



# Lipid biomarker temperature proxy responds to abrupt shift in the bacterial community composition in geothermally heated soils

Cindy De Jonge<sup>a,\*</sup>, Dajana Radujković<sup>a</sup>, Bjarni D. Sigurdsson<sup>b,c</sup>, James T. Weedon<sup>a,d</sup>, Ivan Janssens<sup>a</sup>, Francien Peterse<sup>e</sup>

<sup>a</sup> University of Antwerp, Dept. of Biology, 2610 Wilrijk, Belgium

<sup>b</sup> The Agricultural University of Iceland, Fac. of Agricultural and Environmental Sciences, 311 Borgarnes, Iceland

<sup>c</sup> Global Ecology Unit CREA-FCSC-UAB, CSIC, Bellaterra (Catalonia) E-08193, Spain

<sup>d</sup> Vrije Universiteit Amsterdam, Dept. of Ecological Science, 1081 HV Amsterdam, the Netherlands

<sup>e</sup> Utrecht University, Dept. of Earth Sciences, 3584 CD Utrecht, the Netherlands

## ARTICLE INFO

### Article history:

Received 15 April 2019

Received in revised form 10 July 2019

Accepted 12 July 2019

Available online 15 July 2019

### Keywords:

Biomarker lipid

Soil brGDGTs

Bacterial community

## ABSTRACT

Specific soil bacterial membrane lipids, branched glycerol dialkyl glycerol tetraethers (brGDGTs), are used as an empirical proxy for past continental temperatures. Their response to temperature change is assumed to be linear, caused by physiological plasticity of their, still unknown, source organisms. A well-studied set of geothermally warmed soils (Iceland) shows that the brGDGT fingerprint only changes when the soil mean annual temperature is warmer than 14 °C. This sudden change in brGDGT distribution coincides with an abrupt shift in the bacterial community composition in the same soils. Determining which bacterial OTUs are indicative for each soil cluster shows that Acidobacteria are possible brGDGT producers, together with representatives from the Actinobacteria, Chloroflexi, Gemmatimonadetes and Proteobacteria. Projecting the lipid fingerprint of the cold and warm bacterial communities onto the global soil calibration dataset creates two distinct soil clusters, in which brGDGTs respond differently to temperature and, especially, soil pH. We show that, on a local scale, a community effect rather than physiological plasticity can be the primary driver of the brGDGT-based paleothermometer along large temperature gradients. This threshold response needs to be taken into account when interpreting brGDGT-based paleo-records.

© 2019 Elsevier Ltd. All rights reserved.

## 1. Introduction

Branched glycerol dialkyl glycerol tetraethers (brGDGTs; [Supplementary Fig. S1](#)) are bacterial membrane-spanning lipids that are abundant in soils, fresh and saline water columns, and sediments ([Schouten et al., 2013](#)). Their structural diversity ([Sinninghe Damsté et al., 2000](#); [Weijers et al., 2006](#); [De Jonge et al., 2013](#), [Ding et al., 2016](#)) includes variability in the number of methylations (4–6) and the position ( $\alpha$  and/or  $\omega 5$  or  $\omega 6$  or  $\omega 7$ ) of the outer methyl branch. Following internal cyclisation, one or two cyclopentane moieties can be formed. The degree of methylation of 5-methyl brGDGTs can be summarized using the “Methylation of 5-Methyl Branched Tetraethers” (MBT'<sub>5ME</sub> [Eq. (1)]) ratio. In a set of globally distributed surface soils, this MBT'<sub>5ME</sub> ratio varies with prevailing mean annual air temperature ([De Jonge](#)

[et al., 2014](#)). This empirical relation and its transfer function [Eq. (2)] allow the use of brGDGTs as a paleo-temperature proxy ([Weijers et al., 2007a](#)), reconstructing continental climate from distributional changes in downcore paleosoils or sedimentary archives (e.g., [Weijers et al., 2007b](#); [Peterse et al., 2011](#); [Inglis et al., 2017](#)). Although it is one of the few proxies that can be used to quantitatively reconstruct continental climate change over geologic time scales, the value of the records is currently limited by the accuracy of the existing empirical transfer functions. Despite various attempts to reduce the error on these functions ([Peterse et al., 2012](#); [De Jonge et al., 2014](#); [Naafs et al., 2017a](#)), the average error on the calibration has not been reduced below ~4 °C.

The lack of further progress can be attributed to our poor understanding of the mechanism behind the observed temperature dependence. It was initially proposed that brGDGT-producers introduce brGDGTs with contrasting structures in the cell membrane to ensure optimal membrane fluidity and permeability under different pH and temperature conditions ([Weijers et al.,](#)

\* Corresponding author at: NO G 57, Sonneggstrasse 5, 8092 Zurich, Switzerland.  
E-mail address: [cindy.dejonge@erdw.ethz.ch](mailto:cindy.dejonge@erdw.ethz.ch) (C. De Jonge).

2007a). Alternatively, or complementarily, the distributional changes can also be explained by shifts in the composition of brGDGT producers, which indirectly cause changes in the brGDGT fingerprint. Unfortunately, little is known about the phylogenetic identity and diversity of brGDGT producing organisms, hampering the identification of the exact mechanisms. From the stereochemistry of the glycerol moieties we do know that brGDGTs are produced by bacteria (Weijers et al., 2006), and their stable isotopic composition indicates a heterotrophic lifestyle of their producers (Pancost and Sinninghe Damsté, 2003; Oppermann et al., 2010; Weber et al., 2015). Furthermore, the brGDGT concentration in a peat core (Weijers et al., 2009) and pH-controlled plots (Peterse et al., 2010; Bartram et al., 2014) correlates with the abundance of the diverse phylum Acidobacteria. After one single brGDGT compound was detected in an Acidobacterial culture (Sinninghe Damsté et al., 2011, 2014, 2018), it has been assumed that Acidobacteria contain the source organisms of brGDGTs.

In the absence of representative producers of the complete suite of fifteen brGDGT lipids used in paleoenvironmental proxies, laboratory incubations cannot be used to determine the mechanism of their temperature dependence. Furthermore, manipulations of natural soils to study brGDGT-temperature relations are challenging due to the long turnover time of brGDGTs (8–41 years; Weijers et al., 2010; Huguet et al., 2017). Therefore, we targeted five well-characterized soil temperature gradients, caused by below-ground geothermal warming (Sigurdsson et al., 2016). Geothermal gradients are an excellent natural laboratory, as the variation in possibly confounding environmental factors is less pronounced than in comparable large-scale temperature gradients (e.g., altitudinal or latitudinal transects), allowing the effects of temperature (both direct and indirect) to be isolated (Peterse et al., 2009; Woodward et al., 2010; O’Gorman et al., 2014). Analysis of both brGDGT lipids and the bacterial community composition along a temperature gradient in the same soil transect allows us, for the first time, to determine whether membrane adaptation or community change is the dominant mechanism underlying the brGDGT-paleothermometer at this site.

## 2. Materials and methods

### 2.1. Site description and soil sampling

The ForHot study site in Iceland (Hveragerði [64.008°N, 21.178°W; 83–168 m a.s.l.], Fig. 1) comprises 90 permanent research plots that are designed to study the effect of warming on above- and below-ground ecosystem processes. The presence of subsurface geothermal warming causes local temperature gradients, of up to 20 °C above the ambient Icelandic soil temperature. Of these permanent research plots, 30 plots were established in grasslands that have been subjected to geothermal warming (5 transects with warming increasing between plots A to F) for at least 50 years (Fig. 1). The plots furthest from the fault (A) have experienced no geothermal warming and have a mean annual soil temperature of 5 °C. The other plots in each transect were chosen to exhibit different warming intensities over a distance of ±50 m (B–F; Table 1). The transect at the ForHot study site spans a mean annual measured soil temperature of 5–25 °C (averaged between 23/10/2014 and 23/10/2016). The seasonal range in soil temperatures at 10 cm depth is about 12 °C in all soils, both unwarmed and warmed (Sigurdsson et al., 2016). During the study period the minimum soil temperature in winter was 0 °C. The abiotic parameters measured in situ in the permanent research plots (2 × 2 m) are the soil moisture (±monthly) and the soil temperature (hourly) (Sigurdsson et al., 2016). The soil temperature data

reported in Table 1 is averaged yearly (Mean Soil Temperature). The soil moisture data reported are a yearly average, based on 22 to 26 measurements over the period of 10/04/2013 and 30/08/2016. In July 2016, 30 samples were taken just outside the permanent research plots, at 20 cm distance of the soil temperature logger. A soil corer (diameter 2 cm) was used to sample the top 10 cm of the soil, of which the 5–10 cm depth interval was stored and transported frozen. This depth interval accords with the depth of the soil temperature sensor (10 cm deep). The soil samples were freeze-dried, ground with mortar and pestle and sieved over a 2 mm sieve. The pH of the soil samples was measured on 1 g of freeze-dried soil in deionized water (1:2.5 wt ratio), in duplicate. The mean measurement error is 0.01 pH units.

### 2.2. GDGT analysis

5 g of each soil sample (dry wt) was extracted (4×) following a modified Bligh and Dyer method using a mixture of methanol (MeOH), dichloromethane (DCM), and a phosphate-buffer (2:1:0.8, v/v/v) (Pitcher et al., 2009), which resulted in the Bligh and Dyer extract (BDE). The extract was separated over an activated SiO<sub>x</sub> column, where the core lipid GDGTs were collected using ethyl acetate/MeOH (1:1, v/v). Prior to analysis for GDGTs, 98 ng of C<sub>46</sub> internal standard was added to the BDE (for structural information see Huguet et al., 2006), after which an aliquot was dissolved in *n*-hexane/isopropanol (99:1, v/v) and filtered over a 0.45 μm PTFE filter. These fractions were injected on an Agilent 1260 high performance liquid chromatograph (HPLC) system coupled to an Agilent 6130 single quadrupole mass spectrometer (MS). The chromatography, detection, and compound-peak integration were performed according to previously published methods (Hopmans et al., 2016).

In order to summarize the GDGT distributions, and the reconstructed mean annual air temperature (MAT<sub>rec</sub>), the MBT<sub>5ME</sub> ratio and the MAT<sub>rec</sub> were calculated as previously defined (De Jonge et al., 2014). In the following formulae the compound names refer to the structures in Supplementary Fig. S1:

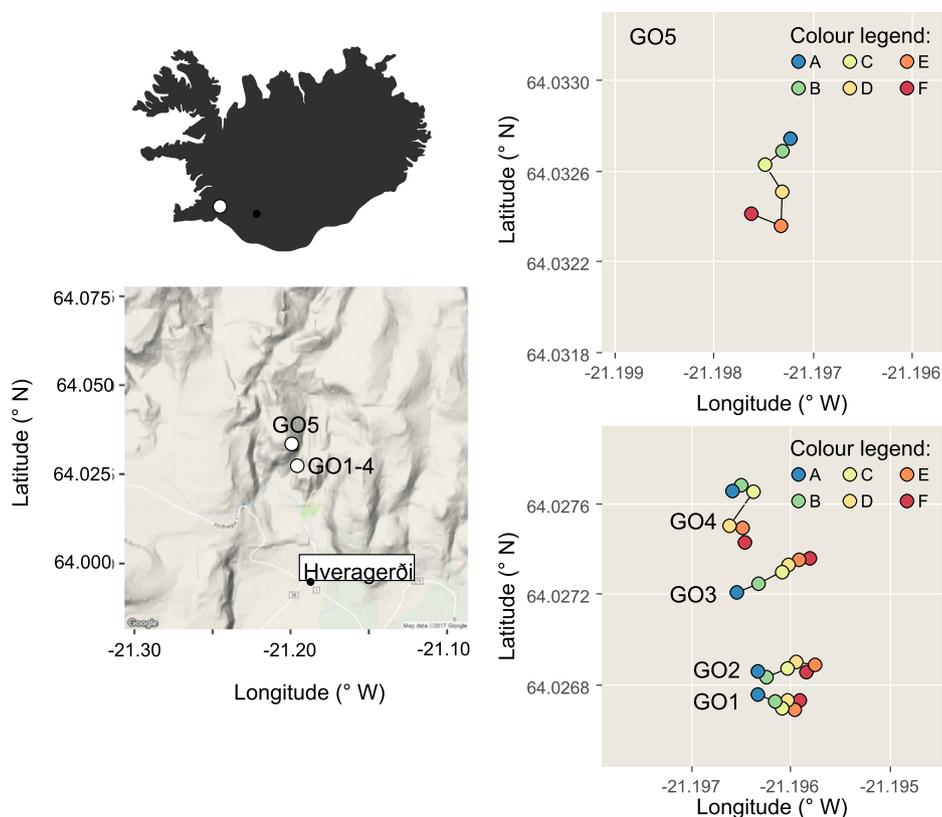
$$\text{MBT}_{5\text{ME}} = (\text{Ia} + \text{Ib} + \text{Ic}) / (\text{Ia} + \text{Ib} + \text{Ic} + \text{IIa} + \text{IIb} + \text{IIc} + \text{IIIa}) \quad (1)$$

$$\text{MAT}_{\text{rec}} = -8.57 + 31.45 \times \text{MBT}_{5\text{ME}} \quad (2)$$

### 2.3. Bacterial community analysis

The bacterial DNA extraction, amplification, and sequencing was performed on soils sampled at the ForHot sites in 2015 (described in Radujković et al., 2018). Although the DNA and lipid data are derived from soil samples taken a year apart, the slow turnover time of core lipid brGDGTs should result in negligible year-to-year variation (Weijers et al., 2010; Huguet et al., 2017), and we expect the lipid signature in 2015 to be very similar to the signature in 2016. To avoid the effect of inter-annual changes on the lipid signal, we focus on core lipids only.

The raw sequence data was deposited in the SRA database with the accession number SRP099121 (Radujković et al., 2018). For this study, we used the SILVA v123 taxonomy (Quast et al., 2012), requiring the reanalysis of the raw sequence data produced by Radujković et al. (2018). The first part of the bioinformatics analysis was performed using USEARCH (v8.0.1517; Edgar, 2013). After paired end reads were merged, 76.9% of the pairs were successfully assembled into contigs. The resulting FASTQ files were quality filtered using a threshold of 0.05 total expected errors for all bases in the read (1.3 M, 93.7% converted). Prior to operational taxonomic unit (OTU) clustering, the obtained sequences were



**Fig. 1.** Situation of the ForHot study site in Iceland (insert), and distribution of the samples and transects. Color coding indicates an approximate level of warming (A–F, where F shows the strongest degree of warming), with six levels of warming per transect.

**Table 1**

Measured mean soil temperature (MST, yearly average), soil pH, soil moisture (SM, yearly average), the scores of the sites on the first and second ordination axis following an NMDS analysis of the bacterial community (BC, NMDS1 and NMDS2), and the brGDGT-based ratios  $MBT'_{5ME}$  [Eq. (1)] and CI [Eq. (3)].

Site	MST (°C)	pH	SM (%)	BC (NMDS1)	BC (NMDS2)	$MBT'_{5ME}$	CI
1A	4.9	5.5	52	−0.43	0.02	0.45	0.43
1B	5.9	5.9	50	−0.51	0.03	0.55	0.54
1C	7.2	6.2	49	−0.22	−0.19	0.46	0.48
1D	9.3	6.2	57	−1.02	0.27	0.41	0.43
1E	11.1	6.3	48	n.d.	n.d.	0.4	0.37
1F	20.2	6.9	44	n.d.	n.d.	0.69	0.67
2A	5.4	5.3	56	n.d.	n.d.	0.52	0.52
2B	5.7	5.6	56	n.d.	n.d.	0.41	0.38
2C	7.6	5.8	36	−0.7	−0.16	0.63	0.62
2D	8.4	5.6	58	n.d.	n.d.	0.61	0.61
2E	11.1	6.3	45	n.d.	n.d.	0.53	0.49
2F	19.0	6.3	44	n.d.	n.d.	0.85	0.84
3A	5.9	5.7	54	−0.72	−0.09	0.48	0.47
3B	6.4	5	49	−0.28	0.12	0.5	0.47
3C	7.8	5.8	55	−0.58	−0.03	0.5	0.47
3D	8.6	5.7	37	−0.49	−0.04	0.49	0.47
3E	12.3	5.8	46	n.d.	n.d.	0.54	0.52
3F	21.0	5.8	36	2.97	0.00	0.72	0.72
4A	6.0	5.7	54	−0.64	0.07	0.49	0.47
4B	5.6	6	66	−0.58	−0.02	0.46	0.46
4C	6.9	5.4	57	−0.40	−0.04	0.54	0.51
4D	8.7	5.4	51	−0.47	−0.05	0.56	0.54
4E	13.5	5.9	41	−0.49	−0.12	0.58	0.57
4F	20.5	5.6	45	2.28	0.19	0.87	0.87
5A	6.8	7	61	−1.28	0.03	0.46	0.46
5B	7.2	6.4	66	−1.03	0.02	0.44	0.42
5C	11.6	5.7	64	−0.5	0.13	0.5	0.48
5D	10.2	5.4	52	−0.14	−0.01	0.52	0.51
5E	13.7	5.8	40	−0.79	−0.07	0.53	0.51
5F	25.0	5	41	6.06	−0.06	0.66	0.65

n.d. indicates not determined.

dereplicated, and singletons (34.2% of sequences) were discarded. OTUs were defined using the UPARSE-OTU algorithm, where the OTU “radius” is default 3.0 (97% sequence similarity). After chimera removal (leaving 8549 non-chimeric OTUs) all original reads were mapped to the non-chimeric OTUs using the USEARCH algorithm with global alignments with the identity threshold of 0.97, yielding an OTU Table. The OTUs were aligned against the core alignment from SILVA v123 (Quast et al., 2012), resulting in 8354 OTUs. This alignment was filtered, and with unaligned OTUs removed, the taxonomy was assigned using the rdp classifier (Wang et al., 2007), trained on the SILVA taxonomy. The downstream analyses and statistical analyses were conducted in R v3.2.2 (R core team, 2015) using the “phyloseq”, “rarefy” and “vegan” packages.

The reads were rarefied at a depth of 7193 reads, retaining 5967 OTUs. To limit our analysis to the more abundant OTUs, the most abundant 90% of sequences was retained for further analysis (2898 OTUs, 6998 reads). Ward’s minimum variance method, implementing the Ward clustering criterion (Ward, 1963), was used for clustering the sites based on the similarity in their OTU distribution, resulting in 2 clusters. We employed the package “indspecies” in R (Cáceres and Legendre, 2009) to determine those OTUs that are typical for one of these clusters, based on their relative abundance and their relative frequency of occurrence (“indicator species”). P-values were corrected using Sidak’s correction for multiple testing ( $p < 0.05$ ). Following up on this, the orders that were characterized by a higher number of indicator species than what is expected from random subsampling (permutation = 999,  $p < 0.05$ ), were identified (Table 3).

#### 2.4. Statistical analysis – linear correlations and models

The squared Pearson correlation coefficient ( $r^2$ ), and p-value of correlation test between paired samples are reported when performing linear regressions, conducted in R v3.2.2 (R core team, 2015). A multiple linear regression model was fitted to test whether the coefficients soil pH and soil moisture (in addition to the mean annual soil temperature (MST)) can result in a better regression with  $MBT'_{5ME}$  values in ForHot soils ( $MBT'_{5ME} \sim MST + \text{soil pH}$ ) or ( $MBT'_{5ME} \sim MST + SM$ ) or ( $MBT'_{5ME} \sim MST + \text{soil pH} + SM$ ). To select linear vs stepwise models along a temperature gradient, both were estimated using maximum likelihood estimation (package “bbmle”; Bolker et al., 2017). For this approach, the score on NMDS axis 1 (global Multidimensional Scaling using monoMDS based on Bray-Curtis distances) is used as an estimate of the similarity of the bacterial community composition. Whether a linear or stepwise model fits the temperature dependency of the  $MBT'_{5ME}$  better, was determined based on the Akaike information criterion parameter (AIC; Akaike, 1973). To determine the unique fraction of the variance in brGDGTs explained by soil pH, measured mean annual soil moisture and MST at the ForHot sites, a partial redundancy analysis (pRDA) was employed. The same approach was used for soil pH, modeled moisture, or mean annual temperature in a dataset of global soils. Both on the local and global scale, the r-value of significant linear correlations (based on p-value cut-off of 0.05 or 0.10) between the environmental parameter and brGDGT relative abundance are reported. To compare whether the fractional abundance of Global Cluster Cold and Global Cluster Warm soils show a different linear dependency (both intercept and slope) with pH and MAT, the cluster ID was introduced as an interacting factor in a linear model [ $Ia \sim \text{Cluster ID} * \text{Soil pH}$ ] and [ $Ia \sim \text{MAT} * \text{Cluster ID} + \text{Cluster ID} * \text{Soil pH}$ ], where an asterisk

indicates an interaction term, using the “stats” package in R (R core team, 2015).

### 3. Results

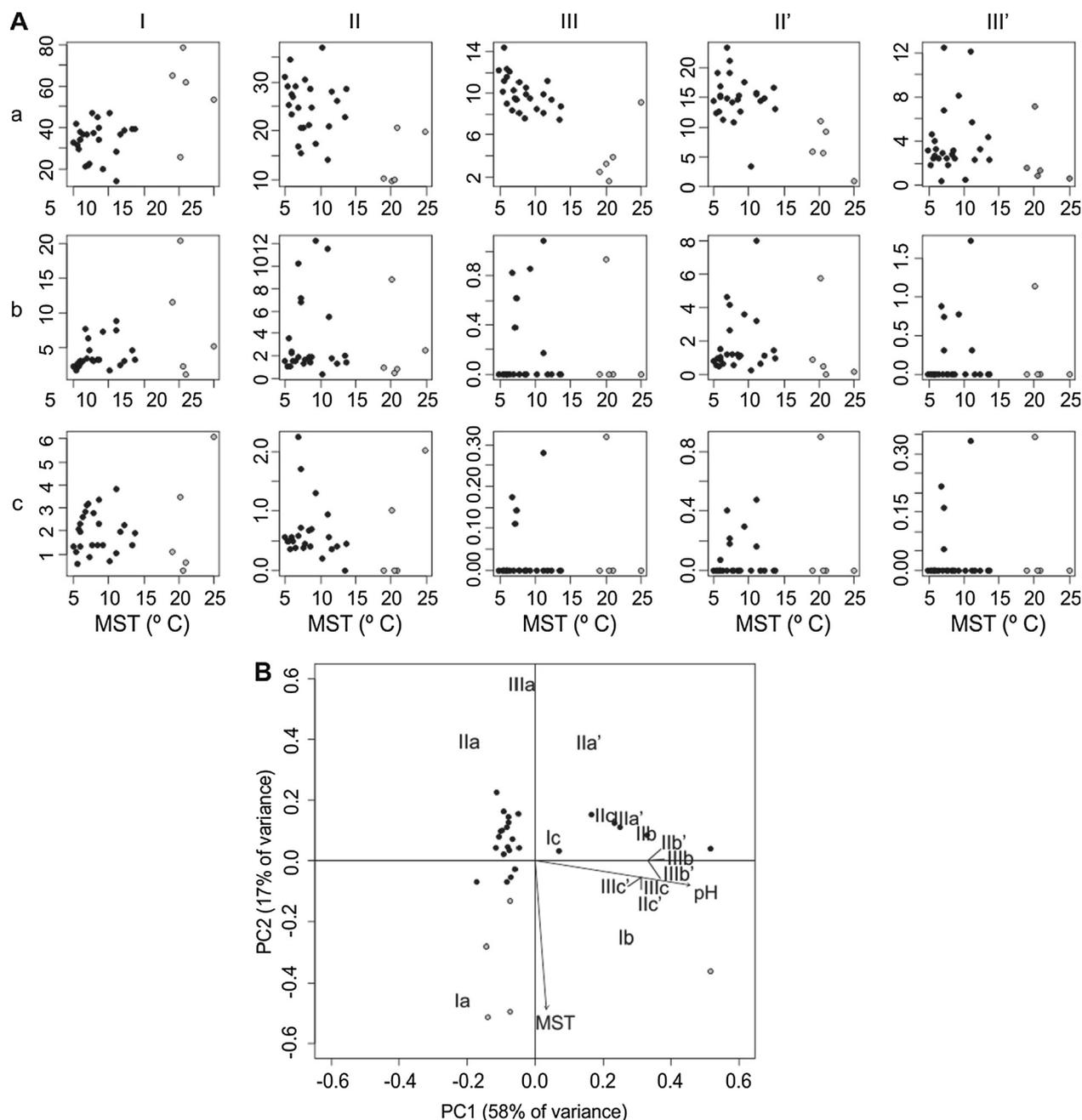
BrGDGTs were detected in all soils from the 30 grassland plots at the ForHot study site (Supplementary Table S1 and Fig. 2). BrGDGT Ia shows the widest range in fractional abundance (14% and 78%), spanning most of the variability that is present in the global soil dataset (10–99%). 5-Methyl compounds with two and three branches were also present; pentamethylated IIa varied between 9.7% and 37% and hexamethylated IIIa between 1.6% and 14%. This translates into  $MBT'_{5ME}$  values [Eq. (1)] ranging between 0.40 and 0.87 (Fig. 3A; Table 1), which cover 60% of the spread in  $MBT'_{5ME}$  values on a global scale (0.21–1.00; De Jonge et al., 2014; Ding et al., 2015; Xiao et al., 2015; Yang et al., 2015; Lei et al., 2016; Naafs et al., 2017a,b). The variation in brGDGT distribution is summarized using a principal component analysis (PCA; Fig. 2B). Those brGDGTs previously identified as pH-sensitive compounds (5Me brGDGTs with cyclopentane groups [Ib, IIb, IIc, IIIb, IIIc] and 6Me brGDGTs [IIa', IIb', IIc', IIIa', IIIb', IIIc']) plot along PC1 (principal component 1), which explains 58% of the variance. The second PC (17% of variance) reflects the increase of brGDGT Ia, relative to the brGDGTs IIa and IIIa. The yearly averaged MST, soil pH and SM were determined for each plot (see Materials and methods) and reported in Table 1. When plotted a posteriori in the ordination space, soil pH varies with PC1, and MST varies with PC2.

The extended global soil dataset is a compilation of brGDGT distributions previously reported in De Jonge et al. (2014), Ding et al. (2015), Xiao et al. (2015), Yang et al. (2015), Lei et al. (2016), and Naafs et al. (2017a,b). Selected brGDGT-based proxies have been calculated for this extended global soil dataset (Fig. 5). For 22 out of the 30 soils analysed for brGDGT lipids, DNA was extracted and amplified successfully (Radujković et al., 2018). Following the bio-informatic approach described above, the Proteobacteria (30%), Acidobacteria (28%), Actinobacteria (11%) and Chloroflexi (10%) are the phyla that represent >10% of the OTU counts at the ForHot site. Based on an NMDS analysis (Table 1), the main variation in the community composition is related to soil temperature (r between NMDS1 and MST values = 0.87 [ $p < 0.01$ ]), followed by soil pH (r between NMDS2 and soil pH values = -0.46,  $p < 0.05$ ). Parameters that describe the phylogenetic shift in the bacterial community, are reported in Table 3 and Figs. 4 and 5.

### 4. Discussion

#### 4.1. The effect of temperature on brGDGTs

To assess the influence of temperature on the brGDGT signal in the ForHot soils, brGDGTs distributions were translated into temperature for each plot using the  $MBT'_{5ME}$  index and the global transfer function (De Jonge et al., 2014) [Eq. (2)]. Comparison of brGDGT-derived temperatures with the mean annual measured soil temperature for each plot shows that reconstructed temperatures follow the ForHot geothermal gradients (Fig. 3A;  $r^2 = 0.76$ ,  $p < 0.01$ ). Nevertheless, the relation still has a large error (RMSE = 4.2 °C). On a global scale, a number of confounding factors have been proposed to explain the discrepancy between reconstructed and measured temperatures, such as the offset between soil and air temperature, seasonality, soil moisture availability, and pH (Weijers et al., 2007a; Peterse et al., 2012; De Jonge et al., 2014; Naafs et al., 2017a,b). In the ForHot soils the use of



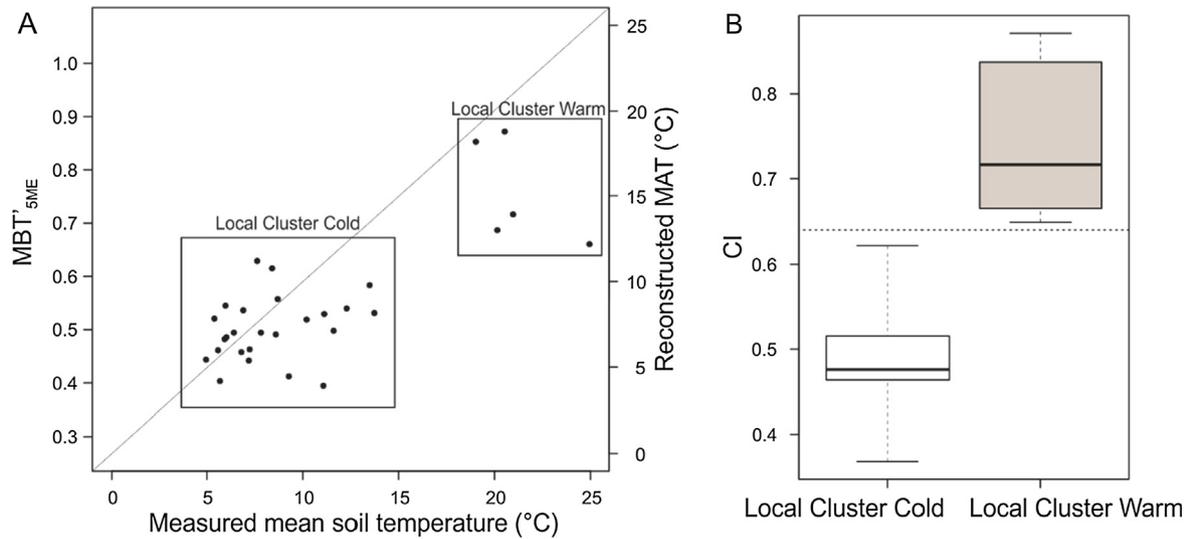
**Fig. 2.** The y-axes show the percentage (%) of the 15 brGDGTs as encountered in the ForHot soils (Supplementary Table S1). Compounds below detection limit are plotted as 0. Samples with a mean annual soil temperature >14 °C have been coloured in grey. Nominal numbers and letters refer to the structures shown in Supplementary Fig. S1.

in situ soil temperature measurements, rather than interpolated weather station data on a global scale, does not substantially reduce the uncertainty associated with the proxy. Also, removing the effect of differences in the seasonal amplitude of soil temperature and limiting the impact of frozen days (Sigurdsson et al., 2016) does not seem to contribute to significantly more accurate temperature reconstructions. Finally, to assess the possible influence of soil pH (De Jonge et al., 2014) and soil moisture (Dang et al., 2016) on the variation of  $MBT'_{5ME}$  at ForHot, they were, together with mean soil temperature, included in a linear model. Both soil pH and soil moisture were not significant in explaining variation in the  $MBT'_{5ME}$  values ( $p > 0.1$ ), both as separate factors or combined with the soil temperature. Hence, whereas soil temperature

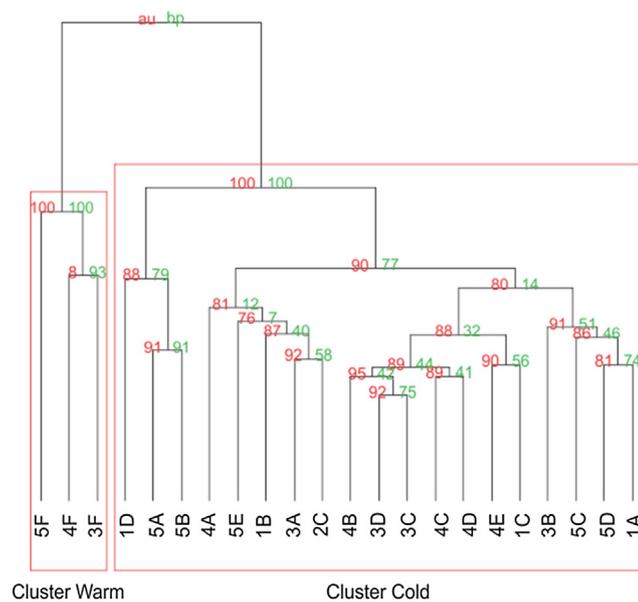
is the main abiotic factor influencing brGDGTs distributions in ForHot soils, the larger error (RMSE = 4.2 °C) in this dataset cannot be attributed to other abiotic soil parameters.

#### 4.2. Direct comparison of brGDGT distributions and bacterial community composition

To describe changes in the bacterial community composition, the ForHot soils have been clustered based on their bacterial community composition (operational taxonomic unit (OTU) distribution). This leads to one cluster with all soils <14 °C (Cluster Cold,  $n = 19$ ), and another cluster that is comprised of soils >14 °C (Cluster Warm,  $n = 3$ ; Fig. 4). This community shift is significant,

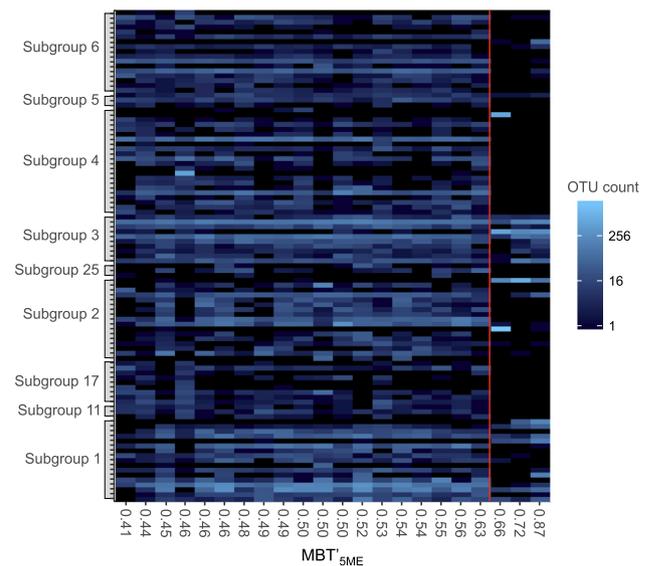


**Fig. 3.** BrGDGT-based indices in the ForHot soils. (A) Scatterplot of the measured soil temperature values against the  $MBT'_{5ME}$  index values [Eq. (1)], with the reconstructed  $MBT'_{5ME}$ -based temperatures indicated [Eq. (2)] for geothermally heated soil transects at ForHot, Iceland. (B) Boxplot of the Community Index (CI [Eq. (3)]) values of the ForHot soils, for Local Cluster Cold (white) and Local Cluster Warm (grey). The cut-off value of 0.64 is indicated.



**Fig. 4.** Dendrogram based on the OTU distribution in ForHot soils, with bootstrap and AU values indicated, showing the two clusters under discussion.

as 80% of the OTUs are unique for each cluster (shown for Acidobacteria; Fig. 5). This confirms the results from Radujković et al. (2018), who had already shown that the bacterial community composition in the ForHot soils shifts abruptly in soils with a mean soil temperature  $>14$  °C. Evaluation of the relative distribution of brGDGTs per soil cluster shows that all soils in Cluster Cold have  $MBT'_{5ME}$  index values  $<0.64$ , while the soils in Cluster Warm all have  $MBT'_{5ME}$  index values  $>0.66$  (Fig. 3A). Unexpectedly, the  $MBT'_{5ME}$  values within each cluster remain relatively stable and thus do not correlate with soil temperature (Cluster Cold,  $n = 25$ ,  $p = 0.19$ ; Cluster Warm,  $n = 5$ ,  $p = 0.25$ ), although the soils in the Cold and Warm clusters cover a temperature range of 9 °C and 6 °C, respectively. This is in contrast to the gradual increase in  $MBT'_{5ME}$  values with mean annual air temperature that is observed in a large dataset of globally distributed soils (i.e. De Jonge et al.,



**Fig. 5.** The fractional abundance of the Acidobacterial (top 75% of Acidobacterial sequences depicted) in the ForHot soils, ordered by subgroup and plotted along increasing  $MBT'_{5ME}$  [Eq. (1)] values. The red line indicates the boundary between Cluster Cold and Cluster Warm soils. This figure illustrates that very few OTUs are shared across samples of the two clusters. Across all phyla, 80% of the OTUs are unique to each cluster. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2014). At ForHot, significantly different  $MBT'_{5ME}$  values (Welch Two Sample  $t$ -test:  $p < 0.01$ , where sample 1: soils  $>14$  °C and sample 2: soils  $<14$  °C) are only encountered after crossing the temperature threshold of 14 °C that coincides with the observed shift in bacterial community composition. Indeed, a stepwise increase with soil temperature fits the  $MBT'_{5ME}$  values better than a linear increase (AIC of  $-71$  (stepwise) vs  $-64$  (linear); Table 2). A similar stepwise change with temperature is only recognized in the bacterial community composition (AIC of 52 (stepwise) vs 59 (linear)).

Due to the absence of a temperature response within a uniform bacterial community (mean soil temperature  $>14$  °C), it seems that physiological plasticity does not result in an increase in  $MBT'_{5ME}$

**Table 2**

Squared Pearson correlation coefficient ( $r^2$ ), the p-value and the Akaike Information criterion (AIC) of the linear model between the mean annual measured soil temperature in the ForHot soils and selected lipid, biotic and abiotic parameters (all values reported in Table 1), fitted using maximum likelihood estimation. The AIC of the stepwise model fitted on the same data is also reported<sup>a</sup>. A lower AIC value (in bold) indicates a better model fit.

	Linear model		Stepwise model <sup>a</sup>	
	Adj. $r^2$	p-value	AIC	AIC
MBT <sub>5ME</sub>	0.56	<0.01	-64	<b>-71</b>
BC (NMDS1)	0.75	<0.01	59	<b>52</b>
pH	-0.03	>0.71	n.d.	n.d.
SM (%)	0.3	<0.01	<b>440</b>	545

<sup>a</sup> The stepwise model is a model with one breakpoint defined at 14 °C, and the slope in each segment constrained to be 0, and intercepts on either side of the breakpoint estimated using maximum likelihood estimation.

values with increasing temperatures at ForHot. Here, soil bacteria produce significantly increased MBT<sub>5ME</sub> values only when their composition has been changed, suggesting that a shift in the composition of the bacterial community is needed for the brGDGT distribution to change with temperature.

A more detailed look at the phylogenetic nature of the shift in the bacterial community may reveal a link between specific clades and the observed threshold in MBT<sub>5ME</sub> values. Based on the relative frequency of occurrence of each OTU in each cluster, we can identify “indicator species” (corrected p-value < 0.05) for each of the two soil clusters (package “indspecies”; Cáceres and Legendre, 2009), resulting in the identification of 230 cold-indicative OTUs, and 72 warm-indicative OTUs, with a large phylogenetic diversity (Table 3). This large phylogenetic diversity fits with earlier reports that the response of bacterial OTUs to temperature changes is poorly conserved on both deep (i.e. Phylum and Subphylum) and shallow (i.e. Order) phylogenetic levels, both on a local (DeAngelis et al., 2015) and global scale (Oliverio et al., 2017). On the other hand, Sinnighe Damsté et al. (2011, 2014, 2018) observed that the potential to produce brGDGT precursor *iso*-diabolic acid, and the position of methylation of *iso*-diabolic acid, may be controlled by the phylogenetic affiliation with Acidobacterial subgroups. This indicates that the production of brGDGTs is possibly phylogenetically conserved. We report those clades of indicator species that show a phylogenetically conserved response between clusters in Table 3. Those orders that also show a phylogenetically conserved response to the change in brGDGT distribution, are considered especially good candidates to contain producers of brGDGT lipids (Table 3, in bold).

The Acidobacteria is currently the only phylum that has been shown to produce brGDGTs and/or precursor lipids (Weijers et al., 2006, 2010; Sinnighe Damsté et al., 2011, 2014, 2018), and are indeed present in both Warm and Cold clusters at ForHot (Fig. 5). Table 3 shows that, using the criteria outlined above, subgroups 1, 2, 4, 7 and 11 are good candidates for containing brGDGT producers in the Cluster Cold soils, while subgroup 1, 13 and 15 are good candidates for containing brGDGT producers responsible for the Cluster Warm signal. Subgroup 1, which also contains the only Acidobacterial cultures (n = 2) where a brGDGT compound was encountered in vitro, contains nine indicator OTUs of the Cluster Cold soils, while also containing four indicator OTUs for the Cluster Warm soils. As the OTUs from this subgroup show a clear temperature dependency, which is what we would expect from the producer of brGDGTs lipids, Acidobacteria subgroup 1 remains a good candidate for containing members that are relevant for producing the brGDGTs signal in the natural environment. It also illustrates the need for an OTU-level analysis, as OTUs within this

subgroup have contrasting responses to soil warming (Fig. 5). Although the presence of complete brGDGT lipids is currently limited to cultures of Acidobacteria subgroup 1, other subdivisions of Acidobacteria (subgroups 1, 3, 4 and 6) have been shown to produce possible precursor compounds of brGDGT lipids (Sinnighe Damsté et al., 2011, 2014, 2018). Of these subgroups, subgroup 4 is significantly increased in typically cold soils, reflecting an environmental control on these producers of *iso*-diabolic mono-glycerol ethers. The even larger diversity of Acidobacterial subgroups (subgroups 7 and 11) that are good candidates for containing brGDGT producers possibly reflects a phylogenetically more diverse source of the environmental brGDGT signal. This could also explain the larger diversity of brGDGT compounds encountered in the environment (n = 15), compared to the signal produced by subgroup 1 cultures in vitro (n = 1).

Although the presence of the brGDGT (precursor) lipids has so far only been reported for Acidobacteria, this could reflect the fact that the harsh extraction protocol that is required to release brGDGTs and their precursors (acid hydrolysis of the cell material) is not commonly performed when describing the membrane lipids of bacterial cultures (Sinnighe Damsté et al., 2018). Hence, other bacterial phyla could possibly also contain brGDGT producers and are therefore included in Table 3. Those phyla that contain clades (Class or Order) that are good candidates for containing producers of brGDGT lipids are Actinobacteria, Chloroflexi, Gemmatimonadetes, Proteobacteria and Verrucomicrobia. Recently, the distribution of the gene responsible for the formation of mono-glycerol compounds was also found in a number of Delta-Proteobacteria (Sinnighe Damsté et al., 2018). However, very few OTUs of this subphylum were identified as bio-indicators for Cluster Cold or Cluster Warm soils (Table 3).

#### 4.3. Projecting the effect of the bacterial community shift on the global calibration dataset

Similar to the phylogenetic variability observed at ForHot, warm- and cold-indicative bio-indicator OTUs have also been identified on a global scale. Using both laboratory experiments and an environmental dataset of 210 soils (mean annual temperature between -3.2 °C and 22.8 °C), Oliverio et al. (2017) describe both warm- and cold bio-indicator OTUs within Acidobacteria and Proteobacteria, as well as the presence of mainly warm-responsive OTUs within the Actinobacteria. These are changes that are therefore observed both along the ForHot geothermal gradient (this study) and on a global scale (Oliverio et al., 2017). Oliverio et al. (2017) report that the strongest phylogenetic signal in the temperature response of OTUs is seen within the phylum Acidobacteria. This pattern is mostly associated with the Acidobacterial families Koribacteraceae and Acidobacteriaceae (both within Acidobacteriales, commonly known as Acidobacteria Subgroup 1; Ward et al., 2009; Thrash and Coates, 2010). This agrees with the temperature effect on Acidobacterial subgroup 1 OTUs at the ForHot geothermal gradient (Table 3).

There is thus a good agreement between the impact of changing soil temperature on the microbial community composition in the geothermally warmed soils and when comparing soils across global-scale climatic gradients. Therefore, we postulate that the effect of a community shift on the brGDGT distribution could also be visible in a global soil brGDGT lipid dataset. Here, we assume that observations on a local scale can be extrapolated to the global scale but wish to highlight that these findings should be corroborated by further studies.

To test whether the global scale confirms our local observation of community-dependent lipid signatures, we first identified the

**Table 3**  
Summary of the complete phylogenetic diversity of the bio-indicator OTUs of Local Cluster Cold and Local Cluster Warm. The count of OTUs per phylogenetic clade is reported. Those clades that are significantly enriched in bio-indicators are indicated in bold.

Taxonomy			Cold bio-indicators (n)	Warm bio-indicators (n)
Unknown			1	6
<i>Acidobacteria</i>				
	Acidobacteria	<b>Acidobacteriales (SG1)</b>	<b>9</b>	<b>4</b>
		SG2	<b>6</b>	2
		SG3	3	1
		SG4	<b>9</b>	
		SG5	1	
		SG6	11	2
		SG11	<b>3</b>	
		SG13	1	<b>2</b>
		SG15		<b>1</b>
		SG17	3	
	Holophagae	SG7	<b>9</b>	
	SG 22		4	
<i>Actinobacteria</i>				
	Unknown		1	
	Acidimicrobiia	Acidimicrobiales	<b>10</b>	1
	Actinobacteria	Unknown	1	
		Pseudonocardiales	2	
		Micrococcales	1	
		Micromonosporales	1	
		Frankiales	2	
		Propionibacteriales	1	1
		Catenulisporales		<b>1</b>
		Corynebacteriales		1
	MB-A2-108		<b>6</b>	
	Thermoleophilia	Gaiellales	<b>7</b>	1
		Solirubrobacterales	1	1
<i>Armatimonadetes</i>				1
<i>Bacteroidetes</i>				
	Cytophagia	Cytophagales	1	
	Sphingobacteriia	Sphingobacteriales	6	2
<i>Chlamydiae</i>				
	Chlamydiae	Chlamydiales		1
<i>Chlorobi</i>				
	Chlorobia	Chlorobiales		1
<i>Chloroflexi</i>				
	Unknown		4	<b>3</b>
	Anaerolineae	Anaerolineales	5	1
	Chloroflexia	Chloroflexales	3	
	Gitt-GS-136		<b>1</b>	
	JG37-AG-4		2	
	KD4-96		<b>5</b>	
	Ktedonobacteria	Ktedonobacterales	2	<b>6</b>
		Thermogemmatissporales		<b>1</b>
	S085		1	
	SHA-26		<b>1</b>	
	Elev-1554			<b>1</b>
	TK10		2	
<i>Cyanobacteria</i>				
	Melainabacteria	Obscuribacterales	1	
	Chloroplast	Loliumperenne		1
<i>Firmicutes</i>				
	Bacilli	Bacillales	1	1
<i>Gemmatimonadetes</i>				
	Gemmatimonadetes	Gemmatimonadales	<b>10</b>	
<i>Latescibacteria</i>			6	
<i>Nitrospirae</i>				
	Nitrospira	Nitrospirales	7	
<i>Planctomycetes</i>				
	OM190		1	
	Planctomycetacia	Planctomycetales	3	1
<i>Proteobacteria</i>				
	unknown		2	
	Alphaproteobacteria	Unknown		1
		Rhizobiales	<b>11</b>	3
		Caulobacterales	2	<b>1</b>
		Rhodospirillales	2	
	ARKICE-90	Unknown	<b>1</b>	
	Betaproteobacteria	Unknown	2	
		Burkholderiales	1	
		SC-1-84	5	2
		Nitrosomonadales	8	
	Deltaproteobacteria	Unknown	1	
		Desulfurellales	1	
		Myxococcales	<b>4</b>	3
		GR-WP33-30	9	2
	Gammaproteobacteria			1
		Xanthomonadales	4	3
		Legionellales		1
		Chromatiales		1
		Oceanospirillales		<b>1</b>
		Alteromonadales		<b>2</b>
<i>Saccharibacteria</i>				1
<i>TM6</i>			1	2
<i>Verrucomicrobia</i>				
	OPB35soil group		9	3
	Opiritatae	Opiritatales		<b>2</b>
	S-BQ2-57soil group		2	
	Spartobacteria	Chthoniobacterales	10	1
WD272				1

lipid signature unique to soils within the Icelandic Cluster Cold and Cluster Warm. Based on the PCA analysis, we determine that the change in brGDGTs distribution in Cluster Warm soils is reflected primarily by an increase in the relative amount of brGDGT Ia, compared to that of brGDGTs IIa and IIIa (Fig. 2B). To summarize this increase in a ratio, we define a “Community Index”.

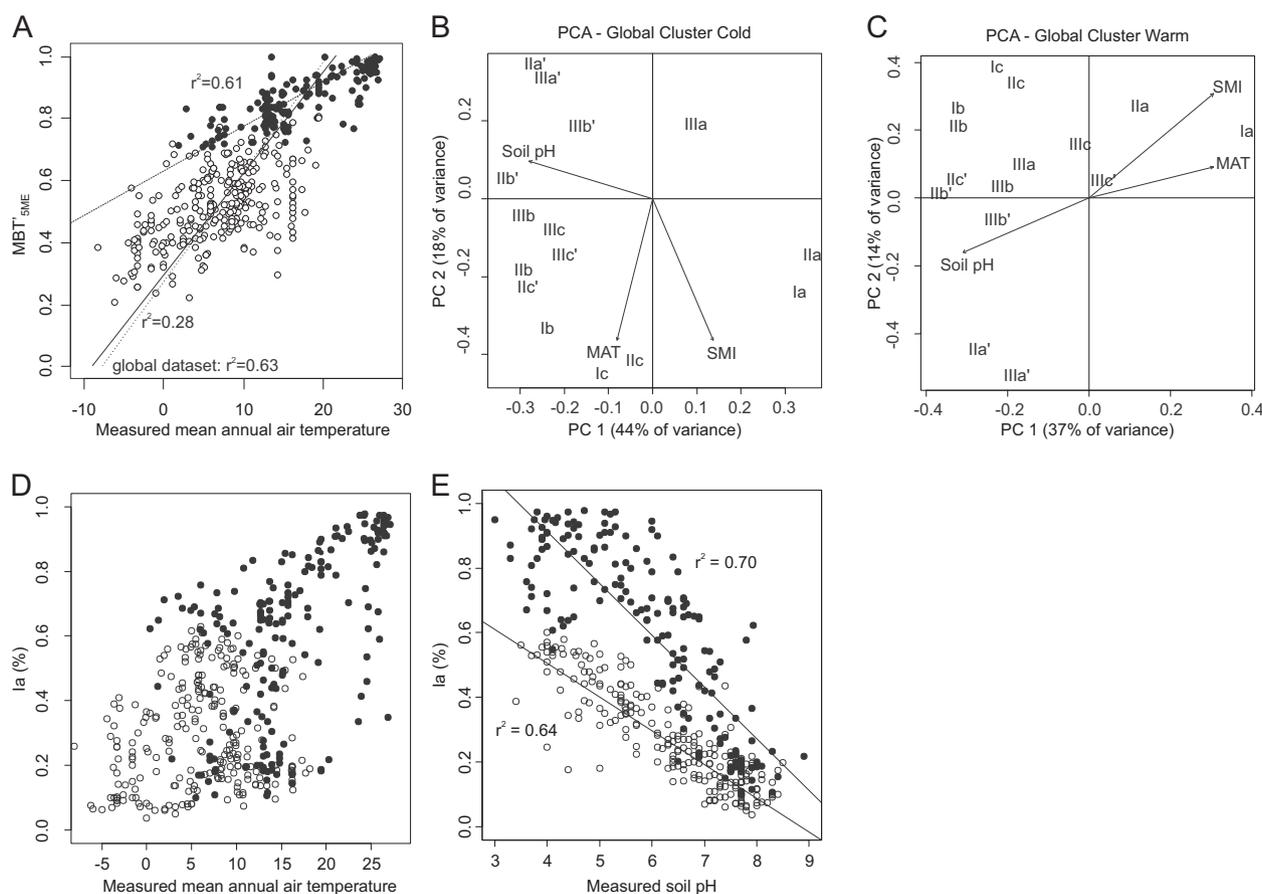
The community index is calculated as:

$$CI = Ia / (Ia + IIa + IIIa) \quad (3)$$

The CI is similar to the  $MBT'_{SME}$  ratio but excludes pH-sensitive compounds, as these plot on a separate principle component at the ForHot site (Fig. 2B), reflecting a different environmental dependency. Based on the variation in CI values along the geothermal warming transect (Fig. 3B), we choose 0.64 as a threshold value to identify two groups on a global scale. This value is based on only one study of geothermally warmed soils ( $n = 30$ ), where the Cluster Warm soils are represented by only five soils and should therefore be interpreted as a first estimate. A sensitivity analysis of this threshold ratio is shown in Supplementary Fig. S2. Subsequently, the CI was applied to previously published datasets (De Jonge et al., 2014; Ding et al., 2015; Xiao et al., 2015; Yang et al., 2015; Lei et al., 2016 [as compiled by Naafs et al., 2017a]; Naafs et al., 2017b). This results in two groups (hereafter referred to as ‘Global Clusters’), a Global Cluster Cold consisting of 251 soils and peats with a large range in measured mean annual air temperature (between  $-8.3$  °C and  $18.2$  °C) and pH values varying between 3.4

and 9.3, and a Global Cluster Warm, containing 195 soils and peats with, on average, higher measured mean annual air temperatures (between  $0.4$  °C and  $27.1$  °C) and pH values varying between 3.0 and 8.9 (Fig. 6A).

The identification of two clusters in the global soil dataset now allows for evaluation of the environmental drivers that explain the variation in brGDGTs within the clusters (Fig. 6B, C). Within both clusters the variation in brGDGTs is mainly controlled by soil pH ( $var_{unique} = 28\%$  and  $16\%$  for Global Cluster Cold and Global Cluster Warm, respectively), with a smaller fraction explained by the mean annual air temperature ( $3\%$  and  $7\%$  for Global Cluster Cold and Global Cluster Warm, respectively; Table 4). Remarkably, a strong cluster-specific correlation between the major brGDGT Ia and soil pH is observed ( $r = -0.86$  and  $-0.83$  for Global Cluster Cold and Global Cluster Warm, respectively; Fig. 6D). The dependency of brGDGT Ia on soil pH is significantly different between Cluster Cold and Cluster Warm soils, with significantly different intercepts ( $p < 0.001$ ) and different slopes between both clusters ( $p < 0.01$ ). The linear model of both clusters are:  $Ia (\%) = 91 - 10 \times pH$  for Global Cluster Cold,  $n = 251$ , and  $Ia (\%) = 156 - 16.1 \times pH$  for Global Cluster Warm,  $n = 195$ . When taking the MAT into account (as Ia also correlates significantly with the MAT in the Global Cluster Warm), the MAT values do not contribute significantly to the model ( $0.05 < p < 0.1$ ), while the dependency between brGDGT Ia and pH is still significantly different (significantly different intercepts ( $p < 0.001$ ) and slopes ( $p < 0.01$ )). The model shows that there



**Fig. 6.** (A) The  $MBT'_{SME}$  values plotted against mean annual air temperature (MAT) for soils in the global soil dataset, with samples of Global Cluster Cold in white and samples of Global Cluster Warm in black, (black lines indicate cluster-specific correlations, the dotted line indicates the correlation for the whole dataset). (B and C) The ordination spaces based on the standardized brGDGT abundances, with environmental parameters soil pH, MAT and modelled soil moisture (SMI) for (B) Global Cluster Cold and (C) Global Cluster Warm. (D and E) The fractional abundance of brGDGT compound Ia, with Global Cluster Cold values in white and Global Cluster Warm values in black, (D) plotted against measured soil pH and (E) against the mean annual air temperature.

**Table 4**  
The proportion of uniquely explained variance of the fractional abundance of the 15 brGDGT compounds in the Global Clusters Cold and Warm by measured soil pH, measured mean annual air temperature (MAT), and modelled soil moisture index (SMI) and in the Local (Icelandic) Clusters Cold and Warm by measured soil pH, measured mean annual soil temperature (MST) and measured mean annual soil moisture (SM). For each cluster and each environmental variable, the Pearson correlation coefficient (r) and the associated p-value of significant correlations with brGDGT compounds is reported. The significance level is at 0.05, except for Local Cluster Warm, where it is 0.10. For each compound the best correlation is indicated in bold.

Global Cluster Cold (n = 251)	Var <sub>Unique</sub>		Ia	Ib	Ic	IIa	IIb	IIc	IIIa	IIIb	IIIc	IIa'	IIb'	IIc'	IIIa'	IIIb'	IIIc'
Soil pH	28%	r	<b>-0.86</b>	<b>0.38</b>		<b>-0.85</b>	<b>0.56</b>			<b>0.61</b>	<b>0.39</b>	<b>0.80</b>	<b>0.71</b>	<b>0.50</b>	<b>0.77</b>	<b>0.34</b>	<b>0.34</b>
		p	<b>0.00</b>	<b>0.00</b>		<b>0.00</b>	<b>0.00</b>			<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
MAT	3%	r	0.15	0.37	<b>0.38</b>			0.11	<b>-0.36</b>	0.16				0.19			<b>-0.15</b>
		p	0.02	0.00	<b>0.00</b>			0.08	<b>0.00</b>	0.01				0.00			0.02
SMI	2%	r	0.40		0.26	0.32		<b>0.22</b>	<b>-0.15</b>		<b>-0.15</b>	<b>-0.47</b>	<b>-0.14</b>		<b>-0.39</b>	<b>-0.15</b>	
		p	0.00		0.00	0.00		<b>0.00</b>	0.04		0.03	0.00	0.05		0.00	0.04	
Global Cluster Warm (n = 195)																	
Soil pH	15%	r	<b>-0.83</b>	<b>0.52</b>	<b>0.21</b>	-0.21	<b>0.53</b>	0.23	0.44	<b>0.32</b>		<b>0.80</b>	<b>0.61</b>	<b>0.46</b>	<b>0.61</b>	<b>0.29</b>	
		p	<b>0.00</b>	<b>0.00</b>	<b>0.01</b>	0.01	<b>0.00</b>	0.00	0.00	<b>0.00</b>		<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>
MAT	7%	r	0.55		0.16	<b>-0.55</b>	-0.28	-0.16	<b>-0.68</b>	-0.14		<b>-0.39</b>	<b>-0.25</b>	<b>-0.18</b>	<b>-0.26</b>	<b>-0.27</b>	
		p	0.00		0.02	<b>0.00</b>	0.00	0.02	<b>0.00</b>	0.05		0.00	0.00	0.01	0.00	0.00	0.00
SMI	1%	r	0.45				-0.19		<b>-0.22</b>			<b>-0.45</b>	<b>-0.18</b>	<b>-0.19</b>	<b>-0.44</b>	<b>-0.15</b>	
		p	0.00				0.01		0.00			0.00	0.02	0.01	0.00	0.04	
Local Cluster Cold (n = 25)																	
Soil pH	40%	r	<b>-0.69</b>	<b>0.80</b>		<b>-0.71</b>	<b>0.79</b>	<b>0.73</b>	-0.40	<b>0.73</b>	<b>0.63</b>	<b>0.73</b>	<b>0.72</b>	<b>0.79</b>	<b>0.43</b>	<b>0.66</b>	<b>0.60</b>
		p	<b>0.00</b>	<b>0.00</b>		<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	0.05	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.00</b>	<b>0.03</b>	<b>0.00</b>
MST	2%	r							<b>-0.53</b>								
		p							<b>0.01</b>								
SM	1%	r						0.46									
		p						0.02									
Local Cluster Warm (n = 5)																	
Soil pH	57%	r										<b>0.84</b>	<b>0.81</b>		<b>0.84</b>		
		p										<b>0.07</b>	<b>0.09</b>		<b>0.07</b>		
MST	39%	r						<b>0.83</b>	<b>0.94</b>								
		p						<b>0.08</b>	<b>0.02</b>								
SM	8%	r				<b>-0.88</b>											
		p				<b>0.05</b>											

is a significant interaction between MAT and soil pH ( $p < 0.05$ ), which means that the slope of soil pH with the brGDGT Ia increases in warmer soils. This fits with our observation in the global soil dataset that there is a different pH-dependency in those soils that are generally warmer (Fig. 6E).

The observation that the soils from both clusters show a different dependency of brGDGT Ia on soil pH cannot be fully explained by the originally proposed mechanism of membrane adaptation of unchanged bacterial communities. Bacterial membrane homeostasis would not require a different contribution of the major compound brGDGT Ia under similar pH conditions between soil clusters, an offset that is especially clear at low pH conditions (Fig. 6D). Instead, the different environmental dependency of the brGDGTs Global Cluster Cold and Global Cluster Warm can be explained by a bacterial community shift, as observed along the ForHot geothermal gradient. However, on a global scale a differential response to pH and MAT of an unchanged bacterial community between these both clusters cannot be ruled out unequivocally without paired DNA data.

The strong effect of pH on brGDGT Ia is unexpected, as this compound was previously interpreted as the main temperature-sensitive brGDGT lipid (Fig. 6E) in publications where soils from both clusters were considered as a single group (Weijers et al., 2007a; De Jonge et al., 2014; Naafs et al., 2017a,b). However, for the Global Cluster Cold, no strong correlation between brGDGT Ia and mean annual air temperature was found ( $r = 0.15$ ; Table 4). Within Cluster Cold, temperature only correlates with the variation

in brGDGTs Ic and IIIa ( $r = 0.38$  and  $-0.36$ , respectively), and these weak correlations drive the slight correlation between  $MBT'_{5ME}$  index values and the mean annual air temperature ( $MAT_{CC} = -8.71 + 30.24 \times MBT'_{5ME}$ ,  $r^2 = 0.28$ ) within the cold cluster soils. Within Global Cluster Warm, several brGDGTs correlate with temperature, resulting in a stronger correlation between  $MBT'_{5ME}$  index values with the mean annual air temperature ( $MAT_{CW} = -28.67 + 53.46 \times MBT'_{5ME}$ ,  $r^2 = 0.61$ ). This is however partly caused by the correlation between pH and mean annual air temperature values in this subset of soils ( $r = -0.37$ ), and only brGDGTs IIa and IIIa show a correlation with temperature that is better than that with soil pH ( $r$  of  $-0.55$  and  $-0.68$ ; Table 4). Furthermore, the different environmental dependencies of the brGDGTs within each cluster result in cluster-specific correlations between the  $MBT'_{5ME}$  values and MAT.

#### 4.4. The bacterial community effect on brGDGTs: implications for paleoclimate reconstruction

In the geothermally warmed soils, we have observed that the  $MBT'_{5ME}$  captures large temperature changes only when the bacterial community shows a strong change in composition (Fig. 3A). Therefore, the presence of two distinct brGDGT-producing bacterial communities can have an effect on paleoclimate reconstructions. The driving factors that cause bacterial community shifts in soils that experience a change in temperature thus need to be better constrained to increase the reliability of brGDGT-based tem-

perature reconstructions. Previous studies have shown that changes in bacterial community composition can already occur after a year-long temperature increase of both +1 °C and +2 °C (Xiong et al., 2014), although a temperature increase does not always cause a community shift on this time scale (Weedon et al., 2017). Other research (DeAngelis et al., 2015) concluded that more than a decade of substantial warming was needed to significantly affect microbial community composition in a temperate forest (5 °C warming) and sub-Arctic heathlands (4 °C warming), respectively. This delayed effect of temperature change on the microbial community composition (Rinnan et al., 2007; DeAngelis et al., 2015) has been attributed to the strong indirect effects of warming on communities via changes to vegetation and soil properties (Classen et al., 2015, Xiong et al., 2016). Also, at the ForHot geothermal gradients, temperature possibly has an indirect effect on the bacterial community shift (Radujković et al., 2018). The indirect effect of temperature on the bacterial community composition can explain why, on a global scale, mean annual air temperature as such does not automatically determine whether a soil contains a Cluster Cold or Warm brGDGT-producing community. Specifically, soils at mean annual air temperatures of 0–18 °C, i.e. mid-latitudes, can be encountered containing a lipid distribution with a typically cold, as well as with a typically warm fingerprint (Fig. 6A).

The implications for a paleoclimate proxy are two-fold. Firstly, the local scale shows that a change in soil temperature strongly affects the brGDGT signature (i.e. between Cluster Cold and Cluster Warm) only after a change in the brGDGT-producing communities. This mechanism can cause the brGDGTs response to temperature change to lag in time, or to only happen when temperature change reaches a threshold. Although our interpretation on a global scale should be interpreted with caution, we hypothesize that the impact of a bacterial community shift can also be observed. Within each soil cluster the MBT<sub>5ME</sub> does vary with the MAT, possibly caused by indirect effects of temperature. A moderate increase in temperature that does not cause the bacterial community to shift (from Global Cluster Cold to Global Cluster Warm), can thus be reconstructed using the MBT<sub>5ME</sub>. However, the MBT<sub>5ME</sub> shows a different response to MAT in Global Cluster Cold and Cluster Warm soils. As the correlation of the MBT<sub>5ME</sub> index values with MAT in Global Cluster Warm soils is distinct from that of the global calibration dataset (Fig. 6A), changes in the MBT<sub>5ME</sub> record possibly reflect larger temperature shifts. This also means that the MBT<sub>5ME</sub> proxy is less sensitive to temperature changes in Global Cluster Warm soils.

The CI index shows potential to reconstruct whether the brGDGT-producing bacterial community in paleoclimate records falls within Cluster Cold or Cluster Warm. However, it is not clear to what extent Cluster Cold and Cluster Warm soils are geographically separated. Therefore, a lake or marine sediment could receive brGDGT lipids from both sources. In these geological records, changing contributions of different communities can influence the MBT<sub>5ME</sub> proxy values.

The stepwise increase of the MBT<sub>5ME</sub> values along a geothermal gradient contrasts with the linear temperature response of the MBT<sub>5ME</sub> index assumed so far. To test for this mechanism at different sites and on a global scale, future efforts should use paired DNA analysis to identify the environmental variables and threshold values that determine the brGDGT variation within and between soil clusters, to develop a robust proxy that is applicable and well-understood across the whole temperature range.

## 5. Conclusions

A local study site was selected specifically to determine the direct effect of temperature on soil brGDGTs lipid distributions.

Here, the MBT<sub>5ME</sub> only responded to the temperature increase after the bacterial community changed. With the insight that the bacterial community shift observed on the local scale is also present on a global scale, we propose extrapolating the Cluster Cold and Cluster Warm lipid signature to the global scale. The different environmental dependency of brGDGTs lipids in each Global Cluster tentatively confirms the presence of two bacterial communities that produce a different brGDGTs fingerprint.

The strong link between the composition of the bacterial community and brGDGT lipids in geothermally heated soils, in combination with the non-uniform response to temperature of brGDGT distributions in Global Cluster Cold and Warm communities, indicates that a change in the community of brGDGT producers, is partly responsible for determining the variability of a brGDGT-based paleothermometer on a global scale. The identification of a bacterial community imprint on this empirical proxy is a crucial step in unraveling the biotic and abiotic drivers of brGDGT signatures in soils, with as ultimate goal to improve the accuracy of brGDGT-based proxies. In the meantime, however, the effect of bacterial community changes on the MBT<sub>5ME</sub> index needs to be taken into account when interpreting brGDGT-based paleotemperature reconstructions.

## Declaration of Competing Interest

None.

## Acknowledgements

We thank Erik Verbruggen for helpful discussions and Dominika Kasjaniuk for laboratory assistance. We also thank two anonymous reviewers for their critical reading and constructive comments on an earlier version of this manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement n° 707270/WISLAS).

## Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.orggeochem.2019.07.006>.

## References

- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, N., Csaki, F. (Eds.), *Proceedings of the Second International Symposium on Information Theory*. Akademiai Kiado, pp. 267–281.
- Bartram, A.K., Jiang, X., Lynch, M.D.J., Masella, A.P., Nicol, G.W., Dushoff, J., Neufeld, J. D., 2014. Exploring links between pH and bacterial community composition in soils from the Craibstone Experimental Farm. *FEMS Microbiology Ecology* 87, 403–415.
- Bolker, B., R Core Team, 2017. *bbmle: Tools for General Maximum Likelihood Estimation*. R Package Version 1, 20.
- Cáceres, M.D., Legendre, P., 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology* 90, 3566–3574.
- Classen, A.T., Sundqvist, M.K., Henning, J.A., Newman, G.S., Moore, J.A.M., Cregger, M.A., Moorhead, L.C., Patterson, C.M., 2015. Direct and indirect effects of climate change on soil microbial and soil microbial-plant interactions: What lies ahead? *Ecosphere* 6, 1–21.
- 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, 24–36.
- DeAngelis, K.M., Pold, G., Topçuoğlu, B.D., Diepen, V.A.L.T., Varney, R.M., Blanchard, J.L., Melillo, J., Frey, S.D., 2015. Long-term forest soil warming alters microbial communities in temperate forest soils. *Frontiers in Microbiology* 6, 104.
- 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 hexamethylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC-MS<sup>2</sup>, GC-MS and GC-SMB-MS. *Organic Geochemistry* 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. *Geochimica et Cosmochimica Acta* 141, 97–112.
- Ding, S., Xu, Y., Wang, Y., He, Y., Hou, J., Chen, L., He, J.-S., 2015. Distribution of branched glycerol dialkyl glycerol tetraethers in surface soils of the Qinghai-Tibetan Plateau: implications of brGDGTs-based proxies in cold and dry regions. *Biogeosciences* 12, 3141–3151.
- 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, 52–58.
- Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* 10, 996–998.
- Hopmans, E.C., Schouten, S., Sinninghe Damsté, J.S., 2016. The effect of improved chromatography on GDGT-based palaeoproxies. *Organic Geochemistry* 93, 1–6.
- Huguet, C., Hopmans, E.C., Febo-Ayala, W., Thompson, D.H., Sinninghe Damsté, J.S., Schouten, S., 2006. An improved method to determine the absolute abundance of glycerol dibiphytanyl glycerol tetraether lipids. *Organic Geochemistry* 37, 1036–1041.
- Huguet, A., Meador, T.B., Laggoun-Défarge, F., Könneke, M., Wu, W., Derenne, S., Hinrichs, K.-U., 2017. Production rates of bacterial tetraether lipids and fatty acids in peatland under varying oxygen concentrations. *Geochimica et Cosmochimica Acta* 203, 103–116.
- Inglis, G.N., Collinson, M.E., Riegel, W., Wilde, V., Farnsworth, A., Lunt, D.J., Valdes, P., Robson, B.E., Scott, A.C., Lenz, O.K., Naafs, B.D.A., Pancost, R.D., 2017. Mid-latitude continental temperatures through the early Eocene in western Europe. *Earth and Planetary Science Letters* 460, 86–96.
- 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, 83–94.
- Naafs, B.D.A., Gallego-Sala, A.V., Inglis, G.N., Pancost, R.D., 2017a. Refining the global branched glycerol dialkyl glycerol tetraether (brGDGT) soil temperature calibration. *Organic Geochemistry* 106, 48–56.
- 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., 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., Marchant, R., McClymont, E., Pontevedra-Pombal, X., Ponton, C., Pourmand, A., Rizzuti, A.M., Rochefort, L., Schellekens, J., De Vleeschouwer, F., Pancost, R.D., 2017b. Introducing global peat-specific temperature and pH calibrations based on brGDGT bacterial lipids. *Geochimica et Cosmochimica Acta* 208, 285–301.
- O'Gorman, E.J., Benstead, J.P., Cross, W.F., Friberg, N., Hood, J.M., Johnson, P.W., Sigurdsson, B.D., Woodward, G., 2014. Climate change and geothermal ecosystems: natural laboratories, sentinel systems, and future refugia. *Global Change Biology* 20, 3291–3299.
- Oliverio, A.M., Bradford, M.A., Fierer, N., 2017. Identifying the microbial taxa that consistently respond to soil warming across time and space. *Global Change Biology* 23, 2117–2129.
- Oppermann, B.I., Michaelis, W., Blumenberg, M., Frerichs, J., Schulz, H.M., Schippers, A., Beaubien, S.E., Krüger, M., 2010. Soil microbial community changes as a result of long-term exposure to a natural CO<sub>2</sub> vent. *Geochimica et Cosmochimica Acta* 74, 2697–2716.
- Pancost, R.D., Sinninghe Damsté, J.S., 2003. Carbon isotopic compositions of prokaryotic lipids as tracers of carbon cycling in diverse settings. *Chemical Geology* 195, 29–58.
- Peterse, F., Schouten, S., van der Meer, J., van der Meer, M.T.J., Sinninghe Damsté, J.S., 2009. Distribution of branched tetraether lipids in geothermally heated soils: Implications for the MBT/CBT temperature proxy. *Organic Geochemistry* 40, 201–205.
- 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, 1171–1175.
- Peterse, F., Prins, M.A., Beets, C.J., Troelstra, S.R., Zheng, H., Gu, Z., Schouten, S., Sinninghe Damsté, J.S., 2011. Decoupled warming and monsoon precipitation in East Asia over the last deglaciation. *Earth and Planetary Science Letters* 301, 256–264.
- 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. *Geochimica et Cosmochimica Acta* 96, 215–229.
- Pitcher, A., Hopmans, E.C., Schouten, S., Sinninghe Damsté, J.S., 2009. Separation of core and intact polar archaeal tetraether lipids using silica columns: Insights into living and fossil biomass contributions. *Organic Geochemistry* 40, 12–19.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41, D590–D596.
- R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Radujković, D., Verbruggen, E., Sigurdsson, B.D., Leblans, N.I.W., Janssens, I.A., Vicca, S., Weedon, J.T., 2018. Prolonged exposure does not increase soil microbial community compositional response to warming along geothermal gradients. *FEMS Microbial Ecology* 94. <https://doi.org/10.1093/femsec/fix174>.
- Rinnan, R., Michelsen, A., Bååth, E., Jonasson, S., 2007. Fifteen years of climate change manipulations alter soil microbial communities in a subarctic heath ecosystem. *Global Change Biology* 13, 28–39.
- Schouten, S., Hopmans, E.C., Sinninghe Damsté, J.S., 2013. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review. *Organic Geochemistry* 54, 19–61.
- Sigurdsson, B.D., Leblans, N.I.W., Dauwe, S., Gudmundsdóttir, E., Gundersen, P., Gunnarsdóttir, G.E., Holmstrup, M., Ilieva-Makulec, K., Katterer, T., Marteinsdóttir, B.-S., Maljanen, M., Oddsdóttir, E.S., Ostonen, I., Penuelas, J., Poeplau, C., Richter, A., Sigurdsson, P., Van Bodegom, P., Wallander, H., Weedon, J., Janssens, I., 2016. Geothermal ecosystems as natural climate change experiments: the ForHot research site in Iceland as a case study. *Icelandic Agricultural Sciences* 29, 53–71.
- Sinninghe Damsté, J.S., Hopmans, E.C., Pancost, R.D., Schouten, S., Geenevasen, J.A.J., 2000. Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments. *Chemical Communications* 17, 1683–1684.
- Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Weijers, J.W.H., Foessel, 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, 4147–4154.
- Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Foessel, B.U., Wüst, P.K., Overmann, J., Tank, M., Bryant, D.A., Dunfield, P.F., Houghton, K., Stott, M.B., 2014. Ether- and ester-bound *iso*-diabolic acid and other lipids in members of *Acidobacteria* subdivision 4. *Applied and Environmental Microbiology* 80, 5207–5218.
- Sinninghe Damsté, J.S., Rijpstra, W.I.C., Foessel, B.U., Huber, K.J., Overmann, J., Nakagawa, S., Kim, J.J., Dunfield, P.F., Dedysh, S.N., Villanueva, L., 2018. An overview of the occurrence of ether- and ester-linked *iso*-diabolic acid membrane lipids in microbial cultures of the *Acidobacteria*: Implications for brGDGT paleoproxies for temperature and pH. *Organic Geochemistry* 124, 63–76.
- Thrash, J.C., Coates, J.D., 2010. Family I. *Acidobacteriaceae* fam. nov. In: Krieg, N.R., Staley, J.T., Brown, D.R., Hedlund, B.P., Paster, B.J., Ward, N.L., Ludwig, W., Whitman, W.B. (Eds.), *Bergey's Manual of Systematic Bacteriology*, 2nd ed., 4, Springer, New York.
- Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new Bacterial taxonomy. *Applied and Environmental Microbiology* 73, 5261–5267.
- Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58, 236–244.
- Ward, N.L., Challacombe, J.F., Janssen, P.H., Henrissat, B., Coutinho, P.M., Wu, M., Xie, G., Haft, D.H., Sait, M., Badger, J., Barabote, R.D., Bradley, B., Brettin, T.S., Brinkac, L.M., Bruce, D., Creasy, T., Daugherty, S.C., Davidsen, T.M., DeBoy, R.T., Detter, J. C., Dodson, R.J., Durkin, A.S., Ganapathy, A., Gwinn-Giglio, M., Han, C.S., Khouri, H., Kiss, H., Kothari, S.P., Madupu, R., Nelson, K.E., Nelson, W.C., Paulsen, I., Penn, K., Ren, Q., Rosovitz, M.J., Selengut, J.D., Shrivastava, S., Sullivan, S.A., Tapia, R., Thompson, L.S., Watkins, K.L., Yang, Q., Yu, C., Zafar, N., Zhou, L., Kuske, C.R., 2009. Three genomes from the phylum *Acidobacteria* provide insight into the lifestyles of these microorganisms in soils. *Applied and Environmental Microbiology* 75, 2046–2056.
- 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, 118–129.
- Weedon, J.T., Kowalchuk, G.A., Aerts, R., Frerichs, S., Rölting, W.F.M., van Bodegom, P. M., 2017. Compositional stability of the Bacterial community in a climate-sensitive sub-Arctic peatland. *Frontiers in Microbiology* 8, 317.
- Weijers, J.W.H., Schouten, S., Hopmans, E.C., Geenevasen, J.A.J., David, O.R.P., Coleman, J.M., Pancost, R.D., Sinninghe Damsté, J.S., 2006. Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits. *Environmental Microbiology* 8, 648–657.
- 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. *Geochimica et Cosmochimica Acta* 71, 703–713.
- Weijers, J.W.H., Schefuss, E., Schouten, S., Sinninghe Damsté, J.S., 2007b. Coupled thermal and hydrological evolution of tropical Africa over the last deglaciation. *Science* 315, 1701–1704.
- Weijers, J.W.H., Panoto, E., van Bleijswijk, J., Schouten, S., Rijpstra, W.I.C., Balk, M., Stams, A.J.M., Sinninghe Damsté, J.S., 2009. Constraints on the biological source (s) of the orphan branched tetraether membrane lipids. *Geomicrobiology Journal* 26, 402–414.
- 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.
- Woodward, G., Dybbjær, J.B., Olafsson, J.S., Gíslason, G.M., Hannesdóttir, E.R., Friberg, N., 2010. Sentinel systems on the razer's edge: effects of warming on Arctic geothermal stream ecosystems: impacts of warming on Icelandic stream ecosystems. *Global Change Biology* 16, 1979–1991.
- 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, 56–60.

- Xiong, J., Sun, H., Peng, F., Zhang, H., Xue, X., Gibbons, S.M., Gilbert, J.A., Chu, H., 2014. Characterizing changes in soil bacterial community structure in response to short-term warming. *FEMS Microbiological Ecology* 89, 281–292.
- Xiong, Q., Pan, K., Zhang, L., Wang, Y., Li, W., He, X., Luo, H., 2016. Warming and nitrogen deposition are interactive in shaping surface soil microbial communities near the alpine timberline zone on the eastern Qinghai-Tibet Plateau, southwestern China. *Applied Soil Ecology* 101, 72–83.
- 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<sup>+</sup>) in soils from an altitudinal transect at Mount Shennongjia. *Organic Geochemistry* 82, 42–53.