) for unequivocal assignments in a Suwannee River fulvic acid standard, where color indicates number of *m/z* values. The positioning and number of *m/z* values from unequivocal assignments can be compared to equivocal assignments to indicate which compositions have the highest reliability, where reliable equivocal compositions will most frequently lie or overlap within the same regions as unequivocal assignments in the same class. Black lines are to indicate a diagonal range between -10 and 10 DBE – O, and the green line indicates regions where equivocal assignments would be unreliable due to their high hydrophobicity (e.g., black carbon components). Reprinted by permission from Copyright Clearance Center: Springer Nature, Analytical and Bioanalytical Chemistry, Understanding molecular formula assignment of Fourier transform ion cyclotron resonance mass spectrometry data of natural organic matter from a chemical point of view, Peter Herzsprung, Oct. 31 2014.[20]

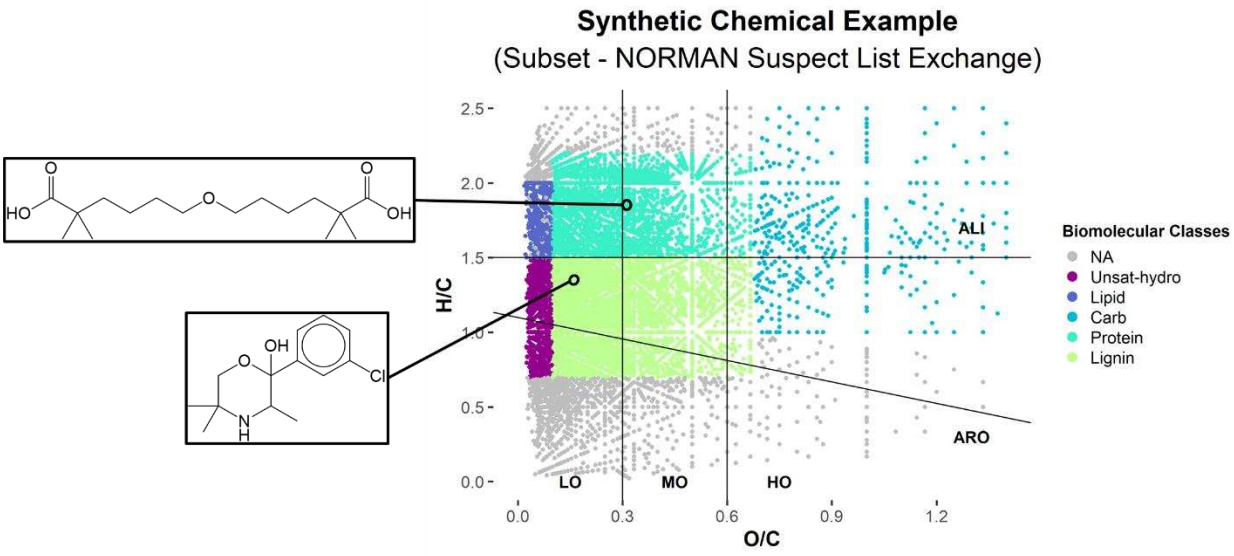


Fig. S5 NORMAN suspect list exchange[21] plotted onto a VKD with 2 of the molecules plotted to demonstrate that assignments unrelated to SOM are still plotted in regions pertaining to specific biomolecular assignments (regions are based on Hockaday 2009)[22]

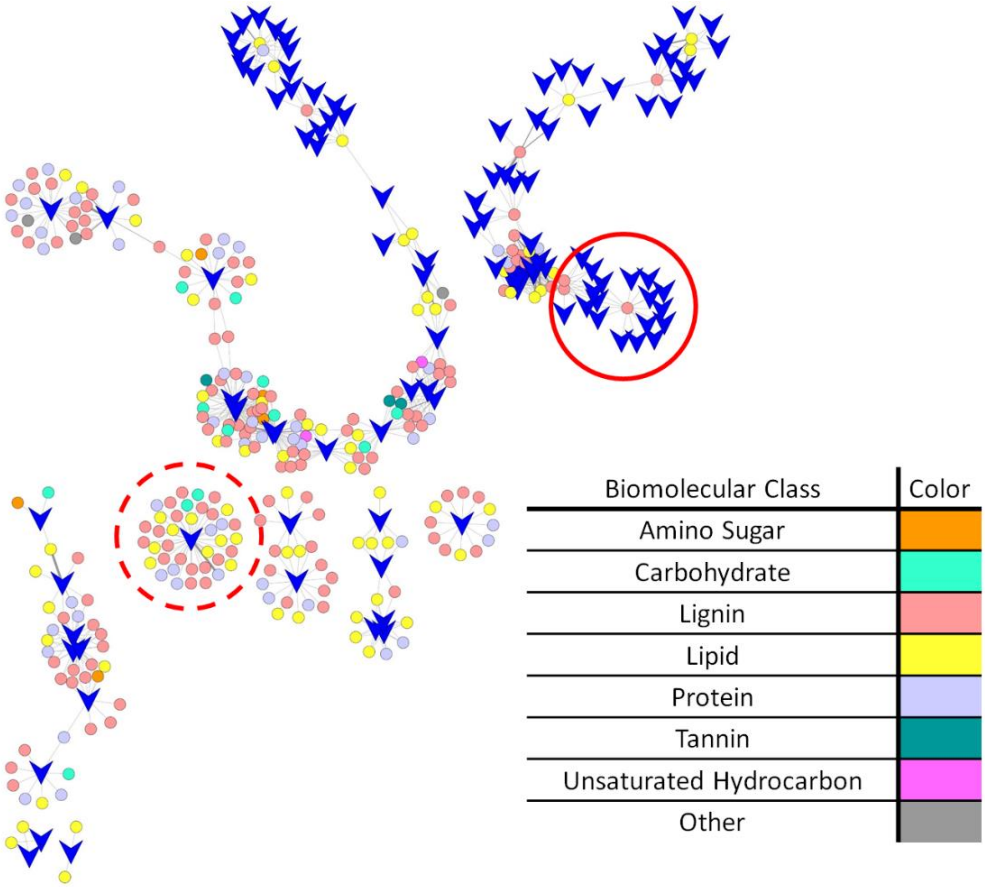


Fig. S6 Correlation networks between microbial communities (arrowhead) and metabolomic features identified from peaks (colored dots) from FTICR-MS data in peat soil. Pairing metabolomics by FTICR MS with amplicon sequencing can be used to identify how different microbial populations are related by specific biomolecular classes. Arrowheads linked to the same dot indicates that microbial populations are each related through the same peak (solid red circle), and dots surrounding an arrowhead indicates that the associated peaks are unique to that microbial population (dotted red circle). Unpublished data from the Tfaily Lab.

SECTION 2. COMPLEMENTARY SPECTROSCOPIC TECHNIQUES

Molecular characterization of SOM composition and reactivity often relies on combining FTICR MS with complementary analyses that provide additional information on bulk properties, elemental composition, molecular structure, or spatial distribution. The requirements for a 'clean' sample devoid of nonvolatile salts and the challenge of converting analytes into gas phase ions

means that there are significant biases that affect signal intensity beyond concentration. The complex nature of SOM makes identifying the diverse range of organic compounds (e.g., lipids, lignin, sugar, fermentation byproducts, charred residues) present in different soil types challenging to analyze. Furthermore, while FTICR MS detects molecular features, direct infusion mass spectra cannot by itself distinguish between structural isomers. Fortunately, additional information from other techniques can be used to attribute molecular formula to functional classes and to particular metabolites. FTICR MS is thus a central component of the analytical suite needed to understand metabolisms and organic matter transformations in soils. Here, examples utilizing spectroscopic techniques will be presented to illustrate how to utilize them to obtain robust conclusions.

Spectroscopic analyses that distinguish the functional groups and bonding environment of SOM components can provide significant quantitative bulk information that complements the molecular composition information of FTICR MS. NMR is particularly complementary because the major functional groups that comprise organic matter are often well resolved in NMR spectra, although the removal of paramagnetic elements such as iron in the samples is required. In previous studies combining these techniques to study marine and wetland dissolved organic matter, Hertkorn *et al*.[23,24] demonstrated that solid phase extracted DOM exhibited NMR shifts characteristic of aliphatic and carboxyl groups, while FTICR MS showed the abundance of diverse suites of unsaturated CHO, CHNO, CHOS molecules. They synthesized this information into the concept that carboxylic rich alicyclic molecules comprise a major component of DOM. Two-dimensional correlation analysis has been used to link FTICR MS and NMR data across a salinity transect by Abdulla *et al*. 2014.[25] The results of these analyses suggested that CHO compounds correlate strongly with carboxyl-rich compounds observed by H^1 NMR while heteropolysaccharides correlate with mass peaks which contain heteroatoms such as nitrogen,

phosphorous and sulfur. DiDonato *et al.* 2016[26] combined solid state ^{13}C NMR and FTICR MS to compare humic acids derived from well aerated versus poorly drained soils. They observed that aerated soils had a higher proportion of condensed mass spectral features, which was corroborated by the higher proportion of aromatic carbon observed by NMR. Together, these results suggested major differences in aliphatic, lignin-like, cyclic, and condensed aromatic classes across soil types (Fig. S7).

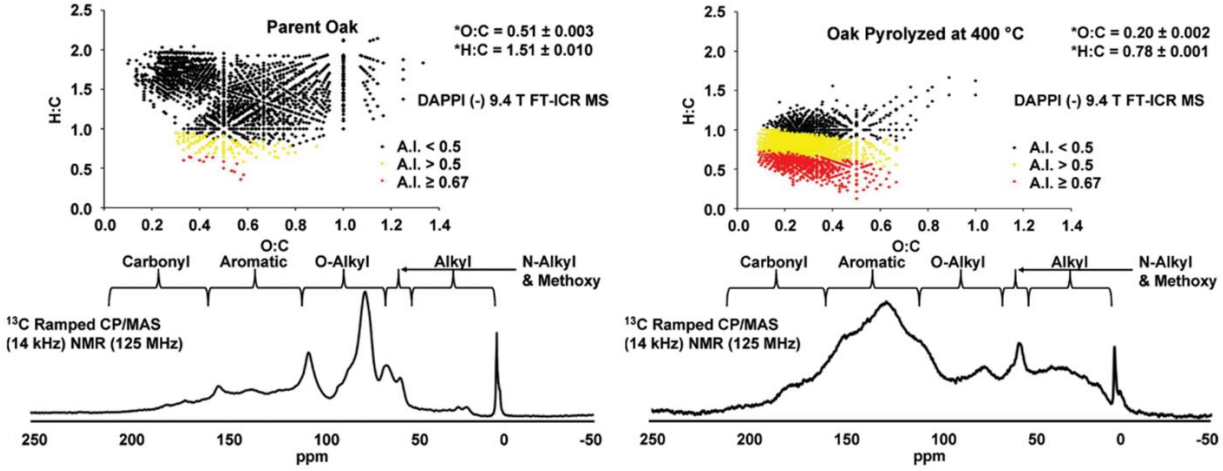


Fig. S7 Comparisons between unburnt (parent) and pyrolyzed (400ºC) oak (*Quercus laurifolia*) using (-)DAPPI FTICR MS with solid-state ^{13}C NMR to corroborate and describe changes to features during pyrolysis. Reproduced from Podgorski *et al*.[27]

Optical measurements have been used and continue to be used to infer organic matter quality through correlations with bulk molecular weight, aromaticity, and humic/fulvic and amino acid content in numerous studies.[28,29] Fourier transform infrared spectroscopy (FTIR) and fluorescence excitation emission spectroscopy (EEMS) can resolve organic functional groups including amide, ester, carbohydrate, phenol, aliphatic and carboxyl functionalities in DOM, and can thus corroborate some compound classes attributed to FTICR MS signals. Hodgkins *et al*.[30]

used FTIR and FTIRCR MS to compare peat organic matter structure across a thaw gradient. FTICR MS analysis suggested a trend from larger O/C rich molecules to smaller compounds with lower O/C ratios across the thaw gradient (Fig. S8).

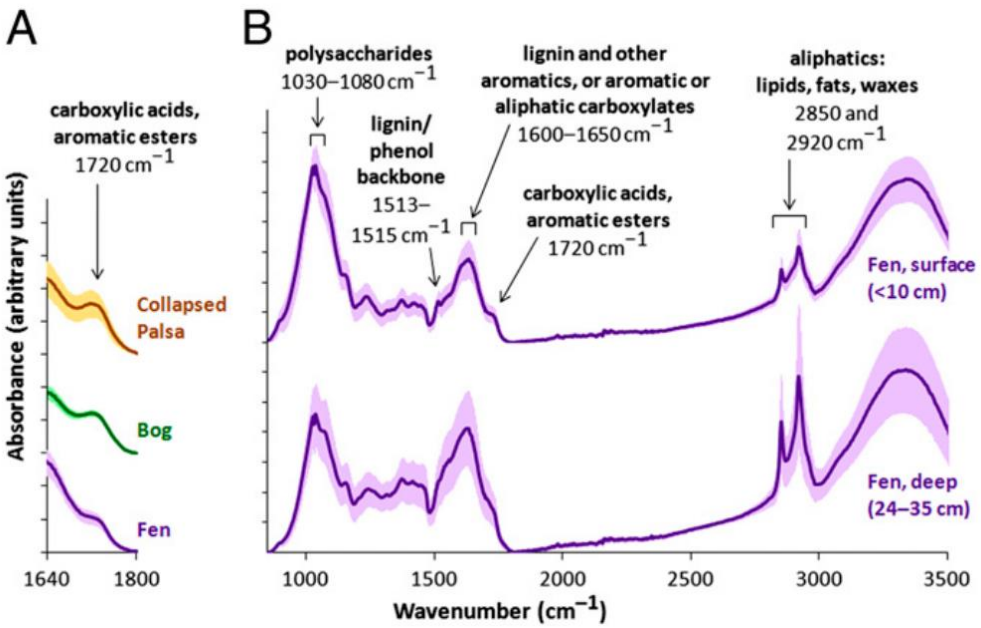


Fig. S8 Average peat FTIR absorption spectra, with between-site standard errors as shaded areas. Spectra are stacked (i.e., absorbance = 0 at each apparent baseline) and sized to the same vertical scale. (A) Organic acids in incubated peat samples (preincubation) decrease with thaw from collapsed palsas (n = 2) to bogs (n = 3) to fens (including both sedge only and Sphagnum + sedge; n = 4). (B)In fens, humification leads to spectral changes between surface (n = 3) and deep (n = 3) peat. Reproduced from Hodgkins *et al*.[30]

This was consistent with a decrease in hydroxyl rich carbon relative to aliphatic and aromatic in the FTIR spectra suggesting degradation of carbohydrates. Malik *et al*.[31] used FTICR MS and FTIR to characterize the organic matter released from soils after chloroform fumigation extraction, which is commonly used to estimate soil microbial biomass. FTIR analysis confirmed the enrichment of aliphatic CH bonds in the fumigated DOM, which together suggested that fumigation largely increased the recovery of membrane lipids. EEMS detects aromatic fluorescent components of DOM, and these signals are often attributed to functional classes with similar

signatures such as lignin phenols, quinones, proteinaceous residues incorporating tryptophan or tyrosine, and refractory humic and fulvic acids. FTICRMS has been used to provide information on the potential molecular identities of these fluorescent agents by correlating the fluorescent signals to the components observed in FTICR MS spectra.[32] EEMS has also been used to differentiate between terrestrial, microbial protein-like, and marine organic matter inputs.[29,33] Together with absorption spectroscopy, they represent inexpensive and fast tools that allow for high throughput. Thus, high-resolution sampling in both space and time is possible. Both techniques can be done in situ and may be used by remote sensing allowing high-frequency, real-time data collection for understanding ecosystem properties.[34,35] While these measurements are highly suitable for looking at a broad range of extracted material[36] and oxidation states that are more or less bioavailable,[37] these compounds represent only a small fraction of the total organic matter pool.[38–40] Thus, for studies that involve microbial communities and microbial pathways, we recommend complimenting optical measurements with either of the above mentioned spectrometric and spectroscopic techniques that allow us to look for individual compounds.

SECTION 3. MASS DIFFERENCE-BASED NETWORK ANALYSIS

Mass difference-based network analysis is similar to KMD analysis in that it links detected ions by characteristic mass differences, such as the mass of the methylene subunit (CH_2 = 14.01565 amu) or the mass change from nitrosylation (HNO = 31.005814 amu),[167,189] but multiple subunits can be used simultaneously. When plotted as one or more interconnected networks, the resulting plot can seem exceedingly complex because the number of "nodes" is equal to the number of detected ions, and because each node can be connected to other nodes through one or several mass-difference-based "edges" (i.e., subunits) (Fig. S9). However, subgraphs can be extracted using a

subset of the mass differences,[167] and nodes or edges can be colored by sample type, element counts, or other characteristics,[190] to focus on specific features. Mass difference-based network analyses have already been used for diverse purposes including formula assignment,[167] untargeted metabolite profiling,[191,192] and reaction pathway identification.[193] To be clear, the assignment of an elemental composition is insufficient to identify a metabolite, and the occurrence of a specific mass difference can be coincidental in a complex sample.

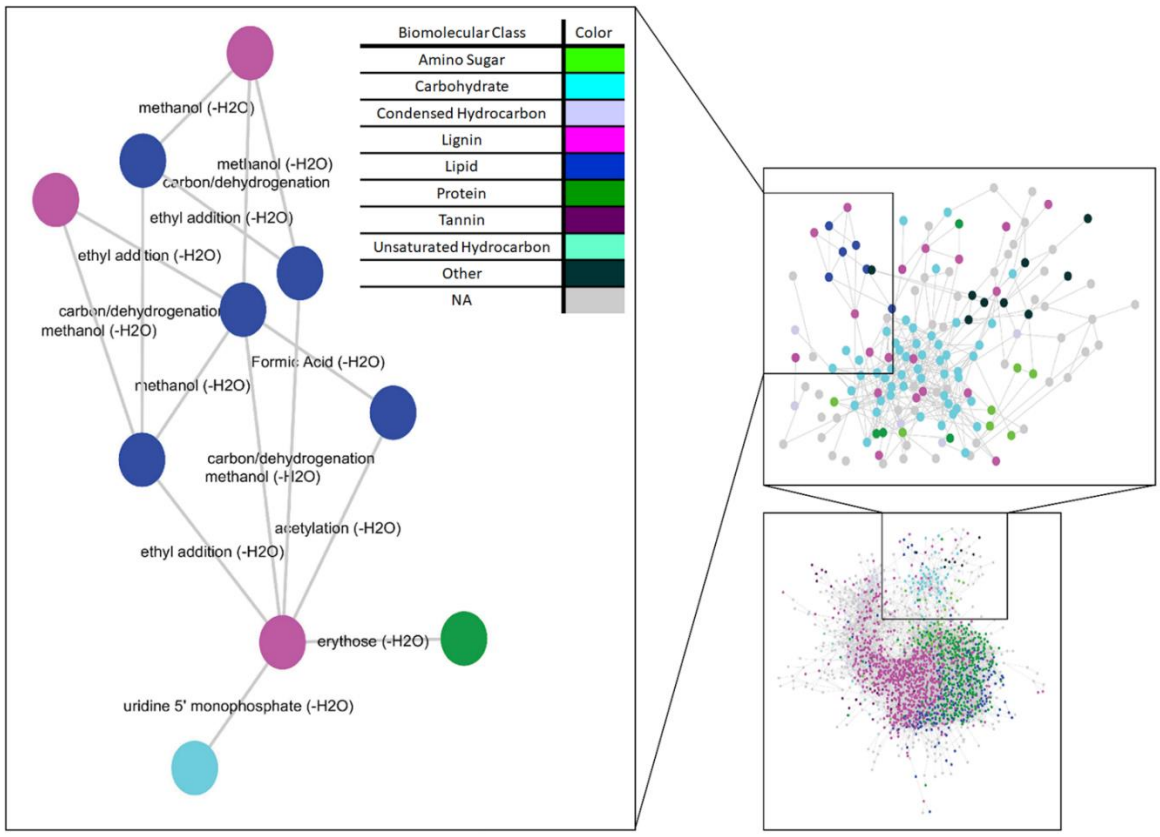


Fig. S9 Abiotic reactions at low pH occurring in cellulose-like compounds produced by Sphagnum (peat moss), plotted using mass difference-based network analysis to study how plant leachates degrade in soils. Each assigned FTICR MS peak (circle) is colored by a biomolecular class approximation using van Krevelen plot regions and related through different reactions (lines). Produced with data obtained from Fudyma *et al*.[194]

REFERENCES

(1) Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; et al. PubChem 2019 Update: Improved Access to Chemical Data. *Nucleic Acids Res.* **2019**, *47* (D1), D1102–D1109. https://doi.org/10.1093/nar/gky1033.

(2) MOSES. Altamira LLC, USA and Molecular Networks GmbH, Germany 2020.

(3) Moncelli, M. R.; Becucci, L.; Guidelli, R. The Intrinsic PKa Values for Phosphatidylcholine, Phosphatidylethanolamine, and Phosphatidylserine in Monolayers Deposited on Mercury Electrodes. *Biophys. J.* **1994**, *66* (6), 1969–1980. https://doi.org/10.1016/S0006-3495(94)80990-7.

(4) Fox, P. M.; Nico, P. S.; Tfaily, M. M.; Heckman, K.; Davis, J. A. Characterization of Natural Organic Matter in Low-Carbon Sediments: Extraction and Analytical Approaches. *Org. Geochem.* **2017**, *114* (Supplement C), 12–22. https://doi.org/10.1016/j.orggeochem.2017.08.009.

(5) Tfaily, M. M.; Chu, R. K.; Toyoda, J.; Tolić, N.; Robinson, E. W.; Paša-Tolić, L.; Hess, N. J. Sequential Extraction Protocol for Organic Matter from Soils and Sediments Using High Resolution Mass Spectrometry. *Anal. Chim. Acta* **2017**, *972*, 54–61. https://doi.org/10.1016/j.aca.2017.03.031.

(6) Guigue, J.; Mathieu, O.; Lévêque, J.; Mounier, S.; Laffont, R.; Maron, P. A.; Navarro, N.; Chateau, C.; Amiotte-Suchet, P.; Lucas, Y. A Comparison of Extraction Procedures for Water-Extractable Organic Matter in Soils. *Eur. J. Soil Sci.* **2014**, *65* (4), 520–530. https://doi.org/10.1111/ejss.12156.

(7) Stevenson, F. J. *Humus Chemistry: Genesis, Composition, Reactions*; John Wiley & Sons, 1994.

(8) Lopez-Sangil, L.; Rovira, P. Sequential Chemical Extractions of the Mineral-Associated Soil Organic Matter: An Integrated Approach for the Fractionation of Organo-Mineral Complexes. *Soil Biol. Biochem.* **2013**, *62*, 57–67. https://doi.org/10.1016/j.soilbio.2013.03.004.

(9) Jones, D. L.; Willett, V. B. Experimental Evaluation of Methods to Quantify Dissolved Organic Nitrogen (DON) and Dissolved Organic Carbon (DOC) in Soil. *Soil Biol. Biochem.* **2006**, *38* (5), 991–999. https://doi.org/10.1016/j.soilbio.2005.08.012.

(10) Liang, C.; Schimel, J. P.; Jastrow, J. D. The Importance of Anabolism in Microbial Control over Soil Carbon Storage. **2017**. https://doi.org/10.1038/nmicrobiol.2017.105.

(11) Wagai, R.; Mayer, L. M. Sorptive Stabilization of Organic Matter in Soils by Hydrous Iron Oxides. *Geochim. Cosmochim. Acta* **2007**, *71* (1), 25–35. https://doi.org/10.1016/j.gca.2006.08.047.

(12) Chao, T. T.; Zhou, L. Extraction Techniques for Selective Dissolution of Amorphous Iron Oxides from Soils and Sediments. *Soil Sci. Soc. Am. J.* **1983**, *47* (2), 225–232. https://doi.org/10.2136/sssaj1983.03615995004700020010x.

(13) Coward, E. K.; Thompson, A. T.; Plante, A. F. Iron-Mediated Mineralogical Control of Organic Matter Accumulation in Tropical Soils. *Geoderma* **2017**, *306*, 206–216. https://doi.org/10.1016/j.geoderma.2017.07.026.

(14) Lv, J.; Huang, Z.; Christie, P.; Zhang, S. Reducing Reagents Induce Molecular Artifacts in the Extraction of Soil Organic Matter. *ACS Earth Sp. Chem.* **2020**, acsearthspacechem.0c00194. https://doi.org/10.1021/acsearthspacechem.0c00194.

(15) Tfaily, M. M.; Chu, R. K.; Tolić, N.; Roscioli, K. M.; Anderton, C. R.; Paša-Tolić, L.;

Robinson, E. W.; Hess, N. J. Advanced Solvent Based Methods for Molecular Characterization of Soil Organic Matter by High-Resolution Mass Spectrometry. *Anal. Chem.* **2015**, *87* (10), 5206–5215. https://doi.org/10.1021/acs.analchem.5b00116.

(16) Flerus, R.; Koch, B. P.; Schmitt-Kopplin, P.; Witt, M.; Kattner, G. Molecular Level Investigation of Reactions between Dissolved Organic Matter and Extraction Solvents Using FT-ICR MS. *Mar. Chem.* **2011**, *124* (1–4), 100–107. https://doi.org/10.1016/j.marchem.2010.12.006.

(17) Schnitzer, M.; Schuppli, P. Method for the Sequential Extraction of Organic Matter from Soils and Soil Fractions. *Soil Sci. Soc. Am. J.* **1989**, *53* (5), 1418–1424. https://doi.org/10.2136/sssaj1989.03615995005300050019x.

(18) Young, R.; Avneri-Katz, S.; McKenna, A.; Chen, H.; Bahureksa, W.; Polubesova, T.; Chefetz, B.; Borch, T. Composition-Dependent Sorptive Fractionation of Anthropogenic Dissolved Organic Matter by Fe(III)-Montmorillonite. *Soil Syst.* **2018**, *2* (1), 14. https://doi.org/10.3390/soilsystems2010014.

(19) Koch, B. P.; Dittmar, T.; Witt, M.; Kattner, G. Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter. *Anal. Chem.* **2007**, *79* (4), 1758–1763. https://doi.org/10.1021/ac061949s.

(20) Herzsprung, P.; Hertkorn, N.; von Tümpling, W.; Harir, M.; Friese, K.; Schmitt-Kopplin, P. 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.* **2014**, *406* (30), 7977–7987. https://doi.org/10.1007/s00216-014-8249-y.

(21) NORMAN Network (2021). NORMAN Suspect List Exchange (NORMAN-SLE) https://www.norman-network.com/nds/SLE/.

(22) Hockaday, W. C.; Purcell, J. M.; Marshall, A. G.; Baldock, J. A.; Hatcher, P. G. Electrospray and Photoionization Mass Spectrometry for the Characterization of Organic Matter in Natural Waters: A Qualitative Assessment. *Limnol. Oceanogr. Methods* **2009**, *7* (1), 81–95. https://doi.org/10.4319/lom.2009.7.81.

(23) Hertkorn, N.; Benner, R.; Frommberger, M.; Schmitt-Kopplin, P.; Witt, M.; Kaiser, K.; Kettrup, A.; Hedges, J. I. Characterization of a Major Refractory Component of Marine Dissolved Organic Matter. *Geochim. Cosmochim. Acta* **2006**, *70* (12), 2990–3010. https://doi.org/10.1016/j.gca.2006.03.021.

(24) Hertkorn, N.; Harir, M.; Koch, B. P.; Michalke, B.; Schmitt-Kopplin, P. High-Field NMR Spectroscopy and FTICR Mass Spectrometry: Powerful Discovery Tools for the Molecular Level Characterization of Marine Dissolved Organic Matter. *Biogeosciences* **2013**, *10* (3), 1583–1624. https://doi.org/10.5194/bg-10-1583-2013.

(25) Abdulla, H. A. N.; Hatcher, P. G. Dynamics of Dissolved Organic Matter: A View from Two Dimensional Correlation Spectroscopy Techniques. *J. Mol. Struct.* **2014**, *1069* (1), 313–317. https://doi.org/10.1016/j.molstruc.2014.03.038.

(26) DiDonato, N.; Chen, H.; Waggoner, D.; Hatcher, P. G. Potential Origin and Formation for Molecular Components of Humic Acids in Soils. *Geochim. Cosmochim. Acta* **2016**, *178*, 210–222. https://doi.org/10.1016/j.gca.2016.01.013.

(27) Podgorski, D. C.; Hamdan, R.; McKenna, A. M.; Nyadong, L.; Rodgers, R. P.; Marshall, A. G.; Cooper, W. T. Characterization of Pyrogenic Black Carbon by Desorption Atmospheric Pressure Photoionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2011**, *84* (3), 1281–1287. https://doi.org/10.1021/ac202166x.

(28) Weishaar, J. L.; Aiken, G. R.; Bergamaschi, B. A.; Fram, M. S.; Fujii, R.; Mopper, K. Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. *Environ. Sci. Technol.* **2003**, *37* (20), 4702–4708. https://doi.org/10.1021/es030360x.
(29) McKnight, D. M.; Boyer, E. W.; Westerhoff, P. K.; Doran, P. T.; Kulbe, T.; Andersen, D. T. Spectrofluorometric Characterization of Dissolved Organic Matter for Indication of Precursor Organic Material and Aromaticity. *Limnol. Oceanogr.* **2001**, *46* (1), 38–48. https://doi.org/10.4319/lo.2001.46.1.0038.
(30) Hodgkins, S. B.; Tfaily, M. M.; McCalley, C. K.; Logan, T. A.; Crill, P. M.; Saleska, S. R.; Rich, V. I.; Chanton, J. P. Changes in Peat Chemistry Associated with Permafrost Thaw Increase Greenhouse Gas Production. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (16), 5819–5824. https://doi.org/10.1073/pnas.1314641111.
(31) Malik, A. A.; Roth, V.-N.; Ebert, M. H.; Tremblay, L.; Dittmar, T.; Gleixner, G. Linking Molecular Size, Composition and Carbon Turnover of Extractable Soil Microbial Compounds. **2016**. https://doi.org/10.1016/j.soilbio.2016.05.019.
(32) Herzsprung, P.; Osterloh, K.; Von Tümpling, W.; Harir, M.; Hertkorn, N.; Schmitt-Kopplin, P.; Meissner, R.; Bernsdorf, S.; Friese, K. Differences in DOM of Rewetted and Natural Peatlands-Results from High-Field FT-ICR-MS and Bulk Optical Parameters. *Sci. Total Environ.* **2017**, *586*, 770–781. https://doi.org/10.1016/j.scitotenv.2017.02.054.
(33) Coble, P. G. Characterization of Marine and Terrestrial DOM in Seawater Using Excitation-Emission Matrix Spectroscopy. *Mar. Chem.* **1996**, *51* (4), 325–346. https://doi.org/10.1016/0304-4203(95)00062-3.
(34) Aiken, G. Fluorescence and Dissolved Organic Matter. In *Aquatic Organic Matter Fluorescence*; Cambridge University Press, 2014; pp 35–74. https://doi.org/10.1017/cbo9781139045452.005.
(35) Ruhala, S. S.; Zarnetske, J. P. Using In-Situ Optical Sensors to Study Dissolved Organic Carbon Dynamics of Streams and Watersheds: A Review. *Sci. Total Environ.* **2017**, *575*, 713–723. https://doi.org/10.1016/j.scitotenv.2016.09.113.
(36) Kelleher, B. P.; Simpson, A. J. Humic Substances in Soils: Are They Really Chemically Distinct? *Environ. Sci. Technol.* **2006**, *40* (15), 4605–4611. https://doi.org/10.1021/es0608085.
(37) Lehmann, J.; Kleber, M. The Contentious Nature of Soil Organic Matter. *Nature* **2015**, *528* (7580), 0–8. https://doi.org/10.1038/nature16069.
(38) Schmidt, M. W. I. I.; Torn, M. S.; Abiven, S.; Dittmar, T.; Guggenberger, G.; Janssens, I. A.; Kleber, M.; Kögel-Knabner, I.; Lehmann, J.; Manning, D. A. C. C.; et al. Persistence of Soil Organic Matter as an Ecosystem Property. *Nature* **2011**, *478* (7367), 49–56. https://doi.org/10.1038/nature10386.
(39) Rosario-Ortiz, F. L.; Korak, J. A. Oversimplification of Dissolved Organic Matter Fluorescence Analysis: Potential Pitfalls of Current Methods. *Environ. Sci. Technol.* **2017**, *51* (2), 759–761. https://doi.org/10.1021/acs.est.6b06133.
(40) Wünsch, U. J.; Acar, E.; Koch, B. P.; Murphy, K. R.; Schmitt-Kopplin, P.; Stedmon, C. A. The Molecular Fingerprint of Fluorescent Natural Organic Matter Offers Insight into Biogeochemical Sources and Diagenetic State. *Anal. Chem.* **2018**, *90* (24), 14188–14197. https://doi.org/10.1021/acs.analchem.8b02863.

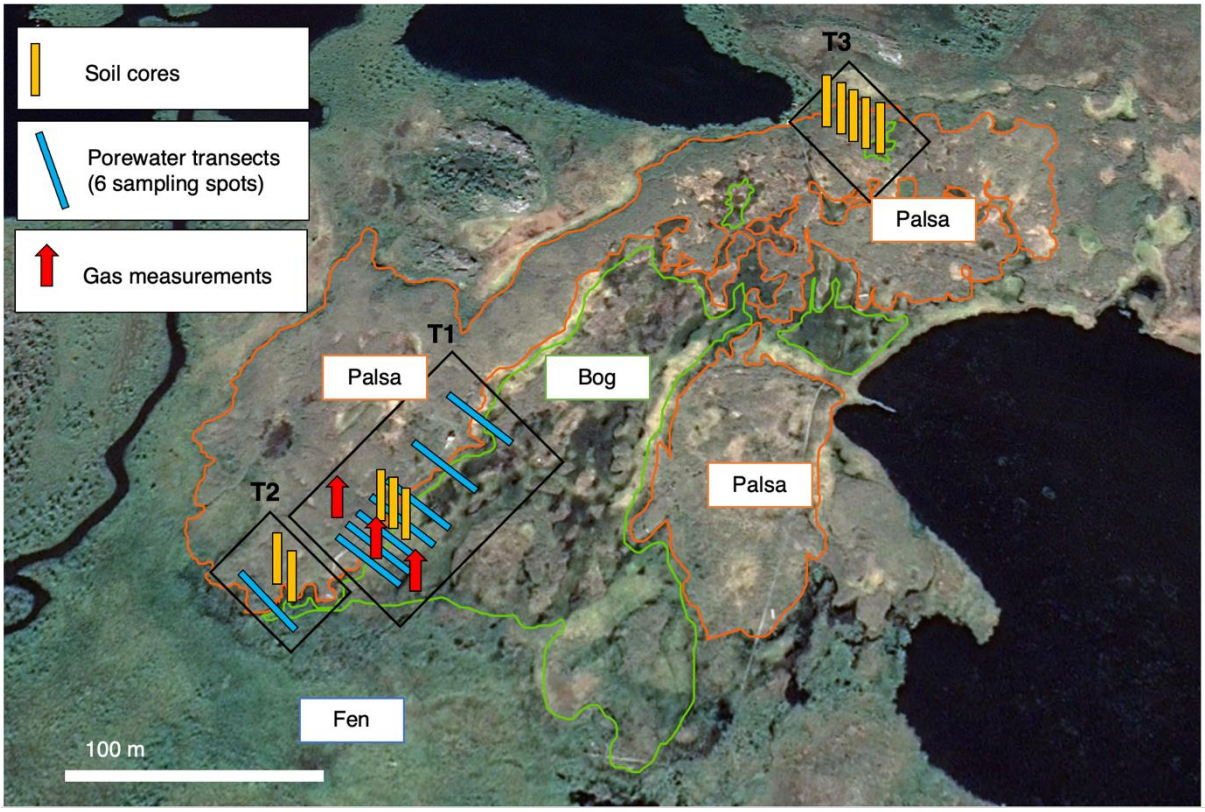


FIGURE S1. FIELD SITE AND SAMPLE COLLECTION.

Soil cores (yellow), porewater samples (blue) and gas samples (red) were taken in three transects (T1, T2 and T3) along palsa hillslopes into bog at Stordalen mire, Abisko (Sweden). Background picture was taken by GoogleEarth in 2019.

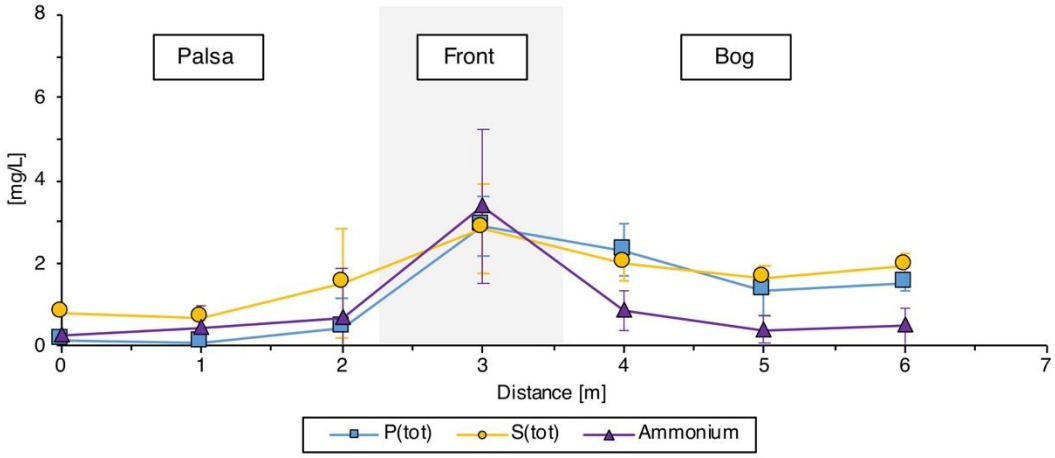


FIGURE S5. AQUEOUS TOTAL PHOSPHOROUS (P(TOT)), AQUEOUS TOTAL SULFUR (S(TOT)) AND AMMONIUM CONCENTRATIONS IN THE POREWATER ALONG THE PALSA HILLSLOPE (30 CM DEPTH, TRANSITION ZONE).

Aqueous concentrations are reported in mg/L from palsa (0-2.7 m) to bog (2.7-7 m). Reported values represent the average of six sampling spots for eight palsa hillslopes (0-2.3 m) to collapsing front (2.3-3.6 m) to bog (3.6-7 m), sampled in June/July (see also SI, Figures S1 and Figure S4). Error bars represent the standard deviation of eight palsa hillslopes (0-2.3 m) to collapsing front (2.3-3.6 m) to bog (3.6-7 m), sampled in June/July (see also SI, Figure S1).

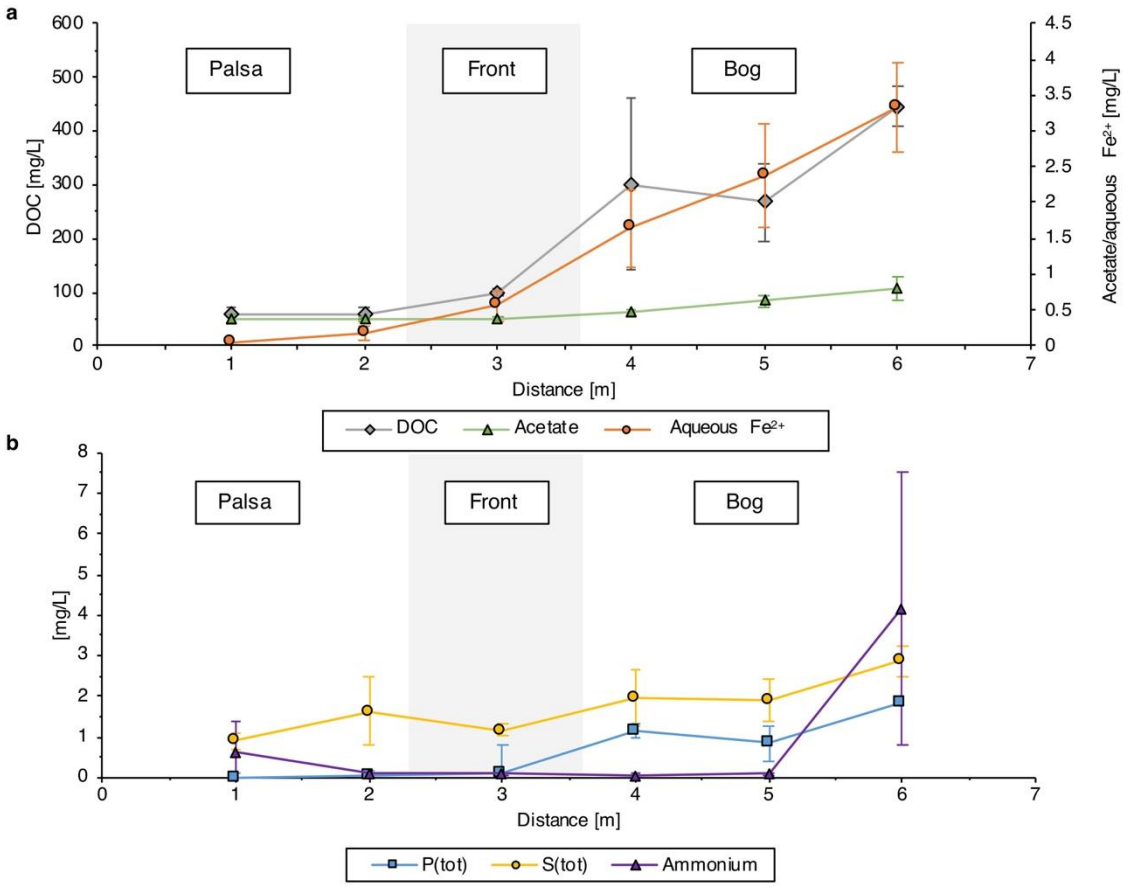


FIGURE S6. POREWATER ANALYSIS ALONG THE PALSA HILLSLOPE (60 CM, MINERAL HORIZON).

a, Dissolved organic carbon (DOC), acetate and aqueous Fe^{2+} in mg/L and b, aqueous total phosphorous (P(tot)), aqueous total sulfur (S(tot)) and ammonium concentrations along the collapsing palsa hillslope into bog. Reported values represent the average of six sampling spots for eight palsa hillslopes (0-2.3 m) to collapsing front (2.3-3.6 m) to bog (3.6-7 m), sampled in June/July (see also SI, Figure S1). Error bars represent the standard deviation of eight palsa hillslopes (0-2.3 m) to collapsing front (2.3-3.6 m) to bog (3.6-7 m), sampled in June/July (see also SI, Figure S1).

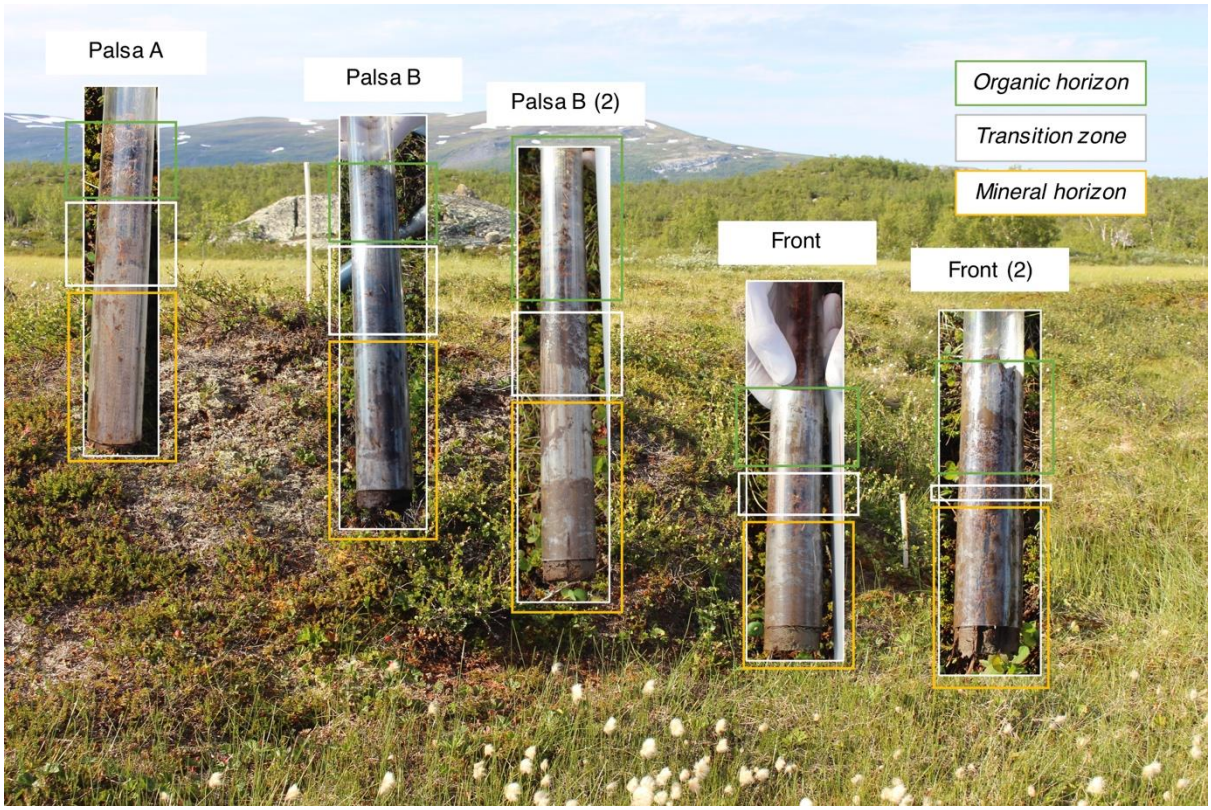


FIGURE S7. CORING ALONG PALSA HILLSLOPES.

Ten soil cores were taken along different palsa hillslopes to capture spatial heterogeneity of iron-carbon associations along the peatland mire. In transect 3 (shown here, see also Figure S1), five cores were taken along a palsa hillslope towards the collapsing front into bog (68°21'27.33"N, 19° 3'1.91"E), immediately frozen in liquid nitrogen and stored at -80°C until analysis of iron-carbon associations and of present and active microbial community.

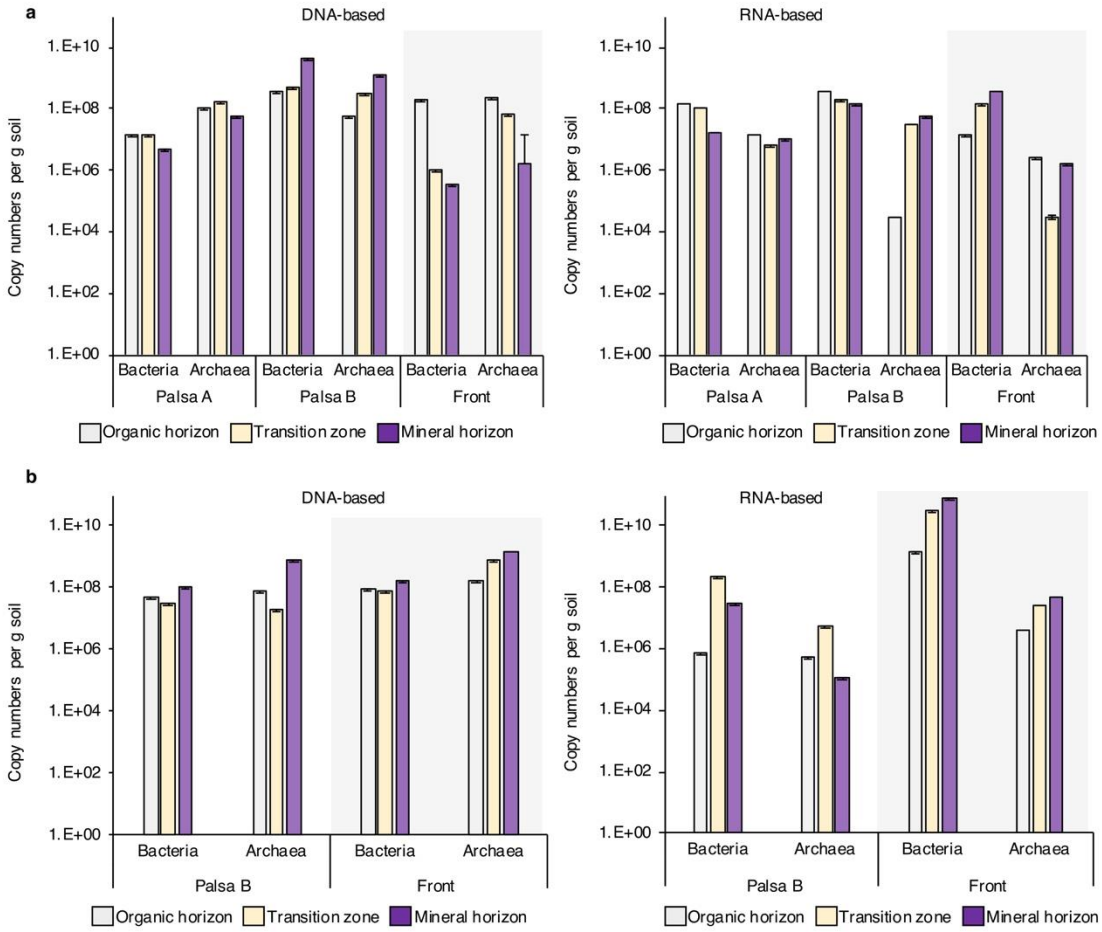


FIGURE S8. ABUNDANCE OF BACTERIA AND ARCHAEA

(copy numbers based on qPCR analysis specific for 16S rRNA genes; DNA-based on the left, RNA-based on the right). a, and b, show replicate analysis for Palsa A, Palsa B and Front (a), and Palsa A and Front (b).

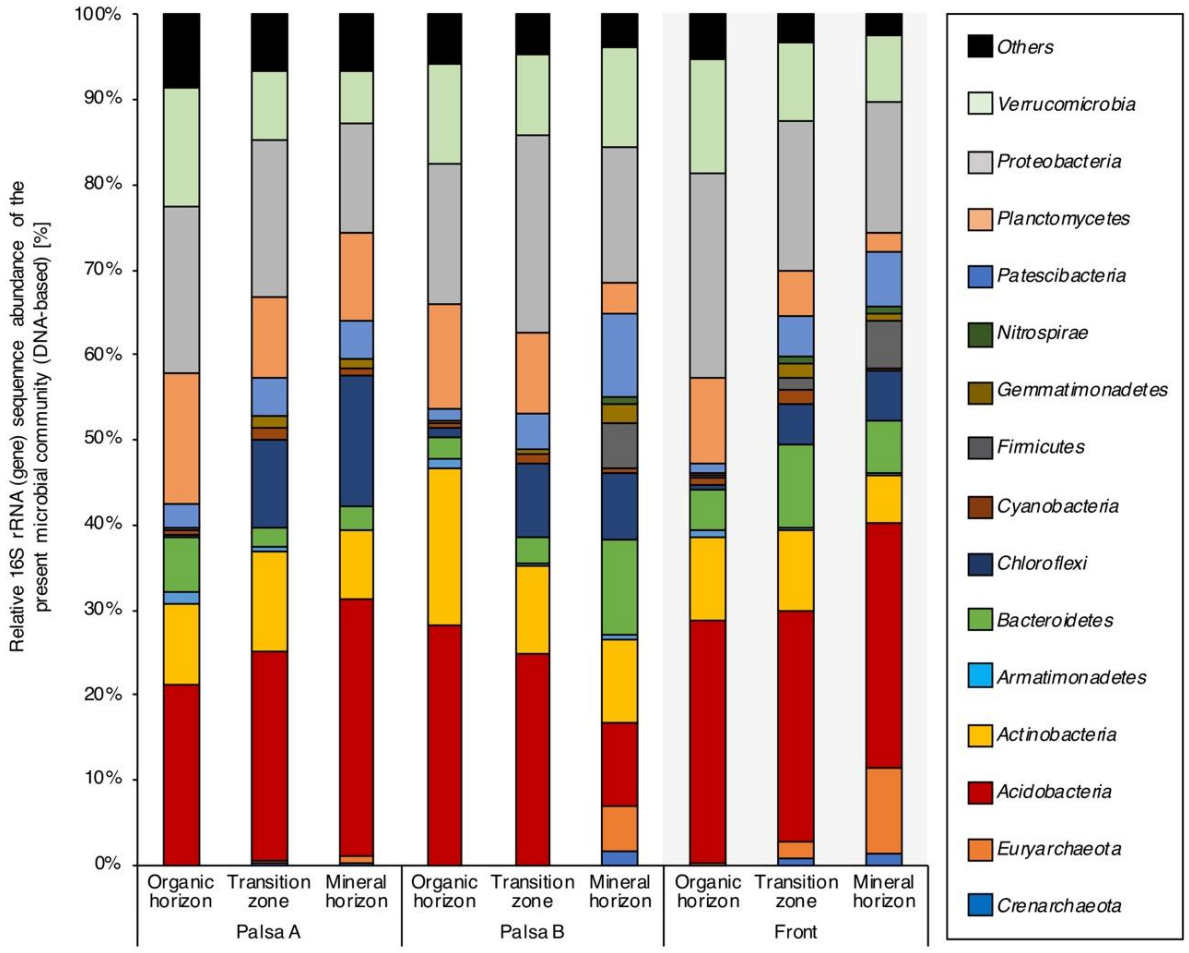


FIGURE S9 TAXONOMIC IDENTIFICATION OF THE MICROBIAL COMMUNITIES ALONG THE PALSA HILLSLOPE

based on 16S rRNA gene amplicon analysis (DNA-based). Data were averaged among triplicate analysis of each soil horizon (organic horizon, transition zone, mineral horizon).

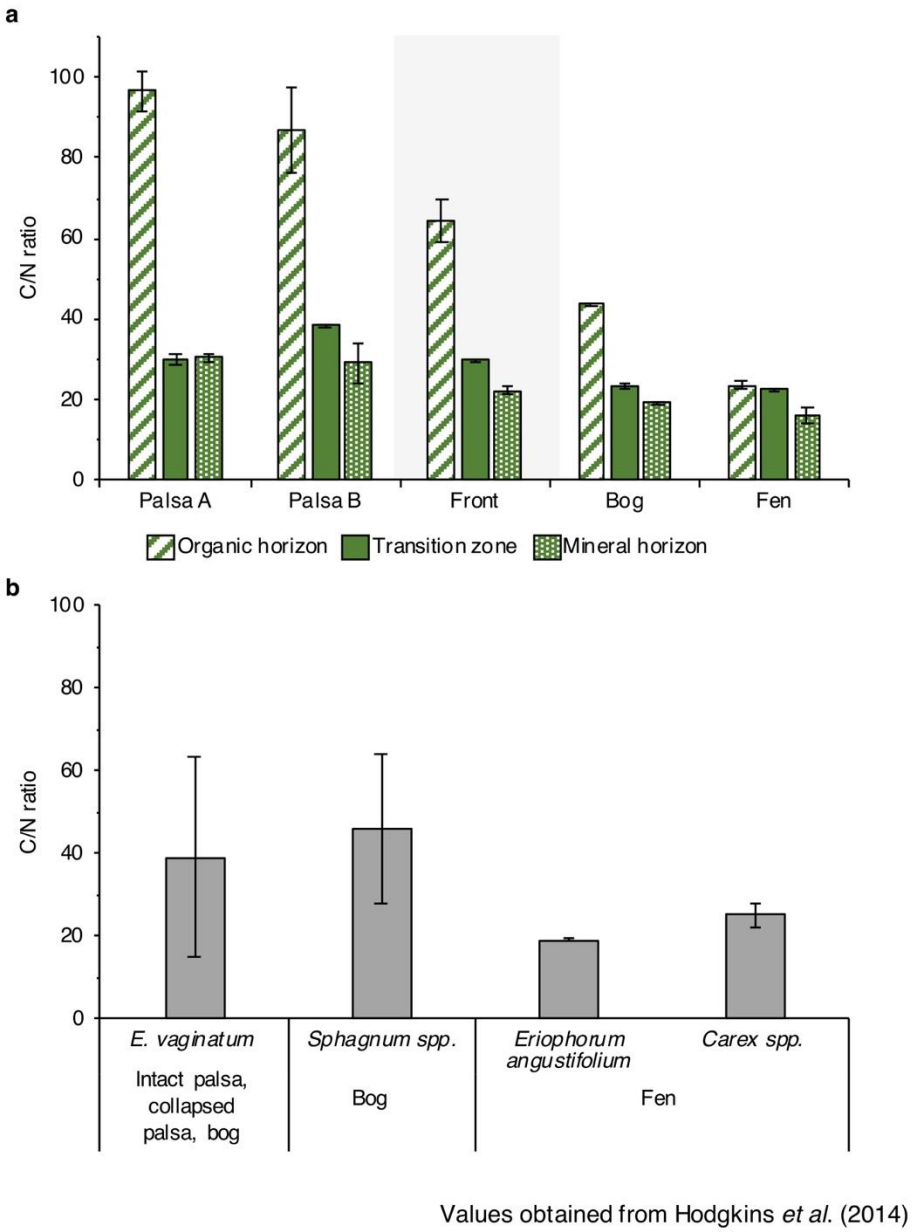


FIGURE S10. C/N WEIGHT RATIOS OF A, SOIL SAMPLES OF DISTINGUISHED SOIL HORIZON ALONG PALSA HILLSLOPE AND THAW GRADIENT, IN COMPARISON TO B, LIVING PLANT SAMPLES OF DOMINANT STORDALEN SPECIES (MODIFIED AND ADAPTED FROM HODGKINS ET AL.2). Reported values represent the average of triplicate analysis of cores Palsa A, Palsa B, Front (transect 1, Figure S1) and Bog C and Fen E, which were previously puplished[3]. Error bars represent the standard deviation of triplicate analysis of cores Palsa A, Palsa B, Front (transect 1, Figure S1), one bog (Bog C) and one fen core (Fen E) (see also Patzner *et al.*[3]).

TABLE S1. OVERVIEW OF IRON(II)-OXIDIZING MICROORGANISMS THAT WERE CROSS-CHECKED IN THE 16S RRNA AMPLICON GENE SEQUENCING RESULTS (DNA- AND RNA-BASED) IN THIS STUDY (adapted from Otte *et al.*[4] and Weber *et al.*[5], see also Dinh *et al.*[6] and Berg *et al.*[7]).

Iron(II)-oxidizing microorganisms (species or strains)
Acidianus brierleyi
Acidiplasma aeolicum; A. cupricumulans
Acidimicrobium ferrooxidans sp. strain DSM 10331
Acidithiobacillus ferrooxidans sp. strain ATCC 23270
Acidovorax sp. strains 2AN, BoFeN1, BrG1; *A. delafieldii*; *A. ebreus* strain TPSY
Alicyclobacillus disulidooxidans; A. tolerans
Aquabacterium sp. strains BrG2, HidR2
Azoarcus sp. strain ToN1
Azospira sp. strain TR1; *A. oryzae*
Bradyrhizobium japonicum sp. strains 22, in8p8, wssI4
Candidatus Brocadia sinica
Candidatus Scalindua sp.
Chlorobium luteolum DSM273; *C. ferrooxidans* sp. strain KoFox
Chromobacterium violaceum sp. strain 2002
Citrobacter freundii sp. strain PXL1
Comamonas sp. strain MPI12
Crenothrix sp. #
Cupriavidus necator sp. strains A5-1, ss1-6-6
Dechlorobacter hydrogenophilus sp. strain LT-1
Dechloromonas sp.; D. agitata sp. strains CKB, is5; *D. aromatica* sp. strains RCB, UWNR4; *D. suillum* sp. strain PS
Dechlorospirillum sp. strain M1
Denitromonas indolicum
Desulfitobacterium frappieri sp. strain G2
Desulfobacterium corrodens (Dinh *et al.*, 2004)
Escherichia coli sp. strain E4
Ferrimicrobium acidiphilum
Ferriphaselus amnicola (Siderooxydans)
Ferrithrix thermotolerans
Ferritrophicum radicicola
Ferroglobus placidus sp. strain AEDII12DO
Ferroplasma acidarmanus sp. strain fer1
Ferrovibrio denitrificans
Gallionella capsiferriformans; G. ferruginea; G. strain ES-2
Geobacter metallireducens sp. strain GS-15
Geothrix spp.
Hoeflea marina; *H. siderophila* sp. strain Hf1
Hyphomicrobium sp.

(Table continues on next page)

Hyphomonas sp.
Klebsiella-like sp. strain FW33AN
Leptospirillum ferrooxidans; L. ferriphilum
Leptothrix cholodnii; L. discophora
Magnetococcus sp. (Berg *et al*., 2016)
Magnetospirillum bellicus sp. strain VDY
Marinobacter aquaeolei sp. strain VT8
Mariprofundus ferrooxidans sp. strains PV-1, RL-1, JV-1, GSBS
Metallosphaera sedula sp. strain J1
Nocardioides sp. strain In31
Paracoccus sp.*; P. denitrificans; P. ferrooxydans* sp. strain BDN-1*; P. pantotrophus; P. versutus*
Parvibaculum sp. strain MBN-A2
Pedomicrobium spp.
Propionivibrio militaris sp. strain MP
Pseudogulbenkiania ferrooxidans sp. strain 2002
Pseudomonas sp. strains LP-1, SZF15; *P. stutzeri*
Ralstonia solanacearum sp. strain in4ss52
Rhodanobacter sp. strain MPN-A3
Rhodobacter sp. strain SW2*; R. ferrooxydans*
Rhodomicrobium vannielii
Rhodopseudomonas palustris strain TIE-1
Rhodovulum sp.*; R. iodosum; R. robiginosum*
Rubrivivax group sp. strains BrG4, BrG5
Siderocapsa sp.
Sideroxydans paludicola; S. lithotrophicus sp. strain ES-1
Sphaerotilus natans sp. strain DSM 6575
Sulfobacillus spp.
Sulfolobus acidocaldarius
Thauera aromatica sp. strain AR-1
Thermomonas sp. strain BrG3
Thiobacillus denitrificans
Thiodictyon sp.
Thiomicrospira denitrificans
Zixibacteria sp. strain RBG-1

Crenothrix is most often associated with methanotrophy but there are also hints for Fe(II) oxidation.

TABLE S2. OVERVIEW OF IRON(III)-REDUCING MICROORGANISMS THAT WERE CROSS-CHECKED IN THE 16S RRNA AMPLICON GENE SEQUENCING RESULTS IN THIS STUDY (adapted from Otte *et al.*[4] and Weber *et al.*[5], see also Berg *et al.*[7], Li *et al.*[8], Holmes *et al.*[9], Finneran *et al.*[10]).

Iron(III)-reducing microorganisms (species or strains)
Acidithiobacillus ferrooxidans
Aeromonas hydrophila
Albidoferax ferrireducens
Alteromonas hydrophila
Anaeromyxobacter sp.
Bacillus infernus
Clostridium sp.
Deferribacter thermophilus
Desulfobacter propionicus
Desulfobacterium sp.
Desulfobulbus spp.
Desulfococcus spp.
Desulfotalea spp.
Desulfotomaculum sp. (Berg *et al*., 2016)
Desulfovibrio sp. *(*Li *et al., 2016)*
Desulfuromonas spp.
Desulfuromusa spp.
Ferribacterium limneticum
Ferrimonas balearica
Ferroglobus placidus
Geobacter spp.
Geoglobus sp.
Geothrix fermentans
Geovibrio ferrireducens
Malonomonas sp. (Holmes *et al*., 2004)
Myxococcales sp.
Pantoea agglomerans sp. strain Sp1
Pelobacter sp.
Pseudomonas sp.
Pseudorhodoferax (Berg *et al*., 2016)
Pyrobaculum sp.
Rhodobacter sp.
Rhodoferax sp. (Finneran *et al*., 2003)
Shewanella colwelliana
Sinorhodobacter sp.
Sulfurospirillum barnesii
Thermoterrabacterium ferrireducens
Thermotoga maritima
Thermus sp. strain SA01
Thiobacillus ferrooxidans

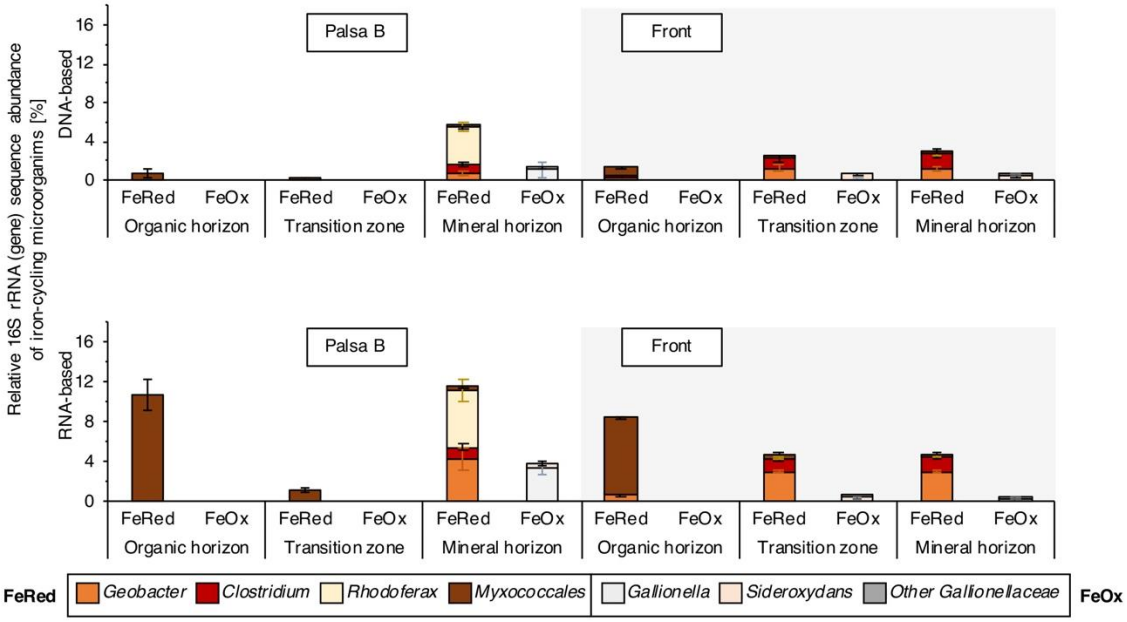
TABLE S3. OVERVIEW OF METHANOGENIC MICROORGANISMS THAT WERE CROSS-CHECKED IN THE 16S RRNA AMPLICON GENE SEQUENCING RESULTS IN THIS STUDY (see also Kim&Whitman[11] and Mondav *et al*.[12]).

Methanogenic microorganisms (genera, species or strains)
Bathyarchaeia
Methanobacterium spp.
Methanobrevibacter spp.
Methanocaldococcus spp.
Methanocella spp.
Methanococcoides spp.
Methanococcus spp.
Methanocorpusculum spp.
Methanoculleus spp.
Methanoflorens spp.*, M. stordalenmirensis (Mondav et al., 2014)*
Methanofollis spp.
Methanogenium spp.
Methanohalobium spp.
Methanohalophilus spp.
Methanoignis spp.
Methanolacinia spp.
Methanolinea spp.
Methanolobus spp.
Methanomassillicoccaceae spp.
Methanomethylovorans spp.
Methanomicrobium spp.
Methanoplanus spp.
Methanopyrus spp.
Methanoregula spp.
Methanosaeta spp.
Methanosalsum spp.
Methanosarcina spp.
Methanosphaera spp.
Methanosphaerula spp.
Methanospirillum spp.
Methanothermobacter spp.
Methanothermococcus spp.
Methanothermus spp.
Methanothrix spp.
Methanotorris spp.
Methermicoccus spp.

TABLE S4. OVERVIEW OF METHANOTROPHIC MICROORGANISMS THAT WERE CROSS-CHECKED IN THE 16S RRNA AMPLICON GENE SEQUENCING RESULTS (DNA- AND RNA-BASED) IN THIS STUDY (see also Jiang *et al.*[13] and Singelton *et al.*[14]).

Methanotrophic microorganisms (species or strains)
Acidimethylosilex spp.
Clonothrix spp., *Clonothrix fusca*
Crenotrhix spp., *Crenothrix polyspora*
Methlyosinus spp.
Methyloacida spp.
Methylobacter spp.
Methylobacterium spp.
Methylocaldum spp.
Methylocapsa spp.
Methylocella spp., *Methylocella silvestris*
Methylococcus spp.
Methylocystis spp.
Methylokorus spp.
Methylomonas spp.
Methylosphaera spp.
Methylothermus spp.
Methylovirgula spp.
Rhodoblastus spp.
Roseiarcus spp.

a Iron-cycling microorganisms along the palsa hillslope



b Methane-cycling microorganisms along the palsa hillslope

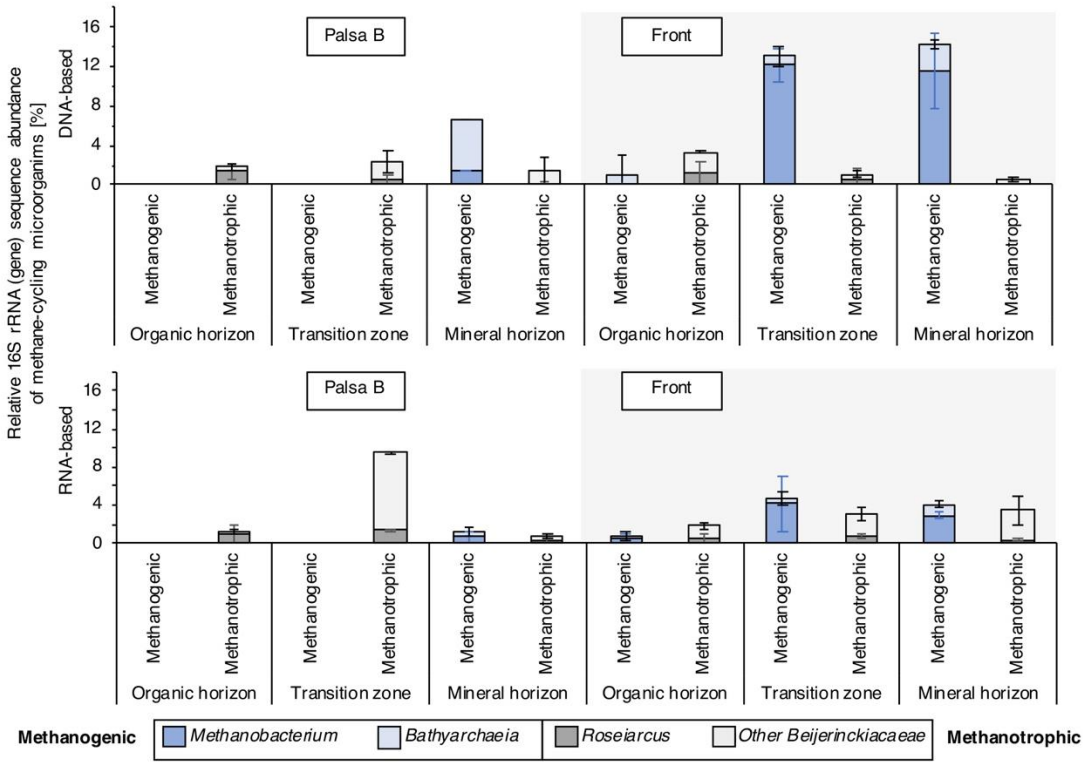


FIGURE S11. RELATIVE ABUNDANCE OF 16S RRNA (GENE) SEQUENCE OF A, IRON- AND B, METHANE CYCLING MICROORGANISMS (DNA- AND RNA-BASED) IN REPLICATE CORES (PALSA B AND FRONT) ALONG THE PALSA HILLSLOPE (TRANSECT 3, FIGURE 1).

All iron- and methane-cycling taxa below 0.1% relative 16S rRNA gene sequencing abundance are not illustrated in this figure. For absolute abundance of bacteria and archaea based on qPCR analysis, specific for 16S rRNA gene (based on DNA) and Fe- and methane-cycling microorganisms, analyzed in these study (adapted from Otte *et al*.[4] and Weber *et al.*[5]), see also Figure S8. Reported values and error bars represent the average and standard deviation of triplicate analysis of each soil horizon (organic horizon, transition zone and mineral horizon) along the palsa hillslope.

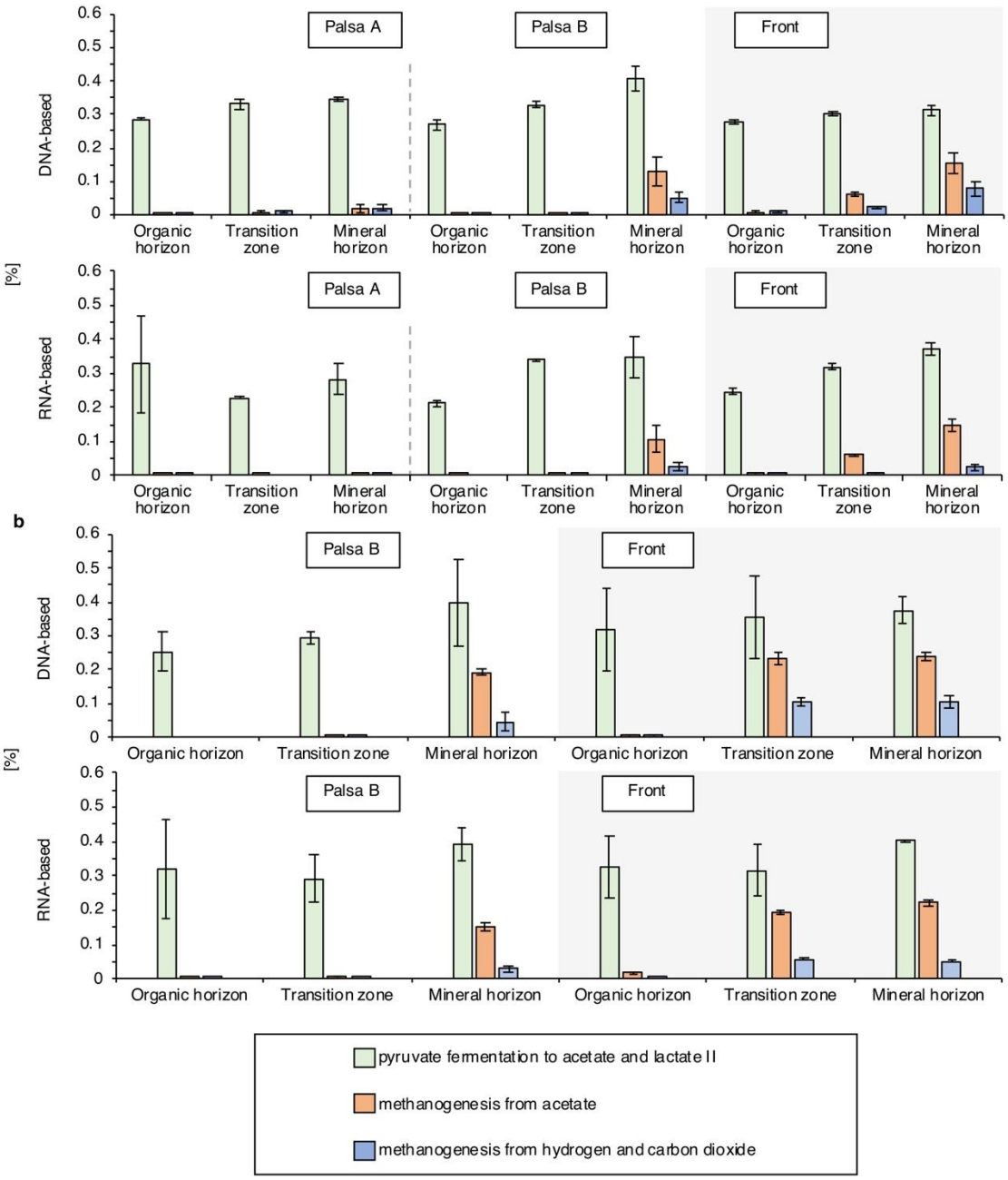


FIGURE S12. PREDICTED METABOLIC PATHWAYS OF THE PRESENT (DNA-BASED) AND POTENTIALLY ACTIVE (RNA-BASED) MICROBIAL COMMUNITIES ALONG THE PALSA HILLSLOPE IN A, CORES PRESENTED IN THE MAIN TEXT (TRANSECT 3, FIGURE S1) AND B, REPLICATE CORES.

Analyses were performed with MetaCyc Metabolic Pathway Database. Reported values and error bars represent the average and standard deviation of triplicate analysis of each soil horizon (organic horizon, transition zone and mineral horizon) along the palsa hillslope.

TABLE S5. ELEMENTAL COMPOSITION PERCENTAGE AND NUMBER OF FT-ICR-MS ASSIGNED FORMULA.

Values are derived from total assigned formula tables. Not available (N/A) values for CHNOS formula in porewater samples and salt control extractions are due to lack of reliable formula series detected in those samples. A comparison of the soil extractions using salt or dithionite solutions indicates that the extracts are not identical, but provides little to no evidence of CHOS molecular artifacts formed through reactions with dithionite, as reported by Lv *et al.*[15].

Sample Type	Sample	Extraction	CHO	CHNO	CHOS	CHNOS
			(% abundance / # formula)			
Soil extraction	Palsa A, Transition zone	Dithionite	74.19%/7486	11.21%/2887	13.76%/2018	0.84%/319
		Salt control	58.67%/4952	7.08%/1716	34.25%/6609	N/A
	Front, Transition zone	Dithionite	76.12%/7466	9.25%/2668	14.42%/2881	0.21%/98
		Salt control	76.22%/6536	7.08%/1716	13.79%/2249	N/A
Porewater	Palsa A	N/A	89.79%/9009	7.81%/3374	2.40%/991	N/A
	Front	N/A	87.41%/9072	10.80%/4172	1.79%/984	N/A

TABLE S6. ABSOLUTE VALUES OF IRON AND CARBON IN LOCATIONS PALSA A, PALSA B AND FRONT OF THE CORES REPORTED IN THE MAIN TEXT.

Errors of the dithionite/citrate extractable a, iron (reactive Fe, control corrected) and b, carbon (carbon bound to reactive iron, control corrected) represent a combined standard deviation of sodium chloride bicarbonate extractable a, iron and b, carbon, b, citrate blank and dithionite/citrate extractable a, iron and b, carbon (not control corrected).

	Palsa A				Palsa A			
	Reactive Fe [mg/g soil]	Error bars	Associated OC [mg/g soil]	Error bars	Leachable Fe [mg/g soil]	Error bars	Leachable OC [mg/g soil]	Error bars
Organic horizon	0.20	0.12	1.42	1.42	0.05	0.07	2.26	0.45
Transition zone	10.04	0.07	83.69	10.04	0.09	0.07	4.06	0.00
Mineral horizon	3.76	0.48	47.21	14.30	0.05	0.03	2.00	0.19
	Palsa B				Palsa B			
	Reactive Fe [mg/g soil]	Error bars	Associated OC [mg/g soil]	Error bars	Leachable Fe [mg/g soil]	Error bars	Leachable OC [mg/g soil]	Error bars
Organic horizon	0.67	0.07	4.80	0.52	0.00	0.00	3.05	0.71
Transition zone	4.61	0.78	76.60	16.89	0.00	0.00	2.76	0.00
Mineral horizon	4.22	0.09	25.08	4.22	0.00	0.00	0.85	0.00
	Front				Front			
	Reactive Fe [mg/g soil]	Error bars	Associated OC [mg/g soil]	Error bars	Leachable Fe [mg/g soil]	Error bars	Leachable OC [mg/g soil]	Error bars
Organic horizon	0.59	0.30	0.00	0.00	0.07	0.10	2.01	1.11
Transition zone	3.22	0.06	40.88	24.76	0.31	0.01	3.67	0.25
Mineral horizon	3.46	1.40	11.60	8.54	0.63	0.59	1.53	0.47

*Reactive Fe = dithionite extractable Fe - salt extractable Fe (leachable OC)

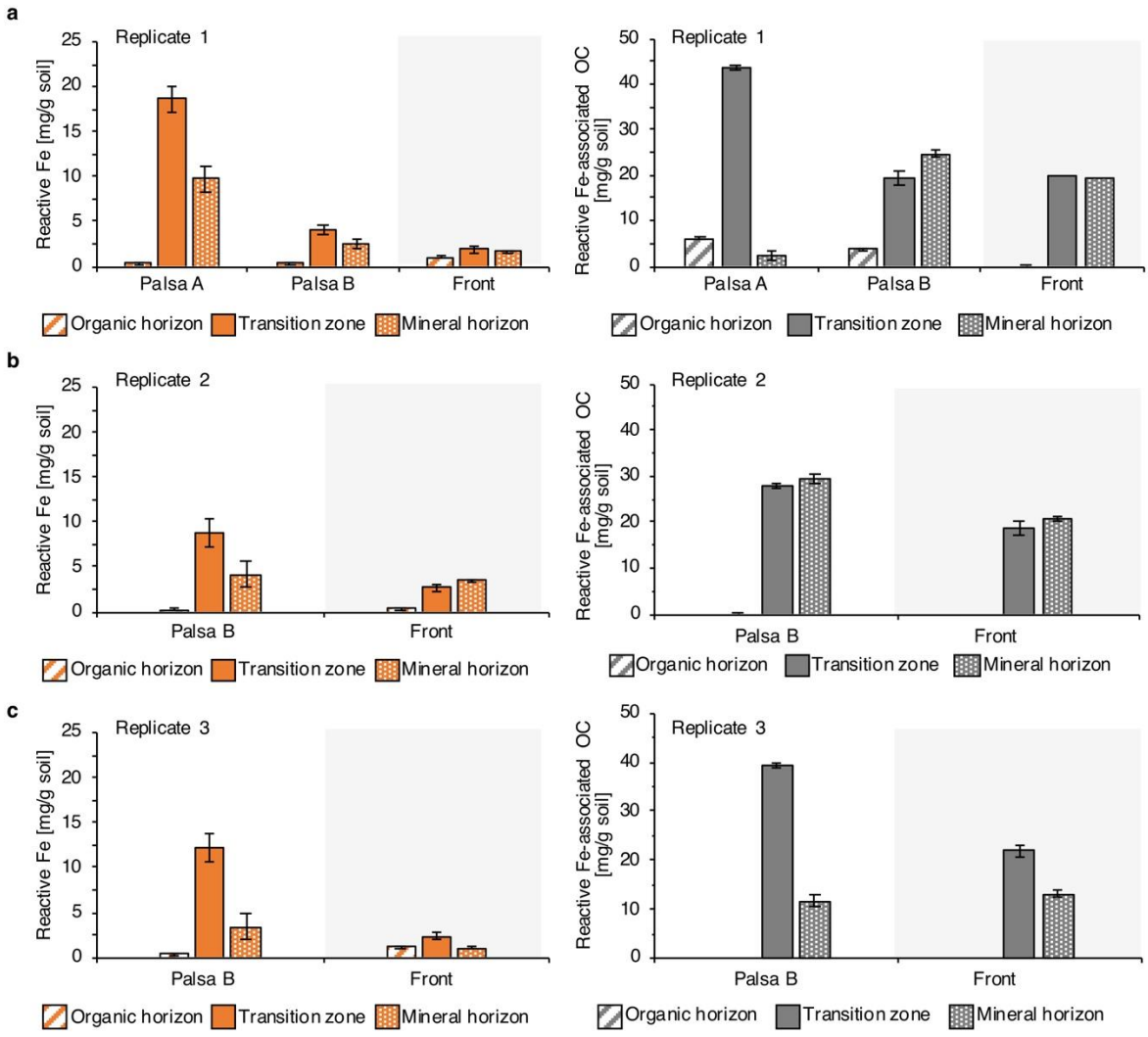


FIGURE S13. ORGANIC CARBON (OC, GREY) ASSOCIATED WITH REACTIVE IRON (FE, ORANGE) ALONG THREE REPLICATE PALSA HILLSLOPES: A, REPLICATE 1 (TRANSECT 3), B, REPLICATE 2 (TRANSECT 3) AND C, REPLICATE 3 (TRANSECT 2).

Reactive Fe and associated organic carbon along the palsa hillslope (Palsa A, Palsa B and Front) per each soil horizon (organic horizon, transition zone and mineral horizon) [mg/g] decreases towards the collapsing front. Palsa A shows the highest reactive Fe and associated organic carbon in intact permafrost soils. Along the palsa hillslope towards the collapsing front, reactive Fe and associated OC are lost in the solid phase. Reactive Fe reported values are the average of sodium dithionite citrate duplicate extractions of each soil horizon, control corrected by a sodium chloride bicarbonate extractable Fe (leachable Fe). Associated OC reported values are the average of sodium dithionite citrate extractions of each soil horizon, control corrected by the citrate background and the sodium chloride bicarbonate extractable OC (leachable OC) (see also Material and Methods). Cores were taken in July 2019 (see Figure S1). Error bars of the reactive Fe represent a combined standard deviation of sodium chloride bicarbonate extractable iron and

dithionite/citrate extractable Fe. Errors of the associated carbon represent a combined standard deviation of citrate blank, sodium chloride bicarbonate extractable OC and dithionite/citrate extractable OC.

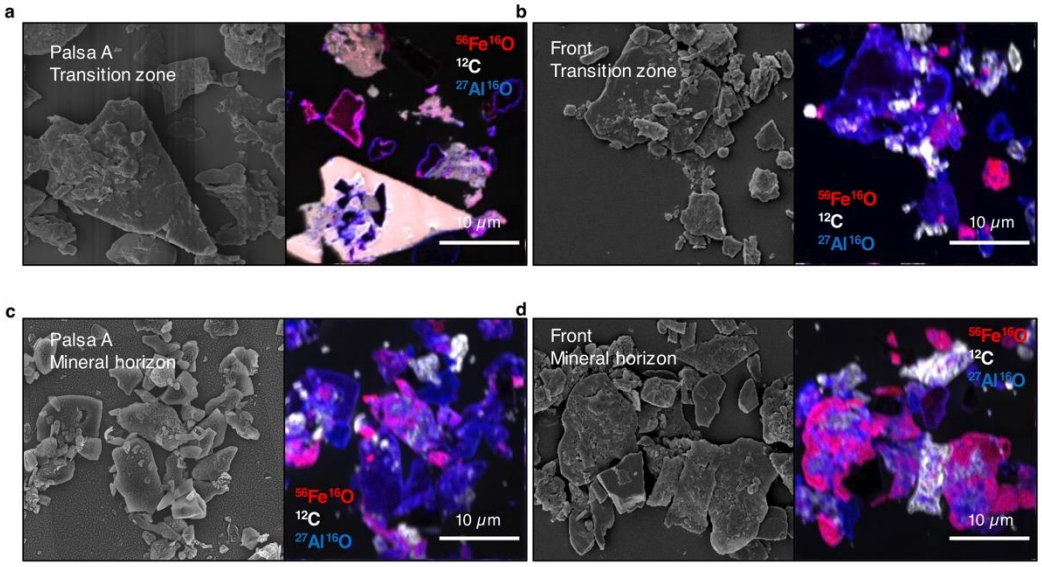


FIGURE S14. CORRELATIVE SCANNING ELECTRON MICROSCOPY AND NANOSCALE SECONDARY ION MASS SPECTROMETRY (NANOSIMS) OF FINE FRACTION OF PALSA SOIL HORIZONS ALONG THE PALSA HILLSLOPE.

Fine particle analysis of a, Palsa A transition zone; b, Front transition zone; c, Palsa A mineral horizon and d, Front mineral horizon (transect 1, Figure S1). Seven detectors were used during nanoSIMS measurements for $^{12}C^-$, $^{16}O^-$, $^{12}C^{14}N^-$, $^{31}P^-$, $^{32}S^-$, $^{27}Al^{16}O^-$ and $^{56}Fe^{16}O^-$ and $^{27}Al^{16}O$. Here, $^{12}C^-$ (white), $^{56}Fe^{16}O^-$ (red) and $^{27}Al^{16}O^-$ (blue) are overlaid in a single composite image.

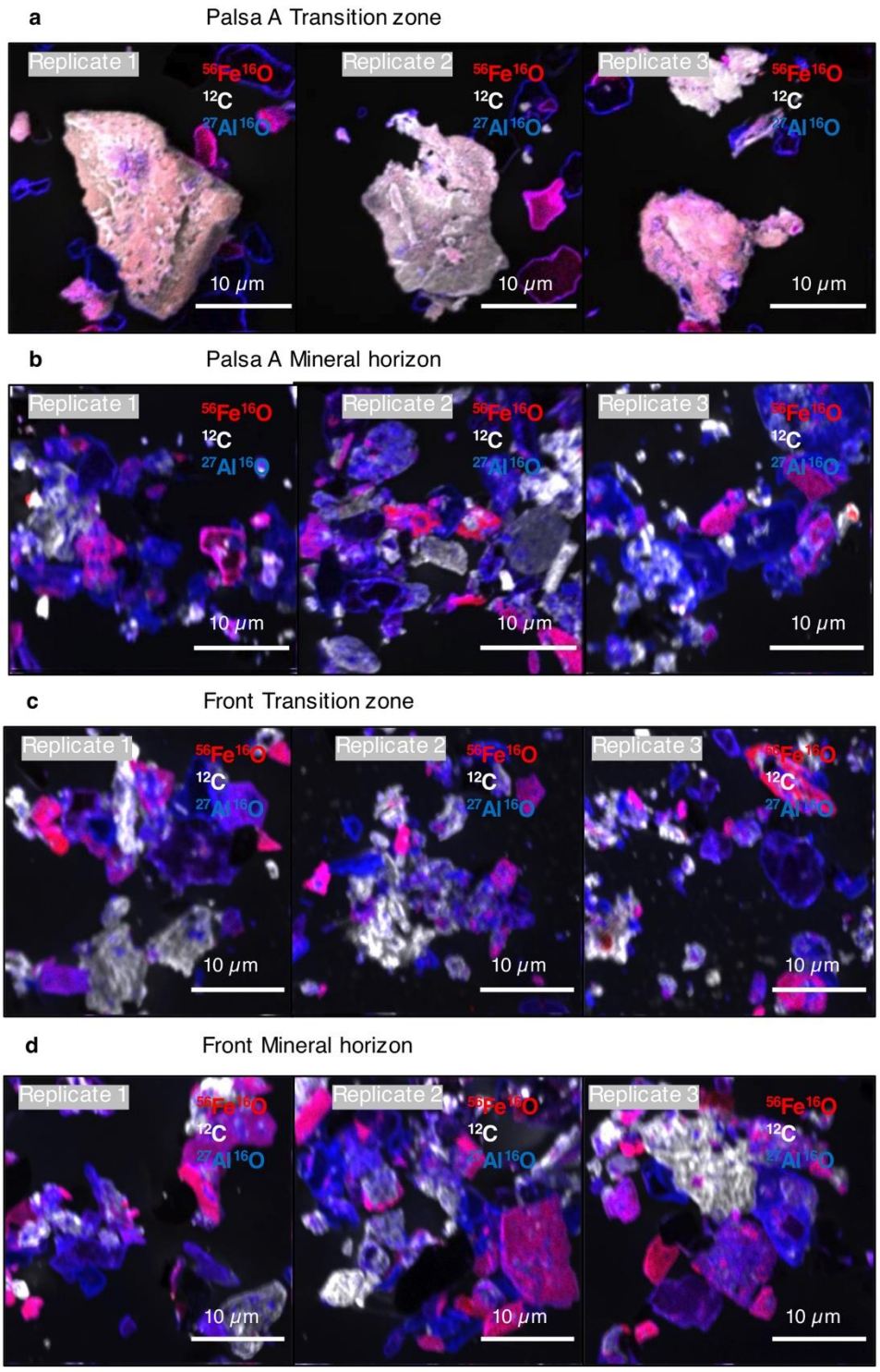


FIGURE S15. REPLICATE ANALYSIS OF NANOSCALE SECONDARY ION MASS SPECTROMETRY (NANOSIMS) OF FINE FRACTION:

a, Palsa A transition zone; b, Palsa A mineral horizon; c, Front transition zone and d, Front mineral horizon (transect 1, Figure S1). Seven detectors were used during nanoSIMS measurements for for $^{12}C^-$, $^{16}O^-$, $^{12}C^{14}N^-$, $^{31}P^-$, $^{32}S^-$, $^{27}Al^{16}O^-$ and $^{56}Fe^{16}O^-$ and $^{27}Al^{16}O$. Here, $^{12}C^-$ (white), $^{56}Fe^{16}O^-$ (red) and $^{27}Al^{16}O^-$ (blue) are overlaid in a single image. In total, four representative fine particles were analyzed with nanoSIMS.

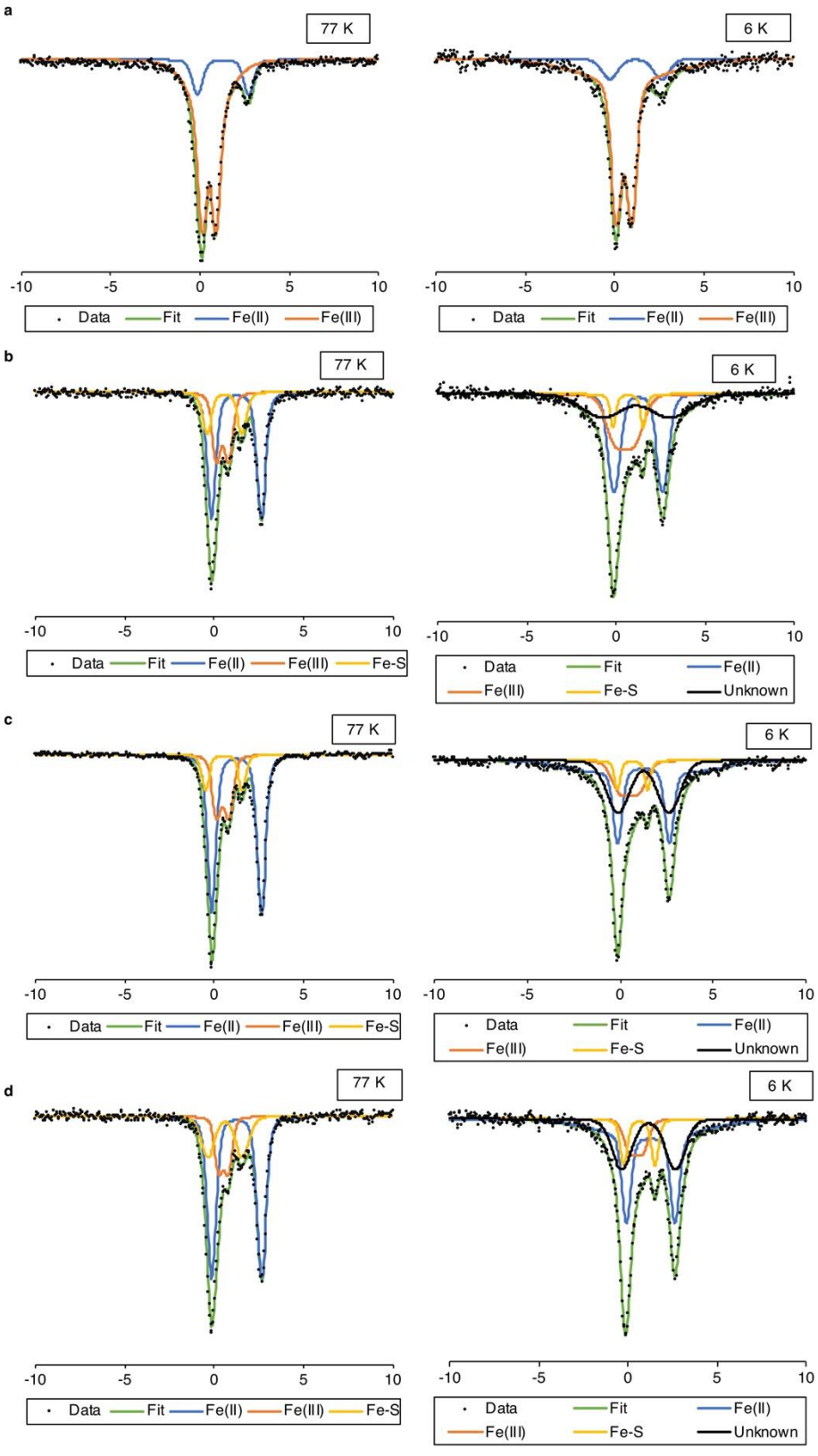

a
77 K
6 K
-10 -5 0 5 10
Data Fit Fe(II) Fe(III)
b
77 K
6 K
-10 -5 0 5 10
Data Fit Fe(II) Fe(III) Fe-S
Data Fit Fe(II) Fe(III) Fe-S Unknown
c
77 K
6 K
-10 -5 0 5 10
Data Fit Fe(II) Fe(III) Fe-S
Data Fit Fe(II) Fe(III) Fe-S Unknown
d
77 K
6 K
-10 -5 0 5 10
Data Fit Fe(II) Fe(III) Fe-S
Data Fit Fe(II) Fe(III) Fe-S Unknown

FIGURE S16. MÖSSBAUER SPECTROSCOPY ANALYSIS AT 77 K AND 6 K OF THE PRESENT FE MINERALS ALONG THE THAW GRADIENT:

a, Palsa A transition zone; b, Palsa A mineral horizon; c, Bog (Bog C^3) mineral horizon; d, Fen (Fen E^3) mineral horizon.

TABLE S7. MÖSSBAUER SPECTROSCOPY PARAMETERS (MEASURED AT 77 AND 6 K)

derived from fitting spectra obtained for Palsa A transition zone and mineral horizon, Bog (Bog C^3) and Fen (Fen E^3) mineral horizon.

Sample	Components	CS[a] (mm/s)	ΔE$_{Q}$ [b] (mm/s)	σ(Δ) [c] (mm/s)	B$_{hf}$ [d] (T)	RA[e] (%)	± (%)	χ2 [f]
77 K								
Palsa A								
Transition zone	Fe(II)	1.28	2.87	0.45	0.00	15.30	0.10	1.77
	Fe(III)	0.50	1.18	1.01	2.06	84.70	0.10	
Palsa A								
Mineral horizon	Fe(II)	1.24	2.79	0.39	0.00	54.46	0.97	0.88
	Fe(III)	0.46	0.67	0.35	0.26	27.45	0.91	
	Fe-S	0.58	1.96	0.40	0.00	18.09	0.93	
Bog (Bog C, see Patzner *et al*., 2020)								
Mineral horizon	Fe(II)	1.25	2.78	0.37	0.00	62.87	0.38	3.47
	Fe(III)	0.47	0.68	0.37	0.31	24.73	0.34	
	Fe-S	0.54	2.07	0.30	0.00	12.40	0.34	
Fen (Fen E, see Patzner *et al*., 2020)								
Mineral horizon	Fe(II)	1.25	2.79	0.44	0.00	61.30	0.21	0.61
	Fe(III)	0.52	0.54	0.31	0.41	17.30	0.18	
	Fe-S	0.61	1.93	0.67	0.03	21.40	0.22	
6 K								
Palsa A								
Transition zone	Fe(II)	1.17	2.90	0.80	0.00	12.30	0.16	0.85
	Fe(III)	0.52	2.17	2.65	2.26	87.70	0.16	
Palsa A								
Mineral horizon	Fe(II)	1.25	2.73	0.58	0.00	38.10	0.25	0.81
	Fe-S	0.68	1.68	0.19	0.00	6.50	0.16	
	Fe(III)	0.49	1.19	0.79	0.64	26.90	0.27	
	Unknown	1.13	3.87	2.06	0.29	28.50	0.30	
Bog (Bog C, see Patzner *et al*., 2020)								
Mineral horizon	Fe(II)	1.26	4.07	2.55	1.50	42.10	0.77	0.75
	Fe-S	0.66	1.64	0.22	0.00	6.00	0.18	
	Fe(III)	0.48	1.07	0.70	0.60	15.90	0.34	
	Unknown	1.26	2.77	1.13	0.08	36.00	0.76	
Fen (Fen E, see Patzner *et al*., 2020)								
Mineral horizon	Fe(II)	1.28	3.65	2.62	2.03	49.20	0.54	0.54
	Fe-S	0.64	1.74	0.27	0.00	10.70	0.27	
	Fe(III)	0.47	0.76	0.48	0.55	11.50	0.27	
	Unknown	1.16	3.02	0.98	0.02	28.50	0.58	

a CS = center shift

b ΔEQ = quadrupole splitting;

c σ(Δ) = standard deviation of quadrupole splitting component d Bhf = hyperfine magnetic field

e RA = Relative abundance

f χ2 = goodness of fit

REFERENCES

1 Rinne, J., ICOS Sweden Ecosystem eco time series (ICOS Sweden), Abisko-Stordalen Palsa Bog, 2018-12-31–2019-12-31, https://hdl.handle.net/11676/s5oBzukX_FaXpHU___86QasO) (2021).

2 Hodgkins, S. B. *et al.* Changes in peat chemistry associated with permafrost thaw increase greenhouse gas production. *Proc. Natl. Acad. Sci. USA* **111**, 5819-5824, (2014).

3 Patzner, M. S. *et al.* Iron mineral dissolution releases iron and associated organic carbon during permafrost thaw. *Nat Commun* **11**, 6329 (2020).

4 Otte, J. M. *et al.* The distribution of active iron-cycling bacteria in marine and freshwater sediments is decoupled from geochemical gradients. *Method Enzymol* **20**, 2483-2499 (2018).

5 Weber, K. A., Achenbach, L.A., Coates, J.D. Microorganisms pumping iron: anaerobic microbial iron oxidation and reduction. *Nat. Rev. Microbiol.* **4**, 752-764 (2006).

6 Dinh, H. T., Kuever, J., MussBmann, M., Hassel, A.W. Iron corrosion by novel anaerobic microorganisms. *Nature* **427**, 829 (2004).

7 Berg, J. S., Michellod, D., Pjevac, P., Martinez-Perez, C., Buckner, C.R., Hach, P.F. Instensive cryptic microbial iron cycling in the low iron water column of the meromictic Lake Cadagno. *Method Enzymol* **18**, 5288-5302 (2016).

8 Li, X., Zhang, W., Liu, T., Chen, L., Chen, P., Li, F. Changes in the composition and diversity of microbial communities during anaerobic nitate reduction and Fe(II) oxidation at circumneutral pH in paddy soil. *Soil Biol Biochem* **94**, 70-79 (2016).

9 Holmes, D., Bond, D., O'neil, R., Reimers, C., Tender, L., Lovely, D. Microbial communities associated with electrodes harvesting electricity from a variety of aquatic sediments *Microb Ecol* **48**, 178-190 (2004).

10 Finneran, K. T., Johnsen, C.V., Lovely, D.R. Rhodoferax ferrireducens sp. nov., a psychotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). *Inst J Syst Evol Microbiol* **53**, 669-673 (2003).

11 Kim, W., Whitman, W.B. Methanogens. *Encyclopedia of Food Microbiology (Second Edition), Academic Press, ISBN 9780123847331*, 602-606 (2014).

12 Mondav, R. *et al.* Discovery of a novel methanogen prevalent in thawing permafrost. *Nat Commun* **5**, 3212 (2014).

13 Jiang, H., Chen, J.C., Murrell, P., Jiang, C., Zhang, C., Xing, X.-H., Smith, T.J. . Methanotrophs: Multifunctional bacteria with promising applications in environmental bioengineering. *Comprehensive Biotechnology (Second Edition), Academic Press, ISBN 9780080885049*, 249-262 (2011).

14 Singelton, C. M. *et al.* Methanotrophy across a natural permafrost thaw environment *The ISME Journal* **12**, 2544-2558 (2018).

15 Lv, J., Huang, Z., Christie, P., Zhang, S. Reducing reagents induce molecular artifacts in the extraction of soil organic matter. *ACS Earth and Space Chem* **4** (11), 1913-1919 (2020).

APPENDIX D - SUPPLEMENTARY INFORMATION FOR CHAPTER 5

SECTION 1. SAMPLE LOCATIONS, DATES, AND COLLECTION

We conducted our research within the Napa River Watershed, California, U.S. (1,103 km^2), where viticulture comprises ~13% of the watershed area (Figure 1; Table S1). Forests (26% of watershed area) and shrubland/grasslands (37% of watershed area) surround vineyard agriculture, providing a natural contrast in S biogeochemistry between agricultural S inputs and background S sources (e.g., geologic and atmospheric).

To assess S chemistry at the agricultural plot, or "field" scale, we collected soil porewater at four sites contrasting forest-grasslands and vineyard agriculture: a woodland-grassland hillslope (C1; G-PW), a forested hillslope (C2 headwaters; F-PW), and two vineyards receiving 22-40 kg S ha^{-1} yr^{-1} S^0 applications (C2 and C3, V-PW1 and V-PW2, respectively; Table S1). At each site, we installed four pre-rinsed (18.2 MΩ cm^{-1} deionized water) tension lysimeters (SoilMoisture, Inc.) to 25-53 cm depth (installed to resistance) across 0.1-4.0 ha to capture spatial heterogeneity within a field. Soil porewater samples were collected after one purge into acid-washed (pH 2 hydrochloric acid) and ultra-high purity (UHP)-N_2-flushed polypropylene 1 L bottles (Prenart Equipment ApS, Denmark) to preserve porewater redox conditions. We stored porewater samples on ice in transit, and then filtered them through 0.45 μm polyethersulfone filters (Pall Supor) into acid-washed, combusted (5 hrs at 450°C), and N_2-purged amber glass bottles in a UHP-N_2-purged glove bag. Porewater samples were shipped on ice to the University of Colorado, Boulder and then were stored at 4°C until processing.

We collected receiving surface water runoff samples to test whether vineyard DOS was detectable beyond sites of agricultural S application. During the first campaign, we collected stream water runoff draining C1 (G-RO), stream water from the headwaters and just downstream

of a vineyard in C2 (HF F-RO and M-RO1, respectively), stream water downstream of a vineyard culvert outflow in C3 (M-RO2), and the Napa River (HF N-RO, Table S1). Many streams were dry during the second field collection, but we were able to resample the headwaters of C2 (LF F-RO), the Napa River (LF N-RO), and collected vineyard culvert runoff, which we considered to be a vineyard S endmember (V-RO). For each sample, we pumped (Geotech Geopump) stream water through 0.45 μm capsule filters (Geotech) into separate bottles for chemical analyses and DOM isolation as follows: 10 L HDPE pre-rinsed containers for DOM isolation, 125 mL pre-combusted glass amber bottles for DOC and UV analyses, and HDPE bottles for S analyses (500 mL to 1 L acid-washed). Stream water samples were stored and shipped on wet ice in transit and then stored at 4°C until analyses.

Table S1. Sample locations and dates

Sample Name	Sample Dates[a]	Sample Code[b]	Catch.[b]	Area (ha)	Land use/land cover (%)[d]				Geology[e]	Soil Series[f]
					Forest	Shrubland/ Grassland	Vineyard	Other		
Grassland soil porewater	12/06/2019 - 12/14/2019	G-PW	C1	~0.15	0	100	0	0	Franciscan Complex, metagraywacke	Sobrante loam
Grassland stream runoff	12/08/2019	G-RO	C1	1,970[c]	64.0	35.0	0.1	1.3	Franciscan Complex, metagraywacke	Sobrante loam, some Kidd loam and Henneke gravelly loam
Forest soil porewater	12/06/2019 - 12/14/2019	F-PW	C2	~0.15	100	0	0	0	Great Valley Complex sandstone and shale	Felton gravelly loam or silt loam
High flow forest stream runoff	12/07/2019	HF F-RO	C2	260[c]	72.0	28.0	0	0	Franciscan Complex greywacke and mélange, Great Valley Complex sandstone and shale	Lodo-Maymen-Felton association, Forward silt loam
Low flow forest stream runoff	02/29/2020	LF F-RO	C2	260[c]	72.0	28.0	0	0	Franciscan Complex greywacke and mélange, Great Valley Complex sandstone and shale	Lodo-Maymen-Felton association, Forward silt loam
Vineyard soil porewater 1	12/05/2019 - 12/14/2019	V-PW1	C2	~2.0	0	0	100	0	Alluvial fan deposits	Bale clay loam/loam

Mixed stream runoff 1	12/07/2019	M-RO1	C2	383[c]	67.8	23.8	7.8	0.6	Franciscan Complex greywacke and mélange, Great Valley Complex sandstone/shale, Sonoma Volcanics andesite to basalt lava flow, alluvial fan deposits	Lodo-Maymen-Felton association, Forward silt loam, Bale clay loam/loam
Vineyard soil porewater 2	12/04/2019 - 12/10/2019	V-PW2	C3	~2.5	0	0	100	0	Alluvial fan deposits	Clear Lake clay
Vineyard 2 culvert runoff	03/01/2020	V-RO	C3	N/A	0	0	100	0	Alluvial fan deposits	Clear Lake clay
Mixed stream runoff 2	12/08/2019	M-RO2	C3	568[c]	56.3	17.1	23.2	3.4	Franciscan Complex greywacke and mélange, Great Valley Complex sandstone/shale, Sonoma Volcanics andesite to basalt lava flow, alluvial fan deposits	Felton gravelly loam, Hambright rock outcrop complex, Bale clay loam, Clear Lake clay
High flow Napa River runoff	12/09/2019	HF N-RO	NA	46,596[c]	40.9	38.2	13.2	7.7	Mixed	Mixed
Low flow Napa River runoff	02/29/2020	LF N-RO	NA	46,596[c]	40.9	38.2	13.2	7.7	Mixed	Mixed

[a] Sample date format: MM/DD/YYYY

[b] Sample codes and sub-catchment ("Catch.") designations correspond to locations in Figure 1

[c] Watershed areas estimated using U.S.G.S. (2016b)

[d] Land use/land cover designations are as follows: “Forest” = deciduous, evergreen, and mixed forest; “Shrubland/grassland” = shrub/scrub and grassland/herbaceous; “Vineyard” = cultivated crops; “Other” = open water, developed land, and/or wetlands. Percentages are calculated as a fraction of total catchment area. Data are from Yang et al. (2018).

[e] Graymer et al. (2007)

[f] O’Geen et al. (2017)

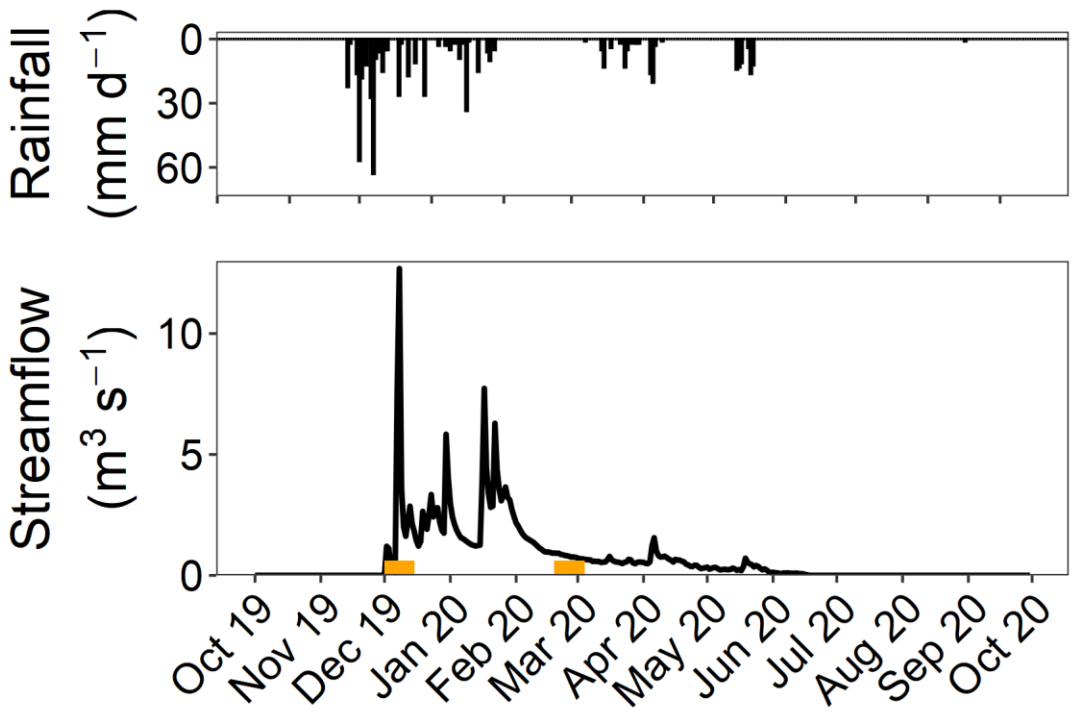


Figure S1. Study year climatology. Water year 2020 St. Helena 4WSW daily precipitation (top panel; GHCND #USC00047646, ncdc.noaa.gov) and U.S. Geological Survey Napa River NR Napa CA station streamflow (bottom panel; #11458000; U.S.G.S. 2016b). Yellow boxes denote field sampling campaigns, which targeted the first significant rainfall event of the 2019-2020 wet season and low flow conditions later in the season.

SECTION 2. CHEMICAL ANALYSES

We analyzed soil porewater and surface water samples for aqueous S and C chemistry (Table S2). Dissolved organic carbon (DOC) was measured by persulfate oxidation (OI Analytical, model 700; detection limit 0.2 mg L^{-1}, relative percent difference (RPD) of sample duplicates < 3.2%) or combustion catalytic oxidation (Shimadzu TOC-L/TNM-L analyzer; detection limit 0.04 mg L^{-1}, RPD < 5%). We quantified SO_4^{2-} concentrations by ion chromatography (Dionex; detection limit 0.2 mg S L^{-1}, RPD ≤ 5%) and total dissolved sulfur (TDS) by inductively-coupled plasma mass spectrometry (ICP-MS; Thermo Finnigan Element2; detection limit 0.05 mg S L^{-1}; RPD ≤ 10%). Organic S concentrations were not detectable, with

differences between total dissolved S and SO_4^{2-} within instrument error (Table S2). UV-vis absorption spectra were collected at room temperature using a 1 cm quartz cuvette (Agilent Technologies, model 8453) and the absorbance at 254 nm was used to calculate $SUVA_{254}$.

For stable isotope analysis, we precipitated $BaSO_4$ with excess $BaCl_2$ after sample acidification to pH 2-4 and boiling (Carmody et al., 1998; Hinckley et al., 2008) and collected $BaSO_4$ on 0.45 μm Whatman mixed cellulose ester filters, which we then dried overnight at 60°C. Following Xie et al. (2016), we combusted solid $BaSO_4$ for 2 hrs at 800°C to remove organic co-precipitates and ensure pure inorganic S for stable isotope analysis.

After DOM isolation (detailed below), we analyzed the PPL-extracted DOM for carbon and sulfur content ([C]-DOM and [S]-DOM, respectively) and δ^{34}S-DOM. We measured [C]-DOM on a carbon-hydrogen-nitrogen (CHN) elemental analyzer (Thermo Finnigan Flash) with <1.6 % RPD across duplicate standards and measured [S]-DOM alongside δ^{34}S-DOM measurements.

For S stable isotope analysis, solid $BaSO_4$ and freeze-dried PPL extracts were packed into tin capsules and analyzed for $^{34}S/^{32}S$ and %S on a Flash IRMS elemental analyzer coupled with a Delta V Plus isotope ratio mass spectrometer (Thermo Fisher Scientific EA IsoLink) at the University of California Berkeley Center for Stable Isotope Biogeochemistry. In parallel, we isolated DOM from two samples collected in the Florida Everglades, which validated that there was not isotopic fractionation during PPL extraction (see Supplementary Information Section 9).

Table S2. Aqueous dissolved chemistry of surface and soil porewater samples.

Sample Code	Land use/land cover	Sample Type	DOC (mg C L^{-1})	$SUVA_{254}$ (L mg C^{-1} m^{-1})	TDS (mg S L^{-1})	SO_4^{2-} (mg S L^{-1})
F-PW	forest/grassland	soil porewater	515; 43.7[a]	1.79	2.97	3.05
HF F-RO	forest/grassland	surface water	11.8	3.51	6.56	6.27

LF F-RO	forest/grassland	surface water	1.58	2.47	8.02	9.16
G-PW	forest/grassland	soil porewater	9.42	1.36	0.648	0.996
G-RO	forest/grassland	surface water	13.6	3.74	3.45	3.22
V-PW1	vineyard	soil porewater	11.17	1.44	5.05	4.83
V-PW2	vineyard	soil porewater	12.1	3.32	14.5	14.5
V-RO	vineyard	surface water	1.14	2.33	24.2	38.5
M-RO1	mixed	surface water	13.9	3.53	5.22	4.74
M-RO2	mixed	surface water	5.02	2.70	24.5	26.7
HF N-RO	mixed	surface water	7.80	3.19	8.67	8.50
LF N-RO	mixed	surface water	2.00	2.41	12.4	13.3

[a] The two soil porewater subsamples collected at F-PW had contrasting [DOC]; both are reported for completeness.

Table S3. Elemental composition of PPL-extracted DOM

Sample Code	Land use/land cover	Sulfur (mmol S (g DOM)$^{-1}$)	Carbon (mmol C (g DOM)$^{-1}$)	Atomic C:S	Atomic S:C (x 10^{-3})
F-PW	forest/grassland	0.12	44.21	366	2.73
HF F-RO	forest/grassland	0.10	41.61	421	2.38
LF F-RO	forest/grassland	0.17	42.25	247	4.05
G-PW	forest/grassland	0.11	40.69	354	2.82
G-RO	forest/grassland	0.11	42.39	383	2.61
V-PW1	vineyard	0.18	40.65	229	4.37
V-PW2	vineyard	0.31	39.45	126	7.95
V-RO	vineyard	0.20	41.32	207	4.83
M-RO2	mixed	0.12	42.55	347	2.88
M-RO1	mixed	0.11	41.86	371	2.70
LF N-RO	mixed	0.24	40.70	173	5.80
HF N-RO	mixed	0.20	42.03	208	4.81

SECTION 3. DOM ISOLATION DETAILS

We isolated DOM from soil porewater and stream water samples following Dittmar et al. (2008) with modifications after Tfaily et al. (2012), Pohlabeln and Dittmar (2015), Lv et al., (2016), Stücheli et al. (2018), and Bahureksa et al. (2021). All extraction materials were pre-rinsed with 18.2 MΩ cm^{-1} deionized H_2O and glassware was combusted at 450°C for 5 hrs. Prior to extraction, samples were acidified to pH 1.8-2.2 with trace-metal-clean hydrochloric acid (HCl). Soil porewater samples were acidified in a glove box with de-aired HCl. We conditioned PPL cartridges for extraction with sequential rinses as follows: 1-volume methanol (ACS grade), 1-volume methanol soak for ≥ 4 hours or overnight, 2-volumes 18.2 M-Ω cm^{-1} DI H_2O, 2-volumes methanol, 2-volumes 0.01 M HCl.

During DOM extraction, samples were pumped through the pre-rinsed PPL cartridges at < 40 mL min^{-1}. Samples were passed over multiple cartridges if necessary to ensure ≤ 24 mg

DOC per 1 g PPL sorbent (Dittmar et al., 2008). Soil porewater samples were extracted into an N_2-purged chamber. After sample loading, we removed salts with 2-volume rinses of 0.01 M HCl and then dried cartridges under N_2. We eluted DOM slowly with 1-cartridge volume of methanol into washed and combusted glass amber vials with PTFE-lined caps (8 mL or 60 mL depending on sample volume). We then dried PPL eluents under N_2 and stored samples at -20°C until analyses (Dittmar et al., 2008).

SECTION 4. FT-ICR MS ANALYSIS

We detail additional FT-ICR MS analysis conditions here. All 21 tesla FT-ICR MS mass spectra files are publicly-available via the Open Science Framework (https://osf.io/nh3zc/) at DOI 10.17605/OSF.IO/NH3ZC.

Sample solution was infused via a microelectrospray source (Emmett et al., 1998; 50 μm i.d. fused silica emitter) at 500 nL/min by a syringe pump. Typical conditions for negative ion formation were: emitter voltage, -2.4-2.9 kV;; and heated metal capillary temperature, 350 C. Ions were released *m/z*-dependently by decrease of an auxiliary radio frequency potential between the multipole rods and the end-cap electrode (Kaiser et al., 2013), excited to *m/z*-dependent radius to maximize the dynamic range and number of observed mass spectral peaks (32-64%; Kaiser et al., 2013) and excitation and detection were performed on the same pair of electrodes (Chen et al., 2014). The dynamically harmonized ICR cell in the 21 T FT-ICR is operated with 6 V trapping potential (Kaiser et al., 2013; Boldin et al., 2011). Time-domain transients of 3.1 seconds were acquired with the Predator data station that handled excitation and detection only, initiated by a TTL trigger from the commercial Thermo data station, with 100 time-domain acquisitions averaged for all experiments (Blakney et al., 2011). Peaks with signal

magnitude > 6 times the baseline root-mean-square noise (S/N ≥ 6) at *m/z* 500 were exported to peak lists for formula assignment and data visualization. Experimentally measured masses were converted from the International Union of Pure and Applied Chemistry (IUPAC) mass scale to the Kendrick mass scale (Kendrick et al., 1963) for rapid identification of homologous series for each heteroatom class (i.e., species with the same $C_cH_hN_nO_oS_s$ content, differing only be degree of alkylation; Hughey et al., 2001).

In PetroOrg, formulas were assigned elemental compositions ($C_{1\text{-}100}H_{1\text{-}200}N_{0\text{-}3}O_{1\text{-}35}S_{0\text{-}2}$) with a mass accuracy error ≤ 0.3 ppm. Formula were validated using the following criteria based on Kind and Fiehn (2007) and Koch et al. (2007): the Nitrogen rule, rings plus double bonds (DBEs) ≥ 0 and whole numbers, LEWIS and SENIOR chemical rules, common element ratios, the chemical building block approach, and matching ^{13}C and ^{34}S isotope peaks when detected. Final validated formula assignments totaled 20,250 – 29,128 across samples (Table S5; see Supplementary Information Section 6).

SECTION 5. FT-ICR MS RESULTS

Each FT-ICR MS-measured and validated formula was summarized into a heteroatom class, degree of saturation (double bond equivalents, DBE = number of rings plus double bonds to carbon, DBE = *C*-*h*/2 + *n*/2 + 1; McLafferty et al., 1993) and carbon number, *c*. We also calculated the modified aromaticity index, AImod, nominal carbon oxidation state (NOSC), and compound classes. AImod was calculated after Koch and Dittmar (2006):

$$AImod = \frac{1 + C - 0.5H - 0.5N - 0.5O - S}{C - N - 0.5O - S}$$

where *C*, *H*, *N*, *O*, and *S* are the number of carbon, hydrogen, nitrogen, oxygen, and sulfur atoms in a chemical formula. The average AImod for a sample was calculated as the abundance-weighted average of all assigned formulas. We calculated NOSC after Kroll et al. (2011):

$$NOSC = -\sum_i OS_i \frac{n_i}{c_i}$$

Where the summation is for all non-carbon elements, OS_i is the oxidation state of element *i*, and n_i/c_i is the molar ratio of the element *i* to carbon. NOSC ranges from -4 to 4, with higher values indicating a higher carbon oxidation state and lower values indicating a lower carbon oxidation state. Lastly, assigned formulas were grouped by AImod and O/C and H/C ratios (Table S4) after Poulin et al. (2017).

Table S4. Criteria for the tentative grouping of molecular formula assignments into compound classes based on elemental stoichiometries and AImod.

Compound Class	Criteria
Condensed aromatics	AImod > 0.67
Aromatics/polyphenols	0.5 > AImod > 0.67
Highly unsaturated	AImod < 0.5, H/C < 1.5, O/C < 0.9
Unsaturated aliphatic	H/C = 1.5 – 2.0, O/C < 0.9, N = 0
Saturated fatty acids	H/C ≥ 2.0; O/C ≤ 0.9
Peptide-like	H/C = 1.5 – 2.0, O/C < 0.9, N > 0
Sugar-like	O/C > 0.9

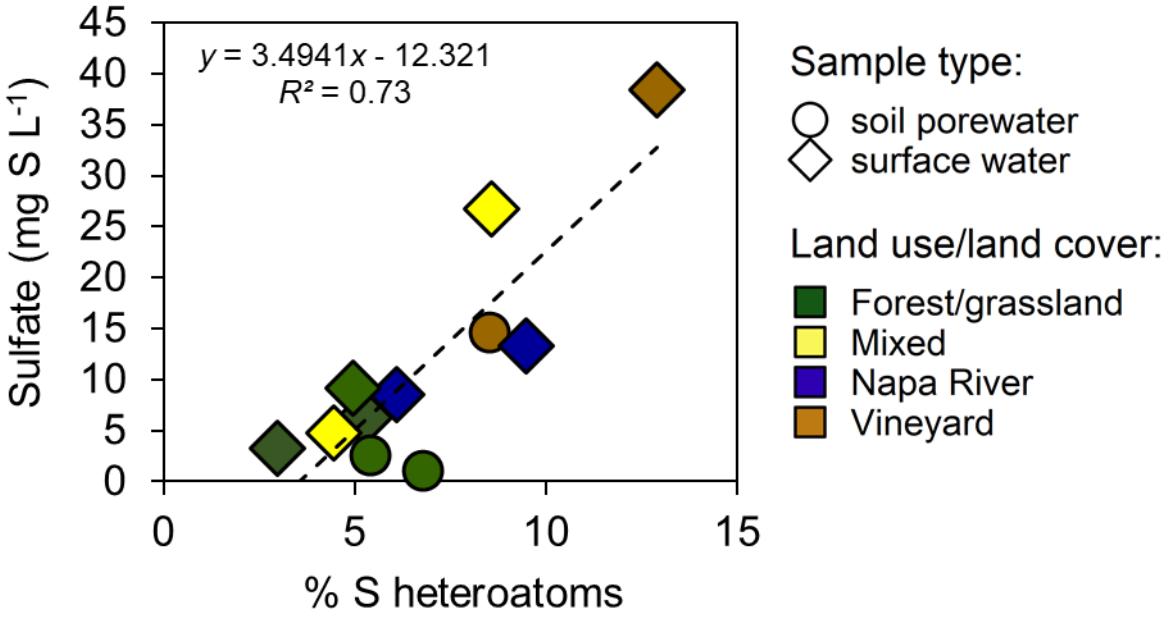


Figure S2. Correlation between the percent of S-containing heteroatom formulas and [SO_4^{2-}].

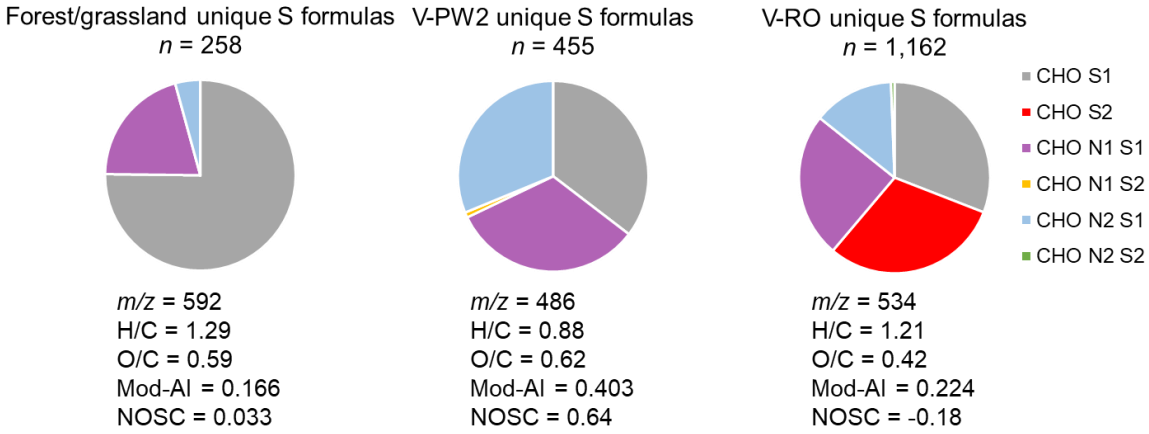


Figure S3. Unique S-containing formulae for forest/grassland samples (G-PW, G-RO, F-PW, HF F-RO, LF F-RO), vineyard soil porewater (V-PW2), and vineyard culvert runoff (V-RO).

SECTION 7. SULFUR K-EDGE XANES SPECTROSCOPY

For S XANES analysis, solid DOM samples were pressed as 6 mm pellets. Replicate scans on pristine material were collected between 2450 – 2530 eV with a 0.2 eV step interval and between 2460 – 2486 eV with a 0.5 eV step interval in pre- and post-edge regions. Spectra were collected under a liquid helium cryostream (~ 100 K in a helium atmosphere, < 0.5 ppm O_2) using a Ge 7 element detector (Canberra X-PIPS 7-Element Silicon Drift Detectors). For samples with low mass, dilutions were performed with 20 μm cellulose. Samples were calibrated against a sodium thiosulfate standard at 2472.02 eV.

Spectra were fit according to Manceau and Nagy (2012) using two arctangent and six Gaussian functions representing exocyclic reduced, heterocyclic reduced, sulfoxide, sulfone, sulfonate, and organosulfate functionalities. We ensured that the initial energies of the two arctangent functions were set below the energies of the sulfoxide and organosulfate Gaussian curves, respectively. Gaussian energies were kept within ± 0.1 eV of recommended values (Manceau and Nagy, 2012). Fits aimed to minimize the normalized sum-squared (NSS) residual:

$$NSS = \frac{\sum(y_{exp} - y_{fit})^2}{\sum y_{exp}^2}$$

where y_{exp} and y_{fit} are experiment and fit values, respectively. Gaussian fit areas were corrected for the change in X-ray absorption cross-section with increasing oxidation state using:

$$y = 0.36841x - 909.97$$

where x is the energy of the absorption maximum of a Gaussian function and y is the scaling factor (Manceau and Nagy, 2012). Spectral decompositions and fit values are shown in Figures S18 and S19 and Tables S8 and S9.

We estimated the spectral precision by analyzing an internal control sample, Everglades F1 HPOA SW (Figure S20; Table S10), which was previously analyzed at the Argonne National Laboratory Advanced Photon Source (APS, beamline 9-BM; Poulin et al.,

2017). Analytical precision, reported as the difference in the relative abundance of a given S functionality between independent measurements at APS and SSRL, is estimated to be within 2% for a given S functionality. Accuracies for Sred and Sox are typically ≤ 4% (Manceau and Nagy, 2012).

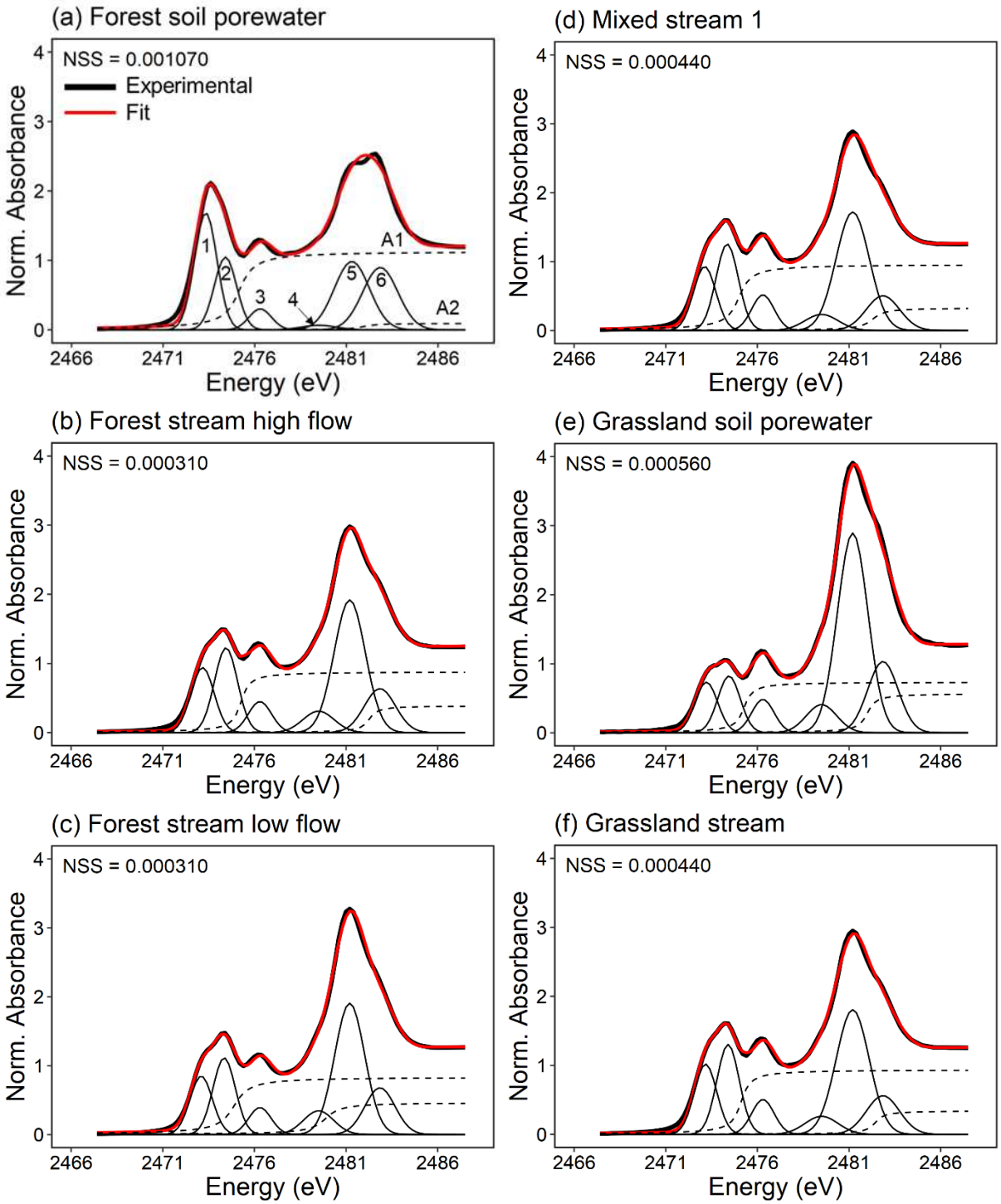


Figure S18. Sulfur K-edge XANES spectral decomposition of PPL-extracted DOM. The normalized sum-squared residual (NSS) is the normalized difference between experimental (black thick line) and fit (red line) spectra. Sulfur functionalities corresponding to the six Gaussian functions are labeled as follows in (a): 1 – exocyclic reduced; 2 – heterocyclic reduced;

3 – sulfoxide; 4 – sulfone; 5 – sulfonate; 6 – organosulfate. Arctangent curves (dashed lines) are labeled as “A1” and “A2” in (a). The parameter values and atomic fractions of each functionality are in Table S8.

Table S8. Gaussian curve fitting (GCF) parameter values and sulfur atomic fractions (%) of PPL-isolated DOM measured spectra. Corresponding GCF curves are shown in Figure S18.

Sample	Function	S functionality	Energy (eV)[a]	Width[a]	Calib. Area[b]	Atomic Fraction (%)
Forest soil porewater (F-PW)	Arc 1		2475.2	**0.60**		
	Arc 2		2481.9	**0.60**		
	G1	Exocyclic S	2473.3	**0.61**	2.33	50.8
	G2	Heterocyclic S	**2474.4**	**0.61**	0.99	21.5
	G3	Sulfoxide	**2476.3**	**0.61**	0.20	4.3
	G4	Sulfone	**2479.5**	**0.96**	0.04	1.0
	G5	Sulfonate	**2481.3**	**0.96**	0.57	12.4
	G6	Organosulfate	**2482.85**	**0.96**	0.46	10.0
High flow forest stream runoff (HF F-RO)	Arc 1		2475.3	**0.35**		
	Arc 2		2482.3	**0.35**		
	G1	Exocyclic S	2473.2	**0.63**	1.34	31.3
	G2	Heterocyclic S	**2474.45**	**0.63**	1.19	27.8
	G3	Sulfoxide	**2476.3**	**0.63**	0.30	7.0
	G4	Sulfone	**2479.5**	**0.85**	0.19	4.4
	G5	Sulfonate	**2481.2**	**0.85**	0.98	22.9
	G6	Organosulfate	**2482.85**	**0.85**	0.29	6.7
Low flow forest stream runoff (LF F-RO)	Arc 1		2474.9	**0.62**		
	Arc 2		2479.9	**0.62**		
	G1	Exocyclic S	2473.1	**0.61**	1.17	29.9
	G2	Heterocyclic S	**2474.35**	**0.61**	1.05	26.8
	G3	Sulfoxide	**2476.3**	**0.61**	0.25	6.5
	G4	Sulfone	**2479.5**	**0.82**	0.20	5.1
	G5	Sulfonate	**2481.2**	**0.82**	0.94	24.1
	G6	Organosulfate	**2482.85**	**0.82**	0.30	7.6
Mixed stream 1 (M-RO1)	Arc 1		2475.0	**0.50**		
	Arc 2		2482.3	**0.50**		
	G1	Exocyclic S	2473.1	**0.61**	1.28	31.0
	G2	Heterocyclic S	**2474.35**	**0.61**	1.18	28.6
	G3	Sulfoxide	**2476.3**	**0.61**	0.33	8.1
	G4	Sulfone	**2479.5**	**0.91**	0.15	3.6
	G5	Sulfonate	**2481.2**	**0.91**	0.94	22.8
	G6	Organosulfate	**2482.85**	**0.91**	0.24	5.9

Table S8 continued

Sample	Function	S functionality	Energy (eV)[a]	Width[a]	Calib. Area[b]	Atomic Fraction (%)
Grassland soil porewater (G-PW)	Arc 1		2475.3	**0.35**		
	Arc 2		2482.0	**0.35**		
	G1	Exocyclic S	2473.2	**0.61**	1.01	23.9
	G2	Heterocyclic S	**2474.5**	**0.61**	0.77	18.3
	G3	Sulfoxide	**2476.3**	**0.61**	0.31	7.4
	G4	Sulfone	**2479.5**	**0.83**	0.24	5.6
	G5	Sulfonate	**2481.2**	**0.83**	1.45	34.1
	G6	Organosulfate	**2482.9**	**0.83**	0.46	10.8
Grassland stream (G-RO)	Arc 1		2475.1	**0.38**		
	Arc 2		2482.4	**0.38**		
	G1	Exocyclic S	2473.2	**0.61**	1.40	32.1
	G2	Heterocyclic S	**2474.4**	**0.61**	1.23	28.1
	G3	Sulfoxide	**2476.3**	**0.61**	0.33	7.5
	G4	Sulfone	**2479.5**	**0.90**	0.17	3.8
	G5	Sulfonate	**2481.2**	**0.90**	0.98	22.4
	G6	Organosulfate	**2482.85**	**0.90**	0.27	6.2

[a] bold values were co-varied and bold and underlined values were fixed during GCF

[b] calculated using generic calibration curve in Manceau and Nagy (2012)

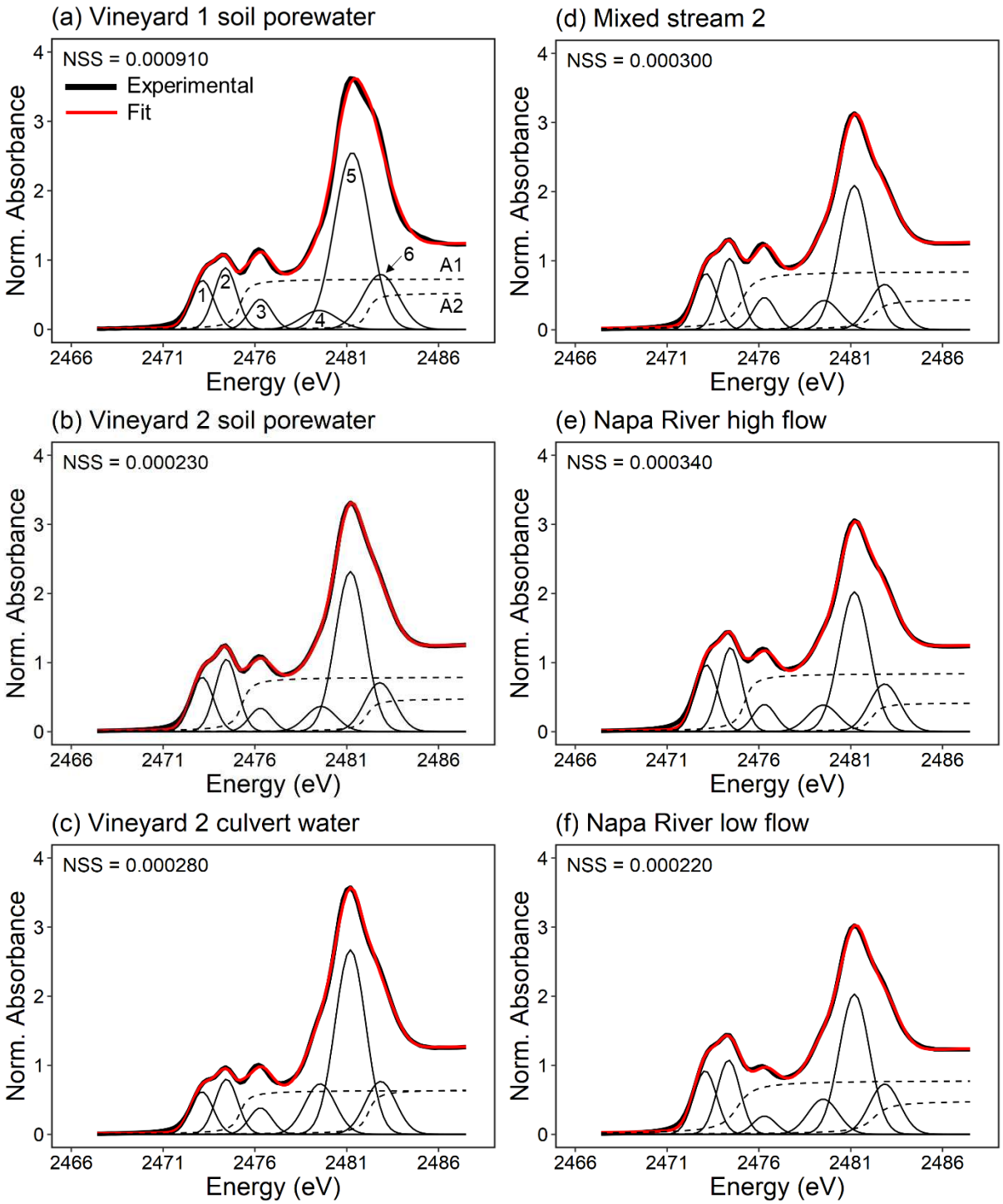


Figure S19. Sulfur K-edge XANES spectral decomposition of PPL-extracted DOM, continued. See Figure S18 for description. The parameter values and atomic fractions of each functionality are in Table S9.

Table S9. Gaussian curve fitting (GCF) parameter values and sulfur atomic fractions (%) of PPL-isolated DOM measured spectra. Corresponding GCF curves are shown in Figure S19.

Sample	Function	S functionality	Energy (eV)[a]	Width[a]	Calib. Area[b]	Atomic Fraction (%)
Vineyard 1 soil porewater (V-PW1)	Arc 1		2475.2	**0.30**		
	Arc 2		2482.1	**0.30**		
	G1	Exocyclic S	2473.1	**0.61**	0.97	23.3
	G2	Heterocyclic S	**2474.4**	**0.61**	0.84	20.0
	G3	Sulfoxide	**2476.3**	**0.61**	0.28	6.7
	G4	Sulfone	**2479.5**	**0.97**	0.19	4.5
	G5	Sulfonate	**2481.3**	**0.97**	1.49	35.6
	G6	Organosulfate	**2482.85**	**0.97**	0.41	9.8
Vineyard 2 soil porewater (V-PW2)	Arc 1		2475.3	**0.35**		
	Arc 2		2482.1	**0.35**		
	G1	Exocyclic S	2473.1	**0.61**	1.09	27.4
	G2	Heterocyclic S	**2474.45**	**0.61**	0.99	24.9
	G3	Sulfoxide	**2476.3**	**0.61**	0.22	5.6
	G4	Sulfone	**2479.6**	**0.82**	0.21	5.4
	G5	Sulfonate	**2481.2**	**0.82**	1.14	28.9
	G6	Organosulfate	**2482.8**	**0.82**	0.31	7.8
Vineyard 2 culvert runoff (V-RO)	Arc 1		2475.2	**0.33**		
	Arc 2		2482.2	**0.33**		
	G1	Exocyclic S	2473.1	**0.61**	0.85	21.5
	G2	Heterocyclic S	**2474.45**	**0.61**	0.75	19.0
	G3	Sulfoxide	**2476.3**	**0.61**	0.25	6.3
	G4	Sulfone	**2479.55**	**0.83**	0.43	10.9
	G5	Sulfonate	**2481.2**	**0.83**	1.33	33.7
	G6	Organosulfate	**2482.85**	**0.83**	0.34	8.6
Mixed stream 2 (M-RO2)	Arc 1		2475.1	**0.50**		
	Arc 2		2482.3	**0.50**		
	G1	Exocyclic S	2473.1	**0.60**	1.10	28.0
	G2	Heterocyclic S	**2474.4**	**0.60**	0.95	24.2
	G3	Sulfoxide	**2476.3**	**0.60**	0.30	7.6
	G4	Sulfone	**2479.55**	**0.83**	0.25	6.3
	G5	Sulfonate	**2481.2**	**0.83**	1.04	26.5
	G6	Organosulfate	**2482.85**	**0.83**	0.29	7.4

Table S9 continued

Sample	Function	S functionality	Energy (eV)[a]	Width[a]	Calib. Area[b]	Atomic Fraction (%)
High flow Napa River runoff (HF N-RO)	Arc 1		2475.2	**0.35**		
	Arc 2		2482.2	**0.35**		
	G1	Exocyclic S	2473.1	**0.61**	1.33	31.2
	G2	Heterocyclic S	**2474.45**	**0.61**	1.15	26.8
	G3	Sulfoxide	**2476.3**	**0.61**	0.26	6.0
	G4	Sulfone	**2479.5**	**0.83**	0.23	5.3
	G5	Sulfonate	**2481.2**	**0.83**	1.01	23.6
	G6	Organosulfate	**2482.85**	**0.83**	0.31	7.1
Low flow Napa River runoff (LF N-RO)	Arc 1		2474.8	**0.62**		
	Arc 2		2482.3	**0.62**		
	G1	Exocyclic S	2473.1	**0.62**	1.29	31.4
	G2	Heterocyclic S	**2474.35**	**0.62**	1.03	25.0
	G3	Sulfoxide	**2476.3**	**0.62**	0.17	4.3
	G4	Sulfone	**2479.5**	**0.82**	0.30	7.2
	G5	Sulfonate	**2481.2**	**0.82**	1.00	24.4
	G6	Organosulfate	**2482.85**	**0.82**	0.32	7.8

[a] bold values were co-varied and bold and underlined values were fixed during GCF

[b] calculated using generic calibration curve in Manceau and Nagy (2012)

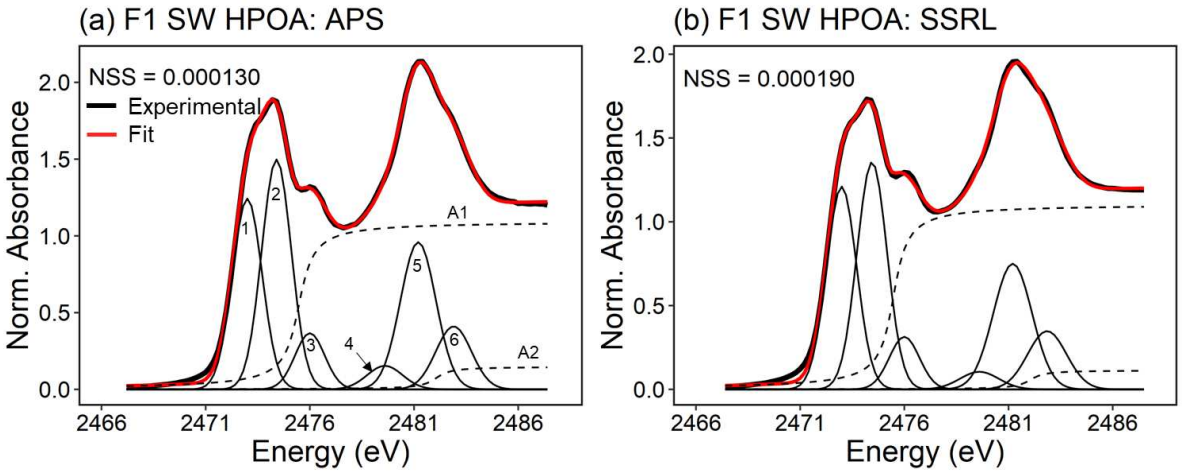


Figure S20. Sulfur K-edge XANES spectral decomposition of F1 SW HPOA analyzed at (a) 9-BM at APS and (b) 4-3 at SSRL. The normalized sum-squared residual (NSS) is the normalized difference between experimental (black thick line) and fit (red line) spectra. Gaussian functions are labeled as follows in (a): 1 – exocyclic reduced; 2 – heterocyclic reduced; 3 – sulfoxide; 4 – sulfone; 5 – sulfonate; 6 – organosulfate. Arctangent curves (dashed lines) are labeled in (a) as "A1" and "A2."

Table S10. Gaussian curve fitting (GCF) parameter values and sulfur atomic fractions (%) of F1 SW HPOA measured spectra. Corresponding GCF curves are shown in Figure S20.

Sample	Function	S functionality	Energy (eV)[a]	Width[a]	Calib. Area[b]	Atomic Fraction (%)
F1 SW HPOA: APS	Arc 1		2475.2	**0.52**		
	Arc 2		2482.2	**0.52**		
	G1	Exocyclic S	2473.0	**0.71**	2.00	42.6
	G2	Heterocyclic S	**2474.4**	**0.71**	1.64	35.1
	G3	Sulfoxide	**2476.4**	**0.71**	0.28	5.9
	G4	Sulfone	**2479.6**	**0.85**	0.09	2.0
	G5	Sulfonate	**2481.3**	**0.85**	0.49	10.5
	G6	Organosulfate	**2482.75**	**0.85**	0.19	4.0
F1 SW HPOA: SSRL	Arc 1		2475.2	**0.50**		
	Arc 2		2482.2	**0.50**		
	G1	Exocyclic S	2473.0	**0.72**	1.97	45.0
	G2	Heterocyclic S	**2474.5**	**0.72**	1.51	34.4
	G3	Sulfoxide	**2476.4**	**0.72**	0.24	5.5
	G4	Sulfone	**2479.6**	**0.93**	0.07	1.6
	G5	Sulfonate	**2481.3**	**0.93**	0.42	9.6
	G6	Organosulfate	**2482.75**	**0.93**	0.17	3.9

[a] bold values were co-varied and bold and underlined values were fixed during GCF

[b] calculated using generic calibration curve in Manceau and Nagy (2012)

SECTION 8. PPL RECOVERY

We calculated PPL carbon recovery after Dittmar et al. (2008), as:

$$\%\ C\ recovered = \frac{[DOC]extr * Vextr}{[DOC]el * F * Vel}$$

where *[DOC]extr* is the concentration of DOC in the filtered, whole water sample pre-extraction, *Vextr* is the sample volume extracted, *[DOC]el* is the concentration of DOC in the PPL eluent (after drying and diluting into water), *F* is the eluent dilution factor, and *Vel* is the volume of PPL eluent. Carbon recovery ranged from 11 – 80%. Excluding the lowest recovery value, the average recovery was 54%. This average is consistent with typical PPL recoveries for riverine DOC (Arellano et al., 2018; Raeke et al., 2016). The lowest recovery was for F-PW, which had

extremely high [DOC] but lower $SUVA_{254}$. Low recovery could be due to a predominance of polar, small molecules in this sample, which are poorly extracted by PPL resin (Raeke et al., 2016). Sulfur recovery calculations were not possible, because TDS measurements in PPL eluent dilutions were at or below instrument detection limits.

Our measurements based on PPL extracts are subject to a few biases. Although PPL sorbents perform generally well across a wide range of samples and chemical classes, PPL can suffer from incomplete elution of higher molecular weight and highly hydrophobic formulas, including low O/C S-containing formulas ($0.1 < O/C < 0.3$; Tfaily et al., 2012; Raeke et al., 2016), and does not extract highly polar molecules well (Minor et al., 2014; Raeke et al., 2016). Furthermore, our FT-ICR MS results only provide information for compounds that are ionizable in negative ESI mode and of a relatively higher molecular weight (> ~150 Da). Although our average PPL carbon recovery was on par with PPL recoveries for riverine DOC, recoveries suggest that a substantial fraction of the total DOM pool was not analyzed in this study.

SECTION 9. Δ^{34}S-DOM STABLE ISOTOPE VALIDATION

This study represents the first measurements of $\delta^{34}S$ on PPL-extracted DOM. Therefore, we tested for stable isotope fractionation during the PPL extraction steps. We extracted DOM from two low-sulfate (< 1 mg $S\text{-}SO_4^{2-}$ L^{-1}) surface water samples collected in the Florida Everglades. $\delta^{34}S$-DOM measurements were 15.0 ± 1.67‰ and 10.9 ± 1.31‰ (average ± standard deviation across sample replicate analyses) for the two PPL extracts, respectively. Both measurements are higher than the range we measured for Napa River Watershed PPL-extracted DOM (0.56 to 8.0‰) and are within the range previously measured for Everglades organic S (Bates et al., 1998). Everglades sample analyses suggest that our narrower range of values

measured in Napa Watershed samples reflect true variability, rather than an artifact of our DOM extraction method.

SECTION 10. $\Delta^{34}S$ RESULTS

Table S11. Inorganic sulfate ($\delta^{34}S\text{-}SO_4^{2-}$) and organic S ($\delta^{34}S$-DOS) isotopes (mean ± 1 SD across duplicate sample runs where available).

Sample Code	Land use/land cover	Sample Type	$\delta^{34}S\text{-}SO_4^{2-}$	$\delta^{34}S$-DOS
F-PW	forest/grassland	soil porewater	N/A	2.9 ± 1.3
HF F-RO	forest/grassland	surface water	-0.5 ± 0.1	4.7 ± 0.7
LF F-RO	forest/grassland	surface water	-1.7 ± 0.5	0.6 ± 1.8
G-PW	forest/grassland	soil porewater	7.5	8.0
G-RO	forest/grassland	surface water	-1.1 ± 2.8	5.2 ± 1.0
V-PW1	vineyard	soil porewater	7.1	4.8
V-PW2	vineyard	soil porewater	11.3 ± 0.4	7.3 ± 1.5
V-RO	vineyard	surface water	12.0 ± 1.8	5.5
M-RO1	mixed	surface water	0.0 ± 0.6	5.6 ± 1.5
M-RO2	mixed	surface water	4.7 ± 0.1	4.0 ± 0.3
HF N-RO	mixed	surface water	6.7 ± 0.1	5.2 ± 1.5
LF N-RO	mixed	surface water	12.0 ± 0.6	3.4 ± 1.1

SECTION 11. PRINCIPAL COMPONENT ANALYSIS

We used principal component analysis (PCA) to synthesize bulk, molecular, S XANES, and S stable isotope datasets (Figure S20). Principal component analysis variables were centered and scaled. Vineyard soil porewater 1 and forest soil porewater samples were excluded from the analysis due to missing values. The PCA analysis showed overall separation between vineyard and forest and grassland samples. PC 1 explained 56.4% of dataset variance and was related to higher relative abundances of CHOS + CHNOS and CHNO formulas, oxidized S species, and

$\delta^{34}S$- SO_4^{2-} values and lower relative abundances of CHO formulas, reduced S species, m/z, $SUVA_{254}$, NOSC, and AI-mod. PC 2 explained an additional 17.8% of dataset variance and was related to higher DOC concentrations and lower SO_4^{2-} and atomic S/C. Vineyard and forest and grassland samples largely separated along PC 1, with the exception of G-PW, suggesting the separation in multivariate space was driven by S applications to vineyard areas. Mixed LULC streams and the Napa River at low and high flow grouped with either vineyard or forest and grassland samples: M-RO1 and HF N-RO grouped with forest and grassland samples, whereas M-RO2 and LF N-RO grouped with vineyard samples.

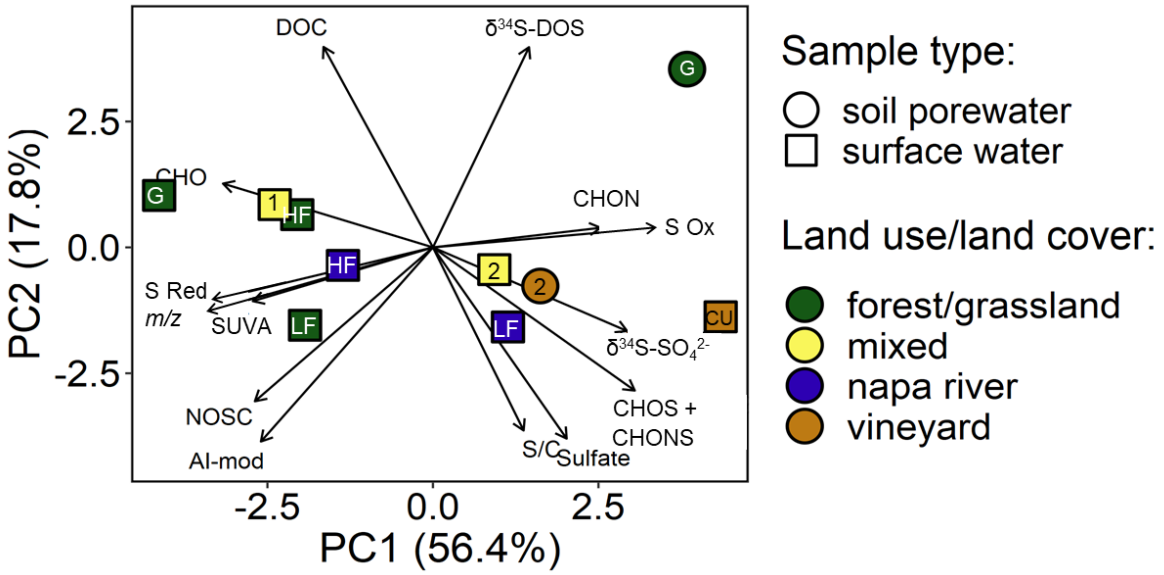


Figure S21. Principal component (PC) analysis synthesizing bulk chemistry, FT-ICR MS, XANES, and S isotope measurements. "S Red" = % reduced S; "S Ox" = % oxidized S; S/C = atomic S/C; "CU" = culvert runoff.

REFERENCES

Arellano, A.R., Bianchi, T.S., Hutchings, J.A., Shields, M.R., Cui, X. (2018). Differential effects of solid-phase extraction resins on the measurement of dissolved lignin-phenols and organic matter composition in natural waters. *Limnol. Oceanogr. Methods* 16, 22–34. https://doi.org/10.1002/lom3.10224

Arguez, A., Durre, I., Applequist, S., Vose, R. S., Squires, M. F., Yin, X., et al. (2012). NOAA's 1981–2010 U.S. Climate Normals: An Overview. *Bull. Am. Meteorol. Soc.* 93, 1687–1697. doi:10.1175/BAMS-D-11-00197.1.

Bahureksa, W.; Tfaily, M. M.; Boiteau, R. M.; Young, R. B.; Logan, M. N.; McKenna, A. M.; Borch, T., Soil Organic Matter Characterization by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR MS): A Critical Review of Sample Preparation, Analysis, and Data Interpretation. *Environ Sci Technol* **2021,** *55*, (14), 9637-9656.

Bates, A.L., Spiker, E.C., Holmes, C.W. (1998). Speciation and isotopic composition of sedimentary sulfur in the. *Chem. Geol.* 146, 155–170.

Blakney, G. T., Hendrickson, C. L., Marshall, A. G. (2011). Predator data station: A fast data acquisition system for advanced FT-ICR MS experiments. *Int. J. Mass Spectrom.* 306(2-3), 246-252.

Boldin, I. A., Nikolaev, E. N. (2011). Fourier transform ion cyclotron resonance cell with dynamic harmonization of the electric field in the whole volume by shaping of the excitation and detection electrode assembly. *Rapid Commun. Mass Spectrom.* 25(1), 122-126.

Carmody, R. W., Plummer, L. N., Busenberg, E., and Coplen, T. B. (1998). Methods for collection of dissolved sulfate and sulfide and analysis of their sulfur isotopic composition. U.S. Geological Survey Open File Report 97-234.

Chen, T., Beu, S. C., Kaiser, N. K., Hendrickson, C. L. (2014). Note: Optimized circuit for excitation and detection with one pair of electrodes for improved Fourier transform ion cyclotron resonance mass spectrometry. *Rev. Sci. Instrum.* 85(6), 0666107/1-066107/3.

Dittmar, T., Koch, B., Hertkorn, N., Kattner, G. (2008). A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Ocean. Methods* 6, 230–235. https://doi.org/10.4319/lom.2008.6.230

Emmett, M. R., White, F. M., Hendrickson, C. L., Shi, S. D.-H., Marshall, A. G. (1998). Application of micro-electrospray liquid chromatography techniques to FT-ICR MS to enable high-sensitivity biological analysis. *J. Am. Soc. Mass Spectrom.* 9(4), 333-340.

Graymer, R.W., Brabb, E.E., Jones, D.L., Barnes, J., Nicholson, R.S., Stamski, R.E. (2007). Geologic map and map database of eastern Sonoma and Western Napa Counties, California. U.S. Geol. Surv. Sci. Investig. Map 2956. Available at: https://pubs.usgs.gov/sim/2007/2956/. [Accessed 10.6.21].

Hinckley, E.-L. S., Kendall, C., and Loague, K. (2008). Not all water becomes wine: Sulfur inputs as an opportune tracer of hydrochemical losses from vineyards. *Water Resour. Res.* 44, 1–14. doi:10.1029/2007WR006672.

Hinckley, E. L. S., Crawford, J. T., Fakhraei, H., and Driscoll, C. T. (2020). A shift in sulfur-cycle manipulation from atmospheric emissions to agricultural additions. *Nat. Geosci.*, 1–8. doi:10.1038/s41561-020-0620-3.

Hughey, C. A., Hendrickson, C. L., Rodgers, R. P., Marshall, A. G., Qian, K. (2001). Kendrick mass defect spectroscopy: A compact visual analysis for ultrahigh-resolution broadband mass Spectra. *Anal. Chem.* 73, 4676-4681.

Kaiser, N. K., McKenna, A. M., Savory, J. J., Hendrickson, C. L., Marshall, A. G. (2013). Tailored ion radius distribution for increased dynamic range in FT-ICR mass analysis of complex mixtures. *Anal. Chem.* 85(1), 265-272.

Kendrick, E. (1963). A mass scale based on CH2 = 14.0000 for high resolution mass spectrometry of organic compounds. *Anal. Chem.* 35(13), 2146-2154.

Kind, T. and Fiehn, O. (2007). Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* 8, 105. doi:10.1186/1471-2105-8-105.

Koch, B.P., Dittmar, T. (2006). From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 20, 926–932. https://doi.org/10.1002/RCM.2386

Koch, B. P., Dittmar, T., Witt, M., and Kattner, G. (2007). Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Anal. Chem.* 79, 1758–1763. doi:10.1021/ac061949s.

Kroll, J.H., Donahue, N.M., Jimenez, J.L., Kessler, S.H., Canagaratna, M.R., Wilson, K.R., Altieri, K.E., Mazzoleni, L.R., Wozniak, A.S., Bluhm, H., Mysak, E.R., Smith, J.D., Kolb, C.E., Worsnop, D.R. (2011). Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol. *Nat. Chem.* 3, 133–139. https://doi.org/10.1038/nchem.948

Lv, J., Zhang, S., Luo, L., and Cao, D. (2016). Solid-phase extraction-stepwise elution (SPE-SE) procedure for isolation of dissolved organic matter prior to ESI-FT-ICR-MS analysis. *Anal. Chim. Acta* 948, 55–61. doi:10.1016/J.ACA.2016.10.038.

Manceau, A., Nagy, K.L. (2012). Quantitative analysis of sulfur functional groups in natural organic matter by XANES spectroscopy. *Geochim. Cosmochim. Acta* 99, 206–223.

https://doi.org/10.1016/j.gca.2012.09.033
McLafferty, F. W., Turecek, F. (1993). *Interpretation of Mass Spectra*, 4th ed. University Science Books: Mill Valley, CA.
Minor, E. C., Swenson, M. M., Mattson, B. M., and Oyler, A. R. (2014). Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. *Environ. Sci. Process. Impacts* 16, 2064–2079. doi:10.1039/C4EM00062E.
O'Geen, A., Walkinshaw, M., Beaudette, D. (2017). SoilWeb: A multifaceted interface to soil survey information. *Soil Sci. Soc. Am. J.* 81, 853–862. https://doi.org/10.2136/sssaj2016.11.0386n
Pohlabeln, A. M. and Dittmar, T. (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.
Poulin, B.A., Ryan, J.N., Nagy, K.L., Stubbins, A., Dittmar, T., Orem, W., Krabbenhoft, D.P., Aiken, G.R. (2017). Spatial dependence of reduced sulfur in Everglades dissolved organic matter controlled by sulfate enrichment. *Environ. Sci. Technol.* 51, 3630–3639. https://doi.org/10.1021/acs.est.6b04142
Raeke, J., Lechtenfeld, O.J., Wagner, M., Herzsprung, P., Reemtsma, T. (2016). Selectivity of solid phase extraction of freshwater dissolved organic matter and its effect on ultrahigh resolution mass spectra. *Environ. Sci. Process. Impacts* 18, 918–927. https://doi.org/10.1039/C6EM00200E
Stücheli, P. E., Niggemann, J., and Schubert, C. J. (2018). Comparison of different solid phase extraction sorbents for the qualitative assessment of dissolved organic nitrogen in freshwater samples using FT-ICR-MS. *J. Limnol.* 77, 400–411. doi:10.4081/jlimnol.2018.1791.
Tfaily, M., Hodgkins, S., Podgorski, D. C., Chanton, J. P., and Cooper, W. T. (2012). Comparison of dialysis and solid-phase extraction for isolation and concentration of dissolved organic matter prior to Fourier transform ion cyclotron resonance mass spectrometry. *Analytical and Bioanalytical Chemistry*. 404(2), 447–457. doi:10.1007/s00216-012-6120-6.
U.S.G.S (2016a). The StreamStats Program, online at http://streamstats.usgs.gov. [Accessed 12/29/20].
U.S.G.S. (2016b). National Water Information System data available on the World Wide Web (USGS Water Data for the Nation). Available at: https://waterdata.usgs.gov. [Accessed 12/1/20].
Xie, L., Spiro, B., and Wei, G. (2016). Purification of $BaSO_4$ precipitate contaminated with organic matter for oxygen isotope measurements ($\delta^{18}O$ and $\Delta^{17}O$). *Rapid Commun. Mass Spectrom.* 30, 1727–1733. doi:10.1002/rcm.7610.
Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S.M., Case, A., Costello, C., Dewitz, J., Fry, J., Funk, M., Granneman, B., Liknes, G.C., Rigge, M., Xian, G. (2018). A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS J. Photogramm. Remote Sens.* 146, 108–123. https://doi.org/10.1016/j.isprsjprs.2018.09.006.