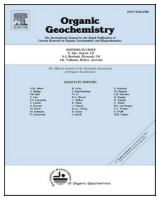
Organic Geochemistry 184 (2023) 104660

Contents lists available at ScienceDirect

Organic Geochemistry

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



Regional vs. global temperature calibrations for lacustrine BrGDGTs in the North American (sub)tropics: Implications for their application in paleotemperature reconstructions

Yanyan Lei [a,*], Damara J. Strong [a], Margarita Caballero [b], Alex Correa-Metrio [c], Liseth Pérez [d], Antje Schwalb [d], Laura Macario-González [e], Sergio Cohuo [f], Socorro Lozano-García [b], Beatriz Ortega-Guerrero [b], Josef P. Werne [a,*]

[a] Department of Geology and Environmental Science, University of Pittsburgh, Pittsburgh, PA 15260, United States
[b] Instituto de Geofisica, UNAM, Ciudad Universitaria, México DF 04510, Mexico
[c] Instituto de Geología, UNAM, Ciudad Universitaria, México DF 04510, Mexico
[d] Institut für Geosysteme und Bioindikation, Technische Universität Braunschweig, Braunschweig, Germany
[e] Instituto tecnológico de La Zona Maya, Quintana Roo, Mexico
[f] Instituto tecnológico de Chetumal, Quintana Roo, Mexico

ARTICLE INFO

Associate Editor: Prof Melissa Berke

Keywords:
brGDGTs
Biomarker
Temperature
Lacustrine
Southern North America continent
Tropics

ABSTRACT

Branched glycerol dialkyl glycerol tetraethers (brGDGTs) have shown great promise in lacustrine temperature reconstructions across different continents. While brGDGTs have been reported from many different regions and global brGDGT-temperature calibrations have been developed with various methods, southern North America remains an understudied area with little available data. In this study, we analyzed 101 lake surface sediment samples across Mexico and Central America and compared their distributions with those in other lacustrine systems. Nine major brGDGTs were found in all samples. We investigated the relationships between the distribution of the fractional abundances of the nine major brGDGTs and temperature and developed regional calibrations for Mean Annual Temperature using three different approaches, including a novel machine learning method – Ridge Regression. All the regional calibrations provide similar results with very close error ranges (RMSE = 3.1 °C). The majority of global brGDGT-temperature calibrations tend to reconstruct lower temperatures when it is below 15 °C. Interestingly, regional brGDGT calibrations appear to reduce the "cold bias", but the various global and regional calibrations tested here are not significantly different in their predictive capability.

1. Introduction

The tropics play a significant role in the global climate system and understanding mechanistic linkages between the low and mid-high latitudes is crucial for evaluating the cause(s) of abrupt climate change. In the tropics, temperature is among the most important climate parameters to be reconstructed quantitatively and robustly. Firstly, temperature is closely linked to atmosphere–ocean circulation and thus has a major influence on continental precipitation. Moreover, temperature variation in low latitudes is generally smaller than in high-latitude regions. Reconstructing paleoclimate change is an extensively applied method to study climate mechanisms. However, this work is particularly challenging in the Southern North American (sub)tropics due in part to a lack of rigorous paleotemperature proxies and long natural archives. Therefore, developing a robust calibration for emerging paleotemperature proxies will significantly contribute to our understanding of the amplitudes of paleoclimate change in this area.

In recent decades, a suite of membrane lipids produced by yet unknown heterotrophic bacteria, branched glycerol dialkyl glycerol tetraethers (brGDGTs), has received increasing interest as a biomarker for temperature reconstruction (Martínez-Sosa et al., 2021; De Jonge et al., 2014a; Weijers et al., 2007). BrGDGTs generally contain straight alkyl chains with various numbers of methyl groups (ranging from 4 to 6) and cyclopentane moieties (ranging from 0 to 2), with methyl groups positioned on the 5th (5-methyl) or 6th (6-methyl) carbon of the alkyl chains (Fig. S1, Weijers et al., 2006; De Jonge et al., 2014). The microbial

* Corresponding authors.
E-mail addresses: yal113@pitt.edu (Y. Lei), jwerne@pitt.edu (J.P. Werne).

https://doi.org/10.1016/j.orggeochem.2023.104660
Received 29 July 2022; Received in revised form 15 August 2023; Accepted 16 August 2023
Available online 24 August 2023

producers of brGDGTs are not fully identified, though there is evidence for Acidobacteria as source organisms (Sinninghe Damsté et al., 2011; Chen et al., 2022; Halamka et al., 2023). These microbes adjust the composition and structure of the lipids through internal cyclization and varying degrees of methylation to modify membrane permeability and rigidity, therefore helping them adapt to ambient environmental conditions (Naafs et al., 2021; Schouten et al., 2013). Indeed, a relationship between temperature and the degree of methylation has been confirmed in recent culture studies (Chen et al., 2022; Halamka et al., 2023).

BrGDGTs were first found in peatland and soils, and in the subsequent decade, they have been discovered in a variety of aquatic environments, including marine, lacustrine, and riverine (Naafs et al., 2017; Sinninghe Damsté, 2016; Weber et al., 2015; De Jonge et al., 2014a; De Jonge et al., 2014b; Schouten et al., 2013; Blaga et al., 2010; Liu et al., 2010; Peterse et al., 2009; Tierney and Russell, 2009; Weijers et al., 2007; Weijers et al., 2006; Sinninghe Damsté et al., 2000). A series of empirical studies focused on soil environments revealed that the distribution of brGDGTs has a systematic relationship to one or more environmental factors, such as mean annual air temperature (MAAT) and/or pH. Specifically, two indices were proposed to describe the distribution of brGDGTs: 1) the degree of methylation of brGDGTs (denoted by the MBT index and various redefined versions) and 2) the degree of cyclization of brGDGTs (denoted by the CBT index and various redefined versions). They were shown to change in response to ambient temperature and pH in soils (Dearing Crampton-Flood et al., 2020; Naafs et al., 2017; Dang et al., 2016; Xiao et al., 2015; Weijers et al., 2007). Similar correlations between brGDGTs and these environmental factors were also observed when brGDGTs were discovered in the lacustrine environment (Blaga et al., 2010; Tierney et al., 2010).

As more and more empirical studies have been done, three major turning points have appeared. First, the distribution of brGDGTs is seemingly different between lacustrine and soil/peat environments, implying different source organisms (Martínez-Sosa et al., 2021; Russell et al., 2018; Naafs et al., 2017; Tierney et al., 2010). It is shown that in lacustrine environments, temperature is the dominant control on the distributions of brGDGTs and pH has a weak correlation with brGDGTs (Martínez-Sosa et al., 2021; Raberg et al., 2021; Russell et al., 2018; Loomis et al., 2014a).

Second, with the separation of 5- and 6-methyl brGDGT isomers (De Jonge et al., 2014a), it is demonstrated that the 5-methyl brGDGTs are the main groups responding to temperature change, and 6-methyl brGDGTs are likely to introduce scatter in the calibrations for temperature reconstruction (Martínez-Sosa et al., 2021; Raberg et al., 2021; Zhao et al., 2021; Russell et al., 2018; Yang et al., 2015; De Jonge et al., 2014a). Therefore, by excluding 6-methyl brGDGTs, several studies focused on the lacustrine environment have successfully applied 5-methyl brGDGT-based proxies using lake sediment samples to reconstruct temperature with less scatter, achieving RMSE < 2.5 °C (Zhao et al., 2021; Russell et al., 2018).

Finally, statistical analyses have had an increasing influence on calibration developments, introducing more rigorous models for most lacustrine temperature reconstructions at various scales (Martínez-Sosa et al., 2021; Raberg et al., 2021; Loomis et al., 2012). For example, rather than using brGDGT indices to develop temperature calibration, another approach that shows promising results in temperature reconstructions with lower statistical errors (RMSE as low as 2.14 °C) applies statistical associations of the fractional abundance of brGDGTs (Russell et al., 2018; Sun et al., 2016; Loomis et al., 2012). In the past decade, brGDGTs have been applied to develop multiple regional lacustrine calibrations using distinct statistical models and applied in paleoclimate downcore reconstructions (Zhao et al., 2021; Ning et al., 2019; Russell et al., 2018; Wang et al., 2016; Peterse et al., 2014; Fawcett et al., 2011; Sun et al., 2011; Tierney et al., 2010).

These regional calibrations are limited in universal applications (e.g., they rarely predict temperature well in other regions), though they appear robust in their regional application. More research and data have been published over the past decade, eventually allowing the development of global lacustrine calibrations (Martínez-Sosa et al., 2021; Raberg et al., 2021). However, these global calibrations also often fail to predict reasonable temperatures downcore (Miller et al., 2018; Lindberg et al., 2022; Parish et al., 2023).

Water temperature in lakes is driven by a complex heat budget. It is often difficult to isolate a single flux because of the availability of most of the hydro-meteorological variables. Therefore, air temperature is utilized for calibrations, and it is assumed these two temperatures (air temperature and lake surface water temperature) are approximately linearly related or can be converted (Piccolroaz et al., 2013; Crisp and Howson, 1982). In some calibration development studies, warm-season temperature or summer temperature was shown to outperform MAAT, especially in the global calibrations, with a mean temperature of months above freezing (MAF) yielding the highest performance by reducing the statistical errors (Zhao et al., 2021; Cao et al., 2020). There is not yet a well-acknowledged fundamental explanation for why this is so, though it is currently thought that brGDGT production is likely to be suppressed during freezing seasons, thereby producing a seasonal bias (Cao et al., 2020; Pearson et al., 2011).

The Neotropics in southern North America are undergoing significant climate change, and so far, there is no long climate record in this region. Fortunately, in recent decades, two long lacustrine sequences were successfully recovered from Lake Petén Itzá, Guatemala and Lake Chalco, Mexico (Valero-Garcés et al., 2021; Brown et al., 2019; Kutterolf et al., 2016; Hodell et al., 2006), which will begin to fill in the blanks. Quantitative temperature reconstructions from these systems will provide important data for regional paleoclimate studies. Nevertheless, it should be noted that the existing brGDGT global calibrations for lacustrine environment have only included 3 samples from Mexico and Central America. In this study, we aim to 1) provide 101 lake sediment samples from Mexico and Central America to fill the void of regional brGDGT samples in the calibrations; 2) discover the most robust calibrations for future regional paleotemperature reconstructions utilizing lacustrine sediment in this region; and 3) introduce a machine learning method utilized in model development.

2. Samples and methods

2.1. Study area and sampling

Surface sediment samples (n = 101) were collected from 83 lakes and 4 cenotes (Perry et al., 1995) across Mexico and Central America, covering an area from 88°53′21″W to 104°44′12″W and 13°16′59″N to 21°22′10″N (Fig. 1). The lacustrine varieties include crater lakes, sinkholes, and reservoirs spanning a range of elevations from 0 to 2852 m a.s.l.. Samples were collected over multiple field trips. Most samples were collected using an Ekman dredge during the wet season (from June to mid-Oct). After retrieving from the lakes, samples were stored in glass jars and transported back to the lab, where they were stored frozen until analysis.

We divide the study area into two subregions: 1) the Mexican Plateau (denoted as MP) and 2) the Yucatán Peninsula and Central America (denoted as YCA; Fig. 1, Fig. S2, and S3). In the MP, the climate is characterized as warm and arid (Fig. S2), while in the YCA, it is warm and subhumid (Fig. S3). In both regions, the sample sites have near-freezing conditions for at most a few days a year, so all 12 months can be considered “Months Above Freezing”. Thus, in our region MAF and MAAT can be used interchangeably.

On the MP, temperature exhibits a mild seasonal pattern with a MAAT of 18.5 °C. The plateau reaches its highest temperature in May. It is generally dry year-round and receives most of its precipitation in summer, starting from June and lasting until mid-October (Fig. S2). Six of the 26 samples from the MP were collected from sites with elevations $<$1000 m and MAAT (MAF) over 25 °C; for most sample sites, altitudes ranged from 1000 to 2644 m a.s.l. with relatively lower MAAT (MAF)

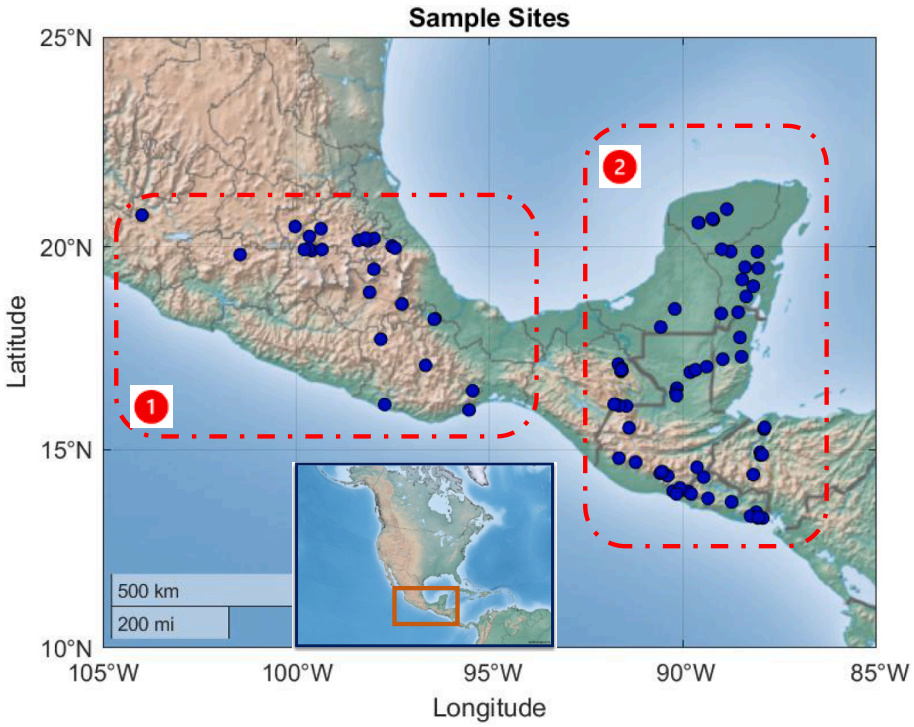


Fig. 1. Map of sampling sites. Blue dots are lakes where sediment samples were collected. Red dashed rectangles numbered 1 and 2 indicate the two subregions termed the Mexican Plateau (MP) and the Yucatán Peninsula and Central America (YCA), respectively. The inset shows the region of the study area in orange rectangle box in relation to the North American Continent. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(Fig. S2, Table 1).

MAAT (MAF) in the YCA area is warmer than the MP by ~ 5 °C, ranging from 17.1 °C to 27.7 °C with an average of 23.8 °C. Like the MP, May is the warmest month. It has a seasonal precipitation pattern, with the wet season lasting from May to October (Fig. S3, Table 1). Compared to the MP, the YCA has ~500 mm more annual average precipitation. Seventy-five samples were collected in the YCA. The topography in this study area changes from very flat in the Yucatán Peninsula in the northeast to mountainous in the Cordillera mountains in the southwest. Most of the samples were collected from low elevation areas, though a few were collected from lakes at 1000 to 2644 m a.s.l. in the Cordillera mountains with relatively lower MAAT (MAF) (Fig. S3).

2.2. Sample preparation and analysis

Samples were freeze-dried and homogenized using a mortar and pestle and rinsed with deionized water. Total Lipid Extracts (TLE) of sediment samples were obtained using an Accelerated Solvent Extractor (Dionex ASE 350) with 9:1 (v:v) methanol (MeOH):dichloromethane (DCM) at 100 °C and 7.6 × 10^6 Pa pressure. Activated copper beads were used to remove inorganic sulfur from each sample. A fraction of each TLE was transferred to a silica gel column and eluted using 9:1 (v:v) hexane(HEX):DCM, ethyl acetate, and MeOH to collect the apolar fraction, target polar fraction (which contains brGDGTs), and remaining polar fraction, respectively. The target polar fractions were re-dissolved using 99:1 (v:v) HEX:isopropanol (IPA) and filtered through 0.45 μm PTFE filters prior to instrumental analysis using atmospheric pressure chemical ionization/high-performance liquid chromatography-mass spectrometry (APCI/HPLC–MS). All polar fractions were analyzed at the University of Pittsburgh following the chromatographic and equipment configuration described by Hopmans et al. (2016). Analyses were performed using full scan mode. Peak areas were selected to target brGDGTs with *m/z* 1050, 1048, 1046, 1036, 1034, 1032, 1022, 1020, 1018, and integrated manually.

2.3. Environmental data collection

Various environmental parameters have been proven to exert controls on both the abundance and distribution of brGDGTs. Yet in most lacustrine-focused studies, temperature remains the primary control (Martínez-Sosa et al., 2021; Zhao et al., 2021; Russell et al., 2018). We acknowledge that water conductivity, redox, salinity and pH are important environmental factors in lacustrine brGDGT studies and know these data will enhance our understanding of the potential bias in using brGDGTs to reconstruct temperatures in lakes (Martínez-Sosa et al., 2021; Wang et al., 2021; Weber et al., 2018; Foster et al., 2016). Unfortunately, there were not enough water chemistry data for all the lakes in the study area, so temperatures will be the primary environmental data on which we focus. For samples collected in Mexico, temperature data were collected from the climatological statistical information published on the Government of Mexico website (https://smn.conagua.gob.mx/es/climatologia/informacion-climatologica/informacion-estadistica-climatologica). We used monthly average temperature data from the nearest station (all were within 5 km) to calculate mean annual air temperature (MAAT, which is considered equivalent to MAF for our samples). Depending on the stations' operation status, monthly temperatures varied over multiple years based on the availability of the data (10–40 years). Due to a lack of weather station data or *in situ* measured temperature, data for Guatemala were collected from the International Research Institute for Climate and Society data library (https://iridl.ldeo.columbia.edu/). Monthly average temperatures from 1981 to 2010 were used to calculate MAAT (again, equivalent to MAF). In El Salvador, Belize, and Honduras, 10 m air temperatures were obtained from the Global Land Data Assimilation System (GLDAS) due to a lack of weather station data and were averaged over thirty years from 1980 to 2010.

Table 1
Environmental information and fractional abundance of brGDGTs. MAAT, MAP and PET are Mean Annual Air Temperature, Mean Annual Precipitation and Potential Evapotranspiration, respectively.

	Longitude	Latitude	Alt(m)	MAAT (°C)	MAP (mm)	PET (mm)	Ia	Ib	Ic	IIa	IIa'	IIb	IIb'	IIIa	IIIa'
Apastepeque	−88.74	13.69	511	23.8	1841	NA	0.46	0.15	0.02	0.12	0.15	0.02	0.04	0.01	0.02
Metapan	−89.47	14.31	450	23.8	1841	NA	0.45	0.13	0.01	0.09	0.19	0.02	0.08	0.01	0.03
Ocotalito	−91.60	16.94	930	22.0	1940	1343	0.40	0.17	0.03	0.13	0.15	0.04	0.04	0.02	0.02
Ocotalito	−91.60	16.94	930	22.0	1940	1343	0.51	0.19	0.03	0.09	0.10	0.03	0.04	0.01	0.01
Caldera	−90.59	14.41	1790	25.8	1880	NA	0.34	0.10	0.04	0.20	0.15	0.05	0.03	0.05	0.04
Cenote colac	−88.87	20.91	11	26.8	973	1710	0.31	0.21	0.06	0.09	0.14	0.03	0.12	0.01	0.03
Negritos super	−87.94	13.28	102	23.8	1841	NA	0.68	0.11	0.01	0.09	0.07	0.01	0.01	0.00	0.01
La Nopala	−99.66	20.25	2360	14.3	663	1615	0.18	0.02	0.00	0.19	0.33	0.02	0.02	0.06	0.18
Yalahau	−89.22	20.66	2	26.5	887	1673	0.37	0.13	0.01	0.09	0.26	0.02	0.08	0.01	0.04
Belize	−88.49	17.30	1	25.1	2041	NA	0.57	0.11	0.01	0.10	0.14	0.02	0.03	0.01	0.02
El Pino	−90.39	14.34	1038	23.1	1604	NA	0.52	0.14	0.02	0.14	0.10	0.03	0.03	0.00	0.01
Espino	−89.87	13.95	689	23.8	1841	NA	0.53	0.14	0.02	0.13	0.10	0.03	0.03	0.02	0.01
Jucutuma	−87.90	15.51	27	23.5	1949	NA	0.53	0.12	0.02	0.12	0.14	0.01	0.04	0.00	0.01
Ocom	−88.05	19.47	1	26.5	681	1480	0.55	0.17	0.05	0.07	0.10	0.02	0.01	0.01	0.02
Laguna San José	−89.01	18.37	118	24.0	1456	1451	0.48	0.16	0.03	0.13	0.12	0.04	0.04	0.01	0.00
La Avispa	−95.53	15.99	15	27.8	1095	1985	0.41	0.15	0.01	0.10	0.21	0.03	0.05	0.01	0.03
Cuetzalan	−97.53	20.01	1045	20.2	4171	1001	0.48	0.06	0.01	0.20	0.13	0.02	0.01	0.04	0.03
Laguna Grande	−97.27	18.61	2514	13.4	762	1286	0.17	0.04	0.01	0.18	0.21	0.04	0.05	0.12	0.17
Laguna Kaná	−88.40	19.50	5	23.1	1407	1524	0.36	0.21	0.04	0.11	0.15	0.03	0.06	0.01	0.02
Lequi	−91.46	16.08	1500	17.5	2487	1052	0.30	0.07	0.01	0.09	0.19	0.02	0.04	0.09	0.19
Yojoa	−88.00	14.93	637	23.5	1949	NA	0.70	0.06	0.02	0.10	0.08	0.01	0.01	0.01	0.01
Aramuaca	−88.10	13.43	96	23.8	1841	NA	0.72	0.08	0.01	0.08	0.08	0.01	0.01	0.01	0.01
Magdalena	−91.40	15.54	2864	18.4	1721	NA	0.25	0.07	0.02	0.25	0.10	0.05	0.04	0.14	0.08
Rosari	−90.16	16.53	126	26.4	2092	NA	0.36	0.28	0.05	0.09	0.11	0.04	0.05	0.01	0.01
Laguna Yaxha	−89.39	17.06	164	25.5	1667	NA	0.40	0.14	0.02	0.13	0.16	0.05	0.05	0.02	0.03
Lago Yaxha	−91.58	16.97	963	22.0	1940	1343	0.39	0.21	0.04	0.10	0.12	0.06	0.05	0.01	0.02
Laguna Lourdes	−100.03	20.49	1891	17.7	401	2193	0.42	0.02	0.00	0.33	0.15	0.01	0.00	0.04	0.03
San Jose Aguilar	−89.01	18.37	120	24.0	1456	1451	0.41	0.17	0.02	0.16	0.14	0.03	0.03	0.01	0.02
Salpeten	−89.68	16.98	105	25.4	2012	NA	0.31	0.27	0.08	0.07	0.13	0.05	0.05	0.01	0.03
Paiasquito	−91.75	16.13	1462	17.5	2487	1052	0.35	0.12	0.02	0.12	0.19	0.05	0.07	0.03	0.06
Cenote Sabak Ha	−89.59	20.58	18	25.4	1100	1817	0.27	0.22	0.18	0.06	0.06	0.06	0.12	0.02	0.01
Santa Elena	−99.60	19.91	2585	14.6	917	1623	0.23	0.06	0.01	0.25	0.19	0.07	0.02	0.10	0.06
Yosocuta	−97.82	17.74	1511	20.7	741	2280	0.30	0.08	0.01	0.15	0.24	0.04	0.07	0.04	0.07
Supitlan	−98.39	20.15	2135	15.2	546	1481	0.27	0.07	0.01	0.24	0.20	0.05	0.03	0.06	0.07
Laguna verde	−89.79	13.89	1609	23.8	1841	NA	0.66	0.02	0.00	0.22	0.06	0.01	0.00	0.04	0.00
La Huaracha	−99.69	19.96	2644	15.6	761	1712	0.27	0.04	0.01	0.29	0.17	0.05	0.01	0.09	0.06
Olomega	−88.06	13.29	96	23.8	1841	NA	0.60	0.12	0.01	0.12	0.10	0.01	0.02	0.01	0.01
Las Pozas	−90.17	16.35	146	26.3	2231	NA	0.34	0.22	0.02	0.14	0.14	0.07	0.04	0.00	0.02
Comandador	−90.25	13.96	20	27.7	1552	NA	0.51	0.17	0.02	0.10	0.10	0.02	0.05	0.01	0.02
Ipala	−89.64	14.56	1495	23.2	1300	NA	0.41	0.08	0.01	0.21	0.14	0.04	0.04	0.04	0.04
Magdalena	−91.39	15.54	2852	18.4	1721	NA	0.30	0.04	0.01	0.26	0.11	0.03	0.02	0.13	0.10
Presa Benito Juarez	−95.44	16.46	139	27.5	671	2600	0.33	0.15	0.01	0.08	0.22	0.02	0.12	0.01	0.04
Petexbatún super	−90.19	16.42	120	26.3	2231	NA	0.47	0.19	0.03	0.11	0.11	0.04	0.04	0.00	0.00
Las pozas	−90.17	16.34	152	26.3	2231	NA	0.50	0.18	0.02	0.08	0.11	0.04	0.04	0.00	0.01
Amarillo	−91.60	16.98	850	22.0	1940	1343	0.41	0.17	0.02	0.12	0.13	0.05	0.05	0.02	0.02
Esmeralda	−91.73	16.12	1473	17.5	2487	1052	0.28	0.14	0.03	0.18	0.16	0.07	0.07	0.04	0.04
Tejocotal	−98.14	20.14	2143	14.6	1677	1019	0.34	0.06	0.01	0.27	0.13	0.05	0.02	0.08	0.04
Atitlán	−91.22	14.68	1556	22.7	2181	NA	0.25	0.11	0.02	0.15	0.22	0.06	0.06	0.05	0.08
Emiliano Zapata	−88.47	19.20	23	26.6	1102	1296	0.55	0.14	0.01	0.13	0.09	0.02	0.03	0.00	0.01
Miguel Hidalgo	−88.37	18.79	31	23.5	1078	1676	0.44	0.21	0.03	0.10	0.13	0.03	0.03	0.01	0.01
Teremendo	−101.45	19.81	2058	17.2	718	1467	0.21	0.05	0.00	0.05	0.28	0.05	0.28	0.01	0.07
Yojoa Punto	−87.98	14.86	639	23.5	1949	NA	0.44	0.19	0.04	0.10	0.12	0.04	0.06	0.00	0.01
Yala	−91.65	16.09	1460	17.5	2487	1052	0.30	0.12	0.02	0.16	0.16	0.07	0.08	0.04	0.05
Chicabal	−91.66	14.79	2726	17.5	1732	NA	0.31	0.14	0.04	0.29	0.07	0.02	0.01	0.10	0.02
Laguna la Perdida	−90.58	18.03	49	26.5	1488	1616	0.55	0.17	0.03	0.09	0.09	0.03	0.03	0.01	0.01
Sabanita	−88.57	18.40	38	25.7	1068	1606	0.67	0.08	0.00	0.12	0.09	0.02	0.01	0.01	0.00
Chan Laguna	−90.21	18.48	67	26.4	1254	1580	0.47	0.19	0.02	0.11	0.12	0.03	0.04	0.00	0.01
Chanmico	−89.35	13.78	477	23.8	1841	NA	0.47	0.16	0.01	0.08	0.17	0.04	0.06	0.00	0.01
Crooked Tree Lagoon	−88.53	17.78	2	25.1	2041	NA	0.54	0.16	0.02	0.10	0.11	0.02	0.03	0.00	0.01
Miguel Aleman	−96.42	18.24	78	25.7	2557	1388	0.39	0.12	0.01	0.10	0.19	0.03	0.08	0.02	0.04
Ticamaya	−87.89	15.55	17	23.5	1949	NA	0.44	0.17	0.03	0.12	0.15	0.03	0.05	0.01	0.01
El Estudiante	−96.66	17.09	1640	21.3	765	1945	0.39	0.20	0.02	0.11	0.17	0.02	0.06	0.01	0.02
Jocotal	−88.25	13.34	26	23.8	1841	NA	0.55	0.14	0.01	0.11	0.14	0.02	0.03	0.01	0.00
Tenango	−97.99	20.20	1306	17.1	2033	1050	0.44	0.06	0.00	0.31	0.06	0.02	0.02	0.05	0.02
Xi-bana	−91.67	17.12	630	22.0	1940	1343	0.34	0.16	0.03	0.13	0.17	0.04	0.07	0.03	0.04
Yegüey	−97.72	16.12	24	27.0	1104	1965	0.43	0.15	0.02	0.15	0.14	0.03	0.04	0.02	0.02
Balam	−91.78	16.13	1457	17.1	1227	1297	0.55	0.08	0.01	0.12	0.14	0.02	0.04	0.02	0.04
Belize	−88.97	17.24	78	25.1	2041	NA	0.69	0.07	0.00	0.06	0.14	0.01	0.02	0.00	0.00
Finca de Escamilla	−90.53	14.45	1200	25.8	1880	NA	0.37	0.06	0.01	0.08	0.35	0.03	0.03	0.01	0.06
Naha	−91.60	16.98	830	22.0	1940	1343	0.37	0.13	0.02	0.14	0.16	0.07	0.05	0.03	0.03
San Juanico	−99.79	19.93	2630	13.4	739	1644	0.21	0.05	0.01	0.22	0.25	0.05	0.03	0.09	0.10
Cenote Yumku	−89.61	20.58	16	25.4	1100	1817	0.27	0.09	0.01	0.05	0.05	0.19	0.22	0.04	0.08
Apastepeque	−88.74	13.69	509	23.8	1841	NA	0.50	0.17	0.02	0.11	0.12	0.03	0.03	0.01	0.00

(*continued on next page*)

Table 1 (*continued*)

	Longitude	Latitude	Alt(m)	MAAT (°C)	MAP (mm)	PET (mm)	Ia	Ib	Ic	IIa	IIa'	IIb	IIb'	IIIa	IIIa'
Miguel Aleman	−96.44	18.23	78	25.7	2557	1388	0.36	0.12	0.02	0.14	0.17	0.04	0.08	0.03	0.04
Cenote Oxolá	−89.24	20.68	18	26.5	887	1673	0.23	0.30	0.12	0.04	0.08	0.05	0.14	0.01	0.01
Quexil	−89.81	16.92	120	25.4	2012	NA	0.62	0.14	0.01	0.08	0.10	0.01	0.01	0.01	0.01
Laguna Señor	−88.08	19.88	3	26.3	1185	1546	0.58	0.11	0.02	0.15	0.06	0.04	0.01	0.01	0.01
San Juan Bautista	−90.08	14.04	1285	22.8	1664	NA	0.73	0.01	0.00	0.19	0.05	0.00	0.00	0.01	0.01
Santa Ana	−98.21	20.21	2200	15.0	641	1487	0.43	0.03	0.00	0.27	0.12	0.03	0.01	0.07	0.03
Yojoa	−87.95	14.88	640	23.5	1949	NA	0.68	0.06	0.01	0.12	0.09	0.01	0.01	0.01	0.01
Alfajayucan	−99.36	20.44	1877	17.0	451	2021	0.22	0.03	0.00	0.10	0.39	0.03	0.04	0.04	0.15
Chacchoben	−88.18	19.04	6	25.8	1505	1561	0.65	0.13	0.01	0.09	0.08	0.01	0.02	0.01	0.01
Centro Colorada	−103.99	20.77	1366	19.2	997	1669	0.35	0.06	0.00	0.25	0.17	0.03	0.03	0.05	0.06
Calderas	−90.59	14.41	1800	25.8	1880	NA	0.26	0.26	0.02	0.15	0.16	0.03	0.04	0.03	0.05
Chiligatoro	−88.18	14.38	1925	23.5	1949	NA	0.86	0.04	0.00	0.01	0.01	0.01	0.01	0.04	0.01
Chichancanab	−88.77	19.88	1	25.5	1194	1561	0.38	0.20	0.04	0.11	0.14	0.04	0.06	0.01	0.02
Caballo	−98.11	18.89	2158	16.4	754	1873	0.20	0.03	0.00	0.20	0.32	0.02	0.02	0.07	0.14
Lacaudon	−91.59	17.02	812	22.0	1940	1343	0.44	0.14	0.02	0.14	0.13	0.04	0.05	0.01	0.02
Las pozas	−90.17	16.34	152	26.3	2231	NA	0.52	0.25	0.02	0.04	0.10	0.02	0.05	0.00	0.01
Quexil	−89.81	16.92	120	25.4	2012	NA	0.57	0.17	0.02	0.07	0.10	0.03	0.03	0.01	0.01
Waha	−91.60	16.98	850	22.0	1940	1343	0.40	0.20	0.04	0.12	0.11	0.07	0.04	0.02	0.01
Chicabal	−91.66	14.79	2742	17.5	1732	NA	0.44	0.01	0.00	0.42	0.03	0.01	0.00	0.10	0.00
Este centro Amatitlán	−90.55	14.45	1204	25.8	1880	NA	0.33	0.06	0.00	0.05	0.40	0.02	0.05	0.01	0.08
Gruta San Miguel	−89.00	19.93	32	26.0	1100	1899	0.34	0.23	0.09	0.07	0.10	0.05	0.08	0.01	0.01
Olomega	−88.06	13.31	66	23.8	1841	NA	0.51	0.14	0.02	0.08	0.16	0.02	0.05	0.00	0.02
Teometitla	−97.99	19.45	2550	14.5	647	1730	0.23	0.04	0.00	0.38	0.16	0.02	0.02	0.09	0.06
Grande litoral	−90.17	13.89	5	27.7	1552	NA	0.75	0.02	0.00	0.15	0.06	0.00	0.00	0.01	0.01
Laguna Chica	−97.27	18.60	2459	13.4	762	1286	0.15	0.07	0.03	0.20	0.17	0.12	0.04	0.12	0.12
Requena	−99.34	19.93	2132	16.1	540	1766	0.32	0.04	0.04	0.24	0.20	0.03	0.02	0.06	0.05
La Soledad	−97.45	19.97	732	20.2	4171	1001	0.32	0.06	0.01	0.23	0.16	0.03	0.03	0.08	0.07
Salpeten	−89.68	16.98	105	25.4	2012	NA	0.46	0.16	0.04	0.11	0.12	0.03	0.04	0.01	0.03

2.4. BrGDGT-based proxy calculations

BrGDGTs can be categorized into three different compound groups based on structural characteristics: tetramethylated (Ia, Ib, and Ic), pentamethylated (IIa, IIb, and IIc, and hexamethylated (IIIa, IIIb, and IIIc), which can be separated based on the number of methyl branches on the alkyl chains (Fig. S1). Similarly, comparing the cyclization of brGDGTs, they can be categorized into three different groups, compounds with 0 (Ia, IIa, and IIIa), 1 (Ib, IIb, and IIIb), and 2 (Ic, IIc, and IIIc) cyclopentane moieties.

Two modified proxies are used herein: 1) MBT'$_{5ME}$, defined as the degree of methylation of 5-methyl brGDGTs:

$$MBT'_{5ME} = \frac{Ia + Ib + Ic}{Ia + Ib + Ic + IIa + IIb + IIc + IIIa}$$

which shows the relative abundance of tetramethylated groups in all compounds (De Jonge et al., 2014a) and 2) the isomer ratio (IR):

$$IR = \frac{IIa' + IIb' + IIc' + IIIa'}{IIa + IIa' + IIb + IIb' + IIc + IIc' + IIIa + IIIa'}$$

which expresses the fractional abundance of the penta- and hexamethylated 6-methyl brGDGT isomers compared to the total of penta- and hexamethylated brGDGTs (De Jonge et al., 2015).

2.5. Statistical analysis

BrGDGTs are all shown as standardized fractional abundances for all calculations. Statistical analyses were performed using MATLAB Version 9.8 (R2020a) to develop suitable calibrations for the regional datasets. We performed a series of statistical models for calibration developments, including stepwise regression, ridge regression, and ordinary least squares regression. In all models, p-values of 0.05 were specified as the threshold for model validation. In addition, correlation coefficients (r values) and Root Mean Squared Error (RMSE) were calculated. Principal component analysis (PCA) and Redundancy Analysis (RDA) were performed using Canoco 5 for better visualization and summary of the distribution of brGDGTs (expressed in fractional abundance) and the explanatory environmental variables. For all statistical analyses, compounds that were below the detection limit (mainly IIIb and IIIb') in most of the samples were excluded to avoid bias. To assess significant differences in predicted temperatures among the previously published temperature calibration models and those in the current study, we performed a heteroscedastic one-way ANOVA for trimmed means.

3. Results

3.1. Distributions of brGDGTs and environmental controls

Most of the brGDGTs were found in all sediment samples analyzed in the present study. The structures and fractional abundance of the 9 major brGDGTs (Ia, Ib, Ic, IIa, IIa', IIb, IIb', IIIa, IIIa') can be found in Supplemental Material and Table 1. In most samples, one pentamethyl brGDGT and its isomer (IIc and IIc') are significantly less abundant than the other major brGDGTs; integrating and quantifying them would likely introduce bias in calculating the fractional abundance. In addition, the abundances of some hexamethyl brGDGTs with cyclopentane moieties, including IIIb, IIIb' and IIIc and IIIc' are frequently below the instrumental detection limit and, even if detected, comprise <1 % of total abundance. Thus, these six compounds were excluded from calculating the total abundance of brGDGTs.

A principal components analysis (PCA) with the remaining nine major brGDGTs as the response variables shows a total variation of 799. The first two axes explain 77 % of the variance, with the first axis explaining 44 % and the second axis explaining 33 % (Fig. 2a). BrGDGT Ia has significant positive loading while IIIa' has a negative loading on axis 1, whereas most of the cyclic and 6-methyl compounds showed a significant negative loading on axis 2 (Fig. 2a).

A redundancy analysis (RDA) summarizes the part of the variation in brGDGT distribution explained by environmental variables, including Mean Annual Temperature (MAT, analogous to MAAT/MAF), Mean Annual Precipitation (MAP), and Evapotranspiration (Fig. 2b). The result shows a total variation of 321, and explanatory variables account for 78.3 % (Fig. 2b). The first two axes cumulatively explain 68 % of the variation. MAT is significantly correlated with the first axis and shows a strong negative correlation with IIIa, while other environmental factors did not contribute significantly to the variation of brGDGT distribution.

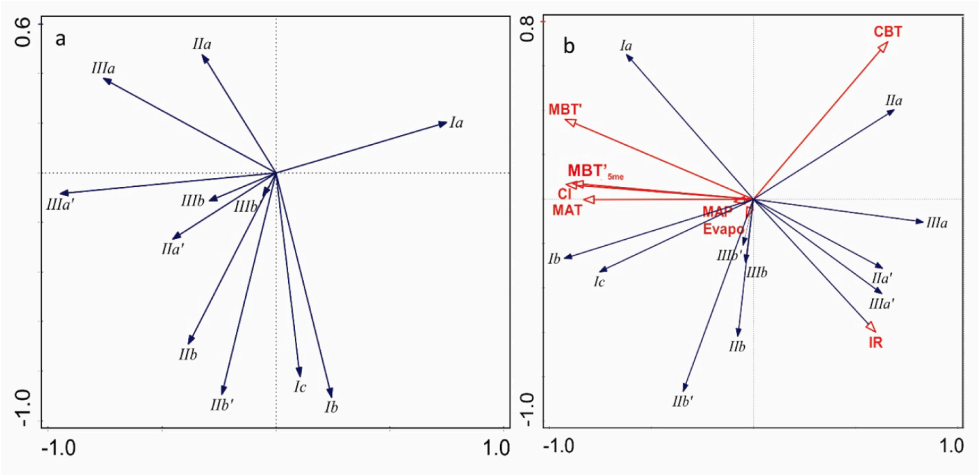


Fig. 2. Biplots of (a) the principal component analysis (PCA) and (b) Redundancy Analysis (RDA) based on the fractional abundance of brGDGTs. Environmental variables included in the constrained ordination are Mean Air Temperature (MAT; considered to be Mean Annual Air temperature and Months Above Freezing Temperature for out study sites), Mean Annual Precipitation (MAP) and evapotranspiration (Evapo).

3.2. Mexican Plateau (MP)

BrGDGT Ia is the most abundant in over half of the samples, and IIa or IIa' is the second most abundant compound (Fig. S2). Overall, Ia comprises 31 % of total brGDGT abundance, though it spans a wide range from 15 to 48 % (Fig. S2). In general, cyclic (Ib, Ib', IIb, and IIb') and bicyclic (Ic) brGDGTs are less abundant compared to the noncyclic ones (Ia, IIa, IIa', IIIa, and IIIa'). IR of brGDGTs ranged from 0.22 to 0.84, with an average of 0.51. The total fractional abundance of all 6-methyl brGDGTs (IIa'+IIIb'+IIIa') is ~ 31 %.

3.3. Yucatán Peninsula and Central America (YCA)

BrGDGT Ia has the greatest fractional abundance, accounting for 47 % of the total abundance. The proportion of IIa, IIa', and Ib are very close, about 12–14 % of the total brGDGT abundance (Fig. S3). The rest of the brGDGTs comprise <15 % of the total abundance combined (Fig. S3). Overall, the distribution of brGDGTs in the YCA is different from that in the MP, with Ia showing higher dominance in the YCA. IR of brGDGTs ranged from 0.6 to 0.87, with an average of 0.52. The fractional abundance of total 6-methyl brGDGTs is around 19 %.

3.4. Comparison of global and regional calibrations

The heteroscedastic one-way ANOVA for trimmed means revealed no significant difference ($F_{trimmed\text{-}means}(4, 657.56) = 0.15$, $p = 0.96$, $CI_{95\%}[0.02, 0.09]$, $n = 2{,}197$) in the predicted temperatures among the four global temperature calibration models (i.e., Raberg et al.., 2021 Full set, Martinez-Sosa et al., 2021, global ridge regression (this study), global stepwise forward selection (this study)). We also observed no significant difference in the predicted temperatures for the sample set spanning Mexico and Central America (MCA; $F_{trimmed\text{-}means}(5, 164.98) = 1.75$, $p = 0.13$, $CI_{95\%}[0.10, 0.30]$, $n = 603$) or the sample set from the Mexican Plateau (MP; $F_{trimmed\text{-}means}(7, 51.21) = 2.19$, $p = 0.05$, $CI_{95\%}[0.22, 0.48]$, $n = 208$). However, the heteroscedastic one-way ANOVA for trimmed means revealed there was a significant difference ($F_{trimmed\text{-}means}(7, 152.29) = 3.49$, $p = 1.70e\text{-}03$, 95 % [0.20, 0.46], $n = 592$) in the predicted temperatures for the Yucatan and Central America samples between at least two of the seven (three global, three regional, and two subregional) temperature calibration models (Fig. S4). A subsequent Yuen's trimmed means test showed there were significant differences in the predicted temperatures between the global ridge regression (this study) and each of the following, MCA SFS (this study), Raberg et al., 2021 (Full set), YCA MBT'5ME (this study), and YCA SFS (this study) (Fig. S4). No significant differences were found among the other temperature calibration models.

4. Discussion

4.1. Lacustrine brGDGT Distributions

The difference in the distribution of brGDGTs between lacustrine and soil environments has been well recognized since the last decade (Martínez-Sosa et al., 2021; Russell et al., 2018; Weber et al., 2015; Buckles et al., 2014; Blaga et al., 2010; Tierney and Russell, 2009). Initially, brGDGTs were identified in terrestrial environments, including peats and soils, and brGDGTs in lake sediments were thought to be derived from catchment soils (Weijers et al., 2006; Hopmans et al., 2004). However, studies focused on lacustrine environments have increasingly shown that brGDGTs in lake sediments and water columns have distinct distributions (Martínez-Sosa et al., 2021; Zhao et al., 2021; Dang et al., 2018; Russell et al., 2018; Buckles et al., 2014; Loomis et al., 2011; Tierney and Russell, 2009). Furthermore, a novel 5/6-methyl brGDGT, which elutes between the known 5- and 6-methyl compounds IIIa and IIIa' during HPLC–MS analysis, was found in lakes but not soils from the catchment (Weber et al., 2015). These studies provide circumstantial evidence for *in situ* brGDGT production in lacustrine environments, in the sediment and/or water column (Martínez-Sosa et al., 2021; van Bree et al., 2020). This study primarily compares our results with lake sediment samples from other regions. In the study area, 9 major brGDGTs were detected with a relative abundance distribution pattern similar to those reported in other studies of lacustrine surface sediments (Zhao et al., 2021; Russell et al., 2018).

The isomer ratios (IR) in the two subregions are similar, with MP 0.51 and the YCA 0.52. These are both higher than the average IR value of 0.24 in the global soil dataset (De Jonge et al., 2014a), 0.24 in Lake 578 in Greenland (Zhao et al., 2021), 0.43 in Central European lakes (Weber et al., 2018), 0.22 in Canadian and Icelandic Lakes (Raberg et al., 2021) and 0.42 in African Lakes (Russell et al., 2018). The only similar IR value was from a suite of Chinese lakes (with a broad range, 0.35–0.88, mean = 0.52 (Dang et al., 2018)). In contrast to the generally agreed assumption that temperature is mainly reflected by 5-methyl brGDGTs, in Chinese lakes, the 6-methyl brGDGTs also appear sensitive to temperature (Dang et al., 2018), and in Lake Garba Guracha (Ethiopia) the inclusion of the 6-methyl isomers improved temperature reconstruction (Bittner et al., 2022). Despite having higher IR values, the total fractional abundance of 6-methyl brGDGTs is still low in our sample set (0.30–0.63 with an average of 0.22). Only 4 lakes are

dominated by 6-methyl brGDGTs. In addition, there is no apparent spatial pattern or strong relationship between IR and temperature (Fig. S5).

It has been shown that IR in saline lakes is systematically higher than in freshwater lakes, and salinity affects the relative abundance of 5-methyl brGDGTs (Wang et al., 2021). Some of the lakes in our study are saline (Sigala et al., 2017; Pérez et al., 2011); for example, Laguna Grande (IR = 0.55) and Laguna Chichancanab (IR = 0.57). We acknowledge that salinity can affect the relative abundance of 5-methyl brGDGTs versus their isomers. However, due to the lack of detected isomers, salinity effect correction was not performed (Wang et al., 2021). Novel brGDGTs discovered in other lakes, such as the 5/6-methyl isomer of hexamethylated brGDGT (denoted as IIIa“) (Weber et al., 2015) and 7-methyl brGDGTs which are typically found in saline lakes (Ding et al., 2016) were not consistently detected either. Another possible reason for the relatively high IR values could be pH since the relative amount of 6- to 5-methyl brGDGTs appears to be highly correlated with pH in the global soils dataset (De Jonge et al., 2014a). However, this possibility remains unclear as controlled lacustrine microcosms had an inconclusive result for pH influence in IR (Martínez-Sosa and Tierney, 2019).

A significant difference was observed when comparing the distribution of brGDGTs (categorized as tetra-, penta-, and hexamethylated) in Mexico and Central America with other regions worldwide (Fig. 3). In our study area, tetramethylated brGDGTs have a noticeable dominance in most samples, and most hexamethylated brGDGT abundances are <10 %. Previous studies have reported an increase of hexamethylated brGDGTs in the sediment of high-elevation lakes (Loomis et al., 2011, 2014), indicating that hexamethylated brGDGTs were favored by lower average temperatures. Maintenance of the fluidity and permeability of the membrane during changing environments relies on the microbes' ability to rapidly adjust the membrane lipid composition, so it is reasonable to see changes in methylation with changing MAAT. Thus, it seems likely that the lower fractional abundance in our samples is mainly due to the higher temperatures of our study sites relative to the global dataset.

Indeed, a study investigating adaptations of bacterial membranes to variations in temperature, pH, and pressure demonstrated that the presence of membrane-spanning ether-lipids and methyl branches has a striking relationship with growth conditions (e.g., temperature) (Siliakus et al., 2017). Moreover, in molecular dynamics simulation experiments, specifically in archaeal-like membranes, the presence of methyl groups was shown to confer greater fluidity (Chugunov et al., 2014). It is suggested that the methyl group disturbs the packing order of the hydrocarbon chains of the membrane lipid. Thus, at low temperatures, bacteria could potentially maintain fluidity partially by adjusting the methyl groups (Vinçon-Laugier et al., 2017), which is consistent with the observations in global lacustrine brGDGT distributions (Fig. 3). Two recent experiments on molecular dynamics simulations of membranes consisting of brGDGTs and cultured strain Ellin6076 have confirmed the temperature sensitivity of the degree of methylation in brGDGTs (Chen et al., 2022; Naafs et al., 2021). More studies aimed at identifying and culturing the brGDGT-specific bacterial producer(s) will greatly benefit the explanation of the brGDGT distribution patterns across temperature gradients.

In the RDA plot (Fig. 2b), MAT has a large negative loading on axis 1, which explains ~ 68 % of the variance, implying that MAT is a major controlling factor but also that there are contributions from other environmental factors in the brGDGT distribution. We acknowledge that pH, one of other environmental factors widely recognized to influence brGDGT distribution (Martínez-Sosa et al., 2021; Duan et al., 2020; Naafs et al., 2017; Xiao et al., 2015), could contribute to the distribution. However, in most lacustrine-focused studies, pH is not the primary control (Martínez-Sosa et al., 2021; Zhao et al., 2021; Russell et al.,

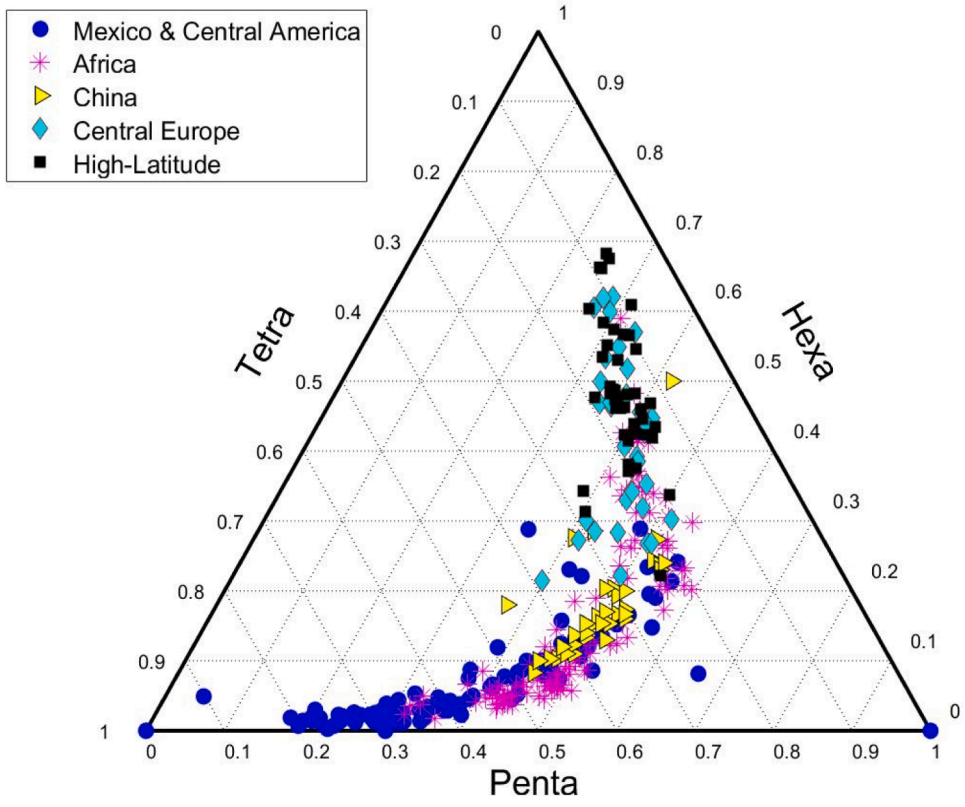


Fig. 3. Ternary plot of brGDGT distributions (categorized in three groups: tetramethylated, pentamethylated, and hexamethylated) across different regions including Mexico and Central America (blue dots, this study), East Africa (pink asterisks) (Russell et al., 2018), China (yellow triangles) (Dang et al., 2018; De Jonge et al., 2014a), Central Europe (cyan diamonds) (Martínez-Sosa et al., 2021; Weber et al., 2018), and high latitude areas including Canada and Iceland (black squares) (Raberg et al., 2021). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2018; Loomis et al., 2014a). Future studies on other environmental factors such as pH and water conductivity and their influences on brGDGT distributions in Mexico and Central American lake sediments will enhance our understanding of environmental controls and potential bias in the brGDGT-MAT relationship.

4.2. Development of Lacustrine brGDGT-Temperature Calibrations

4.2.1. Global lacustrine calibrations

Two recent studies have developed global lacustrine calibrations with different statistical approaches: a Bayesian calibration, which has an $R^2 = 0.82$ and RMSE = 2.9 °C (Martínez-Sosa et al., 2021), and a revised fractional abundance calibration, which has an $R^2 = 0.91$ and RMSE = 1.97 °C (Raberg et al., 2021). Both of these calibrations were developed using Months Above Freezing (MAF) temperature instead of MAAT (though we note that in our samples MAAT and MAF are equivalent). Because these calibrations generally included lakes at high latitudes, using MAF will help to reduce the offset from tropical lakes to cold higher latitude lakes. Indeed, brGDGT distributions predicted MAF better than MAAT in both statistical models. When applying these calibrations to the sample set in this study, both of the calibrations yield consistent results close to each other (within each other's RMSE; Fig. 4). However, they appeared to have a better prediction at temperatures > 20 °C while underestimating the temperature at sites with MAF < 15 °C (Fig. 4). Indeed, when the temperature is below 15 °C, the MAF is not within the error range of both temperature calibrations, implying a cold bias starting at ~ 15–20 °C and potentially increasing towards colder regions (Fig. 4).

This cold bias could be related to samples from high latitude regions used in the development of global calibrations. As discussed previously, samples from colder areas are likely to have significantly more hexamethylated brGDGTs, and this relationship is nonlinear. Indeed, there is a logarithmic correlation between 5-methyl hexamethylated brGDGTs (IIIa, IIIb, and IIIc) and temperature (Fig. S6). Therefore, this non-linear increase in the relative amount of hexamethylated brGDGTs will result in lower MBT values, reflecting colder reconstructed temperatures. Thus, it is not surprising to see the global calibrations introduce bias in the lower range of temperature reconstructions (Fig. 4).

Both of the newly developed global calibrations show that using MAF instead of MAAT can yield better model fitting (Martínez-Sosa et al., 2021; Raberg et al., 2021), ideally helping to address potential problems from seasonality bias at higher latitudes and elevations. It is suggested that during winter, microbial activity will likely be reduced, leading to lower production of brGDGTs during this time (Dearing Crampton-Flood et al., 2020; Miller et al., 2018; Loomis et al., 2014b). A recent study revealed that brGDGTs did not record temperatures below freezing due to ice coverage on lake surfaces in winter (Cao et al., 2020). However, previous studies have suggested that no seasonality is observed in soils (Lei et al., 2016; Weijers et al., 2011). If there were seasonality influences, the reason some previous studies did not show the effect could possibly be that the microbial activity will be reduced only when reaching a threshold. Furthermore, because the turnover time of brGDGTs in soils is long (Huguet et al., 2014; Peterse et al., 2010; Weijers et al., 2010), any change that is not significant enough will likely be obscured by the build-up of brGDGTs in the soil over decades.

4.2.2. Reevaluation of global lacustrine calibrations

Previously developed global calibrations included very few samples from the North American tropics, where temperature reconstructions are really needed. To fill in the blanks and improve the global

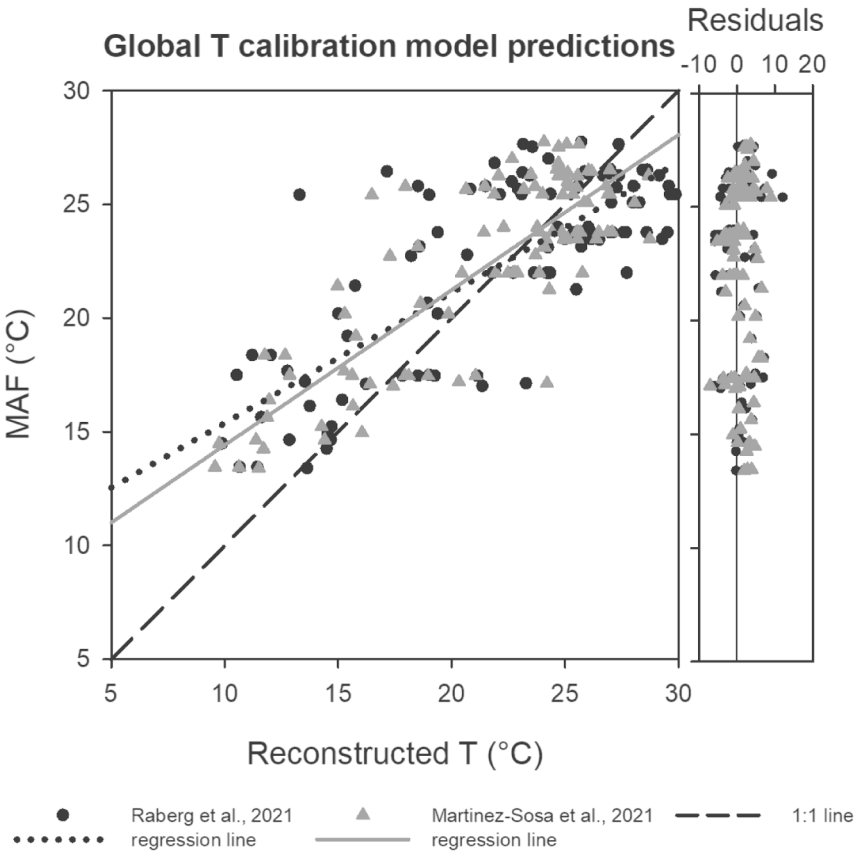


Fig. 4. Reconstructed temperature in the study area from the two newly developed global calibrations (Martinez-Sosa et al., 2021; Raberg et al., 2021) against instrumental Months Above Freezing temperature (MAF). The X axis indicates the reconstructed temperatures, the Y axis is the instrumental MAF. The dotted line is the reference line for regression between MAF and reconstructed temperature using the Raberg et al. (2021) revised fractional abundance calibration and the solid grey line is the regression line using the Martinez-Sosa et al. (2021) Bayesian calibration. The residual plot indicates that both calibrations tend to reconstruct slightly lower temperatures compared to the instrumental temperatures, particularly at lower temperatures.

calibrations, as well as to evaluate the cause of the cold bias at lower temperatures in the previous global calibrations, we summarized the published data of lacustrine brGDGT surface sediment samples and combined these with our sample set to investigate their distribution pattern along temperature gradients at a global scale (Fig. 5). For samples from mid-high latitude areas including Central Europe (Weber et al., 2018), Canada and Iceland (Raberg et al., 2021), and China (Cao et al., 2020; Ning et al., 2019; Dang et al., 2018; Qian et al., 2019), we used MAF rather than MAAT following the better statistical results from the global calibration. The ternary plot (Fig. 5) characterizes the brGDGT distribution pattern of the global sample set. Along the tetramethylated brGDGT fractional abundance axis, the brGDGTs fall into a wide distribution range (Fig. 5). On the pentamethylated fractional abundance axis, a relatively linear distribution for sample groups with higher temperatures (red - yellow dots) was observed (highlighted with a red circle). The linear pattern disappeared along the pentamethylated axis for samples with colder temperatures but appeared along another axis, the hexamethylated fractional abundance axis (blue-green dots).

In this study, we performed a new machine learning parsimonious model – ridge regression (Hoerl and Kennard, 1970) - to evaluate how the distribution of brGDGTs reflects the temperature (grouped by methyl number, e.g. tetramethylated, pentamethylated and hexamethylated brGDGTs are denoted as 'Tetra', 'Penta', and 'Hexa'). Ridge regression is a model tuning method performing L2 regularization (Schreiber-Gregory, 2018). When the issue of multicollinearity occurs in predictor variables (in this case, fractional abundance of brGDGTs, Fig. S5), and variances are large, the result is predicted values that are far away from the actual values. Ridge regression can avoid overfitting by adding an L2 penalty, which equals the square of the magnitude of coefficients. Because all coefficients are reduced by the same factor, it will not perform feature selections; thus, none of the predictor variables will be eliminated (unlike L1 regularization which might result in sparse models). In the ridge regression model, 5-methyl brGDGTs categorized by their number of methyl groups were used because they have been shown to have the strongest response to temperature (Martínez-Sosa et al., 2021; Raberg et al., 2021; Russell et al., 2018). Additionally, to better compare with the previous global calibrations in which MBT'_{5ME} is used, keeping the predictor brGDGTs consistent will help with parallel evaluations.

The resulting calibration is:

$$MAF = 15.2 + 16.6 \times Tetra - 18.7 \times Penta - 18.8 \times Hexa$$

$$R^2 = 0.83,\ RMSE = 3.1^{\circ}C,\ n = 440$$

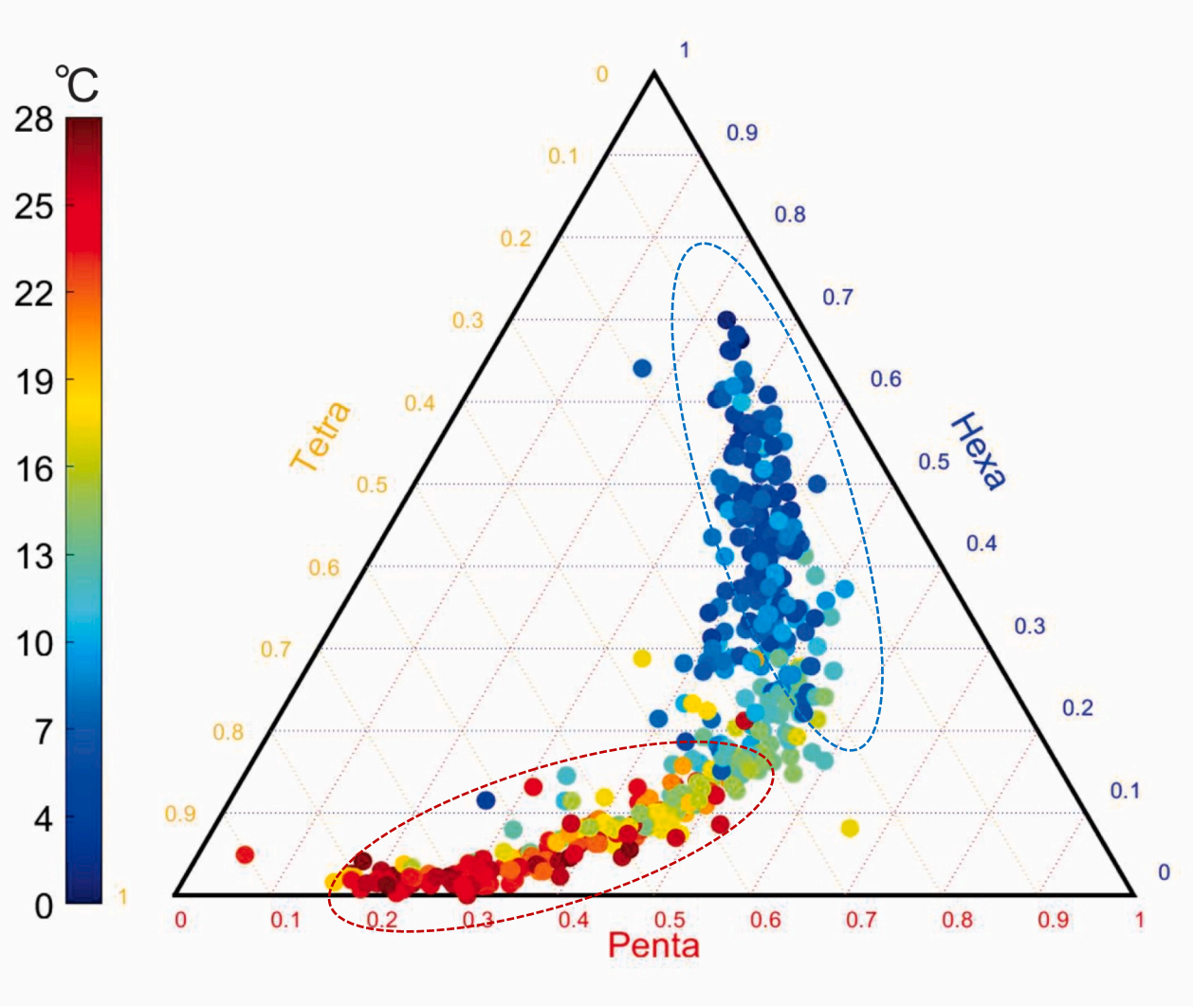


Fig. 5. Ternary plot showing the global lacustrine sample distribution of tetra-, penta-, and hexamethylated brGDGTs with temperature gradient (°C; temperature increases from blue to red). For colder regions, MAF instead of MAAT were used. Plot includes data from this study and previously published studies and references therein (Martínez-Sosa et al., 2021; Raberg et al., 2021; Dang et al., 2018; Russell et al., 2018; Weber et al., 2018). Dashed circles indicated the warm group (red) and cold group (blue) of samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

As mentioned above, ridge regression did not involve feature selection. Therefore, we performed a stepwise selection with the same set of predictor variables:

$$MAF = 15.3 + 16.6 \times Tetra - 18.7 \times Penta - 19.5 \times Hexa$$

$$R^2 = 0.83,\ RMSE = 3.1\ ^{\circ}C,\ n = 440$$

These two calibrations are almost identical in terms of coefficients and statistical parameters. Results of this analysis support the assertion that the methyl branches play important roles in the membrane structure modification in response to temperature change (Fig. 6). We also perform the same ridge regression model with 6-methyl isomers included ($R^2 = 0.71$ and RMSE = 4.0), and the model with only 5-methyl brGDGTs performs better. Our results further confirm that 5-methyl isomers are more sensitive to temperature change (Chen et al., 2022; Naafs et al., 2021).

With more data from Mexico and Central America added to the global sample set, we also performed an ordinary least squares model on MBT'_{5ME} and MAF:

$$MAF = 0.3 + 28.2 \times MBT'_{5ME}$$

$$R^2 = 0.83,\ RMSE = 3.1\ ^{\circ}C,\ n = 440$$

This calibration is comparable with the two previous global calibrations developed (Table 2, Martínez-Sosa et al., 2021; Raberg et al., 2021). The new calibration exhibits a similar higher relative fit (R^2) yet lower absolute fit (RMSE). It is notable that with our NA (sub)tropics regional data included, the global calibration RMSE increased to 3.1 °C (Table 2). The temperatures of our study sample sites are predominantly at the upper end of the temperature range in the global sample set. Thus, the differences in these calibrations is likely related at least in part to the increase in the number of warm sites added, as well as the extension of the temperature range to warmer temperatures.

We further investigate the potential difference in the relationship between brGDGT distributions and temperature between cold and warm regions by dividing the global dataset into a higher temperature group and a lower temperature group, based on the distributions observed in Fig. 5 which shows distinct relationships between the relative abundance of tetra-, penta- and hexa-methylated brGDGTs at higher vs lower temperatures. The higher temperature group (Fig. 5, red circle) exhibits a general distribution pattern with low relative abundance of hexamethylated brGDGTs. There is a moderate reciprocal relationship between the relative abundance of tetramethylated and pentamethylated brGDGTs. The lower temperature group, in contrast, has low fractional abundance of tetramethylated brGDGTs (Fig. 5). The cold temperature subset displayed a homogenous distribution pattern along the hexamethylated fractional abundance axis, meaning that hexamethylated brGDGTs are the major group explaining the distribution, followed by the pentamethylated group. Thus, our data support the hypothesis that the methylation of brGDGTs is different at lower vs higher temperatures, consistent with the findings of De Jonge et al. (2019) in soils. Though our study does not identify any mechanisms that may be driving the observation, such as a change in microbial community or other (e.g. environmental) factors, it suggests an avenue for future research.

4.2.3. *Developing region-specific calibrations for Mexico and Central America*

To explore whether regional calibration could improve on the global models, we developed calibrations using region-specific sample sets. In our study site, the distribution of brGDGTs is relatively homogenous (Fig. 3). Therefore, we applied the widely accepted method of using MBT'_{5ME} to develop a regional temperature calibration (Fig. 7a), resulting in the following calibration:

$$MAF = 6.2 + 22.3 \times MBT'_{5ME}$$

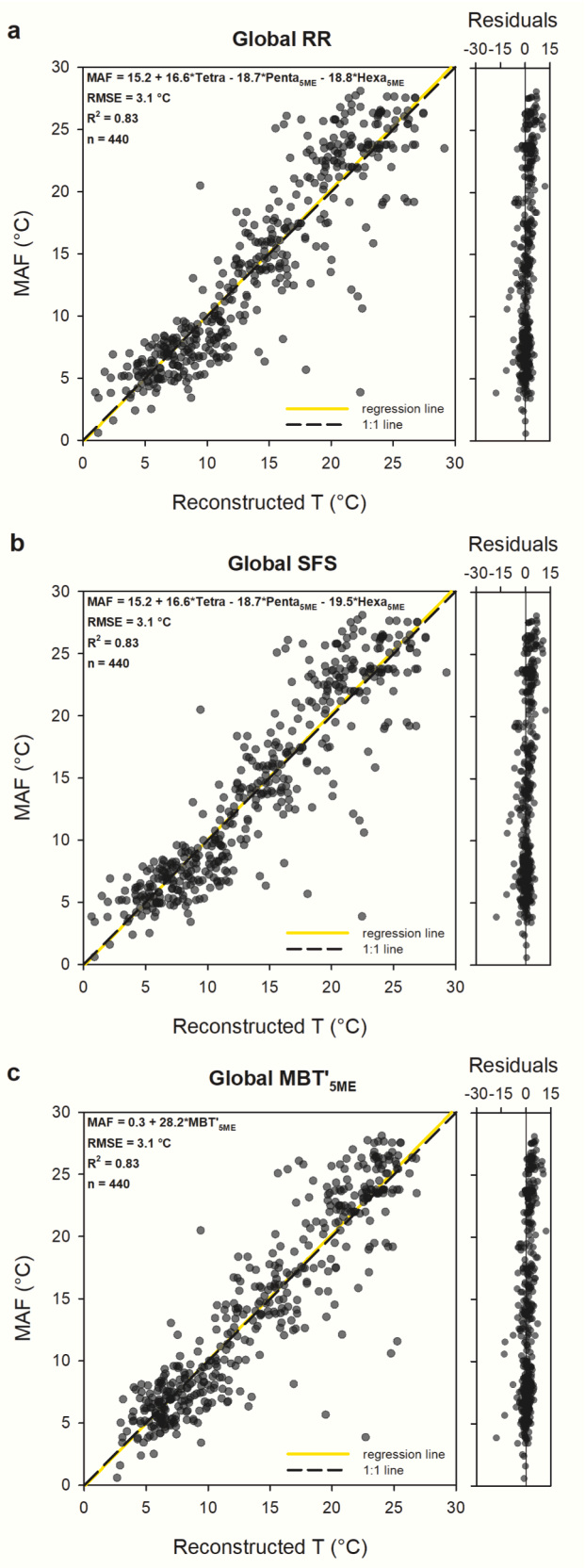


Fig. 6. Global calibrations incorporating new data from Mexico and Central America. A) Ridge regression on 5-methyl tetra-, penta- and hexamethylated brGDGTs. B) Stepwise forward selection regression. C. MBT'_{5ME} based regression. In all plots, individual samples are grey dots, darker-appearing dots indicate higher data density (i.e., the dots overlap and are darker).

Table 2
Comparison among MBT'$_{5me}$ based temperature calibrations from this study, East African Lakes and the previous published global sample sets.

	Slope	Intersept	RMSE (°C)	R^2
Global (This study)	28.2	0.3	3.1	0.83
Regional (This study)	22.3	6.2	2.3	0.68
East African Lake (Russell et al., 2018)	32.42	−1.21	2.44	0.92
Global (Martínez-Sosa et al., 2021)	33.3	−2.5	2.9	0.82
Global (Raberg et al., 2021;)	30.4	−0.5	2.14	0.89

$R^2 = 0.68$, RMSE = 2.3 °C, n = 101

This calibration looks slightly different compared to the most recent global and African lakes calibrations (Table 2, Martínez-Sosa et al., 2021; Raberg et al., 2021; Russell et al., 2018). Those calibrations have coefficients ranging from 30.4 to 32.42, higher than the new regional one with 22.32. This does not necessarily mean that our MBT'$_{5ME}$ calibration explains less of the variance in temperature. It could be because in the study area, the temperature range is narrower compared to that in the African lakes calibration (and of course global scale as well). Thus the explanatory ability of MBT'$_{5ME}$ is constrained when utilizing a regional dataset. Despite having a lower R^2 compared to other calibrations, the new model successfully enhanced its prediction ability at lower temperatures; however, a robust anova indicates that the regional calibration is not significantly different from the global calibrations (Fig. S4).

In addition to the linear regression, we also applied a multivariate linear regression model with stepwise forward selection (SFS) to find the best subset of the predictor variables that improve the calibration most significantly (Fig. 7b). This method has been used to significantly improve the statistical errors of lacustrine brGDGT calibrations in other regions (Russell et al., 2018; Loomis et al., 2012; Pearson et al., 2011). The SFS model result for our regional study is:

$$MAF = 19.34 + 7.55 \times Ia + 19.08 \times Ib - 9.19 \times IIa - 39.58 \times IIIa$$

$R^2 = 0.69$, RMSE = 2.3 °C, n = 101

The result from our MBT'$_{5ME}$ calibration had a very similar statistical error to our SFS model, unlike in other studies where SFS models generally exhibited a better fit with lower RMSE (Russell et al., 2018; Loomis et al., 2012). All the variables in our SFS calibration are 5-methyl brGDGTs, further confirming that in our study area, 5-methyl brGDGTs are more responsive to changes in MAF. F-statistics revealed that the coefficients of all the brGDGTs in this model were significant ($p < 0.05$), indicating that none of the above variables are overfitted.

Ridge regression performed with 5-methyl brGDGTs categorized in different methyl groups (Fig. 7c) shows a similar result compared to the SFS model:

$$MAF = 20.6 + 7.6 \times Tetra - 10.9 \times Penta - 35.2 \times Hexa$$

$R^2 = 0.66$, RMSE = 2.3 °C, n = 101

The performance of these two models is similar, comparing the coefficients of the explanatory variables in the ridge regression model and SFS model. However, R^2 does not necessarily show the predictive ability of the model. Even though the ridge regression model is similar to the SFS model, the regression of reconstructed temperature appears to be slightly off intercept with the reference line.

4.2.4. Reducing scatter in the calibration

We observe a large scatter in both the SFS and MBT'$_{5ME}$ calibrations,

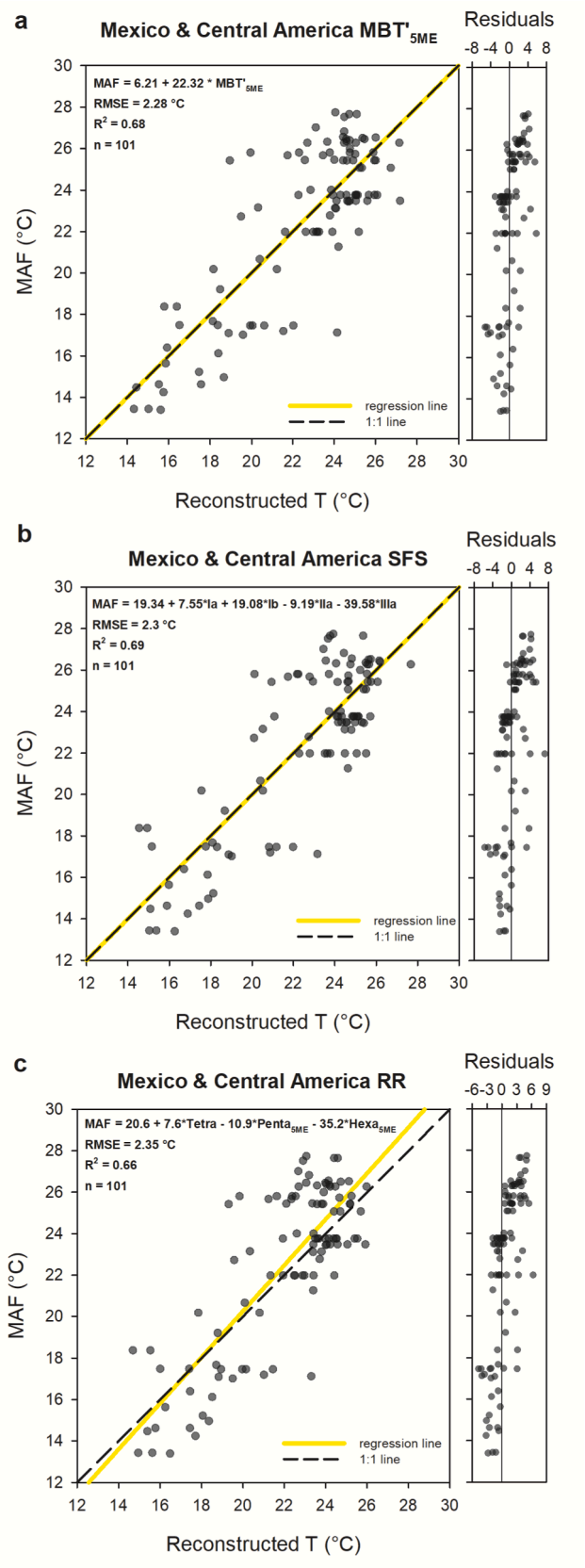


Fig. 7. Regional (Mexico & Central America) temperature calibration based on a) MBT'$_{5ME}$, b) Stepwise Forward Selection (SFS), and c) ridge regression. In all plots, individual samples are grey dots, darker-appearing dots indicate higher data density.

yet RDA shows that air temperature is the primary environmental factor influencing MBT'$_{5ME}$ (Fig. 2b). This observation, in conjunction with the difference in brGDGT distribution between the MP and the YCA subregions (Fig. S2 and S3), suggests that the scatter in the MBT'$_{5ME}$ calibrations may be due to the internal diversity within our study region. Here we perform the same statistical analyses for these two subsets of samples separately.

Surprisingly, MBT'$_{5ME}$ did not have a reasonable correlation with MAF in the YCA area (R^2 0.39, Fig. 8b). Previous analyses showed that brGDGT IIIa and/or IIIa' were more strongly correlated with temperature than any other brGDGT (Martínez-Sosa et al., 2021; Raberg et al., 2021; Loomis et al., 2012; Tierney et al., 2010). In studies where MBT'$_{5ME}$-based calibrations are robust, the fractional abundance of IIIa has shown a significant correlation with MAF (Dang et al., 2018; Russell et al., 2018; Loomis et al., 2012). In the YCA area, no individual brGDGTs exhibit a significant correlation with MAF, and the relative abundance of IIIa only accounts for 2 %, making it one of the least abundant brGDGTs (the other is IIIa', 2 %, Fig. S3). This might contribute to a lack of correlation between brGDGTs and MAAT in this region. While brGDGT-based calibration appears unfavored in this subregion, it implies that brGDGTs might not necessarily respond to temperature change primarily, and other environmental factors, such as pH, lake water alkalinity, redox, or water depth, could significantly influence the distribution of brGDGTs in lakes in this region (Schoon et al., 2013). More study is needed to address this specific problem.

A mixed source of brGDGTs could also be another reason why brGDGTs in the YCA area do not capture a temperature gradient. A few of the samples (n = 4) were collected from sinkholes, known as cenotes (Perry et al., 1995), throughout the Yucatán Peninsula and southern Campeche area. These cenotes are connected to the groundwater systems, in which it is highly likely that *in situ* brGDGT production exists that is distinct from the other lacustrine environments. Indeed, many of these cenotes are connected to ocean waters, which could be influencing the brGDGT composition in unanticipated ways. For example, cenotes are often stratified, with marine-connected saline waters underlying fresh waters, with stratification leading to anoxia in the saline deep waters. Recent studies have documented that redox conditions exert influence on brGDGT distributions, and thereby temperature

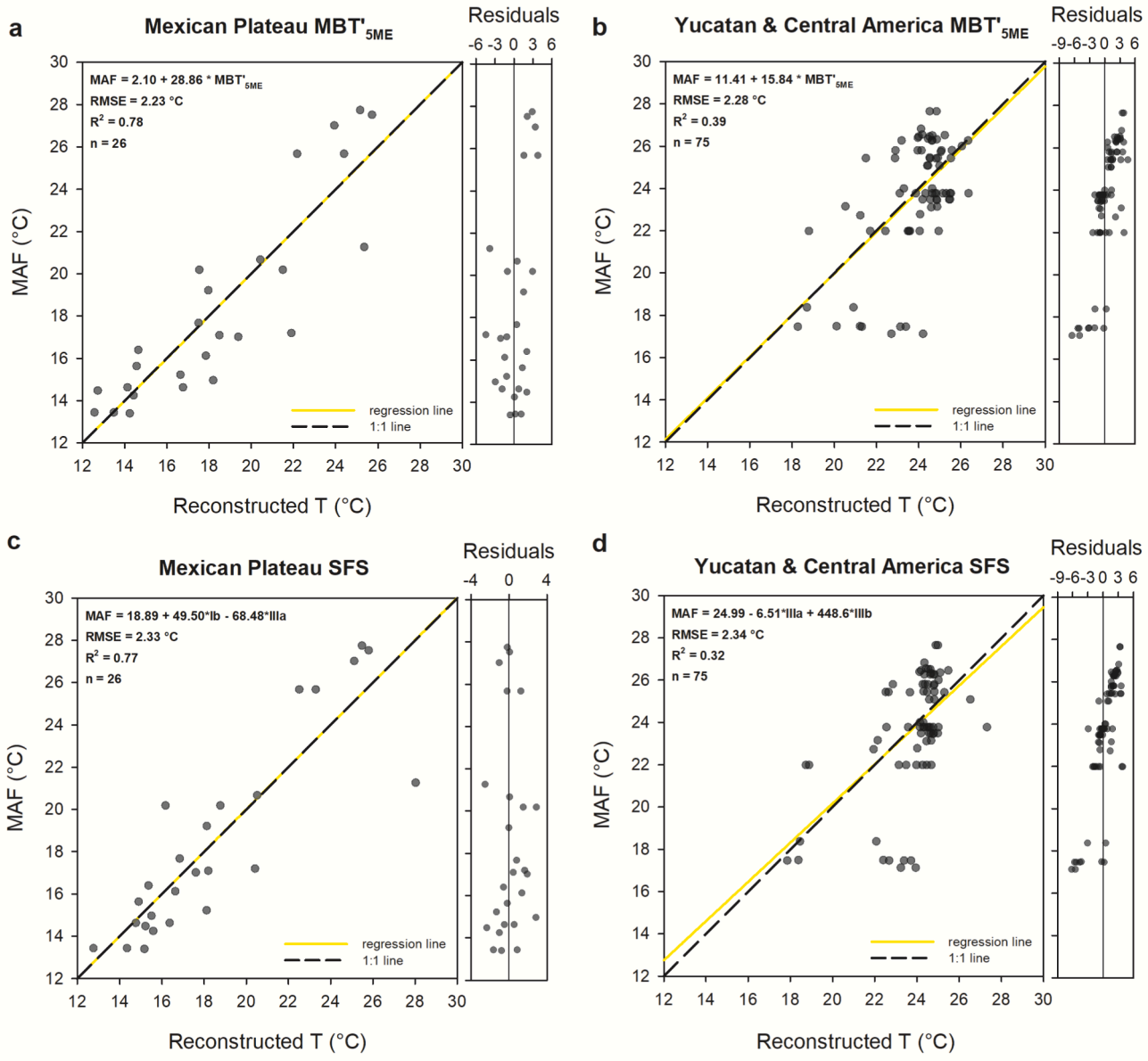


Fig. 8. Temperature calibrations for the Mexican Plateau and YCA regions based on MBT'5ME (a and b, respectively) and Stepwise Forward Selection (SFS; c and d respectively). In all plots, individual samples are grey dots, darker-appearing dots indicate higher data density.

reconstructions (Wu et al., 2021; van Bree et al., 2020; Yao et al., 2020; Weber et al., 2018).

In contrast, in samples from the Mexican Plateau, brGDGTs are primarily controlled by MAF, and by excluding samples from the YCA, the MTB'$_{5ME}$ -based calibration is improved to an $R^2 = 0.78$ with a low RMSE = 2.2 °C.:

$$MAF = 2.1 + 28.9 \times MBT'_{5ME} (n = 26)$$

Despite the similarity in all the above reported RMSEs (all < 2.30 °C), this model focused only on the Mexican Plateau subregion has the best relative fit with an $R^2 = 0.78$ (Fig. 8a).

We also performed an SFS analysis for a fractional-abundance-based calibration using only samples from the MP (Fig. 8c). The two selected brGDGTs are Ib and IIIa, showing significant coefficients ($p < 0.05$) tested by F-statistics. The calibration is:

$$MAF = 18.9 + 49.5 \times Ib - 68.5 \times IIIa$$

$$R^2 = 0.77, RMSE = 2.33\,^{\circ}C, n = 26$$

Like other SFS models in this study, only 5-methyl brGDGTs were selected in the calibration. However, in the SFS models, only two brGDGTs were selected. Both calibrations developed using samples from only the MP provide better results compared to previous calibrations that included samples from the YCA.

4.3. Implications for Continental Paleotemperature Reconstructions in Southern North America

All RMSE values from the different models/calibrations developed herein are similar. Indeed, other than the regional calibration focused on the MP, there is no statistically significant difference among the various calibrations. Therefore, when reconstructing temperature downcore, we recommend performing validation experiments to select the best models. For example, as discussed above, it is highly possible that brGDGTs may not be suitable for paleotemperature reconstructions using sediment cores from at least some lakes in the Yucatán Peninsula area, where groundwater influences can be prominent. In addition, for areas that are likely to have a mixture of brGDGT sources, we urge caution in applying any of these calibrations because the brGDGTs in sediments can represent a combination of different environmental signals from a large catchment area (De Jonge et al., 2015; Zell et al., 2014), thereby introducing significant bias into the reconstructed temperatures.

We suggest that for temperature construction in the Mexican Plateau region, the new regional calibrations based on the MBT'$_{5ME}$ proxy and fractional abundance of brGDGTs can be applied to cross-compare (RMSE of 2.2–2.3 °C). While the global calibrations may have better R^2, they are likely to introduce bias when reconstructing lower temperatures (<15 °C) in the study area. In a broader regional scale, the distribution pattern of brGDGTs could vary; therefore, the fractional abundance-based model could introduce bias in another region where brGDGT distribution is different. The MBT'$_{5ME}$ proxy expresses a sum of the brGDGT distribution pattern and is shown to correspond to temperature globally. Therefore, within the MP region, MBT'$_{5ME}$ calibrations are recommended for paleotemperature reconstruction.

5. Conclusions

In this work, we reevaluated the global lacustrine surface sediment sample set with 101 new lacustrine surface sediment samples from Mexico and Central America. With the new subset of samples introduced to the global dataset, the statistical parameters in the new calibrations altered slightly, but variation among the different calibrations was not significant. This highlighted the regional difference in brGDGT distribution and therefore temperature calibration in the study area. A new machine learning model was introduced for developing brGDGT-temperature calibrations. The result from the new ridge regression model is similar to other global calibrations developed using SFS and linear regression models, implying that the methyl group distribution accounts significantly for temperature variation. We also provided a series of new regional calibrations that address the cold bias observed in samples from Mexico and Central America when applying global calibrations. While all calibrations yield similar results, the MBT'$_{5ME}$ -MAF model using samples from MP provides the best relative fit for temperature reconstructions in this region.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the United States National Science Foundation (Grant P2C2-1804429 to JPW) and the University of Pittsburgh grants (CRDF and Hewlett International Grants to JPW). Sampling in the study was supported by the Universidad Nacional Autónoma de México. We thank Dr T. Elliott Arnold for technical assistance, and Dr. M.D. O'Beirne for assistance with statistical analysis. We thank the two anonymous reviewers for providing helpful and constructive comments. The results published here use data generated by International Research Institute for Climate and Society data library (http://iridl.ldeo.columbia.edu/).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.orggeochem.2023.104660.

References

Bittner, L., De Jonge, C., Gil-Romera, G., Lamb, H.F., Russell, J.M., Zech, M., 2022. A Holocene temperature (brGDGT) record from Garba Guracha, a high-altitude lake in Ethiopia. Biogeosciences 19. https://doi.org/10.5194/bg-19-5357-2022.

Blaga, C.I., Reichart, G.J., Schouten, S., Lotter, A.F., Werne, J.P., Kosten, S., Mazzeo, N., Lacerot, G., Sinninghe Damsté, J.S., 2010. Branched glycerol dialkyl glycerol tetraethers in lake sediments: Can they be used as temperature and pH proxies? Organic Geochemistry 41. https://doi.org/10.1016/j.orggeochem.2010.07.002.

Bree, V., Loes, G.J., Peterse, F., Baxter, A.J., De Crop, W., Van Grinsven, S., Villanueva, L., Verschuren, D., Sinninghe Damsté, J.S., 2020. Seasonal variability and sources of in situ brGDGT production in a permanently stratified African crater lake. Biogeosciences 17. https://doi.org/10.5194/bg-17-5443-2020.

Brown, E.T., Caballero, M., Cabral Cano, E., Fawcett, P.J., Lozano-García, S., Ortega, B., Pérez, L., Schwalb, A., Smith, V., Steinman, B.A., Stockhecke, M., Valero-Garcés, B., Watt, S., Wattrus, N.J., Werne, J.P., Wonik, T., Myrbo, A.E., Noren, A.J., O'Grady, R., Schnurrenberger, D., Abarca, R.M., Beltrán, A.O., Caballero, C., Cappio, L., Cossio, R., Ferland, T., Hesse, K., Kallmeyer, J., Kumar, D., Leon, S.G., Martínez, I., Noriega, C.A., Preusser, F., Rawson, H., Soler, A.M., Sosa-Nájera, S., Villeda, D.A., Zeeden, C., 2019. Scientific drilling of Lake Chalco, Basin of Mexico (MexiDrill). Scientific Drilling 26. https://doi.org/10.5194/sd-26-1-2019.

Buckles, L.K., Weijers, J.W.H., Verschuren, D., Sinninghe Damsté, J.S., 2014. Sources of core and intact branched tetraether membrane lipids in the lacustrine environment: Anatomy of Lake Challa and its catchment, equatorial East Africa. Geochimica et Cosmochimica Acta 140. https://doi.org/10.1016/j.gca.2014.04.042.

Cao, J., Rao, Z., Shi, F., Jia, G., 2020. Ice formation on lake surfaces in winter causes warm-season bias of lacustrine brGDGT temperature estimates. Biogeosciences 17. https://doi.org/10.5194/bg-17-2521-2020.

Chen, Y., Zheng, F., Yang, H., Yang, W., Wu, R., Liu, X., Liang, H., Chen, H., Pei, H., Zhang, C., Pancost, R.D., Zeng, Z., 2022. The production of diverse brGDGTs by an Acidobacterium allows a direct test of temperature and pH controls on their distribution. bioRxiv.

Chugunov, A.O., Volynsky, P.E., Krylov, N.A., Boldyrev, I.A., Efremov, R.G., 2014. Liquid but durable: Molecular dynamics simulations explain the unique properties of archaeal-like membranes. Scientific Reports 4. https://doi.org/10.1038/srep07462.

Crisp, D.T., Howson, G., 1982. Effect of air temperature upon mean water temperature in streams in the north Pennines and English Lake District. Freshwater Biology 12. https://doi.org/10.1111/j.1365-2427.1982.tb00629.x.

Dang, X., Yang, H., Naafs, B.D.A., Pancost, R.D., Xie, S., 2016. Evidence of moisture control on the methylation of branched glycerol dialkyl glycerol tetraethers in semi-arid and arid soils. Geochimica et Cosmochimica Acta 189. https://doi.org/10.1016/j.gca.2016.06.004.

Dang, X., Ding, W., Yang, H., Pancost, R.D., Naafs, B.D.A., Xue, J., Lin, X., Lu, J., Xie, S., 2018. Different temperature dependence of the bacterial brGDGT isomers in 35 Chinese lake sediments compared to that in soils. Organic Geochemistry 119. https://doi.org/10.1016/j.orggeochem.2018.02.008.

De Jonge, C., Hopmans, E.C., Zell, C.I., Kim, J.H., Schouten, S., Sinninghe Damsté, J.S., 2014a. Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate reconstruction. Geochimica et Cosmochimica Acta 141. https://doi.org/10.1016/j.gca.2014.06.013.

De Jonge, C., Stadnitskaia, A., Hopmans, E.C., Cherkashov, G., Fedotov, A., Sinninghe Damsté, J.S., 2014b. In situ produced branched glycerol dialkyl glycerol tetraethers in suspended particulate matter from the Yenisei River, Eastern Siberia. Geochimica et Cosmochimica Acta 125. https://doi.org/10.1016/j.gca.2013.10.031.

De Jonge, C., Stadnitskaia, A., Hopmans, E.C., Cherkashov, G., Fedotov, A., Streletskaya, I.D., Vasiliev, A.A., Sinninghe Damsté, J.S., 2015. Drastic changes in the distribution of branched tetraether lipids in suspended matter and sediments from the Yenisei River and Kara Sea (Siberia): Implications for the use of brGDGT-based proxies in coastal marine sediments. Geochimica et Cosmochimica Acta 165. https://doi.org/10.1016/j.gca.2015.05.044.

Dearing Crampton-Flood, E., Tierney, J.E., Peterse, F., Kirkels, F.M.S.A., Sinninghe Damsté, J.S., 2020. BayMBT: A Bayesian calibration model for branched glycerol dialkyl glycerol tetraethers in soils and peats. Geochimica et Cosmochimica Acta 268. https://doi.org/10.1016/j.gca.2019.09.043.

Ding, S., Schwab, V.F., Ueberschaar, N., Roth, V.N., Lange, M., Xu, Y., Gleixner, G., Pohnert, G., 2016. Identification of novel 7-methyl and cyclopentanyl branched glycerol dialkyl glycerol tetraethers in lake sediments. Organic Geochemistry 102. https://doi.org/10.1016/j.orggeochem.2016.09.009.

Duan, Y., Sun, Q., Werne, J.P., Yang, H., Jia, J., Wang, L., Xie, H., Chen, F., 2020. Soil pH Dominates the Distributions of Both 5- and 6-Methyl Branched Tetraethers in Arid Regions. Journal of Geophysical Research: Biogeosciences 125. https://doi.org/10.1029/2019JG005356.

Fawcett, P.J., Werne, J.P., Anderson, R.S., Heikoop, J.M., Brown, E.T., Berke, M.A., Smith, S.J., Goff, F., Donohoo-Hurley, L., Cisneros-Dozal, L.M., Schouten, S., Damsté, J.S.S., Huang, Y., Toney, J., Fessenden, J., Woldegabriel, G., Atudorei, V., Geissman, J.W., Allen, C.D., 2011. Extended megadroughts in the southwestern United States during Pleistocene interglacials. Nature 470. https://doi.org/10.1038/nature09839.

Foster, L.C., Pearson, E.J., Juggins, S., Hodgson, D.A., Saunders, K.M., Verleyen, E., Roberts, S.J., 2016. Development of a regional glycerol dialkyl glycerol tetraether (GDGT)-temperature calibration for Antarctic and sub-Antarctic lakes. Earth and Planetary Science Letters 433. https://doi.org/10.1016/j.epsl.2015.11.018.

Halamka, T.A., Raberg, J.H., McFarlin, J.M., Younkin, A.D., Mulligan, C., Liu, X.L., Kopf, S.H., 2023. Production of diverse brGDGTs by Acidobacterium Solibacter usitatus in response to temperature, pH, and O2 provides a culturing perspective on brGDGT proxies and biosynthesis. Geobiology 21. https://doi.org/10.1111/gbi.12525.

Hodell, D., Anselmetti, F., Brenner, M., Ariztegui, D., Anselmetti, F., Alfaro, G., Ariztegui, D., Blijdenstein, J., Brenner, M., Bronnimann, C., Brady, K., Bush, M., Chapron, E., Curtis, J., Endsley, E., Escobar, J., Gallup, C., Gamble, V., Gilli, A., Girardclos, S., Grzesik, D., Hodell, D., Hofmann, R., Islebe, G., Mueller, A., Noren, A., Perez, L., Thevenon, F., 2006. The Lake Petén Itzá Scientific Drilling Project. Scientific Drilling 1. https://doi.org/10.2204/iodp.sd.3.02.2006.

Hoerl, A.E., Kennard, R.W., 1970. Ridge Regression: Applications to Nonorthogonal Problems. Technometrics 12. https://doi.org/10.1080/00401706.1970.10488635.

Hopmans, E.C., Weijers, J.W.H., Schefuß, E., Herfort, L., Sinninghe Damsté, J.S., Schouten, S., 2004. A novel proxy for terrestrial organic matter in sediments based on branched and isoprenoid tetraether lipids. Earth and Planetary Science Letters 224. https://doi.org/10.1016/j.epsl.2004.05.012.

Hopmans, E.C., Schouten, S., Sinninghe Damsté, J.S., 2016. The effect of improved chromatography on GDGT-based palaeoproxies. Organic Geochemistry 93. https://doi.org/10.1016/j.orggeochem.2015.12.006.

Huguet, A., Francez, A.J., Jusselme, M.D., Fosse, C., Derenne, S., 2014. A climatic chamber experiment to test the short term effect of increasing temperature on branched GDGT distribution in Sphagnum peat. Organic Geochemistry 73. https://doi.org/10.1016/j.orggeochem.2014.05.010.

Kutterolf, S., Schindlbeck, J.C., Anselmetti, F.S., Ariztegui, D., Brenner, M., Curtis, J., Schmid, D., Hodell, D.A., Mueller, A., Pérez, L., Pérez, W., Schwalb, A., Frische, M., Wang, K.L., 2016. A 400-ka tephrochronological framework for Central America from Lake Petén Itzá (Guatemala) sediments. Quaternary Science Reviews 150. https://doi.org/10.1016/j.quascirev.2016.08.023.

Lei, Y., Yang, H., Dang, X., Zhao, S., Xie, S., 2016. Absence of a significant bias towards summer temperature in branched tetraether-based paleothermometer at two soil sites with contrasting temperature seasonality. Organic Geochemistry 94. https://doi.org/10.1016/j.orggeochem.2016.02.003.

Lindberg, K.R., Daniels, W.C., Castañeda, I.S., Brigham-Grette, J., 2022. Biomarker proxy records of Arctic climate change during the Mid-Pleistocene transition from Lake El'gygytgyn (Far East Russia). Climate of the Past 18. https://doi.org/10.5194/cp-18-559-2022.

Liu, X.L., Leider, A., Gillespie, A., Gröger, J., Versteegh, G.J.M., Hinrichs, K.U., 2010. Identification of polar lipid precursors of the ubiquitous branched GDGT orphan lipids in a peat bog in Northern Germany. Organic Geochemistry 41. https://doi.org/10.1016/j.orggeochem.2010.04.004.

Loomis, S.E., Russell, J.M., Sinninghe Damsté, J.S., 2011. Distributions of branched GDGTs in soils and lake sediments from western Uganda: Implications for a lacustrine paleothermometer. Organic Geochemistry 42. https://doi.org/10.1016/j.orggeochem.2011.06.004.

Loomis, S.E., Russell, J.M., Ladd, B., Street-Perrott, F.A., Sinninghe Damsté, J.S., 2012. Calibration and application of the branched GDGT temperature proxy on East African lake sediments. Earth and Planetary Science Letters 357–358. https://doi.org/10.1016/j.epsl.2012.09.031.

Loomis, S.E., Russell, J.M., Eggermont, H., Verschuren, D., Sinninghe Damsté, J.S., 2014a. Effects of temperature, pH and nutrient concentration on branched GDGT distributions in East African lakes: Implications for paleoenvironmental reconstruction. Organic Geochemistry 66. https://doi.org/10.1016/j.orggeochem.2013.10.012.

Loomis, S.E., Russell, J.M., Heureux, A.M., D'Andrea, W.J., Sinninghe Damsté, J.S., 2014b. Seasonal variability of branched glycerol dialkyl glycerol tetraethers (brGDGTs) in a temperate lake system. Geochimica et Cosmochimica Acta 144. https://doi.org/10.1016/j.gca.2014.08.027.

Martínez-Sosa, P., Tierney, J.E., 2019. Lacustrine brGDGT response to microcosm and mesocosm incubations. Organic Geochemistry 127. https://doi.org/10.1016/j.orggeochem.2018.10.011.

Martínez-Sosa, P., Tierney, J.E., Stefanescu, I.C., Dearing Crampton-Flood, E., Shuman, B.N., Routson, C., 2021. A global Bayesian temperature calibration for lacustrine brGDGTs. Geochimica et Cosmochimica Acta 305, 87–105. https://doi.org/10.1016/j.gca.2021.04.038.

Miller, D.R., Helen Habicht, M., Keisling, B.A., Castañeda, I.S., Bradley, R.S., 2018. A 900-year New England temperature reconstruction from in situ seasonally produced branched glycerol dialkyl glycerol tetraethers (brGDGTs). Climate of the Past 14. https://doi.org/10.5194/cp-14-1653-2018.

Naafs, B.D.A., Inglis, G.N., Zheng, Y., Amesbury, M.J., Biester, H., Bindler, R., Blewett, J., Burrows, M.A., del Castillo Torres, D., Chambers, F.M., Cohen, A.D., Evershed, R.P., Feakins, S.J., Gałka, M., Gallego-Sala, A., Gandois, L., Gray, D.M., Hatcher, P.G., Honorio Coronado, E.N., Hughes, P.D.M., Huguet, A., Könönen, M., Laggoun-Défarge, F., Lähteenoja, O., Lamentowicz, M., Marchant, R., McClymont, E., Pontevedra-Pombal, X., Ponton, C., Pourmand, A., Rizzuti, A.M., Rochefort, L., Schellekens, J., De Vleeschouwer, F., Pancost, R.D., 2017. Introducing global peat-specific temperature and pH calibrations based on brGDGT bacterial lipids. Geochimica et Cosmochimica Acta 208. https://doi.org/10.1016/j.gca.2017.01.038.

Naafs, B.D.A., Oliveira, A.S.F., Mulholland, A.J., 2021. Molecular dynamics simulations support the hypothesis that the brGDGT paleothermometer is based on homeoviscous adaptation. Geochimica et Cosmochimica Acta 312. https://doi.org/10.1016/j.gca.2021.07.034.

Ning, D., Zhang, E., Shulmeister, J., Chang, J., Sun, W., Ni, Z., 2019. Holocene mean annual air temperature (MAAT) reconstruction based on branched glycerol dialkyl glycerol tetraethers from Lake Ximenglongtan, southwestern China. Organic Geochemistry 133. https://doi.org/10.1016/j.orggeochem.2019.05.003.

Parish, M.C., Du, X., Bijaksana, S., Russell, J.M., 2023. A brGDGT-Based Reconstruction of Terrestrial Temperature From the Maritime Continent Spanning the Last Glacial Maximum. Paleoceanography and Paleoclimatology 38. https://doi.org/10.1029/2022PA004501.

Pearson, E.J., Juggins, S., Talbot, H.M., Weckström, J., Rosén, P., Ryves, D.B., Roberts, S.J., Schmidt, R., 2011. A lacustrine GDGT-temperature calibration from the Scandinavian Arctic to Antarctic: Renewed potential for the application of GDGT-paleothermometry in lakes. Geochimica et Cosmochimica Acta 75. https://doi.org/10.1016/j.gca.2011.07.042.

Pérez, L., Bugja, R., Lorenschat, J., Brenner, M., Curtis, J., Hoelzmann, P., Islebe, G., Scharf, B., Schwalb, A., 2011. Aquatic ecosystems of the Yucatán Peninsula (Mexico), Belize, and Guatemala. Hydrobiologia 661. https://doi.org/10.1007/s10750-010-0552-9.

Perry, E., Marin, L., McClain, J., Velazquez, G., 1995. Ring of Cenotes (sinkholes), northwest Yucatan, Mexico: its hydrogeologic characteristics and possible association with the Chicxulub impact crater. Geology 23. https://doi.org/10.1130/0091-7613(1995)023<0017:ROCSNY>2.3.CO;2.

Peterse, F., Kim, J.H., Schouten, S., Kristensen, D.K., Koç, N., Sinninghe Damsté, J.S., 2009. Constraints on the application of the MBT/CBT palaeothermometer at high latitude environments (Svalbard, Norway). Organic Geochemistry 40. https://doi.org/10.1016/j.orggeochem.2009.03.004.

Peterse, F., Nicol, G.W., Schouten, S., Sinninghe Damsté, J.S., 2010. Influence of soil pH on the abundance and distribution of core and intact polar lipid-derived branched GDGTs in soil. Organic Geochemistry 41. https://doi.org/10.1016/j.orggeochem.2010.07.004.

Peterse, F., Vonk, J.E., Holmes, R.M., Giosan, L., Zimov, N., Eglinton, T.I., 2014. Branched glycerol dialkyl glycerol tetraethers in Arctic lake sediments: Sources and implications for paleothermometry at high latitudes. Journal of Geophysical Research: Biogeosciences 119. https://doi.org/10.1002/2014JG002639.

Piccolroaz, S., Toffolon, M., Majone, B., 2013. A simple lumped model to convert air temperature into surface water temperature in lakes. Hydrology and Earth System Sciences 17. https://doi.org/10.5194/hess-17-3323-2013.

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. https://doi.org/10.1016/j.orggeochem.2018.12.007.
Raberg, J.H., Harning, D.J., Crump, S.E., De Wet, G., Blumm, A., Kopf, S., Geirsdóttir, Á., Miller, G.H., Sepúlveda, J., 2021. Revised fractional abundances and warm-season temperatures substantially improve brGDGT calibrations in lake sediments. Biogeosciences 18. https://doi.org/10.5194/bg-18-3579-2021.
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. https://doi.org/10.1016/j.orggeochem.2017.12.003.
Schoon, P.L., De Kluijver, A., Middelburg, J.J., Downing, J.A., Sinninghe Damsté, J.S., Schouten, S., 2013. Influence of lake water pH and alkalinity on the distribution of core and intact polar branched glycerol dialkyl glycerol tetraethers (GDGTs) in lakes. Organic Geochemistry 60. https://doi.org/10.1016/j.orggeochem.2013.04.015.
Schouten, S., Hopmans, E.C., Sinninghe Damsté, J.S., 2013. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. Organic Geochemistry. https://doi.org/10.1016/j.orggeochem.2012.09.006.
Schreiber-Gregory, D.N., 2018. Ridge Regression and multicollinearity: An in-depth review. Model Assisted Statistics and Applications 13. https://doi.org/10.3233/MAS-180446.
Sigala, I., Caballero, M., Correa-Metrio, A., Lozano-García, S., Vázquez, G., Pérez, L., Zawisza, E., 2017. Basic limnology of 30 continental waterbodies of the Transmexican Volcanic Belt across climatic and environmental gradients. Boletin de la Sociedad Geologica Mexicana 69. https://doi.org/10.18268/BSGM2017v69n2a3.
Siliakus, M.F., van der Oost, J., Kengen, S.W.M., 2017. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. Extremophiles. https://doi.org/10.1007/s00792-017-0939-x.
Sinninghe Damste, 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. https://doi.org/10.1039/b004517i.
Sinninghe Damsté, Jaap S., Rijpstra, W.I.C., Hopmans, E.C., Weijers, J.W.H., Foesel, B.U., Overmann, J., Dedysh, S.N., 2011. 13,16-Dimethyl octacosanedioic acid (iso-Diabolic Acid), a common membrane-spanning lipid of Acidobacteria subdivisions 1 and 3. Applied and Environmental Microbiology 77. doi:10.1128/AEM.00466-11.
Sinninghe Damsté, Jaap 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. doi:10.1016/j.gca.2016.04.033.
Sun, Q., Chu, G., Liu, M., Xie, M., Li, S., Ling, Y., Wang, X., Shi, L., Jia, G., Lü, H., 2011. Distributions and temperature dependence of branched glycerol dialkyl glycerol tetraethers in recent lacustrine sediments from China and Nepal. Journal of Geophysical Research: Biogeosciences 116. https://doi.org/10.1029/2010JG001365.
Sun, C.J., Zhang, C.L., Li, F.Y., Wang, H.Y., Liu, W.G., 2016. Distribution of branched glycerol dialkyl glycerol tetraethers in soils on the Northeastern Qinghai-Tibetan Plateau and possible production by nitrite-reducing bacteria. Science China Earth Sciences 59. https://doi.org/10.1007/s11430-015-0230-2.
Tierney, J.E., Russell, J.M., 2009. Distributions of branched GDGTs in a tropical lake system: Implications for lacustrine application of the MBT/CBT paleoproxy. Organic Geochemistry 40. https://doi.org/10.1016/j.orggeochem.2009.04.014.
Tierney, J.E., Russell, J.M., Eggermont, H., Hopmans, E.C., Verschuren, D., Sinninghe Damsté, J.S., 2010. Environmental controls on branched tetraether lipid distributions in tropical East African lake sediments. Geochimica et Cosmochimica Acta 74. https://doi.org/10.1016/j.gca.2010.06.002.
Valero-Garcés, B., Stockhecke, M., Lozano-García, S., Ortega, B., Caballero, M., Fawcett, P., Werne, J.P., Brown, E., Najera, S.S., Pearthree, K., McGee, D., Hodgetts, A.G.E., Martínez, R., 2021. Stratigraphy and Sedimentology of the Upper Pleistocene to Holocene Lake Chalco Drill Cores (Mexico Basin). https://doi.org/10.1007/978-3-030-66576-0_14.
Vinçon-Laugier, A., Cravo-Laureau, C., Mitteau, I., Grossi, V., 2017. Temperature-dependent alkyl glycerol ether lipid composition of mesophilic and thermophilic sulfate-reducing Bacteria. Frontiers in Microbiology 8. https://doi.org/10.3389/fmicb.2017.01532.
Wang, H.Y., Dong, H.L., Zhang, C.L., Jiang, H.C., Liu, W.G., 2016. A 12-kyr record of microbial branched and isoprenoid tetraether index in Lake Qinghai, northeastern Qinghai-Tibet Plateau: Implications for paleoclimate reconstruction. Science China Earth Sciences. https://doi.org/10.1007/s11430-015-5213-4.
Wang, H., Liu, W., He, Y., Zhou, A., Zhao, H., Liu, H., Cao, Y., Hu, J., Meng, B., Jiang, J., Kolpakova, M., Krivonogov, S., Liu, Z., 2021. Salinity-controlled isomerization of lacustrine brGDGTs impacts the associated MBT5ME' terrestrial temperature index. Geochimica et Cosmochimica Acta 305. https://doi.org/10.1016/j.gca.2021.05.004.
Weber, Y., De Jonge, C., Rijpstra, W.I.C., Hopmans, E.C., Stadnitskaia, A., Schubert, C.J., Lehmann, M.F., Sinninghe Damsté, J.S., Niemann, H., 2015. Identification and carbon isotope composition of a novel branched GDGT isomer in lake sediments: Evidence for lacustrine branched GDGT production. Geochimica et Cosmochimica Acta 154. https://doi.org/10.1016/j.gca.2015.01.032.
Weber, Y., Damsté, J.S.S., Zopfi, J., De Jonge, C., Gilli, A., Schubert, C.J., Lepori, F., Lehmann, M.F., Niemann, H., 2018. Redox-dependent niche differentiation provides evidence for multiple bacterial sources of glycerol tetraether lipids in lakes. In: Proceedings of the National Academy of Sciences of the United States of America 115. https://doi.org/10.1073/pnas.1805186115.
Weijers, J.W.H., Schouten, S., Spaargaren, O.C., Sinninghe Damsté, J.S., 2006. Occurrence and distribution of tetraether membrane lipids in soils: Implications for the use of the TEX86 proxy and the BIT index. Organic Geochemistry 37. https://doi.org/10.1016/j.orggeochem.2006.07.018.
Weijers, J.W.H., Schouten, S., van den Donker, J.C., Hopmans, E.C., Sinninghe Damsté, J.S., 2007. Environmental controls on bacterial tetraether membrane lipid distribution in soils. Geochimica et Cosmochimica Acta 71. https://doi.org/10.1016/j.gca.2006.10.003.
Weijers, J.W.H., Wiesenberg, G.L.B., Bol, R., Hopmans, E.C., Pancost, R.D., 2010. Carbon isotopic composition of branched tetraether membrane lipids in soils suggest a rapid turnover and a heterotrophic life style of their source organism(s). Biogeosciences 7. https://doi.org/10.5194/bg-7-2959-2010.
Weijers, J.W.H., Steinmann, P., Hopmans, E.C., Schouten, S., Sinninghe Damsté, J.S., 2011. Bacterial tetraether membrane lipids in peat and coal: Testing the MBT-CBT temperature proxy for climate reconstruction. Organic Geochemistry 42. https://doi.org/10.1016/j.orggeochem.2011.03.013.
Wu, J., Yang, H., Pancost, R.D., Naafs, B.D.A., Qian, S., Dang, X., Sun, H., Pei, H., Wang, R., Zhao, S., Xie, S., 2021. Variations in dissolved O2 in a Chinese lake drive changes in microbial communities and impact sedimentary GDGT distributions. Chemical Geology 579. https://doi.org/10.1016/j.chemgeo.2021.120348.
Xiao, W., Xu, Y., Ding, S., Wang, Y., Zhang, X., Yang, H., Wang, G., Hou, J., 2015. Global calibration of a novel, branched GDGT-based soil pH proxy. Organic Geochemistry 89–90. https://doi.org/10.1016/j.orggeochem.2015.10.005.
Yang, H., Lü, X., Ding, W., Lei, Y., Dang, X., Xie, S., 2015. The 6-methyl branched tetraethers significantly affect the performance of the methylation index (MBT') in soils from an altitudinal transect at Mount Shennongjia. Organic Geochemistry 82. https://doi.org/10.1016/j.orggeochem.2015.02.003.
Yao, Y., Zhao, J., Vachula, R.S., Werne, J.P., Wu, J., Song, X., Huang, Y., 2020. Correlation between the ratio of 5-methyl hexamethylated to pentamethylated branched GDGTs (HP5) and water depth reflects redox variations in stratified lakes. Organic Geochemistry 147. https://doi.org/10.1016/j.orggeochem.2020.104076.
Zell, C., Kim, J.H., Hollander, D., Lorenzoni, L., Baker, P., Silva, C.G., Nittrouer, C., Sinninghe Damsté, J.S., 2014. Sources and distributions of branched and isoprenoid tetraether lipids on the Amazon shelf and fan: Implications for the use of GDGT-based proxies in marine sediments. Geochimica et Cosmochimica Acta 139. https://doi.org/10.1016/j.gca.2014.04.038.
Zhao, B., Castañeda, I.S., Bradley, R.S., Salacup, J.M., de Wet, G.A., Daniels, W.C., Schneider, T., 2021. Development of an in situ branched GDGT calibration in Lake 578, southern Greenland. Organic Geochemistry 152. https://doi.org/10.1016/j.orggeochem.2020.104168.