JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2022JG007126

Particulate Organic Matter Mobilization and Transformation Along a Himalayan River Revealed by ESI-FT-ICR-MS

Johanna Menges[1,2], Niels Hovius[1,3], Stefanie Poetz[1], Helena Osterholz[4,5], and Dirk Sachse[1]

[1]GFZ German Research Centre for Geosciences, Potsdam, Germany, [2]MARUM - Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany, [3]Department of Geosciences, University of Potsdam, Potsdam, Germany, [4]Leibniz Institute for Baltic Sea Research Warnemuende, Rostock, Germany, [5]Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky University, Oldenburg, Germany

Key Points:

- Organic matter sourcing and transformations in a Himalayan river studied by FT-ICR-MS measurements of solvent extractable lipids
- Identification of up to 10^2 indicator mass formulas for different organic matter sources in the landscape using indicator species analysis
- Mass formulas preserved during incorporation of litter into topsoil but selectively lost during entrainment of sources into the river

Correspondence to:
J. Menges,
jmenges@marum.de

Citation:
Menges, J., Hovius, N., Poetz, S., Osterholz, H., & Sachse, D. (2022). Particulate organic matter mobilization and transformation along a Himalayan river revealed by ESI-FT-ICR-MS. *Journal of Geophysical Research: Biogeosciences*, *127*, e2022JG007126. https://doi.org/10.1029/2022JG007126

Received 4 AUG 2022
Accepted 22 NOV 2022

Author Contributions:
Conceptualization: Johanna Menges, Niels Hovius, Stefanie Poetz, Dirk Sachse
Data curation: Johanna Menges
Formal analysis: Johanna Menges, Helena Osterholz
Funding acquisition: Niels Hovius, Dirk Sachse
Investigation: Johanna Menges
Project Administration: Dirk Sachse
Resources: Stefanie Poetz, Dirk Sachse
Software: Helena Osterholz
Supervision: Niels Hovius, Stefanie Poetz, Dirk Sachse
Visualization: Johanna Menges
Writing – original draft: Johanna Menges

Abstract Tracing pathways and transformations of particulate organic carbon from landscape sources to oceanic sinks is commonly done using the isotopic composition or biomarker content of particulate organic matter (POM). However, similarity of source characteristics and complex mixing in rivers often preclude a robust deconvolution of individual contributions. Moreover, these approaches are limited in detecting organic matter transformations. This impedes understanding of carbon cycling. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) can simultaneously identify many molecular formulas from mixtures of organic matter, and provide direct information on its compositional variability. Here, we investigate how FT-ICR-MS can give insight into POM dynamics on a landscape scale, focusing on the trans-Himalayan Kali Gandaki River, Nepal. Using molecular information, we identify source tracers in the solvent extractable lipid fraction of riverine POM, finding up to 102 indicative molecular formulas for individual sources. Further, we assess molecular transformations of the lipid fraction of POM during its transfer from litter into topsoil, and onwards into the river. A large number of shared mass formulas and a well-preserved isoprenoidal patterns suggest efficient incorporation of litter into topsoil. In contrast, we observe a selective loss of mass formulas and a preferential export of formulas with low double bond equivalents and a low nominal oxidation state of carbon after organic matter entrainment in the river. Our results demonstrate the potential of FT-ICR-MS for source-to-sink studies, allowing detailed organic matter source characterization and discrimination, and tracking of molecular transformations along organic matter pathways spanning different spatial and temporal scales.

Plain Language Summary The transfer of organic matter (OM) by rivers from landscape sources into the ocean followed by its burial in marine sediments is an important carbon sink. Therefore, OM is often traced along this journey using its isotopic or biomarker composition. But contributions of OM sources to river sediments can be difficult to estimate because of similar source characteristics, mixing of many sources and changes of the molecular composition along the way. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) is a novel method able to identify many molecular formulas from OM mixtures at once providing direct information about their molecular composition. Here, we investigate how FT-ICR-MS contributes to understanding the transport and transformation of particulate OM focusing on a Himalayan river in Nepal. We use the molecular information to identify tracers for individual OM sources in the landscape. We then assess molecular transformations during the transfer of litter into topsoil, and onwards into the river. Our data suggest efficient incorporation of litter into topsoil, but we observe a selective loss of molecular formulas upon entrainment of sources into the river. Our results reveal that FT-ICR-MS is useful for detailed source characterization and tracking of molecular transformations along OM pathways.

1. Introduction

On geological timescales, organic carbon fluxes into and out of rock, soil and biological reservoirs influence carbon dioxide concentrations in the atmosphere, modulating global climate. One important pathway within this global carbon cycle is the transfer of particulate organic carbon (POC) via erosion and riverine transport from terrestrial sources to ocean basins (Hedges, 1992). This transfer constitutes a geological carbon sink if the buried organic carbon is recently assimilated plant or soil organic carbon (OC_{bios}) (Berner, 1982; Derry & France-Lanord, 1997; Galy et al., 2007; Hilton et al., 2008). In contrast, the reburial of rock derived petrogenic organic carbon (OC_{petro}) does not affect the atmospheric carbon content (Galy, Beyssac, et al., 2008; Hilton et al., 2011; Sparkes et al., 2020) but its remineralization constitutes a carbon source to the atmosphere

Writing – review & editing: Johanna Menges, Niels Hovius, Stefanie Poetz, Helena Osterholz, Dirk Sachse

(Hedges, 1992; Petsch, 2014). However, most of the POC never reaches the ocean as it is either stored in intermediate reservoirs (e.g., Galy & Eglinton, 2011b) or remineralized in transit. Remineralization of OC_{bios} can occur within soils (Bellamy et al., 2005; Oades, 1988), during riverine transport (Cole et al., 2007) or floodplain storage (Scheingross et al., 2021) and in the ocean (Blair & Aller, 2012; Ittekkot, 1988). Similarly, it has recently been shown that OC_{petro} can be rapidly incorporated into soil organic matter in an active mountain range where it is more accessible for remineralizing processes (Hemingway et al., 2018). To understand which components of the organic matter pool at the Earth's surface reach the ocean and are eventually buried, it is thus essential to trace the sourcing and transport as well as transformation and remineralization along the complex path of organic matter from source to sink.

Lipids form an important part of the organic matter in plants and soils and play a significant role during the incorporation of plant material into soil organic matter (Kögel-Knabner, 2002). Due to their recalcitrance (resistance to biotic decomposition), lipids are preserved in marine sediments and ancient sedimentary rocks (e.g., Brocks et al., 1999), forming a part of the organic carbon that is transferred between the different reservoirs. In addition, lipids carry important information along the continent-ocean continuum: lipid biomarkers have been used extensively for reconstruction of paleo-environmental conditions (e.g., Bai et al., 2009; Rach et al., 2014; Zachos et al., 2006) and for determination of organic matter sources (e.g., Feng et al., 2013; Galy et al., 2011; Hoffmann et al., 2016; Ponton et al., 2014). Changes in lipid composition and a partial replacement during transfer have been observed in rivers (Freymond et al., 2018; Galy, France-Lanord, & Lartiges, 2008) and the arctic ocean (Bröder et al., 2018), highlighting that their source-to-sink pathways must be assessed with care. In this context, there is high interest in the identification of additional biomarkers, associated with previously untraceable organic matter sources such as for example, deep soil organic carbon, and in complementing established biomarkers so that different fractions of organic matter and their transformations can be traced with higher confidence. However, common measurements of lipid biomarkers via gas or high-performance liquid chromatography-mass spectrometry only allow the detection of compounds within a narrow mass range after several steps of extraction, separation and purification which leads to a significant time investment for each analysis (e.g., Goñi & Montgomery, 2000; Rach et al., 2020). These factors prevent the rapid and systematic screening for new potential biomarkers, as well as an assessment of a larger suite of lipid compounds.

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) has emerged as a powerful tool to resolve previously "uncharacterizable" natural organic matter on a molecular level (Marshall & Rodgers, 2004; Reemtsma, 2009). Combined with electrospray ionization (ESI), a low-fragmentation "soft" ionization method able to generate molecular ions also for large molecules (Kebarle & Verkerk, 2009), the extraordinary mass accuracy of FT-ICR-MS allows the assignment of discrete molecular formulas to hundreds to thousands of molecular ions without prior chromatographic separation (e.g., Hughey et al., 2002). In contrast to a targeted approach where a compound of interest is isolated in the laboratory prior to analysis, the large data sets of molecular formulas (MFs) yielded by single FT-ICR-MS measurements on complex mixtures can be mined for a range of marker compounds (Bell & Blais, 2019). However, since such MF data sets inform only about the exact mass of compounds, they often do not allow the determination of the exact molecular structure. Therefore, detected MFs alone cannot unambiguously confirm the presence of a compound. Instead, the exact structure of present compounds must be determined in a second step, using fragmentation mass spectrometry after compound isolation (Bell & Blais, 2019). Thus, FT-ICR-MS data from complex mixtures cannot be used directly to detect biomarkers and reveal their exact structure. Rather, this approach constitutes a powerful tool for an initial screening to identify marker MFs and giving direction to further analysis. In addition, as FT-ICR-MS is capable of resolving compounds at pico-to nanomolar concentrations for example, in dissolved organic matter (Repeta, 2015), it can indicate the presence of compounds in low concentrations, which would have to be concentrated prior to the analysis using conventional methods.

ESI-FT-ICR-MS measurements have been used to characterize crude oils (Cho et al., 2015; Qian et al., 2001; Teräväinen et al., 2007), solvent extractable organic matter in sedimentary rocks (Poetz et al., 2014), oceanic and freshwater dissolved organic matter (Hertkorn et al., 2006; Koch et al., 2005; Qi et al., 2022; Sleighter et al., 2009; Wagner et al., 2015), humic and fulvic acids (Koch et al., 2007; Stenson et al., 2003) and soil organic matter (Tfaily et al., 2015, 2017), and its potential to screen for marker compounds has been highlighted recently (Bell & Blais, 2019). Previously, ESI-FT-ICR-MS has been used successfully to identify marker compounds for marine dissolved organic matter (Kujawinski et al., 2009) and trace dissolved terrigenous compounds toward the ocean (Medeiros et al., 2016). FT-ICR-MS has also been used to detect transformation of dissolved organic matter

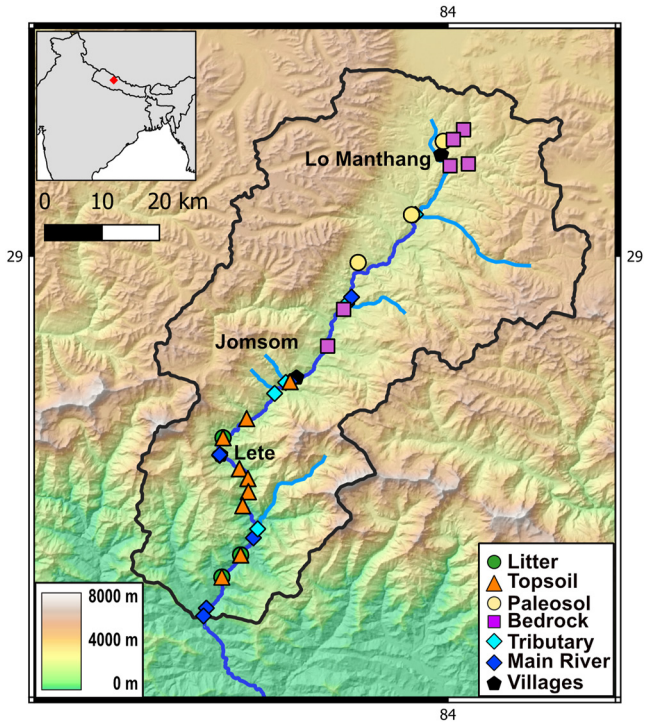

Figure 1. Overview of study area. Hillshaded digital elevation model of the upper KG valley with sample locations: litter (green), modern soils (orange), paleosols (yellow), bedrock (violet), tributary streams (light blue), main river (dark blue).

during decomposition in lake water (Kellerman et al., 2015; Mostovaya et al., 2017), in floodplains (Boye et al., 2017) and during sediment incubation (Pracht et al., 2018).

Here, we employ FT-ICR-MS for the first time to trace the lipid component of particulate organic matter through a natural landscape. For this study, we selected a well-studied trans-Himalayan river catchment that hosts several organic matter sources ranging from active, vegetated soils and paleosols (Menges et al., 2019; Saijo & Tanaka, 2002) to organic rich black shales (Garzanti, 1999). These different sources of modern and aged OC_{bios} as well as OC_{petro} have been shown to get mobilized by the Kali Gandaki River based on bulk isotope analyses (Galy & Eglinton, 2011a; Menges et al., 2020). We performed ESI-FT-ICR-MS measurements on solvent extractable lipids of the particulate organic matter from a sample set of these modern and aged OC_{bios} and OC_{petro} sources and a sample set of the sediments transported by the river and its tributaries. We then used the obtained data set twofold: First, we screened the data set for source specific markers that can trace organic matter through different reservoirs. Second, we investigated potential changes in molecular composition along transformation pathways including the transfer of organic matter from litter into topsoil and the uptake and transfer of different types organic matter by the river system.

2. Materials and Methods

2.1. Study Area

The Kali Gandaki (KG) River in central Nepal is an important tributary of the Ganges river system with a catchment of ~12,000 km^2 and an elevation gradient from 6,000 to 200 masl (meter above sea level). The KG River originates on the Tibetan plateau and cuts through all the major units of the Himalayan orogen. This study focuses on the upper KG valley (4,550 km^2) connecting the edge of the Tibetan Plateau with the High Himalayan mountain range (Figure 1). About 80% of the annual precipitation is provided by the Indian Summer Monsoon reaching the central Himalayas between June and September leading to a high seasonality of water and sediment discharge (Andermann et al., 2012). The high relief of the High Himalayan mountain range blocks the transfer of moisture (Bookhagen & Burbank, 2010) leading to a precipitation gradient from around 2,000 mm/y in front of the orographic barrier to around 250 mm/y in the uppermost river reaches (Andermann et al., 2011). The distribution of vegetation mirrors the elevation and climatic gradients with dense deciduous mountain forest transitioning upward into coniferous forest and then dwarf shrubland in the dry Tibetan part of the catchment. Soils in the High Himalayan part of the catchment, dominated by deciduous and coniferous forest, are predominantly Cambisols ranging from chromic over gleyic to eutric (Dijkshoorn & Huting, 2009). Soils do not currently form in the central part of the dry Northern valley, but abundant paleosol horizons have been reported (Menges et al., 2019; Saijo & Tanaka, 2002). Geologically, the upper part of the KG catchment consists of a graben structure filled with Neogene and Quaternary sediments (Adhikari & Wagreich, 2011; Fort et al., 1982) and underlain by Paleozoic to Mesozoic marine shales of the Tethyan Sedimentary Sequence (Bordet et al., 1971; Garzanti, 1999; Godin, 2003). The Paleozoic section of the Tethyan Sedimentary Sequence comprises massive limestones as well as calcareous shales and pelites (Garzanti, 1999), while the Mesozoic formations consist of Triassic calcareous shales, Jurassic limestones and black shales, and Lower Cretaceous conglomerates and sandstones (Bordet et al., 1971; Garzanti, 1999). The Lupra Formation, Bagung Formation, and Jomsom Formation of this succession contain black shales rich in organic matter, ranging from 0.2% to 2.6% total organic carbon (Menges et al., 2020) and outcrop mainly in the Northern part of the upper catchment (Garzanti, 1999). The High Himalaya of the KG catchment consists of the Greater Himalayan Sequence of para- and orthogneiss, in the lower part (Larson & Godin, 2009; Stocklin, 1980), and migmatites and leucogranites in the upper part (Hodges et al., 1996; Le Fort, 1975). The high peaks of the Dhaulagiri and Annapurna are capped by massive and strongly recrystallized carbonates of the Tethyan Sedimentary Sequence (Bordet et al., 1981).

The KG River rises in the Tibetan part of the catchment as a braided stream occupying a string of extensive braid plains surrounded by sparsely vegetated, gentle hills that are partly covered by paleosol deposits. It then reaches the coniferous forest at around 2,800 masl and from a knickpoint at around 2,500 masl, the KG River traverses the High Himalaya carving a narrow and steep gorge through metamorphic rocks of the Greater Himalayan Sequence between the Dalaughiri and Annapurna. Analyses of bulk stable carbon and nitrogen and radioactive carbon isotopes in the particulate organic matter exported from the KG valley show that OC_{petro} derived from the Jurassic black shales in the upper valley dominate the POC load with secondary inputs from topsoils and paleosols (Galy & Eglinton, 2011a; Menges et al., 2020). Since the KG River is fast flowing, turbulent and sediment charged, riverine production was suggested to be negligible and no changes reflecting decomposition were detected using bulk organic carbon isotope data (Menges et al., 2020).

2.2. Sampling

For this study we used 53 samples from a larger sample set collected in the upper KG catchment during consecutive monsoon seasons in 2015 and 2016 representing different OC sources and river sediments as described in Menges et al. (2020) (Figure 1). Specifically, we used 16 samples representative of OC_{bios} including 3 litter, 10 topsoil and 3 paleosol samples, and 9 samples representing OC_{petro} from bedrock to approximate the major organic matter sources. To trace the sourcing and export of particulate organic matter from the catchment we used 28 river sediment samples, 18 from the main river and 10 from 5 different tributaries. In short, river sediment samples comprise materials collected from river banks and suspended load. Suspended load samples were filtered on the same day through PES filters (pore size 0.2 µm). Bedrock was sampled from outcrops of black shales of the Lupra, Bagung and Jomsom Formations (Tethyan Sedimentary Sequence) within the main river valley. Paleosol samples were collected in the central river valley from visibly organic-rich horizons of paleosol sections and paleosol remnants (Menges et al., 2019). Vegetated topsoils were sampled along the KG River between Jomsom (2,759 masl) and Galeshwar (887 masl). Topsoils (0–10 or 20 cm depth) were sampled after the removal of the litter/humus layer and at three locations (NP15-7, NP16-45 and NP16-47) the litter layer was sampled separately (Menges et al., 2020). Litter, soil and rock samples were dried immediately after return to the lab in Potsdam at 60°C in a drying oven. Sediment samples were frozen at −18°C until drying at 60°C in the drying oven. All samples were ground in a vibratory disk mill using a steel grinding set which was thoroughly cleaned with water, ethanol and acetone between samples. Rock samples were covered with combusted aluminum foil to avoid contamination and crushed with a hammer to a size suitable for the grinding mill. One rock sample (NPCK-4) was crushed in a rock crusher as it was too hard for manual crushing.

2.3. Analytical Methods

The lipid fraction was extracted from ground, solid samples using a modified Soxhlet procedure and organic solvents. Between 25 and 40 g of ground and homogenized sample were weighed into a pre-extracted cellulose tube and covered with glass wool (pre-extracted with dichloromethane). Each cellulose tube was then placed in a clean Soxhlet extractor and extracted with 200 ml of dichloromethane:methanol (99:1, v:v) at 50% power for 24 hr. This extraction method is commonly applied in biomarker studies (Bell & Blais, 2019). After extraction the solvent was evaporated and extract yields were determined gravimetrically. Extracts were then diluted with methanol:toluene (1:1, v:v) to a final concentration of 100 µg/ml. 10 µl of a concentrated aqueous NH_3 solution was added to each sample to facilitate negative ion formation.

Mass spectrometric analyses were performed on a 12 T FT-ICR mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany) applying electrospray ionization in negative ion mode (Apollo II ESI source, Bruker) as described in Poetz et al. (2014) at GFZ Potsdam, Germany. External mass calibration was done by using an external calibration mixture for ESI in negative mode containing fatty acids and modified polyethylene glycols (Poetz et al., 2014). Each mass spectrum was then internally calibrated using compounds that contain two or three oxygen atoms. Detected masses of calibrated spectra were matched across all samples and MFs were assigned within the mass range of 156–900 Da to peaks above twice the method detection limit (Riedel & Dittmar, 2014) with a mass error <0.5 ppm allowing $C_{0–100}H_{0–200}O_{0–20}N_{0–2}S_{0–2}Na_{0–1}$ at ICBM, Carl von Ossietzky University Oldenburg, Germany. Na^+ was included into the formula assignment since it is known to form adducts in negative ESI mode (Schug & McNair, 2002). The measured spectra of three rock samples and one river sediment sample that showed low total monoisotopic abundances were not processed further. All other samples were normalized to

their total monoisotopic abundances to create a relative abundance data set. We then generated a second data set which only indicated the presence or absence of MFs within each sample by assigning 1 to identified MFs and 0 to MFs which were not present in that sample. We refer to this transformation of the data as the presence-absence data set.

2.4. Statistical Processing

2.4.1. Molecular Metrics

While the determination of a specific molecular structure is not possible using broadband ESI-FT-ICR-MS measurements, there are a number of molecular metrics that can be calculated directly from the MF. First, the number of carbon atoms in a molecule provides information about its size.

Then, MFs of the samples were grouped based on their heteroatomic content into five different groups: CHO, CHON, CHN, CHSO and CHNSO. Further, frequently used ratios are the number of oxygen over carbon atoms within each MF (O/C ratio) and the equivalent for hydrogen (H/C ratio). Using these parameters, the compositional variability can be presented in van Krevelen diagrams (Kim et al., 2003).

Double bond equivalents (DBEs) quantify the degree of unsaturation in a molecule (Pellegrin, 1983) and were calculated from the number of atoms within each MF as follows:

$$\mathrm{DBE} = 1 + \tfrac{1}{2}(2\,\#\mathrm{C} - \#\mathrm{H} + \#\mathrm{N} + \#\mathrm{P})$$

where # denotes the number of atoms present of the elements C, H, N, or P, respectively.

An additional proxy obtainable from a MF without further structural information is the nominal oxidation state of carbon (NOSC), which describes the average oxidation state of the carbon atoms within a molecule (LaRowe & Van Cappellen, 2011). The NOSC of organic compounds has been linked to the Gibbs energies of half reactions describing the complete mineralization of these compounds (LaRowe & Van Cappellen, 2011). The oxidation of organic compounds with a higher NOSC is on average more thermodynamically profitable and therefore preferred if all other parameters are equal (LaRowe & Van Cappellen, 2011). Hence, compounds with high NOSC values are generally interpreted to have low stability, and compounds with low NOSC values are interpreted to be relatively stable. We have calculated the NOSC (LaRowe & Van Cappellen, 2011) using the following expression, adjusted for the elements allowed during formula assignment:

$$\mathrm{NOSC} = -((-z + 4(\#\mathrm{C}) + (\#\mathrm{H}) - 3(\#\mathrm{N}) - 2(\#\mathrm{O}) - (\#\mathrm{S}))/(\#\mathrm{C})) + 4$$

where # denotes the number of atoms present of the elements C, H, N, respectively and z corresponds to the net charge of the measured ion.

To test for significant differences of molecular metrics for groups of samples, we verified whether the assumptions for a one-way ANOVA were met using the Shapiro Wilk normality test (p-value <0.05) and the Levene's test (p-value <0.05) for the homogeneity of variances. None of the compared groups had equal variances. If the data was normally distributed, we used a Welch ANOVA followed by a Games-Howell test for pairwise comparison. If the first assumption of normality was also not met, we performed the non-parametric Kruskal-Wallis test followed by a Dunn's test for pairwise comparison with adjusted p-values using the Bonferroni correction method.

To assess the compositional relation between the different samples, we performed non-metric multi-dimensional scaling (NMDS) based on Jaccard Dissimilarities for presence-absence data (Figure 4a, stress = 0.112) and Bray-Curtis Dissimilarities for relative abundance data (Figure 4b, stress = 0.142). Jaccard Dissimilarities place equal weight on all MFs whereas Bray-Curtis Dissimilarities place more weight on highly abundant MFs. The bedrock sample JM26 was excluded from NMDS because of its dissimilarity to all other measured samples (see Section 3.1).

2.4.2. Indicator Species Analysis

To identify potential marker compounds for the different source types (Section 3.3) we used Indicator Species Analysis (ISA), which allows detection of marker compounds frequently present and sufficiently abundant within one specific group (Dufrêne & Legendre, 1997). ISA was developed by Dufrêne and Legendre (1997) for analysis of ecological species assemblages and applied to FT-ICR-MS data of marine and soil dissolved organic matter

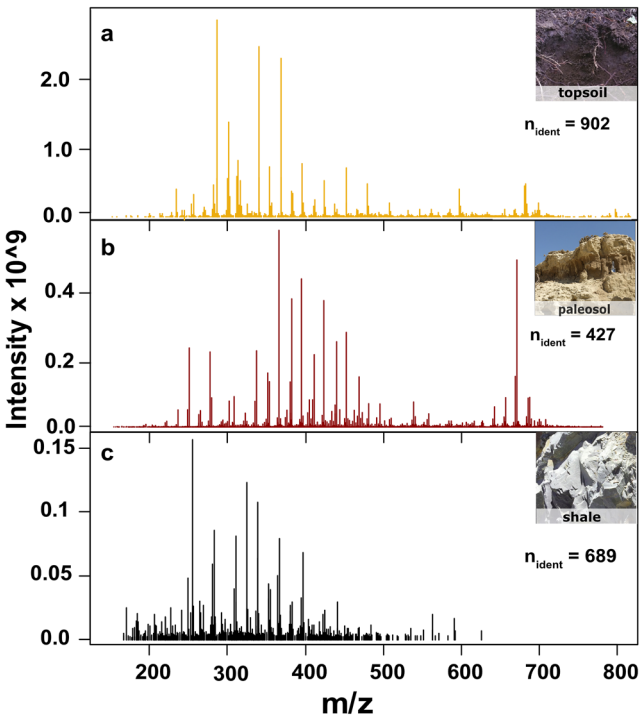

Figure 2. Reproduced mass spectra for the measured solvent extractable lipid fractions of topsoil sample NP16_24PS (a), paleosol sample NP16_24PS (b) and bedrock sample NPCK04 (c) with the number of identified molecular formulas per sample (n_{ident}).

(Kujawinski et al., 2009; Ohno et al., 2016). ISA calculates the indicator value of species as the product of the relative frequency and relative average abundance in predefined groups (Dufrêne & Legendre, 1997). We employed ISA to identify indicator MFs for modern OC_{bios} in the form of topsoils, aged OC_{bios} in the form of paleosol samples and OC_{petro}. These OC types have been shown to be distinguishable based on their bulk carbon and nitrogen isotopic composition and are actively exported by the KG River (Menges et al., 2020). ISA was performed using the multipatt function in the indicspecies package (De Cáceres et al., 2010; De Cáceres & Legendre, 2009, ver. 1.7.8) on the presence-absence and then the relative abundance MF data set. *P*-values were calculated based on 999 randomizations, indicator MFs with a p-value below 0.05 and an indicator value above 0.7 were accepted. We then examined the occurrence of the identified indicator MFs in the river sediment data set.

2.4.3. Molecular Richness

To assess the compositional differences between source types and river sediment samples in more detail (see Section 4.3), we aggregated the assigned MFs from all samples belonging to each sample type into one composite list of MFs representing the full molecular diversity of that sample type. As a measure for molecular diversity we used the molecular richness, which is the number of different MFs detected within one sample type, similar to species richness as an indicator of biodiversity in ecological studies (Legendre & Legendre, 2012). Since ranges of distributions are often dominated by outliers, we used the 95% range when reporting them.

3. Results and Discussion

The lipid extracts of river sediments and different source materials from the KG catchment analyzed via FT-ICR-MS revealed a highly variable composition between sample types (e.g., Figure 2). For the whole data set consisting of 49 samples with a sufficient total monoisotopic abundance (see methods), a total of 4,111 MFs were detected, with 117–1,398 MFs per sample (mean = 506 ± 251) (see Menges et al., 2022).

3.1. Molecular Composition of Organic Matter in Different Sample Types

The OC_{bios} samples included topsoil, paleosol and litter samples. Of these, fewer MFs were detected in paleosol samples compared to topsoil and litter (Table 1). All OC_{bios} samples were dominated by CHO over CHON MFs. In addition, litter samples and 8 topsoil samples contained a minor number of CHOS MFs (<2%). Regarding average molecular metrics such as single molecule O/C and H/C, DBEs, and NOSC (Figure 3), we observed a clear trend from litter to topsoil and paleosol samples. Litter samples were characterized by a mean C number of 27.7 ± 0.3 and a mean O number of 4.5 ± 0.6, resulting in an O/C ratio of 0.17 ± 0.02 and a NOSC value of −1.31 ± 0.10. Paleosol samples had a significantly higher mean C number of 31.5 ± 1.3 and a lower O number of 3.4 ± 0.4 resulting in a significantly lower O/C ratio of 0.11 ± 0.01 and a lower NOSC of −1.64 ± 0.01. Topsoil samples were slightly more variable than paleosol samples but fell broadly between litter and paleosols. They had a mean C number of 29.3 ± 0.5 and an O number of 4.1 ± 0.4 resulting in an O/C ratio of 0.15 ± 0.02 and a NOSC of −1.41 ± 0.16. All litter, topsoil and paleosol samples shared 120 MFs representing 32%–90% of the cumulative relative abundance, rendering the OC_{bios} samples the most coherent group of endmembers.

By comparison, the number and type of detected MFs in the six bedrock samples from the Jurassic black shale formations representing OC_{petro} were more diverse (Table 1, Figure 3). These samples shared only 38 MFs representing between 4% (R26, see below) and 66% of the cumulative relative abundance. Three samples (R15b, R01 and R02) contained only low numbers of assigned MFs, ranging from 171 to 290, while higher numbers of MFs (647–689) were assigned for the other three (R04, R16, R26). Bedrock samples were markedly dominated by CHO MFs with minor contributions of CHNO MFs (2%–14%) and of CHSO MFs (2%–20%). Five of the six bedrock samples were characterized by high O/C ratios ranging from 0.14 to 0.23. H/C ratios and NOSC of

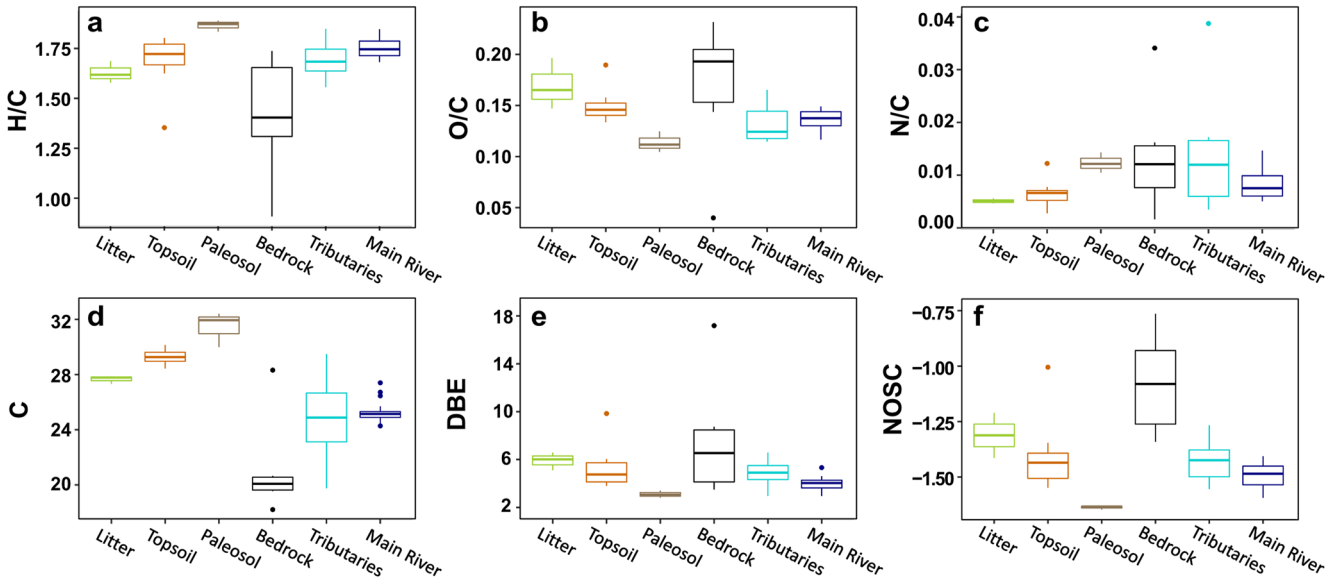

Figure 3. Boxplots of non-weighted mean values for (a) H/C, (b) O/C, (c) N/C ratio, as well as number of (d) C, (e) DBE, and (f) NOSC of the molecular formulas for each sample grouped by sample material.

these five bedrock samples were variable ranging from 1.3 to 1.7 for H/C and from −1.4 to −0.9 for NOSC. The nitrogen-dominated bedrock sample JM26 stood out, with a higher mean C-number of 28 and a low average O/C ratio of 0.04, since most identified formulas were oxygen free. This sample had about 40% each of CHN MFs and CHNO MFs, a very low H/C ratio, probably due to the presence of highly aromatic compounds such as carbazoles which are typical constituents of fossil organic matter (Hughey et al., 2002), and a higher NOSC (−0.8) than the other bedrock samples.

Main river and tributary sediment extracts were dominated by CHO-containing MFs with contribution of CHON MFs, with the exception of sample 16–23F. This sample was dominated by CHNO MFs (63%) with a secondary contribution of CHO (35%). The average molecular metrics of tributary sediment samples fell between those of litter, soil and bedrock samples, but showed a high variability (Figure 3). The main river sediment samples fell within the range of the tributary samples with decreased variability for all parameters except NOSC, which was lower in some main river samples than in any of the tributary samples.

The clear separation of average molecular metrics of the different sample types (e.g., topsoils and bedrock) indicates that the analytical method is suited to the detection of differences in molecular composition between the sampled organic matter pools. In addition, gradual compositional changes from litter to topsoil and paleosol samples provides evidence for potential transformation over time or during the transfer of organic matter between them.

3.2. Relation of the Molecular Composition Between Sample Types

Since FT-ICR-MS data contains hundreds to thousands of MFs per sample, it is necessary to reduce the information so that the molecular composition can be visualized and interpreted. Therefore, we conducted non-metric multi-dimensional scaling (NMDS) to evaluate the similarity in molecular composition of lipids among the different sampled organic matter pools (Figure 4). We conducted this analysis with dissimilarity matrices calculated for the presence-absence data and also for the relative abundance MF data sets (see methods). While analysis based on relative abundance data places more weight on the highly abundant MFs, using the presence-absence data considers each MF with equal weight. In the corresponding ordination space, all sample types formed distinct clusters, except the tributary samples. The first axis (second for presence-absence data) clearly separated OC_{petro} and OC_{bios} samples with river sediment samples falling largely between these source types. The second axis (first for presence-absence data) further separated litter, topsoil and paleosol samples within OC_{bios} while OC_{petro} samples formed a broader cluster. Both NMDS analyses revealed generally similar groupings of the organic

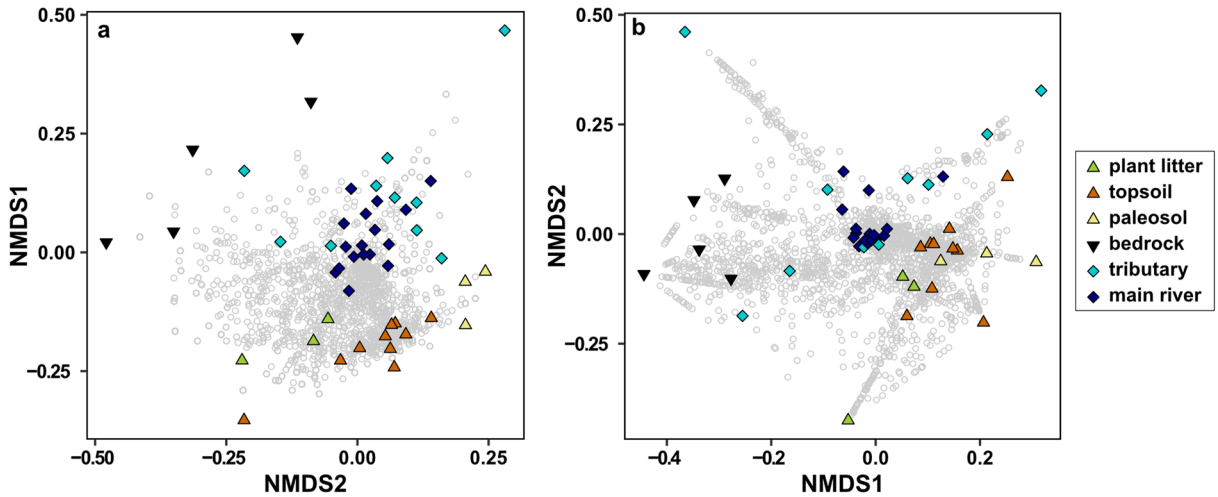

Figure 4. Multivariate analysis of molecular formulas (MFs) using non-metric multidimensional scaling (NMDS). NMDS was based on Jaccard dissimilarities using presence-absence data (a, stress = 0.112) and Bray-Curtis dissimilarities using relative abundance data (b, stress = 0.142). MFs from all samples are represented by gray circles and analyzed samples are colored by sample type.

matter extracts, suggesting that the differentiation between the different sample types is not only evident in the highly abundant MFs but also in the larger number of MFs with low abundances. Only for OC_{petro} and main river samples the results based on presence-absence data showed a greater variability, indicating that in these pools the less abundant MFs were more dissimilar. The position of the cluster of sediment samples from the main river at equal distances from all source types suggests contributions from the sources to river sediments consistent with previous work (Menges et al., 2020). In contrast, tributary samples showed a larger variability with several

Table 1
Molecular Composition of the Lipid Fraction by Source

	Biospheric			Petrogenic	Sediment	
	Litter (n = 3)	Topsoil (n = 10)	Paleosol (n = 3)	Bedrock (n = 6)	Tributaries (n = 10)	Main river (n = 17)
All MF	679–1,083	571–1,398	323–602	117–689	246–474	197–797
	(823 ± 226)	(792 ± 239)	(451 ± 141)	(433 ± 269)	(328 ± 86)	(421 ± 146)
CHO	547–928	380–1,262	184–362	80–616	88–415	138–514
	(689 ± 208)	(636 ± 246)	(279 ± 90)	(286 ± 222)	(232 ± 96)	(319 ± 109)
CHON	103–127	111–205	132–235	11–264	16–160	30–264
	(116 ± 12)	(150 ± 29)	(167 ± 59)	(67 ± 100)	(81 ± 49)	(89 ± 50)
CHN	0–2	0–4	1–3	0–284	0–17	0–3
	(1.0 ± 1.0)	(1.3 ± 1.1)	(1.0 ± 2.0)	(114 ± 62)	(4.9 ± 5.6)	(0.4 ± 0.8)
CHOS	3–18	0–20	0–3	2–25	1–17	4–19
	(10.7 ± 7.5)	(3.8 ± 5.9)	(1.3 ± 1.5)	(17 ± 9)	(8.9 ± 4.1)	(12.0 ± 3.9)
DBE	5.90 ± 0.75	5.29 ± 1.78	3.10 ± 0.3	7.70 ± 5.10	4.89 ± 1.09	4.00 ± 0.57
O/C	0.17 ± 0.02	0.15 ± 0.02	0.11 ± 0.01	0.17 ± 0.07	0.13 ± 0.02	0.14 ± 0.01
H/C	1.63 ± 0.05	1.69 ± 0.13	1.87 ± 0.03	1.41 ± 0.31	1.69 ± 0.09	1.75 ± 0.05
m/z	451.3 ± 4.45	469.6 ± 6.4	495.7 ± 25.3	337.3 ± 30.3	393.3 ± 44.9	403.6 ± 13.6
NOSC	−1.31 ± 0.10	−1.41 ± 0.16	−1.64 ± 0.01	−1.08 ± 0.23	−1.43 ± 0.09	−1.49 ± 0.06

Note. Given are ranges (mean ± sd) for biospheric (n = 16), petrogenic (n = 6) and sediment (n = 27) samples for the total number of detected molecular formulas and subgroups CHO, CHON, CHN, CHOS. Mean ± sd are provided for DBE, NOSC and molar ratios (O/C, H/C) and *m/z*.

Table 2
Number and Molecular Composition of Indicator Molecular Formulas (MFs) for Particulate Organic Matter Sources for Presence-Absence (pa) and Relative Abundance (rel. ab.) Data Sets

	Topsoils		Paleosols		Bedrock		Topsoil + paleosol	
	pa	rel. ab.	pa	rel. ab.	pa	rel. ab.	pa	rel. ab.
n	103	83	63	92	47	78	240	221
CHO	81	64	17	24	26	52	150	141
CHON	22	19	43	64	3	7	90	80
CHN	-	-	2	2	-	-	-	-
CHOS	-	-	-	1	18	19	-	-
DBE	5.0 ± 4.1	5.2 ± 4.4	3.5 ± 4.1	3.2 ± 3.8	4.1 ± 3.1	5.0 ± 2.9	3.2 ± 2.6	3.0 ± 2.5
O/C	0.16 ± 0.07	0.16 ± 0.06	0.10 ± 0.09	0.10 ± 0.08	0.23 ± 0.01	0.22 ± 0.09	0.11 ± 0.05	0.11 ± 0.05
H/C	1.73 ± 0.28	1.74 ± 0.27	1.86 ± 0.30	1.88 ± 0.26	1.59 ± 0.41	1.50 ± 0.37	1.88 ± 0.16	1.88 ± 0.16
NOSC	−1.43 ± 0.32	−1.43 ± 0.29	−1.62 ± 0.40	−1.65 ± 0.35	−1.12 ± 0.35	−1.07 ± 0.32	−1.65 ± 0.19	−1.66 ± 0.19
m/z	517 ± 130	528 ± 116	547 ± 161	558 ± 149	295 ± 91	292 ± 89	532 ± 105	529 ± 106

Note. Given are numbers for all detected indicator MFs and subgroups CHO, CHON, CHN, CHOS. Mean ± sd are provided for DBE, NOSC and molar ratios (O/C, H/C) and *m/z*.

samples defining trends away from any documented source type, indicating that the catchment may contain additional organic matter sources that have not been sampled in this study, such as for example, aquatic production in very small streams and ponds.

3.3. Fingerprinting of Organic Matter Sources

The finding of distinct clusters of organic matter sources and sediment samples in multivariate space suggests a high potential for organic matter tracing. Similar to biomarker compounds that can be traced back to their source organism even after traveling long distances (e.g., Goñi et al., 2008), a multi-dimensional molecular fingerprint based on FT-ICR-MS data may provide insights into OC transport and transformation processes.

To test for potential groups of such indicator compounds we performed indicator species analysis (ISA) (Dufrêne & Legendre, 1997). ISA identifies MFs that are frequently present with a sufficiently high abundance within a pre-defined group of end-member samples and are thus characteristic for this group (Dufrêne & Legendre, 1997). We note that when using the MF data set, different chemical compounds sharing the same molecular formula cannot be distinguished and the exact structure of a characteristic compound cannot be determined. We therefore focus on groups of MFs which have similar structural characteristics such as the number of DBE's or heteroatoms.

We performed ISA for the three different sources of organic matter to the river that were previously shown to get eroded based on bulk isotopes. These sources comprise (a) modern OC_{bios} derived from topsoils, (b) pre-aged OC_{bios} derived from paleosols, and (c) OC_{petro} derived from bedrock (Menges et al., 2020). Bedrock sample R26 was excluded based on the highly different composition, dominated by CHN MFs. The identified indicator MFs are therefore linked to CHO dominated bedrock samples.

ISA identified indicator MFs for the three sources and for a combination of topsoils and paleosols (Table 2). This fourth group represents MFs that are common in topsoils and paleosols but not in bedrock, therefore representing ubiquitous compounds in OC_{bios}. Indicator MFs for the different organic matter sources occupy different regions in the van Krevelen plots (Figure 5) with indicator MFs for paleosols having highest H/C and lowest O/C values followed by indicator MFs for topsoils. Indicator MFs for bedrock form two clusters, one with H/C values below 1.5 and the other with H/C values higher than 2. Mean molecular metrics of groups of indicator MFs varied significantly between several different sources. Indicator MFs for OC_{petro} had significantly higher O/C ratios, lower NOSC and lower *m/z* ratios than all other groups. They also had a significantly lower H/C ratio than indicator MFs for paleosols and ubiquitous OC_{bios}, but not topsoils. Within the OC_{bios} data set, the topsoil indicator MFs had significantly different DBE, H/C, O/C, and NOSC values than the indicator MFs for paleosols and indicator MFs for topsoils and paleosols combined.

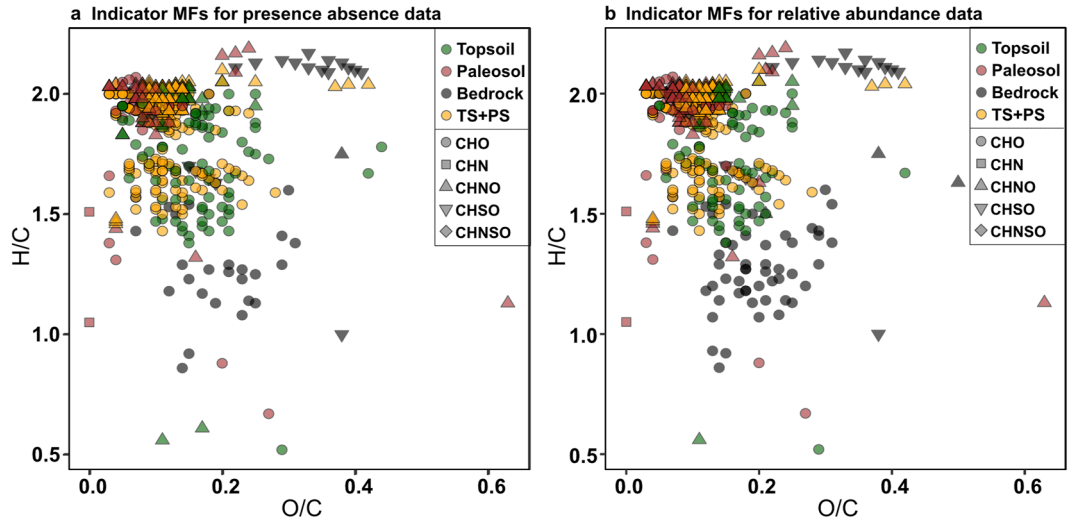

Figure 5. van Krevelen plots of indicator molecular formulas for different particulate organic matter sources for (a) presence-absence data and (b) relative abundance data set. Colors indicate the different organic matter sources (green-Topsoil, red-Paleosol, black- Bedrock, yellow-Topsoil and Paleosol combined) and symbols related to heteroatomic content.

Further examination of the different heteroatomic groups yielded a number of differences: Indicator MFs for topsoils and paleosols, representing ubiquitous compounds in OC_{bios}, were mainly O_2 to O_4 MFs as well as N_1O_4 and N_1O_5 MFs with 1–3 DBEs. More specifically, topsoil indicator MFs had high oxygen content (mainly $O_{3\text{-}5}$), often with high DBEs ranging from 6 to 9 and C-numbers above 27. Indicator MFs for paleosols were dominated by N_1O_2 to N_1O_5 MFs (38 MFs) with 1–6 DBEs. In addition, 17 single CHO MFs belonging to O_1 to O_6 were identified as paleosol indicator MFs. Indicator MFs for OC_{petro} belonged mainly to the O_2 to O_5 groups (24 MFs) and were characterized by low C-numbers between 10 and 22 and mainly high DBEs between 4 and 10. A second group of OC_{petro} indicator MFs comprised $S_1O_{4\text{-}9}$ MFs with C-numbers ranging from 12 to 22 and no DBEs (13 MFs) and $S_1O_{3\text{–}4}$ MFs with 4–5 DBEs (5 MFs). While the exact numbers of identified indicator MFs differed between the presence-absence and relative abundance data sets, the characteristic composition of indicator MFs for each source was remarkably similar. Given the ionization effects of ESI, we restricted the further investigation to indicator MFs derived from the presence-absence data set.

Next, we determined how many of the indicator MFs for each organic matter source could be detected in each river sediment sample. Previous work on particulate organic matter mobilization using bulk isotopes showed that all organic matter sources considered in this study are exported by the KG River (Galy & Eglinton, 2011a; Menges et al., 2020). We found between 2% and 45% of indicator MFs for modern OC_{bios} in samples from the main river, with a mean of 20 ± 11%, between 0% and 32% of indicator MFs for pre-aged OC_{bios} with a mean of 6 ± 6%, and between 5% and 75% of indicator MFs for ubiquitous OC_{bios} (TS + PS) with a mean of 44 ± 16%. For OC_{petro} between 4% and 49% of indicator MFs with a mean of 29 ± 12% were detected in the river sediment samples. These observations reveal that indicator MFs for all three organic matter sources could be generally detected in river sediment samples, suggesting, in unison with previous studies (Galy & Eglinton, 2011a; Menges et al., 2020), that material from all three sources reaches the river. At the same time, the number of detected indicator MFs in the river sediment samples was significantly lower than the number of indicator MFs identified for each specific source, and highly variable between river sediment samples. These differences may be in part due to charge competition effects leading to an overrepresentation of compounds with high ionization efficiencies in ESI-FT-ICR-MS measurements (Kebarle & Verkerk, 2009). Thus, in different complex mixtures, the same compound can be detected at different relative abundances. However, since the total ion abundance in the sediment spectra was below a critical level and we simply consider whether MFs are identified, this should only play a minor role. Furthermore, for OC_{petro}, it is likely that the large molecular variability, potentially associated with the wide range of depositional facies within the Tethyan Sedimentary Sequence (Garzanti, 1999), gives rise to a lower number of identified indicator MFs and therefore an even lower number of indicator MFs detected in the river sediments. Another possible contributing factor may be that changes of the molecular composition, due to chemical transformations during mobilization and riverine transport, reduce the presence of source indicator

MFs in river sediment mixtures. While lipids are generally considered to be recalcitrant compounds (e.g., Lorenz et al., 2007), this may not systematically apply to compounds that are only present in low abundances and overlooked using conventional methods. This highlights the need for more high-resolution studies in which a wider range of compounds can be considered and their behavior understood.

3.4. Transformation of Organic Matter Along Different Pathways

By comparing samples along known transformation pathways, we have tested whether it is possible to detect molecular transformations of particulate organic matter via applying FT-ICR-MS analyses that are not discernible using analytical methods focusing on bulk OC measurements or on a limited set of compounds.

3.4.1. Transformation Pathway I: Incorporation of Litter Into Soil

Previous studies have suggested that characteristic patterns of wax lipids and isoprenoids are preserved as fresh plant material is incorporated into soil organic matter in grassland and forest soils, and partially degraded in the upper soil horizons (Otto & Simpson, 2005; Van Bergen et al., 1998). Focusing on three sample pairs of litter and topsoil obtained from the same sites (NP15-7L and −7S, NP16-45L and −45S, NP16-47L and −47S), we have perused the FT-ICR-MS data for evidence of preservation and loss of compounds by comparing their MF molecular metrics (Figure 6).

Of the 1,005 to 1,400 different MFs detected in the litter—topsoil sample pairs, 28 ± 7% were specific to the litter sample and 31 ± 7% MFs were specific to the topsoil sample. The remaining 42 ± 1% MFs were present in both litter and topsoil at a site. The cumulative relative abundance of the shared MFs was 89 ± 3% in the litter samples and 91 ± 3% in the topsoil samples, indicating that most of the extractable lipid fraction found in litter is also present in soil organic matter. MFs only found in litter were characterized partly by higher O/C ratios and lower H/C ratios than shared MFs and topsoil-specific MFs. While we did not observe a shift in the range of carbon numbers, we found a high frequency of MFs with carbon numbers 20, 30 and 40 in the shared MFs, which most probably correspond to isoprenoidal compounds derived from higher plants (Gershenzon & Croteau, 1993). This suggests that these compounds efficiently make their way from litter into topsoil. The distributions of DBEs had similar ranges for litter-specific, topsoil-specific and shared MFs for each of the litter—topsoil pairs. However, MFs unique to the litter samples had unimodal distributions with a maximum at 6–7 DBEs, shared MFs had bimodal distributions with a primary mode at 1–2 DBEs and a secondary mode at 6–7 DBEs and MFs specific to topsoils had a unimodal distribution peaking at 2–3 DBEs (Figure 6). This suggests that MFs with higher DBEs in the topsoil originated from litter input, and that this input is partially lost or transformed for example, by degradation or saturation, which results in lowering of DBEs. Meanwhile, MFs with lower DBEs seem to form newly within the topsoil or result from degradation of higher DBE compounds. Similarly, the modes of NOSC, an indicator of oxidation state, were different with litter-specific MFs peaking at higher NOSC values and topsoil-specific MFs peaking at lower NOSC values for each of the litter—topsoil pairs (Figure 6). This is consistent with microbial oxidation taking place within the upper soil horizon and preferentially removing higher NOSC compounds (LaRowe & Van Cappellen, 2011).

The significant number of shared MFs and their high relative abundance, as well as the well-preserved typical isoprenoidal pattern of carbon numbers is congruent with the active incorporation of plant-derived extractable lipids into the topsoil. The comparison of the molecular metrics of the MFs revealed small, but clearly detectable shifts toward a lower O/C ratio, lower DBEs and a lower NOSC in soil compared to the overlying litter, as expected with a partial degradation of litter compounds upon integration into the soil material, and showing that FT-ICR-MS molecular metrics are a suitable means of detecting molecular transformations along the litter to topsoil pathway.

3.4.2. Transformation Pathway II: Fluvial Mobilization and Transfer of Organic Matter

A further opportunity for transformation of organic matter occurs when organic matter is transferred and transported in the river network. Particulate organic matter export in the upper KG River is dominated by bedrock derived OC_{petro}, but both aged and modern OC_{bios} have also been shown to contribute to the river load based on bulk isotope data (Galy & Eglinton, 2011a; Menges et al., 2020). Hence, topsoils, paleosols and bedrock contribute extractable lipids to the river sediment load. We have compared the molecular composition of each source type with the molecular composition of the organic matter in main river sediments to investigate potential

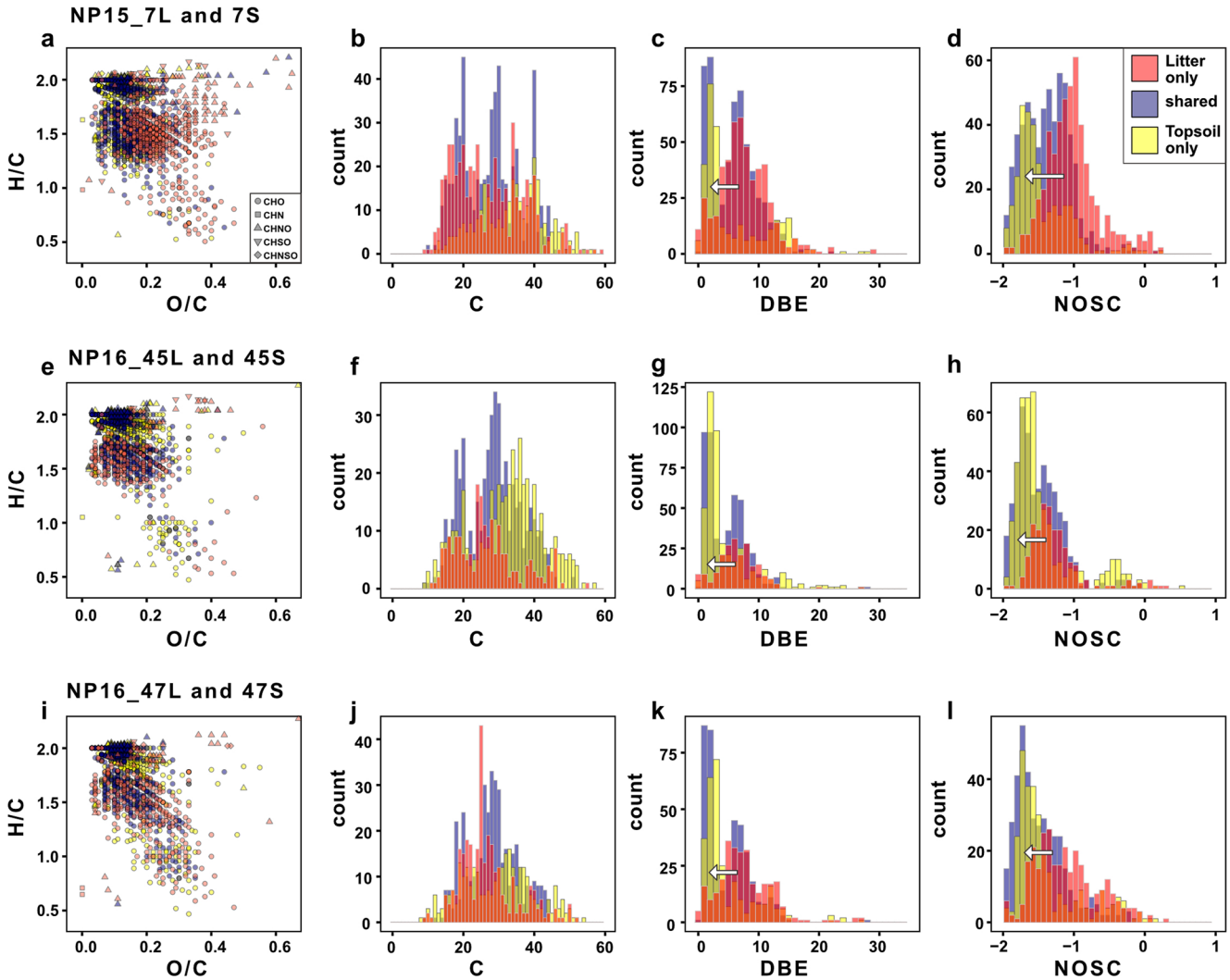

Figure 6. Transformation pathway 1: Summary of all identified molecular formulas (MFs) which only occurred in litter (red), only occurred in topsoils (yellow) and were shared between both sample types (blue). (a) Plots show the detected MFs in van Krevelen space and (b) the distribution of number of c-atoms, (c) double bound equivalents (DBE), and (d) nominal oxidation state of carbon (NOSC) for shared and specific MFs.

changes on molecular level due to organic matter mobilization and riverine transport over distances of $<10^2$ km. For this analysis we used aggregated data sets representing the full molecular richness observed for a given sample type (see Section 2.4.3).

Since the KG River receives organic matter from all documented sources (OC_{petro}, modern and aged OC_{bios}) we expected the molecular richness (i.e., the number of different MFs for each sample type) of main river sediment to be comparably high. However, while the total molecular richness in all source samples combined was 4,031, the total molecular richness of the main river sediment samples was only 1,147. Of these, 1,067 MFs were shared between sources and main river samples and 80 MFs were specific to main river sediment samples. While these 80 MFs are a minor component in number and relative abundance, 41 of those MFs were CHNO MFs hinting at a potential aquatic origin (Meyers, 1997). This highlights the potential of FT-ICR-MS to identify specific marker compounds in further studies.

The majority of the 4,031 MFs present in organic matter source samples were specific to sources (2,964 MFs), and not found in any main river sediment samples. While the MFs shared between source and river samples included those with highest relative abundances, the cumulative relative abundance of source MFs not found in river samples ranged from 10% to around 30% in the different source samples (and above 60% for the bedrock sample R26). Even larger is the loss in molecular richness with the number of "lost" MFs representing 63% of all

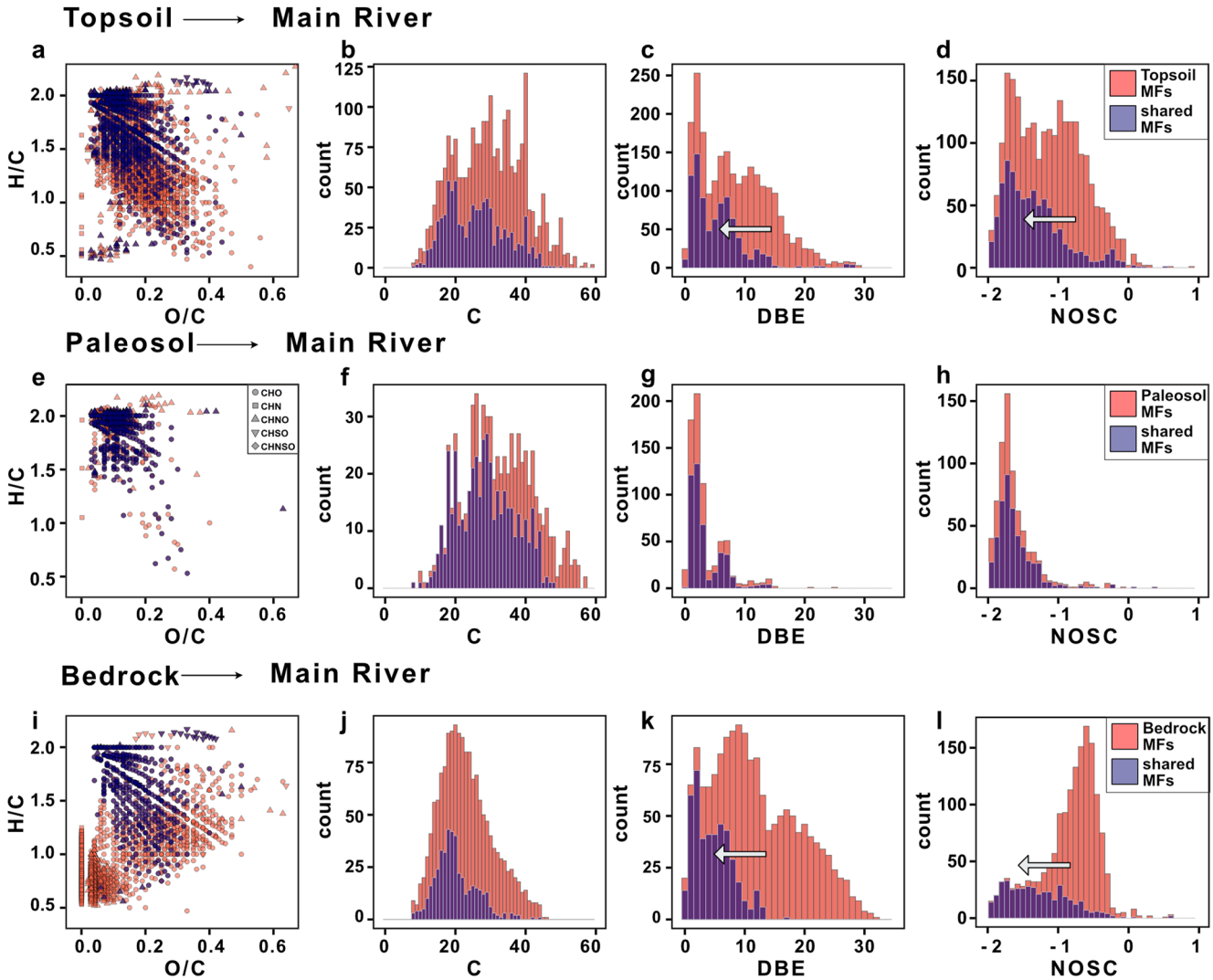

Figure 7. Transformation pathway 2: Summary of all identified molecular formulas (MFs) within the different sources (red) and MFs that could also be detected in main river sediment samples (blue) for (a–d) modern OC_{bios} from topsoils, (e–h) aged OCbios from paleosols, and (i–l) bedrock derived OC_{petro}. (a, e, and i) Plots show the detected MFs in van Krevelen space and (b, f, and g) the distribution of number of c-atoms, (c, g, and k) double bond equivalents (DBE), and (d, h, and l) nominal oxidation state of carbon (NOSC).

MFs detected in topsoils, 37% of all MFs in paleosols and 72% of all MFs in bedrock. Below, we consider changes in molecular metrics associated with this loss of MFs in the transition between individual organic matter sources and the main river sediment load (Figure 7).

Topsoil MFs covered a large range of H/C and O/C ratios in comparison to MFs found in both topsoils and main river sediments, which clustered at higher H/C values and slightly lower O/C values. Similarly, the large range of carbon numbers observed in topsoil MFs was reduced for MFs shared between topsoils and main river sediment samples. For DBEs and NOSC, MFs shared between topsoils and main river sediment samples had a much smaller range in DBEs, and distributions strongly skewed toward low values for DBEs and NOSC. MFs found in both paleosols and main river sediments exhibited lower carbon numbers than MFs found only in paleosols, but the distributions of DBEs, NOSC and oxygen content were very similar between all paleosol MFs and those shared with main river sediments. Bedrock MFs were largely characterized by low H/C and O/C ratios and also contained a number of oxygen-free CHN MFs with low H/C ratios which were not found in main river sediment samples. Shared MFs between bedrock and main river sediment samples had much higher H/C values. Also, the range in carbon numbers was reduced for MFs shared between bedrock and main river sediments. Strongly skewed distributions were observed for the DBEs and NOSC of MFs shared between bedrock and main river

sediments, similar to the observations for topsoils. All observed trends in molecular metrics were still evident when restricting the data set to MFs that occurred in at least two samples and when comparing individual source samples and main river sediment samples.

Taken together, we observed a significant loss of MFs found in source samples associated with sediment entrainment in and transport by the main river. This loss was associated with strong differences in molecular metrics between lost and persistent MFs for topsoils and bedrock but less for paleosols. For topsoils and bedrock, lost MFs were characterized by higher DBEs and NOSC, whereas MFs surviving transfer into the main river sediment had much lower DBEs and NOSC values. Paleosols did not contain high DBE and NOSC MFs in the first place, and for this source we did not observe a strong shift upon transfer into the main river. These observations suggest that the apparent loss of MFs between OC sources and river sediments is driven by a mechanism that is selective with respect to the reported molecular metrics. Potential reasons for these apparent transformations can be factors inherent in our method, as well as natural processes. Methodological issues could include sampling bias and charge competition effects during the measurements. A sampling bias could result in a loss of MFs if the sampled sources do not actually represent the material transferred to the river at the time of sampling. Although we cannot exclude this possibility, our main river sediment data set integrates samples taken at a time of maximum geomorphic activity across the catchment from many locations along the KG River. Therefore, we consider it to capture the compositional variety of material available to the river, and exported from the catchment. Separately, the dilution of compounds below the detection limit, or signal suppression within the ESI source leading to masking of compounds with lower ESI response factors (Reemtsma, 2009) could also explain a partial loss of MFs. If the loss of MFs in our study were driven mainly by dilution and signal suppression, this should be visible in a covariation of molecular metrics and the relative abundance of MFs that showed strong shifts along the source to river pathway. We did not observe such covariation or any clear trend when comparing the relative abundances of MFs with NOSC, DBEs or C for the different source samples in our data set. Additionally, the river sediment samples exhibited a lower diversity in composition and had a lower total absolute abundance when measured in the same concentration compared to soil samples, making significant signal suppression unlikely. Therefore, we do not consider dilution and/or signal suppression to be a major cause of the observed pattern.

Natural processes that could result in the observed pattern of apparent molecular transformations range from small changes in structure or number of atoms, to full remineralization of organic molecules into their constituent inorganic components. Possible mechanisms could be a direct degradation of organic compounds mediated by microorganisms or degradation following the dissociation of organic compounds from the sedimentary matrix. The dissociation of terrestrial lipid biomarkers from the sedimentary matrix and subsequent degradation has been suggested to occur in estuaries, caused by strong changes in salinity and priming by fresh organic matter (Zhu et al., 2013), even though lipids are generally insoluble in water (Dinel et al., 1990). Since we are not aware of strong changes in salinity or availability of fresh organic matter in the upper KG valley, we do not anticipate this process to play a major role. Regarding the direct degradation of lipid compounds, recent work by Graham et al. (2017) has shown that physically bound OC, defined as compounds extractable by chloroform, can serve as a substrate for microbial metabolism within river sediments under certain conditions. The authors suggested that aerobic microbial metabolism can result in a depletion of thermodynamically favorable OC from the physically bound OC pool in rivers where the absence of riparian vegetation is limiting access to weakly bound or dissolved, and therefore readily bioavailable organic matter (Graham et al., 2017). In our data set, the observed trends in molecular metrics toward a lower NOSC are consistent with preferential removal of thermodynamically favorable compounds (LaRowe & Van Cappellen, 2011). Moreover, the dry upper KG River has no riparian vegetation and low dissolved organic carbon concentrations. Therefore, microbially mediated degradation of parts of the lipid fraction is a plausible mechanism explaining our observations.

Our analysis highlights that untargeted FT-ICR-MS analysis can reveal previously unknown shifts in molecular composition along organic matter pathways in the environment. This is an important step forward since the dynamics of organic matter are important for many different processes but challenging to study.

4. Conclusion

In this study, we used ESI-FT-ICR-MS measurements to generate detailed insights into organic matter dynamics on a landscape scale. We analyzed different organic matter sources and river sediment samples from a well-studied headwater catchment of a large Himalayan river. The high-resolution molecular data revealed clear characteristics

shared by samples from individual organic matter pools (except for tributary samples). They also showed systematic differences in molecular composition of organic matter between the different sampled pools. This confirms the suitability of FT-ICR-MS data on extractable lipids for the characterization and differentiation of particulate organic matter types in the landscape.

Using indicator species analysis, we have identified up to 10^2 indicator molecular formulas (MF) for each of the sampled organic carbon sources. Comparing indicator MFs between the sampled environments, we have found that while varying fractions of source indicator MFs were detectably transferred into the river load, a significant loss of indicator MFs occurred between source and river. These findings highlight the potential of FT-ICR-MS data to screen for new biomarkers, and to characterize in detail the particulate organic carbon sources that are currently eroded in the catchment.

Moreover, the detailed molecular data enable the identification of transformations of organic compounds in the landscape. We have focused on the fate of the lipid fraction along two important pathways. Lipids are formed in life vegetation. The large number of MFs present in litter as well as underlying topsoil, and the well-preserved isoprenoidal pattern of carbon numbers in topsoil MFs, suggests an efficient incorporation of litter derived compounds into topsoil. In contrast, our data show the erosion of different organic matter sources and their transfer into the river is accompanied by a large decrease in molecular richness. Moreover, for all sources, MFs shared with sediments sampled along the main river were characterized by low DBEs and a low NOSC, while MFs with higher DBEs and NOSC were selectively lost in the transfer from sources to the river load. We tentatively attribute this to microbially mediated direct degradation of compounds, but cannot exclude other mechanisms. These insights from FT-ICR-MS data are consistent with, and add detail to previous studies on organic matter dynamics in the region, and demonstrate the potential of FT-ICR-MS data for exploration of organic matter transformations. Additional information can be potentially be obtained by including measurements in ESI positive mode or atmospheric pressure photoionization (APPI) which has the advantage of a better concentration reponse relationship (Kauppila et al., 2017).

Taken together, FT-ICR-MS analysis can add depth and detail to the study of organic matter cycling in natural landscapes. While direct insight into molecular composition and reaction pathways is limited for complex organic mixtures, the ultra-high resolution of FT-ICR-MS can be employed to generate large MF-level data sets permitting the identification of indicator MFs and the detailed specification of the distribution of molecular metrics. Both open unique windows on the world of organic matter dynamics in the Earth surface system, and allow the observation and integration of the large natural variability of organic matter compounds and the detection of general patterns of molecular transformations on a landscape scale.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Research data associated with this article are archived on a public repository and can be accessed at Menges et al. (2022, http://doi.org/10.5880/GFZ.4.6.2022.002).

Acknowledgments
We thank Cornelia Karger and Anke Kaminsky (GFZ, Potsdam) for support during laboratory work and FT-ICR-MS analyses, and Markus Reich (GFZ, Potsdam) and Ed Tipper (University of Cambridge) for custom filter units used to collect samples for this study. This work was supported by the Helmholtz Impuls und Vernetzungsfond. Field work was pursued with GFZ expedition funding. Open Access funding enabled and organized by Projekt DEAL.

References

Adhikari, B. R., & Wagreich, M. (2011). Facies analysis and basin architecture of the Thakkhola-Mustang Graben (Neogene-Quaternary), central Nepal Himalaya. *Austrian Journal of Earth Science*, *104*(1), 66–80.

Andermann, C., Bonnet, S., Crave, A., Davy, P., Longuevergne, L., & Gloaguen, R. (2012). Sediment transfer and the hydrological cycle of Himalayan rivers in Nepal. *Comptes Rendus Geoscience*, *344*(11–12), 627–635. https://doi.org/10.1016/j.crte.2012.10.009

Andermann, C., Bonnet, S., & Gloaguen, R. (2011). Evaluation of precipitation data sets along the Himalayan front. *Geochemistry, Geophysics, Geosystems*, *12*(7), Q07023. https://doi.org/10.1029/2011GC003513

Bai, Y., Fang, X., Nie, J., Wang, Y., & Wu, F. (2009). A preliminary reconstruction of the paleoecological and paleoclimatic history of the Chinese Loess Plateau from the application of biomarkers. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *271*(1–2), 161–169. https://doi.org/10.1016/j.palaeo.2008.10.006

Bell, M., & Blais, J. M. (2019). "-Omics" workflow for paleolimnological and geological archives: A review. *Science of the Total Environment*, *672*, 438–455. https://doi.org/10.1016/j.scitotenv.2019.03.477

Bellamy, P. H., Loveland, P. J., Bradley, R. I., Lark, R. M., & Kirk, G. J. D. (2005). Carbon losses from all soils across England and Wales 1978–2003. *Nature*, *437*(7056), 245–248. https://doi.org/10.1038/nature04038

Berner, R. A. (1982). Burial of organic carbon and pyrite sulfur in the modern ocean: It geochemical and environmental significance. *American Journal of Science*, *282*(4), 451–473. https://doi.org/10.2475/ajs.282.4.451

Blair, N. E., & Aller, R. C. (2012). The fate of terrestrial organic carbon in the marine environment. *Annual Review of Marine Science*, *4*(1), 401–423. https://doi.org/10.1146/annurev-marine-120709-142717

Bookhagen, B., & Burbank, D. W. (2010). Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge. *Journal of Geophysical Research*, *115*(3), 1–25. https://doi.org/10.1029/2009JF001426

Bordet, P., Colchen, M., Krummenacher, D., Lefort, P., Mouterde, R., & Remy, M. (1971). *Recherches geologiques dans l'Himalaya du Nepal, region de la Thakkhola*. Éditions du Centre National de la Recherche Scientifique.

Bordet, P., Colchen, M., Le Fort, P., & Pêcher, A. (1981). The geodynamic evolution of the Himalaya - Ten Years of research in central Nepal Himalaya and some other regions. *Geodynamics Series*, *3*, 149–168. https://doi.org/10.1029/GD003p0149

Boye, K., Noël, V., Tfaily, M. M., Bone, S. E., Williams, K. H., Bargar, J. R., & Fendorf, S. (2017). Thermodynamically controlled preservation of organic carbon in floodplains. *Nature Geoscience*, *10*(6), 415–419. https://doi.org/10.1038/ngeo2940

Brocks, J. J., Logan, G. A., Buick, R., & Summons, R. E. (1999). Archean molecular fossils and the early rise of eukaryotes. *Science*, *285*(5430), 1033–1036. https://doi.org/10.1126/science.285.5430.1033

Bröder, L., Tesi, T., Andersson, A., Semiletov, I., & Gustafsson, Ö. (2018). Bounding cross-shelf transport time and degradation in Siberian-Arctic land-ocean carbon transfer. *Nature Communications*, *9*(1), 806. https://doi.org/10.1038/s41467-018-03192-1

Cho, Y., Ahmed, A., Islam, A., & Kim, S. (2015). Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics. *Mass Spectrometry Reviews*, *34*(2), 248–263. https://doi.org/10.1002/mas.21438

Cole, J. J., Prairie, Y. T., Caraco, N. F., McDowell, W. H., Tranvik, L. J., Striegl, R. G., et al. (2007). Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems*, *10*(1), 171–184. https://doi.org/10.1007/s10021-006-9013-8

De Cáceres, M., & Legendre, P. (2009). Associations between species and groups of sites: Indices and statistical inference. *Ecology*, *90*(12), 3566–3574. https://doi.org/10.1890/08-1823.1

De Cáceres, M., Legendre, P., & Moretti, M. (2010). Improving indicator species analysis by combining groups of sites. *Oikos*, *119*(10), 1674–1684. https://doi.org/10.1111/j.1600-0706.2010.18334.x

Derry, L., & France-Lanord, C. (1997). Organic carbon burial forcing of the carbon cycle from Himalayan erosion. *Nature*, *390*(6655), 65–67. https://doi.org/10.1038/36324

Dijkshoorn, K., & Huting, J. (2009). *Soil and terrain database for Nepal*, *ISRIC – World Soil Information*. Wageningen.

Dinel, H., Schnitzer, M., & Mehuys, G. R. (1990). Soil lipids: Origin, nature, content, decomposition, and effect on soil physical properties. In J. M. Bollag & G. Stotzky (Eds.), *Soil biochemistry* (Vol. 6, pp. 397–429). Routledge.

Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, *67*(3), 345–366. https://doi.org/10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2

Feng, X., Vonk, J. E., van Dongen, B. E., Gustafsson, Ö., Semiletov, I. P., Dudarev, O. V., et al. (2013). Differential mobilization of terrestrial carbon pools in Eurasian Arctic river basins. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(35), 14168–14173. https://doi.org/10.1073/pnas.1307031110

Fort, M., Freytet, P., & Colchen, M. (1982). Structural and sedimentological evolution of the Thakkhola Mustang graben (Nepal Himalayas). *Zeitschrift für Geomorphologie*, *42*, 75–98.

Freymond, C. V., Kündig, N., Stark, C., Peterse, F., Buggle, B., Lupker, M., et al. (2018). Evolution of biomolecular loadings along a major river system. *Geochimica et Cosmochimica Acta*, *223*, 389–404. https://doi.org/10.1016/j.gca.2017.12.010

Galy, V., & Eglinton, T. I. (2011a). Protracted storage of biospheric carbon in the Ganges-Brahmaputra basin. *Nature Geoscience*, *4*(12), 843–847. https://doi.org/10.1038/ngeo1293

Galy, V., Eglinton, T. I., France-Lanord, C., & Sylva, S. (2011). The provenance of vegetation and environmental signatures encoded in vascular plant biomarkers carried by the Ganges-Brahmaputra rivers. *Earth and Planetary Science Letters*, *304*(1–2), 1–12. https://doi.org/10.1016/j.epsl.2011.02.003

Galy, V., France-Lanord, C., & Lartiges, B. (2008). Loading and fate of particulate organic carbon from the Himalaya to the Ganga-Brahmaputra delta. *Geochimica et Cosmochimica Acta*, *72*(7), 1767–1787. https://doi.org/10.1016/j.gca.2008.01.027

Galy, V. V., Beyssac, O., France-Lanord, C., & Eglinton, T. (2008). Recycling of graphite during Himalayan erosion: A geological stabilization of carbon in the crust. *Science*, *322*(5903), 943–945. https://doi.org/10.1126/science.1161408

Galy, V. V., & Eglinton, T. I. (2011b). Protracted storage of biospheric carbon in the Ganges–Brahmaputra basin. *Nature Geoscience*, *4*(12), 843–847. https://doi.org/10.1038/ngeo1293

Galy, V. V., France-Lanord, C., Beyssac, O., Faure, P., Kudrass, H., & Palhol, F. (2007). Efficient organic carbon burial in the Bengal fan sustained by the Himalayan erosional system. *Nature*, *450*(7168), 407–410. https://doi.org/10.1038/nature06273

Garzanti, E. (1999). Stratigraphy and sedimentary history of the Nepal Tethys Himalaya passive margin. *Journal of Asian Earth Sciences*, *17*(5–6), 805–827. https://doi.org/10.1016/S1367-9120(99)00017-6

Gershenzon, J., & Croteau, R. B. (1993). Terpenoid biosynthesis: The basic pathway and formation of monoterpenes, sesquiterpenes, and diterpenes. In T. S. Moore (Ed.), *Lipid metabolism in plants*. CRC Press.

Godin, L. (2003). Structural evolution of the Tethyan sedimentary sequence in the Annapurna area, central Nepal Himalaya. *Journal of Asian Earth Sciences*, *22*(4), 307–328. https://doi.org/10.1016/S1367-9120(03)00066-X

Goñi, M. A., Monacci, N., Gisewhite, R., Crockett, J., Nittrouer, C., Ogston, A., et al. (2008). Terrigenous organic matter in sediments from the Fly River delta-clinoform system (Papua New Guinea). *Journal of Geophysical Research*, *113*(F1), F01S10. https://doi.org/10.1029/2006JF000653

Goñi, M. A., & Montgomery, S. (2000). Alkaline CuO oxidation with a microwave digestion system: Lignin analyses of geochemical samples. *Analytical Chemistry*, *72*(14), 3116–3121. https://doi.org/10.1021/ac991316w

Graham, E. B., Tfaily, M. M., Crump, A. R., Goldman, A. E., Bramer, L. M., Arntzen, E., et al. (2017). Carbon inputs from Riparian vegetation limit oxidation of physically bound organic carbon via biochemical and thermodynamic processes. *Journal of Geophysical Research: Biogeosciences*, *122*(12), 3188–3205. https://doi.org/10.1002/2017JG003967

Hedges, J. I. (1992). Global biogeochemical cycles: Progress and problems. *Marine Chemistry*, *39*(1–3), 67–93. https://doi.org/10.1016/0304-4203(92)90096-S

Hemingway, J. D., Hilton, R. G., Hovius, N., Eglinton, T. I., Haghipour, N., Wacker, L., et al. (2018). Microbial oxidation of lithospheric organic carbon in rapidly eroding tropical mountain soils. *Science*, *360*(6385), 209–212. https://doi.org/10.1126/science.aao6463

Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., et al. (2006). Characterization of a major refractory component of marine dissolved organic matter. *Geochimica et Cosmochimica Acta*, *70*(12), 2990–3010. https://doi.org/10.1016/j.gca.2006.03.021

Hilton, R. G., Galy, A., Hovius, N., Chen, M.-C., Horng, M.-J., & Chen, H. (2008). Tropical-cyclone-driven erosion of the terrestrial biosphere from mountains. *Nature Geoscience*, *1*(11), 759–762. https://doi.org/10.1038/ngeo333

Hilton, R. G., Galy, A., Hovius, N., Horng, M. J., & Chen, H. (2011). Efficient transport of fossil organic carbon to the ocean by steep mountain rivers: An orogenic carbon sequestration mechanism. *Geology*, *39*(1), 71–74. https://doi.org/10.1130/G31352.1

Hodges, K. V., Parrish, R. R., & Searle, M. P. (1996). Tectonic evolution of the central Annapurna range, Nepalese Himalayas. *Tectonics*, *15*(6), 1264–1291. https://doi.org/10.1029/96TC01791

Hoffmann, B., Feakins, S. J., Bookhagen, B., Olen, S. M., Adhikari, D. P., Mainali, J., & Sachse, D. (2016). Climatic and geomorphic drivers of plant organic matter transport in the Arun River, E Nepal. *Earth and Planetary Science Letters*, *452*, 104–114. https://doi.org/10.1016/j.epsl.2016.07.008

Hughey, C. A., Rodgers, R. P., Marshall, A. G., Qian, K., & Robbins, W. K. (2002). Identification of acidic NSO compounds in crude oils of different geochemical origins by negative-ion electrospray Fourier transform ion cyclotron resonance mass spectrometry. *Organic Geochemistry*, *33*(7), 743–759. https://doi.org/10.1016/s0146-6380(02)00038-4

Ittekkot, V. (1988). Global trends in the nature of organic matter in river suspensions. *Nature*, *332*(6163), 436–438. https://doi.org/10.1038/332436a0

Kauppila, T. J., Syage, J. A., & Benter, T. (2017). Recent developments in atmospheric pressure photoionization-mass spectrometry. *Mass Spectrometry Reviews*, *36*(3), 423–449. https://doi.org/10.1002/mas.21477

Kebarle, P., & Verkerk, U. H. (2009). Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrometry Reviews*, *28*(6), 898–917. https://doi.org/10.1002/mas.20247

Kellerman, A. M., Kothawala, D. N., Dittmar, T., & Tranvik, L. J. (2015). Persistence of dissolved organic matter in lakes related to its molecular characteristics. *Nature Geoscience*, *8*(6), 454–457. https://doi.org/10.1038/ngeo2440

Kim, S., Kramer, R. W., & Hatcher, P. G. (2003). Graphical method for analysis of ultrahigh-resolution braodband mass spectra of natural organic matter, the van Krevelen diagram. *Analytical Chemistry*, *75*(20), 5336–5344. https://doi.org/10.1021/ac034415p

Koch, B. P., Dittmar, T., Witt, M., & Kattner, G. (2007). Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Analytical Chemistry*, *79*(4), 1758–1763. https://doi.org/10.1021/ac061949s

Koch, B. P., Witt, M., Engbrodt, R., Dittmar, T., & Kattner, G. (2005). Molecular formulae of marine and terrigenous dissolved organic matter detected by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Geochimica et Cosmochimica Acta*, *69*(13), 3299–3308. https://doi.org/10.1016/j.gca.2005.02.027

Kögel-Knabner, I. (2002). The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. *Soil Biology and Biochemistry*, *34*(2), 139–162. https://doi.org/10.1016/S0038-0717(01)00158-4

Kujawinski, E. B., Longnecker, K., Blough, N. V., Del Vecchio, R., Finlay, L., Kitner, J. B., & Giovannoni, S. J. (2009). Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochimica et Cosmochimica Acta*, *73*(15), 4384–4399. https://doi.org/10.1016/j.gca.2009.04.033

LaRowe, D. E., & Van Cappellen, P. (2011). Degradation of natural organic matter: A thermodynamic analysis. *Geochimica et Cosmochimica Acta*, *75*(8), 2030–2042. https://doi.org/10.1016/j.gca.2011.01.020

Larson, K. P., & Godin, L. (2009). Kinematics of the Greater Himalayan sequence, Dhaulagiri Himal: Implications for the structural framework of central Nepal. *Journal of the Geological Society*, *166*(1), 25–43. https://doi.org/10.1144/0016-76492007-180

Le Fort, P. (1975). Himalayas: The collided range. Present knowledge of the continental arc. *American Journal of Science*, *275*(1), 1–44.

Legendre, L., & Legendre, P. (2012). *Numerical ecology* (3rd ed.). Elsevier.

Lorenz, K., Lal, R., Preston, C. M., & Nierop, K. G. J. (2007). Strengthening the soil organic carbon pool by increasing contributions from recalcitrant aliphatic bio(macro)molecules. *Geoderma*, *142*(1–2), 1–10. https://doi.org/10.1016/j.geoderma.2007.07.013

Marshall, A. G., & Rodgers, R. P. (2004). Petroleomics: The next grand challenge for chemical analysis petroleum and mass spectrometry: Divergent. *Accounts of Chemical Research*, *37*(1), 53–59. https://doi.org/10.1021/ar020177t

Medeiros, P. M., Seidel, M., Niggemann, J., Spencer, R. G. M., Hernes, P. J., Yager, P. L., et al. (2016). Novel molecular approach for tracing terrigenous dissolved organic matter into the deep ocean. *Global Biogeochemical Cycles*, *30*(5), 689–699. https://doi.org/10.1002/2015gb005320

Menges, J., Hovius, N., Andermann, C., Dietze, M., Swoboda, C., Cook, K. L., et al. (2019). Late Holocene landscape collapse of a trans-Himalayan dryland: Human impact and aridification. *Geophysical Research Letters*, *46*(23), 13814–13824. https://doi.org/10.1029/2019GL084192

Menges, J., Hovius, N., Andermann, C., Lupker, M., Haghipour, N., Märki, L., & Sachse, D. (2020). Variations in organic carbon sourcing along a trans-Himalayan river determined by a Bayesian mixing approach. *Geochimica et Cosmochimica Acta*, *286*, 159–176. https://doi.org/10.1016/j.gca.2020.07.003

Menges, J., Hovius, N., Poetz, S., Osterholz, H., & Sachse, D. (2022). FT-ICR-MS data used to trace variations of organic carbon sourcing along a trans-Himalayan river, central Nepal [Dataset]. GFZ Data Services. https://doi.org/10.5880/GFZ.4.6.2022.002

Meyers, P. A. (1997). Organic geochemical proxies of paleoceanographic. *Organic Geochemistry*, *27*(5), 213–250. https://doi.org/10.1016/s0146-6380(97)00049-1

Mostovaya, A., Hawkes, J. A., Dittmar, T., & Tranvik, L. J. (2017). Molecular determinants of dissolved organic matter reactivity in lake water. *Frontiers in Earth Science*, *5*, 1–13. https://doi.org/10.3389/feart.2017.00106

Oades, J. M. (1988). The retention of organic matter in soils. *Biogeochemistry*, *5*(1), 35–70. https://doi.org/10.1007/BF02180317

Ohno, T., Sleighter, R. L., & Hatcher, P. G. (2016). Comparative study of organic matter chemical characterization using negative and positive mode electrospray ionization ultrahigh-resolution mass spectrometry. *Analytical and Bioanalytical Chemistry*, *408*(10), 2497–2504. https://doi.org/10.1007/s00216-016-9346-x

Otto, A., & Simpson, M. J. (2005). Degradation and preservation of vascular plant-derived biomarkers in grassland and forest soils from Western Canada. *Biogeochemistry*, *74*(3), 377–409. https://doi.org/10.1007/s10533-004-5834-8

Pellegrin, V. (1983). Molecular formulas of organic compounds: The nitrogen rule and degree of unsaturation. *Journal of Chemical Education*, *60*(8), 626. https://doi.org/10.1021/ed060p626

Petsch, S. T. (2014). Weathering of organic carbon. In *Treatise on Geochemistry* (2nd ed.), (Vol. 12, pp. 217–238). Elsevier. https://doi.org/10.1016/B978-0-08-095975-7.01013-5

Poetz, S., Horsfield, B., & Wilkes, H. (2014). Maturity-driven generation and transformation of acidic compounds in the organic-rich Posidonia Shale as revealed by electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Energy & Fuels*, *28*(8), 4877–4888. https://doi.org/10.1021/ef500688s

Ponton, C., West, A. J., Feakins, S. J., & Galy, V. (2014). Leaf wax biomarkers in transit record river catchment composition. *Geophysical Research Letters*, *41*(18), 6420–6427. https://doi.org/10.1002/2014GL061328.Received

Pracht, L. E., Tfaily, M. M., Ardissono, R. J., & Neumann, R. B. (2018). Molecular characterization of organic matter mobilized from Bangladeshi aquifer sediment: Tracking carbon compositional change during microbial utilization. *Biogeosciences*, *15*(6), 1733–1747. https://doi.org/10.5194/bg-15-1733-2018

Qi, Y., Xie, Q., Wang, J.-J., He, D., Bao, H., Fu, Q.-L., et al. (2022). Deciphering dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS): From bulk to fractions and individuals. *Carbon Research*, *1*(1), 3. https://doi.org/10.1007/s44246-022-00002-8

Qian, K., Robbins, W. K., Hughey, C. A., Cooper, H. J., Rodgers, R. P., & Marshall, A. G. (2001). Resolution and identification of 3000 crude acids in heavy petroleum by negative-ion microelectrospray high field Fourier transform ion cyclotron resonance mass spectrometry. *Energy & Fuels*, *15*(11), 1505–1511. https://doi.org/10.1021/ef010111z

Rach, O., Brauer, A., Wilkes, H., & Sachse, D. (2014). Delayed hydrological response to Greenland cooling at the onset of the Younger Dryas in western Europe. *Nature Geoscience*, *7*(1), 1–4. https://doi.org/10.1038/ngeo2053

Rach, O., Hadeen, X., & Sachse, D. (2020). An automated solid phase extraction procedure for lipid biomarker purification and stable isotope analysis. *Organic Geochemistry*, *142*, 103995. https://doi.org/10.1016/j.orggeochem.2020.103995

Reemtsma, T. (2009). Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry. Status and needs. *Journal of Chromatography A*, *1216*(18), 3687–3701. https://doi.org/10.1016/j.chroma.2009.02.033

Repeta, D. J. (2015). Chapter 2 - Chemical characterization and cycling of dissolved organic matter. In D. A. Hansell & C. A. Carlson (Eds.), *Biogeochemistry of marine dissolved organic matter* (2nd ed., pp. 21–63). Academic Press. https://doi.org/10.1016/B978-0-12-405940-5.00002-9

Riedel, T., & Dittmar, T. (2014). A method detection limit for the analysis of natural organic matter via Fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry*, *86*(16), 8376–8382. https://doi.org/10.1021/ac501946m

Saijo, K., & Tanaka, S. (2002). Paleosols of middle Holocene age in the Thakkhola basin, central Nepal, and their paleoclimatic significance. *Journal of Asian Earth Sciences*, *21*(3), 323–329. https://doi.org/10.1016/S1367-9120(02)00079-2

Scheingross, J. S., Repasch, M. N., Hovius, N., Sachse, D., Lupker, M., Fuchs, M., et al. (2021). The fate of fluvially-deposited organic carbon during transient floodplain storage. *Earth and Planetary Science Letters*, *561*, 116822. https://doi.org/10.1016/j.epsl.2021.116822

Schug, K., & McNair, H. M. (2002). Adduct formation in electrospray ionization. Part 1: Common acidic pharmaceuticals. *Journal of Separation Science*, *25*(12), 759–766. https://doi.org/10.1002/1615-9314(20020801)25:12<759::aid-jssc760>3.0.co;2-m

Sleighter, R. L., McKee, G. A., & Hatcher, P. G. (2009). Direct Fourier transform mass spectral analysis of natural waters with low dissolved organic matter. *Organic Geochemistry*, *40*(1), 119–125. https://doi.org/10.1016/j.orggeochem.2008.09.012

Sparkes, R. B., Hovius, N., Galy, A., & Liu, J. T. (2020). Survival of graphitized petrogenic organic carbon through multiple erosional cycles. *Earth and Planetary Science Letters*, *531*, 115992. https://doi.org/10.1016/j.epsl.2019.115992

Stenson, A. C., Marshall, A. G., & Cooper, W. T. (2003). Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Analytical Chemistry*, *75*(6), 1275–1284. https://doi.org/10.1021/ac026106p

Stocklin, J. (1980). Geology of Nepal and its regional frame. *Journal of the Geological Society*, *137*(1), 1–34. https://doi.org/10.1144/gsjgs.137.1.0001

Teräväinen, M. J., Pakarinen, J. M. H., Wickström, K., & Vainiotalo, P. (2007). Comparison of the composition of Russian and North Sea crude oils and their eight distillation fractions studied by negative-ion electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry: The effect of suppression. *Energy and Fuels*, *21*(1), 266–273. https://doi.org/10.1021/ef060294v

Tfaily, M. M., Chu, R. K., Tolić, N., Roscioli, K. M., Anderton, C. R., Paša-Tolić, L., et al. (2015). Advanced solvent based methods for molecular characterization of soil organic matter by high-resolution mass spectrometry. *Analytical Chemistry*, *87*(10), 5206–5215. https://doi.org/10.1021/acs.analchem.5b00116

Tfaily, M. M., Chu, R. K., Toyoda, J., Tolić, N., Robinson, E. W., Paša-Tolić, L., & Hess, N. J. (2017). Sequential extraction protocol for organic matter from soils and sediments using high resolution mass spectrometry. *Analytica Chimica Acta*, *972*, 54–61. https://doi.org/10.1016/j.aca.2017.03.031

Van Bergen, P. F., Nott, C. J., Bull, I. D., Poulton, P. R., & Evershed, R. P. (1998). Organic geochemical studies of soils from the Rothamsted Classical Experiments—IV. Preliminary results from a study of the effect of soil pH on organic matter decay. *Organic Geochemistry*, *29*(5–7), 1779–1795. https://doi.org/10.1016/S0146-6380(98)00188-0

Wagner, S., Riedel, T., Niggemann, J., Vähätalo, A. V., Dittmar, T., & Jaffé, R. (2015). Linking the molecular signature of heteroatomic dissolved organic matter to watershed characteristics in world rivers. *Environmental Science & Technology*, *49*(23), 13798–13806. https://doi.org/10.1021/acs.est.5b00525

Zachos, J. C., Schouten, S., Bohaty, S., Quattlebaum, T., Sluijs, A., Brinkhuis, H., et al. (2006). Extreme warming of mid-latitude coastal ocean during the Paleocene-Eocene Thermal Maximum: Inferences from TEX_{86} and isotope data. *Geology*, *34*(9), 737–740. https://doi.org/10.1130/G22522.1

Zhu, C., Wagner, T., Talbot, H. M., Weijers, J. W. H., Pan, J. M., & Pancost, R. D. (2013). Mechanistic controls on diverse fates of terrestrial organic components in the East China Sea. *Geochimica et Cosmochimica Acta*, *117*, 129–143. https://doi.org/10.1016/j.gca.2013.04.015