SCIENCE CHINA
Earth Sciences

SPECIAL TOPIC: **Advances in organic proxies for research in climate and environmental changes**
• RESEARCH PAPER •

July 2016 Vol.59 No.7: 1353–1367
doi: 10.1007/s11430-016-5268-x

Branched tetraether lipids in Chinese soils: Evaluating the fidelity of MBT/CBT proxies as paleoenvironmental proxies

ZHENG FengFeng[1], ZHANG ChuanLun[1*], CHEN YuFei[1], LI FuYan[1], MA CenLing[1], PU Yang[2], ZHU YuanQing[1,3], WANG YongLi[4] & LIU WeiGuo[5]

[1] *State Key Laboratory of Marine Geology, Tongji University, Shanghai 200092, China;*
[2] *School of Geography and Remote Sensing, Nanjing University of Information Science & Technology, Nanjing 210044, China;*
[3] *Earthquake Administration of Shanghai Municipality, Shanghai 200062, China;*
[4] *Key Laboratory of Gas Geochemistry, Institute of Geology and Geophysics, Chinese Academy of Sciences, Lanzhou 730000, China;*
[5] *State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 710075, China*

Received September 6, 2015; accepted December 31, 2015; published online March 8, 2016

Abstract The ubiquitous occurrence of branched glycerol dialkyl glycerol tetraethers (brGDGTs) in soils has allowed development of new proxies for reconstruction of past climate and environment. The methylation and cyclization degrees of brGDGTs, expressed as MBT and CBT, respectively, are reported to be mainly controlled by mean annual air temperature (MAAT) and soil pH. However, the brGDGT-derived temperatures and soil pH scatter widely when data from different environmental conditions are considered. In this study, we collected over 300 soil samples from China, which are representative of humid (Xishuangbanna, Guangzhou, and Shanghai), semi-arid (Dongying) and semi-arid/arid (Lanzhou, Tibetan Plateau) regions. Collectively we have the most extensive dataset that broadly characterizes the distribution of brGDGTs according to climate zones in China. The overall data demonstrate that the MBT/CBT derived temperatures better match the measured MAATs in humid and non-alkaline regions than those from regions of low MAP (<400 mm/yr) and above neutral soil pH (>7.0–7.5). Similarly, CBT describes soil pH much better in humid and non-alkaline soils than in semi-arid/arid and alkaline soils; the semi-arid/arid and alkaline soils tend to show a positive correlation between soil pH and CBT, which contradicts that in the humid and non-alkaline soils. While soil pH, MAAT and mean annual precipitation (MAP) are dominating factors controlling the brGDGT distribution across all climate zones, conductivity, total organic carbon and total nitrogen, as well as soil water content can also play an important role locally. Removing brGDGT-II resulted in a revised CBT index that provides more accurate estimation of pH, especially in semi-arid/arid and alkaline soils. The overall Chinese dataset demonstrates that continental air temperature derived from brGDGT-proxies can vastly deviate from real measurements and should be used with extreme caution in paleo-climate or -environment studies.

Keywords brGDGTs, MAAT, Soil pH, MAP, Chinese soils, Tibetan Plateau, Chinese Loess Plateau

Citation: Zheng F F, Zhang C L, Chen Y F, Li F Y, Ma C L, Pu Y, Zhu Y Q, Wang Y L, Liu W G. 2016. Branched tetraether lipids in Chinese soils: Evaluating the fidelity of MBT/CBT proxies as paleoenvironmental proxies. Science China Earth Sciences, 59: 1353−1367, doi: 10.1007/s11430-016-5268-x

1. Introduction

Development of liquid chromatography-mass spectrometry technology allowed the separation and quantification of branched glycerol dialkyl glycerol tetraethers (brGDGTs; Figure 1) that are presumably produced by bacteria (Weijers et al., 2006). Studies of brGDGTs in a number of soil environments demonstrated that the distribution of brGDGTs can be described by the methylation index of branched tet-

*Corresponding author (email: archaeazhang_1@tongji.edu.cn)

raethers (MBT) and the cyclisation ratio of branched tetraethers (CBT), which are mainly governed by mean annual air temperature (MAAT) and/or soil pH, respectively (Weijers et al., 2007a). This has aided research in paleoclimate or global change by providing new proxies for paleocontinental temperatures or soil pH. For example, MBT/CBT-estimated MAAT and CBT-estimated soil pH witnessed a coupled thermal and hydrological evolution of tropical Africa over the last deglaciation (Weijers et al., 2007b) and helped explain the extended megadroughts in the southwestern United States during Pleistocene interglacials (Fawcett et al., 2011). Similarly, enhanced aridity associated with significant uplift of the Tibetan Plateau in the Late Miocene was evidenced by CBT index and the ratio between total archaeal GDGTs and brGDGTs ($R_{i/b}$) (Xie et al., 2012).

However, the fidelity of MBT/CBT-derived proxies in paleo records has suffered from the fact that the calculated temperature and soil pH scatter widely in modern soil environments. Mounting evidence shows that soil water content (SWC) or mean annual precipitation (MAP) are additional factors that may strongly impact the distribution of brGDGTs in arid and semiarid environments (Peterse et al., 2012; Dirghangi et al., 2013; Wang et al., 2014). Further unknown are the sources of brGDGTs because no definitive microorganisms (presumably bacteria) have been identified, which produce all brGDGTs observed from the natural environments. These and other uncertainties have led to multiple revisions of the original MBT/CBT-MAAT calibration, with each revision trying to explain the mechanisms causing the calibration error. For example, Peterse et al. (2012) re-calibrated the previous MBT-CBT proxies using approximately 280 samples globally and claimed that large scatter in the calibration cannot be fully explained by local factors or by seasonality. Anderson et al. (2014) tried to determine the sources of calibration error by using *in situ* monitoring of soil temperature along an elevation gradient across the Eastern Cordillera of Colombia and concluded that local and transient variation in factors such as soil moisture and nutrient availability causes the scattering of the calculated temperatures.

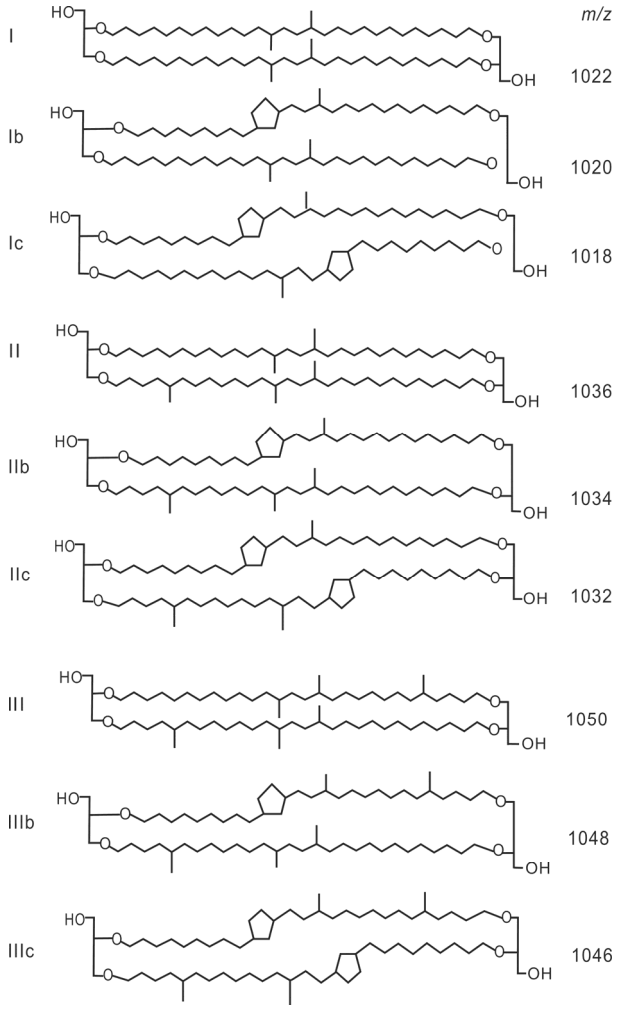

Figure 1 The molecular structures and protonated masses of brGDGTs identified from Chinese soil samples.

The vast land in China not only has distinct climate zones from south to north but also is characterized by the high Tibetan Plateau and the Chinese loess regions that are unique to the East Asia climate zones. The occurrence and distribution of brGDGTs in some of these environments have been reported by a rapid growing number of publications in recent years (e.g., Wang et al., 2012; Zhang et al., 2012; Liu et al., 2013; Yang et al., 2014; Ding et al., 2015). In particular, Yang et al. (2014) proposed new calibration models for Chinese soils using over 100 samples collected from different climatic regions, particularly the semiarid and arid loess regions. Ding et al. (2015) focused on the cold and dry regions of the Tibetan Plateau and recalibrated the MBT/CBT-MAT relationship that is constrained by a root mean square error (RMSE) of 4.2°C. Both studies suggested that local calibrations are necessary for estimating MAAT or soil pH in semiarid and arid regions of China. This highlights the need for caution when interpreting MBT/CBT proxy in loess-paleosol sequence studies and begs for a thorough evaluation of its application in diverse Chinese soil environments, particularly the semi-arid and arid soils on the Tibetan Plateau and the Chinese Loess Plateau.

We thus have performed an extensive sampling (over 300 samples) from different regions of China, which covered a wide range of temperature, pH, vegetation, and soil water content. The objectives of this study were to demonstrate the role of locality or geography on the distribution of brGDGTs and validate the fidelity of brGDGT-based proxies using the most comprehensive dataset from the Chinese soils.

2. Material and methods

2.1 Sample collection and measurements of basic soil properties

Soil samples were collected between 2011 and 2013 from six regions in China (Xishuangbanna, Guangzhou, Shanghai, Dongying, Lanzhou and Tibetan Plateau) (Figure 2). Detailed information for each sample is provided in Appendix Table 1 (available at http://earth.scichina.com). At each site, the upper 5 cm soil was collected for each sample, which was homogenized after removing leaves and plants or plant roots. Samples were stored on dry ice in the field and at −20°C in the lab until further treatment.

Soil pH was measured following Chu et al. (2010) with the following modification. The frozen sample was first thawed at room temperature, and then about 4 g of it was mixed with ultra-pure water at the ratio of 1∶2.5 (w/v, g/mL). The mixture was shaken rigorously for 30 min and centrifuged at 3000 r/min for 10 min. The supernatant was measured by a pH meter (METTLER TOLEDO) and pH was recorded as the average of three measurements. The standard deviation for triplicate measurements was ±0.05 pH units.

Soil conductivity was measured in filtered soil extract by using a portable conductivity/salinity meter (METTLER TOLEDO). Briefly, ca. 5 g of thawed fresh soil was mixed with ultra-pure water at the ratio of 1∶5 (w/v, g/mL) and rigorously shaken for 30 min. The mixture was centrifuged at 20000 g for 10 min and the analytical error for duplicate measurement was <0.01 ms/cm for most soils and <0.10 ms/cm for saline soils (>1.00 ms/cm).

Soil water content (SWC) was determined by the difference in weight before and after drying a portion of the soil sample at 105°C for 24 h (Nocita et al., 2013; Guntiñas et al., 2012).

Total organic carbon (TOC) and total nitrogen (TN) were determined following a modified procedure of Zhang et al. (2012). Briefly, an aliquot of freeze-dried soil sample was treated with 10% HCl to remove bicarbonate. The residues were subsequently washed with distilled water until neutral

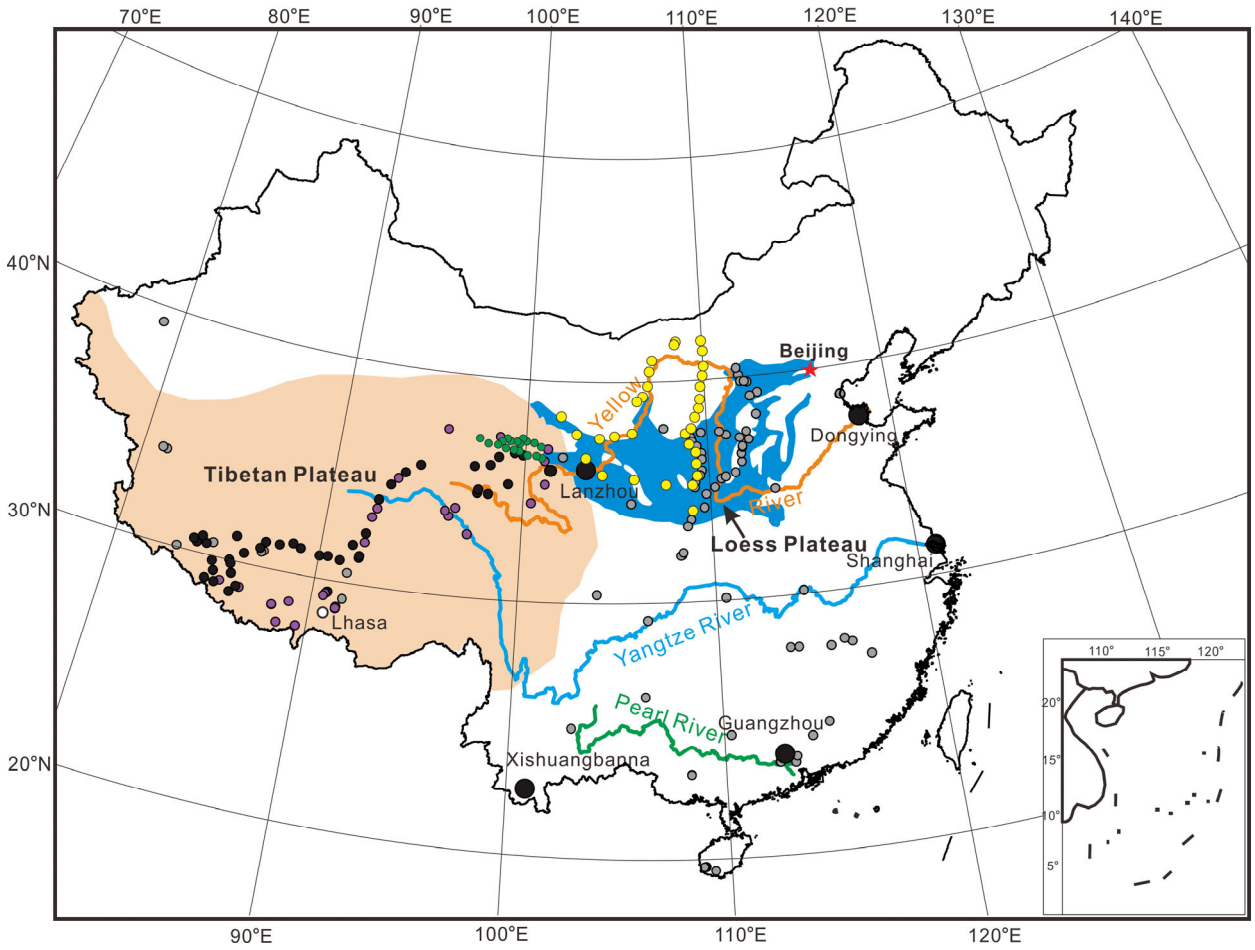

Figure 2 Map of China showing the soil sampling sites. Black solid circles are the sampling sites in this study, with the larger-sized circles representing multiple samples (≥30 samples) collected at each site. Grey solid circles represent samples from Yang et al. (2014) and references therein. Purple solid circles represent samples from Ding et al. (2015). Yellow solid circles are samples from Wang et al. (2014) and Green ones are samples from Sun et al. (2016).

pH and dried at 45°C. Approximately 10 mg residues were used to measure TOC and TN by a Vario Cube Elemental Analyzer. Duplicate measurements were conducted for every 11 samples and the standard deviation was 0.08% for TOC and 0.006% for TN.

2.2 Data collection for MAAT and MAP

The Chinese climate dataset for MAAT and MAP was obtained between 1981 and 2010 from the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/). Subsequently, the MAAT and MAP values from climate stations were interpolated to the whole land area of China using the Kriging method with the ArcGIS software (Version 10.2). The MAAT and MAP values of each sampling site were then extracted using Spatial Analyst Tools in the ArcGIS software. Following Ding et al. (2015), we also extracted MAAT and MAP for sites from previous studies of Günther et al. (2014) and Yang et al. (2014). Overall, 127 out of the total 131 sites had a strong correlation with the reported data (R^2=0.94 for MAAT and R^2=0.95 for MAP), suggesting that our method was reliable to obtain actual climate data for exploring the relationship between environmental variables and distribution of brGDGTs in Chinese soils.

2.3 Analysis of GDGTs

For lipid extraction, 5–25 g freeze-dried and powered soils were extracted (5×) using an accelerated solvent extractor (ASE350, Dionex) with a mixture solvent of dichloromethane (DCM): methanol (MeOH) (9∶1, *v*/*v*) at 100°C for 5 min. The total extracts were dried under a slow N_2 gas flow and then separated into apolar and polar fractions over an activated SiO_2 column by Hexane: DCM (9:1, *v*/*v*) and MeOH: DCM (1∶1, *v*/*v*), respectively. A synthesized C_{46} GDGT was added as an internal standard. The polar fraction containing GDGTs were re-dissolved in hexane/isopropanol (99∶1, *v*/*v*) and filtered through 0.45 μm PTFE filters. Samples were dried again under N_2 gas prior to analysis.

Analysis and quantification of GDGTs were achieved using Agilent 1200 series High Performance liquid chromatograph, coupled to an atmospheric pressure chemical ionization/triple quadrupole mass spectrometer (HPLC/APCI-QQQ-MS), equipped with an automatic injector and chemstation manager software. The analytical procedure followed Schouten et al. (2007) that was modified by Zhang et al. (2012). Briefly, 10-μL polar fractions were injected and peak separation was accomplished by using an Alltech Prevail Cyano column (150 mm×2.1 mm, 3 μm). The lipids were eluted isocratically with 90% A and 10% B (A for *n*-hexane, B for *n*-hexane/isopropanol 9∶1, *v*/*v*) in the first 5 min, followed by a linear gradient to 18% B in 45 min. The column was flushed with 100% B for 10 min and then re-equilibrated with 90% A and 10% B before the next run. The MS condition followed Schouten et al. (2007). A single ion monitoring mode was selected to detect $[M+H]^+$ (m/z=1050, 1048, 1046, 1036, 1034, 1032, 1022, 1020, 1018) corresponding to brGDGT structure in Figure 1. Quantification was achieved by comparing peak area of $[M+H]^+$ and internal standard.

The methylation and cyclization degrees of brGDGTs are defined following Weijers et al. (2007a):

$$\mathrm{MBT}=(\mathrm{I}+\mathrm{Ib}+\mathrm{Ic})/(\mathrm{I}+\mathrm{Ib}+\mathrm{Ic}+\mathrm{II}+\mathrm{IIb}+\mathrm{IIc}+\mathrm{III}+\mathrm{IIIb}+\mathrm{IIIc}), \quad (1)$$

$$\mathrm{CBT}=-\mathrm{Log}(\mathrm{Ib}+\mathrm{IIb}/(\mathrm{I}+\mathrm{II})). \quad (2)$$

A revised methylation index (MBT') is from Peterse et al. (2012), which includes only seven brGDGTs:

$$\mathrm{MBT'}=(\mathrm{I}+\mathrm{Ib}+\mathrm{Ic})/(\mathrm{I}+\mathrm{Ib}+\mathrm{Ic}+\mathrm{II}+\mathrm{IIb}+\mathrm{IIc}+\mathrm{III}). \quad (3)$$

Roman numbers refer to the structure of brGDGTs (Figure 1).

2.4 Statistical analysis

Redundancy analysis (RDA) was performed to determine the relationship between the fractional abundance of brGDGTs and environmental variables. Partial RDA was performed to assess the contribution of unique and independent explanatory power of these environmental variables to the distribution of brGDGTs at a single site. The RDA and partial RDA were performed using CANOCO software (windows version 4.5) following Tierney et al. (2010) and Yang et al. (2014). Correlation coefficients and *P*-values of brGDGTs and environmental variables were obtained by analysis of bivariate correlation using SPSS software (Version 21). The linear and multiple regressions of brGDGT indices and fractional abundance of brGDGTs were also calibrated using SPSS software.

3. Results

3.1 Physicochemical properties of Chinese soils

A total of 311 soil samples were collected from six regions in China, with each region having 30–64 individual samples collected at different locations (Table 1, Figure 2). Based on the range of MAP (Guo et al., 2007), Tibetan Plateau and Lanzhou were grouped as semi-arid (200–500 mm/yr) or arid (MAP<200 mm/yr) regions, Dongying as semi-arid region and the other places as humid regions (MAP>800 mm/yr) (Table 1).

The MAAT was above 20°C in Xishuangbanan (20.2–21.9°C) and Guangzhou (21.0–22.1°C) in the humid regions; in the semi-arid or arid regions, MAAT ranged from 15.4 to 15.5°C in Shanghai, 13.1 to 13.3°C in Dongy-

Table 1 Summary of sample information that divides the regions into semi-humid (MAP>800 mm/yr), semi-arid (200–800 mm/yr), and arid (<200 mm/yr) regions according to Guo et al. (2007)[a)]

	Site	#Sa	MAP (mm/yr)	MAAT (°C)	pH ranges	SWC (%)	Conductivity (ms/cm)	TOC (%)	TN (%)
Humid sites	Xishuangbanna	30	1178.10–1362.35	20.17–21.90	4.58–8.23	–	–	0.18–10.38	0.04–0.50
	Guangzhou	64	1496.03–1588.95	21.02–22.07	4.63–7.81	1.58–31.45	0.02–0.24	0.05–2.40	0.04–0.23
	Shanghai	63	920.39–935.5	15.43–15.50	7.12–8.52	4.49–69.46	0.07–0.58	0.50–10.24	0.07–0.44
Semi-arid sites	Dongying	48	426.14–435.59	13.07–13.29	7.45–9.39	3.15–50.16	0.11–17.00	0.09–2.89	0.04–0.26
Semi-arid/arid sites	Lanzhou	40	167.27–301.65	8.77–9.56	7.66–9.13	1.80 – 30.82	0.05–3.46	0.09–2.82	0.03–0.18
	Tibetan Plateau	66	112.14–384.34	−4.87–6.62	6.15–9.03	1.18–47.68	0.02–3.34	0.22–4.97	0.05–0.56

a) "Sa" = samples. "-" = not available.

ing, 8.8 to 9.6°C in Lanzhou (the Chinese Loess Plateau) and −4.9 to 6.6°C on the Tibetan Plateau (Table 1).

Soil pH varied from below 5.0 to above 7.0 in Xishuangbanan (4.6–8.2) and Guangzhou (4.6–7.8). The other regions were characterized by circumneutral (6.2–7.5) to strongly alkaline (reaching 9.4) soil pH (Table 1). SWC ranged from less than 5% to nearly 70%, with no clear distinction between the semi-arid/arid and the humid/semi-humid regions. The soil conductivity varied widely; the highest value (17.0 ms/cm) occurred in Dongying that was characterized by saline soils, which was followed by Tibetan Plateau and Lanzhou with the higher end values around 3.5 ms/cm. The highest soil conductivity in Shanghai and Guangzhou was less than 0.6 ms/cm. Conductivity measurement was not performed for Xishuangbanna samples because they were not properly preserved.

Total organic carbon varied between 0.05% and 10.38% with samples from Shanghai and Xishuangbanna having the highest values. Total nitrogen values were all less than 1.0% (Table 1).

3.2 Abundance and distribution of brGDGTs in Chinese soils

brGDGTs were detected in all soil samples ranging from less than 1.0 ng/g dwt soil to >2000.0 ng/g dwt soil (Appendix Table 2). The higher concentrations occurred at the pH range from 5.0 to 8.5 and relatively lower concentrations at pH>8.5 and pH<5.0 (Figure 3a). When specified to individual sites, samples from the humid regions (Xishuangbanna, Guangzhou, Shanghai) had the same range of variation between the 5 and 95 percentiles. However, Xishuangbanna had a relatively lower median value (44.1 ng/g dwt soil) of brGDGTs than that of Guangzhou (74.6 ng/g dwt soil) or Shanghai (72.1 ng/g dwt soil) (Figure 3b). The 5 and 95 percentile range of brGDGT concentrations was narrow and similar among the semi-arid and arid regions (Dongying, Lanzhou, Tibetan Plateau) with the median value below 11.0 ng/g dwt soil in all three regions (Figure 3b). These results demonstrate that brGDGT production is generally lower in semi-arid and arid regions than in relatively humid regions in China.

The nine brGDGT compounds (Figure 1) were present in the majority of the Chinese soils; however, brGDGTs-IIIb and -IIIc were below detection limit in some samples (Appendix Table 2), which mostly occurred in humid soils. In general, the relative abundance of brGDGT-I decreased from close to 70% of total brGDGTs in humid region (Xishuangbanan) to less than 10% in the arid region (Tibetan Plateau); in the opposite, brGDGT-II (12–38%) and brGDGT-III (1–32%) increased from humid to arid regions (Figure 4). The sum of brGDGT-Ib and brGDGT-Ic (5–27%) also increased from Xishuangbanna to Shanghai in the humid regions, but decreased from Dongying to Lanzhou and Tibetan Plateau in the semi-arid or arid regions; the sum of brGDGT-IIb and brGDGT-IIc varied similarly in the humid regions (5–20%), but remained relatively constant (15–20%) in the semi-arid and arid regions (Figure 4). In all regions, the sum of brGDGT-IIIb and brGDGT- IIIc did not exceed 3% (Figure 4).

Redundancy analysis (RDA) using relative abundance of brGDGTs also separated the samples according to regions with samples from Xishuangbanna and Guangzhou grouping together; whereas, samples from other regions had a distinct pattern for each region (Figure 5a). Furthermore, RDA identified three major variables controlling the relative abundances of brGDGTs among all samples. brGDGT-I had significant positive correlations with MAP (R^2=0.71, P=0.000) and MAAT (R^2 = 0.59, P=0.000) and significant negative correlation with soil pH (absolute value of R^2 = 0.71, P=0.000). Each of brGDGT-II, -III and -IIIb showed significant negative correlation with MAP (R^2>0.43, P< 0.001) and MAAT (R^2>0.31, P<0.001) and significant positive correlation with soil pH (R^2>0.25, P<0.001). On the other hand, less significant correlations existed between relative abundances of cyclic brGDGTs and the environmental variables (Table 2).

The relative abundances of brGDGTs-I, -Ib, and -Ic aligned in the same direction as MAP and MAAT and the relative abundances of other compounds aligned in the same direction as pH on the RDA plot (Figure 5a). Similarly, Pearson correlation analysis revealed significant positive correlations between these compounds (excluding brGDGT-IIc and brGDGT-IIIc) and soil pH and negative correlations with MAP and MAAT (Table 2). The axis 1 in

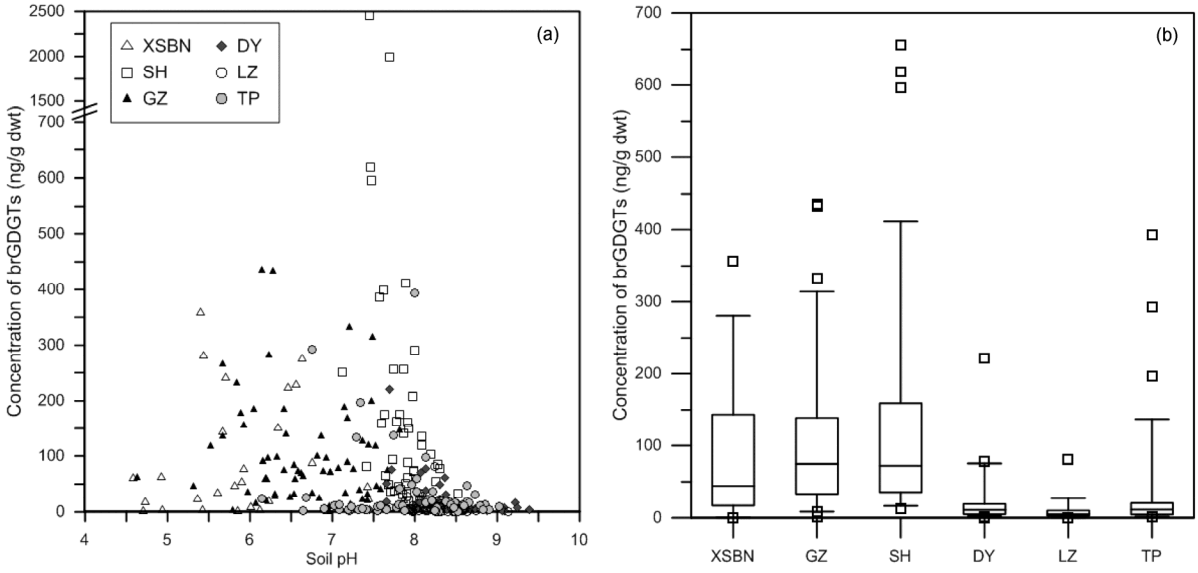

Figure 3 Variation of brGDGT concentration with increasing soil pH (a). "Box and whiskers plot" of brGDGT concentrations with each box representing the interquartile range (IQR) of the sample (b). Lines extending from below and above the box represent 5 and 95 percentiles of the data, respectively. Data not included within these percentiles are plotted as outliers. Samples SH-19 and SH-62 had exceptionally high concentration of brGDGTs (1991.20 and 2449.95 ng/g dwt, respectively) and was not shown in the "Box and whiskers plot" diagram.

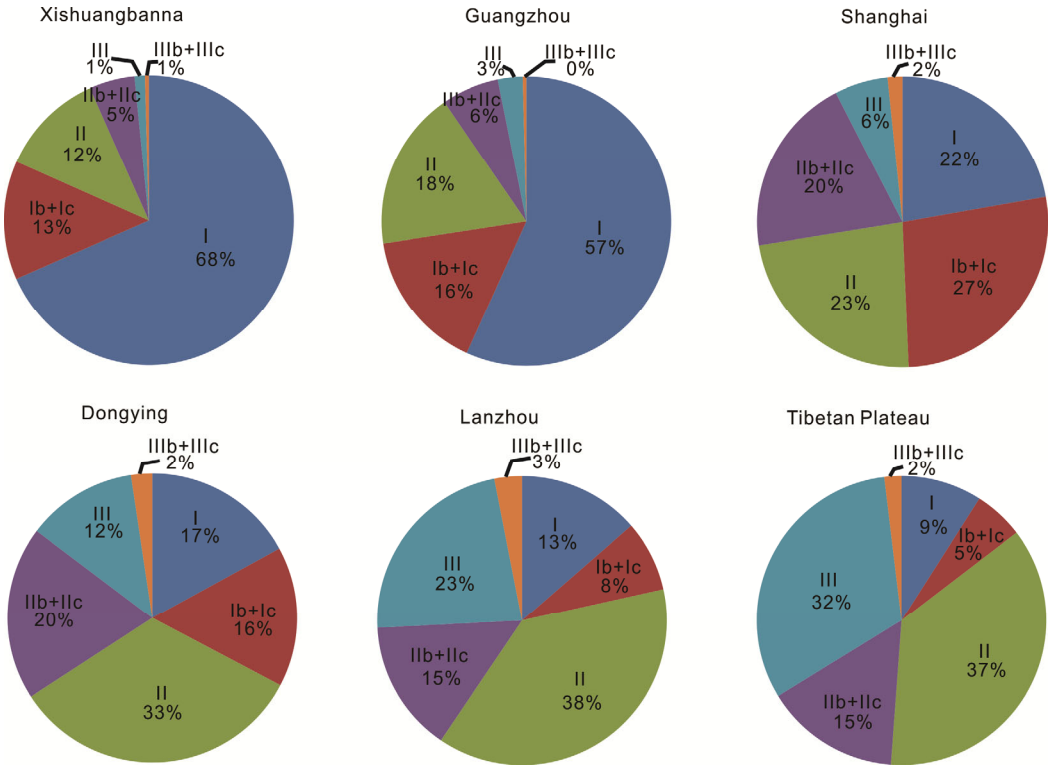

Figure 4 Distribution of fractional abundance of brGDGTs (averaged from all samples from each region) in Chinese soils.

the RDA biplot (Figure 5a and b) primarily captured the gradient in MAAT and MAP. Therefore, we concluded that the difference in brGDGTs among different regions was primarily related to MAAT and MAP, whereas the difference in brGDGTs in a given region was to some extent caused by soil pH.

When restricted to individual regions, samples in Xishuangbannan, Guangzhou, Shanghai and Dongying had invariant MAAT and MAP (Appendix Table 1), which thus exerted no impact on the relative abundances of brGDGTs on the RDA plot (Appendix Figure 1a–d). Among these regions, soil pH was the single variable that showed significant negative impact on brGDGT distribution in Xishuangbanna and Guangzhou, but not in Shanghai or Dongying

Table 2 Correlation coefficients (R) and P values for fractional abundances and proxies of brGDGTs and environmental variables, Pearson correlation >0.5 and P<0.05 were shown in bold[a)]

		Correlations											brGDGTs (ng/g dwt)
		I	Ib	Ic	II	IIb	IIc	III	IIIb	IIIc	MBT	CBT	
MAAT	Pearson correlation	**0.767**	0.447	0.46	**–0.721**	–0.424	0.047	**–0.862**	**–0.562**	0.006	**0.883**	0.236	0.182
	P-value	0.000	0.000	0.000	0.000	0.000	0.412	0.000	0.000	0.919	0.000	0.000	0.001
MAP	Pearson Correlation	**0.841**	0.295	0.317	**–0.78**	**–0.564**	–0.144	**–0.724**	**–0.657**	-0.169	**0.892**	0.355	0.210
	P-value	0.000	0.000	0.000	0.000	0.000	0.012	0.000	0.000	0.003	0.000	0.000	0.000
Soil pH	Pearson Correlation	**-0.841**	0.12	–0.013	**0.625**	**0.668**	0.172	**0.508**	**0.647**	0.105	**–0.74**	**–0.688**	–0.139
	P-value	0.000	0.039	0.820	0.000	0.000	0.003	0.000	0.000	0.071	0.000	0.000	0.016
SWC	Pearson Correlation	0.161	0.157	0.081	–0.187	–0.064	0.007	–0.178	–0.026	0.040	0.195	–0.018	0.495
	P-value	0.009	0.010	0.188	0.002	0.299	0.906	0.003	0.678	0.517	0.001	0.773	0.000
Conductivity	Pearson Correlation	–0.094	0.031	0.147	0.119	0.009	0.099	–0.002	–0.061	–0.015	–0.055	–0.055	–0.084
	P-value	0.134	0.617	0.018	0.057	0.882	0.113	0.970	0.327	0.816	0.377	0.378	0.178
TN	Pearson Correlation	0.030	0.075	–0.072	–0.084	–0.017	–0.234	0.023	–0.007	–0.206	0.045	0.010	0.413
	P-value	0.604	0.195	0.215	0.145	0.763	0.000	0.691	0.902	0.000	0.434	0.864	0.000
TOC	Pearson Correlation	0.074	0.083	–0.068	–0.086	–0.052	–0.214	–0.040	–0.081	–0.175	0.088	0.036	0.263
	P-value	0.202	0.151	0.242	0.134	0.370	0.000	0.488	0.159	0.002	0.126	0.537	0.000

a) Pearson correlation >0.5 and P<0.05 are shown in bold.

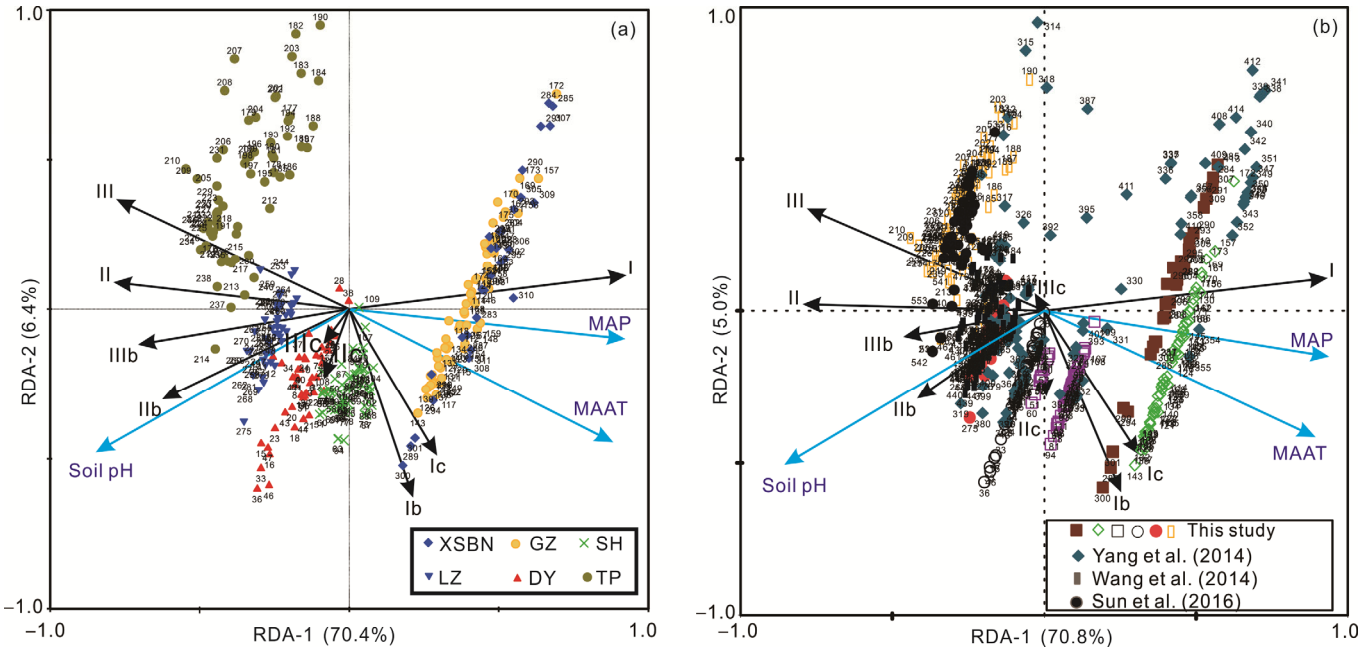

Figure 5 Redundancy analysis (RDA) showing the distribution of brGDGTs from different regions of China and their relationships with the three major environmental variables (MAP, MAAT and soil pH) in this study (a) and the combination of this study and others (b).

(Tables 3 and 4). In Dongying, conductivity had significant positive correlation with brGDGT-I (R^2=0.30, P<0.001) whereas TOC and TN had significant positive correlation with brGDGT-II (R^2=0.64 and 0.51, respectively and P< 0.001) (Appendix Figure 1d). The relationship between environmental variables and brGDGTs was weak (R^2<0.25) in Shanghai and in Lanzhou (Table 3). Despite the large land space range in Tibetan Plateau, fractions of brGDGTs and environmental variables only had poor correlations (Appendix Figure 1f and Table 3).

The impact of soil pH on the distribution of brGDGTs was further analyzed based on the division of humid vs.

Table 3 Results of RDA showing the relationship between environmental variables and fractional abundances of brGDGTs in all sites and individual sites[a)]

Environmental variables	All sites		XSBN		GZ		SH		DY		LZ		TP	
	Axis1	Axis2	Axis1	Axis2	Axis1	Axis2	Axis1	Axis2	Axis1	Axis2	Axis1	Axis2	Axis1	Axis2
MAAT	**0.84**	–0.29	–	–	–	–	–	–	–	–	0.46	–0.14	–0.06	0.39
MAP	**0.88**	–0.07	–	–	–	–	–	–	–	–	0.30	0.05	0.17	–0.08
Soil pH	**–0.81**	–0.31	**0.81**	0.13	**–0.82**	–0.03	0.07	–0.11	–0.02	–0.19	0.31	0.06	–0.46	–0.20
SWC	–0.12	–0.09	–	–	–0.25	0.14	0.41	–0.08	–0.07	–0.33	0.05	0.47	–0.29	–0.08
Conductivity	–0.12	–0.07	–	–	–0.2	0.3	0.35	–0.09	–0.12	**0.58**	0.15	–0.16	–0.37	0.20
TN	0.08	–0.04	–0.21	0.45	–0.11	0.12	0.39	0.01	**0.79**	0.01	–0.10	–0.13	–0.11	–0.10
TOC	0.11	–0.03	–0.14	0.32	–0.05	0.05	0.4	0.16	**0.71**	0.03	–0.03	–0.09	–0.14	–0.03
Cum.% variance, brGDGTs	70.4	76.4	59.8	60.9	62.9	63.8	22.8	24.3	29	42.5	28.3	32.3	25.1	32.2
Cum.% variance, brGDGTs-env	90.9	99.1	97.7	99.5	98.5	99.8	88.7	94.6	64.8	95.1	79.9	94.7	75.1	96.3

a) "–", data not available. XSBN, Xishuangbannal; GZ, Guangzhou; SH, Shanghai; DY, Dongying; LZ, Lanzhou; TP, Tibetan Plateau. *P*-value<0.05 are shown in bold.

Table 4 Results of partial RDA test of samples from Xishuangbanna (XSBN), Guangzhou (GZ), Shanghai (SH) and Dongying (DY) [a)]

Environmental variables	Covariable(s)	XSBN		GZ		SH		DY	
		% Variance	*p*-value	% Variance	*p*-value	% Variance	*p*-value	% Variance	*p*-value
Conductivity	None	–	–	4.30%	0.098	**9.10%**	0.008	**11.70%**	0.006
	Others	–	–	0.70%	0.292	**5.20%**	0.029	**9.80%**	0.004
TN	None	4.50%	0.242	1.20%	0.368	**11.50%**	0.002	**28.60%**	0.002
	Others	2.60%	0.206	1.80%	0.09	1.70%	0.264	**6.40%**	0.008
TOC	None	2.30%	0.404	0.30%	0.71	**12.10%**	0.004	**23.10%**	0.002
	Others	1.50%	0.34	0.10%	0.798	2.70%	0.131	0.50%	0.841
Soil pH	None	**57.30%**	0.002	**60.70%**	0.002	0.60%	0.73	1.90%	0.41
	Others	**53.30%**	0.002	**55%**	0.002	1.60%	0.258	1.10%	0.428
SWC	None	–	–	5.70%	0.06	**12.80%**	0.001	4.40%	0.11
	Others	–	–	0.10%	0.858	**6.80%**	0.018	2.80%	0.074

a) 'None' represents RDA performed with given variable as the sole constraining variable and no co-variables included, and 'others' indicate RDA performed with the given variable as the sole constraining variable and all other variables treated as co-variables. *P*-values<0.05 are shown in bold. "–", data not available.

non-humid regions. In the humid regions (e.g., Xishuangbanna, Guangzhou, and Shanghai), brGDGT-I correlated significantly with soil pH; whereas, all other brGDGTs showed positive correlations (Appendix Figure 2). In non-humid regions (e.g., Dongying, Lanzhou, and Tibetan Plateau), no significant correlation was observed between soil pH and any of the brGDGTs (Appendix Figure 2).

3.3 Regional and local variability in brGDGT proxies

When samples from all regions were considered, MBT index displayed an increasing trend with MAAT when it was >5°C and with MAP, and decreased with soil pH. However, the range of MBT can be large at a given MAAT-, MAP- or soil pH value (Figure 6a–c). In fact, MBT spanned a large range for both MAAT and MAP in each region (Appendix Table 2; Figure 6a, b); it varied continuously with soil pH and became much more scattered when soil pH was >7.5, except for the samples from the Tibetan Plateau that seemed to have a different slope of variation between MBT and soil pH (Figure 6c).

CBT, on the other hand, showed no consistent pattern with change in MAAT or MAP; in both cases, the peak values of CBT seemed to decrease with MAAT or MAP from Tibetan Plateau to Shanghai; whereas, a large range of CBT occurred in a narrow range of MAAT or MAP in Guangzhou or Xishuangbanna (Figure 6d, e). The CBT scattered much less at soil pH<7.0 than >7.0 (Figure 6(f)). The largest variability of CBT index occurred in soils from Xishuangbanna and Guangzhou, which had large ranges of soil pH (4.58–8.23 and 4.63–7.82, respectively; Appendix Table 2). Furthermore, CBT values from more alkaline soils (Tibetan Plateau, Lanzhou and Dongying) were higher than those in Shanghai (Appendix Table 2).

4. Discussions

4.1 Factors controlling the abundance and distribution of brGDGTs

Soil pH was observed to be an important factor controlling the abundance of brGDGTs in soils (Weijers et al., 2007a; Peterse et al., 2010). Because bacteria potentially producing brGDGTs are thought to be heterotrophs, the abundance of organic carbon is also considered to control the abundance

of brGDGTs (Weijers et al., 2010). This is consistent with some of the observations in China; for example, total brGDGTs correlated significantly with soil pH in the Tibetan Plateau soils (Ding et al., 2015) and with organic carbon or water content in alkaline soils surrounding Lake Qinghai (Wang et al., 2013). In our own samples, no significant correlations (R^2<0.25) were observed between total brGDGTs and environmental variables (soil pH, MAAT, MAP, soil water content, total organic carbon, etc.), suggesting that the absolute concentration of brGDGTs is controlled by multiple variables. In general, however, regional variation in brGDGT abundance appeared to be controlled by climate zones, with higher abundance occurring mostly in humid regions (Figure 3b).

In terms of the relative abundance of brGDGTs, although all nine brGDGT compounds can be detected in our samples, two of them (brGDGTs-IIIb and -IIIc) are often missing or below 1% in some of the samples. This is consistent with previous findings as summarized in Peterse et al. (2012), Yang et al. (2014) and Ding et al. (2015). While our data agree with Peterse et al. (2012) that non-cyclic brGDGTs predominate in the soil, we also observe that relative abundance of brGDGTs-II and -III co-vary and increase when brGDGT-I decreases from humid region to semiarid or arid regions (Figure 4a, b).

4.2 Variables causing proxy uncertainties

Variables potentially impacting proxies of brGDGTs have been examined and summarized in detail in Peterse et al. (2012) for global soils, which were further substantiated by Anderson et al. (2014) for the montane regions, and Yang et al. (2014) and Ding et al. (2015) for Chinese Loess Plateau and Tibetan Plateau, respectively. Collectively, a large number of variables have been identified, which can be broadly grouped as instrumental/analytical, environmental, and biological.

Zech et al. (2012) highlighted that poor separation of the HPLC chromatograms resulted in clumping of the different compounds as one brGDGT structure and a cleaning up procedure improved the separation of the compounds that caused a 2.7°C difference from the uncleaned procedure. Recently, with the use of ultra-high performance liquid chromatography (UHPLC) or improved liquid chromatography method, two isomers of brGDGTs with 6-methyl isomer were separated from 5-methyl brGDGTs in peat and soils (De Jonge et al., 2013, 2014; Yang et al., 2015), which may revise the MBT-MAAT or CBT-pH relationship. This effort, unfortunately, has not been practiced popularly in most labs and should be advocated in future studies.

MAP was already observed to correlate with MBT in

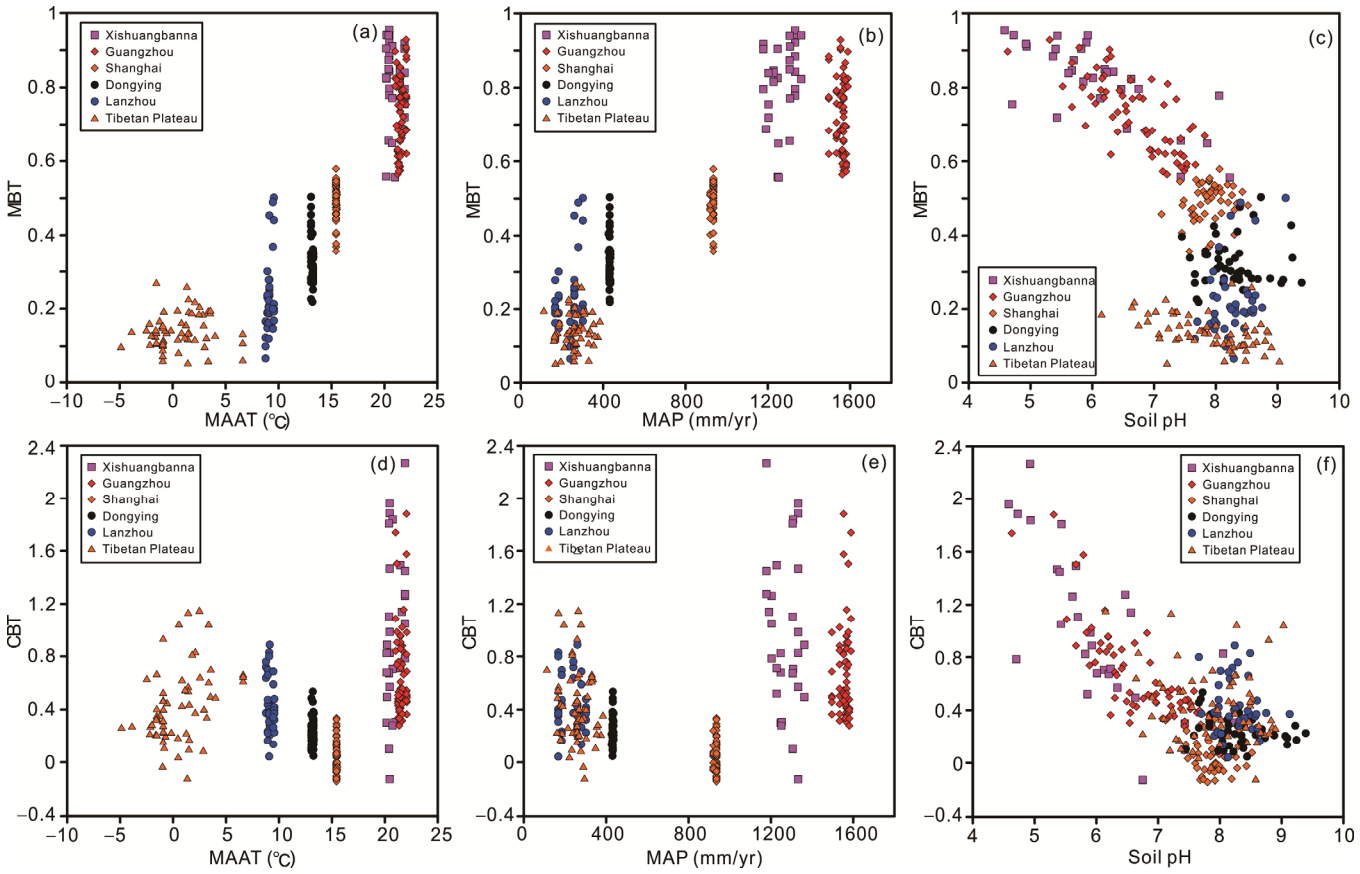

Figure 6 Change of MBT and CBT with MAAT, MAP and soil pH based on observation of all samples from this study.

Weijers et al. (2007a), which is substantiated by the composite dataset in this study (Figure 7b); The correlation between MAP and MBT has also been considered as co-variation between MAAT and MAP and Weijers et al. (2007a) argued that the impact of MAP on MBT lacks a physiological mechanism because increased precipitation is unlikely to cause changes in lipid membrane composition. Meanwhile, MAP or soil water content also has been observed to show a good correlation with CBT in semi-arid regions (Wang et al., 2014), which can be explained by the mechanism that brGDGT-producing bacteria need to maintain denser packing of membrane lipids in order to keep intracellular water from over-evaporation. This can be accomplished by formation of cyclopentyl moieties (Weijers et al., 2007a; Wang et al., 2014). However, the correlation between MAP and CBT is not significant for the composite dataset, which shows large variation at MAP>1200 mm/yr or <500 mm/yr (Figure 7e).

Seasonal response of brGDGTs to temperature has been considered as one variable causing MBT/CBT proxy variation (Shanahan et al., 2013, Peterse et al., 2014), which means that brGDGT producers grow in relatively warmer season and hibernating in extremely cold environments, like Tibetan Plateau or polar region where a long frost period lasts and only a short window of warm season exists for biological production. However, seasonal bias of brGDGT derived temperature is not observed in mid latitude soils (Weijers et al., 2011).

Difference between MAAT and soil temperature also has been proposed as one of the source errors for MBT/CBT variation (Weijers et al., 2007a; Peterse et al., 2012). However, Anderson et al. (2014) observed that the difference of the RMSE between *in situ* temperature and the linearly interpolated temperature is minor (only 0.29°C). Instead, the mismatch between *in situ* soil temperature and MAATs from nearby weather station records is probably the source of scatter in the regional and global calibrations.

Other environmental variables such as change in redox, nutrients, or vegetation cover also have been suggested to affect the MBT/CBT proxies (Peterse et al., 2009b; Huguet et al., 2010, 2012; Loomis et al., 2011; Zech et al., 2012). In particular, human activities on agricultural soil management (e.g., flooding practices or fertilizer enrichment) could also have effect on MBT/CBT proxies. For example, Mueller-Niggemann et al. (2015) observed that paddy soils might have lower MBT values and calculated MAATs than adjacent upland soils. This phenomenon also existed in our samples from Shanghai. The paddy and wetland soils typically had lower MBT values and higher CBT values; On the other hand, Farmland had the higher MBT values and lower

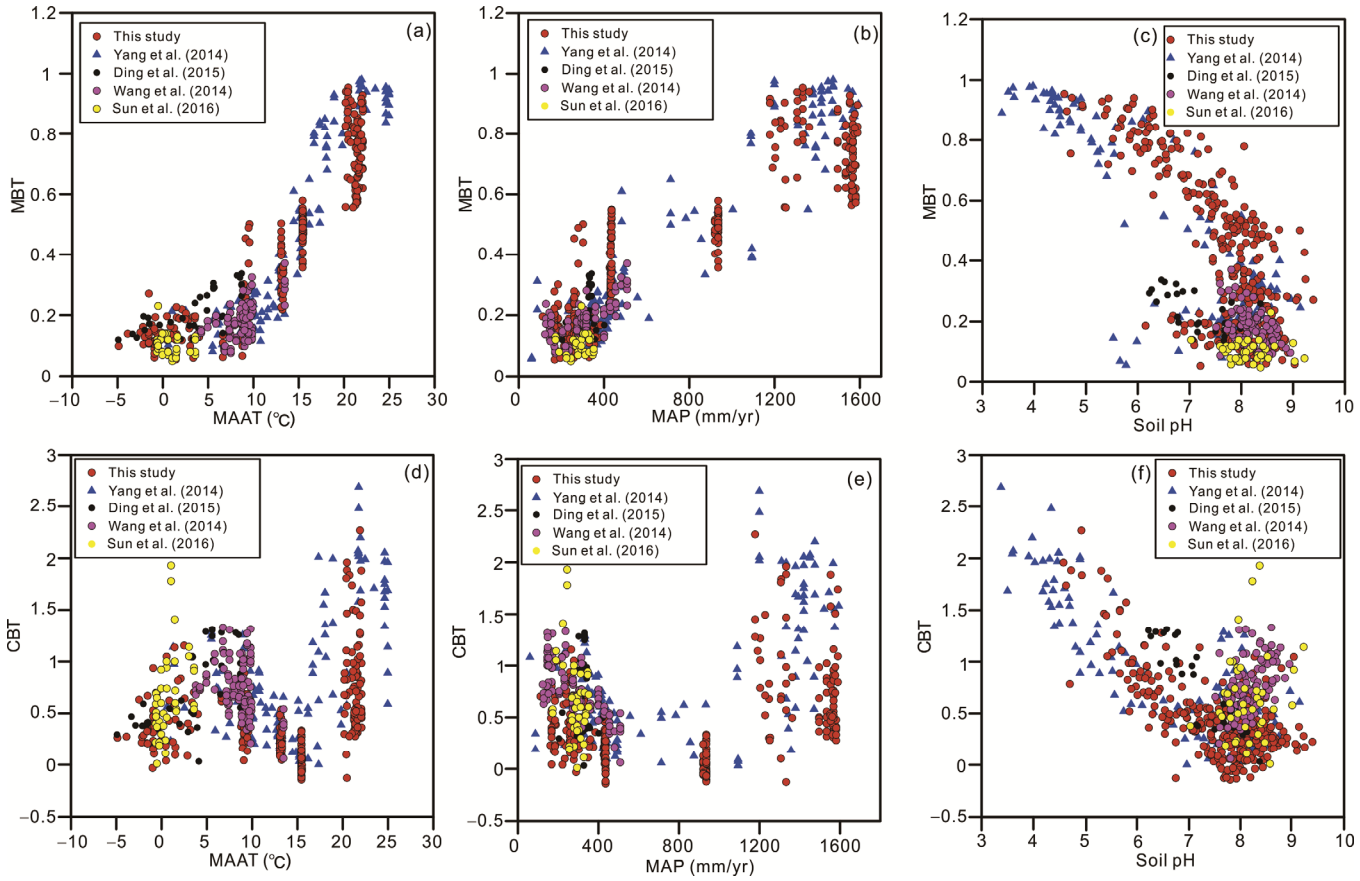

Figure 7 Change of MBT and CBT with MAAT, MAP and soil pH based on observations from this and other studies in China.

CBT values. However, this was not the case in soils from Dongying, where saline wetland soils had the highest MBT values (Appendix Figure 3). This indicated that the impact of other environmental variables on MBT/CBT proxies could be complex and alternative controls might exist.

Furthermore, Bacteria that produce brGDGTs as membrane lipids in alkaline soils may be different from those in acidic soils because bacterial community structure can shift from acid soils to alkaline soils at the phylum and sub-phylum levels. For example, *Acidobacteria* (subgroup 1) dominate in low pH soils, whereas *Proteobacteria* commonly dominate in more alkaline soils (Rousk et al., 2010). Sun et al.[1)] observed that nitrite-reducing bacteria belonging to the *Proteobacteria* may be responsible for the production of brGDGTs in the alkaline soils on the Tibetan Plateau. Other *Proteobacteria* and *Bacteroidetes* of thermophilic origin were also proposed to be the possible biological sources of brGDGTs in hot springs (Zhang et al., 2013; Li et al., 2014).

Most studies, however, have focused on soil pH or MAAT and other environmental variables for the cause of variation in MBT/CBT proxies. The impact of MAAT and soil pH has been extensively discussed (Peterse et al., 2012; Yang et al., 2014; Ding et al., 2015, among others) and this study only summaries data from Chinese soils (Figure 7). Collectively, the new data from this study cover the whole range of variation in both MBT and CBT that have been previously reported (Wang et al., 2014; Yang et al., 2014; Ding et al., 2015; Sun et al.[1)]). For MBT, the composite dataset highlights the lack of change in MBT with temperature when it is less than 5°C (Figure 7a) and the variation in MBT is more significant when pH increases (Figure 7c). CBT shows large variation with temperature when it is above 15 °C; between −5 and 15°C, which shows a big hump with a peak value near 5°C (Figure 7d). The pattern of variation between CBT and soil pH did not change between the new dataset and the composite dataset (Figures 6f and 7f).

Although the collective data from this study and previous ones (Yang et al., 2014; Ding et al., 2015) identify MAP, soil pH and MAAT as major controlling factors affecting the MBT/CBT proxy in Chinese soils (Figure 5b), MAP or soil pH could be the reason for causing the calculated MAAT to be deviated from measured or recorded MAAT over a particular region. For example, MBT/CBT-derived MAAT is considerably overestimated in the Tibetan Plateau; whereas it tended to be underestimated in Guangzhou and Xishuangbanna as soil pH increased (Figure 8a). This apparently also corresponds to low MAP in the Tibetan Plateau and high MAP in Guangzhou and Xishuangbanna (Figure 8b). The intertwined effect of soil pH and MAP on MAAT estimates is difficult to delineate, which may be further impacted by other environmental variables such as SWC, TOC/TN, and conductivity (Appendix Figure 1).

4.3 Approaches improving the fidelity of MBT/CBT proxies

Great efforts were made to reduce the large errors in MBT/CBT-derived MAATs since Weijers et al. (2007a) (Table 5). Peterse et al. (2012) revised the MBT/CBT-MAAT calibration of Weijers et al. (2007a) using an extended data set of 278 globally distributed soils. These samples are from North-America, France, Chile, the Netherlands, Egypt and Uganda, as well as samples that have already been published (Bendle et al., 2010, the Amazon Fan; Peterse et al., 2009b, Gongga Mountain; Peterse et al., 2009a, high latitude environments). However, their efforts do not improve the estimates of soil pH and MAAT based on the MBT' and CBT indices (Table 5). They also tried the least squares multiple linear regression analysis, which assigns weighing factors to the fractional abundance of the individual brGDGTs to obtain the equation with the highest coefficient of determination (Peterse et al., 2012). Although the method improves the calibration in lake studies (Tierney et al., 2010; Pearson et al., 2011), it also only marginally improves the revised MBT' and CBT indices in soils (Peterse et al., 2012).

The regional calibration seems to be more effective in reducing the calibration error (Table 5). For example, Bendle et al. (2010) better constrained their data in estimating the MAAT variation in the past 37 kyr BP in the Amazon Fan using a regional calibration (Table 5) instead of the global calibration of Weijers et al. (2007a). Anderson et al. (2014) reduced the RMSE from 5°C of the global calibration to 3°C in a local calibration in Eastern Cordillera of Colombia. Yang et al. (2014) further reduced the RMSE to less than 2°C in a local calibration of the Chinese Loess Plateau. Other studies also advocated local calibration in paleo-climate studies (Damste et al., 2008; Dirghangi et al., 2013; Ding et al., 2015). Regional calibration also seemed to improve the RMSE for soil pH estimation on the Tibetan Plateau (Ding et al., 2015), which are 0.3 pH units rather than 0.7–0.9 pH units in the global calibration (Table 5) (Peterse et al., 2012).

In this study, the low (<0.3) MBT values occur in soils from very cold and dry regions (MAAT ranges from −4.87 to 6.62°C and MAP ranges from 112.14 to 384.34 mm/yr) and exhibits little variability with MAATs <5°C, similar to observation by Yang et al. (2014) who conclude that the MBT index cannot distinguish change in MAAT in arid and cold environment. The impact of precipitation on MBT/CBT proxy has also been observed in other studies (Dirghangi et al., 2013; Anderson et al., 2014; Wang et al., 2014). This indicates that MAP may be an important variable that constrains the applicability of MBT index in global soils. We thus evaluate the global dataset of Peterse et al. (2012) according to the range of MAP that is largely representative of the semi-arid/arid (MAP 0–500 mm/yr), semi-humid (500–800 mm/yr) and humid (>800 mm/yr) regions;

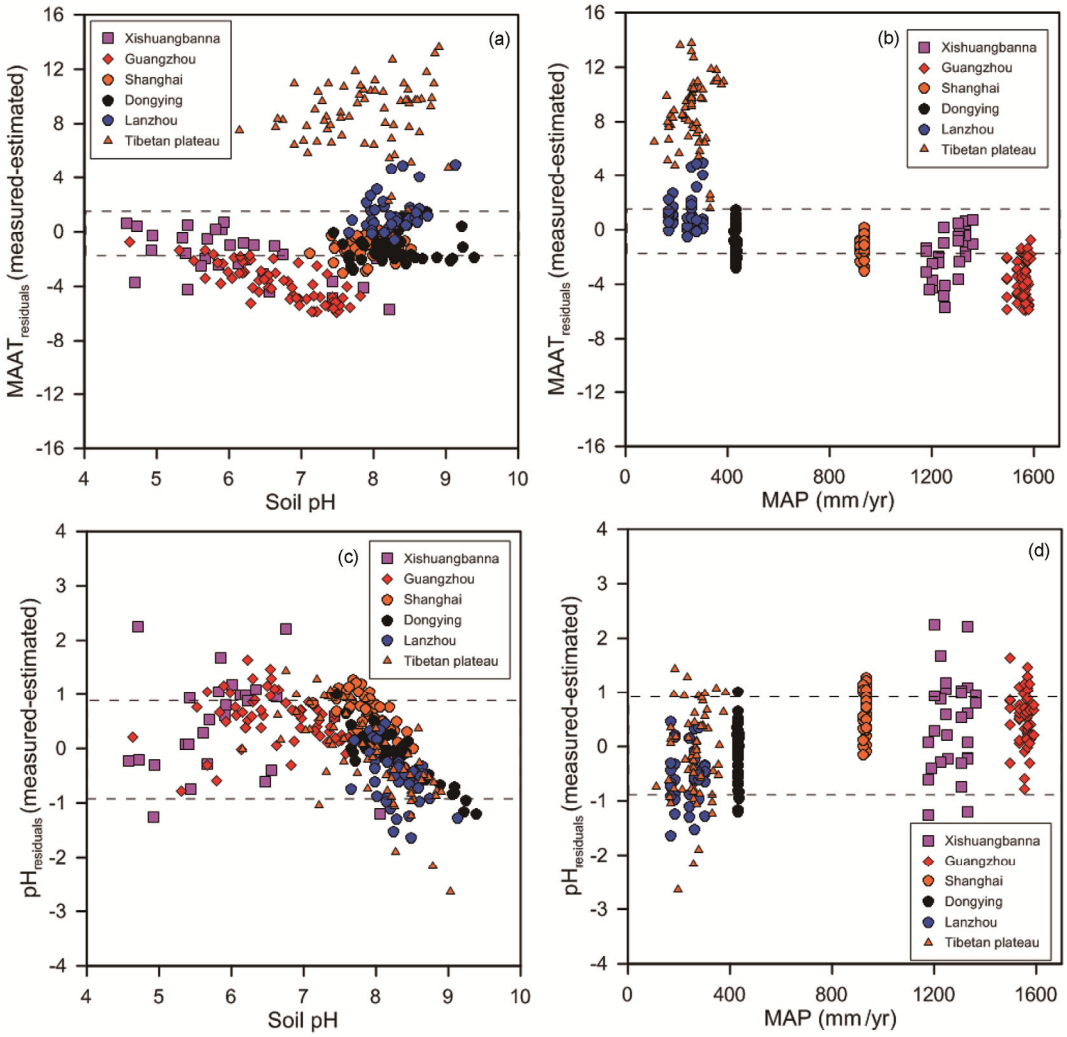

Figure 8 Variations in MAAT and pH residuals (estimated–measured) plotted against soil pH and MAP using the Chinese calibration (Yang et al., 2014). The upmost and bottom dashed lines represent the RMSE of Chinese calibration from Yang et al. (2014).

Table 5 Summary of global and regional calibrations of brGDGTs proxies and MAAT or soil pH[a)]

	Equations	RMSE	Locations	References
Global	MBT=0.122+0.187×CBT+0.020×MAAT (n=114, R^2=0.77)	4.8°C	Global	Weijers et al., 2007
	MAT=–0.64+22.9×MBT' (n=176, R^2=0.47)	5.7°C	Global	Peterse et al., 2012
	MAT=0.81–5.67×CBT+31.0×MBT' (n=176, R^2=0.59)	5.0°C	Global	Peterse et al., 2012
Regional	MBT=0.187+0.083×CBT+0.025+MAAT (R^2=0.91)	–	Amazon	Bendle et al., 2010
	MBT=0.093+0.21×CBT+0.025×MAAT (n=16, R^2=0.82)	–	Mt. Kilimanjaro	Damste et al., 2008
	MAAT=1.2+22.3×MBT'+1.5×CBT (n=31, R^2=0.69)	3.1°C	Colombia	Anderson et al., 2014
	MAAT=29.1–0.017×I–0.61×Log(Ib)–3.34×Log(Ic)–0.34×II-0.11×Log(IIb)+ 0.44×Log(IIc)–0.067×III (n=31, R^2=0.77)	2.9°C	Colombia	Anderson et al., 2014
	MAAT=10.7×MBT'+5.5 (n=67, R^2=0.60)	–	USA	Dirghangi et al., 2013
	MBT=0.069 ×MAAT–0.5 (n=130, R^2=0.85)	1.9°C	Chinese	Yang et al., 2014
	MAAT=7.5+16.1×MBT–1.2×CBT (n=126, R^2=0.86)	1.8°C	Chinese	Yang et al., 2014
	MAAT=20.9–13.4×II–17.2×III–17.5×IIb+11.2×Ib (n=120, R^2=0.87)	1.7°C	Chinese	Yang et al., 2014
Global	CBT=3.33–0.38×pH (n=114, R^2=0.70)	0.7 pH units	Global	Weijers et al., 2007
	pH=7.90–1.97×CBT (n=176, R^2=0.70)	0.8 pH units	Global	Peterse et al., 2012
	pH=8.49–0.043×I +0.013×Ib+0.019×Ic–0.037×II+0.045×IIb–0.18×IIc +0.02×III–1.97×CBT (n=176, R^2=0.72)	0.7 pH units	Global	Peterse et al., 2012
	pH=8.13–1.89×CBT (n=557, R^2=0.64)	0.9 pH units	Global	Yang et al., 2014
Regional	CBT=4.23–0.58×pH (R^2=0.75)	–	Amazon	Bendle et al., 2010
	pH=8.68–2.21×CBT (n=124, R^2=0.70)	0.9 pH units	Chinese	Yang et al., 2014
	pH=8.33–1.43×CBT (R^2=0.80)	0.3 pH units	QTP of China	Ding et al., 2015

a) "–"= not available.

the results showed that the MBT-MAAT relationship is significantly improved (R^2=0.69 vs. 0.39) in the humid region (Figure 9a). We also proposed a linear regression model between MAAT and MBT/CBT indices for humid region:

$$\text{MAAT}=-4.07+28.94\times\text{MBT}-0.42\times\text{CBT}$$
$$(R^2 = 0.70,\ n = 139,\ \text{RMSE} = 4.7°\text{C}). \quad (4)$$

This suggests that the global calibration is perhaps most reliable in regions where MAP is >800 mm/yr. Under other climate conditions, however, a local calibration may be necessary for estimating MAAT.

The dataset of Chinese soils has shown that CBT index appears to be unable to distinguish pH variations in arid and alkaline soils in China and CBT values positively correlate or flatten off with soil pH (Xie et al., 2012; Wang et al., 2014; Yang et al., 2014). As a result, this change in CBT-pH relationship will hamper the use of CBT as a paleo-pH proxy in alkaline and arid soils (Xie et al., 2012; Yang et al., 2014). The large scatters in arid and alkaline soils are ascribed to MAAT (Yang et al. 2014) or soil moisture/MAP (Wang et al., 2014), which serve as dominant environmental variables affecting CBT index instead of soil pH in these soils. However, we notice that the response of individual brGDGTs to soil pH can be different in humid and no-humid soils. Thus, the impact of individual brGDGT on the CBT index is examined by removing either of them from it. The results showed that the relationship is improved significantly when brGDGT-II is removed (Appendix Figure 4). Therefore, we define rCBT to describe the relationship of soil pH and brGDGT distribution:

$$\text{rCBT}=-\text{Log}((\text{Ib}+\text{IIb})/\text{I}). \quad (5)$$

Incorporated with previously published dataset of Chinese soils (Günther et al., 2014; Wang et al., 2014; Sun et al.[1)]; Yang et al., 2014 and references therein), the relationship between rCBT and soil pH is summarized by the following equation:

$$\text{pH}=7.75-1.58\times\text{rCBT}$$
$$(R^2 = 0.67,\ n=539,\ P<0.001,\ \text{RMSE}=0.64). \quad (6)$$

The relationship of CBT-pH was significantly improved (R^2=0.67 vs. 0.38) by removing brGDGT-II and the revised CBT index was linearly fitting with soil pH in all soil pH ranges. The coefficient determination can further improve if we use dataset of Yang et al. (2014):

$$\text{pH}=7.74-1.82\times\text{rCBT}$$
$$(R^2=0.83,\ n=108,\ P<0.001,\ \text{RMSE}=0.70). \quad (7)$$

Overall, this revised CBT index can describe the pH change for all pH values, which confirms that soil pH affects the degree of cyclisation of brGDGTs widely in soils.

Another method to express the relationship of soil pH and distribution of brGDGTs is the multiple regressions of fractional abundance of brGDGTs and soil pH, which has been used to obtain more accurate estimation of MAAT or soil pH (Tierney et al., 2010; Pearson et al., 2011, 2012; Yang et al., 2014). We applied this method using the available Chinese dataset and obtained the following equation:

$$\text{pH}=9.94-2.21\times\text{III}-5.86\times\text{IIIb}-20.33\times\text{IIIc}-1.35\times\text{II}-1.79\times\text{IIb}-4.25\times\text{IIc}-5.54\times\text{I}-0.13\times\text{Ib}-3.54\times\text{Ic}$$
$$(R^2=0.77,\ n=404,\ P<0.001,\ \text{RMSE}=0.59). \quad (8)$$

The new fit shows similar determination coefficients (0.77 vs. 0.72) and RMSE (0.59 vs. 0.70) with global cali-

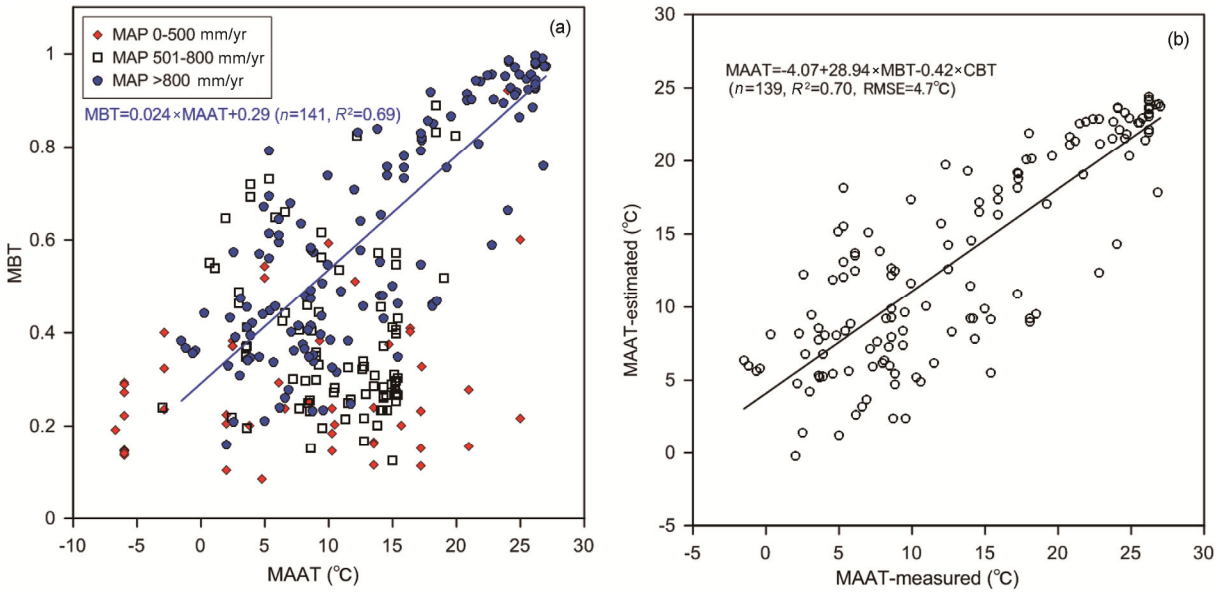

Figure 9 The relationship between MBT and MAAT under different MAP ranges in global dataset of Peterse et al. (2012) (a) and recalibration of the MBT/CBT-MAAT correlation using dataset with MAP>800 mm/yr in Peterse et al. (2012) (b).

brations of Peterse et al. (2012), suggesting that multiple regressions are promising in obtaining the highest coefficient determination for the estimation of soil pH. On the other hand, for humid regions, we propose an improved fitting model including the relative abundances of 5 compounds (III, II, IIb, I, Ib) of brGDGTs that are strongly correlated (R^2 >0.5) with soil pH (IIIb was excluded because it was absent in >20% of humid soils) (Appendix Figure 2). They are log-transformed (except for brGDGT-I) so that they would vary linearly with soil pH. Compounds below detection limit are linearly interpolated before being log-transformed.

$$\text{pH}=9.16-0.45\times\text{Log(III)}+1.00\times\text{Log(II)}+0.61\times\text{Log(IIb)}-2.13\times\text{I}+0.52\times\text{Log(Ib)}\ (R^2=0.88,\ n=189,\ P<0.001,\ \text{RMSE}=0.45\ \text{pH units}). \quad (9)$$

5. Conclusions

The MBT/CBT proxy based on the ubiquitous presence of branched glycerol dialkyl glycerol tetraethers is now widely used in estimating paleo-continental air temperature and soil pH; its application, however, has been hampered by the large errors in global calibrations. The extensive dataset from China allows us to draw the following conclusions regarding the distribution of brGDGTs and the fidelity of MBT/CBT proxy in Chinese soils.

(1) The abundance and distribution of brGDGTs are characterized by climate zones in China with high total brGDGTs occurring in humid regions; brGDGTs-I, -Ib and -Ic also dominate in more humid regions; whereas, brGDGTs-II, -IIb, -IIc, and -III are predominant in semi-arid and arid regions.

(2) While MAAT, soil pH and MAT are the controlling factors affecting the distribution of brGDGTs over all the regions studied, soil water content, conductivity, total organic carbon can play important roles in a particular region.

(3) The composite Chinese dataset from this study and previous publications results in a best fit of revised CBT with soil pH in Chinese soils: pH=7.75−1.58×rCBT (R^2 = 0.67, n=539, P<0.001, RMSE=0.64). In addition, a fitting model of soil pH and fractional abundance of brGDGT was established for humid regions in China:

pH=9.16−0.45×Log(III)+1.00×Log(II)+0.61×Log(IIb)−2.13×I+0.52×Log(Ib) (R^2=0.88, n=189, P<0.001, RMSE= 0.45 pH units).

(4) The global temperature calibration of Peterse et al. (2012) is best applicable in humid regions where MAP is >800 mm/yr:

MAAT=−4.07+28.94×MBT−0.42×CBT (R^2 = 0.70, n= 139, RMSE=4.7°C).

This study and others also demonstrate that regional calibrations of MAAT outperform global calibrations by about 2.8°C reduction of the RMSE (from about 5.1°C to 2.3°C, in average; Table 5), which indicates that development or selection of a regional calibration may be necessary for paleoclimate studies in a particular environment.

Acknowledgements *We would like to thank two anonymous reviewers for their comments, which improved the original manuscript. Yingqin Wu, Bangqi Hu and Joanna Zhang are thanked for assistance in field sampling. We greatly appreciate the comments and advice from Dr. Tommy Phelps during the preparation of the manuscript. This study was supported by the National Natural Science Foundation of China (Grant Nos. 41373072 and 40873011), Shanghai Bureau of Science and Technology (Grant No. 13JC1405200) and the National Thousand Talents Program through the State Key Laboratory of Marine Geology at Tongji University.*

References

Anderson V J, Shanahan T M, Saylor J E, Horton B K, Mora A R. 2014. Sources of local and regional variability in the MBT'/CBT paleotemperature proxy: Insights from a modern elevation transect across the Eastern Cordillera of Colombia. Org Geochem, 69: 42–51

Bendle J A, Weijers J W H, Maslin M A, Sinninghe Damsté J S, Schouten S, Hopmans E C, Boot C S, Pancost R D. 2010. Major changes in glacial and Holocene terrestrial temperatures and sources of organic carbon recorded in the Amazon fan by tetraether lipids. Geochem Geophys Geosyst, 11: Q12007

Chu H, Grogan P. 2010. Soil microbial biomass, nutrient availability and nitrogen mineralization potential among vegetation-types in a low arctic tundra landscape. Plant Soil, 329: 411–420

Damste J S S, Ossebaar J, Schouten S, Verschuren D. 2008. Altitudinal shifts in the branched tetraether lipid distribution in soil from Mt. Kilimanjaro (Tanzania): Implications for the MBT/CBT continental palaeothermometer. Org Geochem, 39: 1072–1076

Ding S, Xu Y, Wang Y, He Y, Hou J, Chen L, He J S. 2015. Distributions of glycerol dialkyl glycerol tetraethers in surface soils of Qinghai-Tibetan Plateau: Implications of GDGT-based proxies in cold and dry regions. Biogeosci Discuss, 12: 481–513

Dirghangi S S, Pagani M, Hren M T, Tipple B J. 2013. Distribution of glycerol dialkyl glycerol tetraethers in soils from two environmental transects in the USA. Org Geochem, 59: 49–60

De Jonge C, Hopmans E C, Stadnitskaia A, Rijpstra W I C, Hofland R, Tegelaar E, Sinninghe Damsté J S. 2013. Identification of novel penta- and hexa-methylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC-MS^2, GC-MS and GC-SMB-MS. Org Geochem, 54: 78–82

De Jonge C, Hopmans E C, Zell C I, Kim J-H, Schouten S, Sinninghe Damsté J S. 2014. Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate reconstruction. Geochim Cosmochim Acta, 141: 97–112

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, Sinninghe Damste J 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: 518–521

Guo Z H, Liu X M, Xiao W F, Wang J L, Meng C. 2007. Regionalization and integrated assessment of climate resource in China based on GIS (in Chinese). Resour Sci, 6: 2–9

Günther F, Thiele A, Gleixner G, Xu B, Yao T, Schouten S. 2014. Distribution of bacterial and archaeal ether lipids in soils and surface sediments of Tibetan lakes: Implications for GDGT-based proxies in saline high mountain lakes. Org Geochem, 67: 19–30

Guntiñas M E, Leirós M C, Trasar-Cepeda C, Gil-Sotres F. 2012. Effects of moisture and temperature on net soil nitrogen mineralization: A laboratory study. Euro J Soil Biol, 48: 73–80

Huguet A, Fosse C, Metzger P, Fritsch E, Derenne S. 2010. Occurrence

and distribution of extractable glycerol dialkyl glycerol tetraethers in podzols. Org Geochem, 41: 291–301

Huguet A, Wiesenberg G L B, Gocke M, Fosse C, Derenne S. 2012. Branched tetraether membrane lipids associated with rhizoliths in loess: Rhizomicrobial overprinting of initial biomarker record. Org Geochem, 43: 12–19

Li F, Zhang C L, Wang S, Chen Y, Sun C, Dong H, Li W, Klotz M G, Hedlund B P. 2014. Production of branched tetraether lipids in Tibetan hot springs: A possible linkage to nitrite reduction by thermotolerant or thermophilic bacteria? Chem Geol, 386: 209–217

Liu W G, Wang H Y, Zhang C L L, Liu Z H, He Y X. 2013. Distribution of glycerol dialkyl glycerol tetraether lipids along an altitudinal transect on Mt. Xiangpi, NE Qinghai-Tibetan Plateau, China. Org Geochem, 57: 76–83

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. Org Geochem, 42: 739–751

Mueller-Niggemann C, Utami S R, Marxen A, Mangelsdorf K, Bauersachs T, Schwark L. 2015. Distribution of tetraether lipids in agricultural soils-differentiation between paddy and upland management. Biogeosci Discuss, 12: 16709–16754

Nocita M, Stevens A, Noon C, van Wesemael B. 2013. Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. Geoderma, 199: 37–42

Pearson E J, Juggins S, Talbot H M, Weckstrom J, Rosen 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. Geochim Cosmochim Acta, 75: 6225–6238

Peterse F, Kim J H, Schouten S, Kristensen D K, Koc N, Sinninghe Damsté J S. 2009a. Constraints on the application of the MBT/CBT palaeothermometer at high latitude environments (Svalbard, Norway). Org Geochem, 40: 692–699

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. Org Geochem, 41: 1171–1175

Peterse F, van der Meer J, Schouten S, Weijers J W H, Fierer N, Jackson R B, Kim J-H, Sinninghe Damsté J S. 2012. Revised calibration of the MBT-CBT paleotemperature proxy based on branched tetraether membrane lipids in surface soils. Geochim Cosmochim Acta, 96: 215–229

Peterse F, van Der Meer M, Schouten S, Jia G, Ossebaar J, Blokker J, Sinninghe Damsté J S. 2009b. Assessment of soil *n*-alkane δD and branched tetraether membrane lipid distributions as tools for paleoelevation reconstruction. Biogeosciences, 6: 2799–2807

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. J Geophys Res, 119: 1738–1754

Rousk J, Baath E, Brookes P C, Lauber C L, Lozupone C, Caporaso J G, Knight R, Fierer N. 2010. Soil bacterial and fungal communities across a pH gradient in an arable soil. ISME J, 4: 1340–1351

Schouten S, Huguet C, Hopmans E C, Kienhuis M V, Sinninghe Damsté J S. 2007. Analytical methodology for TEX_{86} paleothermometry by high-performance liquid chromatography/atmospheric pressure chemical ionization-mass spectrometry. Anal Chem, 79: 2940–2944

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

Sun C J, Zhang C L L, Li F Y, Wang H Y, Liu W G. Distribution of branched glycerol dialkyl glycerol tetraethers in soils on the Northeastern Qinghai-Tibetan Plateau and possible production by nitrite-reducing bacteria. Sci China Earth Sci, doi: 10.1007/S11430-015-0230-2

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. Geochim Cosmochim Acta, 74: 4902–4918

Wang H Y, Liu W G, Zhang C L. 2014. Dependence of the cyclization of branched tetraethers (CBT) on soil moisture in the Chinese Loess Plateau and the adjacent areas: Implications for palaeorainfall reconstructions. Biogeosciences, 11: 10015–10043

Wang H Y, Liu W G, Zhang C L L, Liu Z H, He Y X. 2013. Branched and isoprenoid tetraether (BIT) index traces water content along two marsh-soil transects surrounding Lake Qinghai: Implications for paleo-humidity variation. Org Geochem, 59: 75–81

Wang H Y, Liu W G, Zhang C L L, Wang Z, Wang J X, Liu Z H, Dong H L. 2012. Distribution of glycerol dialkyl glycerol tetraethers in surface sediments of Lake Qinghai and surrounding soil. Org Geochem, 47: 78–87

Weijers J W H, Schouten S, van den Donker J C, Hopmans E C, Sinninghe Damsté J S. 2007a. Environmental controls on bacterial tetraether membrane lipid distribution in soils. Geochim Cosmochim Acta, 71: 703–713

Weijers J W H, Schefuss E, Schouten S, Sinninghe Damste J S. 2007b. Coupled thermal and hydrological evolution of tropical Africa over the last deglaciation. Science, 315: 1701–1704

Weijers J W H, Schouten S, Hopmans E C, Geenevasen J A, David O R, Coleman J M, Pancost R D, Sinninghe Damste J S. 2006. Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits. Environ Microbiol, 8: 648–657

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: 2959–2973

Weijers J W H, Bernhardt B, Peterse F, Werne J P, Dungait J A J, Schouten S, Sinninghe Damsté J S. 2011. Absence of seasonal patterns in MBT-CBT indices in mid-latitude soils. Geochim Cosmochim Acta, 75: 3179–3190

Xie S C, Pancost R D, Chen L, Evershed R P, Yang H, Zhang K X, Huang J H, Xu Y D. 2012. Microbial lipid records of highly alkaline deposits and enhanced aridity associated with significant uplift of the Tibetan Plateau in the Late Miocene. Geology, 40: 291–294

Yang H, Pancost R D, Dang X, Zhou X, Evershed R P, Xiao G, Tang C, Gao L, Guo Z, Xie S. 2014. Correlations between microbial tetraether lipids and environmental variables in Chinese soils: Optimizing the paleo-reconstructions in semi-arid and arid regions. Geochim Cosmochim Acta, 126: 49–69

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. Org Geochem, 82: 42–53

Zech R, Gao L, Tarozo R, Huang Y. 2012. Branched glycerol dialkyl glycerol tetraethers in Pleistocene loess-paleosol sequences: Three case studies. Org Geochem, 53: 38–44

Zhang C L, Wang J, Dodsworth J A, Williams A J, Zhu C, Hinrichs K U, Zheng F, Hedlund B P. 2013. *In situ* production of branched glycerol dialkyl glycerol tetraethers in a great basin hot spring (USA). Front Microbiol, 4: 181

Zhang C L, Wang J, Wei Y, Zhu C, Huang L, Dong H. 2012. Production of branched tetraether lipids in the lower pearl river and estuary: Effects of extraction methods and impact on bGDGT proxies. Front Microbiol, 2: 274