

# Comprehensive assessment of dissolved organic matter processing in the Amazon basin

Siyu Li

Vollständiger Abdruck der von der TUM School of Life Sciences der Technischen Universität  
München zur Erlangung des akademischen Grades einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Jürgen Geist

Prüfer der Dissertation:

1. apl. Prof. Dr. Philippe Schmitt-Kopplin
2. Prof. Dr. Florian Einsiedl

Die Dissertation wurde am 07.06.2023 bei der Technischen Universität München eingereicht und  
durch die TUM School of Life Sciences am 22.09.2023 angenommen.

## Acknowledgements

First and foremost, I would like to express my heartfelt and sincere gratitude to my supervisors Prof. Dr. Philippe Schmitt-Kopplin and Dr. Norbert Hertkorn for providing me with valuable study and work opportunities with state-of-art instruments, patient guidance to the interesting research, supervision of my thesis, and great efforts for my future career. Many thanks to Prof. Dr. Philippe Schmitt-Kopplin for his wide research areas and creative ideas. Particular thanks to Dr. Norbert Hertkorn for his expertise on NMR, brilliant ideas, nice discussions, endless support and encouragement. His wealth of knowledge broadened my horizons and encouraged me to be always be curious and open-minded.

Special thanks to Dr. Mourad Harir for the guidance on FT-ICR MS measurements, patient guidance and discussions on data analysis and writing. Thanks to Prof. Dr. Michael Gonsior and Dr. Alex Enrich-Prast for taking the time and giving nice feedback as external members of my thesis committee. Additionally, I would like to thank Prof. Dr. David Bastviken for his valuable comments during the revision of the manuscripts.

Thanks to Prof. Dr. Philippe Schmitt-Kopplin, Dr. Norbert Hertkorn, Dr. Mourad Harir, and Dr. Juliana Valle for sampling in the 2014 Amazon campaign. Their diligent work provided the basis of my PhD research project. Thanks to Dr. Fausto Machado-Silva and Dr. Roberta Peixoto for their cooperation on the Amazon study, bacterial and methane data, and nice conversations. They helped me to link the composition of dissolved organic matter with microbial processes and provided insightful suggestions for the revision of manuscripts. I would like to thank Dr. Miao Zhang for her nice cooperation on the South China Sea project. This experience helped me open the door to natural organic matter research.

Great thanks to BGC colleagues, with whom I have learnt a lot and enjoyed myself. The Covid-19 pandemic kept me away from China for three years, but their company has made me never feel lonely in Munich. Thanks to Philippe Diedrich for valuable discussions on NMR spectroscopy and research projects. Thanks to my office roommates, Dr. Stefan Pieczonka, Dr. Nina Sillner, and Leopold Weidner, who constantly discussed research ideas with me. Thanks to Yingfei Yan, who gave me nice introduction on statistical analysis methods, and R coding. Thanks to Michelle Berger for her nice suggestions and company during thesis writing. Thanks to Dr. Franco Moritz for his introduction to Netcalc software, mass peak annotation, and calculation of databases. Thanks to Astrid Bösl for her help with visa and contract. Thanks to Dr. Silke Heinzmann for the guidance and help on the NMR autosampler. Thanks to the technical assistance from Brigitte Look and Jenny Uhl. Further thanks to Dr. Alesia Walker, Dr. Basem Kanawati, Dr. Marianna Lucio, Xin Zhang, Eva-Maria Harrieder, Dr. Thuy Thanh Truong, Dr.

Lance Robert Buckett, Micheal Becker, Habiba Selmi, and many others with whom I've shared nice moments together.

I am grateful for my parents, who always support me unconditionally. Thanks to my friends in Germany, with whom I shared the PhD journey. We shared delicious and joyful moments together. Also, thanks to my relatives and friends in China for their constant support during this journey.

Finally, I appreciate the 48-month financial support from China Scholarship Council (CSC) during my studies.

## Summary

The Amazon River catchment is the largest river system in the world, responsible for ~20% of the global freshwater discharge; it exports ~35 Tg organic carbon annually to the Atlantic Ocean. Dissolved organic matter (DOM) in the Amazon River plays a crucial role in ecosystem processes, such as nutrient availability, growth efficiency of aquatic organisms, the decomposition rate of plant debris, and ultimately, the carbon cycle. Polydispersity, heterogeneity, and temporal dynamics of DOM have made its molecular characterization a long-standing challenge. Previous studies on the origin and fate of organic matter in the Amazon River included stable isotope and radioisotope analysis, quantification of targeted compounds, optical spectrometry, mass spectrometry, <sup>13</sup>C NMR spectroscopy, among others. However, a comprehensive molecular understanding of the reactivity and bioavailability of Amazon River DOM, and an assessment of its biotic and abiotic processing is lacking at present, especially in tributaries with different water types and in major confluence zones.

This thesis provides a comprehensive molecular understanding of the processing of DOM molecules in the Amazon River main stem (the Amazonas River), selected main tributaries (Solimões, Negro and Tapajós rivers), as well as the confluences Solimões / Negro (SN) and Amazonas / Tapajós (AT). Ultrahigh resolution negative and positive electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI[±] FT-ICR MS) provided complementary coverage of DOM molecules and demonstrated distinct diversity and susceptibility of processing for molecules with nitrogen and sulfur heteroatoms (CHO, CHNO, and CHOS). 800 MHz <sup>1</sup>H nuclear magnetic resonance spectroscopy (NMR) provided quantification of major atomic environments typical of processed aqueous DOM, including aliphatic substructures which provide only indirect signatures in mass spectra and optical spectra. Both MS and NMR spectra revealed clear distinction of DOM composition and structures in different types of waters in the Amazon basin. DOM in the Negro River contained more polyphenolic and other highly oxygenated compounds, whereas DOM in Solimões and Amazonas rivers contained more nitrogen and sulfur compounds (CHNO and CHOS). Less oxygenated compounds (CHO and CHNO) prevailed in the Tapajós river DOM. Large proportions of polyphenolic compounds and other highly oxygenated compounds (CHO and CHNO) were removed at the confluence of Solimões and Negro rivers due to sorption to mineral particles, as well as biotic and abiotic degradation of DOM molecules. Moreover, we found that heterotrophic bacterial production was stimulated downstream of the confluence of Solimões and Negro rivers and in the Tapajós River, which had significant correlation with DOM composition. Our mixing and incubation experiment conducted on the two sets of endmember river waters (the Solimões and Negro rivers; the Amazonas and Tapajós rivers) showed major distinction due to different effects of mineral sorption and microbial degradation. Mixing experiments showed a clear reversal of DOM molecular alterations within five days for SN and AT mixing experiments,



indicating remarkably fast microbial processing of DOM promoted by tropical temperatures. However, NMR spectra indicated large scale processing of about all DOM substructures in SN mixing, whereas primarily alkyl groups were processed in AT mixing experiments. This implies considerable distinction of DOM processing at SN and AT confluences, with less thorough alteration of molecules at the AT confluence.

## Kurzfassung

Das Einzugsgebiet des Amazonas ist das größte Flusssystem der Welt und für ca. 20 % des weltweiten Süßwassereintrags in die Ozeane verantwortlich; es exportiert jährlich ~35 Tg organischen Kohlenstoff in den Atlantischen Ozean. Gelöste organische Substanz (DOM) im Amazonas ist an kritischen Ökosystemprozessen beteiligt, wie z. B. der Nährstoffverfügbarkeit, der Wachstumseffizienz von Wasserorganismen, der Zersetzungsrate von Pflanzenresten und letztendlich dem Kohlenstoffkreislauf. Polydispersität, Heterogenität und zeitliche Dynamik von DOM machen seine molekulare Charakterisierung seit vielen Jahren zu einer erheblichen Herausforderung. Frühere Studien zum Ursprung und Verbleib von organischem Material im Amazonasbecken umfassten u.a. stabile Isotopen- und Radioisotopenanalysen, Quantifizierung von Zielverbindungen, optische Spektrometrie, Massenspektrometrie, <sup>13</sup>C NMR Spektroskopie. Allerdings fehlt bis heute ein umfassendes molekulares Verständnis der Reaktivität und Bioverfügbarkeit von Amazon River DOM, sowie eine Bewertung seiner biotischen und abiotischen Verarbeitung, insbesondere in Nebenflüssen mit unterschiedlichen Wassertypen und in großen Zusammenflusszonen.

Diese Arbeit liefert ein umfassendes molekulares Verständnis der Veränderungen von DOM-Molekülen im Hauptstamm des Amazonas (Amazonas), ausgewählten Hauptzuflüssen (Flüsse Solimões, Negro und Tapajós) sowie den Zusammenflüssen Solimões / Negro (SN) und Amazonas / Tapajós (AT). Ultrahochauflösende negative und positive Elektrospray-Ionisations-Fourier-Transformations-Ionenzyklotronresonanz-Massenspektrometrie (ESI[±] FT-ICR MS) lieferte eine komplementäre Abdeckung von DOM Molekülen und zeigte eine deutliche Diversität der Verarbeitung für Moleküle mit Stickstoff und Schwefel Heteroatomen (CHO, CHNO, und CHOS). Die 800 MHz <sup>1</sup>H Kernresonanzspektroskopie (NMR) ermöglichte die Quantifizierung der wichtigsten atomaren Umgebungen, die für molekular diversifiziertes wässriges DOM typisch sind, einschließlich aliphatischer Substrukturen, die nur indirekte Signaturen in Massenspektren und optischen Spektren liefern. Sowohl MS als auch NMR Spektren zeigten eine klare Unterscheidung der DOM Zusammensetzung und Strukturen in verschiedenen Wassertypen im Amazonasbecken. DOM im Fluss Negro enthielt mehr polyphenolische und andere stark sauerstoffhaltige Verbindungen, während DOM in den Flüssen Solimões und Amazonas mehr Stickstoff und Schwefel Verbindungen (CHNO und CHOS) enthielt. Im DOM des Tapajós Flusses herrschten weniger sauerstoffhaltige Verbindungen (CHO und CHNO). Große Anteile polyphenolischer Verbindungen und anderer stark sauerstoffhaltiger Verbindungen (CHO und CHNO) wurden am Zusammenfluss der Flüsse Solimões und Negro aufgrund der Sorption an Mineralpartikel sowie des biotischen und abiotischen Abbaus von DOM Molekülen entfernt. Darüber hinaus fanden wir heraus, dass die heterotrophe Bakterienproduktion stromabwärts des Zusammenflusses der Flüsse Solimões und Negro sowie im Fluss Tapajós stimuliert wurde,

was einen signifikanten Zusammenhang mit der DOM Zusammensetzung aufwies. Unser Misch- und Inkubationsexperiment, das an den beiden Gruppen von Zusammenflüssen (den Flüssen Solimões und Negro sowie den Flüssen Amazonas und Tapajós) durchgeführt wurde, zeigte große Unterschiede aufgrund unterschiedlicher Auswirkungen der Mineralsorption und des mikrobiellen Abbaus. Mischexperimente zeigten bei SN und AT Mischexperimenten eine deutliche Umkehrung der molekularen DOM Veränderungen innerhalb von fünf Tagen, was auf eine bemerkenswert schnelle mikrobielle Verarbeitung von DOM hinweist, die durch tropische Temperaturen gefördert wird. Allerdings deuteten NMR-Spektren darauf hin, dass nahezu alle DOM Unterstrukturen beim SN Mischen in großem Maßstab verarbeitet wurden, während bei AT Mischexperimenten hauptsächlich Alkylgruppen verarbeitet wurden. Dies impliziert einen erheblichen Unterschied in der DOM Verarbeitung an SN und AT Konfluenzen, mit einer weniger gründlichen Veränderung der Moleküle an der AT Konfluenz.

## Table of Contents

List of publications .....	x
List of figures .....	xi
List of tables .....	xii
List of abbreviations .....	xiii
1. General introduction.....	1
1.1 DOM.....	2
1.1.1 Definition of DOM .....	2
1.1.2 Composition of DOM .....	2
1.1.3 Function of DOM .....	6
1.2 Dynamics of DOM in the Amazon Basin.....	10
1.2.1 General information of the Amazon Basin .....	10
1.2.2 DOM in different Amazon waters .....	12
1.2.3 Allochthonous and autochthonous sources .....	17
1.2.4 Factors regulating DOM properties .....	18
1.3 Structural characterization of Amazon DOM.....	27
1.3.1 Biomarker analysis .....	30
1.3.2 Elemental analysis .....	32
1.3.3 Isotope analysis and radiocarbon analysis .....	33
1.3.4 Optical spectroscopy.....	35
1.3.5 Mass spectrometry .....	36
1.3.6 NMR spectroscopy .....	37
1.4 Scope of this work .....	40
2. Material and methods .....	42
2.1 Sampling and site locations .....	42
2.2 Sample processing of mixing experiment .....	45
2.3 FT-ICR mass spectrometry.....	46
2.4 <sup>1</sup> H NMR spectroscopy .....	47
2.5 Water characterization.....	47
2.6 Dark carbon fixation (DCF) and heterotrophic bacterial production (HBP) .....	48
2.7 Statistical analysis .....	49
3. Results .....	51

3.1 Article 1: Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy .....	51
3.2 Article 2: Distinct non-conservative behavior of dissolved organic matter after mixing Solimões/Negro and Amazon/Tapajós River waters.....	52
4. General discussion and outlook .....	53
4.1 Molecular insights into DOM in the Amazon River and its tributaries.....	53
4.2 DOM dynamics after riverine mixing .....	56
4.3 Environmental implications.....	60
4.4 Conclusion.....	63
5. References .....	64
6. Appendix .....	92
6.1 Article 1 + licenses for copyrighted content.....	92
6.2 Article 2 + licenses for copyrighted content.....	104
6.3 Supplementary information for article 1 .....	118
6.4 Supplementary information for article 2 .....	152

## List of publications

### Included in this thesis

Siyu Li, Mourad Harir, Philippe Schmitt-Kopplin, Michael Gonsior, Alex Enrich-Prast, David Bastviken, Juliana Valle, Fausto Machado-Silva, Norbert Hertkorn\* (2022). Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy. *Science of the total environment* 857, 159620.

Siyu Li, Mourad Harir, Philippe Schmitt-Kopplin, Fausto Machado-Silva, Michael Gonsior, David Bastviken, Alex Enrich-Prast, Juliana Valle, Norbert Hertkorn (2023). Distinct non-conservative behavior of dissolved organic matter after mixing different types of Amazon River waters. *ACS Environmental Science & Technology water* 3, 2083-2095.

### Additional publications

Miao Zhang, Siyu Li (equal contribution), Norbert Hertkorn, Mourad Harir, Dongfeng Xu, Philippe Schmitt-Kopplin, Ying Wu (2022). Substantial biogeochemical and biomolecular processing of dissolved organic matter in an anticyclonic eddy in the Northern South China Sea down to bathypelagic depths. *Frontiers in Marine Science* 9, 902728.

Siyu Li, Mourad Harir, David Bastviken, Philippe Schmitt-Kopplin, Michael Gonsior, Alex Enrich-Prast, Juliana Valle, Norbert Hertkorn (2023). Dearomatization drives complexity generation in freshwater dissolved organic matter. (under review)

## List of figures

Figure 1 Clearwater, whitewater, and blackwater river colors. ....	12
Figure 2 The distribution of major whitewater, blackwater, and clearwater rivers in the Amazon basin, adapted from Junk et al.2011(Junk et al., 2011).....	13
Figure 3 Maps of the sampling sites in the Amazon River. “SM” (grey) was sampled in the Solimões River upstream and close to the Negro-Solimões River mixing zone. “MZ” (black) was sampled at the very mixing zone. “A1R/A2R” (dark red) were sampled in the Amazon River upstream of the Madeira River inflow, while the other Amazon River samples (red) were sampled after the Madeira River inflow. “AM3Ra” and “AM9Ra” (magenta) were sampled two months later than the other water samples in the Amazon River. ....	43

## List of tables

Table 1 Molecular characteristics of DOM from different origin. ....	4
Table 2 Distinct geophysical and chemical characteristics, as well as DOM composition of whitewater, blackwater, and clearwater rivers in the Amazon basin. ....	14
Table 3 Water characteristics and DOM composition in different water bodies in the Amazon basin. .....	16
Table 4 The characterization approaches of DOM. ....	29
Table 5 Description of the sampling sites and sampling information. “SM” (grey) was sampled in the Solimões River close to the mixing zone of the Negro River and the Solimões River. “MZ” (black) was sampled at the mixing zone of the Negro River and the Solimões River. “A1R/A2R” (dark red) were sampled in the Amazon River upstream of the Madeira River inflow, while the other Amazon River samples (red) were sampled after the Madeira River inflow. “AM3Ra” and “AM9Ra” (magenta) were sampled two months later than the other water samples in the Amazon River. The order of the rows in the table from top to bottom represents the order of sampling point locations in each river from upstream to downstream. ....	43
Table 6 Description of samples in mixing and incubation experiments. ....	45



## List of abbreviations

1d: one day

5d: five days

A: the Amazonas River

A+T: the Amazonas and Tapajós Rivers

A-DOM: solid phase extracted dissolved organic matter from the Amazonas River

$B_0$ : main static magnetic field

BOD: biochemical oxygen demand

CDOM: chromophoric dissolved organic matter

CH<sub>4</sub>: methane

CHO: molecular formulae consist of carbon, hydrogen, and oxygen atoms

CHNO: molecular formulae consist of carbon, hydrogen, nitrogen, and oxygen atoms

CHOS: molecular formulae consist of carbon, hydrogen, oxygen, and sulfur atoms

C/N: carbon-to-nitrogen ratio

CI: chemical ionization

CO<sub>2</sub>: carbon dioxide

Cond.: water conductivity

CRAM: carboxylic-rich alicyclic molecules

d1: relaxation delay

DBE: double bond equivalent

DBE/C: double bond equivalent per carbon

DO: dissolved oxygen

DOC: dissolved organic carbon

DOM: dissolved organic matter

DCF: dark carbon fixation

DEPT: distortionless enhancement by polarization transfer

DESI: desorption electrospray ionization

DIC: dissolved inorganic carbon

EEM: excitation-emission matrix

ESI: electrospray ionization

ESI[±]: negative and positive electrospray ionization

FI: field ionization

FN: filtered Negro River water

FS: filtered Solimões River water

FT-ICR: Fourier transformation ion cyclotron resonance

GHGs: greenhouse gases

HBP: heterotrophic bacterial production

HCA: hierarchical clustering analysis

H/C: hydrogen-to-carbon ratio

H<sub>2</sub>O: water

ID: isotope dilution

LI: leucine incorporated

MS: mass spectrometry

*m/z*: mass-to-charge ratio

MZ: the mixing zone of the Solimões and Negro Rivers

N: the Negro River

N-DOM: solid phase extracted dissolved organic matter from the Negro River

N/C: nitrogen-to-carbon ratio

NH<sub>4</sub><sup>+</sup>: ammonium

NO<sub>3</sub><sup>-</sup>: nitrate

NO<sub>2</sub><sup>-</sup>: nitrite

NOM: natural organic matter

NMDS: non-metric multidimensional scaling

NMR: nuclear magnetic resonance

O/C: oxygen-to-carbon ratio

O<sub>2</sub>: oxygen gas

PARAFAC: parallel factor analysis

pCO<sub>2</sub>: partial pressure of CO<sub>2</sub>

PI: photo ionization

PCA: principal component analysis

POC: particulate organic carbon

POM: particulate organic matter

PO<sub>4</sub><sup>3-</sup>: phosphate

PPL: styrene-divinylbenzene polymer

SPE: solid phase extraction

S: the Solimões River

S+N: the Solimões and Negro Rivers

SN: suspended solids in the Negro River water

SS: suspended solids in the Solimões River water

S/C: sulfur-to-carbon ratio

S-DOM: solid phase extracted dissolved organic matter from the Solimões River

T: the Tapajós River

T-DOM: solid phase extracted dissolved organic matter from the Tapajós River

Temp.: temperature

TCA: trichloroacetic acid

TOC: total organic carbon

UV: ultraviolet

UV-vis: ultraviolet and visible spectrophotometer

WRT: water residence time

XANES: X-ray absorption near edge spectroscopy

# 1. General introduction

Freshwater dissolved organic matter (DOM) is very abundant in terrestrial ecosystems and ranges among the most complex mixtures known on Earth (Einsiedl et al., 2007; Hertkorn et al., 2008). DOM plays pivotal roles in local and global ecosystems as food and energy sources for heterotrophic microorganisms and is an important compartment of dynamic carbon and nutrient cycling. The complex nature of DOM in the composition is attributed to the different relative contributions of various DOM sources in combination with the subsequent biogeochemical transformations. Owing to this complexity, the mechanistic and molecular understanding of DOM processing in freshwater ecosystems is seriously underdeveloped.

The Amazon River catchment is the largest river system in the world, responsible for ~20 % of the global freshwater discharge and exports ~35 Tg organic carbon (of which 60-70 % is dissolved) annually to the Atlantic Ocean (Moreira-Turcq et al., 2003a). Considering the magnitude of the Amazon River basin, an ecosystem-wide understanding of DOM molecular composition and structure is needed to comprehend the various and probably distinct key biogeochemical processes in its key sections.

A pretreatment to increase sample concentration is typically necessary for most of the DOM analytical measurements except for optical methods. Solid phase extraction (SPE) based on PPL sorbents has been frequently used for this purpose (Dittmar et al., 2008; Hertkorn et al., 2016; Herzsprung et al., 2017a; Raeke et al., 2016). Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry (MS) coupled to electrospray ionization (ESI) enables the characterization of SPE-DOM with very high mass accuracy and mass resolution to determine its chemical composition with excellent coverage of thousands of  $m/z$  ions, each of which probably represents a mixture of different isomers (Hertkorn et al., 2008; Hertkorn et al., 2013; Schmitt-Kopplin et al., 2019a). So far, the ESI source was largely operated in the negative ion mode in most DOM studies because DOM mainly contains functional groups such as alcohols, carboxylic acids, cyanides, nitrates, sulfonates, and phosphates, which can lose a proton in the negative ion mode (Liu et al., 2021). However, the use of positive ESI mode can be very beneficial and complementary for DOM studies, as nitrogen-containing compounds get better ionized in positive ESI mode (Hertkorn et al., 2008; Koch and Dittmar, 2016; Kujawinski et al., 2004; Kujawinski et al., 2002).

While FT-ICR mass spectrometry has considerable selectivity depending on sample handling and ionization method applied, NMR on the other hand allows for a successive acquisition of multiple informative complementary NMR spectra (Hertkorn et al., 2002a; Simpson et al., 2002; Simpson et al., 2011). 800 MHz one-dimensional  $^1\text{H}$  NMR spectroscopy provided near-quantitative coverage of major atomic environments in DOM composition with excellent S/N ratio and resolution, enabling quantification of key DOM structural units including aliphatic and carboxyl-rich alicyclic molecules (CRAM), oxygenated aliphatics including carbohydrates, olefins, and aromatics (Hertkorn et al., 2016; Li et al.,

2016). Higher dimensional NMR spectra of DOM provide improved detail information in determining molecular structures, particularly from mixtures of rather sizable molecules like DOM (Hertkorn et al., 2016; Powers et al., 2019).

## **1.1 DOM**

### **1.1.1 Definition of DOM**

Natural organic matter (NOM) is an exceedingly complex mixture of organic materials in natural environments (e.g., water, soil, sediment, and atmosphere) generated through the combined action of biotic and abiotic pathways. Present in soil and within all marine and freshwater sources, NOM constitutes one of the Earth's largest carbon reservoirs, comparable to atmospheric CO<sub>2</sub> (624 and 750 Tg, respectively) (Carlson and Hansell, 2015). In addition, NOM is an exceedingly complex mixture of organic molecules with a large diversity of heterogeneous molecular motifs, intermolecular interactions, and molecular weights. In comparison with biomolecules, NOM has higher overall abundance on Earth, and a more extensive molecular heterogeneity and polydispersity. The quantity and molecular diversity of NOM exceeds that of all living matter by several orders of magnitude in all ecosystems. The composition and chemical structure of polydisperse and molecularly heterogeneous NOM reflect the respective ecosystem history, characteristics, and processes.

Dissolved organic matter (DOM) is operationally defined as a fraction of NOM which passes through a certain filter (pore size: 0.1-1.0 μm, mostly 0.45 μm), and those materials retained on the surface of the filter being particulate organic matter (POM) (Verdugo et al., 2004). Using this operationally but arbitrary separation, DOM includes colloids and small particles (some living) and a wide range of hydrophilic and hydrophobic molecules. In addition, the proportion of DOM and POM can change, for instance, when the water samples for analysis are shaken, because of particle formation, or disruption (Pignatello, 2012). In the main channel of the Amazon River, Quay et al. found that 20% of the total carbon was particulate organic carbon, 30% was dissolved organic carbon, and 50% was dissolved inorganic carbon (Quay et al., 1992).

### **1.1.2 Composition of DOM**

#### **1.1.2.1 Structures of DOM**

DOM is an extremely complex mixture of organic compounds that collectively exhibits a nearly continuous range of properties (size-reactivity continuum) (Amon and Benner, 1996a). Generally, high-molecular weight (>1000 Da) and low-molecular weight (<1000 Da) fractions constitute around 25% and 75% of total DOM, respectively (Benner and Amon, 2015). DOM is composed mainly of carbon, which is referred

to as dissolved organic carbon (DOC), but also includes hydrogen, and oxygen, and minor contributions from heteroatoms (e.g., nitrogen, sulfur, and phosphorus) (Hertkorn et al., 2008), as well as trace amounts of various cations including calcium, potassium, magnesium, and metals including aluminum, iron, zinc, and copper (Borggaard et al., 2019; Hirose, 2007; Philippe and Schaumann, 2014a). DOM constituents consist of molecules such as aromatics, aliphatics, and derivatives of lignins, tannins, carbohydrates, amino acids, peptides, lipids, carotenoids, terpenoids, and CRAM (Hertkorn et al., 2006), and their quantities vary significantly in different water systems and at different times (Jaffé et al., 2008; Lee et al., 2018; Repeta et al., 2002; Xu and Guo, 2017).

Terrestrial polyphenols like lignin- and tannin-like compounds are rather refractory and remain largely unaltered in water bodies for years (McDonald et al., 2004). Their coloration creates “blackwater” rivers, such as the Negro River (Janzen, 1974). Additionally, lignin phenols are exclusively derived from vascular plants and have been utilized to trace the transport of terrestrial DOM in rivers and oceans in many studies (Bao et al., 2015; Hedges et al., 1994; Hernes, 2003). The structural changes of lignins during chemical and biological degradation include preferential loss of methoxylated structural units and oxidative degradation of lignin side chains (Argyropoulos and Sun, 1996; Janusz et al., 2017; Lanzalunga and Bietti, 2000). CRAM is comprised of a complex mixture of carboxylated and fused alicyclic structures and incorporate major refractory components of terrestrial and marine DOM (Hertkorn et al., 2006). CRAM show structural similarities to sterols and hopanoids, and are actively involved in ecosystem processes such as complexation with metals, aggregation, and formation of marine gels (Hertkorn et al., 2006).

In contrast to apparently refractory compounds, other more labile fractions of DOM, like simple sugars and amino acids are rapidly utilized and transformed by the biota for metabolism or growth (Moran and Zepp, 1997). Carbohydrates make up one of the most abundant DOM pools, accounting for up to 50% of the total DOM components in surface marine waters and less than 30% in deep marine waters (Benner et al., 1992). Carbohydrates are constituted of polyhydroxy aldehydes and ketone, serving diverse functions in DOM activities such as storage of energy, structural components, important component of coenzymes, backbone of the genetic molecule, formation of mucilaginous aggregates (Biddanda and Benner, 1997). Moreover, the proportions of certain neutral sugars, in particular glucose and the deoxysugars, can be used as indicator of the diagenetic state of the extracted DOM (Engbrodt and Kattner, 2005). Amino acids contain amide, amine, and carboxyl functional groups, along with aliphatic sidechains. Amino acids actively participate in protein synthesis, neurotransmitter transport and biosynthesis and have been widely applied as molecular indicators for bioavailable DOM in marine and groundwater systems (Benner and Kaiser, 2010).

Lipids make up a small portion of DOM and play a significant role in the DOM pool like storing energy, signaling, and acting as structural components of cell membranes. Lipids have been widely employed as

biomarkers for terrigenous and planktonic organic matter in various aquatic and sedimentary environments. Lipids consist of structurally heterogeneous groups of hydrophilic and hydrophobic small molecules such as fatty acids, glycerolipids, glycerophospholipids, sphingolipids, saccharolipids, polyketides, sterol lipids and prenol lipids, and are chemically highly saturated, with high proportions of methylene carbon on average, and less oxygenated than other main classes of biomolecules. Preferential degradation of carbohydrates and proteinaceous compounds and preferential accumulation of lipid-derived alkyl compounds with increasing age of DOM were found in several studies (Olk et al., 2019).

### 1.1.2.2 DOM composition in different ecosystems

The proportion of different constituents of DOM is commonly variable in different ecosystems. Considerable variance is observed in structures among freshwater DOM of different origins such as rivers, estuaries, lakes, and mangroves, and refer to e.g., lignins, tannins, lipids, and aromatics; especially highly oxygenated tannins (O/C ratio: 0.67-1.0; H/C ratio: 0.6-0.8) are depleted in coastal DOM (Sleighter and Hatcher, 2008). However, marine DOM was found to share ~1/3 of formulae with terrestrial DOM, which might be attributed to the presence of refractory DOM or arise from isomers with different chemical structures for the same formulae (Hertkorn et al., 2007a; Koch et al., 2005; Sleighter and Hatcher, 2008). The averaged elemental ratios (H/C ratio: 1.01-1.09, O/C ratio: 0.53-0.56) in wetland DOM shift to more oxygenated and less saturated compared in lake DOM (H/C ratio: 1.08-1.16, O/C ratio: 0.42-0.43) (Koch et al., 2005). In general, soil DOM showed average O/C ratios of 0.28-0.47 (mean 0.40), average H/C ratios of 1.30-1.74 (mean 1.50) and molecular weights of 380-417 Da (mean 395 Da) (Ohno et al., 2010).

The detailed general molecular characteristics of different DOM samples are summarized in Table 1.

*Table 1 Molecular characteristics of DOM from different origin*

<b>DOM type</b>	<b>Sources of DOM</b>	<b>Characteristics of DOM</b>	<b>Influencing factors of DOM</b>
River DOM	Autochthonous material includes self-maintenance and metabolism of biota, and natural decay, whereas allochthonous matter is leached from plants and soils, as well from human discharge (Kördel et al., 1997).	Main DOM structural components consist of aliphatics, CRAM, carbohydrates, lipids, peptides, and small amounts of aromatics (Sleighter and Hatcher, 2008). Processed riverine DOM was thought to be relatively stable, representing the degraded remains of vascular plant materials aged in soils (Hedges et al., 1997).	Significant seasonal and spatial DOM variations affected by land use cover and precipitation; loss of DOM occurs by sorption on mineral surfaces and microbial breakdown (Singh et al., 2014). DOM presents a more terrestrial character at higher discharge rates, and more relevant microbial imprint at lower discharge rates (Raymond and Spencer, 2015).

Lake DOM	Besides its own autochthonous OC production, lakes and wetlands receive allochthonous input directly from terrestrial sources and via rivers and streams.	Lake DOM includes lipids, peptides, carbohydrates, N- and S-containing compounds, lignin- and tannin-like compounds (Goldberg et al., 2015; Gonsior et al., 2013a; Kellerman et al., 2014). The oxidized aromatic compounds are preferentially removed during organic matter mineralization (photo- or biodegradation), while rather saturated aliphatic and N-containing compounds are more likely to persist in aquatic systems due to their recalcitrant characteristics (Goldberg et al., 2015).	Precipitation, inflow and outflow of tributaries, water residence time and temperature (Goldberg et al., 2015), light, oxygen, microbial community affect DOM. More proteins are produced in spring or summer time, a result of higher metabolic activity following higher temperatures (Goldberg et al., 2015). Due to the prolonged water residence time, lake DOM is more prone to photochemistry, flocculation and further sedimentation. There is an impressive number of ubiquitous features in different lake DOMs (Kellerman et al., 2014).
Soil DOM	Nearly all soil DOM derives from plant polyphenols, metabolism within food chains, photosynthesis, including recently photosynthates – from leaf litter to root exudates – and decomposition and metabolic by-products of microbiologically processed soil organic matter (Bolan et al., 2011).	Soil DOM consists mainly of low molecular weight carboxylic acids, amino acids, carbohydrates, and fulvic acids (Bolan et al., 2011). Carboxylic acids represent less than 10% of total DOM in most soil solutions and fulvic acids are typically the most abundant fractions of DOM (Bolan et al., 2011).	pH, OC content, microbial activity, temperature, and moisture (Bolan et al., 2011), microbial community, soil properties, root biomass (Roth et al., 2014), and mineral particle attachment (Kögel-Knabner and Amelung, 2021), spatial heterogeneity and temporal variability. Phenolic tannin compounds may be more abundant at a lower pH (Roth et al., 2014; Roth et al., 2015). DOM in the upper soil profile has a vegetation-type signature, with lignin-derived phenols and plant-derived carbohydrates, while most subsoil DOM is of microbial origin (Kaiser and Kalbitz, 2012; Schmidt et al., 2011).
Groundwater DOM	Surface plant litter and soil are important sources of groundwater DOM column (Einsiedl et al., 2007). Lignin phenols, known biomarkers of plant-derived DOM, were substantially removed by mineral interaction (Shen et al., 2015). Small and hydrophilic carbohydrates and amino acids eluted faster to the groundwater.	Lower molecular weights compared to surface DOM (Shen et al., 2015). Depletion of oxygen resulted in decrease of aromatic molecules and carbohydrates together with increase of aliphatic compounds (Einsiedl et al., 2007). Contains small portions of both bioavailable and bacterial-derived DOM (Shen et al., 2015).	Depending on metabolic rates and contact time between water and sediments, organic matter mineralization proceeds via a set of microbial mediated alternative terminal electron-accepting processes, which strongly influence biogeochemical constituents and metabolic rates in the subsurface. Interactions between DOM and minerals during infiltration determine the selectivity of compounds retention within the soils (Hedges et al., 1994; Qualls and Haines, 1992).



Sediment DOM	Sediment DOM in coastal zones is a mix between terrestrial-derived material and marine-derived DOM (Chen and Hur, 2015). Freshwater sediments tend to present higher concentrations of DOC and is dominated by terrestrial-derived and microbial sources (Chen and Hur, 2015).	Sediment DOM includes a wide range of molecules, such as lipids, lignins, proteins, humic- and fulvic-like compounds, unsaturated hydrocarbons, condensed aromatics, and carbohydrates (Li et al., 2022). DOM samples from different systems carry a large number of unique CHNO and CHOS formulae, reinforcing the heterogeneity of the sediments along a given gradient (Schmidt et al., 2017; Wang et al., 2013; Xu et al., 2016).	DOM in the water column can form aggregates and particles through microbial and photochemical processes, later sinking into the sediments. Most of DOM accumulated in deeper sediment layers is believed to be overall recalcitrant (Burdige and Komada, 2015).
Marine DOM	Most marine DOM has four main sources: the transportation by river and wind (insignificant with respect to total amounts, except in coastal regions of marginal seas), production by the decay of dead organisms, excretion by algae/plants and excretion by marine animals (Kördel et al., 1997). Euphotic zone is the principal site of marine organic matter production. There are complementary mechanisms such as extracellular release by phytoplankton, grazer-mediated release and excretion, release via cell lysis, solubilization of detrital and sinking particles, and release from prokaryotes (Carlson and Hansell, 2015).	Amino acids and proteins, simple sugars, vitamins, fatty acids, polysaccharides, CRAM, lignin-like compounds, and very complex degradation products of unknown origin, such as humic substances and black carbon (Koch et al., 2005). Approximately 1/3 of the formulae were common between marine and terrestrial DOM, which might be attributed to the presence of refractory DOM or arise from isomers with different chemical structures at the same formulae (Koch et al., 2005; Sleighter and Hatcher, 2008).	Marine DOM can be transformed via biotic or abiotic pathways (Flerus et al., 2012). Both light availability for photosynthesis and the penetration of UV radiation within the marine environment are influenced by abiotic photochemical transformations

### 1.1.3 Function of DOM

DOM is major reactant in and a product of many diverse biogeochemical processes (Leenheer and Croué, 2003). DOM provides organic substrates, nutrients, trace elements, and energy sources to microbes, acts as electron donor and acceptor (Carlson and Hansell, 2015), and carries the dominant flux of carbon and nutrients in many aquatic ecosystems (Battin et al., 2009; Bernhardt and Likens, 2002; Cole et al., 2007;

Fisher and Likens, 1973; Meyers and Lallier-Vergès, 1999). DOM is involved in a wide range of ecosystems services in aquatic ecosystems, such as e.g., soil aggregation and erosion control, acid-base chemistry, mobilization and export of nutrients, bioavailability and ecotoxicology of heavy metals, and photochemistry of natural waters (Bolan et al., 2011). DOM is responsible for sustaining an important part of food webs, both via direct energy transfer and nutrient availability (Bolan et al., 2011; Carlson and Hansell, 2015). Low-molecular-weight (<1000 Da) molecules can be transported across cell membranes, whereas high-molecular-weight (>1000 Da) molecules usually require enzymatic cleavage prior to uptake (Weiss et al., 1991). DOM has other important functions including 1) binding with other organic molecules (e.g., lignin binding with cellulose, hemicellulose, or pectin; tannin binding with protein or pigments) as structural components for plants or influencers for microbial metabolism (Dashtban et al., 2010; Kraus et al., 2003); 2) binding with metallic ions (e.g., tannin or CRAM binding with iron) required for organism growth. DOM is able to interact with different contaminants (such as heavy metals) modifying their bioavailability and toxicity (Kungolos et al., 2006); 3) blocking light penetration thus protecting microbes from absorbing harmful UV radiation and controlling primary production (Dong et al., 2021; Häder et al., 2015; Nelson and Siegel, 2013; Núñez-Pons et al., 2018; Patel-Sorrentino et al., 2004; Sulzberger and Durisch-Kaiser, 2009); 4) scavenging radicals that are harmful for organisms (Dong et al., 2021; Remucal et al., 2020; Zhang and Parker, 2018); 5) controlling levels of dissolved oxygen, nitrogen, phosphorus, sulfur (Ksionzek et al., 2016), numerous toxic and nutrient trace metals, acidity, and salinity (Leenheer and Croué, 2003), and thus shaping the composition of microbial community (Blanchet et al., 2017; Killips and Killips, 2013). Numerous experimental and field studies have proven the correlative relationships between DOM quality, microbial community structure, and metabolic potential (Alonso-Sáez et al., 2012; Carlson and Hansell, 2015; Landa et al., 2013).

### **1.1.3.1 Ecological functions of DOM**

The DOM fraction is an important active and bioavailable C source for microbial biomass, besides sequestering and storing carbon in deep layers. In long-term studies, De Troyer et al. (De Troyer et al., 2011), Fröberg et al. (Fröberg et al., 2003), and Hagedorn et al. (Hagedorn et al., 2004) found that organic carbon from plant residues did not accumulate in the DOC pool; instead, it was mostly released as CO<sub>2</sub>. However, Uselman et al. (Uselman et al., 2007) found that during high rainfall and low temperatures, a larger fraction of <sup>14</sup>C from plant litter was lost as DOC, translocated or leached, rather than released as CO<sub>2</sub>, probably favoring leaching more than microbial metabolism. Deng et al. (Deng et al., 2017) showed that DOC leaching from the litter layer to topsoil in a subtropical forest was the major cause of rain-induced soil CO<sub>2</sub> pulse; consequently, there is great concern with a potential DOC contribution to increasing CO<sub>2</sub> release in tropical soils, due to the increase in DOC fluxes by accelerated microbial activity.

The functional groups in DOM, such as carboxylic and phenolic units, participate in many chemical reactions in soil such as organic metal complexation, increasing the ion adsorption rate and metal detoxification (Franchini et al., 2003; Roberts, 2006). These acidic groups participate in exchangeable Al complexes in the soil solution, making it nontoxic to plants (Amaral et al., 2004; Franchini et al., 2003). Therefore, in tropical soils, these organic acids can compete with other ions, such as phosphate ions, for adsorption sites (Andrade et al., 2003; Jones, 1998).

Humic substances have a great potential to form water-soluble complexes with toxic metals and organics, affecting the bioavailability of these toxicants and potentially increasing their transportation range (Killops and Killops, 2013). Humic compounds have been recognized as an efficient transport for metals and nutrients (Lipczynska-kochany, 2018).

Generally, DOM originating from plant residues is easily degradable and quickly decomposed rather than leached through deeper soil horizons (Amery et al., 2007). The effects of soil pH on DOM mobility remain uncertain. Nonetheless, Tipping and Woof (Tipping and Woof, 1990) reported reduced adsorption capacity at high pH values with an increase in DOM mobilization. Consequently, small increases in the soil pH led to higher amounts of mobilized soil organic matter.

DOM dynamics and processes are mainly affected by adsorption in the soil mineral phase; it is more strongly adsorbed by Fe and Al oxides, higher specific surface area clays, and polyvalent cations. Moreover, DOM is important in nutrient cycling and distribution in the profile, in phosphate availability, and in complexation of Al and Fe, heavy metals, and pollutants. In tropical conditions, DOM is possibly strongly adsorbed by Fe and Al oxides; however, fast production and changes of DOC can boost CO<sub>2</sub> emission.

### **1.1.3.2 Interactions of DOM with elements**

Due to its high mobility, the DOM movement is significant to the cycling and distribution of nutrients, such as N and P (Veum et al., 2009) and Fe and Al complexes (Fujii et al., 2009) in ecosystems. While N cycling is controlled mainly by biological reactions, both abiotic and biotic processes control P and S dynamics (Bolan et al., 2011).

The change in concentration of nitrogen and DOM caused changes in phytoplankton biomass and community structure. The effects of nitrogen enrichment were depended on nutrient limitation status of the boreal lake, whereas the effects of DOM enrichment were observed across lakes regardless of nutrient limitation status, suggesting that increasing DOM may have widespread effects on aquatic systems (Daggett et al., 2015).

Phosphate limitation greatly influenced both the amount and the composition of the secreted DOM

molecules by heterotrophic bacteria. Under phosphate limitation, the composition of the exo-metabolome changed during bacterial growth, showing an increase in phenols and polyphenolic compounds, while under phosphate surplus conditions the secreted DOM molecules were mainly peptides and highly unsaturated compounds (Romano et al., 2014).

Dissolved organic sulfur constitutes a large sulfur reservoir that actively involved in biogeochemical cycles (Ksionzek et al., 2016). Sulfur-containing compounds, such as dimethyl sulfide, dimethyl sulfoniopropionate, carbonyl sulfide, dimethyl sulfoxide, dimethyl disulfide, carbon disulfide, methane thiol, cysteine, glutathione, phytochelatins and methionine can be formed during photochemical transformation of DOM (Mopper, 2002).

The levels of natural organic chlorine typically exceed levels of chloride in most soils and is therefore clearly of high importance for continental chlorine cycling (Svensson et al., 2022). Formation of chlorinated organic compounds may affect the degradation of soil organic matter and thus the carbon cycle with implications for the ability of forest soils to sequester carbon, whilst the occurrence of potentially toxic chlorinated organic compounds in groundwater aquifers is of concern with regard to water quality (Clarke et al., 2009). Transformations of marine chloride to non-volatile organochlorine through biotic and abiotic pathways were observed. The organochlorine was revealed to present primarily in concentrated aliphatic forms consistent with lipid chlorination, along with a more diffuse aromatic fraction (Leri et al., 2015). In addition, organochlorine that mainly contained one or two chloride atoms was observed during chlorination along water treatment (Zhang et al., 2012).

Iron is an important micronutrient for plants and microorganisms in the ecosystem, and is found to be strongly associated with DOM. The ferric iron could be stabilized by DOM as small iron oxide colloids (able to pass a 0.45  $\mu\text{m}$  dialysis membrane), which affects its bioavailability. In addition, a positive correlation of DOM and iron concentration was observed, and the coupling of DOM with iron fluxes could imply that DOM exports was caused in part by increased iron reduction (Knorr, 2013). Reactive Fe and organic carbon are intimately associated in soils and sediments. In acid forest soils, Al and Fe can form relatively stable complexes with DOM and may thus be mobilized and transported in the soil profile (Marschner and Kalbitz, 2003). The interaction of metal ions with organic matter is mainly affected by redox potential and pH. The oxidation state of a metal ion and the availability of anions that may compete with organic moieties for the metal are controlled by these factors. For example, lowering the acidity to pH  $\sim$ 2 can release most of the humic-bound iron (Fe), because carboxylic group become protonated to COOH making it difficult for metal ions to coordinate (Killops and Killops, 2013).

Previous studies have assumed that complexed DOM is stabilized against biodegradation due to toxic effects from the metals, especially Al, but experimental data indicated that complexation of potentially toxic metals by DOM would not necessarily reduce its biodegradability but instead could even enhance

microbial activity by reducing the free metal ion concentrations and thus their toxicity (Marschner and Kalbitz, 2003). Both inhibiting, enhancing and, null effects of Al additions on the mineralization of different DOM fractions from forest floor solutions have been reported (Jandl and Sletten, 1999).

## **1.2 Dynamics of DOM in the Amazon Basin**

### **1.2.1 General information of the Amazon Basin**

Amazon is the world's largest tropical rainforest ecosystem with a rich mosaic of different ecological conditions and organic matter sources like big rivers, streams, standing waters, and ground waters. The Amazon basin drains a vast and heterogeneous region of northern South America, encompassing large areas of seven nations and important population centers such as Manaus, Porto Velho, Iquitos, Cusco, and La Paz with very dissimilar environments, climates, and ecosystems (Gibbs, 1967b; Junk et al., 2006; Mayorga and Aufdenkampe, 2002; Richey et al., 2002b; Salo et al., 1986). The Amazon River is 6771 km long (Quay et al., 1992; Wissmar et al., 1981), 5 km in width, averages 20 m in depth (maximum 100 m) and drains an area  $6.3 \times 10^6$  km<sup>2</sup>. The Amazon River is the largest river system in the world with respect to both drainage area and volume of discharge (Gibbs, 1967b), responsible for ~20% of the global freshwater discharge ( $\sim 1.75 \times 10^5$  m<sup>3</sup>/s), exporting of  $3.15 \times 10^8$  Mg sediment and ~35 Tg organic carbon annually to the Atlantic Ocean, playing a significant continental and global role in biogeochemical cycles (Moreira-Turcq et al., 2003a). The climatic zones of the Amazon Basin extend from tropical rain forest to arctic mountainous, and comprise one of the widest ranges found in any single river basin in the world (Gibbs, 1967b). Mean annual temperatures range from the low temperature (<15°C) of the northern Andes to the high temperatures (26-28°C) of the tropical rain forests. Despite the acceleration of human impact in the form of deforestation, pasture and soya plantations, mining and oil extraction, dam construction, and urbanization, larger tracts remain relatively pristine, and the main river channels remain largely unaffected by engineering (Athayde et al., 2019; Castello et al., 2013; Confalonieri et al., 2014; Constantine et al., 2014; Correa et al., 2008; Crespo-Lopez et al., 2021; Latrubesse et al., 2017; Lees et al., 2016; Richey et al., 2002a; Sponsel, 1986; Swierk and Madigosky, 2014).

The Amazon originates in the Andean Cordillera, a region of high relief developed mainly in thinly bedded sedimentary and volcanic rocks. The combination of steep slopes and weak rocks favors channel incision, rapid mass wasting, and high sediment yields. After leaving the Andean foothills, the tributaries of the Amazon cross the adjoining foreland basin where they deposit large volumes of sediment, and then converge to flow along downward regions of low relief and gentle gradients, mantled with deep saprolite and dense equatorial forest. Afterwards, the rivers in Amazon trough are bordered by a plain of about 90,000 km<sup>2</sup> between São Paulo de Olivença and Óbidos, including unconsolidated fluvial and lacustrine sands, silts, and clays.

Heavy seasonal rainfall, concentrated in the Eastern Andes and the Northwest area of the Basin is

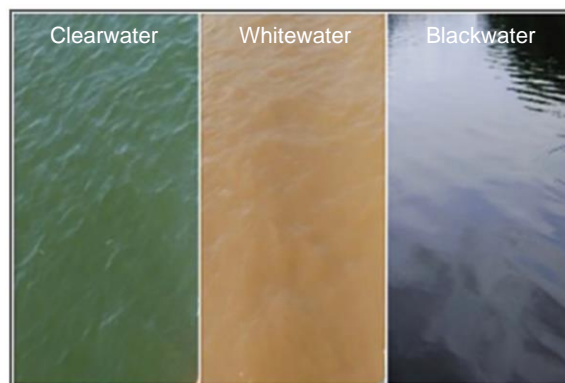
responsible for annual floods (Benner et al., 1995; Davidson et al., 2012; Fassoni - Andrade et al., 2021). Seasonal flooding is characteristic of the Amazon River and many other tropical rivers. In the Amazon basin, the lowest flood stage occurs in August and September, while the highest flood stage occurs in April and May, which commences in the southern portion of the basin and progresses northward (Espinoza et al., 2022; Gibbs, 1967b). As large areas within the Amazon basin are mostly flat, water swells outside riverbanks into floodplains in high-water-season and facilitates the deposition and mobilization of e.g., terrestrial materials, creating new water bodies, such as ponds and oxbow lakes.

The Amazon Basin is one of the most biodiverse areas in the world for most taxonomic groups (Zapata-Ríos et al., 2022). Diversity varies geographically, with some groups being more diverse in the Amazonian lowlands, whereas others thrive in the Andes. Scientists have not been able to estimate, to the nearest order of magnitude, the number of species in the Amazon. Although the Amazon includes one of the largest forests in the world, it is also one of the least known biologically. Documenting its biodiversity is challenging because of its immense size, heterogeneity, and limited access.

The diversity of geological, climatic, and biological characteristics throughout the Amazon basin generates a strong gradient in erosion and weathering regimes, which in turn affects the quality of the water transported by rivers (Bouchez et al., 2017).

## 1.2.2 DOM in different Amazon waters

### 1.2.2.1 Classification by water colors



*Figure 1 Clearwater, whitewater, and blackwater river colors.*

The Amazon River connects with several major tributaries of distinct chemical characteristics along the flow path. In the 1950s, Sioli developed a scientific classification of Amazon waters into three types according to their appearance (Figure 1), determined by the geomorphological and pedologic conditions in the headwaters, as well as physical and chemical transformations during transportation (Junk, 2001; Silva et al., 2013). This classification has been supported by botanists and limnologists, who found corresponding differences in the occurrence of tree species, aquatic macrophytes, and aquatic biota, such as snails, bivalves, and other species (Junk et al., 2011). 1) “Blackwaters” that are relatively acidic (pH ~ 5), low in total cations, and rich in DOM, such as the Negro, Jutaí, Tefé, and Coari rivers (Figure 2). It has been suggested that drainage areas of black water systems are characterized by moist, acidic, hydric soils that allow for the leaching of terrestrially derived plant matter, like lignins, tannins, and other plant materials that also contribute to the chromophoric dissolved organic matter (CDOM) (Leenheer, 1980; Raymond and Spencer, 2015). 2) “Whitewaters” that show a near-neutral pH and are relatively rich in total cations and in suspended sediments, such as the Solimões, Madeira, Juruá, and Purus rivers (Figure 2). The “white” color reflects a high mineral particle loads due to drainage and erosion of calcic sedimentary deposits originating largely from the Andes mountains. 3) “Clearwaters” that exhibit low suspended sediment loads, and high light transparency, such as the Tapajós, Trombetas, and Xingu rivers (McClain and Naiman, 2008) (Figure 2). “Clearwaters” drain kaolinite clays and contain high concentrations of iron and aluminum oxides that may adsorb humic acids. “Clearwaters” are less light-limiting and can support higher phytoplankton biomass, if the generally low nutrient levels are elevated (Costa et al., 2012; Richey et al., 1990).

More recent studies indicated that the chemical composition of waters in the Amazon basin is far more complex than envisioned by Sioli (Ríos-Villamizar et al., 2013), as many rivers and streams are actually

"mixed waters" that do not fit into the three classical categories, resulting from the mixing of lower-order tributaries with different physicochemical characteristics (Gibbs, 1967a; Ríos-Villamizar et al., 2013). For example, many of Amazon rivers have white water during most time of the year but have clear water in low flow season. Also, some black water rivers become white in rainy season when they gain more sediment from the surroundings. However, Sioli's classification remains widely used because of its simplicity.

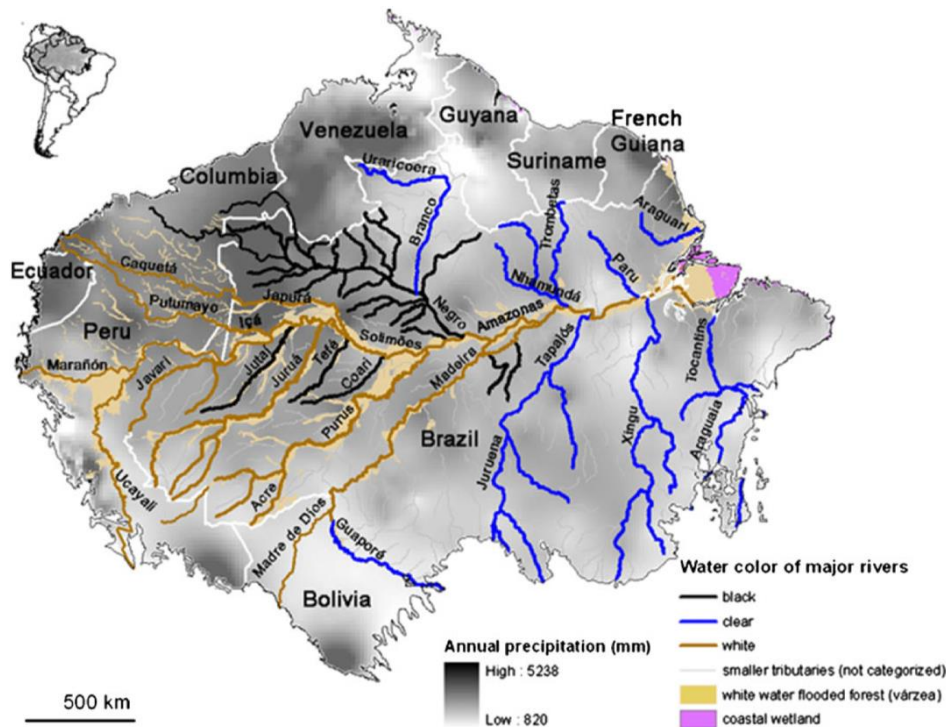


Figure 2 The distribution of major whitewater, blackwater, and clearwater rivers in the Amazon basin, adapted from Junk et al.2011 (Junk et al., 2011).

Different water types in the Amazon basin have been shown to have distinct DOM compositions (Gonsior et al., 2016; Quay et al., 1992). “Blackwaters” such as the Negro River are abundant in high molecular weight and diverse hydrogen-deficient molecules, indicating substantial oxidative processing of ultimately plant-derived polyphenols. In contrast, unique signatures found in “Whitewaters” were defined by presumably labile sulfur- and nitrogen-containing molecules in the river system. “Clearwaters” such as the Tapajós River showed a lower organic molecular complexity, a higher average H/C elemental ratio, and a lower average O/C elemental ratio compared to “Blackwaters” and “Whitewaters” (Gonsior et al., 2016).



Table 2 Distinct geophysical and chemical characteristics, as well as DOM composition of whitewater, blackwater, and clearwater rivers in the Amazon basin.

Water types	Geophysical and chemical characteristics	DOM composition
Whitewaters	drain from the Andes Mountains; have relatively high concentration of nutrient-rich sediments (Gonsior et al., 2016; Quay et al., 1992); the mainly alkaline earth metals and carbonates determine the muddy coloration of whitewaters; relatively high electrical conductivity; pH $6.6 \pm 0.2$	comprised of diverse CHNO and CHOS molecules; low molecular weight, low H/C ratio (Gonsior et al., 2016)
Blackwaters	originate in the central-northern rainforest, have high concentration of humic substances, which determine the reddish-brown color (Gibbs, 1967a); poor in suspended sediments and nutrients (Quay et al., 1992); pH $4.5 \pm 0.9$	lower proportions of CHNO and CHOS compounds, a large diversity of polyphenolic-type compounds with high O/C (0.5–0.8) and low H/C (0.4–0.7) ratios, extended and conjugated $\pi$ systems, unique aromatic high molecular weight compounds (Gonsior et al., 2016)
Clearwaters	drain the neighboring Brazilian shield region (craton, i.e., precambrian rocks); poor in sediments and organic matter (Quay et al., 1992); transparent; larger pH range between 5 and 7	more abundant in highly aliphatic, rather saturated CHO compounds

### 1.2.2.2 Classification by water bodies

The Amazon River is a net heterotrophic ecosystem in which respiration far exceeds primary production (Hedges et al., 1994), where most CO<sub>2</sub> outgassing originates from allochthonous carbon (Abril et al., 2014; Mayorga et al., 2005; Richey et al., 2002a). Solar radiation exposure is limited due to high CDOM content in black rivers and turbidity from high sediment loads in white rivers, whereas clear rivers show comparatively higher primary production (Gagne-Maynard et al., 2017; Moreira-Turcq et al., 2003a). Productivity in the white Andean streams is very low because of the turbidity associated with the high suspended loads. The black rivers draining the forest also are rather unproductive presumably because of their acidity (pH < 5) and high spectral absorbance. However, clear rivers become more turbid during flooding season, leading to lower rate of primary production (Gagne-Maynard et al., 2017). DOM in Amazon main channel is overall more processed and refractory than in the other water bodies in the

Amazon basin (Aufdenkampe et al., 2007; Ward et al., 2013).

DOM in flooded forest had more catchment input compared to that in Amazon main stem. Water exchange between main channels and flooded forest waters had significant impacts on their respective DOM concentration and composition, especially during the rainy season as considerable input of DOM from terrestrial sources enters rivers in large quantities. Aquatic and terrestrial macrophytes and flooded forests in the floodplain are likely sources of labile organic carbon into the main channel. Moreover, bacterial, and photochemical processing and mineralization as well as chemoselective sorption to minerals are the defining processes determining DOM biogeochemistry in the Amazon catchments (Amon and Benner, 1996b; Armanious et al., 2014; Melo et al., 2019; Moreira-Turcq et al., 2003a; Pérez et al., 2011).

DOM in lakes and wetlands arises partly from autochthonous primary production and also from allochthonous input from catchment, river, and streams. The persistence of DOM by sorption, flocculation, sedimentation, and burial is accompanied by the substantial processing of DOM and emission of greenhouse gases (GHGs) (Abril et al., 2014; Constantine et al., 2014; Mayorga et al., 2005; Pérez et al., 2011; Tranvik et al., 2009). DOM in lakes is more abundant in N-/P- containing compounds due to higher contribution of algal biomass (Ward et al., 2015). Additionally, clear seasonal patterns in DOM composition in lakes in Amazon floodplain systems through the flood pulse were observed, which affected the amounts of autochthonous and allochthonous inputs and consequently the extent of humification, molecular weight, and aromaticity of the DOM (Affonso et al., 2011; Melo et al., 2019). Besides flooding, other factors like precipitation, inflow and outflow of tributaries and streams, water residence time (WRT), light, nutrients, and temperature also lead to shifts of DOM composition in Amazon lakes (Kent et al., 2004; Newton et al., 2006; Shade et al., 2007).

Stream DOM is least processed compared to DOM in the other water bodies. DOM from soils enters the aquatic environment via headwater streams. Thereafter, it is gradually transformed, removed by sedimentation, and mineralized. Production, transformation, and consumption of organic compounds are strongly dependent on the source and character of DOM (e.g., initial DOM composition, N content, and labile nature), as well as the microbial community (D'Andrilli et al., 2019). Although different stream DOM samples differ in compositions due to specific environmental drivers and biogeochemical processes, they were found to share several ubiquitous features. This could reflect that fast processing at tropical temperature converges from e.g. distinct groundwaters and headwaters in small water bodies toward higher order rivers (Vannote et al., 1980).

A comparison of water characteristics and DOM composition in different water bodies in the Amazon Basin (e.g., river, flooded forest waters, lakes, and streams) is shown in Table 3.

Table 3 Water characteristics and DOM composition in different water bodies in the Amazon basin.

Water types	Water characteristics	DOM composition
Rivers	high turbidity; bacterial biomass is higher than phytoplankton biomass; high levels of microbial respiration (Wissmar et al., 1981)	high abundance of refractory DOM compounds like polyphenols and CRAM; DOM presents a more terrestrial character during flooded periods, which contrasts more microbial imprint in low-water-seasons due to increase of WRT (Bouwman et al., 2013; Ellis et al., 2012; Moreira-Turcq et al., 2003a)
Flooded forest waters	lower turbidity than main stem; may be disturbed by anthropogenic factors (e.g., deforestation) (Ríos-Villamizar et al., 2016); may have abundant vegetation (Puhakka et al., 1992)	high variability of DOM in hydrological cycles (Coppola et al., 2019b); may have elevated amounts of black carbon due to biomass burning (Colicchio et al., 2011; Metcalfe et al., 2007)
Lakes	phytoplankton biomass is higher than bacterial biomass; high concentration of nutrients, dissolved oxygen, and chlorophyll <i>a</i> than in main channel of the Amazon River (Wissmar et al., 1981); seasonal transport of biolabile organic to Amazon River main channel (Moreira - Turcq et al., 2013), which would be rapidly consumed (de Melo et al., 2020); receive a considerable input of allochthonous materials from the inundated forest and river inflow during high-water-season	contain distinct DOM pools: a large pool of more refractory DOM from C3 allochthonous plants sources (terrestrial carbon from uplands, riverine inputs, leached from flooded forests and soils), and a small pool of labile DOM compounds derived from C4 macrophytes and phytoplankton, which are consumed and rapidly turned over by heterotrophic bacteria (Mortillaro et al., 2016; Waichman, 1996)
Streams	oligotrophic; shaded by canopy cover; low rates of photosynthesis (Vannote et al., 1980); almost the totality of carbon standing stock is allochthonous and composed by microbial degraded dead leaves (Abelho, 2001; Tank et al., 2010); proximity to terrestrial sources; shorter WRT; minimal	high proportion of soil-derived organic matter; characterized by presence of both microbial-derived unsaturated compounds ( $H/C= 1.0-1.5$ ) and plant-derived (poly)phenolic compounds ( $H/C < 1$ ; $O/C > 0.5$ ) (Hawkes et al., 2018)

extent of transformation and high molecular diversity in headwaters (Hawkes et al., 2018)
---

### 1.2.3 Allochthonous and autochthonous sources

The source of DOM in the Amazon waters is either allochthonous (i.e., produced in the terrestrial environment and carried to the aquatic system) or autochthonous (i.e., produced within the water body) (Cai et al., 1988). Allochthonous organic carbon may originate from natural sources or (minor in Amazon basin) from man-made origin in terrestrial environments, including domestic and industrial wastes. Organic matter from dead leaves, macrophytes, and planktonic organisms is carried by runoff into rivers and lakes, where they are degraded by fungi and bacteria followed by the releasing of POM and DOM into the surrounding waters (Vannote et al., 1980). The lignin and stable-carbon isotopic compositions of these fractions indicate tree leaves and other nonwoody tissues from C3 land plants as predominant sources (Ertel et al., 1986; Hedges et al., 2000). In view of the relatively unpolluted nature of the Amazon drainage basin, the contribution of terrestrial organic carbon of man-made origin is insignificant on average. However, many scientists suggested that damming, deforestation, climate change, and agricultural sources may already influence DOM characteristics and water quality downstream in the Amazon basin (Dalmagro et al., 2019; Howarth et al., 1996; Roebuck et al., 2020).

One of the world's largest run-of-river dams, Santo Antônio, constructed on the Madeira River, had a larger effect on the water chemistry and thermal structure of back-flooded tributaries than on the mainstem river within the reservoir. The evidence of incursion of mainstem water into the back-flooded tributaries was indicated by conductivity and dissolved inorganic carbon. The organic matter input and subsequent mineralization increased in the lateral valleys after damming, as indicated by higher concentrations of total organic carbon (TOC), DOC, biochemical oxygen demand (BOD), and partial pressure of CO<sub>2</sub> (pCO<sub>2</sub>) (Arantes et al., 2023).

In contrast, autochthonous organic carbon is produced within the water body from sources like alive or dead planktonic biomass. Additionally, the production of exudates and secretions by living organisms is an important autochthonous input of DOM. Biota produce labile compounds by metabolic processes that pass through their cell/body walls or are excreted into the surrounding water; those are often rapidly utilized by heterotrophs. Higher aquatic primary productivity occurs principally in the Amazon floodplain lakes (referred to as várzeas in the Portuguese language) (Saatchi et al., 2007; Sawakuchi et al., 2017) and in the plume of the Amazon River (Cai et al., 1988; Smith Jr and Demaster, 1996) compared to the Amazon River main channel. The extensive floodplain areas where “white rivers” deposit their nutrient-rich sediment (várzeas) are very fertile and covered by highly productive terrestrial and aquatic herbaceous communities (Junk et al., 2011). In contrast, the floodplains of “black rivers” (igapós) feature low fertility and scarce

terrestrial/aquatic herbaceous plants.

Compared with the Amazon River main stem and the blackwater/whitewater tributaries, clearwater tributaries such as the Tapajós River have more abundant algal biomass and a higher rate of in situ primary production (Moreira-Turcq et al., 2003a). Several studies have found very low chlorophyll *a* values in the whitewater rivers and high chlorophyll *a* in clear waters (Abril et al., 2014; Gagne-Maynard et al., 2017). In situ primary production of C4 grasses or algal material was identified as the key autochthonous source of DOM in Tapajós River (Quay et al., 1992; Ward et al., 2016). Membrane lipids are known products of microbial degradation of phytoplankton and might serve as potential precursors of DOM (Harvey et al., 2006). Phytoplankton blooms provide new molecules relevant to the DOM pool in Amazon plume waters (Medeiros et al., 2015).

#### **1.2.4 Factors regulating DOM properties**

Longitudinal patterns of riverine DOM, in terms of both concentration and composition, are controlled by numerous environmental drivers including connectivity with surrounding wetlands, lateral inputs from tributaries, and shifts in dominant land cover. Once in the aquatic ecosystem, terrestrial DOM is exposed to in-stream processing such as microbial and photochemical processing, and DOM-mineral interactions that usually operate simultaneously and lead to the removal and the transformation of DOM during its transport (Lundeen et al., 2014; Subdiaga et al., 2019).

Abrupt changes of environmental conditions, either in space (“hot spots”, e.g., at interfaces connecting soils, wetlands, sediments, and rivers) or time (“hot moments”, e.g., during algal blooms or flooding events, or upon injection of the easily degradable substrate) influence the dynamics of carbon pools in natural systems (Kuzyakov and Blagodatskaya, 2015; Marín-Spiotta et al., 2014). Mixing phenomena in aqueous media combine both spatial and temporal aspects in a unique way and occur in several settings in nature, most prominent at river confluences or at the boundary of riverine and marine systems. In the case of continental-scale rivers like the Amazon River, such aggregation occurs noticeably at confluences between tributaries with distinct characteristics. For instance, non-conservative behaviour of DOM and intensified biogeochemical processes at the junction of the Amazon’s major blackwater tributary, the Negro River, and its parent, the Solimões River were proved by many studies, likely due to the large gradients in pH, conductivity, DOC, and particle load (Moreira-Turcq et al., 2003b; Simon et al., 2019).

##### **1.2.4.1 Microbial processing**

Allochthonous, terrestrially-derived DOC has been frequently believed to be largely recalcitrant in freshwaters, merely being passively transported rather than transformed in the aquatic environment.

However, DOM has been shown recently to represent an important source of bioavailable carbon, fueling aquatic heterotrophic ecosystem processes, particularly in streams and rivers where the DOC concentrations are high in headwaters (Billett et al., 2010).

DOM has a key role as the prime source of carbon and energy to sustain heterotrophic microbial communities. Organic carbon is consumed and CO<sub>2</sub> is released during heterotrophic metabolism (D'Andrilli et al., 2019). The predominance of respiration over phytoplankton production in the Amazon River indicated that the river system is net heterotrophic (Benner et al., 1995). This could be due to the metabolic activity of heterotrophic microorganisms stimulated by adequate upstream loading of terrestrially produced organics that support microbial growth, fine sediments with sufficient surface area to promote attachment and growth, and shading of phytoplankton by suspended matter that reduces photosynthetic activity (Wissmar et al., 1981).

According to the size-reactivity model (Amon and Benner, 1996a; Benner and Amon, 2015), biological degradation plays a dominant role in shaping the size spectrum of organic matter. The first step in DOM biodegradation is the breakdown of high molecular weight macromolecules into either monomeric or oligomeric low molecular weight DOM, the first being more labile and rapidly degrading into CO<sub>2</sub> or to be used in methanogenesis, while the latter is more refractory and more readily accumulates (Conrad, 1999). Compared to photochemical degradation, microbial degradation has been described as more selective and generally has a general trend of both consumption and production of low molecular weight compounds (i.e., 250-800 Da) (Reader et al., 2015). Biodegradation has been shown to selectively remove oxygenated unsaturated compounds, producing oxygen-rich small molecules CO<sub>2</sub> and water (H<sub>2</sub>O) and organic compounds with low O/C and high H/C ratios (Kim et al., 2006; Medeiros et al., 2015).

In the Amazon basin, blackwater and whitewater rivers supported similar bacterial abundances and rates of bacterial production and respiration, but with an apparent distinction in the factors limiting microbial growth and respiration: Organic carbon limited bacterial growth and respiration in whitewater rivers, whereas carbon, nitrogen, and phosphorus were colimiting in the Negro River (Benner et al., 1995). Clearwater river Tapajós showed ~40% lower depth-integrated respiration rate compared to the turbid Amazon mainstem. Substrate addition experiments indicated that the relatively low rates of bacterial activity in Amazon River water were caused by carbon limitation (Amon and Benner, 1996b). Moreover, high levels of microbial respiration were observed in the Amazon River and near its mouth, resulting in the depletion of organic carbon substrates that are reactive on timescales faster than fluvial transport and the persistence of less reactive DOC (Ward et al., 2015).

In surface waters, the complex mixture of DOM compounds represents the main substrate and energy source for heterotrophic bacteria (Azam et al., 1983). The diversity and coupled functional attributes of the inherently diverse natural bacterial communities observed in streams, lakes and higher order rivers define

their potential to process different types of organic matter, and in general, such heterotrophic activity is a major sink for the aquatic DOM pool (Guillemette et al., 2013; Tranvik et al., 2009). Multiple interactions between heterotrophic bacteria and DOM shape the apparent composition of both key ecosystem components. The availability and composition of organic substrates favor specific bacterial groups and in this way shape bacterial community composition and community metabolism, and vice versa (Carlson and Hansell, 2015; Guillemette et al., 2016; Kritzberg et al., 2006; Romera-Castillo et al., 2011). Previous studies have shown that adding DOM of different quantities and qualities to natural microbial assemblages yields significant differential responses among the major and minor groups of bacterioplankton (Alonso-Sáez et al., 2012; Landa et al., 2013; Nelson and Siegel, 2013). Moreover, bacteria do not only consume and degrade DOM, but also produce and release an array of autochthonous organic compounds during cell growth, division, and death (Kawasaki and Benner, 2006), thereby influencing the availability, composition, and biogeochemical cycling of carbon in the biosphere (Battin et al., 2008; Osterholz et al., 2016).

Bacterial production and respiration were carbon limited in the Amazon mainstem, indicating that the bulk of relatively abundant DOM and POM were of limited bioavailability (Benner et al., 1995). Although increased DOC concentration generally leads to an increase in the respiration response, the bioavailability of the organic matter has an important influence. The terrestrially sourced organic matter was shown to be highly labile and readily metabolized (Ide et al., 2017). In contrast, microbial-derived DOC seemed to be more refractory, suggesting a prior transformation of the labile material in these samples (Cooper et al., 2016; Ide et al., 2017). Compared to much of the bulk organic carbon mobilized at high water, terrestrial derived lignin-like compounds were used inefficiently under these conditions (for example, a greater proportion was remineralized to CO<sub>2</sub> than fixed into microbial biomass). Although bulk organic carbon concentrations increase significantly with increasing discharge, most of these compounds are probably fixed as microbial biomass rather than respired to CO<sub>2</sub> (reflected by low bulk respiration rates and high bacterial abundance), whereas the degradation of lignins and tannins may provide a significant source of the respired CO<sub>2</sub> during high-water conditions (Ward et al., 2013).

#### **1.2.4.2 Photochemical processing**

Although microbial degradation of DOM in aquatic environments is highly important for the understanding of organic carbon cycling, most studies investigate transformations occurring along a natural gradient, in which photochemical reactions play a critical role as well (Ciuraru et al., 2015; Gonsior et al., 2014; Mostofa et al., 2007; Niu and Croué, 2019; Wilske et al., 2020; Xie et al., 2004). Photochemical degradation is one of the most significant abiotic reactions of NOM, and often produces oxygen-rich small molecules like CO, CO<sub>2</sub>, and H<sub>2</sub>O alongside oxygenated and aliphatic DOM molecules. However, a

previous study indicated that photochemical processes account for a relatively small fraction of carbon mineralization in the Amazon River (Amon and Benner, 1996b). In situ ultraviolet-oxidation experiments performed in the Amazon suggest that photo-oxidation accounts for roughly 0.05% of CO<sub>2</sub> outgassing rates, while depth-integrated respiration has been suggested to be the primary source of CO<sub>2</sub> in the Amazon River (Ward et al., 2013).

UV radiation will simultaneously stimulate and inhibit microbial activity in natural environments. Photooxidation of DOM will break down natural DOM and release oxygenated aliphatic molecules, which in turn could be respired by bacteria (Amador et al., 1989). In contrast, UV radiation may inhibit microbial activity due to direct UV damage to the organism (Paul et al., 2012) and by toxic intermediates produced during photooxidation of the natural DOM (Lund, 1994).

Photobleaching of DOM, a recognized process in aquatic environments, could be enhanced by the forthcoming higher levels of UV-radiation. That is concerning because DOM acts as a screening filter for aqueous microorganisms against excessive and potentially harmful UV-radiation (Brinkmann et al., 2003). Thus, changing DOM molecular properties in aquatic environments could threaten its protection to several microbial communities.

Photoproduction of small oxygenated molecules including photomineralization to CO<sub>2</sub> that are or resemble biochemical metabolites will enhance microbial activity (Gonsior et al., 2014; Hertkorn et al., 2006; Schmitt-Kopplin et al., 1998). Coupled photochemical and microbial DOM processing is context-dependent and structure-selective, and is subject to synergy and competition (Huang et al., 2023). DOM can be degraded by sunlight into a variety of photoproducts that stimulate the growth and activity of microorganisms in aquatic environments. A number of laboratory studies using bacterial bioassay approaches have shown that the photochemical breakdown of DOM can stimulate biomass production or activity by 1.5- to 6-fold (Moran and Zepp, 1997). The biologically available photoproducts from DOM in surface waters were proved to have important biological roles. For instance, in a continental shelf system, full exposure of surface seawater to sunlight for one summer day can produce DOM photodegradation products equivalent to >20% of the bacterial carbon demand (Moran and Zepp, 1997). Likewise, 30% of the bacterial nitrogen demand can be met by photodegradation of the nitrogen components of DOM (Moran and Zepp, 1997).

Previous studies have shown that UV exposure alters the molecular composition of DOM (Gonsior et al., 2009; Kujawinski et al., 2009). The chemical composition of DOM dictates whether photochemical degradation will produce bio-available or bio-resistant compounds (Carlson and Hansell, 2015). The production of reactive oxygen species during photodegradation, for example, has been proposed as a mechanism that forms partially oxidized DOM, further reducing its bioavailability (Baltar et al., 2013).



Biologically refractory DOC has unusual high photoreactivity (Amon and Benner, 1996b). CDOM is more sensitive to light-mediated degradation, due to its capability of absorbing more radiation (Altare and Vione, 2023; Moran and Zepp, 1997). Via absorption of sunlight, CDOM greatly influences the underwater light field and subsequently affects photoecological processes, including primary production and microbial activities (Moran and Zepp, 2000). “Clear waters” offer larger light transparency than the other types of Amazon waters, and showed lower concentration of CDOM, the most readily photodegraded component of DOM molecules. In contrast, CDOM is the most abundant DOM fraction in “Blackwaters”. In a previous study, sunlight exposure over 27 h showed that at least 15% of Rio Negro DOM was photoreactive (Amon and Benner, 1996b). The rate of photochemical consumption of DOC was ~7 fold greater than bacterial DOC utilization in Rio Negro surface waters; however, integrated over the entire water column, microbial remineralization was the dominant process for oxygen and DOC consumption (Amon and Benner, 1996b). The suspended sediments and cations in “White waters” contribute to shielding from sunlight (Ríos-Villamizar et al., 2020), and transform DOM through mineral particle adsorption and complexation (Subdiaga et al., 2019), thereby affecting bioavailability as well (Schmitt-Kopplin et al., 1998).

Experiments relating the effect of light-induced DOM degradation and metal toxicity in aquatic environments have been performed. It was expected that DOC concentrations would decrease with UVB irradiation, and that metal toxicity would be enhanced following the decrease of DOC concentration. However, significant effects have only been demonstrated in brown waters, while changes were less evident in the clear-water samples (Winch et al., 2002).

The photo-transformation of DOM has caused the loss of lignin, the decrease in O/C ratio, and a concomitant increase in saturation during photodegradation experiments (Gonsior et al., 2019; Seidel et al., 2015), which can be explained by photochemical decarboxylation reactions (Amon and Benner, 1996b).

It has been proposed that combined photochemical and biological degradation processes preferentially remove oxidized, aromatic compounds, whereas reduced, aliphatic and N-containing compounds persist in aquatic systems, either because they are resistant to degradation or tightly cycled (Kellerman et al., 2015). Therefore, DOM might be degraded along a gradient from aromatic to aliphatic compounds, and from a high to a low nominal oxidation state of carbon.

Photochemical reactions in oxic environments contribute to oxygenation of DOM molecules, mainly by introduction of hydroxyl and carboxy functional groups (Gonsior et al., 2014; Hertkorn et al., 2006; Schmitt-Kopplin et al., 1998); heteroatoms N and S in DOM molecules introduce alternative selectivity in photoprocessing of DOM molecules (Harir et al., 2022). In addition, polyols were found during photo-transformation of DOM in oligotrophic surface ocean waters, which accounted for 2% of total molecular signatures in <sup>1</sup>H NMR spectra (Gonsior et al., 2014). Under oxygen atmosphere, structures of lignic and lipidic origin were most photolabile in comparison with carbohydrates, alkylbenzenes, or N-containing

structures that accumulated in the system. Under nitrogen atmosphere the acids remained fairly stable (Schmitt-Kopplin et al., 1998). Under UV/Vis irradiation, indirect photolysis reactions were suggested to be the major pathway in DOM degradation (Schmitt-Kopplin et al., 1998; Stubbins et al., 2010). Aromatic compounds were found to be most photoactive, with 90% being lost upon irradiation. The refractory condensed aromatics were significantly attenuated whereas CRAM actually shifted towards more refractory regions instead of getting removed (Stubbins et al., 2010). Besides, many photoproduct oxygenated small molecules will readily integrate into biochemical pathways and foodwebs (Gonsior et al., 2014; Huang et al., 2023), and transient particulate organic carbon, making traceability to specific mechanisms difficult.

#### **1.2.4.3 DOM-mineral interaction**

Strong DOM-mineral interactions could alter the reactivity of organic molecules towards increased resistance to degradation. The physical protection of organic matter at interior mineral surfaces provides alternative pathways that foster the recalcitrance of DOM. Sorption results directly in the retention of DOM and the subsequent decrease of its mobility and further degradation. However, complexation can result in the formation of both soluble and insoluble DOM-metal complexes (enhanced and retarded movement, respectively), thereby affecting movement and degradation in different ways (Bolan et al., 2011).

There are generally four types of interactions between organic matters and minerals: ligand exchange or sorption of small molecules to mineral surfaces; sorption of large molecules such as proteins to several mineral surface contact points; aggregation (i.e., chemical bonding or van der Waals attractions); and occlusion, which offers the most substantial protection of organic matter (Kaiser and Guggenberger, 2003).

In general, more hydrophobic and/or N-containing molecules are preferentially adsorbed on sediments and may have longer WRT in the Amazon basin (Ertel et al., 1986). Thus, nitrogen-rich DOM produced by microbial degradation may be strongly adsorbed by minerals and exported by the water body in particulate form (Hedges et al., 1986b).

The continual transfer of waters, nutrients, and sediments occurs through hyporheic zone, which has major control on the residence time of organic matter compounds in freshwaters (Tank et al., 2010). Sediments can accumulate nutrients over time and can also be an autochthonous DOC source (Chen and Hur, 2015; Wang et al., 2014). Laboratory and field studies of organic matter adsorption by soil minerals indicate that uptake can be extensive and highly selective for hydrophobic and nitrogen-containing constituents (Hedges et al., 1986b). Dissolved nitrogen-rich degradation products may be so strongly adsorbed by fine-grained minerals that they ultimately are exported largely in particulate form by soil erosion, as opposed to redissolution.

#### 1.2.4.4 River mixing

DOM can undergo a variety of biogeochemical reactions in mixing regions (e.g., river plumes, estuaries) that ultimately determine both the concentration and composition of DOM reaching the ocean (Raymond and Spencer, 2015).

Solimões (white water) and Negro (black water) combine in Manaus to form the Amazon River (turbid water). Ertel et al. assigned 70% of degraded humic and fulvic acids fluxes at the Solimões and Negro Rivers (S+N) confluence to Negro inputs (Ertel et al., 1986; Johannsson et al., 2017). Madeira (dissolved organic matter-DOC ~ 5.8 mg/L, particulate organic matter (POC) ~ 0.83 mg/L) joins the Amazon River downstream at ~140 km east of Manaus (do Nascimento et al., 2015). Tapajós River drains into the Amazon River at the town of Santarem (~620 km from Manaus). The confluences of major Amazon Basin Rivers can be regarded as effluents of entire aquatic ecosystems, and mixing causes abrupt changes in environmental conditions including temperature, density, flow characteristics, pH, concentrations of ions and (mineral and organic) particles, DOM composition and molecular structure (Simon et al., 2019). Steep gradients develop for all these features in mixing zones, with distinct spatial and temporal evolution of interdependent variables. Opportunities arise for many non-conservative effects upon riverine mixing, making those places hot spots for organic matter processing and microbial metabolism, influencing the dynamics of carbon pools in natural systems (Farjalla, 2014; McClain et al., 2003). Mixing of different water types initiates complex biogeochemical processing, including alteration of aquatic microbial communities and food webs (Lynch et al., 2019), DOM itself (Bianchi and Ward, 2019), and of river bulk characteristics.

The Tapajós River hosts higher proportions of cyanobacteria (~12%) compared to the Amazon River (~5%), where actinobacteria and betaproteobacteria dominate the microbial community (Chen and Hur, 2015). “Positive priming effect” was suggested to happen when algal-rich Tapajós waters mix with turbid Amazon main stem waters, accelerating the decomposition of more recalcitrant organic matter (Seidel et al., 2015).

Ecological models and biogeochemical characterization assume river mixing zones as hotspots of microbial activity. Most studies report distinct non-conservative behavior of DOM after riverine mixing affected by factors like phase partitioning (Moreira-Turcq et al., 2003b; Seidel et al., 2016), respiration (Abril et al., 2014; Mayorga et al., 2005), trace metal complexation (Simon et al., 2019), and discharge (Ward et al., 2018). However, the effects of the physical mixing of different Amazon basin riverine waters on the composition and structure of its DOM at molecular resolution remain ill-constrained.

#### 1.2.4.5 Seasonal variation

Large-scale events such as flooding can have a major impact on the cycling of carbon and nutrients in these aquatic ecosystems: 1) As the sediment in the river water settles out in lakes after the incursion, transparency increases and phytoplankton blooms occur; 2) As the river level drops, lake water, and associated plankton and vegetative production products drain into the river mainstream providing possible organic energy supply for the white-water Amazon. Lowland drainage basins and floodplains are active sources of carbon for the main channel of the river, especially during falling water (Quay et al., 1992). The aquatic and terrestrial macrophytes and flooded forests in the floodplain are likely sources of labile organic carbon input to the main channel (Quay et al., 1992).

The extent of DOM decomposition depends on the WRT of the aquatic ecosystem. In large rivers, WRT varies spatially and seasonally, increasing in reservoirs and lakes compared to river channels, being higher during low flow compared to high flow (Lambert et al., 2016).

Large amounts of saturated fatty acids were concomitantly recorded with high CO<sub>2</sub> in rivers during the high water season (Mortillaro et al., 2012). Contrastingly, fatty acid markers that may be attributed in this ecosystem to aquatic plants and cyanobacteria were correlated with higher O<sub>2</sub>, chlorophyll a, and phaeopigments in floodplains, due to a high primary production during low water season (Mortillaro et al., 2012). Decreasing concentrations of unsaturated fatty acids, that characterize labile organic matter, were recorded during high water season, from upstream to downstream, which was attributed to organic matter retention and the extent of flooded forest in floodplains.

O<sub>2</sub> is the preferential electron acceptor and is used for organic matter oxidation until its depletion. In aquatic environments, the presence/absence of O<sub>2</sub> in the water column depends on seasonal stratification (Glissman et al., 2004; Huttunen et al., 2001). Whenever thermal stratification occurs in the water column, mixing between the upper and the bottom layers is impeded and, therefore, O<sub>2</sub> does not reach to deeper levels. Thus, the concentration of O<sub>2</sub> close to the sediments is highly influenced by water column stratification (Liikanen et al., 2002).

Bacterial growth efficiencies were consistently higher during high-water stages of the hydrograph when river-varzea interactions were maximal, suggesting that varzea inputs of substrates of greater bioavailability than the bulk of riverine materials and the changes in the quality of organic materials in the river can lead to marked responses by the microbiota. Bacterial growth and respiration appear to be supported by moderately labile (> 1 d) oligomeric substrates that are continuously supplied by lateral exchange with the surrounding varzea (Benner et al., 1995).

#### 1.2.4.6 GHGs emission

Just a few years ago, it was the common view that rivers passively transport soil-derived NOM to the sea, but now it is largely acknowledged that most of the terrestrial organic carbon entering freshwater systems is either respired locally or buried in sediments, and only a fraction of DOM is discharged into the ocean. And it has become clear that inland waters process large amounts of organic carbon and must be considered in strategies to mitigate climate change (Du et al., 2016; Xenopoulos et al., 2021).

Inland waters and wetlands globally are estimated to receive inputs of organic and inorganic carbon from the terrestrial environment that are on the order of  $5.7 \text{ Pg y}^{-1}$ , of which only about  $0.9 \text{ Pg y}^{-1}$  is transported to oceans. Thus, rivers and wetlands act as temporary storage of carbon in sediments, and cycle carbon back to the atmosphere in the form of  $\text{CO}_2$  and  $\text{CH}_4$  (Dalmagro et al., 2017).  $\text{CO}_2$  and  $\text{CH}_4$  are the dominant GHGs emitted from tropical rivers. Emission rates are governed by numerous factors including water depth, temperature, dissolved gas concentration, and vegetative cover. More recent evidence suggests that the magnitude of wet-dry cycles may have a significant effect on GHG emissions (Bass et al., 2014).

Under aerobic conditions,  $\text{CO}_2$  is produced by the breakdown of organic material via heterotrophic processes and concurrently consumed via autotrophic pathways, with the balance between the two being a primary dictator of fluxes. Aerobic conditions in the sediment lead to the consumption of methane via methanotrophic bacteria.  $\text{CO}_2$  and  $\text{CH}_4$  are also produced under anaerobic conditions via anaerobic respiration and methanogenesis, respectively. In freshwater environments, anaerobic mineralization of organic carbon has been estimated to account between 20 - 60% of total carbon mineralization rates (Boon and Mitchell, 1995; Hamilton et al., 1995; Utsumi et al., 1998; Ward et al., 2017), from which up to 80% of the anaerobic activity is carried on by methanogenic bacteria. These values exceed the methane emissions from the ocean (Ward et al., 2017).

NOM occupies intermediate carbon oxidation states between fully oxidized  $\text{CO}_2$  and fully reduced  $\text{CH}_4$ . In addition, NOM is a substrate for microorganisms producing GHGs. In this process, NOM composition and structure will change accordingly.

The balance between gross primary production and respiration rate is a key measure of the net metabolic state of ecosystems (Gagne-Maynard et al., 2017). Amazon River outgasses nearly an equivalent amount of  $\text{CO}_2$  as the rainforest sequesters on an annual basis due to microbial decomposition of terrigenous and aquatic organic matter (Richey et al., 2002b). Amazon mainstem waters are depleted in dissolved  $\text{O}_2$  (50-75% of equilibrium saturation values) and supersaturated by up to 24-fold in dissolved  $\text{CO}_2$ . Amazon River is a net heterotrophic ecosystem in which respiration far exceeds primary production (Hedges et al., 1994), and most  $\text{CO}_2$  outgassing originates from allochthonous carbon (Abril et al., 2014; Mayorga et al., 2005; Richey et al., 2002b). Respiration of contemporary organic matter (less than five years old) originating on

land and near rivers is the dominant source of excess carbon dioxide that drives outgassing in medium to large rivers (Ward et al., 2013). Lignin and other terrestrially derived macromolecules contributed significantly to carbon dioxide outgassing from inland waters (Ward et al., 2013). A small, rapidly cycling pool of organic carbon of labile organic matter may be responsible for the large carbon fluxes from land to water to atmosphere in the humid tropics (Devol et al., 1995; Mayorga et al., 2005).

Additionally, anaerobic processes and root respiration in floodplain waters can produce CO<sub>2</sub> without commensurate O<sub>2</sub> depletion. The excess CO<sub>2</sub> in the Amazon River is perhaps introduced by lateral exchange with the floodplain where anaerobic processes and root respiration can produce CO<sub>2</sub> without commensurate O<sub>2</sub> depletion (Devol et al., 1995).

About 20% of Amazon wetlands are permanently or seasonally inundated, and these aquatic habitats are a major CH<sub>4</sub> source (Melack et al., 2004). During high-water seasons, river levels swell significantly and seasonally flood large areas of otherwise dry forest. Biomass begins to rot in oxygen-depleted water in these wetlands and produces methane. Wetlands have a stock of organic matter that provides fuel for respiration and allow this process to progress to the point where methanogenesis is the only remaining option, thus increasing CH<sub>4</sub> production (Dalmagro et al., 2018).

It has been shown that CO<sub>2</sub> production from the sediments increases with temperature for both oxic and anoxic conditions and that higher temperatures also enhance the transport of nutrients from the sediment to the water column, accelerating the eutrophication process and promoting CH<sub>4</sub> production in anoxic sediments (Liikanen et al., 2002).

Theoretically, higher CO<sub>2</sub> concentration in the atmosphere positively feeds back the photosynthetic activity of autotrophic organisms, increasing their consumption of CO<sub>2</sub> (Killops and Killops, 2013). However, higher photosynthetic rates, ultimately, promote DOM production (Carlson and Hansell, 2015). Increased availability of DOC could also lead to increased proportions of carbon mineralization and further GHGs production (Conrad et al., 2020).

### **1.3 Structural characterization of Amazon DOM**

Despite the importance of aquatic DOM and much work on DOM characterization, a comprehensive understanding of the general chemical structures of DOM and the variations of these structures under different environmental conditions remains elusive. Polydispersity, heterogeneity, and temporal dynamics of DOM have made its molecular characterization a long-standing challenge (Hertkorn et al., 2008).

The chemically diverse DOM samples require multi-level complementary analytical approaches for their comprehensive characterization; on the other hand, complementary analysis of DOM samples generates

large data sets, and their joint assessment enables in-depth elucidation of DOM characteristics (Hertkorn et al., 2007b). Common DOM characterization approaches are listed in Table 4.

Measurements of isotopic values (Sanderman et al., 2009; Zigah et al., 2012) on bulk DOM samples provide information on a large portion of the DOM pool, but do not give detailed structural information. In general, two main types of analytical methods have been applied for the molecular level characterization of NOM: optical and organic structural spectroscopy (Dauwe and Middelburg, 1998). Optical spectroscopy such as UV-vis absorption (Helms et al., 2008; Macdonald and Minor, 2013) and fluorescence (Miller and McKnight, 2010; Murphy et al., 2008) relies on the presence and chemical environments of  $\pi$ - and n-electrons and provides swift, sensitive, and low-cost determination of chromophoric component of NOM and coarse distinction of NOM origin and evolution at bulk-resolution (Stedmon and Bro, 2008). Three-dimensional excitation-emission matrix (EEM) fluorescence spectroscopy has been widely used to provide distinctive “fingerprints” for NOM of various origins at however ill-defined relationships between spectra and NOM molecular structures (Gonsior et al., 2013b; Rodriguez et al., 2014; Stubbins et al., 2014; Zhao et al., 2017). On the other hand, structural analytical methods such as Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) that allow for the assignment of thousands of molecular formulas out of complex mixtures (Hertkorn et al., 2008; Herzsprung et al., 2016; Herzsprung et al., 2017a; Herzsprung et al., 2017b; Kamjunke et al., 2017; Koch et al., 2007; Valle et al., 2017), and nuclear magnetic resonance (NMR) spectroscopy that allows for structural and quantitative analysis of key substructures in extremely complex organic mixtures like NOM, exhibit better defined relationships between original data and NOM structural features (Hertkorn et al., 2004; Hertkorn et al., 2002b; Perdue et al., 2007). Moreover, joint assessment of FT-ICR MS and NMR data has enabled remarkable distinction of e.g., freshwater, marine, atmospheric and extraterrestrial NOM (Hertkorn et al., 2013; Hertkorn et al., 2015; Schmitt-Kopplin et al., 2010), and concepts to interconnect optical properties with NOM molecular features have recently emerged (Kellerman et al., 2015; Zhang et al., 2014).

Table 4 The characterization approaches of DOM.

Analytical technique	Function	Advantages	Drawbacks
Elemental analysis	Elemental composition and elemental ratios	Universality and the ease of operation	No in-depth coverage of structural features
Isotope analysis	Isotope abundance	Tracing the heritage of DOM and apparent ages ( $^{14}\text{C}$ )	No in-depth resolution of molecular diversity
UV-vis spectroscopy	Functional groups, chemical bonds	Facile sample preparation, without special handling and ease of operation	Highly selective towards $\text{sp}^2$ carbon chemical environments; highly influenced by solvent, pH, temperature, and high electrolyte concentration
FT-IR spectroscopy	Functional groups	High sensitivity; requires little amounts of sample	some selectivity towards $\text{sp}^2$ carbon chemical environments; but appreciable coverage of aliphatic units; low structural resolution within functional groups
Three-dimensional EEM fluorescence spectroscopy	Fluorescence characteristic of chromophores in DOM; source and structure of DOM	High sensitivity and non-destructive method	Highly selective towards $\text{sp}^2$ carbon chemical environment; need to consider environmental factors including temperature, pH, metal ions, and oxidants to diminish the quenching effect
FT-ICR mass spectrometry	Molecular formulae	exceptional resolution: a wide range of ionization techniques (ESI, CI, PI, DESI, FI, etc.) for mixtures is available under specifically adapted conditions	Selective in ionization; isomer differentiation is a nontrivial task; quantification is difficult, even for identical molecules in mixtures



Nuclear magnetic resonance (NMR)	provides invaluable information about the detailed molecular and chemical structure of the measured compounds; provides short-range atomic order of NMR-active isotopes (e.g., $^1\text{H}$ , $^{13}\text{C}$ , $^{15}\text{N}$ , $^{31}\text{P}$ )	most powerful tool for structural analysis, without any prior selection and non-destructive, with little sample preparation; quantitative, if the experiment is carried out appropriately; 2D experiments are necessary to overcome the problems of signal overlap and annotation of the signals to chemical structures	relative insensitivity compared with other analytical techniques; intricate physics and chemistry of intra- and intermolecular interactions in complex mixtures may interfere with the direct relationship between chemical shift and molecular structure, because of relaxation-induced variable line widths, may affect quantification; near-identical chemical shifts do not necessarily imply similar chemical structures
----------------------------------	---	---	---

Previous studies on the origin and fate of organic matter in the Amazon River included stable isotopic (Mortillaro et al., 2011; Quay et al., 1992), and radioisotope analysis (Hedges et al., 1986b) to assess sources and age of organic carbon. Quantification of targeted compounds involved humic substances (Ertel et al., 1986), saccharides (Saliot et al., 2001), and amino acids (Hedges et al., 1994) among others. Elemental analysis, optical spectroscopy (Martinez et al., 2015), mass spectrometry (Gonsior et al., 2016), and  $^{13}\text{C}$  NMR spectroscopy (Hedges et al., 1992) were applied to understand the composition of Amazon organic matter.

Due to the individual capabilities and limitations of every characterization method, different studies detect distinct components of DOM depending on the needs and methods employed, making the comparison between studies often difficult. A comprehensive molecular understanding of the reactivity and bioavailability of Amazon DOM is lacking at present, especially in tributaries with different water types and in major confluence zones.

### 1.3.1 Biomarker analysis

Direct isolation and analysis of individual compounds and compound classes from filtered aqueous samples can provide useful information on DOM composition in a biomarker sense, but does not allow characterization of a quantitatively important portion of the DOM pool. The three most common biochemical classes analyzed in this manner are hydrolysable carbohydrates, hydrolysable amino acids, and lignin derived phenols. Hydrolysable carbohydrates have been used to provide comparative source and

diagenetic information in sample sets (Goldberg et al., 2011; Kaiser and Benner, 2012). Hydrolysable amino acids are often used to provide insights into the potential bioavailability of DOM and the extent of degradation the DOM may have undergone (Duan and Bianchi, 2007; Yamashita and Tanoue, 2004), using degradation indices based upon multivariate analysis of amino acid distributions within a large suite of natural organic matter samples (Dauwe and Middelburg, 1998; Yamashita and Tanoue, 2003). Lignin-derived phenols are used to indicate terrestrial inputs into DOM samples and the extent of degradation these terrestrial inputs have undergone (Meyers-Schulte and Hedges, 1986; Osburn and Stedmon, 2011; Woods et al., 2011).

Hedges et al. observed the greatest compositional differences in the Amazon River system were among the coarse (>63  $\mu\text{m}$ ), fine (<63  $\mu\text{m}$ ), and ultrafiltered dissolved organic fractions (>1,000 Daltons) (Hedges et al., 1994). All coarse particulate fractions exhibited aldose and amino acid compositions like those of angiosperm tree leaves. Coarse particulate organic materials, although the least degraded of the three fractions, had lost appreciable carbohydrates and had immobilized excess nitrogen of apparent bacterial origin. Fine particulate materials had lower total aldose yields and glucose percentages than coarse counterparts. Fine particles gave greater total yields of amino acids, characterized by high ratios of basic vs. acidic components. Coexisting dissolved organic materials recovered by ultrafiltration yielded the lowest amounts of aldoses, among which deoxy sugars were concentrated. Dissolved fractions gave extremely low yields of amino acids in mixtures that were enriched in nonprotein components and in acidic vs. basic molecules. These yield and composition patterns are consistent with a “regional chromatography” model in which highly degraded leaf material is solubilized and then partitioned between soil minerals and water during transport to the river, resulting in suspended fine particulate organic materials of soil origin that are nitrogen-rich and coexisting dissolved organic substances that are nitrogen-poor.

Ertel found that all the aquatic humic and fulvic acids from nine mainstem and seven major tributary sites in the Amazon River Basin had clearly recognizable lignin components at levels (8 and 3% of the carbon) suggesting a predominantly allochthonous source (Ertel et al., 1986). Compositional characteristics of processed lignins were dominated by diagenetic transformations, which included preferential loss of methoxylated structural units and oxidative degradation of lignin side chains. Fulvic acids showed consistently lower lignin levels, lower lignin phenol methoxylation, higher acid / aldehyde ratios, and higher C/N ratios than coexisting humic acids, all indicative of greater aerobic degradation of the fulvic acid fraction.

Hedges et al. reported lignin compositions of typical plant tissues in the Amazon basin and found that POM has distinct lignin compositions in different grain-size fractions (Hedges et al., 1986a). Particulate organic matter transported in the coarse size fraction of the mainstem, and its major tributaries is composed of recently formed and well-preserved tree leaf debris along with some wood. Organic matter in the fine size

fraction was comparatively old, degraded, and rich in immobilized nitrogen and derives primarily from soils.

Feng et al. characterized lignin phenols in soils, river POM, and DOM across a 4 km elevation gradient in the Madre de Dios River system, Peru, as well as in marine sediments to investigate the source-to-sink evolution of lignin (Feng et al., 2016). They found more oxidized lignin in organic horizons relative to mineral horizons in soils. The oxidized lignin signature was maintained during transfer into rivers, and lignin was a relatively constant fraction of bulk organic carbon in soils and riverine POM. Lignin in DOM became increasingly oxidized downstream, indicating active transformation of dissolved lignin during transport, especially in the dry season. In contrast, POM accumulated undegraded lignin downstream during the wet season, suggesting that terrestrial input exceeded in-river degradation.

### **1.3.2 Elemental analysis**

The determination of the elemental compositions and relevant elemental ratios provides the most fundamental bulk information for DOM characterization (Kördel et al., 1997).

Based on studies of the Amazon and other major rivers, more than half of organic matter in transport occurs either as highly oxidized, nitrogen-poor (atomic C/N = 15-85) dissolved humic substances or nitrogen-rich (C/N = 10-13) organic matter associated with fine mineral grains (Ertel et al., 1986; Hedges et al., 1986a; Meybeck, 1982). An additional major dissolved component is too hydrophilic to be isolated onto resins together with dissolved humic substances and is as yet poorly characterized. Fresh plant debris appears to be a relatively small component of the total organic matter mixture (Hedges et al., 1986a).

Higher plants generally have greatly considerably higher C/N ratios compared to microalgae. The C/N ratio of the more labile fraction of DOM decreased with depth, driven by a switch in the dominant source of DOM from phytoplankton to heterotrophic biomass (Zakem and Levine, 2019). For the Amazon River, seasonal shifts in particulate C/N ratios indicated a greater abundance of algal and microbial-derived POC during low and falling water compared to a larger abundance of higher plant tissue during rising and high water. Fine POM from the Amazon mainstem was characterized by uniformly high organic nitrogen levels (atomic C/N = 11) and low concentrations of highly oxidized lignin, all of which are typical of soil organic matter (Ward et al., 2015).

Moreover, C/N values are also affected by remineralization and sorption processes along the river continuum. For example, nitrogenous compounds such as amino acids generally sorb to the particulate phase to a greater extent than carbon-rich compounds (Ward et al., 2015). Dissolved fulvic acids from the Amazon River system are less hydrophobic and more nitrogen-poor (atomic C/N, 40 to 85) than the

corresponding humic acids (C/N, 15 to 60) (Ertel et al., 1986). Laboratory and field studies of organic matter adsorption by soil minerals indicate that uptake can be extensive and highly selective for hydrophobic and nitrogen-containing constituents (Ertel et al., 1986). Thus, fulvic acids should be adsorbed less strongly and pass more rapidly through the soil than humic acids of the same source. This model suggests that the hydrophilic fraction of the total dissolved organic material (which cannot be isolated on resin) may have an even shorter residence time in the Amazon basin than the coexisting fulvic acid. Dissolved nitrogen-rich degradation products may be so strongly adsorbed by fine grained minerals that they ultimately are exported largely in particulate form by soil erosion, as opposed to redissolution.

Ríos-Villamizar et al. reported that the total N and total P allowed to distinguish the three classical water types (white, black and clear) and other water bodies with intermediate position (Ríos-Villamizar et al., 2013). Additionally, the distribution of alkali and alkaline-earth metals and major anions is especially useful for the distinction of whitewater, blackwater and clearwater categories. Higher variability is shown by water bodies that not fit inside these three classic categories. Therefore, many rivers and streams have to be considered as “mixed waters” resulting from the influence of lower order tributaries with different physico-chemical properties of their waters (Ríos-Villamizar et al., 2013).

Gledhill et al. determined elemental stoichiometries of aluminum, iron, copper, nickel, zinc, cobalt, and manganese associated with a fraction of the DOM pool isolated by solid-phase extraction at ambient pH from the Amazon plume (Gledhill et al., 2022). The biogeochemistry of the trace metals bound to the SPE extracted DOM component in the Amazon plume was found to be determined by the chemical nature of the trace metals and not by that of the DOM (Gledhill et al., 2022).

### **1.3.3 Isotope analysis and radiocarbon analysis**

Isotope analysis identifies the isotope signatures in the chemical compounds and can be used to trace the sources and understand the dynamic cycling processes in DOM studies. The two major isotopes used in DOM studies are  $^{13}\text{C}$  and  $^{15}\text{N}$ , whose ratios may be altered by biological and geophysical processes (Guo and Sun, 2009). Carbon has three natural isotopes  $^{12}\text{C}$ ,  $^{13}\text{C}$  and  $^{14}\text{C}$ , with the abundance of 98.89 %, 1.11 % and  $10^{-10}$  %, respectively.  $\delta^{13}\text{C}$  isotope is widely used in DOM studies for trace its origin based on the assumption that DOM from different origin has a distinct  $\delta^{13}\text{C}$  isotope value. A general trend is observed for DOM with different sources: the value ranges from -35‰ to -25‰ for terrestrial derived DOM; it is in the range of -18‰ and -8‰ for marsh macrophytes; and intermediate for the marine phytoplankton (Guo and Sun, 2009). The variation in the values is mainly attributed to the isotope fractionation during carbon metabolism in the plants in which  $\delta^{13}\text{C}$  commonly becomes progressively enriched but also depends on the isotope composition in different source materials. Nitrogen has two natural isotopes  $^{14}\text{N}$  and  $^{15}\text{N}$ , with the abundance of 99.63 % and 0.37 %, respectively.  $\delta^{15}\text{N}$  isotope is used for tracing and distinguishing nitrogen

sources and monitoring certain nitrogen-dependent biogeochemical processes such as nitrification and microbial processes. Generally, the  $\delta^{15}\text{N}$  values range from -10‰ to 20‰ in different DOM samples (Guo and Sun, 2009). The radioactive  $^{14}\text{C}$  with a half-life of  $5730\pm 40$  years is used to determine the apparent age in organic materials by radiocarbon dating method. It serves as a proxy for DOM ages with the time scale of ~50,000 years based on the assumption that negligible exchange of organic carbon is involved in the studied ecosystem.

Previous studies, using isotopic and elemental analyses, revealed that there are apparently two distinct DOM pools in the Amazon rivers system: a large pool of more refractory DOM from C3 allochthonous plants sources (terrestrial carbon from uplands, riverine inputs, leached from flooded forests and soils), and a small pool of labile DOM compounds derived from C4 macrophytes and phytoplankton, which are consumed and rapidly turned over by heterotrophic bacteria (Ertel et al., 1986; Hedges et al., 1986b; Keil et al., 1997; Mayorga et al., 2005; Mortillaro et al., 2016; Waichman, 1996).

$\delta^{13}\text{C}$ -DOC values are generally slightly lower than  $\delta^{13}\text{C}$ -POC values, with limited natural variation (Aufdenkampe et al., 2007; Ellis et al., 2012; Hedges et al., 1994; Mayorga et al., 2005). Variance in  $\delta^{13}\text{C}$ -DOC is smaller than that of  $\delta^{13}\text{C}$ -POC, possibly consistent with a limited contribution of in situ aquatic primary production and the advanced degradation status of the river-transported DOM pool (Hedges et al., 1994; Quay et al., 1992), which is also generally more recent in origin (Marwick et al., 2015; Mayorga et al., 2005). Variability in  $\delta^{13}\text{C}$ -DOC/POC in tributaries and along downstream gradients of the Amazon River has typically been attributed to differences in plant inputs from forests, and floodplains (Ward et al., 2015). The stable isotopic composition of POC and DOC indicated a replacement of highland forest-derived organic carbon with lowland and floodplain-derived organic carbon from Óbidos to the mouth (Ward et al., 2015). Likewise, lignin phenol signatures showed an increase in the degradation state of vascular plant-derived organic carbon from Óbidos to the mouth (Ward et al., 2015).

During low water season, when the connectivity with the river mainstem was restricted,  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$  signatures of suspended POM from open floodplain lakes (Várzea) of the central Brazilian Amazon basin showed a major contribution of autochthonous material from phytoplankton and C3 aquatic plants, and the C4 aquatic phanerogam contribution appeared to be weak, although these plants were the most abundant macrophyte in the Várzea. Moreover, the fatty acid compositions of suspended POM from the Amazon River exhibited significant seasonal differences, suggesting a seasonal contribution of autochthonous material produced in Várzea to the Amazon River (Mortillaro et al., 2011).

A small range of  $\delta^{13}\text{C}$ -DOC variation was observed throughout the hydrological cycle upstream of the confluence of the Negro River (-29.2 to -29.5‰) and the Solimões River (-28.7 to -29.0‰), whereas a much larger one (-28.0 to -34.6‰) was found in the lower reach of the river, as the proportion of open lakes increased downstream in the floodplains. The low variability in the upper reaches suggests

homogeneous DOC sources from upland soils and flooded forest, while lower  $\delta^{13}\text{C}$ -DOC values recorded in the lower reach mainstem at high and falling waters can be attributed to a greater export of plankton-derived  $^{13}\text{C}$ -depleted DOC from flooded lakes (Albéric et al., 2018). Higher  $\delta^{13}\text{C}$ -DOC values measured in the Rio Madeira and the associated flooded lakes ( $-26.5$  to  $-28.8\%$ ) may reflect the imprint from upland headwaters and a weaker density of flooded forest in the watershed (Albéric et al., 2018).

Mayorga et al. analyzed  $\Delta^{14}\text{C}$  in the  $\text{CO}_2$  from the Amazon and showed that most of it originated from rapid recycling of young organic carbon (Mayorga et al., 2005). The organic matter respired in the Amazon River has been shown to be relatively young, with a radiocarbon age of less than five years (Ward et al., 2013). However, the entire pool of lignin and related phenolic compounds in the Amazon River could completely overturn in only 2-3 weeks (Ward et al., 2013).

Dissolved humic and fulvic acids and particulate organic materials transported by the Amazon River all contain bomb-produced  $\Delta^{14}\text{C}$  (“bomb” carbon derived from atmospheric testing of thermonuclear devices after 1954 (Nydal and Lövseth, 1983), indicating the relatively rapid turnover of the parent carbon pools. However, the  $\Delta^{14}\text{C}$  contents of the coexisting carbon forms are measurably different and may reflect varying degrees of retention by soils in the drainage basin.

The proportion of black carbon like polycyclic aromatic structures decreases downstream in the Amazon River but show marked spatial variability in abundance.  $\Delta^{14}\text{C}$  values of dissolved black carbon molecular markers in the Amazon River implied dynamic sources and cycling in a manner that is incongruent with bulk DOC. Moreover, the dissolved black carbon flux from the Amazon River was estimated to be composed of predominately young dark black carbon (Coppola et al., 2019a).

### **1.3.4 Optical spectroscopy**

The application of UV-visible spectroscopy provides the rough estimates of DOM concentration, sources, and certain functional groups. It boasts the advantages of low cost, ease of use and satisfactory reliability at high sensitivity, rendering it one of the most widely used techniques in DOM characterization. But it is highly selective to the chromophoric DOM instead of the overall DOM pool, and can be easily affected by ionic strength and pH changes (Minor et al., 2014). Fluorescence spectroscopy measures the fluorescence properties of DOM samples. Similar to UV-visible spectroscopy, it is efficient as well as powerful to collect DOM information like rough quantification of the concentration, characterization of certain functional groups and monitoring the reactivity, but only works on chromophoric DOM (Minor et al., 2014). Excitation-emission matrix (EEM) fluorescence spectroscopy, as one of the most widely applied technique in the structure assessment of DOM, acquires a range of emission spectra from the excitation spectra. EEM spectra provide rich information in terms of fluorescence indices and fluorescence peaks (Yu et al., 2020).

Optical (absorbance and fluorescence) analysis revealed a clear seasonal pattern in DOM composition through the flood pulse in the Amazon River, which affected the amounts of autochthonous and allochthonous inputs and consequently the extent of humification, molecular weight, and aromaticity of the DOM (Melo et al., 2019). Bacterial community composition was tightly coupled to DOM fluorescence, which was also driven by differences across the hydrological cycle. DOM concentration tended to be relatively stable throughout the year, whereas DOM composition showed clear changes (Melo et al., 2019).

The Amazon River is a major source of terrestrially derived DOM to the Atlantic Ocean. Cao et al. reported that the distribution and dynamics of optically active fraction of DOM were largely controlled by physical mixing along the Amazon River-ocean continuum, with the exception of the protein-like component whose distribution was driven by factors other than dilution (Cao et al., 2016). The on-board water incubation experiments with photochemical and microbial alterations revealed that photochemistry was primarily responsible for the remineralization of the optically active fraction of DOM whereas the contribution of microbial transformation to this fraction was minor (Cao et al., 2016).

Patel-Sorrentino et al. measured the concentration of TOC and the fluorescence intensity before and after UV-visible irradiation on the waters from sampling sites along the Negro River for a given time. The results demonstrated that when photodegradation happened, the dominant reactions involved O<sub>2</sub> and reactive oxygen species, and the fluorescent sites were among the first to be altered. Fluorescence and complexing site densities increased after a short irradiation time and decreased after long irradiation times, indicating that a change of the macromolecular structure activates previously inactive fluorophores and complexing sites prior to photodegradation (Patel-Sorrentino et al., 2004).

### **1.3.5 Mass spectrometry**

Mass spectrometry (MS) generates ions from inorganic/organic compounds with appropriate methods and separates those ions according to the mass-to-charge ratio ( $m/z$ ) and detects them qualitatively and quantitatively by their  $m/z$  and abundance. The application of MS in DOM studies provides the information on molecular weights and chemical composition that can be further interpreted into different groups of compounds (Kujawinski et al., 2002; Mopper et al., 2007; Nebbioso and Piccolo, 2013). Structural analytical methods such as FT-ICR MS allow for the assignment of thousands of molecular formulae out of complex mixtures and are direct experimental evidence of the immense complexity of DOM (Hertkorn et al., 2008; Hertkorn et al., 2016; Hertkorn et al., 2013; Herzsprung et al., 2016; Herzsprung et al., 2017a; Herzsprung et al., 2017b; Kamjunke et al., 2017; Koch et al., 2007; Valle et al., 2017).

Medeiros et al. investigated the molecular composition and transformations of DOM in the Amazon River-ocean continuum using ultrahigh resolution mass spectrometry and geochemical and biological tracers

(Medeiros et al., 2015). They found that the composition of terrigenous Amazon DOM is surprisingly stable in the adjacent coastal shelf even under different discharge regimes. Alterations in DOM composition are observed in the plume associated with the addition of new organic compounds by phytoplankton and with bacterial and photochemical transformations. Their study further showed that changes in the plume's DOM composition were due to photochemistry and bacterial degradation as well as the addition of DOM compounds by phytoplankton (Medeiros et al., 2015).

Seidel et al. molecularly characterized Amazon River DOM via ultrahigh-resolution FT-ICR MS, determined DOM stable carbon isotopes, bacterial abundance and production, phytoplankton biomass and composition<sup>252</sup>. The incubation experiments with turbid water from close to the river mouth suggested that photo- and bio-alteration leave significant molecular and carbon isotopic imprints on the terrigenous DOM. Nonetheless, quantitative removal by bio- and photo-degradation appeared to proceed at a relatively slow pace because no significant dissolved organic carbon decrease was found within the five days of the incubation (Seidel et al., 2015). However, 9-30% of DOC was lost after the five-day incubation with less turbid water from the intermediate and outer plume, suggesting that the introduction of reactive algal DOM in the intermediate plume may thus have primed the microbial degradation of terrigenous DOM (Bianchi, 2011; Seidel et al., 2015). Additionally, sorption of terrigenous DOM to sinking particles acts as an important DOC sink in the Amazon plume (Seidel et al., 2015).

FT-ICR MS combined with EEM fluorescence spectroscopy and parallel factor analysis (PARAFAC) of Amazon River waters revealed a large proportion of ubiquitous DOM molecular compositions but also unique area-specific molecular signatures (Gonsior et al., 2016). Unique to the DOM of the Rio Negro area was the large abundance of high molecular weight, diverse hydrogen-deficient and highly oxidized molecular ions deviating from known lignin or tannin compositions, indicating substantial oxidative processing of these ultimately plant-derived polyphenols indicative of these black waters. In contrast, unique signatures in the Madeira/Jamari area were defined by presumably labile sulfur- and nitrogen-containing molecules in this white-water river system. Waters from the Tapajós main stem did not show any substantial unique molecular signatures relative to those present in the Rio Madeira and Rio Negro, which implied a lower organic molecular complexity in this clear water tributary, even after mixing with the main stem of the Amazon River (Gonsior et al., 2016).

### **1.3.6 NMR spectroscopy**

Nuclear magnetic resonance (NMR) spectra provide isotope-specific, stand-alone *de novo* perception of extended atomic environments in organic molecules or mixtures of those; the most suitable isotopes in DOM research are <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N and <sup>31</sup>P (Deron et al., 2002; Gigliotti et al., 2002; Hatcher et al., 1980; Hertkorn et al., 2016). The information concerning close range atomic order in molecules is complementary



to other organic structural spectroscopy that represent molecules (mass spectrometry) (Hertkorn et al., 2006; Hertkorn et al., 2016; Hertkorn et al., 2013), functional groups (infrared spectroscopy) (Chen et al., 2019; Gigliotti et al., 2002; Wu et al., 2022) and small atomic units (XANES and other X-ray based methods) (Fan et al., 2014; Karlsson et al., 2005; Xia et al., 2022). NMR has been widely used for DOM structural characterization with both qualitative identification and quantitative measurement of the functional groups. NMR spectra are in particular suited to detect and analyze aliphatic and alkyl systems, which leave commonly inconspicuous signatures in other analytical methods. Different types of NMR spectra are available, and among them solution-state and solid-state NMR are extensively utilized. Generally, solid-state NMR is performed on the solid samples like sediments and soils, while solution-state NMR is applied to liquid samples such as the extracted DOM, some NMR spectroscopy is even possible for DOM samples at natural conditions without pretreatment<sup>18</sup>. Solution-state NMR offers the advantage over solid-state NMR to allow multidimensional NMR spectroscopy with higher resolution for in-depth structural elucidation of relationships between groups of atomic nuclei (Hertkorn et al., 2015; Hertkorn et al., 2002a; Hertkorn et al., 2007b; Koch et al., 2008; Mopper et al., 2007; Simpson et al., 2011).

800 MHz <sup>1</sup>H NMR spectroscopy provided the quantification of major atomic environments in DOM composition with excellent S/N ratio and resolution, enabling quantification of key DOM structural units including CRAM, oxygenated aliphatics including carbohydrates, olefins, and aromatics (Hertkorn et al., 2016; Hertkorn et al., 2013; Powers et al., 2019; Simpson et al., 2011). Moreover, 1D NMR showed better coverage of purely aliphatic structures not resolvable in mass spectra. NMR offers valuable constraints of hydrogen and carbon atomic environments not available by FT mass spectra and is important for providing structural information and obtaining credible conclusions.

Supportive of <sup>1</sup>H NMR, <sup>13</sup>C NMR provides a wider range of chemical shifts with reduced spectral overlaps and allows for the observation of ketone and carboxylate functional groups and other quaternary carbon atoms that are not directly accessible by <sup>1</sup>H NMR spectroscopy (Simpson et al., 2011). Furthermore, DEPT (distortionless enhancement by polarization transfer) experiment by <sup>13</sup>C NMR spectroscopy enables to differentiate different carbon multiplicities (CH, CH<sub>2</sub>, and CH<sub>3</sub>). However, a certain disadvantage of <sup>13</sup>C NMR spectra at low sample availability is the low sensitivity due to only 1.1% <sup>13</sup>C carbon nuclide at natural abundance and low gyromagnetic ratio of <sup>13</sup>C (~ 1/4 of <sup>1</sup>H).

A 2D NMR spectrum is obtained with two Fourier transformations on a matrix of data, and cross peaks that correlate information on one axis with data on the other are generated in a 2D spectrum. 2D NMR experiments detect networks of spin-spin couplings and spatial proximity, necessary to overcome the problems of signal overlap and enabling annotation of the NMR resonances to chemical structures (Simpson et al., 2011).

NMR spectroscopy of terrigenous humic substances isolated from the Amazon River system indicated high

aromaticity and lignin-derived methoxy carbon, as well as aliphatic and carboxyl carbon (Hedges et al., 1992). In contrast, NMR spectroscopic analyses of marine humic substances isolated from the Pacific Ocean indicated an abundance of carboxylated aliphatic molecules with a highly branched structure and low unsaturated carbon (Hedges et al., 1992). Based on these observations, riverine dissolved humic substances appear to comprise only a small fraction of seawater humic substances and therefore must be efficiently and rapidly removed from the ocean.

## 1.4 Scope of this work

Polydispersity, heterogeneity, and temporal dynamics of DOM have made its molecular characterization a long-standing challenge. FT-ICR mass spectrometry of DOM allows for an unprecedented resolution of thousands of molecular formulae direct out of the DOM mixture, however with considerable selectivity depending on sample handling and ionization method applied (Flerus et al., 2012; Hertkorn et al., 2013; Koch et al., 2005; Kujawinski et al., 2002; Kujawinski et al., 2009). High-field NMR spectroscopy on the other hand allows for quantification of major atomic environments in DOM molecules with excellent S/N ratio and resolution, enabling quantification of key DOM structural units including aliphatic and carboxyl-rich alicyclic molecules (CRAM), oxygenated aliphatics including carbohydrates, olefins, and aromatics (Hertkorn et al., 2007b; Simpson et al., 2002; Simpson et al., 2011). A main objective of the thesis is to provide a comprehensive coverage of molecular evolution of Amazon DOM processing using complementary ESI[±] mass spectrometry and NMR spectroscopy.

Repeated attempts have been made to describe the origin and fate of particulate and dissolved organic matter throughout the Amazon River and its tributaries using a variety of methodologies (Mortillaro et al., 2011; Quay et al., 1992). As highlighted in the prior chapters, due to the inherent complexity of DOM, comprehensive molecular understanding of the reactivity, bioavailability, and processing of DOM in the Amazon River is lacking at present, especially in tributaries with different water types and in major confluence zones. Thus, this thesis aims to gain better understanding of Amazon basin DOM processing. The study here is to reveal effects of major processes on the compositional space and structural characteristics of black, white, and clear water systems. A particular emphasis is laid on revealing major compositional and structural changes of DOM at the confluences of the Solimões and Negro Rivers (S+N), and the Amazon and Tapajós Rivers (A+T).

Many studies reported distinct non-conservative behavior of DOM after riverine mixing affected by factors like phase partitioning (Amon and Benner, 1996a; Sleighter and Hatcher, 2007), respiration (Benner and Amon, 2015; Philippe and Schaumann, 2014b), trace metal complexation (Schmitt-Kopplin et al., 2019b), and discharge (Borggaard et al., 2019). Ecological models and biogeochemical characterization assume river mixing zones as hotspots of microbial activity. However, DOM degradation pathways are still poorly addressed in river networks with molecular precision, preventing an accurate estimate of carbon sources and sinks. Moreover, the effects of physical mixing of different Amazon basin riverine waters on the composition and structure of its DOM at molecular resolution remain ill-constrained. Another major goal of this work is to understand how DOM molecular features change following controlled mixing and incubation of the endmember river waters with different biogeochemical characteristics. This thesis also

aims to investigate the biotic and abiotic factors that lead to the DOM changes after water mixing and incubation.

## 2. Material and methods

### 2.1 Sampling and site locations

The Amazon River main stem originates in Andean mountains springs and for most of its flow path meanders across the lowland of Amazon tropical rainforest until its mouth in Atlantic Ocean, while collecting major tributaries on its way. The main stem of the Amazon River is named Ucayali while flowing in Peru and when it crosses Brazil's boundary it changes the name to Solimões. Downstream of the confluence with the Negro River in Manaus (Amazonas state, Brazil) it changes the name again to Amazon River and keeps this name until its mouth on Atlantic Ocean in Brazil's coast. There are many large tributaries with different water characteristics. Three fundamental types of water were already described by Alfred Russel Wallace in 1853 in his trip to the Amazon basin, and these were called as white, clear and blackwaters (McClain and Naiman, 2008). The color varies depending on whether dissolved organic matter (black water) or suspended sediment (white water) is predominant (Moreira-Turcq et al., 2003b). The clear water rivers are transparent and supply a high phytoplankton production (Richey et al., 1990). Both white and clear waters have more aquatic plants and floating meadows than black waters. In contrast, the black water Negro River is nearly devoid of such vegetation. Solimões contributed the most carbon to the Amazon River, about  $500 \text{ kg C s}^{-1}$  during the sampling high water period (Moreira-Turcq et al., 2003b). Madeira is the largest and Negro is the second largest tributary in Amazon watershed. These two rivers together contribute more than 30% of discharge of the Amazon River main stem. Madeira is a primary tributary of Amazon in terms of sediment input.

We collected water samples at 31 sites in the Amazon River main stem eastwards, including Solimões (white water), Negro (black water), Amazon (turbid water), and Tapajós (clear water) between April 2<sup>nd</sup> and May 25<sup>th</sup> in 2014, during a high-water period with extremely high discharge volume. The localities accessed were located near Manacapuru and Manaquery in Solimões, Castanheira and Santa Maria in Negro, Manaus, Urucará and Óbidos in Amazon River and Alter do Chão and Aramanai in the Tapajós River. Detail sampling information see Figure 3, Table 5. We obtained water samples by boat just below the surface, typically 10 L per sample. Solid phase extraction (SPE) of the water samples was done at the field immediately after sampling. The water column DOM was extracted by a previously described SPE method using 1g cartridges of PPL resin (Dittmar et al., 2008). The eluates were stored in the freezer (-20°C) until mass spectrometry analysis and nuclear magnetic resonance spectroscopy analysis.

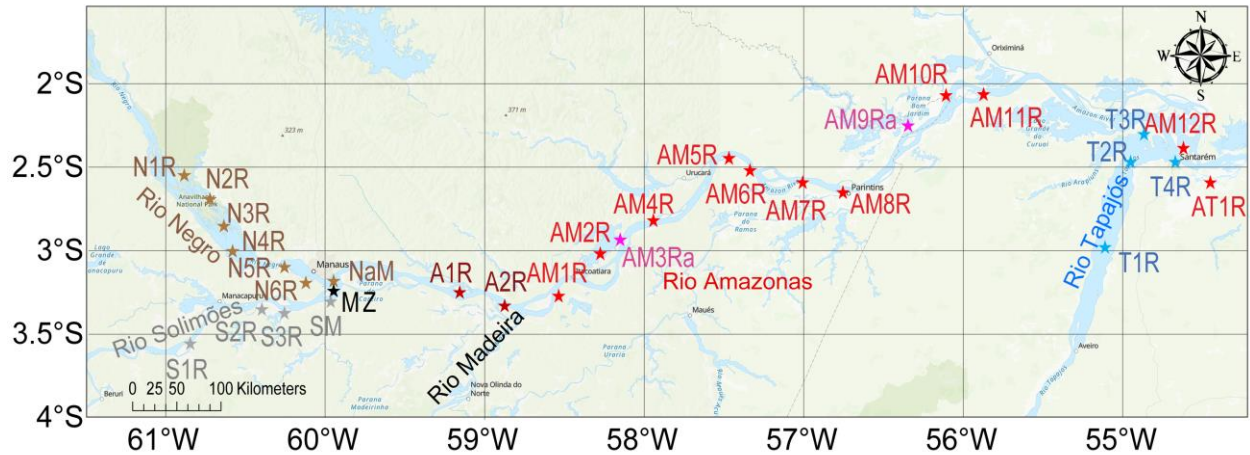


Figure 3 Maps of the sampling sites in the Amazon River. “SM” (grey) was sampled in the Solimões River upstream and close to the Negro-Solimões River mixing zone. “MZ” (black) was sampled at the very mixing zone. “A1R/A2R” (dark red) were sampled in the Amazon River upstream of the Madeira River inflow, while the other Amazon River samples (red) were sampled after the Madeira River inflow. “AM3Ra” and “AM9Ra” (magenta) were sampled two months later than the other water samples in the Amazon River.

Table 5 Description of the sampling sites and sampling information. “SM” (grey) was sampled in the Solimões River close to the mixing zone of the Negro River and the Solimões River. “MZ” (black) was sampled at the mixing zone of the Negro River and the Solimões River. “A1R/A2R” (dark red) were sampled in the Amazon River upstream of the Madeira River inflow, while the other Amazon River samples (red) were sampled after the Madeira River inflow. “AM3Ra” and “AM9Ra” (magenta) were sampled two months later than the other water samples in the Amazon River. The order of the rows in the table from top to bottom represents the order of sampling point locations in each river from upstream to downstream.

River	Sample	Longitude (°W)	Latitude (°S)	Water type
Solimões (S)	S1R	-60.6845	-3.3414	white
Solimões	S2R	-60.1116	-3.3209	white
Solimões	S3R	-60.0957	-3.3136	white
Solimões close to SN mixing zone	SM	-59.8949	-3.1647	turbid
Negro (N)	N1R	-60.8198	-2.5740	black
Negro	N2R	-60.6865	-2.7267	black

<b>Negro</b>	N3R	-60.5509	-2.8719	black
<b>Negro</b>	N4R	-60.5583	-2.9440	black
<b>Negro</b>	N5R	-60.1810	-3.0958	black
<b>Negro</b>	N6R	-60.0902	-3.1242	black
<b>Negro</b>	NaM	-59.9126	-3.1292	black
<b>SN mixing zone</b>	MZ	-59.8958	-3.1301	turbid
<b>Amazon River (A) before Madeira River inflow</b>	A1R	-60.1596	-2.7933	turbid
<b>Amazon River before Madeira River inflow</b>	A2R	-58.7621	-3.2992	turbid
<b>Amazon River after Madeira River inflow</b>	AM1R	-58.5192	-3.1802	turbid
<b>Amazon River</b>	AM2R	-58.5513	-3.1492	turbid
<b>Amazon River</b>	AM3Ra	-58.1626	-2.9750	turbid
<b>Amazon River (out of the river main stem because of too much water)</b>	AM4R	-58.2533	-2.9530	turbid
<b>Amazon River (sampled downstream of the Uatumã River inflow)</b>	AM5R	-57.5167	-2.3797	turbid
<b>Amazon River</b>	AM6R	-57.4771	-2.4972	turbid
<b>Amazon River</b>	AM7R	-57.2811	-2.4676	turbid
<b>Amazon River</b>	AM8R	-57.2811	-2.4676	turbid
<b>Amazon River</b>	AM9Ra	-56.3486	-2.2473	turbid
<b>Amazon River</b>	AM10R	-55.7717	-2.0033	turbid
<b>Amazon River</b>	AM11R	-55.4316	-2.0180	turbid
<b>Amazon River</b>	AM12R	-54.7185	-2.4041	turbid
<b>Amazon River</b>	AT1R	-54.4518	-2.4916	turbid
<b>Tapajós</b>	T1R	-55.1013	-2.9714	clear
<b>Tapajós</b>	T2R	-54.9001	-2.4315	clear
<b>Tapajós</b>	T3R	-54.8332	-2.2742	clear
<b>Tapajós</b>	T4R	-54.7219	-2.4056	clear

## 2.2 Sample processing of mixing experiment

The four endmember samples *S* (S3R), *N* (NaM), *A* (AM12), and *T* (T4R) were used for laboratory mixing and incubation experiments (Table 6). Mixing and incubation experiments with unfiltered endmembers were conducted on the bank or ship directly after sampling by mixing variable ratios of different waters (80:20, 60:40, 50:50, 40:60, 20:80, total volume 500 mL) in Pyrex bottle followed by one day (1d) and five days (5d) dark incubation at *in situ* temperature (30°C) and SPE in the field. We also made a mixture of identical volumes of *S* and *N*, and let it stand in dark for 30 min, named S50N50\_30min.

Moreover, the endmember samples *S* and *N* were separated into “filtered water” and “filtered suspended solids” based on the treatment of the endmember samples before mixing. The waters and suspended solids were isolated by filtration of identical volumes of water from each river, using 0.2 µM Whatman GF/F glass fiber (precombusted at 450°C). The mixture of filtered Negro water and filtered Solimões suspended solids was named FNSS; the mixture of filtered Solimões water and filtered Negro suspended solids was named FSSN; the mixture of filtered Solimões water and filtered Negro water was named FSFN. FNSS and FSSN were left to sit in the dark for one day at *in situ* temperature (30°C) and/or on ice (0°C).

Table 6 Description of samples in mixing and incubation experiments.

Sample	Water mixture	Incubation time
FSFN	50% Solimões + 50% Negro	0 minutes
S50N50_30min	50% Solimões + 50% Negro	30 minutes
S20N80_1d	20% Solimões + 80% Negro	one day
S40N60_1d	40% Solimões + 60% Negro	one day
S50N50_1d	50% Solimões + 50% Negro	one day
S60N40_1d	60% Solimões + 40% Negro	one day
S80N20_1d	80% Solimões + 20% Negro	one day
FSSN_B-1d	Solimões filtered water + Negro filtered suspended solids	one day
FSSN_S-1d	Solimões filtered water + Negro filtered suspended solids	one day
FNSS_B-1d	Negro filtered water + Solimões filtered suspended solids	one day
FNSS_S-1d	Negro filtered water + Solimões filtered suspended solids	one day
S20N80_5d	20% Solimões + 80% Negro	five days
S40N60_5d	40% Solimões + 60% Negro	five days
S50N50_5d_1	50% Solimões + 50% Negro	five days



S50N50_5d_2	50% Solimões + 50% Negro	five days
S50N50_5d_3	50% Solimões + 50% Negro	five days
S60N40_5d	60% Solimões + 40% Negro	five days
S80N20_5d	80% Solimões + 20% Negro	five days
A20T80_1d	20% Amazon + 80% Tapajós	one day
A40T60_1d	40% Amazon + 60% Tapajós	one day
A50T50_1d	50% Amazon + 50% Tapajós	one day
A60T40_1d	60% Amazon + 40% Tapajós	one day
A80T20_1d	80% Amazon + 20% Tapajós	one day
A20T80_5d	20% Amazon + 80% Tapajós	five days
A40T60_5d	40% Amazon + 60% Tapajós	five days
A50T50_5d	50% Amazon + 50% Tapajós	five days
A60T40_5d	60% Amazon + 40% Tapajós	five days
A80T20_5d	80% Amazon + 20% Tapajós	five days

### 2.3 FT-ICR mass spectrometry

Negative and positive electrospray ionization (ESI[±]) Fourier transform ion cyclotron resonance mass spectra (FT-ICR MS) were acquired using a 12T Bruker Solarix mass spectrometer (Bruker Daltonics, Bremen, Germany) and an Apollo II electrospray ionization (ESI) source (Hertkorn et al., 2016). Nebulizer gas pressure, drying gas pressure and the source heater temperature were 138 kPa, 103 kPa and 200 °C, respectively. The spectra were acquired with a time domain of 4 MW. For each sample, 500 broadband scans were accumulated in ESI[−] FT-ICR mass spectra, while 300 broadband scans were accumulated in ESI[+] FT-ICR mass spectra. All spectra were first externally calibrated on clusters of arginine in MeOH (0.57 mmol/L) and internally calibrated using appropriate reference mass lists of common natural organic matter molecules, reaching accuracy values lower than 500 ppb. Data processing was done using Compass Data Analysis 4.1 (Bruker, Bremen, Germany) and formula assignment by an in-house made software (NetCalc) (Tziotis et al., 2011). The molecular formula assignments were based on the following elements:  $^1\text{H}_{0-200}$ ,  $^{12}\text{C}_{0-100}$ ,  $^{16}\text{O}_{0-80}$ ,  $^{32}\text{S}_{0-3}$ ,  $^{14}\text{N}_{0-3}$  as well as the  $^{13}\text{C}_{0-1}$  and  $^{34}\text{S}_{0-1}$  isotopomers. The generated formulae were validated by setting sensible chemical constraints [N rule, O/C ratio < 1, H/C ratio <  $2n+2$  ( $\text{C}_n\text{H}_{2n+2}$ )]. Restriction on nitrogen atoms  $\leq 4$  and sulfur atoms  $\leq 2$  were applied based on previous studies (Hertkorn et al., 2013; Schmitt-Kopplin et al., 2010). Final elemental formulae were generated and categorized into groups containing CHO, CHNO, CHOS, and CHNOS molecular series, which were used to reconstruct the

group-selective mass spectra. The intensity of mass peaks was normalized (total amplitude of assigned mass peaks = 100% for single sample). The average H/C, O/C, N/C, S/C atomic ratios, DBE/C (Double bond equivalent per carbon), mass-to-charge ratios ( $m/z$ ) were computed from the intensity-weighted average of molecular formulae (Hertkorn et al., 2016).

## 2.4 $^1\text{H}$ NMR spectroscopy

All  $^1\text{H}$  NMR spectra were acquired with a Bruker Advance III NMR spectroscopy operating at 800.35 MHz ( $B_0 = 18.8$  Tesla) at 283 K from redissolved solids in  $\text{CD}_3\text{OD}$  (99.95%  $^2\text{H}$ ; Merck) with Bruker standard pulse sequences using 3.0 mm Bruker MATCH tubes. The reference  $^1\text{H}$  NMR chemical shift of  $\text{HD}_2\text{COD}$  was 3.3 ppm.  $^1\text{H}$  NMR spectra were recorded under solvent suppression with presaturation and 1 ms spin-lock (noesypr1d), 5 s acquisition time, 5 s relaxation delay (d1), typically 1024 scans, and 1 Hz exponential line broadening.  $^1\text{H}$  NMR section integrals were obtained by using the software AMIX at 0.01 ppm resolution, with exclusions of HDO and  $\text{HD}_2\text{COD}$  NMR resonances.

Difference  $^1\text{H}$  NMR spectra used full data point resolution ( $1.5 \times 10^{-4}$  ppm) and were computed from area normalized NMR spectra ( $\delta_{\text{H}}$ : 0.5-10.0 ppm), with exclusion of methanol ( $\delta_{\text{H}}$ : 3.2-3.4 ppm) and water section chemical shifts ( $\delta_{\text{H}}$ : 4.8-5.5 ppm) by Bruker AMIX software; data tables were imported into EXCEL tables, and difference NMR spectra were computed in EXCEL and exported to Adobe Illustrator. Vertical axes denote fraction of NMR resonance amplitude with respect to total  $^1\text{H}$  NMR integral (100%).

## 2.5 Water characterization

Water conductivity (Cond.), temperature (Temp.) and dissolved oxygen (DO) were measured in situ with portable instruments (Hanna Instruments, Metrohm electrode and PRO-ODO YSI). For the inorganic dissolved nutrients, the water was sampled with polypropylene bottles, filtered in silica filters (GF/F, Whatman) and placed in amber bottles conditioned in freezer to be analyzed in the laboratory. Ammonium ( $\text{NH}_4^+$ ) concentrations were determined by colorimetric analysis with flux injection (FIA, FiaStar 5000). Nitrate ( $\text{NO}_3^-$ ) and nitrite ( $\text{NO}_2^-$ ) were measured using nitrate reduction and phosphate ( $\text{PO}_4^{3-}$ ) using molybdenum analytical method (Valderrama, 1981). The dissolved organic carbon (DOC) was determined by persulfate digestion analyzed using high-temperature catalytic oxidation method (Sievers InnovOx analyzer, GE) (Gonsior et al., 2019). The total organic carbon (TOC) used the same analytical method as DOC without the water filtration.

The dissolved inorganic carbon (DIC) of water was determined by the acidified headspace method (Åberg and Wallin, 2014). During the field work, 25 mL glass vials were closed with a rubber stopper and prefilled with 0.3 µL of concentrated H<sub>3</sub>PO<sub>4</sub> to acidify the sample to pH ~2. After that, 12.5 mL of water samples without bubbles were injected into the vial with a 0.6 mm needle. Three vials were closed in the field to sample the atmospheric air used in the calculations. Back in the laboratory, the headspace samples were measured by gas chromatography (7890A, Agilent Technologies, USA). Because the water was inoculated in a closed acidified ecosystem, all the DIC was transferred to CO<sub>2</sub>. The CO<sub>2</sub> headspace was measured and the CO<sub>2</sub> that remained in the water sample was equilibrated with the headspace (Henri's Law). DIC was the sum of both CO<sub>2</sub> concentration subtracted by the CO<sub>2</sub> in the atmosphere (Åberg and Wallin, 2014).

## 2.6 Dark carbon fixation (DCF) and heterotrophic bacterial production (HBP)

DCF was estimated by the incorporation of H<sup>14</sup>CO<sub>3</sub><sup>-</sup> in a known time similar to photosynthesis, but in dark conditions (Nielsen, 1965; Santoro et al., 2013). We took care to start the incubations at night, 4 hours after the sunset, to wait the photosynthesis be negligible avoiding dark carbon incorporation by photo-system II (Nielsen, 1960). The incubations were performed by the addition of 10 µL of NaH<sup>14</sup>CO<sub>3</sub> (Perkin Elmer, specific activity of 52.5 mCi mmol<sup>-1</sup>) in 50 mL of river water in five dark (amber) bottles. The incubations were stopped after four hours with Formalin at final concentration of 3.7%. Two controls were prepared containing Formalin (3.7 % at final concentration) before adding the <sup>14</sup>CO<sub>2</sub> to take out the artifacts' values. In lab, the samples were filtered twice inside a fume hood with Sartorius filtration kit. The first filtration was with silica filters GF/F (Whatman) and the second with cellulose acetate filters (0.2 µm pore, Sartorius Stedim Biotech). Both filters were washed with 1M HCl and dried to ensure the elimination of the non-assimilated <sup>14</sup>CO<sub>2</sub>. Filters were placed in scintillation vials with scintillation liquid (Optiphase HiSafe, PerkinElmer) in dark conditions to be analyzed in radiocarbon assay in scintillation counter 24h later to avoid chemiluminescence. Finally, DCF was computed as:

$$\text{DCF } (\mu\text{gC L}^{-1}\text{h}^{-1}) = (\text{Net DPM}_{\text{sample}} - \text{Net DPM}_{\text{control}}) \times \text{DPM}_{\text{added}}^{-1} \times 1.05 \times \text{DIC} \times t^{-1}$$

where DCF was the biomass production or carbon assimilation, DPM was the radioactivity measured by decay per minute of the sample, control and the added into incubations, 1.05 was the isotope discrimination factor, DIC was the dissolved inorganic carbon and t was time (Peterson, 1980).

HBP was determined by the measurement of protein synthesis rates using radio-labelled <sup>3</sup>H-leucine (D. C. Smith, 1992). This method measures the leucine incorporated by the bacteria in a known time after its addition. The methodology had some adjustments for the Amazon waters conditions according to previous

studies (Benner et al., 1995; Farjalla, 2014). Lower leucine concentration and less time was important to not overestimate the bacterial production. 150  $\mu\text{L}$  of  $^3\text{H}$ -leucine with a final radioactivity of 10 Ci  $\text{mmol}^{-1}$  and a leucine final concentration of 20 nM was added to 1.5 mL of river water into five Eppendorf tubes. The temperature was constant and similar to river water during the incubations that were performed in dark conditions. Formalin was added at final concentration of 3.7% to stop the incubation after 45 minutes. In addition, two controls were prepared with the same concentration of formalin before the leucine addition. At the end, the samples were placed in styrofoam containers to be analyzed in the laboratory.

In the lab, the bacterial protein was extracted using cold trichloroacetic acid (TCA) and ethanol. This method consists in the protein precipitation discarding the lighter non-incorporated radio-labelled leucine in the supernatant. Finally, the pellet was resuspended in the scintillation liquid (Optiphase HiSafe, PerkinElmer) with vortex and placed in dark for 24h after being analyzed in the radioassay with the scintillation counter. To compute the bacterial carbon production, we used the equation:

$$\text{HBP } (\mu\text{gC L}^{-1}\text{h}^{-1}) = \text{LI } (\mu\text{mols Leu L}^{-1}\text{h}^{-1}) \times 131.2 \times 0.073^{-1} \times 0.86^{-1} \times \text{ID}$$

where HBP was bacterial production, LI was the leucine incorporated and ID was the isotope dilution. 131.2 was the molecular weight of leucine, 0.073 was the proportion of leucine in total protein and 0.086 was the ratio of cellular carbon to protein (Simon and Azam, 1989).

## 2.7 Statistical analysis

PCA was performed using Simca-P (version 11.5, UmetricsAB, Umeå, Sweden) to identify the dominant modes of variability in DOM composition of AZ-DOM. FT-ICR mass spectra were arranged with the samples as observations and the normalized peak amplitudes of the assigned FT-ICR mass peaks as the response variables.  $^1\text{H}$  NMR spectra were arranged with the samples as observations and the NMR resonances as the response variables (800 MHz  $^1\text{H}$  NMR,  $\text{CD}_3\text{OD}$ , area-normalized from 0.5-10.0 ppm; 0.01 ppm bucket resolution; with the exclusion of residual water and methanol NMR resonances). Before multivariate statistics were performed, the response variables were centered and scaled to unit variance. The based weight was computed as  $1/\text{sqrt}$  (standard deviation of the response variables).

HCA was performed using the Hierarchical Clustering Explorer 3.0 (HCE; <http://www.cs.umd.edu/hcil/multi-cluster/>). Average Linkage (UPGMA) method was used to cluster the dataset and Euclidean distance was used as the similarity/distance measure. Based on the HCA, we used the “profile search” tool from HCE 3.0, choosing a search method (model-based), a distance measure

(Spearman's  $r$ ) and a threshold (0.9).

Spearman correlation analysis ( $p < 0.05$ ,  $r^2 > 0.5$ ) was performed using the R statistical platform with between normalized intensities of mass peaks in A-DOM samples downstream of the Madeira inflow (from AM1R to AM10R) and the fluvial distance of the sampling sites from sampling site AM1R. We used the same method to find significant correlation between AZ-DOM composition based on ESI[ $\pm$ ] FT-ICR mass spectra and water/bacterial parameters.

Non-metric multidimensional scaling (NMDS) is particularly useful if the analyzed species (molecular formulae) are not linearly responding to environmental gradients (Ramette, 2007). We performed NMDS using the R statistical platform with the package *vegan* to determine correlations of environmental variables with DOM molecular composition in the Amazon River continuum. The environmental data used for NMDS included pH, dissolved oxygen (DO), conductivity (Cond.), temperature (Temp.), DOC, POC, TOC, as well as inorganic nutrients such as ammonium ( $\text{NH}_4^+$ ), nitrate ( $\text{NO}_3^-$ ), nitrite ( $\text{NO}_2^-$ ), and phosphate ( $\text{PO}_4^{3-}$ ). These physicochemical variables were centered and scaled for ordination. All NMDS models were run on Bray-Curtis distance matrixes and their statistical significance was tested with 999 permutations. We applied NMDS analysis for the DOM samples with the data of bacterial production and carbon fixation.

## 3. Results

### 3.1 Article 1: Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy

**Title:** Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy

**Journal:** Science of The Total Environment, Volume 857, Part 3

**DOI:** 10.1016/j.scitotenv.2022.159620

**Authors:** Siyu Li, Mourad Harir, Philippe Schmitt-Kopplin, Michael Gonsior, Alex Enrich-Prast, David Bastviken, Juliana Valle, Fausto Machado-Silva, Norbert Hertkorn

**Content:** Rivers are natural biogeochemical systems shaping the fates of dissolved organic matter (DOM) from leaving soils to reaching the oceans. This study focuses on Amazon basin DOM processing employing negative and positive electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI[±] FT-ICR MS) and nuclear magnetic resonance spectroscopy (NMR) to reveal effects of major processes on the compositional space and structural characteristics of black, white and clear water systems. These include non-conservative mixing at the confluences of (1) Solimões and the Negro River, (2) the Amazon River and the Madeira River, and (3) in-stream processing of Amazon River DOM between the Madeira River and the Tapajós River. The Negro River (black water) supplies more highly oxygenated and high molecular weight compounds, whereas the Solimões and Madeira Rivers (white water) contribute more CHNO and CHOS molecules to the Amazon River main stem. Aliphatic CHO and abundant CHNO compounds prevail in Tapajos River DOM (clear water), likely originating from primary production. Sorption onto particles and heterotrophic microbial degradation are probably the principal mechanisms for the observed changes in DOM composition in the Amazon River and its tributaries.

**Contributions:** David Bastviken, Alex Enrich-Prast, Michael Gonsior, Philippe Schmitt-Kopplin, and Norbert Hertkorn designed research. Mourad Harir, Philippe Schmitt-Kopplin, Juliana Valle, Fausto Machado-Silva, David Bastviken, Michael Gonsior, and Norbert Hertkorn participated in the sampling. [Siyu Li](#) and Mourad Harir acquired FT-ICR mass spectra. [Siyu Li](#) and Norbert Hertkorn acquired NMR spectra. Fausto Machado-Silva conducted HBP and DCF measurements. Data interpretation was performed by all authors. [Siyu Li](#) wrote the first draft of the manuscript. All authors provided significant contributions to the final manuscript. About 60% of the work was done by [Siyu Li](#).

## 3.2 Article 2: Distinct non-conservative behavior of dissolved organic matter after mixing Solimões/Negro and Amazon/Tapajós River waters

**Title:** Distinct non-conservative behavior of dissolved organic matter after mixing Solimões/Negro and Amazon/Tapajós River waters

**Journal:** ACS ES&T water

**DOI:** 10.1021/acsestwater.2c00621

**Authors:** Siyu Li, Mourad Harir, Philippe Schmitt-Kopplin, Fausto Machado-Silva, Michael Gonsior, David Bastviken, Alex Enrich-Prast, Juliana Valle, Norbert Hertkorn

**Content:** Positive and negative electrospray ionization FT-ICR mass spectra and  $^1\text{H}$  NMR revealed major compositional and structural changes of dissolved organic matter (DOM) after mixing two sets of river waters in Amazon confluences: the Solimões and Negro Rivers (S+N), and the Amazon and Tapajós Rivers (A+T). We also studied the effects of water mixing ratios and incubation time on composition and structure of DOM molecules. NMR spectra demonstrated large-scale structural transformations in case of S+N mixing, with gain of pure and functionalized aliphatic units, and loss of all other structures after one-day incubation. A+T mixing resulted in comparatively minor structural alterations, with major gain of small aliphatic biomolecular binding motifs. Remarkably, structural alterations from mixing to one-day incubation were in essence reversed from one-day to five-day incubation for both S+N and A+T mixing experiments. Heterotrophic bacterial production (HBP) in endmembers S, N, and S+N mixtures remained near  $0.03 \mu\text{gC L}^{-1}\text{h}^{-1}$  whereas HBP in A, T, A+T were about five times higher. High rates of dark carbon fixation (DCF) took place at S+N mixing in particular. In-depth biogeochemical characterization revealed major distinction of DOM biogeochemical changes and temporal evolution at these key confluence sites within the Amazon basin.

**Contributions:** David Bastviken, Alex Enrich-Prast, Michael Gonsior, Philippe Schmitt-Kopplin, and Norbert Hertkorn designed research. Mourad Harir, Philippe Schmitt-Kopplin, Juliana Valle, Fausto Machado-Silva, David Bastviken, Michael Gonsior, and Norbert Hertkorn participated in the sampling. Siyu Li and Mourad Harir acquired FT-ICR mass spectra. Siyu Li and Norbert Hertkorn acquired NMR spectra. Fausto Machado-Silva conducted HBP and DCF measurements. Data interpretation was performed by all authors. Siyu Li wrote the first draft of the manuscript. All authors provided significant contributions to the final manuscript. About 60% of the work was done by Siyu Li.

## 4. General discussion and outlook

### 4.1 Molecular insights into DOM in the Amazon River and its tributaries

Terrestrial aquatic systems - such as rivers, lakes and wetlands - act as transporters, storages and reactors of organic matter (Battin et al., 2009; Aufdenkampe et al., 2011; Abril et al., 2014). The Amazon River is particularly important because its large freshwater discharge contributes over  $200,000 \text{ m}^3 \text{ s}^{-1}$  (Ward et al., 2015) and it exports 22-27 Tg of DOC annually (Richey et al., 1990). The mainstem of the Amazon River experiences peak discharge during May/June and minimal discharge in November/December (Richey et al., 1990). Low light penetration due to high turbidity and high concentrations of light-absorbing chromophoric DOM are limiting phytoplankton production within the Amazon River system (Rai and Hill, 1984; Wissmar et al., 1981). Most of the river organic matter is therefore derived from land plants with additional autochthonous organic matter inputs from adjacent floodplains and floodplain lakes (Moreira-Turcq et al., 2013).

Dissolved organic matter (DOM) forms a highly complex mixture of thousands of individual molecules interacting with other coexisting solutes and non-dissolved constituents (Kujawinski, 2011; Gonsior et al., 2016). Our joint molecular characterization of SPE-DOM in the Amazon main stem and tributaries by ESI[±] FT-ICR MS and  $^1\text{H}$  NMR provided improved resolution of the alterations of DOM along the Amazon River, and demonstrated distinct reactivity for CHO, CHNO, and CHOS molecules. The molecular features of Amazon DOM can serve as biogeochemical tracers and provide compelling information on their source, reactivity, and processing. ESI[-] FT-ICR MS preferentially detected high-mass oxygen-rich CHO compounds, whereas ESI[+] FT-ICR MS primarily detected aliphatic CHNO compounds (Hertkorn et al., 2013). Our previous study on the SPE-DOM in Amazon main stem and tributaries showed that nearly half of CHO and CHNO  $m/z$  ions were found in both ESI modes at near-average H/C and O/C ratios, representing a large number of isomeric molecules that as a whole were resistant to rapid alteration.

We found a group of more saturated CHO compounds was unique to the Tapajós River SPE-DOM, likely resulting from abundant algal biomass and higher primary production in Tapajós (Moreira-Turcq et al., 2003a), in line with previous work that had in situ primary production of C4 grasses or algal material identified as the key autochthonous source of DOM in Tapajós River (Quay et al., 1992; Ward et al., 2016). Membrane lipids are known products of microbial degradation of phytoplankton and might serve as potential precursors of DOM (Harvey et al., 2006). Phytoplankton blooms provide new molecules relevant to the DOM pool in Amazon plume waters (Medeiros et al., 2015).

CHNO molecules represent important potentially biolabile components of DOM that indicate land cover,



microbial processing, and possibly abiotic nitrogen incorporation (Bernhardt and Likens, 2002; Sleighter et al., 2014). ESI[−] FT-ICR MS primarily ionizes carboxyl-carrying DOM molecules (like e.g., CRAM, carboxyl-rich alicyclic molecules (Hertkorn et al., 2006), which are abundant in common freshwater DOM as observed e.g., by <sup>13</sup>C NMR spectra and titration (Ritchie and Perdue, 2003). ESI[+] FT-ICR MS directly detects N-containing molecules which not necessarily carry carboxylic groups. The complementary of detection of N-containing molecules by ESI[+] and ESI[−] MS is helpful to reveal so far exclusive structural features of CHNO compounds in the Amazon River DOM and allows compositional and structural distinction related to origin and processing of DOM. The abundance of less oxygenated CHNO compounds that mostly ionized in ESI[+] FT-ICR MS was highest in the Amazon River SPE-DOM downstream of the Madeira inflow and then declined rapidly; these more aliphatic compounds likely had a shorter residence time, in agreement with higher general biogeochemical processing of heteroatom-containing molecules (Ksionzek et al., 2016). Gonsior et al. (2016) and Simon et al. (2019) found differences in higher *m/z* molecular formulae and in numbers of CHNO formulae in the Negro waters (Gonsior et al., 2016; Simon et al., 2019). Our study has a better coverage of the *m/z* range and the quantity of N-containing compounds compared to the two previous studies. The differences in different studies could be explained by several aspects. First, our study used ESI[±] FT-ICR MS and provided better coverage of molecular signatures. Second, the number of CHNO compounds would be affected by season. Gonsior et al. (2016) took their samples in May 2013, after the end of the rain season (“hot and wet”, daily precipitation events) and the beginning maximum of the flood pulse (Vidal et al., 2015). The start of the flood peak will presumably differ largely from its end (initial disturbance vs ongoing/ending disturbance), for example in terms of dissolved constituents leached from rewetted soils. Simon et al. (2019) took samples during the dry period (“hot and humid”, non-daily precipitation), at the end of the maximum of the flood peak (July 2014) and its lowest extent (October 2013), when connectivity between floodplains and mainstem gradually decreases and upstream flushing of soils by daily precipitation events is reduced. In comparison, the samples from this study were taken between April 2nd and May 25th in 2014, during a high-water period with extremely high discharge volume.

Higher content of S-containing compounds was found in the Solimões River SPE-DOM and in the Amazon River SPE-DOM downstream of the Madeira inflow, which may have resulted from incorporation of sulfur into organic matter by dissimilatory sulfate-reducing bacteria and/or by abiotic sulfurization reactions under anoxic conditions in sediment-rich waters like Solimões and Madeira (Luek et al., 2017; Sinninghe Damste and De Leeuw, 1990; Schmidt et al., 2009), but anthropogenic sources cannot be excluded (Latrubesse et al., 2017). Furthermore, DOM sulfurization processes are likely to be more intense under more anaerobic conditions during low-water periods relative to the high-water period we sampled.

Unlike mass spectrometry, which has intrinsic limitations like structure-dependent ionization efficiency (Hertkorn et al., 2008),  $^1\text{H}$  NMR more comprehensively covered all structures containing non-exchangeable hydrogen atoms, although the huge diversity of hydrogen atomic environments in DOM produced a considerable overlap of NMR resonances. The Negro River SPE-DOM showed a higher degree of oxygenation relative to the Solimões/Amazon/Tapajós-DOM in a very extensive range of aliphatic and aromatic carbon chemical environments. These molecular features were transferred to the Amazon River and became strongly attenuated ~150 km downstream. Moreover, the Amazon River SPE-DOM downstream of the Madeira River inflow and the Uatumã River inflow showed higher relative abundance of pure and functionalized aliphatic protons, presumably indicative of higher proportions of more microbially processed terrestrial DOM (Lechtenfeld et al., 2015; Schmidt et al., 2009).

Our study provides an unprecedented level of detailed molecular information about the different types of waters in the Amazon basin. The blackwater Negro River SPE-DOM showed higher average  $m/z$  ratios and higher aromaticity compared to DOM of other rivers, in agreement with previous studies (Gonsior et al., 2016; Ertel et al., 1986). High molecular weight aromatic structures probably represent the polar fraction of lignin and tannin degradation products, e.g., through biotic and abiotic oxidation (Johannsson et al., 2017; Waggoner et al., 2015). Moreover,  $^1\text{H}$  NMR revealed higher proportions of  $=\text{CH}$  and  $\text{C}_{\text{ar}}\text{H}$  units ( $\delta_{\text{H}} > 5.2$  ppm) and remotely ( $\text{OCC}\text{H}$ ) and directly oxygenated ( $\text{OCH}$ ) aliphatic molecules ( $\delta_{\text{H}} > 2.4$  ppm) in the Negro River DOM. Low retention of organic matter in bleached soils of the Negro basin (Fritsch et al., 2009; Bardy et al., 2011; Quesada et al., 2011) could explain apparently fresher and more aromatic riverine DOM in blackwater catchments. The SPE-DOM in the whitewater Solimões and Amazon Rivers showed higher average N/C and S/C ratios and had larger contribution of moderately unsaturated and saturated compounds. Suspended particles from whitewater Amazonian rivers can be enriched in nitrogenous compounds (Aufdenkampe et al., 2001) and may transport more sorbed organic compounds in general (Kleber et al., 2007), as opposed to the low amount of suspended sediment carried by the Rio Negro. Moreover, >90% of the sediment load of the Rio Negro is composed of kaolinite clay minerals with low adsorption capacity (Irion et al., 1991). The clearwater Tapajós River SPE-DOM showed higher proportions of CHNO compounds, and of methylene- and alkyl-rich aliphatic structures than the other endmember DOM, in line with higher proportions of primary production, and microbial processing.  $^1\text{H}$  NMR spectra of the Tapajós River SPE-DOM showed a distinct visual appearance owing to a few sharp and abundant resonances representing aliphatic molecules with limited branching, acetic acid ( $\delta_{\text{H}} \sim 1.88$  ppm), and formic acid ( $\delta_{\text{H}} \sim 8.08$  ppm), with an overall  $^1\text{H}$  NMR integral below ~2%. The strong contribution of small aliphatic molecules may originate from general proximity of its DOM molecules to the structural space of biomolecules and metabolites, including some biodegradation of fresh materials as suggested by Roth et al.

(Roth et al., 2019).

## 4.2 DOM dynamics after riverine mixing

Abrupt changes of environmental conditions, either in space (“hot spots”, e.g., at interfaces connecting soils, wetlands, sediments and rivers) or time (“hot moments”, e.g., during algal blooms or flooding events, or upon injection of easily degradable substrate) influence the dynamics of carbon pools in natural systems (Marín-Spiotta et al., 2014; Kuzyakov and Blagodatskaya, 2015; Ward et al., 2017). For example, mixing zones such as river confluences and river plumes are hypothesized to be hotspots for organic matter cycling due to processes such as “priming effects,” whereby the decomposition of less reactive organic matter is stimulated by the presence of highly reactive material such as algal exudates (Bianchi, 2011; Bianchi et al., 2015). Mixing phenomena in aqueous media combine both spatial and temporal aspects in a unique way and occur in several settings in nature, most prominent at river confluences or at the boundary of riverine and marine systems (Simon et al., 2019). Although there is a solid body of literature on biogeochemical processes and their effects on elementary mass balances at aquatic boundaries, most of this work has been conducted in estuaries. There is relatively little known about how physical mixing of different riverine water bodies and disequilibria that may evolve during this mixing process and its influence on DOM molecular composition.

Solimões-Negro confluence ranks among the largest on Earth and provides non-conservative mixing and exemplary spatial and temporal DOM processing on grand-scale. High DOC with large proportions of hydroxyl and phenolic sites in Negro water (Duarte et al., 2016; Gonsior et al., 2016) and abundant sediment load in Solimões water (Aucour et al., 2003; Moreira-Turcq et al., 2003b; Subdiaga et al., 2019) present favorable conditions for structure-selective sorption of organic matter on mineral surfaces (Aufdenkampe et al., 2001; Moreira-Turcq et al., 2003b; Subdiaga et al., 2019). Abiotic reactions such as photochemical processing (oxidation and mineralization) and metal-dependent complexation, coagulation and redox chemistry will have significant impact on the DOM composition in the Amazon River (Schmitt-Kopplin et al., 1998; Aucour et al., 2003). Structure-selective sorption of DOM on mineral surfaces (Subdiaga et al., 2019; Li et al., 2023) has removed substantial proportions of polyphenols present in Negro waters together with minerals present in Solimões waters. This removal of light absorbing and light scattering ingredients in the Solimões-Negro mixing zone opens previously unavailable opportunities for photo-processing of DOM molecules. However, many photoproduct oxygenated small molecules will readily integrate in biochemical pathways and foodwebs (Gonsior et al., 2014; Huang et al., 2023) and transient particulate organic carbon, making traceability to specific mechanisms difficult.

We observed conservative behavior of DOC concentration downstream of the Solimões-Negro confluence. This is a surprising finding, since DOC is generally described as reactive in the confluence zone of the Negro and Solimões (Aucour et al., 2003; Leenheer et al., 1980; Moreira-Turcq et al., 2003; Aufdenkampe et al., 2001; Ertel et al., 1986; Merschel et al., 2007). However, the conservative mixing of DOC concentration was in line with two previous studies (Simon et al., 2019; Merschel et al., 2017), which suggest that organic nanoparticulate or colloidal matter did not readily coagulate to form larger aggregates during mixing (Aufdenkampe et al., 2001; Pérez et al., 2011; Merschel et al., 2017). Another reason suggested by Simon et al. is an occupation of mineral sorption sites by DOM molecules (Simon et al., 2019). The largely conservative DOC behavior is supported by artificial sorption experiments conducted by Pérez et al. (2011). The authors found that the hydrophobic fraction (sorbed to XAD-8 resin, “HPO” fraction) of Rio Negro DOC, besides sorption of ~80% of DOC (initial ca. 18 mg C/L) to fresh synthetic goethite, was not fractionated in terms of aromaticity during the process, even under large excess of DOC. Contrastingly, the aromaticity of the relatively hydrophilic fraction (XAD-8 effluent sorbed onto XAD-4 resin, “transphilic”, “TPH” fraction) decreased upon sorption. Similar high levels of DOC were sorbed. Altogether, this indicated a preferential sorption of relatively hydrophilic aromatic compounds onto sediment at high DOC loads. A second sorption experiment with an isolated organic matter fraction from the Amazon–Curuai floodplain and organic-free natural sediment from the Curuai Várzea Lake showed much lower sorption (5–20% DOC), however, initial DOC levels were also lower in these experiments (ca. 4.2 mg C/L). The authors did not observe any fractionation in terms of aromaticity in these second experiments either. Under the natural confluence settings, with DOC concentrations in the range of the second experiment and mineral surfaces in Rio Solimões already conditioned with organic matter, sorption or fractionation of new DOM (i.e., Rio Negro DOM) may be even further reduced. We conducted artificial sorption experiment and found that the molecular characteristics of the Solimões River water SPE-DOM was nearly identical to the SPE-DOM from the mixture of the Solimões River water and the filtered mineral particles from the Negro River. Likewise, the molecular characteristics of the Negro River water SPE-DOM were nearly identical to the SPE-DOM from the mixture of the Negro River water and the filtered mineral particles from the Solimões River. These results strongly supported Pérez’s finding and proved that the mineral surfaces and DOM in the Solimões and/or Negro Rivers were in equilibrium before mixing.

We found considerable decrease of dissolved oxygen (~30%) and POC (~90%) concentrations and concomitant enhanced microbial activities just downstream of Solimões-Negro confluence, in agreement with the previous report for mixing zones in the Amazon basin (Farjalla, 2014). The large depletion of POC at Solimões-Negro mixing zone, in addition to large number of molecular signatures associated with POC in the Amazon River downstream suggest that the DOC-POC transition may have significant effects on

DOM composition, especially at confluences (Subdiaga et al., 2019; Moreira-Turcq et al., 2003b).

Previous studies suggested overall low aquatic bacterial metabolism for the Amazon waters, but confluences could be hot spot areas of high bacterial metabolism, which are associated with changes in organic matter quality (Farjalla et al., 2014). Large proportions of polyphenolic compounds in the Negro DOM were removed from the DOM pool when Negro met Solimões. The less oxygenated, lower molecular weight CHO/CHNO compounds associated with higher HBP might be preferentially assimilated and converted into new bacterial biomass. Furthermore, we found for the first time that DCF contributed to microbial biomass production to the same or even higher extent than heterotrophic processes just downstream of the Solimões-Negro mixing zone, suggesting that the inorganic carbon pathway may be as important as organic carbon degradation when river waters with distinct characteristics converge. In addition, abiotic reactions such as photochemical processing (oxidation and mineralization) and metal complexation could also have important impact on changes of DOM composition in the Amazon River downstream (Amon and Benner, 1996b; Aucour et al., 2003).

The Tapajós River hosts higher proportions of cyanobacteria (~12%) compared to the Amazon River (~5%), where actinobacteria and betaproteobacteria dominate the microbial community (Doherty et al., 2017). “Positive priming effect” was suggested to happen when algal-rich Tapajós waters mix with turbid Amazon main stem waters, accelerating the decomposition of more recalcitrant organic matter (Seidel et al., 2015). The molecular transformation of CHO, CHNO, and CHOS molecules towards lower  $m/z$ , lower O/C ratios with concomitant higher H/C ratios in A+T SPE-DOM during the five-day incubation is consistent with a simple notion of DOM degradation from larger to smaller molecules, which was expressed in the familiar size-reactivity continuum model (Benner et al., 2015), but here expands to relatively small DOM molecules and/or metabolites. Biodegradation might explain the strong contribution of small molecules, but recent findings suggest that contributions of low molecular weight constituents in FT-ICRMS data may be related to “fresh” or freshly decomposing material (Roth et al., 2016). Moreover,  $^1\text{H}$  NMR spectra indicated polyphenolic molecules were largely removed in A+T SPE-DOM after five-day incubation, suggesting rapid processing of terrestrial organic matter (e.g., leached from litter decay processes) that had been largely hydrologically mobilized in aquatic-terrestrial transition zones (Bertassoli et al., 2017; Melo et al., 2019). This conforms to a study showing that the breakdown of vascular plant-derived organic matter to  $\text{CO}_2$  increased six-fold at Amazon-Tapajós confluence compared to that in the Amazon River (Ward et al., 2016).

Most studies report distinct non-conservative behavior of DOM after riverine mixing affected by factors like phase partitioning (Seidel et al., 2016; Moreira-Turcq et al., 2003), respiration (Mayorga et al., 2005; Abril et al., 2014), trace metal complexation (Simon et al., 2019), and discharge (Ward et al., 2018).

However, the effects of physical mixing of different Amazon basin riverine waters on the composition and structure of its DOM at molecular resolution remain ill-constrained. We observed distinct structural evolution after mixing the two sets of waters: the Solimões and Negro Rivers (S+N) and the Amazon and Tapajós Rivers (A+T). In A+T, a suite of labile lipid-like molecules with prominent alkyl- and methylene-rich carboxylic acids of limited structural diversity, and few oxygenated aliphatic and aromatic structural motifs had been produced after one-day incubation and largely disappeared after five-day incubation. In comparison, the structural changes after one-day incubation of S+N DOM comprised a much larger structural diversity of molecules, covering the entire range of aliphatic (CCCH<sub>2</sub>, OCC<sub>2</sub>H, OCH<sub>2</sub> units) and C<sub>sp2</sub>H units (aromatics and olefins) with a smooth distribution of all structural motifs (atomic environments), implying a more thorough diversification of atomic environments in S+N compared with A+T following mixing and incubation. This probably reflects an overall larger distinction of DOM molecules in original Solimões and Negro waters compared with Amazon and Tapajós waters, originating from the high relative abundance of highly oxygenated aromatic structures in Negro waters (Ertel et al., 1986; Johannsson et al., 2017).

The structural evolution from one-day to five-day incubation was approximately a reversal of the structural changes from original samples to one-day incubation for both S+N and A+T mixing experiments: A+T samples diversified their pure and remotely oxygenated aliphatic molecular structures and lost methylene- and alkyl-rich aliphatic molecules with simple branching motifs. For S+N, the evolution of all aliphatic units from one-day to five-day incubation appeared as very congruent reversal compared with evolution from the average of S+N samples to one-day incubation. This fast “recovery” behavior of DOM processing after temporal mixing reflected remarkably fast DOM processing for both S+N and A+T, and has likely been promoted by tropical temperatures (~30°C).

Most studies report non-conservative effects upon riverine mixing (i.e., deviating from simple dilution), which affect phase partitioning and are frequently linked to distinct changes in DOM composition (Aucour et al., 2003; Leenheer et al., 1980; Moreira-Turcq et al., 2003; Aufdenkampe et al., 2001; Ertel et al., 1986; Merschel et al., 2007). A study by Simon and coworkers observed a conservative behavior of SPE-DOM in the natural and experimental mixing of Solimões and Negro River waters based on negative ionization mode FT-ICR MS (Simon et al., 2019), opposing the presumed non-conservative behavior of DOM in our results. Simon et al. sampled during the dry period, while the samples from this study were taken during a maximum flood pulse. The different sampling season presumably affects composition and structure of dissolved constituents leached from rewetted soils and may influence the endmember water characteristics and microbial community (McClain et al., 2008). Moreover, Simon and coworkers mixed samples under continuous shaking, while we mixed initially and then let the sample settle. The differences in sample

treatment may affect DOM composition, for instance through sorption and desorption between DOM and particles. Besides, the combination of SPE and ESI[-] FT-ICR MS favors observance of carboxylic acids and does not cover the entire spectrum of dissolved organic molecules (Hawkes et al., 2016; Raeke et al., 2016; Li et al., 2017). Our joint use of both complementary electrospray modes FT-ICR MS and <sup>1</sup>H NMR methods and the exhaustive analysis provide valuable DOM molecular information with improved detail.

Ward et al. (2019) conducted experimental mixing on the Amazon and Tapajós Rivers waters and observed that regardless of what water was used, respiration rates are directly linked to the velocity at which samples are rotated. It was hypothesized that this linkage between velocity and aquatic metabolism occurs as a result of interactions between suspended particles, dissolved constituents, and free-living and particle bound microbes considering that the presence of particles have been shown to play an important role in microbial activity and priming effects. The amplified and diversified microbial activity as tributaries mixed would result in enhanced respiration rates, and likely the fast processing of organic carbon. Moreover, the desorption of nitrogen-rich molecules as Amazon River suspended sediments are diluted, providing both nutrients and reactive substrates that could drive priming effects (Ward et al., 2019; Aufdenkampe et al., 2001).

### **4.3 Environmental implications**

Aquatic ecosystems represent conduits in the global carbon cycle that connect the terrestrial biosphere with the oceans and atmosphere (Cole et al., 2007; Aufdenkampe et al., 2011). On a global scale, roughly 75% of the 5.7 Pg C transported by rivers annually is emitted to the atmosphere (Wehrli, 2013), with the majority of this evasive flux occurring in the tropics (Raymond et al., 2013). Microbial degradation (Ward et al., 2013) and photochemical oxidation (Opsahl and Benner, 1998) are the primary pathways for the conversion of terrestrially derived organic carbon to carbon dioxide in the aquatic environment (Battin et al., 2008). The relative importance of these processes is controlled in part by biogeochemical parameters derived from the surrounding watershed along with physical river properties. For example, in the blackwater Rio Negro, roughly 15% of the DOC has been shown to be photoreactive (Amon and Benner, 1996). However, a large fraction of Amazon Rivers is sediment rich and deep relative to light penetration depths, resulting in a minimal contribution of photooxidation to basin scale CO<sub>2</sub> production (less than 1%) relative to depth-integrated respiration (Remington et al., 2011).

In the Amazon River, the presence of large lowland tributaries and productive floodplains provides ample opportunities for “priming effect” to occur within the basin, effectively reducing the amount of organic carbon that would have otherwise been delivered to the ocean. For example, roughly 95% of the terrestrially

derived carbon that enters the Amazon River is outgassed to CO<sub>2</sub>, meaning that less than 5% of the organic carbon mobilized by the system is transported to the ocean (Richey et al., 1992; Ward et al., 2013). While active floodplains may be important with respect to the priming of terrestrially derived organic carbon, they also maintain high levels of in situ organic carbon along river continuums (Ward et al., 2015).

Multistep processing of DOM along the flow path and mixing zones of many ungauged Amazon River tributaries implies that the carbon atoms that are eventually delivered to the Atlantic Ocean have repeatedly changed their atomic environments on their journey (Roth et al., 2019). However, general congruence of biochemical pathways across all organisms, and similarities of fundamental abiotic processing regimes in the extended tropical ecosystem of the Amazon basin will rather realign thermodynamic and kinetic boundary conditions at locations of exceptional importance, leading to thorough redirection for synthesis and degradation of DOM molecules. This is more likely than an unspecific continual entropy-driven diversification of DOM molecules from land to sea. Overall, DOM evolution in the Amazon Basin will diversity beyond trajectories of simple synthesis and degradation as described in the standard river continuum concept (Vannone et al., 1980), and/or small river catchments (Kamjunke et al., 2019).

According to the “regional chromatography” hypothesis (Ertel et al., 1986; Devol et al., 2001), the majority of reactive components within the DOC pool are effectively removed before reaching higher-order confluences such as the S+N confluence. The process could also lead to occupancy of sorption sites at minerals, rendering both educts (minerals and DOM) non-reactive (in line with findings by Pérez et al. (2011)). Hence, we hypothesized that the individual endmember SPE-DOM had reached a steady state under given biogeochemical conditions. However, the assemblages of biogeochemically distinct river waters initiated rapid and widespread molecular alteration of DOM upon mixing, in part by fast abiotic reactions. This averaged physiochemical properties and reassembled microbial community composition, with potentially considerable effects on DOM processing and mineralization, as well as the composition and reactivity of remaining organic molecules (Schmidt et al., 2011). Pronounced changes in DOM composition during incubation suggest substantial diagenetic modification of organic matter, even if the complete mineralization of DOM may not be significant (Valle et al., 2017; Einsiedl et al., 2007). Nevertheless, continuous oxidation of the residual DOM molecules is likely to generate oxygen-rich CO<sub>2</sub> and H<sub>2</sub>O molecules in the river confluence and subsequent river channels, contributing to a nominal reduction of leftover DOM molecules even in case of oxidative processing (Einsiedl et al., 2007).

More than 80% of river length of the Amazon River system is connected to the smallest streams (first and second order streams; McClain and Elsenbeer, 2001). It seems certain that the role of smaller tributaries and lower order confluences will be underestimated when only large confluences are considered; the non-



conservative processes might already have taken place further upstream or even directly at the soil/water interface (Guinoiseau et al., 2017; Ward et al., 2017).

Rivers provide a direct link between terrestrial and marine carbon cycles. However, very little terrestrial DOM appears to accumulate in the global ocean (Eadie et al., 1978; Meyers-Schulte et al., 1986; Opsahl et al., 1997). Incubation experiments with turbid water from close to the river mouth suggested that photo- and bio-alteration leave significant molecular and carbon isotopic imprints on the terrigenous DOM. Nonetheless, quantitative removal by bio- and photo-degradation appeared to proceed at a relatively slow pace because no significant dissolved organic carbon decrease was found within the five days of the incubation (Seidel et al., 2015). However, 9-30% of DOC was lost after the five-day incubation with less turbid water from intermediate and outer plume, suggesting that the introduction of reactive algal DOM in the intermediate plume may thus have primed the microbial degradation of terrigenous DOM (Seidel et al., 2015; Bianchi et al., 2011). Sorption of terrigenous DOM to sinking particles acts as an important DOC sink in the Amazon plume (Seidel et al., 2015). These results, along with our results, demonstrate that the DOM along the Amazon River-to-ocean continuum was not in a steady state as suggested by “regional chromatography” hypothesis. Moreover, river confluences and less turbid offshore plume are likely important carbon sinks in the Amazon River due to biotic processes (e.g., photo-degradation and bio-degradation) and abiotic processes (e.g., sorption to particles).

The higher heterotrophic bacterial activity in A+T compared with S+N suggested that the molecular changes in A+T SPE-DOM during incubation referred more to bioavailable alkyl-rich molecules than in case of S+N that were more oxygenated on average. Moreover, DCF rates noticeably increased in both mixing experiments and amounted up to ~95% of the total heterotrophic bacterial production after the mixing, indicating that non-photosynthetic carbon fixation can represent a substantial contribution to an autochthonous source of organic matter in river confluences (Jiao et al., 2010). DCF had been reported to contribute approximately 80% of the total heterotrophic bacterial production in Swedish boreal lake sediments (Santoro et al., 2013). Moreover, it has been shown that the heterotrophic metabolism of bacteria is suddenly intensified after input of organic matter (Baltar et al., 2019). The water-particle interface in river confluences could be sites of intensive biogeochemical activity, creating steep chemical gradients that provide a microenvironment with high chemolithotrophic rate (Jorcin et al., 2015). The simultaneous presence of oxidized and reduced organic molecules supplies energy and substrates needed to support DCF by chemoautotrophic microbes.

## 4.4 Conclusion

Complementary ESI[±] mass spectrometry and NMR spectroscopy provided comprehensive coverage of molecular evolution of DOM for exemplary processes in the Amazon basin, and distinct reactivity for CHO, CHNO and CHOS molecules. The Negro River (black water) supplies more highly oxygenated and high molecular weight compounds, whereas the Solimões and Madeira Rivers (white water) contribute more CHNO and CHOS molecules to the Amazon River main stem. Aliphatic CHO and abundant CHNO compounds prevail in Tapajós River DOM (clear water), likely originating from primary production. Sorption onto particles and heterotrophic microbial degradation are probably the principal mechanisms for the observed changes in DOM composition in the Amazon River and its tributaries.

NMR spectra demonstrated large-scale structural transformations in case of the Solimões and Negro Rivers (S+N) mixing, with gain of pure and functionalized aliphatic units, and loss of all other structures after one-day incubation. The Amazon and Tapajós Rivers (A+T) mixing resulted in comparatively minor structural alterations, with major gain of small aliphatic biomolecular binding motifs. Remarkably, structural alterations from mixing to one-day incubation were in essence reversed from one-day to five-day incubation for both S+N and A+T mixing experiments. In-depth biogeochemical characterization revealed major distinction of DOM biogeochemical changes and temporal evolution at these key confluence sites within the Amazon basin.

Relevant molecular resolution concerning fate, transport, and transformation of DOM in natural and engineered systems is not attained without use of complementary high-resolution structure-selective analytical and biogeochemical methods. So far elusive structural features of CHNO compounds in Amazon DOM, and DOM in other ecosystems/biomes, will become better constrained by comparative analysis of CHO and CHNO compounds in ESI[+] and ESI[-] mass spectra for individual samples and processes. Joint ESI[±] mass spectrometry should become commonplace for further studies of DOM processing in ecosystems because strong ionization selectivity in ESI[-] mass spectra provides conceptual assessments with serious intrinsic limitations. NMR spectra offers valuable constraints of hydrogen and carbon atomic environments not available by FT mass spectra and is important for providing structural information and obtaining credible conclusions.

Ecological models and biogeochemical characterization have established river mixing zones as hotspots of microbial activity, but molecular DOM degradation pathways are still poorly addressed in river networks, preventing an accurate estimate of carbon sources and sinks. We observed extensive but molecularly distinct structural alterations of DOM upon mixing and follow-up incubation of Solimões/Negro on one hand and Tapajós/Amazon waters on the other hand. Molecular distinction of DOM structural evolution will apply in essence to all river confluences, and educated modelling eventually has to take these mechanistic details of DOM processing into account.

## 5. References

- Abelho, M. From litterfall to breakdown in streams: a review. *Sci. World J.* **1**, 656-680, doi: 10.1100/tsw.2001.103 (2001).
- Åberg, J. and Wallin, M. Evaluating a fast headspace method for measuring DIC and subsequent calculation of pCO<sub>2</sub> in freshwater systems. *Inland Waters* **4**, 157-166, doi: 10.5268/IW-4.2.694 (2014).
- Abril, G. *et al.* Amazon River carbon dioxide outgassing fuelled by wetlands. *Nature* **505**, 395-398, doi: 10.1038/nature12797 (2014).
- Affonso, A. *et al.* Water quality changes in floodplain lakes due to the Amazon River flood pulse: Lago Grande de Curuaí (Pará). *Braz. J. Biol.* **71**, 601-610, doi: 10.1590/S1519-69842011000400004 (2011).
- Albéric, P. *et al.* Variation of the isotopic composition of dissolved organic carbon during the runoff cycle in the Amazon River and the floodplains. *Compt. Rend. Geosci.* **350**, 65-75, doi: 10.1016/j.crte.2017.11.001 (2018).
- Alonso-Sáez, L. *et al.* Bacterial uptake of low molecular weight organics in the subtropical Atlantic: Are major phylogenetic groups functionally different? *Limnol. Oceanogr.* **57**, 798-808, doi: 10.4319/lo.2012.57.3.0798 (2012).
- Altare, N. & Vione, D. Photochemical Implications of Changes in the Spectral Properties of Chromophoric Dissolved Organic Matter: A Model Assessment for Surface Waters. *Molecules* **28**, 2664, doi: 10.3390/molecules28062664 (2023).
- Amador, J. A. *et al.* Sequential photochemical and microbial degradation of organic molecules bound to humic Acid. *Appl. Environ. Microbiol.* **55**, 2843-2849, doi: 10.1128/aem.55.11.2843-2849.1989 (1989),
- Amaral, A. *et al.* Cover plant residues and mobility of dissolution products of surface applied lime. *Revista Brasileira de Ciência do Solo* **28**, 115-123, doi: 10.1590/S0100-06832004000100012 (2004).
- Amery, F. *et al.* The copper-mobilizing-potential of dissolved organic matter in soils varies 10-fold depending on soil incubation and extraction procedures. *Environ. Sci. Technol.* **41**, 2277-2281, doi: 10.1021/es062166r (2007).
- Amon, R. M. W. & Benner, R. Bacterial utilization of different size classes of dissolved organic matter. *Limnol. Oceanogr.* **41**, 41-51, doi: 10.4319/lo.1996.41.1.0041 (1996a).
- Amon, R. M. W. & Benner, R. Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River system. *Geochim. Cosmochim. Ac.* **60**, 1783-1792, doi: 10.1016/0016-7037(96)00055-5 (1996b).

- Andrade, F. *et al.* Addition of organic and humic acids to Latosols and phosphate adsorption effects. *Revista Brasileira de Ciência do Solo* **27**, 1003-1011, doi: 10.1590/S0100-06832003000600004 (2003).
- Arantes, C. C. *et al.* Large-scale hydropower impacts and adaptation strategies on rural communities in the Amazonian floodplain of the Madeira River. *J. Environ. Manage.* **336**, 117240, doi: 10.1016/j.jenvman.2023.117240 (2023).
- Argyropoulos, D. S. & Sun, Y. Photochemically induced solid-state degradation, condensation, and rearrangement reactions in lignin model compounds and milled wood lignin. *Photochem. Photobiol.* **64**, 510-517, doi: 10.1111/j.1751-1097.1996.tb03098.x (1996).
- Armanious, A. *et al.* Dissolved organic matter adsorption to model surfaces: adlayer formation, properties, and dynamics at the nanoscale. *Environ. Sci. Technol.* **48**, 9420-9429, doi: 10.1021/es5026917 (2014).
- Aucour, A.M. *et al.* The Amazon River: behaviour of metals (Fe, Al, Mn) and dissolved organic matter in the initial mixing at the Rio Negro/Solimões confluence. *Chem. Geol.* **197**, 271-285, doi: 10.1016/S0009-2541(02)00398-4 (2003).
- Athayde, S. *et al.* Mapping research on hydropower and sustainability in the Brazilian Amazon: advances, gaps in knowledge and future directions. *Curr. Opin. Environ. Sus.* **37**, 50-69, doi: 10.1016/j.cosust.2019.06.004 (2019).
- Aufdenkampe, A. K. *et al.* Sorptive fractionation of dissolved organic nitrogen and amino acids onto fine sediments within the Amazon Basin. *Limnol. Oceanogr.* **46**, 1921-1935, doi: 10.4319/lo.2001.46.8.1921 (2001).
- Aufdenkampe, A. K. *et al.* Organic matter in the Peruvian headwaters of the Amazon: Compositional evolution from the Andes to the lowland Amazon mainstem. *Org. Geochem.* **38**, 337-364, doi: 10.1016/j.orggeochem.2006.06.003 (2007).
- Azam, F. *et al.* The ecological role of water-column microbes in the sea. *Marine Ecol. Prog. Ser. Oldendorf* **10**, 257-263, doi: 10.3354/meps010257 (1983).
- Baltar, F. *et al.* Major effect of hydrogen peroxide on bacterioplankton metabolism in the Northeast Atlantic. *PLoS One* **8**, e61051, doi: 10.1371/journal.pone.0061051 (2013).
- Baltar, F. & Herndl, G. J. Ideas and perspectives: Is dark carbon fixation relevant for oceanic primary production estimates? *Biogeosciences* **16**, 3793-3799, doi: 10.5194/bg-16-3793-2019 (2019).
- Bao, H., Wu, Y. & Zhang, J. Spatial and temporal variation of dissolved organic matter in the Changjiang: Fluvial transport and flux estimation. *J. Geophys. Res.: Biogeosci.* **120**, 1870-1886, doi: 10.1002/2015jg002948 (2015).

- Bardy, M. *et al.* Podzolization and exportation of organic matter in black waters of the Rio Negro (upper Amazon Basin, Brazil). *Biogeochemistry* **106**, 71–88, doi: 10.1007/s10533-010-9564-9 (2011).
- Bass, A. M. *et al.* Carbon Dioxide and Methane Emissions from a Wet-Dry Tropical Floodplain in Northern Australia. *Wetlands* **34**, 619-627, doi: 10.1007/s13157-014-0522-5 (2014).
- Battin, T. J. *et al.* Biophysical controls on organic carbon fluxes in fluvial networks. *Nat. Geosci.* **1**, 95-100, doi: 10.1038/ngeo101 (2008).
- Battin, T. J. *et al.* The boundless carbon cycle. *Nat. Geosci.* **2**, 598-600, doi: 10.1038/ngeo618 (2009).
- Benner, R. & Amon, R. M. The size-reactivity continuum of major bioelements in the ocean. *Annu. Rev. Mar. Sci.* **7**, 185-205, doi: 10.1146/annurev-marine-010213-135126 (2015).
- Benner, R. & Kaiser, K. Biological and photochemical transformations of amino acids and lignin phenols in riverine dissolved organic matter. *Biogeochem.* **102**, 209-222, doi: 10.1007/s10533-010-9435-4 (2010).
- Benner, R. *et al.* Bacterial carbon metabolism in the Amazon River system. *Limnol. Oceanogr.* **40**, 1262-1270, doi: 10.4319/lo.1995.40.7.1262 (1995).
- Benner, R. *et al.* Bulk chemical characteristics of dissolved organic matter in the ocean. *Science* **255**, 1561-1564, doi: 10.1126/science.255.5051.1561 (1992).
- Bernhardt, E. S. & Likens, G. E. Dissolved organic carbon enrichment alters nitrogen dynamics in a forest stream. *Ecology* **83**, 1689-1700, doi: 10.2307/3071988 (2002).
- Bertassoli, D. J. *et al.* The Fate of Carbon in Sediments of the Xingu and Tapajós Clearwater Rivers, Eastern Amazon. *Front. Mar. Sci.* **4**, 44, doi: 10.3389/fmars.2017.00044 (2017).
- Bianchi, T. S. The role of terrestrially derived organic carbon in the coastal ocean: a changing paradigm and the priming effect. *Proc. Natl. Acad. Sci. USA* **108**, 19473-19481, doi: 10.1073/pnas.1017982108 (2011).
- Bianchi, T. S. Positive priming of terrestrially derived dissolved organic matter in a freshwater microcosm system. *Geophys. Res. Lett.* **42**, 5460–5467, doi: 10.1002/2015GL064765 (2015).
- Bianchi, T. S. & Ward, N. D. Editorial: The Role of Priming in Terrestrial and Aquatic Ecosystems. *Front. Earth Sci.* **7**, 321, doi: 10.3389/feart.2019.00321 (2019).
- Biddanda, B. & Benner, R. Carbon, nitrogen, and carbohydrate fluxes during the production of particulate and dissolved organic matter by marine phytoplankton. *Limnol. Oceanogr.* **42**, 506-518, doi: 10.4319/lo.1997.42.3.0506 (1997).
- Billett, M. *et al.* Carbon balance of UK peatlands: current state of knowledge and future research challenges. *Clim. Res.* **45**, 13-29, doi: 10.3354/cr00903 (2010).

- Blanchet, M. *et al.* When riverine dissolved organic matter (DOM) meets labile DOM in coastal ters: changes in bacterial community activity and composition. *Aquat. Sci.* **79**, 27-43, doi: 10.1007/s00027-016-0477-0 (2017).
- Bolan, N. S. *et al.* Dissolved organic matter: biogeochemistry, dynamics, and environmental significance in soils. *Adv. Agronomy* **110**, 1-75, doi: 10.1016/B978-0-12-385531-2.00001-3 (2011).
- Boon, P. I. & Mitchell, A. Methanogenesis in the sediments of an Australian freshwater wetland: Comparison with aerobic decay, and factors controlling methanogenesis. *FEMS Microbiol. Ecology* **18**, 175-190 (1995).
- Borggaard, O. K. *et al.* Potential of dissolved organic matter (DOM) to extract As, Cd, Co, Cr, Cu, Ni, Pb and Zn from polluted soils: a review. *Geoderma* **343**, 235-246, doi: 10.1016/j.geoderma.2019.02.041 (2019).
- Bouchez, J. *et al.* River mixing in the Amazon as a driver of concentration-discharge relationships. *Water Resour. Res.* **53**, 8660-8685, doi: 10.1002/2017WR020591 (2017).
- Bouwman, A. F. *et al.* Nutrient dynamics, transfer and retention along the aquatic continuum from land to ocean: towards integration of ecological and biogeochemical models. *Biogeosci.* **10**, 1-22, doi: 10.5194/bg-10-1-2013 (2013).
- Brinkmann, T. *et al.* Photobleaching of humic rich dissolved organic matter. *Aquat. Sci.* **65**, 415-424, doi: 10.1007/s00027-003-0670-9 (2003).
- Cai, D. L. *et al.* Sources and transport of particulate organic carbon in the Amazon River and estuary. *Estuar. Coast. Shelf Sci.* **26**, 1-14, doi: 10.1016/0272-7714(88)90008-x (1988).
- Cao, F. *et al.* Optical characterization of dissolved organic matter in the Amazon River plume and the Adjacent Ocean: Examining the relative role of mixing, photochemistry, and microbial alterations. *Marine Chem.* **186**, 178-188, doi: 10.1016/j.marchem.2016.09.007 (2016).
- Carlson, C. A. & Hansell, D. A. DOM sources, sinks, reactivity, and budgets. *Biogeochemistry of marine dissolved organic matter.* **3**, 65-126 (2015).
- Castello, L. *et al.* The vulnerability of Amazon freshwater ecosystems. *Conserv. Lett.* **6**, 217-229, doi: 10.1111/conl.12008 (2013).
- Chen, M. & Hur, J. Pre-treatments, characteristics, and biogeochemical dynamics of dissolved organic matter in sediments: A review. *Wat. Res.* **79**, 10-25, doi: 10.1016/j.watres.2015.04.018 (2015).
- Chen, W. *et al.* Characterizing properties and environmental behaviors of dissolved organic matter using two-dimensional correlation spectroscopic analysis. *Environ. Sci. Tech.* **53**, 4683-4694, doi: 10.1021/acs.est.9b01103 (2019).
- Ciuraru, R. *et al.* Photosensitized production of functionalized and unsaturated organic compounds at the air-sea interface. *Sci. Rep.* **5**, 12741, doi: 10.1038/srep12741 (2015).

- Clarke, N. *et al.* The formation and fate of chlorinated organic substances in temperate and boreal forest soils. *Environ. Sci. Pollut. Res.* **16**, 127-143, doi: 10.1007/s11356-008-0090-4 (2009).
- Cole, J. J. *et al.* Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10**, 172-185, doi: 10.1007/s10021-006-9013-8 (2007).
- Colicchio, E. *et al.* Biogeochemistry of Carbon in the Amazonian Floodplains over a 2000-km Reach: Insights from a Process-Based Model. *Earth Interact.* **15**, 1-29, doi: 10.1175/2010ei338.1 (2011).
- Confalonieri, U. E. *et al.* Environmental change and the dynamics of parasitic diseases in the Amazon. *Acta Trop.* **129**, 33-41, doi: 10.1016/j.actatropica.2013.09.013 (2014).
- Conrad, R. Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments. *FEMS Microbiol. Ecol.* **28**, 193-202 (1999).
- Conrad, R. *et al.* Acetate turnover and methanogenic pathways in Amazonian lake sediments. *Biogeosci.* **17**, 1063-1069, doi: 10.5194/bg-17-1063-2020 (2020).
- Constantine, J. A. *et al.* Sediment supply as a driver of river meandering and floodplain evolution in the Amazon Basin. *Nat. Geosci.* **7**, 899-903, doi: 10.1038/ngeo2282 (2014).
- Cooper, K. J. *et al.* Dissolved organic carbon transformations and microbial community response to variations in recharge waters in a shallow carbonate aquifer. *Biogeochem.* **129**, 215-234, doi: 10.1007/s10533-016-0226-4 (2016).
- Coppola, A. I. *et al.* Marked isotopic variability within and between the Amazon River and marine dissolved black carbon pools. *Nat. Commun.* **10**, 4018, doi: 10.1038/s41467-019-11543-9 (2019).
- Correa, S. B. *et al.* A comparison of flooded forest and floating meadow fish assemblages in an upper Amazon floodplain. *J. Fish Biol.* **72**, 629-644, doi: 10.1111/j.1095-8649.2007.01752.x (2008).
- Correa, S.B. *et al.* A comparison of flooded forest and floating meadow fish assemblages in an upper Amazon floodplain. *J. Fish Biol.* **72**, 629-644, doi: 10.1111/j.1095-8649.2007.01752.x (2008).
- Costa, M. P. F. *et al.* Spatial and temporal variability of light attenuation in large rivers of the Amazon. *Hydrobiol.* **702**, 171-190, doi: 10.1007/s10750-012-1319-2 (2012).
- Crespo-Lopez, M. E. *et al.* Mercury: What can we learn from the Amazon? *Environ. Int.* **146**, 106223, doi: 10.1016/j.envint.2020.106223 (2021).
- D'Andrilli, J. *et al.* DOM composition alters ecosystem function during microbial processing of isolated sources. *Biogeochem.* **142**, 281-298, doi: 10.1007/s10533-018-00534-5 (2019).
- Daggett, C. T. *et al.* Effects of increased concentrations of inorganic nitrogen and dissolved organic matter on phytoplankton in boreal lakes with differing nutrient limitation patterns. *Aquatic Sci.* **77**, 511-521, doi: 10.1007/s00027-015-0396-5 (2015).

- Dalmagro, H. J. *et al.* Spatial patterns of DOC concentration and DOM optical properties in a Brazilian tropical river-wetland system. *J. Geophys. Res. Biogeosci.* **122**, 1883-1902, doi: 10.1002/2017jg003797 (2017).
- Dalmagro, H. J. *et al.* Carbon biogeochemistry of a flooded Pantanal forest over three annual flood cycles. *Biogeochem.* **139**, 1-18, doi: 10.1007/s10533-018-0450-1 (2018).
- Dalmagro, H. J. *et al.* Streams with Riparian Forest buffers versus impoundments differ in discharge and DOM characteristics for pasture catchments in Southern Amazonia. *Water* **11**, 390, doi: 10.3390/w11020390 (2019).
- Dashtban, M. *et al.* Fungal biodegradation and enzymatic modification of lignin. *Int. J. Biochem. Mol. Biol.* **1**, 36-50, PMID: 21968746 (2010).
- Dauwe, B. & Middelburg, J. J. Amino acids and hexosamines as indicators of organic matter degradation state in North Sea sediments. *Limnol. Oceanogr.* **43**, 782-798, doi: 10.4319/lo.1998.43.5.0782 (1998).
- Davidson, E. A. *et al.* The Amazon basin in transition. *Nature* **481**, 321-328, doi: 10.1038/nature10717 (2012).
- D. C. Smith, F.A. A simple, economical method for measuring bacterial protein synthesis rates in seawater using <sup>3</sup>H-leucine. *Mar. Microb. food webs* **6**, 107-111 (1992).
- de Melo, M. G. *et al.* Metals, n-Alkanes, Hopanes, and Polycyclic Aromatic Hydrocarbon in Sediments from Three Amazonian Streams Crossing Manaus (Brazil). *Chem.* **2**, 274-292, doi: 10.3390/chemistry2020018 (2020).
- De Troyer, I. *et al.* Tracing the source and fate of dissolved organic matter in soil after incorporation of a <sup>13</sup>C labelled residue: A batch incubation study. *Soil Biol. Biochem.* **43**, 513-519, doi: 10.1016/j.soilbio.2010.11.016 (2011).
- Deng, Q. *et al.* Rain-induced changes in soil CO<sub>2</sub> flux and microbial community composition in a tropical forest of China. *Sci. Rep.* **7**, 5539, doi: 10.1038/s41598-017-06345-2 (2017).
- Deron, A. *et al.* NMR studies of the bishydroxy bisphosphonate synthesis from o-phthalic aldehyde and diethyl phosphite. *Heteroat. Chem.* **13**, 157-164, doi: 10.1002/hc.10014 (2002).
- Devol, A. H. & Hedges, J. I. Organic matter and nutrients in the mainstem Amazon River. *The biogeochemistry of the Amazon basin* **15**, 275-306, doi: 10.1093/oso/9780195114317.003.0018 (2001).
- Devol, A. H. *et al.* Seasonal variation in chemical distributions in the Amazon (Solimões) River: A multiyear time series. *Glob. Biogeochem. Cyc.* **9**, 307-328, doi: 10.1029/95gb01145 (1995).



- Dittmar, T. *et al.* A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr.* **6**, 230-235, doi: 10.4319/lom.2008.6.230 (2008).
- do Nascimento, D. R. *et al.* Provenance of sands from the confluence of the Amazon and Madeira rivers based on detrital heavy minerals and luminescence of quartz and feldspar. *Sediment. Geol.* **316**, 1-12, doi: 10.1016/j.sedgeo.2014.11.002 (2015).
- Dong, Y. *et al.* Photochemical origin of reactive radicals and halogenated organic substances in natural waters: A review. *J. Hazard. Mater.* **401**, 123884, doi: 10.1016/j.jhazmat.2020.123884 (2021).
- Du, Y. *et al.* Photochemical reactivities of dissolved organic matter (DOM) in a sub-alpine lake revealed by EEM-PARAFAC: An insight into the fate of allochthonous DOM in alpine lakes affected by climate change. *Sci. Total Environ.* **568**, 216-225, doi: 10.1026/j.scitotenv.2016.06.036 (2016).
- Duan, S. & Bianchi, T. S. Particulate and dissolved amino acids in the lower Mississippi and Pearl Rivers (USA). *Mar. Chem.* **107**, 214-229, doi: 10.1016/j.marchem.2007.07.003 (2007).
- Duarte, R.M. *et al.* Dissolved organic carbon from the upper Rio Negro protects zebrafish (*Danio rerio*) against ionoregulatory disturbances caused by low pH exposure. *Sci. Rep.* **6**, 20377, doi: 10.1038/srep20377 (2016).
- Eadie, B. *et al.* Some observations on the stable carbon isotope composition of dissolved and particulate organic carbon in the marine environment. *Geochimica et Cosmochimica acta* **42**, 1265-1269, doi: 10.1016/0016-7037(78)90120-5 (1978).
- Einsiedl, F. *et al.* Rapid biotic molecular transformation of fulvic acids in a karst aquifer. *Geochim. et Cosmochim. Acta* **71**, 5474-5482, doi: 10.1016/j.gca.2007.09.024 (2007).
- Ellis, E. E. *et al.* Factors controlling water-column respiration in rivers of the central and southwestern Amazon Basin. *Limnol. Oceanogr.* **57**, 527-540, doi: 10.4319/lo.2012.57.2.0527 (2012).
- Engbrodt, R. & Kattner, G. On the biogeochemistry of dissolved carbohydrates in the Greenland Sea (Arctic). *Org. Geochem.* **36**, 937-948, doi: 10.1016/j.orggeochem.2004.12.007 (2005).
- Ertel, J. R. *et al.* Dissolved humic substances of the Amazon River system1. *Limnol. Oceanogr.* **31**, 739-754, doi: 10.4319/lo.1986.31.4.0739 (1986).
- Espinoza, J.-C. *et al.* The new historical flood of 2021 in the Amazon River compared to major floods of the 21<sup>st</sup> century: Atmospheric features in the context of the intensification of floods. *Weather Clim. Extremes* **35**, 100406, doi: 10.1016/j.wace.2021.100406 (2022).
- Fan, J.-X. *et al.* Effect of iron oxide reductive dissolution on the transformation and immobilization of arsenic in soils: New insights from X-ray photoelectron and X-ray absorption spectroscopy. *J. Hazard. Mater.* **279**, 212-219, doi: 10.1016/j.jhazmat.2014.06.079 (2014).
- Farjalla, V. F. *et al.* Nutrient limitation of bacterial production in clear water Amazonian ecosystems. *Hydrobiologia* **489**, 197-205, doi: 10.1023/A:1023288922394 (2002).

- Farjalla, V.F. Are the mixing zones between aquatic ecosystems hot spots of bacterial production in the Amazon River system? *Hydrobiologia* **728**, 153-165, doi: 10.1007/s10750-014-1814-8 (2014).
- Fassoni-Andrade, A. C. *et al.* Amazon hydrology from space: scientific advances and future challenges. *Rev. Geophys.* **59**, e2020RG000728, doi: 10.1029/2020RG000728 (2021).
- Feng, X. *et al.* Source to sink: Evolution of lignin composition in the Madre de Dios River system with connection to the Amazon basin and offshore. *J. Geophys. Res. Biogeo.* **121**, 1316-1338, doi: 10.1002/2016JG003323 (2016).
- Fisher, S. G. & Likens, G. E. Energy flow in Bear Brook, New Hampshire: an integrative approach to stream ecosystem metabolism. *Ecol. Monogr.* **43**, 421-439, doi: 10.2307/1942301 (1973).
- Flerus, R. *et al.* A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosci.* **9**, 1935-1955, doi: 10.5194/bg-9-1935-2012 (2012).
- Franchini, J. *et al.* Organic composition of green manure during growth and its effect on cation mobilization in an acid Oxisol. *Commun. Soil Sci. Plant Anal.* **34**, 2045-2058, doi: 10.1081/CSS-120023237 (2003).
- Fritsch, E. *et al.* Organic complexation and translocation of ferric iron in podzols of the Negro River watershed. Separation of secondary Fe species from Al species. *Geochimica et Cosmochimica Acta* **73**, 1813–1825, doi: 10.1016/j.gca.2009.01.008 (2009).
- Fröberg, M. *et al.* Contributions of Oi, Oe and Oa horizons to dissolved organic matter in forest floor leachates. *Geoderma* **113**, 311-322, doi: 10.1016/S0016-7061(02)00367-1 (2003).
- Fujii, K. *et al.* Fluxes of dissolved organic carbon in two tropical forest ecosystems of East Kalimantan, Indonesia. *Geoderma* **152**, 127-136, doi: 10.1016/j.geoderma.2009.05.028 (2009).
- Gagne-Maynard, W. C. *et al.* Evaluation of Primary Production in the Lower Amazon River Based on a Dissolved Oxygen Stable Isotopic Mass Balance. *Front. Mar. Sci.* **4**, 26, doi: 10.3389/fmars.2017.00026 (2017).
- Gibbs, R. J. The Geochemistry of the Amazon River System: Part I. The Factors that Control the Salinity and the Composition and Concentration of the Suspended Solids. *Geol. Soc. Am. Bull.* **78**, 1203, doi: 10.1130/0016-7606(1967)78[1203:tgotar]2.0.co;2 (1967).
- Gigliotti, G., Kaiser, K., Guggenberger, G. and Haumaier, L. 2002. Differences in the chemical composition of dissolved organic matter from waste material of different sources. *Biology and Fertility of Soils* **36**, 321-329.
- Gledhill, M. *et al.* Trace metal stoichiometry of dissolved organic matter in the Amazon plume. *Sci. Adv.* **8**, eabm2249, doi: 10.1126/sciadv.abm2249 (2022).

- Glissman, K. *et al.* Methanogenic pathway and archaeal community structure in the sediment of eutrophic Lake Dagow: effect of temperature. *Microb. Ecol.* **48**, 389-399, doi: 10.1007/s00248-003-2027-2 (2004).
- Goldberg, S. J. *et al.* Refractory dissolved organic nitrogen accumulation in high-elevation lakes. *Nat. Commun.* **6**, 6347, doi: 10.1038/ncomms7347 (2015).
- Goldberg, S. J. *et al.* Systematic removal of neutral sugars within dissolved organic matter across ocean basins. *Geophys. Res. Lett.* **38**, 17, doi: 10.1029/2011GL048620 (2011).
- Gonsior, M. *et al.* Photochemical production of polyols arising from significant photo-transformation of dissolved organic matter in the oligotrophic surface ocean. *Mar. Chem.* **163**, 10-18, doi: 10.1016/j.marchem.2014.04.002 (2014).
- Gonsior, M. *et al.* Photochemically induced changes in dissolved organic matter identified by ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry. *Environ. Sci. Technol.* **43**, 698-703, doi: 10.1021/es8022804 (2009).
- Gonsior, M. *et al.* The chemodiversity of algal dissolved organic matter from lysed *Microcystis aeruginosa* cells and its ability to form disinfection by-products during chlorination. *Wat. Res.* **155**, 300-309, doi: 10.1016/j.watres.2019.02.030 (2019).
- Gonsior, M. *et al.* Depth-dependent molecular composition and photoreactivity of dissolved organic matter in a Boreal Lake under winter and summer conditions. *Biogeosci.* **10**, 8949-8975, doi: 10.5194/bg-10-8949-2013 (2013).
- Gonsior, M. *et al.* Chemodiversity of dissolved organic matter in the Amazon Basin. *Biogeosci.* **13**, 4279-4290, doi: 10.5194/bg-13-4279-2016 (2016).
- Guillemette, F. *et al.* Selective consumption and metabolic allocation of terrestrial and algal carbon determine allochthony in lake bacteria. *ISME* **10**, 1373-1382, doi: 10.1038/ismej.2015.215 (2016).
- Guillemette, F. *et al.* Differentiating the degradation dynamics of algal and terrestrial carbon within complex natural dissolved organic carbon in temperate lakes. *J. Geophys. Res.: Biogeosci.* **118**, 963-973, doi: 10.1002/jgrg.20077 (2013).
- Guinoiseau, D. *et al.* Zinc and copper behavior at the soil-river interface: new insights by Zn and Cu isotopes in the organic-rich Rio Negro basin. *Geochimica et Cosmochimica Acta* **213**, 178-197, doi: 10.1016/j.gca.2017.06.030 (2017).
- Guo, L. & Sun, M.-Y. *Isotope composition of organic matter in seawater.* **6**, 97-124, doi: 10.1201/9781420073072.ch6 (2009).
- Häder, D. P. *et al.* Effects of UV radiation on aquatic ecosystems and interactions with other environmental factors. *Photochem. Photobiol. Sci.* **14**, 108-126, doi: 10.1039/C4PP90035A (2015).

- Hagedorn, F. *et al.* A  $^{13}\text{C}$  tracer study to identify the origin of dissolved organic carbon in forested mineral soils. *Eur. J. Soil Sci.* **55**, 91-100, doi: 10.1046/j.1365-2389.2003.00578.x (2004).
- Hagerman, A. E. *et al.* Tannins and lignins. *Herbivores: their interactions with secondary plant metabolites.* **1**, 355-388, doi: 10.1016/B978-0-12-597183-6.50015-2 (1991).
- Hamilton, S. K. *et al.* Oxygen depletion and carbon dioxide and methane production in waters of the Pantanal wetland of Brazil. *Biogeochem.* **30**, 115-141, doi: 10.1007/BF00002727 (1995).
- Harir, M. *et al.* Molecular and spectroscopic changes of peat-derived organic matter following photo-exposure: Effects on heteroatom composition of DOM. *Sci. Total Environ.* **838**, 155790, doi: 10.1016/j.scitotenv.2022.155790 (2022).
- Harvey, H. R. *et al.* Impact of DOM composition on bacterial lipids and community structure in estuaries. *Aquat. Microb. Ecol.* **42**, 105-117, doi: 10.3354/ame042105 (2006).
- Hawkes, J. A. *et al.* Molecular alteration of marine dissolved organic matter under experimental hydrothermal conditions. *Geochim. Cosmochim. Ac.* **175**, 68-85, doi: 10.1016/j.gca.2015.11.025 (2016).
- Li, Y, J. A. *et al.* Regional diversity of complex dissolved organic matter across forested hemiboreal headwater streams. *Sci. Rep.* **8**, 1-11, doi: 10.1038/s41598-018-34272-3 (2018).
- Hedges, J. I. *et al.* Organic matter in Bolivian tributaries of the Amazon River: A comparison to the lower mainstream. *Limnol. Oceanogr.* **45**, 1449-1466, doi: 10.4319/lo.2000.45.7.1449 (2000).
- Hedges, J. I. *et al.* Organic carbon-14 in the Amazon river system. *Science* **231**, 1129-1131, doi: 10.1126/science.231.4742.1129 (1986).
- Hedges, J. I. *et al.* A comparison of dissolved humic substances from seawater with Amazon River counterparts by  $^{13}\text{C}$ -NMR spectrometry. *Geochim. Cosmochim. Ac.* **56**, 1753-1757, doi: 10.1016/0016-7037(92)90241-a (1992).
- Hedges, J. I. *et al.* Compositions and fluxes of particulate organic material in the Amazon River1. *Limnol. Oceanogr.* **31**, 717-738, doi: 10.4319/lo.1986.31.4.0717 (1986).
- Hedges, J. I. *et al.* Origins and processing of organic matter in the Amazon River as indicated by carbohydrates and amino acids. *Limnol. Oceanogr.* **39**, 743-761, doi: 10.4319/lo.1994.39.4.0743 (1994).
- Hedges, J. I. *et al.* What happens to terrestrial organic matter in the ocean? *Org. Geochem.* **27**, 195-212, doi: 10.1016/s0146-6380(97)00066-1 (1997).
- Helms, J. R. *et al.* Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.* **53**, 955-969, doi: 10.4319/lo.2008.53.3.0955 (2008).

- Hernes, P. J. Photochemical and microbial degradation of dissolved lignin phenols: Implications for the fate of terrigenous dissolved organic matter in marine environments. *J. Geophys. Res.* **108**, C9, doi: 10.1029/2002jc001421 (2003).
- Hertkorn, N. *et al.* Comparative analysis of partial structures of a peat humic and fulvic acid using one- and two-dimensional nuclear magnetic resonance spectroscopy. *J. Environ. Qual.* **31**, 375-387, doi: 10.2134/jeq2002.3750 (2002).
- Hertkorn, N. *et al.* A potentiometric and  $^{113}\text{Cd}$  NMR study of cadmium complexation by natural organic matter at two different magnetic field strengths. *Anal. Chem.* **76**, 6327-6341, doi: 10.1021/ac0400212 (2004).
- Hertkorn, N. *et al.* Nontarget analysis of Murchison soluble organic matter by high-field NMR spectroscopy and FTICR mass spectrometry. *Magn. Reson. Chem.* **53**, 754-768, doi: 10.1002/mrc.4249 (2015).
- Hertkorn, N. *et al.* Natural organic matter and the event horizon of mass spectrometry. *Anal. Chem.* **80**, 8908-8919, doi: 10.1021/ac800464g (2008).
- Hertkorn, N. *et al.* High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Anal. Bioanal. Chem.* **389**, 1311-1327, doi: 10.1007/s00216-007-1577-4 (2007).
- Hertkorn, N. *et al.* Molecular characterization of dissolved organic matter from subtropical wetlands: a comparative study through the analysis of optical properties, NMR and FTICR/MS. *Biogeosci.* **13**, 2257-2277, doi: 10.5194/bg-13-2257-2016 (2016).
- Hertkorn, N. *et al.* High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosci.* **10**, 1583-1624, doi: 10.5194/bg-10-1583-2013 (2013).
- Hertkorn, N. *et al.* Comparative Analysis of Partial Structures of a Peat Humic and Fulvic Acid Using One- and Two-Dimensional Nuclear Magnetic Resonance Spectroscopy. *J. Environ. Qual.* **31**, 375-387, doi: 10.2134/jeq2002.0375 (2002).
- Hertkorn, N. *et al.* Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Ac.* **70**, 2990-3010, doi: 10.1016/j.gca.2006.03.021 (2006).
- Herzprung, P. *et al.* Molecular formula assignment for dissolved organic matter (DOM) using high-field FT-ICR-MS: chemical perspective and validation of sulphur-rich organic components (CHOS) in pit lake samples. *Anal. Bioanal. Chem.* **408**, 2461-2469, doi: 10.1007/s00216-016-9341-2 (2016).
- Herzprung, P. *et al.* High field FT-ICR mass spectrometry data sets enlighten qualitative DOM alteration in lake sediment porewater profiles. *Org. Geochem.* **108**, 51-60, doi: 10.1016/j.orggeochem.2017.03.010 (2017).

- Herzsprung, P. *et al.* Differences in DOM of rewetted and natural peatlands - Results from high-field FT-ICR-MS and bulk optical parameters. *Sci. Total Environ.* **586**, 770-781, doi: 10.1016/j.scitotenv.2017.02.054 (2017).
- Hirose, K. Metal–organic matter interaction: ecological roles of ligands in oceanic DOM. *Appl. Geochem.* **22**, 1636-1645, doi: 10.1016/j.apgeochem.2007.03.042 (2007).
- Howarth, R. W. *et al.* Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. *Biogeochem.* **35**, 75-139, doi: 10.1007/bf02179825 (1996).
- Huang, X. *et al.* Transformation of algal-dissolved organic matter via sunlight-induced photochemical and microbial processes: interactions between two processes. *Environ. Sci. Poll. Res.*, 1-13, doi: 10.1007/s11356-023-26024-2 (2023).
- Huttunen, J. T. *et al.* Greenhouse gases in non-oxygenated and artificially oxygenated eutrophied lakes during winter stratification. *J. Environ. Qual.* **30**, 387-394, doi: 10.2134/jeq2001.302387x (2001).
- Ide, J. i. *et al.* Spatial variations in the molecular diversity of dissolved organic matter in water moving through a boreal forest in eastern Finland. *Sci. Rep.* **7**, 1-12, doi: 10.1038/srep42102 (2017).
- Irion, G. Minerals in rivers. In: Degens, E., Kempe, S., Richey, J. (Eds.), SCOPE 42 – Biogeochemistry of Major World Rivers. John Wiley and Sons, Chichester, pp. 265–282 (1991).
- Jaffé, R. *et al.* Spatial and temporal variations in DOM composition in ecosystems: The importance of long-term monitoring of optical properties. *J. Geophys. Res.: Biogeosci.* **113**, G4, doi: 10.1029/2008JG000683 (2008).
- Jandl, R. & Sletten, R. S. Mineralization of forest soil carbon: interactions with metals. *J. Plant Nutr. Soil Sci.* **162**, 623-629, doi: 10.1002/(SICI)1522-2624(199912)162:6<623::AID-JPLN623>3.0.CO;2-8 (1999).
- Janusz, G. *et al.* Lignin degradation: microorganisms, enzymes involved, genomes analysis and evolution. *FEMS Microbiol. Rev.* **41**, 941-962, doi: 10.1093/femsre/fux049 (2017).
- Janzen, D. H. Tropical blackwater rivers, animals, and mast fruiting by the Diptero­carpaceae. *Biotropica*, **6**, 69-103, doi: 10.2307/2989823 (1974).
- Johannsson, O. E. *et al.* Photo-oxidation processes, properties of DOC, reactive oxygen species (ROS), and their potential impacts on native biota and carbon cycling in the Rio Negro (Amazonia, Brazil). *Hydrobiol.* **789**, 7-29, doi: 10.1007/s10750-016-2687-9 (2017).
- Jones, D. L. Organic acids in the rhizosphere—a critical review. *Plant Soil* **205**, 25-44, doi: 10.1023/A:1004356007312 (1998).
- Junk, W. J. Appraisal of the scientific work of Harald Sioli. *Amazoniana* **16**, 285-297, doi: 10.1896/1413-4705.12.3.153a (2001).

- Junk, W. J. *et al.* Biodiversity and its conservation in the Pantanal of Mato Grosso, Brazil. *Aquat. Sci.* **68**, 278-309, doi: 10.1007/s00027-006-0851-4 (2006).
- Junk, W. J. *et al.* A Classification of Major Naturally-Occurring Amazonian Lowland Wetlands. *Wetlands* **31**, 623-640, doi: 10.1007/s13157-011-0190-7 (2011).
- Kaiser, K. & Kalbitz, K. Cycling downwards—dissolved organic matter in soils. *Soil Biol. Biochem.* **52**, 29-32, doi: 10.1016/j.soilbio.2012.04.002 (2012).
- Kaiser, K. & Guggenberger, G. Mineral surfaces and soil organic matter. *Eur. J. Soil Sci.* **54**, 219-236, doi: 10.1046/j.1365-2389.2003.00544.x (2003).
- Kaiser, K. & Benner, R. Organic matter transformations in the upper mesopelagic zone of the North Pacific: Chemical composition and linkages to microbial community structure. *J. Geophys. Res. Oceans* **117**, C1, doi: 10.1029/2011JC007141 (2012).
- Kamjunke, N. *et al.* A new approach for evaluating transformations of dissolved organic matter (DOM) via high-resolution mass spectrometry and relating it to bacterial activity. *Wat. Res.* **123**, 513-523, doi: 10.1016/j.watres.2017.07.008 (2017).
- Karlsson, T., Persson, P. and Skyllberg, U. 2005. Extended X-ray absorption fine structure spectroscopy evidence for the complexation of cadmium by reduced sulfur groups in natural organic matter. *Environmental science & technology* 39(9), 3048-3055.
- Kawasaki, N. & Benner, R. Bacterial release of dissolved organic matter during cell growth and decline: molecular origin and composition. *Limnol. Oceanogr.* **51**, 2170-2180, doi: 10.4319/lo.2006.51.5.2170 (2006).
- Keil, R. G. *et al.* Loss of organic matter from riverine particles in deltas. *Geochim. Cosmochim. Ac.* **61**, 1507-1511, doi: 10.1016/s0016-7037(97)00044-6 (1997).
- Kellerman, A. M. *et al.* Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. *Nat. Commun.* **5**, 3804, doi: 10.1038/ncomms4804 (2014).
- Kellerman, A. M. *et al.* Persistence of dissolved organic matter in lakes related to its molecular characteristics. *Nat. Geosci.* **8**, 454-457, doi: 10.1038/ngeo2440 (2015).
- Kent, A. D. *et al.* Annual patterns in bacterioplankton community variability in a humic lake. *Microb. Ecol.* **48**, 550-560, doi: 10.1007/s00248-004-0244-y (2004).
- Killops, V. J. & Killops, S. D. *Introduction to Organic Geochemistry*. 2<sup>nd</sup> ed, 406 (2013).
- Kim, S. *et al.* Biodegradable dissolved organic matter in a temperate and a tropical stream determined from ultra-high resolution mass spectrometry. *Limnol. Oceanogr.* **51**, 1054-1063, doi: 10.4319/lo.2006.51.2.1054 (2006).

- Kleber, M. *et al.* A conceptual model of organo-mineral interactions in soils: self-assembly of organic molecular fragments into zonal structures on mineral surfaces. *Biogeochemistry* **85**, 9–24, doi: 10.1007/s10533-007-9103-5 (2007).
- Knorr, K.-H. DOC-dynamics in a small headwater catchment as driven by redox fluctuations and hydrological flow paths—are DOC exports mediated by iron reduction/oxidation cycles? *Biogeosci.* **10**, 891-904, doi: 10.5194/bg-10-891-2013 (2013).
- Koch, B. P. *et al.* Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Anal. Chem.* **79**, 1758-1763, doi: 10.1021/ac061949s (2007).
- Koch, B. P. & Dittmar, T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rap. Commun. Mass Spectrom.* **30**, 250-250, doi: 10.1002/rcm.2386 (2016).
- Koch, B. P. *et al.* Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochim. Cosmochim. Ac.* **69**, 3299-3308, doi: 10.1016/j.gca.2005.02.027 (2005).
- Kögel-Knabner, I. & Amelung, W. Soil organic matter in major pedogenic soil groups. *Geoderma* **384**, 114785, doi: 10.1016/j.geoderma.2020.114785 (2021).
- Kördel, W. *et al.* The importance of natural organic material for environmental processes in waters and soils (Technical Report). *Pure Appl. Chem.* **69**, 1571-1600, 10.1351/pac199769071571 (1997).
- Kraus, T. E. *et al.* Linking chemical reactivity and protein precipitation to structural characteristics of foliar tannins. *J. Chem. Ecol.* **29**, 703-730, doi: 10.1023/a:1022876804925 (2003).
- Kritzberg, E. S. *et al.* Influence of dissolved organic matter source on lake bacterioplankton structure and function—implications for seasonal dynamics of community composition. *FEMS Microbiol. Ecol.* **56**, 406-417, doi: 10.1111/j.1574-6941.2006.00084.x (2006).
- Ksionzek, K. B. *et al.* Dissolved organic sulfur in the ocean: Biogeochemistry of a petagram inventory. *Science* **354**, 456-459, doi: 10.1126/science.aaf7796 (2016).
- Kujawinski, E. B. *et al.* Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochim. Cosmochim. Ac.* **73**, 4384-4399, doi: 10.1016/j.gca.2009.04.033 (2009).
- Kujawinski, E. B. *et al.* The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter. *Org. Geochem.* **33**, 171-180, doi: 10.1016/s0146-6380(01)00149-8 (2002).
- Kujawinski, E. B. *et al.* Probing molecular-level transformations of dissolved organic matter: insights on photochemical degradation and protozoan modification of DOM from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mar. Chem.* **92**, 23-37, doi: 10.1016/j.marchem.2004.06.038 (2004).



- Kujawinski, E. B. *et al.* The impact of microbial metabolism on marine dissolved organic matter. *Annual Review of Marine Science* **3**, 567-599, doi: 10.1146/annurev-marine-120308-081003 (2011).
- Kungolos, A. *et al.* Bioavailability and toxicity of heavy metals in the presence of natural organic matter. *J. Environ. Sci. Health A* **41**, 1509-1517, doi: 10.1080/10934520600754706 (2006).
- Kuzyakov, Y. & Blagodatskaya, E. Microbial hotspots and hot moments in soil: concept & review. *Soil Bio. Biochem.* **83**, 184-199, doi: 10.1016/j.soilbio.2015.01.025 (2015).
- Lambert, T. *et al.* Along-stream transport and transformation of dissolved organic matter in a large tropical river. *Biogeosci.* **13**, 2727-2741, doi: 10.5194/bg-13-2727-2016 (2016).
- Landa, M. *et al.* Changes in bacterial diversity in response to dissolved organic matter supply in a continuous culture experiment. *Aquat. Microb. Ecol.* **69**, 157-168, doi: 10.3354/ame01632 (2013).
- Lanzalunga, O. & Bietti, M. Photo- and radiation chemical induced degradation of lignin model compounds. *J. Photochem. Photobiol. B: Biology* **56**, 85-108, doi: 10.1016/s1011-1344(00)00054-3 (2000).
- Latrubesse, E. M. *et al.* Damming the rivers of the Amazon basin. *Nature* **546**, 363-369, doi: 10.1038/nature22333 (2017).
- Lechtenfeld, O.J. *et al.* Marine sequestration of carbon in bacterial metabolites. *Nat. Commun.* **6**, 6711, doi: 10.1038/ncomms7711 (2015).
- Lee, M. H. *et al.* New insight into the applicability of spectroscopic indices for dissolved organic matter (DOM) source discrimination in aquatic systems affected by biogeochemical processes. *Wat. Res.* **147**, 164-176, doi: 10.1016/j.watres.2018.09.048 (2018).
- Leenheer, J. A. Origin and nature of humic substances in the waters of the Amazon River Basin. *Acta Amaz.* **10**, 513-526, doi: 10.1590/1809-43921980103513 (1980).
- Leenheer, J. A. & Croué, J.-P. Peer Reviewed: Characterizing Aquatic Dissolved Organic Matter. *Environ. Sci. Technol.* **37**, 18A-26A, doi: 10.1021/es032333c (2003).
- Lees, A. C. *et al.* Hydropower and the future of Amazonian biodiversity. *Biodivers. Cons.* **25**, 451-466, doi: 10.1007/s10531-016-1072-3 (2016).
- Leri, A. C. *et al.* A marine sink for chlorine in natural organic matter. *Nat. Geosci.* **8**, 620-624, doi: 10.1038/ngeo2481 (2015).
- Li, S. *et al.* Properties of sediment dissolved organic matter respond to eutrophication and interact with bacterial communities in a plateau lake. *Environ. Pollut.* **301**, 118996, doi: 10.1016/j.envpol.2022.118996 (2022).
- Li, Y. *et al.* Proposed Guidelines for Solid Phase Extraction of Suwannee River Dissolved Organic Matter. *Anal. Chem.* **88**, 6680-6688, doi: 10.1021/acs.analchem.5b04501 (2016).

- Liikanen, A. *et al.* Effects of temperature and oxygen availability on greenhouse gas and nutrient dynamics in sediment of a eutrophic mid-boreal lake. *Biogeochem.* **59**, 269-286, doi: 10.1023/A:1016015526712 (2002).
- Lipczynska-Kochany, E. Humic substances, their microbial interactions and effects on biological transformations of organic pollutants in water and soil: A review. *Chemosphere* **202**, 420-437, doi: 10.1016/j.chemosphere.2018.03.104 (2018).
- Liikanen, A. *et al.* Effects of temperature and oxygen availability on greenhouse gas and nutrient dynamics in sediment of a eutrophic mid-boreal lake. *Biogeochem.* **59**, 269-286, doi: 10.1023/A:1016015526712 (2002).
- Liu, M. Diverse molecular compositions of dissolved organic matter derived from different composts using ESI FT-ICR MS. *J. Environ. Sci.* **99**, 80-89, doi: 10.1016/j.jes.2020.06.011 (2021).
- Luek, J.L. *et al.* Sulfate Reduction in Sediments Produces High Levels of Chromophoric Dissolved Organic Matter. *Sci. Rep.* **7**, 8829, doi: 10.1038/s41598-017-09223-z (2017).
- Lund, V. Ultraviolet irradiated water containing humic substances inhibits bacterial metabolism. *Wat. Res.* **28**, 1111-1116, doi: 10.1016/0043-1354(94)90197-x (1994).
- Lundeen, R. A. *et al.* Environmental Photochemistry of Amino Acids, Peptides and Proteins. *Chimia (Aarau)* **68**, 812-817, doi: 10.2533/chimia.2014.812 (2014).
- Lynch, L. M. *et al.* River channel connectivity shifts metabolite composition and dissolved organic matter chemistry. *Nat. Commun.* **10**, 459, doi: 10.1038/s41467-019-08406-8 (2019).
- Macdonald, M. J. & Minor, E. C. Photochemical degradation of dissolved organic matter from streams in the western Lake Superior watershed. *Aquat. Sci.* **75**, 509-522, doi: 10.1007/s00027-013-0296-5 (2013).
- Marín-Spiotta, E. *et al.* Paradigm shifts in soil organic matter research affect interpretations of aquatic carbon cycling: transcending disciplinary and ecosystem boundaries. *Biogeochem.* **117**, 279-297, doi: 10.1007/s10533-013-9949-7 (2014).
- Marschner, B. & Kalbitz, K. Controls of bioavailability and biodegradability of dissolved organic matter in soils. *Geoderma* **113**, 211-235, doi: 10.1016/S0016-7061(02)00362-2 (2003).
- Martinez, J.-M. *et al.* The optical properties of river and floodplain waters in the Amazon River Basin: Implications for satellite-based measurements of suspended particulate matter. *J. Geophys. Res. Earth Surf.* **120**, 1274-1287, doi: 10.1002/2014jfg003404 (2015).
- Marwick, T. R. *et al.* The age of river-transported carbon: A global perspective. *Glob. Biogeochem. Cyc.* **29**, 122-137, doi: 10.1002/2014GB004911 (2015).
- Mayorga, E. *et al.* Young organic matter as a source of carbon dioxide outgassing from Amazonian rivers. *Nature* **436**, 538-541, doi: 10.1038/nature03880 (2005).

- Mayorga, E. & Aufdenkampe, A. Processing of bioactive elements in the Amazon River system. *The Ecohydrology of South American Rivers and Wetlands* **6**, 1-24 (2002).
- McClain, M. E. & Naiman, R. J. Andean Influences on the Biogeochemistry and Ecology of the Amazon River. *BioScience* **58**, 325-338, doi: 10.1641/b580408 (2008).
- McClain, M. E. *et al.* Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosys.* **6**, 301-312, doi: 10.1007/s10021-003-0161-9 (2003).
- McClain, M. E. *et al.* Terrestrial inputs to Amazon streams and internal biogeochemical processing. In: McClain, M.E., Victoria, R., Richey, J.E. (Eds.), *The Biogeochemistry of the Amazon Basin*. Oxford University Press, New York, pp. 185–208 (2001).
- McDonald, S. *et al.* Analytical chemistry of freshwater humic substances. *Anal. Chim. Acta* **527**, 105-124, doi: 10.1016/j.aca.2004.10.011 (2004).
- Medeiros, P. M. *et al.* Fate of the Amazon River dissolved organic matter in the tropical Atlantic Ocean. *Glob. Biogeochem. Cyc.* **29**, 677-690, doi: 10.1002/2015gb005115 (2015).
- Melack, J. M. *et al.* Regionalization of methane emissions in the Amazon Basin with microwave remote sensing. *Glob. Chang. Biol.* **10**, 530-544, doi: 10.1111/j.1365-2486.2004.00763.x (2004).
- Melo, M. L. *et al.* Linking dissolved organic matter composition and bacterioplankton communities in an Amazon floodplain system. *Limnol. Oceanogr.* **65**, 63-76, doi: 10.1002/lno.11250 (2019).
- Merschel, G. *et al.* Contrasting impact of organic and inorganic nanoparticles and colloids on the behavior of particle-reactive elements in tropical estuaries: an experimental study. *Geochim. Cosmochim. Ac.* **197**, 1-13, doi: 10.1016/j.gca.2016.09.041 (2017).
- Metcalf, D. B. *et al.* Factors controlling spatio-temporal variation in carbon dioxide efflux from surface litter, roots, and soil organic matter at four rain forest sites in the eastern Amazon. *J. Geophys. Res.: Biogeosci.* **112**, G4, doi: 10.1029/2007jg000443 (2007).
- Meybeck, M. Carbon, nitrogen, and phosphorus transport by world rivers. *Am. J. Sci.* **282**, 401-450, doi: 10.2475/ajs.282.4.401 (1982).
- Meyers, P. A. & Lallier-Vergès, E. Lacustrine sedimentary organic matter records of Late Quaternary paleoclimates. *J. PaleoLimnol.* **21**, 345-372, doi: 10.1023/A:1008073732192 (1999).
- Meyers-Schulte, K. J. & Hedges, J. I. Molecular evidence for a terrestrial component of organic matter dissolved in ocean water. *Nature* **321**, 61-63, doi: 10.1038/321061a0 (1986).
- Miller, M. P. & McKnight, D. M. Comparison of seasonal changes in fluorescent dissolved organic matter among aquatic lake and stream sites in the Green Lakes Valley. *J. Geophys. Res. Biogeosci.* **115**, G1, doi: 10.1029/2009JG000985 (2010).

- Minor, E. C. *et al.* Structural characterization of dissolved organic matter: a review of current techniques for isolation and analysis. *Environ. Sc. Proc. Imp.* **16**, 2064-2079, doi: 10.1039/C4EM00062E (2014).
- Mopper, K. *et al.* Advanced instrumental approaches for characterization of marine dissolved organic matter: extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Chem. Rev.* **107**, 419-442, doi: 10.1021/cr050359b (2007).
- Mopper, K. Photochemistry and the cycling of carbon, sulfur, nitrogen and phosphorus. *Biogeochemistry of marine organic matter*, 455-489 (2002).
- Moran, M. A. & Zepp, R. G. Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. *Limnol. Oceanogr.* **42**, doi: 1307-1316 10.4319/lo.1997.42.6.1307 (1997).
- Moran, M. & Zepp, R. UV radiation effects on microbes and microbial processes. *Microb. Ecol. Oceans* **7**, 201-228, D.L.Kirchman (ed.) (2000).
- Moreira-Turcq, P. F. *et al.* Exportation of organic carbon from the Amazon River and its main tributaries. *Hydrol. Proc.* **17**, 1329-1344, doi: 10.1002/hyp.1287 (2003a).
- Moreira-Turcq, P. F. *et al.* Characteristics of organic matter in the mixing zone of the Rio Negro and Rio Solimões of the Amazon River. *Hydrol. Proc.* **17**, 1393-1404, doi: 10.1002/hyp.1291 (2003b).
- Moreira-Turcq, P. *et al.* Seasonal variability in concentration, composition, age, and fluxes of particulate organic carbon exchanged between the floodplain and Amazon River. *Glob. Biogeochem. Cycl.* **27**, 119-130, doi: 10.1002/gbc.20022 (2013).
- Mortillaro, J. M. *et al.* Particulate organic matter distribution along the lower Amazon River: addressing aquatic ecology concepts using fatty acids. *PLoS One* **7**, e46141, doi: 10.1371/journal.pone.0046141 (2012).
- Mortillaro, J. M. *et al.* Fatty acid and stable isotope ( $\delta^{13}\text{C}$ ,  $\delta^{15}\text{N}$ ) signatures of particulate organic matter in the lower Amazon River: Seasonal contrasts and connectivity between floodplain lakes and the mainstem. *Org. Geochem.* **42**, 1159-1168, doi: 10.1016/j.orggeochem.2011.08.011 (2011).
- Mortillaro, J.-M. *et al.* The fate of C4 and C3 macrophyte carbon in central Amazon floodplain waters: Insights from a batch experiment. *Limnologia* **59**, 90-98, doi: 10.1016/j.limno.2016.03.008 (2016).
- Mostofa, K. M. G. *et al.* Photodegradation of fluorescent dissolved organic matter in river waters. *Geochem. J.* **41**, 323-331, doi: 10.2343/geochemj.41.323 (2007).
- Murphy, K. R. *et al.* Distinguishing between terrestrial and autochthonous organic matter sources in marine environments using fluorescence spectroscopy. *Mar. Chem.* **108**, 40-58, doi: 10.1016/j.marchem.2007.10.003 (2008).
- Nebbioso, A. & Piccolo, A. Molecular characterization of dissolved organic matter (DOM): a critical review. *Anal. Bioanal. Chem.* **405**, 109-124, doi: 10.1007/s00216-012-6363-2 (2013).

- Nelson, N. B. & Siegel, D. A. The global distribution and dynamics of chromophoric dissolved organic matter. *Ann. Rev. Mar. Sci.* **5**, 447-476, doi: 10.1146/annurev-marine-120710-100751 (2013).
- Newton, R. J. *et al.* Microbial community dynamics in a humic lake: differential persistence of common freshwater phylotypes. *Environ. Microbiol.* **8**, 956-970, doi: 10.1111/j.1462-2920.2005.00979.x (2006).
- Nielsen, E.S. Dark Fixation of CO<sub>2</sub> and Measurements of Organic Productivity. With Remarks on Chemo-Synthesis. *Physiol. Plant.* **13**, 348-357, doi: 1111/j.1399-3054.1960.tb08037.x (1960).
- Nielsen, E.S. On the Determination of the Activity in <sup>14</sup>C-Ampoules for Measuring Primary Production. *Limnol. Oceanogr.* **10**, R247-R252, doi: 10.4319/lo.1965.10.suppl2.r247 (1965).
- Niu, X.-Z. & Croué, J.-P. Photochemical production of hydroxyl radical from algal organic matter. *Wat. Res.* **161**, 11-16, doi: 10.1016/j.watres.2019.05.089 (2019).
- Núñez-Pons, L. *et al.* UV-protective compounds in marine organisms from the Southern Ocean. *Marine Drugs* **16**, 336, 10.3390/md16090336 (2018).
- Nydal, R. & Lövseth, K. Tracing bomb <sup>14</sup>C in the atmosphere 1962–1980. *J. Geophys. Res. Oceans* **88**, 3621-3642, doi: 10.1029/JC088iC06p03621 (1983).
- Ohno, T. *et al.* Ultrahigh resolution mass spectrometry and indicator species analysis to identify marker components of soil- and plant biomass- derived organic matter fractions. *Environ. Sci. Technol.* **44**, 8594-8600, doi: 10.1021/es101089t (2010).
- Olk, D. C. *et al.* Environmental and Agricultural Relevance of Humic Fractions Extracted by Alkali from Soils and Natural Waters. *J. Environ. Qual.* **48**, 217-232, doi: 10.2134/jeq2019.02.0041 (2019).
- Opsahl, S. and Benner, R. Distribution and cycling of terrigenous dissolved organic matter in the ocean. *Nature* **386**, 480-482, doi: 10.1038/386480a0 (1997).
- Osburn, C. L. & Stedmon, C. A. Linking the chemical and optical properties of dissolved organic matter in the Baltic–North Sea transition zone to differentiate three allochthonous inputs. *Mar. Chem.* **126**, 281-294, doi: 10.1016/j.marchem.2011.06.007 (2011).
- Osterholz, H. *et al.* Deciphering associations between dissolved organic molecules and bacterial communities in a pelagic marine system. *ISME* **10**, 1717-1730, doi: 10.1038/ismej.2015.231 (2016).
- Patel-Sorrentino, N. *et al.* Effects of UV-visible irradiation on natural organic matter from the Amazon basin. *Sci. Total Environ.* **321**, 231-239, doi: 10.1016/j.scitotenv.2003.08.017 (2004).
- Paul, N. D. *et al.* Ecological responses to UV radiation: interactions between the biological effects of UV on plants and on associated organisms. *Physiol. Plant* **145**, 565-581, doi: 10.1111/j.1399-3054.2011.01553.x (2012).
- Perdue, E. M. *et al.* Substitution patterns in aromatic rings by increment analysis. Model development and application to natural organic matter. *Anal. Chem.* **79**, 1010-1021, doi: 10.1021/ac061611y (2007).

- Pérez, M. A. P. *et al.* Dissolved organic matter dynamic in the Amazon basin: Sorption by mineral surfaces. *Chem. Geol.* **286**, 158-168, doi: 10.1016/j.chemgeo.2011.05.004 (2011).
- Peterson, B.J. Aquatic Primary Productivity and the  $^{14}\text{C}$ - $\text{CO}_2$  Method: A History of the Productivity Problem. *Annu. Rev. Ecol. Evol. Syst.* **11**, 359-385, doi: 10.1146/annurev.es.11.110180.002043 (1980).
- Philippe, A. & Schaumann, G. E. Interactions of dissolved organic matter with natural and engineered inorganic colloids: a review. *Environ. Sci. Technol.* **48**, 8946-8962, doi: 10.1021/es502342r (2014).
- Pignatello, J. J. Dynamic interactions of natural organic matter and organic compounds. *J. Soils Sed.* **12**, 1241-1256, doi: 10.1007/s11368-012-0490-4 (2012).
- Powers, L. C. *et al.* Sargassum sp. Act as a Large Regional Source of Marine Dissolved Organic Carbon and Polyphenols. *Glob. Biogeochem. Cycl.* **33**, 1423-1439, doi: 10.1029/2019gb006225 (2019).
- Puhakka, M. *et al.* River Types, Site Evolution and Successional Vegetation Patterns in Peruvian Amazonia. *J. Biogeogr.* **19**, 651, doi: 10.2307/2845707 (1992).
- Qualls, R. G. & Haines, B. L. Biodegradability of dissolved organic matter in forest throughfall, soil solution, and stream water. *Soil Sci. Soc. Am. J.* **56**, 578-586, doi: 10.2136/sssaj1992.03615995005600020038x (1992).
- Quay, P. D. *et al.* Carbon cycling in the Amazon River: Implications from the  $^{13}\text{C}$  compositions of particles and solutes. *Limnol. Oceanogr.* **37**, 857-871, doi: 10.4319/lo.1992.37.4.0857 (1992).
- Quesada, C.A. *et al.* Soils of Amazonia with particular reference to the RAINFOR sites. *Biogeosciences* **8**, 1415–1440, doi: 10.5194/bg-8-1415-2011 (2011).
- Raeke, J. *et al.* Selectivity of solid phase extraction of freshwater dissolved organic matter and its effect on ultrahigh resolution mass spectra. *Environ. Sci. Proc. Imp.* **18**, 918-927, doi: 10.1039/C6EM00200E (2016).
- Rai, H., Hill, G. Primary production in the Amazonian aquatic ecosystem. In: Sioli, H. (Ed.), *The Amazon. Monographiae Biologicae. Springer, Netherlands*, pp. 311–335 (1984).
- Ramette, A. Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol.* **62**, 142-160, doi: 10.1111/j.1574-6941.2007.00375.x (2007).
- Raymond, P. A. & Spencer, R. G. Riverine Dissolved Organic Matter. In: Biogeochemistry of Marine Dissolved Organic Matter. *Academic Press*, 509-533, doi: 10.1016/B978-0-12-405940-5.00011-X (2015).
- Reader, H. E. *et al.* Mass and UV-visible spectral fingerprints of dissolved organic matter: sources and reactivity. *Front. Mar. Sci.* **2**, 88, doi: 10.3389/fmars.2015.00088 (2015).
- Remucal, C. K. *et al.* Molecular-level transformation of dissolved organic matter during oxidation by ozone and hydroxyl radical. *Environ. Sci. Technol.* **54**, 10351-10360, doi: 10.1021/acs.est.0c03052 (2020).

- Repeta, D. J. *et al.* Chemical characterization of high molecular weight dissolved organic matter in fresh and marine waters. *Geochim. Cosmochim. Ac.* **66**, 955-962, doi: 10.1016/S0016-7037(01)00830-4 (2002).
- Richey, J. E. *et al.* Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric CO<sub>2</sub>. *Nature* **416**, 617-620, doi: 10.1038/416617a (2002).
- Richey, J. E. *et al.* Biogeochemistry of carbon in the Amazon River. *Limnol. Oceanogr.* **35**, 352-371, doi: 10.4319/lo.1990.35.2.0352 (1990).
- Ríos-Villamizar, E. *et al.* Chemistry of different Amazonian water types for river classification: a preliminary review. *Water and society ii* **178**, 1117, doi: 10.2495/WS130021 (2013).
- Ríos-Villamizar, E. A. *et al.* Surface water quality and deforestation of the Purus river basin, Brazilian Amazon. *Int. Aqu. Res.* **9**, 81-88, doi: 10.1007/s40071-016-0150-1 (2016).
- Ríos-Villamizar, E. A. *et al.* New insights on the classification of major Amazonian river water types. *Sus. Wat. Res. Manag.* **6**, 1-16, doi: 10.1007/s40899-020-00440-5 (2020).
- Ritchie, J.D. and Perdue, E.M. Proton-binding study of standard and reference fulvic acids, humic acids, and natural organic matter. *Geochimica et cosmochimica acta* **67**, 85-96, doi: 10.1016/s0016-7037(02)01044-x (2003).
- Roberts, S. K. Plasma membrane anion channels in higher plants and their putative functions in roots. *New Phytol.* **169**, 647-666, doi: 10.1111/j.1469-8137.2006.01639.x (2006).
- Rodriguez, F. J. *et al.* A comprehensive structural evaluation of humic substances using several fluorescence techniques before and after ozonation. Part I: structural characterization of humic substances. *Sci. Total. Environ.* **476-477**, 718-730, doi: 10.1016/j.scitotenv.2013.11.150 (2014).
- Roebuck, J. A. *et al.* Controls of Land Use and the River Continuum Concept on Dissolved Organic Matter Composition in an Anthropogenically Disturbed Subtropical Watershed. *Environ. Sci. Technol.* **54**, 195-206, doi: 10.1021/acs.est.9b04605 (2020).
- Romano, S. *et al.* Exo-metabolome of *Pseudovibrio* sp. FO-BEG1 analyzed by ultra-high resolution mass spectrometry and the effect of phosphate limitation. *PloS one* **9**, e96038, doi: 10.1371/journal.pone.0096038 (2014).
- Romera-Castillo, C. *et al.* Net production and consumption of fluorescent colored dissolved organic matter by natural bacterial assemblages growing on marine phytoplankton exudates. *Appl. Environ. Microbiol.* **77**, 7490-7498, doi: 10.1128/AEM.00200-11 (2011).
- Roth, V.-N. *et al.* Ecosystem-specific composition of dissolved organic matter. *Vadose Zone J.* **13**, 7, doi: 10.2136/vzj2013.09.0162 (2014).
- Roth, V. N. *et al.* The molecular composition of dissolved organic matter in forest soils as a function of pH and temperature. *PLoS One* **10**, e0119188, doi: 10.1371/journal.pone.0119188 (2015).

- Roth, V. N. *et al.* Persistence of dissolved organic matter explained by molecular changes during its passage through soil. *Nat. Geosci.* **12**, 755-761, doi: 10.1038/s41561-019-0417-4 (2019).
- Saatchi, S. S. *et al.* Distribution of aboveground live biomass in the Amazon basin. *Glob. Chan. Biol.* **13**, 816-837, doi: 10.1111/j.1365-2486.2007.01323.x (2007).
- Saliot, A. *et al.* Particulate organic carbon, sterols, fatty acids and pigments in the Amazon River system. *Biogeochem.* **53**, 79-103, doi: 10.1023/a:1010754022594 (2001).
- Salo, J. *et al.* River dynamics and the diversity of Amazon lowland forest. *Nature* **322**, 254-258, doi: 10.1038/322254a0 (1986).
- Sanderman, J. *et al.* Linking soils and streams: Sources and chemistry of dissolved organic matter in a small coastal watershed. *Wat. Resour. Res.* **45**, 3, doi: 10.1029/2008WR006977 (2009).
- Santoro, A.L. *et al.* Dark carbon fixation: an important process in lake sediments. *PLoS One* **8**, e65813, doi: 10.1371/journal.pone.0065813 (2013).
- Sawakuchi, H. O. *et al.* Carbon Dioxide Emissions along the Lower Amazon River. *Front. Mar. Sci.* **4**, 76, doi: 10.3389/fmars.2017.00076 (2017).
- Schmidt, F. *et al.* Diagenetic transformation of dissolved organic nitrogen compounds under contrasting sedimentary redox conditions in the Black Sea. *Environ. Sci. Technol.* **45**, 5223-5229, doi: 10.1021/es2003414 (2011).
- Schmidt, F. *et al.* Unraveling signatures of biogeochemical processes and the depositional setting in the molecular composition of pore water DOM across different marine environments. *Geochim. Cosmochim. Ac.* **207**, 57-80, doi: 10.1016/j.gca.2017.03.005 (2017).
- Schmidt, F. *et al.* Molecular characterization of dissolved organic matter in pore water of continental shelf sediments. *Geochim. Cosmochim. Ac.* **73**, 3337-3358, doi: 10.1016/j.gca.2009.03.008 (2009).
- Schmitt-Kopplin, P. *et al.* Structural Changes in a Dissolved Soil Humic Acid during Photochemical Degradation Processes under O<sub>2</sub> and N<sub>2</sub> Atmosphere. *Environ. Sci. & Technol.* **32**, 2531-2541, doi: 10.1021/es970636z (1998).
- Schmitt-Kopplin, P. *et al.* High molecular diversity of extraterrestrial organic matter in Murchison meteorite revealed 40 years after its fall. *Proc. Natl. Acad. Sci. USA* **107**, 2763-2768, doi: 10.1073/pnas.0912157107 (2010).
- Schmitt-Kopplin, P. *et al.* Systems chemical analytics: introduction to the challenges of chemical complexity analysis. *Faraday Discuss.* **218**, 9-28, doi: 10.1039/c9fd00078j (2019).
- Schmitt-Kopplin, P. *et al.* Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh-resolution mass spectrometry and nuclear magnetic resonance spectroscopy: organosulfates as photochemical smog constituents. *Anal. Chem.* **82**, 8017-8026, doi: 10.1021/ac101444r (2010).



- Seidel, M. *et al.* Seasonal and spatial variability of dissolved organic matter composition in the lower Amazon River. *Biogeochem.* **131**, 281-302, doi: 10.1007/s10533-016-0279-4 (2016).
- Seidel, M. *et al.* Molecular-level changes of dissolved organic matter along the Amazon River-to-ocean continuum. *Mar. Chem.* **177**, 218-231, doi: 10.1016/j.marchem.2015.06.019 (2015).
- Shade, A. *et al.* Interannual dynamics and phenology of bacterial communities in a eutrophic lake. *Limnol. Oceanogr.* **52**, 487-494, doi: 10.4319/lo.2007.52.2.0487 (2007).
- Shen, Y. *et al.* Origins and bioavailability of dissolved organic matter in groundwater. *Biogeochem.* **122**, 61-78, doi: 10.1007/s10533-014-0029-4 (2015).
- Silva, M. *et al.* Classification of Amazonian rivers: A strategy for the preservation of these resources. *Holos Environ.* **13**, 163-174, doi: 10.14295/holos.v13i2.7344 (2013).
- Simon, C. *et al.* Riverine mixing at the molecular scale – An ultrahigh-resolution mass spectrometry study on dissolved organic matter and selected metals in the Amazon confluence zone (Manaus, Brazil). *Org. Geochem.* **129**, 45-62, doi: 10.1016/j.orggeochem.2019.01.013 (2019).
- Simon, M. and Azam, F. Protein content and protein synthesis rates of planktonic marine bacteria. *Mar. Ecol. Prog.* **51**, 201-213, doi: 10.3354/meps051201 (1989).
- Simpson, A. J. *et al.* Molecular structures and associations of humic substances in the terrestrial environment. *Naturwiss.* **89**, 84-88, doi: 10.1007/s00114-001-0293-8 (2002).
- Simpson, A. J. *et al.* NMR spectroscopy in environmental research: from molecular interactions to global processes. *Prog. Nucl. Magn. Reson. Spectrosc.* **58**, 97-175, doi: 10.1016/j.pnmrs.2010.09.001 (2011).
- Singh, S. *et al.* Seasonal pattern of dissolved organic matter (DOM) in watershed sources: influence of hydrologic flow paths and autumn leaf fall. *Biogeochem.* **118**, 321-337, doi: 10.1007/s10533-013-9934-1 (2014).
- Sinninghe Damste, J.S. and De Leeuw, J.W. Analysis, structure and geochemical significance of organically-bound sulphur in the geosphere: State of the art and future research. *Org. Geochem.* **16**, 1077-1101, doi: 10.1016/0146-6380(90)90145-p (1990).
- Sleighter, R.L. and Hatcher, P.G. The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter. *J. Mass Spectrom.* **42**, 559-574, doi: 10.1002/jms.1221 (2007).
- Sleighter, R. L. & Hatcher, P. G. Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mar. Chem.* **110**, 140-152, doi: 10.1016/j.marchem.2008.04.008 (2008).
- Sleighter, R.L. *et al.* Evidence of incorporation of abiotic S and N into prairie wetland dissolved organic

- matter. *Environ. Sci. Technol. Lett.* **1**, 345-350, doi: 10.1021/ez500229b (2014).
- Smith Jr, W. O. & Demaster, D. J. Phytoplankton biomass and productivity in the Amazon River plume: correlation with seasonal river discharge. *Cont. Shelf Res.* **16**, 291-319, doi: 10.1016/0278-4343(95)00007-N (1996).
- Sponsel, L. E. Amazon ecology and adaptation. *Ann. Rev. Anthr.* **15**, 67-97, doi: 10.1146/annurev.an.15.100186.000435 (1986).
- Stedmon, C. A. & Bro, R. Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. *Limnol. Oceanogr. Meth.* **6**, 572-579, doi: 10.4319/lom.2008.6.572 (2008).
- Stubbins, A. *et al.* Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. *Limnol. Oceanogr.* **55**, 1467-1477, doi: 10.4319/lo.2010.55.4.1467 (2010).
- Stubbins, A. *et al.* What's in an EEM? Molecular signatures associated with dissolved organic fluorescence in boreal Canada. *Environ. Sci. Technol.* **48**, 10598-10606, doi: 10.1021/es502086e (2014).
- Subdiaga, E. *et al.* Effects of Sorption on Redox Properties of Natural Organic Matter. *Environ. Sci. Technol.* **53**, 14319-14328, doi: 10.1021/acs.est.9b04684 (2019).
- Sulzberger, B. & Durisch-Kaiser, E. Chemical characterization of dissolved organic matter (DOM): A prerequisite for understanding UV-induced changes of DOM absorption properties and bioavailability. *Aqu. Sci.* **71**, 104-126, doi: 10.1007/s00027-008-8082-5 (2009).
- Svensson, T. *et al.* Chlorination of soil organic matter: The role of humus type and land use. *Sci. Total Environ.* **806**, 150478, doi: 10.1016/j.scitotenv.2021.150478 (2022).
- Swierk, L. & Madigosky, S. R. Environmental perceptions and resource use in rural communities of the Peruvian Amazon (Iquitos and vicinity, Maynas Province). *Trop. Cons. Sci.* **7**, 382-402, doi: 10.1177/1940082914007003 (2014).
- Tank, J. L. *et al.* A review of allochthonous organic matter dynamics and metabolism in streams. *J. North Am. Benthol. Soc.* **29**, 118-146, doi: 10.1899/08-170.1 (2010).
- Tipping, E. & Woof, C. Humic substances in acid organic soils: modelling their release to the soil solution in terms of humic charge. *J. Soil Sci.* **41**, 573-586, doi: 10.1111/j.1365-2389.1990.tb00227.x (1990).
- Tranvik, L. J. *et al.* Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceanogr.* **54**, 2298-2314 (2009).
- Tziotis, D. *et al.* Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *Eur. J. Mass Spectrom. (Chichester)* **17**, 415-421, doi: 10.1255/ejms.1135 (2011).

- Uselman, S. M. *et al.* Contribution of root vs. leaf litter to dissolved organic carbon leaching through soil. *Soil Sci. Soc. Am. J.* **71**, 1555-1563, doi: 10.2136/sssaj2006.0386 (2007).
- Utsumi, M. *et al.* Dynamics of dissolved methane and methane oxidation in dimictic Lake Nojiri during winter. *Limnol. Oceanogr.* **43**, 10-17, doi: 10.4319/lo.1998.43.1.0010 (1998).
- Valderrama, J.C. The simultaneous analysis of total nitrogen and total phosphorus in natural waters. *Mar. Chem.* **10**, 109-122, doi: 10.1016/0304-4203(81)90027-X (1981).
- Valle, J. *et al.* Extensive processing of sediment pore water dissolved organic matter during anoxic incubation as observed by high-field mass spectrometry (FTICR-MS). *Wat. Res.* **129**, 252-263, doi: 10.1016/j.watres.2017.11.015 (2017).
- Vannone, R. L. *et al.*, River continuum concept. *Can. J. Fish. Aquat. Sci.* **37**, 37, 130-137, doi : 10.1139/f80-017 (1980).
- Vannote, R. L. *et al.* The River Continuum Concept. *Can. J. Fish. Aquat. Sci.* **37**, 130-137, doi: 10.1139/f80-017 (1980).
- Verdugo, P. *et al.* The oceanic gel phase: a bridge in the DOM–POM continuum. *Mar. Chem.* **92**, 67-85, doi: 10.1016/j.marchem.2004.06.017 (2004).
- Veum, K. S. *et al.* Runoff and dissolved organic carbon loss from a paired-watershed study of three adjacent agricultural watersheds. *Agric. Ecosys. Environ.* **130**, 115-122, doi: 10.1016/j.agee.2008.12.006 (2009).
- Vidal, L.O. *et al.* Hydrological pulse regulating the bacterial heterotrophic metabolism between Amazonian mainstems and floodplain lakes. *Front. Microbiol.* **6**, 1-10, doi: 10.3389/fmicb.2015.01054 (2015).
- Waggoner, D. C. *et al.* Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. *Org. Geochem.* **82**, 69-76, doi: 10.1016/j.orggeochem.2015.02.007 (2015).
- Waichman, A. V. Autotrophic carbon sources for heterotrophic bacterioplankton in a floodplain lake of central Amazon. *Hydrobiol.* **341**, 27-36, doi: 10.1007/bf00012300 (1996).
- Wang, Y. *et al.* Characterization and spacial distribution variability of chromophoric dissolved organic matter (CDOM) in the Yangtze Estuary. *Chemosphere* **95**, 353-362, doi: 10.1016/j.chemosphere.2013.09.044 (2014).
- Wang, Y. *et al.* Revealing sources and distribution changes of dissolved organic matter (DOM) in pore water of sediment from the Yangtze estuary. *PLoS One* **8**, e76633, doi: 10.1371/journal.pone.0076633 (2013).
- Ward, N. D. *et al.* Degradation of terrestrially derived macromolecules in the Amazon River. *Nat. Geosci.* **6**, 530-533, doi: 10.1038/ngeo1817 (2013).

- Ward, N. D. *et al.* The compositional evolution of dissolved and particulate organic matter along the lower Amazon River-Óbidos to the ocean. *Mar. Chem.* **177**, 244-256, doi: 10.1016/j.marchem.2015.06.013 (2015).
- Ward, N. D. *et al.* The reactivity of plant-derived organic matter and the potential importance of priming effects along the lower Amazon River. *J. Geophys. Res.: Biogeosci.* **121**, 1522-1539, doi: 10.1002/2016jg003342 (2016).
- Ward, N. D. *et al.* Where Carbon Goes When Water Flows: Carbon Cycling across the Aquatic Continuum. *Front. Mar. Sci.* **4**, 7, doi: 10.3389/fmars.2017.00007 (2017).
- Ward, N. D. *et al.* Velocity-amplified microbial respiration rates in the lower Amazon River. *Limnol. Oceanogr. Lett.* **3**, 265-274, doi: 10.1002/lol2.10062 (2018).
- Ward, N. D. *et al.* Enhanced Aquatic Respiration Associated With Mixing of Clearwater Tributary and Turbid Amazon River Waters. *Front. Earth Sci.* **7**, 101, doi: 10.3389/feart.2019.00101 (2019).
- Wehrli, B. Conduits of the carbon cycle, *Nature*, **503**, 346–347, doi: 10.1038/503346a (2013).
- Weiss, M. *et al.* Molecular architecture and electrostatic properties of a bacterial porin. *Science* **254**, 1627-1630, doi: 10.1126/science.1721242 (1991).
- Wilske, C. *et al.* Photochemically Induced Changes of Dissolved Organic Matter in a Humic-Rich and Forested Stream. *Water* **12**, 2, doi: 10.3390/w12020331 (2020).
- Winch, S. *et al.* Increased metal bioavailability following alteration of freshwater dissolved organic carbon by ultraviolet B radiation exposure. *Environ. Toxicol. Int. J.* **17**, 267-274, doi: 10.1002/tox.10057 (2002).
- Wissmar, R. C. *et al.* Plankton Metabolism and Carbon Processes in the Amazon River, Its Tributaries, and Floodplain Waters, Peru-Brazil, May-June 1977. *Ecol.* **62**, 1622-1633, doi: 10.2307/1941517 (1981).
- Woods, G. C. *et al.* Evidence for the enhanced lability of dissolved organic matter following permafrost slope disturbance in the Canadian High Arctic. *Geochim. Cosmochim. Ac.* **75**, 7226-7241, doi: 10.1016/j.gca.2011.08.013 (2011).
- Wu, J., Tu, W., Li, C. and He, F. Binding characteristics of copper onto biochar-derived DOM using general, heterospectral and moving-window two-dimensional correlation analyses. *J. Hazard. Mater.* **435**, 129021, doi: 10.1016/j.jhazmat.2022.129021 (2022).
- Xenopoulos, M.A. *et al.* How humans alter dissolved organic matter composition in freshwater: relevance for the Earth's biogeochemistry. *Biogeochem.* **154**, 323-348, doi: 10.1007/s10533-021-00753-3 (2021).

- Xia, X. *et al.* Molecular insights into roles of dissolved organic matter in Cr (III) immobilization by coprecipitation with Fe (III) probed by STXM-Ptychography and XANES spectroscopy. *Environ. Sci. Tech.* **56**, 2432-2442, doi: 10.1021/acs.est.1c07528 (2022).
- Xie, H. *et al.* Photooxidation and its effects on the carboxyl content of dissolved organic matter in two coastal rivers in the southeastern United States. *Environ. Sci. Technol.* **38**, 4113-4119, doi: 10.1021/es035407t (2004).
- Xu, H. & Guo, L. Molecular size-dependent abundance and composition of dissolved organic matter in river, lake and sea waters. *Wat. Res.* **117**, 115-126, doi: 10.1016/j.watres.2017.04.006 (2017).
- Xu, H., Guo, L. & Jiang, H. Depth-dependent variations of sedimentary dissolved organic matter composition in a eutrophic lake: implications for lake restoration. *Chemosphere* **145**, 551-559, doi: 10.1016/j.chemosphere.2015.09.015 (2016).
- Yamashita, Y. & Tanoue, E. Chemical characteristics of amino acid-containing dissolved organic matter in seawater. *Org. Geochem.* **35**, 679-692, doi: 10.1016/j.orggeochem.2004.02.007 (2004).
- Yamashita, Y. & Tanoue, E. Chemical characterization of protein-like fluorophores in DOM in relation to aromatic amino acids. *Mar. Chem.* **82**, 255-271, doi: 10.1016/S0304-4203(03)00073-2 (2003).
- Yu, J. *et al.* Excitation-emission matrix (EEM) fluorescence spectroscopy for characterization of organic matter in membrane bioreactors: Principles, methods and applications. *Front. Environ. Sci. Engin.* **14**, 1-19, doi: 10.1007/s11783-019-1210-8 (2020).
- Zakem, E. J. & Levine, N. M. Systematic variation in marine dissolved organic matter stoichiometry and remineralization ratios as a function of lability. *Glob. Biogeochem. Cyc.* **33**, 1389-1407, doi: 10.1029/2019GB006375 (2019).
- Zapata-Ríos, G. *et al.* Biological Diversity and Ecological Networks in the Amazon. doi:10.55161/dgnm5984 (2022).
- Zhang, F. *et al.* Molecular and structural characterization of dissolved organic matter during and post cyanobacterial bloom in Taihu by combination of NMR spectroscopy and FTICR mass spectrometry. *Wat. Res.* **57**, 280-294, doi: 10.1016/j.watres.2014.02.051 (2014).
- Zhang, H. *et al.* Characterization of low molecular weight dissolved natural organic matter along the treatment trait of a waterworks using Fourier transform ion cyclotron resonance mass spectrometry. *Wat. Res.* **46**, 5197-5204, doi: 10.1016/j.watres.2012.07.004 (2012).
- Zhang, K. & Parker, K. M. Halogen radical oxidants in natural and engineered aquatic systems. *Environ. Sci. Technol.* **52**, 9579-9594, doi: 10.1021/acs.est.8b02219 (2018).
- Zhao, Z. *et al.* Picocyanobacteria and deep-ocean fluorescent dissolved organic matter share similar optical properties. *Nat. Commun.* **8**, 15284, doi: 10.1038/ncomms15284 (2017).

Zigah, P. *et al.* An isotopic ( $\Delta^{14}\text{C}$ ,  $\delta^{13}\text{C}$ , and  $\delta^{15}\text{N}$ ) investigation of the composition of particulate organic matter and zooplankton food sources in Lake Superior and across a size-gradient of aquatic systems. *Biogeosci.* **9**, 3663-3678, doi: 10.5194/bg-9-3663-2012 (2012).

# 6. Appendix

## 6.1 Article 1 + licenses for copyrighted content

Science of the Total Environment 857 (2023) 159620



Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: [www.elsevier.com/locate/scitotenv](http://www.elsevier.com/locate/scitotenv)



### Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy



Siyu Li <sup>a</sup>, Mourad Harir <sup>a,b</sup>, Philippe Schmitt-Kopplin <sup>a,b</sup>, Michael Gonsior <sup>c</sup>, Alex Enrich-Prast <sup>d,e</sup>, David Bastviken <sup>d</sup>, Juliana Valle <sup>a</sup>, Fausto Machado-Silva <sup>f,g</sup>, Norbert Hertkorn <sup>a,\*</sup>

<sup>a</sup> Research Unit Analytical Biogeochemistry, Helmholtz Munich, Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany

<sup>b</sup> Chair of Analytical Food Chemistry, Technische Universität München, Alte Akademie 10, 85354 Freising-Weihenstephan, Germany

<sup>c</sup> University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory, 146 Williams Street, Solomons, MD 20688, United States

<sup>d</sup> Department of Thematic Studies – Environmental Change, Linköping University, SE-581 83 Linköping, Sweden

<sup>e</sup> Institute of Marine Science, Federal University of São Paulo, Santos, Brazil

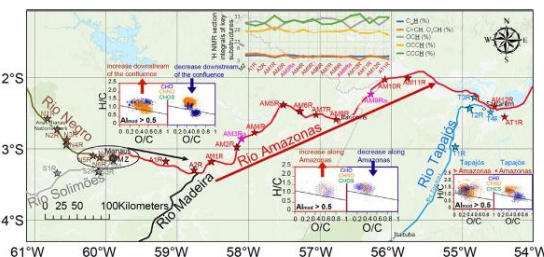
<sup>f</sup> Program in Geosciences – Environmental Geochemistry, Chemistry Institute, Fluminense Federal University, 24020-141 Niterói, Brazil

<sup>g</sup> Department of Environmental Sciences, University of Toledo, Toledo, OH 43606, USA

#### HIGHLIGHTS

- Unveiling DOM molecular and structural characteristics in the Amazon basin
- ESI[±] FT-ICR MS and NMR provided comprehensive coverage of DOM processing.
- Nitrogen containing molecules are critical partners in the cycling of DOM.
- DOM in the Amazon River mixing zones showed non-conservative behavior.
- Physicochemical and microbial processes determine DOM changes in the Amazon basin.

#### GRAPHICAL ABSTRACT



#### ARTICLE INFO

Editor: Shuzhen Zhang

**Keywords:**  
Mixing zone  
Tributaries  
Environmental drivers  
FT-ICR MS  
NMR

#### ABSTRACT

Rivers are natural biogeochemical systems shaping the fates of dissolved organic matter (DOM) from leaving soils to reaching the oceans. This study focuses on Amazon basin DOM processing employing negative and positive electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI[±] FT-ICR MS) and nuclear magnetic resonance spectroscopy (NMR) to reveal effects of major processes on the compositional space and structural characteristics of black, white and clear water systems. These include non-conservative mixing at the confluences of (1) Solimões and the Negro River, (2) the Amazon River and the Madeira River, and (3) in-stream processing of Amazon River DOM between the Madeira River and the Tapajós River. The Negro River (black water) supplies more highly oxygenated and high molecular weight compounds, whereas the Solimões and Madeira Rivers (white water) contribute more CHNO and CHOS molecules to the Amazon River main stem. Aliphatic CHO and abundant CHNO compounds prevail in Tapajós River DOM (clear water), likely originating from primary production. Sorption onto particles and heterotrophic microbial degradation are probably the principal mechanisms for the observed changes in DOM composition in the Amazon River and its tributaries.

**Abbreviations:** S-DOM, PPL-based solid phase extracted dissolved organic matter (SPE-DOM) from the Solimões River; A-DOM, SPE-DOM from the Amazon River; N-DOM, SPE-DOM from the Negro River; T-DOM, SPE-DOM from the Tapajós River.

\* Corresponding author.

E-mail address: [hertkorn@helmholtz-muenchen.de](mailto:hertkorn@helmholtz-muenchen.de) (N. Hertkorn).

<http://dx.doi.org/10.1016/j.scitotenv.2022.159620>

Received 19 July 2022; Received in revised form 15 September 2022; Accepted 17 October 2022

Available online 21 October 2022

0048-9697/© 2022 Published by Elsevier B.V.



## 1. Introduction

The Amazon River catchment is the largest river system in the world, responsible for ~20 % of the global freshwater discharge and exports of ~35 Tg organic carbon (of which 60–70 % is dissolved) annually to the world's oceans (Moreira-Turcq et al., 2003a). The river system is the ultimate driver of aquatic life in the region and a key connector between terrestrial organic carbon and its mineralization (Battin et al., 2009; Cole et al., 2007). Riverine dissolved organic matter in the Amazon River and its tributaries (AZ-DOM) is involved in critical ecosystem processes, such as nutrient availability, growth efficiency of aquatic organisms, decomposition rate of plant debris, and ultimately, the carbon cycle (Drake et al., 2021; Santos-Junior et al., 2020). Therefore, ecosystem-wide understanding of dissolved organic matter (DOM) molecular composition and structures is needed to comprehend key biogeochemical processes in the Amazon basin.

Previous studies on the origin and fate of organic matter (OM) in the Amazon River included stable isotope (Mortillaro et al., 2011; Quay et al., 1992) and radioisotope analysis (Hedges et al., 1986) to assess sources and age of organic carbon. Quantification of targeted compounds involved humic substances (Ertel et al., 1986), saccharides (Salot et al., 2001) and amino acids (Hedges et al., 1994), among others. Optical spectroscopy (Martinez et al., 2015), mass spectrometry (Gonsior et al., 2016) and  $^{13}\text{C}$  NMR spectroscopy (Hedges et al., 1992) were applied to understand the composition of its OM. Incubation experiments measured the rates of primary production (Gagne-Maynard et al., 2017), biological (Benner et al., 1995) and photo-degradation (Amado et al., 2006; Amon and Benner, 1996) in the Amazon River. However, comprehensive molecular understanding of AZ-DOM reactivity and bioavailability, and assessment of its biotic and abiotic processing is lacking at present, especially in tributaries with different water types and in major confluence zones.

Many ungauged tributaries with specific biogeochemical characteristics enter the Amazon River which have long been classified into three types according to their appearance: "Blackwaters" that are relatively acidic (pH ~ 5), low in total cations and rich in DOM, such as the Negro River, "Whitewaters" that show a near-neutral pH and are relatively rich in total cations and in suspended sediments, such as Solimões and Madeira River, and "Clearwaters" which exhibit low suspended sediment loads, and high light transparency, such as Tapajós River (McClain and Naiman, 2008). Different water types in the Amazon basin have been shown to have distinct DOM compositions (Gonsior et al., 2016). Solimões and Negro combine in Manaus to form the Amazon River (turbid water). Madeira (DOC ~ 5.8 mg/L, POC ~ 0.83 mg/L) joins the Amazon River downstream at ~140 km east of Manaus (do Nascimento et al., 2015). Tapajós River drains into the Amazon River at the town of Santarém (~620 km from Manaus). Mixing of different water types initiates complex biogeochemical processing, including alteration of aquatic microbial communities and food webs (Lynch et al., 2019), DOM itself (Bianchi and Ward, 2019), and of river bulk characteristics.

Amazon River is a net heterotrophic ecosystem in which respiration far exceeds primary production (Hedges et al., 1994), and most  $\text{CO}_2$  outgassing originates from allochthonous carbon (Abril et al., 2014; Mayorga et al., 2005; Richey et al., 2002). Clear rivers show comparatively higher primary production, as solar radiation exposure decreases due to high aromatic DOM content in black rivers and turbidity from high sediment loads in white rivers (Moreira-Turcq et al., 2003a). Bacterial and photochemical processing and mineralization as well as chemoselective sorption to minerals are the defining processes determining OM biogeochemistry in the Amazon catchments (Amon and Benner, 1996; Armanious et al., 2014; Moreira-Turcq et al., 2003b; Pérez et al., 2011; Shen, 1999).

Polydispersity, heterogeneity, and temporal dynamics of DOM have made its molecular characterization a long-standing challenge. This study employs complementary negative and positive electrospray ionization ESI [ $\pm$ ] FT-ICR MS to determine chemical composition out of the AZ-DOM with excellent coverage of CHO, CHNO, and CHOS molecules (Hertkorn et al., 2016, 2013). 800 MHz  $^1\text{H}$ NMR spectroscopy provided quantification

of major atomic environments in DOM molecules with excellent S/N ratio and resolution, enabling quantification of key DOM structural units including aliphatic and carboxyl-rich alicyclic molecules (CRAM), oxygenated aliphatics including carbohydrates, olefins, and aromatics (Hertkorn et al., 2013, 2016; Powers et al., 2019; Simpson et al., 2011). Bulk characterization of Amazon waters in conjunction with selected incubation experiments further contributed to dissecting relevant aspects of DOM processing.

## 2. Material and methods

### 2.1. Sampling and site locations

We collected water samples at 31 sites in the Amazon River main stem eastwards, including Solimões (white water), Negro (black water), Amazon (turbid water), and Tapajós (clear water) (Figs. 1, S1 and Table S1) between April 2<sup>nd</sup> and May 25<sup>th</sup> in 2014, during a high-water period with very high discharge volume. Sampling coordinates and information, as well as isolated DOM by PPL-based solid phase extraction (SPE-DOM; Dittmar et al., 2008) of the water samples are described in the Supplementary information (SI).

### 2.2. FT-ICR mass spectrometry

Negative and positive electrospray ionization (ESI [ $\pm$ ]) Fourier transform ion cyclotron resonance mass spectra (FT-ICR MS) were acquired using a 12 T Bruker Solarix mass spectrometer (Bruker Daltonics, Bremen, Germany) and an Apollo II electrospray ionization (ESI) source (Hertkorn et al., 2016). Data processing used Compass Data Analysis 4.1 (Bruker, Bremen, Germany) and formula assignment used an in-house made software (NetCalc) (Tziotis et al., 2011); further details are described in SI. Intensity-weighted average bulk parameters such as  $m/z$  and atomic ratios (e.g., O/C, H/C, S/C, N/C) were computed from ESI [ $\pm$ ] FT-ICR MS of AZ-DOM (Tables 1, S3, S4). A modified aromaticity index ( $AI_{\text{mod}}$ ) was computed by considering only half of the oxygen being present in carbonyl functional groups by the equation:  $AI_{\text{mod}} = (1 + C - 0.5O - S - 0.5H) / (C - 0.5O - S - N - P)$  (Koch and Dittmar, 2006). The total intensity of  $m/z$  ions in each sample that had  $AI_{\text{mod}} > 0.5$ , an indicator of aromatic structures, was computed and shown in Tables. 1, S3, S4.

### 2.3. $^1\text{H}$ NMR spectroscopy

800 MHz  $^1\text{H}$  NMR spectra of AZ-DOM were acquired with a Bruker Avance III NMR spectrometer operating at 800.35 MHz ( $B_0 = 18.8$  Tesla) at 283 K from redissolved solids in  $\text{CD}_3\text{OD}$  as described in SI.

### 2.4. Water characterization

Water conductivity (Cond.), temperature (Temp.) and dissolved oxygen (DO) were measured in situ with a portable instrument (Hanna Instruments, Metrohm electrode and PRO-ODO YSI). Inorganic nutrients, DOC, particulate organic carbon (POC), dissolved inorganic carbon (DIC), and pH were determined by standard procedures as described in SI.

### 2.5. Dark carbon fixation (DCF) and heterotrophic bacterial production (HBP)

DCF and HBP are important microbial processes that supply fresh bio-material to both DOC and POC. DCF was determined by the incorporation of  $^{14}\text{CO}_2$  in a known time, and HBP was determined by the measurement of protein incorporation of radio-labelled  $^3\text{H}$ -Leucine as described in SI.

### 2.6. Statistical analysis

Data mining and the application of multivariate statistics served to analyze these complex sets of complementary data. Principal component analysis (PCA) was performed using Simca-P (version 11.5, UmetricsAB, Umeå, Sweden). Hierarchical Cluster Analysis (HCA) was performed using Hierarchical Clustering Explorer 3.0. Spearman correlation analysis ( $p < 0.05$ ,  $r^2 > 0.5$ )



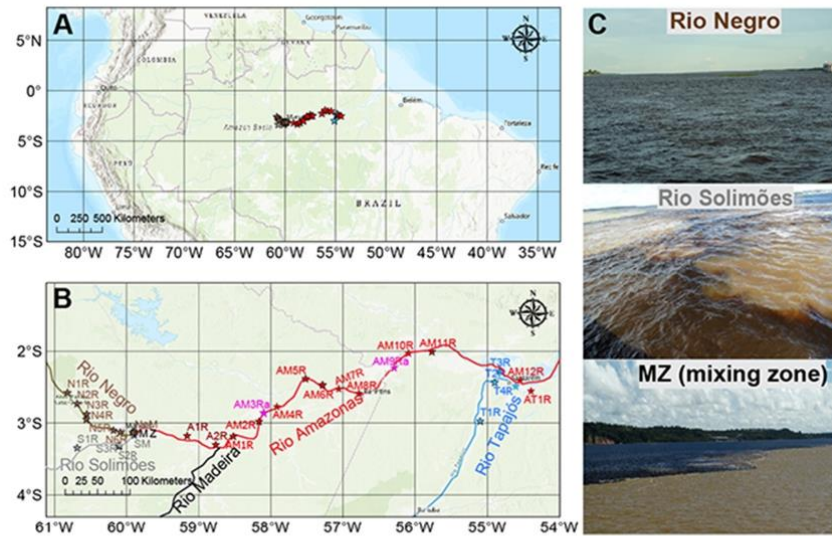


Fig. 1. Maps A, B and own photographs C of the sampling sites in the Amazon River. MZ (black) was sampled at the mixing zone of the Solimões River (grey) and the Negro River (brown). A1R/A2R (dark red) were sampled in the Amazon River upstream of the Madeira River inflow, while the other Amazon River samples (red) were sampled downstream of the Madeira River inflow. AM3Ra and AM9Ra (magenta) were sampled two months later than the other water samples. The enlarged view of panel B is shown in Fig. S1.

was performed using the R statistical platform. Non-metric multidimensional scaling (NMDS) analysis was performed using the R statistical platform with the package vegan to determine correlations of environmental variables with DOM molecular composition in the Amazon River continuum. Expanded descriptions of the statistical analyses are included in SI.

### 3. Results

#### 3.1. High diversity of AZ-DOM revealed by ESI[±] FT-ICR MS and <sup>1</sup>H NMR

ESI[±] FT-ICR MS of all AZ-DOM samples showed distinct distributions of thousands of negative/positive molecular signatures covering the *m/z* range 150–950, indicating considerable ionization selectivity but high complementarity (Figs. 2, S2) (Hertkorn et al., 2008, 2013). Overall, ESI[−] FT-ICR MS preferentially detected high-mass oxygen-rich CHO compounds, whereas ESI[+] FT-ICR MS primarily detected aliphatic CHNO compounds. Nearly half of CHO and CHNO molecular compositions were

found in both ESI[±] modes at near-average H/C and O/C atomic ratios for which the count of feasible isomeric molecules is maximal (Hertkorn et al., 2007). Moreover, CHOS compounds were better ionized in ESI[−] than in ESI[+], and CHNOS compounds were almost absent. All <sup>1</sup>H NMR spectra of AZ-DOM showed smooth bulk envelopes (Fig. S3), reflecting superpositions of millions of atomic environments typical of processed aqueous DOM (Hertkorn et al., 2013).

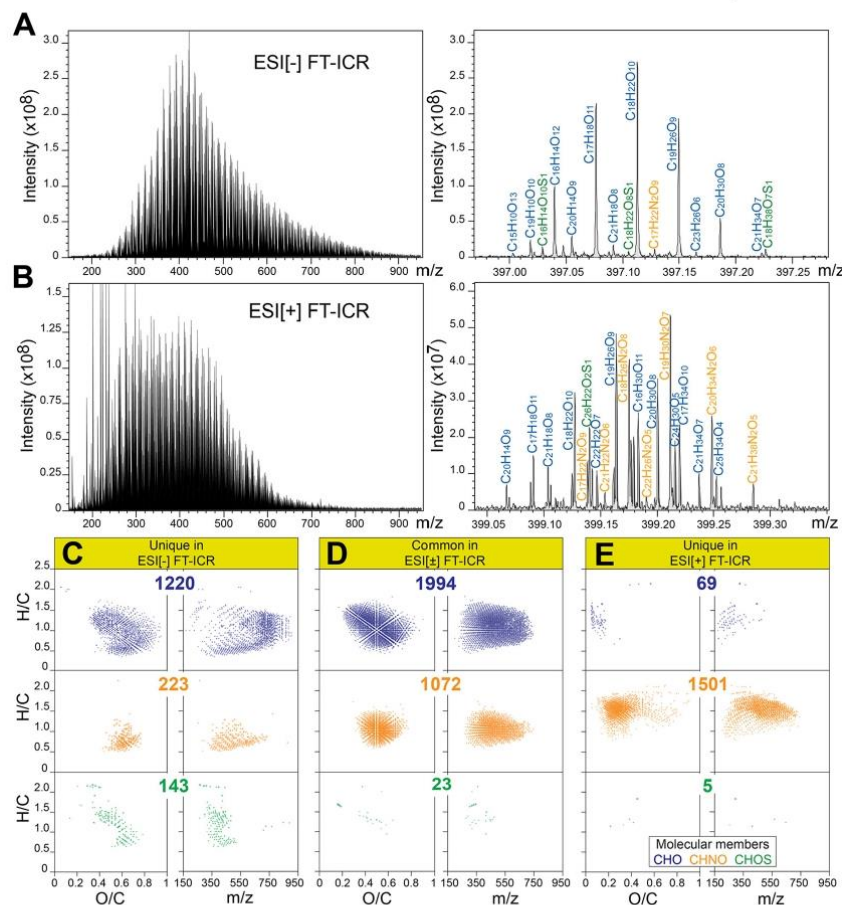
#### 3.2. Decreased abundance of oxygen-rich aromatic DOM molecules downstream of the Solimões-Negro mixing zone

DOC and nitrate (NO<sub>3</sub><sup>−</sup>) concentrations were higher in Negro, whereas POC, DIC, and nitrite (NO<sub>2</sub><sup>−</sup>) concentrations, as well as discharge, conductivity, and pH were higher in Solimões (Table S2). DOC and DIC concentrations in Amazon River ranged between those in Solimões and Negro, consistent with previous studies, whereas ~90 % of POC was lost at the Solimões-Negro (SN)-confluence, much larger than described in previous

Table 1

ESI[±] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned masses occurring in AZ-DOM. This table shows value ranges for DOM in Negro, Solimões, Amazon, and Tapajós main stem, respectively; values for each sample see Tables S3–S4.

FT-ICR MS	Sample	Total counts of mass peaks	CHO %	CHNO %	CHOS %	CHNOS %	<i>m/z</i>	DBE/C	H/C	O/C	N/C × 10 <sup>−2</sup>	S/C × 10 <sup>−5</sup>	Al <sub>mod</sub> > 0.5 %
ESI[−]	Negro	3781–4519	73.2–76.9	21.8–25.2	1.35–1.73	0.00–0.03	477.1–498.3	0.52–0.53	1.04–1.06	0.53–0.55	0.3–0.4	20.7–40.4	14.3–16.7
	Solimões	4759–5043	59.5–63.6	31.0–32.4	5.33–7.20	0.00–0.04	460.2–482.5	0.52–0.53	1.06–1.07	0.54–0.55	0.6–0.8	93.7–112.3	13.7–15.7
	Amazon	4120–5010	63.0–70.3	26.7–30.7	3.05–6.31	0.00–0.34	467.9–488.9	0.51–0.52	1.05–1.07	0.53–0.55	0.5–0.7	43.4–100.5	13.2–15.4
	Tapajós	5276–5898	57.9–62.4	32.1–35.3	3.87–6.93	0.04–2.24	446.6–472.1	0.49–0.52	1.08–1.11	0.51–0.53	0.7–0.8	82.7–120.7	12.0–14.6
	NaM	3781	76.9	21.8	1.35	0.00	498.3	0.53	1.04	0.54	0.3	17.2	16.6
ESI[+]	Negro	4824–5973	28.4–30.9	68.9–71.4	0.05–0.15	0.00–0.04	438.4–454.8	0.39–0.40	1.32–1.36	0.40–0.43	3.5–3.8	2.5–5.3	2.0–2.8
	Solimões	3817–4967	29.5–31.3	68.5–70.2	0.08–0.18	0.00–0.18	425.2–439.2	0.38–0.40	1.34–1.37	0.36–0.41	3.7–3.9	2.3–5.0	2.5–3.3
	Amazon	3477–6160	27.7–34.6	65.1–71.6	0.03–0.61	0.00–1.45	380.89–422.8	0.39–0.41	1.32–1.38	0.36–0.42	3.5–4.3	2.3–16.9	2.4–5.0
	Tapajós	3564–4709	29.6–32.3	67.5–70.3	0.07–0.19	0.02–0.24	386.67–397.6	0.39–0.40	1.35–1.37	0.38–0.39	3.9–4.1	2.8–5.6	3.5–4.5
	NaM	5428	28.9	70.9	0.15	0.02	439.2	0.39	1.36	0.39	3.7	3.4	2.6
MZ	SM	5172	28.5	71.0	0.19	0.27	413.2	0.38	1.38	0.37	4.0	6.2	3.1
	MZ	5808	28.6	71.2	0.17	0.03	437.7	0.38	1.38	0.39	3.8	5.0	2.1



**Fig. 2.** ESI[ $\pm$ ] FT-ICR mass spectra and comparison of their identified molecular compositions exemplified with the Amazon River sample A1R. The A ESI[ $-$ ] and B ESI[ $+$ ] FT-ICR MS spectra showed distinct distributions of mass peaks across a wide mass range ( $m/z$  150–950), as well as signatures at nominal mass 398 confirming relevant ionization selectivity. Respective assignments of molecular compositions are provided for CHO (blue), CHNO (orange), and CHOS (green) molecules. Van Krevelen and mass-edited  $m/z$  diagrams of C the molecular formulae identified only in ESI[ $-$ ] FT-ICR MS; D the molecular formulae identified in both ESI[ $\pm$ ] FT-ICR MS; E the molecular formulae identified only in ESI[ $+$ ] FT-ICR MS. Numbers show counts of assigned formulae.

studies (Moreira-Turcq et al., 2003b). Furthermore, dissolved oxygen (DO) at Solimões-Negro confluence and the Amazon upstream was  $\sim 30\%$  lower than in Solimões and Negro (Table S2), a larger change than previously shown (Quay et al., 1995).

Comparison of MS-derived average parameters revealed cumulative alterations of molecular compositions (Tables 1, S3, S4). The ESI[ $\pm$ ] FT-ICR MS-derived average N/C and S/C atomic ratios were higher in S-DOM than in N-DOM. All MS-derived bulk parameters of DOM close to Solimões-Negro mixing zone (MZ) and A-DOM ranged between those of the N-DOM and S-DOM just upstream of the very mixing zone (NaM and SM). The more downstream A-DOM showed declined average CHO%, O/C and  $m/z$  ratios compared to MZ.

$^1\text{H}$  NMR derived section integrals showed higher content of aromatic ( $\text{C}_{ar}\text{H}$ ;  $\delta_{\text{H}} \sim 7.0\text{--}10.0$  ppm) and olefinic units ( $=\text{CH}$ ;  $\delta_{\text{H}} \sim 5.3\text{--}7.0$  ppm) in N-DOM, while pure aliphatic protons (CCCH units,  $\delta_{\text{H}} \sim 0.5\text{--}1.25$  ppm) were more abundant in S-DOM (Tables 2, S5); direct oxygenated units (OCH;  $\delta_{\text{H}} \sim 3.5\text{--}4.9$  ppm) and remotely oxygenated aliphatic units (OCCH;  $\delta_{\text{H}} \sim 1.9\text{--}3.2$  ppm) showed near equal relative abundance in the

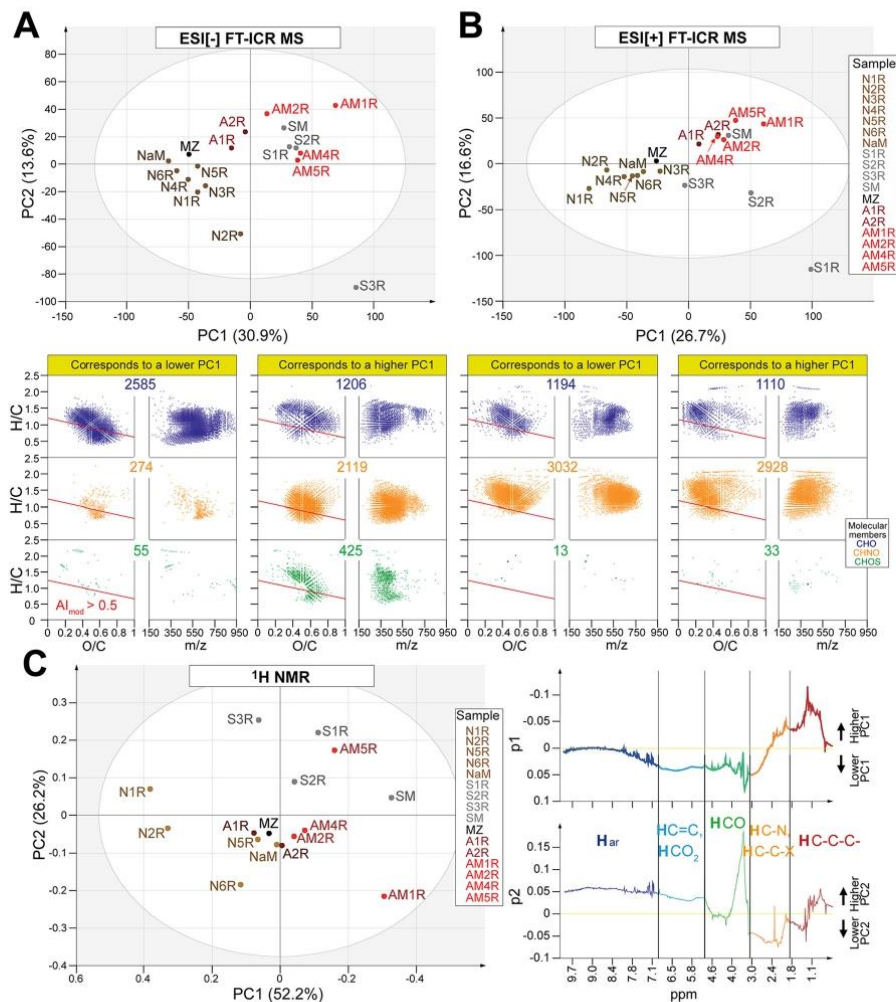
N-DOM and S-DOM just upstream of Solimões-Negro mixing zone.  $^1\text{H}$  NMR section integrals of key substructures in MZ, A1R, and A2R were more similar to N-DOM, while that of more downstream A-DOM were closer to S-DOM.  $^1\text{H}$  NMR section integrals of aromatic, olefinic, and oxygenated protons showed declining trend from A1R to AM1R, in line with our FT-ICR MS results showing declining average O/C, DBE/C, and  $m/z$  ratios, whereas H/C ratios and relative abundance of the total aliphatics ( $\text{OCCCH}_2$ ,  $(\text{CH}_2)_n$ ,  $\text{CCCH}_3$ ;  $\delta_{\text{H}} \sim 0.5\text{--}1.9$  ppm) increased in this reach (Tables S3–S5), suggesting increase of aliphaticity.

We performed unsupervised multivariate statistics on ESI[ $\pm$ ] FT-ICR MS and  $^1\text{H}$  NMR derived compositional and structural features to analyze the alteration of the DOM composition upstream and downstream of Solimões-Negro confluence. In both ESI[ $\pm$ ] FT-ICR MS-based PCA, the first principal components (PC1) showed clear separation as follows: N-DOM  $\sim$  MZ  $<$  A1R/A2R  $<$  S-DOM  $\sim$  more downstream A-DOM  $<$  AM1R (Fig. 3A, B). Van Krevelen and mass-edited H/C diagrams illustrate PCA loadings and show key molecular features differentiating AZ-DOM. The molecular composition with positive/negative loading vector corresponds to higher/lower

**Table 2**

<sup>1</sup>H NMR section integrals (percent of non-exchangeable protons) and key substructures of AZ-DOM (CD<sub>3</sub>OD, exclusion of residual water, and methanol). This table shows value ranges for DOM in Negro, Solimões, Amazon, and Tapajós main stem, respectively; values for each sample see Table S5. The fundamental substructures include aromatics C<sub>ar</sub>H,  $\delta_{\text{H}} \sim 7.0\text{--}10.0$  ppm; olefins = CH, and O<sub>2</sub>CH units,  $\delta_{\text{H}} \sim 5.3\text{--}7.0$  ppm; oxygenated aliphatics OCH units,  $\delta_{\text{H}} \sim 3.2\text{--}4.9$  ppm; "acetate-analogue" and CRAM,  $\delta_{\text{H}} \sim 1.9\text{--}3.2$  ppm; functionalized aliphatics,  $\delta_{\text{H}} \sim 1.35\text{--}1.9$  ppm; polyethylene group,  $\delta_{\text{H}} \sim 1.25\text{--}1.35$  ppm; pure aliphatics,  $\delta_{\text{H}} \sim 0.5\text{--}1.25$  ppm.

$\delta$ ( <sup>1</sup> H) [ppm]	10.0-7.0	7.0-5.3	4.9-3.2	3.2-1.9	1.9-1.35	1.35-1.25	1.25-0.5	1.9-0.5
Key substructures	C <sub>ar</sub> H (%)	HC=C, HCO <sub>2</sub> (%)	HCO (%)	HCCO (%)	HC-C-O (%)	(CH <sub>2</sub> ) <sub>n</sub> (%)	HCCC (%)	Total aliphatic section (%)
Negro	5.0-5.7	5.7-6.9	29.9-33.8	27.1-28.0	13.1-14.6	3.6-4.2	9.8-12.2	27.0-30.4
Solimões	4.6-5.6	4.3-5.6	31.1-33.1	25.7-27.1	14.1-14.8	3.9-4.7	12.2-14.0	30.4-33.3
Amazon	4.5-5.3	3.4-6.0	28.2-32.8	26.5-29.1	14.1-15.9	4.0-4.8	11.5-14.7	29.7-34.2
Tapajós	4.8-4.9	4.0-6.7	27.3-32.8	26.0-28.1	14.2-15.2	4.7-5.4	11.9-15.5	30.8-35.1
NaM	5.5	5.7	30.2	27.6	14.6	4.2	12.2	31.0
SM	4.9	4.0	29.5	27.2	15.1	5.0	14.3	34.4
MZ	5.1	5.8	30.9	27.5	14.7	4.4	11.5	30.6



**Fig. 3.** PCA of AZ-DOM upstream/downstream of the Solimões-Negro mixing zone based on all assigned molecular formulae in ESI[±] FT-ICR MS and <sup>1</sup>H NMR spectra. The upper panels show PCA scatter plots based on A ESI[-] and B ESI[+] FT-ICR MS, respectively. The lower panels of A and B show the van Krevelen and mass-edited H/C diagrams corresponding to lower/higher PC1, with the color code representing elemental composition (i.e., CHO (blue), CHNO (orange), and CHOS (green)). Molecular compositions positioned below the red lines in the van Krevelen diagrams have  $A_{\text{mod}} < 0.5$  (Koch and Dittmar, 2006). Numbers show counts of assigned formulae. The right panel of C shows loading vectors p1/p2 of PC1/PC2, with fundamental substructures within the <sup>1</sup>H NMR spectra indicated.



principal components, respectively. ESI[±] FT-ICR MS-based PCA showed large-scale loss of high molecular weight, more oxygenated and more unsaturated CHO and CHNO DOM molecules at Solimões-Negro confluence. A large proportion of attenuated CHO compounds had  $Al_{mod} > 0.5$  (Koch and Dittmar, 2006), which is an indicator of aromatic structures. The declined CHNO compounds had lower average  $Al_{mod}$  than the CHO compounds, but were highly oxygenated and ionized in both ESI[±] modes. Highly oxygenated CHOS compounds were more abundant in S-DOM/A-DOM than in N-DOM and covered large areas in van Krevelen diagrams, indicative of considerable molecular diversity. For the interpretation of other principal components see Fig. S4. Moreover, HCA separated AZ-DOM alike PCA (Fig. S5). A1R and MZ clustered together with N-DOM in ESI[-] due to higher abundance of polyphenolic CHO-compounds, whereas MZ clustered together with N-DOM in ESI[+] because of more oxygen-rich CHNO compounds ( $m/z$  500–750). A-DOM and S-DOM clustered together with higher content of oxygenated CHO/CHOS compounds and less oxygenated CHNO compounds ( $m/z$  300–550).

<sup>1</sup>H NMR based-PCA showed the structural differences of AZ-DOM proximate to Solimões-Negro mixing zone that were elucidated by NMR loading plots (Fig. 3C). N-DOM showed lower PC1 because of more protons bound to sp<sup>2</sup>-hybridized carbon, as well as OCH units, at the expense of lower relative abundance in pure and remotely oxygenated substituted aliphatic units. Moreover, S-DOM, N1R, and AM5R (sampled just downstream of the Uatumã River inflow; black water) showed higher second principal components (PC2) due to higher content of polycarboxylated aromatic compounds or certain heterocyclic structures, as well as OCH units. AM5R was separated in <sup>1</sup>H NMR based-PCA (Fig. 3C) but not in FT-ICR MS-based PCA (Fig. 3A, B), possibly because the abundant aromatic or heterocyclic structures in AM5R contained few ionizable substituents.

### 3.3. Madeira inflow supplies oxygen-poor CHO molecules and heteroatom-containing compounds to the Amazon River

The Madeira River is an important tributary that supplies a large amount of sediment to the Amazon River. A-DOM just downstream of Madeira (AM1R) showed the highest content of CHNO/CHOS compounds and the lowest average  $m/z$  in A-DOM (Tables S3–S4), and was separated in ESI[±] FT-ICR MS-based PCA (Fig. 3A, B). Furthermore, AM1R was separated in ESI[-] FT-ICR MS-based HCA (Fig. S6-a4) due to higher abundance in oxygenated CHO/CHNO/CHOS compounds ( $m/z$  250–550), and in ESI[+] FT-ICR MS-based HCA (Fig. S6-b3) due to higher abundance in less oxygenated and more saturated CHO/CHNO compounds ( $m/z$  250–550). Moreover, AM4R (sampled out of the river main stem where water overflowed the bank) showed distinction due to high content of unsaturated CHO/CHNO compounds, which likely originated from adjacent floodplains. In addition, A1R and A2R clustered together due to higher abundance of polyphenolic and some oxygenated aliphatic CHO compounds ( $m/z$  400–900) (Fig. S6-a1), and some oxygenated CHNO compounds ( $m/z$  400–750) (Fig. S6-b1).

AM1R showed higher <sup>1</sup>H NMR section integrals of pure and functionalized aliphatic compounds but lower integrals of aromatic, olefinic, and OCH units, compared to other A-DOM (Table S5, Fig. S7). These molecular features caused the distinction of AM1R in NMR-derived PCA (Fig. 3C), opposite to N-DOM. Moreover, AM4R showed higher abundance of OCH units than the other upstream A-DOM (Table S5, Fig. S7).

### 3.4. Changes of DOM composition along the Amazon River

The  $m/z$  ions of each A-DOM in ESI[-] or ESI[+] FT-ICR MS showed similar patterns in van Krevelen and mass-edited H/C diagrams alike A1R (Fig. 2), ~70 % of which were common in all A-DOM (Fig. S8-a3, b3). ESI[±] FT-ICR MS-derived bulk parameters and PCA indicated non-linear trends of DOM composition along the Amazon River (Tables S3–S4, Fig. 4A, B). AM1R was outside the 95 % confidence interval of ESI[-] MS-based PCA (Fig. 4A). AM2R was more similar to A-DOM just downstream of Solimões-Negro confluence (Fig. 4A, B). Moreover, AM12R,

AM3Ra, and AM9Ra showed higher PC1 with higher abundance in polyphenolic and highly oxygenated CHO compounds (Fig. 4A), as well as higher abundance of less oxygenated CHO/CHNO compounds (Fig. 4B). For the interpretation of PC2 see Fig. S9.

<sup>1</sup>H NMR section integrals of A-DOM showed non-continuous trends as well, except that the relative abundance of purely aliphatic CCCH units increased from A1R to AM1R and from AM6R to AM10R (Table S5). A1R and A2R showed higher PC1 in <sup>1</sup>H NMR-based PCA (Fig. 4C), which refers to more abundant functionalized units including alkylated and oxygen-rich aromatic molecules, conjugated double bonds, and various oxygen-rich aliphatics. A-DOM from AM6R to AM10R showed declining PC1 in <sup>1</sup>H NMR-based PCA (Fig. 4C) that were not observed in ESI[±] FT-ICR MS-based PCA (Fig. 4A, B). Furthermore, PC2 separated AM1R from due to higher content of remotely oxygenated and branched aliphatic protons.

Spearman correlation analysis was performed on MS-derived molecular compositions of A-DOM from AM1R to AM10R (Fig. S9). Molecular signatures showing positive/negative correlation with distance from AM1R had similar  $m/z$  but distinct O/C ratios. In both ESI[-] and ESI[+] MS-based analyses, less oxygenated CHO/CHO compounds became depleted downstream (Fig. S10-a2, b2) whereas more oxygenated CHO/CHNO compounds accumulated downstream (Fig. S10-a2, b2).

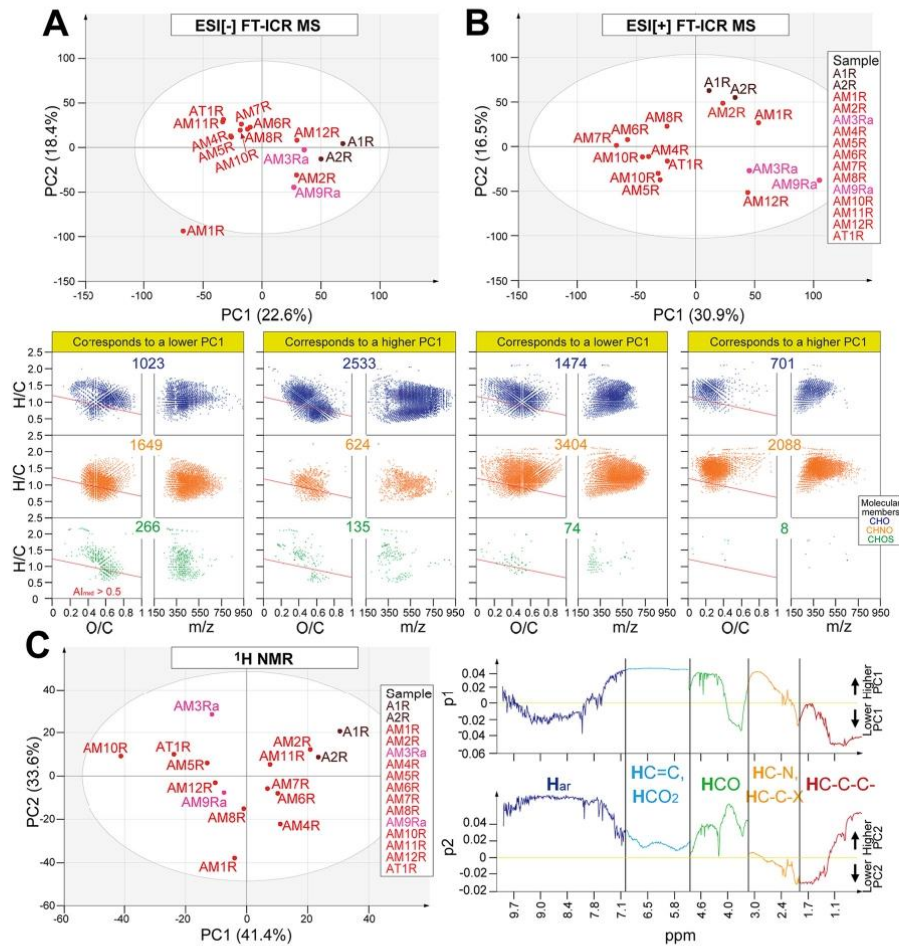
### 3.5. The comparison of S/N/A/T-DOM

The common mass peaks in S/N/A/T-DOM showed similar patterns (Fig. S8), with CHO<sub>14-15</sub>, CHNO<sub>8-9</sub>, and CHO<sub>8-10</sub>S compounds being most abundant (Fig. S11), indicating that AZ-DOM molecules were overall higher oxygenated compared to other riverine DOM (Bae et al., 2011). S-DOM contained ~600 oxygenated CHNO/CHOS compounds not observed in N-DOM (Fig. S12A), while N-DOM contained CHO and CHNO compounds with  $m/z > 450$  not observed in S-DOM (Fig. S12B). Moreover, 124  $m/z$  ions unique to MZ were not found in S-DOM/N-DOM, most of which were CHNO compounds of higher aliphaticity (Fig. S12C). The counts of molecular formulae in each T-DOM in ESI[-] MS were approximately 38 %/27 %/18 % higher than in S/N/A-DOM, respectively. Molecular formulae unique to T-DOM represented less oxygenated CHO/CHNO compounds in a wide  $m/z$  range 300–800 (Fig. S13A), and were positioned in van Krevelen space like abundant mass peaks in exudates from phytoplankton (Medeiros et al., 2015). Meanwhile, counts of molecular formulae in T-DOM from ESI[+] MS were ~20 % lower than in S/N/A-DOM. The ions absent in T-DOM were attributed to CHNO compounds with  $m/z > 600$  (Fig. S13B). Overall, S-DOM had highest content of S-containing compounds and T-DOM had highest content of N-containing compounds (Tables 1, S3, S4). Additionally, T-DOM had lowest average  $m/z$  and O/C, and highest average H/C ratios, consistent with previous studies (Seidel et al., 2016), and indicative of earlier stages of DOM processing.

<sup>1</sup>H NMR spectra of N-DOM were distinct and showed higher levels of oxygenation in all aliphatic and aromatic substructures (Fig. S3). <sup>1</sup>H NMR spectra of S-DOM and T-DOM showed lower proportions of aromatic units and higher content of aliphatic units than A-DOM and N-DOM (Fig. S3).

T-DOM was distinct in ESI[-] FT-ICR MS-based PCA by PC1 due to higher content of more saturated, less oxygenated CHO/CHOS compounds and highly unsaturated CHNO/CHOS compounds (Fig. S14A). However, the distinction of T-DOM did not appear in ESI[+] FT-ICR MS-based PCA (Fig. S14B). Additionally, upstream T-DOM and upstream S-DOM were separated out by PC2 in ESI[-] and ESI[+] FT-ICR MS-based PCA, respectively, due to higher content of less oxygenated compounds (Fig. S15).

<sup>1</sup>H NMR-based PCA (Fig. S16A) showed similar separation in PC1 direction alike ESI[±] FT-ICR MS-based PCA. The <sup>1</sup>H NMR loading spectra showed that the main differences between S/N/A/T-DOM resided in the overall unsaturation and oxygenation, covering the entire aromatic and aliphatic units (Fig. S16B). T4R (very close to Amazon) was outside the 95 % confidence interval because of its huge <sup>1</sup>H NMR resonances representing purely aliphatic protons.



**Fig. 4.** PCA of A-DOM based on all assigned molecular formulae in ESI[±] FT-ICR MS and <sup>1</sup>H NMR spectra. The upper panels show the distribution of A-DOM in PCA scatter plots based on A ESI[-] and B ESI[+] FT-ICR MS, respectively. The lower panels of A and B shows the van Krevelen and mass-edited H/C diagrams corresponding to lower/higher PC1, with the color code representing elemental composition (i.e., CHO (blue), CHNO (orange), and CHOS (green)). Molecular compositions positioned below the red lines in the van Krevelen diagrams have  $A_{I_{mod}} < 0.5$  (Koch and Dittmar, 2006). Numbers show counts of assigned formulae. The right panel of C shows loading vectors p1/p2 of PC1/PC2, with fundamental substructures within the NMR spectra indicated.

### 3.6. Factors regulating AZ-DOM composition

Microbial metabolic processes are suitable biological parameters to understand microbial roles on organic matter decomposition and ecosystem functions (Kamjunke et al., 2015). For instance, HBP refers to assimilatory metabolism consuming preferentially labile DOM like proteins (Seidel et al., 2015). DCF is inorganic light independent C-uptake, mostly performed by chemosynthetic microorganisms, using redox reactions energetic yields to recycle inorganic carbon (Santoro et al., 2013). DCF in A1R was ~15 times higher than in sites upstream of this very mixing zone and A2R (Table S6). Moreover, HBP in A2R was ~6 times of A1R and was >20 times compared to in Negro and Solimões (Table S6). The rate of DCF was ~4 times higher than HBP in A1R, whereas was only 4% of HBP in A2R. These results suggested that the incorporation of inorganic carbon and labile organic carbon could be highly variable and DCF may play an unexpectedly important role for microorganisms at Solimões-Negro mixing zone.

We performed ESI[±] FT-ICR MS and <sup>1</sup>H NMR-based NMDS analysis to elucidate the factors influencing the molecular composition of AZ-DOM (Fig. S17; Table, S7). DOC, POC, DIC, and conductivity differentiated S-DOM and N-DOM, showing that the distinct water characteristics of the two end-member rivers likely result in its SPE-DOM composition. In addition, DOC concentration and pH were the major factors differentiating upstream A-DOM (A1R and A2R) from downstream A-DOM in ESI[-] MS-based NMDS, whereas the other parameters did not show significant correlations (Fig. S17, Table S7). However, pH did not separate downstream A-DOM, suggesting that pH did not have strong effects on DOM composition downstream of the Amazon River (Fig. S17). Inorganic nutrient concentrations did not show significant correlations with AZ-DOM composition (Table S7), likely because of their very low abundances (mostly <1 μM; Table S2). Furthermore, we performed the same NMDS method on the samples for which HBP and DCF were measured (Fig. 5, Table S8) and HBP was found to be associated with the separation of T-DOM from other AZ-DOM.



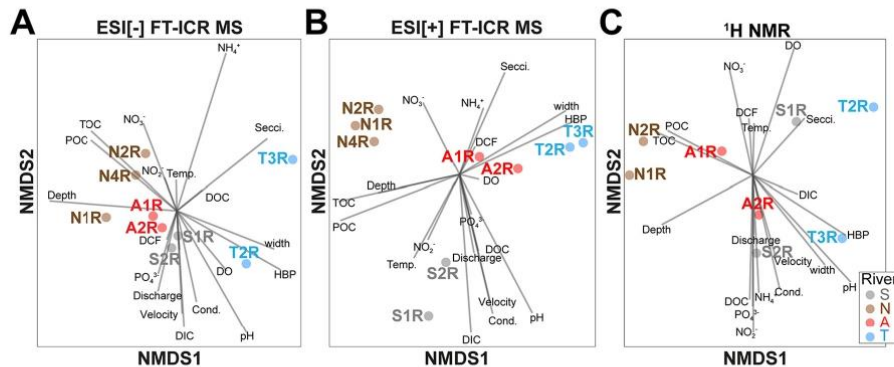


Fig. 5. NMDS analysis using DOM composition in the four Amazon main stem and tributaries based on A, B all assigned molecular formulae in ESI[±] FT-ICR MS, as well as C  $^1\text{H}$  NMR data at 0.01 ppm bucket resolution fitted with geochemical and biological data.

As DOC, POC, pH, and HBP showed more significant impacts on AZ-DOM composition in NMDS results (Tables S7-S8), we further tested their effects on ESI[±] FT-ICR MS-based molecular patterns by Spearman correlation analysis (Fig. S18). Higher POC was significantly correlated with higher abundance of polyphenolic CHO compounds and highly unsaturated CHNO compounds. In addition, pH showed positive correlation with compositions of smaller molecular weight ( $m/z$  200–400), less oxygenated CHO/CHNO and highly oxygenated CHNO/CHOS compounds. The molecular composition that correlated with lower pH included polyphenolic CHO compounds and highly unsaturated oxygen-rich CHO/CHNO compounds, which have potential to inhibit microbial decomposition (Freeman et al., 2001). Moreover, HBP was positively correlated with more oxygen-poor compounds of lower molecular weight (Fig. S18).

#### 4. Discussion

Solid phase extraction (SPE) with PPL resin (Dittmar et al., 2008) isolates the highest diversity of functionalized small organic molecules from freshwater and marine DOM with appreciable yield (Li et al., 2017). SPE of polydisperse and molecularly heterogeneous aqueous DOM is a dynamic process of sorption, desorption and resorption of individual molecules propagating through the PPL column (Li et al., 2016). We have followed strict guidelines during SPE with PPL resin (Li et al., 2016) and regard ecosystem characteristics of actual DOM molecules as well represented in our samples; slight variance of SPE selectivity may induce very minor but not relevant additional distinction.

Our joint molecular characterization of SPE-DOM in the Amazon main stem and tributaries by ESI[±] FT-ICR MS and  $^1\text{H}$  NMR provided improved resolution of the alterations of DOM along the Amazon River, and demonstrated distinct reactivity for CHO, CHNO, and CHOS molecules. The molecular features of AZ-DOM can serve as biogeochemical tracers and provide compelling information on their source, reactivity, and processing.

A group of more saturated CHO compounds was unique to T-DOM, likely resulting from abundant algal biomass and higher primary production in Tapajós (Moreira-Turcq et al., 2003a), in line with previous work that had in situ primary production of C4 grasses or algal material identified as the key autochthonous source of DOM in Tapajós River (Quay et al., 1992; Ward et al., 2016). Membrane lipids are known products of microbial degradation of phytoplankton and might serve as potential precursors of DOM (Harvey et al., 2006). Phytoplankton blooms provide new molecules relevant to the DOM pool in Amazon plume waters (Medeiros et al., 2015). Several studies have found very low chlorophyll  $a$  values in the white water rivers and high chlorophyll  $a$  in clear waters (Gagne-Maynard et al., 2017; Abril et al., 2014).

High abundance of highly oxygenated and highly unsaturated CHNO compounds, which ionized in both ESI[±] FT-ICR MS was observed in all the S/N/A/T-DOM. Detection of CHNO molecules differs in ESI[−] and ESI[+] MS methods because the former primarily ionizes carboxyl-carrying DOM molecules (like e.g. CRAM, carboxyl-rich alicyclic molecules (Hertkorn et al., 2006)), which are abundant in common freshwater DOM as observed e.g. by  $^{13}\text{C}$  NMR spectra and titration (Ritchie and Perdue, 2003). ESI[+] directly detects nitrogen containing molecules which not necessarily carry carboxylic groups. The complementary of “active” detection of N-containing molecules by ESI[+] and “passive” observation in ESI[−] MS allows compositional and structural distinction related to origin and processing of DOM. Hence, CHNO molecular compositions observed in both ESI[±] might originate from different (distributions of) isomers of identical compositions, with some probable emphasis on more recalcitrant nitrogen-containing CRAM. Moreover, the abundance of less oxygenated CHNO compounds that mostly ionized in ESI[+] FT-ICR MS was highest in AMIR and then declined rapidly; these more aliphatic compounds likely had a shorter residence time, in agreement with higher general biogeochemical processing of heteroatom-containing molecules (Ksionzek et al., 2016).

The higher content of S-containing compounds in S-DOM and AMIR may have resulted from incorporation of sulfur into organic matter by dissimilatory sulfate-reducing bacteria and/or by abiotic sulfuration reactions under anoxic conditions in sediment-rich waters like Solimões and Madeira (Luek et al., 2017; Sinnighe Damste and De Leeuw, 1990; Schmidt et al., 2009), but anthropogenic sources cannot be excluded (Latrubesse et al., 2017). Furthermore, DOM sulfuration processes are likely to be more intense under more anaerobic conditions during low-water periods relative to the high-water period we sampled.

Unlike mass spectrometry, which has intrinsic limitations like structure-dependent ionization efficiency (Hertkorn et al., 2008),  $^1\text{H}$  NMR more comprehensively covered all structures containing non-exchangeable hydrogen atoms, although the huge diversity of hydrogen atomic environments in DOM produced a considerable overlap of NMR resonances. N-DOM showed a higher degree of oxygenation relative to S/A/T-DOM in a very extensive range of aliphatic and aromatic carbon chemical environments. These molecular features were transferred to the Amazon River and became strongly attenuated ~150 km downstream. Moreover, A-DOM downstream of the Madeira River inflow and the Uatumã River inflow showed higher relative abundance of pure and functionalized aliphatic protons compared to A-DOM just upstream of these confluences, presumably indicative of higher proportions of more microbially processed terrestrial DOM (Lechtenfeld et al., 2015; Schmidt et al., 2009).

Solimões-Negro confluence ranks among the largest on Earth and provides non-conservative mixing and exemplary spatial and temporal DOM

processing on grand-scale. High DOC with large proportion of hydroxyl and phenolic sites in Negro water (Duarte et al., 2016; Gonsior et al., 2016) and abundant sediment load in Solimões water (Aucour et al., 2003; Moreira-Turcq et al., 2003b; Subdiaga et al., 2019) present favorable conditions for structure-selective sorption of OM on mineral surfaces (Aufdenkampe et al., 2001; Moreira-Turcq et al., 2003b; Subdiaga et al., 2019). Large proportions of polyphenolic compounds in N-DOM were removed from the DOM pool when Negro met Solimões. Many molecular formulae with near average H/C and O/C ratios prevailed in the Amazon River downstream, representing a large number of isomeric molecules that as a whole were resistant to rapid alteration. Continual lateral input of processed terrestrial and autochthonous organic matter from tributaries and floodplains probably contributed to this dynamic equilibrium of DOM synthesis and degradation.

Considerable decrease of DO (~30%) and POC (~90%) concentrations and concomitant increase of HBP and DCF just downstream of Solimões-Negro confluence reflected high microbial activity following mixing, as previously reported for mixing zones in the Amazon basin (Farjalla, 2014). The large depletion of POC at Solimões-Negro mixing zone, in addition to the large number of molecular signatures associated with POC in the Amazon River downstream (Fig. S18) suggest that the DOC-POC transition may have significant effects on DOM composition, especially at confluences. The POC/DOC concentration ratios were ~2 upstream of the Solimões-Negro mixing zone (both Negro and Solimões), while the ratios were ~0.2 downstream of Solimões-Negro mixing zone (Amazon). The loss of POC was probably associated with sedimentation processes, structure selective sorption of DOM molecules to minerals (Subdiaga et al., 2019), favored by a decrease in water velocity and an increase in the depth, as well as OM degradation (Moreira-Turcq et al., 2003b).

However, changes in AZ-DOM composition not only resulted from abiotic processes but were likely associated with considerable bacterial activity, as proved by high HBP rates. Previous studies suggested overall low aquatic bacterial metabolism for the Amazon waters, but confluences could be hot spot areas of high bacterial metabolism, which are associated with changes in organic matter quality (Farjalla, 2014). The less oxygenated, lower molecular weight CHO/CHNO compounds associated with higher HBP might be preferentially assimilated and converted into new bacterial biomass. Furthermore, we found for the first time that DCF contributed to microbial biomass production to the same or even higher extent than heterotrophic processes just downstream of the Solimões-Negro mixing zone, suggesting that the inorganic carbon pathway may be as important as organic carbon degradation when river waters with distinct characteristics converge. In addition, abiotic reactions such as photochemical processing (oxidation and mineralization) and metal complexation could also have important impact on changes of DOM composition in the Amazon River (Amon and Benner, 1996; Aucour et al., 2003).

The inherent complexity of DOM requires multiple analytical dimensions to understand its molecular processing within the carbon cycle. DOM analysis has been conducted by ESI[−] FT-ICR MS with certain preference for selective ionization of polar acidic functional groups (e.g., carboxylic acids), but poor propensity to ionize CHNO compounds lacking carboxyl and/or hydroxyl groups. Moreover, CHNO molecules represent important potentially biolabile components of DOM that indicate land cover, microbial processing, and possibly abiotic nitrogen incorporation (Bernhardt and Likens, 2002; Sleighter et al., 2014). The use of both negative and positive ESI modes for DOM research is highly complementary and beneficial, especially for DOM containing higher proportions of N-containing compounds resulting from higher microbial activity, such as in river confluences, highly autochthonous samples, permafrost thaw, and anthropogenically impacted streams (Spencer et al., 2019; Wagner et al., 2015). Furthermore, high-field proton NMR spectra, despite limited resolution because of massive intrinsic averaging, provided highly complementary quantitative DOM structural information, and enabled clear classification of AZ-DOM. Oxygenated SPE-DOM molecules were more efficiently ionized in ESI[−] FT-ICR MS, unsaturated and nitrogen-containing compounds were better ionized in ESI[+] FT-ICR MS, whereas 1D NMR showed better coverage of purely aliphatic structures not resolvable in mass spectra.

## 5. Conclusion

Complementary ESI[±] mass spectrometry and NMR spectroscopy provided comprehensive coverage of molecular evolution of DOM for exemplary processes in the Amazon basin, and distinct reactivity for CHO, CHNO and CHOS molecules.

So far elusive structural features of CHNO compounds in Amazon DOM, and DOM in other ecosystems/biomes, will become better constrained by comparative analysis of CHO and CHNO compounds in ESI[+] and ESI[−] mass spectra for individual samples and processes.

Joint ESI[±] mass spectrometry should become commonplace for further studies of DOM processing in ecosystems because strong ionization selectivity in ESI[−] mass spectra provides conceptual assessments with serious intrinsic limitations. NMR spectra offer valuable constraints of hydrogen and carbon atomic environments not available by FT mass spectra and are important for providing structural information and obtaining credible conclusions.

## CRedit authorship contribution statement

DB, AEP, MG, PS and NH designed research. JV, FMS, MH, DB, PS, MG, AEP, and NH participated in sampling. SL, JV, MH and NH prepared samples. SL and MH acquired FT-ICR MS spectra. SL and NH acquired NMR spectra. AEP, DB, and NH obtained funding. Data interpretation was performed by all authors. SL, NH, and MH actively participated in writing the manuscript. All authors provided significant contributions to the final manuscript.

## Data availability

Data will be made available on request.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This work was supported by the Alexander von Humboldt Foundation (NH and AEP; Research Linkage Program: Connecting the diversity of DOM and CO<sub>2</sub> and CH<sub>4</sub> production in tropical lakes); the Swedish Research Council for Sustainable Development, FORMAS, grants no. 2013-01077 and 2021-0242, the Swedish Research Council (VR; grant no. 2012-00048), the European Research Council ERC (DB; grant no. 725546), the China Scholarship Council (LS; grant no. 201806360268) Molecular Level Characterization of Natural Organic Matter in Tropical Freshwater and Marine Ecosystems of Worldwide Relevance: Amazon Basin, Pantanal, Red Sea, and several grants from Brazilian Foundations CAPES (Probal), FAPERJ (Cientista do Nosso Estado) and CNPq (Universal and Ciencias sem Fronteiras) to AEP.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at <https://doi.org/10.1016/j.scitotenv.2022.159620>. These data include the Google map of the most important areas described in this article.

## References

- Abril, G., Martinez, J.M., Artigas, L.F., Moreira-Turcq, P., Benedetti, M.F., Vidal, L., Meziane, T., Kim, J.H., Bernardes, M.C., Savoye, N., Deborde, J., Souza, E.L., Alberic, P., Landim de Souza, M.F., Roland, F., 2014. Amazon River carbon dioxide outgassing fuelled by wetlands. *Nature* 505 (7483), 395–398.



- Amado, A.M., Farjalla, V.F., Esteves Fde, A., Bozelli, R.L., Roland, F., Enrich-Prast, A., 2006. Complementary pathways of dissolved organic carbon removal pathways in clear-water Amazonian ecosystems: photochemical degradation and bacterial uptake. *FEMS Microbiol. Ecol.* 56 (1), 8–17.
- Amon, R.M.W., Benner, R., 1996. Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River system. *Geochim. Cosmochim. Acta* 60 (10), 1783–1792.
- Armanious, A., Aeppli, M., Sander, M., 2014. Dissolved organic matter adsorption to model surfaces: adlayer formation, properties, and dynamics at the nanoscale. *Environ. Sci. Technol.* 48 (16), 9420–9429.
- Aucour, A.M., Tao, F.X., Moreira-Turcq, P., Seyler, P., Sheppard, S., Benedetti, M.F., 2003. The Amazon River: behaviour of metals (Fe, Al, Mn) and dissolved organic matter in the initial mixing at the Rio Negro/Solimões confluence. *Chem. Geol.* 197 (1–4), 271–285.
- Aufdenkampe, A.K., Hedges, J.I., Richey, J.E., Krusche, A.V., Llerena, C.A., 2001. Sorptive fractionation of dissolved organic nitrogen and amino acids onto fine sediments within the Amazon Basin. *Limnol. Oceanogr.* 46 (8), 1921–1935.
- Bae, E., Yeo, J.J., Jeong, B., Shin, Y., Shin, K.H., Kim, S., 2011. Study of double bond equivalents and the numbers of carbon and oxygen atom distribution of dissolved organic matter with negative-mode FT-ICR MS. *Anal. Chem.* 83 (11), 4193–4199.
- Battin, T.J., Luysaert, S., Kaplan, L.A., Aufdenkampe, A.K., Richter, A., Tranvik, L.J., 2009. The boundless carbon cycle. *Nat. Geosci.* 2 (9), 598–600.
- Benner, R., Opsahl, S., Chin-Leo, G., Richey, J.E., Forsberg, B.R., 1995. Bacterial carbon metabolism in the Amazon River system. *Limnol. Oceanogr.* 40 (7), 1262–1270.
- Bernhardt, E.S., Likens, G.E., 2002. Dissolved organic carbon enrichment alters nitrogen dynamics in a forest stream. *Ecology* 83 (6), 1689–1700.
- Bianchi, T.S., Ward, N.D., 2019. Editorial: the role of priming in terrestrial and aquatic ecosystems. *Front. Earth Sci.* 7, 321.
- Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., Duarte, C.M., Kortelainen, P., Downing, J.A., Middelburg, J.J., Melack, J., 2007. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* 10 (1), 172–185.
- Dittmar, T., Koch, B., Hertkorn, N., Kattner, G., 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr.* Methods 6 (6), 230–235.
- do Nascimento, D.R., Sawakuchi, A.O., Guedes, C.C.F., Giannini, P.C.F., Grohmann, C.H., Ferreira, M.P., 2015. Occurrence of sands from the confluence of the Amazon and Madeira rivers based on detrital heavy minerals and luminescence of quartz and feldspar. *Sediment. Geol.* 316, 1–12.
- Drake, T.W., Hemingway, J.D., Kurek, M.R., Peucker-Ehrenbrink, B., Brown, K.A., Holmes, R.M., Galy, V., Moura, J.M.S., Mitsuya, M., Wassenaar, L.I., Six, J., Spencer, R.G.M., 2021. The pulse of the Amazon: fluxes of dissolved organic carbon, nutrients, and ions from the world's largest river. *Glob. Biogeochem. Cycles* 35 (4), e2020GB006895.
- Duarte, R.M., Smith, D.S., Val, A.L., Wood, C.M., 2016. Dissolved organic carbon from the upper Rio Negro protects zebrafish (*Danio rerio*) against ionoregulatory disturbances caused by low pH exposure. *Sci. Rep.* 6 (1), 20377.
- Ertel, J.R., Hedges, J.I., Devol, A.H., Richey, J.E., Ribeiro, M.D.N.G., 1986. Dissolved humic substances of the Amazon River system I. *Limnol. Oceanogr.* 31 (4), 739–754.
- Farjalla, V.F., 2014. Are the mixing zones between aquatic ecosystems hot spots of bacterial production in the Amazon River system? *Hydrobiologia* 728 (1), 153–165.
- Freeman, C., Ostle, N., Kang, H., 2001. An enzymic 'latch' on a global carbon store. *Nature* 409 (6817), 149.
- Gagne-Maynard, W.C., Ward, N.D., Keil, R.G., Sawakuchi, H.O., Da Cunha, A.C., Neu, V., Brito, D.C., Da Silva Less, D.F., Diniz, J.E.M., De Matos Valerio, A., Kampel, M., Krusche, A.V., Richey, J.E., 2017. Evaluation of primary production in the lower Amazon River based on a dissolved oxygen stable isotopic mass balance. *Front. Mar. Sci.* 4, 1–12.
- Gonsior, M., Valle, J., Schmitt-Kopplin, P., Hertkorn, N., Bastviken, D., Luek, J., Harir, M., Bastos, W., Enrich-Prast, A., 2016. Chemodiversity of dissolved organic matter in the Amazon Basin. *Biogeochemistry* 13 (1), 4279–4290.
- Harvey, H.R., Dyda, R.Y., Kirchman, D.L., 2006. Impact of DOM composition on bacterial lipids and community structure in estuaries. *Aquat. Microb. Ecol.* 42 (2), 105–117.
- Hedges, J.I., Ertel, J.R., Quay, P.D., Grootes, P.M., Richey, J.E., Devol, A.H., Farwell, G.W., Schmidt, F.W., Salati, E., 1986. Organic carbon-14 in the Amazon river system. *Science* 231 (4742), 1129–1131.
- Hedges, J.I., Hatcher, P.G., Ertel, J.R., Meyers-Schulte, K.J., 1992. A comparison of dissolved humic substances from seawater with Amazon River counterparts by <sup>13</sup>C-NMR spectrometry. *Geochim. Cosmochim. Acta* 56 (4), 1753–1757.
- Hedges, J.I., Cowie, G.L., Richey, J.E., Quay, P.D., Benner, R., Strom, M., Forsberg, B.R., 1994. Origins and processing of organic matter in the Amazon River as indicated by carbohydrates and amino acids. *Limnol. Oceanogr.* 39 (4), 743–761.
- Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., Ketrup, A., Hedges, J.I., 2006. Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* 70 (12), 2990–3010.
- Hertkorn, N., Ruecker, C., Meringer, M., Gugisch, R., Frommberger, M., Perdue, E.M., Witt, M., Schmitt-Kopplin, P., 2007. High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Anal. Bioanal. Chem.* 389 (5), 1311–1327.
- Hertkorn, N., Frommberger, M., Witt, M., Koch, B.P., Schmitt-Kopplin, P., Perdue, E.M., 2008. Natural organic matter and the event horizon of mass spectrometry. *Anal. Chem.* 80 (23), 8908–8919.
- Hertkorn, N., Harir, M., Koch, B.P., Michalke, B., Schmitt-Kopplin, P., 2013. High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeochemistry* 103 (3), 1583–1624.
- Hertkorn, N., Harir, M., Cawley, K.M., Schmitt-Kopplin, P., Jaffé, R., 2016. Molecular characterization of dissolved organic matter from subtropical wetlands: a comparative study through the analysis of optical properties, NMR and FTICR/MS. *Biogeochemistry* 13 (8), 2257–2277.
- Kamjunke, N., Herzsprung, P., Neu, T.R., 2015. Quality of dissolved organic matter affects planktonic but not biofilm bacterial production in streams. *Sci. Total Environ.* 506–507, 353–360.
- Koch, B.P., Dittmar, T., 2006. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 20 (5), 926–932.
- Kšionzek, K.B., Lechtenfeld, O.J., McCallister, S.L., Schmitt-Kopplin, P., Geuer, J.K., Geibert, W., Koch, B.P., 2016. Dissolved organic sulfur in the ocean: biogeochemistry of a petagram inventory. *Science* 354 (6311), 456–459.
- Latrubesse, E.M., Arima, E.Y., Dunne, T., Park, E., Baker, V.R., d'Horta, F.M., Wight, C., Wittmann, F., Zuanon, J., Baker, P.A., Ribas, C.C., Norgaard, R.B., Filizola, N., Ansar, A., Flyvbjerg, B., Stevaux, J.C., 2017. Damming the rivers of the Amazon basin. *Nature* 546 (7658), 363–369.
- Lechtenfeld, O.J., Hertkorn, N., Shen, Y., Witt, M., Benner, R., 2015. Marine sequestration of carbon in bacterial metabolites. *Nat. Commun.* 6 (1), 6711.
- Li, Y., Harir, M., Lucio, M., Kanawati, B., Smirnov, K.S., Koch, B.P., Schmitt-Kopplin, P., Hertkorn, N., 2016. Proposed guidelines for solid phase extraction of Suwannee river dissolved organic matter. *Anal. Chem.* 88 (13), 6680–6688.
- Li, Y., Harir, M., Uhl, J., Kanawati, B., Lucio, M., Smirnov, K.S., Koch, B.P., Schmitt-Kopplin, P., Hertkorn, N., 2017. How representative are dissolved organic matter (DOM) extracts? A comprehensive study of sorbent selectivity for DOM isolation. *Water Res.* 116 (2017), 316–323.
- Luek, J.L., Thompson, K.E., Larsen, R.K., Heyes, A., Gonsior, M., 2017. Sulfate reduction in sediments produces high levels of chromophoric dissolved organic matter. *Sci. Rep.* 7 (1), 8829.
- Lynch, L.M., Sutfin, N.A., Fegol, T.S., Boot, C.M., Covino, T.P., Wallenstein, M.D., 2019. River channel connectivity shifts metabolite composition and dissolved organic matter chemistry. *Nat. Commun.* 10 (1), 459.
- Martinez, J.-M., Espinoza-Villar, R., Armijos, E., Silva Moreira, L., 2015. The optical properties of river and floodplain waters in the Amazon River Basin: implications for satellite-based measurements of suspended particulate matter. *J. Geophys. Res.* 120 (7), 1274–1287.
- Mayorga, E., Aufdenkampe, A.K., Masiello, C.A., Krusche, A.V., Hedges, J.I., Quay, P.D., Richey, J.E., Brown, T.A., 2005. Young organic matter as a source of carbon dioxide outgassing from Amazonian rivers. *Nature* 436 (7050), 538–541.
- McClain, M.E., Naiman, R.J., 2008. Andean influences on the biogeochemistry and ecology of the Amazon River. *BioScience* 58 (4), 325–338.
- Medeiros, P.M., Seidel, M., Ward, N.D., Carpenter, E.J., Gomes, H.R., Niggemann, J., Krusche, A.V., Richey, J.E., Yager, P.L., Dittmar, T., 2015. Fate of the Amazon River dissolved organic matter in the tropical Atlantic Ocean. *Glob. Biogeochem. Cycles* 29 (5), 677–690.
- Moreira-Turcq, P., Seyler, P., Guyot, J.L., Etcheber, H., 2003a. Exportation of organic carbon from the Amazon River and its main tributaries. *Hydro. Process.* 17 (7), 1329–1344.
- Moreira-Turcq, P.F., Seyler, P., Guyot, J.L., Etcheber, H., 2003b. Characteristics of organic matter in the mixing zone of the Rio Negro and Rio Solimões of the Amazon River. *Hydro. Process.* 17 (7), 1393–1404.
- Mortillaro, J.M., Abril, G., Moreira-Turcq, P., Sobrinho, R.L., Perez, M., Meziane, T., 2011. Fatty acid and stable isotope (<sup>813</sup>C, <sup>815</sup>N) signatures of particulate organic matter in the lower Amazon River: seasonal contrasts and connectivity between floodplain lakes and the mainstem. *Org. Geochem.* 42 (10), 1159–1168.
- Pérez, M.A.P., Moreira-Turcq, P., Gallard, H., Allard, T., Benedetti, M.F., 2011. Dissolved organic matter dynamic in the Amazon basin: sorption by mineral surfaces. *Chem. Geol.* 286 (3–4), 158–168.
- Powers, L.C., Hertkorn, N., McDonald, N., Schmitt-Kopplin, P., Del Vecchio, R., Blough, N.V., Gonsior, M., 2019. Sargassum sp. act as a large regional source of marine dissolved organic carbon and polyphenols. *Glob. Biogeochem. Cycles* 33 (11), 1423–1439.
- Quay, P.D., Wilbur, D., Richey, J.E., Hedges, J.I., Devol, A.H., Victoria, R., 1992. Carbon cycling in the Amazon River: implications from the <sup>13</sup>C compositions of particles and solutes. *Limnol. Oceanogr.* 37 (4), 857–871.
- Quay, P.D., Wilbur, D., Richey, J.E., Devol, A.H., Benner, R., Forsberg, B.R., 1995. The <sup>18</sup>O: <sup>16</sup>O of dissolved oxygen in rivers and lakes in the Amazon Basin: determining the ratio of respiration to photosynthesis rates in freshwaters. *Limnol. Oceanogr.* 40 (4), 718–729.
- Richey, J.E., Melack, J.M., Aufdenkampe, A.K., Ballester, V.M., Hess, L.L., 2002. Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric CO<sub>2</sub>. *Nature* 416 (6881), 617–620.
- Ritchie, J.D., Perdue, E.M., 2003. Proton-binding study of standard and reference fulvic acids, humic acids, and natural organic matter. *Geochim. Cosmochim. Acta* 67 (1), 85–96.
- Saliot, A., Mejanelle, L., Scribe, P., Fillaux, J., Pepe, C., Jabaud, A., Dagaut, J., 2001. Particulate organic carbon, sterols, fatty acids and pigments in the Amazon River system. *Biogeochemistry* 53 (1), 79–103.
- Santoro, A.L., Bastviken, D., Gudasz, C., Tranvik, L., Enrich-Prast, A., 2013. Dark carbon fixation: an important process in lake sediments. *PLoS One* 8 (6), e65813.
- Santos-Junior, C.D., Sarmento, H., de Miranda, F.P., Henrique-Silva, F., Logares, R., 2020. Uncovering the genomic potential of the Amazon River microbiome to degrade rainforest organic matter. *Microbiome* 8 (1), 151.
- Schmidt, F., Elvert, M., Koch, B.P., Witt, M., Hinrichs, K.-U., 2009. Molecular characterization of dissolved organic matter in pore water of continental shelf sediments. *Geochim. Cosmochim. Acta* 73 (11), 3337–3358.
- Seidel, M., Yager, P.L., Ward, N.D., Carpenter, E.J., Gomes, H.R., Krusche, A.V., Richey, J.E., Dittmar, T., Medeiros, P.M., 2015. Molecular-level changes of dissolved organic matter along the Amazon River-to-ocean continuum. *Mar. Chem.* 177, 218–231.
- Seidel, M., Dittmar, T., Ward, N.D., Krusche, A.V., Richey, J.E., Yager, P.L., Medeiros, P.M., 2016. Seasonal and spatial variability of dissolved organic matter composition in the lower Amazon River. *Biogeochemistry* 131 (3), 281–302.
- Shen, Y.-H., 1999. Sorption of natural dissolved organic matter on soil. *Chemosphere* 38 (7), 1505–1515.
- Simpson, A.J., McNally, D.J., Simpson, M.J., 2011. NMR spectroscopy in environmental research: from molecular interactions to global processes. *Prog. Nucl. Magn. Reson. Spectrosc.* 58 (3–4), 97–175.



- Sinninghe Damste, J.S., De Leeuw, J.W., 1990. Analysis, structure and geochemical significance of organically-bound sulphur in the geosphere: state of the art and future research. *Org. Geochem.* 16 (4–6), 1077–1101.
- Sleighter, R.L., Chin, Y.-P., Arnold, W.A., Hatcher, P.G., McCabe, A.J., McAdams, B.C., Wallace, G.C., 2014. Evidence of incorporation of abiotic S and N into prairie wetland dissolved organic matter. *Environ. Sci. Technol. Lett.* 1 (9), 345–350.
- Spencer, R.G., Kellerman, A.M., Podgorski, D.C., Macedo, M.N., Jankowski, K., Nunes, D., Neill, C., 2019. Identifying the molecular signatures of agricultural expansion in Amazonian headwater streams. *J. Geophys. Res. Biogeosci.* 124 (6), 1637–1650.
- Subdiaga, E., Orsetti, S., Haderlein, S.B., 2019. Effects of sorption on redox properties of natural organic matter. *Environ. Sci. Technol.* 53 (24), 14319–14328.
- Tziotis, D., Hertkorn, N., Schmitt-Kopplin, P., 2011. Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *Eur. J. Mass Spectrom.* (Chichester) 17 (4), 415–421.
- Wagner, S., Riedel, T., Niggemann, J., Vähätalo, A.V., Dittmar, T., Jaffé, R., 2015. Linking the molecular signature of heteroatomic dissolved organic matter to watershed characteristics in world rivers. *Environ. Sci. Technol.* 49 (23), 13798–13806.
- Ward, N.D., Bianchi, T.S., Sawakuchi, H.O., Gagne-Maynard, W., Cunha, A.C., Brito, D.C., Neu, V., de Matos Valerio, A., da Silva, R., Krusche, A.V., Richey, J.E., Keil, R.G., 2016. The reactivity of plant-derived organic matter and the potential importance of priming effects along the lower Amazon River. *J. Geophys. Res. Biogeosci.* 121 (6), 1522–1539.

Hello Siyu Li,

Thank you for your email to Copyright Clearance Center (CCC) and congratulations on your thesis permission. My name is Christine in the Customer Service Department and I am happy to assist you. The good news is that the publisher Elsevier, uses the CCC's software option, a widget called RightsLink, for self-service permissions. As you are the author hoping to reuse your own work for your thesis, no formal permission is required per Elsevier. Please see the details below to follow the steps so you can also see the messaging in RightsLink:

**Getting Started:**

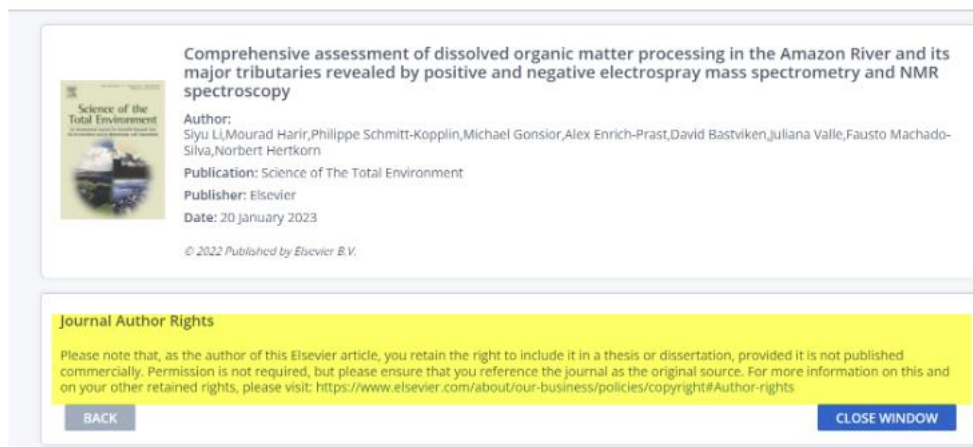
Step 1: Please visit the article abstract on the publisher's website:

<https://www.sciencedirect.com/science/article/pii/S0048969722067201>

Step 2: From the abstract, click 'Get Rights and Content' to open the RightsLink permission page.

Step 3: Please choose your type of use as well as indicating you are the author of the paper.

Step 4: You may obtain a price quote in advance by using the "Quick Price" button. This is the message that you will see:



The screenshot shows a RightsLink interface for an Elsevier article. At the top, there is a thumbnail of the journal cover for 'Science of the Total Environment' and the article title: 'Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy'. Below the title, the author list is provided: 'Siyu Li, Mourad Harir, Philippe Schmitt-Kopplin, Michael Gonsior, Alex Enrich-Prast, David Bastviken, Juliana Valle, Fausto Machado-Silva, Norbert Hertkorn'. The publication information is 'Science of The Total Environment', published by Elsevier on 20 January 2023. A copyright notice at the bottom of the article section reads '© 2022 Published by Elsevier B.V.'. Below the article information is a yellow box titled 'Journal Author Rights' containing the text: 'Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>'. At the bottom of the yellow box are two buttons: 'BACK' and 'CLOSE WINDOW'.

\* Here is the link regarding Author Rights:

For more information on this and on your other retained rights, please visit: <https://www.elsevier.com/about/our-business/policies/copyright#Author-rights>

If you have any further questions please don't hesitate to contact a Customer Account Specialist via email, chat, or by calling 855-239-3415 24 hours a day, seven days a week.

Best,  
Christine Z.  
Customer Account Specialist  
Copyright Clearance Center  
222 Rosewood Drive  
Danvers, MA 01923  
[www.copyright.com](http://www.copyright.com)  
Toll-Free US +1.855.239.3415  
International +1.978-646-2600

## 6.2 Article 2 + licenses for copyrighted content

# Distinct Non-conservative Behavior of Dissolved Organic Matter after Mixing Solimões/Negro and Amazon/Tapajós River Waters

Siyu Li, Mourad Harir, Philippe Schmitt-Kopplin, Fausto Machado-Silva, Michael Gonsior, David Bastviken, Alex Enrich-Prast, Juliana Valle, and Norbert Hertkorn\*

Cite This: *ACS EST Water* 2023, 3, 2083–2095

Read Online

ACCESS |

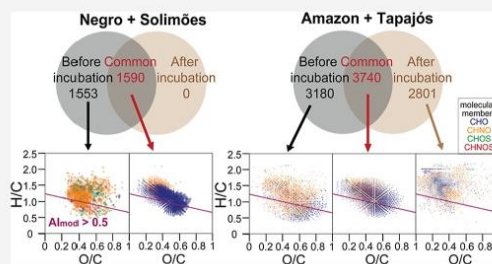
Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Positive and negative electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry and  $^1\text{H}$  NMR revealed major compositional and structural changes of dissolved organic matter (DOM) after mixing two sets of river waters in Amazon confluences: the Solimões and Negro Rivers (S + N) and the Amazon and Tapajós Rivers (A + T). We also studied the effects of water mixing ratios and incubation time on the composition and structure of DOM molecules. NMR spectra demonstrated large-scale structural transformations in the case of S + N mixing, with gain of pure and functionalized aliphatic units and loss of all other structures after 1d incubation. A + T mixing resulted in comparatively minor structural alterations, with a major gain of small aliphatic biomolecular binding motifs. Remarkably, structural alterations from mixing to 1d incubation were in essence reversed from 1d to 5d incubation for both S + N and A + T mixing experiments. Heterotrophic bacterial production (HBP) in endmembers S, N, and S + N mixtures remained near  $0.03 \mu\text{gC L}^{-1} \text{h}^{-1}$ , whereas HBP in A, T, and A + T were about five times higher. High rates of dark carbon fixation took place at S + N mixing in particular. In-depth biogeochemical characterization revealed major distinctions between DOM biogeochemical changes and temporal evolution at these key confluence sites within the Amazon basin.

**KEYWORDS:** FT-ICR MS, NMR, mixing zones, transport, transformation, decomposition, biogeochemical cycling



## INTRODUCTION

The Amazon basin is the largest tropical drainage basin in the world, covering an area of approximately  $7,000,000 \text{ km}^2$ .<sup>1</sup> The Amazon River transports  $\sim 36.1 \text{ Tg C yr}^{-1}$  of organic carbon (of which 60–70% is dissolved) to the Atlantic Ocean.<sup>2</sup> The Amazon River originates in the Andes cordillera, from where it carries large amounts of sediment. Along its flow path, it connects with several major tributaries of distinct chemical characteristics: “Blackwaters” that are relatively acidic (pH  $\sim 5$ ), low in total cations, and rich in dissolved organic matter (DOM), such as the Negro River; “Whitewaters” that show a near-neutral pH and are relatively rich in total cations and suspended sediments, such as the Solimões and Madeira River, and “Clearwaters” that are characterized by low suspended sediment loads, and high light transparency, such as the Tapajós River.<sup>3,4</sup> DOM cycling varies greatly in those chemical environments.<sup>5</sup> “Clearwaters” offer greater light transparency than the other types of Amazon waters and showed lower concentration of chromophoric dissolved organic matter (CDOM), the most readily photodegraded component of DOM molecules. In contrast, CDOM is the most abundant DOM fraction in “Blackwaters”. In a previous study, sunlight exposure over 27 h showed that at least 15% of Rio Negro

DOM was photoreactive.<sup>6</sup> The suspended sediments and cations in “Whitewaters” contribute to shielding from sunlight,<sup>7</sup> and transform DOM through mineral particle adsorption and complexation,<sup>8</sup> thereby affecting bioavailability as well.<sup>9</sup> The confluences of major Amazon basin rivers can be regarded as effluents of entire aquatic ecosystems, and mixing causes abrupt changes in environmental conditions including the temperature, density, flow characteristics, pH, concentration of ions and (mineral and organic) particles, DOM composition, and molecular structure.<sup>10,11</sup> Steep gradients develop for all these features in mixing zones, with distinct spatial and temporal evolution of interdependent variables. Opportunities arise for many non-conservative effects upon riverine mixing, making those places hot spots for organic matter processing and microbial metabolism influencing the dynamics of carbon pools in natural systems.<sup>12,13</sup>

Received: December 7, 2022

Revised: May 4, 2023

Accepted: May 4, 2023

Published: June 12, 2023





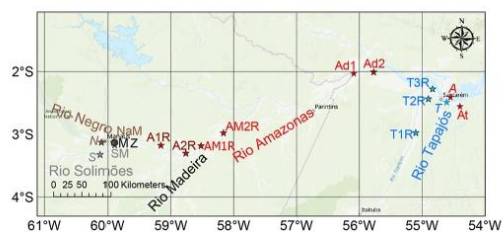
DOM is a highly complex mixture of thousands to millions of individual molecules and is ubiquitous in terrestrial and aquatic ecosystems.<sup>5,10,14–16</sup> DOM is an integral component of food webs, acting as both metabolic waste and substrate to heterotrophs, involving critical ecosystem processes such as nutrient availability and growth efficiency of aquatic organisms, and ultimately the cycling of carbon and other elements.<sup>17,18</sup> Several abiotic processes, such as photochemistry, redox chemistry, mineral sorption, and desorption, affect the composition and structure of DOM molecules.<sup>8,19</sup> Photochemical reactions in oxic environments contribute to oxygenation of DOM molecules, mainly by introduction of hydroxyl and carboxy functional groups;<sup>20–22</sup> heteroatoms N and S in DOM molecules introduce alternative selectivity in photoprocessing of DOM molecules.<sup>23</sup> Photoproduction of small oxygenated molecules, including photomineralization to CO<sub>2</sub>, that are or resemble biochemical metabolites<sup>22</sup> will enhance microbial activity.<sup>24</sup> Coupled photochemical and microbial DOM processing is context-dependent, structure-selective, and subject to synergy and competition.<sup>25</sup>

Most studies report distinct non-conservative behavior of DOM after riverine mixing affected by factors like phase partitioning,<sup>14,26</sup> respiration,<sup>27,28</sup> trace metal complexation,<sup>10</sup> and discharge.<sup>29</sup> However, the effects of physical mixing of different Amazon basin riverine waters on the composition and structure of its DOM at molecular resolution remain ill-constrained. We used ultrahigh-resolution negative and positive mode electrospray ionization (ESI[±]) Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS), as well as <sup>1</sup>H nuclear magnetic resonance spectroscopy (NMR) to study the composition and reactivity of DOM. We investigated the mixing of water samples from two major Amazon mixing zones, i.e., Solimões and Negro Rivers (S + N), and Amazon and Tapajós rivers (A + T). We conducted controlled mixing and incubation experiments with different mixing ratios and times of incubation after mixing, followed by biogeochemical and molecular characterization of the original and cultured samples to study the temporal evolution of DOM molecular composition and structure. We also measured the heterotrophic bacterial production (HBP) and dark carbon fixation (DCF) to assess microbial activities before and after the water mixing. We investigated how DOM molecular features change following mixing of the endmember river waters with different biogeochemical characteristics and investigated the factors that lead to the changes.

## MATERIALS AND METHODS

**Sample Collection.** We collected water samples at 17 sites in the Amazon basin eastward between April 2nd and May 25th in 2014 during an exceptional high-water period (Figure 1, Table S1). We obtained water samples by boat just below the surface and collected 10 L of each sample. Solid-phase extraction (SPE) was carried out in the field immediately after sampling using 1 g cartridges of PPL resin.<sup>30</sup> The SPE-extracted DOM (SPE-DOM) in methanol was stored in the freezer (−20 °C) until FT-ICR MS and <sup>1</sup>H NMR analyses. For more sampling and experimental details, see Supporting Information.

**Sample Processing.** The four endmember samples S, N, A, and T were used for laboratory mixing and incubation experiments (Table S2). Mixing and incubation experiments with unfiltered endmembers were conducted on the bank or ship directly after sampling by mixing variable ratios of



**Figure 1.** Surface water sampling locations. Samples Solimões (S) and Negro (N), Amazon River (A), and Tapajós River (T) refer to endmembers used for laboratory mixing and incubation experiments; all the other samples belong to the regional sample set.

different waters (80:20, 60:40, 50:50, 40:60, and 20:80, total volume 500 mL) in Pyrex bottles, followed by one day (1d) and five days (5d) dark incubation at in situ temperature (30 °C) and SPE in the field. We also made a mixture of identical volumes of S and N and let it stand in dark for 30 min, named S50N50\_30min.

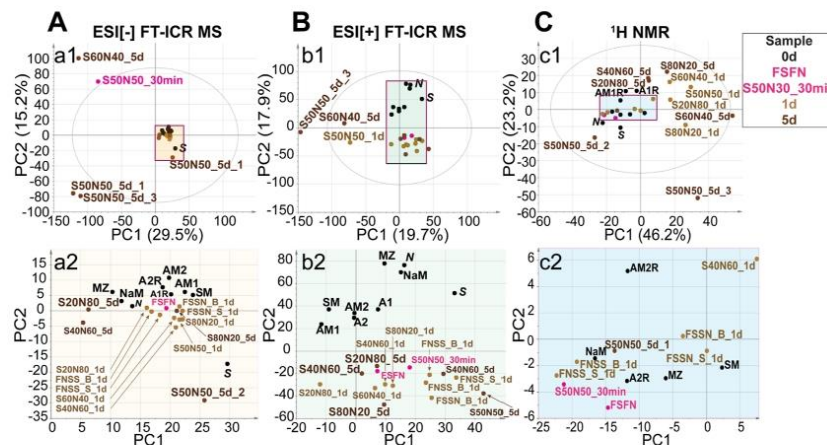
Moreover, the endmember samples S and N were separated into “filtered water” and “filtered suspended solids” based on the treatment of the endmember samples before mixing (Tables S1 and S2). The waters and suspended solids were isolated by filtration of identical volumes of water from each river using 0.2 μM Whatman GF/F glass fiber filter (precombusted at 450 °C). The mixture of filtered Negro water and filtered Solimões suspended solids was named FNSS; the mixture of filtered Solimões water and filtered Negro suspended solids was named FSSN; the mixture of filtered Solimões water and filtered Negro water was named FSN. FNSS and FSSN were left to sit in the dark for 1 day at in situ temperature (30 °C) and/or on ice (0 °C).

**Water Analysis.** Water parameters were measured on part of the samples (Table S5). Water conductivity, temperature, and dissolved oxygen were measured in situ with a portable instrument (Hanna Instruments, Metrohm electrode, and PRO-ODO YSI). Dissolved and particulate organic carbon, inorganic nutrients, and pH were determined by standard procedures, as described in Supporting Information.

**FT-ICR MS Analysis.** ESI[±] FT-ICR mass spectra of SPE-DOM were acquired using a 12 T Bruker Solarix mass spectrometer (Bruker Daltonics, Bremen, Germany) and an Apollo II ESI source.<sup>15</sup> Data processing used Compass Data Analysis 5.0 (Bruker, Bremen, Germany) and formula assignment used an in-house software (NetCalc);<sup>31</sup> and further details are described in Supporting Information.

**<sup>1</sup>H NMR Analysis.** 800 MHz <sup>1</sup>H NMR spectra of SPE-DOM were acquired with a Bruker ADVANCE III NMR spectrometer operating at 800.35 MHz ( $B_0 = 18.8$  tesla) at 283 K from redissolved dried SPE-DOM in CD<sub>3</sub>OD, as described in the Supporting Information. <sup>1</sup>H NMR spectra enabled quantification of key structural units in SPE-DOM, including pure and oxygenated aliphatics, carboxyl-rich alicyclic molecules,<sup>20</sup> olefins, and aromatics.

**Microbial Analysis.** Microbial activity was measured in part of the samples (Table S10). Heterotrophic bacterial production (HBP) refers to the assimilatory metabolism consuming preferentially labile DOM while refractory DOM is mostly respired<sup>32</sup> and was determined by the measurement of protein incorporation of radio-labeled <sup>3</sup>H-Leucine. Dark



**Figure 2.** PCA of assigned SPE-DOM molecular formulae in unmixed and mixed S + N waters pre- and post-incubation, according to sample analysis methods (A) ESI[−] and (B) ESI[+] FT-ICR mass spectrometry and (C)  $^1\text{H}$  NMR spectroscopy. a2/b2/c2 depict an enlarged view of samples in the purple rectangles in a1/b1/c1. Loading vectors are shown in Figure S2.

carbon fixation (DCF) is inorganic light-independent C-uptake, mostly performed by chemosynthetic microorganisms using redox reactions with energetic yields to recycle inorganic carbon,<sup>33</sup> and was determined by the incorporation of  $^{14}\text{CO}_2$ . More details of HBP and DCF measurements are described in Supporting Information.

**Statistical Analysis.** Principal component analysis (PCA) was performed using Simca-P (version 11.5, UmetricsAB, Umeå, Sweden). Hierarchical cluster analysis (HCA) used Hierarchical Clustering Explorer 3.0. Details are described in Supporting Information.

## RESULTS

**Changes in DOM Molecular Composition after S + N Mixing in FT-ICR MS.** Mass spectra of Solimões (S) and Negro River (N) SPE-DOM showed skewed near-Gaussian distributions for thousands of mass peaks ranging from  $m/z$  150–950 (Figure S1), indicative of high molecular diversity.<sup>5,10,11,34</sup> However, mass spectra of S + N SPE-DOM developed recognizable patterns of putative microbial-derived signatures, showed a large-scale shift to lower  $m/z$  in ESI[+] mass spectra (Figure S1), and had a considerably lower number of assigned molecular formulae in ESI[±] mass spectra (Tables S3 and S4). The average  $m/z$  of S + N SPE-DOM was significantly lower than that of S and N SPE-DOM, while its average atomic ratios (e.g., O/C, H/C, S/C, and N/C) varied more slightly after mixing (Tables S3 and S4).

To assess major variation in DOM, we performed PCA on its assigned molecular formulae (Figure 2A,B). S60N40\_5d and S50N50\_5d were separated by the first principal component (PC1) of ESI[−] and ESI[+] MS-based PCA because of the distinct abundance of less oxygenated compounds, in accordance with HCA (see Figure S3 and its footnote). The second principal component (PC2) of ESI[−] MS-based PCA placed S + N between S and other unmixed samples, whereas PC2 of ESI[+] MS-based PCA showed clear separation of river SPE-DOM and incubated S + N SPE-DOM due to the decrease of high molecular weight CHO and CHNO compounds. However, the mixing ratio of endmem-

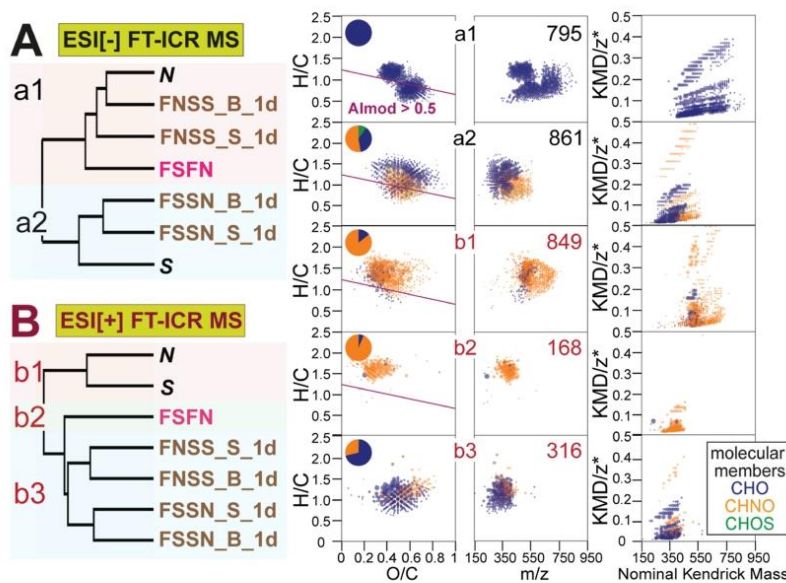
bers (S and N) did not show a link to the PCA distribution of incubated S + N. Moreover, the original samples other than S and N showed higher PC2 in ESI[−] MS-based PCA and were placed between endmembers (S and N) and incubated with S + N in PC2 of ESI[+] MS-based PCA. Furthermore, the 1d- and 5d-incubated samples were not distinct in PCA and HCA (Figures 2A,B and S3), suggesting that S + N SPE-DOM composition did not record a continuous trend in temporal evolution.

We observed a remarkable loss of molecular signatures after S + N mixing, with around half of molecular formulae disappearing immediately after mixing (Figure S4). Additionally, no unique molecular formulae were found in 0d-/1d-/5d-incubated SPE-DOM that were not present in S and N endmembers of SPE-DOM. Hence, no new molecular formulae appeared between 1d and 5d of incubation, which may reflect fast DOM transformation within 1d.

**Changes in DOM Molecular Composition after S + N Mixing in  $^1\text{H}$  NMR Spectra.**  $^1\text{H}$  NMR spectra of S and N SPE-DOM showed smooth bulk envelopes, reflecting superpositions of millions of carbon atomic environments typical of processed aqueous DOM,<sup>15</sup> whereas S + N SPE-DOM showed a large number of additional superimposed small resonances (Figure S5). Although,  $^1\text{H}$  NMR section integrals of half incubated S + N SPE-DOM were placed approximately in the middle of endmember samples S and N (Table S6), the other half showed a higher abundance of pure and functionalized aliphatic protons ( $\delta_{\text{H}} \sim 0.5\text{--}2.35$  ppm), at the expense of oxygenated aliphatic and unsaturated protons ( $\delta_{\text{H}} > 3.1$  ppm) (Figures 2C and S2C and Table S6). The incubated S + N samples were separated in PC2 of  $^1\text{H}$  NMR-PCA due to the different content of aromatic protons ( $\delta_{\text{H}} > 6.6$  ppm) (Figures 2C and S2C).

Difference  $^1\text{H}$  NMR spectra of endmember S and N SPE-DOM revealed a higher abundance of pure aliphatics (CCCH units), a lower relative abundance of aromatic (C<sub>a</sub>H), especially polyphenols, direct (OCH), and remotely oxygenated (OCC<sub>a</sub>H) units in S, compared with N SPE-DOM (Figures S6A and S7A). Furthermore, more extensively oxygenated





**Figure 3.** HCA of assigned SPE-DOM molecular formulae in part of the unmixed and mixed S + N waters pre- and post-incubation; (A) ESI[−] and (B) ESI[+] FT-ICR MS. Van Krevelen, mass-edited H/C, and KMD/z\* diagrams<sup>35,36</sup> show CHO (blue), CHNO (orange), and CHOS (green) molecular classes that were relatively more abundant in clusters a1/a2, b1/b2/b3, respectively. The pie plots depict percentages of the counts of different molecular classes. The numbers show counts of molecular compositions. The molecular formulae positioned below the purple line in the van Krevelen diagrams have a modified aromaticity index ( $Al_{mod}$ )<sup>37</sup> higher than 0.5.

aliphatic and aromatic CH units were more abundant in S than in N SPE-DOM. The river samples A1 and A2 downstream of the S + N confluence showed an increase of methylene and alkyl protons ( $\delta_H \sim 1.2$  ppm), as well as branched and functionalized aliphatics, including simple carboxylic acids ( $\delta_H \sim 1.2$ – $2.1$  ppm), polyols ( $\delta_H \sim 3.2$ – $3.6$  ppm), and a decline of highly oxygenated aliphatics ( $\delta_H > 3.6$  ppm) and  $C_{sp^2}H$  units ( $\delta_H > 6$  ppm) (Figure S8).

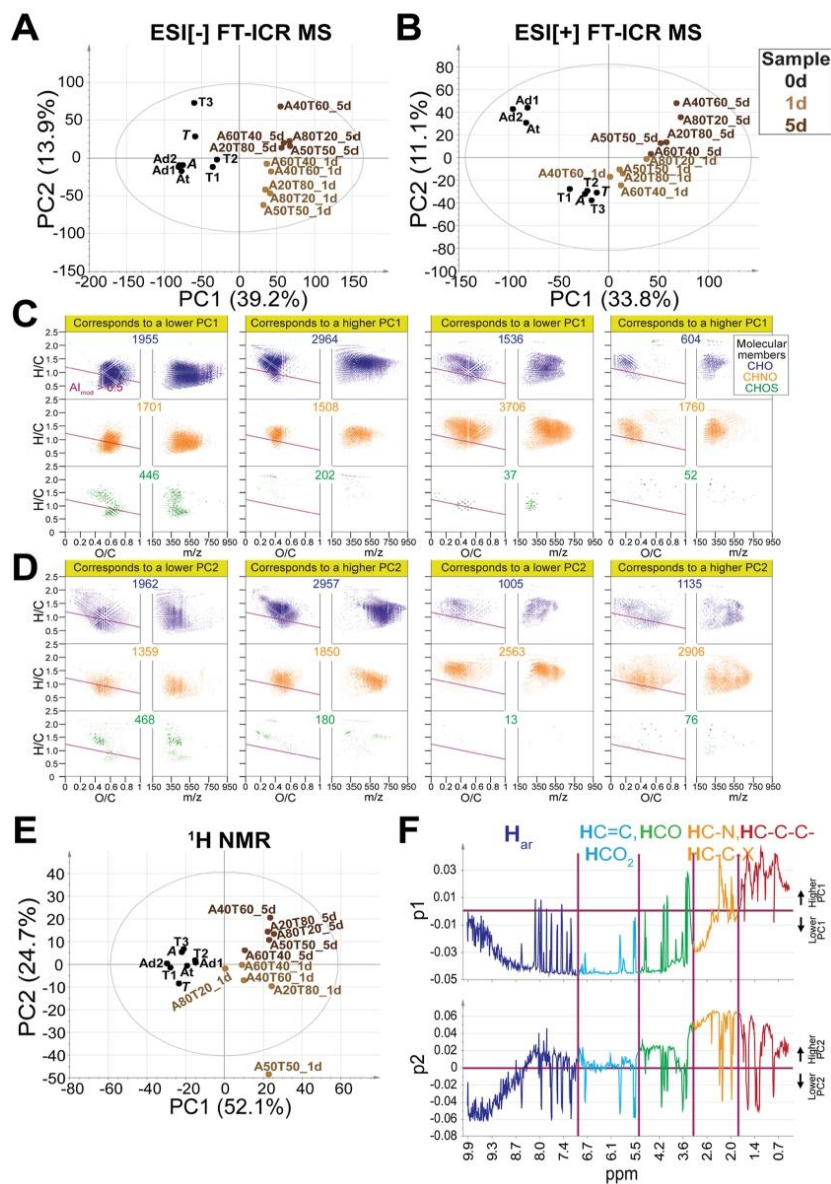
After mixing S and N for 30 min, visible increases of branched purely aliphatic molecules ( $\delta_H$ : 1.3–1.5 ppm), some carboxylic acids ( $\delta_H \sim 2.0$ – $2.4$  ppm), methyl ethers ( $\delta_H$ : 3.6–3.75 ppm) and olefins ( $\delta_H \sim 5.6$ – $6.5$  ppm) were observed (Figures S6B and S7B). After 1d, a very substantial increase of pure and functionalized aliphatics (CCCH and OCCH units, indicative of carboxylic acids), equivalent to  $\sim 50\%$   $^1H$  NMR section integral, had occurred across the board (Figures 5A, S6C); the relative abundance of oxygenated aliphatic units (OCH units) decreased, while a suite of alkyl and carboxylated aromatic molecules increased (Figures 5A, S6C, and S7C). From 1d to 5d, a remarkable reversal of DOM processing had, in essence, restored the initial status of (S + N) SPE-DOM, with an increase in pure and functionalized aliphatics ( $\delta_H \sim 0.5$ – $2.4$  ppm), loss of methoxy ethers and esters ( $\delta_H \sim 3.4$ – $4.0$  ppm), and about all other directly and remotely oxygenated aliphatic molecules ( $\delta_H > 2.4$  ppm) (Figure S6E). Hence, much of the initially produced aliphatic molecules were labile under the given conditions and became either consumed or reoxidized within four days.

**Effects of Particles in S + N Mixing and 1d Incubation.** The clustering in ESI[−] MS-based HCA indicated that the contact between water and suspended

particles had negligible effects on DOM composition (Figure 3A). FSN showed a higher similarity to N than to S, conforming to the higher concentration of DOC (dissolved organic carbon) in Negro (9.96 mg/L) than in Solimões (4.56 mg/L; Table S5). ESI[+] MS-based HCA separated S and N, FSN, 1d incubated S + N into three clusters, with distinction of FSN samples alike ESI[−] derived HCA (Figure 3), suggesting that degradation of some more saturated, lower molecular weight CHNO compounds and synthesis of lower mass ( $m/z < 450$ ) CHO compounds with average H/C and O/C ratios occurred during 1d incubation (Figure 3b2).

**Changes in DOM Molecular Composition after A + T Mixing in FT-ICR MS.** ESI[±] FT-ICR mass spectra of A + T SPE-DOM developed intense mass peaks superimposed to the bulk signature (Figure S9), indicative of efficient production of specific molecular signatures in parallel with large-scale transformation of DOM molecules. The average  $m/z$ , average O/C ratio, and count of molecular formulae substantially declined, whereas the average H/C ratio increased after the 5d-incubation (Tables S7 and S8).

In both ESI[−] and ESI[+] MS-based PCA (Figure 4), PC1 showed a clear trend: Amazon < Tapajós < 1d-incubated A + T < 5d-incubated A + T, due to increased abundance in compounds with a lower O/C ratio, a lower  $m/z$ , and a modified aromaticity index ( $Al_{mod}$ )<sup>37</sup> less than 0.5. Moreover, 1d-incubated A + T < 5d-incubated A + T were separated by PC2 and had different abundances in high molecular weight CHO and CHNO molecules. The molecular patterns indicated extensive transformation from more oxygenated, high molecular weight, higher aromaticity compounds to less oxygenated, low molecular weight, higher aliphaticity compounds during



**Figure 4.** PCA of assigned SPE-DOM molecular formulae in unmixed and mixed A + T waters pre- and post-incubation. Panels (A,B,E) show PCA scatter plots. Panels (C,D) show correlation patterns of PC1 and PC2 loading vectors (p1 and p2) in van Krevelen and mass-edited H/C diagrams, with the color code representing CHO (blue), CHNO (orange), and CHOS (green) molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have a modified aromaticity index ( $AI_{mod}$ )<sup>37</sup> higher than 0.5. Panel F shows the correlation patterns of p1 and p2 in <sup>1</sup>H NMR space, with fundamental molecular structures indicated.

5d-incubation. Additionally, A was separated from other Amazon samples and showed high similarity to Tapajós samples in ESI[+] MS-based PCA. Moreover, the mixing ratio of endmembers (A and T) did not show a link to the PCA

distribution of incubated A + T, which was consistent with the result of mixing S + N.

HCA clearly separated unmixed and mixed A + T samples (Figure S10), in line with PCA results. 1118 oxygen-rich and



polyphenolic CHO/CHNO compounds decreased, whereas 1228 more aliphatic and more saturated CHO/CHNO compounds increased in relative abundance after the 1d-/5d-incubation. KMD/ $z^*$  diagrams revealed extended  $\text{CH}_2$ -based homologous series for CHO compounds that increased/decreased during incubation and shorter ones for CHNO compounds (Figure S10).

We found that nearly half of the molecular formulae disappeared after A + T mixing and 1d-incubation, in line with S + N mixing (Figure S11). The count of molecular formulae that disappeared and newly appeared after A + T mixing were about equal ( $\sim 3000$ ). Furthermore,  $\sim 1000$  molecular signatures remained after 1d and disappeared after 5d (Figure S11a4,b4).

**Changes in DOM Molecular Composition after A + T Mixing in  $^1\text{H}$  NMR.** Incubation caused the emergence of sharp  $^1\text{H}$  NMR resonances of putative biotic origin, which contributed  $\sim 3\%$  to the total  $^1\text{H}$  NMR integral in 1d incubated A + T SPE-DOM and  $\sim 1\%$  in 5d incubated A + T SPE-DOM (Figure S12). The 1d incubation produced strong  $^1\text{H}$  NMR resonances containing methylene-rich aliphatic molecules ( $\delta_{\text{H}} \sim 1.25$  ppm), an admixture of aliphatic branching ( $\delta_{\text{H}} \sim 1.3$ – $1.5$  ppm) that is probably associated with carboxylic acids  $\text{RCOOH}$  ( $\text{HOOC-CH}\alpha\text{-CH}\beta$ ;  $\delta_{\text{H}} \sim 2.3$ – $2.45$  ppm;  $1.6$ – $1.75$  ppm), olefins ( $\delta_{\text{H}} \sim 5.2$ – $5.5$  ppm/ $5.8 \pm 0.1$  ppm), and carboxylated benzene derivatives ( $\delta_{\text{H}} > 7.3$  ppm)<sup>38</sup> (Figures S13B and S14B).

The abundance of purely aliphatic protons (CCCH units;  $\delta_{\text{H}} \sim 0.5$ – $1.9$  ppm) largely increased within 1d from 34% to 42% at the expense of about all other units (Table S9). The carboxylic acids with supposedly more “simple” aliphatic branching motifs that were produced within 1d declined between 1d and 5d (Figure S13D). So, temporal evolution SPE-DOM from 1d to 5d of S + N and A + T shared aspects of synthesis after 1d and receding to the original state after 5d. However, alterations within A + T concerned a rather limited suite of methylene-rich molecules in comparison with S + N which concerned an overall change in different molecular structures.

PC1 and PC2 in  $^1\text{H}$  NMR-based PCA explained  $\sim 77\%$  of the total variance (Figure 4E) and showed similar separation compared to ESI[−] MS-based PCA (Figure 4A). The mixing ratios of endmembers did not show an impact on the molecular composition of incubated A + T SPE-DOM from  $^1\text{H}$  NMR-based PCA.

**Changes in Microbial Activity after S + N and A + T Mixing.** We measured the rates of two microbial metabolic processes in unmixed and mixed waters. HBP refers to assimilatory metabolism consuming preferentially labile DOM-like proteins,<sup>32</sup> and DCF refers to light-independent inorganic carbon uptake mostly performed by chemosynthetic microorganisms.<sup>33</sup> HBP in endmembers S, N, and in the S + N mixing zone, as well as in the one-hour incubated equal volume S + N mixtures, remained nearly constant around  $0.03 \mu\text{gC L}^{-1} \text{h}^{-1}$  (Table S10). However, HBP in A1 and A2 downstream of the S + N confluence showed over 100 times higher HBP rates. The HBP rates in endmembers A, T, and in the A + T mixing zone were around  $0.15 \mu\text{gC L}^{-1} \text{h}^{-1}$ , approximately five times higher than in the S + N mixing zone, in line with previous research that showed high rates of microbial respiration and bacterial growth in the Amazon-Tapajós confluence.<sup>39,40</sup>

The rates of DCF were much higher than HBP in all the water samples, with  $\sim 200$  times higher in S + N confluence

and  $\sim 10$  times higher in A + T confluence. DCF rates were two times higher in N ( $1.66 \mu\text{gC L}^{-1} \text{h}^{-1}$ ) compared to S ( $0.76 \mu\text{gC L}^{-1} \text{h}^{-1}$ ) and increased  $\sim 10$  times in the S + N mixing zone. DCF rates in A and T were  $\sim 0.45 \mu\text{gC L}^{-1} \text{h}^{-1}$  and increased  $\sim 4$  times in the A + T mixing zone.

Our measurements of HBP and DCF in the Amazon River are consistent with rates observed in other aquatic systems worldwide (Table S11). Our observations of biomass production ranging from nanograms to milligrams of carbon per hour are similar to those reported in previous studies conducted in the Amazon and other aquatic systems. Earlier studies in the Amazon have reported HBP varying from 0.1 to  $7 \mu\text{gC L}^{-1} \text{h}^{-1}$ , indicating higher values during high waters<sup>12,41–43</sup> and in mixing zones.<sup>42</sup> We observed our highest HBP in the Amazon River downstream of the confluence of the Tapajós River, which has been associated with increased rates of phytoplankton productivity, including labile organic compounds for microbial growth.<sup>43</sup> On the other hand, we observed the highest DCF in the Amazon River after the confluence of the Negro and Solimões Rivers, where we observed a high concentration of DIC of approximately  $1300 \mu\text{M}$  and high assimilatory ratios. The mixing zone is often reported as a hotspot for microbial processes due to the mixing of microbial communities and river compounds,<sup>12,13</sup> creating opportunities for high rates of carbon assimilation via distinct processes, including chemosynthesis using the energetic yields of oxidation of reduced compounds.<sup>42</sup> Although our survey represents all available measurements in the Amazon River, we found other studies in different aquatic systems where rates vary from nanograms of carbon per hour in lakes.<sup>44–46</sup> High rates of DCF were observed in karstic systems,<sup>47–49</sup> which could be related to the high concentration of inorganic carbon and reduced compounds in soil and groundwater, creating opportunities for microbial growth using the energetic yields of redox reactions to promote chemosynthesis since HBP was observed at lower rates in these systems.<sup>50,51</sup> In river estuaries, turbulent mixing promotes an increase in DCF<sup>47,52,53</sup> and HBP,<sup>54–56</sup> which could also be associated with reduced compounds from bottom sediments meeting oxic surface waters.

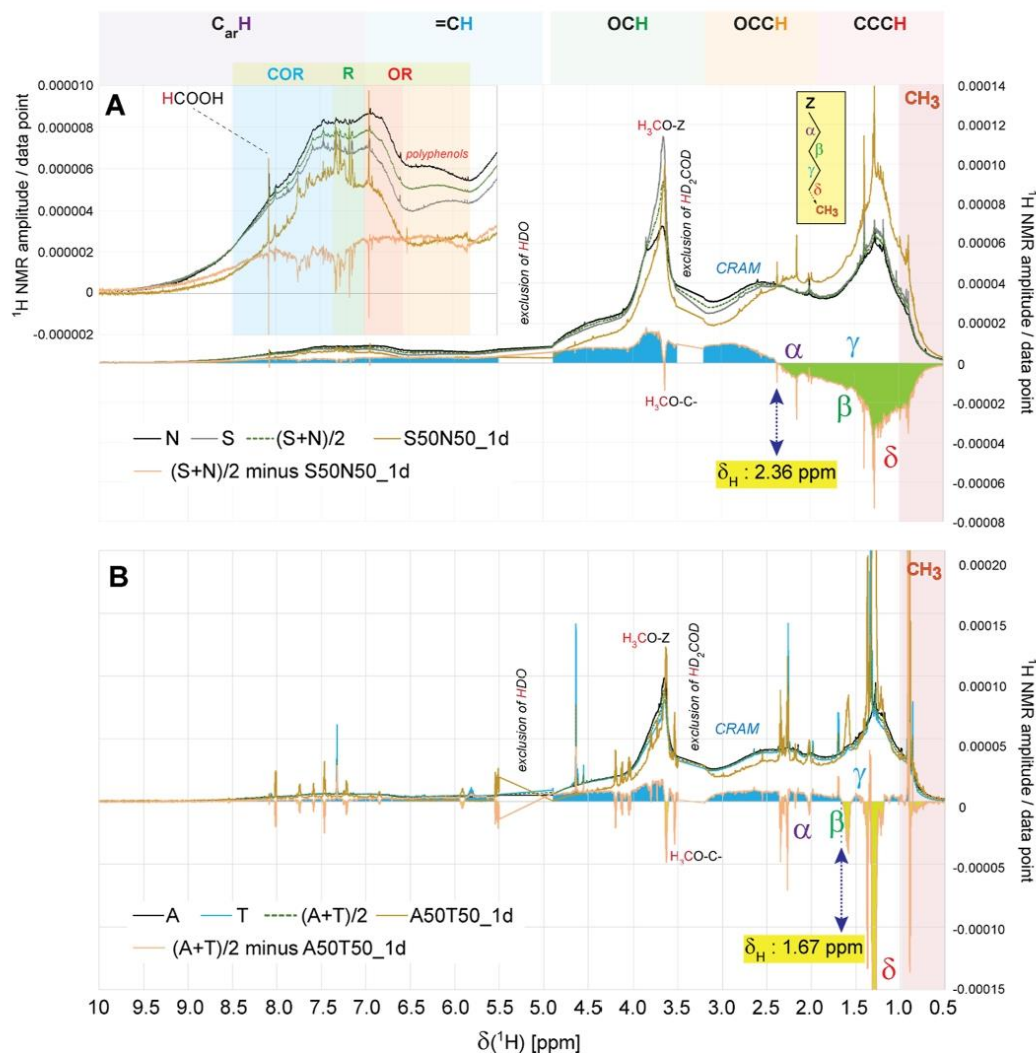
## DISCUSSION

**Molecular Characteristics of Endmembers.** Blackwater N DOM showed higher average  $m/z$  ratios (Tables S3 and S4) and higher aromaticity (Table S6 and Figure S5) compared to DOM of other rivers, in agreement with previous studies.<sup>5,57</sup> Ertel et al. assigned 70% of degraded humic and fulvic acid fluxes at the S + N confluence to Negro inputs.<sup>57,58</sup> High molecular weight aromatic structures probably represent the polar fraction of lignin and tannin degradation products, e.g., through biotic and abiotic oxidation.<sup>58,59</sup> N DOM showed higher proportions of =CH and  $\text{C}_\alpha\text{H}$  units ( $\delta_{\text{H}} > 5.2$  ppm) and remotely (OCCH) and directly oxygenated (OCH) aliphatic molecules ( $\delta_{\text{H}} > 2.4$  ppm) (Table S6 and Figure S5).

S and A SPE-DOM showed higher average N/C and S/C ratios and had a larger contribution of moderately unsaturated and saturated compounds (Tables S3 and S4). Nitrogen-containing compounds could be preferentially sorbed and transported by the sediment-rich waters.<sup>60</sup>

Clearwater T SPE-DOM showed higher proportions of CHNO compounds and methylene- and alkyl-rich aliphatic structures (Figure S13 and Table S9) than other endmember DOM (Tables S7 and S8), in line with higher proportions of





**Figure 5.** Different  $^1\text{H}$  NMR spectra of (A)  $(S + N)/2$  minus S50N50\_1d and (B)  $(A + T)/2$  minus A50T50\_1d, with annotations of key substructures, demonstrate the fundamental distinction in molecular evolution of S + N and A + T after 1d-incubation (cf. text). Incubation of S + N caused alterations across all atomic environments, whereas incubation of A + T referred primarily to a limited set of basic methylene- and alkyl-rich aliphatic molecules (cf. text). Yellow-shaded  $\delta_{\text{H}}$  denotes a significant crossover of aliphatic molecules which were consumed (blue shade) or produced (green shade) after 1d, denoting a higher molecular diversity and degree of functionalization in newly produced S + N molecules compared with those of A + T.

primary production and microbial processing.<sup>11</sup>  $^1\text{H}$  NMR spectra of T SPE-DOM showed a distinct visual appearance, owing to a few sharp and abundant resonances representing aliphatic molecules with limited branching, acetic acid ( $\delta_{\text{H}} \sim 1.88$  ppm) and formic acid ( $\delta_{\text{H}} \sim 8.08$  ppm), with an overall  $^1\text{H}$  NMR integral below  $\sim 2\%$  (Figure S12). The strong contribution of small aliphatic molecules in T DOM may originate from the general proximity of its DOM molecules to the structural space of biomolecules and metabolites,<sup>11</sup>

including some biodegradation of fresh materials, as suggested by Roth et al.<sup>61</sup>

The endmember samples S, N, A, and T showed distinct characteristics, including the intensity distribution of H/C, O/C, N/C, S/C ratios,  $m/z$ , aliphaticity, and aromaticity, in line with previous reports.<sup>12,19</sup> Furthermore,  $^1\text{H}$  NMR spectra showed the main distinction in relative abundance of methoxy-related substructures in the order  $S > N$  (Figure S6A) and  $A > T$  (Figures S13A and S14A). The main distinction between S

and *N* SPE-DOM largely covered methyl ethers ( $\delta_{\text{H}} \sim 3.6\text{--}3.8$  ppm), whereas the difference between *A* and *T* SPE-DOM covered methyl esters and methyl ethers ( $\delta_{\text{H}} \sim 3.6\text{--}4.1$  ppm).<sup>15</sup> This may imply distinct initial abundance and processing of more “ether-rich” lignins and more “ester-rich” tannins,<sup>62</sup> but the scavenging of reactive DOM molecules by methanol<sup>63</sup> cannot be excluded.

**Distinct Structural Evolution after S + N and A + T Mixing.** In *A* + *T*, a suite of labile lipid-like molecules with prominent alkyl- and methylene-rich carboxylic acids of limited structural diversity and few oxygenated aliphatic and aromatic structural motifs had been produced after 1d and largely disappeared after 5d of incubation. In comparison, the structural changes after 1d incubation of *S* + *N* DOM comprised a much larger structural diversity of molecules, covering the entire range of aliphatic (CCCH, OCCH, and OCH units) and C<sub>sp2</sub>H units (aromatics and olefins) with a smooth distribution of all structural motifs (atomic environments), implying a more thorough diversification of atomic environments in *S* + *N* compared with *A* + *T* following mixing and incubation. This probably reflects an overall larger distinction of DOM molecules in original *S* and *N* waters compared with *A* and *T* waters, originating from the high relative abundance of highly oxygenated aromatic structures in *N* waters (Table S6). Gain was observed for a very wide range of pure and functionalized aliphatic units ( $\delta_{\text{H}} \sim 0.5\text{--}2.36$  ppm), and loss was observed for all other atomic environments in the case of *S* + *N* incubation (Figure S5A). While the rather high proportions of the above-mentioned methylene- and alkyl-rich molecules produced after 1d in the case of *A* + *T* incubation may have masked other alterations on visual inspection of <sup>1</sup>H NMR spectra, appreciable distinction between 1d incubation for *A* + *T* versus *S* + *N* also applied for less intense <sup>1</sup>H NMR resonances. The crossover points for difference spectra “endmember minus 1d incubation” was displaced from  $\delta_{\text{H}} \sim 2.36$  ppm (*S* + *N*) to  $\delta_{\text{H}} \sim 1.67$  ppm (*A* + *T*), implying that less oxygenated aliphatic molecules were preferentially degraded after 1d incubation in the case of *A* + *T*. The structural evolution from 1d to 5d incubation was in essence a reversal of the structural changes from original samples to 1d incubation for both *S* + *N* and *A* + *T* mixing experiments: *A* + *T* samples diversified their pure and remotely oxygenated aliphatic molecular structures ( $\delta_{\text{H}} \sim 0.5\text{--}3$  ppm) and lost methylene- and alkyl-rich aliphatic molecules with simple branching motifs (Figure S12); at day 5, *A* + *T* DOM comprised higher proportions of pure aliphatic molecules ( $\delta_{\text{H}} \sim 0.8\text{--}1.7$  ppm) and lower shares of methoxy groups ( $\delta_{\text{H}} \sim 3.5\text{--}4.1$  ppm) than the average of *A* + *T* (Figure S12C). For *S* + *N*, the evolution from 1d to 5d incubation appeared as a very congruent reversal compared with the evolution from average *S* + *N* to 1d incubation for all aliphatic units ( $\delta_{\text{H}} \sim 0.5\text{--}5$  ppm (Figure S6). The distinction between the average of *S* + *N* and 5d incubation were very minor, like the small increase of alkyls C<sub>*n*</sub>H<sub>2*n*+1</sub> ( $\delta_{\text{H}} \sim 0.5\text{--}1.0$  ppm) and methoxy groups ( $\delta_{\text{H}} \sim 3.4\text{--}3.7$  ppm) (Figure S6D,E). This remarkably fast processing of DOM for both mixing experiments has likely been promoted by tropical temperatures ( $\sim 30$  °C).

The large-scale depletion of high molecular weight compounds after mixing was significant for both *S* + *N* and *A* + *T*, especially for nitrogen-containing compounds that were better ionized in ESI[+] MS (Figure S1 and Table S7). This shift of mass was also pronounced for FSN and SSONS0\_30min, suggesting an immediate change following

mixing. Unlike *A* + *T* SPE-DOM that showed continuous molecular changes over 5d (Figure 4), *S* + *N* SPE-DOM changed immediately after mixing but showed no further continuous trends over 5d (Figure 2). This sudden change in DOM properties probably results from the abiotic, preferential sorption of hydrophilic, polyphenol-rich Negro DOM by Solimões suspended particles.<sup>10,64,65</sup> Additionally, the large depletion of POC (particulate organic carbon) at Solimões-Negro mixing zone suggested that the DOC-POC transition significantly affected the DOM composition in the confluence (Table S5). Coagulation heterogeneity, including the coating of minerals with specific DOM fractions and microbes,<sup>66</sup> likely contributed to the observed heterogeneity in *S* + *N* DOM (Figure 2).

We observed conservative behavior of DOC concentration downstream of the *S* + *N* confluence, in line with two previous studies,<sup>10,67</sup> which suggest that organic nanoparticulate or colloidal matter did not readily coagulate to form larger aggregates during mixing.<sup>67</sup> Another reason suggested by Simon et al. is the occupation of mineral sorption sites by DOM molecules.<sup>10</sup>

The abundant *m/z* ions in incubated *A* + *T* SPE-DOM occupied van Krevelen positions congruent with the abundant mass peaks in phytoplankton and microbial exudates (Figure S10a2,b2).<sup>68,69</sup> The Tapajós River hosts higher proportions of cyanobacteria ( $\sim 12\%$ ) compared to the Amazon River ( $\sim 5\%$ ), where actinobacteria and betaproteobacteria dominate the microbial community.<sup>70</sup> “Positive priming effect” was suggested to happen when algal-rich Tapajós waters mix with turbid Amazon main stem waters, accelerating the decomposition of more recalcitrant organic matter.<sup>32</sup> The molecular transformation of CHO, CHNO, and CHOS molecules toward lower *m/z* and lower O/C ratios with concomitantly higher H/C ratios in *A* + *T* SPE-DOM during 5d incubation is consistent with a simple notion of DOM degradation from larger to smaller molecules, which was expressed in the familiar size-reactivity continuum model,<sup>71</sup> but here, it expands to relatively small DOM molecules and/or metabolites. Moreover, <sup>1</sup>H NMR spectra indicated that polyphenolic molecules were largely removed in *A* + *T* SPE-DOM after 5d (Figure S12B and Table S9), suggesting rapid processing of terrestrial organic matter (e.g., leached from litter decay processes) that had been largely hydrologically mobilized in aquatic–terrestrial transition zones.<sup>72,73</sup> This conforms to a study showing that the breakdown of vascular plant-derived organic matter to CO<sub>2</sub> increased six-fold at the Amazon-Tapajós confluence compared to that in the Amazon River.<sup>74</sup>

**Non-conservative DOM Behavior after Mixing.** A study by Simon and co-workers observed a conservative behavior of SPE-DOM in the natural and experimental mixing of Solimões and Negro river waters based on negative ionization mode FT-ICR MS,<sup>10</sup> opposing the often-presumed non-conservative behavior of DOM and our results. In our study, the mixing ratios of both endmembers (*S* + *N* and *A* + *T*) did not show linear correlations with DOM composition based on ESI[±] FT-ICR MS and <sup>1</sup>H NMR spectra. Simon et al. sampled during the dry period, while the samples from this study were taken during a maximum flood pulse. The different sampling season presumably affects composition and structure of dissolved constituents leached from rewetted soils and may influence the endmember water characteristics and microbial community.<sup>4</sup> Moreover, Simon and co-workers mixed samples under continuous shaking, while we mixed initially and then let



the sample settle, incubated for 1d and/or 5d. The differences in sample treatment may affect DOM composition, for instance through sorption and desorption between DOM and particles. Besides, the combination of SPE and ESI[−] FT-ICR MS favors observation of carboxylic acids and does not cover the entire spectrum of dissolved organic molecules.<sup>75–77</sup> Our joint use of both complementary electrospray modes FT-ICR MS and <sup>1</sup>H NMR methods and the exhaustive analysis provide valuable DOM molecular information with improved detail. ESI[−] FT-ICR MS preferentially detected high-mass oxygen-rich CHO compounds, whereas ESI[+] FT-ICR MS primarily detected aliphatic CHNO compounds.<sup>34</sup> Our previous study on the SPE-DOM in Amazon main stem and tributaries showed that nearly half of CHO and CHNO *m/z* ions were found in both ESI modes at near-average H/C and O/C ratios.<sup>11</sup> Moreover, 800 MHz <sup>1</sup>H NMR spectra provided near-quantitative data of major atomic environments in DOM molecules with excellent S/N ratio and resolution, supplying relevant information about the fate of aliphatic and aromatic moieties that are poorly ionized in ESI mass spectrometry.<sup>78</sup>

**Effects of Biotic and Abiotic Factors on DOM Transformation.** Abiotic reactions, such as photochemical processing (oxidation and mineralization) and metal-dependent complexation, coagulation, and redox chemistry, will have significant impact on the DOM composition in the Amazon River.<sup>21,64</sup> Structure-selective sorption of DOM on mineral surfaces<sup>8,11</sup> has removed substantial proportions of polyphenols present in N waters together with minerals present in S waters. This removal of light absorbing and light scattering ingredients in the S + N mixing zone opens previously unavailable opportunities for photo-processing of DOM molecules. However, many photoproduct oxygenated small molecules will readily integrate in biochemical pathways, foodwebs,<sup>22,25</sup> and transient particulate organic carbon, making traceability to specific mechanisms difficult.

Multistep processing of DOM along the flow path and mixing zones of many ungauged Amazon River tributaries implies that the carbon atoms that are eventually delivered to the Atlantic Ocean have repeatedly changed their atomic environments on their journey.<sup>61</sup> However, general congruence of biochemical pathways across all organisms, and similarities of fundamental abiotic processing regimes in the extended tropical ecosystem of the Amazon basin will rather realign thermodynamic and kinetic boundary conditions at locations of exceptional importance, leading to thorough redirection for the synthesis and degradation of DOM molecules. This is more likely than an unspecific continual entropy-driven diversification of DOM molecules from land to sea. Overall, DOM evolution in the Amazon Basin will diversify beyond simple trajectories of synthesis and degradation, as described in the standard river continuum concept<sup>79</sup> and/or small river catchments.<sup>80</sup> Nevertheless, we expect that a certain set of rather unreactive and difficult to observe aliphatic molecules with structural features resisting expedient biotic and abiotic degradation might become selectively enriched along the Amazon River flow path. Plausible structural features comprise fused and bridged polycyclicalkyls with structural oxygen and nitrogen atoms connected to more than a single carbon atom, overall distant from common biomolecular binding motifs that are subject to expedient enzymatic degradation.

**Environmental Implications.** According to the “regional chromatography” hypothesis,<sup>57,81</sup> the majority of reactive components within the DOC pool are attenuated by minerals

before reaching higher-order confluences, such as the S + N confluence. We hypothesized that individual S, N, A, and T Amazon DOM had reached a steady state under given biogeochemical conditions. However, the assemblages of biogeochemically distinct river waters initiated rapid and widespread molecular alteration of DOM upon mixing, in part by fast abiotic reactions. This averaged physiochemical properties and reassembled microbial community composition, with potentially considerable effects on DOM processing and mineralization, as well as the composition and reactivity of remaining organic molecules.<sup>82</sup> Pronounced changes in DOM composition during incubation suggest substantial diagenetic modification of organic matter, even if the complete mineralization of DOM may not be significant.<sup>83,84</sup> Nevertheless, continuous oxidation of the residual DOM molecules is likely to generate oxygen-rich CO<sub>2</sub> and H<sub>2</sub>O molecules in the river confluence and subsequent river channels, contributing to a nominal reduction of leftover DOM molecules even in the case of oxidative processing.<sup>84</sup>

Rivers provide a direct link between terrestrial and marine carbon cycles. However, very little terrestrial DOM appears to accumulate in the global ocean.<sup>85–87</sup> Incubation experiments with turbid water from close to the river mouth suggested that photo- and bio-alteration leave significant molecular and carbon isotopic imprints on the terrigenous DOM. Nonetheless, quantitative removal by bio- and photo-degradation appeared to proceed at a relatively slow pace because no significant decrease of dissolved organic matter was found within five days of the incubation.<sup>32</sup> However, 9–30% of DOC was lost after the five-day incubation with less turbid water from the intermediate and outer plume, suggesting that the introduction of reactive algal DOM in the intermediate plume may thus have primed the microbial degradation of terrigenous DOM.<sup>32,88</sup> Additionally, sorption of terrigenous DOM to sinking particles acts as an important DOC sink in the Amazon plume.<sup>32</sup> These results, along with our results, demonstrate that the DOM along the Amazon River-to-ocean continuum was not in a steady state as suggested by the “regional chromatography” hypothesis. Moreover, river confluences and a less turbid offshore plume are likely important carbon sinks in the Amazon River due to biotic processes (e.g., biodegradation) and abiotic processes (e.g., photodegradation and sorption to particles).

The higher heterotrophic bacterial activity in A + T compared with S + N (Table S10) suggested that the molecular changes in A + T SPE-DOM during 5d-incubation referred more to bioavailable alkyl-rich molecules than in the case of S + N, which were more oxygenated on average. Moreover, DCF rates noticeably increased in both mixing experiments and amounted up to ~95% of the total HBP after the mixing. The results indicate that non-photosynthetic carbon fixation can represent a substantial contribution to an autochthonous source of organic matter in river confluences.<sup>89</sup> DCF had been reported to contribute approximately 80% of the total HBP in Swedish boreal lake sediments.<sup>33</sup> Moreover, it has been shown that the heterotrophic metabolism of bacteria is suddenly intensified after the input of organic matter.<sup>90</sup> The water–particle interface in river confluences could be sites of intensive biogeochemical activity, creating steep chemical gradients that provide a microenvironment with a high chemolithotrophic rate.<sup>91</sup> The simultaneous presence of oxidized and reduced organic molecules supplies energy and

substrates needed to support DCF by chemoautotrophic microbes.

## CONCLUSIONS

Ecological models and biogeochemical characterization have established river mixing zones as hotspots of microbial activity, but molecular DOM degradation pathways are still poorly addressed in river networks, preventing an accurate estimate of carbon sources and sinks. We observed extensive but molecularly distinct structural alterations of DOM upon mixing and follow-up incubation of Solimões/Negro on the one hand and Tapajós/Amazon waters on the other hand. Molecular distinction of DOM structural evolution will apply in essence to all river confluences, and educated modeling eventually has to take these mechanistic details of DOM processing into account.

Relevant molecular resolution concerning the fate, transport, and transformation of DOM in natural and engineered systems is not attained without the use of complementary high-resolution structure-selective analytical and biogeochemical methods. Characterization of molecularly heterogeneous and polydisperse DOM by complementary structural spectroscopy like, e.g., [±]ESI mass spectrometry and NMR spectroscopy provides distinct aspects of its composition and structure. In this study, [+]ESI mass spectra provided relevant coverage of CHNO molecules, <sup>1</sup>H NMR spectra demonstrated major distinctions in the structural evolution of DOM in the two S + N and A + T mixing zones, not visible in mass spectra. Joint use of structural and classical biogeochemical analysis of DOM molecular evolution will clearly diversify our perception and help us appreciate the specificity of DOM processing according to sampling points. Unifying these informative findings will remain a challenge for integrative modeling studies of the global carbon and other element cycles.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsestwater.2c00621>.

Details of sampling, analyses, interpretations, ESI[-] and ESI[+] FT-ICR mass spectra; loading vectors; water parameters; HCA; common and unique molecular signatures; <sup>1</sup>H NMR spectra; <sup>1</sup>H NMR section integrals; difference <sup>1</sup>H NMR spectra; and HBP and DCF (PDF)

## AUTHOR INFORMATION

### Corresponding Author

**Norbert Hertkorn** – *Research Unit Analytical Biogeochemistry, Helmholtz Munich, Neuherberg 85764, Germany; Department of Thematic Studies—Environmental Change and Biogas Solutions Research Center (BSRC), Linköping University, Linköping SE-581 83, Sweden; Email: norbert.hertkorn@liu.se*

### Authors

**Siyu Li** – *Research Unit Analytical Biogeochemistry, Helmholtz Munich, Neuherberg 85764, Germany; orcid.org/0000-0002-2344-0995*  
**Mourad Harir** – *Research Unit Analytical Biogeochemistry, Helmholtz Munich, Neuherberg 85764, Germany; Chair of Analytical Food Chemistry, Technische Universität München, Freising-Weihenstephan 85354, Germany*

**Philippe Schmitt-Kopplin** – *Research Unit Analytical Biogeochemistry, Helmholtz Munich, Neuherberg 85764, Germany; Chair of Analytical Food Chemistry, Technische Universität München, Freising-Weihenstephan 85354, Germany*

**Fausto Machado-Silva** – *Program in Geosciences—Environmental Geochemistry, Chemistry Institute, Fluminense Federal University, Niteroi 24020-141, Brazil; Department of Environmental Sciences, University of Toledo, Toledo, Ohio 43606, United States*

**Michael Gonsior** – *Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, Maryland 20688, United States; orcid.org/0000-0003-0542-4614*

**David Bastviken** – *Department of Thematic Studies—Environmental Change, Linköping University, Linköping SE-581 83, Sweden; orcid.org/0000-0003-0038-2152*

**Alex Enrich-Prast** – *Department of Thematic Studies—Environmental Change and Biogas Solutions Research Center (BSRC), Linköping University, Linköping SE-581 83, Sweden; Multiuser Unit of Environmental Analysis, University Federal of Rio de Janeiro, Rio de Janeiro 11070-100, Brazil*

**Juliana Valle** – *Research Unit Analytical Biogeochemistry, Helmholtz Munich, Neuherberg 85764, Germany*

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsestwater.2c00621>

### Author Contributions

D.B., A.E.-P., M.G., P.S.-K., and N.H. designed the research. M.H., P.S.-K., J.V., F.M.-S., D.B., M.G., and N.H. participated in the sampling. S.L. and M.H. acquired FT-ICR mass spectra. S.L. and N.H. acquired NMR spectra. F.M.-S. conducted HBP and DCF measurements. Data interpretation was performed by all authors. S.L. wrote the first draft of the manuscript. All authors made significant contributions to the final manuscript. CRediT: Siyu Li conceptualization (equal), data curation (equal), formal analysis (equal), investigation (equal), methodology (lead), validation (equal), visualization (equal), writing-original draft (lead), writing-review & editing (equal); Mourad Harir data curation (supporting), formal analysis (supporting), investigation (supporting), methodology (supporting), project administration (supporting), supervision (supporting), validation (supporting), visualization (supporting), writing-original draft (supporting), writing-review & editing (supporting); Philippe Schmitt-Kopplin conceptualization (equal), data curation (equal), formal analysis (supporting), funding acquisition (equal), investigation (equal), project administration (equal), resources (lead), software (lead), supervision (equal), writing-review & editing (supporting); Fausto Machado-Silva data curation (supporting), formal analysis (supporting), funding acquisition (supporting), investigation (supporting), methodology (supporting), writing-review & editing (supporting); Michael Gonsior investigation (supporting), writing-review & editing (supporting); David Bastviken writing-review & editing (supporting), funding acquisition (supporting); Alex Enrich-Prast conceptualization (equal), funding acquisition (equal), project administration (supporting), writing-review & editing (supporting); Juliana Valle data curation (supporting), writing-review & editing (supporting); Norbert Hertkorn conceptualization (equal), data curation (equal), funding acquisition (equal), investigation (lead),



project administration (lead), supervision (lead), visualization (equal), writing-original draft (supporting), writing-review & editing (lead).

#### Notes

The authors declare no competing financial interest.

#### ACKNOWLEDGMENTS

This work was supported by the Alexander von Humboldt Foundation (Norbert Hertkorn and Alex Enrich-Prast; Research Linkage Program: Connecting the diversity of DOM and CO<sub>2</sub> and CH<sub>4</sub> production in tropical lakes); the Swedish Research Council for Sustainable Development, FORMAS (grant no. 2021-02429) the Swedish Research Council (VR; grant no. 2012-00048), the European Research Council ERC (grant no. 725546), the Chinese Scholarship Council (LS; grant no. 201806360268), and several grants from Brazilian Foundations CAPES (Probal), FAPERJ (Cientista do Nosso Estado) and CNPq (Universal and Ciências sem Fronteiras) to Alex Enrich-Prast. We thank the anonymous reviewer for the fair and balanced review, which helped to improve the manuscript considerably.

#### REFERENCES

- Hess, L. L.; Melack, J. M.; Affonso, A. G.; Barbosa, C.; Gastil-Buhl, M.; Novo, E. M. L. M. Wetlands of the lowland Amazon basin: Extent, vegetative cover, and dual-season inundated area as mapped with JERS-1 synthetic aperture radar. *Wetlands* **2015**, *35*, 745–756.
- Richey, J. E.; Hedges, J. I.; Devol, A. H.; Quay, P. D.; Victoria, R.; Martinelli, L.; Forsberg, B. R. Biogeochemistry of carbon in the Amazon River. *Limnol. Oceanogr.* **1990**, *35*, 352–371.
- Junk, W. J.; Piedade, M. T. F.; Schöngart, J.; Wittmann, F. A classification of major natural habitats of Amazonian white-water river floodplains (várzeas). *Wetlands Ecol. Manage.* **2012**, *20*, 461–475.
- McClain, M. E.; Naiman, R. J. Andean Influences on the Biogeochemistry and Ecology of the Amazon River. *BioScience* **2008**, *58*, 325–338.
- Gonsior, M.; Valle, J.; Schmitt-Kopplin, P.; Hertkorn, N.; Bastviken, D.; Luek, J.; Harir, M.; Bastos, W.; Enrich-Prast, A. Chemodiversity of dissolved organic matter in the Amazon Basin. *Biogeochemistry* **2016**, *13*, 4279–4290.
- Amon, R. M. W.; Benner, R. Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River system. *Geochim. Cosmochim. Acta* **1996**, *60*, 1783–1792.
- Ríos-Villamizar, E. A.; Adeney, J. M.; Piedade, M. T. F.; Junk, W. J. New insights on the classification of major Amazonian river water types. *Sustain. Water Resour. Manag.* **2020**, *6*, 83.
- Subdiaga, E.; Orsetti, S.; Haderlein, S. B. Effects of Sorption on Redox Properties of Natural Organic Matter. *Environ. Sci. Technol.* **2019**, *53*, 14319–14328.
- Duarte, R. M.; Menezes, A. C. L.; da Silveira Rodrigues, L.; de Almeida-Val, V. M. F.; Val, A. L. Copper sensitivity of wild ornamental fish of the Amazon. *Environ. Ecotoxicol. Environ. Saf.* **2009**, *72*, 693–698.
- Simon, C.; Osterholz, H.; Koschinsky, A.; Dittmar, T. Riverine mixing at the molecular scale – An ultrahigh-resolution mass spectrometry study on dissolved organic matter and selected metals in the Amazon confluence zone (Manaus, Brazil). *Org. Geochem.* **2019**, *129*, 45–62.
- Li, S.; Harir, M.; Schmitt-Kopplin, P.; Gonsior, M.; Enrich-Prast, A.; Bastviken, D.; Valle, J.; Machado-Silva, F.; Hertkorn, N. Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy. *Sci. Total Environ.* **2023**, *857*, 159620.
- Farjalla, V. F.; et al. Are the mixing zones between aquatic ecosystems hot spots of bacterial production in the Amazon River system? *Hydrobiologia* **2014**, *728*, 153–165.
- McClain, M. E.; Boyer, E. W.; Dent, C. L.; Gergel, S. E.; Grimm, N. B.; Groffman, P. M.; Hart, S. C.; Harvey, J. W.; Johnston, C. A.; Mayorga, E.; et al. Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems. *Ecosystems* **2003**, *6*, 301–312.
- Seidel, M.; Dittmar, T.; Ward, N. D.; Krusche, A. V.; Richey, J. E.; Yager, P. L.; Medeiros, P. M. Seasonal and spatial variability of dissolved organic matter composition in the lower Amazon River. *Biogeochemistry* **2016**, *131*, 281–302.
- Hertkorn, N.; Harir, M.; Cawley, K. M.; Schmitt-Kopplin, P.; Jaffé, R. Molecular characterization of dissolved organic matter from subtropical wetlands: a comparative study through the analysis of optical properties, NMR and FTICR/MS. *Biogeochemistry* **2016**, *13*, 2257–2277.
- Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M. Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Anal. Chem.* **2008**, *80*, 8908–8919.
- Drake, T. W.; Hemingway, J. D.; Kurek, M. R.; Peucker-Ehrenbrink, B.; Brown, K. A.; Holmes, R. M.; Galy, V.; Moura, J. M.; Mitsuya, M.; Wassenaar, L. I.; et al. The Pulse of the Amazon: Fluxes of Dissolved Organic Carbon, Nutrients, and Ions From the World's Largest River. *Global Biogeochem. Cycles* **2021**, *35*, No. e2020GB006895.
- Santos-Junior, C. D.; Sarmiento, H.; de Miranda, F. P.; Henrique-Silva, F.; Logares, R. Uncovering the genomic potential of the Amazon River microbiome to degrade rainforest organic matter. *Microbiome* **2020**, *8*, 151.
- Lundeen, R. A.; Janssen, E. M. L.; Chu, C.; McNeill, K. Environmental Photochemistry of Amino Acids, Peptides and Proteins. *Chimia* **2014**, *68*, 812–817.
- Hertkorn, N.; Benner, R.; Frommberger, M.; Schmitt-Kopplin, P.; Witt, M.; Kaiser, K.; Kettrup, A.; Hedges, J. I. Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* **2006**, *70*, 2990–3010.
- Schmitt-Kopplin, P.; Hertkorn, N.; Schulten, H. R.; Kettrup, A. Structural Changes in a Dissolved Soil Humic Acid during Photochemical Degradation Processes under O<sub>2</sub> and N<sub>2</sub> Atmosphere. *Environ. Sci. Technol.* **1998**, *32*, 2531–2541.
- Gonsior, M.; Hertkorn, N.; Conte, M. H.; Cooper, W. J.; Bastviken, D.; Druffel, E.; Schmitt-Kopplin, P. Photochemical production of polyols arising from significant photo-transformation of dissolved organic matter in the oligotrophic surface ocean. *Mar. Chem.* **2014**, *163*, 10–18.
- Harir, M.; Cawley, K. M.; Hertkorn, N.; Schmitt-Kopplin, P.; Jaffé, R. Molecular and spectroscopic changes of peat-derived organic matter following photo exposure: Effects on heteroatom composition of DOM. *Sci. Total Environ.* **2022**, *838*, 155790.
- Fudyma, J. D.; Chu, R. K.; Graf Grachet, N.; Stegen, J. C.; Tfaily, M. M. Coupled Biotic-Abiotic Processes Control Biogeochemical Cycling of Dissolved Organic Matter in the Columbia River Hyporheic Zone. *Front. Water* **2021**, *2*, 574962.
- Huang, X.; et al. Transformation of algal-dissolved organic matter via sunlight-induced photochemical and microbial processes: interactions between two processes. *Environ. Sci. Pollut. Res.* **2023**, *30*, 52969–52981.
- Moreira-Turcq, P. F.; Seyler, P.; Guyot, J. L.; Etcheber, H. Characteristics of organic matter in the mixing zone of the Rio Negro and Rio Solimões of the Amazon River. *Hydrol. Processes* **2003**, *17*, 1393–1404.
- Mayorga, E.; Aufdenkampe, A. K.; Masiello, C. A.; Krusche, A. V.; Hedges, J. I.; Quay, P. D.; Richey, J. E.; Brown, T. A. Young organic matter as a source of carbon dioxide outgassing from Amazonian rivers. *Nature* **2005**, *436*, 538–541.
- Abril, G.; Martinez, J. M.; Artigas, L. F.; Moreira-Turcq, P.; Benedetti, M. F.; Vidal, L.; Meziane, T.; Kim, J. H.; Bernardes, M. C.;

- Savoye, N.; et al. Amazon River carbon dioxide outgassing fuelled by wetlands. *Nature* **2014**, *505*, 395–398.
- (29) Ward, N. D.; Sawakuchi, H. O.; Neu, V.; Less, D. F. S.; Valerio, A. M.; Cunha, A. C.; Kampel, M.; Bianchi, T. S.; Krusche, A. V.; Richey, J. E.; et al. Velocity-amplified microbial respiration rates in the lower Amazon River. *Limnol. Oceanogr. Lett.* **2018**, *3*, 265–274.
- (30) Dittmar, T.; Koch, B.; Hertkorn, N.; Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr.: Methods* **2008**, *6*, 230–235.
- (31) Tziotis, D.; Hertkorn, N.; Schmitt-Kopplin, P. Kendrick-Analogous Network Visualisation of Ion Cyclotron Resonance Fourier Transform Mass Spectra: Improved Options for the Assignment of Elemental Compositions and the Classification of Organic Molecular Complexity. *Eur. J. Mass Spectrom.* **2011**, *17*, 415–421.
- (32) Seidel, M.; Yager, P. L.; Ward, N. D.; Carpenter, E. J.; Gomes, H. R.; Krusche, A. V.; Richey, J. E.; Dittmar, T.; Medeiros, P. M. Molecular-level changes of dissolved organic matter along the Amazon River-to-ocean continuum. *Mar. Chem.* **2015**, *177*, 218–231.
- (33) Santoro, A. L.; Bastviken, D.; Gudasz, C.; Tranvik, L.; Enrich-Prast, A. Dark carbon fixation: an important process in lake sediments. *PLoS One* **2013**, *8*, No. e65813.
- (34) Hertkorn, N.; Harir, M.; Koch, B. P.; Michalke, B.; Schmitt-Kopplin, P. High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* **2013**, *10*, 1583–1624.
- (35) Hsu, C. S.; Qian, K.; Chen, Y. C. An innovative approach to data analysis in hydrocarbon characterization by on-line liquid chromatography-mass spectrometry. *Anal. Chim. Acta* **1992**, *264*, 79–89.
- (36) Valle, J.; Harir, M.; Gonsior, M.; Enrich-Prast, A.; Schmitt-Kopplin, P.; Bastviken, D.; Hertkorn, N. Molecular differences between water column and sediment pore water SPE-DOM in ten Swedish boreal lakes. *Water Res.* **2020**, *170*, 115320.
- (37) Koch, B. P.; Dittmar, T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 926–932.
- (38) Lechtenfeld, O. J.; Hertkorn, N.; Shen, Y.; Witt, M.; Benner, R. Marine sequestration of carbon in bacterial metabolites. *Nat. Commun.* **2015**, *6*, 6711.
- (39) Bianchi, T. S.; Ward, N. D. Editorial: The Role of Priming in Terrestrial and Aquatic Ecosystems. *Front. Earth Sci.* **2019**, *7*, 321.
- (40) Ward, N. D.; Sawakuchi, H. O.; Richey, J. E.; Keil, R. G.; Bianchi, T. S. Enhanced Aquatic Respiration Associated With Mixing of Clearwater Tributary and Turbid Amazon River Waters. *Front. Earth Sci.* **2019**, *7*, 101.
- (41) Benner, R.; Opsahl, S.; Chin-Leo, G.; Richey, J. E.; Forsberg, B. R. Bacterial carbon metabolism in the Amazon River system. *Limnol. Oceanogr.* **1995**, *40*, 1262–1270.
- (42) Vidal, L. O.; Abril, G.; Artigas, L. F.; Melo, M. L.; Bernardes, M. C.; Lobão, L. M.; Reis, M. C.; Moreira-Turcq, P.; Benedetti, M.; Tomisielo, V. L.; et al. Hydrological pulse regulating the bacterial heterotrophic metabolism between Amazonian mainstems and floodplain lakes. *Front. Microbiol.* **2015**, *6*, 1054.
- (43) Wissmar, R.; Richey, J. E.; Stallard, R. F.; Edmond, J. M. Plankton metabolism and carbon processes in the Amazon River, its tributaries, and floodplain waters, Peru-Brazil, May–June 1977. *Ecology* **1981**, *62*, 1622–1633.
- (44) Enrich-Prast, A. et al. Chemosynthesis. In *Encyclopedia of Inland Waters*, 2nd ed.; Mehner, T.; Tockner, K., Eds.; Elsevier, 2022; pp 118–135, ISBN 9780128220412.
- (45) Signori, C. N.; Valentin, J. L.; Pollery, R. C. G.; Enrich-Prast, A. Temporal variability of dark carbon fixation and bacterial production and their relation with environmental factors in a tropical estuarine system. *Front. Aquat. Microbiol.* **2018**, *41*, 1089–1101.
- (46) Signori, C. N.; Felizardo, J. P. d. S.; Enrich-Prast, A. Bacterial production prevails over photo- and chemosynthesis in a eutrophic tropical lagoon. *Science* **2020**, *243*, 106889.
- (47) Casamayor, E. O.; García-Cantizano, J.; Mas, J.; Pedrós-Alió, C. Primary production in estuarine oxic/anoxic interfaces: contribution of microbial dark CO<sub>2</sub> fixation in the Ebro River Salt Wedge Estuary. *Mar. Ecol. Prog. Ser.* **2001**, *215*, 49–56.
- (48) Nogueroles, I.; Picazo, A.; Lliros, M.; Camacho, A.; Borrego, C. M. Diversity of freshwater *Epsilonproteobacteria* and dark inorganic carbon fixation in the sulphidic redoxcline of a meromictic karstic lake. *FEMS Microbiol. Ecol.* **2015**, *91*, fiv086.
- (49) Di Nezio, F.; et al. Anoxygenic photo- and chemo-synthesis of phototrophic sulfur bacteria from an alpine meromictic lake. *FEMS Microbiol. Ecol.* **2021**, *97*, fiab010.
- (50) Overmann, J.; Beatty, J. T.; Hall, K. J. Purple sulfur bacteria control the growth of aerobic heterotrophic bacterioplankton in a meromictic salt lake. *Appl. Environ. Microbiol.* **1996**, *62*, 3251–3258.
- (51) Saini, J. S.; Hassler, C.; Cable, R.; Fourquez, M.; Danza, F.; Roman, S.; Tonolla, M.; Storelli, N.; Jacquet, S.; Zdobnov, E. M.; et al. Bacterial, Phytoplankton, and Viral Distributions and Their Biogeochemical Contexts in Meromictic Lake Cadagno Offer Insights into the Proterozoic Ocean Microbial Loop. *Mbio* **2022**, *13*, No. e0005222.
- (52) Andersson, M. G.; Brion, N.; Middelburg, J. Comparison of nitrifier activity versus growth in the Scheldt estuary—a turbid, tidal estuary in northern Europe. *Aquat. Microb. Ecol.* **2006**, *42*, 149–158.
- (53) Bräuer, S. L.; Kranzler, K.; Goodson, N.; Murphy, D.; Simon, H. M.; Baptista, A. M.; Tebo, B. M. Dark carbon fixation in the Columbia River's estuarine turbidity maxima: molecular characterization of red-type cbbL genes and measurement of DIC uptake rates in response to added electron donors. *Estuaries Coasts* **2013**, *36*, 1073–1083.
- (54) Calderón-Paz, J. I.; García-Cantizano, J.; Vaqué, D.; Pedrós-Alió, C. Heterotrophic bacterial production in systems of the northern Spanish Mediterranean Region. *Int. Ver. Theor. Angew. Limnol. Verh.* **1993**, *25*, 739–742.
- (55) Crump, B. C.; Baross, J.; Simenstad, C. Dominance of particle-attached bacteria in the Columbia River estuary, USA. *Aquat. Microb. Ecol.* **1998**, *14*, 7–18.
- (56) Goosen, N. K.; van Rijswijk, P.; Kromkamp, J.; Peene, J. Regulation of annual variation in heterotrophic bacterial production in the Schelde estuary (SW Netherlands). *Aquat. Microb. Ecol.* **1997**, *12*, 223–232.
- (57) Ertel, J. R.; Hedges, J. I.; Devol, A. H.; Richey, J. E.; Ribeiro, M. d. N. G. Dissolved humic substances of the Amazon River system. *Limnol. Oceanogr.* **1986**, *31*, 739–754.
- (58) Johannsson, O. E.; Smith, D. S.; Sadauskas-Henrique, H.; Cimprich, G.; Wood, C. M.; Val, A. L. Photo-oxidation processes, properties of DOC, reactive oxygen species (ROS), and their potential impacts on native biota and carbon cycling in the Rio Negro (Amazonia, Brazil). *Hydrobiologia* **2017**, *789*, 7–29.
- (59) Waggoner, D. C.; Chen, H.; Willoughby, A. S.; Hatcher, P. G. Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. *Org. Geochem.* **2015**, *82*, 69–76.
- (60) Aufdenkampe, A. K.; Hedges, J. I.; Richey, J. E.; Krusche, A. V.; Llerena, C. A. Sorptive fractionation of dissolved organic nitrogen and amino acids onto fine sediments within the Amazon Basin. *Limnol. Oceanogr.* **2001**, *46*, 1921–1935.
- (61) Roth, V. N.; Lange, M.; Simon, C.; Hertkorn, N.; Bucher, S.; Goodall, T.; Griffiths, R. I.; Mellado-Vázquez, P. G.; Mommer, L.; Oram, N. J.; et al. Persistence of dissolved organic matter explained by molecular changes during its passage through soil. *Nat. Geosci.* **2019**, *12*, 755–761.
- (62) Hagerman, A. E.; et al. Tannins and lignins. *Herbivores: Their Interact. Second. Plant Metab.* **1991**, *1*, 355–388.
- (63) Schmitt-Kopplin, P.; Gelencsér, A.; Dabek-Zlotorzynska, E.; Kiss, G.; Hertkorn, N.; Harir, M.; Hong, Y.; Gebefügi, I. Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh resolution mass spectrometry and nuclear magnetic resonance spectroscopy: Organosulfates as photochemical smog constituents. *Anal. Chem.* **2010**, *82*, 8017–8026.



- (64) Aucour, A. M.; Tao, F. X.; Moreira-Turcq, P.; Seyler, P.; Sheppard, S.; Benedetti, M. The Amazon River: behaviour of metals (Fe, Al, Mn) and dissolved organic matter in the initial mixing at the Rio Negro/Solimões confluence. *Chem. Geol.* **2003**, *197*, 271–285.
- (65) Pérez, M. A.; Moreira-Turcq, P.; Gallard, H.; Allard, T.; Benedetti, M. F. Dissolved organic matter dynamic in the Amazon basin: sorption by mineral surfaces. *Chem. Geol.* **2011**, *286*, 158–168.
- (66) Subdiaga, E.; Harir, M.; Orsetti, S.; Hertkorn, N.; Schmitt-Kopplin, P.; Haderlein, S. B. Preferential Sorption of Tannins at Aluminum Oxide Affects the Electron Exchange Capacities of Dissolved and Sorbed Humic Acid Fractions. *Environ. Sci. Technol.* **2020**, *54*, 1837–1847.
- (67) Merschel, G.; Bau, M.; Dantas, E. L. Contrasting impact of organic and inorganic nanoparticles and colloids on the behavior of particle-reactive elements in tropical estuaries: an experimental study. *Geochim. Cosmochim. Acta* **2017**, *197*, 1–13.
- (68) Medeiros, P. M.; Seidel, M.; Ward, N. D.; Carpenter, E. J.; Gomes, H. R.; Niggemann, J.; Krusche, A. V.; Richey, J. E.; Yager, P. L.; Dittmar, T. Fate of the Amazon River dissolved organic matter in the tropical Atlantic Ocean. *Global Biogeochem. Cycles* **2015**, *29*, 677–690.
- (69) Miao, Y.; Lv, J.; Huang, H.; Cao, D.; Zhang, S. Molecular characterization of root exudates using Fourier transform ion cyclotron resonance mass spectrometry. *J. Environ. Sci.* **2020**, *98*, 22–30.
- (70) Doherty, M.; Yager, P. L.; Moran, M. A.; Coles, V. J.; Fortunato, C. S.; Krusche, A. V.; Medeiros, P. M.; Payet, J. P.; Richey, J. E.; Satinsky, B. M.; et al. Bacterial Biogeography across the Amazon River-Ocean Continuum. *Front. Microbiol.* **2017**, *8*, 882.
- (71) Benner, R.; Amon, R. M. The size-reactivity continuum of major bioelements in the ocean. *Ann. Rev. Mar. Sci.* **2015**, *7*, 185–205.
- (72) Bertassoli, D. J.; Sawakuchi, A. O.; Sawakuchi, H. O.; Pupim, F. N.; Hartmann, G. A.; McGlue, M. M.; Chiessi, C. M.; Zabel, M.; Schefuß, E.; Pereira, T. S.; et al. The Fate of Carbon in Sediments of the Xingu and Tapajós Clearwater Rivers, Eastern Amazon. *Front. Mar. Sci.* **2017**, *4*, 44.
- (73) Melo, M. L.; Kothawala, D. N.; Bertilsson, S.; Amaral, J. H.; Forsberg, B.; Sarmento, H. Linking dissolved organic matter composition and bacterioplankton communities in an Amazon floodplain system. *Limnol. Oceanogr.* **2019**, *65*, 63–76.
- (74) Ward, N. D.; Bianchi, T. S.; Sawakuchi, H. O.; Gagne-Maynard, W.; Cunha, A. C.; Brito, D. C.; Neu, V.; de Matos Valerio, A.; da Silva, R.; Krusche, A. V.; et al. The reactivity of plant-derived organic matter and the potential importance of priming effects along the lower Amazon River. *J. Geophys. Res.* **2016**, *121*, 1522–1539.
- (75) Hawkes, J. A.; Hansen, C. T.; Goldhammer, T.; Bach, W.; Dittmar, T. Molecular alteration of marine dissolved organic matter under experimental hydrothermal conditions. *Geochim. Cosmochim. Acta* **2016**, *175*, 68–85.
- (76) Raeke, J.; Lechtenfeld, O. J.; Wagner, M.; Herzsprung, P.; Reemtsma, T. Selectivity of solid phase extraction of freshwater dissolved organic matter and its effect on ultrahigh resolution mass spectra. *Environ. Sci.: Processes Impacts* **2016**, *18*, 918–927.
- (77) Li, Y.; Harir, M.; Uhl, J.; Kanawati, B.; Lucio, M.; Smirnov, K. S.; Koch, B. P.; Schmitt-Kopplin, P.; Hertkorn, N. How representative are dissolved organic matter (DOM) extracts? A comprehensive study of sorbent selectivity for DOM isolation. *Water Res.* **2017**, *116*, 316–323.
- (78) Kujawinski, E. B.; Del Vecchio, R.; Blough, N. V.; Klein, G. C.; Marshall, A. G. Probing molecular-level transformations of dissolved organic matter: insights on photochemical degradation and protozoan modification of DOM from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mar. Chem.* **2004**, *92*, 23–37.
- (79) Vannote, R. L.; Minshall, G. W.; Cummins, K. W.; Sedell, J. R.; Cushing, C. E. The River Continuum Concept. *Can. J. Fish. Aquat. Sci.* **1980**, *37*, 130–137.
- (80) Kamjunke, N.; Hertkorn, N.; Harir, M.; Schmitt-Kopplin, P.; Griebler, C.; Brauns, M.; von Tümpling, W.; Weitere, M.; Herzsprung, P. Molecular change of dissolved organic matter and patterns of bacterial activity in a stream along a land-use gradient. *Water Res.* **2019**, *164*, 114919.
- (81) Devol, A. H.; Hedges, J. I. Organic matter and nutrients in the mainstem Amazon River. *Biogeochem. Amazon Basin* **2001**, *15*, 275–306.
- (82) Schmidt, M. W.; Torn, M. S.; Abiven, S.; Dittmar, T.; Guggenberger, G.; Janssens, I. A.; Kleber, M.; Kögel-Knabner, I.; Lehmann, J.; Manning, D. A. C.; et al. Persistence of soil organic matter as an ecosystem property. *Nature* **2011**, *478*, 49–56.
- (83) Valle, J.; Gonsior, M.; Harir, M.; Enrich-Prast, A.; Schmitt-Kopplin, P.; Bastviken, D.; Conrad, R.; Hertkorn, N. Extensive processing of sediment pore water dissolved organic matter during anoxic incubation as observed by high-field mass spectrometry (FTICR-MS). *Water Res.* **2018**, *129*, 252–263.
- (84) Einsiedl, F.; Hertkorn, N.; Wolf, M.; Frommberger, M.; Schmitt-Kopplin, P.; Koch, B. P. Rapid biotic molecular transformation of fulvic acids in a karst aquifer. *Geochim. Cosmochim. Acta* **2007**, *71*, 5474–5482.
- (85) Eadie, B.; Jeffrey, L.; Sackett, W. Some observations on the stable carbon isotope composition of dissolved and particulate organic carbon in the marine environment. *Geochim. Cosmochim. Acta* **1978**, *42*, 1265–1269.
- (86) Meyers-Schulte, K. J.; Hedges, J. I. Molecular evidence for a terrestrial component of organic matter dissolved in ocean water. *Nature* **1986**, *321*, 61–63.
- (87) Opsahl, S.; Benner, R. Distribution and cycling of terrigenous dissolved organic matter in the ocean. *Nature* **1997**, *386*, 480–482.
- (88) Bianchi, T. S. The role of terrestrially derived organic carbon in the coastal ocean: a changing paradigm and the priming effect. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 19473–19481.
- (89) Jiao, N.; Herndl, G. J.; Hansell, D. A.; Benner, R.; Kattner, G.; Wilhelm, S. W.; Kirchman, D. L.; Weinbauer, M. G.; Luo, T.; Chen, F.; Azam, F. Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nat. Rev. Microbiol.* **2010**, *8*, 593–599.
- (90) Baltar, F.; Herndl, G. J. Ideas and perspectives: Is dark carbon fixation relevant for oceanic primary production estimates? *Biogeochemistry* **2019**, *16*, 3793–3799.
- (91) Jorcín, A.; Nogueira, M. G. Temporal and spatial patterns based on sediment and sediment–water interface characteristics along a cascade of reservoirs (Parapanema River, south-east Brazil). *Lakes Reserv.: Res. Manag.* **2005**, *10*, 1–12.

Dear Siyu Li,

In reference to your message below, the following is from the **American Chemical Society's Policy on Theses and Dissertations**:

This policy addresses permission to include **your article(s)** or portions of text from **your article(s)** in your thesis.

**Reuse/Republication of the Entire Work in Theses or Collections:** Authors may reuse all or part of the Submitted, Accepted or Published Work in a thesis or dissertation that the author writes and is required to submit to satisfy the criteria of degree-granting institutions. Such reuse is permitted subject to the ACS' "[Ethical Guidelines to Publication of Chemical Research](#)". Appropriate citation of the Published Work must be made as follows:

"Reprinted with permission from [COMPLETE REFERENCE CITATION]. Copyright [YEAR] American Chemical Society." Insert the appropriate wording in place of the capitalized words. Citation information may be found after the "Cite this:" heading below the title of the online version and at the bottom of **the first page of the pdf or print version of your ACS journal article**.

If the thesis or dissertation to be published is in electronic format, a direct link to the Published Work must also be included using the [ACS Articles on Request](#) author-directed link.

If your university requires written permission and your manuscript has not yet received a DOI (published ASAP), send a request to [copyright@acs.org](mailto:copyright@acs.org) that includes the manuscript number, the name of the ACS journal, and the date that you need to receive our reply.

If your university requires you to obtain permission for manuscripts in ASAP status or final published articles, you must use the RightsLink permission system. See [RightsLink instructions at http://pubs.acs.org/page/copyright/permissions.html](http://pubs.acs.org/page/copyright/permissions.html) and make requests at the "Rights & Permissions" link under the title of the online version of the article.

**Submission to a Dissertation Distributor:** If you plan to submit your thesis to UMI or to another dissertation distributor, you should not include the unpublished ACS paper in your thesis if the thesis will be disseminated electronically, until ACS has published your paper. After publication of the paper by ACS, you may release the entire thesis (**not the individual ACS article by itself**) for electronic dissemination through the distributor; ACS's copyright credit line should be printed on the first page of the ACS paper.

As such, please consider this email confirmation that your request to use the material as described is granted.

I hope this information is helpful.

Sincerely,

Keri Travis (*she/her*)  
Copyright Specialist  
Office of the Secretary & General Counsel  
American Chemical Society  
1155 Sixteenth Street, NW  
Washington, DC 20036  
202-872-4368  
202-604-2874  
[k\\_travis@acs.org](mailto:k_travis@acs.org)



## 6.3 Supplementary information for article 1

### Supporting Online Information

#### Comprehensive assessment of dissolved organic matter processing in the Amazon River and its major tributaries revealed by positive and negative electrospray mass spectrometry and NMR spectroscopy

##### Authors

Siyu Li<sup>a</sup>, Mourad Harir<sup>a,b</sup>, Philippe Schmitt-Kopplin<sup>a,b</sup>, Michael Gonsior<sup>c</sup>, Alex-Enrich-Prast<sup>d,e</sup>, David Bastviken<sup>d</sup>, Juliana Valle<sup>a</sup>, Fausto Machado-Silva<sup>f,g</sup>, Norbert Hertkorn<sup>a\*</sup>

##### Affiliations

<sup>a</sup> Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit Analytical Biogeochemistry (BGC); Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany.

<sup>b</sup> Chair of Analytical Food Chemistry, Technische Universität München; Alte Akademie 10, 85354, Freising-Weißenstephan, Germany.

<sup>c</sup> University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory; 146 Williams Street, Solomons, Maryland 20688, United States.

<sup>d</sup> Department of Thematic Studies - Environmental Change, Linköping University; SE-581 83 Linköping, Sweden.

<sup>e</sup> Multiuser Unit of Environmental Analysis (UMAA/UFRJ); Federal University of Rio de Janeiro, Rio de Janeiro, Brazil.

<sup>f</sup> Program in Geosciences - Environmental Geochemistry, Chemistry Institute, Fluminense Federal University, 24020-141, Niterói, Brazil.

<sup>g</sup> Department of Environmental Sciences, University of Toledo, Toledo, OH, 43606, USA.

**\*Corresponding Author:** Email: [hertkorn@helmholtz-muenchen.de](mailto:hertkorn@helmholtz-muenchen.de); Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Research Unit Analytical Biogeochemistry (BGC); Ingolstaedter Landstrasse 1, 85764 Neuherberg, Germany.

Number of pages: 34  
Number of Tables: 8  
Number of Figures: 18

## Table of Contents

Details of Material and Methods	page
1. Sampling and site locations.....	3-4
2. FT-ICR mass spectrometry.....	4
3. <sup>1</sup> H NMR spectroscopy.....	4
4. Water characterization.....	4-5
5. Dark Carbon Fixation (DCF) and Heterotrophic Bacterial Production (HBP).....	5-6
6. Statistical Analysis.....	6

## List of Figures and Tables

Fig. S1. Maps of the sampling sites in the Amazon River.

Fig. S2. (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra of in the four Amazon River main stem and tributaries.

Table S1. Description of the sampling sites and sampling information.

Table S2. Water parameters in the Amazon River.

Table S3. ESI[-] FT-ICR MS of DOM in the Amazon River and tributaries (AZ-DOM) derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in AZ-DOM.

Table S4. ESI[+] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in AZ-DOM.

Table S5. <sup>1</sup>H NMR section integrals (percent of non-exchangeable protons) and key substructures of AZ-DOM samples (CD<sub>3</sub>OD, exclusion of residual water, and methanol).

Fig. S3. <sup>1</sup>H NMR spectra of four consolidated AZ-DOM, including DOM in the Negro River (N-DOM), in the Amazon River (A-DOM), in the Solimões River (S-DOM) and in the Tapajós River (T-DOM).

Fig. S4. Principal component analysis (PCA) of AZ-DOM molecular composition around SN mixing zone based on all assigned molecular formulae in ESI[±] FT-ICR MS spectra.

Fig. S5. Hierarchical cluster analysis (HCA) of AZ-DOM composition proximate to the mixing zone of Solimões and Negro (SN) based on all assigned molecular formulae in ESI[±] FT-ICR MS spectra.

Fig. S6. HCA of A-DOM composition proximate to the Madeira River inflow to the Amazon River base on all assigned molecular formulae in ESI[-] (A) and ESI[+] (B) FT-ICR MS spectra.

Fig. S7. <sup>1</sup>H NMR spectra of A-DOM proximate to the Madeira River inflow to the Amazon River

Fig. S8. ESI[±] FT-ICR mass spectra and molecular composition of the common *m/z* ions in A-DOM.

Fig. S9. PCA of A-DOM composition based on all assigned molecular formulae in ESI[±] FT-ICR MS spectra.

Fig. S10. Mass peaks identified in ESI[-] and ESI[+]FT-ICR mass spectra, of which the normalized intensity have increased/declined trend along the Amazon River (along fluvial flow from AM1R to AM11R).

Fig. S11. The number of frequency distribution of different numbers of oxygen atoms per molecules in S-DOM, N-DOM, A-DOM, and T-DOM from (A) ESI[-] and (B) ESI[+]FT-ICR MS.

Fig. S12. Van Krevelen and mass edited H/C diagrams of the molecular formulae that were (A) unique in S-DOM compared to N-DOM, (B) unique in N-DOM compared to S-DOM, and (C) unique in MZ compared to N-DOM and S-DOM.

Fig. S13. Van Krevelen and mass edited H/C diagrams of the molecular formulae that were (A) unique in T-DOM and (B) were not present in T-DOM.

Fig. S14. PCA of AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular formulae in ESI[±] FT-ICR MS.

**Fig. S15. PCA of AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular formulae in ESI[±] FT-ICR MS spectra.**

**Fig. S16. PCA of AZ-DOM composition in the four Amazon main stem and tributaries based on <sup>1</sup>H NMR spectra (800 MHz <sup>1</sup>H NMR, CD<sub>3</sub>OD, area-normalized from 0.5-10.0 ppm; 0.01 ppm bucket resolution; with exclusion of residual water and methanol NMR resonances).**

**Table S6. Heterotrophic bacterial production (HBP) and dark carbon fixation (DCF) in the river waters.**

**Fig. S17. NMDS analysis using AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular formulae in (A, B) ESI[±] FT-ICR mass spectra, as well as (C) <sup>1</sup>H NMR spectra in 0.01 ppm bucket resolution fitted with water parameters, rates of bacterial production and carbon fixation shown as vectors.**

**Table S7. Statistical significance of the correlation of environmental parameters on the DOM molecular composition in Amazon basin (corresponds to Fig. S16) based on ESI[±] FT-ICR MS and <sup>1</sup>H NMR in NMDS model.**

**Table S8. Statistical significance of the correlation of environmental parameters, bacterial production and carbon fixation on the DOM molecular composition in Amazon basin (corresponds to Fig. 5) based on ESI[±] FT-ICR MS and <sup>1</sup>H NMR in non-metric multidimensional scaling (NMDS) model.**

**Fig. S18. Spearman correlation analysis using AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular formulae in ESI[±] FT-ICR MS with DOC, POC, pH, and heterotrophic bacterial fixation (HBP) rate.**

## Details of Material and Methods

### 1. Sampling and site locations

The Amazon River main stem originates in Andean mountains springs and for most of its flow path meanders across the lowland of Amazon tropical rainforest until its mouth in Atlantic Ocean, while collecting major tributaries on its way. The main stem of the Amazon River is named Ucayali while flowing in Peru and when it crosses Brazil's boundary it changes the name to Solimões. Downstream of the confluence with the Negro River in Manaus (Amazonas state, Brazil) it changes the name again to Amazon River and keeps this name until its mouth on Atlantic Ocean in Brazil's coast. There are many large tributaries with different water characteristics. Three fundamental types of water were already described by Alfred Russel Wallace in 1853 in his trip to the Amazon basin, and these were called as white, clear and blackwaters (McClain and Naiman, 2008). The color varies depending on whether dissolved organic matter (black water) or suspended sediment (white water) is predominant (Moreira-Turcq et al., 2003). The clear water rivers are transparent and supply a high phytoplankton production (Richey et al., 1990). Both white and clear waters have more aquatic plants and floating meadows than black waters. In contrast, the black water Negro River is nearly devoid of such vegetation. Solimões contributed the most carbon to the Amazon River, about 500 kg C s<sup>-1</sup> during the sampling high water period (Moreira-Turcq et al., 2003). Madeira is the largest and Negro is the second largest tributary in Amazon watershed. These two rivers together contribute more than 30% of discharge of the Amazon River main stem. Madeira is a primary tributary of Amazon in terms of sediment input.

We collected water samples at 31 sites in the Amazon River main stem eastwards, including Solimões (white water), Negro (black water), Amazon (turbid water), and Tapajós (clear water) between April 2<sup>nd</sup> and May 25<sup>th</sup> in 2014, during a high-water period with extremely high discharge volume. The localities accessed were located near Manacapuru and Manaqueri in Solimões, Castanheira and Santa Maria in Negro, Manaus, Uruará and Óbidos in Amazon River and Alter do Chão and Aramanai in the Tapajós River.



Detail sampling information see Table S1. We obtained water samples by boat just below the surface, typically 10 L per sample. Solid phase extraction (SPE) of the water samples was done at the field immediately after sampling. The water column DOM was extracted by a previously described SPE method using 1 g cartridges of PPL resin (Dittmar et al., 2008). The eluates were stored in the freezer (-20°C) until mass spectrometry analysis and nuclear magnetic resonance spectroscopy analysis.

## 2. FT-ICR mass spectrometry

Negative and positive electrospray ionization (ESI[±]) Fourier transform ion cyclotron resonance mass spectra (FT-ICR MS) were acquired using a 12T Bruker Solarix mass spectrometer (Bruker Daltonics, Bremen, Germany) and an Apollo II electrospray ionization (ESI) source (Hertkorn et al., 2016). Nebulizer gas pressure, drying gas pressure and the source heater temperature were 138 kPa, 103 kPa and 200 °C, respectively. The spectra were acquired with a time domain of 4 MW. For each sample, 500 broadband scans were accumulated in ESI[-] FT-ICR mass spectra, while 300 broadband scans were accumulated in ESI[+] FT-ICR mass spectra. All spectra were first externally calibrated on clusters of arginine in MeOH (0.57 mmol/L) and internally calibrated using appropriate reference mass lists of common natural organic matter molecules, reaching accuracy values lower than 500 ppb. Data processing was done using Compass Data Analysis 4.1 (Bruker, Bremen, Germany) and formula assignment by an in-house made software (NetCalc) (Tziotis et al., 2011). The molecular formula assignments were based on the following elements:  $^1\text{H}_{0-200}$ ,  $^{12}\text{C}_{0-100}$ ,  $^{16}\text{O}_{0-80}$ ,  $^{32}\text{S}_{0-3}$ ,  $^{14}\text{N}_{0-3}$  as well as the  $^{13}\text{C}_{0-1}$  and  $^{34}\text{S}_{0-1}$  isotopomers. The generated formulae were validated by setting sensible chemical constraints [N rule, O/C ratio < 1, H/C ratio <  $2n + 2$  ( $\text{C}_n\text{H}_{2n+2}$ )]. Restriction on nitrogen atoms  $\leq 4$  and sulfur atoms  $\leq 2$  were applied based on previous studies (Hertkorn et al., 2013; Schmitt-Kopplin et al., 2010). Final elemental formulae were generated and categorized into groups containing CHO, CHNO, CHOS, and CHNOS molecular series, which were used to reconstruct the group-selective mass spectra. The intensity of mass peaks was normalized (total amplitude of assigned mass peaks = 100% for single sample). The average H/C, O/C, N/C, S/C atomic ratios, DBE/C (Double bond equivalent per carbon), mass-to-charge ratios ( $m/z$ ) were computed from the intensity-weighted average of molecular formulae (Hertkorn et al., 2016).

## 3. $^1\text{H}$ NMR spectroscopy

All  $^1\text{H}$  NMR spectra were acquired with a Bruker Advance III NMR spectroscopy operating at 800.35 MHz ( $B_0 = 18.8$  Tesla) at 283 K from redissolved solids in  $\text{CD}_3\text{OD}$  (99.95%  $^2\text{H}$ ; Merck) with Bruker standard pulse sequences using 3.0 mm Bruker MATCH tubes. The reference  $^1\text{H}$  NMR chemical shift of  $\text{HD}_2\text{COD}$  was 3.3 ppm.  $^1\text{H}$  NMR spectra were recorded under solvent suppression with presaturation and 1 ms spin-lock (noesypr1d), 5 s acquisition time, 5 s relaxation delay (d1), typically 1024 scans, and 1 Hz exponential line broadening.  $^1\text{H}$  NMR section integrals were obtained by using the software AMIX at 0.01 ppm resolution, with exclusions of HDO and  $\text{HD}_2\text{COD}$  NMR resonances.

## 4. Water characterization

Water conductivity (Cond.), temperature (Temp.) and dissolved oxygen (DO) were measured in situ with portable instruments (Hanna Instruments, Metrohm electrode and PRO-ODO YSI). For the inorganic dissolved nutrients, the water was sampled with polypropylene bottles, filtered in silica filters (GF/F, Whatman) and placed in amber bottles conditioned in freezer to be analyzed in the laboratory. Ammonium ( $\text{NH}_4^+$ ) concentrations were determined by colorimetric analysis with flux injection (FIA, FiaStar 5000). Nitrate ( $\text{NO}_3^-$ ) and nitrite ( $\text{NO}_2^-$ ) were measured using nitrate reduction and phosphate ( $\text{PO}_4^{3-}$ ) using molybdenum analytical method (Valderrama, 1981). The dissolved organic carbon (DOC) was determined by persulfate digestion analyzed using high-temperature catalytic oxidation method (Sievers InnovOx

analyzer, GE) (Gonsior et al., 2019). The total organic carbon (TOC) used the same analytical method as DOC without the water filtration.

The dissolved inorganic carbon (DIC) of water was determined by the acidified headspace method (Åberg and Wallin, 2014). During the field work, 25 mL glass vials were closed with a rubber stopper and pre-filled with 0.3 µL of concentrated H<sub>3</sub>PO<sub>4</sub> to acidify the sample to pH ~2. After that, 12.5 mL of water samples without bubbles were injected into the vial with a 0.6 mm needle. Three vials were closed in the field to sample the atmospheric air used in the calculations. Back in the laboratory, the headspace samples were measured by gas chromatography (7890A, Agilent Technologies, USA). Because the water was inoculated in a closed acidified ecosystem, all the DIC was transferred to CO<sub>2</sub>. The CO<sub>2</sub> headspace was measured and the CO<sub>2</sub> that remained in the water sample was equilibrated with the headspace (Henri's Law). DIC was the sum of both CO<sub>2</sub> concentration subtracted by the CO<sub>2</sub> in the atmosphere (Åberg and Wallin, 2014).

### 5. Dark Carbon Fixation (DCF) and Heterotrophic Bacterial Production (HBP)

DCF was estimated by the incorporation of H<sup>14</sup>CO<sub>3</sub><sup>-</sup> in a known time similar to photosynthesis, but in dark conditions (Nielsen, 1965; Santoro et al., 2013). We took care to start the incubations at night, 4 hours after the sunset, to wait the photosynthesis be negligible avoiding dark carbon incorporation by photosystem II (Nielsen, 1960). The incubations were performed by the addition of 10 µL of NaH<sup>14</sup>CO<sub>3</sub> (Perkin Elmer, specific activity of 52.5 mCi mmol<sup>-1</sup>) in 50 mL of river water in five dark (amber) bottles. The incubations were stopped after four hours with Formalin at final concentration of 3.7%. Two controls were prepared containing Formalin (3.7 % at final concentration) before adding the <sup>14</sup>CO<sub>2</sub> to take out the artifacts values. In lab, the samples were filtered twice inside a fume hood with Sartorius filtration kit. The first filtration was with silica filters GF/F (Whatman) and the second with cellulose acetate filters (0.2 µm pore, Sartorius Stedim Biotech). Both filters were washed with 1M HCl and dried to ensure the elimination of the non-assimilated <sup>14</sup>CO<sub>2</sub>. Filters were placed in scintillation vials with scintillation liquid (Optiphase HiSafe, PerkinElmer) in dark conditions to be analyzed in radiocarbon assay in scintillation counter 24h later to avoid chemiluminescence. Finally, DCF was computed as:

$$DCF (\mu\text{gC L}^{-1}\text{h}^{-1}) = (\text{Net DPM}_{\text{sample}} - \text{Net DPM}_{\text{control}}) \times \text{DPM}_{\text{added}}^{-1} \times 1.05 \times \text{DIC} \times t^{-1}$$

where DCF was the biomass production or carbon assimilation, DPM was the radioactivity measured by decay per minute of the sample, control and the added into incubations, 1.05 was the isotope discrimination factor, DIC was the dissolved inorganic carbon and t was time (Peterson, 1980).

HBP was determined by the measurement of protein synthesis rates using radio-labelled <sup>3</sup>H-leucine (D. C. Smith, 1992). This method measures the leucine incorporated by the bacteria in a known time after its addition. The methodology had some adjustments for the Amazon waters conditions according to previous studies (Benner et al., 1995; Farjalla, 2014). Lower leucine concentration and less time was important to not overestimate the bacterial production. 150 µL of <sup>3</sup>H-leucine with a final radioactivity of 10 Ci mmol<sup>-1</sup> and a leucine final concentration of 20 nM was added to 1.5 mL of river water into five Eppendorf tubes. The temperature was constant and similar to river water during the incubations that were performed in dark conditions. Formalin was added at final concentration of 3.7% to stop the incubation after 45 minutes. In addition, two controls were prepared with the same concentration of formalin before the leucine addition. At the end, the samples were placed in styrofoam containers to be analyzed in the laboratory.

In the lab, the bacterial protein was extracted using cold trichloroacetic acid (TCA) and ethanol. This method consists in the protein precipitation discarding the lighter non-incorporated radio-labelled leucine in the supernatant. Finally, the pellet was resuspended in the scintillation liquid (Optiphase HiSafe, PerkinElmer) with vortex and placed in dark for 24h after being analyzed in the radioassay with the scintillation counter. To compute the bacterial carbon production, we used the equation:

$$\text{HBP} (\mu\text{gC L}^{-1}\text{h}^{-1}) = \text{LI} (\mu\text{mols Leu L}^{-1}\text{h}^{-1}) \times 131.2 \times 0.073^{-1} \times 0.86^{-1} \times \text{ID}$$

where HBP was bacterial production, LI was the leucine incorporated and ID was the isotope dilution. 131.2 was the molecular weight of leucine, 0.073 was the proportion of leucine in total protein and 0.086 was the ratio of cellular carbon to protein (Simon and Azam, 1989).

## 6. Statistical Analysis

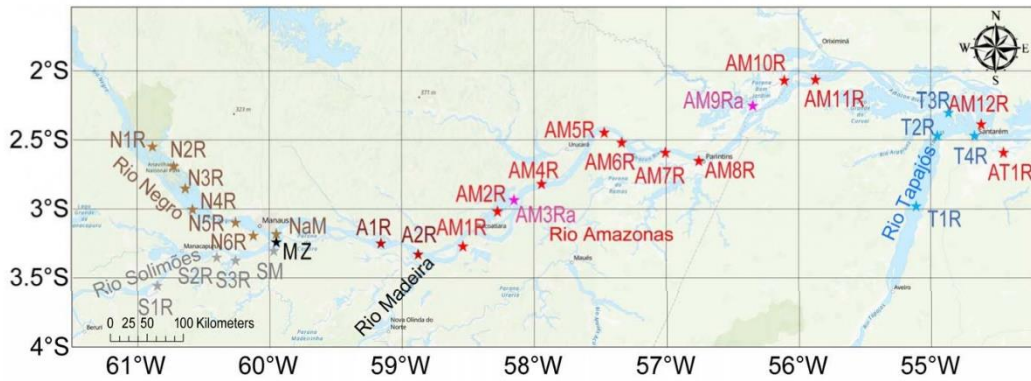
PCA was performed using Simca-P (version 11.5, UmetricsAB, Umeå, Sweden) to identify the dominant modes of variability in DOM composition of AZ-DOM. FT-ICR mass spectra were arranged with the samples as observations and the normalized peak amplitudes of the assigned FT-ICR mass peaks as the response variables. <sup>1</sup>H NMR spectra were arranged with the samples as observations and the NMR resonances as the response variables (800 MHz <sup>1</sup>H NMR, CD<sub>3</sub>OD, area-normalized from 0.5-10.0 ppm; 0.01 ppm bucket resolution; with the exclusion of residual water and methanol NMR resonances). Before multivariate statistics were performed, the response variables were centered and scaled to unit variance. The based weight was computed as 1/sqrt (standard deviation of the response variables).

HCA was performed using the Hierarchical Clustering Explorer 3.0 (HCE; <http://www.cs.umd.edu/hcil/multi-cluster/>). Average Linkage (UPGMA) method was used to cluster the dataset and Euclidean distance was used as the similarity/distance measure. Based on the HCA, we used the “profile search” tool from HCE 3.0, choosing a search method (model-based), a distance measure (Spearman's r) and a threshold (0.9).

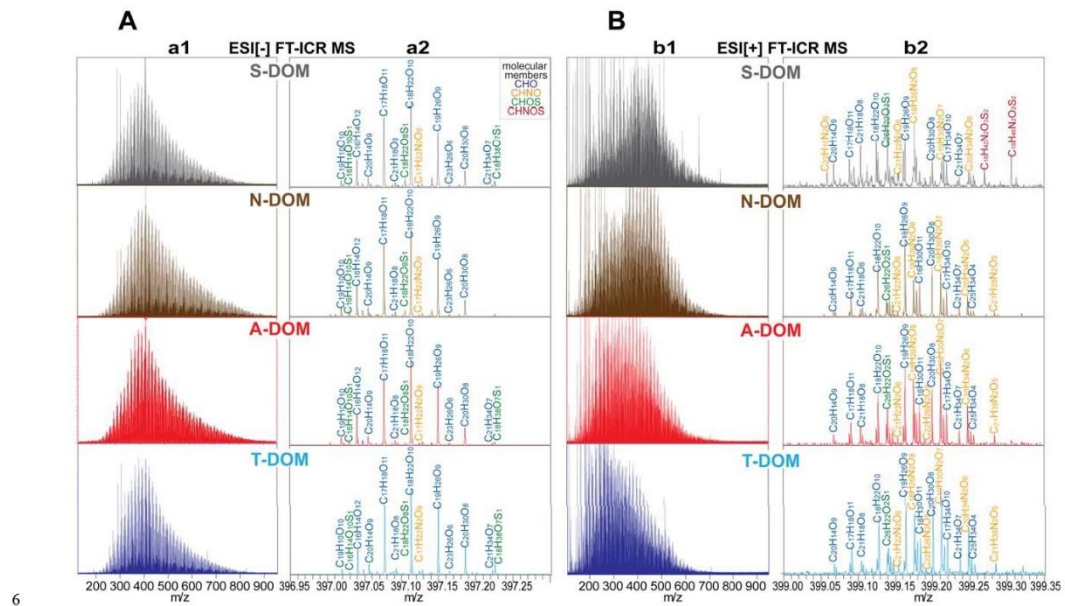
Spearman correlation analysis ( $p < 0.05$ ,  $r^2 > 0.5$ ) was performed using the R statistical platform with between normalized intensities of mass peaks in A-DOM samples downstream of the Madeira inflow (from AM1R to AM10R) and the fluvial distance of the sampling sites from sampling site AM1R. We used the same method to find significant correlation between AZ-DOM composition based on ESI[±] FT-ICR mass spectra and water/bacterial parameters.

Non-metric multidimensional scaling (NMDS) is particularly useful if the analyzed species (molecular formulae) are not linearly responding to environmental gradients (Ramette, 2007). We performed NMDS using the R statistical platform with the package *vegan* to determine correlations of environmental variables with DOM molecular composition in the Amazon River continuum. The environmental data used for NMDS included pH, dissolved oxygen (DO), conductivity (Cond.), temperature (Temp.), DOC, POC, TOC, as well as inorganic nutrients such as ammonium (NH<sub>4</sub><sup>+</sup>), nitrate (NO<sub>3</sub><sup>-</sup>), nitrite (NO<sub>2</sub><sup>-</sup>), and phosphate (PO<sub>4</sub><sup>3-</sup>). These physicochemical variables were centered and scaled for ordination. All NMDS models were run on Bray-Curtis distance matrixes and their statistical significance was tested with 999 permutations. We applied NMDS analysis for the DOM samples with the data of bacterial production and carbon fixation.





1  
 2 **Fig. S1. Maps of the sampling sites in the Amazon River.** “SM” (grey) was sampled in the Solimões River upstream and close to the Negro-Solimões River  
 3 mixing zone. “MZ” (black) was sampled at the very mixing zone. “A1R/A2R” (dark red) were sampled in the Amazon River upstream of the Madeira River  
 4 inflow, while the other Amazon River samples (red) were sampled after the Madeira River inflow. “AM3Ra” and “AM9Ra” (magenta) were sampled two  
 5 months later than the other water samples in the Amazon River.



6  
 7 **Fig. S2.** (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra of in the four Amazon River main stem and tributaries. S-DOM, N-DOM, A-DOM, and T-  
 8 DOM are exemplified with representative samples S1R, N1R, A1R, and T1R. The (A) ESI[-] and (B) ESI[+] FT-ICR MS spectra showed the dissimilar  
 9 distribution of mass peaks across a wide mass range ( $m/z$  150-950), as well as distinct signatures at nominal mass 398 confirming ionization selectivity.  
 10 Respective assignments of molecular compositions are provided for CHO (blue), CHNO (orange), CHOS (green), and CHNOS (red) molecules.



11 **Table S1. Description of the sampling sites and sampling information.** “SM” (grey) was sampled in the  
 12 Solimões River close to the mixing zone of the Negro River and the Solimões River. “MZ” (black) was sampled  
 13 at the mixing zone of the Negro River and the Solimões River. “A1R/A2R” (dark red) were sampled in the  
 14 Amazon River upstream of the Madeira River inflow, while the other Amazon River samples (red) were sampled  
 15 after the Madeira River inflow. “AM3Ra” and “AM9Ra” (magenta) were sampled two months later than the  
 16 other water samples in the Amazon River. The order of the rows in the table from top to bottom represents the  
 17 order of sampling point locations in each river from upstream to downstream.

River	Sample	Longitude (°W)	Latitude (°S)	Water type
Solimões (S)	S1R	-60.6845	-3.3414	white
Solimões	S2R	-60.1116	-3.3209	white
Solimões	S3R	-60.0957	-3.3136	white
Solimões close to SN mixing zone	SM	-59.8949	-3.1647	turbid
Negro (N)	N1R	-60.8198	-2.5740	black
Negro	N2R	-60.6865	-2.7267	black
Negro	N3R	-60.5509	-2.8719	black
Negro	N4R	-60.5583	-2.9440	black
Negro	N5R	-60.1810	-3.0958	black
Negro	N6R	-60.0902	-3.1242	black
Negro	NaM	-59.9126	-3.1292	black
SN mixing zone	MZ	-59.8958	-3.1301	turbid
Amazon River (A) before Madeira River inflow	A1R	-60.1596	-2.7933	turbid
Amazon River before Madeira River inflow	A2R	-58.7621	-3.2992	turbid
Amazon River after Madeira River inflow	AM1R	-58.5192	-3.1802	turbid
Amazon River	AM2R	-58.5513	-3.1492	turbid
Amazon River	AM3Ra	-58.1626	-2.9750	turbid
Amazon River (out of the river main stem because of too much water)	AM4R	-58.2533	-2.9530	turbid
Amazon River (sampled downstream of the Uatumã River inflow)	AM5R	-57.5167	-2.3797	turbid
Amazon River	AM6R	-57.4771	-2.4972	turbid
Amazon River	AM7R	-57.2811	-2.4676	turbid
Amazon River	AM8R	-57.2811	-2.4676	turbid
Amazon River	AM9Ra	-56.3486	-2.2473	turbid
Amazon River	AM10R	-55.7717	-2.0033	turbid
Amazon River	AM11R	-55.4316	-2.0180	turbid
Amazon River	AM12R	-54.7185	-2.4041	turbid
Amazon River	AT1R	-54.4518	-2.4916	turbid
Tapajós	T1R	-55.1013	-2.9714	clear
Tapajós	T2R	-54.9001	-2.4315	clear
Tapajós	T3R	-54.8332	-2.2742	clear
Tapajós	T4R	-54.7219	-2.4056	clear

18

19 **Table S2. Water parameters in the Amazon River.** The environmental parameters include pH, dissolved oxygen (DO) concentration, conductivity (Cond.),  
 20 temperature (Temp.), DOC, POC, TOC, as well as inorganic nutrients concentrations such as ammonium (NH<sub>4</sub><sup>+</sup>), nitrate (NO<sub>3</sub><sup>-</sup>), nitrite (NO<sub>2</sub><sup>-</sup>) and phosphate  
 21 (PO<sub>4</sub><sup>3-</sup>).

22

Sample	Temp. (°C)	pH	Depth (m)	Secci. (m)	Discharge (m <sup>3</sup> /s)	Velocity (m/s)	Width m	DOC (mg/L)	TOC (mg/L)	POC (mg/L)	Cond. (µs/cm)	DO (mg/L)	DIC (µM)	NH <sub>4</sub> <sup>+</sup> (µM)	NO <sub>3</sub> <sup>-</sup> (µM)	NO <sub>2</sub> <sup>-</sup> (µM)	PO <sub>4</sub> <sup>3-</sup> (µM)
S1R	31.0	7.10	30.93	0.2	128235	1.643	2691.86	5.24	17.90	11.38	96.9	6.0	2111	0.42	0.26	0.10	0.10
S2R	29.8	6.97	30.93	0.2	128235	1.643	2691.86	5.44	16.03	10.59	109.0	5.6	2000	0.79	0.10	0.21	0.52
N1R	30.0	4.73	35.08	1.5	28580	0.302	2698.42	4.23	22.50	12.54	12.6	5.9	500	0.68	0.41	0.09	0.10
N2R	29.9	4.88	35.08	1.5	28580	0.302	2698.42	6.34	19.13	12.79	16.4	5.3	500	0.86	0.91	0.08	0.11
N3R	30.5	4.73	NA	NA	NA	NA	NA	9.96	22.50	NA	16.9	4.7	NA	0.68	0.41	0.09	0.10
N4R	30.3	4.95	35.08	1.5	28580	0.302	2698.42	3.26	16.80	13.54	11.3	6.4	451	0.97	2.76	0.21	0.24
N5R	29.9	4.82	NA	NA	NA	NA	NA	7.85	21.83	NA	57.9	6.0	NA	0.64	0.22	0.06	0.15
A1R	28.8	6.35	36.47	0.2	146136	1.472	2716.75	5.36	6.80	1.44	84.9	4.5	1300	0.53	0.94	0.08	0.24
A2R	28.5	6.45	45.86	0.2	251599	1.858	2953.31	4.85	5.82	0.90	79.3	3.9	1000	0.63	0.69	0.05	0.69
AM4R	28.8	6.55	NA	NA	NA	NA	NA	4.68	6.17	1.49	62.4	3.9	NA	0.81	0.76	0.06	0.15
AM5R	29.1	6.65	NA	NA	NA	NA	NA	2.84	3.07	0.23	35.1	6.1	NA	0.69	0.35	0.03	0.11
AM6R	28.7	6.39	NA	NA	NA	NA	NA	5.21	6.01	0.80	57.4	3.7	NA	0.77	0.62	0.10	0.25
AM7R	29.0	6.41	NA	NA	NA	NA	NA	5.32	6.23	0.91	58.4	4.0	NA	NA	0.51	0.10	0.32
AM8R	29.0	6.41	NA	NA	NA	NA	NA	3.77	4.22	0.45	58.1	3.7	NA	0.69	0.54	0.06	0.17
AM10R	27.5	6.49	NA	NA	NA	NA	NA	4.16	4.39	0.23	56.4	5.0	NA	0.88	0.59	0.15	0.22
AM11R	29.1	6.54	NA	NA	NA	NA	NA	4.68	4.87	0.19	50.8	4.9	NA	0.73	0.71	0.12	0.11
AT1R	27.8	6.79	NA	NA	NA	NA	NA	3.63	4.02	0.39	35.0	7.2	NA	0.47	1.55	0.04	0.15
T1R	26.9	5.68	NA	NA	NA	NA	NA	3.11	3.34	NA	36.9	5.8	NA	0.42	0.57	0.10	0.12
T2R	29.5	6.61	9.55	3.5	20736	0.709	2914.85	3.48	3.52	0.17	34.2	7.8	1000	0.41	0.47	0.02	0.11
T3R	29.9	6.82	9.55	3.5	20736	0.709	2914.85	5.51	6.41	0.85	61.0	6.0	681	2.32	0.33	0.13	0.10

23 **Table S3. ESI(-) FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions**  
 24 **present in AZ-DOM.** The bottom four rows show the range of values for the main stem river DOM samples in each area (i.e., Negro, Solimões, Amazon,  
 25 Tapajós). FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular series, computed experimental  
 26  $m/z$  (charge-to-mass ratio of singly charged ions), DBE/C (double bond equivalency per carbon of non-charged molecules), H/C (ratio of hydrogen to carbon  
 27 atoms of non-charged molecules), O/C (ratio of oxygen to carbon atoms of non-charged molecules), N/C (ratio of nitrogen to carbon atoms of non-charged  
 28 molecules), and S/C (ratio of sulfur to carbon atoms of non-charged molecules).

Sample	total counts of mass peaks	CHO%	CHNO%	CHOS%	CHNOS%	$m/z$	DBE/C	H/C	O/C	N/C	S/C $\times 10^5$	$\Delta I_{mod} > 0.5$ %
S1R	4759	63.6	31.0	5.34	0.04	479.7	0.516	1.071	0.540	0.006	96.8	13.7
S2R	4906	62.1	31.5	6.14	0.29	482.5	0.521	1.061	0.546	0.006	93.7	14.3
S3R	5043	59.5	32.4	7.20	0.91	460.2	0.526	1.057	0.550	0.008	112.3	15.7
SM	4408	65.7	29.0	5.08	0.16	479.9	0.521	1.061	0.542	0.006	78.7	14.9
N1R	4167	73.9	24.4	1.73	0.00	498.0	0.530	1.035	0.549	0.003	30.8	16.0
N2R	3889	74.0	24.3	1.62	0.03	477.1	0.524	1.051	0.551	0.004	26.6	14.3
N3R	4519	73.2	25.2	1.62	0.02	483.8	0.522	1.053	0.537	0.003	28.1	15.7
N4R	4349	73.9	24.5	1.63	0.02	490.4	0.526	1.044	0.541	0.003	20.7	16.2
N5R	4428	73.7	24.6	1.69	0.00	485.0	0.517	1.063	0.530	0.003	40.4	15.0
N6R	4233	74.4	23.9	1.68	0.02	496.5	0.529	1.038	0.539	0.003	22.8	16.7
NaM	3781	76.9	21.8	1.35	0.00	498.3	0.528	1.038	0.538	0.003	17.2	16.6
MZ	4424	73.4	24.8	1.79	0.00	492.3	0.521	1.054	0.534	0.003	22.8	15.5
A1R	4675	68.7	27.7	3.55	0.00	488.9	0.522	1.054	0.541	0.005	43.3	15.4
A2R	4468	69.0	27.3	3.65	0.00	485.1	0.516	1.068	0.538	0.005	64.2	14.2
AM1R	4120	63.8	30.4	5.51	0.29	467.9	0.516	1.075	0.540	0.007	100.5	13.9
AM2R	4522	67.2	28.6	4.07	0.07	482.9	0.514	1.072	0.535	0.005	58.5	13.8
AM3Ra	4350	67.7	27.3	4.99	0.02	483.7	0.521	1.058	0.542	0.005	67.0	14.7
AM4R	4415	65.0	29.9	5.07	0.05	474.7	0.516	1.071	0.544	0.006	64.8	13.2
AM5R	4444	66.0	29.8	3.98	0.18	474.3	0.518	1.066	0.543	0.006	48.6	14.0
AM6R	4519	65.1	30.3	4.62	0.04	479.6	0.524	1.051	0.544	0.006	64.8	14.7
AM7R	4426	65.3	29.7	4.88	0.16	480.1	0.519	1.064	0.544	0.006	70.7	13.6
AM8R	4395	64.9	30.2	4.82	0.14	481.7	0.523	1.057	0.547	0.006	78.7	13.2
AM9Ra	4890	64.9	29.9	5.11	0.14	480.3	0.519	1.064	0.547	0.006	67.3	14.2
AM10R	4473	64.1	30.2	5.59	0.16	478.5	0.518	1.067	0.549	0.006	90.8	14.4
AM11R	4515	64.2	30.7	4.94	0.20	476.1	0.515	1.071	0.535	0.006	80.8	13.5
AM12R	4787	65.0	30.2	4.72	0.00	480.2	0.521	1.061	0.541	0.006	54.4	15.8
AT1R	3870	68.2	29.5	2.27	0.03	479.8	0.520	1.062	0.546	0.006	39.3	13.6
T1R	5276	62.4	33.8	3.77	0.04	472.1	0.518	1.067	0.547	0.007	82.7	12.0

<b>T2R</b>	5858	60.9	35.3	3.81	0.09	466.3	0.525	1.052	0.543	0.008	98.9	12.3
<b>T3R</b>	5813	58.7	32.1	6.93	2.24	446.6	0.519	1.064	0.546	0.007	120.7	14.6
<b>T4R</b>	5898	57.9	34.4	6.51	1.22	462.4	0.497	1.109	0.513	0.008	94.2	14.5
<b>S</b>	4759-5043	59.5-63.6	31.0-32.4	5.33-7.20	0.00-0.04	460.2-482.5	0.52-0.53	1.06-1.07	0.54-0.55	0.006-0.008	(93.7-112.3)	13.7-15.7
<b>N</b>	3781-4519	73.2-76.9	21.8-25.2	1.35-1.73	0.00-0.03	477.1-498.3	0.52-0.53	1.04-1.06	0.53-0.55	0.003-0.004	(17.2-40.4)	14.3-16.7
<b>A</b>	4120-5010	63.0-70.3	26.7-30.7	3.05-6.31	0.00-0.34	467.9-488.9	0.51-0.52	1.05-1.07	0.53-0.55	0.005-0.007	(43.4-100.5)	13.2-15.4
<b>T</b>	5276-5898	57.9-62.4	32.1-35.3	3.87-6.93	0.04-2.24	446.6-472.1	0.49-0.52	1.08-1.11	0.51-0.53	0.007-0.008	(82.7-120.7)	12.0-14.6

29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43

44 **Table S4. ESI[+] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions**  
 45 **present in AZ-DOM.** The bottom four rows show the range of values for the main stem river DOM samples in each area (i.e., Negro, Solimões, Amazon,  
 46 Tapajós). FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular series, computed average  
 47 experimental  $m/z$  (charge-to-mass ratio of singly charged ions), DBE/C (double bond equivalency per carbon of non-charged molecules), H/C (ratio of hydrogen  
 48 to carbon atoms of non-charged molecules), O/C (ratio of oxygen to carbon atoms of non-charged molecules), N/C (ratio of nitrogen to carbon atoms of non-  
 49 charged molecules), and S/C (ratio of sulfur to carbon atoms of non-charged molecules).

Sample	total counts of mass peaks	CHO%	CHNO%	CHOS%	CHNOS%	$m/z$	DBE/C	H/C	O/C	N/C	S/C $\times 10^5$	AI <sub>msa</sub> >0.5 %
S1R	3817	31.3	68.5	0.18	0.00	425.2	0.385	1.366	0.363	0.037	5.0	3.3
S2R	4161	29.9	69.9	0.17	0.02	425.3	0.382	1.374	0.376	0.038	3.3	2.5
S3R	4967	29.5	70.2	0.08	0.18	439.2	0.398	1.341	0.409	0.039	2.3	2.7
SM	5172	28.5	71.0	0.19	0.27	413.2	0.382	1.379	0.373	0.040	6.2	3.1
N1R	5973	30.9	68.9	0.12	0.00	454.8	0.399	1.333	0.423	0.035	3.6	2.4
N2R	5342	30.1	69.7	0.13	0.00	446.0	0.405	1.325	0.428	0.036	5.3	2.8
N3R	4824	29.0	70.8	0.12	0.04	438.4	0.385	1.365	0.393	0.038	3.0	2.0
N4R	5466	28.4	71.4	0.15	0.02	445.3	0.389	1.355	0.397	0.037	3.1	2.3
N5R	5575	28.8	71.0	0.13	0.04	448.8	0.388	1.357	0.401	0.037	3.1	2.0
N6R	5727	30.0	69.9	0.05	0.03	442.5	0.393	1.347	0.395	0.036	2.5	2.7
NaM	5428	28.9	70.9	0.15	0.02	439.2	0.389	1.357	0.393	0.037	3.4	2.6
MZ	5808	28.6	71.2	0.17	0.03	437.7	0.379	1.378	0.391	0.038	5.0	2.1
A1R	5159	30.0	69.7	0.12	0.17	422.3	0.388	1.365	0.390	0.039	3.2	2.4
A2R	4949	29.6	70.1	0.20	0.04	417.1	0.387	1.370	0.383	0.041	4.4	2.5
AM1R	5081	27.7	71.6	0.20	0.49	407.2	0.385	1.376	0.376	0.043	5.8	3.0
AM2R	5377	29.2	70.1	0.19	0.48	419.1	0.387	1.368	0.386	0.040	5.5	2.6
AM3Ra	3763	34.6	65.1	0.05	0.19	396.4	0.394	1.359	0.391	0.039	4.6	3.2
AM4R	5471	32.3	67.1	0.27	0.35	415.9	0.411	1.317	0.405	0.035	7.5	4.6
AM5R	5748	29.4	69.4	0.28	0.94	407.3	0.403	1.335	0.412	0.037	9.5	5.0
AM6R	6112	28.4	69.8	0.61	1.13	422.7	0.393	1.361	0.386	0.037	13.3	3.9
AM7R	6160	29.7	68.7	0.45	1.19	422.8	0.412	1.320	0.404	0.037	13.2	4.0
AM8R	5654	28.2	70.4	0.46	0.94	418.7	0.405	1.330	0.414	0.039	10.6	2.9
AM9Ra	3477	31.2	68.7	0.09	0.03	380.9	0.408	1.325	0.420	0.041	3.2	3.7
AM10R	6060	30.4	67.6	0.53	1.45	404.8	0.392	1.358	0.404	0.037	16.9	4.6
AM11R	5845	30.1	68.3	0.46	1.13	415.4	0.385	1.379	0.359	0.037	14.1	4.1
AM12R	3943	32.0	68.0	0.03	0.05	395.0	0.390	1.368	0.380	0.040	2.3	3.7
AT1R	5809	30.2	69.4	0.10	0.29	410.7	0.404	1.335	0.403	0.039	7.4	4.5

<b>T1R</b>	4709	30.5	69.3	0.13	0.08	397.6	0.408	1.327	0.414	0.040	3.9	3.7
<b>T2R</b>	4365	29.6	70.3	0.07	0.02	390.0	0.398	1.352	0.393	0.041	2.8	4.5
<b>T3R</b>	4189	29.8	69.8	0.19	0.24	389.4	0.407	1.330	0.404	0.041	5.6	3.5
<b>T4R</b>	3564	32.3	67.5	0.14	0.06	386.7	0.393	1.363	0.386	0.039	3.5	3.9
<b>S</b>	3817-4967	29.5-31.3	68.5-70.2	0.08-0.18	0.00-0.18	425.2-439.2	0.38-0.40	1.34-1.37	0.36-0.41	0.037-0.039	(2.3-5.0)	2.5-3.3
<b>N</b>	4824-5973	28.4-30.9	68.9-71.4	0.05-0.15	0.00-0.04	438.4-454.8	0.39-0.40	1.32-1.36	0.39-0.43	0.035-0.038	(2.5-5.3)	2.0-2.8
<b>A</b>	3477-6160	27.7-34.6	65.1-71.6	0.03-0.61	0.00-1.45	380.89-422.8	0.39-0.41	1.32-1.38	0.36-0.42	0.035-0.043	(2.3-16.9)	2.4-5.0
<b>T</b>	3564-4709	29.6-32.3	67.5-70.3	0.07-0.19	0.02-0.24	386.67-397.6	0.39-0.40	1.35-1.37	0.38-0.39	0.039-0.041	(2.8-5.6)	3.5-4.5

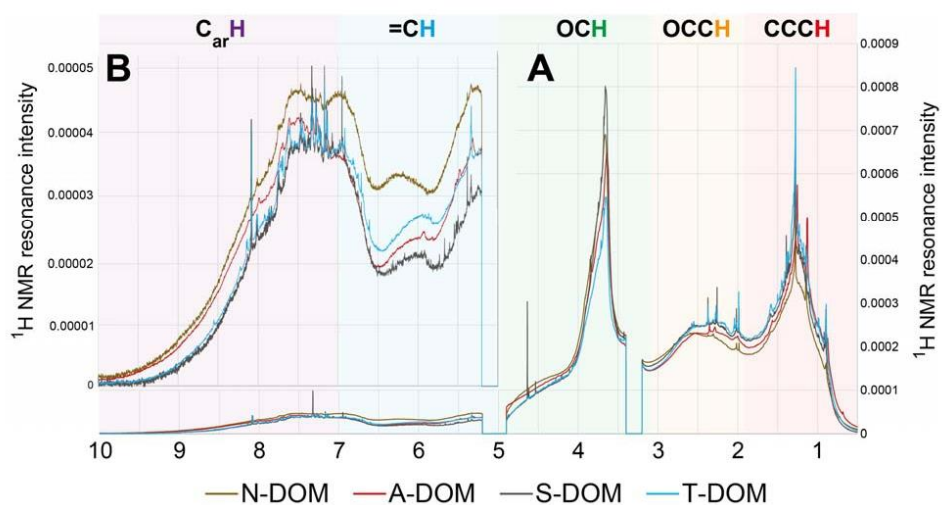


51 **Table S5. <sup>1</sup>H NMR section integrals (percent of non-exchangeable protons) and key substructures**  
 52 **of AZ-DOM samples (CD<sub>3</sub>OD, exclusion of residual water, and methanol).** The bottom four rows  
 53 show the range of values for the main stem river DOM samples in each area (i.e., Negro, Solimões,  
 54 Amazon, Tapajós). The fundamental substructures from lower to higher field (from left to right):  
 55 aromatic (C<sub>ar</sub>H), δ<sub>H</sub> ~ 7.0-10.0 ppm; olefins and dioxygenated aliphatic units (=CH, O<sub>2</sub>CH), δ<sub>H</sub> ~ 5.3-  
 56 7.0 ppm; oxygenated aliphatic units (OCH), including “carbohydrate-like” and methoxy (H<sub>3</sub>CO), δ<sub>H</sub> ~  
 57 3.2-4.9 ppm; remotely functionalized aliphatic units (OCC), including “acetyl” derivatives and  
 58 CRAM (HOOC-CH<sub>α</sub>-R), δ<sub>H</sub> ~ 1.9-3.2 ppm; functionalized aliphatics δ<sub>H</sub> ~ 1.35-1.9 ppm; branched alkyl  
 59 groups (C<sub>n</sub>H<sub>2n+1</sub>), δ<sub>H</sub> ~ 1.25-1.35 ppm; pure aliphatic units (CCC) δ<sub>H</sub> ~ 0.5-1.25 ppm.

δ ( <sup>1</sup> H) [ppm]	10.0-7.0	7.0_5.3	4.9-3.2	3.2-1.9	1.9-1.35	1.35-1.25	1.25-0.5	1.9-0.5
key substructures	C <sub>ar</sub> H (%)	C=CH, O <sub>2</sub> CH (%)	OCH (%)	OCC (%)	OCCC (%)	(CH <sub>2</sub> ) <sub>n</sub> (%)	CCC (%)	total aliphatic section (%)
S1R	5.1	4.3	31.1	26.1	14.7	4.7	14.0	33.3
S2R	4.6	5.6	31.5	27.1	14.8	4.1	12.2	31.1
S3R	5.6	5.2	33.1	25.7	14.1	3.9	12.4	30.4
SM	4.9	4.0	29.5	27.2	15.1	5.0	14.3	34.4
N1R	5.4	6.6	33.8	27.1	13.1	3.6	10.4	27.0
N2R	5.0	6.9	33.1	27.6	13.8	3.8	9.8	27.3
N5R	5.7	5.8	30.6	27.5	14.2	4.2	12.0	30.4
N6R	5.7	6.5	29.9	28.0	14.4	4.1	11.4	29.9
NaM	5.5	5.7	30.2	27.6	14.6	4.2	12.2	31.0
MZ	5.1	5.8	30.9	27.5	14.7	4.4	11.5	30.6
A1R	5.3	6.0	31.5	27.4	14.1	4.0	11.5	29.7
A2R	5.2	5.5	30.6	28.0	14.5	4.2	12.0	30.7
AM1R	4.5	4.7	28.2	29.1	15.9	4.7	12.9	33.5
AM2R	5.3	5.5	30.5	27.8	14.7	4.3	12.1	31.0
AM3Ra	5.3	4.3	31.8	26.8	14.2	4.2	13.5	31.9
AM4R	4.7	5.1	30.8	28.2	14.9	4.3	12.0	31.1
AM5R	4.9	4.4	31.7	26.8	14.5	4.8	13.0	32.3
AM6R	4.8	5.1	32.6	27.2	14.4	4.1	11.8	30.3
AM7R	4.8	5.0	32.5	27.3	14.4	4.1	11.8	30.4
AM8R	4.6	4.8	31.7	27.4	14.8	4.7	12.0	31.5
AM9Ra	4.9	4.6	28.8	28.3	15.2	4.6	13.6	33.5
AM10R	4.8	3.4	31.0	26.6	14.8	4.7	14.7	34.2
AM11R	4.9	5.1	32.8	26.5	14.6	4.3	11.9	30.7
AM12R	4.8	4.6	30.0	27.4	15.1	4.6	13.4	33.1
AT1R	4.9	3.8	32.5	25.8	14.5	4.6	13.8	33.0
T1R	4.9	5.2	32.8	26.3	14.2	4.7	11.9	30.8
T2R	4.9	4.0	29.7	26.5	14.8	5.1	14.8	34.8
T3R	4.8	4.7	28.9	28.1	15.2	5.0	13.3	33.5
T4R	4.8	6.7	27.3	26.0	14.2	5.4	15.5	35.1
S	4.6-5.6	4.3-5.6	31.1-33.1	25.7-27.1	14.1-14.8	3.9-4.7	12.2-14.0	30.4-33.3
N	5.0-5.7	5.7-6.9	29.9-33.8	27.1-28.0	13.1-14.6	3.6-4.2	9.8-12.2	27.0-30.4
A	4.5-5.3	3.4-6.0	28.2-32.8	26.5-29.1	14.1-15.9	4.0-4.8	11.5-14.7	29.7-34.2
T	4.8-4.9	4.0-6.7	27.3-32.8	26.0-28.1	14.2-15.2	4.7-5.4	11.9-15.5	30.8-35.1

60

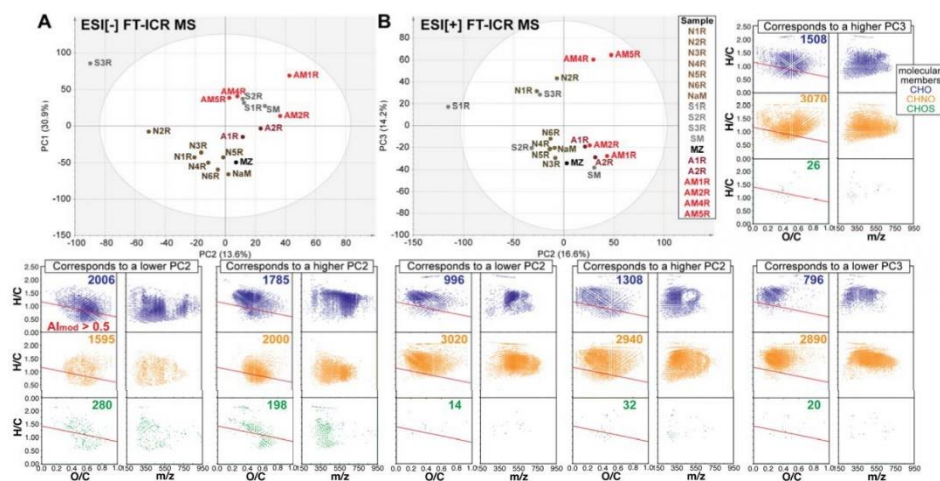




61

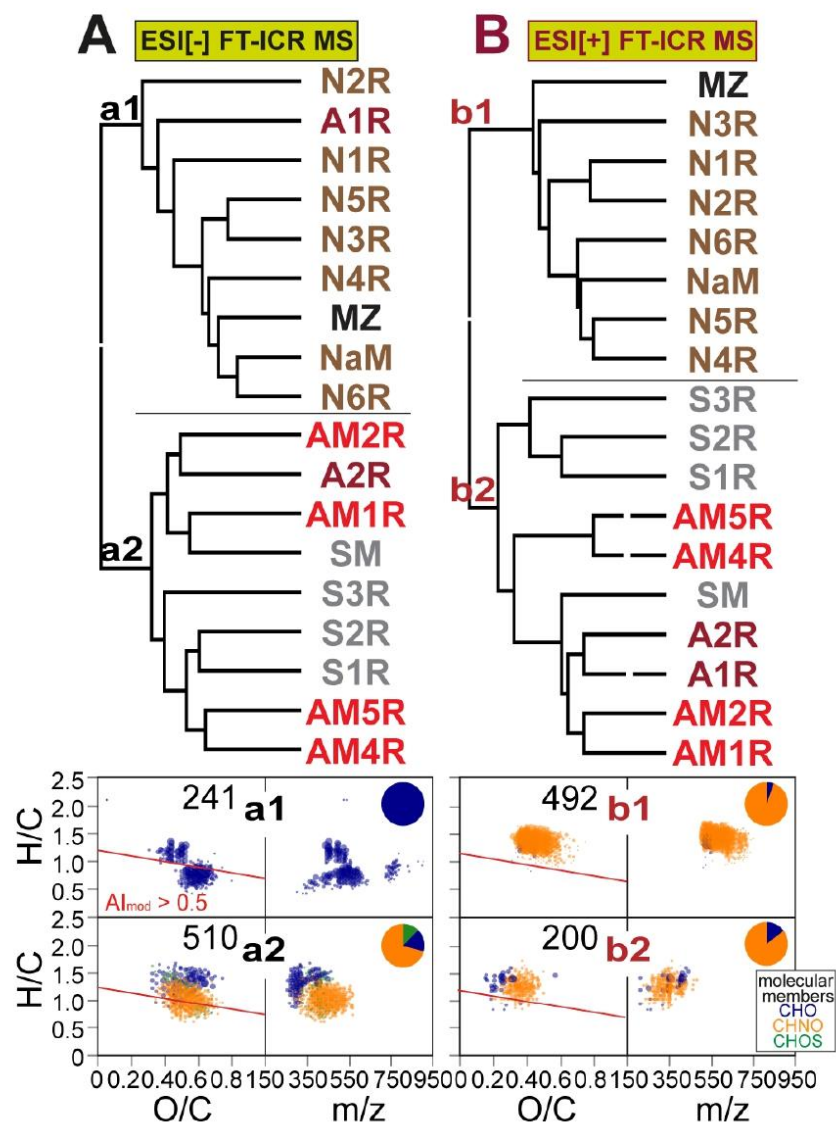
62 **Fig. S3.**  $^1\text{H}$  NMR spectra of four consolidated AZ-DOM, including N-DOM, A-DOM, S-DOM and T-  
 63 **DOM.** (A) Area normalized  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ;  $\delta_{\text{H}} = 0.5\text{-}10.0$  ppm with exclusion of residual  
 64 water and methanol NMR resonances) of four DOM, with (B) vertical expansion of  $C_{\text{sp}_2\text{H}}$  units. For adequate  
 65 representation and good S/N ratio in  $^1\text{H}$  NMR spectra, we have used consolidated samples as follows: S-DOM  
 66 = S1R + S2R + S3R; N-DOM = N1R + N2R + N3R + N4R + N5R + N6R; A-DOM = A1R + A2R + AM1R +  
 67 AM2R + AM4R + AM5R + AM6R + AM7R + AM8R + AM10R + AM11R + AM12R; T-DOM = T1R + T2R  
 68 + T3R; cf. Table S1.

69



70

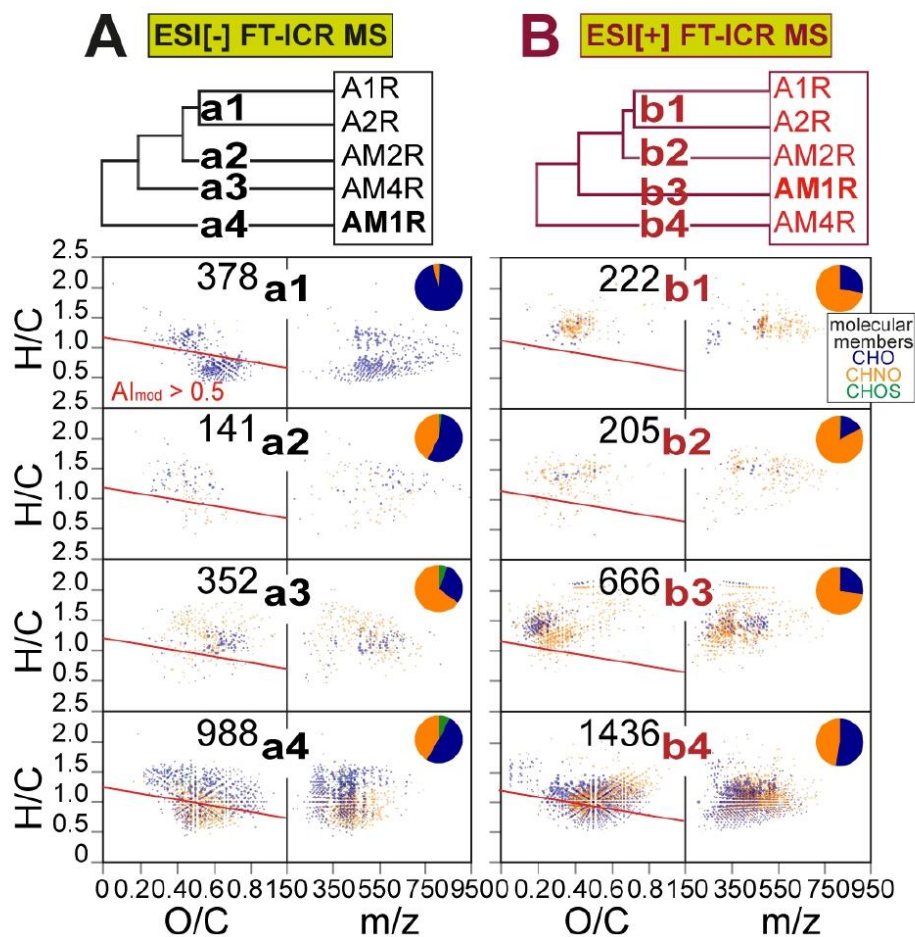
71 **Fig. S4.** PCA of AZ-DOM molecular composition proximate to SN mixing zone based on all assigned molecular formulae in ESI[±] FT-ICR MS spectra. The  
 72 upper left panels show the distribution of AZ-DOM in PCA scatter plot based on (A) ESI[-] and (B) ESI[+]FT-ICR mass spectra, respectively. PC1 is discussed in Fig.  
 73 3. The lower left panel depict van Krevelen and mass-edited H/C diagrams corresponding to lower/higher second principal component (PC2), based on PCA loading  
 74 vectors, with the color representing elemental compositions with the assigned CHO (blue), CHNO (orange), and CHOS (green) molecular series. The most right panel  
 75 show the van Krevelen and mass-edited H/C diagrams corresponding to lower/higher third principal component (PC3) in PCA based on ESI[+] FT-ICR MS. Molecular  
 76 composition positioned below the red line in the van Krevelen diagrams show the compounds with modified aromaticity index (Al<sub>mod</sub>) higher than 0.5 (Koch and  
 77 Dittmar, 2006). The numbers show counts of compounds. PC2 in ESI[±] FT-ICR MS-based PCA explained 13.6%/16.6% of the total variances and showed a similar  
 78 distribution to PC1 (Fig. 3). PC2 in ESI[-] FT-ICR MS-based PCA separated N-DOM from other AZ-DOM (Fig. S4-A), while PC2 in ESI[+] FT-ICR MS-based PCA  
 79 separated A-DOM from S-DOM and N-DOM (Fig. S4-B).



80

81 **Fig. S5. HCA of AZ-DOM composition proximate to the mixing zone of Solimões and Negro (SN) based**  
 82 **on all assigned molecular formulae in ESI[±] FT-ICR MS spectra. Van Krevelen and mass-edited H/C**  
 83 **diagrams show CHO (blue), CHNO (orange), and CHOS (green) molecules that were relatively more abundant**  
 84 **in clusters a1, a2, b1, and b2, respectively. The sizes of bubbles represent the normalized intensity of AZ-DOM**  
 85 **in each cluster. The pie plots depict percentages of the counts of different molecular series. Molecular**  
 86 **composition positioned below the red line in the van Krevelen diagrams show the compounds with modified**  
 87 **aromaticity index ( $AI_{mod}$ ) higher than 0.5 (Koch and Dittmar, 2006). The numbers show counts of compounds.**

Page 18 of 34



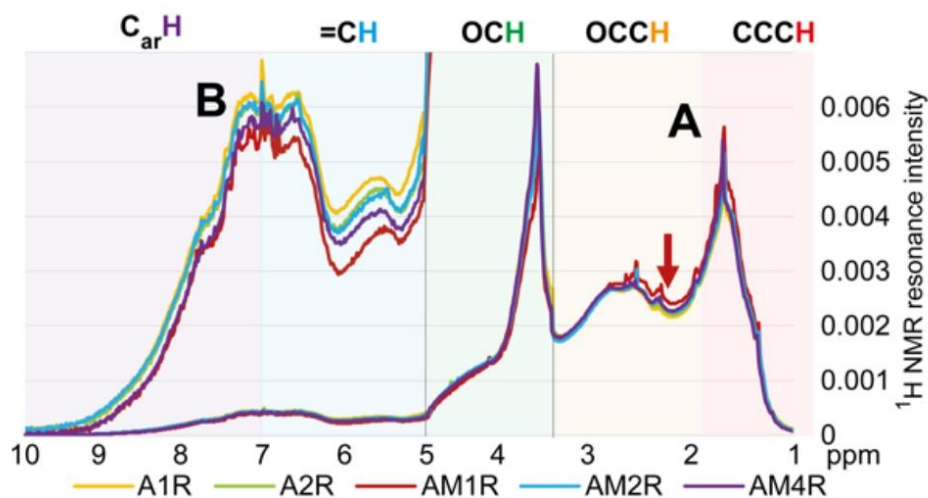
88

89 **Fig. S6. HCA of A-DOM composition proximate to the Madeira River inflow to the Amazon River base**  
 90 **on all assigned molecular formulae in (A) ESI[-] and (B) ESI[+] FT-ICR MS spectra.** Van Krevelen and  
 91 mass-edited H/C diagrams show CHO (blue) and CHNO (orange) and CHOS (green) molecules that were  
 92 relatively more abundant in clusters a1/b1 (A1R and A2R), clusters a2/b2 (AM2R), clusters a3/b4 (AM4R), and  
 93 clusters a4/b3 (AM1R) compared to the A-DOM in all the other clusters, respectively. The sizes of bubbles  
 94 represent the normalized intensity of AZ-DOM in each cluster. The numbers in van Krevelen diagrams show  
 95 counts of identified molecular compositions. The pie plots show the percentages of the counts of different  
 96 molecular series. Molecular composition positioned below the red line in the van Krevelen diagrams show the  
 97 molecules with modified aromaticity index ( $AI_{mod}$ ) higher than 0.5 (Koch and Dittmar, 2006). The numbers  
 98 show counts of assigned molecular compositions.

99

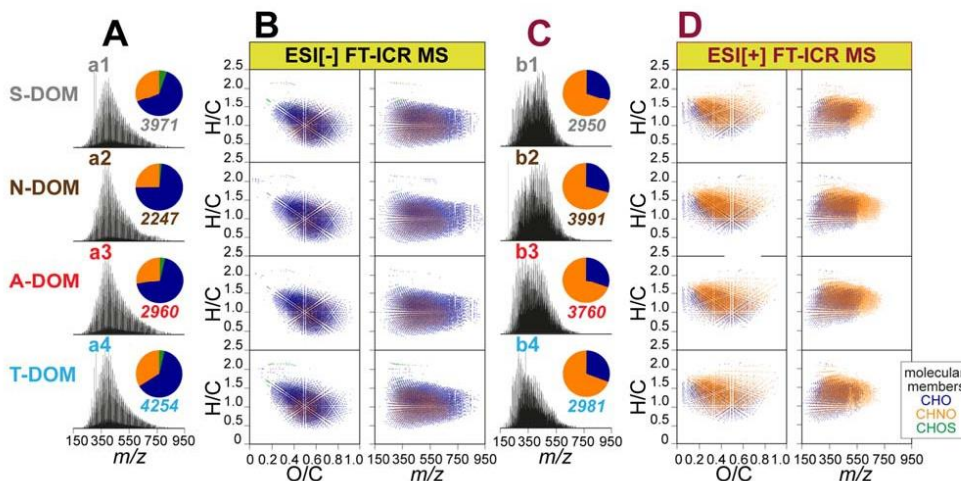
100

101



102

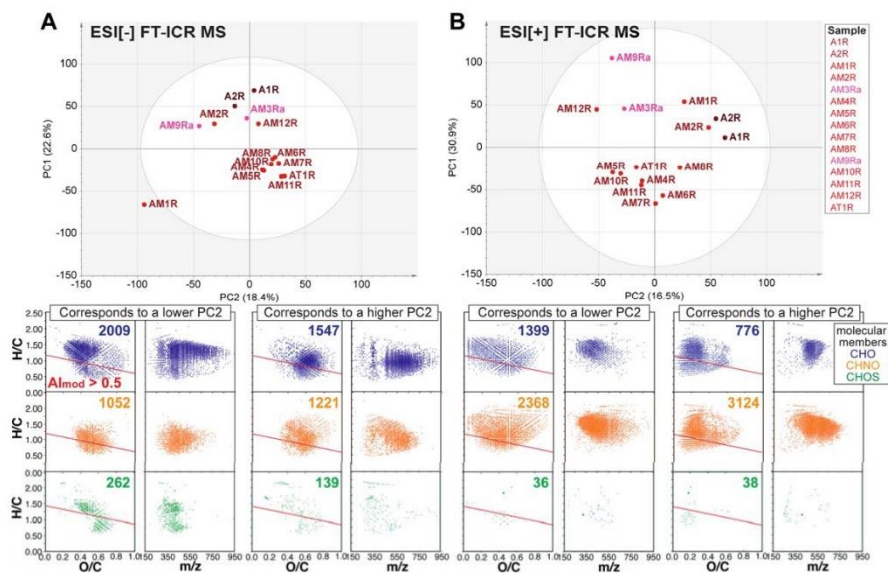
103 **Fig. S7.**  $^1\text{H}$  NMR spectra of A-DOM proximate to the Madeira River inflow to the Amazon River. (A)  
 104 Area normalized  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ;  $\delta_{\text{H}} = 0.5\text{-}10.0$  ppm with exclusion of residual water and  
 105 methanol NMR resonances) of four DOM, with (B) vertical expansion.  
 106



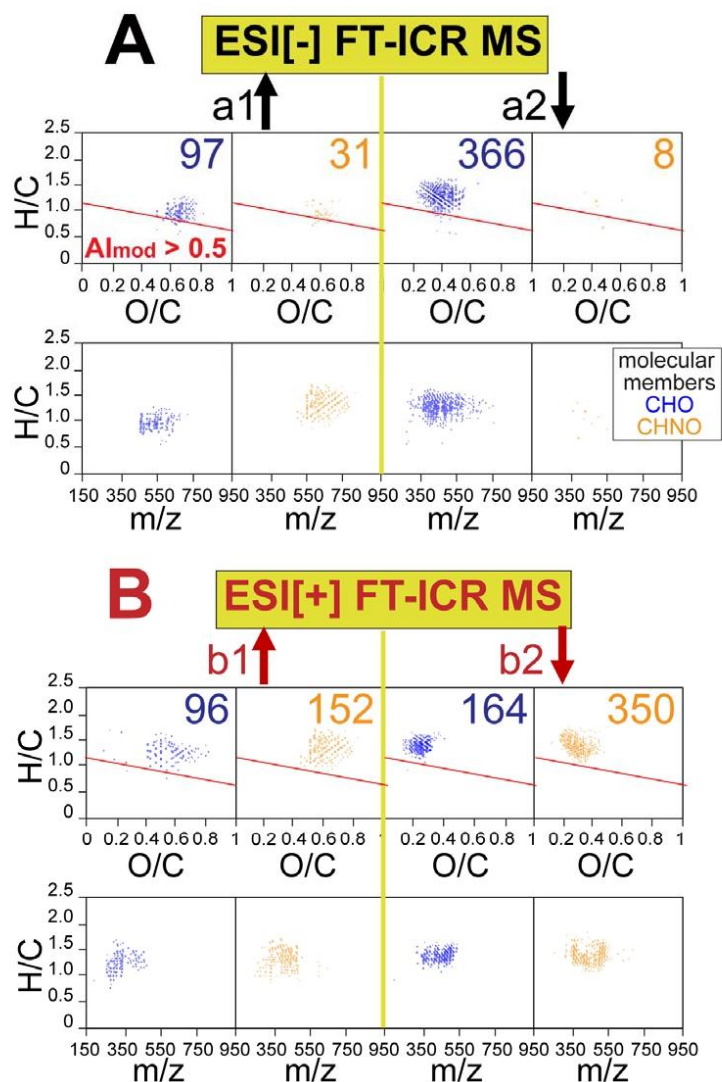
107

108 **Fig. S8.** ESI[ $\pm$ ] FT-ICR mass spectra and molecular composition of the common  $m/z$  ions in A-DOM. The  
 109 (A) ESI[-] and (C) ESI[+] FT-ICR MS spectra showed the dissimilar distribution of mass peaks across a wide  
 110 mass range ( $m/z$  150-950). (B, D) van Krevelen and mass-edited  $m/z$  diagrams. Respective assignments of  
 111 molecular compositions are provided for CHO (blue), CHNO (orange), and CHOS (green) molecules. The  
 112 inserted numbers show counts of assigned formulae.  
 113





114  
 115 **Fig. S9.** PCA of A-DOM composition based on all assigned molecular formulae in ESI[±] FT-ICR MS spectra. The upper left panel show the distribution of AZ-DOM in  
 116 PCA scatter plot based on (A) ESI[-] and (B) ESI[+] FT-ICR MS, respectively. The lower left panel show the van Krevelen and mass-edited H/C diagrams corresponding to  
 117 lower/higher PC2, based on PCA loading vectors, with the color representing elemental compositions with the assigned CHO (blue), CHNO (orange), and CHOS (green)  
 118 molecular series. Molecular composition positioned below the red line in the van Krevelen diagrams show the molecules with modified aromaticity index ( $Al_{mod} > 0.5$ )  
 119 0.5 (Koch and Dittmar, 2006). The numbers show counts of assigned molecular composition. AM9Ra and AM2R showed lower PC2 in ESI[-] FT-ICR MS-based PCA due to  
 120 the higher relative abundance of less unsaturated CHO compounds with a large  $m/z$  range and higher content of low-mass CHNO/CHOS molecules (Fig. S9A). Higher PC2 of  
 121 upstream A-DOM in ESI[+] FT-ICR MS-based PCA arose from higher content of high-mass CHO and CHNO molecules (Fig. S9B).

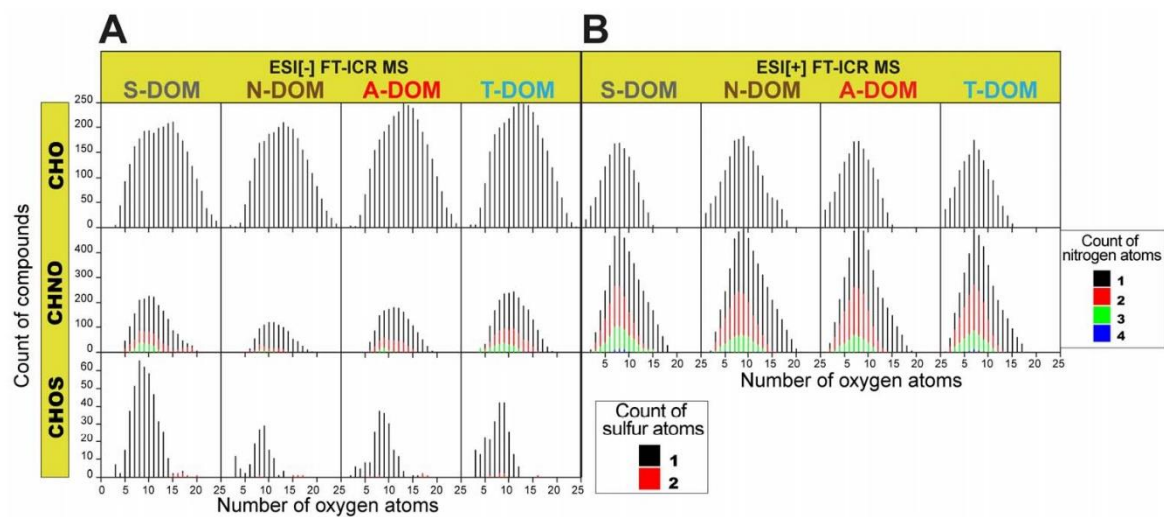


122

123 **Fig. S10.** Mass peaks identified in ESI[-] and ESI[+]FT-ICR mass spectra, of which the normalized  
 124 **intensity have increased/declined trend along the Amazon River (along fluvial flow from AMIR to**  
 125 **AM11R).** The increased/declined compounds were computed by significant correlation between normalized  
 126 intensities of mass peaks and the distances of the sampling point from AMIR ( $p < 0.05$ ,  $r^2 > 0.5$ ). Van Krevelen  
 127 and mass-edited H/C diagrams depict CHO (blue) and CHNO (orange) molecular series. Molecular composition  
 128 positioned below the red line in the van Krevelen diagrams show the compounds with modified aromaticity  
 129 index ( $AI_{mod}$ ) higher than 0.5 (Koch and Dittmar, 2006). The numbers show counts of assigned molecular  
 130 compositions.

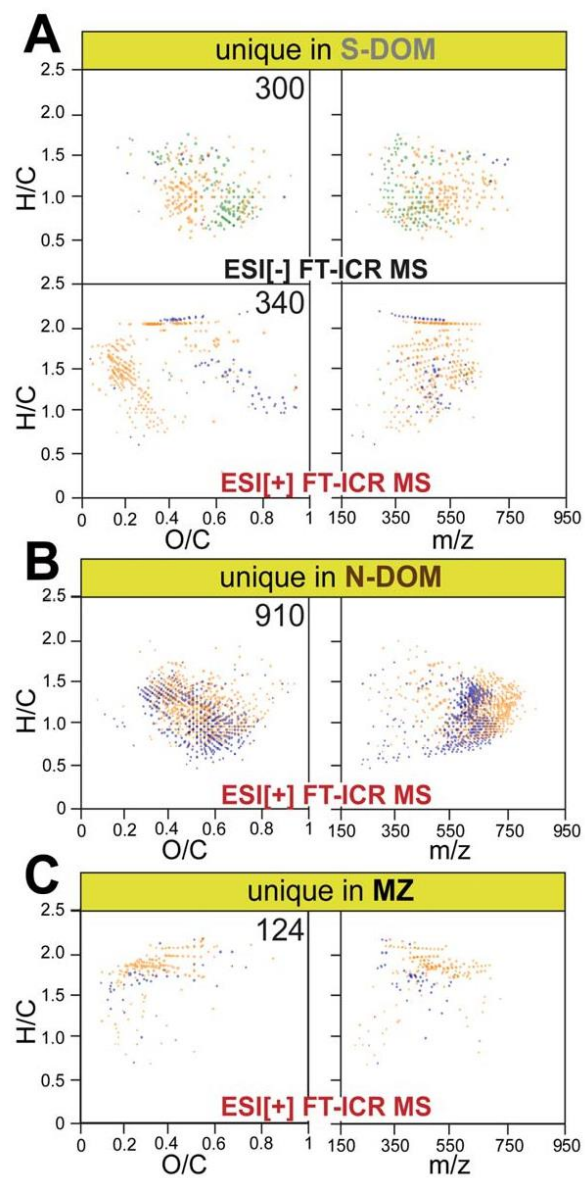


131



132

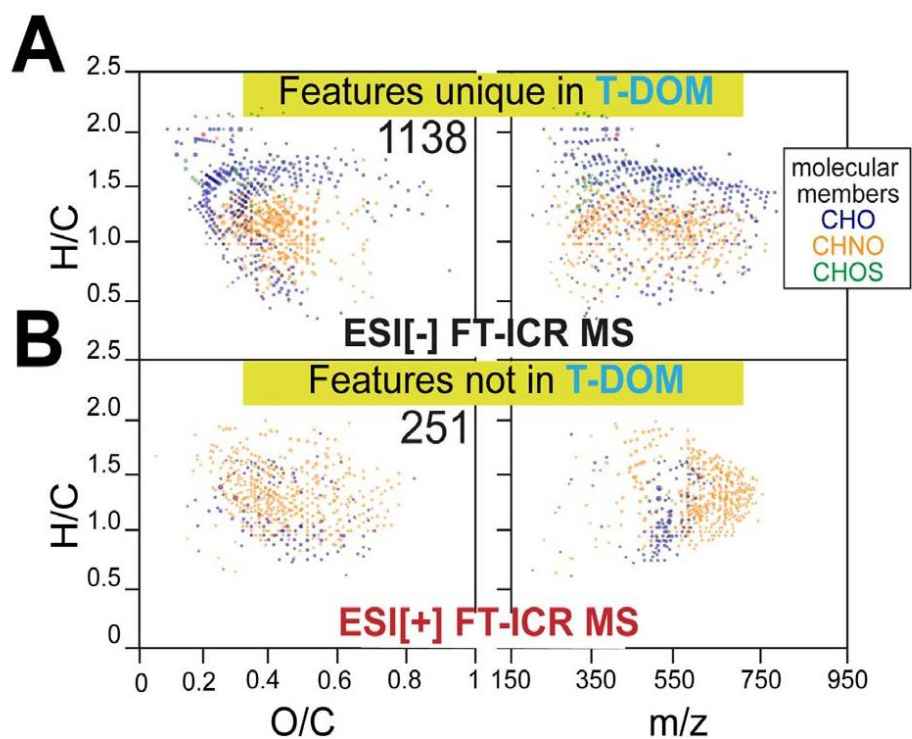
133 **Fig. S11. Frequency distribution of counts of oxygen atoms per molecules in S-DOM, N-DOM, A-DOM, and T-DOM from (A) ESI[-] and (B) ESI[+] FT-ICR MS. The**  
 134 **color shows the count of nitrogen and sulfur atoms. S-DOM, N-DOM, A-DOM, and T-DOM are exemplified with representative samples S1R, N1R, A1R, and T1R.**



135

136 **Fig. S12. Van Krevelen and mass edited H/C diagrams of the molecular formulae that were (A) unique**  
 137 **in S-DOM compared to N-DOM, (B) unique in N-DOM compared to S-DOM, and (C) unique in MZ**  
 138 **compared to N-DOM and S-DOM.** The colors in the diagrams depict CHO (blue), CHNO (orange) and CHOS  
 139 (green) molecular series. The numbers show counts of assigned molecular compositions. There were no  
 140 molecular formulae assigned in ESI[-] FT-ICR MS that were unique in N-DOM and MZ.

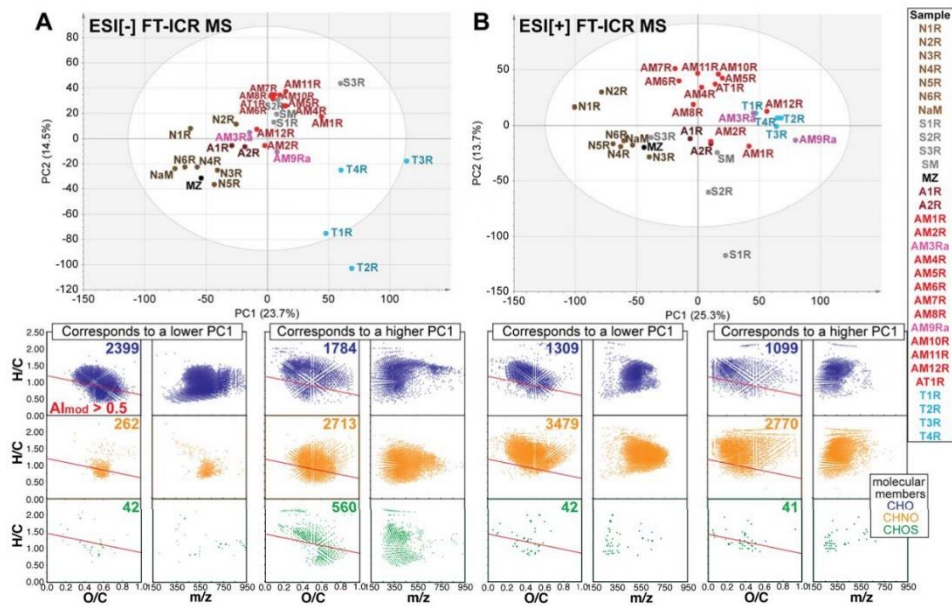
Page 24 of 34



141

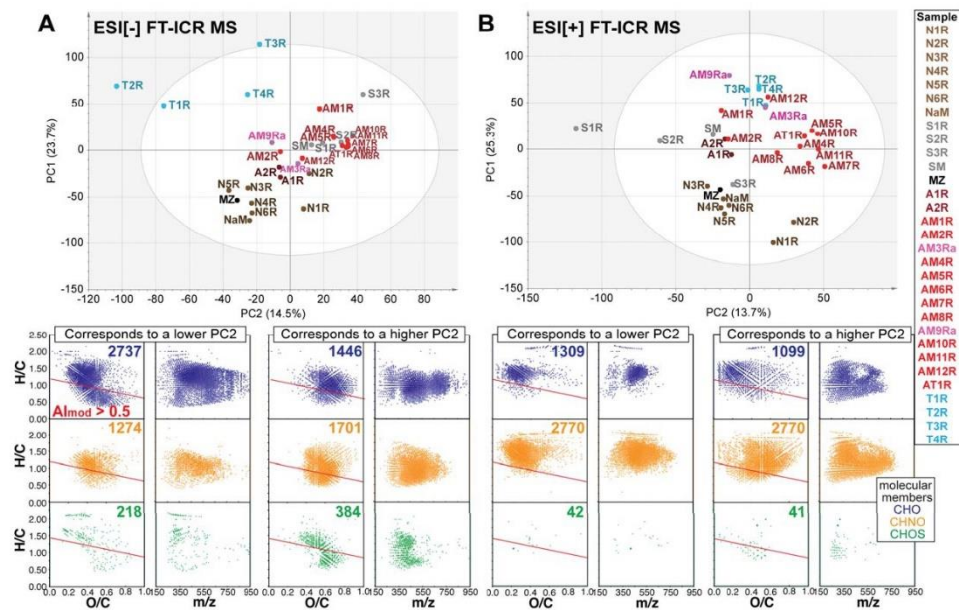
142 **Fig. S13. Van Krevelen and mass edited H/C diagrams of the molecular formulae that were (A) unique**  
 143 **in T-DOM and (B) were not present in T-DOM.** The colors in the diagrams depict CHO (blue), CHNO  
 144 (orange), and CHOS (green) molecular series. The numbers show counts of assigned molecular compositions.  
 145 There were no molecular formulae assigned in ESI[+] FT-ICR MS that were unique in T-DOM, and there were  
 146 no molecular formulae assigned in ESI[-] FT-ICR MS that were not in T-DOM.

147



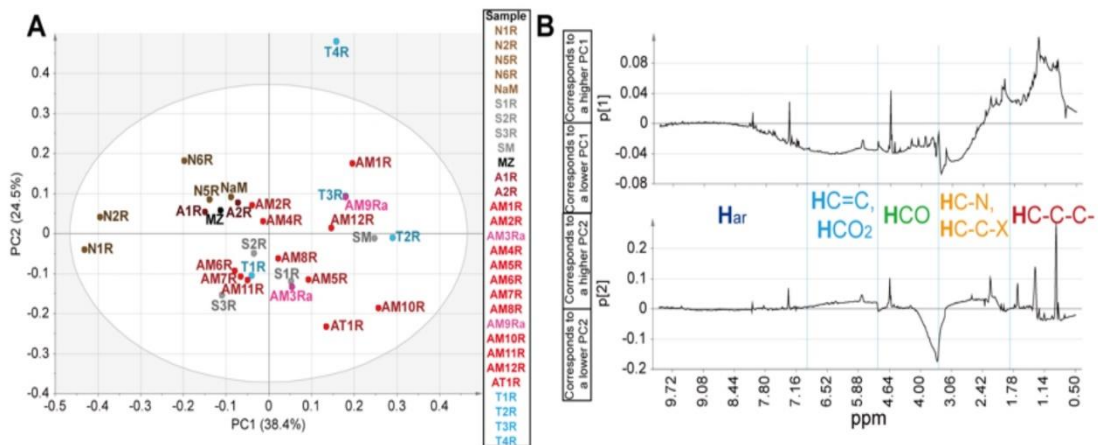
148

149 **Fig. S14.** PCA of AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular formulae in ESI[±] FT-ICR MS. The upper  
 150 panel show the distribution of AZ-DOM in PCA scatter plot based on (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra, respectively. The lower panel depict van Krevelen and  
 151 mass-edited H/C diagrams corresponding to lower/higher PC1, based on PCA loading vectors, with the color representing elemental compositions with the assigned CHO (blue),  
 152 CHNO (orange), and CHOS (green) molecular series. Molecular composition positioned below the red line in the van Krevelen diagrams show the compounds with modified  
 153 aromaticity index ( $AI_{mod}$ ) higher than 0.5 (Koch and Dittmar, 2006). The numbers show counts of assigned molecular compositions.



154

155 **Fig. S15. PCA of AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular formulae in ESI[±] FT-ICR MS spectra.** The  
 156 upper panel show the distribution of AZ-DOM in PCA scatter plot based on (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra, respectively. The lower panel depict van Krevelen  
 157 and mass-edited H/C diagrams corresponding to lower/higher PC2, based on PCA loading vectors, with the color representing elemental compositions with the assigned CHO  
 158 (blue), CHNO (orange), and CHOS (green) molecular series. Molecular composition positioned below the red line in the van Krevelen diagrams show the compounds with  
 159 modified aromaticity index ( $AI_{mod}$ ) higher than 0.5 (Koch and Dittmar, 2006). The numbers show counts of assigned molecular compositions.



160

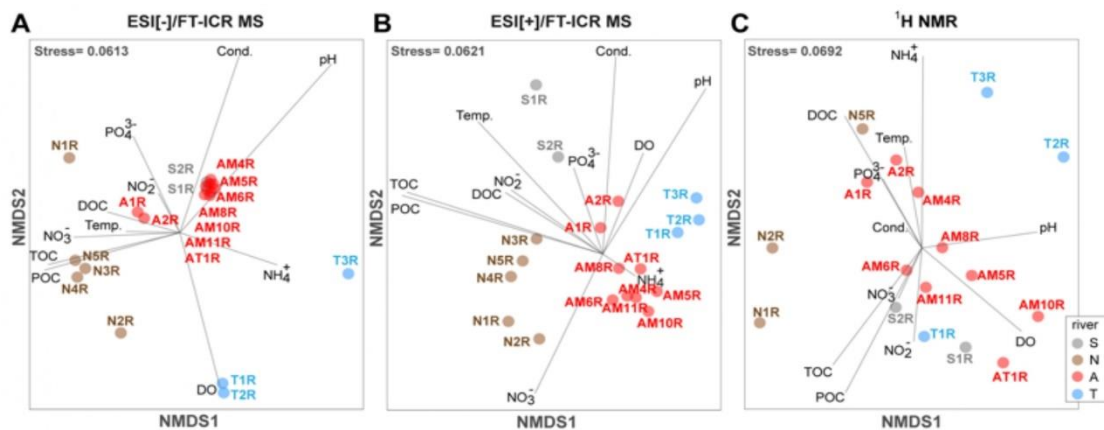
161 Fig. S16. PCA of AZ-DOM composition in the four Amazon main stem and tributaries based on <sup>1</sup>H NMR spectra (800 MHz <sup>1</sup>H NMR, CD<sub>3</sub>OD, area-normalized  
 162 from 0.5-10.0 ppm; 0.01 ppm bucket resolution; with exclusion of residual water and methanol NMR resonances). Panel B show loading vectors p1 of PC1 and  
 163 loading vectors p2 of PC2. The fundamental substructures of the NMR sections are indicated.

164 **Table S6. Heterotrophic bacterial production (HBP) and dark carbon fixation (DCF) in the river waters.**  
 165 SD, Standard Deviation; n, number of replicates. Sample information see [Table S1](#).

Sample	DCF (mean) $\mu\text{gC L}^{-1}\text{h}^{-1}$	DCF (SD) $\mu\text{gC L}^{-1}\text{h}^{-1}$	DCF (n)	HBP (mean) $\mu\text{gC L}^{-1}\text{h}^{-1}$	HBP (SD) $\mu\text{gC L}^{-1}\text{h}^{-1}$	HBP (n)
S1R	0.3085	0.0255	5	1.1933	0.1028	5
S2R	0.8649	0.7165	5	0.9041	0.3810	5
N1R	0.8141	0.2439	6	0.0800	0.0200	5
N4R	0.7715	0.0450	5	0.0946	0.0247	5
A1R	13.2677	1.5780	5	3.5606	0.7900	5
A2R	0.6946	0.8074	5	17.2304	2.3757	5
T2R	2.6292	1.6979	4	18.9508	0.7343	3
T3R	0.4395	0.2414	4	16.7442	0.8419	3

166





167

168 Fig. S17. NMDS analysis using AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular formulae in (A, B) ESI[±]  
 169 FT-ICR mass spectra, as well as (C)  $^1\text{H}$  NMR spectra in 0.01 ppm bucket resolution fitted with water parameters, rates of bacterial production and carbon  
 170 fixation shown as vectors. The environmental parameters used for NMDS included pH, dissolved oxygen (DO) concentration, conductivity (Cond.), temperature  
 171 (Temp.), DOC, POC, TOC, as well as inorganic nutrients concentrations such as ammonium ( $\text{NH}_4^+$ ), nitrate ( $\text{NO}_3^-$ ), nitrite ( $\text{NO}_2^-$ ) and phosphate ( $\text{PO}_4^{3-}$ ). Grey lines  
 172 represent variables with their relative lengths corresponding to their influence on each AZ-DOM.

173 **Table S7. Statistical significance of the correlation of environmental parameters on the DOM molecular**  
 174 **composition in Amazon basin (corresponds to Fig. S17) based on ESI[±] FT-ICR MS and <sup>1</sup>H NMR in**  
 175 **NMDS model.** The environmental and bulk parameters were fitted to NMDS ordination (Bray-Curtis distance  
 176 matrix) of normalized intensity on FT-ICR MS and <sup>1</sup>H NMR (0.01 ppm bucket resolution). The significance of  
 177 the p-values was tested with 999 permutations. Factors printed in bold are significant with p < 0.05.

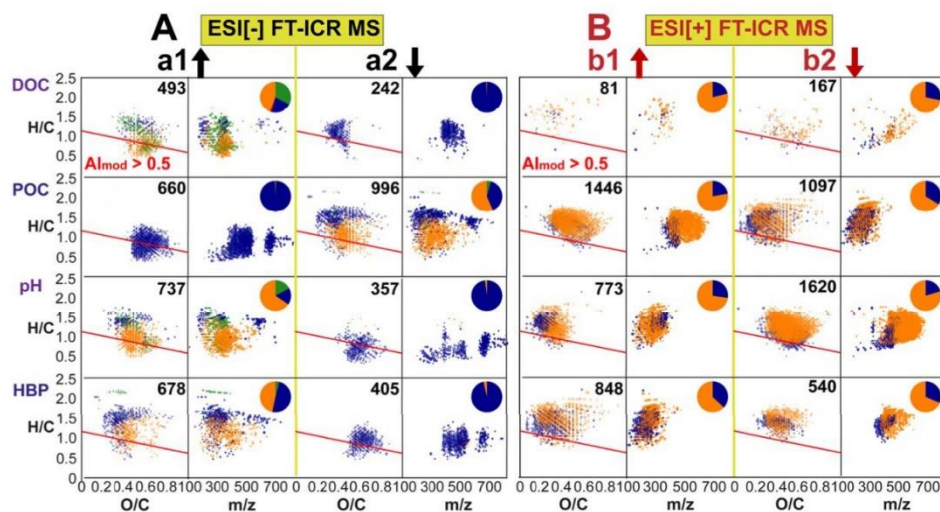
ESI- FT-ICR MS			ESI+ FT-ICR MS			<sup>1</sup> H NMR		
Parameter	r <sup>2</sup>	P-value	Parameter	r <sup>2</sup>	P-value	Parameter	r <sup>2</sup>	P-value
<b>pH</b>	<b>0.7271</b>	<b>0.004</b>	<b>TOC</b>	<b>0.7841</b>	<b>0.001</b>	<b>pH</b>	<b>0.3896</b>	<b>0.049</b>
<b>Cond.</b>	<b>0.3765</b>	<b>0.026</b>	<b>POC</b>	<b>0.8178</b>	<b>0.001</b>	DO	0.3416	0.074
<b>POC</b>	<b>0.3694</b>	<b>0.032</b>	<b>Temp.</b>	<b>0.4667</b>	<b>0.008</b>	POC	0.33	0.078
<b>TOC</b>	<b>0.3472</b>	<b>0.044</b>	<b>pH</b>	<b>0.4616</b>	<b>0.012</b>	TOC	0.3352	0.087
DO	0.2687	0.083	<b>Cond.</b>	<b>0.3647</b>	<b>0.027</b>	DOC	0.3061	0.099
NO <sub>3</sub> <sup>-</sup>	0.2189	0.161	NO <sub>3</sub> <sup>-</sup>	0.2706	0.092	NH <sub>4</sub> <sup>+</sup>	0.2802	0.129
NH <sub>4</sub> <sup>+</sup>	0.1943	0.229	DOC	0.2208	0.172	Temp.	0.0868	0.564
PO <sub>4</sub> <sup>3-</sup>	0.1347	0.276	NO <sub>2</sub> <sup>-</sup>	0.2022	0.197	PO <sub>4</sub> <sup>3-</sup>	0.0817	0.578
DOC	0.1059	0.414	DO	0.1269	0.382	NO <sub>2</sub> <sup>-</sup>	0.0689	0.636
Temp.	0.0567	0.653	PO <sub>4</sub> <sup>3-</sup>	0.0856	0.469	NO <sub>3</sub> <sup>-</sup>	0.0359	0.784
NO <sub>2</sub> <sup>-</sup>	0.0246	0.862	NH <sub>4</sub> <sup>+</sup>	0.0252	0.799	Cond.	0.0102	0.943

178

179 **Table S8. Statistical significance of the correlation of environmental parameters, bacterial production**  
 180 **and carbon fixation on the DOM molecular composition in Amazon basin (corresponds to Fig. 5) based**  
 181 **on ESI[±] FT-ICR MS and <sup>1</sup>H NMR in NMDS model.** The environmental and bulk parameters were fitted to  
 182 NMDS ordination (Bray-Curtis distance matrix) of normalized intensity on FT-ICR MS and <sup>1</sup>H NMR (0.01  
 183 ppm bucket resolution). The significance of the p-values was tested with 999 permutations. Factors printed in  
 184 bold are significant with p < 0.05.

ESI[-] FT-ICR MS			ESI[+] FT-ICR MS			<sup>1</sup> H NMR		
Parameter	r <sup>2</sup>	P-value	Parameter	r <sup>2</sup>	P-value	Parameter	r <sup>2</sup>	P-value
<b>NH<sub>4</sub><sup>+</sup></b>	<b>0.8681</b>	<b>0.009</b>	<b>pH</b>	<b>0.9806</b>	<b>0.001</b>	pH	0.7499	0.052
<b>Depth</b>	<b>0.7465</b>	<b>0.031</b>	<b>POC</b>	<b>0.8841</b>	<b>0.001</b>	HBP	0.5443	0.147
<b>pH</b>	<b>0.6831</b>	<b>0.032</b>	<b>TOC</b>	<b>0.8630</b>	<b>0.003</b>	Depth	0.5017	0.158
HBP	0.5953	0.063	<b>DIC</b>	<b>0.8985</b>	<b>0.003</b>	TOC	0.5550	0.17
TOC	0.5253	0.097	<b>HBP</b>	<b>0.8753</b>	<b>0.013</b>	NO <sub>2</sub> <sup>-</sup>	0.4977	0.176
Secci.	0.5272	0.112	<b>Cond.</b>	<b>0.7462</b>	<b>0.015</b>	Width	0.4421	0.23
POC	0.4854	0.121	<b>Width</b>	<b>0.7953</b>	<b>0.023</b>	DO	0.4145	0.241
Width	0.4731	0.123	Velocity	0.6197	0.054	POC	0.4350	0.271
DIC	0.4213	0.196	Secci.	0.4868	0.123	PO <sub>4</sub> <sup>3-</sup>	0.3415	0.299
Velocity	0.3261	0.309	Depth	0.3892	0.201	DOC	0.3075	0.425
NO <sub>3</sub> <sup>-</sup>	0.2851	0.333	Temp	0.3628	0.257	Cond.	0.2921	0.441
Cond.	0.2717	0.409	NO <sub>3</sub> <sup>-</sup>	0.2573	0.434	NH <sub>4</sub> <sup>+</sup>	0.2764	0.462
DO	0.2239	0.464	Discharge	0.2279	0.458	NO <sub>3</sub> <sup>-</sup>	0.25	0.493
Discharge	0.2146	0.487	DOC	0.2252	0.483	Secci.	0.2117	0.545
PO <sub>4</sub> <sup>3-</sup>	0.1365	0.632	NO <sub>2</sub> <sup>-</sup>	0.2059	0.498	Velocity	0.1652	0.591
NO <sub>2</sub> <sup>-</sup>	0.0404	0.883	NH <sub>4</sub> <sup>+</sup>	0.1895	0.534	DIC	0.1178	0.727
DOC	0.0429	0.897	PO <sub>4</sub> <sup>3-</sup>	0.0637	0.874	Discharge	0.0689	0.838
Temp.	0.0293	0.898	DCF	0.0415	0.901	Temp	0.0542	0.865
DCF	0.03	0.905	DO	0.0226	0.935	DCF	0.0654	0.908

185



187

188 **Fig. S18. Spearman correlation analysis using AZ-DOM composition in the four Amazon main stem and tributaries based on all assigned molecular**  
 189 **formulae in ESI[±] FT-ICR MS with DOC, POC, pH, and heterotrophic bacterial fixation (HBP) rate. The van Krevelen and mass-edited H/C diagrams**  
 190 **show the mass peaks, of which the relative intensities have significant negative/positive correlation ( $p < 0.05$ ,  $r^2 > 0.5$ ), with the color representing elemental**  
 191 **compositions with the assigned CHO (blue), CHNO (orange), and CHOS (green) molecular series and the size of each bubble representing the absolute value**  
 192 **of correlation coefficient  $r$ . The pie plots show the percentages of the counts of different molecular series. Molecular composition positioned below the red line**  
 193 **in the van Krevelen diagrams show the compounds with modified aromaticity index ( $AI_{mod}$ ) higher than 0.5 (Koch and Dittmar, 2006). Numbers show counts**  
 194 **of compounds.**

195 **References**

- 196 Åberg, J. and Wallin, M. 2014. Evaluating a fast headspace method for measuring DIC and  
197 subsequent calculation of pCO<sub>2</sub> in freshwater systems. *Inland Waters* 4 (2), 157-166.
- 198 Benner, R., Opsahl, S., Chin-Leo, G., Richey, J.E. and Forsberg, B.R. 1995. Bacterial carbon  
199 metabolism in the Amazon River system. *Limnol. Oceanogr.* 40 (7), 1262-1270.
- 200 D. C. Smith, F.A. 1992. A simple, economical method for measuring bacterial protein synthesis rates  
201 in seawater using <sup>3</sup>H-leucine. *Mar. Microb. food webs* 6 (2), 107-111.
- 202 Dittmar, T., Koch, B., Hertkorn, N. and Kattner, G. 2008. A simple and efficient method for the solid-  
203 phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr.:*  
204 *Methods* 6 (6), 230-235.
- 205 Farjalla, V.F. 2014. Are the mixing zones between aquatic ecosystems hot spots of bacterial  
206 production in the Amazon River system? *Hydrobiologia* 728 (1), 153-165.
- 207 Gonsior, M., Powers, L.C., Williams, E., Place, A., Chen, F., Ruf, A., Hertkorn, N. and Schmitt-  
208 Kopplin, P. 2019. The chemodiversity of algal dissolved organic matter from lysed  
209 *Microcystis aeruginosa* cells and its ability to form disinfection by-products during chlorination.  
210 *Water Res.* 155 (2019), 300-309.
- 211 Hertkorn, N., Harir, M., Cawley, K.M., Schmitt-Kopplin, P. and Jaffé, R. 2016. Molecular  
212 characterization of dissolved organic matter from subtropical wetlands: a comparative study  
213 through the analysis of optical properties, NMR and FTICR/MS. *Biogeosciences* 13 (8), 2257-  
214 2277.
- 215 Hertkorn, N., Harir, M., Koch, B.P., Michalke, B. and Schmitt-Kopplin, P. 2013. High-field NMR  
216 spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level  
217 characterization of marine dissolved organic matter. *Biogeosciences* 10 (3), 1583-1624.
- 218 Koch, B.P. and Dittmar, T. 2006. From mass to structure: an aromaticity index for high-resolution  
219 mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* 20 (5), 926-932.
- 220 McClain, M.E. and Naiman, R.J. 2008. Andean Influences on the Biogeochemistry and Ecology of  
221 the Amazon River. *BioScience* 58 (4), 325-338.
- 222 Moreira-Turcq, P., Seyler, P., Guyot, J.L. and Etcheber, H. 2003. Exportation of organic carbon from  
223 the Amazon River and its main tributaries. *Hydrol. Process.* 17 (7), 1329-1344.
- 224 Nielsen, E.S. 1960. Dark Fixation of CO<sub>2</sub> and Measurements of Organic Productivity. With Remarks  
225 on Chemo-Synthesis. *Physiol. Plant.* 13 (2), 348-357.
- 226 Nielsen, E.S. 1965. On the Determination of the Activity in <sup>14</sup>C-Ampoules for Measuring Primary  
227 Production. *Limnol. Oceanogr.* 10 (suppl), R247-R252.
- 228 Peterson, B.J. 1980. Aquatic Primary Productivity and the <sup>14</sup>C-CO<sub>2</sub> Method: A History of the  
229 Productivity Problem. *Annu. Rev. Ecol. Evol. Syst.* 11 (1), 359-385.
- 230 Ramette, A. 2007. Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol.* 62 (2), 142-  
231 160.
- 232 Richey, J.E., Hedges, J.I., Devol, A.H., Quay, P.D., Victoria, R., Martinelli, L. and Forsberg, B.R.  
233 1990. Biogeochemistry of carbon in the Amazon River. *Limnol. Oceanogr.* 35 (2), 352-371.
- 234 Santoro, A.L., Bastviken, D., Gudas, C., Tranvik, L. and Enrich-Prast, A. 2013. Dark carbon fixation:  
235 an important process in lake sediments. *PLoS One* 8 (6), e65813.
- 236 Schmitt-Kopplin, P., Gelencser, A., Dabek-Zlotorzynska, E., Kiss, G., Hertkorn, N., Harir, M., Hong,  
237 Y. and Gebefugi, I. 2010. Analysis of the unresolved organic fraction in atmospheric aerosols  
238 with ultrahigh-resolution mass spectrometry and nuclear magnetic resonance spectroscopy:  
239 organosulfates as photochemical smog constituents. *Anal. Chem.* 82 (19), 8017-8026.

- 240 Simon, M. and Azam, F. 1989. Protein content and protein synthesis rates of planktonic marine  
241 bacteria. *Mar. Ecol. Prog.* 51 (3), 201-213.
- 242 Tziotis, D., Hertkorn, N. and Schmitt-Kopplin, P. 2011. Kendrick-analogous network visualisation  
243 of ion cyclotron resonance Fourier transform mass spectra: improved options for the  
244 assignment of elemental compositions and the classification of organic molecular complexity.  
245 *Eur. J. Mass Spectrom. (Chichester)* 17 (4), 415-421.
- 246 Valderrama, J.C. 1981. The simultaneous analysis of total nitrogen and total phosphorus in natural  
247 waters. *Mar. Chem.* 10 (2), 109-122.
- 248



## 6.4 Supplementary information for article 2

### Supporting Information for

#### **Distinct non-conservative behavior of dissolved organic matter after mixing Solimões/Negro and Amazon/Tapajós River waters**

##### AUTHOR NAMES

Siyu Li<sup>a</sup>, Mourad Harir<sup>a,b</sup>, Philippe Schmitt-Kopplin<sup>a,b</sup>, Fausto Machado-Silva<sup>c,d</sup>, Michael Gonsior<sup>e</sup>, David Bastviken<sup>f</sup>, Alex Enrich-Prast<sup>f,g,h</sup>, Juliana Valle<sup>a</sup>, Norbert Hertkorn<sup>a,f\*</sup>

##### AUTHOR ADDRESS

<sup>a</sup> Research Unit Analytical Biogeochemistry, Helmholtz Munich, Ingolstaedter Landstrasse 1, 85764, Neuherberg, Germany.

<sup>b</sup> Chair of Analytical Food Chemistry, Technische Universität München; Alte Akademie 10, 85354, Freising-Weihenstephan, Germany.

<sup>c</sup> Program in Geosciences – Environmental Geochemistry, Chemistry Institute, Fluminense Federal University, 24020-141, Niteroi, Brazil.

<sup>d</sup> Department of Environmental Sciences, University of Toledo, Toledo, OH, 43606, USA.

<sup>e</sup> University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory; Solomons, Maryland 20688, USA.

<sup>f</sup> Department of Thematic Studies – Environmental Change, Linköping University; SE-581 83 Linköping, Sweden.

<sup>g</sup> Biogas Solutions Research Center (BSRC), Linköping University; SE-581 83 Linköping, Sweden.

<sup>h</sup> Multiuser Unit of Environmental Analysis, University Federal of Rio de Janeiro, Rio de Janeiro, Brazil.

Number of pages: 34

Number of Tables: 11

Number of Figures: 14

TABLE OF CONTENTS

<b>Details of Material and Methods</b> .....	<b>5</b>
1. Sample collection.....	5
2. FT-ICR MS analysis.....	6
3. <sup>1</sup> H NMR analysis.....	6
4. Water analysis.....	7
5. Microbial analysis.....	7
6. Statistical Analysis.....	8
<b>Table S1.</b> Description of sample information. ....	<b>8</b>
<b>Table S2.</b> Description of samples in mixing and incubation experiments.....	<b>9</b>
<b>Figure S1.</b> (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra of SPE-DOM in unmixed and mixed S+N water pre- and post- incubation. The FT-ICR mass spectra are exemplified with samples S, N, FSN, S50N50_30min, S50N50_1d, FSSN_B_1d, FNSS_B_1d, and S50N50_5d_1. The ESI[-] and ESI [+] FT-ICR mass spectra in (a1) and (b1) show the regular shaped signal distribution over a wide mass range (m/z 150-950), as well as distinct signatures at nominal neutral mass 400. Respective assignments of molecular compositions are provided for CHO (blue) and CHNO (orange) molecules. ....	<b>10</b>
<b>Table S3.</b> ESI[-] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in SPE-DOM in unmixed and mixed S+N water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental m/z, DBE/C, H/C, N/C, and S/C ratio. ....	<b>11</b>
<b>Table S4.</b> ESI[+] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in SPE-DOM in unmixed and mixed S+N water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental m/z, DBE/C, H/C, N/C, and S/C ratio. ....	<b>12</b>
<b>Figure S2.</b> Loading vectors in PCA of SPE-DOM in unmixed and mixed S+N water pre- and post-incubation based on all assigned molecular formulae in (A, B) ESI[±] FT-ICR mass spectra and (C) <sup>1</sup> H NMR spectra . The scatter plots of the very PCA see Figure 2. Panels A and B show correlation patterns of PC1 and PC2 loading vectors (p1 and p2) in van Krevelen and mass-edited H/C diagrams, with the color code representing CHO (blue), CHNO (orange), and CHOS (green) molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have modified aromaticity index (AI <sub>mod</sub> ) <sup>15</sup> higher than 0.5. Panel C show correlation patterns of p1 and p2 in <sup>1</sup> H NMR space, with fundamental molecular structures indicated. ....	<b>13</b>
<b>Table S5.</b> Water parameters in unmixed and mixed S+N and A+T waters pre- and post- incubation. ....	<b>14</b>
<b>Figure S3.</b> HCA of assigned SPE-DOM molecular formulae in part of the unmixed and mixed S+N water pre- and post- incubation. Van Krevelen, mass-edited H/C, and KMD/z* diagrams <sup>16</sup> show CHO (blue), CHNO (orange), and CHOS (green) molecular classes that were relatively more abundant in clusters a1/a2/a3, b1/b2/b3, respectively. Bubble areas represent the normalized intensities of SPE-DOM in each cluster. The pie plots depict percentages of the counts of different molecular classes. The numbers show counts of compounds. The molecular formulae positioned	



below the purple line in the van Krevelen diagrams have modified aromaticity index ( $AI_{mod}$ )<sup>15</sup> higher than 0.5. .... 15

**Figure S4.** Common and unique molecular signatures in SPE-DOM in the unmixed and mixed S+N water pre- and post- incubation. The numbers show counts of compounds. Van Krevelen and mass-edited H/C diagrams show CHO (blue), CHNO (orange), CHOS (green), and CHNOS (red) m/z ions that were shared before and after incubation (a1 and b1); unique in original water but were not in incubated water (a2 and b2). Not any single compound unique to incubated water did not present in original water. Bubble areas represent the normalized intensity of DOM. The pie plots depict percentages of the counts of different molecular classes. Molecular composition positioned below the purple line in the van Krevelen diagrams shows the compounds with a modified aromaticity index ( $AI_{mod}$ ) higher than 0.5<sup>15</sup>. Panel C shows the normalized intensity of molecular signatures that were degraded during the incubation process (a2 and b2) in DOM of the Amazon mainstem. .... 16

**Figure S5.** <sup>1</sup>H NMR spectra of SPE-DOM in the unmixed and mixed S+N water pre- and post-incubation. The <sup>1</sup>H NMR spectra are exemplified with samples S, N, FSFN, S50N50\_30min, S50N50\_1d, FSSN\_B\_1d, FNSS\_B\_1d, and S50N50\_5d\_1. <sup>1</sup>H NMR spectra are area normalized (800 MHz, CD<sub>3</sub>OD;  $\delta_H = 0.5$ -10.0 ppm with the exclusion of residual water and methanol NMR resonances). The <sup>1</sup>H NMR spectra referred to average abundance are edited by each normalized spectra subtract the average spectra (C). .... 17

**Table S6.** <sup>1</sup>H NMR section integrals (percent of non-exchangeable protons, 800 MHz; CD<sub>3</sub>OD, exclusion of residual water, and methanol) and key substructures of SPE-DOM in the unmixed and mixed S+N water pre- and post- incubation. The fundamental substructures include aromatics C<sub>ar</sub>H,  $\delta_H \sim 7.0$ -10.0 ppm; olefins  $\delta_H \sim 5.3$ -7.0 ppm; oxygenated aliphatic units (OCH) and “carbohydrate-like” and methoxy OCH<sub>3</sub> units  $\delta_H \sim 3.2$ -4.9 ppm; branched aliphatic units (CH<sub>2</sub>)<sub>n</sub>, “acetate-analogue” and CRAM  $\delta_H \sim 1.9$ -3.2 ppm; functionalized aliphatics  $\delta_H \sim 1.35$ -1.9 ppm; polyethylene group  $\delta_H \sim 1.25$ -1.35 ppm, OCC<sub>2</sub>H units; pure aliphatics  $\delta_H \sim 0.5$ -1.25 ppm, CCC<sub>2</sub>H units. .... 18

**Figure S6.** Difference <sup>1</sup>H NMR spectra (800 MHz, CD<sub>3</sub>OD) of Solimões and Negro SPE-DOM according to status of incubation. (A) Solimões minus Negro, with Solimões (S), Negro (N), and Solimões minus Negro (S-N). (B) difference <sup>1</sup>H NMR spectrum after 30 minutes of incubation, with S, N, averaged Solimões and Negro ((S+N)/2), S50N50\_30min, and (S+N)/2 minus S50N50\_30min. (C) difference <sup>1</sup>H NMR spectrum after one day of incubation, with S, N, (S+N)/2, S50N50\_1d, and (S+N)/2 minus S50N50\_1d. (D) difference <sup>1</sup>H NMR spectrum after five days of incubation, with S, N, (S+N)/2, S50N50\_5d, and (S+N)/2 minus S50N50\_5d. (E) difference <sup>1</sup>H NMR spectrum of one day minus five days of incubation, with S50N50\_1d, S50N50\_5d, and (S50N50\_1d minus S50N50\_5d). .... 19

**Figure S7.** Difference <sup>1</sup>H NMR spectra (800 MHz, CD<sub>3</sub>OD) of Solimões and Negro SPE-DOM according to status of incubation, section of unsaturated C<sub>sp2</sub>H protons  $\delta_H \sim 5.5$ -10 ppm in Figure S6. (A) Solimões minus Negro, with Solimões (S), Negro (N), and Solimões minus Negro (S-N). (B) difference after 30 minutes of incubation, with S, N, averaged Solimões and Negro ((S+N)/2), S50N50\_30min, and (S+N)/2 minus S50N50\_30min. (C) difference after one day of incubation, with S, N, (S+N)/2, S50N50\_1d, and (S+N)/2 minus S50N50\_1d. (D) difference <sup>1</sup>H NMR spectrum after five days of incubation, with S, N, (S+N)/2, S50N50\_5d, and (S+N)/2 minus S50N50\_5d. (E) difference <sup>1</sup>H NMR spectrum one day minus five days of incubation, with S50N50\_1d, S50N50\_5d, and (S50N50\_1d minus S50N50\_5d). .... 20

**Figure S8.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Solimões and Negro river waters, entire section  $\delta_{\text{H}} \sim 0.5\text{-}10$  ppm), and proximate Amazon SPE-DOM samples A1 and A2 (sampling locations see Figure 1 and Table S1). (A) (averaged Solimões and Negro) river minus proximate downstream Amazon River sampling points A1 (orange, dotted line) and A2 (pink, dotted line), with Solimões (S, grey), Negro (N, black), and averaged Solimões and Negro ((S+N)/2, dotted green line). (B) same spectra, section of unsaturated  $\text{C}_{\text{sp}^2}\text{H}$  protons ( $\delta_{\text{H}} \sim 5.5\text{-}10$  ppm). ..... 21

**Figure S9.** (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra of SPE-DOM in unmixed and mixed A+T water pre- and post- incubation. The FT-ICR mass spectra are exemplified with samples *A*, *T*, A50T50\_1d, and A50T50\_5d. The ESI[-] and ESI [ + ] FT-ICR mass spectra in (a1) and (b1) show the regular shaped signal distribution over a wide mass range ( $m/z$  150-950), as well as distinct signatures at nominal neutral mass 400. Respective assignments of molecular compositions are provided for CHO (blue) and CHNO (orange) molecules. .... 22

**Table S7.** ESI[-] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in unmixed and mixed A+T water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental  $m/z$ , DBE/C, H/C, N/C, and S/C ratio. .... 23

**Table S8.** ESI[+] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in unmixed and mixed A+T water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental  $m/z$ , DBE/C, H/C, N/C, and S/C ratios. .... 24

**Figure S10.** HCA of assigned SPE-DOM molecular formulae in the unmixed and mixed A+T water pre- and post-incubation. Van Krevelen, mass-edited H/C, and KMD/z\* diagrams<sup>16</sup> show CHO (blue), CHNO (orange), and CHOS (green) molecular classes that were relatively more abundant in clusters a1/a2, b1/b2, respectively. The pie plots depict percentages of the counts of different molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have modified aromaticity index ( $\text{AI}_{\text{mod}}$ )<sup>15</sup> higher than 0.5. .... 25

**Figure S11.** Common and unique molecular signatures in unmixed and mixed A+T water pre- and post- incubation. The numbers show counts of compounds. Van Krevelen and mass-edited H/C diagrams show CHO (blue), CHNO (orange), CHOS (green), and CHNOS (red)  $m/z$  ions that were shared before and after incubation (a1 and b1); unique in original water but were not in incubated water (a2 and b2); unique in incubated water but were not in original water (a3 and b3); unique in one-day incubated water (1d) but were not present in five-day incubated water (5d) (a4 and b4). There was no compound unique to 5d and not present 1d. The pie plots depict percentages of the counts of different molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have modified aromaticity index ( $\text{AI}_{\text{mod}}$ )<sup>15</sup> higher than 0.5. .... 26

**Figure S12.**  $^1\text{H}$  NMR spectra of SPE-DOM in unmixed and mixed A+T water pre- and post-incubation. The  $^1\text{H}$  NMR spectra are exemplified with sample *A*, *T*, A50T50\_1d, and A50T50\_5d.  $^1\text{H}$  NMR spectra are entire region normalized (800 MHz,  $\text{CD}_3\text{OD}$ ;  $\delta_{\text{H}} = 0.5\text{-}10.0$  ppm with exclusion of residual water and methanol NMR resonances). The  $^1\text{H}$  NMR spectra referred to average abundance in the four DOM are edited by each normalized spectra subtract the average spectra of the four DOM (C). .... 27



**Table S9.**  $^1\text{H}$  NMR section integrals (percent of non-exchangeable protons, 800 MHz;  $\text{CD}_3\text{OD}$ , exclusion of residual water, and methanol) and key substructures of SPE-DOM in unmixed and mixed A+T water pre- and post- incubation. The fundamental substructures include aromatics  $\text{C}_{\text{ar}}\text{H}$ ,  $\delta_{\text{H}} \sim 7.0\text{-}10.0$  ppm; olefins  $\delta_{\text{H}} \sim 5.3\text{-}7.0$  ppm; oxygenated aliphatic units ( $\text{OCH}$ ) and “carbohydrate-like” and methoxy  $\text{OCH}_3$  units  $\delta_{\text{H}} \sim 3.2\text{-}4.9$  ppm; branched aliphatic units ( $\text{CH}_2$ )<sub>n</sub>, “acetate-analogue” and CRAM  $\delta_{\text{H}} \sim 1.9\text{-}3.2$  ppm; functionalized aliphatics  $\delta_{\text{H}} \sim 1.35\text{-}1.9$  ppm; polyethylene group  $\delta_{\text{H}} \sim 1.25\text{-}1.35$  ppm,  $\text{OCC}\text{H}$  units; pure aliphatics  $\delta_{\text{H}} \sim 0.5\text{-}1.25$  ppm,  $\text{CCC}\text{H}$  units. .... **28**

**Figure S13.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Amazon and Tapajós river SPE-DOM according to status of incubation (full vertical expansion cf. Figure S14). (A) Amazon minus Tapajós river, with Amazon (A), Tapajós (T), and Amazon minus Tapajós (A-T). (B) difference  $^1\text{H}$  NMR spectrum after one day of incubation, with A, T, averaged Amazon and Tapajós ((A+T)/2), A50T50\_1d, and (A+T)/2 minus A50T50\_1d. (C) difference  $^1\text{H}$  NMR spectrum after five days of incubation, with A, T, (A+T)/2, A50T50\_5d, and (A+T)/2 minus A50T50\_5d. (D) difference  $^1\text{H}$  NMR spectrum one day minus five days of incubation, with A50T50\_1d, A50T50\_5d, and (A50T50\_1d minus A50T50\_5d). .... **29**

**Figure S14.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Amazon and Tapajós river SPE-DOM according to status of incubation. (A) Amazon minus Tapajós river, with Amazon (A), Tapajós (T), and Amazon minus Tapajós (A-T). (B) difference  $^1\text{H}$  NMR spectrum after one day of incubation, with A, T, averaged Amazon and Tapajós ((A+T)/2), A50T50\_1d, and (A+T)/2 minus A50T50\_1d. (C) difference  $^1\text{H}$  NMR spectrum after five days of incubation, with A, T, (A+T)/2, A50T50\_5d, and (A+T)/2 minus A50T50\_5d. (D) difference  $^1\text{H}$  NMR spectrum one day minus five days of incubation, with A50T50\_1d, A50T50\_5d, and (A50T50\_1d minus A50T50\_5d). .... **30**

**Table S10.** HBP and DCF in unmixed and mixed S+N and A+T waters pre- and post- incubation. SD, Standard Deviation; n, number of replicates; N.A., not available. .... **31**

**Table S11.** Ranges of heterotrophic bacterial production (HBP) and dark carbon fixation (DCF) in this study and in the Amazon River other aquatic systems. .... **32**

## Details of Material and Methods

### 1. Sample collection

The Amazon River main stem originate in Andean mountains springs and for most of its flow path meanders across the lowland of Amazon tropical rainforest until its mouth in Atlantic Ocean, while collecting major tributaries on its way. The main stem of the Amazon River is named Ucayali while flowing in Peru and when it crosses Brazil's boundary it changes the name to Solimões. Downstream of the confluence with the Negro River in Manaus (Amazonas state, Brazil) it changes the name again to Amazon and keeps this name until its mouth on Atlantic Ocean in Brazil's coast. There are many large tributaries with different water characteristics. Three fundamental types of water were already described by Alfred Russel Wallace in 1853 in his trip to the Amazon basin, and these were called as white, clear and blackwaters<sup>1</sup>. The color varies depending on whether dissolved organic matter (black water) or suspended sediment (white water) is predominant<sup>2</sup>. The clear water rivers are transparent and supply a high phytoplankton production<sup>3</sup>. Both white and clear waters have more aquatic plants and floating meadows than black waters. In contrast, the black water Negro River is nearly devoid of such vegetation. Solimões contributed the most carbon to the Amazon River, about  $500 \text{ kg C s}^{-1}$  during the sampling high water period<sup>2</sup>. Madeira is the largest and Negro is the second largest tributary in Amazon

watershed. These two rivers together contribute more than 30% of discharge of the Amazon River main stem. Madeira is a primary tributary of Amazon in terms of sediment input. Sampling location coordinates and information of the water samples is described in Tables S1, S2.

### 2. FT-ICR MS analysis

Negative and positive electrospray ionization (ESI[±]) Fourier transform ion cyclotron resonance mass spectra (FT-ICR MS) were acquired using a 12T Bruker Solarix mass spectrometer (Bruker Daltonics, Bremen, Germany) and an Apollo II electrospray ionization (ESI) source<sup>4</sup>. Nebulizer gas pressure, drying gas pressure and the source heater temperature were 138 kPa, 103 kPa and 200 °C, respectively. The spectra were acquired with a time domain of 4 MW. For each sample, 500 broadband scans were accumulated in ESI[-] FT-ICR mass spectra, while 300 broadband scans were accumulated in ESI[+] FT-ICR mass spectra. All spectra were first externally calibrated on clusters of arginine in MeOH (0.57 mmol/L) and internally calibrated using appropriate reference mass lists of common natural organic matter molecules, reaching accuracy values lower than 500 ppb. Data processing was done using Compass Data Analysis 5.0 (Bruker, Bremen, Germany) and formula assignment by an in-house made software (NetCalc)<sup>5</sup>. The molecular formula assignments were based on the following elements: <sup>1</sup>H<sub>0-200</sub>, <sup>12</sup>C<sub>0-100</sub>, <sup>16</sup>O<sub>0-80</sub>, <sup>32</sup>S<sub>0-3</sub>, <sup>14</sup>N<sub>0-3</sub> as well as the <sup>13</sup>C<sub>0-1</sub> and <sup>34</sup>S<sub>0-1</sub> isotopomers. The generated formulae were validated by setting sensible chemical constraints (N rule, O/C < 1, H/C < 2n + 2 (C<sub>n</sub>H<sub>2n+2</sub>)). Restriction on nitrogen atoms ≤ 4 and sulfur atoms ≤ 2 were applied based on previous studies<sup>6,7</sup>. Final elemental formulae were generated and categorized into groups containing CHO, CHNO, CHOS, and CHNOS molecular series, which were used to reconstruct the group-selective mass spectra. The intensity of mass peaks was normalized (total mass peak amplitude of assigned mass peaks = 100% for single sample). The average H/C, O/C, N/C, S/C atomic ratios, DBE/C (Double bond equivalent per carbon), mass-to-charge ratios (*m/z*) were computed from the intensity-weighted average of molecular formulae<sup>4</sup>.

### 3. <sup>1</sup>H NMR analysis

All <sup>1</sup>H NMR spectra were acquired with a Bruker Advance III NMR spectrometer operating at 800.35 MHz (B<sub>0</sub> = 18.8 Tesla) at 283 K from redissolved dried DOM in CD<sub>3</sub>OD (99.95% <sup>2</sup>H; Merck) with Bruker standard pulse sequences using 3.0 mm Bruker MATCH tubes. The reference <sup>1</sup>H NMR chemical shift of HD<sub>2</sub>COD was 3.3 ppm. <sup>1</sup>H NMR spectra were recorded under solvent suppression with presaturation and 1 ms spin-lock (noesypr1d), 5 s acquisition time, 5 s relaxation delay (d1), typically 1024 scans and 1 Hz exponential line broadening. <sup>1</sup>H NMR section integrals were obtained by using the software AMIX at 0.01 ppm resolution, with exclusions of HDO and HD<sub>2</sub>COD NMR resonances.

Difference <sup>1</sup>H NMR spectra used full data point resolution (1.5×10<sup>-4</sup> ppm) and were computed from area normalized NMR spectra (δ<sub>H</sub>: 0.5-10.0 ppm), with exclusion of methanol (δ<sub>H</sub>: 3.2-3.4 ppm) and water section chemical shifts (δ<sub>H</sub>: 4.8-5.5 ppm) by Bruker AMIX software; data tables were imported into EXCEL tables, and difference NMR spectra were computed in EXCEL and exported to Adobe Illustrator. Vertical axes denote fraction of NMR resonance amplitude with respect to total <sup>1</sup>H NMR integral (100%).

### 4. Water analysis

Water conductivity (Cond.), temperature (Temp.) and dissolved oxygen (DO) were measured in situ with portable instruments (Hanna Instruments, Metrohm electrode and PRO-ODO YSI). For the inorganic dissolved nutrients, the water was sampled with polypropylene bottles, filtered in silica filters (GF/F, Whatman) and placed in amber bottles conditioned in freezer to be analyzed in the laboratory. Ammonium (NH<sub>4</sub><sup>+</sup>) concentrations were determined by colorimetric analysis with flux injection (FIA, FiaStar 5000). Nitrate (NO<sub>3</sub><sup>-</sup>) and nitrite (NO<sub>2</sub><sup>-</sup>) were measured using



nitrate reduction and phosphate ( $\text{PO}_4^{3-}$ ) using molybdenum analytical method<sup>8</sup>. The dissolved organic carbon (DOC) was determined by persulfate digestion analyzed using high-temperature catalytic oxidation method (Sievers InnovOx analyzer, GE)<sup>9</sup>. The total organic carbon (TOC) used the same analytical method as DOC without water filtration.

### 5. Microbial analysis

HBP was estimated by the measurement of protein synthesis rates using  $^3\text{H}$ -leucine<sup>10</sup>, previously tested in the Amazon River mixing zones<sup>11</sup>. We added 150  $\mu\text{L}$  of Leucine [ $^3\text{H}$ ] (at the final concentration of 20 nM with the specific radioactivity of 10 Ci  $\text{mmol}^{-1}$ , Perkin Elmer) to 1.5 mL of river water in five replicates and two controls. We used formalin at a final concentration of 3.7% to stop the incubation after 45 minutes and to fix two controls before the leucine addition. In the laboratory, the bacterial protein was extracted using cold trichloroacetic acid (TCA) and ethanol protocol<sup>12</sup>. The precipitated protein pellet was re-suspended in the scintillation liquid (OptiphaseHiSafe, PerkinElmer) and leucine incorporation was calculated as the net decay per minute based in controls and the concentration factor calculated based on the leucine added<sup>12</sup>. Then, we calculate HBP converting leucine incorporation in carbon assimilation using the molecular weight of leucine, the proportion of leucine in total protein, the ratio of cellular carbon to protein, and the intracellular isotope dilution of Leucine<sup>12</sup>.

DCF is a proxy of chemosynthesis that was estimated by the incorporation of dissolved inorganic carbon ( $^{14}\text{C}$ -DIC) as done for radiotracer photosynthesis measurements, but under dark conditions<sup>13</sup>. The incubations were performed by the addition of 10  $\mu\text{L}$  of  $\text{NaH}^{14}\text{CO}_3$  (Perkin Elmer, the specific activity of 52.5 mCi  $\text{mmol}^{-1}$ ) in 50 mL of river water in amber bottles at each sampling site using five replicates and two controls. The incubations were stopped after 4 hours by adding formaldehyde to the final concentration of 3.7 %. In the laboratory, samples extracted using filtration in glass fiber filters (GF/F, Whatman) and cellulose acetate filters (0.2  $\mu\text{m}$  pore, Sartorius Stedim Biotech), washed twice and placed with hydrochloric acid (HCl) 1 M to remove traces of inorganic carbon. Finally, filters activity were counted using scintillation liquid (OptiphaseHiSafe, PerkinElmer) and scintillation counter 24 h later (Tricarb 2800, Perkin Elmer). We estimated dissolved inorganic carbon (DIC) in the water using the acidified headspace method<sup>14</sup> using 12 mL of water sample into small, sealed glass vials (25 mL) prefilled with 50  $\mu\text{L}$  phosphoric acid (20%), which was empirically tested to lower the pH ( $\sim 2$ ). In the laboratory, samples were analyzed for  $\text{CO}_2$  using gas chromatography (Agilent Technologies, USA, 7890A) with a 1.8 m  $\times$  3.175 mm Porapak Q 80/100 column from Supelco, a methanizer converting  $\text{CO}_2$  to methane and determination by a flame ionization detector (FID) by manual injection. The  $\text{CO}_2$  concentration measured in the headspace is then recalculated to DIC because the lowered pH ( $\sim 2$ ) drives the DIC into  $\text{CO}_2$  with partitioning between the water and gas phases according to Henry's law. Then, DCF the fraction of assimilated  $^{14}\text{C}$  was calculated as the sum of both filters times the total DIC levels present. Then, DCF was calculated as the DIC incorporation per time for each sample based on the net decay per minute based on controls and the isotopic discrimination factor<sup>13</sup>.

### 6. Statistical analysis

Principal component analysis (PCA) was performed using Simca-P (version 11.5, UmetricsAB, Umeå, Sweden) to identify the dominant modes of variability in DOM composition of AZ-DOM. FT-ICR Mass spectra were arranged with the samples as observations and the peak areas of the assigned FT-ICR mass peaks as the response variables.  $^1\text{H}$  NMR spectra were arranged with the samples as observations and the NMR resonances as the response variables (800 MHz  $^1\text{H}$  NMR,  $\text{CD}_3\text{OD}$ , area-normalized from 0.5-10.0 ppm; 0.01 ppm bucket resolution; with the exclusion of

residual water and methanol NMR resonances). Before multivariate statistics were performed, the response variables were centered and scaled to unit variance. The based weight was computed as  $1/\sqrt{\text{standard deviation of the response variables}}$ .

Hierarchical Cluster Analysis (HCA) was performed using the Hierarchical Clustering Explorer 3.0 (HCE; <http://www.cs.umd.edu/hcil/multi-cluster/>). Average Linkage (UPGMA) method was used to cluster the dataset and Euclidean distance was used as the similarity/distance measure. Based on the HCA, we used the “profile search” tool from HCE 3.0, choosing a search method (model-based), a distance measure (Pearson's r) and a threshold (0.9).

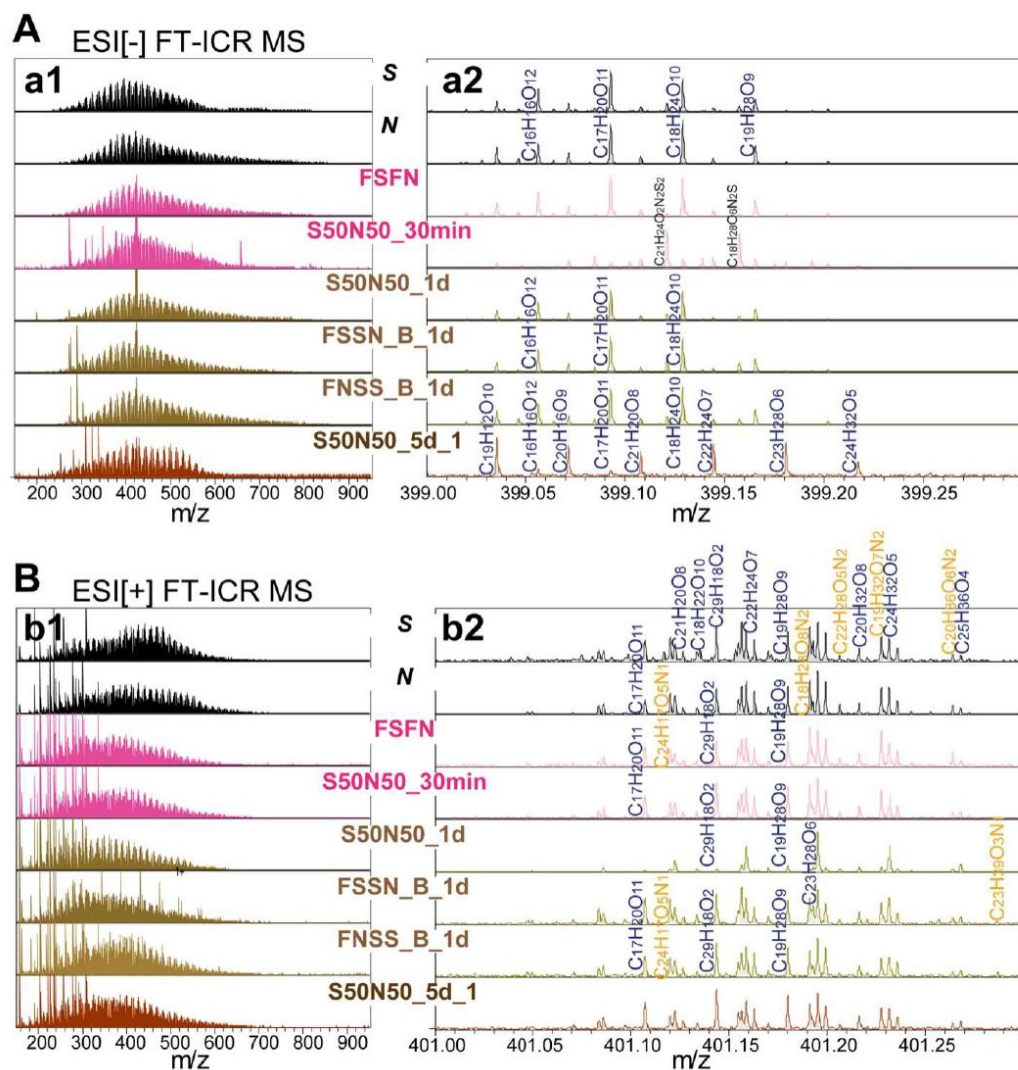
**Table S1.** Description of sample information.

Sample	River	Description	Longitude (°W)	Latitude (°S)	Water type
<i>S</i>	Solimões	used for incubation	-60.0957	-3.3136	white
SM	Solimões	Solimões close to SN mixing zone	-59.8949	-3.1647	turbid
<i>N</i>	Negro	used for incubation	-60.0902	-3.1242	black
NaM	Negro	Negro after Manaus	-59.9126	-3.1292	black
MZ	SN mixing zone	SN mixing zone	-59.8958	-3.1301	turbid
A1	Amazon	before Madeira River inflow	-60.1596	-2.7933	turbid
A2	Amazon	before Madeira River inflow	-58.7621	-3.2992	turbid
AM1	Amazon	after Madeira River inflow	-58.5192	-3.1802	turbid
AM2	Amazon	after Madeira River inflow	-58.5513	-3.1492	turbid
Ad1	Amazon	Amazon downstream	-55.7717	-2.0033	turbid
Ad2	Amazon	Amazon downstream	-55.4316	-2.0180	turbid
<i>A</i>	Amazon	used for incubation	-54.7185	-2.4041	turbid
At	Amazon	a small river branch in Amazon after Tapajós inflow	-54.4518	-2.4916	turbid
T1	Tapajós	Tapajós upstream	-55.1013	-2.9714	clear
T2	Tapajós	Tapajós downstream	-54.9001	-2.4315	clear
T3	Tapajós	Tapajós downstream	-54.8332	-2.2742	clear
<i>T</i>	Tapajós	used for incubation	-54.7219	-2.4056	clear



**Table S2.** Description of samples in mixing and incubation experiments.

Sample	water mixture	incubation time
FSFN	50% Solimões + 50% Negro	0 minutes
S50N50_30min	50% Solimões + 50% Negro	30 minutes
S20N80_1d	20% Solimões + 80% Negro	one day
S40N60_1d	40% Solimões + 60% Negro	one day
S50N50_1d	50% Solimões + 50% Negro	one day
S60N40_1d	60% Solimões + 40% Negro	one day
S80N20_1d	80% Solimões + 20% Negro	one day
FSSN_B-1d	Solimões filtered water + Negro filtered suspended solids	one day at in situ temperature (30 °C)
FSSN_S-1d	Solimões filtered water + Negro filtered suspended solids	one day on ice (0 °C)
FNSS_B-1d	Negro filtered water + Solimões filtered suspended solids	one day at in situ temperature (30 °C)
FNSS_S-1d	Negro filtered water + Solimões filtered suspended solids	one day on ice (0 °C)
S20N80_5d	20% Solimões + 80% Negro	five days
S40N60_5d	40% Solimões + 60% Negro	five days
S50N50_5d_1	50% Solimões + 50% Negro	five days
S50N50_5d_2	50% Solimões + 50% Negro	five days
S50N50_5d_3	50% Solimões + 50% Negro	five days
S60N40_5d	60% Solimões + 40% Negro	five days
S80N20_5d	80% Solimões + 20% Negro	five days
A20T80_1d	20% Amazon + 80% Tapajós	one day
A40T60_1d	40% Amazon + 60% Tapajós	one day
A50T50_1d	50% Amazon + 50% Tapajós	one day
A60T40_1d	60% Amazon + 40% Tapajós	one day
A80T20_1d	80% Amazon + 20% Tapajós	one day
A20T80_5d	20% Amazon + 80% Tapajós	five days
A40T60_5d	40% Amazon + 60% Tapajós	five days
A50T50_5d	50% Amazon + 50% Tapajós	five days
A60T40_5d	60% Amazon + 40% Tapajós	five days
A80T20_5d	80% Amazon + 20% Tapajós	five days



**Figure S1.** (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra of SPE-DOM in unmixed and mixed S+N water pre- and post- incubation. The FT-ICR mass spectra are exemplified with samples *S*, *N*, FSFN, S50N50\_30min, S50N50\_1d, FSSN\_B\_1d, FNSS\_B\_1d, and S50N50\_5d\_1. The ESI[-] and ESI [+] FT-ICR mass spectra in (a1) and (b1) show the regular shaped signal distribution over a wide mass range ( $m/z$  150-950), as well as distinct signatures at nominal neutral mass 400. Respective assignments of molecular compositions are provided for CHO (blue), CHNO (orange), and few CHNOS (black) molecules (a2).

**Table S3.** ESI[-] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in SPE-DOM in unmixed and mixed S+N water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental  $m/z$ , DBE/C, H/C, N/C, and S/C ratios.

Sample	total counts of mass peaks	CHO%	CHNO%	CHOS%	CHNOS%	$m/z$	DBE/C	H/C	O/C	N/C $\times 10^{-2}$	S/C $\times 10^{-3}$
S	5043	59.5	32.4	7.2	1.0	460.2	0.53	1.06	0.55	0.76	1.12
SM	4408	65.7	29.0	5.1	0.2	479.9	0.52	1.06	0.54	0.60	0.79
N	4233	74.4	23.9	1.7	0.0	496.5	0.53	1.04	0.54	0.31	0.23
NaM	3781	76.9	21.8	1.4	0.0	498.3	0.53	1.04	0.54	0.29	0.17
MZ	4424	73.4	24.8	1.8	0.0	492.3	0.52	1.05	0.53	0.33	0.23
A1	4675	68.7	27.7	3.6	0.0	488.9	0.52	1.05	0.54	0.45	0.43
A2	4468	69.0	27.3	3.7	0.0	485.1	0.52	1.07	0.54	0.47	0.64
AM1	4120	63.8	30.4	5.5	0.3	467.9	0.52	1.08	0.54	0.74	1.01
AM2	4522	67.2	28.6	4.1	0.1	482.9	0.51	1.07	0.54	0.55	0.59
FSFN	3719	74.3	23.7	2.0	0.0	488.5	0.53	1.04	0.55	0.40	0.43
S50N50_30min	2284	78.9	19.6	1.5	0.0	477.9	0.45	1.19	0.43	0.35	0.86
S20N80_1d	3777	75.8	23.0	1.2	0.0	492.7	0.53	1.04	0.54	0.33	0.33
S40N60_1d	3886	73.5	24.6	2.0	0.0	487.3	0.53	1.04	0.55	0.37	0.53
S50N50_1d	4188	70.9	26.8	2.3	0.0	485.3	0.53	1.04	0.55	0.43	0.35
S60N40_1d	4157	70.6	26.1	3.3	0.0	484.8	0.53	1.04	0.55	0.45	0.51
S80N20_1d	4205	69.1	27.4	3.5	0.0	479.1	0.52	1.05	0.54	0.48	0.53
FSSN_B_1d	4017	66.7	28.4	4.9	0.0	470.8	0.51	1.08	0.54	0.57	1.18
FSSN_S_1d	3707	67.4	28.6	3.9	0.1	469.5	0.52	1.07	0.55	0.58	0.75
FNSS_B_1d	3689	75.9	22.7	1.4	0.0	489.6	0.52	1.05	0.55	0.31	0.34
FNSS_S_1d	3699	76.2	22.1	1.8	0.0	492.4	0.52	1.05	0.54	0.30	0.64
S20N80_5d	4251	71.1	26.4	2.5	0.1	483.8	0.52	1.05	0.53	0.39	0.35
S40N60_5d	4260	69.6	27.1	3.3	0.0	477.5	0.52	1.05	0.53	0.44	0.42
S50N50_5d_1	1719	93.6	5.4	1.1	0.0	439.6	0.51	1.08	0.41	0.90	0.93
S50N50_5d_2	4229	66.4	27.4	6.2	0.1	455.7	0.54	1.03	0.56	0.51	0.87
S50N50_5d_3	1555	94.0	4.4	1.7	0.0	431.8	0.51	1.07	0.42	0.87	2.17
S60N40_5d	2599	77.0	20.6	2.4	0.1	477.1	0.43	1.23	0.41	0.43	0.70
S80N20_5d	4301	67.5	28.1	4.4	0.0	482.6	0.53	1.05	0.54	0.51	0.61

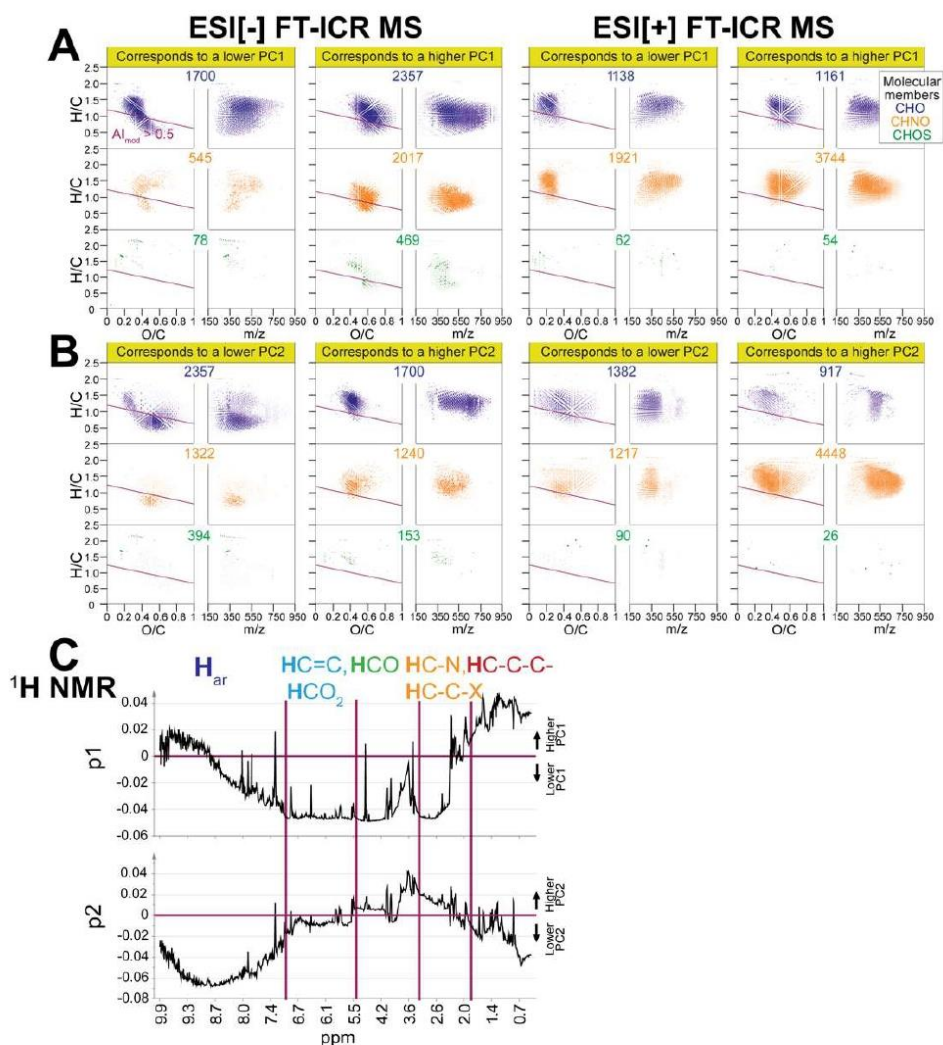
Footnote: 1) FSFN and S50N50\_30min show much less numerous assigned mass peaks, higher %CHO and lower %CHNO molecules; 2) Small decrease of average  $m/z$  at 5d; 3) Some decline of O/C ratio at 5d; 4) The average S/C ratio was higher in FSSN\_B\_1d, showing that more S-containing molecules in the mixture of filtered Solimões River water and suspended solids from the Negro River after incubation at in situ temperature; 5) The lower S/C ratio in FSSN\_S\_1d than in FSSN\_B\_1d was likely because the production of S-containing molecules was slowed down under lower processing.

**Table S4.** ESI[+] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in SPE-DOM in unmixed and mixed S+N water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental  $m/z$ , DBE/C, H/C, N/C, and S/C ratios.

Sample	total counts of mass peaks	CHO%	CHNO%	CHOS%	CHNOS%	$m/z$	DBE/C	H/C	O/C	N/C $\times 10^{-2}$	S/C $\times 10^{-3}$
S	4967	29.5	70.2	0.1	0.2	439.2	0.40	1.34	0.41	3.89	0.02
SM	5172	28.5	71.0	0.2	0.3	413.2	0.38	1.38	0.37	4.05	0.06
N	5727	30.0	69.9	0.1	0.0	442.5	0.39	1.35	0.40	3.63	0.03
NaM	5428	28.9	70.9	0.2	0.0	439.2	0.39	1.36	0.39	3.66	0.03
MZ	5808	28.6	71.2	0.2	0.0	437.7	0.38	1.38	0.39	3.76	0.05
A1	5159	30.0	69.7	0.1	0.2	422.3	0.39	1.37	0.39	3.95	0.03
A2	4949	29.6	70.1	0.2	0.0	417.1	0.39	1.37	0.38	4.11	0.04
AM1	5081	27.7	71.6	0.2	0.5	407.2	0.39	1.38	0.38	4.25	0.06
AM2	5377	29.2	70.1	0.2	0.5	419.1	0.39	1.37	0.39	4.03	0.06
FSFN	3272	35.2	64.7	0.1	0.0	392.9	0.39	1.36	0.38	3.78	0.06
S50N50_30min	3759	35.1	64.8	0.1	0.1	400.0	0.40	1.35	0.40	3.51	0.11
S20N80_1d	2385	39.5	60.3	0.2	0.0	382.6	0.38	1.39	0.36	3.44	0.21
S40N60_1d	2876	39.3	60.3	0.4	0.0	384.9	0.38	1.38	0.38	3.45	0.34
S50N50_1d	3378	46.1	53.6	0.3	0.0	385.0	0.40	1.32	0.31	1.85	0.10
S60N40_1d	2732	37.5	62.4	0.2	0.0	378.8	0.38	1.38	0.37	3.57	0.18
S80N20_1d	3084	39.4	60.4	0.2	0.0	386.7	0.39	1.37	0.38	3.49	0.13
FSSN_B_1d	3837	34.7	64.9	0.2	0.2	396.5	0.41	1.33	0.41	3.62	0.13
FSSN_S_1d	3827	40.4	59.3	0.2	0.1	391.0	0.43	1.29	0.41	2.98	0.12
FNSS_B_1d	3343	39.3	60.6	0.1	0.0	405.9	0.41	1.32	0.42	3.23	0.16
FNSS_S_1d	3355	38.4	61.4	0.2	0.0	400.8	0.40	1.34	0.4	3.25	0.13
S20N80_5d	3424	37.9	61.9	0.2	0.1	401.2	0.38	1.38	0.38	3.42	0.11
S40N60_5d	3058	37.7	62.1	0.2	0.0	388.2	0.37	1.4	0.37	3.43	0.12
S50N50_5d_1	3692	36.6	63.3	0.1	0.0	403.1	0.40	1.34	0.41	3.39	0.08
S50N50_5d_2	3831	43.4	56.4	0.2	0.1	401.0	0.43	1.28	0.42	2.66	0.08
S50N50_5d_3	2771	42.7	57.2	0.0	0.0	388.5	0.38	1.37	0.27	2.90	0.03
S60N40_5d	3238	34.8	63.6	1.5	0.1	393.0	0.35	1.43	0.29	3.11	1.40
S80N20_5d	2820	35.6	61.9	2.3	0.1	378.3	0.39	1.37	0.37	3.70	2.41

Footnote: 1) Strong decrease of counts: original > 1d, 5d; 2) Strong increase of %CHO, decrease of %CHNO and some decrease of N/C ratio from original to 1d, 5d; 3) Decrease of  $m/z$  from original to 1d, 5d; 4) Increase of S/C ratio and %CHOS.



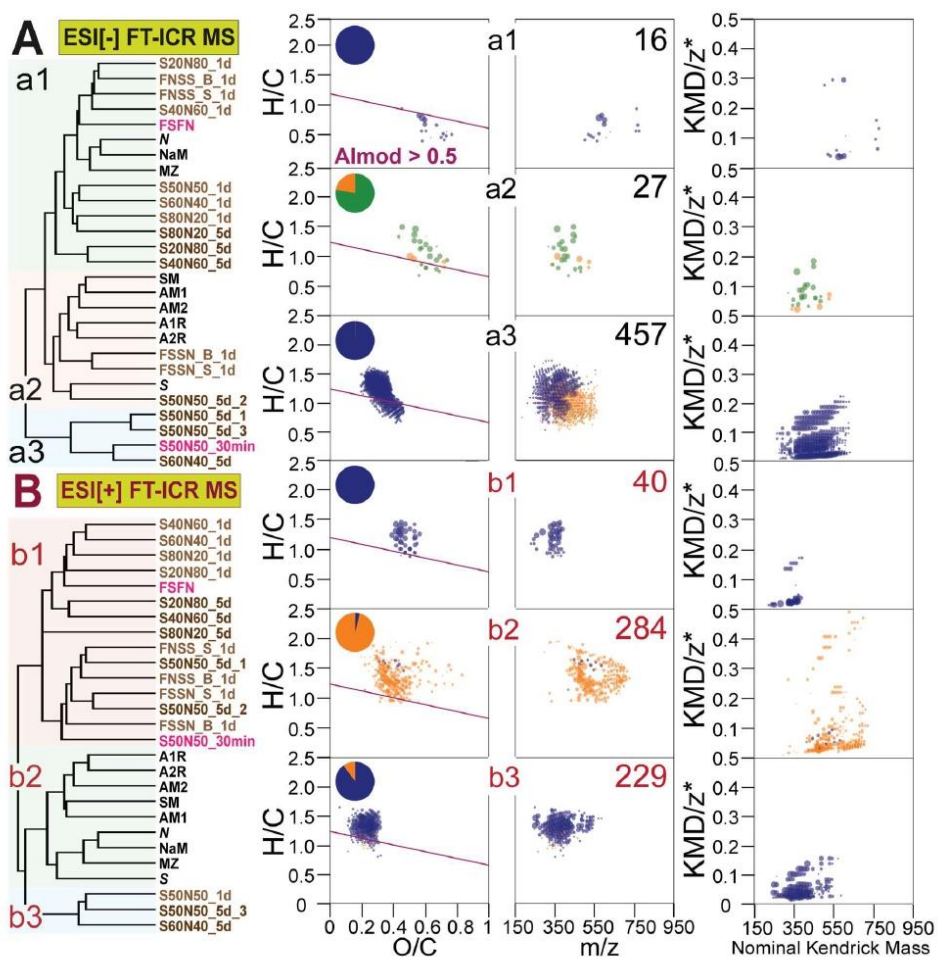


**Figure S2.** Loading vectors in PCA of SPE-DOM in unmixed and mixed S+N water pre- and post-incubation based on all assigned molecular formulae in (A, B) ESI[ $\pm$ ] FT-ICR mass spectra and (C)  $^1H$  NMR spectra. The scatter plots of the very PCA see Figure 2. Panels A and B show correlation patterns of PC1 and PC2 loading vectors (p1 and p2) in van Krevelen and mass-edited H/C diagrams, with the color code representing CHO (blue), CHNO (orange), and CHOS (green) molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have modified aromaticity index ( $AI_{mod}$ )<sup>15</sup> higher than 0.5. Panel C shows correlation patterns of p1 and p2 in  $^1H$  NMR space, with fundamental molecular structures indicated.

**Table S5.** Water parameters in unmixed and mixed S+N and A+T waters pre- and post- incubation.

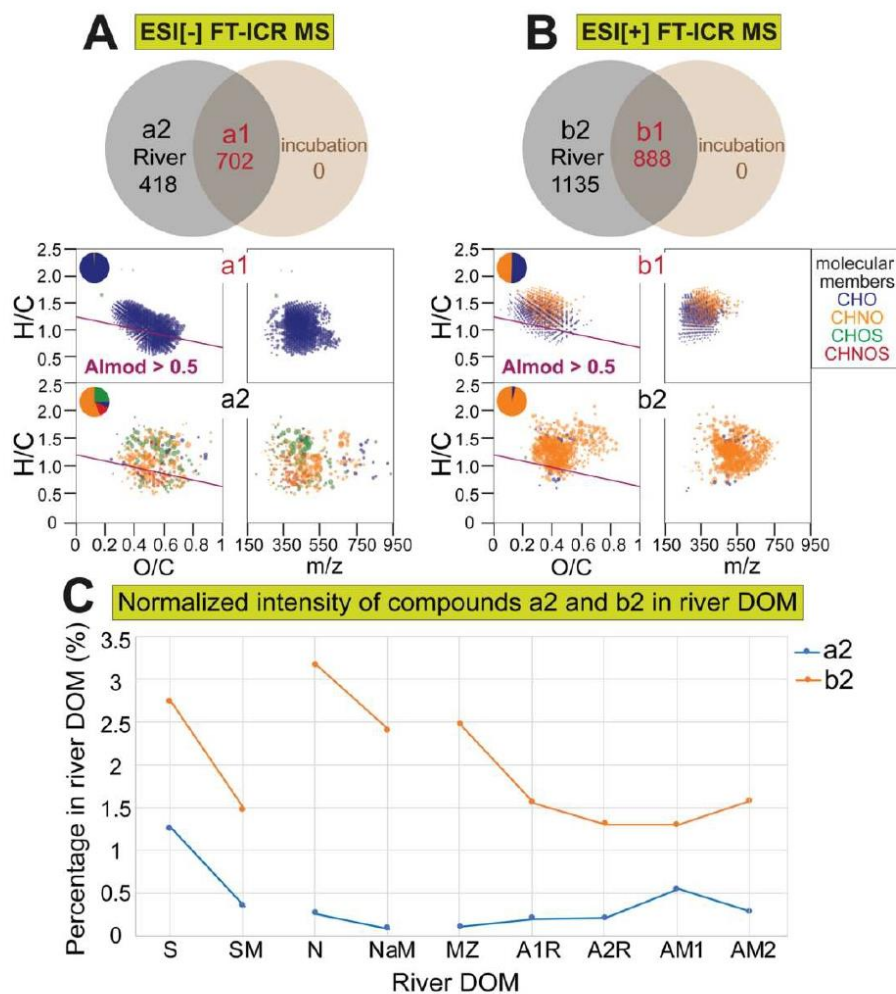
Sample	Temp. (°C)	pH	DOC (mg/L)	POC (mg/L)	TOC (mg/L)	Cond. (mg/L)	DO (mg/L)	NH <sub>4</sub> <sup>+</sup> (μM)	NO <sub>3</sub> <sup>-</sup> (μM)	NO <sub>2</sub> <sup>-</sup> (μM)	PO <sub>4</sub> <sup>3-</sup> (μM)
<i>S</i>	28.0	7.0	4.56	13.94	18.50	84	6.4	0.42	0.26	0.10	0.10
<i>N</i>	29.0	4.7	9.96	12.54	22.50	17	6.2	0.68	0.41	0.09	0.10
<i>A1</i>	28.8	6.4	5.36	1.44	6.80	85	4.5	0.53	0.94	0.08	0.24
<i>A2</i>	28.5	6.4	4.85	0.90	5.82	79	3.9	0.63	0.69	0.05	0.69
<i>Ad1</i>	27.5	6.5	4.16	0.23	4.39	56	5.0	0.88	0.59	0.15	0.22
<i>Ad2</i>	29.1	6.5	4.68	0.19	4.87	51	4.9	0.73	0.71	0.12	0.11
<i>A</i>	29.9	6.8	5.51	0.90	6.41	69	6.0	2.32	0.33	0.13	0.10
<i>At</i>	27.8	6.8	3.63	0.39	4.02	35	7.2	0.47	1.55	0.04	0.15
<i>T1</i>	26.9	5.7	3.11	0.23	3.34	37	5.8	0.42	0.57	0.10	0.12
<i>T2</i>	29.5	6.6	3.48	0.17	3.52	34	7.8	0.41	0.47	0.02	0.11
<i>T3</i>	29.9	6.8	5.51	0.85	6.41	61	6.0	2.32	0.33	0.13	0.10
<i>T</i>	29.2	6.6	7.03	0.17	7.20	55	4.7	2.36	0.71	0.12	0.14
S20N80 1d	27.2	6.79	N.A.	N.A.	N.A.	25.6	N.A.	0.59	0.52	0.08	0.11
S40N60 1d	28.0	7.04	N.A.	N.A.	N.A.	34.8	N.A.	0.62	0.49	0.10	0.13
S60N40 1d	28.1	7.25	N.A.	N.A.	N.A.	49.4	N.A.	0.56	0.44	0.10	0.11
S80N20 1d	N.A.	7.43	N.A.	N.A.	N.A.	60.6	N.A.	0.62	0.49	0.07	0.12
S20N80 5d	29.1	6.48	N.A.	N.A.	N.A.	25.2	N.A.	1.35	0.75	0.13	0.11
S40N60 5d	29.1	6.74	N.A.	N.A.	N.A.	40.9	N.A.	1.23	0.69	0.13	0.11
S60N40 5d	28.7	7.8	N.A.	N.A.	N.A.	57.6	N.A.	1.22	0.71	0.12	0.13
S80N20 5d	29.1	7.18	N.A.	N.A.	N.A.	68.9	N.A.	0.94	0.69	0.12	0.12





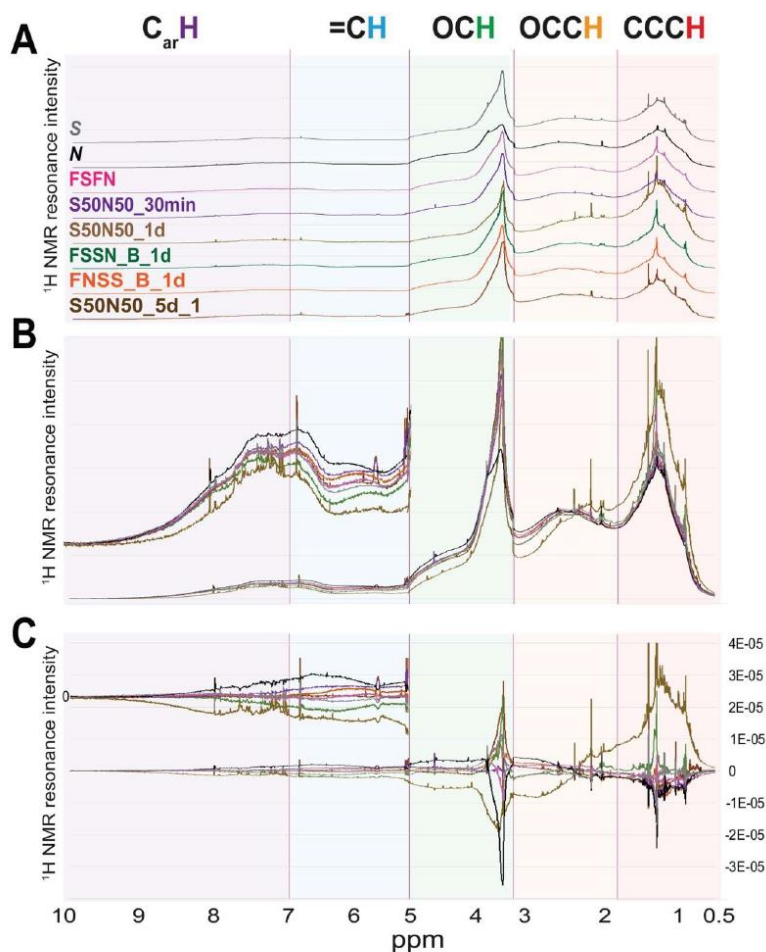
**Figure S3.** HCA of assigned SPE-DOM molecular formulae in part of the unmixed and mixed S+N water pre- and post- incubation. Van Krevelen, mass-edited H/C, and KMD/z\* diagrams<sup>16</sup> show CHO (blue), CHNO (orange), and CHOS (green) molecular classes that were relatively more abundant in clusters a1/a2/a3, b1/b2/b3, respectively. Bubble areas represent the normalized intensities of SPE-DOM in each cluster. The pie plots depict percentages of the counts of different molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have modified aromaticity index ( $AI_{mod}$ )<sup>15</sup> higher than 0.5.

Footnote: 1) ESI[±] MS-derived HCA separated samples in clusters a3 and b3 owing to their higher abundance in less oxygenated CHO compounds, in line with the PCA results (Figure 2AB); 2) ESI[+] MS-derived HCA separated river water samples (b2) and experimentally mixed water samples (b1,b3). SPE-DOM in river waters were more abundant in some CHNO ions ( $m/z$  450-700) than SPE-DOM in mixed waters.



**Figure S4.** Common and unique molecular signatures in SPE-DOM in the unmixed and mixed S+N water pre- and post- incubation. The numbers show counts of compounds. Van Krevelen and mass-edited H/C diagrams show CHO (blue), CHNO (orange), CHOS (green), and CHNOS (red) m/z ions that were shared before and after incubation (a1 and b1); unique in original water but were not in incubated water (a2 and b2). Not any single compound unique to incubated water did not present in original water. Bubble areas represent the normalized intensity of DOM. The pie plots depict percentages of the counts of different molecular classes. Molecular composition positioned below the purple line in the van Krevelen diagrams shows the compounds with a modified aromaticity index ( $AI_{mod}$ ) higher than 0.5<sup>15</sup>. Panel C shows the normalized intensity of molecular signatures that were degraded during the incubation process (a2 and b2) in DOM of the Amazon mainstem.

**Figure S4.** Footnote: 1) Two sets of overlapping CHO molecules with good coverage of the compositional space with (A) average H/C and O/C ratios ( $m/z \sim 350\sim 570$ ) and (B) more unsaturated and highly oxygenated lignin-like molecules ( $m/z \sim 350\sim 620$ ); 2) A set of more unsaturated, low mass ( $m/z < 380$ ) CHO molecules and another set of more saturated CHNO molecules with limited range of oxygenation and mass; 3) Interspersed set of CHNO, CHOS, CHNOS molecules with high chemodiversity covering large areas in van Krevelen diagrams; 4) Two sets of considerably unsaturated CHO molecules with distinct oxygenation.



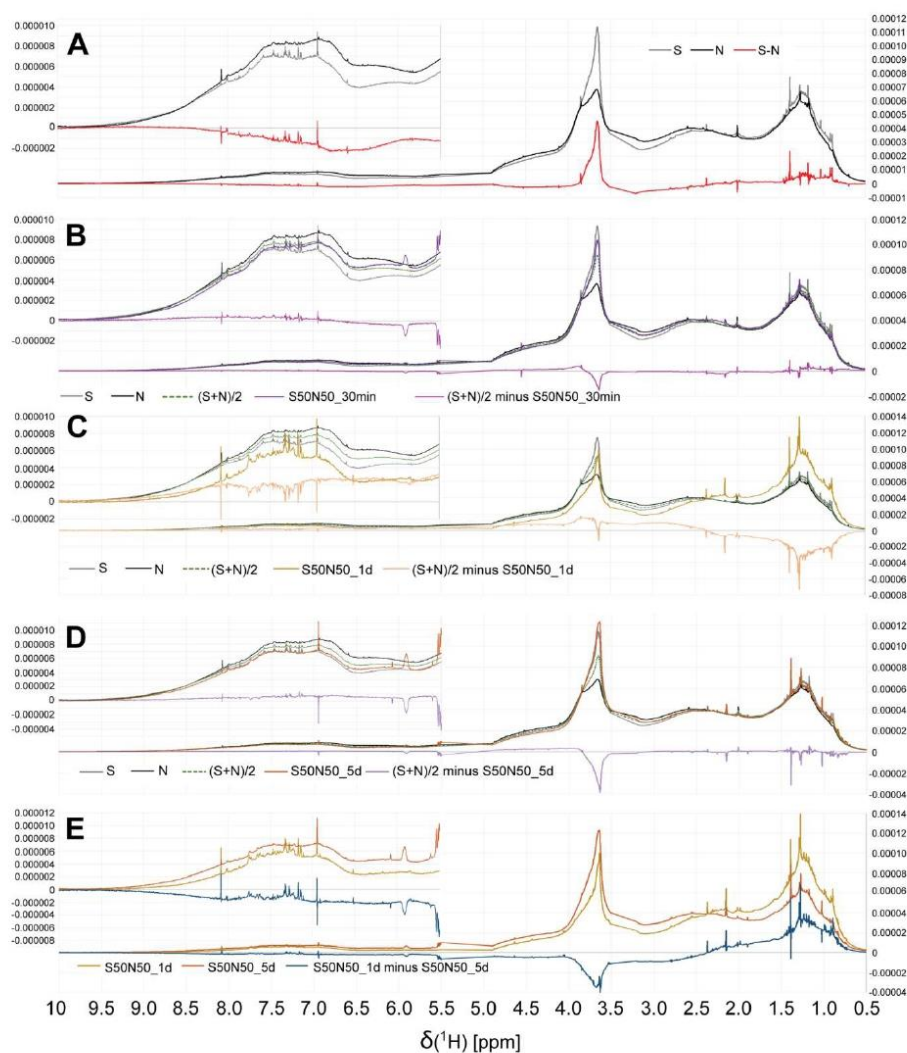
**Figure S5.**  $^1\text{H}$  NMR spectra of SPE-DOM in the unmixed and mixed S+N water pre- and post-incubation. The  $^1\text{H}$  NMR spectra are exemplified with samples *S*, *N*, FSFN, S50N50\_30min, S50N50\_1d, FSSN\_B\_1d, FNSS\_B\_1d, and S50N50\_5d\_1.  $^1\text{H}$  NMR spectra are area normalized (800 MHz,  $\text{CD}_3\text{OD}$ ;  $\delta\text{H} = 0.5\text{--}10.0$  ppm with the exclusion of residual water and methanol NMR resonances). The  $^1\text{H}$  NMR spectra referred to average abundance are edited by each normalized spectra subtract the average spectra (C).

**Table S6.** <sup>1</sup>H NMR section integrals (percent of non-exchangeable protons, 800 MHz; CD<sub>3</sub>OD, exclusion of residual water, and methanol) and key substructures of SPE-DOM in the unmixed and mixed S+N water pre- and post- incubation. The fundamental substructures include aromatics C<sub>ar</sub>H, δ<sub>H</sub> ~ 7.0-10.0 ppm; olefins δ<sub>H</sub> ~ 5.3-7.0 ppm; oxygenated aliphatic units (OCH) and “carbohydrate-like” and methoxy OCH<sub>3</sub> units δ<sub>H</sub> ~ 3.2-4.9 ppm; branched aliphatic units (CH<sub>2</sub>)<sub>n</sub>, “acetate-analogue” and CRAM δ<sub>H</sub> ~ 1.9-3.2 ppm; functionalized aliphatics δ<sub>H</sub> ~ 1.35-1.9 ppm; polyethylene group δ<sub>H</sub> ~ 1.25-1.35 ppm, OCCH units; pure aliphatics δ<sub>H</sub> ~ 0.5-1.25 ppm, CCCH units.

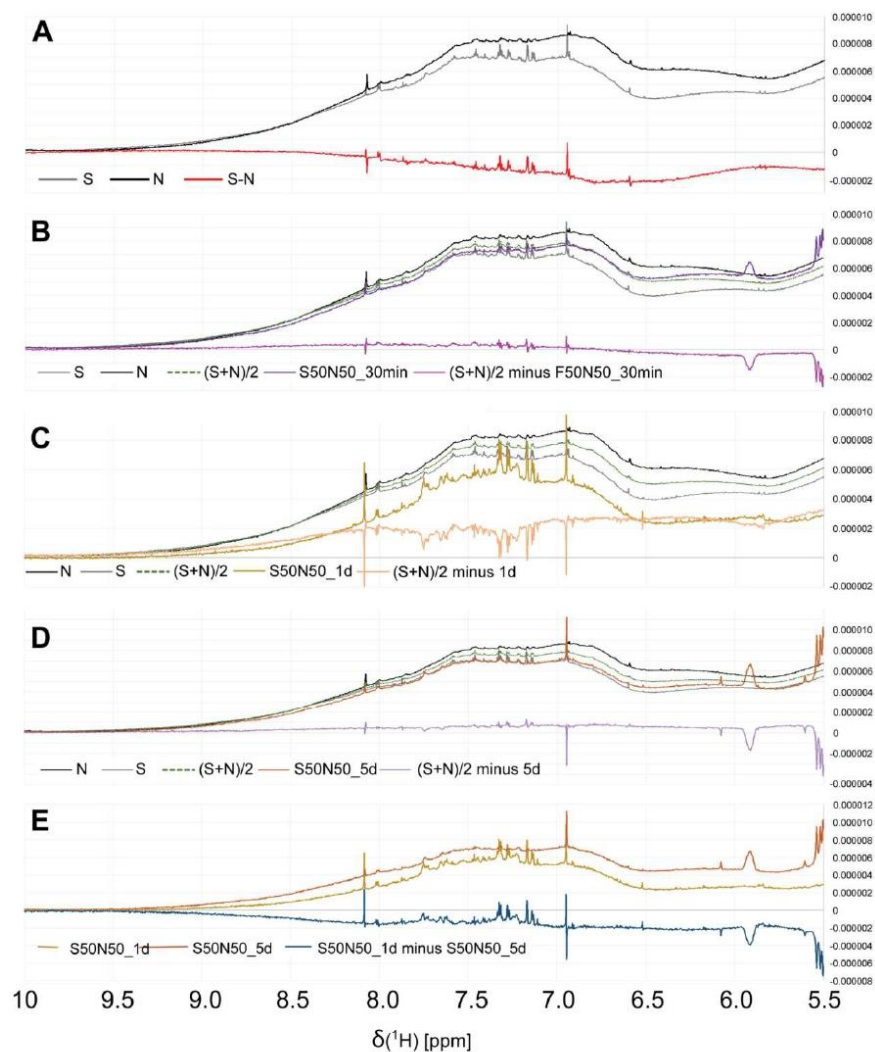
δ ( <sup>1</sup> H) [ppm]	10.0-7.0	7.0-5.3	4.9-3.2	3.2-1.9	1.9-1.35	1.35-1.25	1.25-0.5	1.9-0.5
key substructures	C <sub>ar</sub> H (%)	C=CH, O <sub>2</sub> CH (%) O <sub>2</sub> CH (%)	OCH (%)	OCCH (%)	OCCCH (%)	(CH <sub>2</sub> ) <sub>n</sub> (%)	CCCH <sub>3</sub> (%)	total aliphatic section (%)
S	5.8	4.6	30.3	28.0	14.5	4.0	12.8	31.3
SM	5.1	3.5	26.6	29.6	15.5	5.1	14.6	35.3
N	5.9	5.8	26.8	30.7	14.8	4.2	11.8	30.8
NaM	5.5	5.2	26.5	30.9	15.0	4.5	12.4	31.9
MZ	5.9	4.3	24.6	29.0	16.8	5.5	13.8	36.2
A1	4.6	4.1	25.1	31.7	16.3	4.9	13.3	34.5
A2	5.4	4.8	27.4	30.4	15.1	4.4	12.4	31.9
AM1	4.8	4.5	27.7	30.9	15.3	4.5	12.3	32.1
AM2	5.0	4.8	28.6	30.4	15.0	4.3	11.8	31.2
FSFN	5.6	4.8	28.6	30.0	14.4	4.2	12.3	31.0
S50N50 30min	5.7	5.7	29.2	29.7	14.4	3.9	11.5	29.8
S20N80 1d	4.8	2.7	24.8	28.1	14.3	9.0	16.2	39.6
S40N60 1d	4.7	2.9	28.9	29.5	14.4	6.3	13.2	34.0
S50N50 1d	4.0	2.9	21.7	26.9	18.7	6.0	19.7	44.5
S60N40 1d	4.2	2.3	27.0	28.0	14.3	9.3	14.9	38.5
S80N20 1d	4.7	1.1	27.4	26.1	14.3	9.9	16.5	40.7
FSSN B 1d	5.1	3.9	29.1	28.8	15.1	4.9	13.1	33.1
FSSN S 1d	5.0	3.6	30.2	28.4	14.7	4.6	13.5	32.8
FNSS B 1d	5.6	5.3	29.9	29.6	14.1	4.2	11.3	29.6
FNSS S 1d	5.7	5.4	30.9	29.5	13.7	4.0	10.9	28.6
S20N80 5d	4.5	3.7	26.7	27.9	15.0	6.6	15.6	37.2
S40N60 5d	4.3	3.6	26.9	29.8	15.3	5.7	14.5	35.4
S50N50 5d 1	5.5	5.0	30.3	28.8	14.0	4.0	12.5	30.4
S50N50 5d 2	6.4	5.9	29.0	29.1	14.1	3.8	11.7	29.6
S50N50 5d 3	6.9	2.2	17.2	25.9	17.8	6.9	23.1	47.8
S60N40 5d	3.9	1.9	15.0	25.5	19.0	12.7	22.0	53.7
S80N20 5d	3.9	2.2	28.8	28.0	14.7	7.0	15.5	37.1

Footnote: 1) Nearly unchanged C<sub>ar</sub>H; 2) Considerable decrease of olefins =CH; 3) uneven evolution in general.



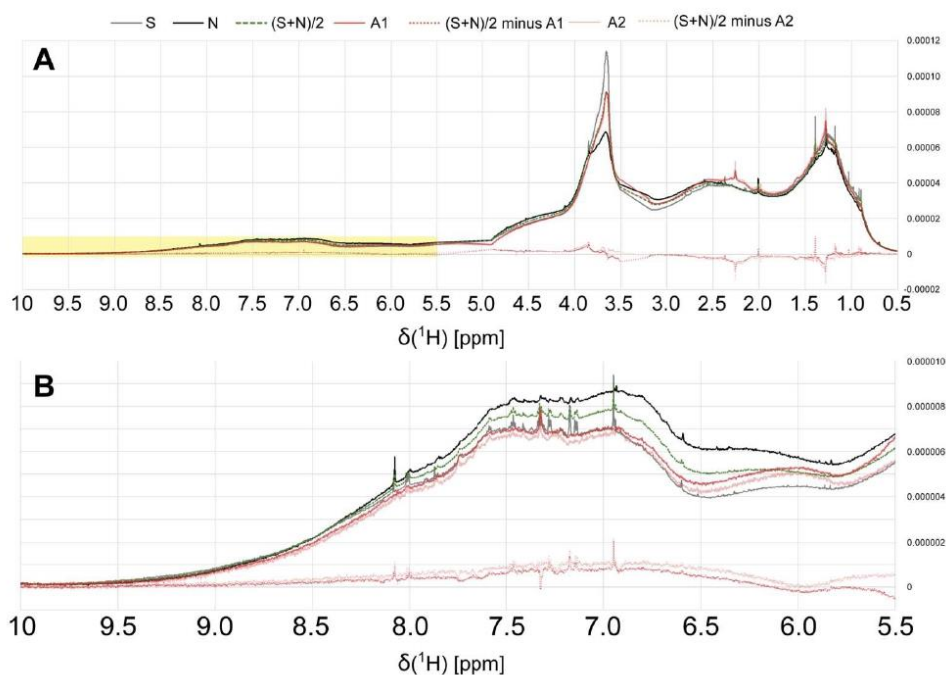


**Figure S6.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Solimões and Negro SPE-DOM according to status of incubation. (A) Solimões minus Negro, with Solimões (S), Negro (N), and Solimões minus Negro (S-N). (B) difference  $^1\text{H}$  NMR spectrum after 30 minutes of incubation, with S, N, averaged Solimões and Negro ( $(\text{S}+\text{N})/2$ ), S50N50\_30min, and  $(\text{S}+\text{N})/2$  minus S50N50\_30min. (C) difference  $^1\text{H}$  NMR spectrum after one day of incubation, with S, N,  $(\text{S}+\text{N})/2$ , S50N50\_1d, and  $(\text{S}+\text{N})/2$  minus S50N50\_1d. (D) difference  $^1\text{H}$  NMR spectrum after five days of incubation, with S, N,  $(\text{S}+\text{N})/2$ , S50N50\_5d, and  $(\text{S}+\text{N})/2$  minus S50N50\_5d. (E) difference  $^1\text{H}$  NMR spectrum of one day minus five days of incubation, with S50N50\_1d, S50N50\_5d, and  $(\text{S50N50\_1d minus S50N50\_5d})$ .

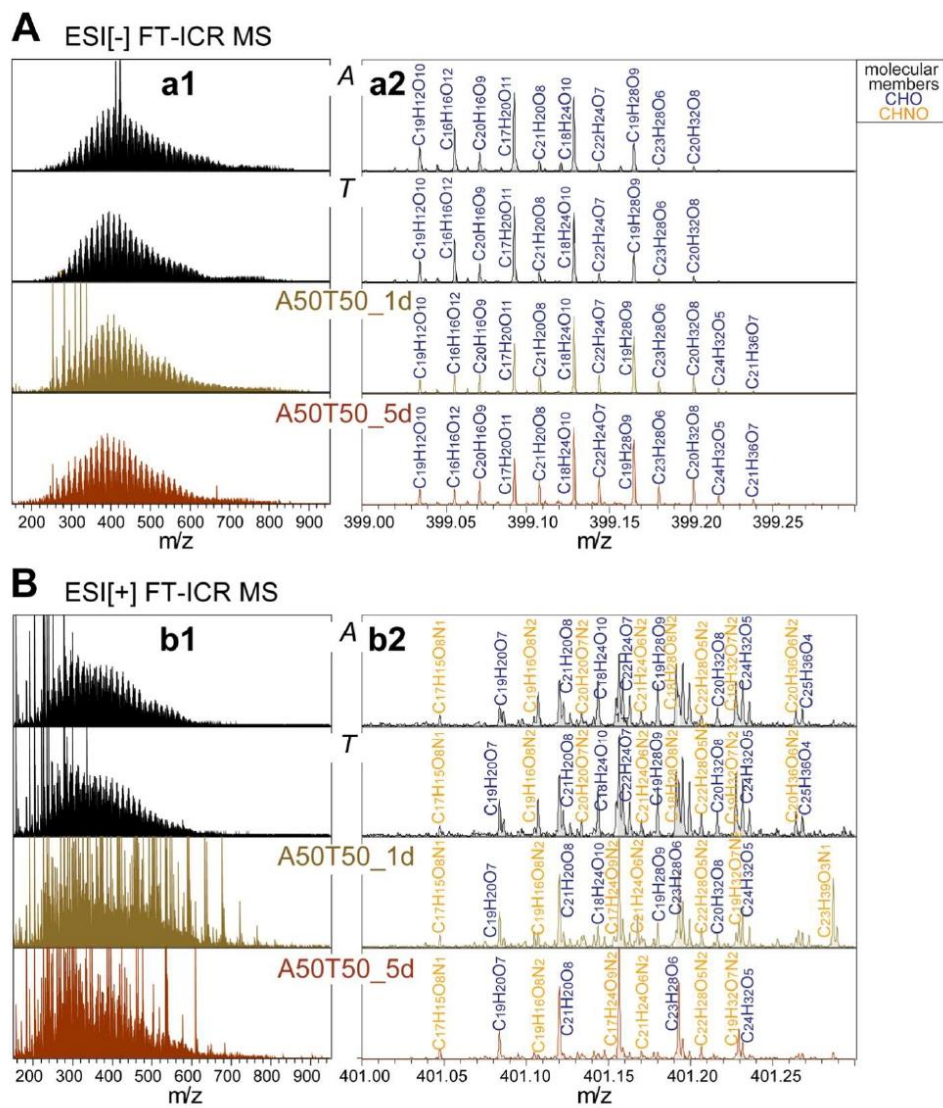


**Figure S7.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Solimões and Negro SPE-DOM according to status of incubation, section of unsaturated  $\text{C}_{\text{sp}^2}\text{H}$  protons  $\delta_{\text{H}} \sim 5.5\text{-}10$  ppm in Figure S6. (A) Solimões minus Negro, with Solimões (S), Negro (N), and Solimões minus Negro (S-N). (B) difference after 30 minutes of incubation, with S, N, averaged Solimões and Negro ((S+N)/2), S50N50\_30min, and (S+N)/2 minus S50N50\_30min. (C) difference after one day of incubation, with S, N, (S+N)/2, S50N50\_1d, and (S+N)/2 minus S50N50\_1d. (D) difference  $^1\text{H}$  NMR spectrum after five days of incubation, with S, N, (S+N)/2, S50N50\_5d, and (S+N)/2 minus S50N50\_5d. (E) difference  $^1\text{H}$  NMR spectrum one day minus five days of incubation, with S50N50\_1d, S50N50\_5d, and (S50N50\_1d minus S50N50\_5d).





**Figure S8.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Solimões and Negro river waters, entire section  $\delta_{\text{H}} \sim 0.5\text{-}10$  ppm, and proximate Amazon SPE-DOM samples A1 and A2 (sampling locations see Figure 1 and Table S1). (A) (averaged Solimões and Negro) river minus proximate downstream Amazon River sampling points A1 (orange, dotted line) and A2 (pink, dotted line), with Solimões (S, grey), Negro (N, black), and averaged Solimões and Negro ((S+N)/2, dotted green line). (B) same spectra, section of unsaturated  $\text{C}_{\text{sp}2}\text{H}$  protons ( $\delta_{\text{H}} \sim 5.5\text{-}10$  ppm).



**Figure S9.** (A) ESI[-] and (B) ESI[+] FT-ICR mass spectra of SPE-DOM in unmixed and mixed A+T water pre- and post- incubation. The FT-ICR mass spectra are exemplified with samples *A*, *T*, A50T50\_1d, and A50T50\_5d. The ESI[-] and ESI[+] FT-ICR mass spectra in (a1) and (b1) show the regular shaped signal distribution over a wide mass range ( $m/z$  150-950), as well as distinct signatures at nominal neutral mass 400. Respective assignments of molecular compositions are provided for CHO (blue) and CHNO (orange) molecules.

**Table S7.** ESI[-] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in unmixed and mixed A+T water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental  $m/z$ , DBE/C, H/C, N/C, and S/C ratios.

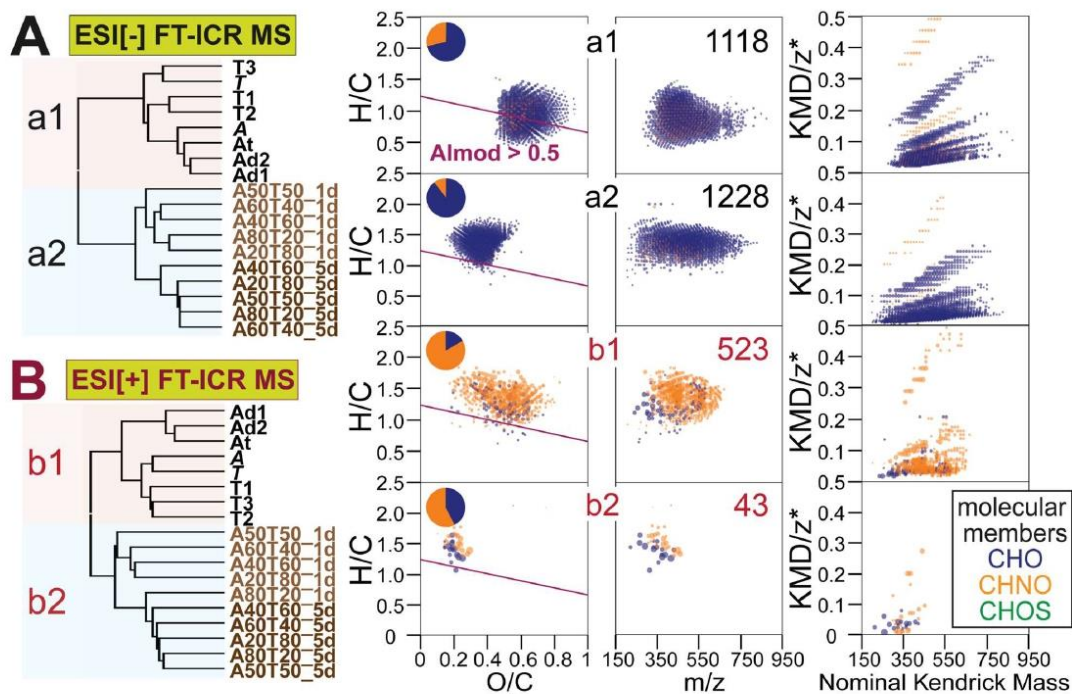
Sample	Total counts of mass peaks	CHO%	CHNO%	CHOS%	CHNOS%	$m/z$	DBE/C	H/C	O/C	N/C $\times 10^{-2}$	S/C $\times 10^{-3}$
Ad1	4473	64.1	30.2	5.6	0.2	478.5	0.52	1.07	0.55	0.53	0.91
Ad2	4515	64.2	30.7	4.9	0.2	476.1	0.52	1.07	0.54	0.54	0.81
A	4787	65.0	30.2	4.7	0.0	480.2	0.52	1.06	0.54	0.55	0.54
At	3870	68.2	29.5	2.3	0.0	479.8	0.52	1.06	0.55	0.58	0.39
T1	5276	62.4	33.8	3.8	0.0	472.1	0.52	1.07	0.55	0.68	0.83
T2	5858	60.9	35.3	3.8	0.1	466.3	0.53	1.05	0.54	0.79	0.99
T3	5813	58.7	32.1	6.9	2.2	446.6	0.52	1.06	0.55	0.71	1.21
T	5898	57.9	34.4	6.5	1.2	462.4	0.50	1.11	0.51	0.76	0.94
A20T80 1d	5713	66.2	30.8	3.1	0.0	478.2	0.45	1.20	0.46	0.61	0.78
A40T60 1d	5186	67.6	29.9	2.5	0.0	469.6	0.46	1.19	0.46	0.58	0.52
A50T50 1d	6210	63.7	32.4	3.9	0.0	475.7	0.44	1.22	0.45	0.63	1.77
A60T40 1d	5072	68.3	28.9	2.7	0.0	463.3	0.46	1.19	0.46	0.56	0.58
A80T20 1d	5837	65.2	31.5	3.3	0.0	480.0	0.45	1.19	0.46	0.61	0.71
A20T80 5d	4730	68.6	28.5	2.9	0.0	455.0	0.45	1.20	0.45	0.58	0.65
A40T60 5d	3694	72.4	24.6	3.0	0.0	427.6	0.43	1.24	0.44	0.67	3.01
A50T50 5d	4773	68.8	28.1	3.2	0.0	453.7	0.44	1.22	0.44	0.56	0.69
A60T40 5d	4690	68.3	28.3	3.4	0.0	451.7	0.45	1.21	0.45	0.58	0.85
A80T20 5d	4699	69.5	27.2	3.3	0.0	450.1	0.44	1.22	0.44	0.55	1.01

Footnote: 1) No clear trend in count of mass peaks and %element ratios when changing mixing ratio; 2)  $m/z$  overall decreases from 1d to 5d incubation; 3) Overall small gain in saturation as expressed by decrease in DBE/C and increase in H/C ratios during 5d incubation; 4) Considerable change from original to 1d and very minor loss of O/C ratio, most probably because of loss of oxygen-rich thermodynamic endmember molecules.

**Table S8.** ESI[+] FT-ICR MS derived counts of mass peaks and intensity-weighted average bulk parameters for all assigned molecular compositions present in unmixed and mixed A+T water pre- and post- incubation. FT-ICR MS derived bulk parameters comprise percentages of counts of CHO, CHNO, CHOS, and CHNOS molecular classes, computed experimental  $m/z$ , DBE/C, H/C, N/C, and S/C ratios.

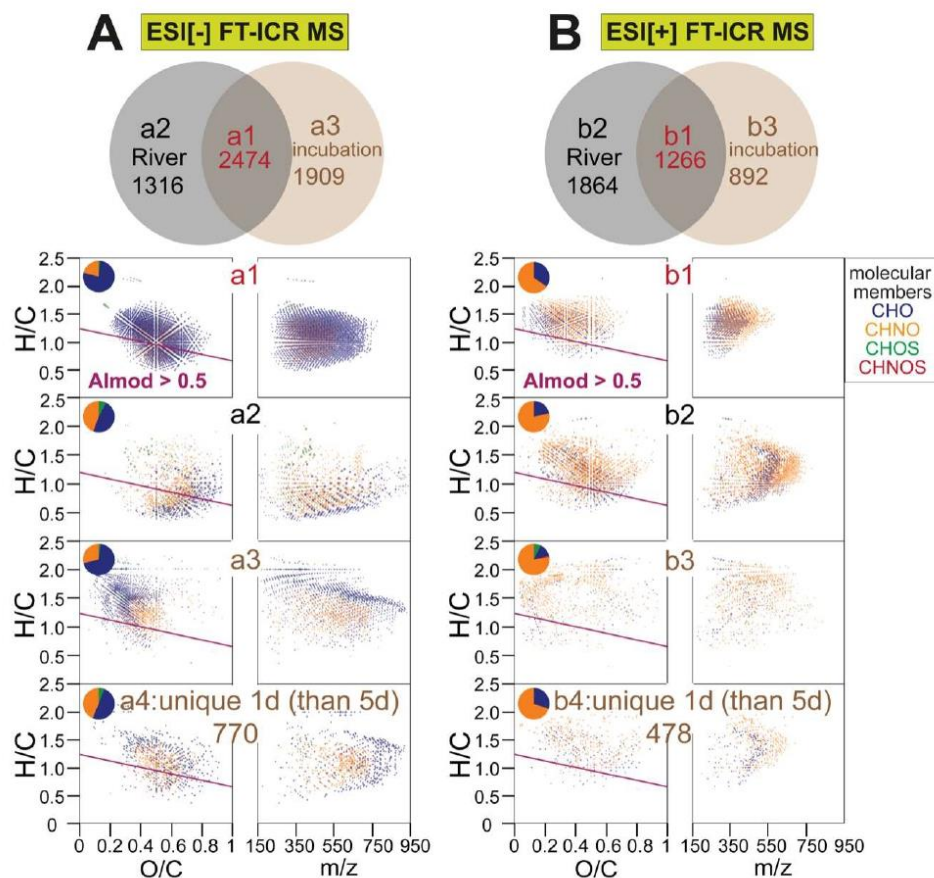
Sample	Total counts of mass peaks	CHO%	CHNO%	CHOS%	CHNOS%	$m/z$	DBE/C	H/C	O/C	N/C $\times 10^{-2}$	S/C $\times 10^{-3}$
Ad1	6060	30.4	67.6	0.5	1.5	404.8	0.39	1.36	0.40	3.71	0.17
Ad2	5845	30.1	68.3	0.5	1.1	415.4	0.39	1.38	0.36	3.72	0.14
A	3943	32.0	68.0	0.0	0.1	395.0	0.39	1.37	0.38	3.95	0.02
At	5809	30.2	69.4	0.1	0.3	410.7	0.40	1.34	0.40	3.87	0.07
T1	4709	30.5	69.3	0.1	0.1	397.6	0.41	1.33	0.41	4.05	0.04
T2	4365	29.6	70.3	0.1	0.0	390.0	0.40	1.35	0.39	4.07	0.03
T3	4189	29.8	69.8	0.2	0.2	389.4	0.41	1.33	0.40	4.08	0.06
T	3564	32.3	67.5	0.1	0.1	386.7	0.39	1.36	0.39	3.91	0.04
A20T80 1d	3382	32.3	67.1	0.5	0.1	380.9	0.37	1.40	0.37	3.69	0.00
A40T60 1d	4070	34.0	65.5	0.4	0.1	387.6	0.39	1.37	0.37	3.66	0.00
A50T50 1d	3580	28.2	71.3	0.5	0.1	422.0	0.32	1.50	0.33	3.63	0.00
A60T40 1d	3131	35.6	63.9	0.4	0.1	376.5	0.38	1.39	0.37	3.60	0.01
A80T20 1d	2698	35.2	63.6	0.9	0.2	375.1	0.38	1.39	0.38	3.87	0.02
A20T80 5d	2574	33.0	65.8	1.0	0.2	362.6	0.38	1.39	0.35	4.24	0.02
A40T60 5d	2319	27.3	71.6	1.0	0.2	376.3	0.40	1.36	0.36	5.12	0.03
A50T50 5d	2578	31.3	67.7	0.9	0.2	364.0	0.39	1.38	0.34	4.12	0.01
A60T40 5d	2927	31.4	67.9	0.7	0.1	362.7	0.39	1.38	0.34	3.95	0.01
A80T20 5d	2015	31.5	67.1	1.4	0.1	357.6	0.38	1.40	0.35	4.55	0.03

Footnote: 1) Counts of assigned mass peaks and average  $m/z$  decrease during incubation (original > 1d > 5d); 2) %CHO increases from original water to 1d, and somewhat decreases from 1d to 5d; 3) While %CHNO remains near constant, %CHOS goes up (but S/C ratio goes down); 4) Average  $m/z$  decline from original to 1d to 5d; 5) Small initial decrease of DBE/C followed by "tiny rebound" from 1d to 5d; Similarly, H/C ratio increases "a little bit" from original to 1d; 6) Small deoxygenation (O/C ratio original > 1d > 5d) may result from loss of oxygen-rich thermodynamic endmember molecules H<sub>2</sub>O and CO<sub>2</sub> upon oxidative degradation of DOM; 7) Small increase of %CHNO and N/C ratio at 5d may reflect higher longevity of CHNO molecules.



**Figure S10.** HCA of assigned SPE-DOM molecular formulae in the unmixed and mixed A+T water pre- and post-incubation. Van Krevelen, mass-edited H/C, and KMD/z\* diagrams<sup>16</sup> show CHO (blue), CHNO (orange), and CHOS (green) molecular classes that were relatively more abundant in clusters a1/a2, b1/b2, respectively. The pie plots depict percentages of the counts of different molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have modified aromaticity index ( $AI_{mod}$ )<sup>15</sup> higher than 0.5.

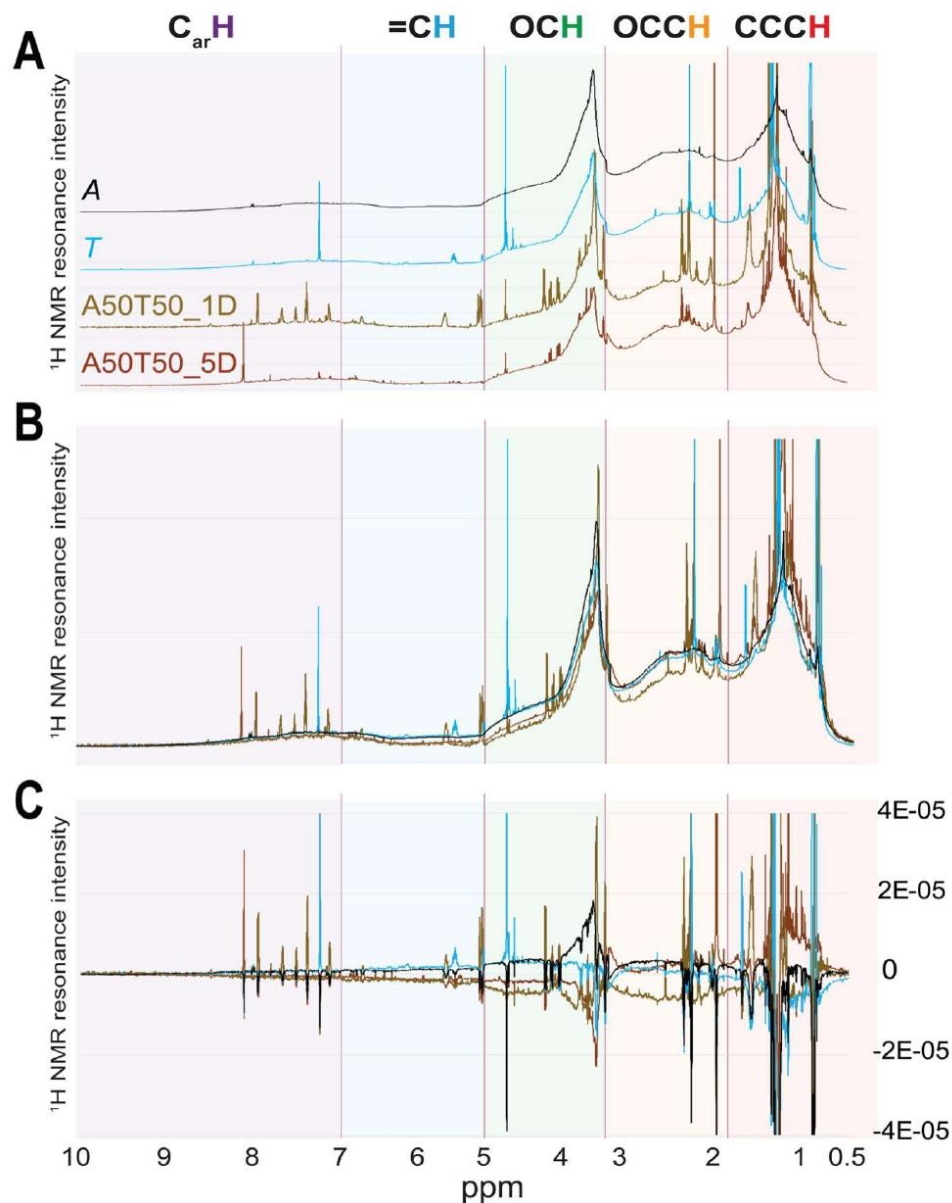




**Figure S11.** Common and unique molecular signatures in unmixed and mixed A+T water pre- and post-incubation. The numbers show counts of compounds. Van Krevelen and mass-edited H/C diagrams show CHO (blue), CHNO (orange), CHOS (green), and CHNOS (red)  $m/z$  ions that were shared before and after incubation (a1 and b1); unique in original water but were not in incubated water (a2 and b2); unique in incubated water but were not in original water (a3 and b3); unique in one-day incubated water (1d) but were not present in five-day incubated water (5d) (a4 and b4). There was no compound unique to 5d and not present 1d. The pie plots depict percentages of the counts of different molecular classes. The numbers show counts of compounds. The molecular formulae positioned below the purple line in the van Krevelen diagrams have modified aromaticity index ( $AI_{mod}$ )<sup>15</sup> higher than 0.5.

Footnote: (a1, b1) ESI[-] MS denotes more CHO than CHNO compounds of near average H/C and O/C ratio for which number of isomers for given molecular composition in maximal island of stability<sup>17</sup>, chemodiversity of observed CHO molecules is larger than that of CHNO compounds as shown by coverage in van Krevelen and mass-edited H/C diagrams. ESI[+] MS denotes more CHNO than CHO compounds and displacement of CHNO molecules toward higher H/C ratio and higher mass. (a2, b2) ESI[-] MS denotes loss or certain hydrogen-deficient polyphenols, and lower mass CHOS compounds from original waters; ESI[+] MS denotes a surprising large-scale loss of CHNO molecules at near average H/C and O/C ratios. (a3, b3) Synthesis of several long-range series of CHO molecules near H/C ratio ~2, and a higher mass ( $m/z \sim 500-950$ ) group of less oxygenated (O/C ratio < 0.45) molecules.



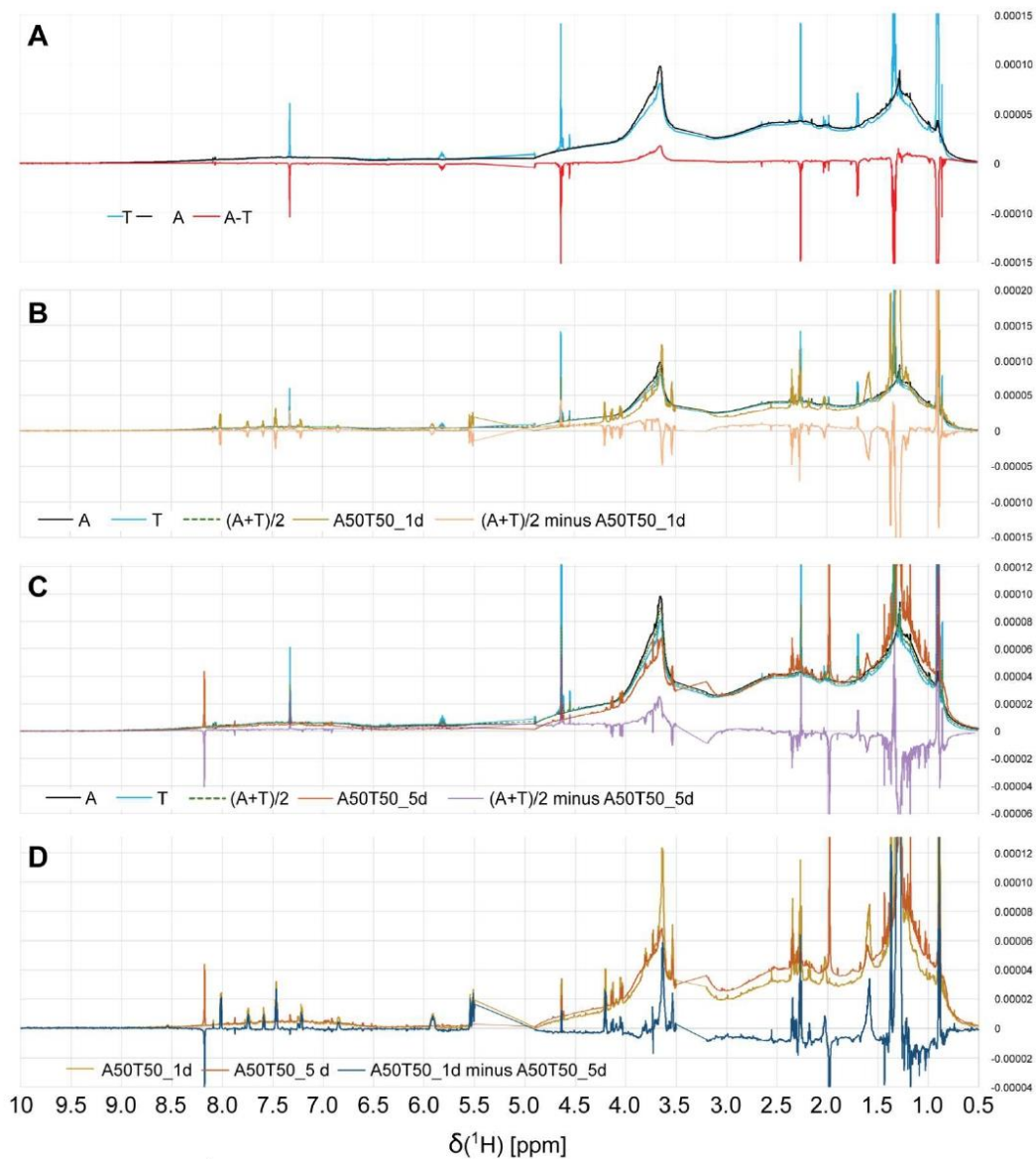


**Figure S12.**  $^1\text{H}$  NMR spectra of SPE-DOM in unmixed and mixed A+T water pre- and post- incubation. The  $^1\text{H}$  NMR spectra are exemplified with sample *A*, *T*, A50T50\_1d, and A50T50\_5d.  $^1\text{H}$  NMR spectra are entire region normalized (800 MHz,  $\text{CD}_3\text{OD}$ ;  $\delta_{\text{H}} = 0.5\text{-}10.0$  ppm with exclusion of residual water and methanol NMR resonances). The  $^1\text{H}$  NMR spectra referred to average abundance in the four DOM are edited by each normalized spectra subtract the average spectra of the four DOM (C).

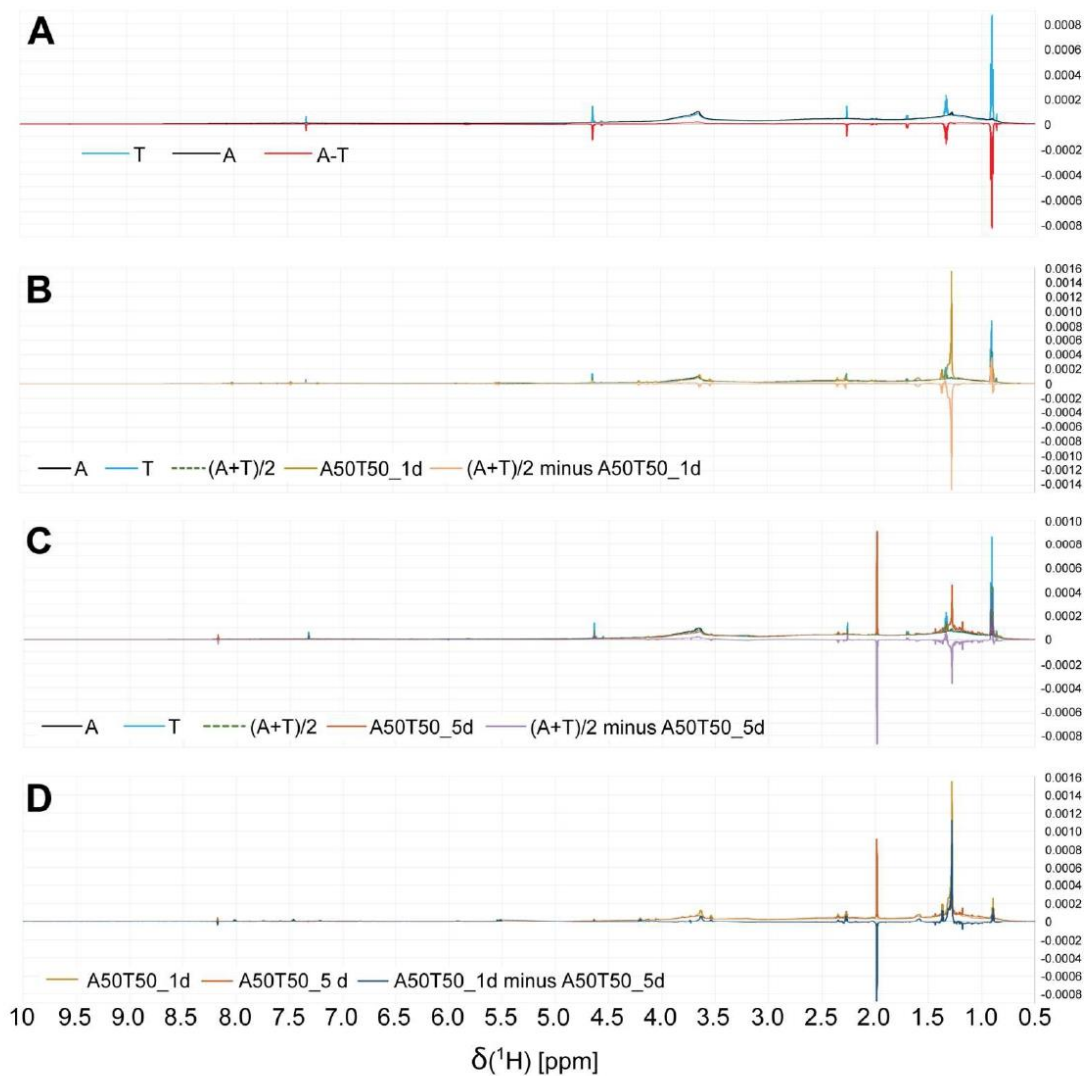
**Table S9.**  $^1\text{H}$  NMR section integrals (percent of non-exchangeable protons, 800 MHz;  $\text{CD}_3\text{OD}$ , exclusion of residual water, and methanol)) and key substructures of SPE-DOM in unmixed and mixed A+T water pre- and post- incubation. The fundamental substructures include aromatics  $\text{C}_{\text{ar}}\text{H}$ ,  $\delta_{\text{H}} \sim 7.0\text{-}10.0$  ppm; olefins  $\delta_{\text{H}} \sim 5.3\text{-}7.0$  ppm; oxygenated aliphatic units ( $\text{OCH}$ ) and “carbohydrate-like” and methoxy  $\text{OCH}_3$  units  $\delta_{\text{H}} \sim 3.2\text{-}4.9$  ppm; branched aliphatic units  $(\text{CH}_2)_n$ , “acetate-analogue” and CRAM  $\delta_{\text{H}} \sim 1.9\text{-}3.2$  ppm; functionalized aliphatics  $\delta_{\text{H}} \sim 1.35\text{-}1.9$  ppm; polyethylene group  $\delta_{\text{H}} \sim 1.25\text{-}1.35$  ppm,  $\text{OCCCH}$  units; pure aliphatics  $\delta_{\text{H}} \sim 0.5\text{-}1.25$  ppm,  $\text{CCCH}$  units.

$\delta$ ( $^1\text{H}$ ) [ppm]	10.0-7.0	7.0 5.3	4.9-3.2	3.2-1.9	1.9-1.35	1.35-1.25	1.25-0.5	1.9-0.5
key substructures	$\text{C}_{\text{ar}}\text{H}$ (%)	$\text{C}=\text{CH}$ , $\text{O}_2\text{CCH}$ (%) $\text{O}_2\text{CCH}$ (%)	$\text{OCH}$ (%)	$\text{OCCCH}$ (%)	$\text{OCCCH}$ (%)	$(\text{CH}_2)_n$ (%)	$\text{CCCH}_3$ (%)	total aliphatic section (%)
Ad1	4.8	3.4	31.0	26.6	14.8	4.7	14.7	34.2
Ad2	4.9	5.1	32.8	26.5	14.6	4.3	11.9	30.7
A	4.8	4.6	30.0	27.4	15.1	4.6	13.4	33.1
At	4.9	3.8	32.5	25.8	14.5	4.6	13.8	33.0
T1R	4.9	5.2	32.8	26.3	14.2	4.7	11.9	30.8
T2R	4.9	4.0	29.7	26.5	14.8	5.1	14.8	34.8
T3R	4.8	4.7	28.9	28.1	15.2	5.0	13.3	33.5
T	4.8	6.7	27.3	26.0	14.2	5.4	15.5	35.1
A20T80_1d	3.4	1.2	24.5	27.9	15.5	12.6	14.9	43.0
A40T60_1d	4.2	2.1	23.6	28.2	15.9	9.3	16.8	42.0
A50T50_1d	4.3	2.1	20.0	25.0	16.2	17.8	14.7	48.6
A60T40_1d	4.3	2.0	24.8	29.3	16.1	8.0	15.4	39.5
A80T20_1d	4.4	3.1	25.0	28.9	15.9	8.8	13.9	38.6
A20T80_5d	3.5	1.7	21.6	30.9	17.1	8.3	16.8	42.2
A40T60_5d	3.6	1.6	21.7	30.3	16.7	6.8	19.2	42.7
A50T50_5d	3.7	2.0	19.9	31.9	16.8	8.3	17.4	42.5
A60T40_5d	4.1	2.2	22.1	29.7	17.0	7.5	17.4	41.8
A80T20_5d	3.5	1.4	22.3	29.3	16.9	8.1	18.5	43.5

Footnote: 1)  $\text{C}_{\text{ar}}\text{H}$  decrease original > 1d > 5d; 2) Olefins strongly decline original > 1d, 5d; 3)  $\text{OCH}$  units decline original > 1d > 5d; 4) Remotely functionalized aliphatic units and carboxylic acids increase original < 1d < 5d; 5)  $(\text{CH}_2)_n$  strongly increases from original to 1d; then somewhat declines; 6) Continual increase of purely aliphatic units including alkyl; 7) strong increase original < 1d < 5d.



**Figure S13.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Amazon and Tapajós river SPE-DOM according to status of incubation (full vertical expansion cf. Figure S14). (A) Amazon minus Tapajós river, with Amazon (A), Tapajós (T), and Amazon minus Tapajós (A-T). (B) difference  $^1\text{H}$  NMR spectrum after one day of incubation, with A, T, averaged Amazon and Tapajós ( $(A+T)/2$ ), A50T50\_1d, and  $(A+T)/2$  minus A50T50\_1d. (C) difference  $^1\text{H}$  NMR spectrum after five days of incubation, with A, T,  $(A+T)/2$ , A50T50\_5d, and  $(A+T)/2$  minus A50T50\_5d. (D) difference  $^1\text{H}$  NMR spectrum one day minus five days of incubation, with A50T50\_1d, A50T50\_5d, and  $(A50T50_1d \text{ minus } A50T50_5d)$ .



**Figure S14.** Difference  $^1\text{H}$  NMR spectra (800 MHz,  $\text{CD}_3\text{OD}$ ) of Amazon and Tapajós river SPE-DOM according to status of incubation. (A) Amazon minus Tapajós river, with Amazon (A), Tapajós (T), and Amazon minus Tapajós (A-T). (B) difference  $^1\text{H}$  NMR spectrum after one day of incubation, with A, T, averaged Amazon and Tapajós ((A+T)/2), A50T50\_1d, and (A+T)/2 minus A50T50\_1d. (C) difference  $^1\text{H}$  NMR spectrum after five days of incubation, with A, T, (A+T)/2, A50T50\_5d, and (A+T)/2 minus A50T50\_5d. (D) difference  $^1\text{H}$  NMR spectrum one day minus five days of incubation, with A50T50\_1d, A50T50\_5d, and (A50T50\_1d minus A50T50\_5d).

**Table S10.** HBP and DCF in unmixed and mixed S+N and A+T waters pre- and post- incubation. SD, Standard Deviation; n, number of replicates; N.A., not available.

Sample	HBP (mean) $\mu\text{gC L}^{-1}\text{h}^{-1}$	HBP (SD) $\mu\text{gC L}^{-1}\text{h}^{-1}$	HBP (n)	DCF (mean) $\mu\text{gC L}^{-1}\text{h}^{-1}$	DCF (SD) $\mu\text{gC L}^{-1}\text{h}^{-1}$	DCF (n)
<i>S</i>	0.0317	0.0049	5	0.7594	0.0929	5
<i>N</i>	0.0284	0.0028	5	1.6634	0.6510	5
S+N mixing zone	0.0274	0.0042	6	6.0521	0.7198	5
S50N50_1h	0.0261	0.0027	5	N.A.	N.A.	N.A.
A1	3.5606	0.7900	5	13.2677	1.5780	5
A2	17.2304	2.3757	5	0.6946	0.8074	5
<i>A</i>	0.1535	0.0212	4	0.4730	0.5498	3
<i>T</i>	0.1492	0.0075	4	0.4395	0.2414	3
A+T mixing zone	0.1535	0.0065	4	1.7905	1.1563	3

**Table S11.** Ranges of heterotrophic bacterial production (HBP) and dark carbon fixation (DCF) in this study and in the Amazon River other aquatic systems.

Ecosystem	HBP ( $\mu\text{gC L}^{-1}\text{h}^{-1}$ )	HBP (Ref.)	DCF ( $\mu\text{gC L}^{-1}\text{h}^{-1}$ )	DCF (Ref.)
Amazon river	<0.001 to 27.2	This Study	<0.001 to 16.0	This Study
Amazon river	0.1 to 7.0	Benner et al., 1995a <sup>18</sup> ; Farjalla, 2014 <sup>11</sup> ; Farjalla et al., 2002 <sup>19</sup> ; Vidal et al., 2015 <sup>20</sup> ; Wissmar et al., 1981 <sup>21</sup>	NA	
Rodrigo de Freitas Lagoon (Brazil)	0.7 to 47.5	Signori et al., 2020 <sup>22</sup>	0.6 to 4.3	Signori et al., 2020 <sup>22</sup>
Guanabara Bay (Brazil)	2.20 to 122.61	Signori et al., 2018 <sup>23</sup>	0.001 to 2.5	Signori et al., 2018 <sup>23</sup>
Boreal lakes (Sweden)	<0.001 to 1.1	Bastviken et al. 2003 <sup>24</sup>	<0.001 to 0.3	Bastviken et al. 2003 <sup>24</sup>
Scheldt estuary (Belgium)	0.2 to 25	Goosen et al., 1997 <sup>25</sup>	0.001 to 0.2	Andersson et al., 2006 <sup>30</sup>
Ebro River estuary (Spain)	0.7 to 5.6	Calderón-Paz et al., 1993 <sup>26</sup>	0.04 to 0.7	Casamayor et al., 2001 <sup>31</sup>
Columbia River estuary (USA)	0.13 to 4.5	Crump et al., 1998 <sup>27</sup>	<0.001 to 2	Brauer et al., 2013 <sup>32</sup>
Karstic Lakes Spain	<0.001 to 8.75	Overman et al., 1996 <sup>28</sup>	0.05 to 41	Casamayor et al., 2001 <sup>31</sup> ; Noguerola et al. 2015 <sup>33</sup>
Lake Cadagno	0.002 to 0.006	Saini et al., 2022 <sup>29</sup>	1.9 to 4.4	Di Nezio et al., 2021 <sup>34</sup>



## References

- 1 McClain, M. E. & Naiman, R. J. Andean Influences on the Biogeochemistry and Ecology of the Amazon River. *BioScience* **2008**, *58*, 325-338.
- 2 Aucour, A. M. et al. The Amazon River: behaviour of metals (Fe, Al, Mn) and dissolved organic matter in the initial mixing at the Rio Negro/Solimões confluence. *Chem. Geol.* **2003**, *197*, 271-285.
- 3 Richey, J. E. et al. Biogeochemistry of carbon in the Amazon River. *Limnol. Oceanogr.* **1990**, *35*, 352-371.
- 4 Dvorski, S. E. et al. Geochemistry of Dissolved Organic Matter in a Spatially Highly Resolved Groundwater Petroleum Hydrocarbon Plume Cross-Section. *Environ. Sci. Technol.* **2016**, *50*, 5536-5546.
- 5 Schmitt-Kopplin, P. et al. Dissolved organic matter in sea spray: a transfer study from marine surface water to aerosols. *Biogeosciences* **2012**, *9*, 1571-1582.
- 6 Schmitt-Kopplin, P. et al. Analysis of the unresolved organic fraction in atmospheric aerosols with ultrahigh-resolution mass spectrometry and nuclear magnetic resonance spectroscopy: organosulfates as photochemical smog constituents. *Anal. Chem.* **2010**, *82*, 8017-8026.
- 7 Hertkorn, N. et al. High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. *Biogeosciences* **2013**, *10*, 1583-1624.
- 8 Valderrama, J. C. The simultaneous analysis of total nitrogen and total phosphorus in natural waters. *Mar. Chem.* **1981**, *10*, 109-122.
- 9 Gonsior, M. et al. The chemodiversity of algal dissolved organic matter from lysed *Microcystis aeruginosa* cells and its ability to form disinfection by-products during chlorination. *Water Res.* **2019**, *155*, 300-309.
- 10 Kirchman, D., K'nees, E. & Hodson, R. Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems. *Appl. Environ. Microbiol.* **1985**, *49*, 599-607.
- 11 Farjalla, V. F. Are the mixing zones between aquatic ecosystems hot spots of bacterial production in the Amazon River system? *Hydrobiologia* **2014**, *728*, 153-165.
- 12 Simon, M. & Azam, F. Protein content and protein synthesis rates of planktonic marine bacteria. *Mar. Ecol. Prog. Ser.* **1989**, *51*, 201-213.
- 13 Santoro, A. L. et al. Dark carbon fixation: an important process in lake sediments. *PLoS One* **2013**, *8*, e65813.
- 14 Åberg, J. & Wallin, M. Evaluating a fast headspace method for measuring DIC and subsequent calculation of pCO<sub>2</sub> in freshwater systems. *Inland Waters* **2014**, *4*, 157-166.
- 15 Koch, B. P. & Dittmar, T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* **2016**, *30*, 250-250.
- 16 Hsu, C. S., Qian, K. & Chen, Y. C. An innovative approach to data analysis in hydrocarbon characterization by on-line liquid chromatography-mass spectrometry. *Anal. Chim. Acta.* **1992**, *264*, 79-89.
- 17 Hertkorn, N. et al. High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Anal. Bioanal. Chem.* **2007**, *389*, 1311-1327.
- 18 Benner, R., S. et al. Bacterial carbon metabolism in the Amazon River system. *Limnol. Oceanogr.* **1995**, *40*, 1262-1270.

- 19 Farjalla, V. F. et al. Nutrient limitation of bacterial production in clear water Amazonian ecosystems. *Hydrobiologia* **2002**, *489*, 197-205.
- 20 Vidal, L. O. et al. Hydrological pulse regulating the bacterial heterotrophic metabolism between Amazonian mainstems and floodplain lakes. *Frontiers in Microbiology* **2015**, *6*, 1054.
- 21 Wissmar, R. et al. Plankton metabolism and carbon processes in the Amazon River, its tributaries, and floodplain waters, Peru-Brazil, May-June 1977. *Ecology* **1981**, *62*, 1622-1633.
- 22 Signori, C. N. et al. Bacterial production prevails over photo-and chemosynthesis in a eutrophic tropical lagoon. *Estuarine, Coastal and Shelf Science* **2020**, *243*, 106889.
- 23 Signori, C. N. et al. Temporal variability of dark carbon fixation and bacterial production and their relation with environmental factors in a tropical estuarine system. *Estuaries and Coasts* **2018**, *41*, 1089-1101.
- 24 Bastviken, D. et al. Methane as a source of carbon and energy for lake pelagic food webs. *Ecology* **2003**, *84*, 969-981.
- 25 Goosen, N. K. et al. Regulation of annual variation in heterotrophic bacterial production in the Schelde estuary (SW Netherlands). *Aquatic Microbial Ecology* **1997**, *12*, 223-232.
- 26 Calderón-Paz, J. I. et al. Heterotrophic bacterial production in systems of the northern Spanish Mediterranean Region. *Internationale Vereinigung für theoretische und angewandte Limnologie: Verhandlungen* **1993**, *25*, 739-742.
- 27 Crump, B. C. et al. Dominance of particle-attached bacteria in the Columbia River estuary, USA. *Aquatic Microbial Ecology* **1998**, *14*, 7-18.
- 28 Overmann, J. et al. Purple sulfur bacteria control the growth of aerobic heterotrophic bacterioplankton in a meromictic salt lake. *Applied and Environmental Microbiology* **1996**, *62*, 3251-3258.
- 29 Saini, J. S. et al. Bacterial, Phytoplankton, and Viral Distributions and Their Biogeochemical Contexts in Meromictic Lake Cadagno Offer Insights into the Proterozoic Ocean Microbial Loop. *Mbio.* **2022**, *13*, e00052-22.
- 30 Andersson, M. G. et al. Comparison of nitrifier activity versus growth in the Scheldt estuary—a turbid, tidal estuary in northern Europe. *Aquatic Microbial Ecology* **2006**, *42*, 149-158.
- 31 Casamayor, E. O. et al. Primary production in estuarine oxic/anoxic interfaces: contribution of microbial dark CO<sub>2</sub> fixation in the Ebro River Salt Wedge Estuary. *Marine Ecology Progress Series* **2001**, *215*, 49-56.
- 32 Bräuer, S. L. et al. Dark carbon fixation in the Columbia River's estuarine turbidity maxima: molecular characterization of red-type cbbL genes and measurement of DIC uptake rates in response to added electron donors. *Estuaries and coasts* **2013**, *36*, 1073-1083.
- 33 Noguerola, I. et al. Heterotrophic bacterial production in systems of the northern Spanish Mediterranean Region. *FEMS microbiology ecology* **2015**, *91*, fiv086.
- 34 Di Nezio, F. et al. Anoxygenic photo-and chemo-synthesis of phototrophic sulfur bacteria from an alpine meromictic lake. *FEMS microbiology ecology* **2021**, *97*, fiab010.