Geoderma 432 (2023) 116404

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Molecular diversity and the fate of biochemical fractions of eucalypt tissues in soil

Luís F.J. Almeida [a], Ivan F. Souza [a,*], Luís C.C. Hurtarte [b], Pedro P.C. Teixeira [c], Thiago M. Inagaki [c,e], Ivo R. Silva [a], Carsten W. Mueller [c,d]

[a] Departamento de Solos, Avenida P.H. Rolfs, s/n Campus Universitário, Universidade Federal de Viçosa, Viçosa, MG CEP 36570-900, Brazil
[b] European Synchrotron Radiation Facility, Beamline ID21, Grenoble 38100, France
[c] Chair of Soil Science, School of Life Sciences, Technical University of Munich, Emil-Ramann-Strasse 2, Freising 85354, Germany
[d] Department of Geosciences and Natural Resource Management, University of Copenhagen, Øster Voldgade 10, DK-1350 Copenhagen K, Denmark
[e] Norwegian Institute of Bioeconomy Research (NIBIO), Department of Biogeochemistry and Soil Quality, Høgskoleveien 7, 1430 Ås, Norway

ARTICLE INFO

Handling Editor: Mariluz Cayuela

Keywords:
Proximate analysis
^{13}C
Substrate biochemistry
^{13}C-CP/MAS-NMR spectroscopy
Molecular diversity
Microbial respiration

ABSTRACT

The molecular diversity of the source substrate has been regarded as a significant controller of the proportion of plant material that is either mineralized or incorporated into soil organic matter (SOM). However, quantitative parameters to express substrate molecular diversity remain elusive. In this research, we fractionated leaves, twigs, bark, and root tissues of ^{13}C-enriched eucalypt seedlings into hot water extractables (HWE), total solvent (acetone) extractables (TSE), a cellulosic fraction (CF), and the acid unhydrolyzable residue (AUR). We used ^{13}C NMR spectroscopy to obtain a molecular diversity index (MDI) based on the relative abundance of carbohydrate, protein, lignin, lipid, and carbonyl functional groups within the biochemical fractions. Subsequently, we obtained artificial plant organs containing fixed proportions (25%) of their respective biochemical fractions to be incubated with soil material obtained from a Haplic Ferralsol for 200-days, under controlled temperature (25 ± 1 °C) and moisture adjusted to 70–80% of the soil water holding capacity. Our experimental design was a randomized complete block design, arranged according to a factorial scheme including 4 plant organs, 4 biochemical fractions, and 3 blocks as replicates. During the incubation, we assessed the evolution of CO_2 from the microcosms after 1, 2, 3, 4, 7, 10, 13, 21, 28, 38, 45, 70, 80, 92, 112, 148, 178 and 200 days from the start of the incubation. After the incubation, soil subsamples were submitted to a density fractionation to separate the light fraction of SOM (LFOM) i.e., with density <1.8 g cm^{-3}. The heavy fraction remaining was submitted to wet-sieving yielding the sand-sized SOM (SSOM) and the mineral-associated SOM (MAOM), with particle-size greater and smaller than 53 μm, respectively. We found that HWE and AUR exhibited comparatively higher MDIs than the TSE and CF. During the incubation, HWE and CF were the primary sources of ^{13}C-CO_2 from all plant organs and after 92 days, the respiration of the TSE of bark and roots increased. Otherwise, the AUR contributed the least for the release of ^{13}C-CO_2. There were no significant relationships between the MDI and the amount of ^{13}C transferred into the LFOM or SSOM. Otherwise, the transfer of ^{13}C into the MAOM increased as a linear-quadratic function of MDI, which in turn was negatively correlated with the total ^{13}C-CO_2 loss. Overall, the MDI exerted a stronger control on the ^{13}C-labeled MAOM than on ^{13}C-CO_2 emissions, highlighting the need to improve our ability to distinguish and quantify direct plant inputs from those of microbial origin entering soil C pools.

1. Introduction

Over the last decades, the long-standing paradigm that the accumulation of organic substrates that are rather resistant to decomposition drives the differentiation of soil organic matter (SOM) pools has been increasingly challenged (Grandy and Neff, 2008; Hall et al., 2020; Schmidt et al., 2011). Otherwise, it has been increasingly recognized that products of microbial anabolism have a significant role as building blocks controlling SOM formation (Cotrufo et al., 2013; Kopittke et al., 2017; Liang et al., 2017). Moreover, it has been proposed that for substrates showing high molecular diversity and multiple distinctive chemical bonds, their microbial respiration should require complex

* Corresponding author.
E-mail addresses: ivanfrsouza@gmail.com, ivan.souza@ufv.br (I.F. Souza).

https://doi.org/10.1016/j.geoderma.2023.116404
Received 3 September 2022; Received in revised form 19 February 2023; Accepted 20 February 2023
Available online 6 March 2023

enzymatic systems (Lehmann et al., 2020). In this context, it remains unclear whether high substrate molecular diversity would imply direct transfer of C from plant material into SOM (i.e., without or with little microbial intermediation) or simply high rates of CO_2 release as microbes spend more energy to produce extracellular enzymes. Furthermore, owing to the relatively complex chemistry of plant material, quantitative parameters to express substrate molecular diversity remain elusive. Consequently, the connection between substrate molecular diversity and its fate in soils remains poorly constrained, i.e., microbial respiration or entry into SOM pools.

Proximate analysis has long been used to isolate and quantify specific biochemical fractions of plant materials (Osono et al., 2008; Preston and Trofymow, 2000; Ryan et al., 1990; Sjöberg et al., 2004). Usually, the proximate analysis includes stepwise treatments of plant materials to isolate soluble materials either in water (e.g., polar substances) or organic solvents (e.g., non-polar substances), as well as structural components (e.g., cellulose, hemicellulose, and lignin) (Preston and Trofymow, 2015). As a result, the proximate analysis may help improve the inference of the fate of whole plant tissues in soil by characterizing their specific biochemical fractions (i.e., non-structural *versus* structural) (Cotrufo et al., 2015). Although the interest in evaluating the fate of specific biochemical fractions of plant material in soils has been increasing (Almeida et al., 2018; Cotrufo et al., 2022; Haddix et al., 2016), these studies remain comparatively less frequent in the literature. Furthermore, owing to the inherently empirical nature of the proximate analysis, the fractions obtained from such procedures require detailed characterization (Kögel-Knabner and Rumpel, 2018; Preston et al., 2009; Zech et al., 1987). Ideally, the characterization scheme for the specific biochemical fractions of plant material should go beyond the partition of their constituents into either non-structural or structural. Therefore, a stepwise fractionation of plant material and a detailed characterization of its biochemical fractions should improve our ability to determine their fate in soil.

An assessment of the major functional groups within specific biochemical fractions obtained from plant material may be achieved by using solid-state ^{13}C nuclear magnetic resonance (NMR) spectroscopy coupled to cross-polarization magic angle spinning (CP/MAS) (Kögel-Knabner and Rumpel, 2018; Preston and Trofymow, 2015). Typically, CP/MAS-NMR elucidates the relative abundance of C as alkyl C, N-alkyl/methoxy C, O-alkyl C, aromatic C, and carbonylic C groups (Nelson and Baldock, 2005). Additionally, Nelson and Baldock (2005) have proposed a molecular mixing model (MMM) according to which NMR data can be used to estimate the nominal oxidation state of C (NOSC) as well as the contents of carbohydrate, protein, lipid, lignin, and carbonyl within a given organic material. Therefore, the data obtained by CP/MAS-NMR spectroscopy and the MMM could help determine to what extent the fate of C in soils, i.e., CO_2 emissions or SOM buildup, can be impacted by the molecular diversity of biochemical fractions of plant material. Moreover, this is particularly relevant owing to the increasing interest to attribute the role of direct contribution of plant-derived C to SOM pools (Angst et al., 2021; Chen et al., 2022; Huang et al., 2020).

In this study, we obtained hot water extractables (HWE), total solvent extractables (TSE), a cellulosic fraction (CF), and the acid unhydrolyzable residue (AUR) from eucalypt plants submitted to ^{13}C labeling. These fractions were thoroughly characterized using the NMR-MMM approach and the data obtained was used to calculate their molecular diversity index. Subsequently, we obtained artificial plant organs (leaves, twigs, bark, and roots) formed by 25% of each biochemical fraction to be incubated into soil material derived from a Haplic Ferralsol. Our incubation lasted 200 days, during which we monitored the evolution of ^{13}C-CO_2 from the microcosms. After the incubation, SOM fractions were separated according to a density fractionation followed by a particle-size fractionation. Finally, we used correlation and regression analysis to determine the relationship between the molecular diversity index of the biochemical fractions and their fate in soil, i.e., mineralization to CO_2 and/or entry into SOM pools.

2. Methods

2.1. Plant growth and isotopic labeling

For this study, we grew clonal eucalypt seedlings (hybrids of *Eucalyptus grandis* × *E. urophylla*) in a growth chamber (448.0 dm^3) under constant aeration. The seedlings were placed into polyethylene vessels (3.5 L capacity), and all nutrients were supplied via Clark nutrient solutions (Clark, 1975). Seedlings were divided in two groups; one group was submitted to a pulse-labeling with ^{13}C-CO_2 (i.e., labeled group) and the other with ^{12}C-CO_2 (i.e., unlabeled group). This procedure was done to avoid potential differences caused by the enhanced CO_2 concentration on the plant growth between the unlabeled and labeled groups. Briefly, the labeled group of plants was taken into a labeling chamber (16 weeks in total) and submitted to a ^{13}C-CO_2 enriched atmosphere obtained by dissolving $Na_2^{13}CO_3$ (99 atom%) into an H_2SO_4 solution (Machado et al., 2011). Likewise, the unlabeled controls were submitted to the pulse-labeling with a non-isotopically enriched Na_2CO_3 (instead of the $Na_2^{13}CO_3$).

After 16 weeks, the seedlings were harvested and separated into leaves, twigs, bark, and roots. All plant organs were dried to constant weight under a forced-draft oven at 60 ℃, ground into a Wiley mill (Thomas Model T4275.B05) and stored for chemical fractionation and subsequent analysis. Total C, N, and ^{13}C content of the plant organs are given in Table 1.

2.2. Isolation of the biochemical fractions from plant organs

Subsamples of leaves, twigs, bark, and roots from the isotopically labeled and unlabeled plant material were chemically fractioned into four operationally defined "biochemical fractions", namely: hot water extractables (HWE), total solvent (acetone) extractables (TSE), a cellulosic fraction (CF) and the acid unhydrolyzable residue (AUR). Both TSE and HWE were isolated according to the procedures described by Ryan et al. (1990), except that the TSE was isolated using acetone after the removal of the HWE. Briefly, 2.0 g of each milled plant organ (particle-size $\leq$ 0.425 mm) were extracted in Soxhlets with 150.0 mL of deionized water at 100 ℃ for 6 h to obtain the HWE. Likewise, 2.0 g of the residue obtained after the HWE extraction were extracted in the Soxhlets with 150.0 mL of acetone at 56 ℃ for 6 h to obtain the TSE. After the extraction, HWE and TSE were freeze-dried and stored. Subsequently, the residue-free of extractives (HWE and TSE) was fractioned into either CF or AUR in separate procedures. Accordingly, the CF was obtained by treating 1.0 g of the residue with a mixture of ethanol and aqueous HNO_3 (65%, v/v) at a ratio of 4:1 at 85 ℃ for 1 h. This procedure is referred to as the Kürschner-Hoffer method (Pettersen, 1984; Reid and Lynch, 1937). This extraction was repeated four times, and the residue generated was treated with an aqueous solution of KOH 25% (m/v) at 100 ℃ for 1 h. We isolated the AUR fraction by treating 1.0 g of the residue-free extractives with 6 mL of aqueous H_2SO_4 (72%, v/v) at 30 ℃ in a water bath for 1 h. Subsequently, we added 140 mL of distilled water, and the temperature was raised to 100 ℃ and kept constant for 4 h (Pettersen, 1984). Both CF and AUR were rinsed with ultra-pure water,

Table 1
Mean and standard error for total carbon, nitrogen, C:N ratio, and isotopic data expressed in $\delta^{13}C_{\text{-PDB}}$ and ^{13}C-atom% of each plant organ (leaves, twigs, bark, and roots) of eucalypt seedlings.

Plant Organ	C, g kg^{-1}	N, g kg^{-1}	C:N	$\delta^{13}C_{\text{-PDB}}$	^{13}C-atom %
Leaves	440.9 ± 3.9	18.3 ± 1.0	24	425.1 ± 10.9	1.6 ± 0.0
Twigs	429.0 ± 10.4	7.6 ± 0.5	56	359.6 ± 31.9	1.5 ± 0.0
Bark	416.4 ± 1.7	5.2 ± 0.1	80	319.1 ± 16.7	1.4 ± 0.0
Roots	435.6 ± 16.2	10.4 ± 0.5	42	377.9 ± 35.1	1.5 ± 0.0

freeze-dried, and stored. Although obtaining only four biochemical fractions may not be sufficient to represent the actual chemistry of plant material, we stress that our approach was in line with other studies available in the literature (Fulton-Smith and Cotrufo, 2019; Preston and Trofymow, 2015; Soong et al., 2015).

2.3. Characterization of the biochemical fractions

Subsamples of the biochemical fractions (HWE, TSE, CF, and AUR) were analyzed in an isotope ratio mass spectrometer – IRMS (ANCA-GSL, 20-20, Sercon, Crewe-UK) to determine their total contents of C, N, C:N ratios, and ^{13}C enrichment (Table 2). Isotopic data were referenced to the V-PDB international standard and expressed in ^{13}C-atom% notation. Subsequently, we used solid-state ^{13}C cross-polarization magic angle spinning (CP/MAS) NMR spectroscopy (Bruker DSX 200 spectrometer, Bruker BioSpin GmbH) to determine the major functional groups within the biochemical fractions. Accordingly, the samples were filled into zirconium dioxide rotors, and the measurements were conducted at a rotation speed of 6.8 kHz, with a 'magic angle' around 54.74°, with the acquisition time of 0.01024 s and a delay time of 1.0 s. All spectra were integrated along seven chemical shift regions: 0–45 ppm (Alkyl C), 45–60 ppm (N-Alkyl/Methoxyl C), 60–95 ppm (O-Alkyl C), 95–110 ppm (Di-O-Alkyl C), 110–145 ppm (Aromatic C), 145–165 ppm (Phenolic C) and 165–215 ppm (Amide/Carboxylic C). In combination with the total C and N contents of each sample analyzed (Table 2), our NMR data were used to estimate the relative abundance of carbohydrate, protein, lignin, lipid, and carbonyl using the molecular mixing model (Nelson and Baldock, 2005). We further refer to this method in this text as NMR-MMM.

2.4. Quantitative parametrization of molecular diversity

We used the relative abundance of the five chemical groups (carbohydrates, protein, lignin, lipids, and carbonyl) identified by the NMR-MMM approach (Nelson and Baldock, 2005) to obtain a molecular diversity index (MDI) for each biochemical fraction as follows

$$\mathrm{MDI} = \frac{\mathrm{MCG}}{\mathrm{PCG}}$$

in which MCG refers to the sum of the relative abundances of the four minor chemical groups, and PCG refers to the relative abundance of the predominant chemical group. For example, for a given biochemical fraction in which carbohydrates were predominant, the MDI was calculated by dividing the sum of the relative proportion of the remaining four minor chemical groups (i.e., protein, lignin, lipids, and carbonyl) by the relative proportion of carbohydrates. As a result, the calculated MDI would have been inversely proportional to the relative abundance of the main chemical group, i.e., carbohydrates. Thus, the overall molecular diversity would increase as the proportion of the four minor constituents increased relative to that of carbohydrates.

2.5. Incubation experiment

Our incubation experiment was set according to a randomized complete block design with three replicates and arranged according to a 4 × 4 factorial scheme, i.e., four artificial plant organs (leaves, twigs, bark, and roots) and four biochemical fractions (HWE, TSE, CF, and AUR). These artificial plant organs included fixed proportions (i.e., 25%) of each biochemical fraction (HWE, TSE, CF, and AUR) corrected for their C content. We opted for using these fixed proportions because our primary objective was to evaluate the mineralization and the incorporation of each biochemical fraction into SOM fractions. We recognize that following their fractionation, each biochemical fraction would not necessarily be mineralized in the soil at the same rate as if they were within the original tissue. This is because in many plant tissues, materials such as cellulose may be part of structures physically associated with lignin or other resistant components (Kögel-Knabner, 2002). Thus, our experimental set up would restrict comparisons between the behavior of natural and artificial plant organs. For the mixture of the biochemical fractions, only one of the four was isotopically labeled with ^{13}C for a given artificial plant organ. For instance, for the artificial leaf material, the ^{13}C-enriched leaf-HWE was mixed with the other three fractions (TSE, CF, and AUR) obtained from leaf tissues of unlabeled plants. The same strategy was used for the other biochemical fractions of leaves and the other artificial organs. This approach allowed us to determine to what extent the respiration of a specific biochemical fraction would be impacted by its inherent molecular properties in the presence of other biochemical fractions. All artificial plant organs were added to soil samples at 10.0 mg C g^{-1} of soil.

We incubated the artificial plant organs mixed into soil subsamples of the Ap horizon (0–20 cm depth) of a sandy-loam Haplic Ferralsol (IUSS Working Group WRB, 2015). The soil material was taken from an area within the Atlantic Forest biome at Paula Cândido, Minas Gerais, Brazil. At this location, the vegetation includes pastures of the genus *Brachiaria* also known as *Urochloa*, eucalypt plantations, as well as remnants of the original forest vegetation. For this study, soil material was collected at 4 points across the pastureland distanced about 1.0 km from one another. At each selected point, four samples were taken from

Table 2
Mean and standard error for total carbon, nitrogen, C:N ratio and isotopic data expressed in $\delta^{13}C_{\text{-PDB}}$ and ^{13}C-atom% of each biochemical fraction (BF) extracted from leaves, twigs, bark, and root tissues of eucalypt plants.

Plant organ	BF	C, g kg^{-1}	N, g kg^{-1}	C:N	$\delta^{13}C_{\text{-PDB}}$	^{13}C-atom%
Leaves	HWE	421.2 ± 10.3	4.2 ± 0.2	100	324.4 ± 22.6	1.5 ± 0.0
	TSE	694.4 ± 2.9	4.6 ± 0.6	152	504.4 ± 43.0	1.7 ± 0.0
	CF	427.7 ± 11.1	2.5 ± 1.0	173	501.9 ± 28.6	1.7 ± 0.0
	AUR	588.2 ± 4.0	41.6 ± 2.4	14	441.5 ± 36.3	1.6 ± 0.0
Twigs	HWE	397.8 ± 6.4	4.7 ± 0.3	85	319.0 ± 12.0	1.5 ± 0.0
	TSE	731.9 ± 3.6	3.1 ± 0.2	239	379.7 ± 6.2	1.5 ± 0.0
	CF	420.3 ± 3.3	0.3 ± 0.1	1401	361.6 ± 21.3	1.5 ± 0.0
	AUR	570.8 ± 5.5	22.4 ± 1.1	26	382.4 ± 7.8	1.5 ± 0.0
Bark	HWE	407.9 ± 3.7	2.6 ± 0.1	157	276.1 ± 1.7	1.4 ± 0.0
	TSE	737.5 ± 10.7	3.4 ± 0.3	217	334.1 ± 12.9	1.5 ± 0.0
	CF	408.6 ± 1.8	0.5 ± 0.1	817	368.0 ± 18.7	1.5 ± 0.0
	AUR	560.9 ± 6.8	18.3 ± 1.5	31	312.9 ± 10.9	1.5 ± 0.0
Roots	HWE	400.9 ± 2.6	5.6 ± 0.6	72	302.2 ± 26.9	1.4 ± 0.0
	TSE	738.9 ± 2.4	2.0 ± 0.1	369	373.9 ± 25.8	1.5 ± 0.0
	CF	406.3 ± 3.3	1.2 ± 0.0	339	326.9 ± 33.4	1.5 ± 0.0
	AUR	552.1 ± 3.0	16.0 ± 2.0	34	342.1 ± 32.2	1.5 ± 0.0

HWE: hot water extractables; TSE: total solvent extractables; CF: cellulosic fraction; AUR: acid unhydrolyzable residue.

points oriented along the 4 cardinal axes (i.e., N, S, E, W) distanced about 100.0 m from the center. After collecting these samples at these four points, the soil material obtained was mixed to form a composite sample to make a sure a homogeneous soil material was used for the incubation. This soil had a pH (in H_2O) of 5.7 and 4.6 (in KCl); P content was 3.60 mg dm^{-3} soil; exchangeable K, Ca, and Mg contents were about 0.22, 1.8, and 1.2 ($cmol_c$ dm^{-3}), respectively; Total sand, clay, and silt contents were 520, 310 and 170 g kg^{-1} soil, respectively. Kaolinite was the predominant clay mineral (220 g kg^{-1} soil); the original soil organic carbon content was 22.2 g kg^{-1}, and $\delta^{13}C$ was − 16.0 ‰. Water holding capacity (WHC) was 0.2 g g^{-1} soil, and during the incubation of the samples, their water content was kept at 70.0–80.0% of WHC. These samples remained incubated in a room under controlled temperature 25 ± 1 °C.

For the incubation, we placed 20.0 g of the soil material (particle-size < 2.0 mm) in 0.57 L air-tight glass jars with lids containing a rubber septum, which allowed sampling of the headspace to monitor the evolution of C-CO_2. Throughout the incubation experiment, the first two assessments of soil respiration were made after 6 and 12 h from the beginning of the incubation. Additional assessments were made after 1, 2, 3, 4, 7, 10, 13, 21, 28, 38, 45, 70, 80, 92, 112, 148, 178 and 200 days when the incubation was concluded. For each sampling, we collected an aliquot of 100-mL air from the jars' headspace and analyzed the gas sample to determine $^{13}C/^{12}C$-CO_2 concentrations in a cavity ring-down spectrometer (CRDS, G2131-i, Picarro, Sunnyvale, CA). After each sampling, we vented the jars with ambient air, checked soil moisture content, and kept the experimental units closed until the subsequent assessment. This soil respiration data set were grouped into the following time intervals: 0–2; 2–10; 10–38; 38–92; 92–200 days, with each interval including four individual samplings.

2.6. Density and particle-size fractionation of soil organic matter

After the incubation, soil subsamples were taken from each replicate and submitted to physical fractionation in two steps. In the first step, we used a density fractionation to isolate the light fraction organic matter (LFOM), which was done by adding 5.0 g of air-dried soil subsamples into a beaker containing 25.0 mL of a sodium iodide (NaI) solution (1.80 g cm^{-3}). These subsamples were fully dispersed under a sonication treatment (Heinemann Branson Sonifier 250), with the application of ultrasound at a rate equivalent to 240 J mL^{-1}. After the dispersion, we centrifuged the soil material and collected the floating LFOM using a 2.0 μm pore nylon filter coupled to a vacuum system. The LFOM and the heavy fraction (>1.80 g cm^{-3}) were thoroughly rinsed with deionized water to remove the NaI. In the second step, the dispersed soil heavy fraction was wet-sieved in a 53-μm to yield the sand-sized organic matter (SSOM, particle size > 53 μm) as well as the organic matter within the clay and silt-sized particles, henceforth referred to as mineral associated organic matter (MAOM, particle < 53 μm). After concluding the physical fractionation, all samples were dried under forced air circulation at 45 °C for seven days, weighed, and milled to achieve a particle size < 0.149 μm (100 mesh). All samples were analyzed using an IRMS as described above.

2.7. Isotopic partitioning in CO_2 and soil organic matter fractions

The proportion of ^{13}C-labeled substrates released as CO_2 throughout the incubation or retained within SOM fractions was calculated as follows

$$f = \frac{(\delta tbf - \delta co)}{(\delta bf - \delta co)}$$

In which *f* is the proportion of C derived from a given biochemical fraction; *δtbf* is the $\delta^{13}C$ of the sample (for both CO_2 or SOM fractions) from the treatments receiving isotopically labeled biochemical fractions; *δbf* is the $\delta^{13}C$ of each biochemical fraction used for the incubation, and *δco* is the $\delta^{13}C$ of the control treatments without plant C addition (for both CO_2 or SOM fractions).

Based on the *f* values calculated above, we estimated the recovery rate (RR) of the C transferred into SOM fractions as follows

$$RR = \frac{f \times Cc}{Ci} \times 100$$

Cc is the total C mass within a given SOM fraction at the end of the incubation, and *Ci* is the total C mass added to the samples at the beginning of the experiment. Total C-CO_2 loss was calculated as the difference between the initial C input and total C mass remaining in SOM fractions (LFOM, SSOM, and MAOM).

2.8. Statistics

Initially, we used a two-way analysis of variance (ANOVA) to determine the effect of the interaction between the plant organs and their biochemical fractions on the relative abundance of the chemical groups (i.e., contents of carbohydrates, lignin, protein, lipids, and carbonyl) and the chemical properties (NOSC, N:C, H:C, and O:C). In all tests run, this interaction was significant (data not shown). In the next step, we used one-way ANOVA to compare the relative abundance of the chemical groups and the chemical properties of a given biochemical fraction among the plant organs (Supplementary Tables S1 and S2). We verified the assumptions of normality (Shapiro-Wilk) and homogeneity of variance (Levene's test) and used data transformation whenever necessary to meet these requirements. We also checked the distribution of residuals for each ANOVA. All means values were compared using Tukey's test ($p < 0.05$). We also used correlation analysis to identify the relationship between the results obtained from the NMR-MMM, the respiration of the biochemical fractions for each sampling interval (0–2; 2–10; 10–38; 38–92; 92–200 days), and the proportion of the C added retained within the LFOM, SSOM, and MAOM. The significance level for Pearson's correlation coefficients was set at $p < 0.05$. We also used linear regressions to assess the relationship between substrate MDI and the partitioning of C derived from each biochemical fraction into SOM or CO_2. For the regressions, we used the Shapiro-Wilk to verify normally and Spearman Rank Correlation to verify the homogeneity of variance. All statistical analyses were performed using Statistica STATSOFT 2013 (Palo Alto, CA, USA), and Sigma Plot 11.0 (San Jose, CA, USA).

3. Results

3.1. Functional groups of the biochemical fractions and molecular diversity

Our NMR-MMM data indicates that each biochemical fraction's composition varied according to the plant organ source (Fig. 1). Generally, materials extracted within the HWE fraction were dominated by lignin (48.1–61.3%) and carbohydrates (31.0–42.8%), whereas protein (2.0–5.7%), carbonylic C (2.1–5.6%), and lipids (1.0–3.0%) were less abundant (Fig. 1a). Besides, the most evident difference observed for the HWE among organs was that roots contained more carbohydrates and less lignin (Fig. 1a; $p < 0.05$). Within the TSE, lipids (62.7–90.7%) and lignin (7.9–27.9%) were predominant, with minor proportions of proteins (0.9–2.2%) as shown in Fig. 1b. In this case, the TSE of roots contained more lignin and fewer lipids than the other organs (Fig. 1b; $p < 0.05$).

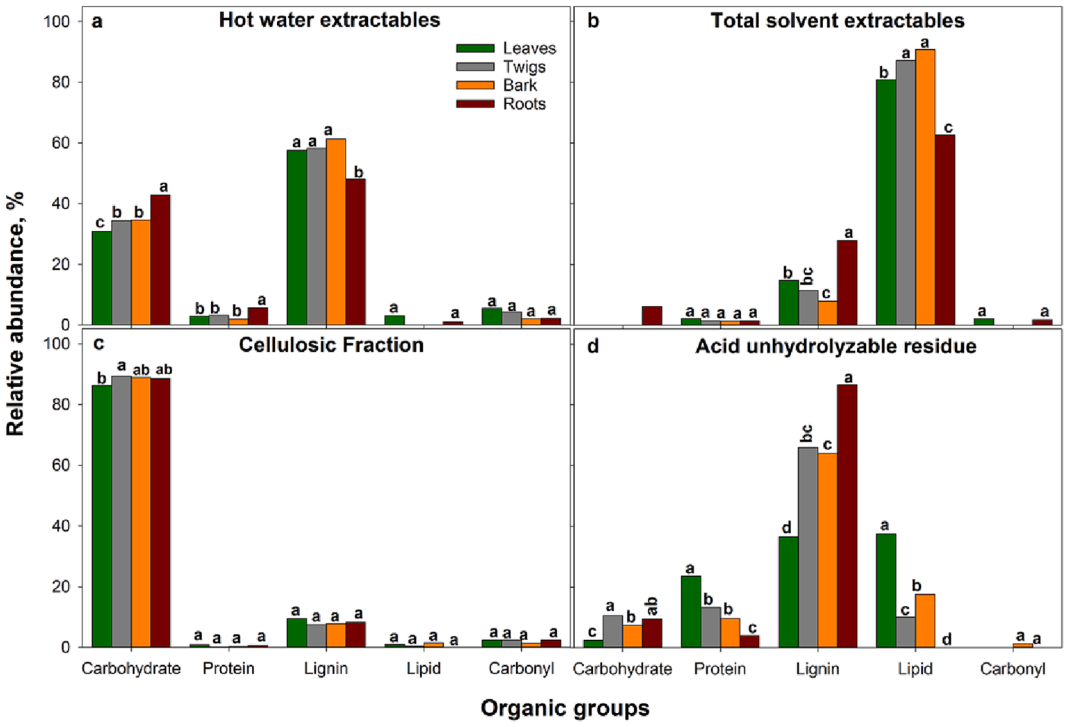

Fig. 1. Relative abundance of the main organic groups (carbohydrate, protein, lignin lipid, and carbonyl) within the biochemical fractions extracted from leaves, twigs, bark, and roots of eucalypt plants as determined by NMR spectroscopy and the Molecular Mixing Model (Nelson and Baldock, 2005); a) hot water extractables; b) total solvent extractables; c) cellulosic fraction; d) acid unhydrolyzable residue. For a given organic group within each biochemical fraction, means followed by the same lowercase letter do not differ at $p < 0.05$ (Tukey's test).

For the structural fractions, CF was predominantly composed of carbohydrates (86.3–89.6%) and lignin (7.4–9.5%), with a small proportion of carbonylic C varying between 1.4 and 2.4% (Fig. 1c). Generally, the CF was the most homogeneous fraction among the plant organs (Fig. 1c), whereas the AUR fraction tended to be the most heterogeneous within and among plant organs (Fig. 1d). As expected, the AUR was dominated by lignin in twigs (65.9%), bark (64.0%), and roots (86.6%), whereas in leaves, lignin accounted for about 36.5% (Fig. 1d). In addition, the AUR of leaves also had substantial amounts of proteins (23.6%) and lipids (36.5%). The other organs' AUR also contained proteins and lipids at varying amounts between 3.9–13.3% and 10.1–17.6%, respectively. Only for roots, no lipids were detectable in their AUR using the NMR-MMM approach (Fig. 1b). Carbohydrates accounted for a minor proportion of the AUR, varying from 2.4 to 10.6% (Fig. 1d).

We used the results obtained for the major chemical groups shown in Fig. 1 (carbohydrate, protein, lipids, lignin, and carbonyl) to obtain a molecular diversity index (MDI) based on the NMR-MMM (Fig. 2). According to our estimates, both TSE and CF showed comparatively low MDI values (varying from 0.10 up to 0.24), consistent with their predominant lipid and carbohydrate contents, respectively. In contrast, the HWE and AUR showed comparatively high MDI values (varying from 0.52 up to 1.67), except for the AUR of roots, which MDI was 0.15 (Fig. 2). Hence, there was a considerable difference in the MDI values obtained for the AUR of leaves and roots. On the other hand, the HWE fraction of roots showed the highest MDI (1.08), which values were comparable between leaves, twigs, and bark, varying between 0.63 and 0.74 (Fig. 2).

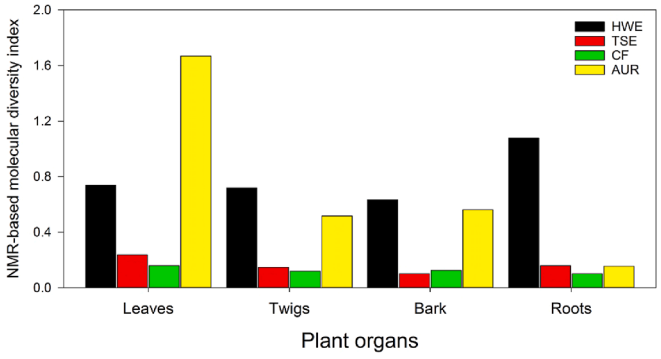

Fig. 2. NMR-based molecular diversity index (MDI^{Y}) of the hot water extractables (HWE), total solvent extractables (TSE), cellulosic fraction (CF), and acid unhydrolyzable residue (AUR) obtained from leaves, twigs, bark, and roots of eucalypt plants;

3.2. *Properties of the biochemical fractions*

Based on the NMR-MMM approach, we also obtained information for NOSC, N:C, H:C, and O:C for all biochemical fractions (Fig. 3). In terms of NOSC, the CF showed the highest averaged values (–0.02), followed by HWE (–0.20) and AUR (–0.49), whereas the TSE showed the lowest averaged values (–1.27), as shown in Fig. 3a. Additionally, we observed significant differences in the NOSC values of TSE and AUR depending on the plant organ source. As such, TSE and AUR showed higher NOSC values in roots compared to other plant organs, whereas the TSE of twigs and bark and the AUR of leaves had the lowest NOSC values overall (Fig. 3a; $p < 0.05$). Furthermore, the constituents of the AUR had the highest N:C ratio, followed by HWE, TSE, and CF (Fig. 3b). For the HWE and AUR specifically, their N:C ratio was significantly affected by plant organs. Accordingly, the HWE of roots showed a higher N:C ratio than the other plant organs, whereas the highest N:C ratio in the AUR was observed for leaves and the lowest values occurred in roots (Fig. 3b; $p < 0.05$). Further, the H:C ratio was highest in TSE (1.83), intermediate in CF (1.61), and similar for HWE (1.36) and AUR (1.37) (Fig. 3b). Finally, the CF showed the highest O:C ratio (>0.75), and the lowest values occurred for the constituents of the TSE (about 0.25). For compounds

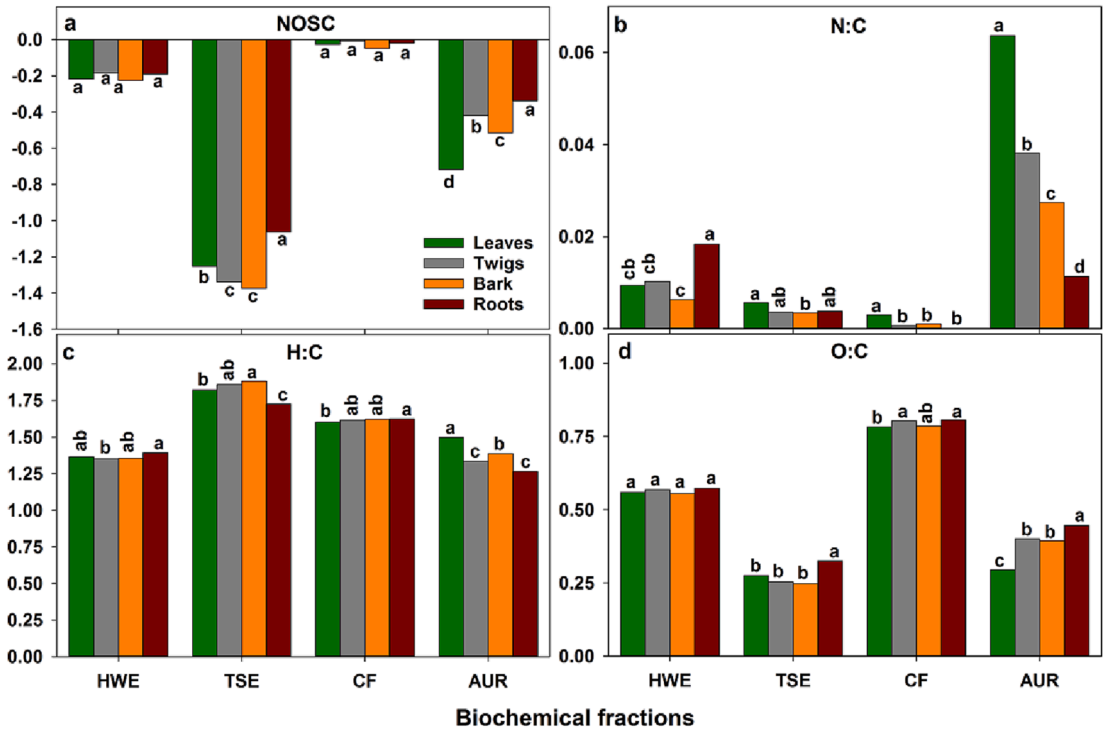

Fig. 3. Chemical characteristics of the biochemical fractions extracted from leaves, twigs, bark, and roots of eucalypt plants as determined by NMR spectroscopy and the Molecular Mixing Model (Nelson and Baldock, 2005); a) Nominal oxidation state of carbon (NOSC); b) nitrogen to carbon ratio (N:C); c) hydrogen to carbon ratio (H:C); and d) oxygen to carbon ratio (O:C). HWE — hot water extractables, TSE — total solvent extractables, CF — cellulosic fraction, and AUR — acid unhydrolyzable residue. For a given biochemical fraction, means followed by the same lowercase letter do not differ at $p < 0.05$ (Tukey's test).

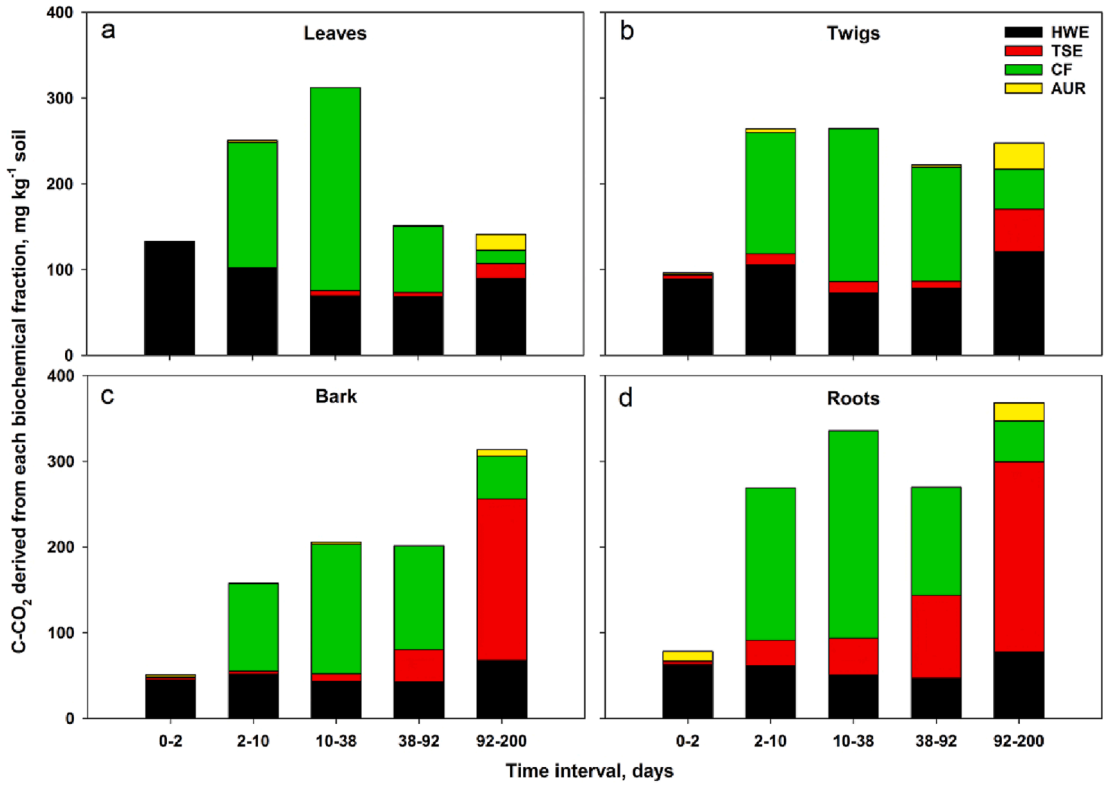

Fig. 4. Total C-CO_2 was derived from each biochemical fraction (mg kg^{-1} soil) for 200 days of the incubation experiment. Each plant organ (leaves, twigs, bark, and roots) refers to the combination of 25% of each biochemical fraction, namely, hot water extractables (HWE), total solvent extractables (TSE), cellulosic fraction (CF), and acid unhydrolyzable residue (AUR) obtained from leaves, twigs, bark, and roots of eucalypt plants. For each time interval, soil gas was sampled in 4 distinct samplings, with 3 replicates for each treatment. Each bar shows the sum of the averaged CO_2 concentration for these samples.

within the HWE, their O:C ratios varied between 0.50 and 0.55 for all plant organs, whereas the O:C ratios of the AUR were about 0.28 for leaves and up to 0.45 for roots (Fig. 3b).

3.3. Microbial respiration on each biochemical fraction

During the first two days of the incubation experiment, soil microorganisms mineralized mostly the HWE fraction for all plant organs (Fig. 4). For the subsequent time interval (2–92 days), the contribution of HWE to total soil respiration either remained unchanged or was reduced. In contrast, the contribution of C-CO_2 derived from the CF increased and became predominant (Fig. 4). After 92 days of incubation, the HWE and CF fractions remained the primary C source to soil decomposers, except for bark and roots, for which the TSE became the predominant C-CO_2 source over the last 108 days of incubation (Fig. 4b-c). The substrates added via AUR underwent much lower microbial respiration than other biochemical fractions (Fig. 4).

3.4. Properties of the biochemical fractions and their respiration rates

For the first four samplings (0.25–2 days), the release of C-CO_2 from the biochemical fractions was only positively correlated with carbonylic C groups (Table 3; $r = 0.7$, $p < 0.001$). For the subsequent time interval (2–10 days), the positive correlation between carbonylic C and the respiration of the biochemical fractions remained significant ($r = 0.64$; $p < 0.01$) but not afterwards (Table 3). Bulk properties (C and N contents) showed significant ($p < 0.05$) and negative correlations with the release of C-CO_2 from the biochemical fractions for the interval of the incubation between 2 and 92 days. Simultaneously, the C:N ratio was positively correlated with the respiration of the biochemical fractions (Table 3; $p < 0.05$). The release of C-CO_2 from the biochemical fractions also showed strong and positive correlations with variables obtained from the NMR-MMM approach. Among these variables, carbohydrates showed the highest correlations coefficients with the release of CO_2 (0.91, 0.94, and 0.80) for the time intervals between 2 and 10, 10–38, and 38–92 days, respectively (Table 3). Other variables showed less consistent correlations with the release of C-CO_2 from the isotopically labeled substrates. For instance, NOSC showed strong and positive correlation with C-CO_2 emissions only for the time intervals between 2–10 ($r = 0.71$, $p < 0.002$) and 10–38 days ($r = 0.62$, $p < 0.010$) of the incubation (Table 3). Otherwise, lignin content was negatively correlated ($r = –0.51$ and –0.52, $p < 0.05$) with the respiration of the biochemical fractions for the time interval between 10–38 and 38–92 days (Table 3; $p < 0.05$). Generally, none of the variables included in our analysis showed significant correlations with C-CO_2 emissions from the biochemical fractions for the time interval between 112 and 200 days (Table 3). Considering the total C-CO_2 loss (i.e., the difference between the initial C input and the amount of C remaining in SOM fractions), we observed significant correlations with all variables included in the analysis, except for H:C ratios (Table 3).

Overall, the respiration rates of the biochemical fractions in the soil may be more representative of their chemical properties relative to their physical structure. This is mostly because the fractionation used minimized the physical arrangement between the biochemical fractions within the natural plant materials.

3.5. Transfer of C from the biochemical fractions into soil organic matter as a function of substrate molecular diversity

For the first step of this analysis, we determined the proportions of the ^{13}C-labeled biochemical fractions that were transferred into all SOM pools considered in our study (Table 4). Generally, the SSOM showed the lowest retention of C derived from all biochemical fractions added for the incubation. On average, this particle-size fraction retained about 0.2% of the initial C input. In contrast, the transfer of ^{13}C-labeled organics into the LFOM and the MAOM was much more relevant. Within the LFOM, there was higher retention of the initial C input added via TSE and AUR compared to the HWE or the CF (Table 4). Among the chemical groups identified by the MMM, lipids showed the strongest correlation with the retention of C within the LFOM (Table 5). Otherwise, there was a higher contribution of C derived from the HWE and AUR for the MAOM fraction. In this case, the retention of C was strongly correlated to protein and lignin contents (Table 5).

In the second step, we used regressions to infer the relationship between substrate MDI and the proportion of the initial C input that was either retained in SOM fractions or lost as CO_2 (Fig. 5). Accordingly, we found no significant relationships between the NMR-based MDI and the retention of C within the LFOM or the SSOM (Fig. 5a-b). In contrast, we observed a linear-quadratic increase in the retention of C derived from the biochemical fractions within the MAOM (Fig. 5c; $p < 0.01$). Otherwise, the proportion of the initial C input lost via CO_2 decreased as a linear function of the MDI (Fig. 5d; $p < 0.05$). Therefore, the MDI

Table 3
Pearson's correlation coefficients (r) between selected properties of the biochemical fractions[‡], their contribution to C–CO_2 emissions for each time interval (0–2, 2–10, 10–38, 38–92, and 92–200 days) of the incubation, and the total C-CO_2 loss[†] expressed as a percentage (%) of the initial C input. Correlation coefficients significant at $p < 0.05$ are indicated in bold typing.

Data source	Time interval	0–2 days	2–10 days	10–38 days	38–92 days	92–200 days	Total C-CO_2 loss[†]
	Variable	**C-CO_2, mg kg^{-1} soil**					**%**
Bulk properties	Carbon, g kg^{-1}	–0.46; p = 0.072	**–0.75; p = 0.001**	**–0.61; p = 0.012**	**–0.51; p = 0.043**	0.31; p = 0.237	**–0.54; p = 0.000**
	Nitrogen, g kg^{-1}	–0.19; p = 0.489	**–0.53; p = 0.035**	–0.50; p = 0.050	**–0.63; p = 0.009**	–0.38; p = 0.143	**–0.72; p = 0.000**
	C:N ratio	–0.26; p = 0.329	0.49; p = 0.053	**0.54; p = 0.033**	**0.72; p = 0.002**	0.05; p = 0.855	**0.58; p = 0.000**
NMR*	NOSC$^{\varphi}$	0.31; p = 0.248	**0.71; p = 0.002**	**0.62; p = 0.010**	0.46; p = 0.071	–0.43; p = 0.101	**0.46; p = 0.001**
	N:C ratio	–0.07; p = 0.791	–0.45; p = 0.078	–0.46; p = 0.072	**–0.57; p = 0.022**	–0.34; p = 0.193	**–0.65; p = 0.000**
	H:C ratio	–0.48; p = 0.057	–0.07; p = 0.811	0.09; p = 0.728	0.19; p = 0.484	0.40; p = 0.130	0.17; p = 0.235
	O:C ratio	0.14; p = 0.594	**0.90; p = 0.000**	**0.88; p = 0.000**	**0.74; p = 0.001**	–0.29; p = 0.282	**0.74; p = 0.000**
	Carbohydrate, %	0.00; p = 0.999	**0.91; p = 0.000**	**0.94; p = 0.000**	**0.80; p = 0.000**	–0.23; p = 0.386	**0.79; p = 0.000**
	Protein, %	–0.10; p = 0.708	–0.47; p = 0.068	–0.47; p = 0.070	**–0.57; p = 0.022**	–0.34; p = 0.199	**–0.65; p = 0.000**
	Lignin, %	0.45; p = 0.078	–0.35; p = 0.179	**–0.51; p = 0.043**	**–0.52; p = 0.039**	–0.23; p = 0.399	**–0.49; p = 0.000**
	Lipid, %	–0.35; p = 0.179	**–0.57; p = 0.021**	–0.47; p = 0.068	–0.31; p = 0.236	0.45; p = 0.080	**–0.31; p = 0.031**
	Carbonyl, %	**0.79; p = 0.000**	**0.64; p = 0.008**	0.41; p = 0.113	0.42; p = 0.105	–0.03; p = 0.916	**0.51; p = 0.000**
	MDI$^{\gamma}$	0.39; p = 0.134	–0.22; p = 0.411	–0.36; p = 0.167	–0.36; p = 0.169	–0.16; p = 0.565	**–0.35; p = 0.014**

[‡] Hot-water extractables (HWE), total solvent extractables (TSE), cellulosic fraction (CF), and the acid unhydrolyzable residue (AUR) extracted from leaves, twigs, bark, and roots of eucalypt plants;

[†] Calculated as the difference between the total initial C input and the amount of C retained with the soil organic matter fractions at the end of the incubation experiment.

[φ] NOSC: Nominal oxidation state of carbon atoms;

[*] Data obtained from nuclear magnetic resonance (NMR) spectroscopy and the Molecular Mixing Model (MMM);

[γ] NMR-based molecular diversity index (MDI), see section 2.4 for further details.

Table 4
Average values (n = 3 replicates) for total C content and the C derived from the biochemical fractions (BF-C) extracted from leaves, twigs, bark, and roots of eucalypt plants within the light fraction organic matter (LFOM), particulate organic matter (POM), and mineral-associated organic matter (MAOM), as well as the respective recovery rate (RR[†]) for the BF-derived C retained within each fraction.

Plant organ	BCF	Total LFOM C	BF-derived C LFOM	RR	Total POM C	BF-derived C POM	RR	Total MAOM C	BF-derived C MAOM	RR
		g kg^{-1} soil		%	g kg^{-1} soil		%	g kg^{-1} soil		%
Leaves	HWE	4.70	0.01	0.4	0.83	0.01	0.4	20.60	0.71	28.4
Leaves	TSE	4.61	1.16	46.3	0.53	0.00	0.1	20.33	0.26	10.3
Leaves	CF	5.24	0.03	1.2	0.61	0.00	0.2	21.77	0.28	11.3
Leaves	AUR	4.14	0.92	36.7	0.57	0.01	0.3	22.64	0.75	30.0
Twigs	HWE	3.27	0.06	2.3	0.48	0.01	0.3	20.56	0.70	28.1
Twigs	TSE	3.13	0.74	29.5	0.48	0.00	0.0	19.13	0.38	15.0
Twigs	CF	3.57	0.05	1.9	0.53	0.01	0.2	21.03	0.32	12.9
Twigs	AUR	2.91	0.76	30.4	0.52	0.00	0.1	22.36	0.85	33.9
Bark	HWE	2.55	0.00	0.0	0.66	0.01	0.2	23.02	0.86	34.6
Bark	TSE	2.93	0.50	20.0	0.86	0.01	0.2	23.86	0.67	26.8
Bark	CF	2.82	0.07	2.8	1.26	0.01	0.2	21.43	0.37	14.8
Bark	AUR	2.93	0.81	32.3	0.63	0.01	0.6	20.15	0.80	32.1
Roots	HWE	4.30	0.00	0.0	0.63	0.01	0.2	24.43	0.94	37.6
Roots	TSE	4.19	0.26	10.5	1.18	0.01	0.2	24.23	0.59	23.8
Roots	CF	3.94	0.02	2.1	0.75	0.01	0.6	23.13	0.55	21.9
Roots	AUR	3.27	0.89	35.8	0.79	0.02	1.0	21.62	0.66	26.4

[†] Values expressed as a percentage of the initial C input.

Table 5
Pearson's correlation coefficients (*r*) between the properties of the biochemical fractions (BF)[‡] as inferred from their bulk properties (variables 1–3), nuclear magnetic resonance (NMR) spectroscopy, and the Molecular Mixing Model (MMM) (variables 4–13), and the proportions (%) of the initial C input that remained within the light fraction organic matter (LFOM), particulate organic matter (POM), and mineral-associated organic matter (MAOM). Significant correlation coefficients at $P < 0.05$ are indicated in bold typing.

Variable	C recovery LFOM, %	C recovery POM, %	C recovery MAOM, %
1. Carbon, g kg^{-1}	**0.68; p = 0.000**	−0.09; p = 0.561	−0.14; p = 0.355
2. Nitrogen, g kg^{-1}	**0.58; p = 0.000**	0.08; p = 0.586	**0.41; p = 0.004**
3. C:N ratio	**−0.37; p = 0.010**	−0.16; p = 0.279	**−0.52; p = 0.000**
4. NOSC[φ]	**−0.58; p = 0.000**	0.13; p = 0.386	0.11; p = 0.451
5. N:C ratio	**0.45; p = 0.001**	−0.01; p = 0.970	**0.50; p = 0.000**
6. H:C ratio	0.13; p = 0.369	**−0.29; p = 0.047**	**−0.60; p = 0.000**
7. O:C ratio	**−0.70; p = 0.000**	0.01; p = 0.927	−0.23; p = 0.122
8. Carbohydrates	**−0.69; p = 0.000**	−0.07; p = 0.643	**−0.35; p = 0.014**
9. Proteins	**0.48; p = 0.001**	−0.01; p = 0.928	**0.47; p = 0.001**
10. Lignin	0.17; p = 0.239	**0.34; p = 0.018**	**0.68; p = 0.000**
11. Lipids	**0.49; p = 0.000**	−0.18; p = 0.209	−0.24; p = 0.094
12. Carbonyl	**−0.56; p = 0.000**	0.02; p = 0.905	−0.02; p = 0.889
13. MDI[γ]	0.08; p = 0.606	0.02; p = 0.908	**0.59; p = 0.000**

[‡] Hot-water extractables (HWE), total solvent extractables (TSE), cellulosic fraction (CF), and the acid unhydrolyzable residue (AUR) extracted from leaves, twigs, bark, and roots of eucalypt plants;

[φ] NOSC: Nominal oxidation state of carbon atoms; Variables 1–3 were determined by the results of bulk analysis of the biochemical fractions and variables 4–12 were based on the molecular mixing model (MMM) used to estimate the contents of carbohydrates, protein, lignin, lipids, and carbonyl.

[γ] NMR-based molecular diversity index (MDI); see section 2.4 for further details.

provided a metric sensitive to the distinctive properties of the biochemical fractions and their fate in soil, either entering the MAOM pool or being mineralized to CO_2.

4. Discussion

4.1. Fractionation of plant material and the molecular diversity of its biochemical fractions

Due to the complex chemistry of plant material, its characterization has long been accomplished through stepwise extractions in proximate analysis (Preston et al., 1997; Ryan et al., 1990). Lately, there has been renewed interest in such procedures owing to the perceived relevance of improving our understanding of the relationship between chemical or physical properties of plant material and the formation of SOM (Cotrufo et al., 2022). Usually, proximate analysis is applied to characterize whole plant material before (or after) incubation experiments under field or laboratory conditions (Fulton-Smith and Cotrufo, 2019; Preston and Trofymow, 2015; Shay et al., 2018; Soong et al., 2015). In our study, the sequential procedure allowed us to extract polar and non-polar compounds (HWE and TSE, respectively) and the structural components of plant material (i.e., CF and AUR). However, only categorizing the constituents of each biochemical fraction into either 'non-structural' or 'structural' materials would have limited our inference on their molecular diversity.

Based on the NMR-MMM approach, carbohydrates were predominant within the cellulosic fraction and also relevant for the hot water extractables fraction (Fig. 1), suggesting the predominance of fibrillar crystalline cellulose as well as non-cellulosic polysaccharides (i.e., hemicelluloses) in the former (Kögel-Knabner, 2002) and non-structural carbohydrates in the latter (Landhäusser et al., 2018). Non-structural carbohydrates usually include monosaccharides, disaccharides, polysaccharides, and oligosaccharides (Quentin et al., 2015). In contrast to the predominance of carbohydrates within the cellulosic fraction (i.e., above 80%), the hot water extractables also had proportionally more lignin, varying from approximately 45 to 60% of the total (Fig. 1a,c). Possibly, the hot water treatment might have dissolved monolignols that are precursors of lignin (Frei, 2013), in line with previous findings showing a wide range of phenolic compounds within the water-soluble fraction of eucalypt tissues (Chapuis-Lardy et al., 2002; Santos et al., 2011). Besides, it has been demonstrated that monolignol esters and several phenolic compounds are formed in the phenylpropanoid pathway that forms lignin (del Río et al., 2020).

Lipids and lignin were also relevant for the total solvent extractables and the acid unhydrolyzable residue, with lipids predominating in the former and aromatics concentrated in the latter (Fig. 1c,d). As a result, the molecular diversity index of the AUR was comparatively higher than that of the TSE (Fig. 2). Possibly, the major sources of lipids of the total

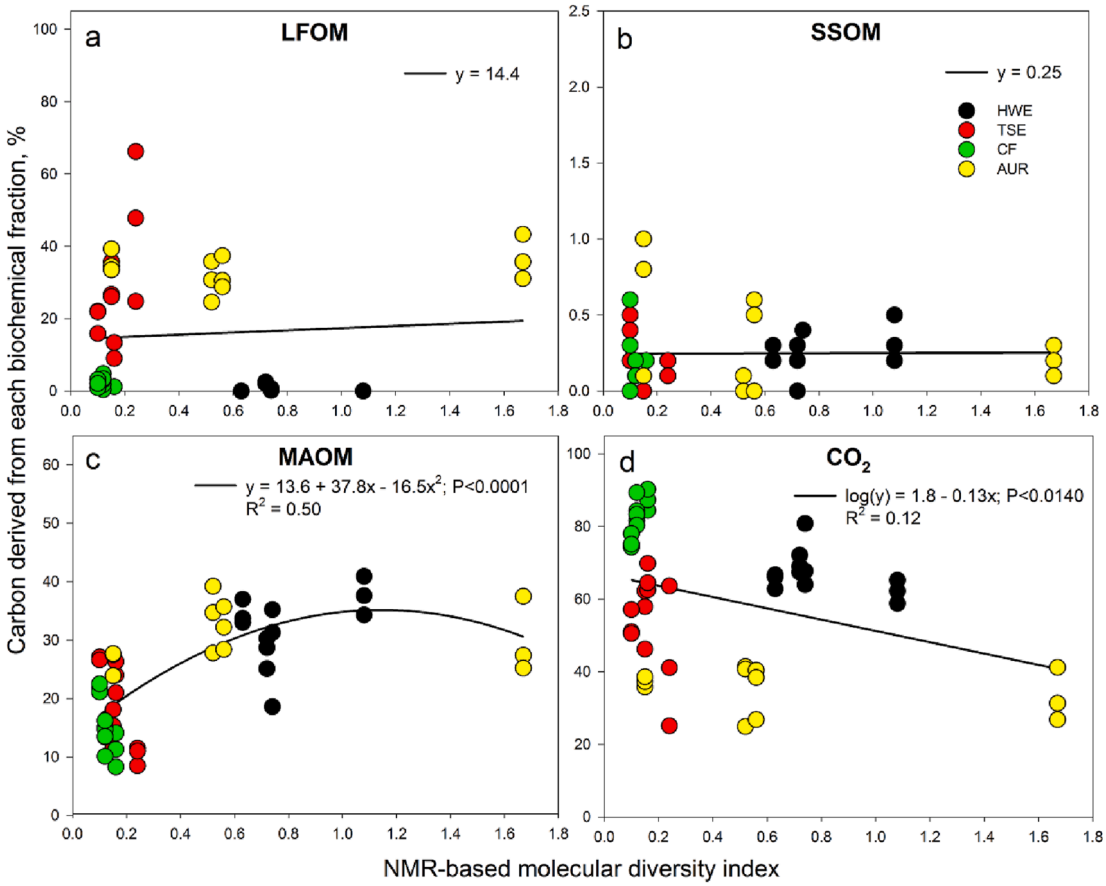

Fig. 5. Relationship between the NMR-based molecular diversity index and the partitioning (percent, %) of total C input into a) light fraction organic matter; b) sand-sized organic matter (SSOM), c) mineral-associated organic matter (MAOM), and d) total CO_2 losses, calculated as the difference between the initial C input and the amount of C retained within SOM fractions. HWE: Hot water extractables (HWE), TSE: total solvent extractables (TSE), CF: cellulosic fraction (CF), and AUR: acid unhydrolyzable residue (AUR) extracted from leaves, twigs, bark, and roots of total eucalypt plants. There were no significant relationships between MDI and the transfer of C from the biochemical fractions into the SSOM or the LFOM pools.

solvent extractables was cutin, whereas lignin and suberin should have been more relevant for the acid unhydrolyzable residue (Fich et al., 2016; Graça, 2015; Serra and Geldner, 2022). Although cutin is predominantly formed by aliphatic C, suberin includes approximately 30% of aliphatic C and about 40–45% of phenolic C (Kolattukudy, 1981). Consistent with the relevance of suberin in roots structure (Harman-Ware et al., 2021), the acid unhydrolyzable residue of roots showed comparatively higher concentration of aromatic C than all aboveground tissues (Fig. 1d). Besides, the relatively high molecular diversity index of the AUR also reflected its relatively high protein content, especially in leaves (Fig. 1d). This is in line with the N:C ratio (Fig. 3b) and the N contents varying from 16.0 to 41.6 g kg^{-1} depending on the plant organ for the acid unhydrolyzable residue fraction (Table 2). Clearly, these results show that the chemical composition of the AUR depends on the plant organ source, hence it would have been misleading to consider this fraction as a proxy for lignin. Similar findings have been reported for the tissues of seven species of temperate plants showing high N concentration within their AUR fraction (Preston and Trofymow, 2015). In fact, these authors observed that the acid unhydrolyzable residue had proportionally more total N than total phenolics or condensed tannins. Hence, whenever plant tissues are fractionated into specific biochemical fractions, determining their molecular diversity should contribute to improve the inference on their fate in soils.

Therefore, proximate analysis can be interesting to fractionate bulk plant material into specific biochemical fractions that may be tentatively grouped into 'non-structural' or 'structural' constituents. However, a simple designation of the hot water and the total solvent extractables as 'non-structural' as well as the cellulosic fraction and the acid unhydrolyzable residue as 'structural' materials undoubtedly would have minimized the distinct chemical properties of their constituents as inferred from the NMR-MMM. Critically, this approach also provides solid parameters to express qualitative attributes into quantitative terms.

4.2. *Molecular diversity and the fate of the biochemical fractions in soil*

Specific aspects might contribute to linking the molecular properties of the biochemical fractions (i.e., MDI) and their partitioning into microbial respiration to CO_2 or SOM pools, including: i) substrate solubility in water; ii) SOM formation via direct transfer (e.g., sorption) of C from the biochemical fractions *versus* or microbial-driven incorporation; iii) the microbial energy investment-return (i.e., energy input *versus* output) to metabolize each C source; and iv) the concentration of N-rich materials. Although these factors should exert joint effects on the fate of plant C in soil, some of them may be prevalent under specific constraints.

Firstly, the water solubility of the constituents of the hot water extractables was certainly relevant for their microbial respiration, as outlined by the strong correlation between CO_2 evolution and carbonylic groups, particularly during the first assessments (Table 3; $r = 0.79$; $p < 0.05$). In contrast, the likely low solubility of the lipid-rich total solvent extractables fraction probably led to its accumulation within the light fraction organic matter (Table 4 and Table 5), despite showing a relatively low molecular diversity index (Fig. 2). Thus, the polarity of the organic compounds with carbonylic groups and their solubility in

water should have favored microbial assimilation of the hot water extractables and the entry of microbial (by)products into the mineral-associated organic matter (Cotrufo et al., 2015; Kleber et al., 2015). Yet, it is not possible to distinguish the contribution of microbial (by) products from direct transfers of soluble components into the mineral-associated organic matter. Given the significant contribution of the HWE for microbial respiration (Fig. 4) and the mineral-associated organic matter pool as well (Table 4), the high molecular diversity index of this biochemical fraction is consistent with the persistence of SOM as a result of functional complexity (Lehmann et al., 2020). According to this reasoning, simpler substrates are rapidly mineralized, whereas complex molecules accumulate in SOM pools. However, the fate of the hot water extractables in our study indicates that functional complexity may not be only a consequence of microbial substrate selection, but also should reflect the formation of products of microbial anabolism.

Generally, microbial respiration increased with decreasing molecular diversity index (Table 3; Fig. 5d), and the release of CO_2 between the second and the 92nd day of the incubation indicates that microbes selected in favor of substrates probably yielding high energy return (Gunina and Kuzyakov, 2022). Indeed, high microbial respiration rates were strongly correlated with carbohydrate contents, high O:C ratios, and high NOSC values (Table 3; Fig. 3). This clearly points out to the fact that these plant residue foster a high substrate uptake and assimilation rates at a relatively low cost in energy investment (Gunina and Kuzyakov, 2022; Lehmann et al., 2020). Such high uptake was expected given the low MDI of the cellulosic fraction but inconsistent with the comparatively high molecular diversity index of the hot water extractables (Fig. 2). Yet, between 28.1 and 37.6% of the C added via HWE remained within the mineral-associated organic matter, whereas these values varied between 11.3 and 21.9% for the cellulosic fraction (Table 4). Under the assumption that microbial-derived (by)products entered the mineral-associated organic matter in both treatments, the efficiency of conversion was probably much smaller for the cellulosic fraction than for the hot water extractables.

Among the biochemical fractions studied, the acid unhydrolyzable residue included a mixture of aromatics, lipids, N-rich materials, and some carbohydrates, all of which combined to yield high MDI values (Fig. 1d; Fig. 2). Thus, more energy would have to be used to produce extracellular enzymes to break down these compounds (Almeida et al., 2018; Malik et al., 2019; Ramin and Allison, 2019). Additionally, N-rich materials within the acid unhydrolyzable residue could have been used in microbial anabolism to a greater extent than catabolism (Kallenbach et al., 2016). This aspect aligns with weak or negative correlations between CO_2 evolution and N:C ratios or protein contents (Table 3). Furthermore, the retention of N-rich compounds of the acid unhydrolyzable residue onto mineral surfaces could render these materials less susceptible to enzymatic attack (Jilling et al., 2018; Kleber et al., 2007; Kopittke et al., 2017). Although it is difficult to pinpoint the specific processes involved, the relatively high molecular diversity of the AUR would be consistent with its fate in soil: relevant contributions for the light fraction organic matter and mineral-associated organic matter, and little contributions for CO_2 emissions (Fig. 4; Table 4). This trend indicates that the microbial respiration of substrates showing high molecular diversity such as those within the AUR would be less energetically rewarding (Lehmann et al., 2020). Instead, such materials are more likely to accumulate in soil, through direct transfer into SOM pools or microbial anabolism. Generally, our data indicates that substrate molecular diversity may be more relevant to explain the entry of plant-derived materials into SOM pools than the release of CO_2.

Overall, despite the relevance of the molecular diversity index to predict the fate of the biochemical fractions in the soil in our study (Fig. 5), the most evident limitations of this index were that: i) it was not valuable to distinguish the respiration rates of the TSE obtained from distinct plant organs during the late stages of the incubation (Fig. 4), ii) both HWE and AUR had similar MDI, but contrasting respiration rates (Fig. 4); and iii) it was not relevant to explain the accumulation of C derived from the biochemical fractions within the sand-sized organic matter or the light fraction organic matter (Fig. 5). Thus, future studies are warranted to improve our comprehension of the relationship between substrate molecular diversity, its assimilation and respiration by microorganisms, as well as their entry into SOM pools.

5. Conclusions

Our results highlight the relevance of exploring quantitative approaches such as the NMR-MMM pairing to express substrate molecular diversity, which is overwhelmingly seen through qualitative lenses. Another critical issue is that the use of the NMR-MMM approach helps to constrain the chemical composition of biochemical fractions obtained from plant materials. This aspect was particularly relevant for the evaluation of the behavior of the acid unhydrolyzable residue fraction, which contribution to SOM was very significant. Importantly, the relevant proportion of N-rich materials (e.g., proteins) within the acid unhydrolyzable residue certainly underlined its strong contribution to SOM, particularly the mineral-associated C pool. Furthermore, substrate molecular diversity may be more relevant to predict the entry of plant-derived materials into SOM pools than their release in the form of CO_2. Therefore, quantitative parameters to express substrate molecular may be critical to quantify direct plant inputs from those of microbial origin entering soil C pools. This distinction should contribute to our ability to design specific strategies to maximize soil C storage while minimizing soil CO_2 fluxes.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

We thank the financial support provided by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), process/grant number 88887.356500/2019-00. We thank Maria Greiner for her help in the lab, and Isabel Prater for her assistance with the NMR spectroscopy at TU Munich, and the two anonymous reviewers for their valuable insights that helped to correct and improve our manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.geoderma.2023.116404.

References

Almeida, L.F.J., Hurtarte, L.C.C., Souza, I.F., Soares, E.M.B., Vergütz, L., Silva, I.R., 2018. Soil organic matter formation as affected by eucalypt litter biochemistry — Evidence from an incubation study. Geoderma 312, 121–129. https://doi.org/10.1016/j.geoderma.2017.10.004.

Angst, G., Mueller, K.E., Nierop, K.G.J., Simpson, M.J., 2021. Plant- or microbial-derived? A review on the molecular composition of stabilized soil organic matter. Soil Biol. Biochem. 156, 108189 https://doi.org/10.1016/j.soilbio.2021.108189.

Chapuis-Lardy, L., Contour-Ansel, D., Bernhard-Reversat, F., 2002. High-performance liquid chromatography of water-soluble phenolics in leaf litter of three Eucalyptus hybrids (Congo). Plant Sci. 163, 217–222. https://doi.org/10.1016/S0168-9452(02)00099-7.

Chen, S., Feng, X., Lin, Q., Liu, C., Cheng, K., Zhang, X., Bian, R., Liu, X., Wang, Y., Drosos, M., Zheng, J., Li, L., Pan, G., 2022. Pool complexity and molecular diversity

shaped topsoil organic matter accumulation following decadal forest restoration in a karst terrain. Soil Biol. Biochem. 166, 108553 https://doi.org/10.1016/j.soilbio.2022.108553.
Clark, R.B., 1975. Characterization of phosphatase of intact maize roots. J. Agric. Food Chem. 23, 458–460. https://doi.org/10.1021/jf60199a002.
Cotrufo, M.F., Wallenstein, M.D., Boot, C.M., Denef, K., Paul, E., 2013. The microbial efficiency-matrix stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? Glob. Chang. Biol. 19, 988–995. https://doi.org/10.1111/gcb.12113.
Cotrufo, M.F., Soong, J.L., Horton, A.J., Campbell, E.E., Haddix, M.L., Wall, D.H., Parton, W.J., 2015. Formation of soil organic matter via biochemical and physical pathways of litter mass loss. Nat. Geosci. 8, 776–779. https://doi.org/10.1038/ngeo2520.
Cotrufo, M.F., Haddix, M.L., Kroeger, M.E., Stewart, C.E., 2022. The role of plant input physical-chemical properties, and microbial and soil chemical diversity on the formation of particulate and mineral-associated organic matter. Soil Biol. Biochem. 168, 108648 https://doi.org/10.1016/j.soilbio.2022.108648.
del Río, J.C., Rencoret, J., Gutiérrez, A., Elder, T., Kim, H., Ralph, J., 2020. Lignin monomers from beyond the canonical monolignol biosynthetic pathway: Another brick in the wall. ACS Sustain. Chem. Eng. 8, 4997–5012. https://doi.org/10.1021/acssuschemeng.0c01109.
Fich, E.A., Segerson, N.A., Rose, J.K.C., 2016. The plant polyester cutin: Biosynthesis, structure, and biological roles. Annu. Rev. Plant Biol. 67, 207–233. https://doi.org/10.1146/annurev-arplant-043015-111929.
Frei, M., 2013. Lignin: characterization of a multifaceted crop component. Sci. World J. 2013, 1–25. https://doi.org/10.1155/2013/436517.
Fulton-Smith, S., Cotrufo, M.F., 2019. Pathways of soil organic matter formation from above and belowground inputs in a Sorghum bicolor bioenergy crop. GCB Bioenergy 11, 971–987. https://doi.org/10.1111/gcbb.12598.
Graça, J., 2015. Suberin: The biopolyester at the frontier of plants. Front. Chem. 3, 1–11. https://doi.org/10.3389/fchem.2015.00062.
Grandy, A.S., Neff, J.C., 2008. Molecular C dynamics downstream: The biochemical decomposition sequence and its impact on soil organic matter structure and function. Sci. Total Environ. 404, 297–307. https://doi.org/10.1016/j.scitotenv.2007.11.013.
Gunina, A., Kuzyakov, Y., 2022. From energy to (soil organic) matter. Glob. Chang. Biol. 28 (7), 2169–2182.
Haddix, M.L., Paul, E.A., Cotrufo, M.F., 2016. Dual, differential isotope labeling shows the preferential movement of labile plant constituents into mineral-bonded soil organic matter. Glob. Chang. Biol. 22, 2301–2312. https://doi.org/10.1111/gcb.13237.
Hall, S.J., Ye, C., Weintraub, S.R., Hockaday, W.C., 2020. Molecular trade-offs in soil organic carbon composition at continental scale. Nat. Geosci. 13 (10), 687–692.
Harman-Ware, A.E., Sparks, S., Addison, B., Kalluri, U.C., 2021. Importance of suberin biopolymer in plant function, contributions to soil organic carbon and in the production of bio-derived energy and materials. Biotechnol. Biofuels 14, 75. https://doi.org/10.1186/s13068-021-01892-3.
Huang, W., Ye, C., Hockaday, W.C., Hall, S.J., 2020. Trade-offs in soil carbon protection mechanisms under aerobic and anaerobic conditions. Glob. Chang. Biol. 26, 3726–3737. https://doi.org/10.1111/gcb.15100.
IUSS Working Group WRB, 2015. World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps, World Soil Resources Reports No. 106. doi:10.1017/S0014479706394902.
Jilling, A., Keiluweit, M., Contosta, A.R., Frey, S., Schimel, J., Schnecker, J., Smith, R.G., Tiemann, L., Grandy, A.S., 2018. Minerals in the rhizosphere: overlooked mediators of soil nitrogen availability to plants and microbes. Biogeochemistry 139, 103–122. https://doi.org/10.1007/s10533-018-0459-5.
Kallenbach, C.M., Frey, S.D., Grandy, A.S., 2016. Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls. Nat. Commun. 7, 13630. https://doi.org/10.1038/ncomms13630.
Kleber, M., Eusterhues, K., Keiluweit, M., Mikutta, C., Mikutta, R., Nico, P.S., 2015. Mineral–organic associations: Formation, properties, and relevance in soil environments, in: Advances in Agronomy. Elsevier Ltd, pp. 1–140. doi:10.1016/bs.agron.2014.10.005.
Kleber, M., Sollins, P., Sutton, R., 2007. A conceptual model of organo-mineral interactions in soils: self-assembly of organic molecular fragments into zonal structures on mineral surfaces. Biogeochemistry 85, 9–24. https://doi.org/10.1007/s10533-007-9103-5.
Kögel-Knabner, I., 2002. The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter. Soil Biol. Biochem. 34, 139–162. https://doi.org/10.1016/S0038-0717(01)00158-4.
Kögel-Knabner, I., Rumpel, C., 2018. Advances in molecular approaches for understanding soil organic matter composition, origin, and turnover: A historical overview. In: Advances in Agronomy. Elsevier Inc., pp. 1–48. https://doi.org/10.1016/bs.agron.2018.01.003
Kolattukudy, P.E., 1981. Structure, biosynthesis, and biodegradation of cutin and suberin. Annu. Rev. Plant Physiol. 32, 539–567. https://doi.org/10.1146/annurev.pp.32.060181.002543.
Kopittke, P.M.M., Hernandez-Soriano, M.C.C., Dalal, R.C.C., Finn, D., Menzies, N.W.W., Hoeschen, C., Mueller, C.W.W., 2017. Nitrogen-rich microbial products provide new organo-mineral interactions for the stabilization of soil organic matter. Glob. Chang. Biol. 24, 1–9. https://doi.org/10.1111/gcb.14009.
Landhäusser, S.M., Chow, P.S., Dickman, L.T., Furze, M.E., Kuhlman, I., Schmid, S., Wiesenbauer, J., Wild, B., Gleixner, G., Hartmann, H., Hoch, G., McDowell, N.G., Richardson, A.D., Richter, A., Adams, H.D., Mencuccini, M., 2018. Standardized protocols and procedures can precisely and accurately quantify non-structural carbohydrates. Tree Physiol. 38 (12), 1764–1778.
Lehmann, J., Hansel, C.M., Kaiser, C., Kleber, M., Maher, K., Manzoni, S., Nunan, N., Reichstein, M., Schimel, J.P., Torn, M.S., Wieder, W.R., Kögel-Knabner, I., 2020. Persistence of soil organic carbon caused by functional complexity. Nat. Geosci. 13 (8), 529–534.
Liang, C., Schimel, J.P., Jastrow, J.D., 2017. The importance of anabolism in microbial control over soil carbon storage. Nat. Microbiol. 2, 17105. https://doi.org/10.1038/nmicrobiol.2017.105.
Machado, D.N., Novais, R.F., da Silva, I.R., Loureiro, M.E., Milagres, J.J., Soares, E.M.B., 2011. Enriquecimento e alocação de ^{13}C em plantas de eucalipto. Rev. Bras. Cienc. do Solo 35, 857–866. https://doi.org/10.1590/S0100-06832011000300020.
Malik, A.A., Puissant, J., Goodall, T., Allison, S.D., Griffiths, R.I., 2019. Soil microbial communities with greater investment in resource acquisition have lower growth yield. Soil Biol. Biochem. 132, 36–39. https://doi.org/10.1016/j.soilbio.2019.01.025.
Nelson, P.N., Baldock, J.A., 2005. Estimating the molecular composition of a diverse range of natural organic materials from solid-state ^{13}C NMR and elemental analyses. Biogeochemistry 72, 1–34. https://doi.org/10.1007/s10533-004-0076-3.
Osono, T., Takeda, H., Azuma, J.I., 2008. Carbon isotope dynamics during leaf litter decomposition with reference to lignin fractions. Ecol. Res. 23, 51–55. https://doi.org/10.1007/s11284-007-0336-5.
Pettersen, R.C., 1984. The chemical composition of wood. In: R.M.R. (Ed.), The chemistry of solid wood. ACS Publications, Washington, pp. 57–126. https://doi.org/10.1021/ba-1984-0207.ch002.
Preston, C.M., Nault, J.R., Trofymow, J.A., 2009. Chemical changes during 6 years of decomposition of 11 litters in some Canadian forest sites. Part 2. ^{13}C abundance, solid-state 13C NMR spectroscopy and the meaning of "lignin". Ecosystems 12, 1078–1102. https://doi.org/10.1007/s10021-009-9267-z.
Preston, C.M., Trofymow, J.A., 2000. Variability in litter quality and its relationship to litter decay in Canadian forests. Can. J. Bot. 78 (10), 1269–1287.
Preston, C.M., Trofymow, J.A., 2015. The chemistry of some foliar litters and their sequential proximate analysis fractions. Biogeochemistry 126, 197–209. https://doi.org/10.1007/s10533-015-0152-x.
Preston, C.M., Trofymow, J.A., Sayer, B.G., Niu, J., 1997. ^{13}C nuclear magnetic resonance spectroscopy with cross-polarization and magic-angle spinning investigation of the proximate-analysis fractions used to assess litter quality in decomposition studies. Can. J. Bot. 75, 1601–1613. https://doi.org/10.1139/b97-872.
Quentin, A.G., Pinkard, E.A., Ryan, M.G., Tissue, D.T., Baggett, L.S., Adams, H.D., Maillard, P., Marchand, J., Landhäusser, S.M., Lacointe, A., Gibon, Y., Anderegg, W.R.L., Asao, S., Atkin, O.K., Bonhomme, M., Claye, C., Chow, P.S., Clément-Vidal, A., Davies, N.W., Dickman, L.T., Dumbur, R., Ellsworth, D.S., Falk, K., Galiano, L., Grünzweig, J.M., Hartmann, H., Hoch, G., Hood, S., Jones, J.E., Koike, T., Kuhlmann, I., Lloret, F., Maestro, M., Mansfield, S.D., Martínez-Vilalta, J., Maucourt, M., McDowell, N.G., Moing, A., Muller, B., Nebauer, S.G., Niinemets, Ü., Palacio, S., Piper, F., Raveh, E., Richter, A., Rolland, G., Rosas, T., Joanis, B.S., Sala, A., Smith, R.A., Sterck, F., Stinziano, J.R., Tobias, M., Unda, F., Watanabe, M., Way, D.A., Weerasinghe, L.K., Wild, B., Wiley, E., Woodruff, D.R., 2015. Non-structural carbohydrates in woody plants compared among laboratories. Tree Physiol. 35, 1146–1165. https://doi.org/10.1093/treephys/tpv073.
Ramin, K.I., Allison, S.D., 2019. Bacterial tradeoffs in growth rate and extracellular enzymes. Front. Microbiol. 10, 1–10. https://doi.org/10.3389/fmicb.2019.02956.
Reid, J.D., Lynch, D.F.J., 1937. Cellulose analysis. A comparison of three principal methods. Ind. Eng. Chem. Anal. Ed. 9, 570–573. https://doi.org/10.1021/ac50116a010.
Ryan, M.G., Melillo, J.M., Ricca, A., 1990. A comparison of methods for determining proximate carbon fractions of forest litter. Can. J. For. Res. 20, 166–171. https://doi.org/10.1139/x90-023.
Santos, S.A.O., Freire, C.S.R., Domingues, M.R.M., Silvestre, A.J.D., Neto, C.P., 2011. Characterization of phenolic components in polar extracts of Eucalyptus globulus Labill. bark by high-performance liquid chromatography–mass spectrometry. J. Agric. Food Chem. 59, 9386–9393. https://doi.org/10.1021/jf201801q.
Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A.C., Nannipieri, P., Rasse, D.P., Weiner, S., Trumbore, S.E., 2011. Persistence of soil organic matter as an ecosystem property. Nature 478, 49–56. https://doi.org/10.1038/nature10386.
Serra, O., Geldner, N., 2022. The making of suberin. New Phytol 235 (3), 848–866.
Shay, P.E., Peter Constabel, C., Trofymow, J.A., 2018. Evidence for the role and fate of water-insoluble condensed tannins in the short-term reduction of carbon loss during litter decay. Biogeochemistry 137, 127–141. https://doi.org/10.1007/s10533-017-0406-x.
Sjöberg, G., Nilsson, S.I., Persson, T., Karlsson, P., 2004. Degradation of hemicellulose, cellulose and lignin in decomposing spruce needle litter in relation to N. Soil Biol. Biochem. 36, 1761–1768. https://doi.org/10.1016/j.soilbio.2004.03.010.
Soong, J.L., Parton, W.J., Calderon, F., Campbell, E.E., Cotrufo, M.F., 2015. A new conceptual model on the fate and controls of fresh and pyrolized plant litter decomposition. Biogeochemistry 124, 27–44. https://doi.org/10.1007/s10533-015-0079-2.
Zech, W., Johansson, M.-B., Haumaier, L., Malcolm, R.L., 1987. CPMAS ^{13}C NMR and IR spectra of spruce and pine litter and of the Klason lignin fraction at different stages of decomposition. Z. Pflanzenernährung Bodenkd. 150, 262–265. https://doi.org/10.1002/jpln.19871500413.