Nature Geoscience

Roth, V.-N., Lange, M., Simon, C., Hertkorn, N., Bucher, S., Goodall, T., Griffiths, R.I., Mellado-Vázquez, P.G., Mommer, L., Oram, N.J., Weigelt, A., Dittmar, T. & Gleixner, G. (2019).

Persistence of dissolved organic matter explained by molecular changes during its passage through soil.

Nature Geoscience, 12(9), 755-761.

https://doi.org/10.1038/s41561-019-0417-4

https://www.nature.com/articles/s41561-019-0417-4

Persistence of dissolved organic matter explained by molecular changes during its passage through soil

Authors

Vanessa-Nina Roth[1, 2], Markus Lange[1,*], Carsten Simon[1], Norbert Hertkorn[3], Sebastian Bucher[1], Timothy Goodall[4], Robert I. Griffiths[4], Perla G. Mellado-Vázquez[1,5], Liesje Mommer[6], Natalie J. Oram[6,7], Alexandra Weigelt[8,9], Thorsten Dittmar[10,11], Gerd Gleixner[1]

Affiliations

[1] Max Planck Institute for Biogeochemistry Jena Germany

[2] Thüringer Landesamt für Umwelt, Bergbau und Naturschutz Jena Germany

[3] Helmholtz Zentrum München, German Research Center for Environmental Health, Research Unit Analytical Biogeochemistry (BGC) Neuherberg Germany

[4] Centre for Ecology and Hydrology Wallingford/Bangor UK

[5] Politechnic University of Sinaloa Mazatlan Mexico

[6] Plant Ecology and Nature Conservation Group Wageningen University Wageningen The Netherlands

[7] Department of Soil Quality Wageningen University Wageningen The Netherlands

[8] Institute of Biology, Leipzig University Leipzig Germany

[9] Systematic Botany and Functional Biodiversity, Institute of Biology, University of Leipzig Leipzig Germany

[10] Research Group for Marine Geochemistry (ICBM-MPI Bridging Group), University of Oldenburg, Institute for Chemistry and Biology of the Marine Environment (ICBM) Oldenburg Germany

[11] Helmholtz Institute for Functional Marine Biodiversity University of Oldenburg Oldenburg Germany

*Corresponding author: mlange@bgc-jena.mpg.de

Abstract

Dissolved organic matter affects fundamental biogeochemical processes in the soil such as nutrient cycling and organic matter storage. The current paradigm is that processing of dissolved organic matter converges to recalcitrant molecules (those that resist degradation) of low molecular mass and high molecular diversity through biotic and abiotic processes. Here we demonstrate that the molecular composition and properties of dissolved organic matter continuously change during soil passage and propose that this reflects a continual shifting of its sources. Using ultrahigh-resolution mass spectrometry and nuclear magnetic

resonance spectroscopy, we studied the molecular changes of dissolved organic matter from the soil surface to 60 cm depth in 20 temperate grassland communities in soil type Eutric Fluvisol. Applying a semi-quantitative approach, we observed that plant-derived molecules were first broken down into molecules containing a large proportion of low-molecular-mass compounds. These low-molecular-mass compounds became less abundant during soil passage, whereas larger molecules, depleted in plant-related ligno-cellulosic structures, became more abundant. These findings indicate that the small plant-derived molecules were preferentially consumed by microorganisms and transformed into larger microbial-derived molecules. This suggests that dissolved organic matter is not intrinsically recalcitrant but instead persists in soil as a result of simultaneous consumption, transformation and formation.

Main

Dissolved organic matter (DOM) fluxes are an important component of the global carbon cycle[1,2] that enable cycling and distribution of carbon and nutrients[3]. DOM, initially leached from decomposing plant material or exuded directly (for example, as carbohydrates and organic acids) from roots, changes its characteristics during its transport through the soil[3,4]. The study of DOM may thus provide information about a range of processes within the zone from vegetation down to groundwater, known as the Critical Zone[5,6]. Organic matter production, its degradation, reprocessing, storage and its transport downward all take place in the Critical Zone[7]. These processes have been identified as the cause of rapid changes in terrestrial DOM properties as it passes through the soil profile and finally enters the groundwater. However, a general consensus on how these processes affect the molecular DOM properties during its downward transport is lacking. There has been debate for many years about the chemical nature of DOM molecules, namely whether there are molecules being intrinsically recalcitrant[8,9]. Recalcitrant molecules are assumed to be refractory and not easily decomposed because of their molecular properties. As a consequence of intrinsic recalcitrance, the same molecular structures containing the same atoms remain in the system. However, recent findings suggest that the turnover of soil organic matter with soil depth is largely controlled by its accessibility[10], concentration[11], bioavailability and biodegradability[12,13] but not primarily affected by any intrinsic recalcitrance of DOM molecules[8,14]. Instead, the concept of persistence[15] suggests that DOM molecules are

continuously found in soil because they are constantly recycled, transformed and assembled. As a consequence of persistence, the same molecular structures that contain atoms from different sources are found in the system. Vertical transport of DOM through the soil column is accompanied by a series of sorption and desorption processes together with microbial processing[4,12,16]. Aspects of structural changes of DOM with depth are visible even to the naked eye owing to preferential loss of light-absorbing structures (that is, coloured organic matter), changing the colour of DOM from yellowish-brown to transparent[7,17]. These changes are independent of the decrease of dissolved organic carbon (DOC) concentration. At the level of the compound classes, the hydrophilic fraction and aromatic C and lignin phenols[17,18] – that is, plant-derived organic molecules[19] – decrease with soil depth. In turn, altered and reprocessed microbial-derived molecules increase in abundance[4,16]. The stepwise degradation of lignin to water-soluble compounds[20] and then possibly to extended aromatic compounds[21] is therefore suggested to be an important contribution to DOM formation[22].

Sorption processes themselves change the composition of DOM during soil passage: for instance, lignin degradation products are more efficiently adsorbed than carbohydrates in, for example, clays and other reactive minerals[23]. Mineral-bound organic molecules can be remobilized by percolating surface-reactive plant-derived DOM[4] or acidic root exudates[24]. These changes in DOM are accompanied by microbial processing of organic molecules[4,16]. Although the microbial community is potentially able to decompose the vast majority of DOM molecules, the molecular properties of DOM strongly influence the efficiency of its degradation[13]. Thus, the microbial community adapts to and degrades key DOM components depending on their molecular properties at specific depths. This microbial degradation leads to changes in the DOM composition and structure with soil depth[12]. However, microbial processing involves not only consumption and degradation, but probably also the build-up of new molecules[25-27]. Applying non-targeted ultrahigh-resolution analytical techniques allows us to capture molecular-level information about DOM for a vast variety of molecules. This enhanced level of information describes the alteration of molecular functionalities and the molecular mass of DOM along gradients and provides access to a more detailed molecular-level understanding of the processes than has been previously available. This insight is needed to link processes that are shown to affect DOM degradation, transformation and

(re)synthesis during passage through the Critical Zone and to reveal the relative importance of the proposed mechanisms.

In this study we examine the molecular composition of soil DOM from 20 grassland communities from the surface to a depth of 60 cm using ultrahigh-resolution mass spectrometry (electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, ESI-FT-ICR-MS) and nuclear magnetic resonance spectroscopy (^{1}H NMR). The composition and properties of DOM identified by these ultrahigh-resolution methods are related to known environmental drivers of DOM composition, namely: (1) root standing biomass, a proxy indicative of plant-derived inputs, (2) microbial biomass and genetic diversity as proxies for the community of decomposers and recyclers and (3) the clay, carbon and nitrogen contents in the soil as a proxy for soil physical constraints. We hypothesize that the key drivers of the formation and transformation of DOM are mainly inputs from plants and their microbial conversion products, having even stronger effects than physical processes such as sorption to minerals.

Chemical DOM characteristics at different soil depths

In all soil profiles, DOC concentrations decreased with soil depth ($R2 = 0.41$; P value < 0.001), with the largest decrease observed at depths between 30 cm and 60 cm (Supplementary Table 1). Concentrations of dissolved organic nitrogen (DON) did not change with soil depth (P value $= 0.445$), and consequently the DON/DOC ratio increased. Over all samples taken from the 20 grassland communities, 4,264 molecular formulae of DOM compounds were identified (see Methods). Along the soil profile, the molecular composition of DOM continuously changed as the relative abundance of low-mass DOM molecules (mass-to-charge ratio $m/z = 150–275$) decreased with depth in comparison with those in the mid-molecular-mass range ($m/z = 300–450$) (Fig. 1a–d). Thus, the molecular composition of DOM became more dissimilar with larger distance between sampling depths (Supplementary Table 2). However, most of the detected molecules occurred at all depths, but with varying relative abundance. Therefore, the shift of molecular properties during downward transport, as revealed by differential mass spectra (see Methods), was maximal between DOM at depths of 10 cm and 60 cm (Fig. 1e–g; Table 1). In the differential mass spectra, based on the complete spectra (‘All compounds’ in Table 1), mainly aromatic CHO

compounds (containing only carbon, hydrogen and oxygen atoms, with aromaticity index $AI_{mod} > 0.54$; refs.[28]) of low molecular mass became considerably less abundant with depth.

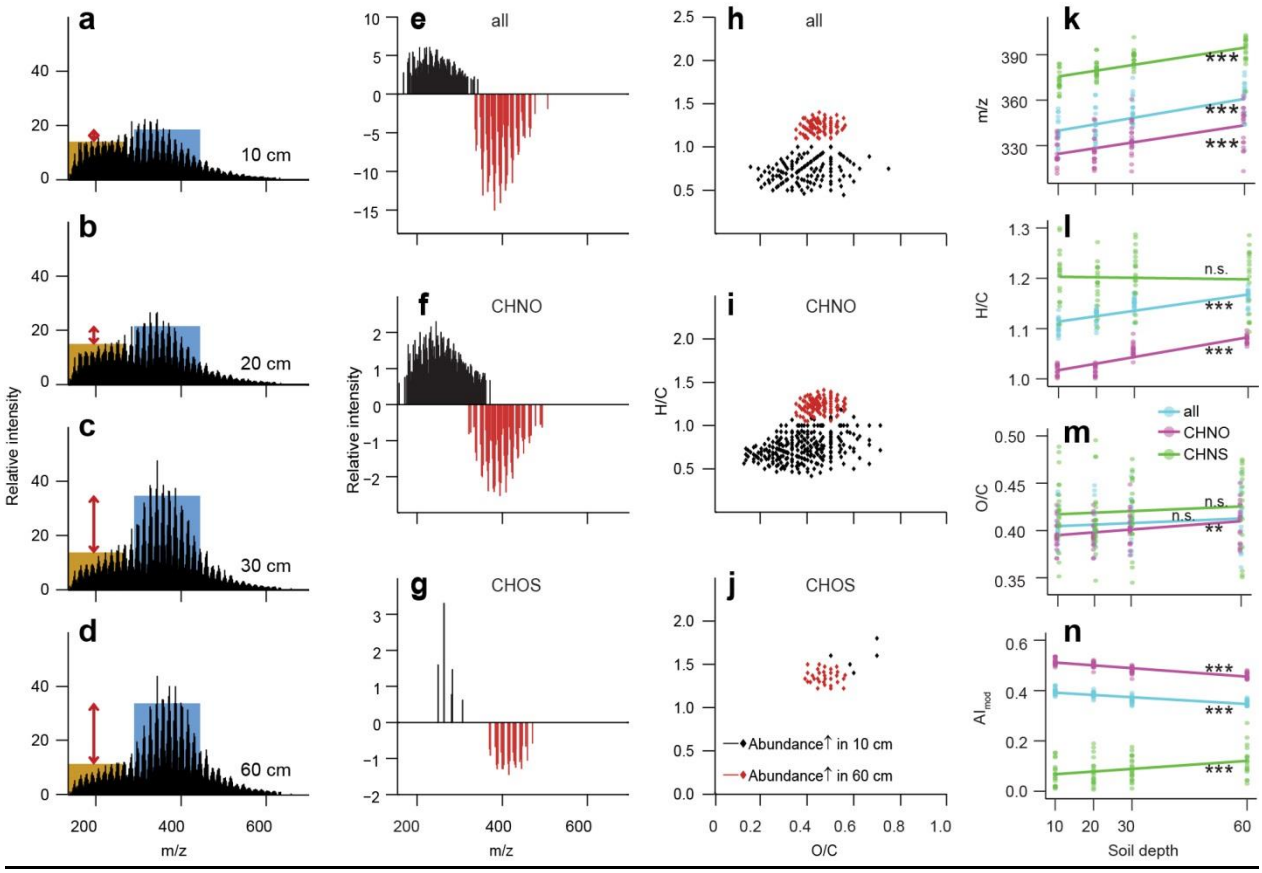

Fig. 1: Molecular changes in soil DOM based on FT-ICR mass spectra. a–d, DOM mass spectra at soil depths of 10 cm, 20 cm, 30 cm and 60 cm. The intensity distribution along the mass axis is bimodal. The intensity maximum for the low-molecular-mass range is highlighted yellow, and for the mid-molecular-mass range it is highlighted blue. **e–g**, Differential mass spectra of DOM sampled at soil depths of 10 cm and 60 cm (Methods). **h–j**, The corresponding van Krevelen diagrams. **k–n**, Linear regressions of the soil depth effect on *m*/*z*, H/C, O/C and AI_{mod} of DOM (***P* < 0.01, ****P* < 0.001, n.s., not significant, based on likelihood-ratio tests, see Methods). Values are based on the categories in Table 1: all compounds, only CHNO compounds or only CHOS compounds.

Compounds that were more abundant at 60 cm were mainly CHO compounds, too, but were characterized by remarkably similar, narrow O/C and H/C ranges in the mid-molecular-mass range (Fig. 1h), in which the number of isomers for any given molecular compositions is maximal[29]. These trends were also observed when considering only CHNO compounds (containing nitrogen besides C, H and O atoms) and, less specifically, when considering only CHOS compounds (containing sulfur besides C, H and O atoms; Table 1; Fig. 1i,j). The shift in abundances was remarkably consistent around *m*/*z* 300 for the complete spectra, the

CHNO compounds and the CHOS compounds, although the CHOS compounds were less numerous (Table 1).

^{1}H solution NMR spectra confirmed the changes in DOM molecular composition with depth that had been demonstrated by FT-ICR-MS. The relative proportions of aromatic and olefinic as well as those of O-alkyl carbon (indicative of carbohydrates and methoxy functional groups) decreased with depth, whereas the amount of saturated groups (aliphatic, sp^3-hybridized carbon) increased with depth (Supplementary Fig. 1). The relative proportions of acetate analogues and carboxyl-rich alicyclic molecules were nearly constant along the depth profile.

Table 1: Molecular properties derived from differential mass spectra shown Fig. 1e-g. Properties are shown for all, CHNO and CHOS compounds.

Compounds	Depth	m/z	No of formulae	H/C ratio	O/C ratio	AI_{mod}
all	10	165-341	125	0.44-1.00	0.15-0.75	0.45-0.85
	60	335-507	77	1.10-1.40	0.35-0.56	0.14-0.39
CHNO	10	154-370	251	0.42-1.22	0.13-0.71	0.36-1.00
	60	320-494	104	1.05-1.41	0.35-0.56	0.16-0.45
CHOS	10	247-275	5	1.40-1.80	0.50-0.70	0.00-0.23
	60	369-471	31	1.22-1.50	0.39-0.56	0.00

Drivers of the DOM transformation during soil passage

The decline of the low-molecular-mass compounds towards more prominent mid-molecular-mass compounds over depth (for the categories in Table 1: all compounds, only CHNO compounds and only CHOS compounds) was reflected by a large increase in the weighted mean of *m*/*z* (Fig. 1k; Supplementary Table 3). The shift in molecular mass was accompanied by a considerable decrease in abundance of unsaturated molecules, as indicated by the weighted mean of the H/C ratio and a decrease of the weighted mean of DOM aromaticity in all compounds and CHNO compounds (Fig. 1l,n). The saturation of CHOS compounds did not change with soil depth, whereas AI_{mod} increased substantially (Fig. 1l,n). The weighted mean of the O/C ratio of the CHNO compounds slightly increased with depth, whereas the weighted mean of the O/C ratio of all compounds and CHOS compounds did not change with depth (Fig. 1m).

The effect of depth on the molecular properties and composition of DOM has mostly been accounted for by the decrease of root biomass, soil organic carbon and soil nitrogen along the soil profile, but not by the soil clay content (Supplementary Fig. 2). However, the decrease of root standing biomass, organic carbon and nitrogen along the soil profile were highly inter-correlated (Supplementary Table 4). To estimate the relative importance of potential drivers of the DOM transformation, we compared the influences of (1) plant inputs (root standing biomass), (2) soil properties (contents of clay, organic carbon, nitrogen) and (3) the microbial community (bacterial and fungal biomass and their respective genetic diversity). To eliminate the inter-correlation among the variables with depth, this estimation was performed for the topsoil. Variation partitioning showed that the chosen parameters accounted for the molecular DOM variation between 67% and 46% (Supplementary Fig. 3). The soil microbial community was the most important predictor (explaining on average 25% of the molecular DOM variation) followed by the soil properties (19%) and root biomass (10%, Fig. 2a). In addition, a large part of the variation (13%) was jointly explained by the microbial community and root biomass. Bacteria generally explained more of the variability than fungi (Fig. 2b; Supplementary Fig. 3). However, more than 40% of the DOM variation remains unexplained (Fig. 4, residuals), suggesting that other environmental variables contribute to DOM variability. This unexplained DOM variability might be related to the chemical composition of soil clay, the exudate and root chemistry (influenced by plant composition and diversity) or to the composition of the community of decomposers and higher trophic levels. Furthermore, a significantly positive correlation of the chemodiversity with the bacterial diversity ($r = 0.46, P = 0.048$) but not with the fungal diversity ($r = -0.36, P = 0.127$) confirmed the strong relation of the bacterial community to molecular DOM composition. Interestingly, this important relationship was not related to the diversity of DOM compounds (calculated as the Shannon index) but to the dominance structure of the DOM samples (Simpson's index[30]). This positive relation indicates that a more diverse bacterial community results in a more evenly distributed DOM composition (Supplementary Fig. 4).

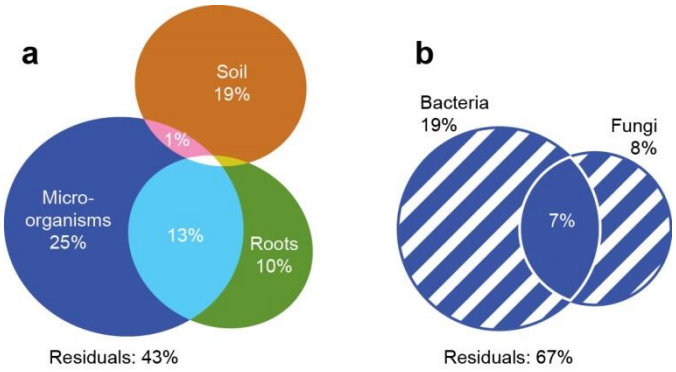

Fig. 2: Variation partitioning for potential drivers of DOM transformation. a, Variation explained by soil, roots and microbial community. **b**, Variation explained by bacteria and fungi only. Diagrams show the average explained variations of *m*/*z*, H/C, O/C, AI_{mod} and molecular composition (see detailed results in Supplementary Fig. 2). Variation partitioning is based on data gathered at topsoil (Methods).

A semi-quantitative approximation of the abundances of individual compounds was applied to evaluate whether the increased weighted mean of *m*/*z* with depth is a result of the disappearance of the low-molecular-mass compounds or of new formation of mid-molecular-mass compounds. Therefore, FT-ICR-MS signal intensities of individual molecular formulae were scaled to DOC concentrations[31]. This approximation indicated that the increase of the weighted mean of *m*/*z* with soil depth was caused by both an absolute decline in abundance of low-molecular-mass compounds and by an absolute increase in abundance of mid- and high-molecular-mass compounds (Fig. 3a). Although the abundance of low-molecular-mass compounds continuously decreased with soil depth, the mid- and high-molecular-mass compounds became more abundant between depths of 10 cm and 30 cm and declined below 30 cm. Unexpectedly, the shift in the ratio of intensities of low- to mid-molecular-mass compounds (I_{mid}/I_{low}) was strongly related to the degradation state of DOM ($R^2 = 0.71$), described as the degradation index I_{deg} (see Methods[32]) (Fig. 3b). Thus, both DOM degradation increased (higher I_{deg}) and DOM molecular mass increased (higher I_{mid}/I_{low}) with soil depth.

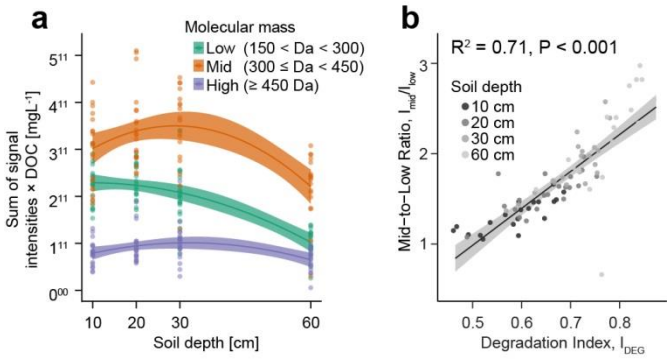

Fig. 3: Shift of molecular DOM masses during soil passage. a, Sum of absolute FT-ICR-MS signal intensities for formulae of low, mid and high molecular masses. **b**, The ratio of relative intensities for small- to mid-molecular-mass compounds (I_{mid}/I_{low}) is related to an established degradation index of DOM (I_{deg}, ref.[32]). Shaded areas represent 0.95% confidence intervals.

Mechanisms underlying the persistence of DOM

The results presented here strongly indicate that there is microbial processing, but not production of recalcitrant products, as DOM moves through soils and sediments. The continual shift of molecular composition and properties of DOM during its soil passage furthermore indicates an ongoing transformation of DOM at the molecular level (Fig. 4). The most unexpected change along the soil profile was the shift from low-molecular-mass to mid-molecular-mass compounds, which seems to be a general pattern, independent of land use and soil (see Supplementary Fig. 5 for supplementary grassland and forest sites). At 10 cm depth, the most abundant compounds are characterized by a high level of unsaturation, related to polycyclic aromatics ($AI_{mod} \geq 0.67$), including condensed combustion-derived dissolved black carbon (C atoms $n > 15$), and other highly aromatic compounds ($0.69 \geq AI_{mod} \geq 0.54$)[22,28]. The highly aromatic compounds may include polyphenols and polycyclic aromatics with aliphatic side chains[28]. Oxygen-poor black carbon (such as polycyclic aromatic hydrocarbon) as a precursor of these compounds may be of anthropogenic origin because the field site is close to an area that was formerly highly industrialized. Microbial alteration of these compounds was shown in a recent study conducted at the same site[33]. Those authors reported considerable amounts of polycyclic aromatic compounds that undergo biodegradation, resulting in low-molecular-mass oxygenated polycyclic aromatic hydrocarbons. The most abundant compounds at 60 cm depth are characterized as highly unsaturated but not aromatic compounds ($AI_{mod} < 0.54$ and $H/C < 1.5$), which could include residues of microorganisms and degradation products of soil

organic matter as well as carboxylic-rich alicyclic molecules[34]. Plant-derived DOM is considered to be more unsaturated and aromatic than microbial-derived substances because of the higher lignin content in plant tissues and other secondary aromatic plant metabolites[35,36]. In consequence, our data suggest that the observed changes in molecular properties are probably driven by a fast microbial turnover[26] of plant-derived DOM during downward transport[4,18], which has a high degree of relative unsaturation and aromaticity (Fig. 4). The results of both FT-ICR-MS and NMR spectra suggest that aromatic substances are degradable[20] and could be used as indicators of fresh, plant-derived organic matter, or early products of its decomposition. The even greater decrease in abundance of molecular functional groups similar to carbohydrates is as expected[17]. These functional groups might be derived from plant products, such as carbohydrates and cellulose.

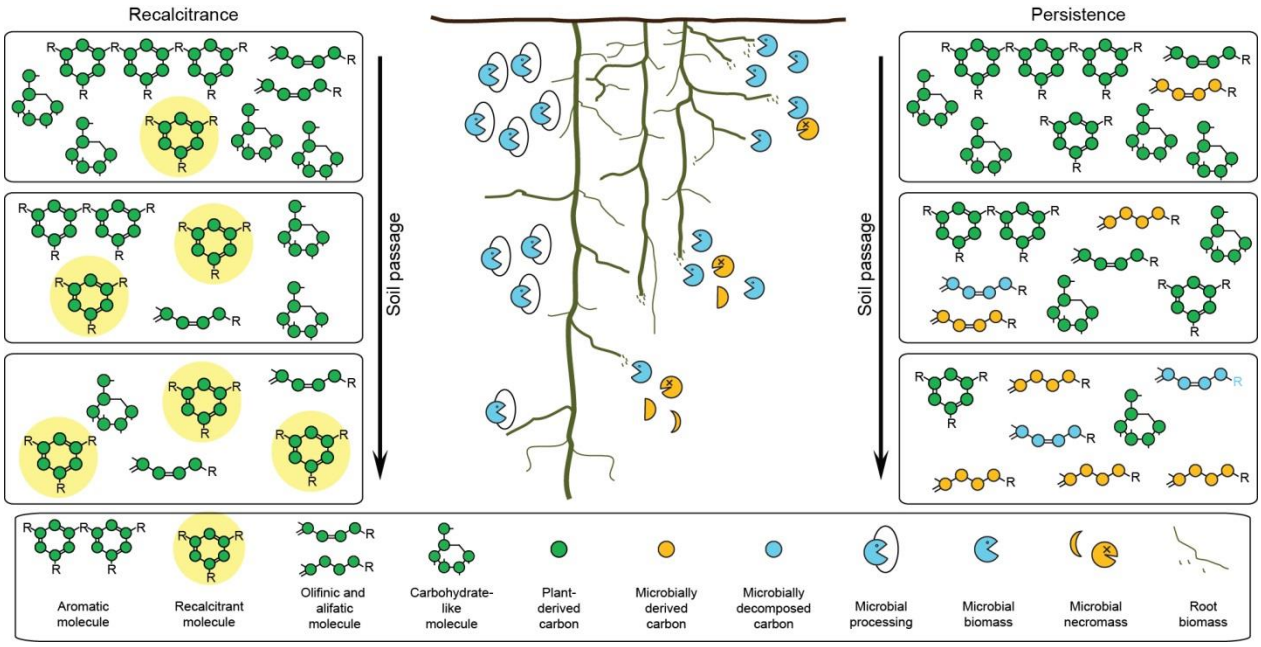

Fig. 4: Proposed mechanisms for spatial and temporal evolution of DOM molecular structures during soil passage. Our results indicate that chemical recalcitrance is not the primary mechanism that preserves small DOM molecules from decomposition. Instead we found that the decomposition of DOM molecules increases with depth. Our findings suggest that the persistence of DOM molecules in soil is due to microbial transformation and that DOM consumption is accompanied by formation of new microbial-derived compounds. The consumption, transformation and production by microorganisms initially lead to a preferential degradation of large plant-derived polymers (such as lignin), partial mineralization and transformation into a diverse suite of small molecules that are subsequently consumed by the soil microbial community. As indicated by the decreasing aromaticity and unsaturation with soil depth, DOM molecules found in deeper soils are mainly of microbial origin and are either decomposition products or remnants of bacterial necromass.

In our study, the soil clay content had only minor effects on the molecular properties of DOM during its soil passage. This contrasts with many studies reporting a strong impact of soil mineralogy on DOM[3,4,12,15]. This discrepancy might be attributed to the fact that the soil clay content is a good proxy for mineral stabilization on regional or global scales[15], but not on local scales as in our study. The scale-dependency of the mineral-DOM relation is supported when comparing the changes in the DOM mass spectra with soil depth between the Jena site and the supplementary sites (see Methods). Although the depth effect found at the Jena site (Fig. 1) was confirmed by DOM analyses from supplementary grassland and forest sites (Supplementary Fig. 5), the mass change from small compounds to mid-molecular-mass compounds was more pronounced at the Jena site. This might be due to a finer grain size at the Jena site than at the supplementary site, which has a very sandy soil with probably low sorption capacity (see Methods). This indicates (1) a general pattern that the simultaneous decrease of small compounds (150–275 Da) and increase of mid-size compounds (300–450 Da) is independent of the soil properties, and (2) that the soil properties affect the strength of this mass shift, indicating the microbial−mineral interaction during the DOM soil passage[4]. The generality of the mass shift, independent of soil properties, is in line with a recent study showing that the biochemical composition of mineral-retained organic matter was similar across four different classes of clay minerals[37]. However, describing the soil on the basis of its texture is a relatively broad parameter with which to assess mineral−organic interactions in soils. Given that the strength of mineral−organic interactions is also reflected by the soil organic carbon content[38], we have statistically considered the soil organic carbon content in our analyses, as this probably controls the mineral−organic interactions. However, we acknowledge that additional laboratory experiments (such as our incubation experiment, Supplementary Fig. 6) could also provide detailed mechanistic insights into the mineral−organic interactions to which DOM is exposed during its passage through the soil profile.

Although plant-derived inputs in soil water decreased rapidly with soil depth[16], we found that root standing biomass, which is a proxy indicative of belowground plant-derived carbon inputs, influenced DOM molecular properties along the entire soil profile. This finding supports the observation that microbial processing and reworking is responsible for changes in the molecular properties of DOM with depth[4,16,19]. The shared variation in DOM accounted for by the microbial community and root biomass, and the prominent role of the

microbial community in explaining DOM variation indicate a strong microbial imprint and transformation of plant-derived inputs into the DOM pool. Recent findings have shown that the microbial community also adapts to changes in DOM characteristics at different soil depths[12]. This highlights the dual role of microorganisms in the carbon cycle in soils, decomposing organic matter while simultaneously recycling and producing new molecules. The shift in the abundance of DOM molecules around 300 Da further suggests a selective removal of molecules smaller than 300 Da. Such size-selective processes are well known from the transport of small molecules across membranes of plant roots and soil microorganisms. The outer membranes of Gram-negative bacteria, as well as mitochondria and plastids, contain water-filled transport channels, so-called porins[39]. Molecules of <600 Da can diffuse through porins with a size-dependent diffusion rate that allows for faster uptake of smaller molecules[40]. Thus, the selective removal of small molecules from fresh DOM might be related to root and microbial uptake through porins[41]. However, because of the increased aliphaticity with soil depth, we propose that microbial uptake is the main driver of the shift in the molecular mass of DOM. The microbial processing of DOM may further explain the decrease in its concentrations with depth, while the concentrations of DON remained at the same level[17]. This suggests a microbial recycling of N or preservation of N-containing molecular structures, while C is partly mineralized. In agreement with the absolute increase of the mid-molecular-mass compounds in the upper 30 cm of the soil profile (Fig. 3), this strongly points to the formation of new microbial DOM products with higher average masses during soil passage and not to simple dilution of plant-derived DOM in deeper soil layers because of decreased inputs of plant material. Recently, it has been demonstrated that fresh DOM from primary producers is also characterized by low-molecular-mass compounds, whereas the so-called ‘refractory’ DOM is clearly shifted to higher masses[42]. The shift to higher molecular masses indicates that the decline of small molecules is related to the progressive degradation of plant-derived DOM and microbial resynthesis of new DOM molecules. These molecules might have high proportions of peptides or proteins, thereby explaining the approximately constant abundance of DON. We performed a supplementary plant decomposition experiment in which DOM extracted from fresh plant material (Supplementary Fig. 6) was dominated by large, source-specific compounds with a relatively high molecular mass (>500 Da). These large plant-derived

compounds decreased during the first weeks of decomposition. In addition, the abundance of low-molecular-mass compounds and decomposition-related compounds increased.

Increasing production of low-molecular-mass compounds during the early decomposition of plant-derived carbon supports the common opinion that large biopolymers are rapidly degraded during decomposition[43,44]. Our results do not support the paradigm of the build-up of humic polymers[15]. Instead, we suggest that the plant material is extracellularly decomposed to smaller molecules, which are then consumed and in part mineralized or transformed to larger microbial-derived molecules that form a secondary pool of soil organic matter and DOM (Fig. 4). Our observations support the conclusion that low-molecular-mass compounds contain early decomposition products that might be closely related to the source materials[45]. The emerging higher-molecular-mass compounds indicate the formation of new molecules from the decomposer community, namely microbial tissues, fragments or products and not random polymerization products. Consequently, the processing of DOM does not primarily lead to recalcitrant molecules, but DOM persists in the Critical Zone because it is constantly decomposed, recycled and newly formed (Fig. 4). The vertical transport of DOM through the soil is assumed to be central to the functioning of soil, for example, in the formation of soil organic carbon along the soil profile[46]. Many soil processes, important for soil functioning, are driven or mediated by microorganisms. DOM is therefore both a source of nutrients and energy for microorganisms in deeper soils[12] and its molecular characteristics are shaped by the microorganisms above. However, DOM in deeper horizons might be primarily a function of the processes at that depth, that is, with depth DOM composition may become more independent of soil passage, reflecting instead the complex biogeochemical processes at depth[4]. Therefore the interplay between soil microorganisms and DOM denotes the key to understanding the functioning of the Critical Zone, and ultrahigh-resolution mass spectrometry provides the means of disentangling this interaction.

Methods

Field sites and sampling

Soil water samples were collected on a semi-natural grassland with different grassland communities, being part of The Jena Experiment[47]. The site, located in Jena, Germany (50° 55′ N, 11° 35′ E, altitude 130 m) close to the Saale River, was converted from grassland to arable field in the early 1960s and ploughed to a depth of about 30 cm. The field site with

its current management as grassland was established in 2002. The soil of the field site is classified as Eutric Fluvisol (FAO-Unesco 1997), developed from loamy fluvial sediments up to 2 m thick, almost free of stones. As expected for a lowland river floodplain setting, sand content correlates with distance from the Saale river ($r = 0.95$). Close to the river, the topsoil consists of sandy loam, gradually changing into a silt loam with increasing distance to the river. Although soil texture varied considerably among the entire field site, the variation in pH (7.1–8.4), soil organic carbon (5–33 g C kg^{-1}) and total soil nitrogen concentrations (1.0–2.7 g N kg^{-1}) was smaller. This study was carried out on a subset ($n = 20$) of all grassland communities (plots, $n = 80$), that were fully equipped at depths of 10 cm, 20 cm, 30 cm and 60 cm with glass suction plates (pore size 1−1.6 µm, 1 cm thickness, 12 cm in diameter; from UMS, installed in April 2002)[16]. The investigated plots were aligned parallel to the river, with soil being dominated by silt (57.3 ± 5.0 s.d.), whereas the portions of clay (22.7 ± 2.8 s.d.) and sand (20.0 ± 7.5 s.d.) were relatively similar. Soil water was sampled on 7 May 2014 after two weeks of continuous sampling. The sampling bottles were evacuated to a negative pressure of 250 mbar, so that the suction pressure was approximately 50 mbar above the actual soil water tension. Thus, only the soil leachate[48] was cumulatively collected for two weeks.

In addition to the main site in Jena, soil water was analysed from a site located in the northeast of Germany (Linde, Brandenburg 52° 32′ N 12° 39′ E, altitude 45 m), with different soil and land-use properties. This site is characterized by a relatively acidic soil pH (forest 5.78 ± 0.3, grassland 5.05 ± 0.41). Soils have developed on aeolian sands and show mainly features of podzols and cambisols[49]. In November 2014, suction plates, identical to those at the Jena site, were installed on a grassland, and on stands of oak (*Quercus robur*) and pine (*Pinus sylvestri*), at depths of 5 cm, 10 cm, 20 cm, 30 cm and 60 cm. Samples were taken in February 2016. Water sample treatment was conducted as described for Jena samples.

Soil water sample preparation

On subsamples of the soil water samples pH, concentration of DOC (highTOC analyser, Elementar Analysensysteme) and dissolved organic nitrogen (DON) were measured. DON was quantified by subtracting the amount of inorganic nitrogen NH_4^+ (ICS-5000, Thermo Fisher Scientific), NO_2^- and NO_3^- (Dionex DX-500, Thermo Fisher Scientific) from the

amount of total bound nitrogen (TN-100, a1 environsciences). The remaining samples were acidified to pH 2 (HCl, pro analysi) and stored at 2 °C until DOM was concentrated and desalted by solid phase extraction[50]. In brief, we used Agilent Bond Elute PPL solid phase extraction cartridges (1 g) that were soaked in methanol overnight. Prior to extraction, the cartridges were rinsed with ultrapure water, methanol and ultrapure water acidified with HCl to pH 2. Considering the respective DOC concentrations determined from the subsamples, the volume of soil water for extraction was adjusted to load 2 mg organic carbon onto the columns. Acidified ultrapure water was stored in the same type of bottles as the soil pore water samples, to be used as procedural blanks. After loading the solid phase extraction cartridges with sample, the cartridges were rinsed with acidified ultrapure water and dried with nitrogen. The DOM extracts were eluted with methanol. At each day of extraction a process blank extract was produced. The average extraction efficiency for soil water DOM was 69% ± 6% s.d. on a carbon basis for the samples from the main side and 79% ± 7 s.d. for samples from the supplementary site. Redundancy analysis confirmed that the extraction efficiency had no influence on the molecular composition of DOM (explained variation = 0.7%, pseudo-F = 0.6, P = 0.722).

Assessment of soil carbon and nitrogen, roots and the microbial community

Soil samples were taken in April 2012 at each of the plots at the main site to a depth of 1 m using a machine-driven soil corer (inner diameter of 8.7 cm, Cobra, Eijkelkamp Agrisearch Equipment). Soil cores were segmented into 5 cm depth sections. Air-dried soil samples were milled and subsamples were analysed for organic C and total N with a Vario Max and a Vario EL (Elementar Analysensysteme), respectively. Organic C was determined as the difference between the total C content and the inorganic C content measured after heating the sample to 450 °C for 16 h in a muffle furnace[51].

Root standing biomass was sampled in 2014, following the protocol of previous years[52]. For the root sampling campaign, three soil cores (inner diameter 4.0 cm, Eijkelkamp Agrisearch Equipment) per plot were taken to a depth of 40 cm. Cores were segmented into five layers: 0–5 cm, 5–10 cm, 10–20 cm, 20–30 cm and 30–40 cm, which were pooled in the field. Samples were stored at 4 °C until they were washed over a 0.2 mm sieve, which took place within 7 days. Clean root biomass was dried at 70 °C for 72 h and weighed. Root biomass was calculated as milligrams dry mass per cubic centimeter[52].

The microbial community was assessed using the phospholipid fatty acids (PLFA) method of identifying bacterial and fungal biomass, and by sequencing for evaluating bacterial and fungal diversity. For both measures identical soil samples were used. In early May 2012, three soil samples per plot were taken with a corer (inner diameter 4.8 cm, Eijkelkamp Agrisearch Equipment) to a depth of 5 cm, and pooled. Within 48 h after sampling the soil was kept at 4 °C, sieved to 2 mm, remains of roots were manually removed and the samples were stored at −20 °C until further sample processing. PLFA were extracted according to the method of Bligh and Dyer[53] as modified by Kramer and Gleixner[54,55]. As indicator for fungal biomass PLFA 18:2ω6 was used[56,57]. The bacterial biomass was calculated as sum of the PLFA markers 15:0i, 15:0a, c15:0n, 16:0i, c16:1ω7c, c17:0br, 17:0i, 17:0a, 17:1ω8, 17:0cy, 18:1ω7 and 19:0cy[56].

For sequencing analyses of bacterial and fungal communities DNA was extracted from 0.3 g of soil using the MoBIO PowerSoil-htp 96 Well DNA Isolation kit according to the manufacturer's protocols. The dual indexing protocol of Kozich et al.[58] was used for Illumina MiSeq sequencing of the V3-V4 hypervariable regions of the bacterial 16 S rRNA gene (primers 341F[59] and 806R[60]); and the ITS2 region for fungi (fITS7f and ITS4r primer sequences[61]). Amplicon concentrations were normalized using the SequalPrep Normalization Plate Kit (Thermo Fisher Scientific) prior to sequencing on the Illumina MiSeq using V3 chemistry. Fungal ITS sequences were analysed using PIPITS[62]) with default parameters as described. Similar approaches were used for analyses of bacterial sequences, using PEAR (http://sco.h-its.org/exelixis/web/software/pear) for merging forward and reverse reads, quality filtering using FASTX tools (hannonlab.cshl.edu), chimera removal with VSEARCH_UCHIME_REF and clustering to 97% operational taxonomic units (OTUs) with VSEARCH_CLUSTER (http://github.com/torognes/vsearch). Both bacterial and fungal OTU abundance tables were resampled to a minimum of 4,000 reads per sample prior to calculating indices of diversity (the Shannon index and Simpson's index on the basis of OTUs) using the 'diversity'function in the 'vegan' package[63].

FT-ICR-MS

For the FT-ICR-MS measurements extract aliquots were diluted to 20 mg l^{-1} organic carbon in ultrapure water/methanol (1:1). The Bruker Solarix FT-ICR-MS (15 tesla) at the University of Oldenburg (Germany) was used. Samples were continuously injected into the

ESI source with a flow rate of 120 µl h^{-1} and an ESI needle voltage of −4 kV in negative ionization mode. 500 single scans with an ion accumulation time of 0.2 s were recorded over a mass range of *m*/*z* 150 to 2,000 and added to one spectrum. An in-house mass reference list was used for internal calibration.

To consider only those masses that had been measured to be relevant for statistical analyses, several criteria, also described in ref.[64], were applied. First, only *m*/*z* values with a signal-to-noise ratio of the maximum of each *m*/*z*value ($S/N_{max,i} > 5$) were considered[65]. To determine $S/N_{max,i}$ the maximum relative intensity of each *m*/*z* value was divided by the noise. Second, only *m*/*z* values were kept that were detected in more than one sample. Third, all *m*/*z*values with an S/N_{blanks} ratio ≥ 20 were removed. To determine S/N_{blanks} the average of signal intensity across all measured blanks was divided by the noise. Fourth, *m*/*z* of low intensity ($S/N_{max} < 20$) were removed if they were detected in less than 20% of the measurements. Mass-to-charge ratios with assigned molecular formulae meeting all the above criteria (4,264 different *m*/*z* values) were isolated from the remaining *m*/*z* list and normalized to the sum of intensities. For molecular formula assignment (C, H, O, N, S and P) an in-house algorithm that is based on Koch et al.[66] and Stenson et al.[67] was used. Ions with $m/z > 660$ were not detected in our samples. Only singly charged ions were considered. In consequence, the *m*/*z* values represent the molecular mass (in dalton) of the detected ions. Matlab R2013a (The MathWorks) and R[68] were used for data preparation and evaluation.

NMR

Solid phase extraction DOM was pooled depthwise to samples of 3.5 mg DOC, with equal shares of each individual sample (20 samples for 10−30 cm, 14 samples for 60 cm), resulting in one representative sample for each depth. ^{1}H NMR spectra of solid phase extracted (Agilent Bond Elute PPL) DOM were obtained with a Bruker Avance III 800 MHz NMR spectrometer (typically 0.5 mg in 300 µg deuterated methanol (CD3OD); Bruker 3 mm sealed MATCH (multiple adjustable tube clamp holder) tubes). Proton spectra were acquired at 283 K with a 5 mm z-gradient ^{1}H / ^{13}C / ^{15}N / ^{31}P QCI cryogenic probe (in CD_3OD, Merck, 99.95% ^{2}H). One-dimensional ^{1}H NMR spectra were recorded using the first increment of the presaturation-NOESY (nuclear Overhauser effect spectroscopy) sequence; solvent suppression with presaturation and spin-lock, 5 s acquisition time, 15 s

relaxation delay, typically 64–512 scans, 1 ms mixing time, 1 Hz exponential line broadening.

Data analysis

Weighted means of formula-based characteristics (m/z, AI_{mod} and H/C) were calculated as the sum of the product of the individual information (m/z_i, $AI_{mod,i}$ or H/C_i) and relative intensity I_i divided by the sum of all intensities (for example $m/z_{wm} = \mathrm{sum}(m/z_i I_i)/\Sigma(I_i)$). H/C provides information about the saturation and the modified aromaticity index (AI_{mod}) was used to estimate the aromaticity of individual formulae ($AI_{mod} \leq 0.54$ non-aromatic, $0.54 < AI_{mod} < 0.69$ aromatic, $AI_{mod} \geq 0.69$ condensed aromatic)[28]. To relate mass changes to processes, the shift in masses were examined in relation to a degradation index (I_{deg}, ref.[32]). Therefore the normalized intensities of formulae negatively related to degradation ($I_{deg,neg}$: $C_{21}H_{26}O_{11}$, $C_{17}H_{20}O_9$, $C_{19}H_{22}O_{10}$, $C_{20}H_{22}O_{10}$, $C_{20}H_{24}O_{11}$) and formulae being positively related to degradation ($I_{deg,pos}$: $C_{13}H_{18}O_7$, $C_{14}H_{20}O_7$, $C_{15}H_{22}O_7$, $C_{15}H_{22}O_8$, $C_{16}H_{24}O_8$) were summed. I_{deg} was calculated based on equation (1)

$$I_{deg}=\sum(I_{deg,neg})/\sum(I_{deg,neg}+I_{deg,pos})I_{deg}=\sum(I_{deg,neg})/\sum(I_{deg,neg}+I_{deg,pos}) \quad (1)$$

Furthermore, the median intensity maximum of the low-molecular-mass (m/z 150−275) I_{low} and mid-molecular-mass (m/z 300−450) I_{mid} range was introduced. For this purpose 1% of the intensities ranked according to intensity for the low- and mid-molecular mass range was identified. The median of the identified intensities was set as I_{low} and I_{mid}, respectively. To avoid interference from overlap, the range m/z 275−300 was neglected. To compare spectra with respect to the intensity of their low- and mid-molecular-mass ranges the ratio I_{mid}/I_{low} was used.

To display differences between the mass spectra of samples from depths of 10 cm and 60 cm, differential mass spectra were calculated. The differential mass spectrum is the result of subtracting the relative mass peak intensities of the 10 cm measurement from those of the 60 cm measurement. Resulting positive intensities represent formulae that are of higher abundance in 10 cm samples whereas negative intensities indicate formulae that are of higher abundance in 60 cm ones. To consider only formulae that were of much higher abundance for each depth, a threshold value (10% of the median of 1% of the highest absolute intensities)[64] was introduced. This threshold value was determined for each differential mass spectrum individually and the average value (1.35, s.d. = 0.06) was applied

to all differential mass spectra. The differential mass spectra were calculated for each plot separately. To focus on the common depth dependent trend of all plots, those formulae that occurred in 90% of all differential mass spectra were identified.

The analyses for weighted means (*m*/*z*, H/C, AI_{mod}) and differential mass spectra were performed for all molecular formulae (4,264 different formulae) but also for subsets of formulae that either contained all formulae with at least 1 N (1,704 formulae) or 1 S (511 formulae). The datasets based on different formulae were named 'All compounds', 'CHNO compounds' or 'CHOS compounds'.

To investigate the general differences in the molecular DOM composition at each soil depth, the Bray−Curtis distances[69] between each depth for each plot were calculated. The averaged Bray-Curtis distances between each depth are given in Supplementary Table 2. Linear mixed-effects models were used considering the repeated measurements on the same plot along the soil profile to statistically test whether depth-dependent effects exist independently of the plot identity (the 'lme' function in the nlme package[70]).

Variation partitioning analyses[71] were performed based on the comparison of variance explained by linear models including every possible combination of variables being proxies for mineralogy (clay content), plant derived C (root standing biomass) and microbial communities (biomass and genetic diversity). A series of seven models was fitted for each bacterial and fungal community to extract the unique and shared variance for each combination of variables (mineralogy only, plant C only, microbial community only, mineralogy + plant C, mineralogy + microbial community, plant C + microbial community, and all predictors together). Venn diagrams with two factors were displayed using the compute.Venn function in Vennerable; Venn diagrams with three factors were displayed using Euler APE for Windows[72]. To relate the bacterial and fungal diversity to the chemodiversity of DOM compounds, the Shannon index and Simpson's index were calculated based on the measured DOM compounds (molecular formulae) and their relative intensities (ion abundances) using the 'diversity' function in the 'vegan' package[63]. The relations between microbial diversity and chemodiversity were determined using the Pearson correlation coefficient (the cor.test function[68] in R[68]).

Supplementary incubation experiment and Orbitrap DOM measurements

Incubation experiment

Dried and ground plant shoot material (*Bromus erectus, Leucanthemum vulgare* agg., *Medicago varia*) were inoculated with a mixture of sand (Sigma-Aldrich, baked at 500 °C for 4 h), based on published experimental setups[73,74]. Plant and soil material were sampled in May 2016, at the Jena site[47]. Mineral soil from depth <10 cm was floated (200 g and 400 g soil per litre, depending on C content) in ultrapure water, for one hour and the supernatant suspension was filtered (0.7 µm, Whatman glass microfibre filters, grade GF/F) to obtain the aqueous inoculum. Incubations were done in 50 ml polyethene (PE) Falcon tubes filled with 5 g of pure sand (drainage) and 10.5 g of sand mixed with plant material (plant-to-sand ratio of 1:20). The contribution of inoculum solution to C stocks was <0.1%. Samples were kept at 80% water-holding capacity over the course of the experiment. Tube lids were closed but air exchange was allowed. Incubations were done in triplicate for each time point (0 weeks, 1 week, 2 weeks and 3 weeks after inoculation) and blank incubations were carried along. Samples were extracted three times with 30 ml ultrapure water and ultracentrifuged after each extraction step (3,500 min^{-1}). The unified supernatant was vacuum-filtered (700 hPa, 0.7 µm GF/F), acidified to pH 2 (HCl, Roth) and extracted by solid phase extraction[50].

DOM analyses of supplementary soil water samples and incubation experiment

Methanol extracts were diluted to 20 mg l^{-1} DOC with ultrapure water and directly infused into an Orbitrap Elite ultrahigh resolution mass spectrometer (Thermo Fisher Scientific) equipped with a negative-mode ESI source. Instrumental settings for the incubation sample set were: Flow rate, 7 µl min^{-1}; spray voltage, 2.65 kV; source fragmentation, 40 eV; capillary temperature, 275 °C; S-Lens RF level, 70%; automatic gain control setting, 1E6; accumulation time, maximum 100 ms; scan range, 150–1,500 *m*/*z*; transient length, 0.8 ms; nominal resolution, 240.000, scans collected, 300. These settings were slightly changed to measure the soil water sample set from Linde (changed parameters: scan range, 150–1,000 *m*/*z*; transient length, 1.6 ms; nominal resolution, 480.000, scans collected, 100).

All measurements were done within days and in random order. In-house reference samples[75] were used to check instrumental stability. External calibration was done every day according to the manufacturer's protocol. Raw data were averaged in Xcalibur (Thermo Fisher Scientific), transformed with Proteo Wizard[76] and further processed in mmass[77]. Peak picking was done at 80% peak height and S/N of 5. Further processing followed similar rules

as described in Methods section 'FT-ICR-MS'. In addition, a peak occurrence filter was applied to the data. In the Linde soil water sample dataset, only peaks were included that were detected in more than 10% of samples. In the incubation dataset duplicate measurements were used and only those signals were kept that had been detected in both measurements[75].

Acknowledgements

We thank U. Gerighausen for sampling and K. Klapproth for technical support with FT-ICR-MS measurements. This work was supported by the Zwillenberg-Tietz Stiftung and the Deutsche Forschungsgemeinschaft as part of the Critical Zone Observatory 'AquaDiva' (CRC 1076) and the Jena Experiment (FOR 1451, GL 262/14 and GL 262/19). The International Max Planck Research School for Global Biogeochemical Cycles (IMPRS-gBGC) provided the funding for the PhD scholarships of P.G.M.-V. and C.S.

Author contributions

V.-N.R. and G.G. conceived and designed the study. M.L., V.-N.R. and G.G. wrote the main text. V.-R.N. and T.D. measured and processed MS data, N.H. obtained the NMR data. M.L. and V.-R.N. analysed the data. V.-N.R. and S.B. performed the supplementary decomposition experiment and C.S. measured, processed and analysed the data from the supplementary sites. L.M., N.J.O. and A.W. provided root standing biomass data, P.G.M.-V. provided data on microbial biomass, and R.I.G. and T.G. provided data on microbial diversity. All authors reviewed and edited the manuscript.

Data availability

The compiled dataset used in our analyses is available at https://doi.org/10.17617/3.28 and root standing biomass at https://doi.org/10.1594/PANGAEA.880324. The raw data are available from the corresponding author (M.L.) on request.

Code availability

The codes used for this study are available on request.

Competing interests

The authors declare no competing interests.

References

1. Battin, T. J. et al. The boundless carbon cycle. *Nat. Geosci.* **2**, 598–600 (2009).

2. Roulet, N. & Moore, T. R. Environmental chemistry. Browning the waters. *Nature* **444**, 283–284 (2006).

3. Kalbitz, K., Solinger, S., Park, J. H., Michalzik, B. & Matzner, E. Controls on the dynamics of dissolved organic matter in soils: a review. *Soil Sci.* **165**, 277–304 (2000).

4. Kaiser, K. & Kalbitz, K. Cycling downwards—dissolved organic matter in soils. *Soil Biol. Biochem.* **52**, 29–32 (2012).

5. Brantley, S. L., Goldhaber, M. B. & Ragnarsdottir, K. V. Crossing disciplines and scales to understand the Critical Zone. *Elements* **3**, 307–314 (2007).

6. Li, L. et al. Expanding the role of reactive transport models in critical zone processes. *Earth Sci. Rev.* **165**, 280–301 (2017).

7. Sanderman, J., Baldock, J. A. & Amundson, R. Dissolved organic carbon chemistry and dynamics in contrasting forest and grassland soils. *Biogeochemistry* **89**, 181–198 (2008).

8. Marschner, B. et al. How relevant is recalcitrance for the stabilization of organic matter in soils? *J. Plant Nutr. Soil Sc.* **171**, 91–110 (2008).

9. von Luetzow, M. et al. Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions—a review. *Eur. J. Soil. Sci.* **57**, 426–445 (2006).

10. Dungait, J. A. J., Hopkins, D. W., Gregory, A. S. & Whitmore, A. P. Soil organic matter turnover is governed by accessibility not recalcitrance. *Glob. Change Biol.* **18**, 1781–1796 (2012).

11. Don, A., Roedenbeck, C. & Gleixner, G. Unexpected control of soil carbon turnover by soil carbon concentration. *Environ. Chem. Lett.* **11**, 407–413 (2013).

12. Leinemann, T. et al. Multiple exchange processes on mineral surfaces control the transport of dissolved organic matter through soil profiles. *Soil Biol. Biochem.* **118**, 79–90 (2018).

13. Marschner, B. & Kalbitz, K. Controls of bioavailability and biodegradability of dissolved organic matter in soils. *Geoderma* **113**, 211–235 (2003).
14. Fontaine, S. et al. Stability of organic carbon in deep soil layers controlled by fresh carbon supply. *Nature* **450**, 277–280 (2007).
15. Schmidt, M. W. I. et al. Persistence of soil organic matter as an ecosystem property. *Nature* **478**, 49–56 (2011).
16. Steinbeiss, S., Temperton, V. M. & Gleixner, G. Mechanisms of short-term soil carbon storage in experimental grasslands. *Soil Biol. Biochem.* **40**, 2634–2642 (2008).
17. Kaiser, K., Guggenberger, G. & Haumaier, L. Changes in dissolved lignin-derived phenols, neutral sugars, uronic acids, and amino sugars with depth in forested Haplic Arenosols and Rendzic Leptosols. *Biogeochemistry* **70**, 135–151 (2004).
18.
Gleixner, G., Poirier, N., Bol, R. & Balesdent, J. Molecular dynamics of organic matter in a cultivated soil. *Org. Geochem.* **33**, 357–366 (2002).
19. Gleixner, G. Soil organic matter dynamics: a biological perspective derived from the use of compound-specific isotopes studies. *Ecol. Res.* **28**, 683–695 (2013).
20. Klotzbücher, T., Kalbitz, K., Cerli, C., Hernes, P. J. & Kaiser, K. Gone or just out of sight? The apparent disappearance of aromatic litter components in soils. *Soil* **2**, 325–335 (2016).
21. Waggoner, D. C., Chen, H., Willoughby, A. S. & Hatcher, P. G. Formation of black carbon-like and alicyclic aliphatic compounds by hydroxyl radical initiated degradation of lignin. *Org. Geochem.* **82**, 69–76 (2015).
22. DiDonato, N., Chen, H., Waggoner, D. & Hatcher, P. G. Potential origin and formation for molecular components of humic acids in soils. *Geochim. Cosmochim. Acta* **178**, 210–222 (2016).
23. Saidy, A. R., Smernik, R. J., Baldock, J. A., Kaiser, K. & Sanderman, J. The sorption of organic carbon onto differing clay minerals in the presence and absence of hydrous iron oxide. *Geoderma* **209–210**, 15–21 (2013).
24. Keiluweit, M. et al. Mineral protection of soil carbon counteracted by root exudates. *Nat. Clim. Change* **5**, 588–595 (2015).
25. Lange, M. et al. Plant diversity increases soil microbial activity and soil carbon storage. *Nat. Commun.* **6**, 6707 (2015).

26. Liang, C., Schimel, J. P. & Jastrow, J. D. The importance of anabolism in microbial control over soil carbon storage. *Nat. Microbiol.* **2**, 17105 (2017).
27. Miltner, A., Bombach, P., Schmidt-Bruecken, B. & Kaestner, M. SOM genesis: microbial biomass as a significant source. *Biogeochemistry* **111**, 41–55 (2012).
28. Koch, B. P. & Dittmar, T. From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Commun. Mass Spectrom.* **20**, 926–932 (2006); erratum **30**, 250 (2016).
29. Hertkorn, N. et al. High-precision frequency measurements: indispensable tools at the core of the molecular-level analysis of complex systems. *Anal. Bioanal. Chem.* **389**, 1311–1327 (2007).
30. Magurran A. E. *Measuring Biological Diversity* (Blackwell, 2004).
31. Seidel, M. et al. Molecular-level changes of dissolved organic matter along the Amazon river-to-ocean continuum. *Mar. Chem.* **177**, 218–231 (2015). **Part 2**.
32. Flerus, R. et al. A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* **9**, 1935–1955 (2012).
33. Bandowe, B. A. M. et al. Plant diversity enhances the natural attenuation of polycyclic aromatic compounds (PAHs and oxygenated PAHs) in grassland soils. *Soil Biol. Biochem.* **129**, 60–70 (2019).
34. Hertkorn, N. et al. Characterization of a major refractory component of marine dissolved organic matter. *Geochim. Cosmochim. Acta* **70**, 2990–3010 (2006).
35. Einsiedl, F. et al. Rapid biotic molecular transformation of fulvic acids in a karst aquifer. *Geochim. Cosmochim. Acta* **71**, 5474–5482 (2007).
36. Fellman, J. B., D'Amore, D. V. & Hood, E. Fluorescence characteristics and biodegradability of dissolved organic matter in forest and wetland soils from coastal temperate watersheds in southeast Alaska. *Biogeochemistry* **88**, 169–184 (2008).
37. Sanderman, J., Maddern, T. & Baldock, J. Similar composition but differential stability of mineral retained organic matter across four classes of clay minerals. *Biogeochemistry* **121**, 409–424 (2014).
38. Rasmussen, C. et al. Beyond clay: towards an improved set of variables for predicting soil organic matter content. *Biogeochemistry* **137**, 297–306 (2018).
39. Jap, B. & Walian, P. Structure and functional mechanism of porins. *Physiol. Rev.* **76**, 1073–1088 (1996).

40. Nikaido, H. Transport across the bacterial outer membrane. *J. Bioenerg. Biomembr.* **25**, 581–589 (1993).

41. Lehmann, J. & Kleber, M. The contentious nature of soil organic matter. *Nature* **528**, 60–68 (2015).

42. Osterholz, H., Niggemann, J., Giebel, H.-A., Simon, M. & Dittmar, T. Inefficient microbial production of refractory dissolved organic matter in the ocean. *Nat. Commun.* **6**, 7422 (2015).

43. Amon, R. M. W. & Benner, R. Bacterial utilization of different size classes of dissolved organic matter. *Limnol. Oceanogr.* **41**, 41–51 (1996).

44. Riedel, T., Zak, D., Biester, H. & Dittmar, T. Iron traps terrestrially derived dissolved organic matter at redox interfaces. *Proc. Natl Acad. Sci. USA* **110**, 10101–10105 (2013).

45. Benk, S. A., Li, Y., Roth, V.-N. & Gleixner, G. Lignin dimers as potential markers for ^{14}C-young terrestrial dissolved organic matter in the Critical Zone. *Front. Earth Sci.* **6**, 168 (2018).

46. Neff, J. C. & Asner, G. P. Dissolved organic carbon in terrestrial ecosystems: synthesis and a model. *Ecosystems* **4**, 29–48 (2001).

47. Roscher, C. et al. The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. *Basic Appl. Ecol.* **5**, 107–121 (2004).

48. Scheffer F. & Schachtschabel P. *Lehrbuch der Bodenkunde* (Spektrum Akademischer Verlag, 2002).

49. Bauriegel, A., Kühn, D., Schmidt, R., Hering, J. & Hannemann, J. *Bodenübersichtskarte des Landes Brandenburg im Maßstab 1:300 000* (Kleinmachnow, Landesamt für Geowissenschaften und Rohstoffe, 2001).

50. Dittmar, T., Koch, B., Hertkorn, N. & Kattner, G. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol. Oceanogr. Methods* **6**, 230–235 (2008).

51. Steinbeiss, S. et al. Plant diversity positively affects short-term soil carbon storage in experimental grasslands. *Glob. Change Biol.* **14**, 2937–2949 (2008).

52. Ravenek, J. M. et al. Long-term study of root biomass in a biodiversity experiment reveals shifts in diversity effects over time. *Oikos* **123**, 1528–1536 (2014).

53. Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).

54. Kramer, C. & Gleixner, G. Variable use of plant- and soil-derived carbon by microorganisms in agricultural soils. *Soil Biol. Biochem.* **38**, 3267–3278 (2006).
55. Mellado-Vázquez, P. G., Lange, M. & Gleixner, G. Soil microbial communities and their carbon assimilation are affected by soil properties and season but not by plants differing in their photosynthetic pathways (C3 vs. C4). *Biogeochemistry* **142**, 175–187 (2019).
56. Frostegard, A. & Baath, E. The use of phospholipid fatty acid analysis to estimate bacterial and fungal biomass in soil. *Biol. Fertil. Soils* **22**, 59–65 (1996).
57. Zelles, L. Identification of single cultured micro-organisms based on their whole-community fatty acid profiles, using an extended extraction procedure. *Chemosphere* **39**, 665–682 (1999).
58. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
59. Muyzer, G., Dewaal, E. C. & Uitterlinden, A. G. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**, 695–700 (1993).
60. Yu, Y., Lee, C., Kim, J. & Hwang, S. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol. Bioeng.* **89**, 670–679 (2005).
61. Ihrmark, K. et al. New primers to amplify the fungal ITS2 region—evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiol. Ecol.* **82**, 666–677 (2012).
62. Gweon, H. S. et al. PIPITS: an automated pipeline for analyses of fungal internal transcribed spacer sequences from the Illumina sequencing platform. *Methods Ecol. Evol.* **6**, 973–980 (2015).
63. Oksanen J. et al. vegan: Community ecology package. R package version 2.5-3 https://cran.r-project.org/web/packages/vegan/index.html (2015).
64. Malik, A. A. et al. Linking molecular size, composition and carbon turnover of extractable soil microbial compounds. *Soil Biol. Biochem.* **100**, 66–73 (2016).

65. Pohlabeln, A. M. & Dittmar, T. Novel insights into the molecular structure of non-volatile marine dissolved organic sulfur. *Mar. Chem.* **168**, 86–94 (2015).
66. Koch, B. P., Dittmar, T., Witt, M. & Kattner, G. Fundamentals of molecular formula assignment to ultrahigh resolution mass data of natural organic matter. *Anal. Chem.* **79**, 1758–1763 (2007).
67. Stenson, A. C., Marshall, A. G. & Cooper, W. T. Exact masses and chemical formulas of individual Suwannee River fulvic acids from ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectra. *Anal. Chem.* **75**, 1275–1284 (2003).
68. R Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2016).
69. Bray, J. R. & Curtis, J. T. An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27**, 326–349 (1957).
70. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Development Core Team. nlme: Linear and nonlinear mixed effects models. R package version 3.1-137 https://cran.r-project.org/web/packages/nlme (2016).
71. Legendre, P. & Legendre, L. *Numerical Ecology* Vol. 20 (Elsevier, 1998).
72. Micallef, L. & Rodgers, P. eulerAPE: Drawing area-proportional 3-Venn diagrams using ellipses. *PLOS ONE* **9**, e101717 (2014).
73. Hunt, J. F. & Ohno, T. Characterization of fresh and decomposed dissolved organic matter using excitation-emission matrix fluorescence spectroscopy and multiway analysis. *J. Agric. Food Chem.* **55**, 2121–2128 (2007).
74. Merritt, K. A. & Erich, M. S. Influence of organic matter decomposition on soluble carbon and its copper-binding capacity. *J. Environ. Qual.* **32**, 2122–2131 (2003).
75. Simon, C., Roth, V.-N., Dittmar, T. & Gleixner, G. Molecular signals of heterogeneous terrestrial environments identified in dissolved organic matter: a comparative analysis of orbitrap and ion cyclotron resonance mass spectrometers. *Front. Earth Sci.* **6**, 138 (2018).
76. Chambers, M. C. et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918–920 (2012).
77. Strohalm, M., Kavan, D., Novak, P., Volny, M. & Havlicek, V. mMass 3: A cross-platform software environment for precise analysis of mass spectrometric data. *Anal. Chem.* **82**, 4648–4651 (2010).