Using Multi-Homologue Plant-Wax Carbon Isotope Signatures to Reconstruct Tropical Vegetation Types

C. Häggi[1*], D. J. Bertassoli Jr.[2,3], T. K. Akabane[3], R. T. So[1], A. O. Sawakuchi[3], C. M. Chiessi[3], V. R. Mendes[4], C. A. Jaramillo[5], S. J. Feakins[1]

[1] Department of Earth Sciences, University of Southern California, 3651 Trousdale Pkwy, Los Angeles CA-90089, USA

[2] School of Arts, Sciences and Humanities, University of São Paulo, Av. Arlindo Bettio 1000, 03828-000 São Paulo SP, Brazil

[3] Institute of Geosciences, University of São Paulo, Rua do Lago 562, 05508-080 São Paulo SP, Brazil

[4] Institute of Marine Science, Federal University of São Paulo, Rua Dr. Carvalho de Mendonça 144, Santos 110-070, Brazil

[5] Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancón, Panama

Corresponding author: Christoph Häggi (christoph.haeggi@gmx.ch)

*Now at: Intercantonal Laboratory Schaffhausen, Mühlentalstrasse 188, 8200 Schaffhausen, Switzerland

Key Points:

- *n*-Alkane and fatty acid carbon isotope compositions were studied to provide proxy endmembers for tropical South American vegetation types
- Forest vegetation types show a narrow carbon isotope range facilitating the detection of minor savanna incursions into rainforest
- The offset of the carbon isotope composition of different *n*-alkane homologues can be used to differentiate savannas and shrublands

Abstract

The stable carbon isotope composition ($\delta^{13}C$) of plant components such as plant wax biomarkers is an important tool for reconstructing past vegetation. Plant wax $\delta^{13}C$ is mainly controlled by photosynthetic pathways, allowing for the differentiation of C_4 tropical grasses and C_3 forests. Proxy interpretations are however complicated by additional factors such as aridity, vegetation density, elevation, and the considerable $\delta^{13}C$ variability found among C_3 plant species. Moreover, studies on plant wax $\delta^{13}C$ in tropical soils and plants have focused on Africa, while structurally different South American savannas, shrublands and forests remain understudied. Here, we analyze the $\delta^{13}C$ composition of long-chain *n*-alkanes and fatty acids from tropical South American soils and leaf litter to assess the isotopic variability in each vegetation type and to investigate the influence of climatic features on $\delta^{13}C$. Rainforests and open vegetation types show distinct values, with rainforests having a narrow range of low $\delta^{13}C$ values (*n*-C_{29} *n*-alkane: $-34.5^{+0.9}_{-0.6}$‰ $\left(Q_{25}^{75}\right)$; Suess-effect corrected) allowing for the detection of even minor incursions of savanna into rainforests (^{13}C-enriched). While Cerrado savannas and semi-arid Caatinga shrublands grow under distinctly different climates, they can yield indistinct $\delta^{13}C$ values for most compounds. Cerrado soils and litter show pronounced isotopic spreads between the *n*-C_{33} and *n*-C_{29} alkanes, while Caatinga shrublands show consistent values across the two homologues, thereby enabling the differentiation of these vegetation types. The same multi-homologue isotope analysis can be extended to differentiate African shrublands from savannas.

Plain Language Summary

The reconstruction of past vegetation dynamics is key for the understanding of the impact of future climate variability on ecosystems. One of the most widely used tools to reconstruct past vegetation from sediment deposits are plant waxes – comparably stable molecules that form the wax coating of leaves. The ratio of heavier and lighter carbon isotopes preserved in plant waxes can be used to differentiate between rainforest and tropical savanna vegetation. This method has been frequently applied in African vegetation types. Other tropical regions such as South America, which have different vegetation structure, remain understudied. In our study, we characterize the plant wax carbon isotope composition of the major tropical South American

vegetation types. One of the complications of the method in both African and South American vegetation types is that (semi-) arid shrublands and savannas show similar plant wax carbon isotope values. To further differentiate between arid shrublands and savannas, we show that the comparison of the carbon isotope values from different plant waxes can be useful both in Africa and South America.

1. Introduction

Tropical South America is home to diverse tropical vegetation types such as the Amazon and Atlantic rainforests, the Cerrado savanna, the Llanos savanna and the Caatinga shrublands, which include the most biodiverse rainforest and savanna biomes on Earth (Jenkins et al., 2013; Kier et al., 2009). Reconstructions of past tropical vegetation dynamics are essential to understand both the sensitivity of tropical vegetation types to climate variability and the origin of regional biodiversity (Brienen et al., 2015; Cox et al., 2013; Häggi et al., 2017; Salati et al., 1979; Zemp et al., 2017). Aside from the study of plant microfossils (pollen and phytoliths) in sedimentary archives, the main tool used in these reconstructions is the carbon isotope composition ($\delta^{13}C$) of plant organic matter (Cerling et al., 2011; Feakins et al., 2013; Huang et al., 2000; Schefuß et al., 2005). This approach is based on the different carbon fixation mechanisms used by tropical forest and grass species. While most plants, including tropical forest species, use the C_3 metabolism that involves direct carbon fixation through the Calvin cycle, the C_4 metabolism has an added additional step of pre-concentration of CO_2 through the Hatch-Slack cycle prior to fixation via the Calvin cycle (Kortschak et al., 1965; Slack & Hatch, 1967). The C_3 pathway discriminates against the heavy isotope ^{13}C, preferentially taking up ^{12}C. Due to the higher efficiency of carbon fixation by C_4 metabolism, heavy isotope discrimination is less pronounced in C_4 plants, which leads to contrasting stable carbon isotope compositions in plant organic matter from tropical grasses and trees (O'Leary, 1981; Rieley et al., 1991).

The $\delta^{13}C$ compositions of soil organic matter, soil carbonates and tooth enamel have been widely used for reconstructions of past vegetation from terrestrial settings (Cerling et al., 1997; Cerling et al., 2011). The $\delta^{13}C$ composition of refractory plant components such as plant wax biomarkers can also be used for applications in marine, lacustrine and aeolian sediment archives that integrate a larger catchment (Huang et al., 2001; W. Liu et al., 2005; Schefuß et al., 2005). The

most commonly used plant wax biomarkers are long-chain *n*-alkanes and long-chain fatty acids that form the wax coating of leaves (Eglinton & Hamilton, 1967) and their isotope compositions remains stable during deposition and post depositional degradation (Häggi et al., 2021). During their synthesis, fractionation yields lower $\delta^{13}C$ values than bulk organic matter (Collister et al., 1994; Rieley et al., 1991). Detailed studies on the $\delta^{13}C$ composition of plant wax homologues of varying chain lengths from individual C_3 and C_4 species found a larger range of isotope values for C_3 plants than for C_4 plants (Boom et al., 2014; Garcin et al., 2014; Krull et al., 2006; Rommerskirchen et al., 2006; Vogts et al., 2009; Wu et al., 2017). In addition to the dominant impact of the photosynthetic metabolism, climate conditions and altitude have been found to act as secondary forcing mechanisms controlling plant wax $\delta^{13}C$ compositions, with aridity and higher altitudes leading to higher $\delta^{13}C$ values (Ceccopieri et al., 2021; Diefendorf et al., 2010; Wu et al., 2017).

The $\delta^{13}C$ composition of plant waxes has also been studied in soils, as well as in marine and lacustrine sediments covering gradients between tree-domainted and grass-dominated vegetation. In mixed vegetation types, the *n*-C_{29} *n*-alkane, produced in large concentrations by tropical trees, yields lower $\delta^{13}C$ values values than the *n*-C_{31} and *n*-C_{33} *n*-alkanes evenly produced by both trees and grasses (Douglas et al., 2012; Garcin et al., 2014; Rommerskirchen et al., 2003; Schwab et al., 2015; D. Zhang et al., 2021). Hence, the $\delta^{13}C$ composition of *n*-C_{29} is particularly sensitive to limited tree vegetation in a grass-dominated, while the $\delta^{13}C$ of *n*-C_{33} can better detect limited grass expansions in a forest-dominated biome (Garcin et al., 2014). The study of multiple plant wax homologues has thereby the potential to elucidate the vegetation structure in greater detail than the $\delta^{13}C$ of bulk organic matter. The variance in the isotope composition of plant waxes of different chain-length has also been suggested as a tool for assessing past plant biodiversity (Magill et al., 2019). Calibration studies on plant wax biomarker $\delta^{13}C$ from tropical savanna and forest areas have so far focused on Africa and there are only few other studies from Central America (Douglas et al., 2012) and northern Australia (Krull et al., 2006). There are, however, no systematic studies on the savanna and shrubland vegetation types of tropical South America.

South American savanna vegetation types show marked differences to African and Australian counterparts. For instance, they exist under more humid climate conditions than African and

Australian savannas and evolved different responses to fire and herbivory disturbances (Dantas & Pausas, 2013; Hirota et al., 2011; Lehmann et al., 2014). This resulted in a pattern where African savanna vegetation features higher stems more adapted to herbivory disturbance, while South American savanna vegetation has thicker bark as an adaptation to high fire activity (Dantas & Pausas, 2013; Lehmann et al., 2014). The unique vegetation structure of the South American Cerrado savanna and Caatinga shrubland and the lack of plant-wax isotope data from these vegetation types presents a knowledge gap that has implications for the accurate interpretation of plant wax signals recorded in marine, lacustrine and riverine sediment cores from the region (Bertassoli et al., 2019; Ferreira et al., 2022; Fornace et al., 2016; Häggi et al., 2017; Mulitza et al., 2017; Reis et al., 2022).

In tropical South America, the study of the $\delta^{13}C$ composition of plant waxes in leaf, soil and fluxival sediment samples has focused on the Amazon rainforest. On a transect across the Andes-Amazon, Wu et al. (2017), found a trend towards higher $\delta^{13}C$ values with increasing altitude. The transport of plant wax compounds and their $\delta^{13}C$ values has been studied in a broader area covering the Amazon River main stem and some of its tributaries (Bertassoli et al., 2022; Feakins et al., 2018; Häggi et al., 2016). The study of plant waxes derived from open vegetation types is so far limited to marine sediments offshore the Caatinga shrubland in northeastern Brazil, where long-chain *n*-alkane $\delta^{13}C$ values were higher than in the Amazon plume (Häggi et al., 2016). Yet, there is no detailed information on the $\delta^{13}C$ composition of plant waxes from different savanna and shrubland vegetation types in tropical South America.

To fill this gap, we studied the $\delta^{13}C$ composition of long-chain *n*-alkanes, long-chain fatty acids and bulk organic matter from leaf-litter and soil samples across all major vegetation types of tropical South America, including the Amazon and Atlantic rainforests, the Cerrado savanna, the Caatinga shrubland and the savannas and riparian forests of the Llanos. We first assess the relative distribution of plant wax homologues, to test if assumptions on different distributions between C_3 and C_4 plants observed elsewhere also hold in tropical South America. We then report the $\delta^{13}C$ endmember values of the different tropical South American vegetation types. Finally, we use the $\delta^{13}C$ composition of plant wax homologues of different chain-length to

further assess differences between South American savanna and shrubland vegetation types and compare our findings to African tropical vegetation.

2. Materials and methods

2.1.Study area

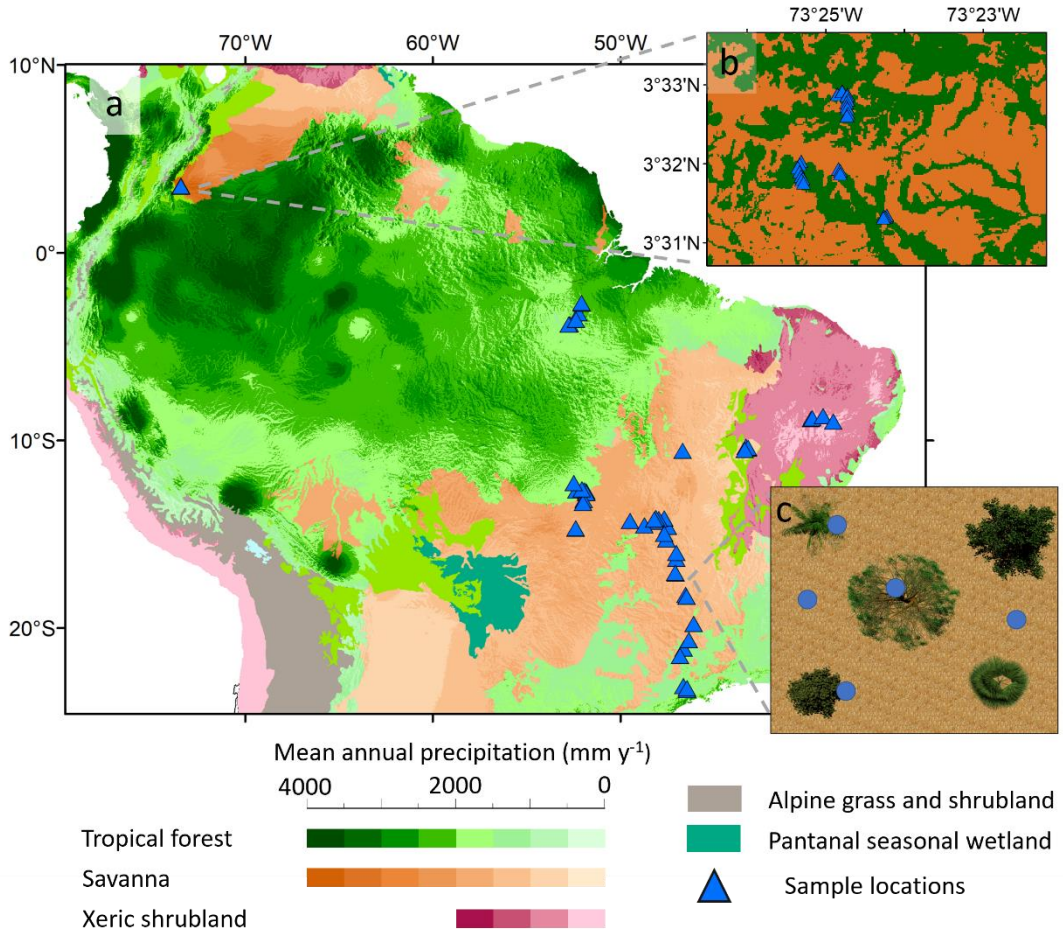


Figure 1. Study area and sampling strategy. a) Vegetation and precipitation map of tropical South America (Karger et al., 2017; Olson et al., 2001). Coloring indicates the major biomes of tropical South America (Olson et al., 2001), while color shades represent mean annual precipitation (MAP). Note the separate shading for each biome. Sample locations are given in blue triangles. b) Close-up view of the sampling strategy in the Llanos where transects over the savanna-riparian forest mosaic were studied. c) Schematic overview of how the five samples at each site in the Cerrado savanna were taken in proximity to trees, shrubs and grasses.

The study area covers all major biomes in tropical South America including the Atlantic rainforest, the Amazon rainforest, the Cerrado savanna types of varying tree density, the Llanos savanna and the Caatinga shrubland. The Atlantic rainforest can be found along the eastern coast of South America (Fig. 1) extending from temperate to tropical latitudes, and consists of a humid rainforest vegetation that becomes gradually drier inland, where it merges into Cerrado savanna vegetation (De Azevedo, 1950; Oliveira-Filho & Fontes, 2000). The areas close to the equator between 10°S and 5°N are mostly covered by the Amazon rainforest (Fig. 1). The most humid parts of the Amazon rainforest are found in the northwestern parts of the Amazon Basin, while in the south and east less humid conditions persist (Karger et al., 2017). Cerrado spreads between the Amazon and the Atlantic rainforests and consists of a continuum of savanna vegetation of varying density (Fig. S1) (da Silva & Bates, 2002). The densest physiognomic variety is the Cerradão, which is a dry forest vegetation type that contains little to no grasses (Eiten, 1972). The Cerrado sensu stricto is a mixed savanna with trees and interspersed grasses (Goodland, 1971). The Campo Cerrado is a more open Cerrado variety with fewer trees and shrubs that are smaller in size than in Cerrado sensu stricto (Goodland, 1971). The open most varieties are Campo Sujo and Campo Limpo, which are open grass-dominated vegetation types that contain at most a few small shrubs (Goodland, 1971). The distribution of these Cerrado physiognomies depends on proximity to rivers and creeks and on soil properties that define water availability through the dry season (Ruggiero et al., 2002). To the north of the Amazon Basin (3 to 7°N), the Llanos savanna stretches between the Andes and the mouth of the Orinoco River (Fig. 1). The Llanos is also made up of a variety of different types of varying tree density (Blydenstein, 1967). The Llanos also contains the Riparian Forest, a type of forest that is only found along the river margins and represents a subset of the Amazonian Rainforest where water is available all year long via water-table (Jaramillo, 2023). Caatinga shrubland vegetation (also called xerophytic forest) is found in northeastern Brazil and constitutes the driest of the studied vegetation types (Fig. 1). Caatinga vegetation is mostly made up of deciduous shrubs and trees but never with a continuous canopy, as well as succulents and grasses (da Silva & Lacher, 2020; de Queiroz, 2006). In contrast to the Cerrado vegetation types, understory grass cover is less dense and sometimes absent from Caatinga shrublands (Lloyd et al., 2008).

Mean annual precipitation (MAP) in the study area varies between 500 mm y^{-1} in the Caatinga of northeastern Brazil and 3500 mm y^{-1} in the northwestern Amazon and the Llanos Basin (Fig. 1) (Karger et al., 2017). While Caatinga shrublands are characterized by distinctly lower precipitation values than the surrounding Cerrado savanna, the boundary between the Llanos savanna and the Amazon rainforest in the northern portion of the continent features continuously high MAP values (Fig. 1). Rather than on MAP, the boundary between Llanos savanna and rainforest is heavily dependent on the duration of wet spells during the dry season (Hoyos et al., 2022). Likewise, the boundary between the Cerrado and the drier southern part of the Amazon rainforest is also not defined by a distinct MAP boundary (Fig. 1).

Mean annual temperatures vary between 18 °C in portions of the Atlantic rainforest areas and 27 °C in the Caatinga shrublands (Karger et al., 2017). The Caatinga shrubland, therefore, features both the warmest and driest conditions in the study area. The low temperatures in the Atlantic rainforest sites are related to its latitudinal position and elevated altitude in the mountainous areas of southeastern Brazil.

2.2.Sampling

Soil and leaf litter samples were collected during three field campaigns to the different South American vegetation types (Häggi, Naafs, et al., 2023). Soil samples consist of the upper five centimeters of the A-horizon, while leaf litter samples consist of loose-leaf material found on top of the sampled soils. Samples from the Atlantic rainforest, the Cerrado, and the southern Amazon rainforest were collected in April 2019. Sampling was conducted on sites featuring undisturbed primary vegetation. Special care was taken to avoid locations where invasive species such as grasses of the genus *Brachiaria* were present (Ratter et al., 1997). Locations with signs of reforestation were also avoided. Cerrado sampling sites were selected to cover the full density range of this vegetation type. Five locations sampled Campo Sujo and Campo Limpo, both open Cerrado types. Seven locations featured the Cerrado Sensu Stricto and five locations covered the more open Campo Cerrado. Five locations featured Cerradão, the tree-dominated, closed canopy Cerrado type. Five locations were sampled from both the Amazon rainforest and the Atlantic rainforest. For each sampling site, five samples were collected from an area of up to 15x15 meters. In Cerrado areas, sampling was conducted in locations both close and distal from trees.

This strategy was applied because previous studies on savanna vegetation in Africa and Australia found that the bulk $\delta^{13}C$ isotope composition of soils varied depending on their proximity to trees (Bird et al., 2004; Wynn, 2007). Soil samples from the Llanos Basin were collected in February 2019. The sample set consists of transects starting in riverine back swamps and connecting riparian forest vegetation with savanna vegetation. Soil samples from Caatinga shrubland vegetation samples were taken in March 2020 and followed the same sampling protocol described for Cerrado vegetation . In addition, we also analyzed soil and litter samples from 6 isolated sites from the interior of the Amazon rainforest and 2 isolated sites from the Caatinga shrublands collected between 2016 and 2019.

2.3.Laboratory procedures

2.3.1. Bulk organic analyses

Total organic carbon (TOC) concentration and carbon isotopic composition ($\delta^{13}C$ OC) were analyzed at the Kentucky Stable Isotope Geochemistry Laboratory (KSIGL) at the University of Kentucky using a Costech 4010 Elemental Analyzer coupled via a Conflo IV to a Thermo Finnigan DELTAplus XP Isotope Ratio Mass Spectrometer (IRMS). Freeze-dried, sieved (1 mm mesh size) and homogenized soil samples (~50 mg) were decarbonized with 2 mL 1 M HCl at 50°C for 2 hours (with agitation every 30 minutes). This step was repeated as needed. Samples were then rinsed with deionized water (three times or until pH reached 6-7), centrifuged, freeze-dried and reweighed to assess carbonate loss. Leaf litter samples were freeze-dried and powdered. Samples were weighed into pre-cleaned Sn capsules with 1-3 mg of WO_3 to promote combustion. TOC concentrations were determined relative to the known TOC concentrations of USGS64 (TOC = 32 %) and accounted for carbonate loss where applicable. All carbon isotope values were normalized to the Vienna Pee Dee Belemnite (VPDB) scale using two internationally accepted reference materials (RMs): USGS64 glycine ($\delta^{13}C$ = –40.81‰) and USGS41 L-glutamic acid (+37.63‰). In each analytical session, precision and accuracy were assessed with multiple blind analyses of a matrix-matched RM (NIST 8704 Buffalo River Sediment; $\delta^{13}C$ = –19.86‰ per inter-lab comparison). The accuracy of the NIST 8704 standard was 0.1% for TOC analyses and 0.1‰ for $\delta^{13}C$ analyses. The precision was 0.1% for TOC analyses and 0.2‰ for $\delta^{13}C$ analyses. Sample reproducibility averaged 0.21 % (TOC) and 0.26 ‰ ($\delta^{13}C$).

2.3.2. Lipid extraction and separation

For lipid analysis, at the University of Southern California, samples were extracted using a Dionex Instruments ASE350 accelerated solvent extractor for two cycles of 15 minutes at 100 °C and 1500 psi using dichloromethane (DCM): methanol (MeOH) 9: 1 as solvent. Total lipid extracts (TLE) were dried under a stream of N_2. TLEs were separated into neutral and acid fractions over LC-NH_2 gel columns by subsequent elution in DCM: isopropanol 2: 1 and ethyl ether: formic acid 25: 1. The neutral fraction was further separated into apolar and polar fractions using silica gel columns and hexane and MeOH and DCM as subsequent solvents. The apolar fraction containing the *n*-alkanes was further cleaned over $AgNO_3$ coated silica gel columns using hexane as solvent to remove unsaturated compounds. The acid fraction was methylated at 70 °C for 12 hours using MeOH: HCL 95: 5. The isotope composition of MeOH was -36.8 ± 0.2 ‰ vs. VPDB. The methylated fatty acid methyl esters (FAMEs) were liquid-liquid extracted using MiliQ water and hexane and transferred over Na_2SO_4 columns to remove residual water. Dried FAME fractions were then cleaned over silica gel columns using DCM as solvent. As an additional cleaning step, FAME fractions were cleaned over $AgNO_3$ coated silica gel columns using hexane and DCM as subsequent solvents.

2.3.3. Lipid quantification

FAMEs and *n*-alkanes were quantified using an Agilent 6890 gas chromatograph coupled with a mass spectrometer (Agilent 5973) and flame ionization detector (GC–MS/FID). Compounds were identified by comparison to an external standard solution as well as their mass spectra. Quantification was achieved by comparison of the integrated peak areas to external standard solutions of known composition and concentration. The average chain length (ACL_{Alk} and ACL_{FAME}) and the carbon preference index (CPI_{Alk} and CPI_{FAME}), that is a measure of the odd vs. even chain length distribution (Cranwell, 1981), are calculated for this study as follows:

$$ACL_{Alk27-33} = \frac{27 * n\text{-}C_{27} + 29 * n\text{-}C_{29} + 31 * n\text{-}C_{31} + 33 * n\text{-}C_{33}}{n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31} + n\text{-}C_{33}} \quad (1)$$

$$ACL_{Alk29-33} = \frac{29 * n\text{-}C_{29} + 31 * n\text{-}C_{31} + 33 * n\text{-}C_{33}}{n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31} + n\text{-}C_{33}} \quad (2)$$

$$ACL_{FAME26-32} = \frac{26 * n\text{-}C_{26} + 28 * n\text{-}C_{28} + 30 * n\text{-}C_{30} + 32 * n\text{-}C_{33}}{n\text{-}C_{26} + n\text{-}C_{28} + n\text{-}C_{30} + n\text{-}C_{32}} \quad (3)$$

$$ACL_{FAME24-30} = \frac{24 * n\text{-}C_{24} + 26 * n\text{-}C_{26} + 28 * n\text{-}C_{28} + 30 * n\text{-}C_{30}}{n\text{-}C_{24} + n\text{-}C_{26} + n\text{-}C_{28} + n\text{-}C_{30}} \quad (4)$$

$$CPI_{Alk} = 0.5 \times \left(\frac{n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31} + n\text{-}C_{33}}{n\text{-}C_{26} + n\text{-}C_{28} + n\text{-}C_{30} + n\text{-}C_{32}} + \frac{n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31} + n\text{-}C_{33}}{n\text{-}C_{28} + n\text{-}C_{30} + n\text{-}C_{32} + n\text{-}C_{34}}\right) \quad (5)$$

$$CPI_{FAME} = 0.5 \times \left(\frac{n\text{-}C_{26} + n\text{-}C_{28} + n\text{-}C_{30} + n\text{-}C_{32}}{n\text{-}C_{25} + n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31}} + \frac{n\text{-}C_{26} + n\text{-}C_{28} + n\text{-}C_{30} + n\text{-}C_{32}}{n\text{-}C_{27} + n\text{-}C_{29} + n\text{-}C_{31} + n\text{-}C_{33}}\right) (6)$$

While long-chain *n*-alkanes with a chain-length between 27 and 35 carbon atoms are mostly sourced by land plants, alkanes of a shorter chain-length are mostly sourced by microbes. Hence, the ratio of mid- to long-chain *n*-alkanes ($R_{m/l}$) can be used to analyze the relative input of microbial and plant derived compounds

$$R_{m/l\ Alk} = \frac{\sum_{i=21}^{25} n\text{-}C_i}{\sum_{i=26}^{35} n\text{-}C_i} \quad (7)$$

Likewise, we assessed the ratio of mid- to long-chain FAMEs.

$$R_{m/l\ FAME} = \frac{\sum_{i=14}^{23} n\text{-}C_i}{\sum_{i=24}^{34} n\text{-}C_i} \quad (8)$$

2.3.4. Compound-specific carbon isotope analysis

Compound-specific stable $\delta^{13}C$ compositions of long-chain *n*-alkanes and long-chain fatty acids were analyzed by gas chromatography–isotope ratio mass spectrometry (GC-IRMS) using a Thermo Scientific Trace gas chromatograph connected to a Delta V Plus mass spectrometer via an Isolink combustion furnace operated at 1000 °C. $\delta^{13}C$ values were determined by comparison to a reference gas and normalized to the Vienna Pee Dee Belemnite (VPDB) standard by

comparison to an external *n*-alkane standard containing the C_{16} to C_{30} *n*-alkanes and an isotope range between -28.6 and -33.3 ‰ (A3-mix by Arndt Schimmelmann, Indiana University). Measurement uncertainty was determined by calculating the root mean square error (RMSE) of the relationship between known and measured standard values used for normalization. RMSE was 0.19 ± 0.07 ‰ vs. VPDB through the measurement period. Linearity of CO_2 reference gas was checked daily across a 1 to 9 V amplitude range. The isotope composition of FAMEs was corrected for the methyl group added during methylation. The majority of the samples were analyzed in duplicate.

2.3.5. Statistical analysis

Statistical analyses reported in this manuscript were performed with the statistical software R (R_Core_Team, 2022). The violin plots shown throughout the manuscript were created using the R-package 'vioplot' (Adler & Kelly, 2019). Tests performed included Shapiro-Wilk-tests to test for normalcy of the data, and unpaired two-sample Wilcoxon tests to test the significance of differences between vegetation types and between soil and litter samples. The test was selected since Shapiro-Wilk-tests indicated that some of the compared sample sets did not follow a normal distribution. For the same reason we also report uncertainties as interquartile ranges. For the tests between litter and soil samples, the comparison of multiple vegetation types and parameters in our sample set lead to a multiple comparisons problem, where significant differences (i.e., with a significance level of $\alpha < 0.05$) would be expected to arise by chance. To account for this issue, we used the Bonferroni correction, which is a conservative measure to correct significance levels for multiple comparisons:

$$\alpha_{\text{corrected}} = \frac{\alpha_{\text{original}}}{n} \quad (9)$$

Where $\alpha_{\text{corrected}}$ is the corrected significance level, α_{original} the original significance level and *n* the number of comparisons made.

2.3.6. Correction for the Suess-effect

To facilitate the use of our modern calibration data set for paleo vegetation reconstructions we report isotope values after correction for the Suess-effect to yield pre-industrial equivalents. In pre-industrial times (before 1750), the long-term average atmospheric $\delta^{13}C$ was around -6.5‰

vs. VPDB, which shifted to -8.5‰ vs. VPDB in 2019, when most of the samples were collected (Rubino et al. 2013, Keeling et al. 2005). To correct the values from litter samples we assumed recent deposition and used isotope values from the sampling year of 2019 for correction and thereby a shift of 2 ‰ since the pre-industrial:

$$\delta^{13}C_{corr.} = \delta^{13}C_{meas.} + \delta^{13}C_{Suess} \quad (10)$$

Where $\delta^{13}C_{corr.}$ represents the δ^{13}C composition of plant waxes corrected for the Suess-effect. $\delta^{13}C_{meas.}$ represents the analytical δ^{13}C value and $\delta^{13}C_{Suess}$ represents the shift in atmospheric δ^{13}C introduced by the Suess-effect.

Organic matter in soils represents a mixture of different carbon ages. In tropical soils the pre-aging of topsoil organic carbon is typically minor with turnover times of around 10 years. Hence, we used the atmospheric δ^{13}C value from 2009 of -8.2‰ for correction yielding a difference of 1.7‰ compared to the pre-industrial. When using *n*-alkane δ^{13}C values from the literature for comparison, we also corrected for the Suess-effect assuming an average turnover time of 10 years. For samples without reported collection date, sampling in the year prior to publication was assumed.

2.3.7. Canopy Cover

Tree cover data was derived from the 2015 version of the satellite derived tree cover model by Sexton et al. (2013). In the partially deforested areas of the southern Amazon, the tree cover model usually does not provide the sharp pasture/forest boundaries found in the field. Rather the transition is smoothed over several 30m × 30 meters grid cell. Locations in proximity to roads sometimes have the same effects. We adjusted the data to account for this effect by selecting values from the closest grid cell inside the forest proper and outside the artifactual transition zone in these cases.

3. Results

3.1. Bulk organic parameters

Soil TOC is highest under forest vegetation, with TOC values of up to 20% (Fig. 2a). Open savanna types and shrubland have generally lower TOC values of up to 5% (Fig. 2a). The Suess-effect corrected $\delta^{13}C$ compositions of bulk organic matter ($\delta^{13}C$ OC) in soil varies between $-29.0^{+0.6}_{-0.4}$‰ (Q^{75}_{25}) in the Amazon rainforest and $-15.5^{+0.8}_{-0.3}$‰ in Llanos savanna (Figs. 2a; S1). We thereby finnd the expected trend with the lowest values in multistoried rainforests and the highest values in vegetation-dominated by C_4 grasses. We finnd that Cerrado sensu stricto savanna has an intermediate value of $-25.3^{+1.8}_{-1.8}$‰, comparable to the mean value of $-25\ ^{+0.8}_{-0.6}$‰ in Caatinga shrubland (Fig. 2a). Forest sites have comparable $\delta^{13}C$ OC values even though canopy density varies between the dense canopies found in the Amazon and Atlantic rainforests and more open Cerradão dry forests (Fig. 2b).

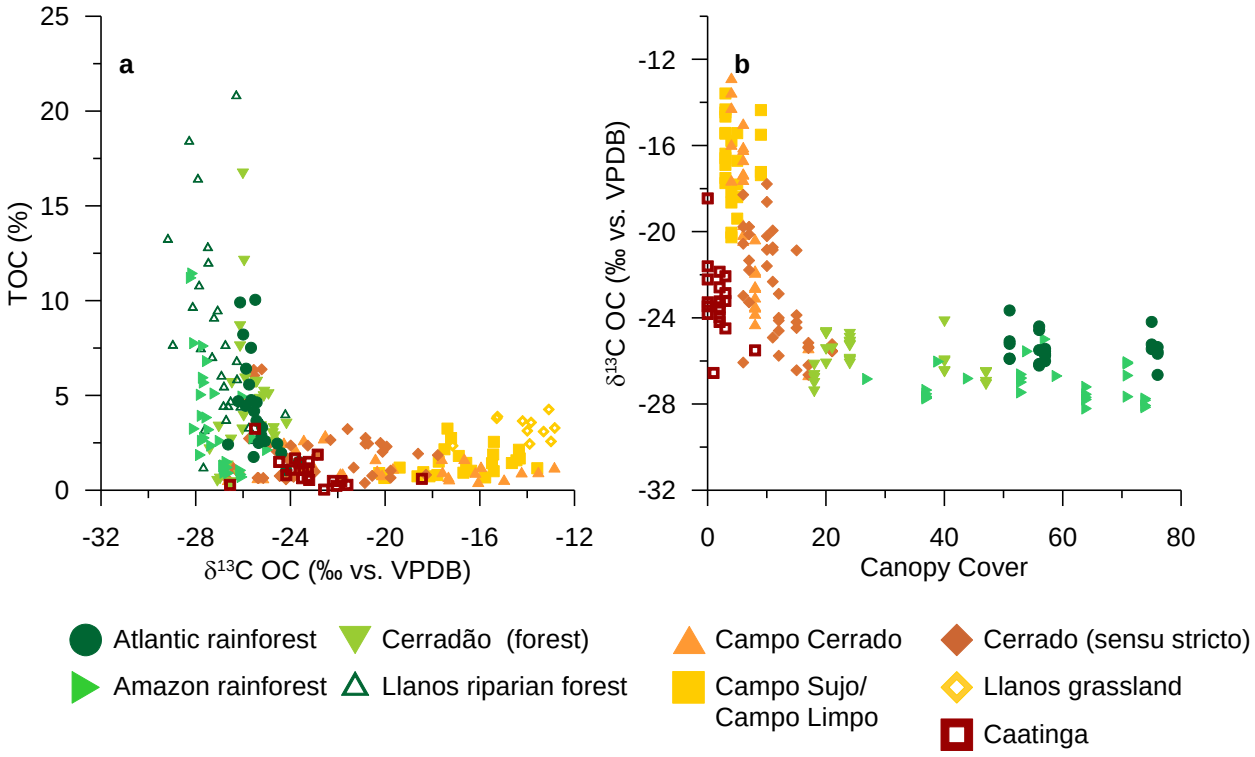


Figure 2. Organic parameters of surface soils from the major tropical South American vegetation types. a) Total organic carbon (TOC) vs. the stable carbon isotope composition of TOC ($\delta^{13}C$ OC) corrected for the Suess-effect. b) Canopy cover inferred from Sexton et al. (2013) vs. $\delta^{13}C$ OC. Since resolution of Sexton et al. (2013) did not allow the differentiation of closeby riparian forest and savanna sites of the Llanos transects, these samples are not plotted in b).

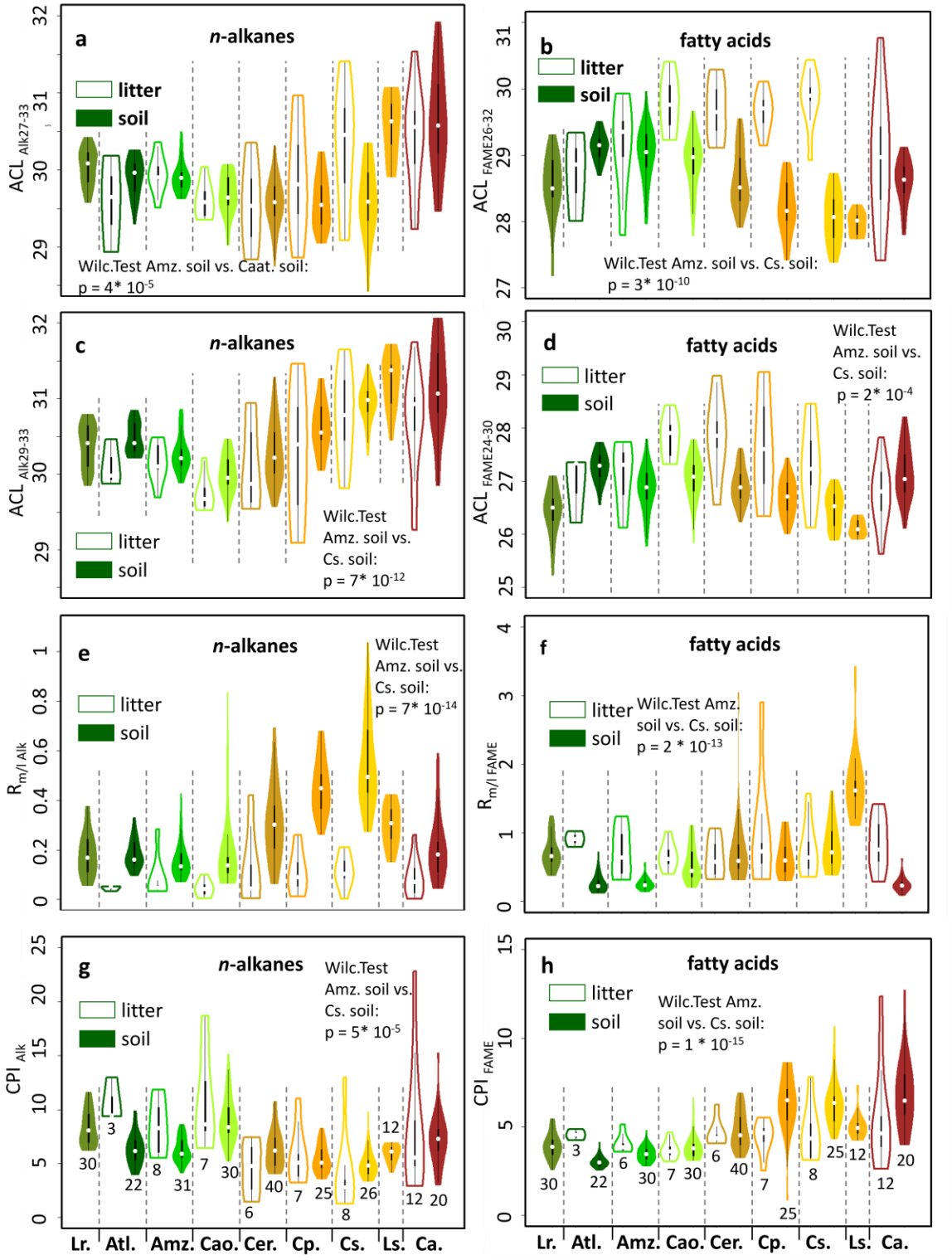
3.2. Plant wax distribution

For $ACL_{Alk27\text{-}33}$ we find consistent values in soil and litter samples for both forest and shrubland vegetation, while there are contrasting trends between soils and litter from the Cerrado vegetation types with higher $ACL_{Alk27\text{-}33}$ in litter than in soil (Fig. 3a). In the Llanos, savanna soils had higher $ACL_{Alk27\text{-}33}$ values than riparian forest (Fig. 3a). In contrast to $ACL_{Alk27\text{-}33}$, $ACL_{Alk29\text{-}33}$ shows a consistent trend towards higher values in both Cerrado soil and litter samples (Fig. 3c). For the $R_{m/l\ Alk}$, there are higher values for soil samples than in litter samples for all studied vegetation types (Figs. 3e, 6b). This discrepancy is again especially pronounced in the open Cerrado vegetation types. CPI_{Alk} values show lower values in savanna soils compared to the values observed for forests (e.g., in the Amazon rainforest $5.92\ ^{+1.23}_{-0.37}$ and in the Campo Sujo and Campo Limpo sites $4.81\ ^{+0.50}_{-0.73}$) (Fig. 3g). There is no consistent trend for CPI_{Alk} between litter and soil samples (Figs. 3g, 6c).

$ACL_{FAME26\text{-}32}$ and $ACL_{FAME24\text{-}30}$ values also show an increasing discrepancy between soil and litter samples for open Cerrado vegetation types with lower values for soil samples ($ACL_{FAME26\text{-}32}$: $29.9\ ^{+0.12}_{-0.08}$ in Campo Sujo/Campo Limpo litter vs. $28.07\ ^{+0.26}_{-0.31}$ in soils) (Fig. 3b, d). The $R_{m/l\ FAME}$ showed relatively consistent values for litter samples without a difference between open and closed vegetation types. In soil samples there is a clear trend towards higher values as the vegetation density decreases (e.g., $0.24^{+0.05}_{-0.04}$ in the Amazon compared to $0.72\ ^{+0.30}_{-0.16}$ in the Campo Sujo/Campo Limpo; Figs. 3f, 6h). Interestingly, the soil $R_{m/l\ FAME}$ values for forest and shrubland soils vegetation types were lower than for the litter samples of the same vegetation type, while in the open vegetation types featured similar values in soil and litter (Fig. 6h). For CPI_{FAME}, there are no distinct trends in the values of litter samples of the different vegetation types. For soil samples, there is a trend from lower values in soils from forest vegetation types (e.g., Amazon rainforest $3.48\ ^{+0.28}_{-0.29}$) to higher values in open vegetation types (e.g., Campo Sujo/Campo Limpo $6.35\ ^{+0.37}_{-1.01}$) and the Caatinga shrublands ($6.48\ ^{+1.46}_{-0.78}$; Fig. 3f).

Overall, our results show that plant wax distributions in forests are more stable between litter and soil samples, while there are greater discrepancies in open savanna vegetation types (Figs. 3, 6, S2, S3).

411



412

413 **Figure 3.** Relative distribution of plant wax long-chain *n*-alkanes and long-chain fatty acids in

414 soil and litter samples from the major tropical South American vegeteation types. a) Average

chain length (ACL) of the *n*-C_{27} to *n*-C_{33} long-chain *n*-alkanes. b) ACL of the C_{26} to C_{32} long-chain fatty acids. c) ACL of the *n*-C_{29} to *n*-C_{33} long-chain *n*-alkanes. d) ACL of the C_{24} to C_{30} long-chain fatty acids. e) $R_{m/l\ Alk}$. f) $R_{m/l\ FAME}$. g) Carbon preference index (CPI) of long-chain *n*-alkanes. h) CPI of long-chain fatty acids. The black lines in the violin plot represent box-whisker plots and the white point the median. The violins represent kernel density plots that are cut off at the data limits. The numbers below the violin plots indicate the number of samples represented by each plot. Abbreviations are as follows:. Lr: Llanos riparian forest. Atl: Atlantic rain forest. Amz: Amazon rainforest. Cao: Cerradão dry forest. Cer: Cerrado sensu stricto. Cp: Campo Cerrado. Cs: Campo Sujo/Campo Limpo. Ls: Llanos savanna. Ca: Caatinga shrublands. From the Llanos samples, only soil material was available. For complete chain length distributions of the litter and soil samples see Figs. S2 and S3.

3.3.Plant wax isotope composition

The $\delta^{13}C$ composition of all plant wax homologues shows a pattern that is similar to the one described for $\delta^{13}C$ OC (Fig. 2), with higher values for more open vegetation types and the lowest values in the Atlantic and Amazon rainforests as well as the Llanos riparian forests (Fig. 4a-d). Litter and soil samples also showed broadly consistent patterns (Fig. 4a-d, Fig. 6d,e,j,k). Aside from the broad general trend towards higher values in open vegetation types, there are some pronounced differences among compounds and homologues, especially *n*-alkanes of varying chain-lengths.

The $\delta^{13}C$ composition of *n*-C_{29} long-chain *n*-alkanes ($\delta^{13}C\ C_{29}$) from Cerrado sensu stricto soils shows, for instance, lower values ($-31.4\ ^{+1.3}_{-1.6}$ ‰) than the $\delta^{13}C$ composition of the *n*-C_{33} *n*-alkanes ($-28.4\ ^{+5.1}_{-2.5}$ ‰) (Fig. 4a). In forest formations, both homologues have comparable values (e.g., Amazon rainforest: $\delta^{13}C\ C_{29}$: $-35.0\ ^{+0.5}_{-0.4}$ ‰; $\delta^{13}C\ C_{33}$: $-35.9\ ^{+1.3}_{-0.5}$ ‰), while in Caatinga shrublands $\delta^{13}C\ C_{29}$ is even higher than $\delta^{13}C\ C_{33}$ ($\delta^{13}C\ C_{29}$: $-28.2\ ^{+1.2}_{-1.3}$ ‰; $\delta^{13}C\ C_{33}$: $-31.2\ ^{+2.4}_{-0.6}$ ‰) (Fig. 4a). This leads to a pattern, where the difference between $\delta^{13}C\ C_{33}$ and $\delta^{13}C\ C_{29}$ ($\delta^{13}C\ C_{33\text{-}29}$) has small or negative values in forests (e.g., Amazon rainforest $-0.9\ ^{+0.9}_{-0.9}$ ‰; Fig. 5a) and Caatinga shrublands ($-2.1\ ^{+1.2}_{-1.5}$ ‰), while open vegetation types have positive values (e.g., Cerrado sensu stricto ($3.2\ ^{+2.3}_{-1.9}$ ‰; Fig. 5a). A similar, albeit less pronounced pattern also arises for the difference between $\delta^{13}C\ C_{31}$ and $\delta^{13}C\ C_{29}$ ($\delta^{13}C\ C_{31\text{-}29}$; Fig. 5b). Leaf litter *n*-alkanes show the same pattern described above (Fig. 6d, e, f).

The $\delta^{13}C$ composition of different long-chain fatty acids homologues (e.g., $\delta^{13}C$ C_{24} and C_{30}) also shows the expected trends, with higher values in more open vegetation types (Fig. 4c, d). In contrast to long-chain *n*-alkanes, differences between different homologues ($\delta^{13}C$ C_{30} and $\delta^{13}C$ C_{24} ($\delta^{13}C$ C_{30-24}), as well as between $\delta^{13}C$ C_{28} and $\delta^{13}C$ C_{24} ($\delta^{13}C$ C_{28-24}) are not significantly distinct between rainforest and open vegetation types (Figs. 5c, d). There is a consistent offset in values between the fatty acids of different homologues with higher values in homologues of shorter chain length. This leads to a pattern where C_{24} compounds in soils have 2.1 $^{+0.8.}_{-0.9}$ ‰ higher values than C_{30} homologues (Fig. 5c). On average, the $\delta^{13}C$ composition of different long-chain fatty acids in litter samples is slightly lower than in soils (Figs. 6j, k, l).

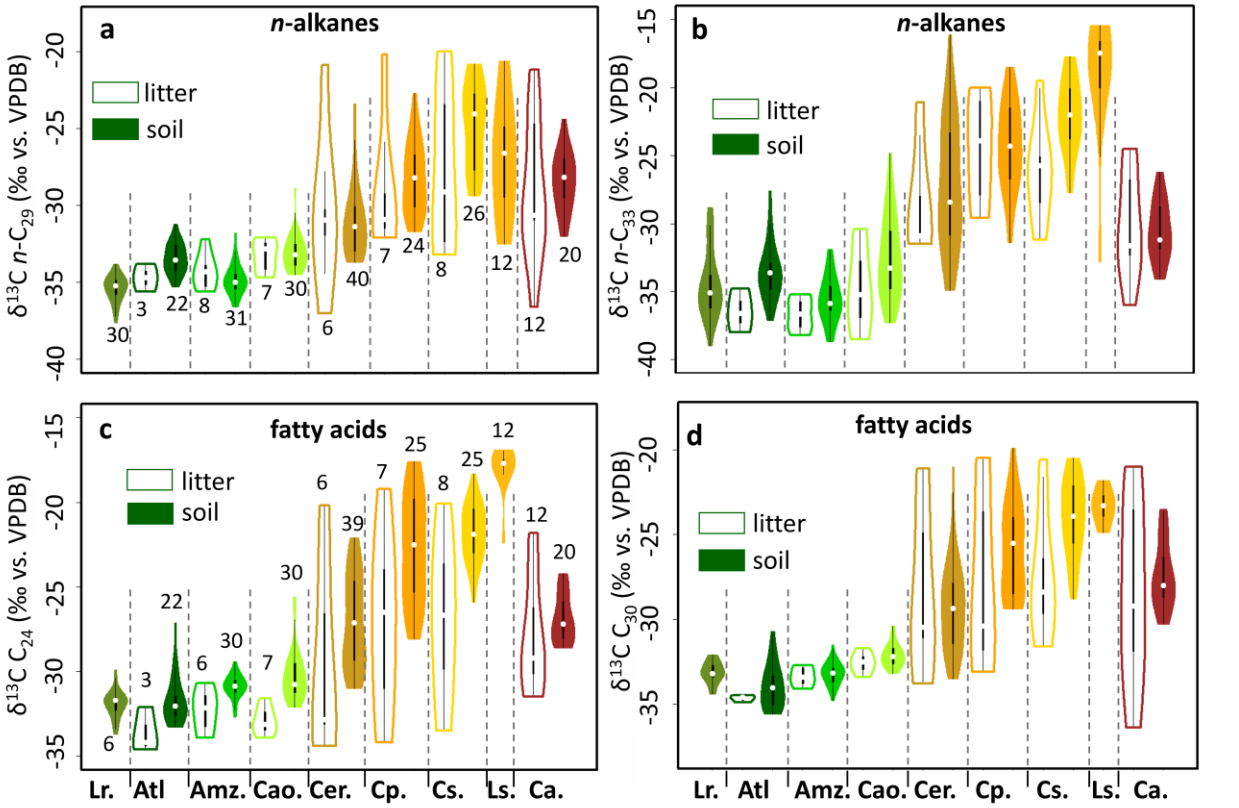


Figure 4. Stable carbon isotope composition ($\delta^{13}C$) of long-chain *n*-alkanes and long-chain fatty acids in soil and litter samples from the major tropical South American vegetation types. a) $\delta^{13}C$ composition of the C_{29} long-chain *n*-alkane. b) $\delta^{13}C$ composition of the C_{33} long-chain *n*-alkane. c) $\delta^{13}C$ composition of the C_{24} long-chain fatty acid d) $\delta^{13}C$ composition of the C_{30} long-chain fatty acid. The black lines in the violin plot represent box-whisker plots and the white point the median. The violins represent kernel density plots that are cut off at the data limits. The numbers adjacent to the violin plots indicate the number of samples represented by each plot. Panels a)

and b) have the same sample numbers as do panels c) and d). Abbreviations are as follows: Lr: Llanos riparian forest. Atl: Atlantic rain forest. Amz: Amazon rainforest. Cao: Cerradão dry forest. Cer: Cerrado sensu stricto. Cp: Campo Cerrado. Cs: Campo Sujo. Ls: Llanos savanna. Ca: Caatinga shrublands. From the Llanos samples, only soil material was available.

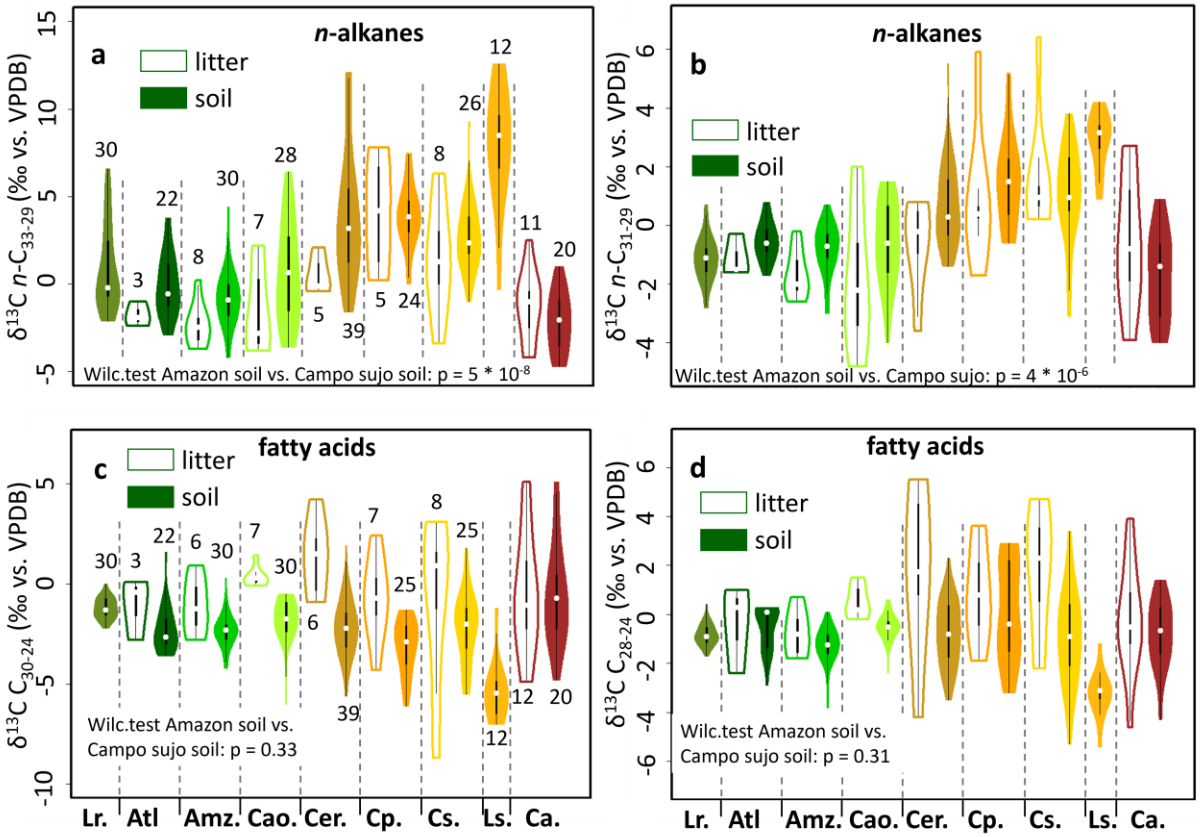


Figure 5. Variations in stable carbon isotope composition ($\delta^{13}C$) among long-chain *n*-alkane and long-chain fatty acid homologues in soil and litter samples from the major tropical South American vegetation types. a) Difference between the $\delta^{13}C$ composition of the *n*-C_{33} and *n*-C_{29} long-chain *n*-alkanes. b) Difference between the $\delta^{13}C$ composition of the *n*-C_{31} and *n*-C_{29} long-chain *n*-alkanes c) Difference between the $\delta^{13}C$ composition of the C_{30} and the C_{24} long-chain fatty acids. d) Difference between the $\delta^{13}C$ composition of the C_{28} and the C_{24} long-chain fatty acids. The black lines in the violin plot represent box-whisker plots and the white point the median. The violins represent kernel density plots that are cut off at the data limits. The numbers below or above the violin plots indicate the number of samples represented by each violin.

Abbreviations are as follows: Lr: Llanos riparian forest. Atl: Atlantic rain forest. Amz: Amazon rainforest. Cao: Cerradão dry forest. Cer: Cerrado sensu stricto. Cp: Campo Cerrado. Cs: Campo Sujo/Campo Limpo. Ls: Llanos savanna. Ca: Caatinga shrublands. From the Llanos samples, only soil material was available.

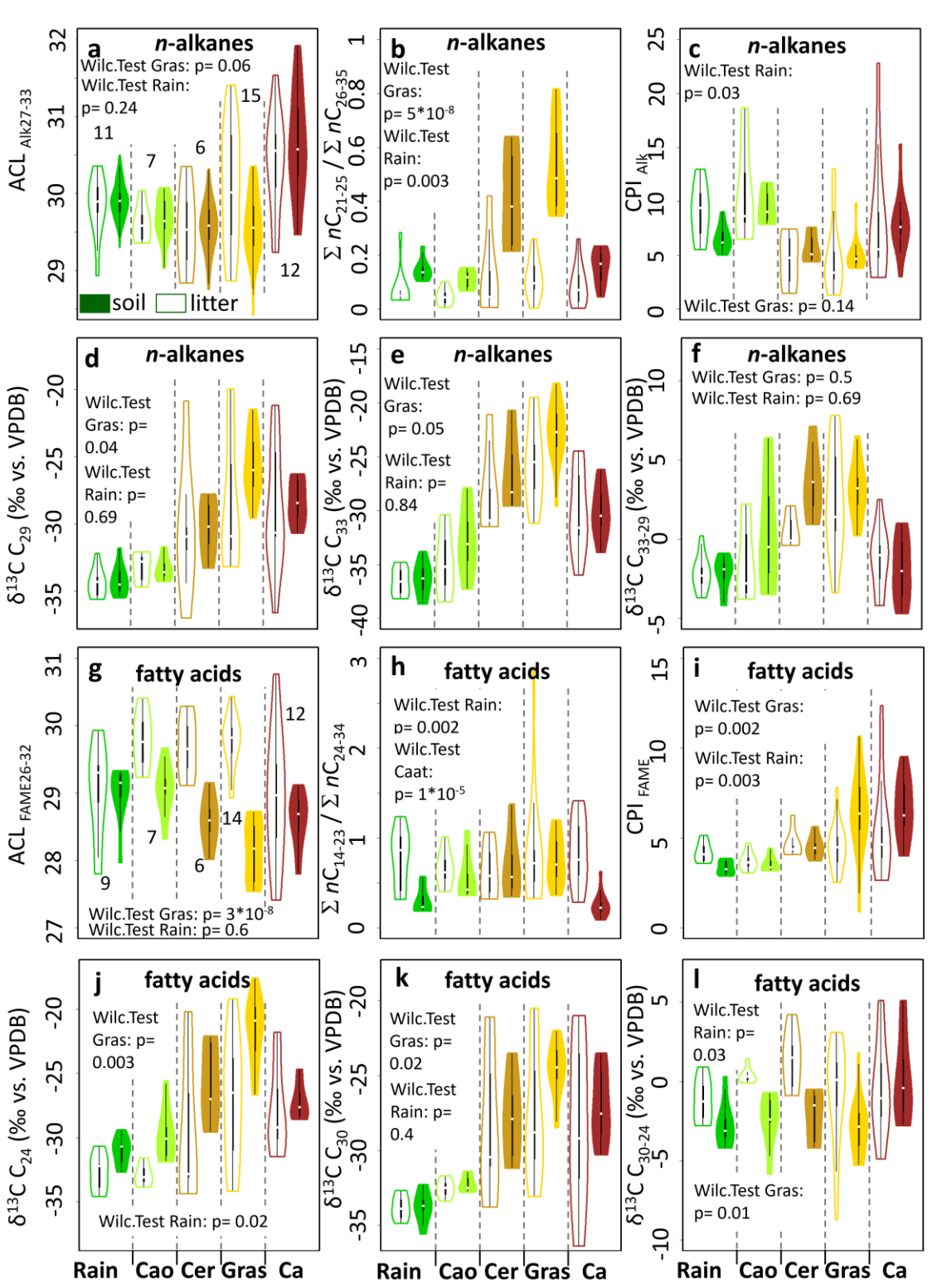
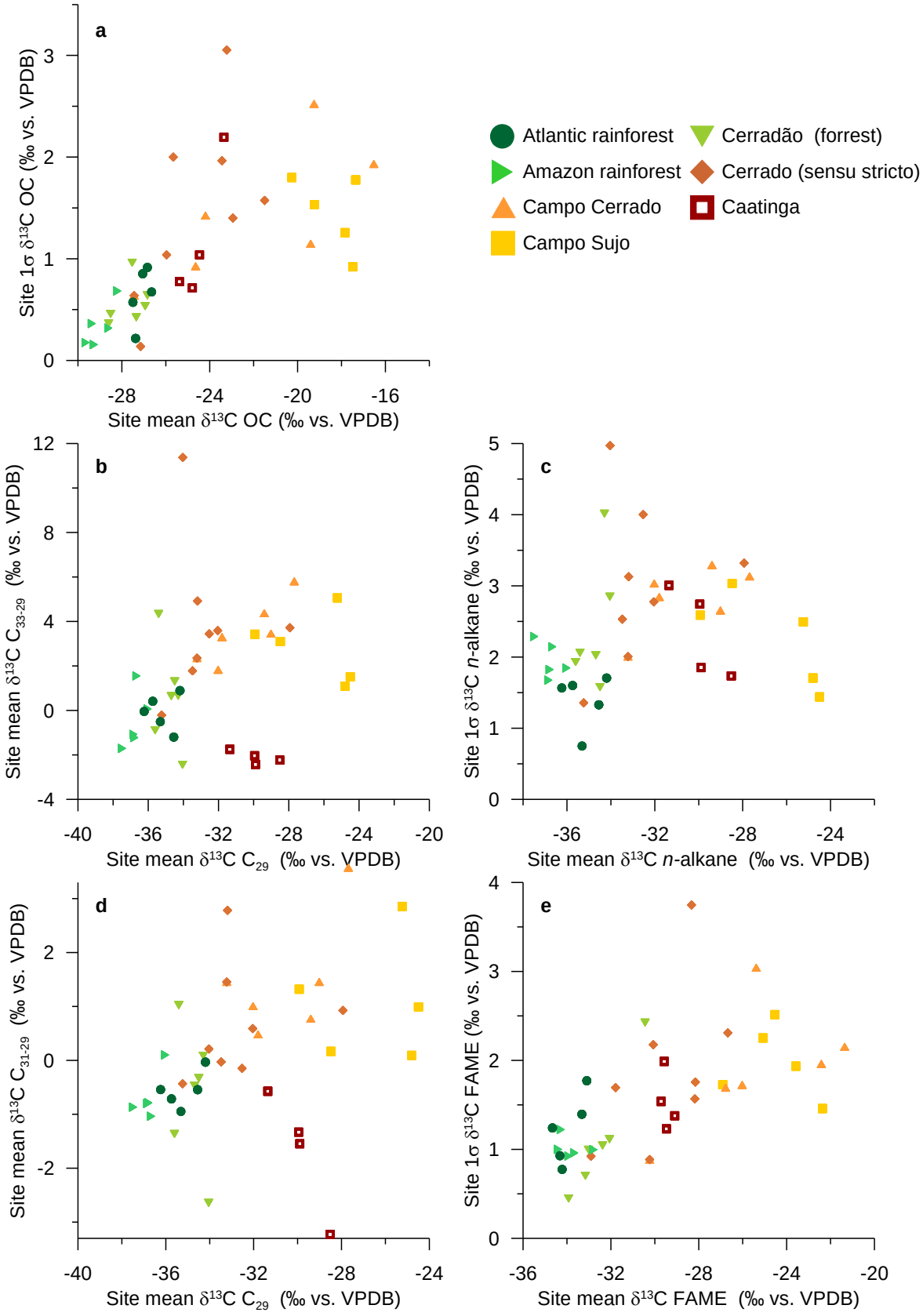


Figure 6. Relationship between litter and soil long-chain *n*-alkanes and long-chain fatty acid distributions and stable carbon isotope compositions ($\delta^{13}C$) in soil and litter samples from the

major tropical South American vegeation types. To allow for a direct comparison between litter and soil, the figure only features soil samples that also have corresponding litter samples. a) Average chain length ($ACL_{Alk27-33}$) of long-chain *n*-alkanes. b) $R_{m/l\ Alk}$. c) Carbon preference index (CPI) of long-chain *n*-alkanes. d) $\delta^{13}C$ composition of the C_{29} long-chain *n*-alkane. e) $\delta^{13}C$ composition of the C_{33} long-chain *n*-alkane f) Difference between the $\delta^{13}C$ composition of the *n*-C_{29} and *n*-C_{33} long-chain *n*-alkanes. g) ACL of long-chain fatty acids ($ACL_{FAME26-32}$). h) $R_{m/l\ FAME}$. i) CPI of long-chain fatty acids. j) $\delta^{13}C$ composition of the C_{24} long-chain fatty acid. k) $\delta^{13}C$ composition of the C_{30} long-chain fatty acid. l) Difference between the $\delta^{13}C$ composition of the C_{24} and the C_{30} long-chain fatty acids. The reported Wilcoxon tests were made between litter and soil samples of the indicated vegetation types. The black lines in the violin plot represent box-whisker plots and the white point the median. The violins represent kernel density plots that are cut off at the data limits. The numbers below the violin plots indicate the number of samples represented by each plot. Abbreviations are Rain: rainforests (Amazon & Atlantic), Cao: Cerradão dry forest. Cer: Cerrado sensu stricto. Gras: Open grass-dominated Cerrado physiognomies (Campo Sujo/Campo Limpo and Campo Cerrado). Ca: Caatinga shrublands. Due to the multiple comparisons conducted for each proxy, the corrected significance level using the Bonferroni correction would be 0.01, since 5 comparisons were conducted for each panel.

3.4. Local variability

The variability of $\delta^{13}C$ values among the five samples taken at each site shows matching trends for $\delta^{13}C$ OC and plant waxes (Fig. 7a, c, e). For the forests, we observe relatively stable values, while there is an increase in variability in the Cerradão dry forest and Caatinga shrubland sites (Fig. 7a, c, e). Variability is maximal for the mixed Cerrado sensu stricto and Campo Cerrado sites and slightly lower again for the Campo Sujo/Campo Limpo sites (Fig.7a, c, e). With the exception of Caatinga shrublands, this pattern is similar to the trends observed for $\delta^{13}C\ C_{33-29}$ and $\delta^{13}C\ C_{31-29}$ (Fig. 7b, d).


a
Site 1σ δ13C OC (‰ vs. VPDB)
Site mean δ13C OC (‰ vs. VPDB)
b
Site mean δ13C C33-29 (‰ vs. VPDB)
Site mean δ13C C29 (‰ vs. VPDB)
c
Site 1σ δ13C n-alkane (‰ vs. VPDB)
Site mean δ13C n-alkane (‰ vs. VPDB)
d
Site mean δ13C C31-29 (‰ vs. VPDB)
Site mean δ13C C29 (‰ vs. VPDB)
e
Site 1σ δ13C FAME (‰ vs. VPDB)
Site mean δ13C FAME (‰ vs. VPDB)
Atlantic rainforest
Amazon rainforest
Campo Cerrado
Campo Sujo
Cerradão (forrest)
Cerrado (sensu stricto)
Caatinga

512

Figure 7. Site specific variation in the stable carbon isotope composition ($\delta^{13}C$) of plant wax *n*-alkanes ($\delta^{13}C$ *n*-alkane), fatty acids ($\delta^{13}C$ fatty acid) and bulk organic carbon ($\delta^{13}C$ OC). a) Standard deviation (1 σ) of all $\delta^{13}C$ OC values at each site. b) Site average of $\delta^{13}C$ $C_{33\text{-}29}$. c) Standard deviation of the $\delta^{13}C$ values among all long-chain *n*-alkane (*n*-$C_{27,29,31,33}$) homologues at each site. d) Site average of $\delta^{13}C$ $C_{31\text{-}29}$. e) Standard deviation of the $\delta^{13}C$ values of all long-chain fatty acid ($C_{24,26,28,30}$) homologues at each site.

4. Discussion

4.1.Bulk organic parameters

The highest TOC contents in surface soil samples from tropical South America are found under forests, while savanna and shrubland samples yielded on average lower TOC (Fig. 2a). This finding is in line with global trends found in previous studies (Minasny et al., 2017; Stockmann et al., 2015). The $\delta^{13}C$ composition of soil organic carbon also shows expected patterns with the highest values in open savannas that hold the highest proportion of C_4 grass species (Campo Sujo and Campo Limpo) and the lowest values in tropical rainforests (Amazon rainforest; Fig. 2a). Cerrado sensu stricto show similar values to the semi-arid Caatinga shrublands, both being intermediate between forests and the most open savannas. The carbon isotope values measured in Caatinga soil samples are consistent with an extensive data set of $\delta^{13}C$ OC from leaves of Caatinga shrubs (Martinelli et al., 2020). In contrast to savannas, where $\delta^{13}C$ values are controlled by the relative input of C_3 and C_4 species, $\delta^{13}C$ values in Caatinga shrublands are also increased by the aridity effects on C_3 plants caused by drier conditions with MAP values around 500 mm y^{-1} found in northeastern Brazil (Karger et al., 2017; Martinelli et al., 2020). Under arid conditions, a decrease in leaf stomatal conductance is thought to lead to a decrease of the intercellular pCO_2 resulting in a decrease in fractionation during carbon fixation (Diefendorf et al., 2010; Kohn, 2010). In addition, CAM plants such as cacti and other succulents are abundant in Caatinga vegetation. While CAM plants can feature a wide range of $\delta^{13}C$ OC values, they tend to be intermediate between C3 and C4 plants (Messerschmid et al., 2021) and thereby also contribute to the observed $\delta^{13}C$ OC in the Caatinga $\delta^{13}C$ data set. The values observed for Caatinga thereby reflect the combination of aridity and photosynthesis effects.

4.2.Relative distribution of plant waxes

4.2.1. Microbial overprint

The relative distributions of plant waxes show substantial shifts from litter to soil samples as well as among different vegetation types (Figs. 3; 6a-c, g-i). The magnitude of the shifts between litter and soils tends to be larger for open savannas than for rainforests and the most significant shifts were observed for the $R_{m/l\ Alk}$ and $ACL_{FAME26-32}$ of the open Cerrado vegetation types (Figs. 3; 6a-c, g-i).

Fresh angiosperm plant material typically has low amounts of mid-chain *n*-alkanes (Bush & McInerney, 2013). The increase in mid-chain *n*-alkanes in soils is therefore likely caused by microbial overprint and can either be the result of the addition of compounds by fresh microbial synthesis (Nguyen Tu et al., 2011) or by the reworking of longer-chain compounds into shorter chain homologues by microbes (Brittingham et al., 2017). While this input is minor for rainforests, as previously reported from soils of the Amazon rainforest and Andean cloud forests (Wu et al., 2019), the encroachment of microbial compounds is more pronounced in soils from the savannas of the Cerrado and the Llanos Basin (Figs. 3c; 6b; S2a-i).

Shifts in the relative distribution of *n*-alkanes due to microbial overprint not only affect mid-chain compounds with a carbon chain-length ≤ 25 but also affect longer chain-length compounds. $ACL_{Alk27-33}$ shows lower values in soils from open Cerrado types than in soils from forest vegetation (Figs. 3a, 6a). This contrasts with the trend found in litter samples as well as with the findings from the Llanos savannas and other tropical savannas in Africa and Australia (Figs. 3a, 6a) (Krull et al., 2006; Rommerskirchen et al., 2006). For $ACL_{Alk29-33}$, which does not take the C_{27} *n*-alkane into account, soils from open Cerrado types show the expected trends towards higher values (Fig. 3c). This indicates that microbial overprint in Cerrado savanna samples can also lead to an increase in the relative contribution of the C_{27} *n*-alkane, while the relative distribution of the longer-chain *n*-alkanes remains relatively stable during incorporation into soils. Further evidence for the stability of the distribution of longer-chain compounds is also provided by the consistent CPI values between litter and soil samples of all studied vegetation types (Figs. 3g, 6c).

$ACL_{FAME26-32}$ and $ACL_{FAME24-30}$ from savannas vegetation types are also affected by microbial overprint and feature lower values in soils than in litter (Figs. 3b, d; 6g). This can be caused by the reworking to compounds of shorter chain-length, the fresh synthesis of shorter-chain homologues, or by the conversion of *n*-alkanes to fatty acids (Ji et al., 2013; Ofiti et al., 2021). In

contrast to *n*-alkanes, fatty acids with shorter chain-length between 16 and 23 carbon atoms can also be dominant in angiosperm leaf material (Almendros et al., 1996; Chikaraishi & Naraoka, 2006), leading to a more complex picture than for long-chain *n*-alkanes (Figs. 3d, 6h). Since shorter-chain length compounds are already present in leaves, microbial turnover during incorporation into soils can both lead to a lowering in $R_{m/l\ FAME}$ resulting from the preferential degradation of shorter-chain fatty acids and to an increase in $R_{m/l\ FAME}$ due to the fresh synthesis of shorter chain-length compounds (Chikaraishi & Naraoka, 2006). The lowering of $R_{m/l\ FAME}$ values in soils from forests and Caatinga shrubland therefore suggest that degradation is the dominant process there (Figs. 3d, 6g,h). For the other vegetation types the two processes seem to be balanced, leading to comparably stable relative contributions of mid-chain compounds (Figs. 3d, 6h). Another interesting aspect of the consequences of microbial overprint can be observed in CPI_{FAME}. While CPI_{FAME} of forests shows the expected trend towards lower values in soils compared to litter, there is a trend towards higher values for CPI_{FAME} in soils of open vegetation types (Fig. 6i). When the relative individual contributions of long-chain fatty acids are considered, it becomes apparent that the odd-chain compounds from savannas show similar contributions for both litter and soil samples (Fig. S3). Rather than a decrease in the relative contribution of odd-chain length compounds, the main cause for the increased CPI_{FAME} values is an enhanced contribution of C_{24}, C_{26} and C_{28} fatty acids (Fig. S3). Soils from open Cerrado vegetation types also show a decrease in C_{32} and C_{34} between soil and litter, which is however not part of the formulation of CPI_{FAME}. Hence, the increase in CPI_{FAME} in soils from savannas is likely the effect of the overall shift towards shorter even-chained homologues (Fig. S3).

Overall, our results consistently show a more pronounced microbial overprint in savannas than in forests. Multiple mechanisms may lead to this pattern. First, grasses produce generally lower amounts of long-chain *n*-alkanes compared to tropical trees leaving savannas more prone to microbial overprint. Second, the studied savannas have soils with relatively low pH values of $5\ ^{+0.3}_{-0.3}$. In acidic soils, *n*-alkanes have been reported to be less well preserved than fatty acids (Bull et al., 2000; Wu et al., 2019), which potentially further reduced *n*-alkane concentrations facilitating overprint. In addition, microbes using alkane hydroxylase can also convert long-chain *n*-alkanes to long-chain fatty acids (Ji et al., 2013), which may have contributed to the relative increase in C_{24}, C_{26} and C_{28} fatty acids in savanna soils (Fig. S3). Again, such mechanism would

606 lead to more pronounced effects in savanna vegetation, that synthesize smaller amounts of plant
607 waxes. Another factor that might additionally lead to the lower average chainlength of fatty acid
608 and *n*-alkanes is the high fire activity in the Cerrado savanna (Pivello, 2011), since soils from fire
609 affected areas tipically yield long-chain *n*-alkane and fatty acid compounds with lower chain
610 length (Eckmeier & Wiesenberg, 2009; Faria et al., 2015).
611
612 Our finding of increased mid-chain *n*-alkane contributions in savanna soils relative to the
613 respective litter samples is corroborated by concentrations of mid-to-long chain compounds in
614 the Serengeti savanna soils in Africa (Zhang et al., 2021). Microbial production of mid-chains in
615 savanna soils has implications for the use of the ratio of mid-to-long chain *n*-alkanes as proxy for
616 aquatic OC contributions in lacustrine sediments (Ficken et al., 2000). As savanna soils carry
617 $R_{m/l\ Alk}$ values similar to those found in lakes with heavy aquatic contributions, this ambiguity
618 has to be considered when interpreting alkane distributions in savanna-bordered aquatic settings.
619 Overprint also complicates interpretation of $ACL_{Alk27\text{-}33}$, which has been proposed as an indicator
620 of climate and vegetation change (Bush & McInerney, 2015; Chen et al., 2022; Rommerskirchen
621 et al., 2006). We therefore recommend using $ACL_{Alk29\text{-}33}$ to avoid the effects of microbial
622 overprint.
623

624 #### 4.2.2. Relation to vegetation and climate

625 Microbial overprint has no significant effect on the relative distribution of long-chain *n*-alkanes
626 with a chain-length ≥ 29. For these compounds, the $ACL_{Alk29\text{-}33}$ shows higher values in more
627 open savannas and in Caatinga compared to forests (Fig. 3c). The trend towards higher values in
628 tropical savannas has been previously observed in Africa and Australia (Krull et al., 2006; J. Liu
629 et al., 2022; Rommerskirchen et al., 2003; Rommerskirchen et al., 2006). Likewise, the higher
630 values for xeric shrublands have also been observed in African shrublands (Carr et al., 2014).
631 $ACL_{Alk29\text{-}33}$ from open Cerrado vegetation types show a larger variability in litter samples than
632 the ones from soils, demonstrating the wide variability that *n*-alkane distributions can have on a
633 species level, as well as the integrating nature of soils (Fig. 3c). For Caatinga shrublands, we also
634 find a large range of $ACL_{Alk29\text{-}33}$ in both litter and soils. This is in line with findings from shrubs
635 and succulents from xeric shrublands in Africa that feature a wide range of different *n*-alkane
636 distributions (Boom et al., 2014; Carr et al., 2014; Feakins & Sessions, 2010). Given that xeric

shrublands are often sparsely vegetated, the large differences introduced by individual plants may be locally preserved in Caatinga soils, whereas greater averaging is expected in Cerrado savanna soils.

On a biome-integrated scale, the *n*-alkane patterns in tropical South American and African savannas and shrublands show remarkable consistency despite the different evolutionary origin of the savannas on these continents (Edwards et al., 2010). This may be the result of convergent adaptation of leaf-wax *n*-alkane traits of the dominant grass and shrubland species to environmental factors (C.-K. Yang et al., 2018), or represent traits deeply rooted in plant phylogeny. Given that there are additional factors such as altitude and climate variables that affect *n*-alkane distributions (Bush & McInerney, 2015; Feakins et al., 2016) we emphasize that the chemotaxonomic significance of our sampling is limited to the study region. In contrast to previous regional studies (Bush & McInerney, 2015), we do not observe an impact of climatic variables on long-chain *n*-alkane distributions. For example, we find that xeric shrubland (MAP ~ 500 mm y^{-1}) has similar $ACL_{Alk29-33}$ values to Llanos savanna with MAP > 2000 mm y^{-1} (Fig. 3c). The absence of a relation to climate variables here is unsurprising, since the various relationships observed in regional studies, even though they may be statistically significant (Chen et al., 2022) have proven too weak and variable to be generalizable for paleoenvironmental research on a global level.

4.3.Plant wax isotope compositions as proxies for vegetation in tropical South America

4.3.1. Stability during incorporation into soil

The leaf litter and soil plant wax $\delta^{13}C$ data follows the expected trend with higher $\delta^{13}C$ values in more open vegetation (Fig. 4). Similar to the pattern observed for the relative distribution of long-chain *n*-alkanes and long-chain fatty acids, we also find that the variability of isotope values in leaf litter is greater than in the underlying soils. Again, the variability is especially pronounced for the savanna and shrubland vegetation types (Figs. 4, 5, 6d-f; j-l). For forests, the plant wax isotope data of long-chain *n*-alkanes and long-chain fatty acids is consistent among leaf litter and soil samples (Fig. 6d-f). For savannas, the large variability found in litter complicates the assessment of the effect of incorporation into soils; although Cerrado litter samples have on average lower values than soils, the difference is not significant due to the large variability found in both populations (Fig. 6d, e, j, k). While there is considerable variability for separate

668 homologues, the consistency in $\delta^{13}C$ C_{33-29} values between soil and litter suggests that the isotope
669 values in savanna and shrubland areas are relatively stable during soil incorporation as well (Fig.
670 6f). The consistency in isotope values found for the forest samples somewhat contrast previous
671 work that found shifts in $\delta^{13}C$ compositions during incorporation of plant waxes into soils
672 (Chikaraishi & Naraoka, 2006; Tu et al., 2004; Wu et al., 2019; Y. Zhang et al., 2017). Evidence
673 from litter bag experiments studying the evolution of plant wax $\delta^{13}C$ compositions during
674 degradation also paint an inconsistent picture, with some reporting effects on the isotope
675 composition, while others do not (Huang et al., 1997; Nguyen Tu et al., 2011). This indicates that
676 the impact of degradation on plant wax $\delta^{13}C$ might have varying outcomes under different
677 conditions. These might include varying vegetation types with as well as different climate
678 conditions or fire activity (Sarangi et al., 2022). We also note that our study focused on the upper
679 5 cm of the sampled soils, and deeper soil horizons might feature the enrichment observed
680 elsewhere.
681

682 **4.3.2. Differentiating South American vegetation types using plant wax isotopes**

683 While C_3 plants can cover a wide range of isotope compositions our dataset shows that both the
684 litter and soil samples from the different forest types of the studied area cover a relatively narrow
685 range of values (Fig. 4 a-d). For instance, the Suess-effect corrected $\delta^{13}C$ C_{29} of the combined
686 Atlantic and Amazon rainforests, the Cerradão dry forest and Llanos riparian forests a show a
687 range of $-34.4\ ^{+1.1}_{-0.7}$ ‰ (Fig. 4a). Values above that range can indicate savanna or shrubland
688 intrusions, even though there are C_3 species that might have higher $\delta^{13}C$. Indeed, the interquartile
689 range of $\delta^{13}C$ C_{29} values found in soils from our Cerrado sensu stricto sample set (i.e.,
690 $-31.4\ ^{+1.3}_{-1.6}$; Fig. 4a) is well within the range covered by some C_3 plants (J. Liu & An, 2020).
691 Although $\delta^{13}C$ C_{29} values in Cerrado sensu stricto can still be in the range of forest samples, $\delta^{13}C$
692 C_{33} values are more sensitive to contributions by grass vegetation. The cause for this difference
693 lays in the different production of homologues by grass and tree taxa. Tree-dominated vegetation
694 types have long-chain *n*-alkane distributions dominated by *n*-C_{29} and *n*-C_{31} homologues with low
695 proportions of *n*-C_{33} (Figs. S2). In contrast, grass-dominated vegetation types have more uniform
696 homologue distribution and comparable contributions of *n*-C_{29}, *n*-C_{31} and *n*-C_{33} (Fig. S2).
697 Therefore, the addition of grasses with a higher C_{33} homologue contributions leads to a more
698 pronounced impact on $\delta^{13}C$ C_{33} values than on $\delta^{13}C$ C_{29} values (Figs. 4b, S4). Indeed, the

Cerradão dry forests, which can be in transition into Cerrado sensu stricto, feature the overall highest relative contribution of *n*-C_{29}, leading to a strong impact of *n*-C_{33} provided by minor grass contributions (Fig. S2). Likewise, the addition of tree derived *n*-alkanes to a grass-dominated environment leads to a more pronounced response in $\delta^{13}C$ C_{29} than in $\delta^{13}C$ C_{33} (Fig. 4a, b). Our observation that $\delta^{13}C$ of different long-chain *n*-alkane homologues show a nuanced response to shifts in vegetation has also been made in savanna areas from Africa, D. (Garcin et al., 2014; Schwab et al., 2015; D. Zhang et al., 2021). While the $\delta^{13}C$ of different long-chain *n*-alkane homologues yield distinct sensitivity to tree or grass input, the $\delta^{13}C$ of long-chain fatty acids show more consistent trends among different homologues (Fig. 4c, d). Fatty acid $\delta^{13}C$ compositions are also more similar to $\delta^{13}C$ OC than long-chain *n*-alkanes (Figs. S4, S5). Since fatty acids have more homogeneous chain-length distributions among different vegetation types, the mixing between tree and grass genera does not lead to shifts that are focused on specific homologues but rather affect all reported compounds.

Our data shows that the $\delta^{13}C$ composition of both long-chain *n*-alkanes and long-chain fatty acids from Cerrado sensu stricto is similar to semi-arid Caatinga shrublands (Fig. 4a-d). Hence, the $\delta^{13}C$ of single plant wax homologues or pooled data of multiple homologues cannot be used to differentiate between shrubland and savanna vegetation types found under distinctly different climate conditions, complicating interpretations of plant-wax records (Boom et al., 2014). As discussed for soil $\delta^{13}C$ OC, the enrichment in Caatinga vegetation is partly the result of drier conditions and CAM photosynthesis (da Silva & Lacher, 2020; J. Liu & An, 2020; Martinelli et al., 2020). Our results therefore indicate that the interpretation of plant wax $\delta^{13}C$ compositions as recorders of relative contributions of C_3 and C_4 plants on a forest-savanna continuum is incomplete, as shrublands may produce similar values to savannas (Goñi et al., 1997; Huang et al., 2000; Schefuß et al., 2005). As xeric shrublands cover substantial tropical areas (Olson et al., 2001), this has major implications for the interpretation of the $\delta^{13}C$ of plant waxes as paleoenvironmental proxies and calibration studies treating $\delta^{13}C$ values in sedimentary archives as binary mixtures between C_3 and C_4 endmembers (Magill et al., 2013; D. Yang & Bowen, 2022).

4.4. Multi-homologue isotope assessment of tropical vegetation structure

4.4.1. Tropical South America

730 While single homologue or bulk organic matter analysis cannot differentiate between Cerrado
731 semsu stricto and semi-arid Caatinga shrublands, our data show substantial differences between
732 the two vegetation types for $\delta^{13}C$ C_{33-29} or $\delta^{13}C$ C_{31-29} (Fig. 5, a, b). Here, samples from Cerrado
733 savanna have substantially higher values than from Caatinga shrubland samples (Fig. 5a, b).
734 Therefore, our results indicate that the $\delta^{13}C$ composition of different long-chain *n*-alkane
735 homologues can be used to differentiate between xeric shrubland and savanna vegetation types
736 (Fig. 5a, b). As discussed above, grass and tree genera in savannas produce different relative
737 amounts of the *n*-C_{29} and *n*-C_{33}. This leads to overall higher values for the grass-dominated $\delta^{13}C$
738 C_{33} and higher $\delta^{13}C$ C_{33-29} values in savanna samples (Fig. 5a). The lower $\delta^{13}C$ C_{33-29} values in
739 Caatinga samples can be explained along similar lines. First, arid shrub and succulent species
740 synthesize *n*-alkanes with highly variable distributions, but overall tend to have longer chain-
741 length than C_3 trees (Boom et al., 2014; Carr et al., 2014; Feakins & Sessions, 2010). Thereby,
742 the mixing of material from different grass, shrub and succulent sources is not affecting specific
743 homologues as in savannas (Fig. 5a, b). Second, the differences in $\delta^{13}C$ composition of leaf-
744 waxes from different plant taxa are likely also less pronounced than between grasses and trees
745 found in Cerrado savanna. C_3 shrubs have elevated $\delta^{13}C$ composition due to aridity (Boom et al.,
746 2014) and succulents using CAM carbon fixation also feature $\delta^{13}C$ values that are between C_4
747 grass and C_3 tree endmembers (Messerschmid et al., 2021). The mixing of different Caatinga
748 sources therefore does not lead to the large differences between $\delta^{13}C$ C_{33} and $\delta^{13}C$ C_{29} observed
749 for the savanna settings, but to more uniform values (Fig. 5a).
750 While long-chain *n*-alkanes feature marked differences in $\delta^{13}C$ C_{33-29} or $\delta^{13}C$ C_{31-29} among
751 different vegetation types, long-chain fatty acids have a relatively homogeneous isotope
752 distribution, and the difference in $\delta^{13}C$ composition is consistent among the studied vegetation
753 types (Fig. 5c, d). An exception to this feature is the Llanos savanna samples, which show the
754 overall lowest $\delta^{13}C$ C_{30-24} and $\delta^{13}C$ C_{28-24} values. Since this is only observed in a very local set of
755 samples, the usefulness of this pattern would need further testing in larger sample sets.
756

757 **4.4.2. Comparison to tropical vegetation of other continents**

758 To expand the applicability of our findings on the differentiation of savannas and xeric
759 shrublands, we also analyzed previously published $\delta^{13}C$ C_{33-29} data from Africa and tropical
760 South and Central America. As studies often report single homologue or pooled values, there is

761 only limited data on $\delta^{13}C$ C_{33} from soils available, especially from xeric shrublands (Fig. 8a).
762 Given this dearth of $\delta^{13}C$ C_{33} data, we also compiled $\delta^{13}C$ $C_{31\text{-}29}$ data which showed a similar,
763 though muted pattern than $\delta^{13}C$ $C_{33\text{-}29}$ in our data set (Fig. 5d).
764 In South and Central America, multi-homologue long-chain *n*-alkane data has been published on
765 soils from mountainous areas in Central America that covered tropical rainforests, dry forests
766 and tropical coniferous forests (Douglas et al., 2012), while another study exclusively covered
767 soils from the Amazon rainforest (Wu et al., 2019). In Africa, Schwab et al. (2015) provided *n*-
768 alkane $\delta^{13}C$ along a transect in Cameroon that covered forests, savanna and xeric shrubland
769 areas, while D. Zhang et al. (2021) covered savannas from the eastern African Serengeti and
770 Magill et al. (2019) also reported data from eastern Africa and South Africa. There has been an
771 extensive study on the major southwest African vegetation types, including the Nama Karoo and
772 Succulent Karoo xeric shrublands (Herrmann et al., 2016), which only reported $\delta^{13}C$ C_{29} and
773 $\delta^{13}C$ C_{31} values.
774 The $\delta^{13}C$ $C_{33\text{-}29}$ trends in the forest-savanna continuum show lower values for forest and open
775 grass-dominated endmembers, while mixed savannas show higher values (Fig. 8a). The
776 magnitude of the increase in $\delta^{13}C$ $C_{33\text{-}29}$ is comparable for African and South American savannas
777 (Fig. 8a). The shrublands also show consistently lower $\delta^{13}C$ $C_{33\text{-}29}$ values in both South America
778 and Africa (Fig. 8a). $\delta^{13}C$ $C_{31\text{-}29}$ trends also show comparable values for Africa and South
779 America and the lowest values in shrubland samples (Fig. 8b). There is however substantial
780 overlap between savanna and shrubland samples that limits the usefulness of $\delta^{13}C$ $C_{31\text{-}29}$
781 compared $\delta^{13}C$ $C_{33\text{-}29}$ in differentiating between these vegetation types (Fig. 8b).
782 Our compilation shows that $\delta^{13}C$ $C_{33\text{-}29}$ can be used to differentiate xeric shrubland and savannas
783 in the major tropical and subtropical areas of Africa and America, despite the different
784 evolutionary history and vegetation structure (Simon et al., 2009). The similarity of the patterns
785 between African and American savannas suggests that the savanna structure and species
786 diversity are not important for $\delta^{13}C$ $C_{33\text{-}29}$ as long as there is a mixture of C_3 tree species with
787 their strongly dominant *n*-C_{29} production with C_4 grasses yielding modestly higher proportions in
788 *n*-C_{33}.
789
790 The isotopic offset between chain lengths are preserved through incorporation into soils in this
791 South American study (Fig. 6f) and there is also evidence from Africa that the isotope

distribution among homologues is retained through aerial and fluvial transport to marine deposition areas (Magill et al., 2019). Hence, there is promise for broad application of $\delta^{13}C$ C_{33-29} as a conservative tracer to differentiate between savannas and xeric shrublands in tropical soils and sediments of lakes and marine settings.

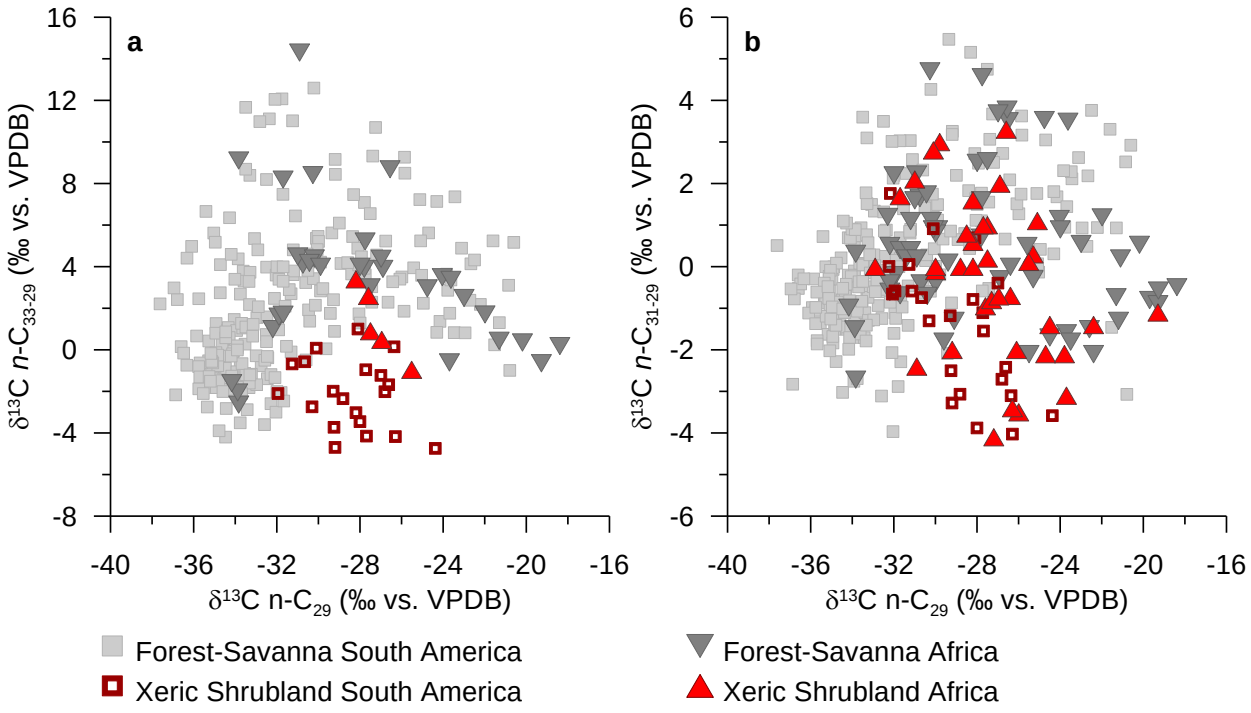


Figure 8. Stable carbon isotope composition ($\delta^{13}C$) variability among different *n*-alkane homologues in tropical vegetation types. a) $\delta^{13}C$ C_{33-29} in tropical vegetation types of Central- and South America (Douglas et al., 2012; Wu et al., 2019), as well as Africa (Magill et al., 2019; Schwab et al., 2015; D. Zhang et al., 2021). b) $\delta^{13}C$ C_{31-29} in tropical vegetation types of Central- and South America (Douglas et al., 2012; Wu et al., 2019), as well as Africa (Herrmann et al., 2016; Magill et al., 2019; Schwab et al., 2015; D. Zhang et al., 2021). All data were corrected for the Suess-effect.

4.5. Local variability

To study the variability of plant wax and $\delta^{13}C$ OC values at each site, we analyzed multiple (usually 5) samples per site. For mixed Cerrado types, we collected samples in the vicinity of trees and shrubs as well as in grass-dominated locations (Fig. 1c).

For $\delta^{13}C$ OC and the $\delta^{13}C$ composition of the different studied plant-wax homologues, we find variability is greatest for mixed savanna sites, while forests show the lowest variability. The most open vegetation types show intermediate values (Fig. 7a, c, e). A similar pattern is found for $\delta^{13}C$ $C_{33\text{-}29}$ both on a sample specific level and averaged over each site (Fig. 7b, 8a). The main contrast between the $\delta^{13}C$ $C_{33\text{-}29}$ and the pattern of bulk $\delta^{13}C$ and plant wax $\delta^{13}C$ variability is found in Caatinga shrublands. While $\delta^{13}C$ $C_{33\text{-}29}$ values are close to zero, $\delta^{13}C$ OC variability is similarly elevated in Caatinga shrublands and Cerrado savannas. While the causes for the discrepancy in $\delta^{13}C$ $C_{33\text{-}29}$ between Caatinga and Cerrado are already discussed above, the lack of such a pattern in the variability of $\delta^{13}C$ OC and site-specific *n*-alkane and fatty acid $\delta^{13}C$ likey arises from the considerable variety of $\delta^{13}C$ values found in different shrubland plants (Boom et al., 2014; Messerschmid et al., 2021).

5. Conclusions

We studied the $\delta^{13}C$ composition of plant-wax long-chain *n*-alkanes and long-chain fatty acids in soils and litter from the major biomes of tropical South America, with the objective to further develop understanding of these proxies for vegetation reconstructions. Our results show that the relative distribution of plant wax compounds (with the exception of long-chain *n*-alkanes $\geq$ *n*-C_{29}) is detectably altered by microbial overturning in soils under open vegetation types. This indicates that the relative distribution of plant wax abundances is not a suitable tool for reconstructions of past vegetation in tropical South America. Conversely, isotope values remain relatively stable during incorporation into soils. Tropical rainforest soils show $\delta^{13}C$ OC and plant wax $\delta^{13}C$ compositions are within a narrow range, allowing the detection of even minor occurrences of open vegetation. The most open vegetation types have the highest $\delta^{13}C$ values, faithfully recording the increased abundance of grasses using the C_4 metabolism. The mixed Cerrado savannas and the Caatinga shrublands show comparable $\delta^{13}C$ values, even though they are distinctly different vegetation types growing under different climates. To differentiate between the Cerrado savannas and the xeric Caatinga shrublands, we use the $\delta^{13}C$ composition of long-chain *n*-alkane homologues of different chain-lengths. While savanna vegetation types typically feature lower $\delta^{13}C$ values for the *n*-C_{29} homologue than for the *n*-C_{33} homologue, Caatinga shrublands show consistent values for all homologues. We further compare our results with published data from African savannas and xeric shrublands. We find that the $\delta^{13}C$

composition among different *n*-alkane homologues in Africa and tropical South America show the same trends in the forest-savanna continuum and also in xeric shrublands. Likewise, the relative distribution of plant waxes among different vegetation types is also consistent between Africa and tropical South America. Both observations indicate a remarkable consistency despite the different vegetation structure and evolutionary history of the savanna biomes of Africa and South America. Overall, our study of tropical South American plant wax carbon isotopes demonstrates that multiple plant wax homologues can be used to gain more detailed information on past vegetation than obtained from single homologue analysis or the averaging of multiple homologues into a single value.

Acknowledgments

This project was funded by a Swiss National Science Foundation (SNF) mobility fellowship (grant P400P2_183856) to CH, supporting fieldwork and the postdoctoral researcher. Laboratory analyses were supported by funding from University of Southern California and the Women in Science and Engineering Program to SF. We acknowledge undergraduate laboratory assistants at USC: Betelhem Assefa, Jonnie Dolan, Lindsay Luchinsky, Dea Kurti, Christopher Rincon and Sharon Tu. Soils were imported under USDA Permit P330-19-00164 to SF. Fieldwork in the eastern Amazon was funded by FAPESP (grant 2016/02656-9). DJB was financially supported by the São Paulo Research Foundation (FAPESP) (grants 2019/24977-0 and 2022/06440-1). CMC acknowledges the financial support from FAPESP (grants 2018/15123-4 and 2019/24349-9), CNPq (grant 312458/2020-7), and the Alfred Wegener Institute for Polar and Marine Research. AOS thanks the financial support from FAPESP (grant 2018/23899-2) and CNPq (grant 307179/2021-4). TKA acknowledges the financial support from FAPESP (grants 2019/19948-0 and 2021/13129-8). VRM thanks the financial support from FAPESP (grant 2022/02957-0) and Serrapilheira (Serra R-2012-38252)

Open Research

The plant wax data presented in the study is available at zenodo.org via https://zenodo.org/doi/10.5281/zenodo.10214559 (Häggi, Bertassoli, et al., 2023).

References

Adler, D., & Kelly, S. T. (2019). vioplot: violin plot. R package version 0.3.4 https://github.com/TomKellyGenetics/vioplot.

Almendros, G., Sanz, J., & Velasco, F. (1996). Signatures of lipid assemblages in soils under continental Mediterranean forests. *European Journal of Soil Science*, *47*(2), 183-196. doi:10.1111/j.1365-2389.1996.tb01389.x

Bertassoli, D. J., Häggi, C., Chiessi, C. M., Schefuß, E., Hefter, J., Akabane, T. K., & Sawakuchi, A. O. (2022). Controls on the distributions of GDGTs and n-alkane isotopic compositions in sediments of the Amazon River Basin. *Chemical Geology*, *594*, 120777. doi:10.1016/j.chemgeo.2022.120777

Bertassoli, D. J., Sawakuchi, A. O., Chiessi, C. M., Schefuß, E., Hartmann, G. A., Häggi, C., et al. (2019). Spatiotemporal Variations of Riverine Discharge Within the Amazon Basin During the Late Holocene Coincide With Extratropical Temperature Anomalies. *Geophysical Research Letters*, *46*(15), 9013-9022. doi:10.1029/2019GL082936

Blydenstein, J. (1967). Tropical Savanna Vegetation of the Llanos of Colombia. *Ecology*, *48*(1), 1-15. doi:10.2307/1933412

Boom, A., Carr, A. S., Chase, B. M., Grimes, H. L., & Meadows, M. E. (2014). Leaf wax n-alkanes and δ13C values of CAM plants from arid southwest Africa. *Organic Geochemistry*, *67*, 99-102. doi:10.1016/j.orggeochem.2013.12.005

Brienen, R. J. W., Phillips, O. L., Feldpausch, T. R., Gloor, E., Baker, T. R., Lloyd, J., et al. (2015). Long-term decline of the Amazon carbon sink. *Nature*, *519*(7543), 344-348. Letter. doi:10.1038/nature14283

Brittingham, A., Hren, M. T., & Hartman, G. (2017). Microbial alteration of the hydrogen and carbon isotopic composition of n-alkanes in sediments. *Organic Geochemistry*, *107*, 1-8. doi:10.1016/j.orggeochem.2017.01.010

Bull, I. D., Bergen, P. F. v., Nott, C. J., Poulton, P. R., & Evershed, R. P. (2000). Organic geochemical studies of soils from the Rothamsted classical experiments—V. The fate of lipids in different long-term experiments. *Organic Geochemistry*, *31*(5), 389-408. doi:10.1016/S0146-6380(00)00008-5

Bush, R. T., & McInerney, F. A. (2013). Leaf wax n-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochimica Et Cosmochimica Acta*, *117*, 161-179. doi:10.1016/j.gca.2013.04.016

Bush, R. T., & McInerney, F. A. (2015). Influence of temperature and C4 abundance on n-alkane chain length distributions across the central USA. *Organic Geochemistry*, *79*, 65-73. doi:10.1016/j.orggeochem.2014.12.003

Carr, A. S., Boom, A., Grimes, H. L., Chase, B. M., Meadows, M. E., & Harris, A. (2014). Leaf wax n-alkane distributions in arid zone South African flora: Environmental controls, chemotaxonomy and palaeoecological implications. *Organic Geochemistry*, *67*, 72-84. doi:10.1016/j.orggeochem.2013.12.004

Ceccopieri, M., Scofield, A. L., Almeida, L., Araújo, M. P., Hamacher, C., Farias, C. O., et al. (2021). Carbon isotopic composition of leaf wax n-alkanes in mangrove plants along a latitudinal gradient in Brazil. *Organic Geochemistry*, *161*, 104299. doi:10.1016/j.orggeochem.2021.104299

Cerling, T. E., Harris, J. M., MacFadden, B. J., Leakey, M. G., Quade, J., Eisenmann, V., & Ehleringer, J. R. (1997). Global vegetation change through the Miocene/Pliocene boundary. *Nature*, *389*, 153. doi:10.1038/38229

Cerling, T. E., Wynn, J. G., Andanje, S. A., Bird, M. I., Korir, D. K., Levin, N. E., et al. (2011). Woody cover and hominin environments in the past 6 million years. *Nature*, *476*, 51. doi:10.1038/nature10306

Chen, G., Li, X., Tang, X., Qin, W., Liu, H., Zech, M., & Auerswald, K. (2022). Variability in pattern and hydrogen isotope composition (δ2H) of long-chain n-alkanes of surface soils and its relations to climate and vegetation characteristics: A meta-analysis. *Pedosphere*, *32*(3), 369-380. doi:10.1016/S1002-0160(21)60080-2

Chikaraishi, Y., & Naraoka, H. (2006). Carbon and hydrogen isotope variation of plant biomarkers in a plant–soil system. *Chemical Geology*, *231*(3), 190-202. doi:10.1016/j.chemgeo.2006.01.026

Collister, J. W., Rieley, G., Stern, B., Eglinton, G., & Fry, B. (1994). Compound-Specific Delta C-13 Analyses of leaf lipids from plants with differing carbon-dioxide metabolisms. *Org. Geochem.*, *21*(6-7), 619-627. doi:10.1016/0146-6380(94)90008-6

Cox, P. M., Pearson, D., Booth, B. B., Friedlingstein, P., Huntingford, C., Jones, C. D., & Luke, C. M. (2013). Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature, 494*(7437), 341-344. doi:10.1038/nature11882

Cranwell, P. A. (1981). Diagenesis of free and bound lipids in terrestrial detritus deposited in a lacustrine sediment. *Org. Geochem., 3*, 79-89. doi:10.1016/0146-6380(81)90002-4

da Silva, J. M. C., & Bates, J. M. (2002). Biogeographic patterns and conservation in the South American Cerrado: A tropical Savanna hotspot. *Bioscience, 52*(3), 225-233. doi:10.1641/0006-3568(2002)052[0225:BPACIT]2.0.CO;2

da Silva, J. M. C., & Lacher, T. E. (2020). Caatinga—South America. In M. I. Goldstein & D. A. DellaSala (Eds.), *Encyclopedia of the World's Biomes* (pp. 554-561). Oxford: Elsevier.

Dantas, V. D., & Pausas, J. G. (2013). The lanky and the corky: fire-escape strategies in savanna woody species. *Journal of Ecology, 101*(5), 1265-1272. doi:10.1111/1365-2745.12118

De Azevedo, A. (1950). Regiões climato-botânicas do Brasil. *Anudrio Brasileiro de Economia Florestal, 11*, 201-232.

de Queiroz, L. P. (2006). The Brazilian Caatinga: phytogeographical patterns inferred from distribution data of the Leguminosae. In *Neotropical Savannas and Seasonally Dry Forests* (pp. 121-157): CRC Press.

Diefendorf, A. F., Mueller, K. E., Wing, S. L., Koch, P. L., & Freeman, K. H. (2010). Global patterns in leaf 13C discrimination and implications for studies of past and future climate. *Proceedings of the National Academy of Sciences, 107*(13), 5738-5743. doi:10.1073/pnas.0910513107

Douglas, P. M. J., Pagani, M., Brenner, M., Hodell, D. A., & Curtis, J. H. (2012). Aridity and vegetation composition are important determinants of leaf-wax δD values in southeastern Mexico and Central America. *Geochimica Et Cosmochimica Acta, 97*, 24-45. doi:10.1016/j.gca.2012.09.005

Eckmeier, E., & Wiesenberg, G. L. B. (2009). Short-chain n-alkanes (C16–20) in ancient soil are useful molecular markers for prehistoric biomass burning. *Journal of Archaeological Science, 36*(7), 1590-1596. doi:10.1016/j.jas.2009.03.021

Edwards, E. J., Osborne, C. P., Strömberg, C. A. E., & Smith, S. A. (2010). The Origins of C4 Grasslands: Integrating Evolutionary and Ecosystem Science. *Science, 328*(5978), 587-591. doi:10.1126/science.1177216

Eglinton, G., & Hamilton, R. J. (1967). Leaf epicuticular waxes. *Science, 156*(3780), 1322-1335. doi:10.1126/science.156.3780.1322

Eiten, G. (1972). The cerrado vegetation of Brazil. *The Botanical Review, 38*(2), 201-341. doi:10.1007/BF02859158

Faria, S. R., De La Rosa, J. M., Knicker, H., González-Pérez, J. A., Villaverde, J., & Keizer, J. J. (2015). Wildfire-induced alterations of topsoil organic matter and their recovery in Mediterranean eucalypt stands detected with biogeochemical markers. *European Journal of Soil Science, 66*(4), 699-713. doi:10.1111/ejss.12254

Feakins, S. J., Levin, N. E., Liddy, H. M., Sieracki, A., Eglinton, T. I., & Bonnefille, R. (2013). Northeast African vegetation change over 12 m.y. *Geology, 41*(3), 295-298. doi:10.1130/g33845.1

Feakins, S. J., Peters, T., Wu, M. S., Shenkin, A., Salinas, N., Girardin, C. A. J., et al. (2016). Production of leaf wax n-alkanes across a tropical forest elevation transect. *Organic Geochemistry, 100*, 89-100. doi:10.1016/j.orggeochem.2016.07.004

Feakins, S. J., & Sessions, A. L. (2010). Crassulacean acid metabolism influences D/H ratio of leaf wax in succulent plants. *Organic Geochemistry, 41*(12), 1269-1276. doi:10.1016/j.orggeochem.2010.09.007

Feakins, S. J., Wu, M. S., Ponton, C., Galy, V., & West, A. J. (2018). Dual isotope evidence for sedimentary integration of plant wax biomarkers across an Andes-Amazon elevation transect. *Geochimica Et Cosmochimica Acta, 242*, 64-81. doi:10.1016/j.gca.2018.09.007

Ferreira, J. Q., Chiessi, C. M., Hirota, M., Oliveira, R. S., Prange, M., Häggi, C., et al. (2022). Changes in obliquity drive tree cover shifts in eastern tropical South America. *Quaternary Science Reviews, 279*, 107402. doi:10.1016/j.quascirev.2022.107402

Ficken, K. J., Li, B., Swain, D. L., & Eglinton, G. (2000). An n-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes. *Organic Geochemistry, 31*(7–8), 745-749. doi:10.1016/S0146-6380(00)00081-4

Fornace, K. L., Whitney, B. S., Galy, V., Hughen, K. A., & Mayle, F. E. (2016). Late Quaternary environmental change in the interior South American tropics: new insight from leaf wax stable isotopes. *Earth and Planetary Science Letters, 438*, 75-85. doi:10.1016/j.epsl.2016.01.007

Garcin, Y., Schefuß, E., Schwab, V. F., Garreta, V., Gleixner, G., Vincens, A., et al. (2014). Reconstructing C3 and C4 vegetation cover using n-alkane carbon isotope ratios in recent lake sediments from Cameroon, Western Central Africa. *Geochimica Et Cosmochimica Acta, 142*, 482-500. doi:10.1016/j.gca.2014.07.004

Goñi, M. A., Ruttenberg, K. C., & Eglinton, T. I. (1997). Sources and contribution of terrigenous organic carbon to surface sediments in the Gulf of Mexico. *Nature, 389*, 275. doi:10.1038/38477

Goodland, R. (1971). A Physiognomic Analysis of the `Cerrado' Vegetation of Central Brasil. *Journal of Ecology, 59*(2), 411-419. doi:10.2307/2258321

Häggi, C., Bertassoli, D. J., Akabane, T. K., So, R. T., Sawakuchi, A. O., Chiessi, C. M., et al. (2023). Plantwax carbon isotope data from tropical South American surface soils [Data set]. Zenodo. https://doi.org/10.5281/zenodo.10214560.

Häggi, C., Chiessi, C. M., Merkel, U., Mulitza, S., Prange, M., Schulz, M., & Schefuß, E. (2017). Response of the Amazon rainforest to late Pleistocene climate variability. *Earth and Planetary Science Letters, 479*, 50-59. doi:10.1016/j.epsl.2017.09.013

Häggi, C., Naafs, B. D. A., Silvestro, D., Bertassoli, D. J., Akabane, T. K., Mendes, V. R., et al. (2023). GDGT distribution in tropical soils and its potential as a terrestrial paleothermometer revealed by Bayesian deep-learning models. *Geochimica Et Cosmochimica Acta, 362*, 41-64. doi:j.gca.2023.09.014

Häggi, C., Pätzold, J., Bouillon, S., & Schefuß, E. (2021). Impact of selective degradation on molecular isotope compositions in oxic and anoxic marine sediments. *Organic Geochemistry, 153*, 104192. doi:10.1016/j.orggeochem.2021.104192

Häggi, C., Sawakuchi, A. O., Chiessi, C. M., Mulitza, S., Mollenhauer, G., Sawakuchi, H. O., et al. (2016). Origin, transport and deposition of leaf-wax biomarkers in the Amazon Basin and the adjacent Atlantic. *Geochimica Et Cosmochimica Acta, 192*, 149-165. doi:10.1016/j.gca.2016.07.002

Herrmann, N., Boom, A., Carr, A. S., Chase, B. M., Granger, R., Hahn, A., et al. (2016). Sources, transport and deposition of terrestrial organic material: A case study from southwestern Africa. *Quaternary Science Reviews, 149*, 215-229. doi:10.1016/j.quascirev.2016.07.028

Hirota, M., Holmgren, M., Van Nes, E. H., & Scheffer, M. (2011). Global Resilience of Tropical Forest and Savanna to Critical Transitions. *Science, 334*(6053), 232-235. doi:10.1126/science.1210657

Hoyos, N., Correa-Metrio, A., Jaramillo, C., Villegas, J. C., & Escobar, J. (2022). Effects of consecutive dry and wet days on the forest–savanna boundary in north-west South America. *Global Ecology and Biogeography, 31*(2), 347-361. doi:10.1111/geb.13432

Huang, Y., Dupont, L., Sarnthein, M., Hayes, J. M., & Eglinton, G. (2000). Mapping of C4 plant input from North West Africa into North East Atlantic sediments. *Geochimica Et Cosmochimica Acta, 64*(20), 3505-3513. doi:10.1016/S0016-7037(00)00445-2

Huang, Y., Eglinton, G., Ineson, P., Latter, P. M., Bol, R., & Harkness, D. D. (1997). Absence of carbon isotope fractionation of individual n-alkanes in a 23-year field decomposition experiment with Calluna vulgaris. *Organic Geochemistry, 26*(7), 497-501. doi:10.1016/S0146-6380(97)00027-2

Huang, Y., Street-Perrott, F. A., Metcalfe, S. E., Brenner, M., Moreland, M., & Freeman, K. H. (2001). Climate Change as the Dominant Control on Glacial-Interglacial Variations in C3 and C4 Plant Abundance. *Science, 293*(5535), 1647-1651. doi:10.1126/science.1060143

Jaramillo, C. (2023). The evolution of extant South American tropical biomes. *New Phytologist, 239*(2), 477-493. doi:10.1111/nph.18931

Jenkins, C. N., Pimm, S. L., & Joppa, L. N. (2013). Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences, 110*(28), 2602-2610. doi:10.1073/pnas.1302251110

Ji, Y., Mao, G., Wang, Y., & Bartlam, M. (2013). Structural insights into diversity and n-alkane biodegradation mechanisms of alkane hydroxylases. *Frontiers in microbiology, 4*. Review. doi:10.3389/fmicb.2013.00058

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., et al. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data, 4*(1), 170122. doi:10.1038/sdata.2017.122

Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibisch, P. L., Nowicki, C., et al. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences, 106*(23), 9322-9327. doi:10.1073/pnas.0810306106

Kohn, M. J. (2010). Carbon isotope compositions of terrestrial C3 plants as indicators of (paleo)ecology and (paleo)climate. *Proceedings of the National Academy of Sciences, 107*(46), 19691-19695. doi:10.1073/pnas.1004933107

Kortschak, H. P., Hartt, C. E., & Burr, G. O. (1965). Carbon Dioxide Fixation in Sugarcane Leaves. *Plant Physiology, 40*(2), 209-213. doi:10.1104/pp.40.2.209

Krull, E., Sachse, D., Mügler, I., Thiele, A., & Gleixner, G. (2006). Compound-specific δ13C and δ2H analyses of plant and soil organic matter: A preliminary assessment of the effects of vegetation change on ecosystem hydrology. *Soil Biology and Biochemistry, 38*(11), 3211-3221. doi:10.1016/j.soilbio.2006.04.008

Lehmann, C. E. R., Anderson, T. M., Sankaran, M., Higgins, S. I., Archibald, S., Hoffmann, W. A., et al. (2014). Savanna Vegetation-Fire-Climate Relationships Differ Among Continents. *Science*, *343*(6170), 548-552. doi:10.1126/science.1247355

Liu, J., & An, Z. (2020). Leaf wax n-alkane carbon isotope values vary among major terrestrial plant groups: Different responses to precipitation amount and temperature, and implications for paleoenvironmental reconstruction. *Earth-Science Reviews*, *202*, 103081. doi:10.1016/j.earscirev.2020.103081

Liu, J., Zhao, J., He, D., Huang, X., Jiang, C., Yan, H., et al. (2022). Effects of plant types on terrestrial leaf wax long-chain n-alkane biomarkers: Implications and paleoapplications. *Earth-Science Reviews*, *235*, 104248. doi:10.1016/j.earscirev.2022.104248

Liu, W., Huang, Y., An, Z., Clemens, S. C., Li, L., Prell, W. L., & Ning, Y. (2005). Summer monsoon intensity controls C4/C3 plant abundance during the last 35 ka in the Chinese Loess Plateau: Carbon isotope evidence from bulk organic matter and individual leaf waxes. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *220*(3), 243-254. doi:10.1016/j.palaeo.2005.01.001

Lloyd, J., Bird, M. I., Vellen, L., Miranda, A. C., Veenendaal, E. M., Djagbletey, G., et al. (2008). Contributions of woody and herbaceous vegetation to tropical savanna ecosystem productivity: a quasi-global estimate†. *Tree Physiology*, *28*(3), 451-468. doi:10.1093/treephys/28.3.451

Magill, C. R., Ashley, G. M., & Freeman, K. H. (2013). Ecosystem variability and early human habitats in eastern Africa. *Proceedings of the National Academy of Sciences*, *110*(4), 1167-1174. doi:10.1073/pnas.1206276110

Magill, C. R., Eglinton, G., & Eglinton, T. I. (2019). Isotopic variance among plant lipid homologues correlates with biodiversity patterns of their source communities. *Plos One*, *14*(2), e0212211. doi:10.1371/journal.pone.0212211

Martinelli, L. A., Nardoto, G. B., Soltangheisi, A., Reis, C. R. G., Abdalla-Filho, A. L., Camargo, P. B., et al. (2020). Determining ecosystem functioning in Brazilian biomes through foliar carbon and nitrogen concentrations and stable isotope ratios. *Biogeochemistry*. doi:10.1007/s10533-020-00714-2

Messerschmid, T. F. E., Wehling, J., Bobon, N., Kahmen, A., Klak, C., Los, J. A., et al. (2021). Carbon isotope composition of plant photosynthetic tissues reflects a Crassulacean Acid Metabolism (CAM) continuum in the majority of CAM lineages. *Perspectives in Plant Ecology, Evolution and Systematics*, *51*, 125619. doi:10.1016/j.ppees.2021.125619

Minasny, B., Malone, B. P., McBratney, A. B., Angers, D. A., Arrouays, D., Chambers, A., et al. (2017). Soil carbon 4 per mille. *Geoderma*, *292*, 59-86. doi:10.1016/j.geoderma.2017.01.002

Mulitza, S., Chiessi, C. M., Schefuß, E., Lippold, J., Wichmann, D., Antz, B., et al. (2017). Synchronous and proportional deglacial changes in Atlantic meridional overturning and northeast Brazilian precipitation. *Paleoceanography*, *32*(6), 622-633. doi:10.1002/2017PA003084

Nguyen Tu, T. T., Egasse, C., Zeller, B., Bardoux, G., Biron, P., Ponge, J.-F., et al. (2011). Early degradation of plant alkanes in soils: A litterbag experiment using 13C-labelled leaves. *Soil Biology and Biochemistry*, *43*(11), 2222-2228. doi:10.1016/j.soilbio.2011.07.009

O'Leary, M. H. (1981). Carbon isotope fractionation in plants. *Phytochemistry*, *20*(4), 553-567. doi:10.1016/0031-9422(81)85134-5

Ofiti, N., Zosso, C., Soong, J., Solly, E., Torn, M., Wiesenberg, G., & Schmidt, M. (2021). Warming promotes loss of subsoil carbon through accelerated degradation of plant-derived organic matter. *Soil Biology and Biochemistry*, *156*, 108185. doi:10.1016/j.soilbio.2021.108185

Oliveira-Filho, A. T., & Fontes, M. A. L. (2000). Patterns of Floristic Differentiation among Atlantic Forests in Southeastern Brazil and the Influence of Climate1. *Biotropica*, *32*(4b), 793-810. doi:10.1111/j.1744-7429.2000.tb00619.x

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., et al. (2001). Terrestrial ecoregions of the worlds: A new map of life on Earth. *Bioscience*, *51*(11), 933-938. doi:10.1641/0006-3568(2001)051[0933:teotwa]2.0.co;2

Pivello, V. R. (2011). The use of fire in the Cerrado and Amazonian rainforest of Brazil: Past and present. *Fire Ecology*, *7*(1), 24-39. doi:10.4996/fireecology.0701024

R_Core_Team. (2022). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Reis, L. S., Bouloubassi, I., Mendez-Millan, M., Guimarães, J. T. F., de Araújo Romeiro, L., Sahoo, P. K., & Pessenda, L. C. R. (2022). Hydroclimate and vegetation changes in southeastern Amazonia over the past ~25,000 years. *Quaternary Science Reviews*, *284*, 107466. doi:10.1016/j.quascirev.2022.107466

Rieley, G., Collier, R. J., Jones, D. M., Eglinton, G., Eakin, P. A., & Fallick, A. E. (1991). Sources of sedimentary lipids deduced from stable carbon-isotope analyses of individual compounds. *Nature, 352*(6334), 425-427. 10.1038/352425a0. doi:10.1038/352425a0

Rommerskirchen, F., Eglinton, G., Dupont, L., Güntner, U., Wenzel, C., & Rullkötter, J. (2003). A north to south transect of Holocene southeast Atlantic continental margin sediments: Relationship between aerosol transport and compound-specific δ13C land plant biomarker and pollen records. *Geochemistry, Geophysics, Geosystems, 4*(12), 1101–1128. doi:10.1029/2003GC000541

Rommerskirchen, F., Plader, A., Eglinton, G., Chikaraishi, Y., & Rullkotter, J. (2006). Chemotaxonomic significance of distribution and stable carbon isotopic composition of long-chain alkanes and alkan-1-ols in C-4 grass waxes. *Organic Geochemistry, 37*(10), 1303-1332. doi:10.1016/j.orggeochem.2005.12.013

Ruggiero, P. G. C., Batalha, M. A., Pivello, V. R., & Meirelles, S. T. (2002). Soil-vegetation relationships in cerrado (Brazilian savanna) and semideciduous forest, Southeastern Brazil. *Plant Ecology, 160*(1), 1-16. doi:10.1023/A:1015819219386

Salati, E., Dall'Olio, A., Matsui, E., & Gat, J. R. (1979). Recycling of water in the Amazon Basin: An isotopic study. *Water Resources Research, 15*(5), 1250-1258. doi:10.1029/WR015i005p01250

Sarangi, V., Roy, S., & Sanyal, P. (2022). Effect of burning on the distribution pattern and isotopic composition of plant biomolecules: Implications for paleoecological studies. *Geochimica Et Cosmochimica Acta, 318*, 305-327. doi:10.1016/j.gca.2021.12.003

Schefuß, E., Schouten, S., & Schneider, R. R. (2005). Climatic controls on central African hydrology during the past 20,000 years. *Nature, 437*(7061), 1003-1006. doi:10.1038/nature03945

Schwab, V. F., Garcin, Y., Sachse, D., Todou, G., Séné, O., Onana, J.-M., et al. (2015). Effect of aridity on δ13C and δD values of C3 plant- and C4 graminoid-derived leaf wax lipids from soils along an environmental gradient in Cameroon (Western Central Africa). *Organic Geochemistry, 78*, 99-109. doi:10.1016/j.orggeochem.2014.09.007

Sexton, J. O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., et al. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *International Journal of Digital Earth, 6*(5), 427-448. doi:10.1080/17538947.2013.786146

Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E. (2009). Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by in situ evolution of adaptations to fire. *Proceedings of the National Academy of Sciences, 106*(48), 20359-20364. doi:10.1073/pnas.0903410106

Slack, C. R., & Hatch, M. D. (1967). Comparative studies on the activity of carboxylases and other enzymes in relation to the new pathway of photosynthetic carbon dioxide fixation in tropical grasses. *The Biochemical journal, 103*(3), 660-665. doi:10.1042/bj1030660

Stockmann, U., Padarian, J., McBratney, A., Minasny, B., de Brogniez, D., Montanarella, L., et al. (2015). Global soil organic carbon assessment. *Global Food Security, 6*, 9-16. doi:10.1016/j.gfs.2015.07.001

Tu, T. T. N., Derenne, S., Largeau, C., Bardoux, G., & Mariotti, A. (2004). Diagenesis effects on specific carbon isotope composition of plant n-alkanes. *Organic Geochemistry, 35*(3), 317-329. doi:10.1016/j.orggeochem.2003.10.012

Vogts, A., Moossen, H., Rommerskirchen, F., & Rullkötter, J. (2009). Distribution patterns and stable carbon isotopic composition of alkanes and alkan-1-ols from plant waxes of African rain forest and savanna C3 species. *Organic Geochemistry, 40*(10), 1037-1054. doi:10.1016/j.orggeochem.2009.07.011

Wu, M. S., Feakins, S. J., Martin, R. E., Shenkin, A., Bentley, L. P., Blonder, B., et al. (2017). Altitude effect on leaf wax carbon isotopic composition in humid tropical forests. *Geochimica Et Cosmochimica Acta, 206*, 1-17. doi:10.1016/j.gca.2017.02.022

Wu, M. S., West, A. J., & Feakins, S. J. (2019). Tropical soil profiles reveal the fate of plant wax biomarkers during soil storage. *Organic Geochemistry, 128*, 1-15. doi:10.1016/j.orggeochem.2018.12.011

Yang, C.-K., Huang, B.-H., Ho, S.-W., Huang, M.-Y., Wang, J.-C., Gao, J., & Liao, P.-C. (2018). Molecular genetic and biochemical evidence for adaptive evolution of leaf abaxial epicuticular wax crystals in the genus Lithocarpus (Fagaceae). *BMC Plant Biology, 18*(1), 196. doi:10.1186/s12870-018-1420-4

Yang, D., & Bowen, G. J. (2022). Integrating plant wax abundance and isotopes for paleo-vegetation and paleoclimate reconstructions: a multi-source mixing model using a Bayesian framework. *Clim. Past, 18*(10), 2181-2210. doi:10.5194/cp-18-2181-2022

Zemp, D. C., Schleussner, C.-F., Barbosa, H. M. J., Hirota, M., Montade, V., Sampaio, G., et al. (2017). Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks. *Nature Communications, 8*(1), 14681. doi:10.1038/ncomms14681

1146 Zhang, D., Beverly, E. J., Levin, N. E., Vidal, E., Matia, Y., & Feakins, S. J. (2021). Carbon isotopic composition of
1147 plant waxes, bulk organics and carbonates from soils of the Serengeti grasslands. *Geochimica Et*
1148 *Cosmochimica Acta, 311*, 316-331. doi:10.1016/j.gca.2021.07.005
1149 Zhang, Y., Zheng, M., Meyers, P. A., & Huang, X. (2017). Impact of early diagenesis on distributions of Sphagnum
1150 n-alkanes in peatlands of the monsoon region of China. *Organic Geochemistry, 105*, 13-19.
1151 doi:10.1016/j.orggeochem.2016.12.007

1152