Biomarker approaches for reconstructing terrestrial environmental change

Gordon N. Inglis[1], Tripti Bhattacharya[2], Jordon D. Hemingway[3,4], Emily H. Hollingsworth[1], Sarah J. Feakins[5], Jessica E. Tierney[6]

[1] *School of Ocean and Earth Science, National Oceanography Centre Southampton, UK*
[2] *Earth and Environmental Sciences, Syracuse University, USA*
[3] *Department of Earth and Planetary Science, Harvard University, USA*
[4] *Geological Institute, Department of Earth Sciences, ETH Zurich, Switzerland*
[5] *Department of Earth Sciences, University of Southern California, USA*
[6] *Department of Geosciences, University of Arizona, USA*

Shortened running title*:* Biomarkers in terrestrial environments

Corresponding author: Gordon N. Inglis

Email: gordon.inglis@soton.ac.uk

Abstract

The response of the terrestrial biosphere to warming remains one of the most poorly understood and quantified aspects of the climate system. One way to test the behaviour of the Earth system in warm climate states is to examine the geological record. The abundance, distribution, and/or isotopic composition of source-specific organic molecules ("biomarkers") have been used to reconstruct terrestrial paleoenvironmental change over a range of geological timescales. Here, we review new or recently improved biomarker approaches for reconstructing: (i) physical climate variables (land temperature, rainfall), (ii) ecosystem state variables (vegetation, fire regime), and (iii) biogeochemical variables (soil residence time, methane cycling). This review encompasses a range of key compound classes (e.g., lipids, lignin, and carbohydrates). In each section, we explore the concept behind key biomarker approaches and discuss their successes as paleoenvironmental indicators. We emphasize that analyzing several biomarkers in tandem can provide unique insights into the Earth System.

Keywords: Biomarkers, lipids, paleoclimate, paleoenvironment, biogeochemistry, proxies

1. Introduction

Human activity is fundamentally altering the chemical composition of the atmosphere and warming the planet. Higher temperatures coupled with changing rainfall patterns will alter vegetation distributions, soil residence time, and a variety of terrestrial biogeochemical processes, resulting in feedbacks that can amplify or reduce greenhouse gas forcing (Rohling et al 2012). Quantifying these feedbacks remains challenging because of large uncertainties associated with land surface processes and their response to carbon cycle perturbations (Luo 2007).

The geological record encompasses a wide range of climate states that can help explore the response of regional climate and terrestrial ecosystems to external forcing (Tierney et al 2020). However, direct instrumental records span only the last few centuries. In deeper time, researchers must rely upon indirect biological, chemical, or physical indicators (“proxies”) preserved within the sedimentary record. Plant fossils (e.g., leaves, pollen, spores) reveal information about past ecosystems, but also inform climate through analysis of leaf physiognomy (e.g., leaf margin analysis; Wilf 1997) and leaf and/or pollen taxonomy. The stable and clumped isotopic composition of terrestrial carbonates (e.g., herbivore teeth, soil carbonates) also provide important constraints into terrestrial ecosystems and climate (e.g., Cerling 1984). However, these archives are often poorly preserved and/or discontinuous, such that additional techniques are often required.

The abundance, distribution, and/or isotope composition of source-specific organic molecules—termed “biomarkers” (Sidebar 1)—preserved in a variety of sedimentary deposits provides an additional window into ancient terrestrial environments (Figure 1). Here, we provide a holistic overview and synthesis of key terrestrial biomarker proxies. We review biomarker approaches for reconstructing: (i) physical climate variables (land temperature, rainfall), (ii) ecosystem state variables (vegetation, fire regime), and (iii) biogeochemical variables (soil residence time, methane cycling). In each section, we explore the concept behind key biomarker approaches—highlighting their limitations and challenges—and discuss their successes as paleoenvironmental indicators. We also highlight emerging terrestrial biomarker proxies and discuss potential for future studies.

2. Reconstructing terrestrial environmental change using organic biomarkers

2.1. Temperature proxies

Terrestrial temperature exerts a major control on vegetation distributions (Section 2.2), the hydrological cycle (Section 2.3) and a variety of biogeochemical cycles (Section 2.4 to 2.6). Yet even for relatively recent, well-studied time intervals (e.g., the Holocene; 12 thousand years (ka) to present), large continental regions lack rigorous temperature constraints (Kaufman et al 2020). Branched glycerol dialkyl glycerol tetraethers (brGDGTs) are increasingly used to reconstruct land temperature. Branched GDGTs are comprised of two *n*-alkyl chains, each containing 4-6 methyl groups and 0-2 cyclopentane moieties (Sinninghe Damsté et al 2000). The stereochemistry of the glycerol units confirms a bacterial—rather than archaeal—source organism, but the exact biological source organism(s) remain(s) debated (Sinninghe Damsté et al 2018). Acidobacteria are a potential source organism, due to their widespread abundance in soil and peat (Weijers et al., 2009). Various Acidobacterial strains synthesize the presumed “building block” of brGDGTs (*iso*-diabolic acid). However, only a few Acidobacterial strains have been directly shown to produce brGDGTs (Sinninghé Damste et al., 2011; 2018; Halamka et al., 2021). Recent work demonstrates that oxygen limitation can trigger brGDGT production in at least one Acidobacterial strain, perhaps explaining why few brGDGTs had previously been identified in the laboratory (Halamka et al., 2021). Using a global soil database, Weijers et al. (2007b) first demonstrated that the number of methyl groups (captured via the methylation of branched tetraethers; MBT) varied as a function of temperature and pH (Weijers et al 2007b), whereas the number of cyclopentane moieties (captured via the cyclization of branched tetraethers; CBT) varied as a function of pH alone (Weijers et al 2007b). This led to the formulation of the MBT/CBT index, which was correlated to mean annual air temperature (MAAT; i.e., measured 2 m above ground as determined from climatological reanalysis data) (Weijers et al 2007b).

The MBT/CBT index was originally applied in a marginal marine setting (Weijers et al 2007a) and revealed that tropical African temperatures increased by ~4°C during the last deglaciation (Weijers et al 2007a). However, later analytical refinements (Becker et al 2013, De Jonge et al 2013, Hopmans et al 2016) found that penta- and hexa-methylated brGDGTs contained methyl groups at the C5 (5-methyl brGDGTs) or C6 position (6-methyl brGDGTs) and that these co-eluted in the original method of Weijers et al (2007b). This led to the formulation of the methylation index of 5-methyl brGDGTs (MBT'_{5ME}). This metric removes the soil pH dependency that was present in the original MBT index (Weijers et al 2007b) and yields more accurate MAAT estimates in soils (De Jonge et al

2014). However, there remains a poor correlation between MBT'$_{5ME}$ and MAAT in soils with a high abundance of 6-methyl brGDGTs (i.e., arid and/or alkaline soils; Dang et al 2016). Some studies have suggested that samples with a high 6- over 5-methyl brGDGT ratio (e.g., IR$_{6ME}$ > 0.5) are not reliable for temperature inference (Dang et al 2016), although in a global analysis of the soil and peat dataset, Crampton-Flood et al., (2020) did not find support for this particular cut-off. Only core tops with a very high IR$_{6Me}$ value (> 0.8) showed a high residual error.

The MBT'$_{5ME}$ index correlates strongly with MAAT in a global peat dataset (Naafs et al 2017) and is comparable to the relationship observed in soils (Crampton-Flood et al 2020). The MBT'$_{5ME}$ index has been increasingly applied in peat archives to reconstruct land temperatures and can reproduce millennial-scale Holocene climate events (i.e., Younger Dryas, Bølling-Allerød) (Zheng et al 2017). However, peat and/or lignite deposits can be subject to dramatic changes in local hydrology which may lead to an apparent change in brGDGT-derived temperature estimates. For instance, brGDGT distributions within a UK lignite deposit imply ~10°C cooling during the onset of the Paleocene-Eocene Thermal Maximum (PETM) (Inglis et al 2019a). This coincided with enhanced waterlogging and is consistent with in situ production of aquatic brGDGTs (see below; Weber et al 2018). The application of MBT'$_{5ME}$ in dynamic terrestrial environments (e.g., lake-mire transitions) should therefore be interpreted cautiously and within a multi-proxy and/or data-modelling framework.

BrGDGTs respond differently in lakes than they do in soils. In particular, the application of mineral soil or peat calibrations in lacustrine settings can yield up to 10°C colder-than-predicted temperatures (Tierney et al 2010). In lake environments, brGDGTs can be produced in situ—possibly by a distinct microbial community (Weber et al 2018)—and regional or global lake-specific calibration models are therefore required (Martínez-Sosa et al 2021). The relationship between MBT'$_{5ME}$ and temperature is stronger in lakes than soils and peats; conversely, the latter have a stronger relationship to pH (Martínez-Sosa et al 2021). This supports the hypothesis that different microbial communities are involved in the production of these compounds. Notably, much like arid and/or alkaline soils, lacustrine brGDGT distributions are distinct in hypersaline and/or alkaline lakes, with higher relative abundance of brGDGT-*Ia* and more 6′ methyl isomers (Martinez-Sosa et al., 2021). Hypersaline and/or alkaline lakes have higher average MBT$_{5ME'}$ values than other lakes for a given temperature, implying that global lake calibrations may overestimate temperature values in such environments. This potentially limits the application of MBT'$_{5me}$ in these environments (Martínez-Sosa et al 2021, Tierney et al 2010).

One of the strengths of the MBT'$_{5ME}$ is that is can be applied in older geological time intervals when other terrestrial temperature proxies (e.g., pollen transfer functions) may be less certain. For example, Tibbett et al (2021) used MBT'$_{5ME}$ to constrain temperatures across the Eocene-Oligocene transition (EOT) and found an abrupt ~4-6°C cooling in East Antarctica (Prydz Bay) prior to the establishment of continental-scale ice sheets (Figure 2). The branched GDGT-inferred cooling is larger, more abrupt and later than that reconstructed from the rock-derived chemical weathering proxies. This is related to catchment sourcing: rock erosion is dominated by the mountainous sectors of catchments, whereas soil-derived biomarkers are dominated by the greater lowland extent of catchments. The two signals together reveal the earlier cooling of the high mountain regions in the rock-derived proxy, whereas the brGDGTs detect the later and more abrupt lowland cooling.

Despite these successes, $\mathrm{MBT'}_{\mathrm{5ME}}$ values can be overprinted by brGDGT production in rivers and/or marine sediments (e.g., upland soils, rivers, marine sediments; Crampton-Flood et al 2018, Zell et al 2014). If these sources were substantial, for example in the lowland Amazon River (Zell et al., 2014), they can bias brGDGT temperature estimates. This can be partially resolved by screening for and excluding and/or correcting for possible marine/riverine overprints (Crampton-Flood et al 2018). Marine settings are typically characterised by a high degree of cyclisation of tetramethylated brGDGTs (i.e., high $\#\mathrm{rings}_{\mathrm{tetra}}$ values; Sinninghé Damste, 2018). This is because marine environments are relatively alkaline compared to soils and peat. As $\#\mathrm{rings}_{\mathrm{tetra}}$ values do not exceed 0.7 in modern soils, higher values (i.e., >0.7) are evidence for marine *in situ* production (Sinninghe Damsté, 2016). Rivers often have a higher contribution of 6-methyl brGDGT isomers (de Jonge et al., 2015), yielding high Isomerisation Ratio values (IR; de Jonge et al., 2015). However, diagnosing riverine input using the IR is challenging as both arid and alkaline soils contain a high abundance of 6-methyl brGDGT isomers (Dang et al, 2016).

The MBT'$_{5ME}$ index also has a theoretical maximum temperature of ~25-30°C in soils and peat (Crampton-Flood et al 2020, De Jonge et al 2014, Naafs et al 2017, Weijers et al 2007b), which limits use of this proxy in greenhouse climates (e.g., the Eocene). Branched GDGT distributions in settings that are warmer than modern remain largely unknown, although both soil (Chen et al 2018) and lake water (Martinez-Sosa et al 2020) incubations indicate that organisms can synthesise abundant brGDGTs when grown at temperatures above 35^{o}C.

Isoprenoidal glycerol dibiphytanyl glycerol tetraethers (isoGDGTs) might help reconstruct temperatures not represented in modern climates. IsoGDGTs are synthesised by archaea and comprised of two isoprenoid side chains containing up to eight cyclopentane moieties (although rarely more than 4 cyclopentane moieties; Schouten et al, 2013 and ref. therein). The number of cyclopentane moieties increases at higher temperatures (De Rosa et al 1980), resulting in a more densely packed and stable membrane. This is the premise of the TEX$_{86}$ surface water temperature proxy (Schouten et al 2002). In some lacustrine settings—generally large and deep lakes—TEX$_{86}$ is correlated to annual mean lake temperature (Powers et al 2004) and can be used to infer continental temperatures (Tierney et al 2008). However, various factors can influence TEX$_{86}$ values in lakes (e.g. input of GDGTs from sources other than Thaumarchaeota) (Blaga et al 2009).

The degree of isoGDGT cyclisation is poorly correlated with temperature in soils (Coffinet et al 2014) and peats (Naafs et al 2018), likely due to the wide range of potential source organisms in these settings. However, recent work identified isoGDGTs with more than four cyclopentane moieties (e.g., GDGT-5 to -7) in tropical (>19.5°C) and acidic (pH < 5.1) peats (Naafs et al 2018). These compounds had previously only been observed in extremophile cultures and/or hyperthermophilic settings (e.g., hot springs) (Schouten et al 2013), suggesting that their occurrence in tropical, acidic peats represents an adaptation to higher temperatures and/or lower pH. Highly-cyclised isoGDGTs (e.g., isoGDGT-5, but also isoGDGT-6) have been detected in early Eocene (~56 to 48 Ma) lignite deposits (Naafs et al 2018). This confirms that mid-latitude temperatures were higher than 19.5°C during the early Eocene, consistent with elevated brGDGT temperature estimates (>22-29°C) from the same samples (Naafs et al 2018). The abundance of isoGDGT-5 in low latitude Eocene lignite deposits is higher than that found in any modern peat, implying MAATs were higher than presently found in the low latitudes. These results suggest that highly-cyclised isoGDGTs (e.g., isoGDGT-5 to -8) could potentially be used to reconstruct temperatures >30°C (c.f., the brGDGT temperature proxy). However, experimental approaches (e.g., mesocosm or microcosm experiments) are required to decipher the relationship between highly-cyclised isoGDGTs and MAAT beyond the modern range. Genomic approaches may provide additional insights into the physiological function of highly-cyclised GDGTs (e.g., Zeng et al., 2019).

2.2. Vegetation proxies

The geological record can provide insights into the response of plant communities to CO_2-induced warming and associated hydroclimate changes. Leaf fossils document local vegetation change across major climate events (e.g. the PETM; Wing et al., 2005). Biomarkers are highly complementary sources of information and provide a more spatially-integrated perspective. The most common biomarkers used to reconstruct changes in the plant community are: i) long-chain *n*-alkyl compounds (e.g., *n*-alkanes, *n*-alcohols, *n*-alkanoic acids) (Eglinton & Hamilton 1967), ii) lignin (e.g., lignin phenols; (Hedges & Mann 1979), and iii) terpenoids (e.g. di- and triterpenoids; Otto & Simoneit 2001)

Long-chain *n*-alkane compounds typically have carbon chain lengths ranging from C_{25} to C_{35} and exhibit high odd-over-even ratios, whereas long-chain *n*-alcohols or *n*-alkanoic acids range from C_{24} to C_{34} and exhibit high even-over-odd ratios. Various factors determine the abundance of *n*-alkanoic acids vs n-alkanes in the sedimentary record. Over short timescales, there can be preferential loss of long-chain *n*-alkanes during soil litter decomposition (Wu et al., 2018). Over longer timescales (i.e., millions of years) or at elevated temperatures, *n*-alkanoic acids are expected to degrade faster due to their functionalized nature. However, laboratory pyrolysis experiments show that they are stable as long as thermal maturity is low (Diefendorf et al., 2015) and remain abundant in some early Cenozoic (Hollis et al., 2021) and late Mesozoic sediments (van Dongen et al., 2006). While both compound classes (*n*-alkanoic acids and *n*-alkanes) are commonly reported separately in paleoclimatic studies, the decision is often guided by absolute abundances, the extent of reworking or petroleum exposure, and/or co-elution with other compounds. To assess this the extent of reworking, we strongly encourage future studies to report thermal maturity ratios (e.g., the carbon preference index (CPI), hopane $\beta\beta/(\alpha\beta+\beta\beta)$ ratios., etc). Where possible dual compound classes can be used to reveal productivity or catchment sourcing differences (e.g. Hemingway et al., 2016, Feakins et al., 2018; 2020).

Studies have looked for patterns in long-chain *n*-alkyl distributions between different plant types (Bush & McInerney 2013, Diefendorf et al 2011) and across climate gradients (Feakins et al 2016b), but recent meta-analyses have found limited value of summary statistics (e.g., average chain length, odd-over-even ratios) as taxonomic or climatic indicators (Bush & McInerney 2013, Diefendorf et al 2011). However, in some low-diversity settings, interpretations of chain length ratios are able to detect plant type variations (Bush & McInerney 2013). For example, in boreal *Sphagnum* moss-dominated peats, the C_{23}/C_{29} or C_{23}/C_{31} *n*-alkane ratio has been used to differentiate between *Sphagnum* moss (dominated by mid-chain homologues; e.g., C_{23}) and woody plants or grasses (which are dominated by long-chain homologues; e.g., C_{29} or C_{31}) and reveals changes in environmental conditions (Nott et al 2000). Certain studies also report a greater predominance of C_{33} and/or C_{35} *n*-alkanes in graminoids (e.g., Bush and McInerney, 2013). As tropical trees produce a substantially higher abundance of C_{29} and C_{31} *n*-alkanes compared to grasses (Garcin et al., 2014), they dominate the sedimentary C_{29} and C_{31} *n*-alkane pool. Thus, longer chain length alkanes (C_{33} and/or C_{35}) may more strongly detect C_4 grass inputs into sediments (e.g., Schefuß et al., 2003). However, as these chain lengths are not taxon-specific, additional evidence from pollen and spores aids interpretations.

Other biomarkers can provide additional information on past vegetation. Lignin—a key component of tree wood and other plant tissues—is a heterogeneous polymer containing a variety of monomeric building blocks (see also Section 2.3 below). The ratio of different lignin monomers (i.e., syringyl, vanillyl, and cinnamyl phenols) can differentiate between non-woody angiosperms, woody angiosperms, and woody gymnosperms (Hedges & Mann 1979). Terpenoids can also distinguish between angiosperms and gymnosperms. Triterpenoids (e.g., ursane, lupane, oleanane) are mostly synthesised by angiosperms, whereas diterpenoids (e.g., abietic acid, abietane, pimarane) are more abundant in gymnosperms, especially evergreens (Diefendorf et al 2012). The relative abundance of di- vs triterpenoids has been applied to reconstruct gymnosperms vs angiosperms abundance in the past, although selective loss of triterpenoids can lead to an overestimation of gymnosperms in the sedimentary record (Diefendorf et al 2014).

Machine learning methods have considerable potential to extract information from vegetation biomarkers, especially the ‘forest’ of peaks present in homologous *n*-alkyl series. Machine learning has detected chemotaxonomic classification power from plant wax distributions, requiring the diagnostic information in molecular distributions across two compound classes (*n*-alkanoic acids and *n*-alkanes) to sufficiently differentiate desert shrub, conifer forest and macrophytes in modern ecosystems and in lake sediments (Peaple et al 2021). These two compound classes are generalist biomarkers, but their proportions and chain length distributions vary across taxa (Diefendorf et al 2011).

Presumably, additional compound classes and specialist biomarkers would add further skill at chemotaxonomic separations. The machine learning approach has advantages over summary metrics such as the average chain length or carbon preference index on a single compound class (usually only the *n*-alkanes) and appears to perform better than linear mixing models on dual compound classes (Gao et al 2011). Machine learning (e.g., non-negative matrix factorization, random forests, neutral networks) can also help untangle source mixing and/or identify end-member distributions (Peaple et al., 2021; Karp et al., 2021; Polissar et al. 2021). However, machine learning is subject to various uncertainties (e.g., input of *n*-alkyl lipids from plants not represented in the training dataset) or from taphonomic alterations between plant and sediment. Machine learning should always be guided by physical and/or chemical knowledge of the relevant biomarkers – otherwise, the outputs may not have clear physical meaning or may be somewhat uncertain. Machine learning may also require training across a wider variety of plant, soil and sedimentary settings and across ecosystems. It remains to be seen whether it will perform well in ecosystems with greater diversity of plant taxa (e.g., rainforest) or lower leaf trait variability. It is also unclear how such machine learning approaches will perform in non-analogue settings beyond the training set, although such non-analogue questions also apply to other proxies. In order to enable further testing of these approaches, we recommend collecting multi-compound plant wax abundances (e.g., *n*-alkanoic acids, *n*-alkanes), including the broadest possible range of chain lengths and the non-dominant odd/even chain lengths to develop information about plant wax in a broader range of species and sedimentary settings.

Another widespread approach for vegetation reconstruction using biomarker methods involves analysing the stable carbon isotopic composition ($\delta^{13}C$) of long-chain *n*-alkyl compounds. The $\delta^{13}C$ value of long-chain *n*-alkyl compounds can readily discern the prevalence of C_3 and C_4 plant types (Huang et al 2001). This results from a large fractionation in plant biosynthesis that differs according to photosynthesis pathway (n.b., the isotopic difference is larger in epicuticular waxes than in bulk tissues). Plant wax from C_4 plants (most tropical grasses and some shrubs) are ^{13}C-enriched (~15 ‰) relative to those produced by C_3 plants. This has been exploited to study the Neogene expansion of C_4 grasslands (Freeman & Colarusso 2001, Polissar et al 2019). Plant wax $\delta^{13}C$ records document major expansion of C_4 grasslands on the African continent ~ 10 Ma (Polissar et al 2019). However, C_4 expansion is not globally uniform and plant wax $\delta^{13}C$ records show that C_4 grassland expansion occurs in the Ganges-Brahmaputra catchment ~3 Ma later (Figure 3) (Karp et al 2021).

Despite the substantial discriminatory difference between photosynthesis pathways, the carbon isotopic composition ($\delta^{13}C$) C_3 plants has a large spread which introduces uncertainties into linear mixing model interpretations of past vegetation change (Diefendorf et al., 2010). As a partial solution, (Cerling et al 2011) proposed a sinusoidal regression between woody cover and $\delta^{13}C$ of soil organic matter, which was adapted for plant waxes (Magill et al 2013). However, this only applies to post-Miocene, low-latitude settings dominated by C_4 grasses. $\delta^{13}C$ variability *within* C_3 plants is also a valuable tool for paleoenvironmental reconstruction. Within C_3 ecosystems there are discernable signals including a primary sensitivity to mean annual precipitation and altitude, but only in wet climates (Wu et al 2017). For further review of carbon isotopes in terrestrial ecosystem reconstructions, other paleoenvironmental interpretations, and areas for future research, readers are referred to a recent review by Diefendorf and Freimuth (2017).

2.3. Hydroclimate proxies

Hydroclimate variability exerts an important influence on vegetation patterns (Section 2.2), soil residence time (Section 2.4) and a variety of other biogeochemical processes (Section 2.5-2.6). However, there are few continuous and direct records of hydrological change in deep time. The hydrogen isotopic composition (δ^2H) of long-chain *n*-alkyl lipids (e.g., *n*-alkanes, *n*-alcohols, *n*-alkanoic acids) can provide important insights into the hydrologic cycle. These compounds have a well-constrained source organism (i.e., higher plants) and offer excellent preservation potential in a variety of sedimentary environments. In addition, the hydrogen is bound to the alkyl carbon making it non-exchangeable in most (<150°C) archives. Sachse et al (2004) first demonstrated an empirical correlation between sedimentary long-chain *n*-alkyl lipid δ^2H (δ^2H_{wax}) and precipitation δ^2H in European lakes. This was later corroborated across a variety of environmental gradients (Balascio et al 2018, Daniels et al 2017, Feakins & Sessions 2010, Garcin et al 2012) and exploited to reconstruct hydroclimate in the geological record (Bhattacharya et al 2018, Fornace et al 2014, Schefuß et al 2005, Tierney et al 2008, Weijers et al 2007a).

δ^2H_{wax} is usually employed as a proxy for the isotopic composition of precipitation (δ^2H_{precip}). This typically integrates a combination of climatic changes including temperature, rainfall amount and character (e.g., stratiform vs. convective rainfall), as well as atmospheric vapor transport. δ^2H_{wax} records thus benefit from interpretations based on isotope-enabled climate model simulations, which can be used to identify the effects of climatic changes on water isotope compositions (Lee et al 2007, Schmidt et al 2007). However, despite the complexity of water isotopes, paleoclimate δ^2H_{wax} records capture larger-scale hydroclimate shifts than local hydrological proxies (c.f. lake levels). The variety of information in δ^2H_{precip} is also an asset; in the mid-latitudes δ^2H_{wax} can be used to distinguish between different seasonal moisture sources (Bhattacharya et al 2018) while in the tropics δ^2H_{wax} reveals general trends in aridity (Tierney & deMenocal 2013) or atmospheric convection (Windler et al 2021). Comparisons between δ^2H_{wax} and ice core or speleothem $\delta^{18}O$ demonstrates that δ^2H_{wax} captures similar information and also records rapid climate changes such as the Younger Dryas cooling (Fornace et al 2014) (Figure 4). Furthermore, δ^2H_{wax} can capture climatic changes in the historical period, as demonstrated by the δ^2H_{wax} record from the Gulf of Aden, which documents increasingly arid conditions over the last 200 years (Tierney et al 2015) (Figure 4). Tandem measurements of $\delta^{13}C_{wax}$ show the presence of the Suess effect, indicating rapid transport of plant waxes from the terrestrial source (the Horn of Africa) to the coastal ocean (Tierney et al 2015). As discussed further below (Section 2.4), such records demonstrate that a young, rapidly overturning plant wax pool allows for rapid climate changes to be recovered despite the presence of a pre-aged and/or reworked plant wax pool.

Many studies use δ^2H_{wax} as a qualitative indicator of climatic changes (e.g., Figure 4). However, quantitative inference of δ^2H_{precip} is possible if reasonable constraints can be made on the apparent fractionation between δ^2H_{wax} and δ^2H_{precip} (= $^2\varepsilon_{wax/w}$; see Sachse et al 2012, Sessions 2016 for a detailed review). Three key factors have the potential to impact $^2\varepsilon_{wax/precip}$: i) soil evaporation, ii) leaf-water transpiration iii) wax biosynthesis (see below).

Although theoretical models have incorporated evaporative ^{2}H-enrichment of soil water into predictions of $^2\varepsilon_{wax/w}$ (Konecky et al 2019, Smith & Freeman 2006), ecohydrological data demonstrates that this effect is minimal. For instance, in drylands, where soil water is readily evaporated and non-bioavailable, woody plants take up rainwater opportunistically via shallow roots and/or seek more consistent water sources by deep rooting strategies (Fan et al 2017), resulting in minimal xylem water ^{2}H-enrichment. In contrast, leaf water ^{2}H-enrichment that occurs during transpiration is more variable, depending on the species and the climatological setting (Daniels et al 2017, Feakins et al 2016a, Feakins & Sessions 2010, Kahmen et al 2013a). In mesic environments where lake water evaporation is minimal, pairing of plant wax with aquatic biomarkers has proved useful for disentangling the effect of transpiration on δ^2H_{wax} (Mügler et al 2008, Rach et al 2014).

An empirical collection of $^2\varepsilon_{wax/w}$ values across species and climatic settings implies that it varies widely (up to ~70 ‰) between plant life forms (i.e., trees, shrubs, forbs, graminoids) and physiological groups (i.e., C_3, C_4, CAM) (Sachse et al 2012). Taken at face value, this suggests that biosynthetic differences have a large impact on apparent fractionation. However, these empirical data combine physiological/biochemical and climatic differences, and separating these influences on fractionation can be challenging (Feakins et al 2016a, Kahmen et al 2013b, Sachse et al 2012). Experimental approaches enable biophysical vs. climatic differences on fractionation to be isolated and have found that ${}^2\varepsilon_{wax/w}$ is relatively consistent between species under controlled environmental settings, with the only major physiological/biochemical difference between grasses (which have a more negative ${}^2\varepsilon_{wax/w}$) and eudicots (Gamarra et al 2016, Gao et al 2014). Since paleoclimate records of δ^2H_{wax} integrate across a landscape, this enables the use of plant community-scale ${}^2\varepsilon_{wax/w}$ values in paleoclimate studies to infer δ^2H_{precip}.

Estimating ${}^2\varepsilon_{wax/precip}$ in the paleoenvironmental record must also account for any seasonal bias relative to annual rainfall. In most environments, plant growth is stimulated by seasonal availability of light and plants often grow using water from the previous season's rainfall. Seasonal studies of modern plant ecohydrology reveal the seasonality of rainfall and plant uptake (Griepentrog et al 2019) and this has been invoked to explain small ${}^2\varepsilon_{wax/precip}$ values observed in alpine gymnosperm flora (Polissar and Freeman, 2010). However, expanded observational and experimental data are required to resolve these uncertainties.

Several strategies are available for improving ${}^2\varepsilon_{wax/w}$ estimates in the paleoenvironmental record. One strategy employs pollen data from the same samples to calculate plant-specific fractionation factors (Feakins 2013), thus opening the possibility of developing region- and time-specific vegetation corrections based upon appropriate modern analogues. When applied to the Miocene (Feakins 2013) and Eocene (Inglis et al 2020), this method can shift inferred δ^2H_{precip} values by >30 ‰. However, pollen-corrected δ^2H_{precip} reconstructions have large uncertainties due to the differential production of pollen between wind and insect-pollinated plants (e.g., pines and grasses are prolific pollen dispersers, whereas tropical forests disperse little pollen). This concern can be qualitatively assessed (e.g., absence of rainforest pollen does not guarantee an absence of rainforest) or quantitatively deconvolved with additional mixing model weighting factors.

A more common strategy for Neogene (~23 Ma to present) climate reconstructions involves measuring *n*-alkyl lipid $\delta^{13}C$ ($\delta^{13}C_{wax}$) values alongside δ^2H_{wax} in order to infer C_3 vs. C_4 plant prevalence (Tipple & Pagani 2010). This is particularly effective in tropical and subtropical regions, where C_4 grasslands are a major component of the ecosystems (Tierney et al 2017, Windler et al 2020) and have a significantly different $^2\varepsilon_{wax/w}$ than C_3 dicots (e.g., Gao et al, 2014). Figure 3 demonstrates how $^2\varepsilon_{wax/w}$ can be inferred via pollen data and lipid $\delta^{13}C$ values (Polissar et al 2021) and used to calculate δ^2H_{w}. Vegetation-corrected δ^2H_w shows a 10 ‰ greater enrichment after 6 Ma than raw δ^2H_{wax} and follows the large C_4 grass expansion that is documented in $\delta^{13}C_{wax}$ (Figure 3). While this is small relative to the large amplitude of change in this particular record, in other settings where δ^2H_{wax} variance is lower, changes in $^2\varepsilon_{wax/w}$ strongly influence overall trends (Tierney et al 2017).

In addition to δ^2H_{wax}, the hydrogen isotopic composition of lignin methoxy groups ($\delta^2H_{methoxy}$) holds promise as a novel paleohydrologic indicator (Keppler et al 2007). Lignin is a key component (30–50%) of wood and other plant tissues and is a heterogeneous polymer containing a variety of monomeric building blocks. Each monomeric building block contains 0, 1 or 2 methoxy groups ($-OCH_3$), which are attractive targets for hydrogen isotope analysis because the hydrogens are in non-exchangeable positions. Methoxy groups can be released from wood (Keppler et al 2007), peat (Lee et al 2019a), and lignite (Lee et al 2019a), and spans a wide isotopic range (-325 to -150 ‰) in tree trunk samples (Keppler et al 2007). However, variability between species and trees may complicate detection of hydrologic signals, as shown in a coastal saltwater to freshwater gradient (Feakins et al 2013). Calibration and process-based understanding of drivers of biological variability in $^2\varepsilon_{methoxy/w}$ remains in the early stages, but recent analytical method development (Greule et al 2008, Lee et al 2019a) has enabled successful applications in Eocene-aged wood (Anhäuser et al 2018) and Neogene-aged sediments (Lee et al 2019b).

2.4. Terrestrial carbon cycling proxies

Carbon export from the terrestrial biosphere is a significant component of the global carbon cycle (Hilton & West 2020). One of the main unquantified processes is the lateral transport of terrestrial organic carbon (OC) along the aquatic continuum from upland terrestrial ecosystems to the ocean. Biomarker abundance (Bianchi et al 2004, Goñi et al 1997), $\delta^{13}C$ values (Feakins et al 2018, Hemingway et al 2016) and radiocarbon (^{14}C) content (Feng et al 2013, Kusch et al 2010) can be used to study changes in the storage, aging, and mobilization of different components of the terrestrial OC cycle. This may encompass plants (e.g., plant wax, lignin) or soils (e.g., brGDGTs), and can reveal the spatial and temporal scale of their integration before delivery to depositional basins (Feng et al 2013, Hemingway et al 2016, Kusch et al 2010).

Plant or soil derived biomarkers transported by rivers are often assumed to reflect a catchment-integrated signal, but the spatial and temporal integration may vary by compound type and/or change through time. Soils and rivers are typically dominated by long-chain *n*-alkanoic acids (Feakins et al., 2018) due to preferential loss of long-chain *n*-alkanes during soil litter decomposition (Wu et al., 2018). Studies also suggest that long-chain *n*-alkanoic acids respond more rapidly to proximal changes in climate, whereas long-chain *n*-alkanes have a greater spatial and temporal range and also have more risk of petrogenic input (Feakins et al 2018, Hemingway et al 2016) (see Section 2.2). In a global river dataset, Eglinton et al (2021) found that the age of exported plant wax (*n*-alkanoic acids) ranged from 1-2ka in the tropics to up to 8ka in the high latitudes. They found the strongest correlation between climatic variables (MAAT and MAP) and fluvial *n*-alkanoic acid ^{14}C ages occurred when using an *e*-folding distance of ~500 km, implying that this is roughly the spatial extent of plant wax *n*-alkanoic acid integration in large river systems. The *e*-folding distance will likely be different for other compound classes (e.g., *n*-alkanes, lignin) and requires additional investigation.

Bomb-spike ^{14}C (produced during nuclear weapons detonation and testing that began in 1945 and increased until 1963) can also be detected in OC, including plant wax biomarkers in soils and sediments. Biomarker $^{14}C/^{12}C$ ratios in soils or river sediments that are higher than modern (natural) concentrations indicate the presence of bomb-produced ^{14}C and can be used to track the mean age of biomarkers in storage or fluvial transit. Using a two-pool mixing model approach, French et al. (2018) estimates that ~80% of the *n*-alkanoic acids in the Bengal Fan have a reservoir age of ~1000 years, with the remainder stored on the landscape for only ~15 yr (French et al 2018). This approach has been extended to other river catchments (including the Mackenzie River, Saanich Inlet, and Cariaco Basin) and shows a similar distribution of ”fast-cycling” vs. “slow-cycling” *n*-alkanoic acids (Vonk et al 2019), with high-latitude (permafrost-dominated) catchments exhibiting the longest carbon storage. However, this approach requires assumptions about the age distributions of each compound class and requires further validation.

Carbon storage on land has profound implications for modern carbon cycle feedbacks and may be illuminated by the study of past climate perturbations. Pairing of δ^2H_{wax} and $^{14}C_{wax}$ (e.g., across the last deglaciation; Figure 5) provides a powerful tool to probe the relationships between climate and the carbon cycle, from the deglaciation into the Anthropocene. In the low-latitudes, there is a negative relationship between δ^2H_{wax} and mean transit times across the deglaciation (reported as $F^{14}R$, or the plant-wax ^{14}C content relative to that of the contemporaneous atmosphere at the time of deposition (Fornace 2016, Hein et al 2020, Schefuß et al 2016) (Figure 5). This implies wetter conditions are associated with shorter OC residence times in both large river systems (Hein et al 2020, Schefuß et al 2016) and in a more restricted lake catchment (Fornace 2016). The observations that large portions of the plant wax pool being “old” (i.e., 1000 yr residence times) and the inference that plant waxes are mobilized more quickly in wetter climates raises the possibility that δ^2H_{wax} records are affected by changing spatiotemporal integration in terrestrial catchments. The effect of “pre-aging” of plant wax would be to smooth, dampen, and delay the original δ^2H_{precip} signature (French et al 2018). Some δ^2H_{wax} records (e.g., the time series from the Congo Basin) have a smooth character that might imply substantial spatiotemporal averaging (Schefuß et al 2005) compared to the smaller and more abrupt changes recorded elsewhere (e.g., Lake Tanganyika; Tierney et al., 2008). However, in many cases δ^2H_{wax} captures rapid climate changes with little apparent delay, including decadal-scale historical climate shifts (Figure 4). Fornace (2016) found plant wax reservoir ages of ~8,000 years in Lake Titicaca, yet δ^2H_{wax} shows rapid changes and millennial-scale events similar to the Huascarán ice core (Figure 4), indicating that the “decadal” plant-wax pool must be very responsive to climate change. In the open ocean, but also in arid (coastal) regions, plant waxes may be exported primarily via aeolian processes (Pagani et al., 2000; Eglinton et al., 2002) (e.g., Gulf of Aden, Figure 4a). Wind transport of plant wax is able to capture seasonal variations in C_3 vs C_4 vegetation (Conte & Weber 2002) and provides another explanation for why δ^2H_{wax} captures rapid climate change.

Beyond the late Pleistocene (ca. 40 to 50 ka), ^{14}C-based dating techniques are not applicable and additional approaches are required to assess terrestrial OC cycling. The ratio of brGDGTs to crenarchaeol (i.e., the Branched-versus-Isoprenoid Tetraether (BIT) index) tracks the amount of terrestrial input in marine environments and ranges between

0 (marine-dominated) and 1 (terrestrial-dominated) (Hopmans et al 2004). The BIT index typically decreases from the river, through the estuary, and into the open shelf (Hopmans et al 2004). However, it can be controlled strongly by crenarchaeol—rather than brGDGT—concentrations, leading to a decoupling between BIT values and other terrestrial OC tracers (e.g., lignin concentrations) (Smith et al 2012). Due to these challenges, brGDGT concentrations (rather than the BIT index) may be a more robust tracer for soil OC (Smith et al 2012) and has shown promise in mountainous catchments (Kirkels et al 2020). However, branched GDGTs degrade more quickly than other soil OC tracers (e.g., long-chain *n*-alkanes, lignin phenols) (Zhu et al 2013) and caution is required when applying this approach (and/or the BIT index) to represent bulk soil OC, especially when the system is characterized by extensive OM degradation (e.g., large floodplains). We therefore argue that future studies should employ multiple proxies when attempting to trace soil OC export (e.g., lignin, *n*-alkyl lipids, brGDGTs).

2.5. Fire proxies

Reconstructing fire regimes usually relies upon the preservation of charcoal. However, charcoal abundance can be over- or under exaggerated by preservation biases (Vachula & Cheung 2021). Biomarkers offer a complementary view on the relationships between fire regime, climate, and vegetation in past terrestrial ecosystems. The most common approach involves analyzing the distribution of polycyclic aromatic hydrocarbons (PAHs), which can form rapidly at high (>300°C) temperatures (i.e., during the combustion of fossil fuels or biomass) or slowly at low (<150°C) temperatures (i.e., during petroleum formation). Previous studies have argued that PAH distributions are linked to changes in fire temperature and/or intensity (e.g., Denis et al. 2012). However, a recent meta-analysis of PAHs in natural burn experiments (Karp et al., 2020) reveals that PAH distributions are linked to burn phase, not temperature. PAHs formed in smoke are typically smaller than those formed in residues. PAH size distributions also reflect transport processes. As smaller PAHs are preferentially emitted into the aerosol phase, they can be aerially transported far from the original combustion source. In contrast, larger PAHs are less volatile and more likely to be stored in soils. This approach was employed by Lyons et al. (2020) to show that PAHs were transported long distances following the Cretaceous/Paleogene asteroid impact.

PAHs also provide insights into the fuel source: combusted vegetation yields PAHs with no alkyl substitutions, whereas fossil carbon (e.g., oil or coal) is associated with one or more alkyl substitutions. This is assessed via the methylphenanthrene to phenanthrene ratio (Yunker et al 2002), non-negative matrix factorization (Karp et al 2021), and/or the alkylated PAH derivative index (APDI; Karp et al, 2018). Positive APDI values (>10; i.e., no alkyl substitutions) imply biomass burning, whereas negative APDI values (< -10; one or more alkyl substitutions) imply a fossil fuel carbon source (Karp et al 2020). However, burned conifers can exhibit negative APDI values and may therefore resemble a fossil fuel carbon source. PAH studies reveal that Neogene fire dynamics differed between continents (Karp et al 2018, Karp et al 2021) and that fire occurrence often coincided with hydrologic shifts and C_4 expansion (Figure 3). Carbon isotopic measurements of specific PAH molecules provide additional information about the fuel source (Karp et al 2021). For example, $\delta^{13}C$ analysis of pyrene indicates an increasing proportion of C_4 grasses burning between 6 and 8 Ma, coeval with a shift in plant wax $\delta^{13}C$ to more C_4-like values (Figure 3).

Levoglucosan and its isomers (mannosan, galactosan) provide further detail about the fire regime. Levoglucosan is a thermal by-product of cellulose or hemicellulose generated during biomass burning and forms at relatively low temperatures (~100 to 400°C) (Simoneit et al 1999). This compound is relatively labile and thus geological applications may be limited to late Quaternary sediments. However, recent methodological advances have lowered the limits of detection (~5 pg) (Schreuder et al 2018) and may enable detection in older sediments and/or settings where oxygen exposure is low (e.g., anoxic peats, water-saturated soils). Benzene polycarboxylic acids (BPCAs; benzene rings with a differing number of carboxylic-acid groups) offer an additional approach and are more likely to represent the bulk pyrogenic carbon pool (Glaser et al 1998). BPCAs form at relatively high temperatures (>300-600°C, but up to 1000°C) and are produced by oxidative chemical degradation of the condensed aromatic phase. BPCA distributions are unable to differentiate between fuel sources (fossil fuel vs. biomass), but improved source apportionment in modern or Holocene-aged samples is possible using ^{14}C dating since different fuel sources (e.g., vegetation, pre-aged soils, fossil carbon) carry unique ^{14}C signatures (Coppola et al 2018). The combined analysis of PAHs, levoglucosan, and BPCAs can differentiate between low (<300°C) and high temperature fires (>300-1000°C) and thus can provide a nuanced perspective on fire dynamics (Hanke et al 2016).

2.6. Methane cycling proxies

Atmospheric methane (CH_4) is a potent greenhouse gas second only to carbon dioxide in its importance to climate change (Dean et al 2018). Higher temperatures and associated changes in rainfall patterns are expected to enhance CH_4 emissions, resulting in further warming. However, there have been few tools to test these predictions. Below we discuss biomarker-based approaches used to study two key aspects of the methane cycle: 1) methanogenesis (Section 2.6.1) and ii) methanotrophy (Section 2.6.2).

2.6.1. Methanogenesis

Methanogens thrive within water-saturated and anoxic environments (e.g., peatlands, permafrost, anoxic lakes) and synthesise a suite of diether- and/or tetraether membrane lipids (Schouten et al 2013). Archaeol (2,3-diphytanyl-O-sn-glycerol) is the most common archaeal lipid in cultured methanogens (Bauersachs et al 2015, Koga et al 1993) and shows promise as an indicator of methanogen biomass (Pancost et al 2011, Zheng et al 2014). Archaeol has been applied to reconstruct methanogen abundance in Holocene-aged peat archives and reveals a minimum in methanogenesis in China between ~6 and 4 ka (Zheng et al, 2014). This suggests that archaeol may be a useful methanogen biomarker in older (> 1 Ma) sediments. However, direct estimates of methanogen biomass from archaeol should be approached with caution as there can be differing concentrations of archaeol per methanogen cell (McCartney et al 2013).The acyclic isoGDGT (i.e., isoGDGT-0) is also abundant in methanogens (Bauersachs et al 2015, Koga et al 1993, Schouten et al 2013) and may provide complementary insights into methanogenesis (e.g., peatlands, permafrost, lakes).

Both isoGDGT-0 and archaeol have diverse source organisms, perhaps limiting their utility as methanogen biomarkers. In contrast, a sub-set of methanogens (e.g., Methanococcus, Methanosarcina) are known to synthesise *sn*-2-hydroxyarchaeol (Koga et al 1993); this is structurally similar to archaeol but contains a hydroxyl group at the C-3 position of the *sn*-2 phytanyl chain (Hinrichs et al 2000). Due to the labile nature of sn-2-hydroxyarchaeol, this biomarker holds promise as a marker for living methanogen biomass (Pancost et al 2011). There is also growing evidence that methanogens may synthesise unusual butanetriol and pentanetriol dibiphytanyl glycerol tetraethers (BDGTs and PDGTs, respectively). BDGTs and PDGTs have been identified in different environments—including wetlands (Blewett et al 2020) —and have been assigned to the methanogen order Methanomassiliicoccales (Becker et al 2016). Analogous to other archaeal lipids (e.g., archaeol, GDGT-0), BDGTs and PDGTs in wetlands increase in concentration below the anoxic layer and are nearly absent from oxygenated layers. This is consistent with observations that only methanogens synthesize BDGTs and PDGTs (Becker et al 2016). These compounds—alongside other 'minor' GDGTs (Bauersachs et al 2015)—are promising methanogen-specific biomarkers that warrant further study.

2.6.2. Methanotrophy

Microbes capable of consuming methane are known as methanotrophs. Aerobic methanotrophs can synthesise diagnostic hopanoids, including 35-aminobacteriohopanepentol (i.e., aminopentol) and 35-aminobacteriohopanetetrol (i.e., aminotetrol) (Rush et al 2016, Talbot & Farrimond 2007). Both compounds are present in terrestrial environments, especially peats (Van Winden et al 2012) and lakes (Talbot &

Farrimond 2007), and can persist in the sedimentary record for >50 Ma under favorable conditions (e.g., anoxia). Both compounds have been used to reconstruct aerobic methanotrophy during the Quaternary (Talbot et al 2014) and early Eocene (Talbot et al 2016), but typically lose their functionalized amino side group during diagenesis.

Analysis of hopanoid $\delta^{13}C$ provides an alternative means to reconstruct methanotrophy. Hopanoid $\delta^{13}C$ is primarily determined by the $\delta^{13}C$ of the substrate, carbon assimilation pathways and an organism's source ecology (Hayes 1993, Pancost & Sinninghe Damsté 2003). Heterotrophic organisms consuming organic substrates (e.g., organic acids, sugars) will typically yield hopanoid $\delta^{13}C$ values which are similar to the food source (ca. -20 to -30 ‰), whereas methanotrophic organisms (consuming methane) will yield hopanoid $\delta^{13}C$ values which are lower (ca. -40 to -60 ‰, but up to -100 ‰) and more variable (Pancost & Damsté 2003). A recent survey of hopanoid $\delta^{13}C$ values in peatlands from different geographic regions shows that the $\delta^{13}C$ composition of the C_{31} hopane—one of the most abundant hopanoids in peats—spans a relatively narrow range (−22 to −32 ‰ VPDB) and is ^{13}C-enriched relative to bulk organic matter and co-occurring plant wax biomarkers (Inglis et al 2019b). This suggests the C_{31} hopane is derived from a heterotrophic source organism(s) and has limited utility as a methanotroph biomarker. In contrast, other hopanoids (e.g., hop-22(29)-ene) yields lower $\delta^{13}C$ values (e.g., up to −45 ‰) (Inglis et al 2019b), suggesting that in some settings this compound is derived from a mixed suite of bacterial sources consuming both ^{13}C-enriched carbohydrates and ^{13}C-depleted, methane-derived CO_2. Low hopanoid $\delta^{13}C$ values (ca. -60 to -80 ‰) have also been found in modern (e.g., Naeher et al 2014), Holocene (e.g., Elvert et al 2016, Naeher et al 2014), and Eocene lakes (e.g., Collister et al 1992, Freeman et al 1990). These values are considerably lower than in wetlands and indicate more vigorous methane consumption in lakes.

Hopanoid $\delta^{13}C$ values have successfully been applied as a proxy for methanotrophy across the Paleocene-Eocene Thermal Maximum (PETM). Pancost et al (2007) first observed a sudden decrease in hopanoid $\delta^{13}C$ values (to -75 ‰) in the Cobham lignite, UK, an ancient wetland deposited during the onset of the PETM (Figure 6). These low values lie outside the modern range (Inglis et al 2019b) and coincided with an increase in methanotroph (e.g., aminoBHPs) (Talbot et al 2016) and methanogen biomarkers (GDGT-0) (Inglis et al 2019a). This implies enhanced methane cycling at this site during the PETM. Recent work provides additional evidence that enhanced methanotrophy occurred the onset of the PETM, with low hopanoid $\delta^{13}C$ values observed in PETM-aged terrestrial/marine deposits from New Zealand (up to -60 ‰) (Inglis et al 2021). Hopanoid $\delta^{13}C$ values rapidly returned to pre-event values even though brGDGT-derived temperatures remain high for the duration of the PETM (Inglis et al, 2021) (Figure 6). This suggests it is the onset of rapid global warming that is particularly disruptive to methane cycling in wetlands, a finding that is particularly concerning given the rapid global warming we are currently experiencing.

Methane can also be oxidised anaerobically (anaerobic oxidation of methane; AOM) by a consortium of anaerobic methanotrophic archaea and sulfate reducing bacteria (SRB) (Hinrichs et al., 1999). Anaerobic methanotrophs produce diagnostic isoGDGT distributions, with a high abundance of isoGDGT-0 to -3 relative to crenarchaeol (Pancost et al 2000). This is captured in the Methane Index (Zhang et al 2011), whereby high values (> 0.5) suggest extensive anaerobic methanotrophy. This ratio is mainly used in marine settings but has utility in terrestrial settings where AOM is elevated (e.g., freshwater wetlands). To confirm the presence of AOM, other biomarkers can be analysed (e.g., ^{13}C-depleted pentamethylicosane and/or crocetane)

Summary Points:

1. **Temperature** – Bacterial-derived branched GDGTs provide insights into terrestrial temperatures up to the maximum theoretical limit of present calibrations (~30°C). Mesocosm experiments and new GDGT proxies may offer insights into higher terrestrial temperatures observed in the geological past.

2. **Plant ecosystems** – Lignin and plant wax molecular compositions and their stable carbon isotopic composition ($\delta^{13}C$) carry signals of plant community composition and paleoenvironmental change. Machine learning of higher plant biomarkers offers promise in terms of reconstructing ecosystem turnover.

3. **Hydroclimate –** Recent advances, including tandem reconstruction of ecosystem change with pollen or plant wax $\delta^{13}C$ and paired analysis of isotope-enabled model simulations, have improved constraints on the interpretation of plant wax $\delta^{2}H$

values in the geological record. Lignin methoxy δ^2H values similarly records the hydrological cycle, but the proxy is in comparatively early stages of development.

4. **Carbon cycling** – Plant - and soil-derived biomarkers provide insights into the storage, aging, and mobilization of different terrestrial OC pools along the terrestrial-aquatic continuum. Pairing of plant wax δ^2H and ^{14}C provides a powerful tool to probe the relationships between climate and the carbon cycle during the late Quaternary.

5. **Fire regime** - The analysis of multiple fire biomarkers provides a holistic perspective on fire dynamics and can reveal the temperature history, fuel source and burn conditions.

6. **Methane cycling** – The δ^{13}C of bacterial hopanoids provides unique (qualitative) insights into methanotrophy throughout the geological record. Coupling this approach alongside archaeal biomarkers for methanogenesis may corroborate detection of changes in the terrestrial methane cycle.

Acknowledgments: GNI is supported by a GCRF Royal Society Dorothy Hodgkin Fellowship (DHF\R1\191178). TB is supported by NSF P2C2 Grant Numbers OCE 1903148 and OCE 2103015 and MRI Grant EAR 2018078. SF is supported by NSF-OPP-1908548. We thank Kate Freeman and one anonymous reviewer whose constructive feedback improved this review. We are grateful to the organic geochemical community whose innovative work inspired this review.

Literature Cited

Anhäuser T, Hook BA, Halfar J, Greule M, Keppler F. 2018. Earliest Eocene cold period and polar amplification-Insights from δ2H values of lignin methoxyl groups of mummified wood. *Palaeogeography, Palaeoclimatology, Palaeoecology* 505: 326-36

Balascio NL, D'Andrea WJ, Anderson RS, Wickler S. 2018. Influence of vegetation type on n-alkane composition and hydrogen isotope values from a high latitude ombrotrophic bog. *Organic Geochemistry* 121: 48-57

Bauersachs T, Weidenbach K, Schmitz RA, Schwark L. 2015. Distribution of glycerol ether lipids in halophilic, methanogenic and hyperthermophilic archaea. *Organic Geochemistry* 83: 101-8

Becker KW, Elling FJ, Yoshinaga MY, Söllinger A, Urich T, Hinrichs K-U. 2016. Unusual butane-and pentanetriol-based tetraether lipids in Methanomassiliicoccus luminyensis, a representative of the seventh order of methanogens. *Applied and Environmental Microbiology* 82: 4505-16

Becker KW, Lipp JS, Zhu C, Liu X-L, Hinrichs K-U. 2013. An improved method for the analysis of archaeal and bacterial ether core lipids. *Organic Geochemistry* 61: 34-44

Bhattacharya T, Tierney JE, Addison JA, Murray JW. 2018. Ice-sheet modulation of deglacial North American monsoon intensification. *Nature Geoscience* 11: 848-52

Bianchi TS, Filley T, Dria K, Hatcher PG. 2004. Temporal variability in sources of dissolved organic carbon in the lower Mississippi River. *Geochimica et Cosmochimica Acta* 68: 959-67

Blaga CI, Reichart G-J, Heiri O, Damsté JSS. 2009. Tetraether membrane lipid distributions in water-column particulate matter and sediments: a study of 47 European lakes along a north–south transect. *Journal of Paleolimnology* 41: 523-40

Blewett J, Naafs B, Gallego-Sala A, Pancost RD. 2020. Effects of temperature and pH on archaeal membrane lipid distributions in freshwater wetlands. *Organic Geochemistry* 148: 104080

Bush RT, McInerney FA. 2013. Leaf wax n-alkane distributions in and across modern plants: implications for paleoecology and chemotaxonomy. *Geochimica et Cosmochimica Acta* 117: 161-79

Cerling TE. 1984. The stable isotopic composition of modern soil carbonate and its relationship to climate. *Earth and Planetary science letters* 71: 229-40

Cerling TE, Wynn JG, Andanje SA, Bird MI, Korir DK, et al. 2011. Woody cover and hominin environments in the past 6 million years. *Nature* 476: 51-6

Chen Y, Zheng F, Chen S, Liu H, Phelps TJ, Zhang C. 2018. Branched GDGT production at elevated temperatures in anaerobic soil microcosm incubations. *Organic Geochemistry* 117: 12-21

Coffinet S, Huguet A, Williamson D, Fosse C, Derenne S. 2014. Potential of GDGTs as a temperature proxy along an altitudinal transect at Mount Rungwe (Tanzania). *Organic Geochemistry* 68: 82-9

Collister JW, Summons RE, Lichtfouse E, Hayes JM. 1992. An isotopic biogeochemical study of the Green River oil shale. *Organic Geochemistry* 19: 265-76

Conte MH, Weber JC. 2002. Plant biomarkers in aerosols record isotopic discrimination of terrestrial photosynthesis. *Nature* 417: 639-41

Coppola AI, Wiedemeier DB, Galy V, Haghipour N, Hanke UM, et al. 2018. Global-scale evidence for the refractory nature of riverine black carbon. *Nature Geoscience* 11: 584-8

Crampton-Flood ED, Peterse F, Munsterman D, Damsté JSS. 2018. Using tetraether lipids archived in North Sea Basin sediments to extract North Western European Pliocene continental air temperatures. *Earth and Planetary Science Letters* 490: 193-205

Crampton-Flood ED, Tierney JE, Peterse F, Kirkels FM, Damsté JSS. 2020. BayMBT: A Bayesian calibration model for branched glycerol dialkyl glycerol tetraethers in soils and peats. *Geochimica et Cosmochimica Acta* 268: 142-59

Dang X, Yang H, Naafs BDA, Pancost RD, Xie S. 2016. Evidence of moisture control on the methylation of branched glycerol dialkyl glycerol tetraethers in semi-arid and arid soils. *Geochimica et Cosmochimica Acta*

Daniels WC, Russell JM, Giblin AE, Welker JM, Klein ES, Huang Y. 2017. Hydrogen isotope fractionation in leaf waxes in the Alaskan Arctic tundra. *Geochimica et Cosmochimica Acta* 213: 216-36

De Jonge C, Hopmans EC, Stadnitskaia A, Rijpstra WIC, Hofland R, et al. 2013. Identification of novel penta- and hexamethylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC–MS2, GC–MS and GC–SMB-MS. *Organic Geochemistry* 54: 78-82

De Jonge C, Hopmans EC, Zell CI, Kim J-H, Schouten S, Damsté JSS. 2014. Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate reconstruction. *Geochimica et Cosmochimica Acta* 141: 97-112

De Rosa M, Esposito E, Gambacorta A, Nicolaus B, Bu'Lock JD. 1980. Effects of temperature on ether lipid composition of Caldariella acidophila. *Phytochemistry* 19: 827-31

Dean JF, Middelburg JJ, Röckmann T, Aerts R, Blauw LG, et al. 2018. Methane feedbacks to the global climate system in a warmer world. *Reviews of Geophysics* 56: 207-50

Diefendorf, A.F., Sberna, D.T. and Taylor, D.W., 2015. Effect of thermal maturation on plant-derived terpenoids and leaf wax n-alkyl components. *Organic Geochemistry*, *89*, pp.61-70.

Diefendorf AF, Freeman KH, Wing SL. 2012. Distribution and carbon isotope patterns of diterpenoids and triterpenoids in modern temperate C3 trees and their geochemical significance. *Geochimica et Cosmochimica Acta* 85: 342-56

Diefendorf AF, Freeman KH, Wing SL. 2014. A comparison of terpenoid and leaf fossil vegetation proxies in Paleocene and Eocene Bighorn Basin sediments. *Organic Geochemistry* 71: 30-42

Diefendorf AF, Freeman KH, Wing SL, Graham HV. 2011. Production of n-alkyl lipids in living plants and implications for the geologic past. *Geochimica et Cosmochimica Acta* 75: 7472-85

Diefendorf AF, Freimuth EJ. 2017. Extracting the most from terrestrial plant-derived n-alkyl lipids and their carbon isotopes from the sedimentary record: A review. *Organic Geochemistry* 103: 1-21

Denis, E.H., Toney, J.L., Tarozo, R., Anderson, R.S., Roach, L.D. and Huang, Y., 2012. Polycyclic aromatic hydrocarbons (PAHs) in lake sediments record historic fire events: Validation using HPLC-fluorescence detection. *Organic Geochemistry*, *45*, 7-17.

Eglinton G, Hamilton RJ. 1967. Leaf epicuticular waxes. *Science* 156: 1322-35

Eglinton TI, Galy VV, Hemingway JD, Feng X, Bao H, et al. 2021. Climate control on terrestrial biospheric carbon turnover. *Proceedings of the National Academy of Sciences* 118

Eglinton, T.I., Eglinton, G., Dupont, L., Sholkovitz, E.R., Montluçon, D. and Reddy, C.M., 2002. Composition, age, and provenance of organic matter in NW African dust over the Atlantic Ocean. *Geochemistry, Geophysics, Geosystems*, *3*, 1-27.

Elvert M, Pohlman JW, Becker KW, Gaglioti B, Hinrichs K-U, Wooller MJ. 2016. Methane turnover and environmental change from Holocene lipid biomarker records in a thermokarst lake in Arctic Alaska. *The Holocene* 26: 1766-77

Fan Y, Miguez-Macho G, Jobbágy EG, Jackson RB, Otero-Casal C. 2017. Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences* 114: 10572-7

Feakins SJ. 2013. Pollen-corrected leaf wax D/H reconstructions of northeast African hydrological changes during the late Miocene. *Palaeogeography, Palaeoclimatology, Palaeoecology* 374: 62-71

Feakins SJ, Bentley LP, Salinas N, Shenkin A, Blonder B, et al. 2016a. Plant leaf wax biomarkers capture gradients in hydrogen isotopes of precipitation from the Andes and Amazon. *Geochimica et Cosmochimica Acta* 182: 155-72

Feakins SJ, Ellsworth PV, Sternberg LdSL. 2013. Lignin methoxyl hydrogen isotope ratios in a coastal ecosystem. *Geochimica et cosmochimica acta* 121: 54-66

Feakins SJ, Peters T, Wu MS, Shenkin A, Salinas N, et al. 2016b. Production of leaf wax n-alkanes across a tropical forest elevation transect. *Organic Geochemistry* 100: 89-100

Feakins SJ, Sessions AL. 2010. Controls on the D/H ratios of plant leaf waxes in an arid ecosystem. *Geochimica et Cosmochimica Acta* 74: 2128-41

Feakins SJ, Wu MS, Ponton C, Galy V, West AJ. 2018. Dual isotope evidence for sedimentary integration of plant wax biomarkers across an Andes-Amazon elevation transect. *Geochimica et Cosmochimica Acta* 242: 64-81

Feng X, Vonk JE, Van Dongen BE, Gustafsson Ö, Semiletov IP, et al. 2013. Differential mobilization of terrestrial carbon pools in Eurasian Arctic river basins. *Proceedings of the National Academy of Sciences* 110: 14168-73

Fornace KL. 2016. *Late Quaternary climate variability and terrestrial carbon cycling in tropical South America*. Massachusetts Institute of Technology

Fornace KL, Hughen KA, Shanahan TM, Fritz SC, Baker PA, Sylva SP. 2014. A 60,000-year record of hydrologic variability in the Central Andes from the hydrogen isotopic composition of leaf waxes in Lake Titicaca sediments. *Earth and Planetary Science Letters* 408: 263-71

Freeman KH, Colarusso L. 2001. Molecular and isotopic records of C4 grassland expansion in the late Miocene. *Geochimica et Cosmochimica Acta* 65: 1439-54

Freeman KH, Hayes J, Trendel J-M, Albrecht P. 1990. Evidence from carbon isotope measurements for diverse origins of sedimentary hydrocarbons. *Nature* 343: 254-6

French KL, Hein CJ, Haghipour N, Wacker L, Kudrass HR, et al. 2018. Millennial soil retention of terrestrial organic matter deposited in the Bengal Fan. *Scientific reports* 8: 1-8

Gamarra B, Sachse D, Kahmen A. 2016. Effects of leaf water evaporative 2H-enrichment and biosynthetic fractionation on leaf wax n-alkane δ2H values in C3 and C4 grasses. *Plant, Cell & Environment* 39: 2390-403

Gao L, Edwards EJ, Zeng Y, Huang Y. 2014. Major evolutionary trends in hydrogen isotope fractionation of vascular plant leaf waxes. *PloS one* 9: e112610

Gao L, Hou J, Toney J, MacDonald D, Huang Y. 2011. Mathematical modeling of the aquatic macrophyte inputs of mid-chain n-alkyl lipids to lake sediments: Implications for interpreting compound specific hydrogen isotopic records. *Geochimica et Cosmochimica Acta* 75: 3781-91

Garcin Y, Schwab VF, Gleixner G, Kahmen A, Todou G, et al. 2012. Hydrogen isotope ratios of lacustrine sedimentary n-alkanes as proxies of tropical African hydrology: insights from a calibration transect across Cameroon. *Geochimica et Cosmochimica Acta* 79: 106-26

Glaser B, Haumaier L, Guggenberger G, Zech W. 1998. Black carbon in soils: the use of benzenecarboxylic acids as specific markers. *Organic geochemistry* 29: 811-9

Goñi MA, Ruttenberg KC, Eglinton TI. 1997. Sources and contribution of terrigenous organic carbon to surface sediments in the Gulf of Mexico. *Nature* 389: 275-8

Greule M, Mosandl A, Hamilton JT, Keppler F. 2008. A rapid and precise method for determination of D/H ratios of plant methoxyl groups. *Rapid Communications in Mass Spectrometry: An International Journal Devoted to the Rapid Dissemination of Up‐to‐the‐Minute Research in Mass Spectrometry* 22: 3983-8

Griepentrog M, De Wispelaere L, Bauters M, Bodé S, Hemp A, et al. 2019. Influence of plant growth form, habitat and season on leaf-wax n-alkane hydrogen-isotopic signatures in equatorial East Africa. *Geochimica et Cosmochimica Acta* 263: 122-39

Halamka, T.A., McFarlin, J., Younkin, A.D., Depoy, J., Dildar, N., and Kopf, S.H. 2021. Oxygen limitation can trigger the production of branched GDGTs in culture. Geochemical Perspectives.19. 36-39. doi: 10.7185/geochemlet.2132

Hanke UM, Eglinton TI, Braun AL, Reddy CM, Wiedemeier DB, Schmidt MW. 2016. Decoupled sedimentary records of combustion: Causes and implications. *Geophysical Research Letters* 43: 5098-108

Hayes JJMG. 1993. Factors controlling 13C contents of sedimentary organic compounds: principles and evidence. 113: 111-25

Hedges JI, Mann DC. 1979. The characterization of plant tissues by their lignin oxidation products. *Geochimica et Cosmochimica Acta* 43: 1803-7

Hein CJ, Usman M, Eglinton TI, Haghipour N, Galy VV. 2020. Millennial-scale hydroclimate control of tropical soil carbon storage. *Nature* 581: 63-6

Hemingway JD, Schefuß E, Dinga BJ, Pryer H, Galy VV. 2016. Multiple plant-wax compounds record differential sources and ecosystem structure in large river catchments. *Geochimica et Cosmochimica Acta* 184: 20-40

Hilton RG, West AJ. 2020. Mountains, erosion and the carbon cycle. *Nature Reviews Earth & Environment* 1: 284-99

Hinrichs K-U, Hayes JM, Sylva SP, Brewer PG, DeLong EF. 1999. Methane-consuming archaebacteria in marine sediments. *Nature* 398: 802-5

Hinrichs KU, Pancost RD, Summons RE, Sprott GD, Sylva SP, et al. 2000. Mass spectra of sn-2-hydroxyarchaeol, a polar lipid biomarker for anaerobic methanotrophy. *Geochemistry, Geophysics, Geosystems* 1

Hopmans EC, Schouten S, Damsté JSS. 2016. The effect of improved chromatography on GDGT-based palaeoproxies. *Organic Geochemistry* 93: 1-6

Hopmans EC, Weijers JW, Schefuß E, Herfort L, Damsté JSS, Schouten S. 2004. A novel proxy for terrestrial organic matter in sediments based on branched and isoprenoid tetraether lipids. *Earth and Planetary Science Letters* 224: 107-16

Huang Y, Street-Perrott FA, Metcalfe SE, Brenner M, Moreland M, Freeman KH. 2001. Climate change as the dominant control on glacial-interglacial variations in C3 and C4 plant abundance. *Science* 293: 1647-51

Inglis GN, Carmichael MJ, Farnsworth A, Lunt DJ, Pancost RD. 2020. A long-term, high-latitude record of Eocene hydrological change in the Greenland region. *Palaeogeography, Palaeoclimatology, Palaeoecology* 537: 109378

Inglis GN, Farnsworth A, Collinson ME, Carmichael MJ, Naafs BDA, et al. 2019a. Terrestrial environmental change across the onset of the PETM and the associated impact on biomarker proxies: A cautionary tale. *Global and Planetary Change* 181: 102991

Inglis GN, Naafs BDA, Zheng Y, Schellekens J, Pancost RD. 2019b. δ13C values of bacterial hopanoids and leaf waxes as tracers for methanotrophy in peatlands. *Geochimica et Cosmochimica Acta* 260: 244-56

Inglis GN, Rohrssen M, Kennedy EM, Crouch EM, Raine JI, et al. 2021. Terrestrial methane cycle perturbations during the onset of the Paleocene-Eocene Thermal Maximum. *Geology* 49: 520-4

Kahmen A, Hoffmann B, Schefuß E, Arndt SK, Cernusak LA, et al. 2013a. Leaf water deuterium enrichment shapes leaf wax n-alkane δD values of angiosperm plants II: Observational evidence and global implications. *Geochimica et Cosmochimica Acta* 111: 50-63

Kahmen A, Schefuß E, Sachse D. 2013b. Leaf water deuterium enrichment shapes leaf wax n-alkane δD values of angiosperm plants I: Experimental evidence and mechanistic insights. *Geochimica et Cosmochimica Acta* 111: 39-49

Karp AT, Behrensmeyer AK, Freeman KH. 2018. Grassland fire ecology has roots in the late Miocene. *Proceedings of the National Academy of Sciences* 115: 12130-5

Karp AT, Holman AI, Hopper P, Grice K, Freeman KH. 2020. Fire distinguishers: Refined interpretations of polycyclic aromatic hydrocarbons for paleo-applications. *Geochimica et Cosmochimica Acta* 289: 93-113

Karp AT, Uno KT, Polissar PJ, Freeman KH. 2021. Late Miocene C4 grassland fire feedbacks on the Indian Subcontinent. *Paleoceanography and Paleoclimatology* 36: e2020PA004106

Kaufman D, McKay N, Routson C, Erb M, Davis B, et al. 2020. A global database of Holocene paleotemperature records. *Scientific data* 7: 1-34

Keppler F, Harper DB, Kalin RM, Meier-Augenstein W, Farmer N, et al. 2007. Stable hydrogen isotope ratios of lignin methoxyl groups as a paleoclimate proxy and constraint of the geographical origin of wood. *New Phytologist* 176: 600-9

Kirkels FM, Ponton C, Galy V, West AJ, Feakins SJ, Peterse F. 2020. From Andes to Amazon: Assessing branched tetraether lipids as tracers for soil organic carbon in the Madre de Dios River system. *Journal of Geophysical Research: Biogeosciences* 125: e2019JG005270

Koga Y, Nishihara M, Morii H, Akagawa-Matsushita M. 1993. Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses. *Microbiological Reviews* 57: 164-82

Konecky B, Noone D, Cobb K. 2019. The influence of competing hydroclimate processes on stable isotope ratios in tropical rainfall. *Geophysical Research Letters* 46: 1622-33

Kusch S, Rethemeyer J, Schefuß E, Mollenhauer G. 2010. Controls on the age of vascular plant biomarkers in Black Sea sediments. *Geochimica et Cosmochimica Acta* 74: 7031-47

Lee H, Feng X, Mastalerz M, Feakins SJ. 2019a. Characterizing lignin: Combining lignin phenol, methoxy quantification, and dual stable carbon and hydrogen isotopic techniques. *Organic Geochemistry* 136: 103894

Lee H, Galy V, Feng X, Ponton C, Galy A, et al. 2019b. Sustained wood burial in the Bengal Fan over the last 19 My. *Proceedings of the National Academy of Sciences* 116: 22518-25

Lee JE, Fung I, DePaolo DJ, Henning CC. 2007. Analysis of the global distribution of water isotopes using the NCAR atmospheric general circulation model. *Journal of Geophysical Research: Atmospheres* 112

Luo Y. 2007. Terrestrial carbon–cycle feedback to climate warming. *Annu. Rev. Ecol. Evol. Syst.* 38: 683-712

Lyons, S.L., Karp, A.T., Bralower, T.J., Grice, K., Schaefer, B., Gulick, S.P., Morgan, J.V. and Freeman, K.H., 2020. Organic matter from the Chicxulub crater exacerbated the K–Pg impact winter. *Proceedings of the National Academy of Sciences*, 117, 25327-25334.

Magill CR, Ashley GM, Freeman KH. 2013. Ecosystem variability and early human habitats in eastern Africa. *Proceedings of the National Academy of Sciences* 110: 1167-74

Martinez-Sosa P, Tierney JE, Meredith LK. 2020. Controlled lacustrine microcosms show a brGDGT response to environmental perturbations. *Organic Geochemistry* 145: 104041

Martínez-Sosa P, Tierney JE, Stefanescu IC, Crampton-Flood ED, Shuman BN, Routson C. 2021. A global Bayesian temperature calibration for lacustrine brGDGTs. *Geochimica et Cosmochimica Acta* 305: 87-105

McCartney C, Bull I, Yan T, Dewhurst R. 2013. Assessment of archaeol as a molecular proxy for methane production in cattle. *Journal of dairy science* 96: 1211-7

Mügler I, Sachse D, Werner M, Xu B, Wu G, et al. 2008. Effect of lake evaporation on δD values of lacustrine n-alkanes: A comparison of Nam Co (Tibetan Plateau) and Holzmaar (Germany). *Organic Geochemistry* 39: 711-29

Naafs B, Rohrssen M, Inglis G, Lähteenoja O, Feakins S, et al. 2018. High temperatures in the terrestrial mid-latitudes during the early Palaeogene. *Nature Geoscience* 11: 766-71

Naafs BDA, Inglis GN, Zheng Y, Amesbury M, Biester H, et al. 2017. Introducing global peat-specific temperature and pH calibrations based on brGDGT bacterial lipids. *Geochimica et Cosmochimica Acta* 208: 285-301

Naeher S, Niemann H, Peterse F, Smittenberg RH, Zigah PK, Schubert CJ. 2014. Tracing the methane cycle with lipid biomarkers in Lake Rotsee (Switzerland). *Organic geochemistry* 66: 174-81

Nott CJ, Xie S, Avsejs LA, Maddy D, Chambers FM, Evershed RP. 2000. n-Alkane distributions in ombrotrophic mires as indicators of vegetation change related to climatic variation. *Organic Geochemistry* 31: 231-5

Otto A, Simoneit BR. 2001. Chemosystematics and diagenesis of terpenoids in fossil conifer species and sediment from the Eocene Zeitz formation, Saxony, Germany. *Geochimica et Cosmochimica Acta* 65: 3505-27

Pagani, M., Freeman, K.H. and Arthur, M.A., 2000. Isotope analyses of molecular and total organic carbon from Miocene sediments. *Geochimica et Cosmochimica Acta*, *64*, 37-49.

Pancost RD, Damsté JSS. 2003. Carbon isotopic compositions of prokaryotic lipids as tracers of carbon cycling in diverse settings. *Chemical Geology* 195: 29-58

Pancost RD, McClymont EL, Bingham EM, Roberts Z, Charman DJ, et al. 2011. Archaeol as a methanogen biomarker in ombrotrophic bogs. *Organic Geochemistry* 42: 1279-87

Pancost RD, Sinninghe Damsté JS. 2003. Carbon isotopic compositions of prokaryotic lipids as tracers of carbon cycling in diverse settings. *Chemical Geology* 195: 29-58

Pancost RD, Sinninghe Damsté JS, de Lint S, van der Maarel MJ, Gottschal JC, Party‡ MSS. 2000. Biomarker evidence for widespread anaerobic methane oxidation in Mediterranean sediments by a consortium of methanogenic archaea and bacteria. *Applied and environmental microbiology* 66: 1126-32

Pancost RD, Steart DS, Handley L, Collinson ME, Hooker JJ, et al. 2007. Increased terrestrial methane cycling at the Palaeocene–Eocene thermal maximum. *Nature* 449: 332-5

Peaple MD, Tierney JE, McGee D, Lowenstein TK, Bhattacharya T, Feakins SJ. 2021. Identifying plant wax inputs in lake sediments using machine learning. *Organic Geochemistry* 156: 104222

Polissar PJ, Rose C, Uno KT, Phelps SR, deMenocal P. 2019. Synchronous rise of African C 4 ecosystems 10 million years ago in the absence of aridification. *Nature Geoscience* 12: 657-60

Polissar, P.J. and Freeman, K.H., 2010. Effects of aridity and vegetation on plant-wax δD in modern lake sediments. *Geochimica et Cosmochimica Acta*, 74, 5785-5797.

Polissar PJ, Uno KT, Phelps SR, Karp AT, Freeman KH, Pensky JL. 2021. Hydrologic Changes Drove the Late Miocene Expansion of C4 Grasslands on the Northern Indian Subcontinent. *Paleoceanography and Paleoclimatology* 36: e2020PA004108

Powers LA, Werne JP, Johnson TC, Hopmans EC, Damsté JSS, Schouten S. 2004. Crenarchaeotal membrane lipids in lake sediments: a new paleotemperature proxy for continental paleoclimate reconstruction? *Geology* 32: 613-6

Rach O, Brauer A, Wilkes H, Sachse D. 2014. Delayed hydrological response to Greenland cooling at the onset of the Younger Dryas in western Europe. *Nature Geoscience* 7: 109-12

Rohling EJ, Sluijs A, Dijkstra HA, Köhler P, van de Wal RS, et al. 2012. Making sense of palaeoclimate sensitivity. *Nature* 491: 683

Rush D, Osborne KA, Birgel D, Kappler A, Hirayama H, et al. 2016. The bacteriohopanepolyol inventory of novel aerobic methane oxidising bacteria reveals new biomarker signatures of aerobic methanotrophy in marine systems. *PLoS One* 11: e0165635

Sachse D, Billault I, Bowen GJ, Chikaraishi Y, Dawson TE, et al. 2012. Molecular paleohydrology: interpreting the hydrogen-isotopic composition of lipid biomarkers from photosynthesizing organisms. *Annual Review of Earth and Planetary Sciences* 40: 221-49

Sachse D, Radke J, Gleixner G. 2004. Hydrogen isotope ratios of recent lacustrine sedimentary n-alkanes record modern climate variability. *Geochimica et Cosmochimica Acta* 68: 4877-89

Schefuß E, Eglinton TI, Spencer-Jones CL, Rullkötter J, De Pol-Holz R, et al. 2016. Hydrologic control of carbon cycling and aged carbon discharge in the Congo River basin. *Nature Geoscience* 9: 687-90

Schefuß, E., Ratmeyer, V., Stuut, J.B.W., Jansen, J.F. and Damsté, J.S.S., 2003. Carbon isotope analyses of n-alkanes in dust from the lower atmosphere over the central eastern Atlantic. *Geochimica et Cosmochimica Acta*, *67*, 1757-1767.

Schefuß E, Schouten S, Schneider RR. 2005. Climatic controls on central African hydrology during the past 20,000 years. *Nature* 437: 1003-6

Schmidt GA, LeGrande AN, Hoffmann G. 2007. Water isotope expressions of intrinsic and forced variability in a coupled ocean-atmosphere model. *Journal of Geophysical Research: Atmospheres* 112

Schouten S, Hopmans EC, Damsté JSS. 2013. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. *Organic geochemistry* 54: 19-61

Schouten S, Hopmans EC, Schefuß E, Damste JSS. 2002. Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures? *Earth and Planetary Science Letters* 204: 265-74

Schreuder LT, Hopmans EC, Stuut J-BW, Damsté JSS, Schouten S. 2018. Transport and deposition of the fire biomarker levoglucosan across the tropical North Atlantic Ocean. *Geochimica et Cosmochimica Acta* 227: 171-85

Sessions AL. 2016. Factors controlling the deuterium contents of sedimentary hydrocarbons. *Organic Geochemistry* 96: 43-64

Simoneit BR, Schauer JJ, Nolte C, Oros DR, Elias VO, et al. 1999. Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. *Atmospheric Environment* 33: 173-82

Sinninghe Damsté JS, Hopmans EC, Pancost RD, Schouten S, Geenevasen JA. 2000. Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments. *Chemical Communications*: 1683-4

Sinninghe Damsté JS, Rijpstra WIC, Foesel BU, Huber KJ, Overmann J, et al. 2018. An overview of the occurrence of ether-and ester-linked iso-diabolic acid membrane lipids in microbial cultures of the Acidobacteria: Implications for brGDGT paleoproxies for temperature and pH. *Organic Geochemistry* 124: 63-76

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Weijers, J.W., Foesel, B.U., Overmann, J. and Dedysh, S.N., 2011. 13, 16-Dimethyl octacosanedioic acid (iso-diabolic acid), a common membrane-spanning lipid of Acidobacteria subdivisions 1 and 3. *Applied and environmental microbiology*, *77*(12), pp.4147-4154.

Smith FA, Freeman KH. 2006. Influence of physiology and climate on δD of leaf wax n-alkanes from C3 and C4 grasses. *Geochimica et Cosmochimica Acta* 70: 1172-87

Smith RW, Bianchi TS, Li X. 2012. A re-evaluation of the use of branched GDGTs as terrestrial biomarkers: Implications for the BIT Index. *Geochimica et Cosmochimica Acta* 80: 14-29

Stogiannidis E, Laane R. 2015. Source characterization of polycyclic aromatic hydrocarbons by using their molecular indices: an overview of possibilities. *Reviews of environmental contamination and toxicology*: 49-133

Talbot HM, Bischoff J, Inglis GN, Collinson ME, Pancost RD. 2016. Polyfunctionalised bio-and geohopanoids in the Eocene Cobham Lignite. *Organic Geochemistry* 96: 77-92

Talbot HM, Farrimond P. 2007. Bacterial populations recorded in diverse sedimentary biohopanoid distributions. *Organic Geochemistry* 38: 1212-25

Talbot HM, Handley L, Spencer-Jones CL, Dinga BJ, Schefuß E, et al. 2014. Variability in aerobic methane oxidation over the past 1.2 Myrs recorded in microbial biomarker signatures from Congo fan sediments. *Geochimica et Cosmochimica Acta* 133: 387-401

Tibbett EJ, Scher HD, Warny S, Tierney JE, Passchier S, Feakins SJ. 2021. Late Eocene record of hydrology and temperature from Prydz Bay, East Antarctica. *Paleoceanography and Paleoclimatology* 36: e2020PA004204

Tierney JE, deMenocal PB. 2013. Abrupt shifts in Horn of Africa hydroclimate since the Last Glacial Maximum. *Science* 342: 843-6

Tierney JE, Pausata FS, deMenocal PB. 2017. Rainfall regimes of the Green Sahara. *Science advances* 3: e1601503

Tierney JE, Poulsen CJ, Montañez IP, Bhattacharya T, Feng R, et al. 2020. Past climates inform our future. *Science* 370

Tierney JE, Russell JM, Eggermont H, Hopmans E, Verschuren D, Damsté JS. 2010. Environmental controls on branched tetraether lipid distributions in tropical East African lake sediments. *Geochimica et Cosmochimica Acta* 74: 4902-18

Tierney JE, Russell JM, Huang Y, Damsté JSS, Hopmans EC, Cohen AS. 2008. Northern hemisphere controls on tropical southeast African climate during the past 60,000 years. *Science* 322: 252-5

Tierney JE, Ummenhofer CC, Demenocal PB. 2015. Past and future rainfall in the Horn of Africa. *Science advances* 1: e1500682

Tipple BJ, Pagani M. 2010. A 35 Myr North American leaf-wax compound-specific carbon and hydrogen isotope record: Implications for C4 grasslands and hydrologic cycle dynamics. *Earth and Planetary Science Letters* 299: 250-62

Vachula R, Cheung A. 2021. Late Neogene surge in sedimentary charcoal fluxes partly due to preservation biases, not fire activity. *Palaeogeography, Palaeoclimatology, Palaeoecology* 567: 110273

Van Winden JF, Talbot HM, Kip N, Reichart G-J, Pol A, et al. 2012. Bacteriohopanepolyol signatures as markers for methanotrophic bacteria in peat moss. *Geochimica et Cosmochimica Acta* 77: 52-61

Vonk JE, Drenzek NJ, Hughen KA, Stanley RH, McIntyre C, et al. 2019. Temporal deconvolution of vascular plant-derived fatty acids exported from terrestrial watersheds. *Geochimica et Cosmochimica Acta* 244: 502-21

Weber Y, Damsté JSS, Zopfi J, De Jonge C, Gilli A, et al. 2018. Redox-dependent niche differentiation provides evidence for multiple bacterial sources of glycerol tetraether lipids in lakes. *Proceedings of the National Academy of Sciences* 115: 10926-31

Weijers JW, Schefuß E, Schouten S, Damsté JSS. 2007a. Coupled thermal and hydrological evolution of tropical Africa over the last deglaciation. *Science* 315: 1701-4

Weijers JW, Schouten S, van den Donker JC, Hopmans EC, Damsté JSS. 2007b. Environmental controls on bacterial tetraether membrane lipid distribution in soils. *Geochimica et Cosmochimica Acta* 71: 703-13

Wilf P. 1997. When are leaves good thermometers? A new case for leaf margin analysis. *Paleobiology*: 373-90

Windler G, Tierney JE, Anchukaitis KJ. 2021. Glacial-interglacial shifts dominate tropical Indo-Pacific hydroclimate during the late Pleistocene. *Geophysical Research Letters*: e2021GL093339

Windler G, Tierney JE, Zhu J, Poulsen CJ. 2020. Unraveling Glacial Hydroclimate in the Indo-Pacific Warm Pool: Perspectives From Water Isotopes. *Paleoceanography and Paleoclimatology* 35: e2020PA003985

Wing SL, Harrington GJ, Smith FA, Bloch JI, Boyer DM, Freeman KH. 2005. Transient floral change and rapid global warming at the Paleocene-Eocene boundary. *Science* 310: 993-6

Wu MS, Feakins SJ, Martin RE, Shenkin A, Bentley LP, et al. 2017. Altitude effect on leaf wax carbon isotopic composition in humid tropical forests. *Geochimica et Cosmochimica Acta* 206: 1-17

Yunker MB, Macdonald RW, Vingarzan R, Mitchell RH, Goyette D, Sylvestre S. 2002. PAHs in the Fraser River basin: a critical appraisal of PAH ratios as indicators of PAH source and composition. *Organic geochemistry* 33: 489-515

Zell C, Kim J-H, Hollander D, Lorenzoni L, Baker P, et al. 2014. Sources and distributions of branched and isoprenoid tetraether lipids on the Amazon shelf and fan: Implications for the use of GDGT-based proxies in marine sediments. *Geochimica et Cosmochimica Acta* 139: 293-312

Zhang YG, Zhang CL, Liu X-L, Li L, Hinrichs K-U, Noakes JE. 2011. Methane Index: A tetraether archaeal lipid biomarker indicator for detecting the instability of marine gas hydrates. *Earth and Planetary Science Letters* 307: 525-34

Zheng Y, Pancost RD, Liu X, Wang Z, Naafs B, et al. 2017. Atmospheric connections with the North Atlantic enhanced the deglacial warming in northeast China. *Geology* 45: 1031-4

Zheng Y, Singarayer JS, Cheng P, Yu X, Liu Z, et al. 2014. Holocene variations in peatland methane cycling associated with the Asian summer monsoon system. *Nature communications* 5: 1-7

Zhu C, Wagner T, Talbot HM, Weijers JW, Pan J-M, Pancost RD. 2013. Mechanistic controls on diverse fates of terrestrial organic components in the East China Sea. *Geochimica et Cosmochimica Acta* 117: 129-43

[Sidebar 1]

Title: **Organic biomarkers**

Organisms produce a wide range of organic compounds, including proteins, carbohydrates and lipids. Organic compounds undergo various structural transformations during diagenesis, but can retain structural or isotopic information that enables insights into their biological precursor organism (“biomarkers”). Treibs (1934) first demonstrated that organic pigments (porphyrins) preserved within ancient soil, shale and coal deposits were derived from chlorophyll-a, a light-harvesting pigment found in all photosynthetic organisms. In general, a robust biomarker must be: (i) diagnostic of some species, taxa, or process; (ii) well preserved over long timescales (10^3 to 10^8 yr); and/or (iii) responsive to environmental perturbations (e.g., temperature).

Summary phrases:

- Biomarkers can be used to reconstruct terrestrial environmental change over a range of geological timescales
- Analyzing several biomarkers in tandem can provide unique insights into the Earth System

Captions

Figure 1 Key terrestrial biomarkers in the environment. Biomarkers provide insights into physical climate variables (land temperature, rainfall), ecosystem state variables (vegetation, fire regime), and biogeochemical variables (soil residence time, methane cycling). Abbreviation: brGDGT, branched glycerol dialkyl glycerol tetraether.

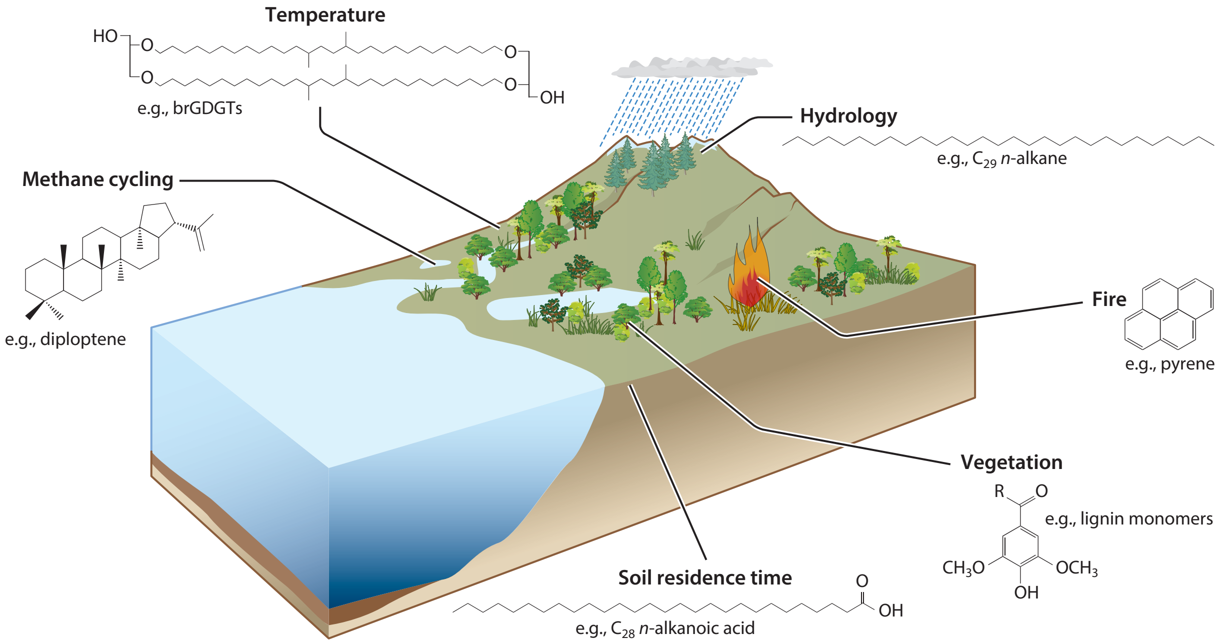
Figure 2 brGDGTs indicate a rapid drop in terrestrial temperatures in Prydz Bay, East Antarctica across the Eocene–Oligocene transition (~34 million years ago). (*a*) Benthic foraminiferal $\delta^{18}O$ values (Westerhold et al. 2020). (*b*) brGDGT-inferred MAAT estimates (Tibbett et al. 2021). The shaded region in panel *b* represents 1σ uncertainty. Abbreviations: brGDGT, branched glycerol dialkyl glycerol tetraether; MAAT, mean annual air temperature.

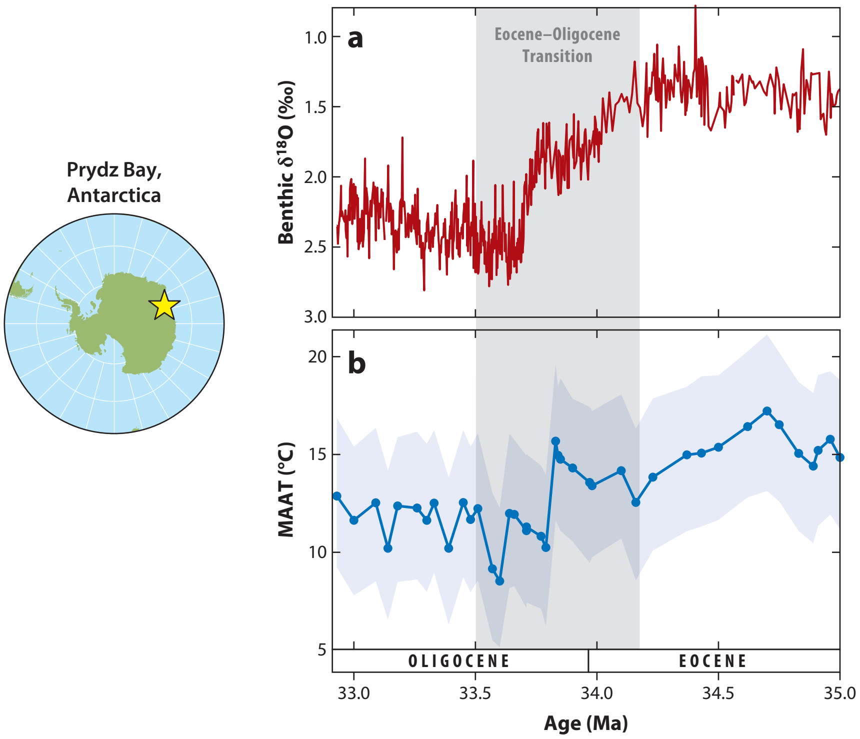
Figure 3 Coupling of hydrology, vegetation, and fire regime on the Indian subcontinent during the late Miocene (~10 million years ago to present). (*a*) δ^2H values from the C_{31} *n*-alkane (Karp et al. 2021) (*black*) and δ^2H values corrected for vegetation fractionation to generate estimates of environmental water δ^2H (*blue*), following Polissar et al. (2021). Colored shaded areas represent 1-sigma errors. (*b*) $\delta^{13}C$ values from pyrene (*dark red*) and C_{31} *n*-alkane (*orange*), both corrected for $\delta^{13}C$ changes in CO_2 (Karp et al. 2021). (*c*) Total pyrogenic PAH concentrations normalized to tetra aromatic β-amyrin derivative, following the method of Karp et al. (2021b). Abbreviation: PAH, polycyclic aromatic hydrocarbons.

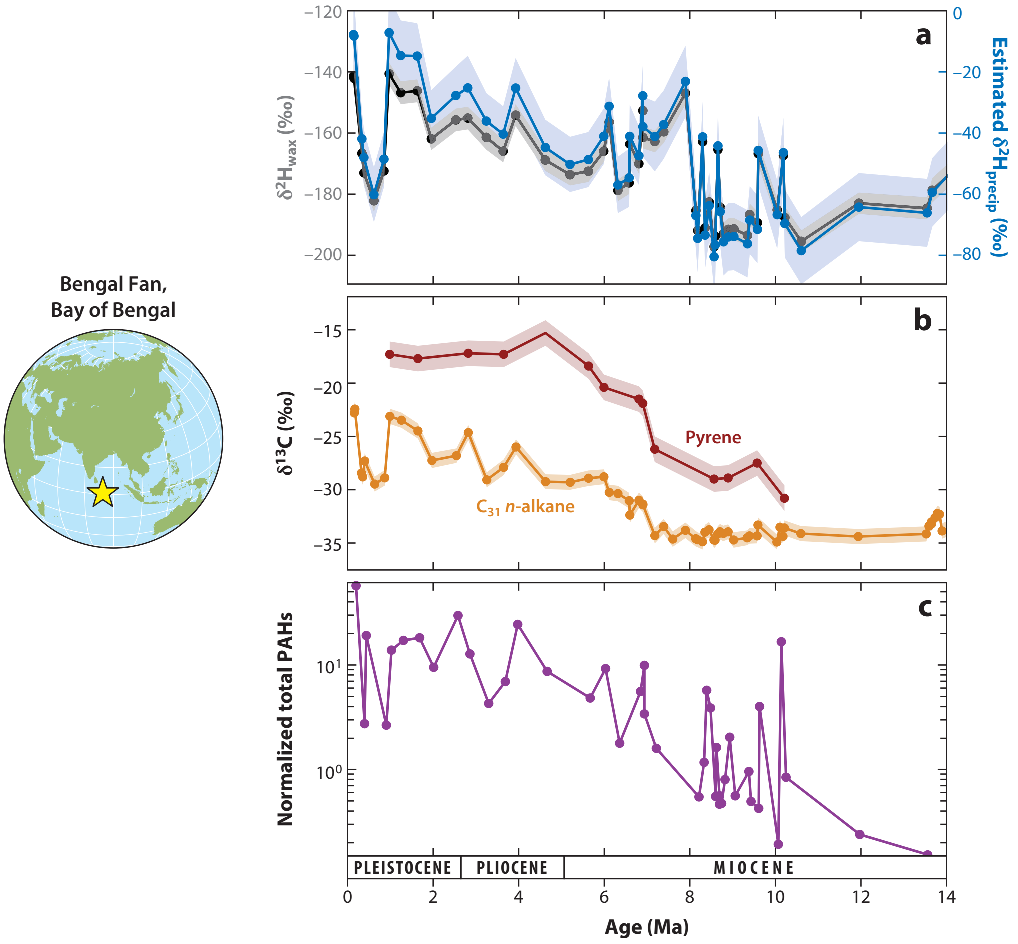
Figure 4 Carbon and hydrogen isotopic compositions of plant wax biomarkers capture rapid climate changes. (*a*) Long-chain *n*-alkanoic acid δ^2H data from the Gulf of Aden (*blue*) track northern hemisphere temperature anomaly (*gray*), indicating regional aridity accompanies twentieth-century warming, and *n*-alkanoic acid δ^{13}C data (*orange*) capture the Suess effect, as represented by δ^{13}C of atmospheric CO_2 (Tierney et al. 2015). (*b*) Long-chain *n*-alkanoic acid δ^2H values from Lake Titicaca capture millennial-scale climate shifts during the last deglaciation (Fornace et al. 2014) that are as rapid—if not more—than the record of δ^{18}O from the Huascarán ice core in Peru (Thompson et al. 1995).

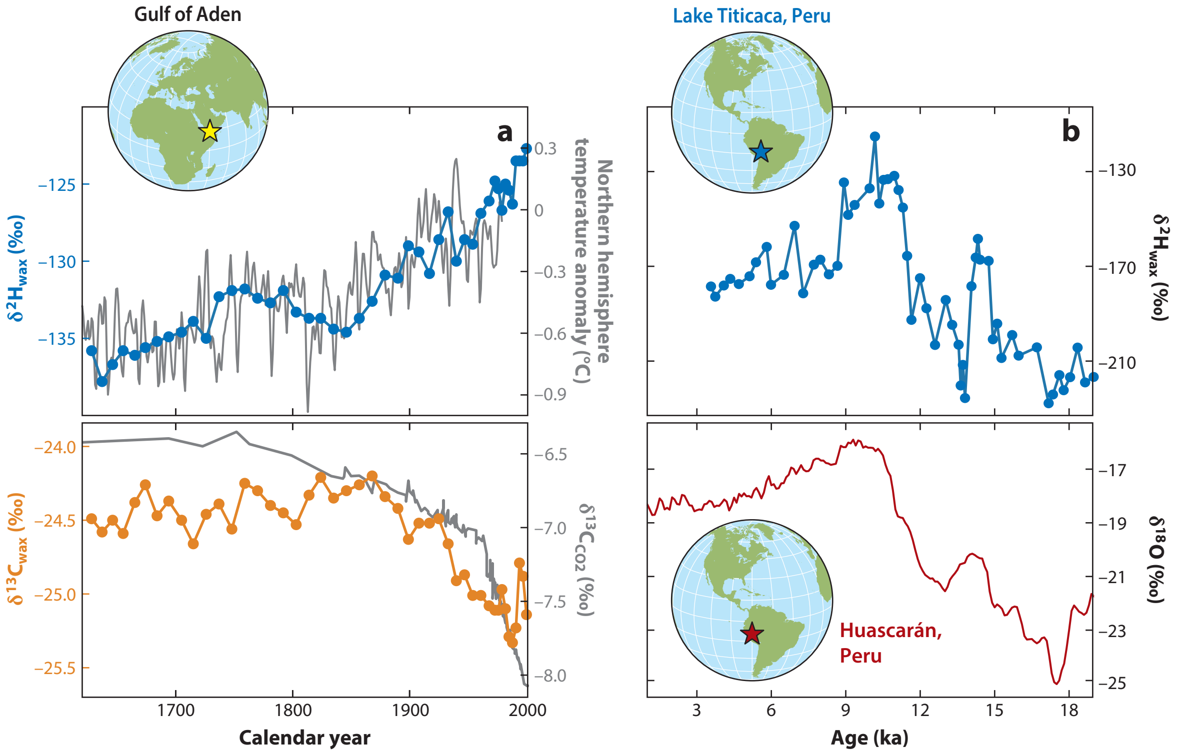
Figure 5 Relationships between lipid ^{14}C and δ^2H_{wax} values since the last deglaciation (~17 thousand years ago to present). Results show a general negative relationship between δ^2H_{wax} and plant-wax mean transit time [reported as the ratio of ^{14}C activity between plant waxes and the contemporaneous atmosphere at the time of deposition (F^{14}R)] and indicate shorter terrestrial residence times under wetter climates. Data from Fornace (2016) (Lake Titicaca), Hein et al. (2020) (Bay of Bengal), and Schefuß et al. (2016) (Congo Fan). Uncertainty for Lake Titicaca is taken as either the propagated analytical uncertainty for combined chain lengths or the maximum-minimum difference between values for different chain lengths, whichever is greater. Uncertainty for other records is taken as that propagated in the original publications.

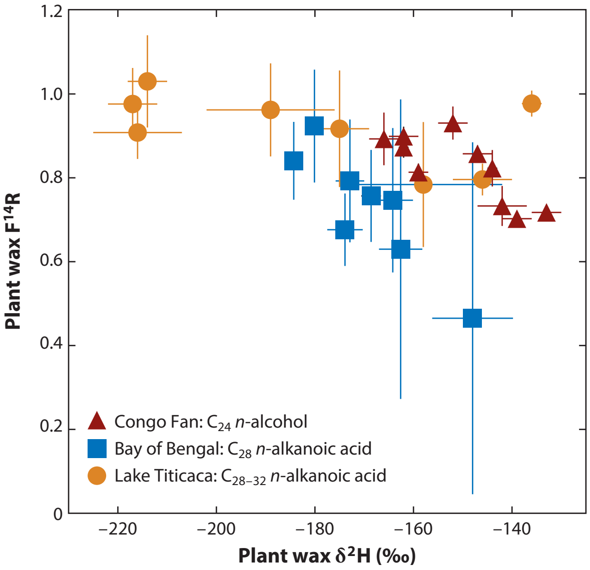
Figure 6 Enhanced terrestrial methane cycling in Otaio River, New Zealand, during the onset of the Paleocene-Eocene Thermal Maximum (~56 million years ago). (*a*) Bulk organic matter δ^{13}C values. (*b*) C_{30} hop-17(21)-ene δ^{13}C values. (*c*) brGDGT-inferred MAAT estimates in marine interbeds only. The light blue shaded region in panel *c* indicates 1σ error. Data from Inglis et al. (2021). Abbreviations: MAAT, mean annual air temperature; brGDGT, branched glycerol dialkyl glycerol tetraether.

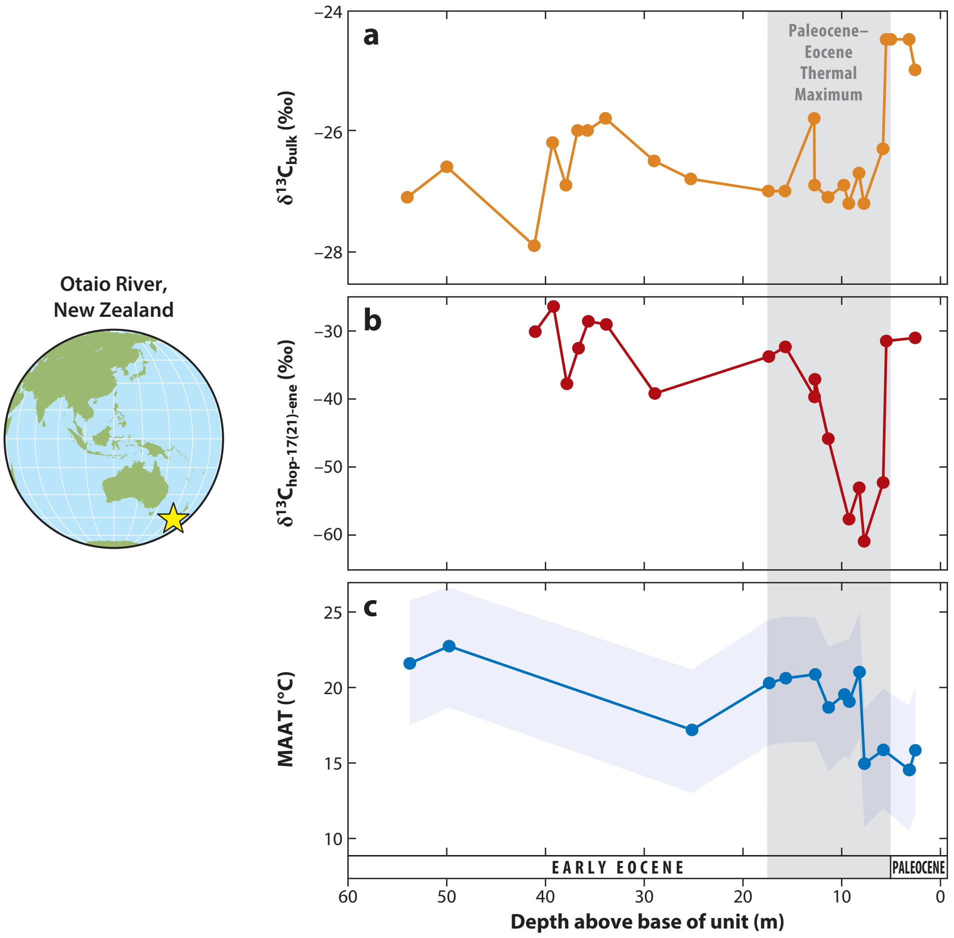

Temperature
HO
O
O
O
O
OH
e.g., brGDGTs
Methane cycling
e.g., diploptene
Hydrology
e.g., C_{29} *n*-alkane
Fire
e.g., pyrene
Vegetation
R
O
e.g., lignin monomers
CH_3O
OCH_3
OH
Soil residence time
O
OH
e.g., C_{28} *n*-alkanoic acid


Prydz Bay,
Antarctica
a
Eocene–Oligocene
Transition
Benthic δ18O (‰)
1.0
1.5
2.0
2.5
3.0
b
MAAT (°C)
20
15
10
5
OLIGOCENE
EOCENE
33.0
33.5
34.0
34.5
35.0
Age (Ma)


Bengal Fan,
Bay of Bengal
a
δ^2H_{wax} (‰)
−120
−140
−160
−180
−200
Estimated δ^2H_{precip} (‰)
0
−20
−40
−60
−80
b
$\delta^{13}C$ (‰)
−15
−20
−25
−30
−35
Pyrene
C_{31} *n*-alkane
c
Normalized total PAHs
10^1
10^0
PLEISTOCENE
PLIOCENE
MIOCENE
0
2
4
6
8
10
12
14
Age (Ma)


Gulf of Aden
a
δ²H_wax (‰)
−125
−130
−135
0.3
0
−0.3
−0.6
−0.9
Northern hemisphere temperature anomaly (°C)
δ¹³C_wax (‰)
−24.0
−24.5
−25.0
−25.5
−6.5
−7.0
−7.5
−8.0
δ¹³C_CO2 (‰)
1700
1800
1900
2000
Calendar year
Lake Titicaca, Peru
b
−130
−170
−210
δ²H_wax (‰)
Huascarán, Peru
−17
−19
−21
−23
−25
δ¹⁸O (‰)
3
6
9
12
15
18
Age (ka)


1.2
1.0
0.8
0.6
0.4
0.2
0
−220
−200
−180
−160
−140
Plant wax F14R
Plant wax δ2H (‰)
Congo Fan: C24 n-alcohol
Bay of Bengal: C28 n-alkanoic acid
Lake Titicaca: C28–32 n-alkanoic acid


Otaio River,
New Zealand
a
b
c
Paleocene–
Eocene
Thermal
Maximum
$\delta^{13}C_{bulk}$ (‰)
–24
–26
–28
$\delta^{13}C_{hop-17(21)-ene}$ (‰)
–30
–40
–50
–60
MAAT (°C)
25
20
15
10
EARLY EOCENE
PALEOCENE
60
50
40
30
20
10
0
Depth above base of unit (m)