https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

Biogeosciences Discussions

Revised fractional abundances and warm-season temperatures substantially improve brGDGT calibrations in lake sediments

Jonathan H. Raberg[1,2], David J. Harning[1,2], Sarah E. Crump[1,*], Greg de Wet[1,†], Aria Blumm[1,‡], Sebastian Kopf[1], Áslaug Geirsdóttir[2], Gifford H. Miller[1], Julio Sepúlveda[1]

1, Department of Geological Sciences and Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO, 80309, USA
2, Faculty of Earth Sciences, University of Iceland, Reykjavík, Iceland

[*]Currently at the Ecology and Evolutionary Biology Department, University of California Santa Cruz, Santa Cruz, CA 95064
[†]Currently at the Department of Geosciences, Smith College, Northampton, MA 01063
[‡]Currently at the Department of Geosciences, University of Arizona, Tucson, AZ 85721

Correspondence to: Jonathan H. Raberg (jonathan.raberg@colorado.edu)

Abstract. Distributions of branched glycerol dialkyl glycerol tetraethers (brGDGTs) are frequently employed for reconstructing terrestrial paleotemperatures from lake sediment archives. Although brGDGTs are globally ubiquitous, the microbial producers of these membrane lipids remain unknown, precluding a full understanding of the ways in which environmental parameters control their production and distribution. Here, we advance this understanding in three ways. First, we present 43 new high-latitude lake sites characterized by low mean annual air temperatures (MATs) and high seasonality, filling an important gap in the global dataset. Second, we introduce a new approach for analyzing brGDGT data in which compound fractional abundances (FAs) are calculated within structural groups based on methylation number, methylation position, and cyclization number. Finally, we perform linear and nonlinear regressions of the resulting FAs against a suite of environmental parameters in a compiled global lake sediment dataset (n = 182). We find that our approach deconvolves temperature, conductivity, and pH trends in brGDGTs without increasing calibration errors from the standard approach. We also find that it reveals novel patterns in brGDGT distributions and provides a methodology for investigating the biological underpinnings of their structural diversity. Warm-season temperature indices outperformed MAT in our regressions, with Months Above Freezing yielding the highest-performing model (adjusted R^2 = 0.91, RMSE = 1.97°C, n = 182). The natural logarithm of conductivity had the second-strongest relationship to brGDGT distributions (adjusted R^2 = 0.83, RMSE = 0.66, n = 143), notably outperforming pH in our dataset (adjusted R^2 = 0.73, RMSE = 0.57, n = 154) and providing a potential new proxy for paleohydrology applications. We recommend these calibrations for use in lake sediments globally, including at high latitudes, and detail the advantages and disadvantages of each.

1 Introduction

Paleotemperature records from lake sediment archives are highly sought after in studies of terrestrial paleoclimate. Bacterial branched glycerol dialkyl glycerol tetraether (brGDGT) lipids have solidified themselves as an important tool in this pursuit (Fig A1; Schouten et al., 2013). First isolated from peat (Sinninghe Damsté et al.,

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021


Biogeosciences Discussions Open Access EGU

2000), these membrane lipids have since been measured in increasingly diverse settings, from marine, soil, lacustrine, and riverine locations (Hopmans et al., 2004; Weijers et al., 2006; Pearson et al., 2011; De Jonge et al., 2014b, respectively) to hot springs, fossil bones, groundwater, deep ocean trenches, and methane seeps (Li et al., 2014; Dillon et al., 2018; Ding et al., 2018; Xiao et al., 2020; Zhang et al., 2020, respectively). Their ubiquity in nature has given them widespread applicability as environmental proxies; brGDGTs have been used to reconstruct temperature in a variety of archives including lake sediments (e.g. de Wet et al., 2016), marine sediments (e.g. Dearing Crampton-Flood et al., 2018), peat (e.g. Zheng et al., 2017), loess (e.g. Lu et al., 2019), and fossil bone (e.g. Zhao et al., 2020). They have additionally been used to reconstruct lake water pH (e.g. Cao et al., 2017). As the microbial producers of brGDGTs remain elusive (Sinninghe Damsté et al., 2018), these paleoclimate reconstructions currently rely on empirical calibrations at both the regional and global level.

At the heart of brGDGT calibrations is the observation that the degree of alkyl-chain methylation and cyclization are correlated to environmental temperature and pH, respectively. These relationships were first quantified by the Methylation and Cyclization of Branched Tetraether indices (MBT and CBT) in a global soil dataset (Weijers et al., 2007). The authors proposed physiological explanations for both connections, positing that an increase in methylation number will enhance membrane fluidity, a desirable trait in cold environments, while a greater number of cyclic moieties could improve proton permeability, an advantageous adaptation at high pH. These physiological responses have precedent in other bacterial lipid classes (Reizer et al., 1985; Beales, 2004; Yuk and Marshall, 2004) and appear to function for brGDGTs as well. However, genomic analyses of environmental samples (Weber et al., 2018; De Jonge et al., 2019; van Bree et al., 2020) have suggested that differences in brGDGT distributions may also stem from shifts in bacterial community composition. Variations in the position of alkyl-chain methylations (Fig. A1; De Jonge et al., 2014a) further complicate the picture, with most studies showing 5-methyl brGDGT isomers to correlate better with temperature than their 6-methyl counterparts (Russell et al., 2018), but others arriving at the opposite result (Dang et al., 2018). These isomeric variations have additionally been shown to correlate with pH in lake sediments (Dang et al., 2016). These discoveries highlight the multifaceted nature of the empirical relationship between brGDGTs and environmental gradients and the need for further study.

Without a clear mechanistic understanding of brGDGTs' dependencies on environmental parameters and no brGDGT-producing model organisms currently available for laboratory experimentation, researchers have relied on statistical methods to construct empirical brGDGT calibrations. The majority of recent calibrations have employed a variety of statistical techniques to construct linear or polynomial regressions using brGDGT fractional abundances (FAs; De Jonge et al., 2014a; Martínez-Sosa et al., 2020b; Pérez-Angel et al., 2020). The fractional abundance fx_i of a compound x_i in a set of n compounds is defined as,

$$fx_i = x_i/(x_1 + x_2 + \cdots + x_n) \quad (1)$$

where any x is the absolute abundance of the given compound in the set. For brGDGTs, these FAs are traditionally calculated using all 15 commonly-measured compounds (Fig. A1),

$$fx_i = x_i/(Ia + Ib + Ic + IIa + IIb + IIc + IIIa + IIIb + IIIc + IIa' + IIb' + IIc' + IIIa' + IIIb' + IIIc') \quad (2)$$

where x_i is any given compound in the denominator. By grouping together all 15 common brGDGTs, this approach makes no prior assumptions about the relationships between the compounds themselves and maximizes the degrees

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021


Biogeosciences Discussions Open Access EGU

of freedom available when exploring a dataset. As relationships between brGDGTs and environment parameters are not yet fully understood, this indiscriminate approach is appropriate. However, by lumping compounds of various types and abundances into the denominator, the approach can also dampen meaningful trends and obscure important relationships, especially for less abundant molecules. This adverse effect has been recognized for other lipid biomarkers and has led to, for example, the exclusion of crenarchaeol from the TEX_{86} index (Schouten et al., 2002) and tetra-unsaturated alkenones from the $U^{K'}_{37}$ index (Prahl and Wakeham, 1987). Numerous ratio-based indices have been developed for brGDGTs that similarly exclude low-abundance (e.g. MBT'; Peterse et al., 2012) or problematic (e.g. MBT'_{5Me}; De Jonge et al., 2014a) compounds. However, a selective approach to fractional abundance calculations has been hitherto unexplored.

On the other side of the calibration equations are the environmental variables that are regressed against brGDGT indices and FAs. Mean annual air temperature (MAT) has been the traditional target of brGDGT calibrations in lake sediments (e.g. Tierney et al., 2010; Loomis et al., 2012). However, it was recognized early on that brGDGT-derived temperatures in cold regions may more accurately reflect warm-season temperatures (Pearson et al., 2011; Sun et al., 2011), an hypothesis that was strongly supported in high-latitude lake sediments (Shanahan et al., 2013; Peterse et al., 2014; Foster et al., 2016). Since the methodological advances that allowed for the separation of 5- and 6-methyl isomers (De Jonge et al., 2014a) and the development of new calibrations, both modern (Hanna et al., 2016; Dang et al., 2018; Cao et al., 2020) and paleo (Super et al., 2018; Thomas et al., 2018; Crump et al., 2019; Harning et al., 2020) studies have continued to support a warm-season bias. Additionally, a recent Bayesian calibration found the mean temperature of Months Above Freezing (MAF) to be the only mode to significantly correlate with brGDGT distributions in a global lake sediment dataset (Martínez-Sosa et al., 2020). However, this warm-season bias has yet to be tested thoroughly in the regions in which it is most pronounced – namely, those with low MAT and high seasonality. As these are the regions that are currently experiencing the most rapid climate change (Landrum and Holland, 2020), their temperature histories are of high interest (Miller et al., 2010) and the quantification of the brGDGT warm-season bias is an important target of study.

Outside of temperature, pH is the most common focus of calibration studies (e.g. Russell et al., 2018). However, numerous other variables including conductivity (Tierney et al., 2010; Shanahan et al., 2013), dissolved oxygen (DO; Colcord et al., 2017; Weber et al., 2018; van Bree et al., 2020; Yao et al., 2020), nutrient availability (Loomis et al., 2014a), and lake mixing regime (Loomis et al., 2014b; van Bree et al., 2020) have been shown to be potentially important controls on brGDGT distributions. Modern calibrations do not currently exist for these environmental variables, largely due to the complexity of the relationships and data limitations.

In this study, we aim to improve lake sediment calibrations for brGDGTs in three ways. First, we extend the global calibration dataset to include high-latitude sites by adding surface sediment from 43 lakes in the Eastern Canadian Arctic, Northern Quebec, and Iceland. Second, we selectively group brGDGTs based on methylation number, methylation position, and cyclization number, and use FAs calculated within these structural sets to deconvolve environmental influences and identify novel patterns in brGDGT distributions. Finally, we analyze the relationship between the compiled global dataset and MAT, four warm-season temperature indices, pH, conductivity, DO, and lake geometry and generate empirical calibrations for use in lake sediments globally.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

2 Methods

2.1 Study sites and sample collection

Surface sediments (0-0.5, 0-1, or 0-2 cm; Ekman box corer or core-top sediment) were collected from 43 lakes (26 from Iceland; 16 from Baffin Island, Arctic Canada; 1 from Northern Quebec; Fig. 1 insets) between 2003 and 2019. For 28 of these lakes, water temperature, pH, conductivity, and DO were measured at the time of sampling during the summers of 2017-2020 using a multiparameter probe (HydroLab HL4, OTT HydroMet). These parameters were additionally measured beneath the lake ice in Feb-May of 2018-2020 for 11 lakes. All but one of these lakes experienced depleted bottom water oxygen levels under the ice relative to ice-free conditions. We therefore assume that our ice-free water chemistry profiles do not capture the minimum DO (DO_{min}) levels experienced by our Canadian and Icelandic lakes and exclude them from our analysis of DO_{min}, with the exception of the site from Northern Quebec, which contained a summer oxycline. All other water chemistry parameters were averaged across all depths and seasons before being used in calibrations.

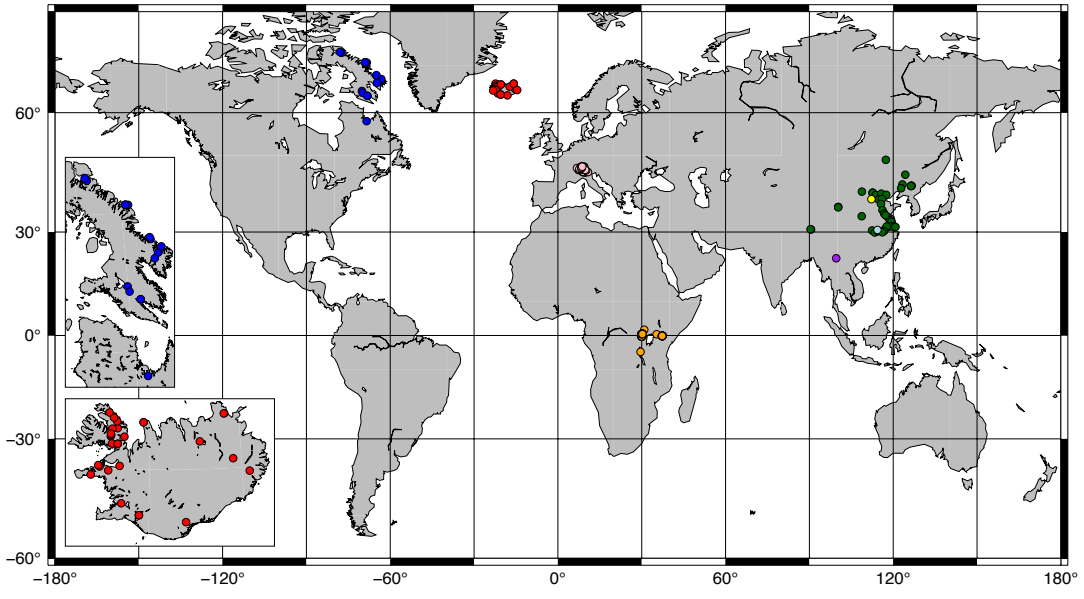


Figure 1. Map of sites included in this study. Red and blue: Iceland and Canada (this study; see insets); Orange: Russell et al. (2018); Green: Dang et al. (2018); Pink: Weber et al. (2018); Yellow: Cao et al. (2020); Light blue: Qian et al. (2019); Purple: Ning et al. (2019).

Previously published data from 36 lakes in Central Europe (Weber et al., 2018), 65 lakes in East Africa (Russell et al., 2018), and 38 lakes in China (Dang et al., 2018; Ning et al., 2019; Qian et al., 2019; Cao et al., 2020), were added to the dataset for a total of 182 data points (Fig. 1). Average brGDGT FAs were used for lakes with multiple surface sediment samples (up to three in some cases). Lake surface areas (SAs) were taken from published datasets or estimated using DigitalGlobe imagery. Lake volumes were estimated by approximating each lake basin as

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

a hemiellipsoid ($volume = 4/3 \times SA \times maximum\ depth$). Water chemistry parameters were taken from the literature where available or else excluded from our analyses.

2.2 Sample extraction and analysis

Roughly 1 g of freeze-dried sediment was extracted using either an accelerated solvent extractor (ASE 200 DIONEX; 10 samples) or a modified Bligh and Dyer (BD; 33 samples) method. We provide a comparison of both extraction methods below. For the ASE method, samples were extracted twice using 9:1 (v:v) dichloromethane:methanol (DCM:MeOH) at 100°C and 2,000 psi. Total lipid extracts (TLEs) were redissolved in 99:1 (v:v) hexane:isopropanol (Hex:IPA) and filtered (0.45 µm, PTFE) before analysis. The remaining 33 samples were extracted using a modified BD procedure (Wörmer et al., 2013). Briefly, sediment was vortexed and sonicated in Mix A (DCM:MeOH:50mM Phosphate buffer (aq., pH 7.4) [1:2:0.8, v:v:v]). The mixture was then centrifuged at 3,000 rpm and 10°C for 10 minutes and the supernatant was collected in a glass separatory funnel. The process was performed twice with Mix A, twice with Mix B (DCM:MeOH:5% Trichloroacetic acid buffer (aq., pH 2) [1:2:0.8, v:v:v]), and once with Mix C (DCM:MeOH [1:5, v:v]). Equal volumes of HPLC-grade water and DCM were added to induce separation. The organic fraction was collected and dried under a nitrogen stream. The aqueous phase was washed once with DCM and the organic fraction was added to the extract. The TLE was then redissolved in 99:1 (v:v) Hex:IPA and filtered (0.45-µm, PTFE) before analysis.

We analyzed brGDGTs using a Thermo Scientific UltiMate 3000 high-performance liquid chromatography instrument coupled to a Q Exactive Focus Orbitrap-Quadrupole high-resolution mass spectrometer (HPLC-MS) via an atmospheric pressure chemical ionization (APCI). We achieved chromatographic separation using a slightly modified version (Crump et al., 2019; Harning et al., 2019; Pérez-Angel et al., 2020) of the HPLC method described by Hopmans et al. (2016). Due to observed deterioration of chromatography over time, we lowered the initial concentration of eluent B from 18% to 14% to maintain optimal separation of the 5- and 6-methyl isomers. A C46 GDGT internal standard (Huguet et al., 2006) was added to the TLE immediately after extraction and was used to quantify brGDGT yields.

2.3 Comparison of ASE and BD Extraction Methods

To ensure that brGDGT distributions were agnostic to our extraction method, we extracted three surface sediments, two suspended particulate matter (SPM) samples (2.5 L lake water filtered onto 0.3 µm glass fiber filters), and three soils from Baffin Island using the ASE and BD methods in parallel (Fig. S1). The mean difference in MBT'_{5Me} (Eq. A3) between the two extraction methods was 0.006 ± 0.004, or 2 ± 2%. This translates to a MBT'_{5Me}-derived temperature difference of 0.2 ± 0.2°C using recent calibrations for soils (Naafs et al., 2017) and lake sediments (Russell et al., 2018), which is well below their respective RMSEs of 5.3°C and 2.14°C (Fig. S2). To test for compound-specific differences, we calculated percent differences in FAs between the two methods. For compounds with FAs > 0.05 in the ASE method, the mean absolute percent difference compared to BD was 4 ± 3%. For the lower abundance compounds (FA ≤ 0.05), this difference was higher (19 ± 17%). No biases in the FA differences were found in either case (difference < standard deviation).

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021


Biogeosciences Discussions Open Access EGU

We further extracted the BD sample residue with the ASE method to determine if any brGDGTs remained after BD extraction. On average, we recovered only an additional 0.8 ± 0.6% brGDGTs. These residual brGDGTs had a similar MBT'$_{5Me}$ to that of original BD extract (mean difference = 0.02 ± 0.01, equivalent to 0.5 ± 0.4°C) and were not present in high enough abundances to significantly affect the overall BD distributions (Fig. S2). We therefore conclude that there are no significant differences between samples extracted with the two methods and treat them identically in the analyses that follow.

2.4 Air Temperatures

Monthly air temperature averages were gathered using the following methods. For nine sites on Baffin Island, one year of *in situ* two-meter air temperature data from five temperature loggers (one- to four-hour resolution, Thermochron iButtons, Maxim Integrated Products) was converted to a 30-year monthly climate normal (1971 to 2000) using a transfer function to relate local data to nearby meteorological stations (Department of Environment, Government of Canada). For the remaining Canadian sites as well as all sites in Iceland, we used the WorldClim database (Fick and Hijmans, 2017) to generate 30-year climate normals for the same time period (1970 to 2000). Monthly temperatures for Central European sites were derived using monthly altitudinal lapse rates constructed from climate normals (1970 to 2013) of 148 meteorological stations (Federal Office of Meteorology and Climatology: MeteoSwiss). Monthly temperature data was not available for the East African lakes. However, the seasonality (standard deviation of monthly temperatures) of these lakes is low (0.5 ± 0.2 °C in the WorldClim database, or < 2% of range of the dataset). We therefore approximate all monthly temperatures to be equivalent to MAT for these lakes. Monthly temperature data from all other studies were either published or provided by the authors.

We used the above monthly air temperatures to calculate MAT and four warm-season temperature indices. Three of these indices represent an average temperature for the warmer portion of the year: mean temperature of Months Above Freezing (MAF), Mean Summer Temperature (MST; mean of June, July, and August in the Northern Hemisphere and December, January, and February in the Southern Hemisphere), and mean Warmest Month Temperature (WMT). These indices capture average temperatures for the warm season, but are unaffected by its duration. We therefore additionally calculate the Summer Warm Index (SWI), defined as the cumulative sum of all monthly temperatures above 0°C. This index represents an important control on vegetation patterns at high latitudes and is a useful alternative to the Growing Degree Days Above 0°C (GDD_0) index when daily temperature data is not available (e.g. Raynolds et al., 2008). For our five *in situ* temperature loggers in the Eastern Canadian Arctic for which sub-daily temperature data is available, GDD_0 and SWI are highly correlated ($R^2 = 0.998$).

2.5 Statistical Methods

To construct calibrations between brGDGT FAs and environmental variables, we used the following method. Each FA was first regressed alone against the environmental variable being investigated. Compounds with a correlation p-value ≥ 0.01 were considered non-significant and removed from further analysis (Pérez-Angel et al., 2020). Fits were then constructed from the remaining compounds using two independent approaches. The first approach was stepwise forward selection/backwards elimination (SFS/SBE) using the MASS package (Venables and

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021


Biogeosciences Discussions Open Access EGU

Ripley, 2002) in R (R Core Team, 2018). This approach finds the best fit by sequentially adding (SFS) or removing (SBE) terms in a generalized linear model and evaluating the resulting fit using the Bayesian Information Criterion (BIC; Schwarz, 1978). The approach is common for constructing brGDGT calibrations (e.g. Dang et al., 2018; Russell et al., 2018), but it is not exhaustive. We therefore additionally used the leaps package (Lumley, 2020) in R to evaluate all possible linear combinations of fitting variables (the "combinatoric" approach, Pérez-Angel et al., 2020). We again used the BIC to select the best fit. For both methods, we imposed the additional criterion that each of the resulting fitting variables must itself be statistically significant ($p < 0.01$). To help avoid overfitting, we additionally used the adjusted R^2 to evaluate calibration performances. Some of our variables (conductivity, depth, surface area to depth, and volume) spanned multiple orders of magnitude. For these variables, we performed regressions against the natural logarithm of the variable.

We applied this calibration procedure to the FAs of each brGDGT structural set and subset defined below (Sect. 3). In the subset-specific calibrations, it was sometimes possible for a single compound to dominate (FA = 1). These samples were generally clear outliers resulting from the low natural abundances of all other members of the subset and they were therefore removed from the subset-specific calibration models. We additionally tested for linear regressions against a number of previously-defined brGDGT indices. A summary of these indices and their definitions is provided in the Appendix.

All correlations reported in the text and figures were significant ($p < 0.01$) except those marked with an asterisk or with a p-value provided. All R^2 values reported in the text and figures are adjusted R^2.

3 Partitioning brGDGTs into structural sets for FA calculations

The structure of brGDGTs can vary in three (currently-observed) ways: methylation number, methylation position, and cyclization number. These variations result in 15 commonly-measured brGDGTs (Fig. A1). The standard approach in brGDGT analysis is to calculate fractional abundances using all 15 of these compounds (Eq. 2). By mathematically mixing brGDGTs with varying methylation number, methylation position, and cyclization number, however, this approach risks convoluting the influences of disparate environmental variables. Here, we present a method of grouping brGDGTs that highlights one type of structural variation (e.g. methylation number) while holding one or both of the others constant. By calculating FAs within each group, we aim to deconvolve the influences of temperature, pH, and other environmental variables on the structural variations of brGDGTs.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

Biogeosciences Discussions Open Access EGU

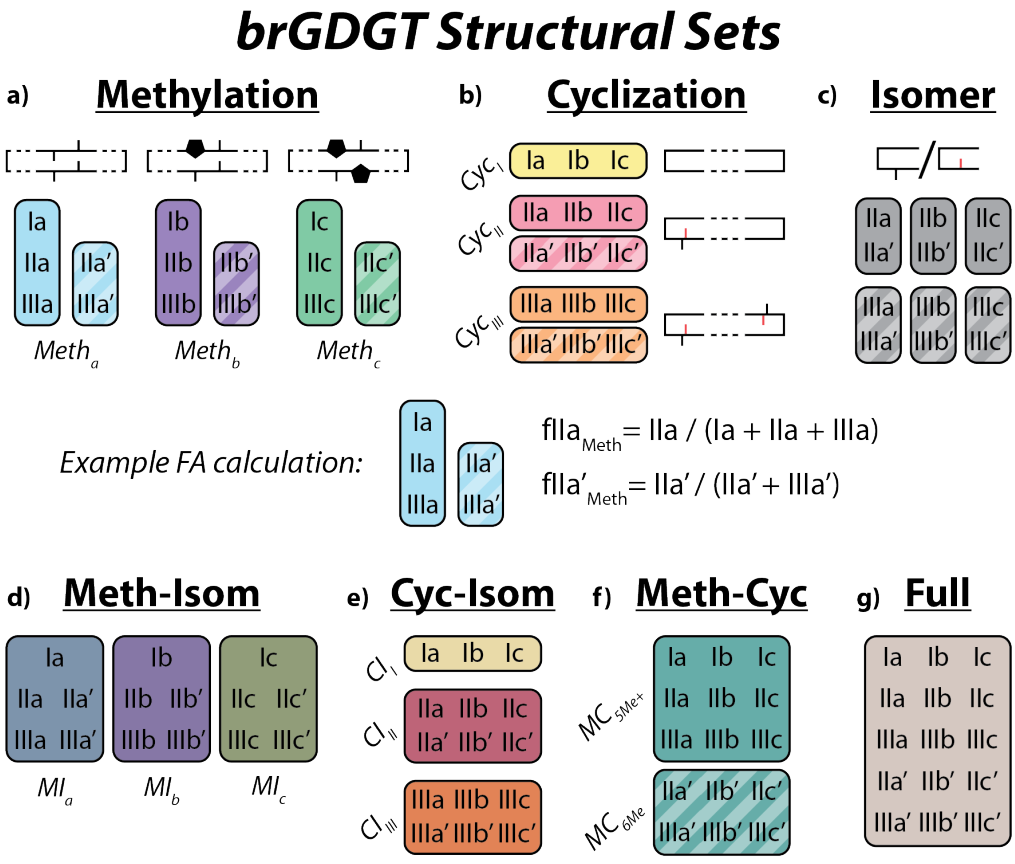


Figure 2. Schematic of the basic (a-c) and combined (d-g) brGDGT structural sets. Fractional abundances are calculated within each boxed group independently (Eq. (3-9) and Table A1). Schematic structures highlight the defining alkyl-chain moieties, with cyclopentane rings filled in for emphasis and C6 methylations denoted in red. Complete structures are available in Fig. A1.

To explore how changes in methylation number alone relate to environmental parameters, we constructed the Methylation (Meth) set (Fig. 2a). This set was generated by grouping brGDGTs with the same number of cyclopentane rings and the same methylation positions. Fractional abundances calculated within the Meth set solely reflect changes in methylation number, while ring number and isomer designation are held constant. Within the Meth set, we defined the Meth-5Me and Meth-6Me subsets as those that contained only 5- and 6-methyl lipids, respectively. As the tetramethylated brGDGTs (Ia, Ib, and Ic) are neither 5-methyl nor 6-methyl isomers, we generated versions of these two subsets that excluded (Meth-5Me, Meth-6Me) and included (Meth-5Me+, Meth-6Me+) these compounds (Fig. S3a-d). We found that the correlation between 5-methyl isomers and temperature was significantly improved when the tetramethylated compounds were included in their FA calculations, while the opposite was true for the 6-methyl compounds (see discussion in Sect. 4.2.2). We therefore grouped the tetramethylated compounds with the 5-methyl compounds (and not with the 6-methyl compounds) when generating the Meth set. FAs for the Meth set were calculated using Eq. (3),

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

$$fxy_{Meth} = xy \Big/ \sum_{n=I}^{III} ny \,;\; fxy'_{Meth} = xy' \Big/ \sum_{n=II}^{III} ny' \quad (3)$$

where *fxy* and *xy* are the fractional and absolute abundances of the brGDGT with Roman numeral *x* (I, II, or III) and alphabet letter *y* (a, b, or c) and tetramethylated compounds are grouped with 5-methyl isomers.

We next defined the Cyclization (Cyc) set to examine the relationship of brGDGT ring number with environmental variables (Fig. 2b). The Cyc set was formed by grouping brGDGTs with the same number and position of methylations. With these variables held constant, variations in the Cyc FAs reflect only variations in the number of cyclopentane moieties. We defined the Cyc-5Me and Cyc-6Me subsets as those containing only 5- and 6-methyl isomers, respectively (Fig. S3e-f). FAs for the Cyc set were calculated using Eq. (4),

$$fxy^{(\prime)}{}_{Cyc} = xy^{(\prime)} \Big/ \sum_{m=a}^{c} xm^{(\prime)} \quad (4)$$

where *fxy* and *xy* are the fractional and absolute abundances of the 5- or 6-methyl brGDGT with Roman numeral *x* (I, II, or III) and alphabet letter *y* (a, b, or c).

Third, we defined the Isomer (Isom) set to isolate changes in the relative abundances of brGDGT isomers (Fig. 2c). The Isom set was constructed by grouping brGDGTs with the same number of methylations and cyclizations. Its FAs are solely a measure of the relative abundances of 5- and 6-methyl isomers, without the convoluting influence of ring or methylation number variations. The isomeric diversity of brGDGTs is large, however, and there are structural variations that are not controlled for within this set. For example, hexamethylated brGDGTs have two isomers: one with methylations on different alkyl chains (e.g. at C5 and C5') and another with methylations on the same chain (e.g. C5 and C24; De Jonge et al., 2013). As these compounds coelute, they cannot be treated independently at this time. Additionally, it is unclear whether brGDGT-IIIa" (Weber et al., 2015), which contains both a C5 and a C6' methylation, should be grouped with 5- or 6-methyl brGDGTs, or whether the three recently-identified 7-methyl brGDGTs (Ding et al., 2016) should be included in the Isom series. As these compounds are rarely reported, we excluded them from our analysis here, but suggest the possibility of expanding the Isom set to include them in the future should more data become available. The Isom FAs were calculated using Eq. (5),

$$fxy^{(\prime)}{}_{Isom} = xy^{(\prime)} \Big/ \sum_{isomers} xy \quad (5)$$

where f*xy* and *xy* are the fractional and absolute abundances of the 5- or 6-methyl brGDGT with Roman numeral *x* (I, II, or III) and alphabet letter *y* (a, b, or c) and "isomers" refers to 5- and 6-methyl brGDGTs. The Isom set contained groups of two compounds each, making their FAs redundant. We therefore used only the 6-methyl FAs in our analysis.

The Meth, Cyc, and Isom sets each allow only one structural component of brGDGTs to vary. It is possible, however, that two structural alterations occur in tandem in response to the same environmental variable. We therefore defined three additional sets that hold one variable constant while allowing the other two to vary. The first is the Meth-Isom (MI) combination set (Fig. 2d). In this set, brGDGTs with the same ring number are grouped together, while both methylation number and position are allowed to vary. The Cyc-Isom (CI; Fig. 2e) is analogously constructed by holding methylation number constant, while the Meth-Cyc (MC; Fig. 2f) set holds methylation position constant (again treating tetramethylated brGDGTs as 5-methyl compounds, Fig. S3g-j). Finally, we defined the Full set (Fig.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

Biogeosciences Discussions Open Access EGU

2g), which takes the standard approach of allowing all three structural characteristics to vary freely by grouping all 15 commonly-measured brGDGTs together. The FAs of the combined sets are calculated using Eq. (6-9),

$$fxy^{(\prime)}{}_{MI} = xy^{(\prime)} \Big/ \sum_{isomers} \sum_{n=I}^{III} ny \quad (6)$$

$$fxy^{(\prime)}{}_{CI} = xy^{(\prime)} \Big/ \sum_{isomers} \sum_{m=a}^{c} xm \quad (7)$$

$$fxy_{MC} = xy \Big/ \sum_{n=I}^{III} \sum_{m=a}^{c} nm; \; fxy'_{MC} = xy' \Big/ \sum_{n=II}^{III} \sum_{m=a}^{c} nm' \quad (8)$$

$$fxy^{(\prime)}{}_{Full} = xy^{(\prime)} \Big/ \sum_{isomers} \sum_{n=I}^{III} \sum_{m=a}^{c} nm \quad (9)$$

where *fxy* and *xy* are the fractional and absolute abundances of the 5- or 6-methyl brGDGT with Roman numeral *x* (I, II, or III) and alphabet letter *y* (a, b, or c) and tetramethylated compounds are treated as 5-methyl isomers. An expanded guide to FA calculations is provided in Table A1.

As a proof of concept, we show that the fractional abundance of just one compound, brGDGT-Ia, can be calculated within different structural sets to provide either a strong temperature or pH correlation, without a strong cross-correlation. When the standard FA is calculated using all 15 compounds (fIa_{Full}), a moderate correlation is found with MAF ($R^2 = 0.61$) and none with pH ($R^2 = 0.04$, $p = 0.014$; Fig. 3). When cyclization is held constant by using the MI set, correlation with MAF increases ($R^2 = 0.75$) while that with pH remains uncorrelated ($R^2 = 0.13$). The temperature correlation increases further when isomer designation is controlled for as well ($fIa_{Meth\text{-}5Me+}$, $R^2 = 0.88$), with an R^2 nearly matching that of MBT'_{5Me} ($R^2 = 0.89$) and the pH correlation remaining low ($R^2 = 0.27$). In contrast, the correlation with temperature disappears for the analogous 6-methyl subset ($fIa_{Meth\text{-}6Me+}$, $R^2 = 0.08$). Finally, allowing only ring number to vary (fIa_{Cyc}) effectively erases the correlation with temperature ($R^2 = 0.18$) and instead provides a correlation with pH that, while modest, is already higher than any reported for a lake sediment calibration to date ($R^2 = 0.59$).

FA	Compounds	MAF R^2	pH R^2
fIa_{Full}	All (15)	0.61	0.04*
fIa_{MI}	Ia, IIa, IIIa, IIa′, IIIa′	0.75	0.13
fIa_{Cyc}	Ia, Ib, Ic	0.18	0.59

FA	Compounds	MAF R^2	pH R^2
$fIa_{Meth\text{-}5Me+}$	Ia, IIa, IIIa	0.88	0.27
$fIa_{Meth\text{-}6Me+}$	Ia, IIa′, IIIa′	0.08	0.07

Figure 3. Adjusted R^2 values for a linear regression of environmental parameters against the fractional abundance (FA) of brGDGT-Ia calculated within different structural sets. Colors denote the strengths of the relationships, from the minimum to the maximum observed coefficients of determination, with white being the median of the dataset. "Compounds" denote all brGDGTs used in the FA calculation. All values are significant ($p < 0.01$) unless marked with an asterisk.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

Biogeosciences Discussions
Open Access
EGU

4 Results and Discussion

4.1 Distributions of brGDGTs in Icelandic and Canadian lake sediments

The new Icelandic and Canadian lake sediments bolster the global dataset on the cold end, extending the lowest MAT from -0.2°C to -18°C and containing 28 of the 30 coldest samples by MAT. The low temperatures are reflected in the brGDGT distributions of these sediments (Fig. 4), which contain on average a higher fIIIa$_{Full}$ and lower fIa$_{Full}$ than the other samples in this dataset. Additionally, the Canadian and Icelandic datasets provide important end-member samples, including those with the highest fIIIa$_{Full}$, fIIIb$_{Full}$, and fIIIc$_{Full}$ and the lowest or second-lowest fIa$_{Full}$, fIb$_{Full}$, and fIc$_{Full}$. The new samples also contain some of the lowest fractional abundances of 6-methyl isomers, providing low end-member values for fIIa'$_{Full}$, fIIb'$_{Full}$, fIIc'$_{Full}$ and below-average values for the hexamethylated 6-methyl brGDGTs.

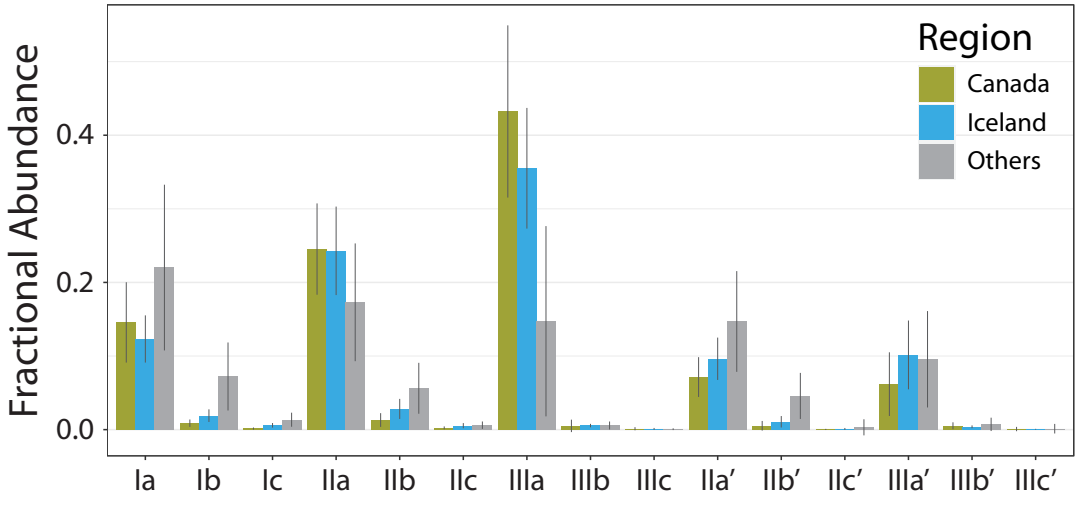
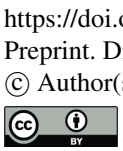


Figure 4. Average fractional abundances (calculated within the standard Full set) of new samples from Canada and Iceland as compared to the average of the other samples in this study (Others).

4.2 Temperature relationships with brGDGTs

In this section, we show that two adjustments can be made to significantly improve the correlations between brGDGT FAs and temperature in lake sediments: 1) replace MAT with a warm-season temperature index and 2) use FAs calculated within the Meth structural set. The effect of these two adjustments on one representative compound, brGDGT-Ia, is shown from left to right in Fig. 5. The relationship between temperature and fIa is improved from a weak but significant correlation with a large error (R^2 = 0.49; RMSE = 7.00°C) to a stronger correlation with a smaller error (R^2 = 0.88; RMSE = 2.42°C). The two adjustments are detailed in the following sections.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021



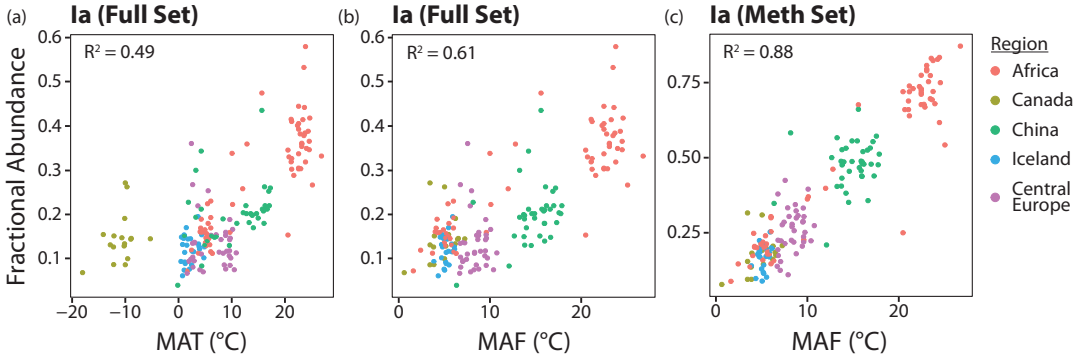


Figure 5. From left to right, the effects of substituting the Months Above Freezing (MAF) warm-season index for Mean Annual Temperature (MAT) and fIa_{Meth} for fIa_{Full} on the relationship between the fractional abundance of brGDGT-Ia and temperature.

4.2.1 Warm-season temperatures outperform MAT

To assess the possibility of a warm-season bias in brGDGT temperature reconstructions, we tested the relationships between brGDGTs and five air temperature indices: MAT, MAF, MST, WMT, and SWI (Sect. 2.4). The correlation between these variables and the FA of one representative compound, fIa_{Meth}, is shown in Fig. 6. In all cases, substituting MAT with a warm-season temperature variable draws samples with strong seasonality into the main body of data. MAF performs best for brGDGT-Ia (R^2 = 0.88), with high-seasonality lakes falling progressively out of alignment for SWI (R^2 = 0.84), MST (R^2 = 0.64), and WMT (R^2 = 0.60). This result was upheld when performing temperature calibrations; MAF outperformed all other measures of temperature examined in this study (Sect. 4.4.1).

A possible explanation for the success of the MAF temperature index is that the activity of brGDGT-producing microbes may be heavily depressed under lake ice and/or in frozen soils (Pearson et al., 2011; Shanahan et al., 2013; Peterse et al., 2014; Cao et al., 2020). However, studies employing sub-seasonal sampling of sediment traps and suspended particulate matter in two mid-latitude lakes have shown that brGDGTs are produced within the water column throughout the year, despite the presence of ice cover (Woltering et al., 2012; Loomis et al., 2014b). These and other studies employing similar sampling techniques (Hu et al., 2016; Weber et al., 2018; van Bree et al., 2020) have additionally found production of brGDGTs to be dependent on the degree and timing of lake mixing versus stratification. Though heightened biological activity may still be the underlying driver of the observed warm-season bias, these depth- and time-resolved studies paint a complex picture of brGDGT production in lakes that precludes a simple explanation. Unfortunately, as knowledge of the timing, extent, and temperature of ice cover and mixing events is lacking for the vast majority of lakes in this study, these effects cannot be tested here. We therefore stress the empirical nature of our MAF calibrations and the need for further study.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

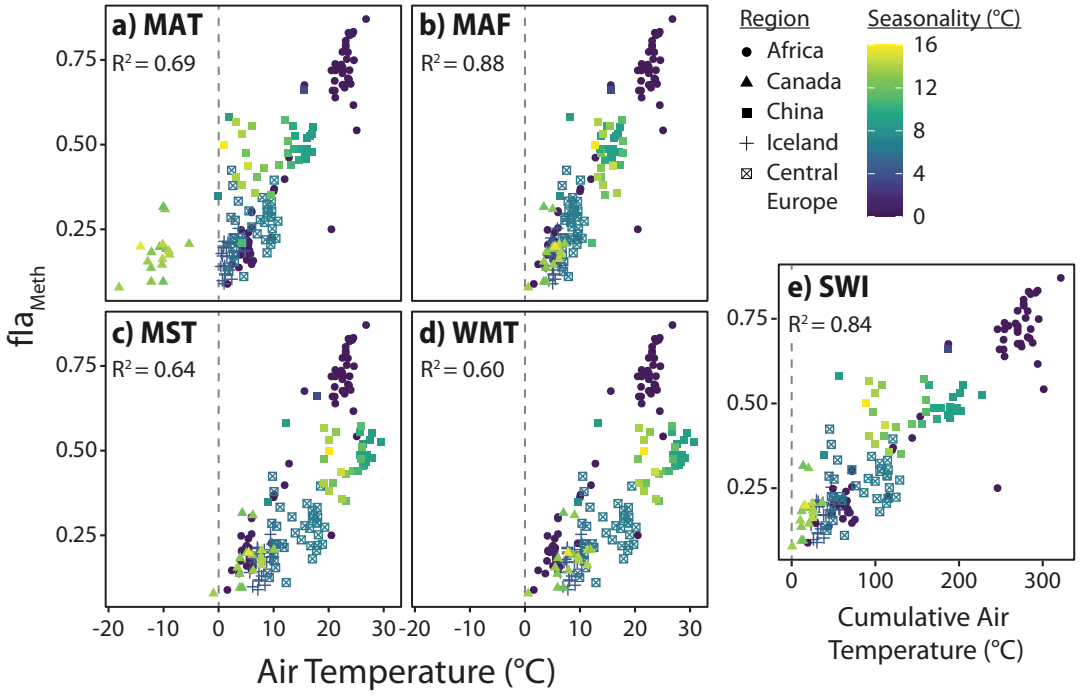
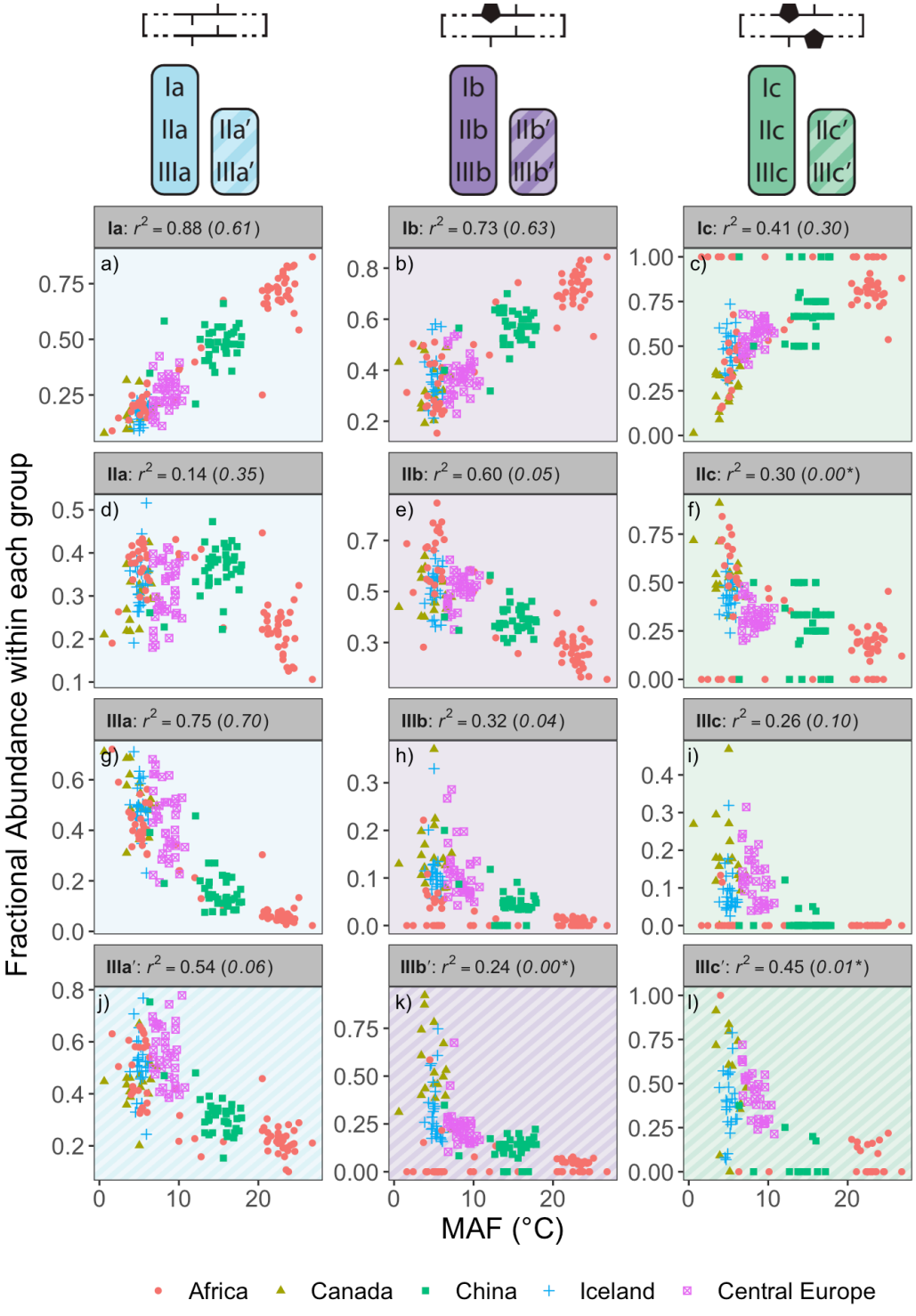


Figure 6. Relationships between fIa_{Meth} and the five air temperature variables tested in this study, a) mean annual temperature (MAT), b) months above freezing (MAF), c) mean summer temperature (MST), d) warmest month temperature (WMT), and e) summer warmth index (SWI). Colors indicate air temperature seasonality (standard deviation of mean monthly temperatures). Dashed lines indicate mean or cumulative temperatures of 0°C for reference. R^2 values for the relationship between fIa_{Meth} and each temperature variable are provided in each subplot.

4.2.2 Temperature and the Methylation set

The Methylation set provided the strongest relationships between brGDGT FAs and MAF (Fig. 7). This set both strengthened existing correlations (e.g. fIa_{Meth}, Fig. 7a) and generated new ones (e.g. $fIIb_{Meth}$, Fig. 7e). Furthermore, many relationships between FAs and temperature were revealed to be qualitatively similar regardless of ring number. For example, the FAs of all tetramethylated compounds (fIa_{Meth}, fIb_{Meth}, fIc_{Meth}) had a strong positive linear relationship with temperature (Fig. 7a-c), while those of the 5-methyl pentamethylated compounds all had a noisier negative relationship (Fig. 7d-f). The hexamethylated FA trends were less clear in part due to their lower abundances, but all were negatively correlated with temperature and nonlinearities were apparent (Fig. 7g-l). These analogous trends show that the number of methylations responds similarly to temperature regardless of the number of cyclopentane rings. Within a paleoclimate lens, this observation opens up the possibility of independent temperature calibrations for un-, mono-, and bicyclized brGDGTs (Sect. 4.4.1). From a biological standpoint, it could imply that methylation and cyclization play their biological roles independently, as the former appears to vary more or less freely of the latter.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021



https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

Figure 7. Relationships between the average air temperature of months above freezing (MAF) and brGDGT FAs calculated within the Meth set. R^2 values are provided for each subplot, with R^2 values for the standard Full FAs given in parentheses for comparison. P-values were < 0.01 except where marked with an asterisk. Note that plots of IIa', IIb', and IIc' are redundant because they exactly mirror those of IIIa', IIIb', and IIIc', respectively, and are therefore not shown.

The brGDGT temperature response appears to be agnostic to methylation position as well (Fig. 7g-i versus j-l), but only when the tetramethylated brGDGTs (Ia, Ib, and Ic) are excluded from 6-methyl FA calculations. The Meth-5Me+ subset (Fig. S3a) showed strong relationships between 5-methyl brGDGTs and MAF ($R^2 \leq 0.88$; Fig. S4). On the other hand, the analogous Meth-6Me+ subset (Fig. S3c) was broadly uncorrelated with MAF ($R^2 \leq 0.29$; Fig. S5). At present, there is no known mechanism whereby an additional methylation at the C5 position would have an influence on membrane physiology in a way that a methylation at C6 would not. At first glance, then, the markedly different responses of the Meth-5Me+ and Meth-6Me+ subsets to changes in temperature do not appear to support a physiological basis for the empirical relationship between temperature and brGDGT methylation number. However, when tetramethylated brGDGTs were excluded from the FA calculations for these compounds (Meth-5Me and Meth-6Me subsets; Figs. S3b and d), statistically significant and qualitatively similar temperature relationships did become visible for both isomer types (Figs. S6-7). An analogous result was found for the MC set (Figs. S3g-j and S8-11). It is not clear at this time why the inclusion of tetramethylated brGDGTs improved temperature correlations for 5-methyl compounds but weakened them for 6-methyl compounds. The discrepancy may imply one or a combination of the following: 1) the isomers are produced by different organisms; 2) the isomers serve distinct biological functions, either in addition to or apart from a temperature response; or 3) the currently-measured tetramethylated brGDGTs (Ia, Ib, and Ic) are not the precursors of 6-methyl brGDGTs. Regardless, this result allows us to combine the higher-performing Meth-5Me+ and Meth-6Me subsets to generate the Meth set (Fig. 2a), which maximizes the temperature responses of all 15 commonly-measured brGDGT (Fig. 7).

While the Meth set highlights the relationships between brGDGT FAs and temperature, it simultaneously weakens those with other environmental variables. FAs calculated in the standard Full set contain conductivity and pH dependencies ($R^2 \leq 0.66$ and 0.50, respectively; Figs S25 and S32) that are greatly reduced in the Meth set ($R^2 \leq$ 0.40 and 0.28, respectively; Figs S19 and S26). This is evidence that many of the conductivity and pH relationships visible in the Full FAs are in fact due to the mathematical mixing of brGDGTs with different cyclization numbers and isomer designations. Holding these variations constant in the Meth FAs largely removes the effects of these environmental variables (e.g. $fIIa_{Meth}$ in Fig. S19d). DO dependencies are weak in both the Full and Meth sets, but slightly weaker in the latter ($R^2 \leq 0.40$ and 0.35, respectively, Figs S39 and S33). The Methylation set thus improves compound-specific correlations with temperature while decreasing their dependencies on other environmental variables.

4.3 Conductivity and pH relationships with brGDGTs

While pH is the traditional secondary target of brGDGT calibrations after temperature, numerous works have suggested that conductivity plays an important role in controlling brGDGT distributions (Tierney et al., 2010; Shanahan et al., 2013). The two variables often plot nearly colinearly in principal component analyses (Shanahan et al., 2013; Dang et al., 2018; Russell et al., 2018), suggesting that they may have similar influences on brGDGT

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021


Biogeosciences Discussions Open Access EGU

distributions. In our dataset, conductivity and pH were moderately correlated ($R^2 = 0.57$) and elicited similar responses in brGDGT FAs. We therefore discuss them together in this section, with an emphasis on the more strongly-correlated of the two, conductivity.

4.3.1 The Isomer set and conductivity

Conductivity provided the strongest compound-specific correlations with brGDGT FAs after temperature. Of the basic brGDGT sets (Meth, Cyc, and Isom), the Isom set had the highest statistical performance (Fig. 8). The FAs of all 6-methyl brGDGTs showed a positive linear correlation with conductivity in this set, with coefficients of determination as high as $R^2 = 0.70$ (Fig. 8d). Furthermore, this positive correlation was broadly independent of both methylation number and cyclization number, indicating that methylation position varies with conductivity irrespective of other structural properties. The 6-methyl brGDGTs were also all positively correlated with pH, but more weakly so ($R^2 \leq 0.54$, Fig. S28). A relationship between brGDGT isomers and pH in lake sediments has been previously observed and quantified by the isomerization of branched tetraethers (IBT, Eq. A14; Ding et al., 2015) and the Isomer Ratio of 6-methyl isomers (IR_{6Me}, Eq. A8; Dang et al., 2016). However, these indices were more closely tied to conductivity (IBT $R^2 = 0.65$, IR_{6Me} $R^2 = 0.66$) than pH (IBT $R^2 = 0.55$, IR_{6Me} $R^2 = 0.49$) in our dataset. These results indicate that isomer abundances are primarily dependent on conductivity, but have some relation to pH as well. Temperature correlations were also present in the Isom subset ($R^2 \leq 0.57$ with MAF, Fig. S14) that were stronger than the inherent correlation between MAF and conductivity in our dataset ($R^2 = 0.48$). While conductivity is the primary control on isomer ratios, temperature may therefore play a secondary role. DO provided little to no correlation in the Isom set ($R^2 \leq 0.38$, Fig. S35).

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

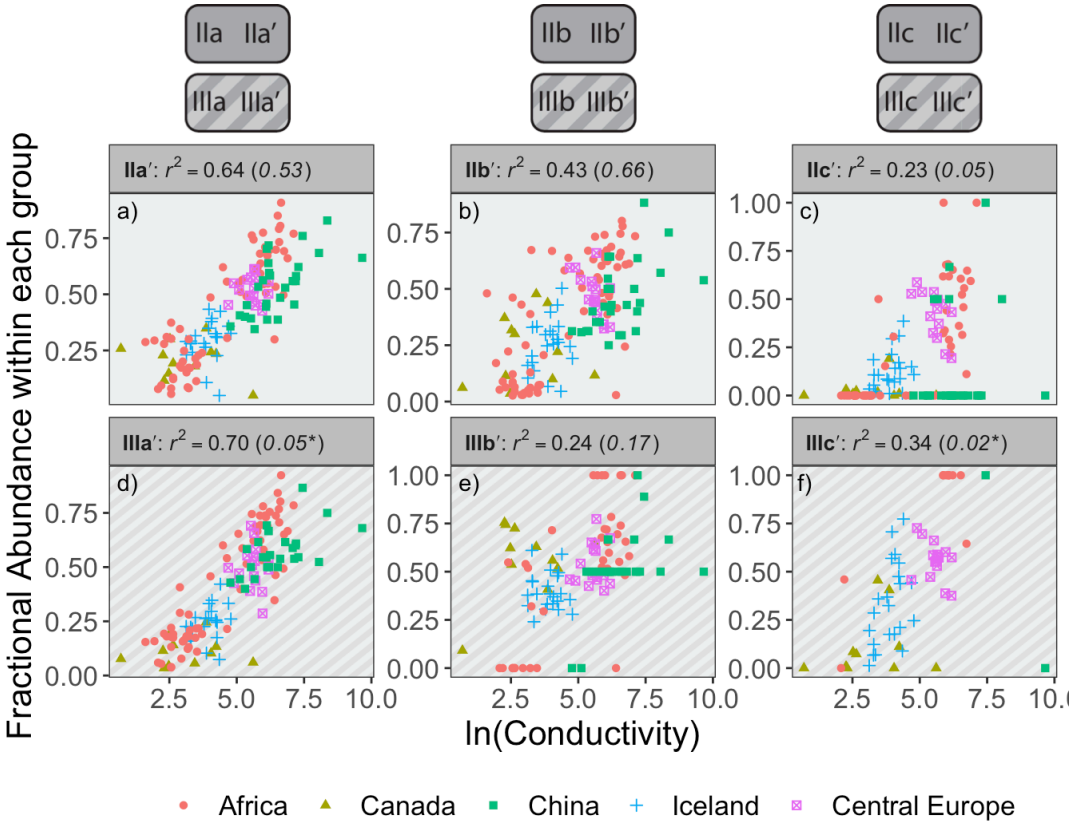


Figure 8. Relationships between the natural logarithm of conductivity and brGDGT FAs calculated within the Isom set. R^2 values are provided for each subplot, with R^2 values for the standard Full FAs given in parentheses for comparison. P-values were < 0.01 except where marked with an asterisk. Note that plots of 5-methyl compounds are redundant because they exactly mirror those of their 6-methyl counterparts; only the 6-methyl FAs are shown.

4.3.2 The Cyclization set and pH

The Cyclization set highlighted the relationship between brGDGTs and pH. As pH increased, all cyclized brGDGTs were found in greater relative abundance (Fig. 9). This result reinforces previous observations that higher ring number is associated with higher pH, as has been quantified by the CBT (Weijers et al., 2007), $\#rings_{tetra}$ (Sinninghe Damsté, 2016), DC (Baxter et al., 2019), and related indices (see Appendix). However, this pH relationship has thus far been demonstrated primarily in soils; the single-compound FAs of the Cyc subset already provide the strongest pH correlations ($R^2 = 0.61$, Fig. 9b) yet reported in lake sediments. The Cyc set also reveals that compounds with structural similarities exhibit analogous responses to pH. For example, all monocyclized brGDGTs show a nonlinear increase with pH (Fig. 9b, e, h, k, n), while uncyclized brGDGTs all exhibit a nonlinear decrease (Fig. 9a, d, g, j, m). These trends are apparent regardless of methylation number or methylation position, suggesting that ring number is broadly independent of both. This independence may imply that alkyl-chain cyclization serves its biological

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

function(s) regardless of the number and position of methylations present. It also allows for the construction of independent pH calibrations for tetra-, penta-, and hexamethylated brGDGTs (Sect. 4.4.2).

Though the Cyc set FAs were most strongly correlated with pH ($R^2 \leq 0.61$), they also exhibited robust relationships with conductivity ($R^2 \leq 0.57$; Fig. S20). All cyclized compounds showed positive correlations with conductivity, and this increase was largely independent of methylation number or position. These results indicate that ring number is primarily dependent on pH, but is correlated with conductivity in a similar manner. The brGDGT indices showed an analogous result; all cyclization indices (CBT and related indices, $\#rings_{tetra}$ and related indices, and DC; see Appendix) correlated most strongly with pH, but also exhibited weaker relationships with conductivity. The Cyc set exhibited little to no correlation with either temperature ($R^2 \leq 0.25$) or DO ($R^2 \leq 0.08$), suggesting that neither of these environmental variables plays an important role in controlling brGDGT cyclization.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

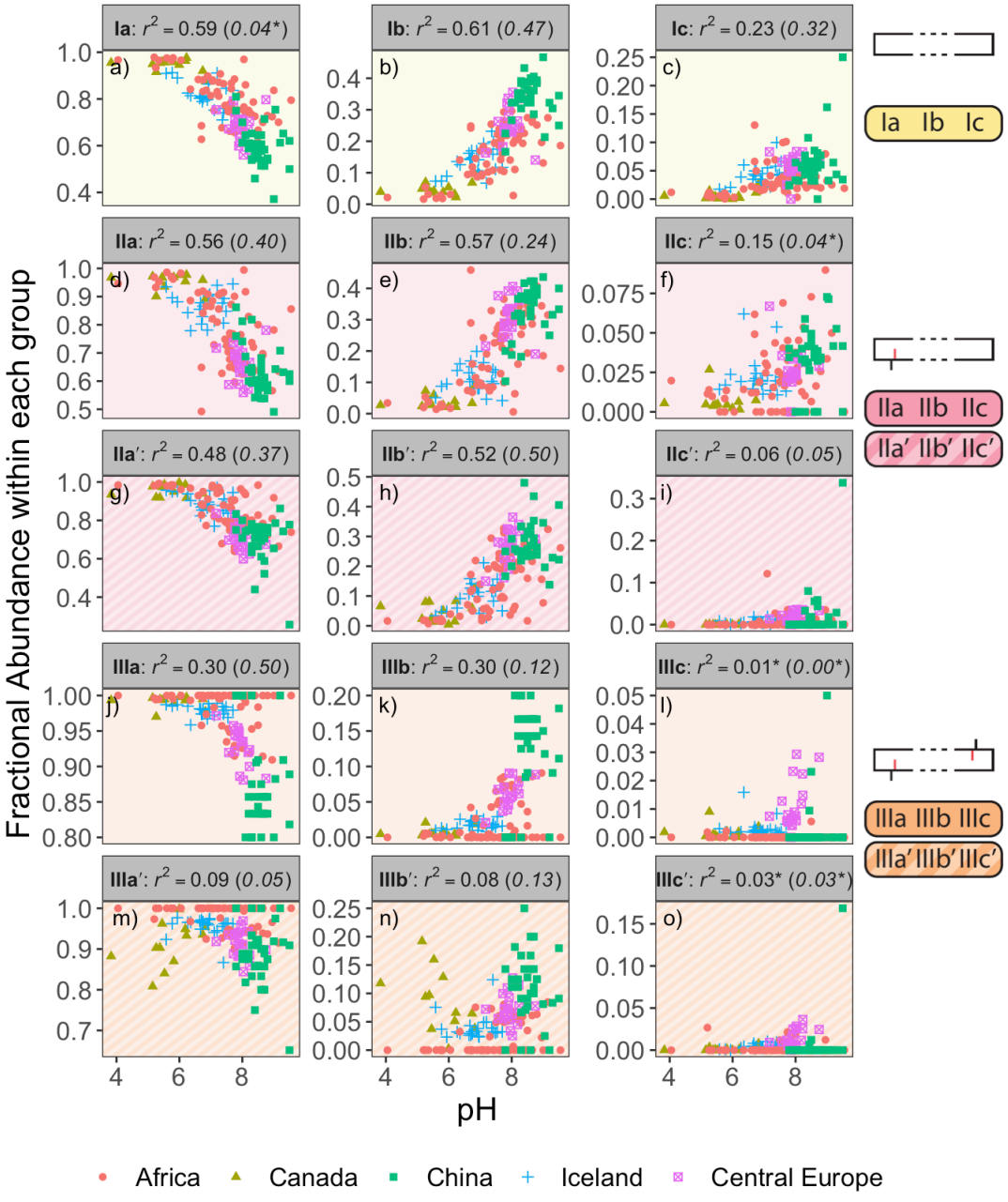


Figure 9. Relationships between pH and brGDGT FAs calculated within the Cyc set. R^2 values are provided for each subplot, with R^2 values for the standard Full FAs given in parentheses for comparison. P-values were < 0.01 except where marked with an asterisk.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021


Biogeosciences Discussions Open Access EGU

4.3.3 The Combined Cyclization-Isomer set strengthens both conductivity and pH trends

Though the strongest conductivity trends were displayed by the Isom set ($R^2 \leq 0.70$), correlations were also present in the Cyc FAs ($R^2 \leq 0.57$, Fig. S20). Similarly, the Cyc set contained the highest pH dependencies ($R^2 \leq 0.61$), but notable relationships were visible in the Isom set as well ($R^2 \leq 0.54$, Fig. S28). To take advantage of all of these conductivity and pH relationships, we therefore used the combined Cyc-Isom set, which holds only methylation number constant while allowing both cyclization number and methylation position to vary (Fig. 2). Both conductivity and pH trends were strengthened in this combination set ($R^2 \leq 0.73$ and 0.62, Fig. S23 and S30, respectively), especially for the uncyclized compounds. However, temperature correlations were also increased in the CI set ($R^2 \leq 0.60$, Fig. S16), a potentially convoluting influence that may not be desired.

4.4 Calibrations

For each set, combined set, and subset defined in Sect. 3, we performed linear and quadratic regressions against temperature, conductivity, pH, dissolved oxygen, and lake geometry variables using SFS/SBE and combinatoric fitting methods. We found temperature and conductivity to provide the strongest empirical calibrations with brGDGTs, followed by pH and dissolved oxygen, and discuss our recommended calibrations below.

4.4.1 Temperature Calibrations

We performed regressions against five temperature variables to generate multiple global-scale calibrations. Of the temperature indices that we tested, MAF provided the fit with the highest statistical significance ($R^2 = 0.91$), followed closely by MST ($R^2 = 0.90$), SWI ($R^2 = 0.89$), and WMT ($R^2 = 0.88$). MAT provided calibrations with high statistical performance as well ($R^2 \leq 0.87$). However, these fits showed clear seasonality biases in their residuals, resulting in substantial over-estimations of MAT for cold sites (Fig. S40). Furthermore, they often relied heavily on low-abundance compounds as fitting variables (IIc, IIb', IIIc, IIIc', IIIb, and IIIb'). We therefore do not recommend a brGDGT MAT calibration and focus our discussion on the warm-season temperature indices, especially MAF.

Methylation number was the single most important structural variable for temperature calibrations. The Meth set, which allowed only methylation number to vary, provided a MAF calibration ($R^2 = 0.90$, Fig. 10b) that was on par with other recent global and regional lake sediment calibrations ($R^2 = 0.85$ to 0.94; Dang et al., 2018; Russell et al., 2018; Martínez-Sosa et al., 2020b). The MI, MC, and Full sets, which additionally allowed for changes in cyclization number and/or methylation position, added little to the calibration performance ($R^2 = 0.90$ to 0.91, Fig. 10a). Furthermore, sets which held methylation number constant – Cyc, Isom, and CI – performed markedly worse ($R^2 = 0.51$, 0.63, and 0.67, respectively, Fig. 10a). These results indicate that effectively all of the temperature dependence of the 15 commonly-measured brGDGTs is captured by methylation number alone.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

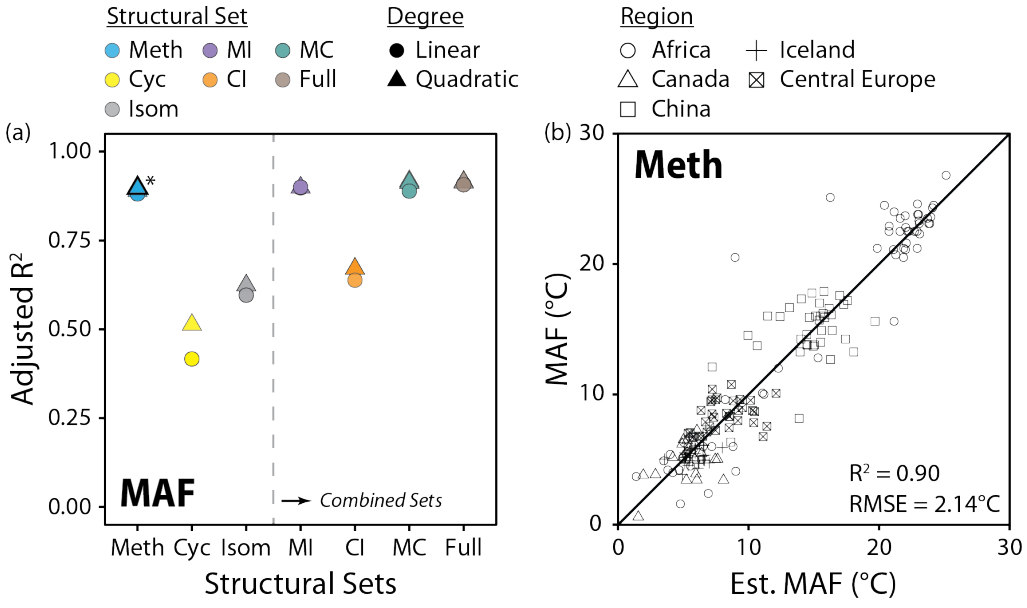


Figure 10. a) Performance (adjusted R^2) of all linear and quadratic fits for the mean air temperature of months above freezing (MAF) and brGDGT fractional abundances (FAs) calculated within the basic (Meth, Cyc, Isom; left of dashed line) and combined (Meth-Isom (MI), Cyc-Isom (CI), Meth-Cyc (MC), and Full sets; right of dashed line) structural sets. Results of both the SFS/SBE and combinatoric fitting methods are plotted. The fit we suggest for general use (Meth set, quadratic, SFS/SBE; Eq. 10) is bolded and marked with an asterisk in a) and plotted in b). "Est. MAF" is the MAF temperature estimated using this suggested fit.

Within the Meth set, independent calibrations were generated for the $Meth_a$, $Meth_b$, and $Meth_c$ subsets as well. These subsets consist of only un-, mono-, and bicylcized compounds, respectively (Fig. 2a). As the FAs of these subsets were calculated independently, we were able to test for temperature calibrations of each subset alone. The $Meth_a$ subset provided the strongest of the subset MAF fits ($R^2 = 0.88$). This calibration has the notable advantage of employing only the 3 most abundant brGDGTs typically found in nature, Ia, IIa, and IIIa, which may allow for temperatures to be reproduced with high fidelity from even organic-lean samples. Furthermore, since it uses only non-cyclized, 5-methyl brGDGTs, it may be less subject to influence by the environmental factors that impact cyclization numbers and isomer ratios. MAF calibrations were also obtained using only the monocyclized ($Meth_b$, $R^2 = 0.79$) and bicyclized ($Meth_c$, $R^2 = 0.74$) brGDGTs. These fits represent, to our knowledge, the first calibrations that make use of only cyclized brGDGTs. This is a noteworthy result as it shows conclusively what can be seen by eye in Fig. 7; the relationship between temperature and methylations is a broad feature of brGDGTs that is present regardless of the number of rings on the carbon backbone. These un-, mono-, and bicyclized calibrations may find use in the case that one or more brGDGTs are suspected to be influenced by variables other than temperature.

Of the brGDGT temperature indices that we tested (see Appendix), MBT'_{5Me} performed best. This index correlated better with MAF ($R^2 = 0.89$) than any other warm-season variable (SWI $R^2 = 0.84$; MST $R^2 = 0.70$; MAT $R^2 = 0.70$; WMT $R^2 = 0.66$). The slope and intercept of the MAF/MBT'_{5Me} calibration (MAF = -0.5 (± 0.4) + 30.4 (± 0.8) * MBT'_{5Me}) were similar to the MAT/MBT'_{5Me} calibration presented by Russell et al. (2018) (MAT = -1.21 + 32.42 * MBT'_{5Me}). This may suggest that MBT'_{5Me}-derived temperatures using the Russell et al. (2018) calibration in

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

cold regions are best considered to reconstruct MAF rather than MAT. Though MBT'$_{6Me}$ was previously found to correlate well with temperature on a regional scale (R^2 = 0.75; Dang et al., 2018), it was not correlated with any temperature variable in our global dataset ($R^2 \leq 0.12$). Finally, we note that the Community Index (De Jonge et al., 2019), which was associated with bacterial community changes in geothermally heated Icelandic soils, is identical to our fIa$_{Meth}$. This may suggest that the strong connection between fIa$_{Meth}$ and MAF could also be driven at least in part by changes in microbial community composition. However, genomic data is not currently available for the majority of the sites in this study to test this hypothesis.

4.4.2 Conductivity and pH Calibrations

Conductivity outperformed pH in our calibrations (R^2 = 0.83 versus 0.74) and was the second most important predictor of brGDGT distributions in our dataset after temperature. Both cyclization number and methylation position were important to the success of these calibrations. The Cyc set, which allowed only cyclization number to change, provided a conductivity fit with R^2 = 0.73 (Fig. 11a). The Isom set, which isolated trends in isomer abundances, generated a slightly stronger calibration with R^2 = 0.76. When both of these structural properties were allowed to vary together in the CI set, the calibration was markedly improved (R^2 = 0.83, Fig. 11b). In contrast, methylation number was a poorer predictor of conductivity alone (Meth set R^2 = 0.65) and did not improve upon the CI correlation in the combined sets (MI R^2 = 0.80; MC R^2 = 0.83; Full R^2 = 0.81). The CI$_{III}$ and CI$_{II}$ subsets also provided conductivity calibrations with relatively high statistical performance (R^2 = 0.75, 0.73). The CI$_I$ subset performed worse (R^2 = 0.65), likely due to the fact that no isomer variations are present in these FAs. These subset-specific fits are the first of their kind, and their success indicates that the relationship between brGDGTs and conductivity is present regardless of methylation number.

In addition to conductivity, the CI set also best captured the relationship between brGDGTs and pH (R^2 = 0.73; Fig. 11c and d). The addition of methylation number variations in the Full set did not substantially improve the calibration (R^2 = 0.74), indicating that the majority of the relationship between brGDGTs and pH is captured by cyclization number and methylation position. Of these two structural variables, cyclization number was more important for pH; the Cyc set calibration (R^2 = 0.67) outperformed those from the Isom and Meth sets (R^2 = 0.59 for both). The CI$_I$, CI$_{II}$, and CI$_{III}$ subsets provided weaker, but significant calibrations with comparable performances to one another (R^2 = 0.67, 0.68, and 0.62, respectively), indicating again that pH relationships are more or less independent of methylation number.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

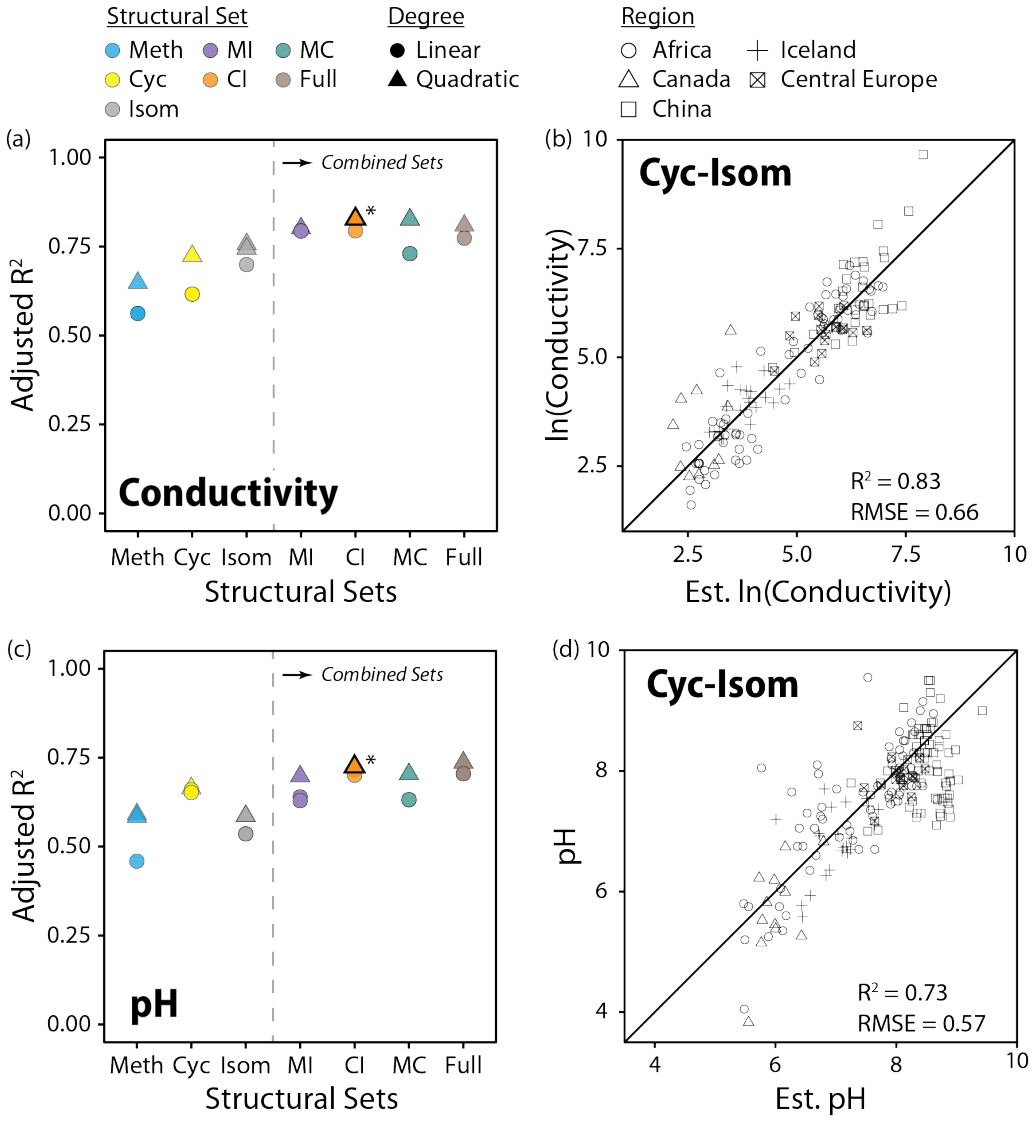


Figure 11. Performance (adjusted R^2) of all linear and quadratic fits for lake water conductivity (a) and pH (c) and brGDGT fractional abundances (FAs) calculated within the basic (Meth, Cyc, Isom; left of dashed lines) and combined (Meth-Isom (MI), Cyc-Isom (CI), Meth-Cyc (MC), and Full sets; right of dashed lines) structural sets. Results of both the SFS/SBE and combinatoric fitting methods are plotted. The fits we suggest for general use (CI set, quadratic, combinatoric, for both variables; Eqs. 12-13) is bolded and marked with an asterisk in a) and c) and plotted in b) and d). "Est. ln(Conductivity)" and "Est. pH" are the natural logarithm of lake water conductivity and pH estimated using these suggested fits.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021


Biogeosciences Discussions Open Access EGU

580 **4.4.3 Dissolved Oxygen and Lake Geometry Calibrations**

581 There is increasing evidence that oxygen availability strongly affects lacustrine brGDGT distributions
582 (Colcord et al., 2017; Weber et al., 2018; van Bree et al., 2020; Yao et al., 2020). We therefore tested for calibrations
583 within our dataset with mean and minimum dissolved oxygen concentration (DO_{mean} and DO_{min}). As lake morphology
584 can be an important predictor of lake oxygen levels (Hutchinson, 1938; Nürnberg, 1995), we also tested the natural
585 logarithms of maximum water depth (Depth), the ratio of lake surface area to maximum depth (SA/D), and
586 approximate lake volume.

587 None of the DO or lake geometry variables generated strong brGDGT calibrations ($R^2 \leq 0.63$; Table S1).
588 The highest-performing fit was provided by the Meth set with DO_{mean} ($R^2 = 0.63$, Fig. S41). Moderate correlations
589 were found with DO_{min} as well (Cyc set, $R^2 = 0.55$). Lake depth alone was a poor predictor of brGDGT distributions
590 ($R^2 = 0.35$), but both volume and the ratio of surface area to depth were found to provide moderate correlations ($R^2 =$
591 0.55 and 0.59, respectively). However, none of these lake morphology variables was itself well-correlated with DO_{mean}
592 or DO_{min} ($R^2 \leq 0.22$) in this dataset, and we therefore cannot explain their relationship with brGDGTs at this time.
593 Additionally, although the HP5 index (Eq. A13) was recently shown to reflect redox conditions via a correlation with
594 lake water depth (Yao et al., 2020), it does not correlate with any of our lake geometry indices ($R^2 \leq 0.02$) and only
595 weakly correlates with DO ($R^2 \leq 0.28$) in this dataset, indicating that it may be primarily useful for within-lake studies.

596 The Meth set provided both the strongest DO_{mean} and MAF calibrations, raising the possibility that DO may
597 have a problematic influence on that calibration's temperature estimates. Individual Meth FAs were weakly correlated
598 with DO_{mean} at best ($R^2 \leq 0.35$), however, and the residuals of the Meth/MAF fit in Eq. 10 showed no correlation with
599 DO_{mean} ($R^2 = 0.01$, $p = 0.2$). Given these weak relationships, we do not see evidence for the influence of DO on
600 temperatures reconstructed with the Meth calibration in our dataset.

601 In light of increasing evidence that oxygen availability strongly affects lacustrine brGDGT distributions, it is
602 perhaps surprising that we do not find a stronger correlation between brGDGTs and DO. However, the effects of DO
603 on brGDGT distributions appear to be highly site-specific. For example, some detailed studies have found elevated
604 levels of brGDGT-IIIa in low oxygen conditions (Weber et al., 2018; Yao et al., 2020), but another found all of the
605 most common brGDGTs *except* IIIa in abundance in the oxygen-depleted hypolimnion (van Bree et al., 2020). A third
606 detailed study found no correlation between brGDGTs and oxygen at all (Loomis et al., 2014b), and no calibration
607 study to date has found regional or global trends. The wide range of possible drivers of DO – mixing regimes,
608 eutrophication state, and ice cover, to name a few – may play a role in the incoherent relationship between brGDGTs
609 and DO in these studies and our own. Additionally, DO measurements are often taken at the time of sampling and are
610 most likely not representative of the annual range. For the Canadian and Icelandic lakes in this study, for example,
611 DO was depleted under lake ice relative to ice-free conditions in 10 out of 11 cases. Few of the lakes in our study have
612 continuous DO monitoring data available, and most are from Central Europe. Therefore, while our study does find
613 significant correlations between brGDGTs and DO, we do not recommend a calibration for general use and instead
614 highlight the need for further study.
615

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

4.4.4 Recommended Calibrations

The Meth, MC, MI, and Full subsets all provided MAF temperature calibrations with comparable R^2 (0.90 to 0.91) and RMSE (1.97 to 2.14°C) values (Fig. 10a, Table S1). However, the Meth set provided the FAs with the strongest compound-specific relationships with MAF ($R^2 \leq 0.88$) in the modern dataset and exhibited little influence ($R^2 \leq 0.49$) from any other variable examined in this study. In contrast, the MC, MI, and Full sets all exhibited stronger relationships with pH and conductivity (Fig. 11a and c). We therefore recommend the highest-performing Meth calibration (Fig. 10b; Eq. 10; n = 182, $R^2 = 0.90$, RMSE = 2.14°C) for general use in lake sediments:

$$\begin{aligned} MAF\ (°C) = {} & 92.9(\pm15.98) + 63.84(\pm15.58) \times fIb^2_{Meth} - 130.51(\pm30.73) \times fIb_{Meth} \\ & - 28.77(\pm5.44) \times fIIa^2_{Meth} - 72.28(\pm17.38) \times fIIb^2_{Meth} - 5.88(\pm1.36) \times fIIc^2_{Meth} \\ & + 20.89(\pm7.69) \times fIIIa^2_{Meth} - 40.54(\pm5.89) \times fIIIa_{Meth} \\ & - 80.47(\pm19.19) \times fIIIb_{Meth} \end{aligned} \quad (10)$$

The Full set MAF calibration provided the highest R^2 and lowest RMSE in our dataset. This fit may be applicable in settings with good conductivity or pH control and may be useful for comparison with previous calibrations. It is therefore provided in Eq. (11) (n = 182, $R^2 = 0.91$, RMSE = 1.97°C):

$$\begin{aligned} MAF\ (°C) = {} & -8.06(\pm1.56) + 37.52(\pm2.35) \times fIa_{Full} - 266.83(\pm98.61) \times fIb^2_{Full} \\ & + 133.42(\pm19.51) \times fIb_{Full} + 100.85(\pm9.27) \times fIIa'^2_{Full} + 58.15(\pm10.09) \times fIIIa'^2_{Full} \\ & + 12.79(\pm2.89) \times fIIIa_{Full} \end{aligned} \quad (11)$$

The calibrations presented in Eqs. (10) and (11) allow for the quantitative reconstruction of warm-season air temperatures from lake sediment archives, including those at high latitudes. The statistical performance of these fits is comparable to recently-published calibrations (Russell et al. (2018): $R^2 = 0.94$, RMSE = 2.14°C, n = 65; Martínez-Sosa et al. (2020): $R^2 = 0.85$, RMSE = 2.8°C, n = 261; Dang et al. (2018): $R^2 = 0.91$, RMSE = 1.10°C, n = 39). Subset-specific calibrations ($Meth_a$, $Meth_b$, and $Meth_c$) are also statistically comparable and are available in the Supplement (Eq. S1-3). We do not recommend an independent MAT calibration due to residual seasonality biases (Fig. S40), though we note that MAT is often identical to MAF in warm or low-seasonality settings.

The CI set generated the highest-performing conductivity calibration ($R^2 = 0.83$, RMSE = 0.66; Fig. 11b, Table S1). The Full, MC, and MI sets provided calibrations that were statistically comparable ($R^2 = 0.80$ to 0.83, RMSE = 0.65 to 0.70), but their FAs contained marked temperature dependencies ($R^2 \leq 0.70$, 0.77, and 0.75, respectively; Figs S18 , S17 and S15). We therefore recommend the top CI fit for use (Eq. 12; n = 143, $R^2 = 0.83$, RMSE = 0.66), and provide subset-specific calibrations for CI_I, CI_{II}, and CI_{III} in the Supplement (Eq. S4-6):

$$\begin{aligned} ln(Cond.) = {} & 6.62(\pm1.01) + 8.87(\pm1.24) \times fIb_{CI} + 5.12(\pm1.54) \times fIIa'^2_{CI} + 10.64(\pm1.88) \times fIIa^2_{CI} \\ & - 8.59(\pm2.21) \times fIIa_{CI} - 4.32(\pm1.46) \times fIIIa'^2_{CI} - 5.31(\pm0.95) \times fIIIa^2_{CI} \\ & - 142.67(\pm36.08) \times fIIIb^2_{CI} \end{aligned} \quad (12)$$

A brGDGT-based paleoconductivity reconstruction has yet to be attempted, but diatom-inferred conductivity records show the potential for this variable to provide valuable insight into changes in lake hydrology. These records have reconstructed changes in precipitation and evaporation balance, lake level fluctuations, meltwater influx events, and the isolation of a lake from the sea (Ng and King, 1999; Yang et al., 2004; Stager et al., 2013). The conductivity

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

652 calibration in Eq. (12) thus enables brGDGTs to be tested as a new alternative or complementary proxy in
653 paleohydrology reconstructions.
654 The Full and CI sets provided calibrations with pH that were statistically comparable to one another (R^2 =
655 0.74 and 0.73, Fig. 11c, Table S1). Given the stronger temperature dependencies of the Full set, we recommend the
656 CI calibration (Fig. 11d; Eq. 13; n = 154, R^2 = 0.73, RMSE = 0.57) for use and provide its subset-specific calibrations
657 (CI_I, CI_{II}, and CI_{III}) in the Supplement (Eq. S7-9):

658 $$pH = 8.93(\pm 0.21) - 3.84(\pm 0.25) \times fIa_{CI}^2 + 2.63(\pm 0.35) \times fIIa'_{CI} \quad (13)$$

659 The residuals of the fit in Eq. (13) have a weak but significant correlation with pH (R^2 = 0.17), causing it to
660 overestimate pH for acidic samples and underestimate it for alkaline ones. A similar bias in pH calibrations has been
661 previously observed in soils (De Jonge et al., 2014a). We therefore caution the use of this calibration in acidic (pH <
662 5) or alkaline (pH > 9) conditions.
663 Previous work has demonstrated the value of brGDGT-derived pH records in studies of terrestrial
664 paleoclimate (Tyler et al., 2010; Cao et al., 2017; Fastovich et al., 2020). However, these studies relied on calibrations
665 generated from soils and/or analyses in which the 5- and 6-methyl isomers were not separated. The fit presented in
666 Eq. (13) may improve such studies by providing a globally-distributed pH calibration in lake sediments using the latest
667 chromatographic methods which improves upon the error of previously available calibrations (RMSE = 0.80, Russell
668 et al., 2018).
669 Dissolved oxygen and lake geometry calibrations generated significant, but statistically weaker fits ($R^2 \leq$
670 0.63). Due to the low R^2 of these calibrations and an incomplete understanding of the relationship between DO and
671 brGDGT distributions, we do not recommend their application at this time. However, the equation for the highest-
672 performing variable, DO_{mean}, is provided in the Supplement for reference (Eq. S10).

673 **5 Conclusions**

674 We have shown that brGDGT structural sets and warm-season temperature indices improve correlations with
675 environmental parameters while advancing our biological understanding of the lipids themselves. Grouping brGDGTs
676 into structural sets based on methylation number, methylation position, and cyclization number elucidated the
677 relationships between environmental variables and brGDGT structures. These sets revealed that methylation number
678 fully captures the relationship between brGDGT distributions and temperature. They also showed the relative
679 abundance of 5- and 6-methyl isomers to be dependent on conductivity and cyclization number to be primarily tied to
680 pH. The deconvolved relationships provided by these subsets allowed for the generation of calibrations with
681 temperature and pH that relied on fewer compounds with robust modern trends in a global dataset. They additionally
682 revealed conductivity to be the second-most important variable in controlling brGDGT distributions and provided a
683 calibration for this oft-overlooked variable, which may find use as a proxy for precipitation/evaporation balances or
684 hydrologic changes.
685 The structural sets also provided insight into the biological underpinnings of brGDGT structural diversity.
686 The Meth, Cyc, and Isom sets gave evidence that methylation number, cyclization number, and methylation position
687 vary more or less independently of one another across environmental gradients. They further revealed that the
688 inclusion of tetramethylated compounds (Ia, Ib, Ic) enhances the temperature dependencies of 5-methyl compounds,

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

but erases those of their 6-methyl counterparts. As the microbial producers of brGDGTs have yet to be identified and cultured, the structural set approach thus provides a valuable tool for investigating controls on brGDGT diversity with a biological lens.

Warm-season temperatures outperformed MAT as the most important predictors of brGDGT distributions. We introduced 43 new lake sediment samples from sites with low MAT and high seasonality. In conjunction with a global dataset, these samples showed a clear warm-season bias in brGDGT temperature relationships, with MAF providing the strongest fits in our dataset. The warm-season bias may suggest a direct or indirect connection to heightened primary productivity in the summer. Alternatively, it may be the result of a more complex relationship with dynamic lake processes such as mixing events. While further study is needed to unravel these complications, the strong empirical calibrations presented here support the use of the brGDGT paleotemperature proxy to quantitatively reconstruct warm-season air temperatures from high-latitude lake sediments.

In summary, the use of brGDGT structural sets and warm-season temperature indices deconvolved relationships between brGDGT structure and environmental gradients, revealed trends with biological implications, and tied brGDGT distributions to warm-season temperatures. Furthermore, they allowed for the construction of improved temperature and pH calibrations as well as the first brGDGT conductivity calibration in a global lake sediment dataset. Our results thus allow brGDGTs to be used to quantitatively reconstruct warm-season air temperatures and lake water conductivity and pH from lake sediment archives and provide a new methodology for the study of brGDGTs in the future.

Appendix

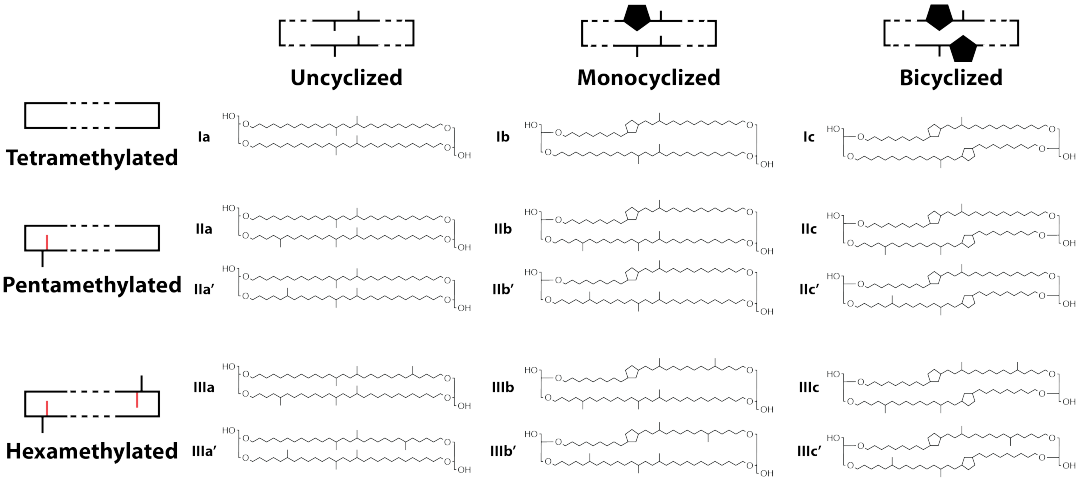


Figure A1. Structures of the 15 commonly-measured brGDGTs along with their schematic representations, with C6 methylations denoted in red in the schematics.

Full and schematic structures of the 15 commonly-measured brGDGTs are provided in Fig. A1. Table A1 details equations for calculating FAs within the structural sets.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

(Sub)set name	(Sub)set compounds	Fractional abundance equation
Full	S = {Ia, Ib, Ic, IIa, IIb, IIc, IIIa, IIIb, IIIc, IIa', IIb', IIc', IIIa', IIIb', IIIc'}	$fxy_{[(sub)set\ name]} = xy/\sum S$
MC-5Me+	S = {Ia, Ib, Ic, IIa, IIb, IIc, IIIa, IIIb, IIIc}	
MC-6Me+	S = {Ia, Ib, Ic, IIa', IIb', IIc', IIIa', IIIb', IIIc'}	
MC-5Me	S = {IIa, IIb, IIc, IIIa, IIIb, IIIc}	
MC-6Me	S = {IIa', IIb', IIc', IIIa', IIIb', IIIc'}	
MC	*Use MC-5Me+ and MC-6Me FAs*	
MI	S_a = {Ia, IIa, IIIa, IIa', IIIa'} S_b = {Ib, IIb, IIIb, IIb', IIIb'} S_c = {Ic, IIc, IIIc, IIc', IIIc'}	$fxy_{[(sub)set\ name]} = xy/\sum S_y$
Meth-5Me+	S_a = {Ia, IIa, IIIa} S_b = {Ib, IIb, IIIb} S_c = {Ic, IIc, IIIc}	
Meth-6Me+	S_a = {Ia, IIa', IIIa'} S_b = {Ib, IIb', IIIb'} S_c = {Ic, IIc', IIIc'}	
Meth-5Me	S_a = {IIa, IIIa} S_b = {IIb, IIIb} S_c = {IIc, IIIc}	
Meth-6Me	S_a = {IIa', IIIa'} S_b = {IIb', IIIb'} S_c = {IIc', IIIc'}	
Meth	*Use Meth-5Me+ and Meth-6Me FAs*	
CI	S_I = {Ia, Ib, Ic} S_{II} = {IIa, IIb, IIc, IIa', IIb', IIc'} S_{III} = {IIIa, IIIb, IIIc, IIIa', IIIb', IIIc'}	$fxy_{[(sub)set\ name]} = xy/\sum S_x$
Cyc-5Me	S_I = {Ia, Ib, Ic} S_{II} = {IIa, IIb, IIc} S_{III} = {IIIa, IIIb, IIIc}	
Cyc-6Me	S_I = {Ia, Ib, Ic} S_{II} = {IIa', IIb', IIc'} S_{III} = {IIIa', IIIb', IIIc'}	
Cyc	*Use Cyc-5Me and Cyc-6Me FAs*	
Isom	S_{IIa} = {IIa, IIa'}; S_{IIIa} = {IIIa, IIIa'} S_{IIb} = {IIb, IIb'}; S_{IIIb} = {IIIb, IIIb'} S_{IIc} = {IIc, IIc'}; S_{IIIc} = {IIIc, IIIc'}	$fxy_{[(sub)set\ name]} = xy/\sum S_{xy}$

Table A1. Equations for calculating FAs within brGDGT subsets, where f*xy* and *xy* are the fractional and absolute abundances of the 5- or 6-methyl brGDGT with Roman numeral *x* (I, II, or III) and alphabet letter *y* (a, b, or c).

We used the following previously-defined brGDGT indices in this study: CBT (Weijers et al., 2007); MBT' (Peterse et al., 2012); MBT'_{5Me}, MBT'_{6Me}, CBT_{5Me}, CBT', and Index1 (De Jonge et al., 2014a); IR_{6Me} (Dang et al., 2016); $\#rings_{tetra}$, $\#rings_{penta\ 5Me}$, and $\#rings_{penta\ 6Me}$ (Sinninghe Damsté, 2016); Degree of Cyclization (DC) as reformulated by (Baxter et al., 2019); HP5 (Yao et al., 2020), isomerization of branched tetraethers (IBT; Ding et al., 2015), community index (CI; De Jonge et al., 2019). Their equations are given below:

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

$$CBT = -log\left(\frac{Ib + IIb + IIb'}{Ia + IIa + IIa'}\right) \quad (A1)$$

$$MBT' = \frac{(Ia + Ib + Ic)}{(Ia + Ib + Ic + IIa + IIb + IIc + IIIa + IIa' + IIb' + IIc' + IIIa')} \quad (A2)$$

$$MBT'_{5Me} = \frac{(Ia + Ib + Ic)}{(Ia + Ib + Ic + IIa + IIb + IIc + IIIa)} \quad (A3)$$

$$MBT'_{6Me} = \frac{(Ia + Ib + Ic)}{(Ia + Ib + Ic + IIa' + IIb' + IIc' + IIIa')} \quad (A4)$$

$$CBT_{5Me} = -log\left(\frac{Ib + IIb}{Ia + IIa}\right) \quad (A5)$$

$$CBT' = -log\left(\frac{Ic + IIa' + IIb' + IIc' + IIIa' + IIIb' + IIIc'}{Ia + IIa + IIIa}\right) \quad (A6)$$

$$Index1 = log\left(\frac{Ia + Ib + Ic + IIa' + IIIa'}{Ic + IIa + IIc + IIIa + IIIa'}\right) \quad (A7)$$

$$IR_{6Me} = \frac{(IIa' + IIb' + IIc' + IIIa' + IIIb' + IIIc')}{(IIa' + IIb' + IIc' + IIIa' + IIIb' + IIIc' + IIa + IIb + IIc + IIIa + IIIb + IIIc)} \quad (A8)$$

$$\#rings_{tetra} = \frac{(Ib + 2 * Ic)}{(Ia + Ib + Ic)} \quad (A9)$$

$$\#rings_{penta\ 5Me} = \frac{(IIb + 2 * IIc)}{(IIa + IIb + IIc)} \quad (A10)$$

$$\#rings_{penta\ 6Me} = \frac{(IIb' + 2 * IIc')}{(IIa' + IIb' + IIc')} \quad (A11)$$

$$DC = \frac{(Ib + 2 * Ic + IIb + IIb')}{(Ia + Ib + Ic + IIa + IIa' + IIb + IIb')} \quad (A12)$$

$$HP5 = \frac{IIIa}{(IIa + IIIa)} \quad (A13)$$

$$IBT = -log\left(\frac{IIa' + IIa'}{IIa + IIIa}\right) \quad (A14)$$

$$CI = \frac{Ia}{(Ia + IIa + IIIa)} \quad (= fIa_{Meth}) \quad (A15)$$

Code and data availability

Biomarker and associated metadata is provided in the Supplement and will be archived at the PANGAEA data repository. Code for generating and plotting set-specific FAs in R is provided in the Supplement. Additional code, data, and calibration equations will be available upon request.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

745 **Author contribution**

746 GHM, JS, ÁG, JHR, SEC, DJH, and GdW designed the study and carried out the sampling. GHM, JS, ÁG, SK, and
747 SEC funded the research. JHR, AB, and DJH performed the laboratory work and processed the HPLC-MS data under
748 the supervision of JS. JHR and SK generated the R code. JHR analyzed the data and interpreted the results. JHR wrote
749 the manuscript with input from all authors. All authors contributed to the article and approved the submitted version.
750

751 **Competing interests**

752 The authors declare that they have no conflict of interest.
753

754 **Acknowledgements**

755 This work was supported by the National Science Foundation (OPP-1737712 to GHM and JS; OPP-1836981to GHM,
756 ÁG, and JS; DDRI-1657743 to GHM and SEC; EAR-1945484 to SK), a Doctoral Grant from the University of Iceland
757 and a project grant from the University of Iceland Research Fund to ÁG, a National Geographic Society Early Career
758 Grant (#CP-019ER-17 to SEC), and the University of Colorado Boulder. We thank the Inuit of Nunavut for permitting
759 access to their land and to sample soils and lake sediment (Scientific Research Licenses 01022 17R-M, 02034 18R-
760 M, 02038 19R-M) and the Qikiqtaani Inuit of Qikiqtarjuaq and Clyde River for assistance in the field. We thank the
761 Nunavut Research Institute for logistical assistance and Polar Continental Shelf Project for air support. Field research
762 in Iceland benefited from the assistance of Ian Holmen, Sveinbjörn Steinthorsson, and Thor Blöndahl. We thank Yuki
763 Weber and Shucheng Xie for providing additional temperature data, James Russell for a valuable discussion, and Lina
764 Pérez-Angel for assistance with the calibration code. We additionally thank Nadia Dildar, Sebastian Cantarero, and
765 Katie Rempfert for laboratory assistance. The manuscript benefited from field assistance and/or discussions with
766 Martha Raynolds, Shawnee Gowan, Helga Bueltmann, Elizabeth Thomas, Devon Gorby, Kayla Hollister, Kurt
767 Lindberg, Nicolò Ardenghi, and Jamie McFarlin.
768

769 **Literature Cited**

770 Baxter, A.J., Hopmans, E.C., Russell, J.M., Sinninghe Damsté, J.S., 2019. Bacterial GMGTs in East African lake
771 sediments: Their potential as palaeotemperature indicators. Geochimica et Cosmochimica Acta 259, 155–169.
772 Beales, N., 2004. Adaptation of Microorganisms to Cold Temperatures, Weak Acid Preservatives, Low pH, and
773 Osmotic Stress: A Review. Comprehensive Reviews in Food Science and Food Safety 3, 1–20.
774 Cao, J., Rao, Z., Jia, G., Xu, Q., Chen, F., 2017. A 15 ka pH record from an alpine lake in north China derived from
775 the cyclization ratio index of aquatic brGDGTs and its paleoclimatic significance. Organic Geochemistry 109,
776 31–46.
777 Cao, J., Rao, Z., Shi, F., Jia, G., 2020. Ice formation on lake surfaces in winter causes warm-season bias of lacustrine
778 brGDGT temperature estimates. Biogeosciences 17, 2521–2536.
779 Colcord, D.E., Pearson, A., Brassell, S.C., 2017. Carbon isotopic composition of intact branched GDGT core lipids

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

780 in Greenland lake sediments and soils. Organic Geochemistry 110, 25–32.
781 Crump, S.E., Miller, G.H., Power, M., Sepúlveda, J., Dildar, N., Coghlan, M., Bunce, M., 2019. Arctic shrub
782 colonization lagged peak postglacial warmth: Molecular evidence in lake sediment from Arctic Canada. Global
783 Change Biology 1–13.
784 Dang, X., Ding, W., Yang, H., Pancost, R.D., Naafs, B.D.A., Xue, J., Lin, X., Lu, J., Xie, S., 2018. Different
785 temperature dependence of the bacterial brGDGT isomers in 35 Chinese lake sediments compared to that in
786 soils. Organic Geochemistry 119, 72–79.
787 Dang, X.Y., Xue, J.T., Yang, H., Xie, S.C., 2016. Environmental impacts on the distribution of microbial tetraether
788 lipids in Chinese lakes with contrasting pH: Implications for lacustrine paleoenvironmental reconstructions.
789 Science China Earth Sciences.
790 De Jonge, C., Hopmans, E.C., Stadnitskaia, A., Rijpstra, W.I.C., Hofland, R., Tegelaar, E., Sinninghe Damsté, J.S.,
791 2013. Identification of novel penta- and hexamethylated branched glycerol dialkyl glycerol tetraethers in peat
792 using HPLC-MS2, GC-MS and GC-SMB-MS. Organic Geochemistry 54, 78–82.
793 De Jonge, C., Hopmans, E.C., Zell, C.I., Kim, J.H., Schouten, S., Sinninghe Damsté, J.S., 2014a. Occurrence and
794 abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate
795 reconstruction. Geochimica et Cosmochimica Acta 141, 97–112.
796 De Jonge, C., Radujković, D., Sigurdsson, B.D., Weedon, J.T., Janssens, I., Peterse, F., 2019. Lipid biomarker
797 temperature proxy responds to abrupt shift in the bacterial community composition in geothermally heated soils.
798 Organic Geochemistry 137. doi:10.1016/j.orggeochem.2019.07.006
799 De Jonge, C., Stadnitskaia, A., Hopmans, E.C., Cherkashov, G., Fedotov, A., Sinninghe Damsté, J.S., 2014b. In situ
800 produced branched glycerol dialkyl glycerol tetraethers in suspended particulate matter from the Yenisei River,
801 Eastern Siberia. Geochimica et Cosmochimica Acta 125, 476–491.
802 de Wet, G.A., Castañeda, I.S., DeConto, R.M., Brigham-Grette, J., 2016. A high-resolution mid-Pleistocene
803 temperature record from Arctic Lake El'gygytgyn: A 50 kyr super interglacial from MIS 33 to MIS 31? Earth
804 and Planetary Science Letters 436, 56–63.
805 Dearing Crampton-Flood, E., Peterse, F., Munsterman, D., Sinninghe Damsté, J.S., 2018. Using tetraether lipids
806 archived in North Sea Basin sediments to extract North Western European Pliocene continental air temperatures.
807 Earth and Planetary Science Letters 490, 193–205.
808 Dillon, J.T., Lash, S., Zhao, J., Smith, K.P., van Dommelen, P., Scherer, A.K., Huang, Y., 2018. Bacterial tetraether
809 lipids in ancient bones record past climate conditions at the time of disposal. Journal of Archaeological Science
810 96, 45–56.
811 Ding, S., Kohlhepp, B., Trumbore, S., Küsel, K., Totsche, K.U., Pohnert, G., Gleixner, G., Schwab, V.F., 2018. In
812 situ production of core and intact bacterial and archaeal tetraether lipids in groundwater. Organic Geochemistry
813 126, 1–12.
814 Ding, S., Schwab, V.F., Ueberschaar, N., Roth, V.-N., Lange, M., Xu, Y., Gleixner, G., Pohnert, G., 2016.
815 Identification of novel 7-methyl and cyclopentanyl branched glycerol dialkyl glycerol tetraethers in lake
816 sediments. Organic Geochemistry 102, 52–58.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

817 Ding, S., Xu, Y., Wang, Y., He, Y., Hou, J., Chen, L., He, J.S., 2015. Distribution of branched glycerol dialkyl glycerol
818 tetraethers in surface soils of the Qinghai-Tibetan Plateau: Implications of brGDGTs-based proxies in cold and
819 dry regions. Biogeosciences 12, 3141–3151.
820 Fastovich, D., Russell, J.M., Jackson, S.T., Williams, J.W., 2020. Deglacial temperature controls on no-analog
821 community establishment in the Great Lakes Region. Quaternary Science Reviews 234, 106245.
822 Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas.
823 International Journal of Climatology 37, 4302–4315.
824 Foster, L.C., Pearson, E.J., Juggins, S., Hodgson, D.A., Saunders, K.M., Verleyen, E., Roberts, S.J., 2016.
825 Development of a regional glycerol dialkyl glycerol tetraether (GDGT)-temperature calibration for Antarctic
826 and sub-Antarctic lakes. Earth and Planetary Science Letters 433, 370–379.
827 Hanna, A.J.M., Shanahan, T.M., Allison, M.A., 2016. Distribution of branched GDGTs in surface sediments from the
828 Colville River, Alaska: Implications for the MBT′/CBT paleothermometer in Arctic marine sediments. Journal
829 of Geophysical Research: Biogeosciences 121, 1762–1780.
830 Harning, D.J., Andrews, J.T., Belt, S.T., Cabedo-Sanz, P., Geirsdóttir, Á., Dildar, N., Miller, G.H., Sepúlveda, J.,
831 2019. Sea Ice Control on Winter Subsurface Temperatures of the North Iceland Shelf During the Little Ice Age:
832 A TEX86 Calibration Case Study. Paleoceanography and Paleoclimatology 34, 1006–1021.
833 Harning, D.J., Curtin, L., Geirsdóttir, Á., D'Andrea, W.J., Miller, G.H., Sepúlveda, J., 2020. Lipid Biomarkers
834 Quantify Holocene Summer Temperature and Ice Cap Sensitivity in Icelandic Lakes. Geophysical Research
835 Letters 47, 1–11.
836 Hopmans, E.C., Schouten, S., Sinninghe, J.S., 2016. The effect of improved chromatography on GDGT-based
837 palaeoproxies. Organic Geochemistry 93, 1–6.
838 Hopmans, E.C., Weijers, J.W.., Schefuß, E., Herfort, L., Sinninghe Damsté, J.S., Schouten, S., 2004. A novel proxy
839 for terrestrial organic matter in sediments based on branched and isoprenoid tetraether lipids. Earth and Planetary
840 Science Letters 224, 107–116.
841 Hu, J., Zhou, H., Peng, P., Spiro, B., 2016. Seasonal variability in concentrations and fluxes of glycerol dialkyl glycerol
842 tetraethers in Huguangyan Maar Lake, SE China: Implications for the applicability of the MBT-CBT
843 paleotemperature proxy in lacustrine settings. Chemical Geology 420, 200–212.
844 Huguet, C., Hopmans, E.C., Febo-Ayala, W., Thompson, D.H., Sinninghe Damsté, J.S., Schouten, S., 2006. An
845 improved method to determine the absolute abundance of glycerol dibiphytanyl glycerol tetraether lipids.
846 Organic Geochemistry 37, 1036–1041.
847 Hutchinson, G.E., 1938. On the Relation between the Oxygen Deficit and the productivity and Typology of Lakes.
848 Internationale Revue der gesamten Hydrobiologie und Hydrographie 36, 336–355.
849 Landrum, L., Holland, M.M., 2020. Extremes become routine in an emerging new Arctic. Nature Climate Change 10.
850 doi:10.1038/s41558-020-0892-z
851 Li, F., Zhang, C.L., Wang, S., Chen, Y., Sun, C., Dong, H., Li, W., Klotz, M.G., Hedlund, B.P., 2014. Production of
852 branched tetraether lipids in Tibetan hot springs: A possible linkage to nitrite reduction by thermotolerant or
853 thermophilic bacteria? Chemical Geology 386, 209–217.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

854 Loomis, S.E., Russell, J.M., Eggermont, H., Verschuren, D., Sinninghe Damsté, J.S., 2014a. Effects of temperature,
855 pH and nutrient concentration on branched GDGT distributions in East African lakes: Implications for
856 paleoenvironmental reconstruction. Organic Geochemistry 66, 25–37.
857 Loomis, S.E., Russell, J.M., Heureux, A.M., D'Andrea, W.J., Sinninghe Damsté, J.S., 2014b. Seasonal variability of
858 branched glycerol dialkyl glycerol tetraethers (brGDGTs) in a temperate lake system. Geochimica et
859 Cosmochimica Acta 144, 173–187.
860 Loomis, S.E., Russell, J.M., Ladd, B., Street-Perrott, F.A., Sinninghe Damsté, J.S., 2012. Calibration and application
861 of the branched GDGT temperature proxy on East African lake sediments. Earth and Planetary Science Letters
862 357–358, 277–288.
863 Lu, H., Liu, W., Yang, H., Wang, H., Liu, Z., Leng, Q., Sun, Y., Zhou, W., An, Z., 2019. 800-kyr land temperature
864 variations modulated by vegetation changes on Chinese Loess Plateau. Nature Communications 10.
865 doi:10.1038/s41467-019-09978-1
866 Lumley, T., 2020. leaps: Regression Subset Selection.
867 Martínez-Sosa, P., Tierney, J.E., Stefanescu, I.C., Dearing Crampton-Flood, E., Shuman, B.N., Routson, C., 2020. A
868 global Bayesian temperature calibration for lacustrine brGDGTs. EarthArXiv [preprint].
869 doi:https://doi.org/10.31223/X5PS3P
870 Miller, G.H., Alley, R.B., Brigham-Grette, J., Fitzpatrick, J.J., Polyak, L., Serreze, M.C., White, J.W.C., 2010. Arctic
871 amplification: Can the past constrain the future? Quaternary Science Reviews 29, 1779–1790.
872 Naafs, B.D.A., Gallego-sala, A. V, Inglis, G.N., Pancost, R.D., 2017. Refining the global branched glycerol dialkyl
873 glycerol tetraether (brGDGT) soil temperature calibration. Organic Geochemistry 106, 48–56.
874 Ng, S.L., King, R.H., 1999. Development of a diatom-based specific conductivity model for the glacioisostatic lakes
875 of Truelove Lowland: Implications for paleoconductivity and paleoenvironmental reconstructions in Devon
876 Island lakes, N.W.T., Canada. Journal of Paleolimnology 22, 367–382.
877 Ning, D., Zhang, E., Shulmeister, J., Chang, J., Sun, W., Ni, Z., 2019. Holocene mean annual air temperature (MAAT)
878 reconstruction based on branched glycerol dialkyl glycerol tetraethers from Lake Ximenglongtan, southwestern
879 China. Organic Geochemistry 133, 65–76.
880 Nürnberg, G.K., 1995. Quantifying anoxia in lakes. Limnology and Oceanography 40, 1100–1111.
881 Pearson, E.J., Juggins, S., Talbot, H.M., Weckström, J., Rosén, P., Ryves, D.B., Roberts, S.J., Schmidt, R., 2011. A
882 lacustrine GDGT-temperature calibration from the Scandinavian Arctic to Antarctic: Renewed potential for the
883 application of GDGT-paleothermometry in lakes. Geochimica et Cosmochimica Acta 75, 6225–6238.
884 Pérez-Angel, L.C., Sepúlveda, J., Molnar, P., Montes, C., Rajagopalan, B., Snell, K., Gonzalez-Arango, C., Dildar,
885 N., 2020. Soil and Air Temperature Calibrations Using Branched GDGTs for the Tropical Andes of Colombia:
886 Toward a Pan-Tropical Calibration. Geochemistry, Geophysics, Geosystems 21, 1–18.
887 Peterse, F., Meer, J. Van Der, Schouten, S., Weijers, J.W.H., Fierer, N., Jackson, R.B., Kim, J., Damste, J.S.S., 2012.
888 Revised calibration of the MBT – CBT paleotemperature proxy based on branched tetraether membrane lipids
889 in surface soils 96, 215–229.
890 Peterse, F., Vonk, J.E., Holmes, R.M., Giosan, L., Zimov, N., Eglinton, T.I., 2014. Branched glycerol dialkyl glycerol

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

tetraethers in Arctic lake sediments: Sources and implications for paleothermometry at high latitudes. Journal of Geophysical Research: Biogeosciences 119, 1738–1754.

Prahl, F.G., Wakeham, S.G., 1987. Calibration of unsaturation patterns in long-chain ketone compositions for palaeotemperature assessment. Nature 330, 367–369.

Qian, S., Yang, H., Dong, C., Wang, Y., Wu, J., Pei, H., Dang, X., Lu, J., Zhao, S., Xie, S., 2019. Rapid response of fossil tetraether lipids in lake sediments to seasonal environmental variables in a shallow lake in central China: Implications for the use of tetraether-based proxies. Organic Geochemistry 128, 108–121.

R Core Team, 2018. R: A language and environment for statistical computing.

Raynolds, M.K., Comiso, J.C., Walker, D.A., Verbyla, D., 2008. Relationship between satellite-derived land surface temperatures, arctic vegetation types, and NDVI. Remote Sensing of Environment 112, 1884–1894.

Reizer, J., Grossowicz, N., Barenholz, Y., 1985. The effect of growth temperature on the thermotropic behavior of the membranes of a thermophilic Bacillus. Composition-structure-function relationships. BBA - Biomembranes 815, 268–280.

Russell, J.M., Hopmans, E.C., Loomis, S.E., Liang, J., Sinninghe Damsté, J.S., 2018. Distributions of 5- and 6-methyl branched glycerol dialkyl glycerol tetraethers (brGDGTs) in East African lake sediment: Effects of temperature, pH, and new lacustrine paleotemperature calibrations. Organic Geochemistry 117, 56–69.

Schouten, S., Hopmans, E.C., Schefuß, E., Sinninghe Damsté, J.S., 2002. Distributional variations in marine crenarchaeotal membrane lipids: A new tool for reconstructing ancient sea water temperatures? Earth and Planetary Science Letters 204, 265–274.

Schouten, S., Hopmans, E.C., Sinninghe Damsté, J.S., 2013. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. Organic Geochemistry 54, 19–61.

Schwarz, G., 1978. Estimating the Dimension of a Model. The Annals of Statistics 6, 461–464.

Shanahan, T.M., Hughen, K.A., Van Mooy, B.A.S., 2013. Temperature sensitivity of branched and isoprenoid GDGTs in Arctic lakes. Organic Geochemistry 64, 119–128.

Sinninghe Damsté, J.S., 2016. Spatial heterogeneity of sources of branched tetraethers in shelf systems: The geochemistry of tetraethers in the Berau River delta (Kalimantan, Indonesia). Geochimica et Cosmochimica Acta 186, 13–31.

Sinninghe Damsté, J.S., Hopmans, E.C., Pancost, R.D., Schouten, S., Geenevasen, J.A.J., 2000. Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments. Chemical Communications 0, 1683–1684.

Sinninghe Damsté, J.S., Rijpstra, W.I.C., Foesel, B.U., Huber, K.J., Overmann, J., Nakagawa, S., Kim, J.J., Dunfield, P.F., Dedysh, S.N., Villanueva, L., 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.

Stager, J.C., Ryves, D.B., King, C., Madson, J., Hazzard, M., Neumann, F.H., Maud, R., 2013. Late Holocene precipitation variability in the summer rainfall region of South Africa. Quaternary Science Reviews 67, 105–120.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

928 Sun, Q., Chu, G., Liu, M., Xie, M., Li, S., Ling, Y., Wang, X., Shi, L., Jia, G., Lü, H., 2011. Distributions and
929 temperature dependence of branched glycerol dialkyl glycerol tetraethers in recent lacustrine sediments from
930 China and Nepal. Journal of Geophysical Research: Biogeosciences 116, G01008.
931 Super, J.R., Chin, K., Pagani, M., Li, H., Tabor, C., Harwood, D.M., Hull, P.M., 2018. Late Cretaceous climate in the
932 Canadian Arctic: Multi-proxy constraints from Devon Island. Palaeogeography, Palaeoclimatology,
933 Palaeoecology 504, 1–22.
934 Thomas, E.K., Castañeda, I.S., McKay, N.P., Briner, J.P., Salacup, J.M., Nguyen, K.Q., Schweinsberg, A.D., 2018.
935 A Wetter Arctic Coincident With Hemispheric Warming 8,000 Years Ago. Geophysical Research Letters 45,
936 10,637-10,647.
937 Tierney, J.E.E., Russell, J.M.M., Eggermont, H., Hopmans, E.C.C., Verschuren, D., Sinninghe Damsté, J.S.S., 2010.
938 Environmental controls on branched tetraether lipid distributions in tropical East African lake sediments.
939 Geochimica et Cosmochimica Acta 74, 4902–4918.
940 Tyler, J.J., Nederbragt, A.J., Jones, V.J., Thurow, J.W., 2010. Assessing past temperature and soil pH estimates from
941 bacterial tetraether membrane lipids: Evidence from the recent lake sediments of Lochnagar, Scotland. Journal
942 of Geophysical Research 115, G01015.
943 van Bree, L., Peterse, F., Baxter, A., De Crop, W., van Grinsven, S., Villanueva, L., Verschuren, D., Sinninghe
944 Damsté, J., 2020. Seasonal variability and sources of in situ brGDGT production in a permanently stratified
945 African crater lake. Biogeosciences Discussions 1–36.
946 Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, Fourth. ed. Springer, New York.
947 Weber, Y., Damsté, J.S.S., Zopfi, J., De Jonge, C., Gilli, A., Schubert, C.J., Lepori, F., Lehmann, M.F., Niemann, H.,
948 2018. Redox-dependent niche differentiation provides evidence for multiple bacterial sources of glycerol
949 tetraether lipids in lakes. Proceedings of the National Academy of Sciences of the United States of America 115,
950 10926–10931.
951 Weber, Y., De Jonge, C., Rijpstra, W.I.C., Hopmans, E.C., Stadnitskaia, A., Schubert, C.J., Lehmann, M.F., Sinninghe
952 Damsté, J.S., Niemann, H., 2015. Identification and carbon isotope composition of a novel branched GDGT
953 isomer in lake sediments: Evidence for lacustrine branched GDGT production. Geochimica et Cosmochimica
954 Acta 154, 118–129.
955 Weijers, J.W.H., Schouten, S., Spaargaren, O.C., Sinninghe Damsté, J.S., 2006. Occurrence and distribution of
956 tetraether membrane lipids in soils: Implications for the use of the TEX86 proxy and the BIT index. Organic
957 Geochemistry 37, 1680–1693.
958 Weijers, J.W.H., Schouten, S., van den Donker, J.C., Hopmans, E.C., Sinninghe Damsté, J.S., 2007. Environmental
959 controls on bacterial tetraether membrane lipid distribution in soils. Geochimica et Cosmochimica Acta 71, 703–
960 713.
961 Woltering, M., Werne, J.P., Kish, J.L., Hicks, R., Sinninghe Damsté, J.S., Schouten, S., 2012. Vertical and temporal
962 variability in concentration and distribution of thaumarchaeotal tetraether lipids in Lake Superior and the
963 implications for the application of the TEX 86 temperature proxy. Geochimica et Cosmochimica Acta 87, 136–
964 153.

https://doi.org/10.5194/bg-2021-16
Preprint. Discussion started: 28 January 2021

965 Wörmer, L., Lipp, J.S., Schröder, J.M., Hinrichs, K.U., 2013. Application of two new LC-ESI-MS methods for
966 improved detection of intact polar lipids (IPLs) in environmental samples. Organic Geochemistry 59, 10–21.
967 Xiao, W., Wang, Y., Liu, Y., Zhang, X., Shi, L., Xu, Y., 2020. Predominance of hexamethylated 6-methyl branched
968 glycerol dialkyl glycerol tetraethers in the Mariana Trench: Source and environmental implication.
969 Biogeosciences 17, 2135–2148.
970 Yang, X., Wang, S., Kamenik, C., Schmidt, R., Shen, J., Zhu, L., Li, S., 2004. Diatom assemblages and quantitative
971 reconstruction for paleosalinity from a sediment core of Chencuo Lake, southern Tibet. Science in China, Series
972 D: Earth Sciences 47, 522–528.
973 Yao, Y., Zhao, J., Vachula, R.S., Werne, J.P., Wu, J., Song, X., Huang, Y., 2020. Correlation between the ratio of 5-
974 methyl hexamethylated to pentamethylated branched GDGTs (HP5) and water depth reflects redox variations
975 in stratified lakes. Organic Geochemistry 147, 104076.
976 Yuk, H.G., Marshall, D.L., 2004. Adaptation of Escherichia coli O157:H7 to pH alters membrane lipid composition,
977 verotoxin secretion, and resistance to simulated gastric fluid acid. Applied and Environmental Microbiology 70,
978 3500–3505.
979 Zhang, Z.X., Li, J., Chen, Z., Sun, Z., Yang, H., Fu, M., Peng, X., 2020. The effect of methane seeps on the bacterial
980 tetraether lipid distributions at the Okinawa Trough. Marine Chemistry 225, 103845.
981 Zhao, J., Huang, Y., Yao, Y., An, Z., Zhu, Y., Lu, H., Wang, Z., 2020. Calibrating branched GDGTs in bones to
982 temperature and precipitation: Application to Alaska chronological sequences. Quaternary Science Reviews
983 240, 106371.
984 Zheng, Y., Pancost, R.D., Liu, X., Wang, Z., Naafs, B.D.A., Xie, X., Liu, Z., Yu, X., Yang, H., 2017. Atmospheric
985 connections with the North Atlantic enhanced the deglacial warming in northeast China. Geology 45, 1031–
986 1034.
987