

Journal of Visualized Experiments

Single-throughput complementary high-resolution analytical techniques for characterizing complex natural organic matter mixtures

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE59035R2
Full Title:	Single-throughput complementary high-resolution analytical techniques for characterizing complex natural organic matter mixtures
Keywords:	Climate; greenhouse gas; complex mixture; Mass spectrometry; thermodynamic; nuclear magnetic resonance; ion cyclotron resonance
Corresponding Author:	Rachel Wilson Florida State University Tallahassee, Florida UNITED STATES
Corresponding Author's Institution:	Florida State University
Corresponding Author E-Mail:	rmwilson@fsu.edu
Order of Authors:	Malak Tfaily Rachel Wilson Heather Brewer Rosalie Chu Heino Heyman David Hoyt Jennifer Kyle Samuel Purvine
Additional Information:	
Question	Response
Please indicate whether this article will be Standard Access or Open Access.	Standard Access (US\$2,400)
Please indicate the city, state/province, and country where this article will be filmed . Please do not use abbreviations.	Richland, WA, USA



THE FLORIDA STATE UNIVERSITY

Earth, Ocean, and Atmospheric Science • Tallahassee, Florida 32306-4320 U.S.A. • (850) 645-1887 • FAX (850) 644-2581

Rachel Marie Wilson, Ph.D.

EMAIL: rmwilson@fsu.edu

Dr. Lyndsay Troyer
Science Editor, *Journal of Visualized Experiments*

09 October 2018

Dear Lyndsay,

Please find enclosed our revised manuscript titled “Single-throughput complementary high-resolution analytical techniques for characterizing complex natural organic matter mixtures” enclosed for consideration in the *Journal of Visualized Experiments*. We appreciate the thoughtful and thorough reviews provided and have substantially revised the manuscript in response to the comments we received. We have added specific details in a number of places throughout the protocol as requested, as well as clarified and condensed several of the steps. We added a panel to figure 2 in response to the request of one of the reviewers to demonstrate changes in specific metabolite classes with depth and added a discussion of the changes in NOSC. In addition, we added text to the introduction to improve the flow from the abstract into the system we use for demonstration purposes (as suggested by one of the reviewers) and added a paragraph outlining the primary goals of the protocol as requested by the editor. In short, we have addressed all of the questions, comments, and suggestions of the editor and reviewers and have revised the manuscript accordingly. We provide a line-by-line response explaining how each of these changes was made in response to the comments we received and we hope that you will find the revised manuscript much improved and suitable for publication.

The results and conclusions of this work have not been submitted elsewhere. All authors have approved the revised submission and declare no conflicting interests.

Thank you for your consideration and we look forward to hearing from you regarding this manuscript.

Sincerely,

Rachel M. Wilson
Earth Ocean and Atmospheric Sciences,
Florida State University
Tallahassee, FL 32306, USA,
Phone number: (850) 645-1887,
email address: rrmwilson@fsu.edu

and

Malak M. Tfaily
University of Arizona
Tucson, AZ, USA
850-274-9385
malak.m.tfaily@gmail.com

TITLE:

Single-throughput Complementary High-resolution Analytical Techniques for Characterizing Complex Natural Organic Matter Mixtures

AUTHORS AND AFFILIATIONS:

Malak M. Tfaily^{1,2*}, Rachel M. Wilson^{3*}, Heather M. Brewer¹, Rosalie K. Chu¹, Heino M. Heyman⁴, David W. Hoyt¹, Jennifer E. Kyle⁵, Samuel O. Purvine¹

¹Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA

²Department of Soil, Water and Environmental Science, University of Arizona, Tucson, AZ, USA

³Department of Earth Ocean and Atmospheric Sciences, Florida State University, Tallahassee, FL, USA

⁴Bruker Daltonics Inc., Billerica, MA, USA

⁵Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA

*These authors contributed equally.

Corresponding Authors:

Malak M. Tfaily (malak.m.tfaily@gmail.com)

Rachel M. Wilson (rmwilson@fsu.edu)

Email Addresses of Co-authors:

David Hoyt (David.Hoyt@pnnl.gov)

Heather Brewer (Heather.Brewer@pnnl.gov)

Heino Heyman (Heino.Heyman@bruker.com)

Jennifer Kyle (Jennifer.Kyle@pnnl.gov)

Rosalie Chu (Rosalie.Chu@pnnl.gov)

Samuel Purvine (Samuel.Purvine@pnnl.gov)

KEYWORDS:

Fourier transform ion cyclotron resonance mass spectrometry, high-resolution analytical techniques, peatlands, dissolved organic matter, microbial decomposition, environmental metabolomics, proteomics

SUMMARY:

This protocol describes a single throughput for complementary analytical and omics techniques culminating in a fully-paired characterization of natural organic matter and microbial proteomics in different ecosystems. This approach permits robust comparisons for identifying metabolic pathways and transformations important for describing greenhouse gas production and predicting responses to environmental change.

ABSTRACT:

Natural organic matter (NOM) is composed of a highly complex mixture of thousands of organic compounds which, historically, proved difficult to characterize. However, to understand the thermodynamic and kinetic controls on greenhouse gas (carbon dioxide [CO₂] and methane [CH₄]) production resulting from the decomposition of NOM, a molecular-level characterization coupled with microbial proteome analyses is necessary. Further, climate and environmental changes are expected to perturb natural ecosystems, potentially upsetting complex interactions that influence both the supply of organic matter substrates and the microorganisms performing the transformations. A detailed molecular characterization of the organic matter, microbial proteomics, and the pathways and transformations by which organic matter is decomposed will be necessary to predict the direction and magnitude of the effects of environmental changes. This article describes a methodological throughput for comprehensive metabolite characterization in a single sample by direct injection Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), gas chromatography mass spectrometry (GC-MS), nuclear magnetic resonance (NMR) spectroscopy, liquid chromatography mass spectrometry (LC-MS), and proteomics analysis. This approach results in a fully-paired dataset which improves statistical confidence for inferring pathways of organic matter decomposition, the resulting CO₂ and CH₄ production rates, and their responses to environmental perturbation. Herein we present results of applying this method to NOM samples collected from peatlands; however, the protocol is applicable to any NOM sample (*e.g.*, peat, forested soils, marine sediments, *etc.*).

INTRODUCTION:

Globally, wetlands are estimated to contain 529 Pg of carbon (C), mostly as organic C buried in peat deposits¹. Currently, such peatlands act as a net C sink, sequestering 29 Tg C y⁻¹ in North America alone¹. However, environmental disturbance such as draining, fires, drought, and warmer temperatures can offset this C sink by increasing organic matter decomposition resulting in increased C losses *via* greenhouse gas (carbon dioxide [CO₂] and methane [CH₄]) production^{1,2}. Climate change may contribute to C loss if warmer temperatures or dryer conditions stimulate faster C decomposition by microorganisms. Alternatively, higher temperatures and air CO₂ concentrations may stimulate primary production to sequester more CO₂ as organic carbon (OC). To what extent and how fast that OC is then decomposed into CO₂ and CH₄ depends on the complex interactions between the electron donor substrates, the availability of electron acceptors, and the microorganisms that mediate the transformation. In many cases, the mechanisms are not well-characterized, thus their response to environmental perturbations is not well-constrained and it remains unclear what the net result of climate change will be on carbon balance in peatland ecosystems.

The complex nature of natural organic matter (NOM) has made even identifying the organic compounds present in the NOM mixtures historically difficult. Recent advances have greatly improved our ability to characterize compounds that traditionally and, to some extent continue to be, regarded as recalcitrant humic or fulvic compounds³⁻⁵. We now understand that many of these compounds are actually microbially available and may be decomposed if a suitable terminal electron acceptor (TEA) is made available^{6,7}. Calculating the nominal oxidation state of the carbon (NOSC) for a compound provides a metric for predicting the potential for decomposition and the energy yield of the TEA required. However, it requires a molecular-level characterization of the

organic matter⁷. NOSC is calculated from the molecular formula *via* the following equation⁷:
$$\text{NOSC} = -((-z + 4(\#C) + (\#H) - 3(\#N) - 2(\#O) + 5(\#P) - 2(\#S)) / (\#C)) + 4$$
, where *z* is the net charge.
NOSC is correlated with the thermodynamic driving force⁸, wherein compounds with higher NOSC are easier to degrade, while compounds with lower NOSC require increasingly energetic TEAs in order to be reduced. Compounds with NOSC less than -2 require a high energy yielding TEAs such as O₂, nitrate or Mn^{IV}, and cannot be degraded by commonly occurring lower energy yielding TEAs such as Fe^{III} or sulfate⁷. This is an important consideration in the waterlogged anoxic conditions found in wetlands where O₂ and other high energy yielding TEAs are scarce⁹ and therefore the degradation of lower NOSC compounds under these conditions are thermodynamically limited. Environmental perturbation can influence the thermodynamic state of the ecosystem through hydrologic changes that influence O₂ (the most energetic electron acceptor), changes in organic substrates and electron acceptors made available by primary production, and to a smaller extent by temperature. An important example of the temperature effects in wetland systems occurs with regard to the trade-off that occurs between homoacetogenesis (*i.e.*, acetate production from CO₂ and H₂) and hydrogenotrophic methanogenesis (*i.e.*, CH₄ production from CO₂ and H₂). At low temperatures it appears that homoacetogenesis is slightly favored, while warmer temperatures favor CH₄ production¹⁰. This temperature effect may have important implications for the response of ecosystems to changing climate, as CH₄ is a much stronger greenhouse gas than CO₂¹¹ and thus increasing production of CH₄ at the expense of CO₂ at warmer temperatures may contribute to a positive feedback with climate warming.

Peatlands produce globally significant quantities of CO₂ and CH₄⁶ *via* microbial respiration of naturally occurring organic matter. The NOSC of the organic carbon substrates determines the relative proportion of CO₂:CH₄ produced which is a critical parameter because of the higher radiative forcing of CH₄ compared to CO₂¹¹, but also because modeling efforts have identified this ratio as a critical parameter for estimating C flux in peatlands¹². In the absence of terminal electron acceptors other than CO₂, it can be shown by electron balance that organic C substrates with NOSC > 0 will produce CO₂:CH₄ > 1, organic C with NOSC = 0 produces CO₂ and CH₄ in equimolar ratio, and organic C with NOSC < 1 will produce CO₂:CH₄ < 1¹³. Decomposition of OC in natural ecosystems is mediated by microorganisms, so that even when degradation of a specific compound is thermodynamically feasible, it is kinetically limited by the activity of microbial enzymes and, under anoxic conditions, by the thermodynamic driving force (*i.e.*, NOSC)⁷. Until now it has been challenging to fully characterize the organic matter because the diversity of compounds present requires different complementary techniques for their characterization. Recent advances have closed the gap; using a suite of analytical techniques we can analyze a large range of organic compounds providing molecular-level characterization and, in some cases quantification, from small primary metabolites like glucose up to 800 Da poly-heterocycles. Previously such large complex molecules would have been characterized simply as lignin-like or tannin-like and assumed to have been recalcitrant. Molecular-level characterization, however, allows the calculation of NOSC for even these large complex molecules. These NOSC values are linearly correlated with the thermodynamic driving force allowing for an assessment of the quality of organic matter available for decomposition, which in many cases reveals that these

complex molecules may actually be microbially degradable even under the anoxic conditions that prevail in wetlands.

Since introduction of O₂ allows organic matter of nearly all naturally observed NOSC values to be decomposed, herein we focus on changes in the organic matter and microbial proteomics which are likely to be the primary drivers in wetland (*i.e.*, limited O₂) systems. However, all the techniques that we will discuss can be applied to organic matter from any ecosystem. Commonly, bulk measurements based on optical and fluorescence analyses have been used to assess organic matter quality^{3,14}. When using bulk measurements such as these, however, fine details are lost as large numbers of molecules are categorized together under generic terms like humics or fulvics. The definitions of these categories are not well-constrained and, in fact, may vary from study to study making comparisons impossible. Further, bulk measurements do not provide the molecular detail necessary for calculating the thermodynamics governing the system and therefore fall short of truly assessing organic matter quality¹⁵.

Individual techniques such as Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), nuclear magnetic resonance (NMR) spectroscopy, gas chromatography mass spectrometry (GC-MS), and liquid chromatography mass spectrometry (LC-MS) do provide such molecular-level detail. While each of these techniques presents its own limitations, they also bring their own strengths that can be leveraged in an integrated approach to achieve the fine molecular detail necessary for quantifying organic matter quality in a rigorous thermodynamic sense. GC-MS is useful for identifying critical small metabolites that are likely to have proximal influence on CO₂ and CH₄ production (*e.g.*, glucose, acetate, *etc.*); however, GC-MS requires verification against a standard and is therefore limited to already known compounds present in the database preventing identification of novel compounds. Furthermore, GC-MS is a semi-qualitative technique allowing inference about the changes in relative concentrations, but not providing the actual concentration information necessary for calculating Gibb's free energies for example. Finally, GC-MS requires derivatization of molecules prior to analysis which limits resolution to compounds smaller than ~400 Da and volatile alcohols are lost during the drying step.

One-dimensional (1D) ¹H liquid state NMR allows highly quantitative characterization of small metabolites (including primary small molecular weight metabolites and volatiles like alcohols, acetate, acetone, formate, pyruvate, succinate, short-chained fatty acids, as well as a range of carbohydrates notoriously absent or compromised from MS-based methods) and their concentrations are particularly useful for calculating thermodynamic parameters. Yet, like GC-MS, 1D NMR of complex mixtures requires standardization relative to a database and therefore does not alone allow easy identification of novel compounds that are likely to be abundant in complex natural and changing ecosystems. Additionally, NMR is less sensitive than the MS-based techniques and therefore, quantitative metabolite profiling is achieved only above 1 μM using NMR systems equipped with helium-cooled cold-probes. Not widely appreciated, some NMR cold-probes are salt-tolerant and allow environmental mixture analysis in the presence of millimolar salt concentrations when used in smaller diameter (< 3 mm outer diameter) sample tubes¹⁶. However, a further complication of NMR is that high amounts of paramagnetic metals

and minerals (*e.g.*, Fe and Mn above 1 - 3 wt%), which can be abundant in upland soils, can broaden spectral features and complicate interpretation of the NMR spectra. Using solid phase extraction (SPE) can aide in the interpretation of both NMR and MS-based metabolomics methods by reducing the mineral salts and increasing spectral quality.

FTICR-MS by direct injection is a highly sensitive technique capable of detecting tens of thousands of metabolites from a single sample, but it does not capture the critical small metabolites such as acetate, pyruvate, and succinate and is notoriously difficult to use for sugars and other carbohydrates¹⁷, nor does it provide quantitative information. However, unlike the other techniques, FTICR-MS excels at identifying and assigning molecular formula to novel compounds and therefore identifies the largest number of compounds providing more molecular information than any of the other described techniques. This is useful, because the molecular information provided by FTICR-MS (and other techniques) can be used to calculate NOSC which is related to the thermodynamic driving force governing the likelihood of certain reactions⁸ and their rates under certain conditions⁷. Furthermore, by coupling FTICR-MS with separation techniques, such as LC together with tandem MS, quantitative structural information can be attained, offsetting some of the disadvantages of this technique. LC-MS is useful for identifying lipid-like compounds and other metabolites that are not well-characterized by any of the other methods. Coupling LC FTICR-MS or LC-MS with a fraction collector and collecting fractions of specific unknowns of interest for structural elucidation by two-dimensional (2D) liquid state NMR is the ideal situation for identifying and quantifying unknown compounds^{18,19}. However, this is a very time-consuming step that could be used if and when needed. Taken individually, each of these techniques provide a different snapshot of the organic matter, and by integrating them, we can achieve a more complete understanding than using any of the techniques in isolation.

While the thermodynamic considerations set the ultimate constraints on what transformations are possible in a system, organic matter decomposition is mediated by microorganisms whose enzyme activities control reaction rates. Thus, fully understanding the controls on organic matter decomposition and ultimately greenhouse gas (CO₂ and CH₄) production from wetlands requires an integrated omics approach to characterizing the microbial enzyme activities as well as the metabolites. In this article, we describe a method for achieving such a comprehensive analysis from a single sample using a sequential approach that results in a fully paired analysis. This approach expands on the metabolite, protein, and lipid extraction (MPlex) protocol in which proteomics was coupled with GC-MS and LC-MS²⁰ to identify small metabolites, proteins, and lipids by incorporating quantitative metabolite information *via* NMR and identification of larger secondary metabolites *via* FTICR-MS. Slightly different to MPlex, we begin the protocol with a water extraction and then use sequential extraction with increasingly non-polar solvents. All extractions are done on a single sample which conserves sample when volumes are limited or difficult to obtain and decreases experimental error introduced through variation among aliquots from heterogeneous sample matrices (*e.g.*, soil and peat) or differences in storage conditions and duration.

Finally, by coupling the OM analyses with proteomics analyses of the microbial community, we can build metabolic networks that describe the pathways and transformation of organic matter

decomposition. This allows us to test specific hypotheses about how perturbations to the system will influence ultimate CO₂ and CH₄ production through alteration to the available organic substrates, electron acceptors, and the microbial communities mediating the reactions *via* the activity of enzyme catalysts.

The overall goal of this method is to provide a single throughput protocol for analyzing metabolites, lipids, and microbial proteins from a single sample thereby creating a fully paired dataset for building metabolic networks while constraining analytical errors.

PROTOCOL:

1. Sequential Extraction of Organic Matter from Soil, Sediments, or Peat

1.1. Collect soil, sediments, or peat *via* coring and divide cores according to the hypothesis being tested (*e.g.*, depth). Store samples in polytetrafluoroethylene coated containers and freeze at -80 °C for storage prior to analysis.

NOTE: Approximately 45 mg C is needed for this protocol. For peat (typically 45% C), 100 mg of dried peat is required. Larger amounts of sample may be needed for low organic samples like mineral or forested uplands soils depending on the C content. Because extraction with organic solvents will pull any polyethylene glycol (PEG) into the extracts, which will negatively affect ionization during step 2.4, it is important to avoid allowing samples to contact plastic at any point during collection, storage or extraction.

1.2. When ready to analyze the samples, freeze dry to constant weight, then grind the samples in a high-speed ball mill using stainless steel grinding balls to homogenize and break up any aggregates.

NOTE: The protocol can be paused at this point and the material stored at -80 °C.

1.3. Using an ethanol-washed stainless-steel utensil, aliquot 50 mg of each of the dried samples into individual 2 mL glass vials. These samples will be sequentially extracted using a series of solvents to consecutively extract increasingly non-polar metabolites from each sample. Add 1 mL of distilled, degassed water (H₂O) to each sample, cap the vials and shake for 2 h on a shaker table.

1.4. After shaking allow the solutions to stand at room temperature (RT) for 20 min, then centrifuge at 15,000 x g for 30 min, decant and save the supernatant from each.

NOTE: These solutions will be injected into the FTICR-MS by direct injection.

1.5. Conduct a Folch extraction²¹ (also known as MPLEx²⁰) on the now water extracted residues by repeating steps 1.3 and 1.4 substituting 1 mL of a -20 °C 2:1 chloroform:methanol mixture for the water in step 1.3.

CAUTION: Chloroform and methanol are both highly flammable and toxic. Use appropriate personal protective equipment (PPE) to avoid skin contact and avoid open flame.

1.6. Carefully separate the two resulting solvent layers, which will be visually distinguishable, for separate analysis by FTICR-MS by using a separation funnel or simply remove the top layer by careful pipetting.

NOTE: The chloroform-containing fraction will be at the bottom while the methanol-containing fraction is less dense and will be on the top.

1.7. Dilute both the chloroform extract (step 1.6) and water extract (step 1.4) 2:1 in methanol to improve the electrospray ionization (ESI) efficiency for FTICR-MS by direct injection.

NOTE: The methanol-containing fraction from step 1.7 does not need to be diluted further in methanol. The methanol layer will be run by direct injection on the FTICR-MS.

2. FTICR-MS Analysis

2.1. 2.1. Calibrate the FTICR spectrometer by directly injecting 100 μL of a tuning solution (see Table of Materials) spanning a mass range of approximately 100 - 1300 Da into the FTICR-MS.

2.2. Prepare the Suwannee River Fulvic Acid standard (see **Table of Materials**) by making a 1 mg mL^{-1} solution in ultrapure filtered water and then diluting the resulting solution to 20 $\mu\text{g mL}^{-1}$ in methanol.

2.3. Direct inject 23 μL of this final solution to the ESI source coupled to the FTICR spectrometer through a syringe pump set to a flow rate of 3.0 $\mu\text{L min}^{-1}$. Set needle voltage to +4.4 kV, Q1 to 150 m/z and glass capillary at 180 $^{\circ}\text{C}$. Inspect resulting spectra using the analysis software (see **Table of Materials**) to confirm the quality of the data.

2.4. Use HPLC grade methanol before running samples and throughout the sampling to monitor carryover. Introduce 23 μL of each extract *via* direct injection to the ESI source coupled to the FTICR spectrometer through a syringe pump set to a flow rate of 3.0 $\mu\text{L min}^{-1}$. Set needle voltage to +4.4 kV, Q1 to 150 m/z and glass capillary at 180 $^{\circ}\text{C}$.

2.5. Adjust ion accumulation time (IAT) for each sample or group of samples to account for variation in C concentration. Typical values are between 0.1 to 0.3 s. Collect 144 scans for each sample, average the scans and then conduct an internal calibration using homologous CH_2 (*i.e.*, 14 Da separation) series.

NOTE: After FTICR-MS analysis, the decision can be made to either combine the water and methanol extracted fractions in order to streamline the remaining steps or the fractions can be kept separate throughout subsequent steps. The advantages and disadvantages of each

approach are described at length in the **Discussion**. If doing so, combine the water and methanol-extracted fractions.

2.6. Dry extracts using a concentrator and save the remainder of the extracts (~ 1 mL) for subsequent GC-MS (water, methanol or water + methanol), LC-MS (chloroform) and NMR (water + methanol) analysis.

NOTE: The protocol can be paused at this point and the material stored at -20 °C.

3. FTICR-MS Data Processing

3.1. Co-align all sample peak lists for the entire dataset to reduce mass shifts and standardize peak assignments using Formularity software²² ahead of formula assignment.

3.2. Use the Formularity software²² to assign molecular formula using a signal to noise ratio greater than 7, mass measurement error < 1 ppm, and allowing C, H, O, N, S, and P while excluding all other elements.

3.3. If multiple candidate formulas are returned for a given mass (frequent above 500 Da) impose constraints consistent with the material being sampled. For example, in peat, typical constraints include: lowest mass error, lowest number of heteroatoms (N, S, P), and when present, P must be in oxidized form (*i.e.*, there must be at least 4 O atoms for each P atom in the formula).

4. Chemical Derivatization for GC-MS

4.1. Prepare blank control samples of HPLC-grade hexane in GC-MS autosampler vials²³. Dissolve 100 mg fatty acid methyl esters (FAMES: C8 - C28) mixture retention time standard in 200 µL of hexane.

4.2. To protect carbonyl groups, add 20 µL of 30 mg mL⁻¹ methoxyamine hydrochloride in pyridine to each of the methanol extracts and water extracts (or combined methanol/water extracts if using) from step 2.6, blanks and FAME calibration samples²³. Seal vials with caps.

CAUTION: Pyridine is both highly toxic and flammable. Wear appropriate PPE to prevent skin or eye contact and avoid open flame. In addition, pyridine volatilizes at RT causing harmful air contamination. Work only in well-ventilated spaces under a fume hood.

4.3. Vortex extracts for 20 s. Sonicate extracts for 60 s. Then, incubate extracts in a centrifuge at 37 °C for 90 min at 100 x g.

NOTE: An excessive amount of carbohydrates or salts can cause metabolites to crystalize after being dried down. Sonicating the samples will help to reconstitute the crystalized metabolites into the derivatizing reagent.

4.4. After incubation, add 80 μL of N-methyl-N-(trimethylsilyl)trifluoroacetamide with 1% trimethylchlorosilane to each sample, vortex extracts for 20 s. Sonicate extracts for 60 s and incubate extracts again in a centrifuge at 37 $^{\circ}\text{C}$ for 30 min at 100 x g.

4.5. Cool extracts to RT (20 - 24 $^{\circ}\text{C}$), then transfer into GC-MS autosampler vials.

5. GC-MS Analysis

5.1. Tune and calibrate MS according to vendor recommendations (see **Table of Materials**) before analysis to make sure the machine read MS data correctly. Check that helium gas pressure is within specified tolerance.

5.2. Separate polar metabolites on a GC column (30 m \times 0.25 mm \times 0.25 μm). Set the oven temperature protocol as follows: (1) initial temperature of 60 $^{\circ}\text{C}$ for 1 min, (2) ramp to 325 $^{\circ}\text{C}$ at a rate of 10 $^{\circ}\text{C min}^{-1}$, and (3) hold at 325 $^{\circ}\text{C}$ for 5 min.

5.3. Analyze extracts using a GC coupled to a single quadrupole MS. Set injection port temperature to a constant 250 $^{\circ}\text{C}$. Inject 1 μL of each derivatized extract in the splitless mode.

6. GC-MS Data Processing

6.1. Inspect all the data files to ensure that they were correctly captured. Pay attention to potential shifts with regards to internal standard retention times and intensities to confirm that the data was captured consistently throughout the analysis.

6.2. Convert the vendor specific MS data format to a general MS format if required. Process raw data files using MetaboliteDetector²⁴ calibrating retention indices based on the FAME internal standards. After aligning the retention times of all data files, continue with deconvolution and finally metabolite identification by matching retention indices and GC-MS spectra against the FiehnLib polar metabolite library²⁵.

6.3. Cross-check remaining unidentified metabolites against the NIST14 GC-MS library using spectral matching. Validate identifications individually to eliminate false identifications and reduce deconvolution errors.

7. Liquid State NMR Analysis

7.1. Dilute the remainder of the water extracts (\sim 300 μL) by 10% (vol/vol) with a 5 mM 2,2-dimethyl-2-silapentane-5-sulfonate- d_6 internal standard. Alternatively, combine water and methanol extracts from step 1 then resubstitute in water. By doing so however, some of the volatile compounds may be lost during the freeze-drying step. Typical final sample volumes are in the range 180 - 300 μL .

7.2. Transfer the mixture into a high-quality 3 mm outer diameter (O.D.) borosilicate glass NMR tube.

7.3. Collect spectra using an NMR spectrometer (ideally at least 600 MHz) equipped with a 5 mm triple resonance salt-tolerant cold probe and a cold-carbon pre-amplifier.

7.4. Collect 1D ^1H spectra using a 1D nuclear Overhauser effect spectroscopy (NOESY) presaturation experiment at 298 K with a 4 s acquisition time, 1.5 s recycle delay, 65,536 complex points, and 512 scans.

7.5. Collect 2D ^1H - ^{13}C heteronuclear single-quantum correlation (HSQC) and ^1H - ^1H total correlation (TOCSY) spectra to help in assigning metabolites and validation.

7.6. Process, assign, and analyze all spectra using the NMR analysis software (see **Table of Materials**) to quantify intensities relative to the internal standard. Identify metabolites by matching chemical shift, J-coupling and intensity information in the samples against the library.

NOTE: The library is further enhanced by custom targeted metabolite standards and metabolite additions made on a regular basis. The protocol can be paused at this point and the material stored at $-20\text{ }^\circ\text{C}$.

8. LC-MS Lipidomics Analysis

8.1. Rewet the dried chloroform extract generated during step 2.5 with 200 μL of methanol.

8.2. Inject 10 μL of each extract into an ultra-performance liquid chromatograph coupled to an Orbitrap mass spectrometer using a reversed phase charged surface hybrid column (3.0 mm \times 150 mm \times 1.7 μm particle size). Set a 34-min gradient (mobile phase A: acetonitrile/ H_2O (40:60) containing 10 mM ammonium acetate; mobile phase B: acetonitrile/isopropanol (10:90) containing 10 mM ammonium acetate) at a flow rate of 250 $\mu\text{L}/\text{min}$. Use both negative and positive ionization modes with higher-energy collision dissociation and collision induced dissociation.

CAUTION: Acetonitrile is an eye, skin, and respiratory irritant. Wear appropriate PPE. In addition, acetonitrile is combustible. Do not allow to contact oxidizing agents, toxic fumes may be produced during burning.

8.3. Upload the raw LC-MS/MS data files into LIQUID²⁶ along with the target file (*i.e.*, list of > 25,000 lipid species) for the respective ionization mode (positive or negative). Process the raw file. Manually validate the resulting identifications by examining the MS/MS spectra for presence of diagnostic ions, if applicable, matching fragment ions (e.g., fatty acyl chains), isotopic separation, mass ppm error of the precursors, and the retention time. Export resulting list of confidently identified lipids as a .tsv file.

NOTE: The protocol can be paused at this point and store the material at -20 °C.

9. Proteomics Analysis

9.1. Extract proteins according to the MPLEx protocol²⁰ from the remainder of the methanol phase resulting from step 2.4, by washing the extract with 20 times the extract volume of additional cold (-20 °C) methanol.

9.2. Use a 1 mL/50 mg silica-based sorbent (see **Table of Materials**) to condition C18 SPE columns with 3 mL methanol, 2 mL of 0.1% trifluoroacetic acid (TFA) in water, followed by the addition of the extract from step 9.1 at a rate no greater than 1 mL/min.

9.3. Following the addition of the sample, wash the column with 4 mL of 95:4.9:0.1 water:acetonitrile:TFA, then allow to dry. Place a 1.5 mL collection tube under the SPE column, and elute the sample with 1 mL of 80:19.9:0.1 methanol:water:TFA.

9.4. Concentrate the extracts to 100 µL under vacuum-assisted freeze dryer, then measure the protein concentration by bicinchoninic acid (BCA) colorimetric assay²⁷ at a wavelength of 562 nm.

9.5. Centrifuge extracts at 10,000 x g for 10 min at 4 °C. Discard the resulting supernatant and dry the remaining pellet under vacuum for 5 min. Resuspend the protein pellet in water to a final concentration of 0.1 µg peptide per µL.

9.6. Add dithiothreitol to a final concentration of 5 mM and incubate at 60 °C for 30 min. Dilute 10-fold with 100 mM NH₄HCO₃/8M urea solution and incubate at 37 °C for 3 h in the presence of 1 mM CaCl₂ and porcine trypsin at a 1:50 enzyme to protein ratio.

NOTE: Protocol may be paused at this point and store material at -20 °C.

9.7. Separate extracts *via* liquid chromatography using an exponential gradient of a 0.1% formic acid in water mobile phase (A) and a 0.1% formic acid in acetonitrile mobile phase (B) at 10 kpsi and 500 nL min⁻¹.

9.8. Introduce resulting eluent into an ESI-coupled mass spectrometer collecting spectra from 400-2000 *m/z* with 100,000 resolution at *m/z* 400 in a linear ion trap (LTQ) Orbitrap mass spectrometer.

9.9. Convert RAW spectra files to the mzML format using msConvert or ProteoWizard accepting all default parameters. Use the universal database search tool MSGFPlus to search resulting proteomes against a targeted protein database of protein sequences predicted from relevant metagenome assembled genomes.

9.10. Append common contaminants (*e.g.*, trypsin, keratin, albumin) and remove all exactly duplicated protein sequences to improve peptide-to-mass-spectrum match statistics. Evaluate the resulting MSGF spectral probability scores to determine which peptide-to-mass-spectrum match is best. Use the Q-Value from MSGFPlus to filter the entire data pile to 1% false discovery rate (FDR).

10. Metabolomics Analysis and Metabolic Network Building

10.1. Compile all metabolite molecular formula identified in steps 3, 6, 7, and 8 into a single database of metabolites present in the samples. Combine these metabolites with the Enzyme Commission (EC) or KEGG Orthology (KO) numbers of enzymes identified in step 9. Search this combined dataset against the KEGG²⁸ database, metabolic pathways section.

10.2. Manually evaluate results to identify most probable pathways and integrate into a metabolic model.

REPRESENTATIVE RESULTS:

We performed the described complementary analysis protocol and compared peat with depth in the S1 bog in the Spruce and Peatlands Response Under Changing Environments (SPRUCE) site in Minnesota, USA. These results are compared to those from a permafrost bog and fen from northern Sweden to show how sites may vary in metabolite and enzyme activities. We identified 3,312 enzymes in the proteomics analysis. An analysis of the enzymes activities with depth reveals that the number of enzymes declines sharply between 15 cm and 45 cm in the SPRUCE bog (**Figure 1**).

While the proteomics results indicate which proteins are expressed, the metabolic data show which reactions are actually occurring. Overall, we identified 67,040 metabolites (including lipids) in all of the peat samples from the combination of FTICR-MS, NMR, GC-MS, and LC-MS analyses. Of these, we were able to assign molecular formula to 15,385 compounds (**Figure 2**). The combined metabolic data spans a range of oxidation states, masses, and compound classes.

This is typically visualized through the use of a van Krevelen diagram in which the atomic H/C ratios of identified formula are plotted against their atomic O/C ratios²⁹. We have included an added dimension to the typical 2D format, depicting NOSC by means of color coding the symbols representing individual formula (**Figure 3 and Figure 4**). With increasing depth in the SPRUCE bog, there is an increase in the total number of large secondary metabolites identified by FTICR-MS, specifically in highly condensed (lower left of the van Krevelen plot) and hydrogenated formulae (top of the plot) (**Figure 4**). Over the same depths there is a decrease in the number of small highly energetic compounds identified by GC-MS and in the lipids (**Figure 5**). This could suggest either that the lipids and small metabolites are consumed in the surface peat before reaching the deeper depths or that decomposition rates in the deep peat are faster than in the surface such that downward advecting compounds are rapidly consumed. Differentiating between these two competing hypotheses requires a process-level understanding of the C cycling at the site. Such a process-level understanding can only be gained by coupling the metabolomics and proteomics

datasets. We accomplish this by cross-validating the combined FTICR-MS, GC-MS, LC-MS, and NMR identified metabolites against the KEGG database. In doing so, we find that the compounds we identified are involved in common metabolic pathways such as tricarboxylic acid (TCA) cycle, glycolysis, and sugar metabolism. Since individual metabolites can be involved in multiple pathways confirmation with the enzymes increases our confidence in assigning pathways (**Figure 6**).

Through this mapping we find evidence of sucrose and starch metabolism in the surface peat, while in the deeper depths pyruvate and other fermentation products build up (**Figure 6**). These results are consistent with the first hypothesis that sugars and other energetic metabolites are degraded in the surface peat and do not reach the deeper peat depths. As can be seen in **Figure 5**, sugars (NOSC = 0) and amino acids ($0 < \text{NOSC} < 1$) are consumed in the surface peat, while lipids ($-2 < \text{NOSC} < -1$) appear to accumulate with depth. This is consistent with expectations based on NOSC values that higher NOSC compounds are more readily degraded while lower NOSC compounds persist in the highly anaerobic (*i.e.*, TEA-limited) conditions of subsurface peat. This approach also successfully distinguishes between soil types from different environments. For example, soil organic matter from boreal peatland appears to be compositionally different than permafrost peatland, as well as within fen and bog within the permafrost region. These results are consistent with a previous study that showed differences in site geochemistry between these habitats³⁰, suggesting that site geochemistry have a big effect on microbial degradation of C below ground.

FIGURE LEGENDS:

Figure 1: Proteomics analysis indicates strong depth stratification in the number of enzymes. Bars indicate averages and error bars indicate one standard deviation of abundances. This suggests that microorganisms are most active in the peat surface.

Figure 2: The number of compounds identified by each technique in the surface peat (< 30 cm) of each habitat as well as the mid (45 cm) and deep (87 cm) peat from the SPRUCE bog.

Figure 3: The van Krevelen diagram (atomic H/C vs. O/C of each identified molecular formula) for the surface (15 cm) SPRUCE bog depth to demonstrate the coverage of compounds characterized by each technique. FTICR-MS allows us to identify the largest number of compounds (small circles), while GC-MS (triangles) is good for differentiating and identifying sugars ($\text{H/C} = 2$ and $\text{O/C} = 1$). NMR (squares) spectroscopy provides quantitative information on energetically important compounds such as sugars, amino acids, pyruvate, *etc.* LC-MS (diamonds) provides information on lipids.

Figure 4: The van Krevelen diagram (atomic H/C vs. O/C of each identified molecular formula) for the deep (87 cm) SPRUCE bog depth to demonstrate the coverage of compounds characterized by each technique.

Figure 5: The relative fraction of different chemical classes identified *via* the various techniques in the different depths. Bars are plotted as averages for each depth \pm one standard deviation. Classes plotted include amino acids, sphingolipids (Cer), glycerophosphocholines (PC), phosphoethanolamines (PE), diacylglyceroltrihomoserine (DGTSA), diacylglycerol (DG), triacylglycerols (TG), and sugars.

Figure 6: Combining results to create a metabolic network. Depth distributions of metabolites identified by NMR (a) and GC-MS (b), and a simplified metabolic map (c) showing a select number of transformations identified at SPRUCE bog. Green boxes indicate metabolites that decrease with depth, brown boxes indicate metabolites that increase with depth. Green connecting arrows indicate enzyme identified in our dataset that mediate the indicated transform (enzyme EC numbers indicated next to arrows). Gray connecting arrows indicate transforms for which enzymes were not identified in our dataset.

DISCUSSION:

The single-throughput, fully-coupled analysis stream used to characterize metabolites and the proteome provides insights into the pathways by which C cycling is occurring in a complex ecosystem. Soil and peat are heterogeneous matrices, and therefore, one of the critical steps of this method occurs in the earliest steps in ensuring that the starting peat or soil material is highly homogenous. It is preferable to grind the sample well as aggregates can reduce extraction efficiency. This is a particular problem for aggregated soils and soils with low C and high mineral contents that may require the use of a stainless-steel ball grinder to adequately homogenize. Because spatial heterogeneity of soils and peat within an experimental site may be high, biological replicates are highly recommended.

This method utilizes three solvents: water, methanol, and chloroform, which bias the types of compounds it is possible to extract. In principle, this sequential extraction method should solubilize compounds covering a wide polarity range. However, these solvents are not optimized for extracting organo-mineral complexes or highly stabilized complexes. If such compounds are of interest, harsher solvents such as strong acids and bases are preferred, but harsher extraction processes could alter the chemistry of the samples. Incorporating such solvents at the end of the extraction sequence may minimize this effect. Additionally, chloroform will inactivate clinically important soil and peat bacterial and viral pathogens, and other pathogenic disease-causing biological agents by dissolving cell membrane lipids. Thus, incorporating chloroform in the extraction protocol will reduce the risks involved in biology studies of samples potentially infected by pathogens from different regions of the world. If there are concerns about dead microbial cells, debris, or particulates following extraction, filtering the extracts through a 0.2 μ m glass fiber filter is recommended.

To streamline the process, the water and methanol extracts could be combined for GC-MS analysis during step 4.1. However, these extracts must remain separated for FTICR-MS analysis due to ionization efficiency issues with the ESI source. The advantage of running the combined extract on the GC-MS is that more metabolites (covering a larger range of polarity) will be identified. The disadvantage of combining the extracts at this point is that one of the major

metabolites important in microbial processing of organic matter is methanol. Traces of methanol will always remain in the combined sample, even after drying, so if methanol is a metabolite of interest, the water extract must be run separately. Having good controls with no analytes will also help in identifying potential contamination in the samples.

In step 7.1 of preparing extracts for NMR analysis, as in the GC-MS analysis, either the water extract alone or the combined water and methanol extracts can be used. The disadvantages of doing this are similar to those enumerated in the preceding paragraph. The advantage of combining the methanol and water extracts is that some of the less polar metabolites that are solubilized in the methanol fraction will be identified. However, because of the freeze-drying step, many more volatile compounds will be lost. This is a particular disadvantage if quantifying volatile fatty acids is a critical component of the experiment.

In this procedure for identifying proteomics, only fully tryptic peptides are searched, thus endogenous peptidase activity and in-source fragmentations will be missed. On the other hand, oxidized methionine may be considered as a post-translational modification for the peptide candidates as this modification commonly occurs during sampling processing and handling. Quantification of enzymes using peptide elution areas can be done, but is outside the scope of this project.

Many recent advances in the analysis of NOM and microbial parameters are providing a wealth of techniques for understanding organic C cycling. By combining these techniques in a single streamlined protocol, we gain a novel view of the processes that are taking place. Coupling the proteomics analysis with the metabolic analyses provides compelling corroborative evidence that a reaction is actually occurring. Metabolic analysis alone is limited in that it informs us which metabolites are higher or lower in concentration in comparisons among sites or after a treatment effect, but can't discern the causes for those changes. For example, we identified declining sugar concentrations with depth in the bog, but without further information it is unclear whether the decline with depth is due to slow inputs or fast degradation in the deep peat. Analysis of the proteome and the declining enzyme expression with depth allows us to reject the second hypothesis. Rather our results are consistent with fast sugar degradation in the surface peat limiting inputs to the deeper depths. These particular insights are only possible when all techniques are considered in tandem as each one provides a unique, but critical piece of the C cycling puzzle.

ACKNOWLEDGMENTS:

We would like to thank J.P. Chanton, J.E. Kostka, and M.M. Kolton for assistance with collecting peat samples. Portions of this work were conducted at the Environmental Molecular Sciences Laboratory, a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research. PNNL is operated by Battelle for the DOE under Contract DE-AC05-76RL01830. This work was supported by the U.S. Department of Energy, Office of Science, and Office of Biological and Environmental research (grants: DE-AC05-00OR22725, DE-SC0004632, DESC0010580, DE-SC0012088, and DE-SC0014416).

DISCLOSURES:

The authors have nothing to disclose.

REFERENCES:

1. Bridgham, S. D., Megonigal, P. J., Keller, J. K., Bliss, N. B., Trettin, C. The carbon balance of North American wetlands. *Wetlands*. **26** (4), 889-916 (2006).
2. Wilson, R. M., *et al.* Greenhouse gas balance over thaw-freeze cycles in discontinuous zone permafrost. *Journal of Geophysical Research: Biogeosciences*. **122** (2), 387-404 (2017).
3. Broder, T., Knorr, K. H., Biester, H. Changes in dissolved organic matter quality in a peatland and forest headwater stream as a function of seasonality and hydrologic conditions. *Hydrology and Earth System Sciences*. **21** (4), 2035-2051 (2017).
4. Ejarque, E., *et al.* Quality and reactivity of dissolved organic matter in a Mediterranean river across hydrological and spatial gradients. *Science of The Total Environment*. **599**, 1802-1812 (2017).
5. Valenzuela, E. I., *et al.* Anaerobic methane oxidation driven by microbial reduction of natural organic matter in a tropical wetland. *Applied and Environmental Microbiology*. **83** (11) e00645-e00617 (2017).
6. Lehmann, J., Kleber, M. The contentious nature of soil organic matter. *Nature*. **528** (7580), 60-68 (2015).
7. Keiluweit, M., Nico, P. S., Kleber, M., Fendorf, S. Are oxygen limitations under recognized regulators of organic carbon turnover in upland soils? *Biogeochemistry*. **127** (2-3), 157-171 (2016).
8. LaRowe, D. E., Van Cappellen, P. Degradation of natural organic matter: A thermodynamic analysis. *Geochimica et Cosmochimica Acta*. **75** (8), 2030-2042 (2011).
9. Wilson, R. M., *et al.* Hydrogenation of organic matter as a terminal electron sink sustains high CO₂: CH₄ production ratios during anaerobic decomposition. *Organic Geochemistry*. **112**, 22-32 (2017).
10. Ye, R., Jin, Q., Bohannan, B., Keller, J. K., Bridgham, S. D. Homoacetogenesis: A potentially underappreciated carbon pathway in peatlands. *Soil Biology and Biochemistry*. **68**, 385-391 (2014).
11. Neubauer, S. C., Megonigal, J. P. Moving beyond global warming potentials to quantify the climatic role of ecosystems. *Ecosystems*, **18** (6), 1000-1013. (2015).
12. Ma, S., *et al.* Data-Constrained Projections of Methane Fluxes in a Northern Minnesota Peatland in Response to Elevated CO₂ and Warming. *Journal of Geophysical Research: Biogeosciences*. **122** (11), 2841-2861 (2017).
13. Conrad, R. Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments. *Federation of European Microbiological Societies Microbiology Ecology*. **28** (3), 193-202 (1999).
14. Cunada, C. L., Lesack, L. F. W., Tank, S. E. Seasonal dynamics of dissolved methane in lakes of the Mackenzie Delta and the role of carbon substrate quality. *Journal of Geophysical Research: Biogeosciences*. **123** (2), 591-609 (2018).
15. 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*. **123** (6), 1790-1795 (2018).

16. Borton, M. A., *et al.* Coupled laboratory and field investigations resolve microbial interactions that underpin persistence in hydraulically fractured shales. *Proceedings of the National Academy of Sciences*. **115** (28), E6585-E659 (2018).
17. Tang, K., Page, J. S., Smith, R. D. Charge competition and the linear dynamic range of detection in electrospray ionization mass spectrometry. *Journal of the American Society for Mass Spectrometry*. **15** (10), 1416-1423 (2004).
18. Boiteau, R. M., *et al.* Structure Elucidation of Unknown Metabolites in Metabolomics by Combined NMR and MS/MS Prediction. *Metabolites*. **8** (1), 8 (2018).
19. Walker, L. R., *et al.* Unambiguous Metabolite Identification in High-throughput Metabolomics by Hybrid 1D NMR/ESI MS Approach. *Magnetic Resonance in Chemistry*. **54** (12), 998-1003 (2016).
20. Nakayasu, E. S., *et al.* MPLEX: a robust and universal protocol for single-sample integrative proteomic, metabolomic, and lipidomic analyses. *MSystems*, **1** (3), e00043-16 (2016).
21. Folch, J., Lees, M., Sloane-Stanley, G. H. Extraction of fatty acid. *Journal of Biological Chemistry*. **226**, 497-509 (1957).
22. Tolic, N., *et al.* Formularity: software for automated formula assignment of natural and other organic matter from ultrahigh-resolution mass spectra. *Analytical Chemistry*. **89** (23), 12659-12665 (2017).
23. Kim, Y. M., *et al.* Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms. *Frontiers in microbiology*. **6**, 209 (2015).
24. Hiller, K., *et al.* MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Analytical Chemistry*. **81** (9), 3429-3439 (2009).
25. Kind, T., *et al.* FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry*. **81** (24), 10038-10048 (2009).
26. Kyle, J. E., *et al.* LIQUID: an-open source software for identifying lipids in LC-MS/MS-based lipidomics data. *Bioinformatics*. **33** (11), 1744-1746 (2017).
27. Smith, P. K., *et al.* Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*. **150** (1), 76-85 (1985).
28. Kanehisa, M. Enzyme annotation and metabolic reconstruction using KEGG. *Protein Function Prediction: Methods and Protocols*. **1611**, 135-145 (2017).
29. Van Krevelen, D. W. Graphical-statistical method for the study of structure and reaction processes of coal. *Fuel*. **29**, 269-284 (1950).
30. Hodgkins, S. B., *et al.* Changes in peat chemistry associated with permafrost thaw increase greenhouse gas production. *Proceedings of the National Academy of Sciences*. **111** (16), 5819-5824 (2014).

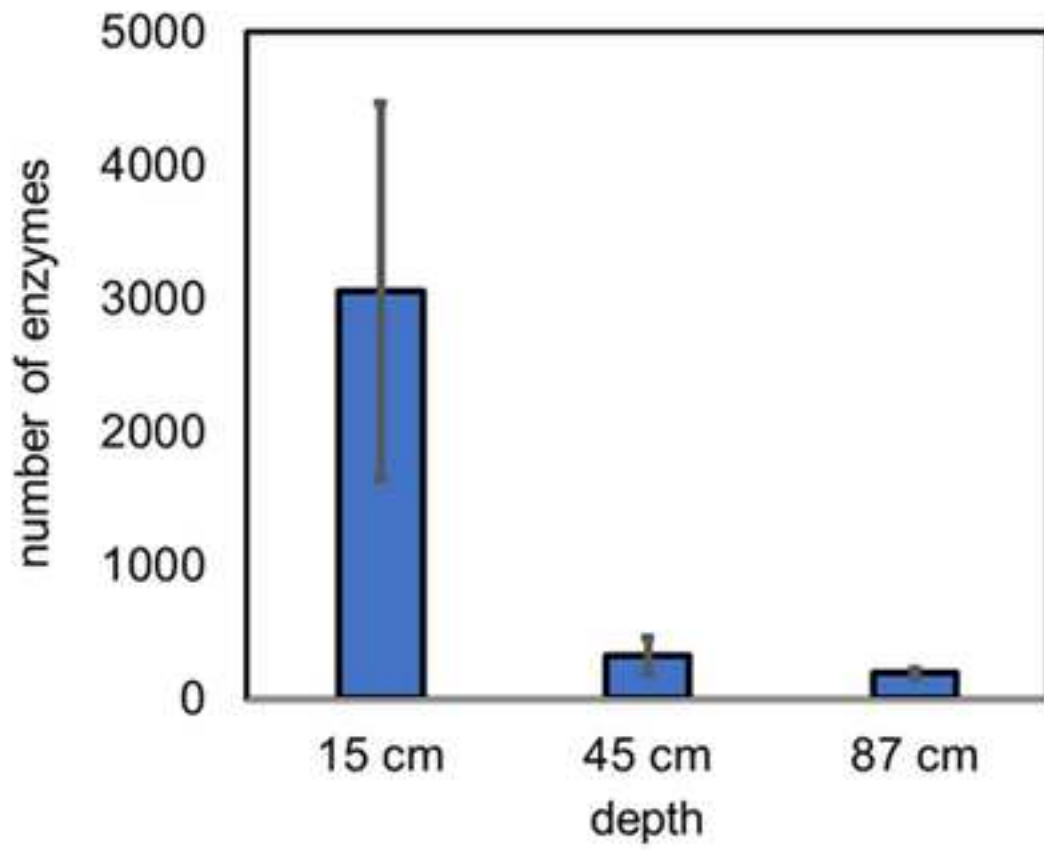


Figure 2

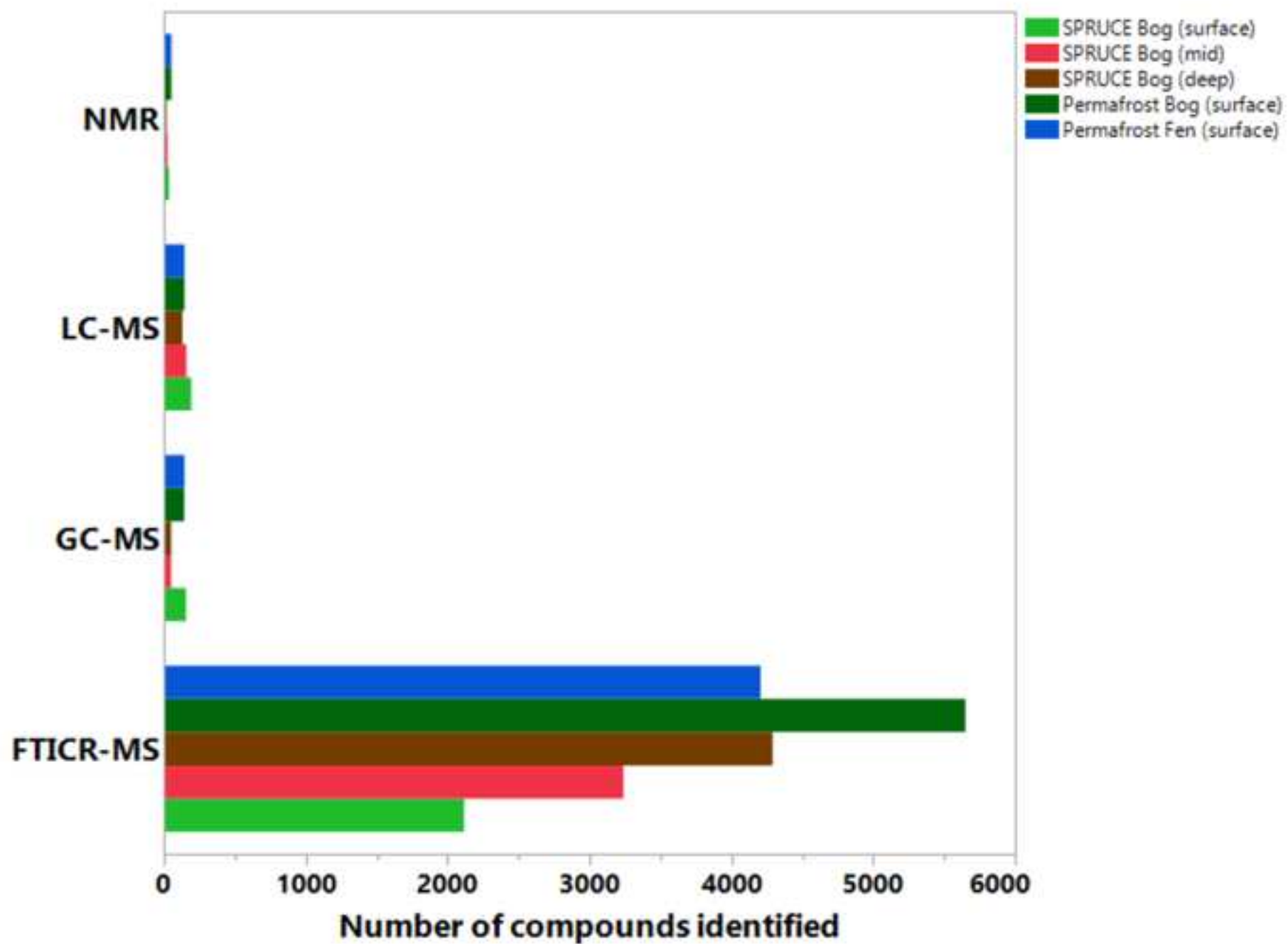


Figure 3

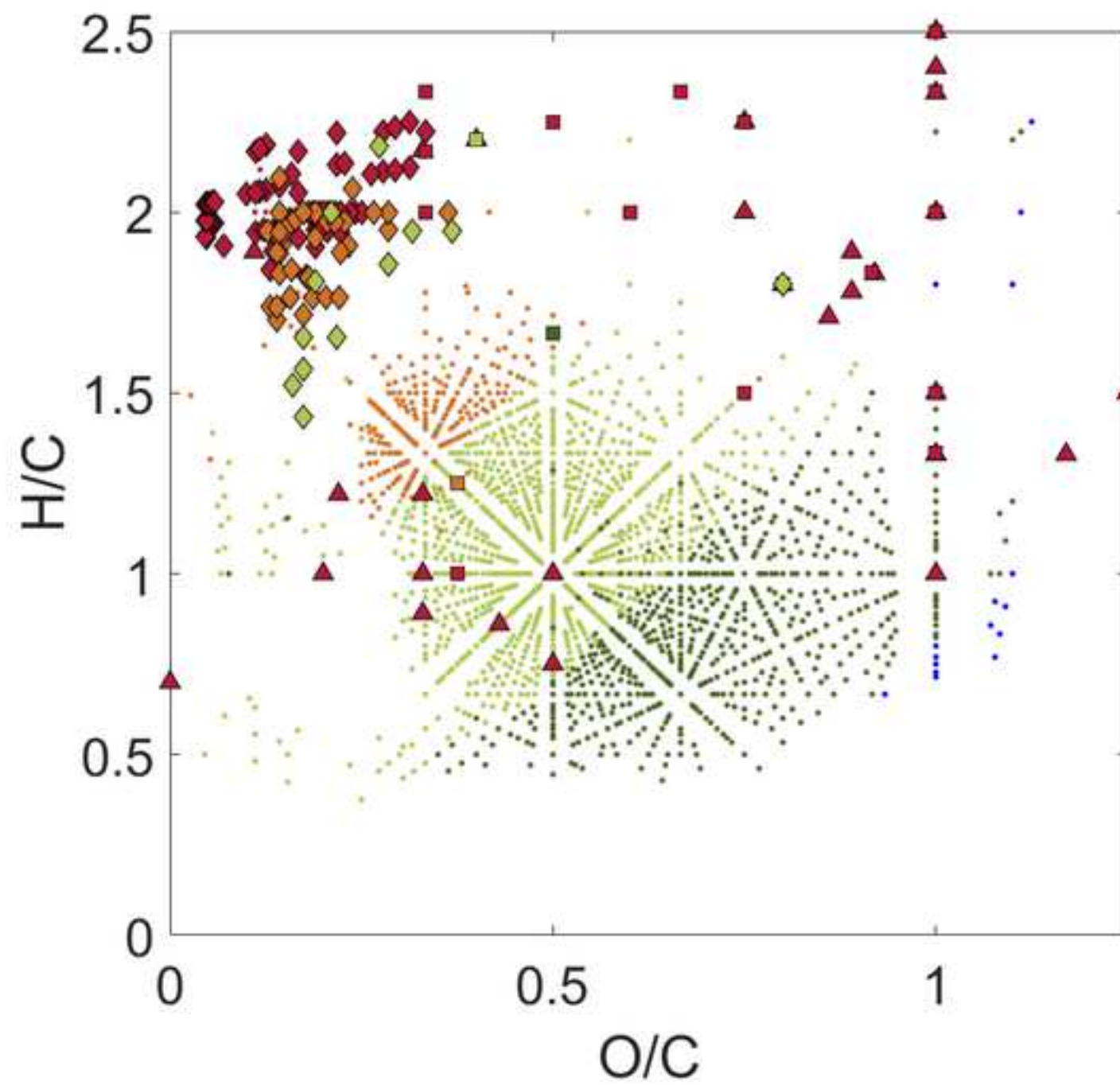


Figure 4

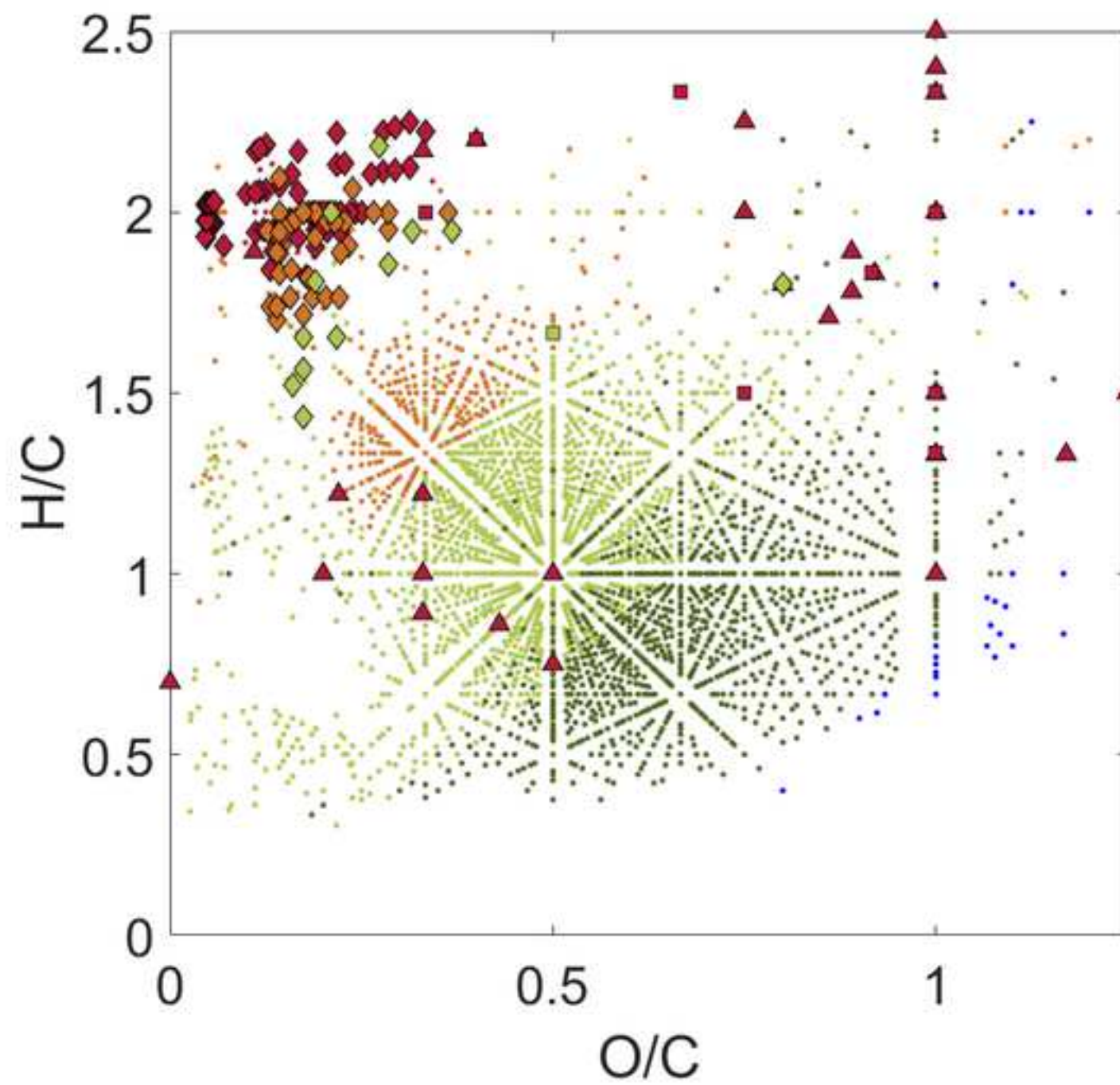
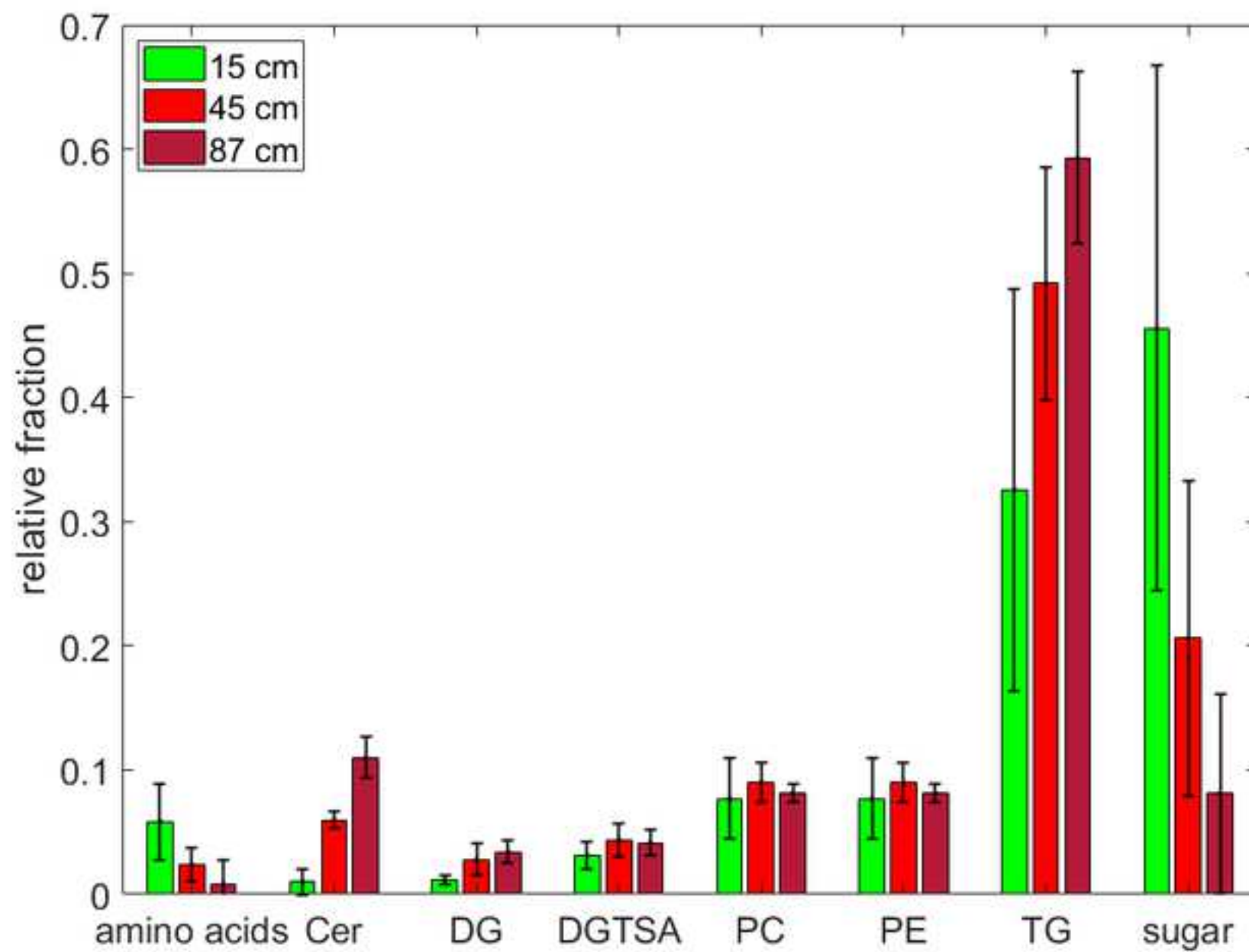
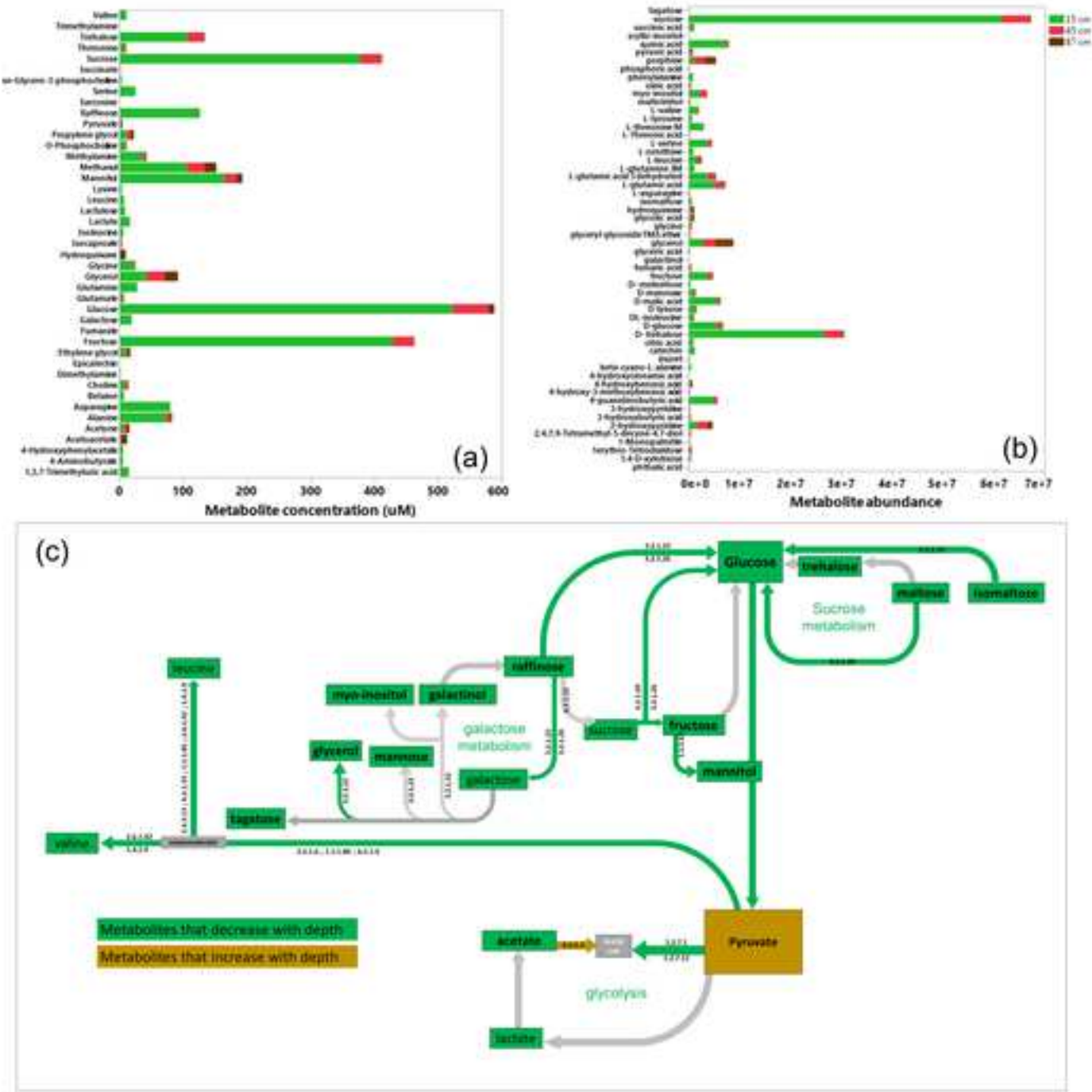


Figure 5





Name of Material/ Equipment	Company	Catalog Number
methoxyamine hydrochloride	Sigma	
5 mm triple resonance salt-tolerant cold probe	Aldrich	226904
capillary GC column HP-5MS column (30 m × 0.25 mm × 0.25 µm)	Bruker	
reversed phase charged surface hybrid column (3.0 mm × 150 mm × 1.7 µm particle size)	Agilent	AG19091S-433
	ThermoFis	
	her	
	VWR	
2 mL glass vials	Internatio	
	nal	46610-722
	VWR	
autosampler vials	Internatio	97055-324;
	nal	9467671
	VWR	
Chloroform	Internatio	
	nal	JT9174-3
	VWR	
Ethanol	Internatio	
	nal	BDH67002.400
	VWR	
methanol	Internatio	
	nal	BDH85681.400
	VWR	
pyridine	Internatio	
	nal	BDH67007.400
	Sigma	
2,2-dimethyl-2-silapentane-5-sulfonate-d6	Aldrich	178837
C8-C24 fatty acid methyl ester	Sigma	
	Aldrich	CRM18918
	Sigma	
N-methyl-N- (trimethylsilyl)trifluoroacetamide	Aldrich	24589-78-4

Suwanee River Fulvic Acid standard	International Humic Substances Society	2S101F
trimethylchlorosilane	Sigma	
Tuning Solution	Aldrich	89595
FTICR-MS analysis software	Agilent	
	Bruker	Compass DataAnalysis
	Pacific Northwest National Laboratory	
Formularity Software	y	Formularity
GC-MS	Agilent	Agilent GC 7890A with Phenomenex
	(Torrance, CA)	
silica-based sorbent	VWR	Strata C18-E (PN 8E)
	International	
NMR TUBE 3MM 8 150 CS5	Varian	KT897820-0008
Varian Direct Drive 600-MHz NMR spectrometer	Inova	Varian Direct Drive
Chenomx NMR Suite 8.3	Chenomx	Chenomx NMR Suite
ultra-performance liquid chromatograph	waters	Aquity UPLC H
	ThermoFisher	
Velos-ETD Orbitrap mass spectrometer	her	Thermo Scientific L

Please remove commercial language and use generic terms instead. Please include product information in the Table of Materials.

Comments/Description

derivitization agent

instrumentation

instrumentation

instrumentation

sample vials

sample vials

solvent

solvent

solvent

solvent

standard

standard

standard

standard

standard

is 4.1

available for download at:

<https://omics.pnl.gov/software/formularity>

h MSD 5975C

-S001-DAK)

NMR tube

NMR spectrometer

NMR software

liquid chromatograph

mass spectrometer



1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article:

Single-throughput complementary high-resolution analytical techniques for characterizing complex natural organic matter mixtures

Author(s):

Malak M Tfaily, Rachel M Wilson, David Hoyt, Heather Brewer, Heino Heyman, Jennifer Kyle, Rosalie Chu

Item 1 (check one box): The Author elects to have the Materials be made available (as described at

<http://www.jove.com/author>) via: ☒ Standard Access ☐ Open Access

Item 2 (check one box):

☐

The Author is NOT a United States government employee.

☒

The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐

The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: “**Agreement**” means this Article and Video License Agreement; “**Article**” means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; “**Author**” means the author who is a signatory to this Agreement; “**Collective Work**” means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; “**CRC License**” means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: <http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode>; “**Derivative Work**” means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; “**Institution**” means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; “**JoVE**” means MyJoVE Corporation, a Massachusetts corporation and the publisher of *The Journal of Visualized Experiments*; “**Materials**” means the Article and / or the Video; “**Parties**” means the Author and JoVE; “**Video**” means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4** and **7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the “Open Access” box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

ARTICLE AND VIDEO LICENSE AGREEMENT

4. Retention of Rights in Article. Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. Grant of Rights in Video – Standard Access. This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. Grant of Rights in Video – Open Access. This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this Section 6 is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. Government Employees. If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such

statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. Likeness, Privacy, Personality. The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

9. Author Warranties. The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

10. JoVE Discretion. If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have

ARTICLE AND VIDEO LICENSE AGREEMENT

full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

11. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's

expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

12. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US\$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

13. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to be one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement required per submission.

CORRESPONDING AUTHOR:

Name:

Rachel M. Wilson

Department:

Earth Ocean and Atmospheric Sciences

Institution:

Florida State University

Article Title:

Single-throughput complementary high-resolution analytical techniques for characterizing complex natural organic matter mixtures

Signature:



Date:

8-31-18

Please submit a signed and dated copy of this license by one of the following three methods:

- 1) Upload a scanned copy of the document as a pdf on the JoVE submission site;
- 2) Fax the document to +1.866.381.2236;
- 3) Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02139

For questions, please email submissions@jove.com or call +1.617.945.9051

Editorial comments:

Changes to be made by the Author(s) regarding the written manuscript:

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.

We have carefully proofread the article and corrected all grammatical and spelling errors.

2. Please rephrase the Introduction to include a clear statement of the overall goal of this method.

We added the following statement of goals to the introduction:

“The overall goal of this method is to provide a single throughput protocol for analyzing metabolites, lipids, and microbial proteins from a single sample thereby creating a fully paired dataset for building metabolic networks while constraining analytical errors.”

3. Please convert centrifuge speeds to centrifugal force (x g) instead of revolutions per minute (rpm).

We converted rpms to centrifugal force throughout the protocol where applicable (i.e. lines 224, 282, 287, & 354).

4. Please add more details to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol. Please ensure you answer the “how” question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action. See examples below:

5. 1.2: Please specify roughly how much soil/peat is needed for the protocol.

We added the following text to allow users to determine the sample size needed for the protocol:
“Approximately 45 mg C is needed for this protocol. For peat (typically 45% C), 100 mg of dried peat are required. Larger amounts of sample may be needed for low organic samples like mineral or forested uplands soils depending on the C content.”

6. 1.3: What is used to grind the samples?

We clarified that one can use a high-speed ball mill to grind the samples in the bolded text as follows:

“When ready to analyze the samples, freeze dry to constant weight, then grind the samples **in a high-speed ball mill using stainless steel grinding balls** to homogenize and break up any aggregates.”

7. 1.5: Is the sample shaken in a shaker?

We clarified that a shaker table may be used for this step:

“Add 1 mL of distilled, degassed water (H₂O) to each sample, cap the vials and shake for 2 h **on a shaker table.**”

8. 1.6: Centrifuge for how long?

We clarified that the samples should be centrifuged for 30 minutes.

9. 1.8: Please describe how to separate the two solvent layers. Are the two layers labeled?

The two layers are not labeled per se, but the chloroform-containing fraction is denser than the methanol containing fraction so it will be the lower layer. We clarified this in the instructions and added text to suggest using a separation funnel or careful pipetting to separate the two layers:

“Carefully separate the two resulting solvent layers, **which will be visually distinguishable**, for separate analysis by FTICR-MS by using a **separation funnel or simply remove the top layer by careful pipetting. The chloroform-containing fraction will be on the bottom while the methanol-containing fraction is less dense and will be on the top.**”

10. 1.9: There are two chloroform extracts from step 1.8, correct? Are they both diluted and analyzed? Please also specify the dilution factor.

There is one chloroform extract and one methanol extract generated during [now] step 1.7. We clarified this as noted in the previous response. We added a dilution factor (2:1) and a note that the methanol fraction does not need to be diluted further in methanol.

11. 2.1: Please provide the composition of tuning solution. If it is purchased, please cite the Table of Materials. Please describe the calibration process. For instance, what volume of tuning solution is needed?

This is a commercially available tuning solution. We added a reference to the table of materials noting that the tuning solution can be purchased from Agilent and added details about volume of solution needed for the tuning.

12. 2.2: How is the Suwannee River Fulvic Acid standard prepared? Is it dissolved in methanol? Please describe how to tune and optimize.

We added text to describe preparing SRFA samples as follows:

“Prepare the Suwannee River Fulvic Acid standard (Table of Materials) by diluting 1 mg mL⁻¹ in milliQ filtered water then diluting the resulting solution to 20 ug mL⁻¹ in methanol. Direct inject 23 uL of this final solution to the ESI source coupled to the FTICR spectrometer through a syringe pump set to a flow rate of 3.0 µL min⁻¹. Set needle voltage to +4.4 kV, Q1 to 150 *m/z* and glass capillary at 180° C. Inspect resulting spectra using the *DataAnalysis* software package 4.2 (Bruker Daltonics) to confirm the quality of the data.”

13. 2.8-2.10: These steps can be combined into one. Please mention why water and methanol-extracted fractions are combined. Please specify how to dry the combined extract.

We clarified that the samples should be freeze dried. We combined steps 2.8-2.10 into a single step. We explained that combining the water and methanol fractions streamlines subsequent steps in the protocol but point the user to the Discussion where the two options are treated in more detail:

“Note: After analyzing by FTICR-MS, the decision can be made to either combine the water and methanol extracted fractions in order to streamline the remaining steps or the fractions can be kept separate throughout subsequent steps. The advantages and disadvantages of each approach are described at length in the Discussion. If doing so, combine the water and methanol-extracted fractions. Dry using a vacufuge/freeze dryer and save the remainder of the extracts for subsequent GC-MS (water, methanol or water + methanol), LC-MS (chloroform) and NMR (water + methanol) analysis.”

14. 4.1: This step seems to be repetitive to what mentioned in 2.8-2.10. (-97.6 °C)?

We revised the protocol to eliminate this repetitive instruction. We also clarified in the earlier step that a vacufuge-freeze dryer can be used to aid in drying down the methanol extracts.

15. 4.2: Please describe how to prepare blanks and FAME calibration samples.

We added additional text to the protocol to better describe blank and FAME preparation:

“Prepare blank control samples of HPLC-grade hexane in GC-MS autosampler vials²². Dissolve 100 mg fatty acid methyl esters mixture retention time standard (FAMEs: C8–C28) in 200 uL hexane.

To protect carbonyl groups, add 20 µL of 30 mg mL⁻¹ methoxyamine hydrochloride in pyridine to each of the methanol extracts and water extracts (or combined methanol/water extracts if using) from step 2.5, blanks and calibration samples. Seal vials with caps.”

16. 6.2: Please describe how to deconvolute and align features.

We added additional information to the GC- MS data processing section:

Inspect all the data files to ensure that they were correctly captured. Pay attention to potential shifts with regards to internal standard retention times and intensities to confirm that that the data was captured consistently throughout the analysis.

Convert the vendor specific MS data format to a general MS format if required. Process raw data files using MetaboliteDetector calibrating retention indices based on the FAME internal standards. After aligning the retention times of all data files, continue with deconvolution and finally metabolite identification by matching retention indices and GC-MS spectra against the FiehnLib polar metabolite library²³.

Cross-check remaining unidentified metabolites against the NIST14 GC-MS library using spectral matching. Validate identifications individually to eliminate false identifications and reduce deconvolution errors.

17. 7.2: Please specify the size of the NMR tube used. What volume of water extracts is used?

We clarified in the protocol that a 3 mm outer diameter (O.D.) NMR tube is used and that the sample volumes are typically in the range of 180 – 300 µL:

Dilute the remainder of the water extracts by 10% (vol/vol) with a 5 mM 2,2-dimethyl-2-silapentane-5-sulfonate-d₆ internal standard. Alternatively, you can combine water and methanol extracts from step 1 then resubstitute in water. By doing so however, you might lose some of the volatile compounds during the freeze-drying step. **Typical final sample volumes are in the range 180-300 µL**

Transfer mixture into a **high-quality 3 mm outer diameter (O.D.) borosilicate glass NMR tube.**

18. 8.4: Please write the text in the imperative tense.

We revised to imperative tense.

19. 8.5: Please describe how this is done. We generally require that software steps be more explicitly explained ('click', 'select', etc.).

We added more detail on the software steps as follows:

“Upload the raw LC-MS/MS data files into LIQUID²⁴ along with the target file (i.e. list of >25,000 lipid species) for the respective ionization mode (positive or negative). Process the raw file. Manually validate the resulting identifications by examining the MS/MS spectra for presence of diagnostic ions, if applicable, matching fragment ions (e.g., fatty acyl chains), isotopic separation, mass ppm error of the precursors, and the retention time. Export resulting list of confidently identified lipids as a .tsv file.”

20. 9.1, 9.5: Although a reference has been provided about the protocol, we still need more specific details for filming. For instance, what extract is washed? What is the ratio of extract/cold methanol? What volume of 100 mM NH₄HCO₃ and 8 M urea is added? What wavelengths are measured for protein concentration?

We added text to this section to clarify that the methanol-containing extract is to be washed in 20 times the volume of cold methanol. We indicated that the pellet should be redissolved in the NH₄HCO₃/Urea solution to a final protein concentration of 1mM (approximately 10-fold).

Finally, we added that the absorbance measurements should be made at 562 nm as follows:

“Extract proteins according to the MPLEx protocol²⁰ from the remainder of the methanol phase resulting from step 2.4, by washing the extract with 20 times the extract volume of additional cold (-20 °C) methanol.

Use a 1 mL/50 mg Phenomenex Strata C18-E (PN 8E-S001-DAK, Torrance, CA), to condition C18 solid phase extraction (SPE) columns with 3 mL MeOH, 2 mL of 0.1% trifluoroacetic acid (TFA) in water, followed by the addition of the extract from step 9.1 at a rate no greater than 1 mL/min. Following the addition of the sample, wash the column with 4 mL of 95:4.9:0.1 water:acetonitrile:TFA, then allow to go dry. Place a 1.5-mL collection tube under the SPE column, and elute the sample with 1.0 mL of 80:19.9:0.1 methanol:water:TFA.

Concentrate the extracts to 100 uL under vacufuge, then measure the protein concentration by bicinchoninic acid (BCA) colorimetric assay²⁵ at a wavelength of 562 nm.

Centrifuge extracts at 10,000 x g for 10 min at 4 °C. Discard the resulting supernatant and dry the remaining pellet under vacuum for 5 min. Resuspend the protein pellet in water to a final concentration of 0.1 ug peptide uL⁻¹. Add dithiothreitol to a final concentration of 5 mM and incubate at 60 °C for 30 min. Dilute 10-fold with 100 mM NH₄HCO₃/8M urea solution and CaCl₂ to a final concentration of 1 mM and incubate at 37 °C for 3 h.”

21. 9.9: This step does not have sufficient details to replicate. Please add more details.

We added additional details about the solid phase extraction process including the following text:

“Use a 1 mL/50 mg Phenomenex Strata C18-E (PN 8E-S001-DAK, Torrance, CA), to condition C18 solid phase extraction (SPE) columns with 3 mL MeOH, 2 mL of 0.1% trifluoroacetic acid (TFA) in water, followed by the addition of the extract from step 9.1 at a rate no greater than 1 mL/min. Following the addition of the sample, wash the

column with 4 mL of 95:4.9:0.1 water:acetonitrile:TFA, then allow to go dry. Place a 1.5-mL collection tube under the SPE column, and elute the sample with 1.0 mL of 80:19.9:0.1 methanol:water:TFA.”

22. 9.10: Please specify the volume of water added.

We added text to indicate that water is added to a final peptide concentration of 0.1 ug mL⁻¹.

23. Please combine some of the shorter Protocol steps so that individual steps contain 2-3 actions and maximum of 4 sentences per step.

We combined many of the protocol steps together reducing the total number of steps from 61 to 41.

24. Please include single-line spaces between all paragraphs, headings, steps, etc.

We added spaces between all of the paragraphs, headings and steps in the manuscript.

25. After you have made all the recommended changes to your protocol (listed above), please highlight 2.75 pages or less of the Protocol (including headings and spacing) that identifies the essential steps of the protocol for the video, i.e., the steps that should be visualized to tell the most cohesive story of the Protocol.

We have highlighted 2.75 pages of the essential steps of the protocol for filming.

26. Please highlight complete sentences (not parts of sentences). Please ensure that the highlighted part of the step includes at least one action that is written in imperative tense. Please do not highlight any steps describing anesthetization and euthanasia.

We checked to make sure we highlighted only complete sentences that included an imperative action. No steps in our protocol involve anesthetization or euthanasia.

27. Please include all relevant details that are required to perform the step in the highlighting. For example: If step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be highlighted.

We checked that steps are complete and details relevant for performing each step are included.

28. Please submit the figures as a vector image file to ensure high resolution throughout production: (.svg, .eps, .ai). Please save any .ai files as a .pdf for submission but maintain .ai files for production purposes. If submitting as a .tif or .psd, please ensure that the image is 1920 pixels x 1080 pixels or 300 dpi.

We submitted images as a .tif file in response to communications we had during the review process requesting .tif files.

29. Figure 1: Please define the error bars in the figure legend.

We added definition of error bars to the figure legend.

30. Figures 2 and 3: Please add a short title for each figure.

We added a title for each of these figures to the figure legend.

31. Please revise the table of the essential supplies, reagents, and equipment to include the name, company, and catalog number of all relevant materials in separate columns in an xls/xlsx file.

We have included this table.

Reviewers' comments:

Reviewer #1:

Manuscript Summary:

Manuscript describes sequential, multimodal spectroscopic and spectrometric approach for the analysis of proteins, primary and secondary metabolites, and lipids within soil, a complex matrix. The technique is applied to only one ecosystem but is ready applied to others. It is generally well written although could adopt a more consistent "voice" throughout.

Major Concerns:

The starting point for extracts in LC-MS and Proteomics is not explicitly described. What earlier step created these extracts?

We clarified in the protocol that the extracts were generated during step 2.4.

Minor Concerns:

- The minor concerns listed below are to increase comprehension of manuscript.
- Consider including some description of the system you are studying in the abstract, perhaps as the last sentence. As currently written, the transition from the abstract to the introduction which immediately talks about the global importance of wetlands is quite abrupt.

We added a final sentence to the abstract to explain that we used the peat samples as a demonstration in this paper but that the method can be applied to other kinds of natural organic matter samples as well.

“Herein we present results of applying this method to natural organic matter samples collected from peatlands, however, the protocol is applicable to any natural organic matter sample (e.g. peat, forested soils, marine sediments, etc).”

We have added the following text to explain why we are interested in characterizing NOSC in peatland systems specifically.:

“Peatlands produce globally significant quantities of CO₂ and CH₄⁶ via microbial respiration of naturally occurring organic matter. The NOSC of the organic carbon substrates determines the relative proportion of CO₂:CH₄ produced which is a critical parameter because of the higher radiative forcing of CH₄ compared to CO₂¹¹, but also because modeling efforts have identified this ratio as a critical parameter for estimating C flux in peatlands¹². In the absence of terminal electron acceptors other than CO₂, it can be shown by electron balance that organic C substrates with NOSC > 0 will produce CO₂:CH₄ > 1, organic C with NOSC = 0, produces CO₂ and CH₄ in equimolar ratio and organic C with NOSC < 1 will produce CO₂:CH₄ < 1¹³.”

-Line 93 - consider replacing "directly related" to correlated.

We have made the suggested change.

-Line 123 - consider replacing "directly" to linearly

We have made the suggested change.

-Paragraph starting on line 127 continues until line 191. Can you identify natural breaks in the discussion? Perhaps a break between bulk and molecular techniques that occurs at line 137, and then breaks between discussion of each molecular technique?

We have made the suggested change to break this paragraph up into several smaller sections.

-In most cases it is easy to follow the flow of different subsample analyses in the protocol. However it is not carried thru the entire protocol. For example, it is unclear where the chloroform extracts come from for the LC-MS analysis. Earlier, in the description this was unambiguously indicated (e.g. step 1.6 on line 230). Can you do this consistently?

We added an indication of the step that generated the extracts in the protocol as requested:
Liquid chromatography mass spectrometry (LC-MS) lipidomics analysis

Dry the remaining chloroform extract **generated during step 2.4** under vacuum. Then rewet the dried organic layer with 200 µL methanol.

-Within the Representative results section the "voice" of the narrative changes from standard journal language to a more of oral one.

-Line 381 Suggest rephrasing "We will compare these result to those from a permafrost bog and fen from northern Sweden..."

We have made the suggested change.

-Line 387 - proteomics shows what proteins are expressed. The potential of microbial community comes from the meta genome.

We meant to indicate that the proteins indicate potential in that they act as the catalysts for transformation, thus the proteins indicate what transformations are metabolically possible. While the transforms in the metabolites indicate which of these possible reactions are actually occurring. We have revised this in the text as follows:

"While the proteomics results indicate which proteins are expressed, the metabolic data shows which reactions are actually occurring."

-Figure 2 - It would be helpful in the text to explicitly say what samples (depth horizons) are plotted in 2b and 2c. It is also difficult to discern the differences for the two depth from the VK plot itself. It might be helpful to have an additional table of the % of sugar, lipid, protein, etc. for the depths described. This would better support the arguments that are presented.

We added depths to the figure legend. We also added a panel to this figure containing a bar chart showing average amino acids, lipid classes, and sugars at the three measured depths.

-Lines 400-410 - I expected that a discussion of preservation of metabolites at depth due oxygen limitation and their NOSC values to be presented here. Is that not observed?

We do observe consumption of high NOSC compounds and preservation of low NOSC compounds in the peat and added the following text to the Representative Results section describing this phenomenon:

“As can be seen in Figure 2d sugars (NOSC = 0) and amino acids ($0 < \text{NOSC} < 1$) are consumed in the surface peat, while lipids ($-2 < \text{NOSC} < -1$) appear to accumulate with depth indicating. This is consistent with expectations based on NOSC values that higher NOSC compounds are more readily degraded while lower NOSC compounds persist in the highly anaerobic (i.e. TEA-limited) conditions of subsurface peat.”

-Line 465 - Is clinically supposed to be chemically?

No, we meant clinical pathogens. The idea being that environmental samples contaminated with medically-relevant microorganisms could be rendered safe by this protocol.

-Line 493 - similar to earlier comments - proteomics reveals the proteins and enzymes that are expressed. The potential comes from metagenome analysis.

Again, we meant to highlight that as catalysts of the reaction the enzymes do not provide the most direct evidence that a reaction is occurring or has occurred only that the microbial enzymes are present, that is what we meant by potential. By coupling that information with the metabolic analyses provides corroborative evidence that the reaction we think should occur has actually in fact taken place. We revised the text as follows:

“Coupling the proteomics analysis with the metabolic analyses provides compelling corroborative evidence that a reaction is actually occurring.”

-Line 503 - missing "of" but critical piece "of " the C cycling puzzle.

We added the missing word.

Reviewer #2:

The manuscript describes the pipeline for co-extraction and analyses of proteins and metabolites from environmental samples. Overall, the manuscript is well presented and details analyses that are becoming more utilized. I have some minor comments and suggestions below:

L42 : Microbial composition is a bit vague here - and also incorrect. None of the analyses detailed here provide either information on the 'community composition', while the detected proteins only offer a partial view of microbial (cellular) composition - and this kind of analysis isn't discussed.

We revised the text to indicate specifically that we measure proteomics and not microbial composition.

L52: 'the system' is vague

We clarified this to “natural ecosystems”.

L112: 'also mediated'? The preceding paragraph described microbial degradation of OM. The whole sentence is somewhat confusing in the context of the earlier paragraphs.

It's not clear to us what the reviewer finds confusing here. Thermodynamics ultimately limits decomposition. Microorganisms control rates through the production of enzymes that act as catalysts. We begin with the discussion of thermodynamic control to define the limits of what is possible. Then we discuss kinetic control to focus in on what is actualized in the system. We removed the word “also” and rearranged the indicated sentence in hopes that this would clarify our meaning.

L154: I would consider NMR to be reasonably sensitive - I think 'only metabolites in very high concentrations can be quantified' is too strong a statement. Perhaps some detection limits could be given here?

This is clearly subjective, so we revised the text to indicate that reliable quantification requires at least 1 uM concentration.

L181: Is metabolites meant here, rather than 'metabolomics'?

We changed this to metabolites.

L366: There are no details provided on how the mass spectra are matched to a database of predicated peptides. This is a critical step (databases selection or design) that should be at least discussed (perhaps in the discussion?)

We added the following text regarding the database matching:

“Convert RAW spectra files to mzML format using msConvert or ProteoWizard 3.0.10200 accepting all default parameters. Use the universal database search tool MSGFPlus v2017.01.13 to search resulting proteomes against a targeted protein database of protein sequences predicted from relevant metagenome assembled genomes. Append common contaminants (e.g. trypsin, keratin, albumin) and remove all exactly duplicated protein sequences to improve peptide-to-mass-spectrum match statistics. Evaluate the resulting MSGF Spectral Probability scores to determine which peptide-to-mass-spectrum match is best. Use the Q-Value from MSGFPlus to filter the entire data pile to 1% false discovery rate (FDR).”

We also added the following caveats on proteomics analysis to the discussion:

In this procedure for identifying proteomics, only fully tryptic peptides are searched, thus endogenous peptidase activity and in-source fragmentations will be missed. On the other hand, oxidized methionine may be considered as a post-translational modification for the peptide candidates as this modification commonly occurs during sampling processing and handling. Quantification of enzymes using peptide elution areas can be done, but was outside the scope of this project.

L383: Are these truly all enzymes? Or are some structural proteins or even proteins of unknown function?

These are enzymes for which Enzyme Commission numbers were able to be identified.

L389: Are these 67,040 unique metabolites? Or are some likely shared across measurement platforms?

67,040 unique metabolites. There is some overlap among techniques, but redundant compounds were not counted in this number.

L500: 'declining enzyme activity' isn't technically accurate here. Enzymes are at lower abundances at depth, but we don't have any information on their actual kinetics.

Point taken. We revised this to declining enzyme expression.

Figures 2 and 3 were of low quality, and so I wasn't able to effectively view or assess them.

We have uploaded revised high resolution tiff files of all of the figures in response to a request from the editor.