Molecular links between whitesand ecosystems and blackwater formation in the Rio Negro watershed

Carsten Simon[1], Tania P. Pimentel[2], Maria Terezinha F. Monteiro[2], Luiz A. Candido[2], Didier Gastmans[3], Heike Geilmann[4], Regison da Costa Oliveira[2], João Batista Rocha[2], Elaine Pires[2], Carlos A. Quesada[2], Bruce R. Forsberg[2,5], Sávio J. Filgueiras Feirrera[2], Hillândia Brandão da Cunha[2] and Gerd Gleixner[1*]

[1] Molecular Biogeochemistry, Max Planck Institute for Biogeochemistry (MPI-BGC), Hans Knöll-Str. 10, 07745 Jena, Germany

[2] Coordenação de Dinâmica Ambiental (CODAM), Instituto Nacional de Pesquisas da Amazônia (INPA), Av. Efigênio Sales 2239, Aleixo, Manaus, Brazil

[3] São Paulo State University (UNESP), Centro de Estudos Ambientais, Av. 24A, 1515, Bela Vista, Rio Claro, São Paulo, Brazil

[4] Stable Isotope Laboratory (BGC-IsoLab), Max Planck Institute for Biogeochemistry, Hans Knöll-Str. 10, 07745 Jena, Germany

[5] Vermont Agricultural and Environmental Laboratory, 163 Admin Dr, Randoph Center, 05602, Vermont, United States

*Corresponding author: Gerd Gleixner (gerd.gleixner@bgc-jena.mpg.de)

ORCID IDs (Simon 0000-0003-3882-3013; Pimentel 0000-0002-0103-9658; Monteiro 0000-0003-2071-1785; Candido 0000-0002-4840-5379; Gastmans 0000-0002-1340-3373; Forsberg 0000-0001-7606-0585; Cunha 0000-0002-6434-8829; Gleixner 0000-0002-4616-0953)

Key Points:

- In blackwater catchments, a major portion of net ecosystem productivity is lost by aquatic C export, holding molecular ecosystem information
- Terrestrial markers in the Rio Negro basin show a molecular imprint of specific upland systems but less so from widespread riparian systems
- We propose that terrestrial whitesand markers could help to constrain land-derived primary production and regional net ecosystem C balances

Abstract

We investigated the role of whitesand ecosystems (WSEs) in blackwater formation to develop novel constraints for the terrestrial carbon export from land to ocean. We used Orbitrap mass spectrometry to identify markers in dissolved organic carbon (DOC) from ground- and surface waters of two contrasting WSEs feeding Rio Negro tributaries, and compared them with known Rio Negro markers. Tributaries were fed by a whitesand riparian valley connected to *terra firme* plateau, and a typical upland whitesand *Campina*. WSE-DOC molecular composition differed by 80% from plateau DOC, which was characterized by reworked, highly unsaturated N- and S-containing molecules. WSE-DOC contained mainly condensed aromatics and polyphenols. WSE samples differed by 10% in molecular DOC composition and also by their isotopic content (^{14}C, ^{18}O, ^{2}H). Upland WSE-DOC was exported by fresh precipitation and had maximum age of 13 years, being five years older than riparian valley WSE-DOC. Unexpectedly, only markers from the upland WSE, which cover a small proportion of the landscape, were identical to Negro markers. Markers of the riparian valley WSE, which are widespread and known for high DOC export, surprisingly showed lower coverage with Negro markers. Our results suggest that terrestrial DOC from upland WSEs is a main source of specific blackwater molecules missing in the ecosystem C balance, whereas C export from the riparian valley and especially from *terra firme* plateaus represents mainly recycled and transformed carbon not directly affecting ecosystem C balance. Our study highlights the potential of high-resolution techniques to constrain carbon balances of ecosystems and landscapes.

1 Introduction

The riverine export of terrestrial dissolved organic carbon (DOC) constitutes a major flux within the boundless carbon cycle that connects land and ocean (Drake, Raymond, et al., 2018; Regnier et al., 2013; Webb et al., 2018). Global estimates indicate a total of 208 ± 28 Tg DOC exported by rivers each year (Dai et al., 2012), mainly coming from large tropical and circumboreal watersheds (Li et al., 2019; Raymond & Spencer, 2014). Tropical rivers account for a major portion of the flux, being equivalent to 62-66 % of global DOC export (Dai et al., 2012; Huang et al., 2012). Moreover, recent modelling efforts show that tropical annual exports have been rising over the past 65 years, with an increase of 10 Tg C compared to 1960 in case of South America (Li et al., 2019). The three tropical rivers with highest discharge, the Amazon (incl. Tocantins), the Congo and the Orinoco, alone deliver 18% of global riverine DOC (Raymond & Spencer, 2014). Due to within-river DOC transformations, actual export of terrestrial organic carbon expected to be even higher (Drake, Raymond, et al., 2018), and especially markers are needed to quantify the original terrestrial part of the carbon export.

Blackwater river basins stand out as hotspots of DOC release in the tropics (Junk et al., 2011). The Rio Negro basin for example, which covers roughly 10% of the area of the Amazon river basin, accounts for an annual DOC export of 5.2 – 6.7 Tg C (Coynel et al., 2005; Guinoiseau et al., 2016) equivalent to 17-23% of the Amazon's total DOC export (Raymond & Spencer, 2014). Tropical blackwater rivers such as the Rio Negro are thus pivotal in understanding the regional carbon cycle and its response to global environmental change (Alvarez-Cobelas et al., 2012; Raymond & Spencer, 2014; Webb et al., 2018).

Qualitative analyses of DOC composition have shown great potential to track processes and source contributions on the watershed scale (Creed et al., 2015; Drake et al., 2019; Hutchins et al., 2017; Riedel et al., 2016; Spencer et al., 2019). Processes such as deforestation (Drake et

al., 2019; James et al., 2019), drainage (Moore et al., 2013) or warming (Drake, Guillemette, et al., 2018) all affect the molecular composition or age of leached DOC. Progress in qualitative DOC measures could thus effectively complement global and regional modelling efforts based on quantitative DOC export data (Alvarez-Cobelas et al., 2012; Lv et al., 2019). However, we do know surprisingly little about potential molecular markers in DOM and their environmental fate. Only a few authors have addressed the traceability and stability of ecosystem imprints within the aquatic land-to-ocean continuum, and its conditions (Hutchins et al., 2017; Medeiros et al., 2016; Roth et al., 2014; Wagner et al., 2019). This gap is due to the limited availability of molecular-level analytical tools, such as ultrahigh-resolution mass spectrometry (FT-MS). FT-MS techniques allow unprecedented molecular insight by resolving thousands of signals within a single DOM sample, which are assigned molecular formulae based on exact mass (Hertkorn et al., 2013). Molecular techniques have now added important detail in character and transformation of ecosystem imprints at all stages of the aquatic continuum (Hutchins et al., 2017; Kellerman et al., 2018; Lynch et al., 2019; Raeke et al., 2017; Roth et al., 2019; Wagner et al., 2019). The information content of traditional terrestrial source markers, namely of lignin phenols, has to be questioned due to findings suggesting their fast turnover in soils (Gleixner et al., 2002; Hernes et al., 2007; Marschner et al., 2008) and within the land-to-ocean continuum (Cao et al., 2018), and potential autochtonous sources in marine systems (Powers et al., 2019). Dedicated sets of novel, traceable ecosystem markers are thus rare and need to be calibrated to complement high-resolution DOC flux data (Cao et al., 2018; Medeiros et al., 2016; Roth et al., 2014). Robust sets of molecular markers could promote better understanding of ecosystem-resolved DOM export dynamics and its drivers. This knowledge is pivotal to evaluate and predict the vulnerability and biogeochemical functionality of watersheds under environmental change scenarios (Abbott et al., 2018; Bernhardt et al., 2018; Jehn et al., 2020; McGuire et al., 2014).

As described above, the Rio Negro basin is one of the world's largest DOC emitters in terms of estimated annual flux (6.7 Tg DOC) and yield (9.7 g DOC m^{-2}), making it a classic "blackwater" river (Coynel et al., 2005; Dai et al., 2012). Scientists early noted the co-occurrence of tropical whitesand ecosystems (WSEs) and blackwater streams, and hypothesized a link between them (Goulding et al., 1988; Janzen, 1974; Junk et al., 2011; Leenheer, 1980; Sioli, 1954). The most characteristic feature of WSEs in these landscapes is their sandy soil, classified as either podzol (2% of Amazon basin area) or arenosol (3%) by the World Reference Base (Quesada et al., 2011). These soils differ largely from the more widespread clayey tropical soils such as ferralsols (32%), acrisols (29%), or plinthosols (9%) that are typically found on plateaus and their slopes (Lucas et al., 2012; Do Nascimento et al., 2004; Quesada et al., 2011). Due to the low water-holding capacity of sand, WSEs are also characterized by specific types of vegetation that differ from highly diverse *terra firme* rainforests which are dominated by large trees. In the central Amazon, these are sclerophyllous, shruby and smaller-tree dominated *Caatinga*, *Campina*, and *Campinarana* forests, known for their unique plant secondary metabolites and adapted decomposer communities (Demarchi et al., 2018; Janzen, 1974; Klinge & Medina, 1979; Vasco-Palacios et al., 2018; Zanchi et al., 2015). WSEs occur as local upland depressions on plateaus, as intersected valleys forming large riparian corridors at the foot of plateau slopes, or in low-elevation terrain in the form of wide peneplains (Montes et al., 2011; Do Nascimento et al., 2004). Roughly, upland and riparian valley WSEs are dominant in the lower Rio Negro basin while peneplains are widely distributed in the upper Rio Negro basin (Montes et al., 2011). It is likely that the molecular composition of DOM resolves the different

environmental conditions, allowing for proper landscape-based DOC source identification and export calculations.

Previous studies that assessed terrestrial sources of Rio Negro DOM largely supported the older hypotheses that WSEs, and more specifically the widely distributed riparian corridors are responsible for the highest amount of carbon export (Bardy et al., 2011; Junk, 1993; Melack & Hess, 2010; Remington et al., 2007). However, this DOC export is mainly controlled by precipitation amount and flooding events (McClain et al., 1997; Remington et al., 2007; Zanchi et al., 2015). This is in line with the finding that in riparian settings, DOC generally shows transport-limited behavior, meaning that its absolute flux scales with discharge (Musolff et al., 2017; Zarnetske et al., 2018). As a result, water passing through the riparian zone continuously leaches existing reserves of processed organic matter (Laudon & Sponseller, 2018; Ledesma et al., 2015; Tiegs et al., 2019). In line with the older hypothesis, McClain and coworkers reported low annual DOC yields for widespread plateau areas (*terra firme* – ferralsol; 2 g DOC m^{-2} yr^{-1}) but large yields for a relatively small *Campina* WSE catchment (40 g DOC m^{-2} yr^{-1}) in the lower Negro basin, north of Manaus. The authors predicted that a WSE molecular DOM signal would thus be easily detectable in higher order rivers within the Rio Negro catchment (McClain et al., 1997), and later studies conducted in the same region corroborated this hypothesis (Remington et al., 2007). DOM from well-developed podzols reflected best the chemical properties of DOM isolates from local groundwater and nearby rivers, showing the fast transit of DOM in well-drained sandy soils with low sorption potential (Bardy et al., 2011; Remington et al., 2007). However, no subsequent markers were identified or tracked in the lower reaches of the stream network to assess their environmental fate or marker potential (Bardy et al., 2011). Such novel markers are however needed to study variations in DOM export and spatiotemporal dynamics in riverine DOM sources within a catchment (Bernhardt et al., 2018; Hutchins et al., 2017; Laudon & Sponseller, 2018).

Recently, watershed-specific molecular DOM signatures of the Rio Negro and other Amazon tributaries (Tapajos, Madeira, Solimões) were reported (Gonsior et al., 2016; Simon et al., 2019) that could serve as potential markers of ecosystem DOC exports due to their largely conservative behavior during mixing (Simon et al., 2019). We here make use of these openly available FT-MS DOM datasets and compare them to groundwater, surface, and soil water DOM measured by Orbitrap FT-MS. We hypothesized that the overall large export of DOM from riparian WSEs in the Rio Negro basin would allow for the retrieval of Rio Negro-specific markers as assumed by previous studies (Bardy et al., 2011; McClain et al., 1997; Remington et al., 2007). We therefore compared an upland *Campina* forest, and a riparian valley system dominated by *Campinarana* forest, both typical for WSE-podzol systems within elevated *terra firme* plateaus north of Manaus. We hypothesized that both WSEs and plateaus would differ in terms of water chemistry and DOC properties, and that the DOM molecular composition would reflect these differences as well, yielding new sets of unique ecosystem markers. We further hypothesized that riparian valley WSE markers would indicate clear overlap with known Rio Negro markers, and could thus serve as complementary proxies of land-derived primary production in the Rio Negro basin.

2 Material and Methods

2.1 Field sites and sampling procedures

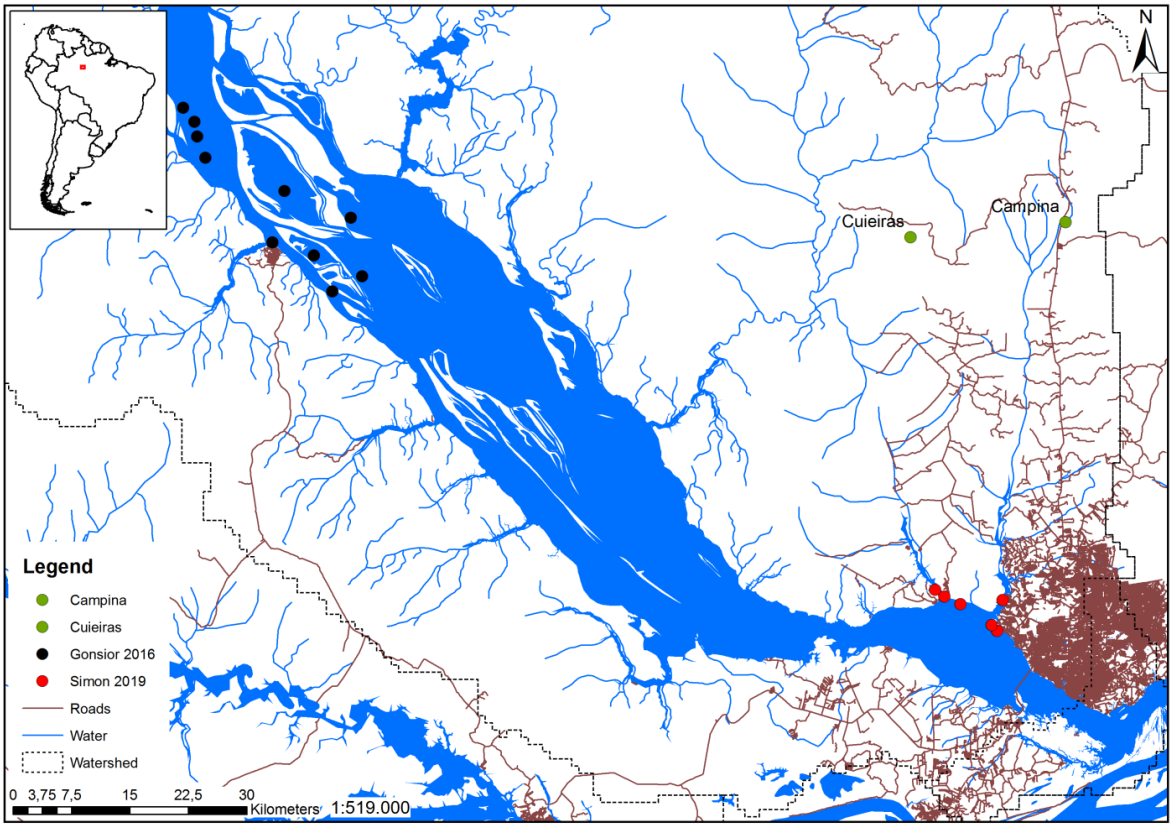

Figure 1. Detail of the lower Rio Negro catchment northwest of Manaus showing sampling sites of available FT-MS studies. We sampled at two locations north of Manaus (green dots; "Cuieiras", "Campina"). Two other FT-MS datasets from blackwater sampling locations are shown, including the Rio Negro and connected lakes (black dots; Gonsior et al., 2016) and the lower reach of the Rio Negro and two of its tributaries close to Manaus (red dots; Simon et al., 2019). We accessed watershed limits and river data as shapefiles from www.ore-hybam.org (Seyler et al., 2009). The headwater stream width is not drawn to scale. Roads and main water bodies were extracted from the OpenStreetMap project (OSM; natural features and roads) and downloaded as shapefiles from www.download.geofabrik.de. Map editor: Marcus Guderle, MPI Jena.

Soil water samples were taken in early November 2017 at the onset of the rainy season in two protected forest reserves under the responsibility of the Instituto Nacional de Pesquisas da Amazônia (INPA) in Manaus, Brazil (**Figure 1**; **Supporting Information Figure S1**). Both reserves, the Reserva Biológica do Cuieiras – ZF2 (2°36'32.67" S, 60°12'33.48" W, at 40–110 m above sea level) and the Reserva Biológica de Campina (2°35'30.26" S, 60°01'48.79" W, at 93–101 m a.s.l.) are located about 60 - 70 km north of Manaus (Marques et al., 2016; Zanchi et al., 2014). We conducted sampling from 31st October – 2nd of November, and all lab procedures followed within three days. The geological setting, landscape structure, forest composition, and soil characteristics have been described in detail by Zanchi et al. (2014). Broad swampy valleys surrounded by elevated plateaus cover about half of the Cuieiras reserve's area (in total, 22735 ha) (Zanchi et al., 2014). The valley soils (podzols, gleysols) differ markedly from the clayey plateau and slope soils (oxisols, ultisols). Bleached quartz sand and high amounts of phenolics characterize the valley's podzols (Marques et al., 2016; Monteiro et al., 2014; Zanchi et al., 2014), and the presence of *Mauritia flexuosa* (Luizão et al., 2004), a palm species indicative of hydromorphic conditions, suggests poor drainage (Junk, 1993). A second-order blackwater stream (Rio Aşu) drains the area, meeting with Rio Cuieiras, Rio Branquinho, and Rio Negro

downstream (Monteiro et al., 2014). The catchment has an area of 660 ha (Monteiro et al., 2014). Monteiro and coworkers report a total annual rainfall of 2806 mm in 2002 and 2004 mm in 2003. Stream discharge was 0.08 – 5.59 $m^3 s^{-1}$ (average, 0.18 $m^3 s^{-1}$) and DOC levels ranged from 3.2 – 15.2 mg C l^{-1} (average, 8.7 ± 3.0 mg C l^{-1}) during that period. The watershed's annual total carbon flux was estimated as 13.3 g C m^{-2} in 2003 (Waterloo et al., 2006) and Monteiro and coworkers estimated a stream DOC flux (Rio Aşu) of 8.7 g C m^{-2} for the same year. Zanchi and coworkers estimated the annual DOC export of the Aşu watershed to lie within the range 9.3 – 22.7 g C m^{-2} (Zanchi et al., 2015). Water samples were taken from piezometers installed across a valley transect (Monteiro et al., 2014) which is maintained and sampled regularly. Piezometers were emptied once before final sampling. The stream was sampled manually, with nitrile gloves, against the direction of flow, using pre-cleaned (acidified ultrapure water, pH2, HCl, Merck EMSURE®, p.a., ACS grade) Nalgene™ polycarbonate bottles (Fisher Scientific, Schwerte, Germany) that were cleaned with the respective sample before final sampling. We also sampled two deep wells on the plateau (35 m and 39 m depth) by lowering an empty, clean sampling bottle on a string until water was reached.

The Reserva Campina is a 900 ha reserve that shows only small relief; poor sandy soils (up to 99% sand) that co-occur with typical but specific forest types, so-called *Campina* and *Campinarana* forests (heath forests; Demarchi et al., 2018) characterize the area. In contrast to the highly diverse plateau (*terra firme*) forests, *Campinas* show much lower species diversity and low canopy (~10 m vs. 25-40 m at Reserva Cuieiras). Bare patches of sand cover ~11% of the area (Zanchi et al., 2014). The headwater area is drained by a single headwater stream that is less than 1m wide and often less than 30 cm deep (McClain et al., 1997) and feeds the Rio Tãruma Açu in the southward direction, meeting with the Rio Negro close by Manaus (**Figure 1**). The catchment area is estimated by 6.5 ha and thus ~100 times smaller than the Rio Aşu catchment (Zanchi et al., 2015). According to the same authors, annual rainfall levels are comparable among both reserves. Surface runoff (from the stream) and estimated groundwater outflow amounted to 485 and 1071 mm in that period, respectively (in sum 1556 mm; as compared to Aşu 1362 mm; Waterloo et al., 2006). Zanchi and coworkers estimated the total watershed's DOC export with 49.2 g C m^{-2}, including groundwater outflow and rainfall, of which the stream exported 15.3 g C m^{-2}. An annual streamflow of 485 mm equals an estimated discharge of 1.0 x 10^{-3} $m^3 s^{-1}$, which is about 180 times lower than average discharge at Rio Aşu in 2002 – 2003 (Monteiro et al., 2014). McClain and coworkers reported annual DOC exports of 40 g C m^{-2} yr (McClain et al., 1997; Zanchi et al., 2015) from the same catchment in the period 1993 – 1994. We took samples at the side slopes of the stream from piezometers installed in 1993, as detailed in McClain et al. (1997). For this purpose, wells were emptied three times and sampled afterward (Zanchi et al., 2015). The stream was sampled as described above.

2.2 Water chemistry

Aliquots of the samples were subjected to TOC analysis in the water laboratory of the Instituto Nacional de Pesquisas da Amazônia (INPA) in Manaus, Brazil (Laboratório de Águas do INPA/ CPRHC – Coordenação de Pesquisas em Recursos Hídricos e Clima). Samples were measured on a total organic carbon analyzer (TOC-VCPH model, Shimadzu, Kyoto, Japan) (Marques et al., 2010; Monteiro et al., 2014). Before extraction, we analyzed samples for pH and electrical conductivity (EC) with a Multi 340i probe system (WTW, Weilheim, Germany).

2.3 Solid-phase extraction

DOM samples were solid-phase extracted (SPE) shortly after sampling at INPA, Manaus (Laboratório de Ecossistemas Aquáticos) using an established protocol (Dittmar et al., 2008). The solid-phase sorbent was a modified styrene-divinylbenzene polymer (PPL Bond Elut™, Agilent, Santa Clara, CA, USA). Samples were acidified to pH 2 with 37% hydrochloric acid (Merck EMSURE®, p.a., ACS grade) before extraction. Solvents used for extraction were ultrapure water, acidified ultrapure water (pH 2, HCl), and ultrapure methanol (Biotec Reagentes Analíticos, p.a., ACS grade). We loaded columns with maximal amounts of 3 mg C. The extraction efficiency (EE) of samples with high DOC concentrations (> 2 mg l^{-1}, n = 11) was always > 60% and on average 71 ± 8%, at a loading ratio of 443 ± 143 (average ± standard deviation; PPL: DOC in mg/ mg). Sample PT6 was an exception (high DOC, low EE: 33%). Samples with low DOC concentrations (< 2 mg l^{-1}, n = 4) showed lower extraction efficiency (23 ± 15%), and loading ratios were generally higher (2400 ± 960) (**Supporting Information Table S1**).

2.4 Water isotopes

Isotopic signatures of water (δ^2H- and δ^{18}O-values) were analyzed by high-temperature conversion-isotope ratio mass spectrometry (HTC-IRMS) in the stable isotope laboratory of the Max Planck Institute for Biogeochemistry (BGC-IsoLab). For method details, we refer the reader to the literature (Gehre et al., 2004). In short, measurements were conducted on a Delta+ XL coupled to a high-temperature furnace via a ConFlow III interface (Thermo Fisher Scientific, Bremen, Germany). One µl of water was injected using an A200S autosampler (CTC Analytics AG, Zwingen, Switzerland), and the furnace temperature was held at 1350 °C. The δ^2H and δ^{18}O values are reported on the VSMOW scale, which is realized by parallel analysis of samples against in-house standards. In-house standards are routinely calibrated against internationally accepted water standards. Daily standard deviations for δ^2H and δ^{18}O measurements are usually better than 1 and 0.1 ‰, respectively. We accessed regional isotope ratio estimates of rain for October and November with the help of the *online isotopes in precipitation calculator* (OIPC2.2, version 3.1, http://wateriso.utah.edu/waterisotopes/; Bowen & Revenaugh, 2003). The values were similar for both sites; -7 ‰ and -17 ‰ (V-SMOW) for δ^2H, and -2.4‰ and -3.8‰ (V-SMOW) for δ^{18}O in October and November, respectively. We estimated the average between both monthly values for each isotope, considering our sampling date at the end of October/ beginning of November.

2.5 Radiocarbon analysis

We transferred an aliquot of methanolic PPL extract equivalent to 0.25 mg C to tin capsules (8 mm diameter, 20 mm height; IVA Analysentechnik, Meerbusch, Germany). The methanol was left to evaporate. Air-dried capsules were combusted in an elemental analyzer and graphitized for radiocarbon analysis on a 3 MV Tandetron ^{14}C-AMS (HVEE, Amersfoort, Netherlands) at the Max Planck Institute for Biogeochemistry in Jena, Germany. Modern (Oxalic Acid II) and ^{14}C-depleted standard materials were carried along for quality control and data corrections (Benk et al., 2018; Steinhof et al., 2017). Graphitization is conducted with Duran glass tubes at a temperature of 550°C in the presence of hydrogen and with an iron catalyst (Steinhof et al., 2017). Radiocarbon concentrations are given as fraction modern ($F^{14}C$), which is the fraction of the standard concentration normalized for δ^{13}C of the oxalic acid standard measured at the same time (with δ^{13}C of -19‰; Trumbore et al., 2016). Δ^{14}C values also take into account radioactive decay of the oxalic acid standard since 1950, which demarks the start of bomb tests that increased the ^{14}C content of the atmosphere (Trumbore et al., 2016). We

calibrated our background-corrected $F^{14}C$ data with atmospheric data published in Graven et al. (2015) to derive DOC age. We used values for the inner-tropical convergence zone (zone SH3). To cover the time point of sampling (2017), we predicted $\Delta^{14}C$ values for 2016 and 2017 based on data from 1997-2015 by linear regression (R^2 = 0.996) and then estimated DOC mean calendar age. The measurement error was < 1 year and < 2 years for repeated measurements (standard deviation).

2.6 Orbitrap measurements and data processing

We conducted Orbitrap measurements at the Max Planck Institute for Biogeochemistry, as described elsewhere (Simon et al., 2018), at a nominal resolution setting of 480.000. Allowable numbers of atoms in formula calculation were as follows: ^{12}C, 0-60; ^{1}H, 0-120; ^{14}N, 0-2; ^{32}S, 0-1; ^{16}O, 1-60, ^{13}C, 0-1. The assignment was done at ± 1ppm tolerance. We defined noise as the smallest peak in the sample set, not including zeros. We then calculated the maximal signal-to-noise ratio (S/N_{max}) of a peak across all samples (without blanks and reference material) and only kept peaks with a value > 5. We calculated the number of matches of each peak across all samples (without blanks and reference) and only kept peaks with more than one match. To exclude contaminants, we calculated a signal-to-noise ratio of blank peaks (only blank samples) and discarded all peaks with values > 20. Finally, we excluded those peaks that were only present in less than 20% of all samples (i.e., < 9 matches) and small (SN_{max} < 20). After the exclusion of peaks outside the usual mass defect range of natural organic compounds (nominal *m/z* in Da, -0.05 mDa … +0.3 mDa), 13658 peaks remained in the dataset in total. Of those, 9893 had a formula assigned, and 3672 included one ^{13}C atom. Formulae having H/C ratios > 2 or DBE-O values (double bond equivalent minus oxygen atoms) > 15 (modulus) or containing the elemental combination N_2S were discarded due to their unlikely occurrence in DOM (Hawkes et al., 2016). In the case of ambiguous peaks with more than one assigned molecular formula, only suggested CHO formulae with a DBE-O of < 10 (modulus) were kept (Herzsprung et al., 2014). Other formula suggestions or ambiguous hits were excluded from the formula pool and kept as "no reference" peaks. ^{13}C-containing formulae that were missing their equivalent monoisotopic (only ^{12}C) formula were excluded. As a last measure, we only considered peaks detected twice in two separate runs for further analysis (Riedel & Dittmar, 2014). The final dataset contained 7705 formulae (of those, 1963 containing a ^{13}C). For comparison of samples, we normalized all mass spectra to the sum of their peak intensities (including all peaks > S/N = 5, also those with no assigned formula). The further analysis of the data focused on the subset of peaks with an assigned monoisotopic formula (n = 5709). The crosstab is available from (**Supporting Information Data Set S1**).

2.7 Statistical analyses: Ecosystem markers

We analyzed the molecular formula data by Principal Coordinate Analysis (PCoA, cmdscale function, stats package, v3.5.1) and post-ordination gradient fitting analysis (envfit function, vegan package, v2.5-2) within the statistical computation environment R Studio (v1.1.453, © 2009-2018 RStudio, Inc.). PCoA was based on Bray-Curtis dissimilarities obtained by the function vegdist from package *vegan* (Oksanen, 2010). The method allows the comparison of samples based on their formula populations. Redundancy within the dataset, i.e., formulae showing similar trends in ion abundance across samples, is effectively reduced and yields a set of coordinates that summarize the variability of the dataset best (Osterholz et al., 2016). We then analyzed the distribution of samples in coordinate space for clustering (indicating similarity among samples) and correlations with specific DOM indices by the envfit

function (at 999 permutations). DOM indices aggregate properties of the molecular formula population of each sample (**Supporting Information Table S2**). We further assessed molecular formulae with a significant (p = 0.05) association to clusters of samples derived from PCoA analyses by Student's t-test of averaged relative ion abundances across samples of each group. We conducted two-sided tests assuming unequal variances.

2.8 Origin of Rio Negro markers

To link headwaters and downstream signals, we compared sets of specific molecular formulae to known Rio Negro-specific markers available from two open-access FT-MS datasets (Gonsior et al., 2016; Simon et al., 2019). The dataset from 2016 compared samples from the Rio Negro and its adjacent lakes (close to Novo Airão, 120 km northwest of Manaus) to distant large rivers (Rio Tapajos, Rio Madeira) to reveal large-scale differences in chemodiversity, i.e., unique signals of each river basin. We extracted the robust Rio Negro fingerprint by selecting unique formulae of the Rio Negro detected across all measurements, with an average ion abundance of at least $5*10^7$ (or ~ 0.01% relative intensity; n = 225; Gonsior et al., 2016). The dataset published in 2019 assessed differences of riverine DOM at the confluence of the Amazon near Manaus (Encontro das Águas). To obtain a robust DOM fingerprint, we extracted formulae that showed a significant positive correlation (Pearson's r, p = 0.05) to the fraction of Rio Negro during mixing in each of the studies' three experiments (Simon et al., 2019; n=299). The datasets were also compared on a general level. Information on this aspect is presented in the Supporting Information (**Supporting Information Text S1**). The combined data from all three studies is available in presence/absence format and with ion abundance information from the Pangaea Data Publisher website (**Supporting Information Data Set S2**).

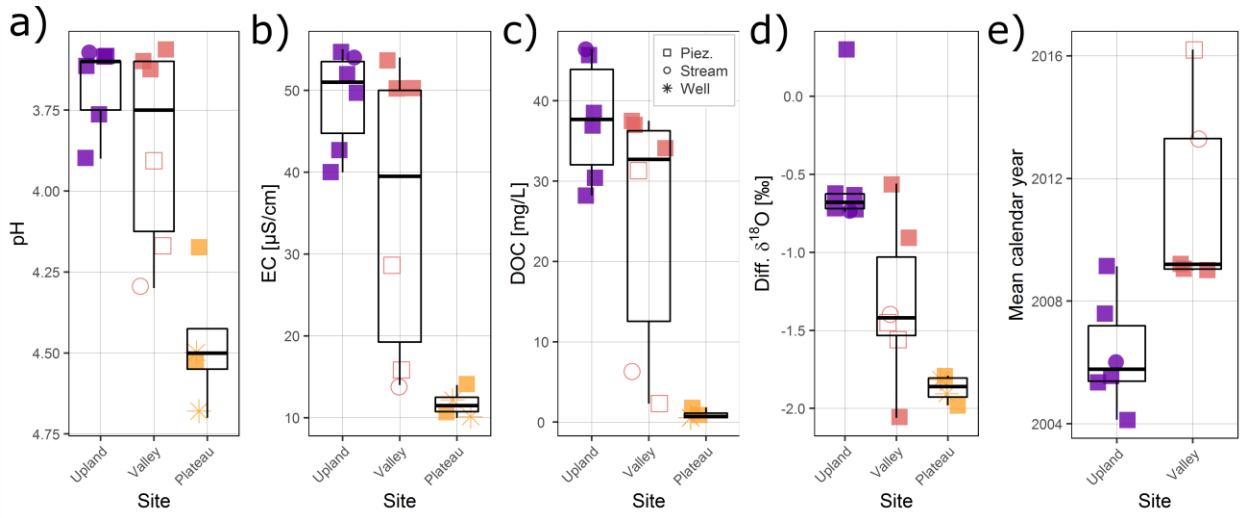

Figure 2. Differences in the water and SPE-DOM properties. Variables shown are a) pH, b) electrical conductivity, c) dissolved organic carbon concentration, d) difference in δ18O values of water compared to regional average precipitation (OIPC estimate), e) mean calendar age of DOC in solid-phase extracts. Samples are grouped into biogeochemical environments (*upland* Campina; riparian *valley*, and *plateau*). Symbols denote sample type (squares, piezometer; star, deep well; circle, stream); open symbols mark "intermediate" valley samples (see main text).

3 Results and discussion

3.1 Site characterization: Biogeochemistry of water and carbon

The general analysis of water chemistry showed clear differences among samples, revealing two major endmembers and one intermediate environment. Strongest differences were apparent between fine- and coarse-textured soils ("Plateau" vs. "Upland"; **Figure 2**, **Supporting Information Table S1**). Acidity (pH), electrical conductivity (EC) and concentrations of dissolved organic carbon (DOC) were 3.7 ± 0.1, 49 ± 6 µS cm^{-1} and 37.7 ± 7.54 mg C l^{-1} in Campina samples (**Figure 2a – c**), and some samples at Reserva Cuieiras indicated the same tendency (PR7, PR8, PR9), suggesting common WSE water properties. Samples from the plateau environment showed slightly higher pH, and low EC and DOC levels (4.5 ± 0.2, 12 ± 2 µS cm^{-1} and 0.9 ± 0.6 mg C l^{-1}). Three samples from Reserva Cuieiras (PR10, PT06, and stream RA, shown by open symbols) indicated "intermediate" levels of all three parameters (pH 3.9 – 4.3, EC 14 – 29 µS cm^{-1} and 2.3 – 31.3 mg C l^{-1}) and were likely plateau-influenced. Magnitudes and correlations of acidity (pH), electrical conductivity (EC), and dissolved organic carbon (DOC) concentration agree with previous reports from similar environments, showing a strong positive correlation of EC, proton concentration, and DOC (Bardy et al., 2011; Monteiro et al., 2014; Do Nascimento et al., 2008). Plateau samples showed higher pH values and stronger variability, probably indicating ongoing buffering by oxides, and the absence of acidic DOC (Do Nascimento et al., 2004, 2008). Relatively low DOC concentrations of Rio Aşu (RA) compared to surrounding valley piezometers indicated dilution by precipitation at the time point of sampling. In contrast, DOC concentration of the stream ("Rio Campina", RC) draining the upland site revealed similarly high DOC levels as the surrounding piezometers, which indicates a more direct contact stream and surrounding soil at the time point of sampling (McClain et al., 1997; Zanchi et al., 2015). Despite similarities in water chemistry of WSE samples ("Valley" and "Upland", **Figure 2a-c**), differences were most apparent in water isotopic composition, and radiocarbon content of DOC (**Figure 2d-e**, **Supporting Information Figure S2**). All samples plotted on the local meteoric water line, resembling the range of expected isotopic composition found in the region. However, water was overall lighter as compared to local precipitation in previous years (**Supporting Information Figure S2**). Despite regional climatological effects, sites differed significantly in water composition. Campina samples showed a very homogenous water isotopic composition except for the most upslope sample P2. In comparison, samples from Reserva Cuieiras showed consistently lighter (more negative) δ^2H and $\delta^{18}O$ values; plateau samples were lightest and showed a very homogenous water isotopic composition. The valley samples indicated variation in isotopic composition within the range of the other samples, indicating mixing between endmembers. Local differences in ecohydrology can explain the heavier isotopic composition of groundwater and stream water in the drier upland Reserva Campina (Zanchi et al., 2014, 2015), and the lighter isotopic composition of water in the valley and its adjacent plateaus (Kunert et al., 2017; Leopoldo et al., 1982). While isotopically heavy samples like upland P2 (most upslope) and valley PR8/ PR9 (high water levels) likely reflected the imprint of recent precipitation, groundwater replenished during previous wet seasons can explain the lighter water isotopic composition (Jasechko & Taylor, 2015; Leopoldo et al., 1982; Miguez-Macho & Fan, 2012b; Tomasella et al., 2007; Zanchi et al., 2015; Zhang et al., 2009; **Supporting Information Figure S3**). While the $\delta^{13}C$ composition of all DOC extracts was roughly similar (average ± standard deviation: -29.40 ± 0.45 ‰, n=13) and reflected a typical C_3 source, their calibrated radiocarbon ages were young – less than 2 to 13 years old at maximum – and differed markedly by five years between both sites, with the upland site showing older DOC (**Figure 2e).** Young radiocarbon ages of tropical DOC are frequently reported from aquatic (Mayorga et al., 2005; Moyer et al., 2013; Ward et al., 2013) and soil systems (James et al.,

2019) and are explained by strong linkages between C fixation, DOC release, and nutrient recycling (Mayorga et al., 2005). The slight but consistent differences in DOC radiocarbon ages between the drier upland and the wetter valley system likely relate to processes that govern the short-term (years to decades) turnover of organic matter, for example litter turnover and soil respiration rates (Zanchi et al., 2011, 2014, 2015).

3.2 DOM characterization by ultrahigh-resolution mass spectrometry

The analysis of molecular DOM composition reflected the separation of samples based on water and DOC properties presented in the previous section (**Figure 3a**). The PCoA separated samples into two main clusters ("plateau" and "WSE") on the first coordinate, which held 78% of molecular variation. Plateau-derived DOM was thus most dissimilar from WSE-DOM, in line with a major control of DOC properties and abundance by soil texture (Remington et al., 2007). Consequently, PCoA 1 was linked to significant (Pearson' r, $p < 0.05$) trends in pH, EC, and DOC (not shown). In line with more subtle differences between WSE samples, the overall explained variability of the second coordinate was smaller (10%). However, WSE sites were clearly separated, and molecular trends thus paralleled differences in water isotopes and radiocarbon age described above. Moreover, the separation of WSE-DOM by sites suggested differing trajectories of DOM processing that seemed to converge to a common DOM composition (note the close proximity of samples P7 and PR7 in **Figure 3a**). Several significant trends in molecular indices derived from the DOM data paralelled the PCoA separation (**Figure 3b-d).** The used descriptors – a-priori defined molecular groups and formula classes that aggregate molecular composition information – are described in **Supporting Information Table S2**, and data is available online from the Pangeae Data Publisher website (**Supporting Information Data Set S1**). WSE-DOM was more oxidized (higher O/C and NOSC), less saturated (lower H/C, higher DBE) and more aromatic (higher AImod) than plateau DOM (**Figure 3b**). Consequently, a-priori-defined molecular groups reflected those trends (**Figure 3c**): Aromatic and oxidized groups (polyphenols, black carbon-like, carbohydrate-like, and O-rich highly unsaturated formulae) were more abundant in WSE samples, while aliphatic, less oxidized groups (Oxygen-poor highly unsaturated markers, unsaturated aliphatics, and peptide-like compounds) were dominant in plateau DOM. Average numbers of C, O, and N atoms per formula followed these major trend as well (**Figure 3b, d**). Simple oxidized formulae (CHO) dominated the upland WSE cluster while the percentage of N- and S- containing formulae (CHNO, CHOS, CHNOS) were more abundant in plateau and plateau-influenced ("intermediate") valley WSE samples. The percentage of CHNO formulae also differentiated WSE-DOM from upland and valley sites on PCoA2. Besides the effect of N-containing formulae, upland WSE-DOM was also heavier in terms of molecular weight (MW, **Figure 3b**) due to more C and O atoms per molecular formula. These findings are in line with known bulk characteristics of DOC endmembers from soils, groundwater, and rivers in the region (Leenheer, 1980; McClain et al., 1997; Remington et al., 2007). In a next step we thus extracted the molecular markers that caused the separation of DOM from different plateau and whitesand environments.

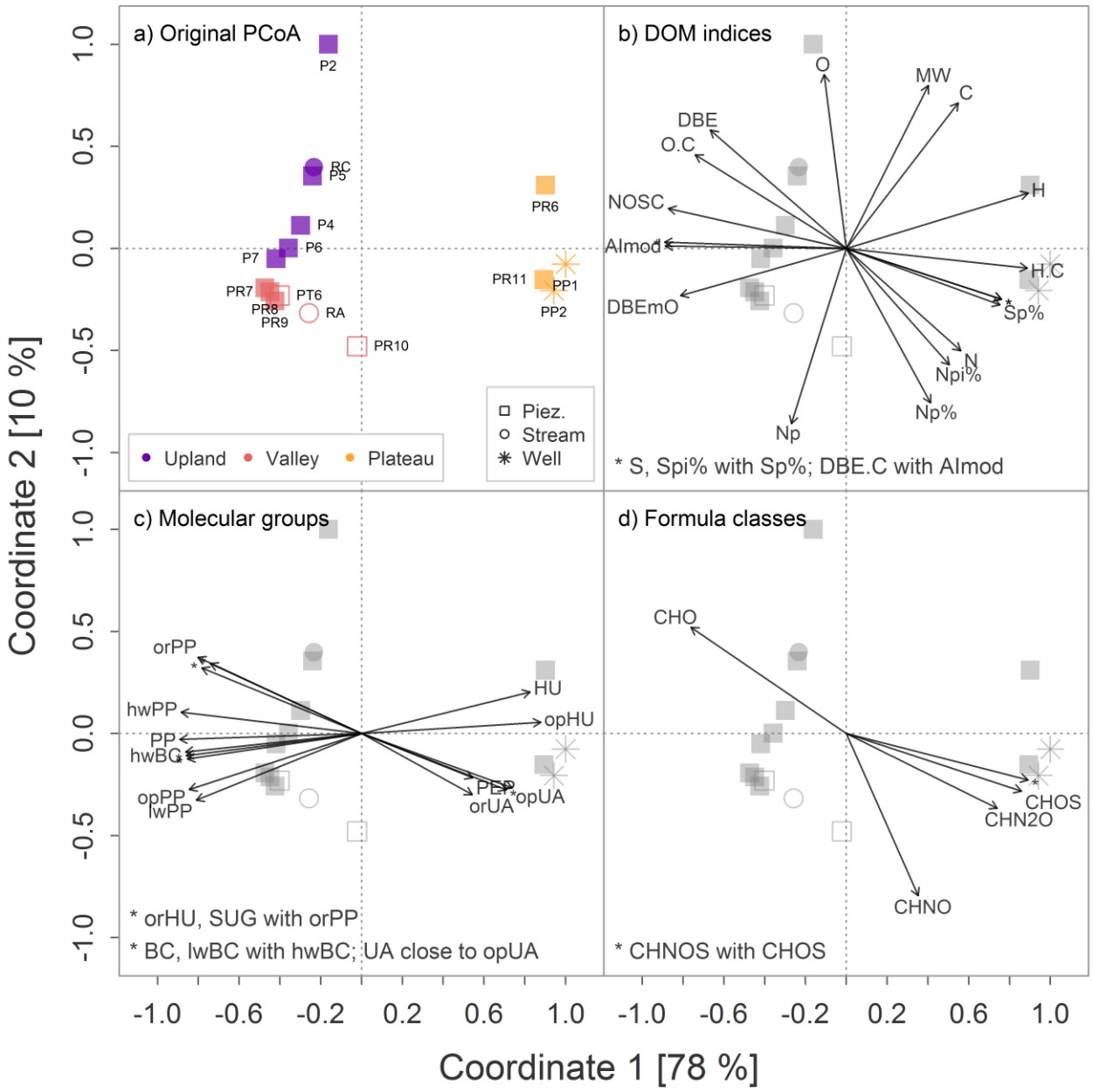

Figure 3. Multivariate analysis based on molecular DOM data. **a)** Separation of samples in a principal coordinate analysis (PCoA) based on Bray Curtis dissimilarity. The plot shows only the first two coordinates, the third coordinate (5% of explained variability, not shown) did not contribute further to separation. Percentages denote the degree of explained variability in DOM molecular composition. Samples are grouped into biogeochemical environments ("Upland", "Valley", "Plateau"). Symbols denote sample type (square, piezometer; star, deep well; circle, stream); open symbols mark "intermediate" valley samples (see main text). **b - d)** Post-ordination gradient fit (function envfit of R package vegan, at 999 permutations) of three different sets of variables based on PCoA separation (same as in **a**). Variables sets are b) DOM indices, c) Molecular groups, and d) Formula classes. Significant correlations (Pearson, $p < 0.05$) with the ordination are shown as arrows. Arrow length corresponds to the strength of correlation and arrows head into the direction of the steepest increase of the respective variable, based on the ordination pattern of samples. Variable abbreviations in b): C, H, O, N, S (average numbers of respective atoms per formula), MW (molecular weight as mass to charge-ratio), AImod (Aromaticity index), DBE, DBEmO, DBE.C (double bond equivalents, DBE minus oxygen, DBE/C ratio), H/C (atomic ratio of hydrogen to oxygen, "saturation axis"), O/C (atomic ratio of oxygen to carbon, "oxidation axis"), NOSC (nominal oxidation state of carbons), Np, Np%, Npi% (number, percentage and relative abundance of N-containing peaks), Sp, Sp%, Spi% (same for S-containing peaks). Abbreviations in c): BC (Polycyclic, condensed aromatics, such as "Black Carbon"), PP (polyphenols), HU (highly unsaturated), PEP (unsaturated, O- and N-containing, such as peptides), UA (unsaturated aliphatics), SUG (very high O content, such as sugars). Prefix to PP, HU & UA:

op (O/C ≤ 0.5); or (O/C > 0.5). Prefix to BC and PP: lw (< 15 C atoms), hw (≥ 15 C atoms). Abbreviations in d): CHO (average number of molecular formulae containing only C, H and O atoms), CHNO, CHN2O, CHOS, CHNOS (formulae containing one N, two N atoms, one S atom, or both one N and S atom).

As expected from the results of the gradient analysis presented in the previous section, we found clear sets of markers for each DOM type that paralleled trends of weight-averaged molecular indices (**Figure 3b-d**). To this end, we separated unique from shared ("common") markers, and thus excluded non-informative formulae (**Figure 4**, **Supporting Information Figure S4**, and **Supporting Information Figure S5**). It is important to note that our definition of "unique" and "common" relates to significant differences in abundance. In fact, most molecular formulae were shared based only on presence (30% of all formulae shared among all three ecosystems, and 83% among WSEs; **Supporting Information Figure S6**). However, besides subsets of formulae indicating less overlap (CHNO, CHOS: **Supporting Information Figure S6**; BC, PP, HU: **Supporting Information Figure S7**), ecosystems differed significantly in abundances of formulae, giving rise to individual "molecular fingerprints" (**Supporting Information Figure S5**). Markers of plateau DOM showed a narrow mass and chemical space distribution as assessed by van Krevelen diagrams (formulae centered at *m/z* 425, O/C < 0.5, and H/C > 1). They were also characterized by high numbers of N- and S-containing formulae (63% of all markers) as compared to "simpler" CHO formulae (**Figure 4a-c**, and **Supporting Information Figure S5a**). Unsaturated, relatively less oxidized nitrogen-containing formulae classified as "highly unsaturated" compounds ("HU", i.e., lignin-like formulae, or carboxyl-rich alicyclic molecules, CRAM) were the most dominant group of markers, representing ~50% of all plateau markers. In contrast, molecular markers common to both WSE-DOM types were dominated by CHO formulae and showed a broad chemical space distribution with molecular weights up to m/z 800, and a distinct center at low m/z (~ 200 Da; **Supporting Information Figure S4a-c**). The most distinct WSE-DOM markers were oxidized, phenolic/ aromatic CHO formulae (74% belonging to molecular groups "orHU", "PP", or "BC"; and 64% belonging to "CHO" class). The distinct markers of plateau and WSE settings documented a major texture effect on DOM properties: Fine-textured soils can cause longer water retention and contact times between minerals, microbes, and water (Marques et al., 2004; Remington et al., 2007) and may thus favor overall lower DOC levels due to intensified decomposition (Marques et al., 2010). Clay particles are also often associated with N-containing compounds (Chassé et al., 2015; Newcomb et al., 2017), and newly synthesized, larger and N-containing microbial compounds can also become dominant during decomposition (Roth et al., 2019). Previous reports on narrow C/N ratios of DOC (~10) in plateau soils support these general findings (McClain et al., 1997). Higher DOC concentration and the dominance of aromatic, oxidized WSE markers may thus represent initial stages of decomposition that also agree with reports on wider C/N ratios in WSEs (>15, up to 60; McClain et al., 1997). These results demonstrate the overall importance of WSEs for the amount and qualtity of exported terrestrial DOC in the lower Rio Negro basin.

We observed clear molecular differences between WSE-DOM from the Cuieiras valley and upland Reserva Campina, as suggested by the explorative PCoA (PCoA 2 in **Figure 3a**), and unique markers of each WSE reflected this divergence (**Figure 4d-f**, and **g-i**). This differentiation was most apparent through a sharp "cutoff" at ~ *m/z* 400 (**Figure 4d, g**). Reserva Cuieiras valley samples were characterized by lower-molecular-weight N-containing formulae which represented 62% of all valley markers (weighted average mass ~ *m/z* 300; **Figure 4d, f**, and **Supporting Information Figure S5b**). Despite their low mass, these formulae showed wide distribution in chemical space, mostly belonging to the classes of highly unsaturated and

polyphenolic compounds ("HU", 47% of all markers, and "PP", 20%; **Figure 4e**). Upland WSE-DOM, on the other hand, was differentiated by a confined cluster of CHO formulae with higher average mass (~ *m/z* 425). Indicative upland CHO formulae concentrated in a chemical space defined by O/C > 0.3 and H/C < 1 and showed distribution across a wide mass range, high aromaticity, and high oxidation state (**Figure 4h, i**). All in all, 57% of upland WSE markers belonged to the molecular groups of "orHU", "PP", or "BC", and 83% were classified "CHO" formulae. Subtle differences in WSE-DOM composition between sites may reflect ecohydrological differences linked to the frequency of drying/ rewetting events (non-saturated/ saturated conditions), which was also suggested by water isotopic composition and radiocarbon data. We expected that valley WSE-DOM fingerprints would reflect in part the lateral flows from adjacent plateau environments, and the presence of large amounts of indicative N-containing markers with relatively high saturation, low oxidation and low molecular weight supports this assumption (compare **Figure 4b,c** and **e,f**). Large differences in DOM composition, however, indicate a loss of the plateau-derived DOM signature upon transit through the valley WSE system before groundwater enters the stream. This finding is remarkable because the riparian zone concentrates the water flux that is sourced from the surrounding plateaus (Miguez-Macho & Fan, 2012a). Simultaneously, this water flux drives the continuous export of young dissolved organic matter from the riparian valleys (Ledesma et al., 2015). However, permanent saturation, i.e., stagnating waters, seem to contribute to the preservation of a wide suite of organic compounds as opposed to the well-drained upland Campina site, and could thus explain the presence of unique, and mainly N-containing, valley WSE-DOM markers. Oxygen depletion in stagnating groundwater limits the turnover of reduced organic matter (Boye et al., 2017). The prevalence of low molecular weight (m/z < 450) N- and S-containing molecular formulae with a relatively high degree of saturation (H/C > 1) and a low degree of oxygenation (O/C < 0.6) could thus also be due to limited DOM uptake caused by unfavorable environmental conditions (low oxygen, low pH, high concentrations of phenolics; Bardy et al., 2011). In contrast, upland WSE-DOM contained unique, highly oxidized, aromatic, and phenolic CHO formulae (**Figure 4g-i**). These molecules likely represent the initial decomposition products of plant material containing large portions of lignin, cellulose, tannin, flavonoids, and terpenoids. Under non-stagnating, well-drained conditions, sandy soils favor the escape of such surface signals to streams because of the low sorption capacity of soil (Remington et al., 2007). However, fast export contrasts with older calibrated radiocarbon ages of DOC at the upland WSE (~11 years) as compared to the valley WSE site. This delay can be explained by slower litter turnover and reduced CO_2 efflux rates as discussed above (Zanchi et al., 2011, 2014). Highly acidic conditions, periods of drought, and an adapted plant community emerge as main drivers of such ecosystem-level differences. It is of special interest that fungi are known to remain active under dry and acidic conditions (Rousk et al., 2010; Vasco-Palacios et al., 2018). Fungal enzymes can potentially alter DOM towards higher-molecular weight, aromatic, and oxidized structures (Waggoner et al., 2015; Zavarzina et al., 2018), all of which seem to be occurring uniquely in upland WSE soils.

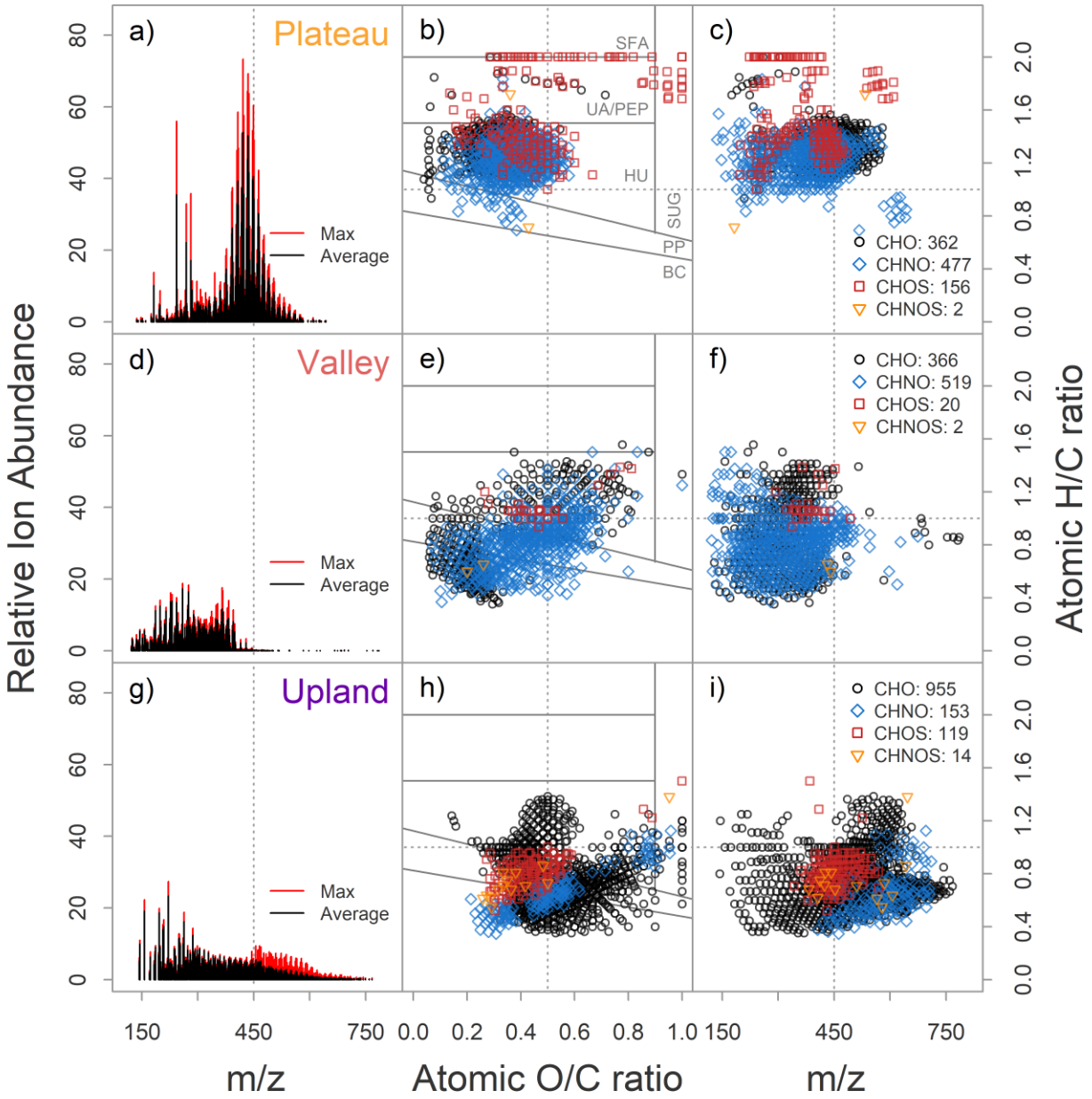

Figure 4. Subsets of molecular formulae showing significant "enrichment" (higher ion abundance), i.e., ecosystem specificity, in the plateau (panels a-c), valley WSE (d-f) and upland WSE samples (g-i). General WSE (valley and upland) markers and non-significant signals (common to all samples) are shown in **Supporting Information Figure S4**. Left column panels (**a, d, g**) show the average (and max) mass spectrum of each ecosystem. Mid column panels (**b, e, h**) show the formula subsets in Van Krevelen space (each dot represents a molecular formula defined by its atomic ratios of hydrogen, H/C, and oxygen to carbon, O/C). Formulae are colored according to classes (see legend and numbers of formulae in panels to the right). The plot is divided by solid lines that mark molecular group categories (see also caption of Figure 3, and **Supporting Information Table S2**). Right column panels (**c, f, i**) show the same formula subsets from mid-row panels in H/C vs. m/z (mass-to-charge) space. Dotted grey lines are for visual guidance and comparison (at m/z = 450, O/C = 0.5, and H/C = 1). Left- and right-column plots share the m/z abscissa (x-axis) while mid- and right-column plots share the H/C ordinate (y-axis).

3.3 Upland *Campina* WSEs are a potential source of indicative Rio Negro markers

We compared the sets of distinct DOM markers of the valley and upland WSEs to known Rio Negro markers in order to gain qualitative insight into the contribution of headwater aquifer and stream DOM to the "integrated" watershed signal of the lower Rio Negro (**Figure 5**).

Although many CHNO and CHOS formulae were part of the specific sets of signals in both WSEs, the overlap to Rio Negro markers was restricted to CHO formulae. The two WSEs differed in terms of potential molecular links. The cluster of heavy, highly oxidized, and aromatic upland WSE markers indicated consistent overlap with known Rio Negro markers (**Figure 5**). Overlapping sets of formulae appeared in a confined area of the Van Krevelen plot (boxes in **Figure 5b**, **d**; ranges: H/C 0.4 – 0.8, O/C 0.4 – 0.8), suggesting robust matching. However, the overlap was caused by different sets of Rio Negro markers, as indicated by the formula's *m/z* (**Figure 5c, f**). One study (Gonsior et al., 2016) revealed molecular links with markers of higher mass (m/z 350 – 650, **Figure 5c**), while another (Simon et al., 2019) found them in the lower mass range (m/z 200 – 500, **Figure 5f**). Nevertheless, all these markers were specific to the upland WSE and restricted to a common type of chemistry, and valley WSE markers contributed to a much lower degree.

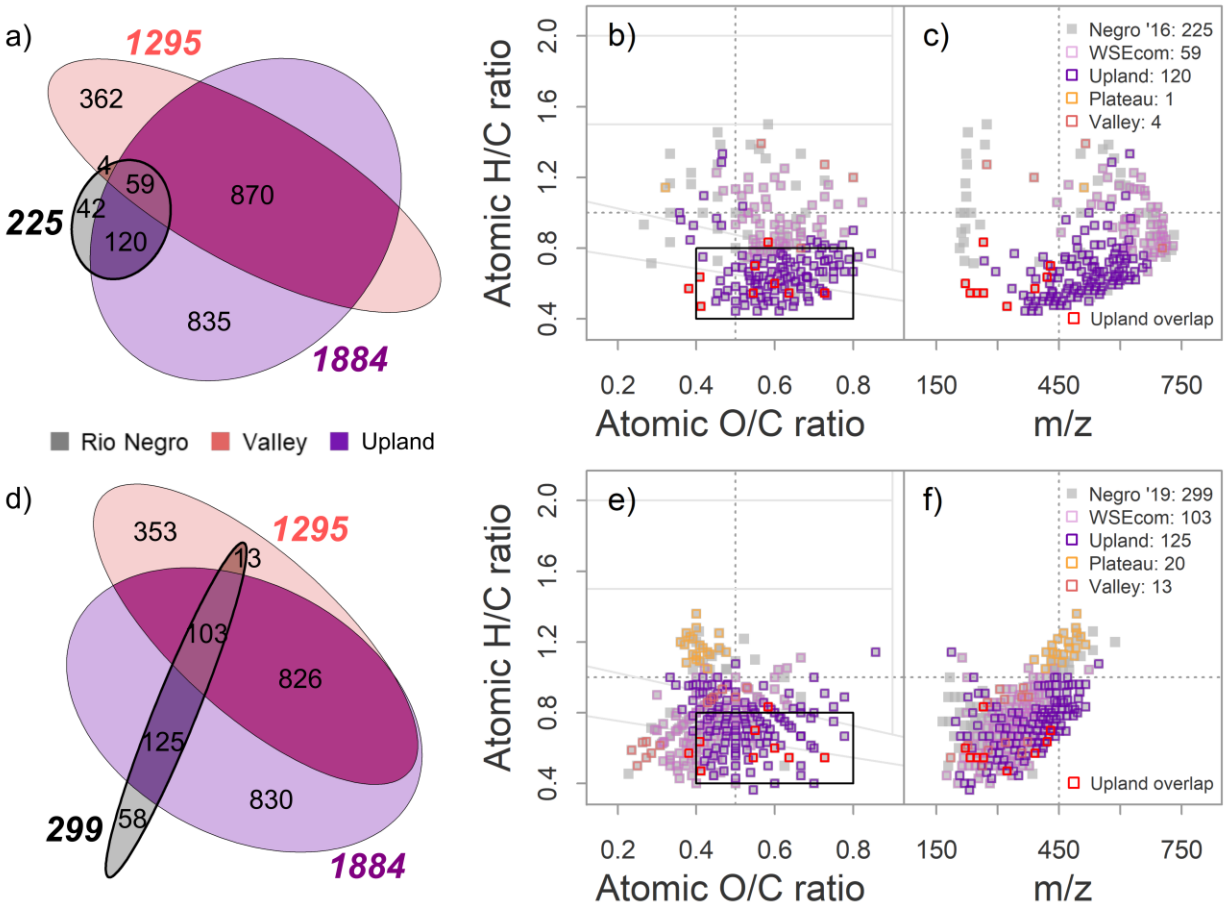

Figure 5. Overlap of WSE-specific formulae (significantly enriched formulae) with two independent sets of Rio Negro DOM markers (**a-c**: data from Gonsior et al., 2016; **d-f**: data from Simon et al., 2019). Left panels (a, d) show the overlap of three sets of markers (Rio Negro, Valley, Upland) in a Venn diagram created with package "eulerr" in RStudio. Overlap indicates common formulae, i.e., shared information. Areas are scaled to number of formulae, which are given for each subset (note color, area may be dissected). Overlap is shown only for CHO formulae because Rio Negro markers were mainly CHO-type, other formula classes showed no distinct overlap. Right panels show Rio Negro markers of each study (grey filled squares) in chemical space as a Van Krevelen plot (b, e) or H/C vs. m/z plot (c, f; similar visualization as in **Figure 4**). Colored symbols show the match between datasets ("overlap") and refer to the bold ellipse in panels a and d, respectively. Plateau markers are additionally added for comparison. Red squares denote the small set of WSE markers that were found to match with both sets of Rio Negro markers. They were uniquely enriched in upland (n = 6) or common to upland and valley WSE's (n = 3). Based on the chemistry of these

nine formulae, the black box denotes the wider area of consistent matching in terms of chemical space (H/C 0.4 – 0.8; O/C 0.4 – 0.8). Independent of the underlying set of Rio Negro markers, the matching rate was higher for upland WSE markers (lilac).

Differences in overlap of markers does not diminish the match in chemical space. Fine-tuning of measurement conditions can easily affect the instrument response in terms of ion abundance patterns and thus, *m/z* range (Hawkes et al., 2016; Hawkes et al., 2020; Simon et al., 2018). Despite sample set and lab/ instrument effects, measurement settings were relatively similar among the two studies and ours, thus supporting the general overlap of the upland WSE marker and Rio Negro marker sets in chemical space (oxidation and saturation state; **Supporting Information Text S1**). Accumulation time of ions before FT-MS analysis is one factor that explains *m/z* shifts (Hawkes et al., 2016; Simon et al., 2018). However, sampling criteria such as location, season, scale, and coverage will also have a strong influence on the resulting sets of markers. While it is not surprising to see large differences in marker sets (**Supporting Information Figure S8** and **Supporting Information Figure S9**), it is remarkable to find small but consistent overlap between studies, pointing toward subsets of markers with specific oxidation and saturation states. It is known that different instruments capture gradients of variation across a sample set even at very slight compositional differences (Hawkes et al., 2016), as may be expected when samples originate from similar environmental contexts as in our study. It is thus highly encouraging to find consistent and robust molecular overlap in three independent FT-MS datasets, being in line with long-standing hypotheses of landscape functioning in the Rio Negro basin (Goulding et al., 1988; Leenheer, 1980), namely that whitesand ecosystems are potential sources of indicative Rio Negro markers. Counter-intuitively, the highly indicative set of upland WSE markers were not only found to overlap with Rio Negro samples from the proximity of its draining higher-order river, the Rio Tarumã Açu (Simon et al., 2019) but also in samples upstream (Gonsior et al., 2016). This observation implies that similar high-molecular-weight oxidized aromatic compounds are exported upstream of Novo Airão, possibly by other upland, *Campina*-covered WSE systems. The DOM fingerprints of these other (extensive) WSE systems, especially in the upper Rio Negro basin (Adeney et al., 2016), remain to be revealed.

We queried the nine potential upland whitesand Rio Negro markers that consistently matched with two independent lists of Rio Negro markers ("Upland overlap", red squares in **Figure 5**) in PubChem to obtain potential structure suggestions (**Supporting Information Table S3**). A query with the ChEBI (chemical entities of biological interest) database yielded only two suggestions that were also found in PubChem. The structural suggestions thus represent only a first indication of potential structures and not necesarilly structures that were identified in DOM or organisms. Pubchem yielded up to 108 hits for formula $C_{10}H_6O_6$ and five at minimum for formula $C_{11}H_6O_8$ (numbers after exclusion of twelve ions, one hydrate and six stereoisomers from the list). In general, structures were highly oxidized, showing on average 2.5 carbonyl groups, 0.97 ether bonds, and 0.65 lactone groups per molecule across formulae. All formulae were classified as Black-Carbon like or polyphenol-like (BC, PP). On average, 68% of suggested structures per formula featured at least one heterocyclic ring, with five and six-membered rings dominating (each 30% on average). Functional group count was strongly correlated with mass (R^2 = 0.84, range = 5.2 – 9.7), same as numbers of carbon double bonds (0.86, 3.2 – 8.4) aromatic rings per molecule (0.87, 0.6 – 2.6) and thus, DBE (0.85, 8 – 16). Suggestions of smaller molecules (< 270 Da, with less than 15 C atoms) were by tendency dominated by naphthalene and chromene structures that are oxygen-poor scaffolds consisting of two rings. Larger molecules > 320 Da (with >15 C atoms) in contrast showed more hits involving at least one benzofuran, phenol or benzenecarboxylic scaffold, which all show a higher degree of

oxidation and functionalization. Chromenes and benzofurans are well known photo- and bioactive plant metabolites (Towers & Hudson, 1987). Knowledge of these trends and tendencies in molecular properties may help to develop tailored chromatographic techniques (LC-MS) for the targeted analysis of these biological relevant ecosystem markers in future (Petras et al., 2017). Structural data used for the above analyses is openly available as supporting data sets (**Supporting Information Data Set S1**, **Supporting Information Data Set S3**).

Groundwater DOM markers revealed a potential direct link between ecosystem-specific headwater and ecosystem-integrated downstream signals in the Rio Negro basin. The consistent overlap of upland WSE and Rio Negro markers was unexpected because our initial hypothesis, namely that typical valley WSE's will show stronger, or at least similar, matching with Rio Negro markers, had to be refused. Highly specific, tannin-like aromatic DOM compounds of the smaller upland WSE watershed represent a potential molecular link between headwaters and the Rio Negro. This finding opposes the assumed role of riparian valley WSEs as main sources of DOM in the Rio Negro watershed (Bardy et al., 2011; McClain et al., 1997). Although Bardy and coworkers showed that upland *Campina* forests areas produce a marked DOM signal traceable to rivers, others found that riparian and wetland WSEs along rivers contributed mainly to DOC exports of amazonian and other watersheds (Dosskey & Bertsch, 1994; McClain et al., 1997; Remington et al., 2007). Our results suggest that there is more detail to add to this simple model. Small watersheds with strong terrestrial-lotic linkages may leave a more significant downstream imprint as expected from absolute annual discharge, areal extent, or DOC export. Our results show that highly specific forest ecosystems potentially leave a distinct imprint within the Rio Negro's exported DOM, although contributing only secondarily to the overall fluxes of water, and maybe, DOC. The smaller upland WSE watershed at Reserva Campina shows a two times lower average annual stream discharge (based on the catchment area) but a two-three times higher annual DOC export (Monteiro et al., 2014; Zanchi et al., 2015). This discrepancy demonstrates that water and carbon cycles may be decoupled on the molecular level and the watershed scale, despite the generally accepted transport-limited, or chemostatic, DOC behavior (Musolff et al., 2017; Zarnetske et al., 2018). Ecosystems may show different behavior in terms of discharge and DOC export (Webb et al., 2018), or release of ecosystem markers (Wagner et al., 2019). Taking into account that hydromorphic soils in this region are thought to cover more than 40% of area (Junk, 1993), their missing imprint is evidence for the importance of unique ecosystems in control of indicative DOM characteristics within the land-to-ocean continuum. Riparian valley corridors likely contribute the major part of the annual DOC export in this region due to constant water supply from adjacent plateaus and their wider spatial distribution (Miguez-Macho & Fan, 2012a, 2012b; Remington et al., 2007), but do only contribute general DOM markers that likely reflect processes operating in many watersheds. In conclusion, our results indicate that riparian valleys may contribute less to the specific signal that discerns the Rio Negro from other watersheds on the molecular level. These specific markers show the potential for qualitative ecosystem recognition further downstream, and may thus serve as new proxies of land-derived primary production in the Rio Negro basin if properly calibrated.

4 Conclusions

This study investigated potential molecular links between indicative DOM markers at two main stages of the land-to-ocean-continuum, headwater catchments and river basins. New sets of markers are needed to better constrain variations in land-derived DOM exports and losses within aquatic systems, especially in the tropics, which account for 62% of riverine DOC

exports. We provide molecular evidence of long-assumed hypothetical links drawn from the simultaneous occurrence of tropical whitesand ecosystems (WSEs) and blackwater rivers at the example of the Rio Negro basin (Goulding et al., 1988, and references therein). Unique whitesand ecosytems contribute mainly to the fast export of DOC, while it is retained and decomposed in the highly productive, widespread plateau systems. Regardless of the wide occurrence of whitesand ecosystems and the known variation in WSE characteristics across the Amazon, there is little appreciation of this knowledge in models of DOC export and information content. This gap is due to a lack of qualitative DOC data. By using FT-MS techniques, we provide unique markers of whitesand ecosystems that can serve as future constraints on the terrestrial portion of aquatic DOC export. However, for such applications further calibration studies are necessary.

Against initial expectation, large valley WSEs may contribute only little to the flux of soluble ecosystem markers found at the river basin scale. In contrast, highly specific upland *Campina* WSEs emerged as their potential source. The twofold larger annual discharge and hundred-fold larger size of the Cuieiras watershed seemingly do not result in the transfer of DOM markers downstream, as expected from conservative (chemostatic) DOC leaching of dominant source layers. Instead, estimates suggest that the two-three times higher annual DOC export of the smaller upland watershed leaves a stronger imprint in terms of traceable WSE markers. This discrepancy stresses the importance of specific ecosystems for DOM information content at higher-order stages of the aquatic continuum, and how it may inform hydrological models that include qualitative DOC data. Future research needs to assess the influence of the large bandwidth of WSE types across the Rio Negro basin, and relate their environmental characteristics (climate, soil, vegetation) and geographical extent with information on quantity (DOC) and quality (molecular composition) of the exported DOM.

The molecular composition of DOM at different stages of the aquatic continuum emerges as a qualitative measure of DOM exports that complements quantitative DOC data. Our study shows the gained information that can be drawn from comparative FT-MS studies when complemented with other types of ecosystem information, especially information on the isotopic composition of water and organic carbon. Integration and nesting of DOM data show great prospects to bridge traditional gaps between soil science, limnology, and hydrology. The interconnected nature and multivariate complexity of DOM shows high potential to allow studying multiple source contributions and processes simultaneously. However, improvements in data integration across FT-MS platforms are required to reveal this information properly, and to reach robust conclusions about sources, fate, and identity of new sets of ecosystem markers in DOM. For this it will also be central to calibrate novel ecosystem markers with environmental variables such as DOC export, discharge, or ecosystem productivity.

Acknowledgments, Samples, and Data

We thank Axel Steinhof and Heike Machts for radiocarbon analyses and data, and Heiko Moossen for support in analyses of water isotopic composition and feedback on ealier versions of the manuscript. This work was accomplished in the framework of the Amazon Tall Tower Observatory (ATTO). We acknowledge funding and support from the Max-Planck-Gesellschaft (MPG), German Bundesministerium für Bildung und Forschung (BMBF), Brazilian Ministry of Science, Technology, Innovation and Communications (MCTIC), Amazonas State Foundation for Research (FAPEAM), Large-scale Biosphere-Atmosphere Experiment of Brazil's National

Institute for Amazon Research (LBA/ INPA), Uatumã Sustainable Development Reserve of Amazonas State's Secretariat of Sustainable Development (SDS/ CEUC/ RDS-Uatumã), and São Paulo Research Foundation (FAPESP). CS received a Ph.D. stipend from the International Max Planck Research School for Global Biogeochemical Cycles (IMPRS-gBGC). We are also grateful to Deutsche Forschungsgemeinschaft (DFG) for financial support as part of CRC 1076 "AquaDiva".

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Abbott, B. W., Gruau, G., Zarnetske, J. P., Moatar, F., Barbe, L., Thomas, Z., et al. (2018). Unexpected spatial stability of water chemistry in headwater stream networks. *Ecology Letters*, *21*, 296–308. https://doi.org/10.1111/ele.12897

Adeney, J. M., Christensen, N. L., Vicentini, A., & Cohn-Haft, M. (2016). White-sand ecosystems in Amazonia. *Biotropica*, *48*, 7–23.

Alvarez-Cobelas, M., Angeler, D. G., Sánchez-Carrillo, S., & Almendros, G. (2012). A worldwide view of organic carbon export from catchments. *Biogeochemistry*, *107*, 275–293. https://doi.org/10.1007/s10533-010-9553-z

Bardy, M., Derenne, S., Allard, T., Benedetti, M. F., & Fritsch, E. (2011). Podzolisation and exportation of organic matter in black waters of the Rio Negro (upper Amazon basin, Brazil). *Biogeochemistry*, *106*, 71–88. https://doi.org/10.1007/s10533-010-9564-9

Benk, S. A., Li, Y., Roth, V.-N., & Gleixner, G. (2018). Lignin Dimers as Potential Markers for 14C-young Terrestrial Dissolved Organic Matter in the Critical Zone. *Frontiers in Earth Science*, 1–9. https://doi.org/10.3389/feart.2018.00168

Bernhardt, E. S., Heffernan, J. B., Grimm, N. B., Stanley, E. H., Harvey, J. W., Arroita, M., et al. (2018). The metabolic regimes of flowing waters. *Limnology and Oceanography*, *63*, 99–118. https://doi.org/10.1002/lno.10726

Bowen, G. J., & Revenaugh, J. (2003). Interpolating the isotopic composition of modern meteoric precipitation. *Water Resources Research*, *39*(10), 1–13. https://doi.org/10.1029/2003WR002086

Cao, X., Aiken, G. R., Butler, K. D., Huntington, T. G., Balch, W. M., Mao, J., & Schmidt-Rohr, K. (2018). Evidence for major input of riverine organic matter into the ocean. *Organic Geochemistry*, *116*, 62–76. https://doi.org/10.1016/j.orggeochem.2017.11.001

Chassé, A. W., Ohno, T., Higgins, S. R., Amirbahman, A., Yildirim, N., & Parr, T. B. (2015). Chemical force spectroscopy evidence supporting the layer-by-layer model of organic matter binding to iron (oxy)hydroxide mineral surfaces. *Environmental Science and Technology*, *49*, 9733–9741. https://doi.org/10.1021/acs.est.5b01877

Coynel, A., Seyler, P., Etcheber, H., Meybeck, M., & Orange, D. (2005). Spatial and seasonal dynamics of total suspended sediment and organic carbon species in the Congo River. *Global Biogeochemical Cycles*, *19*(4), 1–17. https://doi.org/10.1029/2004GB002335

Creed, I. F., McKnight, D. M., Pellerin, B. A., Green, M. B., Bergamaschi, B. A., Aiken, G. R.,

et al. (2015). The river as a chemostat: fresh perspectives on dissolved organic matter flowing down the river continuum. *Canadian Journal of Fisheries and Aquatic Sciences*, *72*, 1272–1285. https://doi.org/10.1139/cjfas-2014-0400

Dai, M., Yin, Z., Meng, F., Liu, Q., & Cai, W.-J. (2012). Spatial distribution of riverine DOC inputs to the ocean: An updated global synthesis. *Current Opinion in Environmental Sustainability*, *4*, 170–178. https://doi.org/10.1016/j.cosust.2012.03.003

Demarchi, L. O., Scudeller, V. V., Moura, L. C., Dias-Terceiro, R. G., LOPES, A., Wittmann, F. K., & Piedade, M. T. F. (2018). Floristic composition, structure and soil-vegetation relations in three white-sand soil patches in central Amazonia. *Acta Amazonica*, *48*(1), 46–56. https://doi.org/10.1590/1809-4392201603523

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. *Limnology and Oceanography: Methods*, *6*, 230–235. https://doi.org/10.4319/lom.2008.6.230

Dosskey, M. G., & Bertsch, P. M. (1994). Forest sources and pathways of organic matter transport to a blackwater stream : A hydrologic approach. *Biogeochemistry*, *24*, 1–19.

Drake, T. W., Raymond, P. A., & Spencer, R. G. M. (2018). Terrestrial carbon inputs to inland waters: A current synthesis of estimates and uncertainty. *Limnology and Oceanography Letters*, *3*, 132–142. https://doi.org/10.1002/lol2.10055

Drake, T. W., Guillemette, F., Hemingway, J. D., Chanton, J. P., Podgorski, D. C., Zimov, N. S., & Spencer, R. G. M. (2018). The ephemeral signature of permafrost Carbon in an Arctic fluvial network. *Journal of Geophysical Research: Biogeosciences*, *123*, 1475–1485. https://doi.org/10.1029/2017JG004311

Drake, T. W., Van Oost, K., Barthel, M., Bauters, M., Hoyt, A. M., Podgorski, D. C., et al. (2019). Mobilization of aged and biolabile soil carbon by tropical deforestation. *Nature Geoscience*, *12*, 541–546. https://doi.org/10.1038/s41561-019-0384-9

Gehre, M., Geilmann, H., Richter, J., Werner, R. A., & Brand, W. A. (2004). Continuous flow 2H/1H and 18O/ 16O analysis of water samples with dual inlet precision. *Rapid Communications in Mass Spectrometry*, *18*, 2650–2660. https://doi.org/10.1002/rcm.1672

Gleixner, G., Poirier, N., Bol, R., & Balesdent, J. (2002). Molecular dynamics of organic matter in a cultivated soil. *Organic Geochemistry*, *33*, 357–366. https://doi.org/10.1016/S0146-6380(01)00166-8

Gonsior, M., Valle, J., Schmitt-Kopplin, P., Hertkorn, N., Bastviken, D., Luek, J., et al. (2016). Chemodiversity of dissolved organic matter in the Amazon Basin. *Biogeosciences*, *13*, 4279–4290. https://doi.org/10.5194/bg-13-4279-2016

Goulding, M., Carvalho, M. L., & Ferreira, E. G. (1988). Blackwaters. In M. Goulding, M. L. Carvalho, & E. G. Ferreira (Eds.), *Rio Negro, rich life in poor water. Amazonian diversity and foodchain ecology as seen through fish communities* (1st editio, pp. 29–36). The Hague: SPB Academic Publishing. https://doi.org/10.1016/0169-5347(89)90067-0

Guinoiseau, D., Bouchez, J., Gélabert, A., Louvat, P., Filizola, N., & Benedetti, M. F. (2016). The geochemical filter of large river confluences. *Chemical Geology*, *441*, 191–203. https://doi.org/10.1016/j.chemgeo.2016.08.009

Hawkes, J. A., Dittmar, T., Patriarca, C., Tranvik, L., & Bergquist, J. (2016). Evaluation of the Orbitrap Mass Spectrometer for the Molecular Fingerprinting Analysis of Natural Dissolved Organic Matter. *Analytical Chemistry*, *88*, 7698–7704. https://doi.org/10.1021/acs.analchem.6b01624

Hawkes, J. A., D'Andrilli, J., Agar, J. N., Barrow, M. P., Berg, S. M., Catalán, N., et al. (2020). An international laboratory comparison of dissolved organic matter composition by high resolution mass spectrometry: Are we getting the same answer? *Limnology and Oceanography: Methods, accepted*, doi: 10.1002/lom3.10364.

Hernes, P. J., Robinson, A. C., & Aufdenkampe, A. K. (2007). Fractionation of lignin during leaching and sorption and implications for organic matter "freshness." *Geophysical Research Letters*, *34*, 1–6. https://doi.org/10.1029/2007GL031017

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. *Biogeosciences*, *10*, 1583–1624. https://doi.org/10.5194/bg-10-1583-2013

Herzsprung, P., Hertkorn, N., von Tümpling, W., Harir, M., Friese, K., & Schmitt-Kopplin, P. (2014). Understanding molecular formula assignment of Fourier transform ion cyclotron resonance mass spectrometry data of natural organic matter from a chemical point of view. *Analytical and Bioanalytical Chemistry*, *406*(30), 7977–7987. https://doi.org/10.1007/s00216-014-8249-y

Huang, T. H., Fu, Y. H., Pan, P. Y., & Chen, C. T. A. (2012). Fluvial carbon fluxes in tropical rivers. *Current Opinion in Environmental Sustainability*, *4*, 162–169. https://doi.org/10.1016/j.cosust.2012.02.004

Hutchins, R. H. S., Aukes, P., Schiff, S. L., Dittmar, T., Prairie, Y. T., & del Giorgio, P. A. (2017). The Optical, Chemical, and Molecular Dissolved Organic Matter Succession Along a Boreal Soil-Stream-River Continuum. *Journal of Geophysical Research: Biogeosciences*, *122*, 2892–2908. https://doi.org/10.1002/2017JG004094

James, J. N., Gross, C. D., Dwivedi, P., Myers, T., Santos, F., Bernardi, R., et al. (2019). Land use change alters the radiocarbon age and composition of soil and water-soluble organic matter in the Brazilian Cerrado. *Geoderma*, *345*, 38–50. https://doi.org/10.1016/j.geoderma.2019.03.019

Janzen, D. H. (1974). Tropical blackwater rivers, animals, and mast fruiting by the Dipterocarpaceae. *Biotropica*, *6*, 69–103. https://doi.org/10.2307/2989823

Jasechko, S., & Taylor, R. G. (2015). Intensive rainfall recharges tropical groundwaters. *Environmental Research Letters*, *10*, 124015.

Jehn, F. U., Bestian, K., Breuer, L., Kraft, P., & Houska, T. (2020). Using hydrological and climatic catchment clusters to explore drivers of catchment behavior. *Hydrology and Earth System Sciences*, *24*, 1081–1100. https://doi.org/10.5194/hess-24-1081-2020

Junk, W. J. (1993). Wetlands of tropical South America. In D. F. Whigham, D. Dykyjová, & S. Hejný (Eds.), *Wetlands of the world: Inventory, ecology and management.* (Vol. 1, pp. 679–739). Heidelberg: Springer Science+Business Media B.V. https://doi.org/10.1007/978-94-015-8212-4_14

Junk, W. J., Piedade, M. T. F., Schöngart, J., Cohn-Haft, M., Adeney, J. M., & Wittmann, F. (2011). A classification of major naturally-occurring Amazonian Lowland Wetlands. *Wetlands*, *31*, 623–640. https://doi.org/10.1007/s13157-011-0190-7

Kellerman, A. M., Guillemette, F., Podgorski, D. C., Aiken, G. R., Butler, K. D., & Spencer, R. G. M. (2018). Unifying Concepts Linking Dissolved Organic Matter Composition to Persistence in Aquatic Ecosystems. *Environmental Science and Technology*, *52*, 2538–2548. https://doi.org/10.1021/acs.est.7b05513

Klinge, H., & Medina, E. (1979). Rio Negro caatingas and campinas, Amazonas States of Venezuela and Brazil. In R. L. Specht (Ed.), *Ecosystems of the World: Heathlands and related shrublands. Ecosystems of the World. A. Descriptive studies* (Vol. 9A, pp. 483–488). Amsterdam: Elsevier.

Kunert, N., Aparecido, L. M. T., Wolff, S., Higuchi, N., Dos Santos, J., De Araujo, A. C., & Trumbore, S. (2017). A revised hydrological model for the Central Amazon: The importance of emergent canopy trees in the forest water budget. *Agricultural and Forest Meteorology*, *239*, 47–57. https://doi.org/10.1016/j.agrformet.2017.03.002

Laudon, H., & Sponseller, R. A. (2018). How landscape organization and scale shape catchment hydrology and biogeochemistry: insights from a long-term catchment study. *Wiley Interdisciplinary Reviews: Water*, *5*, e1265. https://doi.org/10.1002/wat2.1265

Ledesma, J. L. J., Grabs, T., Bishop, K. H., Schiff, S. L., & Köhler, S. J. (2015). Potential for long-term transfer of dissolved organic carbon from riparian zones to streams in boreal catchments. *Global Change Biology*, *21*, 2963–2979. https://doi.org/10.1111/gcb.12872

Leenheer, J. A. (1980). Origin and nature of humic substances in the waters of the Amazon River Basin. *Acta Amazonica*, *10*, 513–526. https://doi.org/10.1590/1809-43921980103513

Leopoldo, P. R., Matsui, E., Salati, E., Franken, W., & Ribeiro, M. de N. G. (1982). Composição isotópica da água de chuva e da água do solo em floresta amazônica do tipo terra firme, região de Manaus. *Acta Amazonica*, *12*, 7–13. https://doi.org/10.1590/1809-43921982123s007

Li, M., Peng, C., Zhou, X., Yang, Y., Guo, Y., Shi, G., & Zhu, Q. (2019). Modeling global riverine DOC flux dynamics from 1951 to 2015. *Journal of Advances in Modeling Earth Systems*, *11*, 514–530. https://doi.org/10.1029/2018MS001363

Lucas, Y., Montes, C. R., Mounier, S., Loustau Cazalet, M., Ishida, D., Achard, R., et al. (2012). Biogeochemistry of an Amazonian podzol-ferralsol soil system with white kaolin. *Biogeosciences*, *9*, 3705–3720. https://doi.org/10.5194/bg-9-3705-2012

Luizão, R. C. C., Luizão, F. J., Paiva, R. Q., Monteiro, T. F., Sousa, L. S., & Kruijt, B. (2004). Variation of carbon and nitrogen cycling processes along a topographic gradient in a central Amazonian forest. *Global Change Biology*, *10*, 592–600. https://doi.org/10.1111/j.1529-8817.2003.00757.x

Lv, S., Yu, Q., Wang, F., Wang, Y., Yan, W., & Li, Y. (2019). A synthetic model to quantify dissolved organic carbon transport in the Changjiang River system: Model structure and spatiotemporal patterns. *Journal of Advances in Modeling Earth Systems*, *11*, 3024–3041. https://doi.org/10.1029/2019MS001648

Lynch, L. M., Sutfin, N. A., Fegel, T. S., Boot, C. M., Covino, T. P., & Wallenstein, M. D. (2019). River channel connectivity shifts metabolite composition and dissolved organic matter chemistry. *Nature Communications, 10*, 459.

Marques, J. D. de O., Libardi, P. L., Teixeira, W. G., & Reis, A. M. (2004). Estudo de parâmetros físicos, químicos e hídricos de um Latossolo Amarelo, na região Amazônica. *Acta Amazonica*, *34*, 145–154. https://doi.org/10.1590/s0044-59672004000200002

Marques, J. D. de O., Teixeira, W. G., Reis, A. M., Cruz Junior, O. F., Batista, S. M., & Afonso, M. A. C. B. (2010). Atributos químicos, físico-hídricos e mineralogia da fração argila em solos do Baixo Amazonas: Serra de Parintins. *Acta Amazonica*, *40*, 01–12. https://doi.org/10.1590/s0044-59672010000100001

Marques, J. D. de O., Luizão, F. J., Teixeira, W. G., Vitel, C. M., & Marques, E. M. de A. (2016). Soil organic carbon, carbon stock and their relationships to physical attributes under forest soils in central Amazonia. *Revista Árvore*, *40*(2), 197–208. https://doi.org/10.1590/0100-67622016000200002

Marschner, B., Brodowski, S., Dreves, A., Gleixner, G., Gude, A., Grootes, P. M., et al. (2008). How relevant is recalcitrance for the stabilization of organic matter in soils? *Journal of Plant Nutrition and Soil Science*, *171*, 91–110. https://doi.org/10.1002/jpln.200700049

Mayorga, E., Aufdenkampe, A. K., Masiello, C. a, Krusche, A. V, Hedges, J. I., Quay, P. D., et al. (2005). Young organic matter as a source of carbon dioxide outgassing from Amazonian rivers. *Nature*, *436*, 538–541. https://doi.org/10.1038/nature03880

McClain, M. E., Richey, J. E., Brandes, J. A., & Pimentel, T. P. (1997). Dissolved organic matter and terrestrial-lotic linkages in the central Amazon basin of Brazil. *Global Biogeochemical Cycles*, *11*(3), 295–311. https://doi.org/10.1029/97GB01056

McGuire, K. J., Torgersen, C. E., Likens, G. E., Buso, D. C., Lowe, W. H., & Bailey, S. W. (2014). Network analysis reveals multiscale controls on streamwater chemistry. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 7030–7035. https://doi.org/10.1073/pnas.1404820111

Medeiros, P. M., Seidel, M., Niggemann, J., Spencer, R. G. M., Hernes, P. J., Yager, P. L., et al. (2016). A novel molecular approach for tracing terrigenous dissolved organic matter into the deep ocean. *Global Biogeochemical Cycles*, *30*, 1–11. https://doi.org/10.1002/2015GB005320

Melack, J. M., & Hess, L. L. (2010). Remote Sensing of the Distribution and Extent of Wetlands in the Amazon Basin. In W. J. Junk, M. T. F. Piedade, F. Wittmann, J. Schöngart, & P. Parolin (Eds.), *Amazonian Floodplain Forests: Ecophysiology, Biodiversity and Sustainable Management. Ecological Studies 210* (pp. 43–59). Heidelberg: Springer Science+Business Media B.V. https://doi.org/10.1007/978-90-481-8725-6

Miguez-Macho, G., & Fan, Y. (2012a). The role of groundwater in the Amazon water cycle: 1. Influence on seasonal streamflow, flooding and wetlands. *Journal of Geophysical Research Atmospheres*, *117*, D15113. https://doi.org/10.1029/2012JD017539

Miguez-Macho, G., & Fan, Y. (2012b). The role of groundwater in the Amazon water cycle: 2. Influence on seasonal soil moisture and evapotranspiration. *Journal of Geophysical Research Atmospheres*, *117*, D15114. https://doi.org/10.1029/2012JD017540

Monteiro, M. T. F., Oliveira, S. M., Luizão, F. J., Cândido, L. A., Ishida, F. Y., & Tomasella, J. (2014). Dissolved organic carbon concentration and its relationship to electrical conductivity in the waters of a stream in a forested Amazonian blackwater catchment. *Plant Ecology and Diversity*, *7*, 205–213. https://doi.org/10.1080/17550874.2013.820223

Montes, C. R., Lucas, Y., Pereira, O. J. R., Achard, R., Grimaldi, M., & Melfi, A. J. (2011). Deep plant-derived carbon storage in Amazonian podzols. *Biogeosciences*, *8*, 113–120. https://doi.org/10.5194/bg-8-113-2011

Moore, S., Evans, C. D., Page, S. E., Garnett, M. H., Jones, T. G., Freeman, C., et al. (2013). Deep instability of deforested tropical peatlands revealed by fluvial organic carbon fluxes. *Nature*, *493*, 660–663. https://doi.org/10.1038/nature11818

Moyer, R. P., Bauer, J. E., & Grottoli, A. G. (2013). Carbon isotope biogeochemistry of tropical small mountainous river, estuarine, and coastal systems of Puerto Rico. *Biogeochemistry*, *112*, 589–612. https://doi.org/10.1007/s10533-012-9751-y

Musolff, A., Fleckenstein, J. H., Rao, P. S. C., & Jawitz, J. W. (2017). Emergent archetype patterns of coupled hydrologic and biogeochemical responses in catchments. *Geophysical Research Letters*, *44*, 4143–4151. https://doi.org/10.1002/2017GL072630

Do Nascimento, N. R., Bueno, G. T., Fritsch, E., Herbillon, A. J., Allard, T., Melfi, A. J., et al. (2004). Podzolization as a deferralitization process: A study of an Acrisol-Podzol sequence derived from Palaeozoic sandstones in the northern upper Amazon Basin. *European Journal of Soil Science*, *55*, 523–538. https://doi.org/10.1111/j.1365-2389.2004.00616.x

Do Nascimento, N. R., Fritsch, E., Bueno, G. T., Bardy, M., Grimaldi, C., & Melfi, A. J. (2008). Podzolization as a deferralitization process: Dynamics and chemistry of ground and surface waters in an Acrisol - Podzol sequence of the upper Amazon Basin. *European Journal of Soil Science*, *59*, 911–924. https://doi.org/10.1111/j.1365-2389.2008.01049.x

Newcomb, C. J., Qafoku, N. P., Grate, J. W., Bailey, V. L., & De Yoreo, J. J. (2017). Developing a molecular picture of soil organic matter-mineral interactions by quantifying organo-mineral binding. *Nature Communications*, *8*, 396. https://doi.org/10.1038/s41467-017-00407-9

Oksanen, J. (2010). Multivariate analysis of ecological communities in R: vegan tutorial. https://doi.org/10.1016/0169-5347(88)90124-3

Osterholz, H., Singer, G., Wemheuer, B., Daniel, R., Simon, M., Niggemann, J., & Dittmar, T. (2016). Deciphering associations between dissolved organic molecules and bacterial communities in a pelagic marine system. *ISME Journal*, *10*, 1717–1730. https://doi.org/10.1038/ismej.2015.231

Petras, D., Koester, I., Da Silva, R., Stephens, B. M., Haas, A. F., Nelson, C. E., et al. (2017). High-resolution liquid chromatography tandem mass spectrometry enables large scale molecular characterization of dissolved organic matter. *Frontiers in Marine Science*, *4*(December), 406. https://doi.org/10.3389/fmars.2017.00405

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. *Global Biogeochemical Cycles*, *33*, 1423–1439. https://doi.org/10.1029/2019GB006225

Quesada, C. A., Lloyd, J., Anderson, L. O., Fyllas, N. M., Schwarz, M., & Czimczik, C. I. (2011). Soils of Amazonia with particular reference to the RAINFOR sites. *Biogeosciences*, *8*, 1415–1440. https://doi.org/10.5194/bg-8-1415-2011

Raeke, J., Lechtenfeld, O. J., Tittel, J., Oosterwoud, M. R., Bornmann, K., & Reemtsma, T. (2017). Linking the mobilization of dissolved organic matter in catchments and its removal in drinking water treatment to its molecular characteristics. *Water Research*, *113*, 149–159. https://doi.org/10.1016/j.watres.2017.01.066

Raymond, P. A., & Spencer, R. G. M. (2014). Riverine DOM. In D. A. Hansell & C. A. Carlson (Eds.), *Biogeochemistry of Marine Dissolved Organic Matter: Second Edition* (2nd ed., pp. 509–533). Cambridge: Academic Press. https://doi.org/10.1016/B978-0-12-405940-5.00011-X

Regnier, P., Friedlingstein, P., Ciais, P., Mackenzie, F. T., Gruber, N., Janssens, I. A., et al. (2013). Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nature Geoscience*, *6*, 597–607. https://doi.org/10.1038/ngeo1830

Remington, S. M., Strahm, B. D., Neu, V., Richey, J. E., & Da Cunha, H. B. (2007). The role of sorption in control of riverine dissolved organic carbon concentrations by riparian zone soils in the Amazon basin. *Soil Science*, *172*, 279–291. https://doi.org/10.1097/ss.0b013e318032ab46

Riedel, T., & Dittmar, T. (2014). A Method Detection Limit for the Analysis of Natural Organic Matter via Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.*, *86*, 8376–82. https://doi.org/10.1021/ac501946m

Riedel, T., Zark, M., Vähätalo, A. V., Niggemann, J., Spencer, R. G. M., Hernes, P. J., & Dittmar, T. (2016). Molecular Signatures of Biogeochemical Transformations in Dissolved Organic Matter from Ten World Rivers. *Frontiers in Earth Science*, *4*, 85. https://doi.org/10.3389/feart.2016.00085

Roth, V.-N., Dittmar, T., Gaupp, R., & Gleixner, G. (2014). Ecosystem-specific composition of dissolved organic matter. *Vadose Zone Journal*, *13*. https://doi.org/http://dx.doi.org/10.2136/vzj2013.09.0162

Roth, V.-N., Lange, M., Simon, C., Hertkorn, N., Bucher, S., Goodall, T., et al. (2019). Persistence of dissolved organic matter explained by molecular changes during its passage through soil. *Nature Geoscience*, *12*, 755–761. https://doi.org/10.1038/s41561-019-0417-4

Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., et al. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *ISME Journal*, *4*, 1340–1351. https://doi.org/10.1038/ismej.2010.58

Seyler, F., Muller, F., Cochonneau, G., Guimarães, L., & Guyot, J. L. (2009). Watershed delineation for the Amazon sub-basin system using GTOPO30 DEM and a drainage network extracted from JERS SAR images. *Hydrological Processes*, *23*(22), 3173–3185. https://doi.org/10.1002/hyp

Simon, C., Roth, V.-N., Dittmar, T., & Gleixner, G. (2018). Molecular Signals of Heterogeneous Terrestrial Environments Identified in Dissolved Organic Matter: A Comparative Analysis of Orbitrap and Ion Cyclotron Resonance Mass Spectrometers. *Frontiers in Earth Science*, *6*(September), 1–16. https://doi.org/10.3389/feart.2018.00138

Simon, C., Osterholz, H., Koschinsky, A., & Dittmar, T. (2019). 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). *Organic Geochemistry*, *129*, 45–62. https://doi.org/10.1016/j.orggeochem.2019.01.013

Sioli, H. (1954). Gewässerchemie und Vorgänge in den Böden im Amazonasgebiet. *Die Naturwissenschaften*, *41*(19), 456–457.

Spencer, R. G. M., Kellerman, A. M., Podgorski, D. C., Macedo, M. N., Jankowski, K. J., Nunes, D., & Neill, C. (2019). Identifying the Molecular Signatures of Agricultural Expansion in Amazonian Headwater Streams. *Journal of Geophysical Research: Biogeosciences*, 1637–1650. https://doi.org/10.1029/2018JG004910

Steinhof, A., Altenburg, M., & Machts, H. (2017). Sample preparation at the Jena 14C laboratory. *Radiocarbon*, *59*, 815–830. https://doi.org/10.1017/RDC.2017.50

Tiegs, S. D., Costello, D. M., Isken, M. W., Woodward, G., McIntyre, P. B., Gessner, M. O., et al. (2019). Global patterns and drivers of ecosystem functioning in rivers and riparian zones. *Science Advances*, *5*, 1–9. https://doi.org/10.1126/sciadv.aav0486

Tomasella, J., Hodnett, M. G., Cuartas, L. A., Nobre, A. D., Waterloo, M. J., & Oliveira, S. M. (2007). The water balance of an Amazonian micro-catchment: The effect of interannual variability of rainfall on hydrological behaviour. *Hydrological Processes*, *22*, 2133–2147. https://doi.org/10.1002/hyp.6813

Towers, G. H. N., & Hudson, J. B. (1987). Potentially useful antimicrobial and antiviral phototoxins from plants. *Photochemistry and Photobiology*, *46*, 61–66. https://doi.org/10.1111/j.1751-1097.1987.tb04736.x

Trumbore, S. E., Sierra, C. A., & Hicks Pries, C. E. (2016). Radiocarbon nomenclature, theory, models, and interpretation: measuring age, determining cycling rates, and tracing source pools. In E. A. G. Schuur, E. R. M. Druffel, & S. E. Trumbore (Eds.), *Climate, Radiocarbon and Change* (1st ed., pp. 45–82). Basel: Springer.

Vasco-Palacios, A. M., Hernandez, J., Peñuela-Mora, M. C., Franco-Molano, A. E., & Boekhout, T. (2018). Ectomycorrhizal fungi diversity in a white sand forest in western Amazonia. *Fungal Ecology*, *31*, 9–18. https://doi.org/10.1016/j.funeco.2017.10.003

Waggoner, D. C., Chen, H., Willoughby, A. S., & Hatcher, P. G. (2015). Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. *Organic Geochemistry*, *82*, 69–76. https://doi.org/10.1016/j.orggeochem.2015.02.007

Wagner, S., Fair, J. H., Matt, S., Hosen, J. D., Raymond, P. A., Saiers, J., et al. (2019). Molecular hysteresis: Hydrologically driven changes in riverine dissolved organic matter chemistry during a storm event. *Journal of Geophysical Research: Biogeosciences*, *124*(4), 759–774. https://doi.org/10.1029/2018JG004817

Ward, N. D., Keil, R. G., Medeiros, P. M., Brito, D. C., Cunha, A. C., Dittmar, T., et al. (2013). Degradation of terrestrially derived macromolecules in the Amazon River. *Nature Geoscience*, *6*, 530–533. https://doi.org/10.1038/ngeo1817

Waterloo, M. J., Oliveira, S. M., Drucker, D. P., Nobre, A. D., Cuartas, L. A., Hodnett, M. G., et

al. (2006). Export of organic carbon in run-off from an Amazonian rainforest blackwater catchment. *Hydrological Processes*, *20*, 2581–2597.

Webb, J. R., Santos, I. R., Maher, D. T., & Finlay, K. (2018). The importance of aquatic carbon fluxes in net ecosystem carbon budgets: A catchment-scale review. *Ecosystems*, *22*, 508–527.

Zanchi, F. B., Waterloo, M. J., Dolman, A. J., Groenendijk, M., Kesselmeier, J., Kruijt, B., et al. (2011). Influence of drainage status on soil and water chemistry, litter decomposition and soil respiration in central Amazonian forests on sandy soils. *Revista Ambiente e Agua*, *6*, 6–29. https://doi.org/10.4136/1980-993X

Zanchi, F. B., Meesters, A. G. C. A., Waterloo, M. J., Kruijt, B., Kesselmeier, J., Luizão, F. J., & Dolman, A. J. (2014). Soil CO2 exchange in seven pristine Amazonian rain forest sites in relation to soil temperature. *Agricultural and Forest Meteorology*, *192–193*, 96–107. https://doi.org/10.1016/j.agrformet.2014.03.009

Zanchi, F. B., Waterloo, M. J., Tapia, A. P., Alvarado Barrientos, M. S., Bolson, M. A., Luizão, F. J., et al. (2015). Water balance, nutrient and carbon export from a heath forest catchment in central Amazonia, Brazil. *Hydrological Processes*, *29*(17), 3633–3648. https://doi.org/10.1002/hyp.10458

Zarnetske, J. P., Bouda, M., Abbott, B. W., Saiers, J., & Raymond, P. A. (2018). Generality of hydrologic transport limitation of watershed organic Carbon flux across ecoregions of the United States. *Geophysical Research Letters*, *45*, 11702–11711. https://doi.org/10.1029/2018GL080005

Zavarzina, A. G., Lisov, A. V., & Leontievsky, A. A. (2018). The role of ligninolytic enzymes laccase and a versatile peroxidase of the white-rot fungus Lentinus tigrinus in biotransformation of soil humic matter: Comparative in vivo study. *Journal of Geophysical Research: Biogeosciences*, *123*, 2727–2742. https://doi.org/10.1029/2017JG004309

Zhang, X.-P., Yang, Z.-L., Niu, G.-Y., & Wang, X.-Y. (2009). Stable water isotope simulation in different reservoirs of Manaus, Brazil, by Community Land Model incorporating stable isotopic effect. *International Journal of Climatology*, *29*, 619–628. https://doi.org/10.1002/joc