

1 **Refining the global branched glycerol dialkyl glycerol tetraether (brGDGT) soil**  
2 **temperature calibration**

3  
4 B.D.A. Naafs<sup>1\*</sup>, A.V. Gallego-Sala<sup>2</sup>, G.N. Inglis<sup>1</sup>, and R.D. Pancost<sup>1</sup>

5  
6 <sup>1</sup>Organic Geochemistry Unit, School of Chemistry and Cabot Institute, University of  
7 Bristol, Bristol, UK

8 <sup>2</sup>Geography, College of Life and Environmental Sciences, University of Exeter,  
9 Exeter, UK

10  
11 \*Corresponding author. E-mail address: [david.naafs@bristol.ac.uk](mailto:david.naafs@bristol.ac.uk)

12  
13 **Abstract**

14 Branched glycerol dialkyl glycerol tetraethers (brGDGTs) are increasingly used to  
15 reconstruct past terrestrial temperature and soil pH. Here we compare all available  
16 modern soil brGDGT data (n=350) to a wide range of environmental parameters to  
17 obtain new global temperature calibrations.

18 We show that soil moisture index (MI), a modeled parameter that also takes  
19 potential evapotranspiration into account, is correlated to the 6-methyl brGDGT  
20 distribution but does not significantly control the distribution of 5-methyl brGDGTs.  
21 Instead, temperature remains the primary control on 5-methyl brGDGTs. We propose  
22 the following global calibrations:  $MAAT_{soil} = 40.01 \times MBT'_{5me} - 15.25$  (n=350,  $R^2 =$   
23  $0.60$ , RMSE = 5.3 °C) and growing degree days above freezing ( $GDD_{0\ soil} = 14344.3$   
24  $\times MBT'_{5me} - 4997.5$  (n=350,  $R^2 = 0.63$ , RMSE = 1779 °C).

25 Recent studies have suggested that factors other than temperature can impact  
26 arid and/or alkaline soils dominated by 6-methyl brGDGTs. As such, we develop new  
27 global temperature calibrations using samples dominated by 5-methyl brGDGTs only  
28 ( $IR_{6me} < 0.5$ ). These new calibrations have significantly improved correlation  
29 coefficients and lower root mean square errors (RMSE) compared to the global  
30 calibrations:  $MAAT_{soil}' = 39.09 \times MBT'_{5me} - 14.50$  (n=177,  $R^2 = 0.76$ , RMSE =  
31  $4.1$  °C) and  $GDD_{0\ soil}' = 13498.8 \times MBT'_{5me} - 4444.5$  (n=177,  $R^2 = 0.78$ , RMSE =  
32  $1326$ ). We suggest that these new calibrations should be used to reconstruct terrestrial

33 climate in the geological past; however, care should be taken when employing these  
34 calibrations outside the modern calibration range.

35

## 36 **1. Introduction**

37 Branched glycerol dialkyl glycerol tetraethers (brGDGTs) are membrane-spanning  
38 lipids produced by bacteria, presumably acidobacteria (Weijers et al., 2009; Sinninghe  
39 Damsté et al., 2011). First discovered in a Dutch peat (Sinninghe Damsté et al., 2000),  
40 brGDGTs are ubiquitous in mesophilic settings (Schouten et al., 2013) such as soils  
41 (Weijers et al., 2006b; Peterse et al., 2012), lakes (Pearson et al., 2011; Schoon et al.,  
42 2013; Li et al., 2016), rivers (De Jonge et al., 2014b), marine sediments (Hopmans et  
43 al., 2004; Fietz et al., 2012), and peat deposits (Weijers et al., 2006a; Huguet et al.,  
44 2010; Zheng et al., 2015; Naafs et al., 2017). Recent advances in analytical methods  
45 (De Jonge et al., 2013; Yang et al., 2015; Hopmans et al., 2016) have revealed the  
46 existence of a wide range of brGDGTs in mineral soils (Fig. S1), varying in the  
47 number of methyl branches (between 4 and 6), the carbon position of methyl branches  
48 (C5 and C6 position), and number of cyclopentane moieties (between 0 and 2).

49 Over the last decade brGDGTs have become of great interest to the organic  
50 geochemistry and paleoclimate communities because their distribution (degree of  
51 cyclisation and methylation) correlates with soil-pH and mean annual air temperature.  
52 This was originally expressed by Weijers et al. (2007b) in the MBT/CBT-proxy in a  
53 global soil dataset and redefined by Peterse et al. (2012) with the MBT'/CBT-proxy:

$$(1) MBT' = \frac{Ia + Ib + Ic}{Ia + Ib + Ic + IIa + IIa' + IIb + IIb' + IIc + IIc' + IIIa + IIIa'}$$

$$(2) CBT = -\log\left(\frac{Ib + IIb + IIb'}{Ia + IIa + IIa'}\right)$$

$$(3) MAT (\text{°C}) = 0.81 - 5.67 \times CBT + 31.0 \times MBT' \quad (n = 176, R^2 = 0.59, RSME = 5.0 \text{ °C})$$

54 Consequently, the brGDGT-based MBT(′)/CBT-proxy has been increasingly applied  
55 to proximal marine and lake sediments as well as loess and paleosols to gain insights  
56 into past terrestrial temperatures (Weijers et al., 2007a; Pancost et al., 2013; Schouten  
57 et al., 2013; Peterse et al., 2014; Lu et al., 2016).

58 De Jonge et al. (2013; 2014a) recently demonstrated that 5-methyl penta- and  
59 hexamethylated brGDGTs, used to calculate the CBT and MBT(′) indices, co-elute  
60 with newly identified 6-methyl brGDGTs. Re-evaluation of the global soil calibration

61 dataset in this context removed the pH dependence upon the degree of methylation of  
62 brGDGTs (De Jonge et al., 2014a) and suggested that the abundance of 6-methyl  
63 brGDGTs is influenced predominantly by pH (De Jonge et al., 2014a; Xiao et al.,  
64 2015). Excluding the 6-methyl brGDGTs from the regressions, De Jonge et al.  
65 (2014a) developed two new types of equations with a dependence of 5-methyl  
66 brGDGTs on temperature alone; one based on the degree of methylation of 5-methyl  
67 branched tetraethers ( $MBT'_{5ME}$ ):

$$(4) MBT'_{5ME} = \frac{(Ia + Ib + Ic)}{(Ia + Ib + Ic + IIa + IIb + IIc + IIIa)}$$

$$(5) MAT = -8.57 + 31.45 \times MBT'_{5ME} \quad (n = 231, R^2 = 0.64, RMSE = 4.9 \text{ } ^\circ C)$$

68 And another based on a multiple linear regression using the relative abundance of  
69 specific 5-methyl brGDGTs ( $MAT_{mr}$ ).

$$(6) MAT_{mr} = 7.17 + 17.1 \times \{Ia\} + 25.9 \times \{Ib\} + 34.4 \times \{Ic\} - 28.6 \times \{IIa\}$$
$$(n = 231, R^2 = 0.67, RMSE = 4.7 \text{ } ^\circ C)$$

70 (Note that the calibration statistics ( $n$ ,  $R^2$ , and RMSE) given in De Jonge et al. (2014a)  
71 were recently corrected (De Jonge et al., 2016)).

72 Although the latest calibrations improved the correlation coefficient and root  
73 mean square error (RMSE) of brGDGT temperature and pH proxies, the relatively  
74 large scatter in the global calibration indicates the potential influence of additional  
75 environmental parameters, such as precipitation and soil moisture content (SMC) on  
76 brGDGT distributions (Weijers et al., 2011; Anderson et al., 2014; Ding et al., 2015;  
77 Xiao et al., 2015; Yang et al., 2015; Dang et al., 2016; Lei et al., 2016). For example,  
78 Dang et al. (2016) recently showed that the soil brGDGT distribution changed along a  
79 short transect with varying soil moisture content.

80 Although previous studies have argued that mean annual temperature (MAT),  
81 soil pH, and precipitation (MAP) are the key environmental parameters influencing  
82 brGDGT distributions in the global soil database, the influences of other parameters  
83 such as growing degree days (GDD) have not been considered in the context of the  
84 global dataset. Further more, although some studies have investigated the impact of  
85 soil moisture on the brGDGT distribution at a regional level (Dirghangi et al., 2013;  
86 Dang et al., 2016), this has not been studied in a global context. This is important as  
87 both soil moisture and GDD are potentially better indicators of the growth  
88 temperature and moisture content experienced by brGDGT-producing bacteria living  
89 in soils (McMaster and Wilhelm, 1997; Gallego-Sala et al., 2010). An additional

90 source of error can derive from the instrumental temperature data used in calibrations,  
91 because this comprises a mix of global databases and local weather station data and  
92 potentially there is an offset between air temperature and soil temperature (Weijers et  
93 al., 2007b; Peterse et al., 2012). In addition, the existing mineral soil calibrations  
94 indicate a bias in the coldest soils with brGDGT-based temperatures up to 15 °C  
95 higher than observed mean annual air temperature (De Jonge et al., 2014a).

96 Here we compile and revisit all available soil brGDGTs data from around the  
97 world and calibrate these against a range of environmental parameters obtained from a  
98 simple bioclimatic model (PeatStash) that calculates bioclimatic variables using long-  
99 term mean monthly values of temperature, precipitation and the fraction of possible  
100 sunshine hours (Gallego-Sala et al., 2010). We use this to assess the environmental  
101 controls on brGDGTs in soils and re-define the global soil temperature proxies.  
102 Although several studies advocate the use of local calibrations (e.g., Ding et al., 2015;  
103 Yang et al., 2015), the earth's climate system was significantly different during the  
104 geological past (e.g., during the Eocene). As such, the application of local calibrations  
105 should remain limited to recent (i.e. Quaternary) sediments when environmental  
106 conditions were likely similar to those covered in the local calibration dataset. For  
107 deep time application, global calibrations are required as these incorporate all  
108 modern-day climate zones.

109

## 110 **2. Material and methods**

### 111 *2.1. Material*

112 We use the distribution of brGDGTs in the global soil dataset compiled by De Jonge  
113 et al., (2014a), based on a sample set generated previously (Weijers et al., 2007b;  
114 Peterse et al., 2012). This is supplemented with data from Chinese soils (Ding et al.,  
115 2015; Xiao et al., 2015; Yang et al., 2015; Lei et al., 2016). Although other data sets  
116 containing modern brGDGT distributions in soils exist, these do not separate the 5-  
117 and 6-methyl brGDGTs and were therefore not included. The revised dataset from De  
118 Jonge et al., (2014a) consists of 239 samples from across the globe. We exclude 13  
119 samples either because the altitude of the soil sample was unknown (Peterse, personal  
120 communication December 2015) or because the altitude in PeatStash for that location  
121 was significantly different compared to that of the soil sample and no altitude  
122 correction could be made. Combined with 27 samples from the Qinghai–Tibetan  
123 Plateau (Ding et al., 2015), 27 samples from across the 400 mm isoline of mean



124 annual precipitation in China (Xiao et al., 2015), 26 samples from Mt. Shennongjia in  
125 China (Yang et al., 2015), and 44 samples from the Henan and Yunnan provinces in  
126 China (Lei et al., 2016), we use a total of 350 soil samples (Fig. 1). The global dataset  
127 consists of data measured in different laboratories. Although there are no data  
128 available for the interlaboratory variation of soil brGDGT-based indices, compiling  
129 data from different laboratories could introduce additional variation.

130

## 131 *2.2. Environmental parameters*

132 We used bioclimatic variables obtained from a simple bioclimatic model (PeatStash).  
133 PeatStash calculates these variables globally with a 0.5 degree spatial resolution  
134 (Gallego-Sala and Prentice, 2013). The calculations are based on long-term mean  
135 climatology data, obtained by interpolating long-term mean weather station  
136 climatology (temperature, precipitation and the fraction of possible sunshine hours)  
137 from around the world for the period 1931-1960 (Climate 2.2 available online  
138 <http://www.pik-potsdam.de/~cramer/climate.html>). If the altitude of the grid cell was  
139 significantly different ( $> 250$  meter) from that reported for a soil sample, a nearby  
140 grid cell with an altitude difference  $< 250$  m was used for comparison. For soil  
141 altitude transects (e.g., Peterse et al., 2009), we generate temperature transects using  
142 PeatStash, but these could not be used to calculate precipitation or moisture index  
143 across transects. The temperature dataset thus consists of 350 soils, whereas the  
144 dataset for MAP and moisture index consists of 275 soils.

145 brGDGT distributions were compared to the following climatological data,  
146 obtained using PeatStash (Gallego-Sala et al., 2010; Gallego-Sala and Prentice, 2013):  
147 mean annual air temperature (MAAT), mean warmest month temperature (MWMT),  
148 mean annual precipitation (MAP), moisture index (MI), and growing degree days  
149 above 0 °C (GDD<sub>0</sub>). The MI is defined as annual precipitation over annual potential  
150 evapotranspiration (P/PET) (Gallego-Sala et al., 2010). MI provides a better measure  
151 of water availability compared to MAP as it takes into account the large difference in  
152 evaporative demand between different climate regimes. Values  $< 1$  are indicative of  
153 dry soils whereas values  $> 3$  are encountered in the wettest areas on earth. GDD<sub>0</sub> is  
154 defined as the yearly cumulated daily average temperature of the daily maximum and  
155 minimum temperature for average temperatures  $> 0$  °C. GDD<sub>0</sub> is a measure of annual  
156 soil heat accumulation and widely used to predict the timing of biological processes  
157 (Kaplan et al., 2003). A high value is indicative of a (sub)tropical climate and a low

158 value for polar/tundra climates. At MAAT > 15 °C, GGD<sub>0</sub> is linearly correlated with  
159 MAAT following 365 (the number of days in a year) multiplied by MAAT (the mean  
160 annual air temperature at a given location).

161

### 162 2.3. Statistical methods

163 Instead of simple linear regression, we use Deming regressions. The advantage of  
164 Deming regressions is that they account for error in both x and y-axis, meaning both  
165 the proxy (e.g. MBT<sub>5me</sub>') and environmental parameter (e.g., MAAT) (Adcock,  
166 1878). For this purpose we used RStudio (RStudio Team, 2015) and the Method  
167 Comparison Regression (MCR) package (Manuilova et al., 2014), which are freely  
168 available to download<sup>1</sup>. The Rscript and data are available in the supplements for  
169 future users. The errors associated with the proxy measurements (e.g. MBT<sub>5me</sub>') and  
170 environmental parameters (e.g., MAAT) are independent and assumed to be normally  
171 distributed. In order to calculate the ratio of their variances (δ), needed to calculate a  
172 Deming regression, we assumed that the standard deviation (σ) of the environmental  
173 data MAAT, GDD<sub>0</sub>, and pH are 1.5 °C, 547.5 °C (365 x 1.5 °C) and 0.25,  
174 respectively. For the brGDGT-based proxies (e.g. MBT<sub>5me</sub>') we assumed a σ of 0.05.  
175 This results in a δ of 0.0011 for the MBT<sub>5me</sub>'/MAAT calibration, 8.3 x 10<sup>-7</sup> for the  
176 MBT<sub>5me</sub>'/GDD<sub>0</sub> calibration, and 0.04 for the pH calibrations (see supplementary  
177 information), respectively. Residuals were calculated using

$$(7) \text{Residual}_y = y_{\text{observed}} - y_{\text{predicted}}$$

178 The root mean square error (RMSE) for y, the predictive error for the  
179 environmental parameter of interests (e.g., MAAT), was calculated using

$$(8) \text{RSME}_y = \sqrt{\frac{\sum_{x=1}^n (y_{x,\text{observed}} - y_{x,\text{predicted}})^2}{n}} \times \frac{n}{df}$$

180 Where *df* stands for degrees of freedom, which in this case is n-1.

181

## 182 3. Results and Discussion

183 Previous studies suggested that the distribution of brGDGTs in soils is controlled  
184 predominantly by MAAT and soil pH (Weijers et al., 2007b; Peterse et al., 2012; De  
185 Jonge et al., 2014a). Other studies have explicitly focused on the pH dependence of  
186 brGDGTs in the global data set, showing that the relative abundance of 6-methyl

---

<sup>1</sup> <https://www.rstudio.com> and <https://cran.r-project.org/web/packages/mcr/index.html>

187 brGDGTs is positively correlated to pH (Ding et al., 2015; Xiao et al., 2015). Guided  
188 by these results we focus on the influence of a range of environmental parameters on  
189 the fractional abundance of brGDGTs.

190

### 191 *3.1 brGDGTs versus temperature*

192 When the fractional abundances of brGDGTs from all 350 samples are plotted versus  
193 mean annual air temperature (MAAT), it is clear that only 5-methyl brGDGTs lacking  
194 cyclopentane moieties (i.e. brGDGT-Ia, -IIa, and -IIIa) are significantly ( $R^2 > 0.2$ ,  
195  $p < 0.001$ ) correlated to MAAT (Fig. 2). brGDGT-Ia is positively correlated with  
196 MAAT ( $R^2 = 0.38$ ,  $p < 0.001$ ), whereas brGDGT-IIa ( $R^2 = 0.20$ ,  $p < 0.001$ ) and -IIIa ( $R^2$   
197  $= 0.36$ ,  $p < 0.001$ ) are negatively correlated with MAAT. These values ( $R^2$  of 0.38,  
198 0.20 and 0.36) are similar to those reported by De Jonge et al. (2014a) for these  
199 compounds with  $R^2$  of 0.34, 0.43, and 0.43, respectively, although the dataset used  
200 here is larger. The correlation between brGDGT-IIa and MAAT is lower in the dataset  
201 used here. Together these results confirm the fundamental dependence of brGDGT  
202 distributions on temperature with lower temperatures associated with a higher degree  
203 of methylation. This was originally proposed by Weijers et al. (2007b) who argued  
204 that additional methyl groups result in a more loose packing of brGDGTs, allowing  
205 bacteria to maintain membrane fluidity at lower temperature, similar to what is seen in  
206 fatty acids synthesized by bacteria (Russell, 1984). 6-methyl brGDGTs and brGDGTs  
207 containing cyclopentane moieties are not significantly correlated to MAAT ( $R^2 <$   
208 0.11).

209

### 210 *3.2 brGDGTs versus precipitation and soil moisture index*

211 As observed by Weijers et al. (2007b), the fractional abundances of several brGDGTs  
212 are also significantly correlated to mean annual precipitation (MAP). In our data set  
213 ( $n = 275$ ) the highest correlations are found for the relative abundance of 5-methyl  
214 brGDGT-Ia ( $R^2 = 0.48$ ,  $p < 0.001$ ) and -IIIa ( $R^2 = 0.21$ ,  $p < 0.001$ ), as well as 6-methyl  
215 brGDGT-IIa' ( $R^2 = 0.35$ ,  $p < 0.001$ ) and -IIIa' ( $R^2 = 0.24$ ,  $p < 0.001$ ) (Fig. 2).

216 Intriguingly, the correlation of brGDGT-Ia with MAP ( $R^2 = 0.48$ ) is higher than that  
217 with MAAT ( $R^2 = 0.38$ ).

218 Crucially, MAP does not reflect the soil moisture content experienced by soil  
219 bacteria as the latter also depends on evaporation and transpiration. To explore this  
220 further, we compared the brGDGTs distribution to the soil moisture index calculated

221 by PeatStash (Fig. 2), although we want to stress that this approach does not take into  
222 local variations (e.g. microtopography, etc) that might be important in determining  
223 soil moisture content for a given sample. A soil moisture index of  $< 1$  indicates that  
224 the potential annual evapotranspiration (the combined effect of evaporation and  
225 transpiration) is higher than annual precipitation and hence a dry soil, whereas a value  
226  $> 1$  indicates the opposite and suggests a wet soil. The correlation of brGDGT-Ia with  
227 soil moisture index is significantly lower ( $R^2 = 0.26, p < 0.001$ ) than that observed for  
228 MAP ( $R^2 = 0.48, p < 0.001$ ), and the correlation for brGDGT-IIIa drops below 0.2.  
229 Therefore, we suggest that the high correlation between MAP and 5-methyl  
230 brGDGTs-Ia is partly due to the correlation between MAP and MAAT ( $R^2 = 0.51$  for  
231 this dataset). In contrast, for 6-methyl brGDGTs (IIa' and IIIa'), the correlation with  
232 MI is 0.29 and 0.20, respectively ( $p < 0.001$ ), and similar to that of MAP. Taken  
233 together this suggests a control of moisture content on the abundance of 6-methyl  
234 brGDGTs in soils. These results are supported by a recent study from Dang et al.  
235 (2016) who demonstrated that the brGDGT distribution in Chinese soils depends on  
236 soil moisture content with a higher amount of brGDGT-IIa' and -IIIa' in dry mineral  
237 soils compared to wet mineral soils.

238

### 239 *3.3 brGDGTs versus warmest mean month temperature*

240 Seasonality is not generally considered to impact the brGDGT distribution in soils,  
241 but there is only limited data to support this. A study of several mid-latitude soils  
242 argued against seasonal changes in brGDGT production as: 1) MBT/CBT-derived  
243 temperatures yield similar temperature estimates throughout the year, and 2) the  
244 concentration of brGDGTs remained constant through the year, indicating a slow  
245 turnover time of brGDGT-producing bacteria on the order of  $\sim 20$  years (Weijers et  
246 al., 2011). However, these results do not preclude a systematic bias in brGDGT  
247 production (and therefore environmental influence) towards a particular season in  
248 high-latitude regions, most likely the warm season as bacterial growth is greater at  
249 higher temperature. Indeed, much higher turnover rates ( $< 2$  years) and a bias in  
250 brGDGT distribution towards the warmest month of the year were recently inferred in  
251 a French peatland (Huguet et al., 2013).

252 To examine the influence of warm season temperature on GDGT distributions  
253 we compared the brGDGT distribution to mean warmest month temperature  
254 (MWMT). The aim was to investigate whether there is a better correlation with

255 MWMT compared to that observed for MAAT (Fig. 3). However, the correlation  
256 coefficients for MWMT (e.g.  $R^2$  for brGDGT-Ia is 0.22) are overall lower than those  
257 for MAAT ( $R^2$  for brGDGT-Ia is 0.38), which would suggest that on a global basis  
258 there is no bias towards the warm season in brGDGT distribution. The same results  
259 are obtained when only samples from soils with MAAT < 5 °C are used (Fig. 3). This  
260 is further supported by a recent study that reported no difference in brGDGT  
261 distribution between Chinese soils with contrasting seasonality (Lei et al., 2016).

262

### 263 *3.4 brGDGTs versus growing degree days above freezing*

264 Although our results do not indicate a global bias in seasonality, the season of  
265 brGDGT production could still be dependent upon latitude. For example, bacteria in a  
266 tropical and temperate soil are likely to grow throughout the year with temperatures  
267 always above freezing, whereas those in a high latitude soil could be heavily biased to  
268 those months when soil temperatures are above zero. Indeed, it is hard to envision that  
269 in the high-latitudes, where winters are characterized by subzero temperatures, that  
270 brGDGTs are produced in equal amounts throughout the year. To explore this further,  
271 we compared the brGDGTs distributions to growing degree days above freezing  
272 ( $GDD_0$ ), a measure of the cumulative temperature (in °C) above zero a soil  
273 experiences over the year.

274 MAAT has previously been considered to reflect the temperature that soil  
275 bacteria experience. However, this does not reflect the time and intensity at which a  
276 given soil remains above freezing. In temperate and polar climates with MAAT < 15  
277 °C,  $GDD_0$  is more indicative of the cumulative heat a soil experiences. For example,  
278 soils from central Kazakhstan and Newfoundland (at 49 °N latitude) both experience  
279 a MAAT of ~3.5 °C. However, Kazakhstan is characterized by a continental climate  
280 with extremely cold winters and hot summers, whereas Newfoundland has a maritime  
281 climate with much less extreme seasonal variation.  $GDD_0$  differentiates the two  
282 climates as the warm summers in Kazakhstan lead to a  $GDD_0$  for this region of  
283 around 2890 °C (cumulative degrees centigrade above zero over one year), much  
284 higher than the value of 1930 °C for Newfoundland.

285 As with MAAT, when  $GDD_0$  is compared to the brGDGT distribution only  
286 5-methyl brGDGTs lacking cyclopentane moieties (i.e. brGDGT-Ia, -IIa, - and IIa)  
287 are significantly ( $R^2 > 0.2$ ) correlated (Fig. 3). brGDGT-Ia is positively correlated  
288 with  $GDD_0$  ( $R^2 = 0.42$ ,  $p < 0.001$ ), whereas brGDGT-IIa ( $R^2 = 0.24$ ,  $p < 0.001$ ) and -IIIa

289 ( $R^2 = 0.38, p < 0.001$ ) are negatively correlated. These  $R^2$  values are slightly higher  
290 than those found for MAAT and higher than those found for MWMT (Fig. 2).

291

### 292 3.5 Redefining temperature calibrations using $MBT'_{5me}$ and Deming regressions

293 Our results confirm that the degree of methylation is significantly correlated with  
294 temperature, either using MAAT or  $GDD_0$ . Following previous studies (De Jonge et  
295 al., 2014a; Ding et al., 2015) we calculated the modified methylation index of 5-  
296 methyl branched tetraethers  $MBT'_{5me}$  (see equation 4) and calibrate this against  
297 MAAT and  $GDD_0$  using Deming regressions (Fig. 4a and 4d).

298 This results in the following two Deming temperature regressions:

$$(9) MAAT_{soil} (^\circ C) = 40.01 \times MBT'_{5me} - 15.25 \quad (n = 350, R^2 = 0.60, RMSE = 5.3 \text{ }^\circ C)$$

$$(10) GDD_{0\ soil} = 14344.3 \times MBT'_{5me} - 4997.5 \quad (n = 350, R^2 = 0.63, RMSE = 1779 \text{ }^\circ C)$$

299 Overall the  $GDD_0$  calibration performs slightly better than the MAAT calibration as it  
300 has a slightly higher  $R^2$ . Although the slope and intercept of  $MAAT_{soil}$  are different,  
301 the  $R^2$  and RMSE are similar compared to those reported for the  $MBT_{5me}$ -MAT  
302 calibration by De Jonge et al. (2014a) ( $n = 231, R^2 = 0.64$ , and  $RMSE = 4.9 \text{ }^\circ C$ ), but  
303 lower compared to those reported by Ding et al. (2015) ( $n = 249, R^2 = 0.70$ , and  
304  $RMSE = 4.7 \text{ }^\circ C$ ). Nonetheless, there is still significant scatter in our revised  
305 calibrations (Fig. 4b and 4e). Interestingly, the calibrations versus MAAT are  
306 characterized by relatively large residuals at lower temperatures (Fig. 4b).  
307 Specifically, the brGDGT distribution ( $MBT_{5me}'$ ) overestimates MAAT at lower  
308 temperatures and may be related to a seasonal production bias in high-latitude sites.  
309 Indeed these low temperature residuals are reduced when  $GDD_0$  is applied instead of  
310 MAAT (Fig. 4e).

311 An additional problem with both calibrations is that  $MBT'_{5me}$  reaches 1 at a  
312 MAAT of  $24.8 \text{ }^\circ C$  and  $GDD_0$  at  $9347 \text{ }^\circ C$ , thereby compromising the application of  
313 these calibrations to (sub)tropical settings both in the recent past but especially in the  
314 geological past when terrestrial temperatures were generally higher (e.g., Huber and  
315 Caballero, 2011).

316

### 317 3.6 Temperature calibrations using multiple linear regressions

318 As  $MBT'_{5me}$  reaches 1 at a relatively low MAAT (22.9 °C) in the temperature  
319 calibration of De Jonge et al. (2014a), the authors suggested that a multiple linear  
320 regression (MLR)-based calibration, based upon the fractional abundances of  
321 brGDGTs-Ia, -Ib, -Ic, and -IIa (eq. 6), was a more suitable choice for paleoclimate  
322 studies. As such, we have also explored the performance of MLRs in our expanded  
323 mineral soil dataset. The optimal MLRs are

$$(11) MAAT_{mlr\ soil} (^{\circ}C) = 19.8 \times \{Ia\} + 31.1 \times \{Ib\} - 23.4 \times \{IIa\} + 4.32 \quad (n = 350, R^2 = 0.62, RMSE = 4.7 ^{\circ}C)$$

$$(12) GDD_{0\ mlr\ soil} = 6152.9 \times \{Ia\} + 8272.1 \times \{Ib\} - 8015.8 \times \{IIa\} + 2509.4 \quad (n = 350, R^2 = 0.68, RMSE = 1319)$$

324 Adding additional compounds inflates the  $p$ -level of the slopes and intercept to values  
325  $> 0.01$ . The advantage of using a MLR model is that the correlations ( $R^2$ ) improve and  
326 RMSEs decrease compared to the  $MBT'_{5me}$  calibrations. However, the MLRs 1) reach  
327 saturation (100% brGDGT-Ia) around 24-25 °C, similar to the  $MBT'_{5me}$  calibration, 2)  
328 do not account for the error in both the proxy and environmental parameter as Deming  
329 regressions do, and 3) are characterized by structural residuals, which are the most  
330 significant at low MAAT (Fig. 4c and 4f). We therefore suggest the  $MBT'_{5me}$   
331 calibrations (eq. 9 and 10) are the optimal global calibrations. However, the amount of  
332 scatter in the calibration remain large, indicating that additional parameters may  
333 influence the total brGDGT distributions.

334

### 335 3.7 Temperature calibration excluding samples dominated by 6-methyl brGDGTs

336 Numerous studies have shown that there is a poor correlation between the methylation  
337 of brGDGTs and MAAT in arid and/or alkaline soils (Peterse et al., 2012; Dirghangi  
338 et al., 2013; Anderson et al., 2014; Zell et al., 2014; Ding et al., 2015; Yang et al.,  
339 2015). The reason for the apparent control of moisture (or other related environmental  
340 parameters) on the abundance of 6-methyl brGDGTs is unknown. Soil moisture is  
341 correlated to pH with dry soils predominantly being alkaline. Given that the  
342 abundance of 6-methyl brGDGTs is highly correlated to pH (see Fig. 5 and  
343 supplementary information), this might explain the correlation. In fatty acids, the  
344 position of methyl groups impacts membrane fluidity, with anteiso chains (methyl on  
345 C2 position) inducing a greater degree of fluidity compared to iso chains (methyl on



346 C1 position) (Denich et al., 2003 and references therein). Potentially the same applies  
 347 to brGDGTs, whereby shifting a methyl group from C5 to C6 leads to a greater  
 348 membrane fluidity in arid and/or alkaline soils. An alternative explanation could be  
 349 that different communities that thrive at different pH and soil moisture content  
 350 produce a different distribution of 5- and 6-methyl brGDGTs.

351 Using a transect of varying soil moisture content (SMC) in Chinese soils,  
 352 Dang et al. (2016) demonstrated that SMC has an impact on the distribution of  
 353 brGDGTs, in particular 6-methyl brGDGTs . Using the ratio of 6- over 5-methyl  
 354 brGDGTs ( $IR_{6me}$ ) they proposed that  $MBT'$  is only significantly correlated to MAAT  
 355 in mineral soils with  $IR_{6me} \leq 0.5$ .

$$(13) IR_{6me}$$

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

356 Based on this observation, we evaluated the influence of  $IR_{6me}$  on the correlation  
 357 coefficient ( $R^2$ ) between  $MBT'_{5me}$  and MAAT (Fig. 6). In the global soil dataset the  
 358 correlation coefficient ( $R^2$ ) between  $MBT'_{5me}$  and MAAT decreases significantly  
 359 from 0.76 to 0.67 when the threshold for  $IR_{6me}$  increases from  $< 0.5$  to  $< 0.6$ , similar  
 360 to observations from the Chinese soil transect. These results suggest that the  
 361 temperature dependence of brGDGTs in soils with a high amount of 6-methyl  
 362 brGDGTs over 5-methyl brGDGTs (mainly arid/alkaline soils) is different.

363 We therefore excluded samples with  $IR_{6me} > 0.5$ . From the total of 350 soil samples;  
 364 roughly half (177) have  $IR_{6me} < 0.5$ , and these are mostly from acidic soils (Fig. 5).  
 365 This yields a significant improvement in the correlations between individual  
 366 brGDGT-Ia, -IIa, and -IIIa and MAAT (Fig. 7;  $R^2$  values of 0.67, 0.68, and 0.51,  
 367 respectively) and leads to significantly improved calibrations with higher  $R^2$  and  
 368 lower RMSEs (Fig. 8). This applies to both Deming (Eq. 14 and 15) and multiple  
 369 linear regressions (Eq. 16 and 17).

$$(14) MAAT_{soil 5me} ( ^\circ C) = 39.09 \times MBT'_{5me} - 14.50 \quad (n = 177, R^2 = 0.76, RMSE = 4.1 ^\circ C)$$

$$(15) GDD_{0 soil 5me} = 13498.8 \times MBT'_{5me} - 4444.5 \quad (n = 177, R^2 = 0.78, RMSE = 1326)$$



$$(16) MAAT_{mlr\ soil\ 5me} (^{\circ}C) = 14.7 \times \{Ia\} - 31.7 \times \{IIa\} + 10.0 \quad (n = 177, R^2 = 0.77, RMSE = 3.8 ^{\circ}C)$$

$$(17) GDD_{0\ mlr\ soil\ 5me} = 4881 \times \{Ia\} - 10112 \times \{IIa\} + 3942 \quad (n = 177, R^2 = 0.82, RMSE = 1079)$$

370 Adding additional compounds to the MLRs does not improve the correlations and  
371 inflates the  $p$ -level of the slopes and intercept to values  $> 0.01$ . As explained  
372 previously, the MLRs do not take the error in both proxy and environmental  
373 parameter into account, only perform slightly better than the  $MBT'_{5me}$  calibrations,  
374 and the MLR calibration of of MAAT (eq. 16) is characterized by structural residuals,  
375 especially at the low temperature end (Fig. 8c). As such we suggest that the  $MBT'_{5me}$   
376 calibrations (eq. 14 and 15) are the best choice for paleoclimate reconstructions.  
377 However, both set of calibrations continue to saturate at temperatures of around 24-  
378 25 °C, implying that their application to past greenhouse climates has to be  
379 undertaken with caution.

380 These improvements over earlier calibrations imply that sample sets need to  
381 be screened for the abundance of significant amounts of 6-methyl brGDGTs prior to  
382 MAAT determinations. This will be particular important for archives from (semi-)  
383 arid regions such as loess and paleosols. It is important to note that in paleoclimate  
384 archives such as marine sediments, brGDGTs might be derived from a mixture of  
385 sources. This means that although these archives overall might be characterized by  
386  $IR_{6me} < 0.5$ , there could be a contribution of soils with  $IR_{6me} > 0.5$ .

387 We envision that the ability to calculate growing degree days will be of  
388 particular interest to climate modelers as  $GDD_0$  is more indicative of the seasonal  
389 temperature cycle than MAAT, especially at the high latitudes. Information about past  
390 seasonal temperatures and summer intensity is not readily available and provides a  
391 clear advantage of our calibration over other temperature proxies (marine and  
392 terrestrial).

393

#### 394 **4. Conclusions**

395 The distribution of brGDGTs in soils has been shown previously to depend on  
396 environmental parameters such as mean annual air temperature (MAAT) and pH, but  
397 significant scatter in the existing calibrations suggests additional controls. Combining  
398 all available data, here we compare the brGDGT distribution to a range of

399 environmental parameters obtained from a globally integrated data set. In agreement  
400 with previous studies, we demonstrate that the distribution of 5-methyl brGDGTs  
401 depends primarily on temperature. Excluding samples from arid and/or alkaline soils  
402 dominated by 6-methyl brGDGTs significantly improves the correlation with  
403 temperature and growing degree days above zero (GDD<sub>0</sub>). Guided by these results we  
404 provide new temperature calibrations. These new regressions have significantly  
405 improved correlation coefficients and lower root mean square errors (RMSE)  
406 compared to the existing global calibrations. We suggest that these new calibrations  
407 should be used to reconstruct terrestrial climate during the geological past, but caution  
408 should be taken when applying these calibrations to past greenhouse periods.

409

#### 410 **Acknowledgements**

411 This research was funded through the advanced ERC grant 'The greenhouse earth  
412 system' (T-GRES, project reference 340923). R.D.P. acknowledges the Royal Society  
413 Wolfson Research Merit Award.

414

#### 415 **References**

- 416 Adcock, R.J., 1878. A Problem in Least Squares. *The Analyst* **5**, 53-54.  
417
- 418 Anderson, V.J., Shanahan, T.M., Saylor, J.E., Horton, B.K., Mora, A.R., 2014.  
419 Sources of local and regional variability in the MBT' /CBT paleotemperature proxy:  
420 Insights from a modern elevation transect across the Eastern Cordillera of Colombia.  
421 *Organic Geochemistry* **69**, 42-51.  
422
- 423 Dang, X., Yang, H., Naafs, B.D.A., Pancost, R.D., Evershed, R.P., Xie, S., 2016.  
424 Direct evidence of moisture control on the methylation of branched glycerol dialkyl  
425 glycerol tetraethers in semi-arid and arid soils. *Geochimica et Cosmochimica Acta*  
426 **189**, 24-36.  
427
- 428 De Jonge, C., Hopmans, E.C., Stadnitskaia, A., Rijpstra, W.I.C., Hofland, R.,  
429 Tegelaar, E., Sinninghe Damsté, J.S., 2013. Identification of novel penta- and  
430 hexamethylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC–  
431 MS<sup>2</sup>, GC–MS and GC–SMB-MS. *Organic Geochemistry* **54**, 78-82.  
432
- 433 De Jonge, C., Hopmans, E.C., Zell, C.I., Kim, J.-H., Schouten, S., Sinninghe Damsté,  
434 J.S., 2014a. Occurrence and abundance of 6-methyl branched glycerol dialkyl  
435 glycerol tetraethers in soils: implications for palaeoclimate reconstruction.  
436 *Geochimica et Cosmochimica Acta* **141**, 97-112.  
437
- 438 De Jonge, C., Stadnitskaia, A., Hopmans, E.C., Cherkashov, G., Fedotov, A.,  
439 Sinninghe Damsté, J.S., 2014b. In situ produced branched glycerol dialkyl glycerol

440 tetraethers in suspended particulate matter from the Yenisei River, Eastern Siberia.  
441 *Geochimica et Cosmochimica Acta* **125**, 476-491.  
442  
443 De Jonge, C., Hopmans, E.C., Zell, C.I., Kim, J.-H., Schouten, S., Sinninghe Damsté,  
444 J.S., 2016. Corrigendum to “Occurrence and abundance of 6-methyl branched  
445 glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate  
446 reconstruction” [Geochim. Cosmochim. Acta 141 (2014) 97–112]. *Geochimica et*  
447 *Cosmochimica Acta* **183**, 276-278.  
448  
449 Denich, T.J., Beaudette, L.A., Lee, H., Trevors, J.T., 2003. Effect of selected  
450 environmental and physico-chemical factors on bacterial cytoplasmic membranes.  
451 *Journal of Microbiological Methods* **52**, 149-182.  
452  
453 Ding, S., Xu, Y., Wang, Y., He, Y., Hou, J., Chen, L., He, J.S., 2015. Distribution of  
454 branched glycerol dialkyl glycerol tetraethers in surface soils of the Qinghai-Tibetan  
455 Plateau: implications of brGDGTs-based proxies in cold and dry regions.  
456 *Biogeosciences* **12**, 3141-3151.  
457  
458 Dirghangi, S.S., Pagani, M., Hren, M.T., Tipple, B.J., 2013. Distribution of glycerol  
459 dialkyl glycerol tetraethers in soils from two environmental transects in the USA.  
460 *Organic Geochemistry* **59**, 49-60.  
461  
462 Fietz, S., Huguet, C., Bendle, J., Escala, M., Gallacher, C., Herfort, L., Jamieson, R.,  
463 et al., 2012. Co-variation of crenarchaeol and branched GDGTs in globally-  
464 distributed marine and freshwater sedimentary archives. *Global and Planetary*  
465 *Change* **92–93**, 275-285.  
466  
467 Gallego-Sala, A.V., Clark, J.M., House, J.I., Orr, H.G., Prentice, I.C., Smith, P.,  
468 Farewell, T., et al., 2010. Bioclimatic envelope model of climate change impacts on  
469 blanket peatland distribution in Great Britain. *Climate Research* **45**, 151-162.  
470  
471 Gallego-Sala, A.V., Prentice, C.I., 2013. Blanket peat biome endangered by climate  
472 change. *Nature Climate Change* **3**, 152-155.  
473  
474 Hopmans, E.C., Weijers, J.W.H., Schefuß, E., Herfort, L., Sinninghe Damsté, J.S.,  
475 Schouten, S., 2004. A novel proxy for terrestrial organic matter in sediments based on  
476 branched and isoprenoid tetraether lipids. *Earth and Planetary Science Letters* **224**,  
477 107-116.  
478  
479 Hopmans, E.C., Schouten, S., Sinninghe Damsté, J.S., 2016. The effect of improved  
480 chromatography on GDGT-based palaeoproxies. *Organic Geochemistry* **93**, 1-6.  
481  
482 Huber, M., Caballero, R., 2011. The early Eocene equable climate problem revisited.  
483 *Climate of the Past* **7**, 603-633.  
484  
485 Huguet, A., Fosse, C., Laggoun-Défarge, F., Toussaint, M.-L., Derenne, S., 2010.  
486 Occurrence and distribution of glycerol dialkyl glycerol tetraethers in a French peat  
487 bog. *Organic Geochemistry* **41**, 559-572.  
488

489 Huguet, A., Fosse, C., Laggoun-Défarge, F., Delarue, F., Derenne, S., 2013. Effects of  
490 a short-term experimental microclimate warming on the abundance and distribution of  
491 branched GDGTs in a French peatland. *Geochimica et Cosmochimica Acta* **105**, 294-  
492 315.

493  
494 Kaplan, J.O., Bigelow, N.H., Prentice, I.C., Harrison, S.P., Bartlein, P.J., Christensen,  
495 T.R., Cramer, W., et al., 2003. Climate change and Arctic ecosystems: 2. Modeling,  
496 paleodata-model comparisons, and future projections. *Journal of Geophysical*  
497 *Research: Atmospheres* **108**.

498  
499 Lei, Y., Yang, H., Dang, X., Zhao, S., Xie, S., 2016. Absence of a significant bias  
500 towards summer temperature in branched tetraether-based paleothermometer at two  
501 soil sites with contrasting temperature seasonality. *Organic Geochemistry* **94**, 83-94.  
502

503 Li, J., Pancost, R.D., Naafs, B.D.A., Yang, H., Zhao, C., Xie, S., 2016. Distribution of  
504 glycerol dialkyl glycerol tetraether (GDGT) lipids in a hypersaline lake system.  
505 *Organic Geochemistry* **99**, 113-124.  
506

507 Lu, H., Liu, W., Wang, H., Wang, Z., 2016. Variation in 6-methyl branched glycerol  
508 dialkyl glycerol tetraethers in Lantian loess–paleosol sequence and effect on  
509 paleotemperature reconstruction. *Organic Geochemistry* **100**, 10-17.  
510

511 Manuilova, E., Schuetzenmeister, A., Model, F., 2014. Method Comparison  
512 Regression. CRAN, <https://cran.r-project.org/web/packages/mcr/index.html>.  
513

514 McMaster, G.S., Wilhelm, W.W., 1997. Growing degree-days: one equation, two  
515 interpretations. *Agricultural and Forest Meteorology* **87**, 291-300.  
516

517 Naafs, B.D.A., Inglis, G.N., Pancost, R.D., 2017. Introducing global peat-specific  
518 temperature and pH calibrations based on brGDGT bacterial lipids. *Geochimica et*  
519 *Cosmochimica Acta* **accepted pending minor revisions**.  
520

521 Pancost, R.D., Taylor, K.W.R., Inglis, G.N., Kennedy, E.M., Handley, L., Hollis, C.J.,  
522 Crouch, E.M., et al., 2013. Early Paleogene evolution of terrestrial climate in the SW  
523 Pacific, Southern New Zealand. *Geochemistry, Geophysics, Geosystems* **14**, 5413-  
524 5429.  
525

526 Pearson, E.J., Juggins, S., Talbot, H.M., Weckström, J., Rosén, P., Ryves, D.B.,  
527 Roberts, S.J., et al., 2011. A lacustrine GDGT-temperature calibration from the  
528 Scandinavian Arctic to Antarctic: Renewed potential for the application of GDGT-  
529 paleothermometry in lakes. *Geochimica et Cosmochimica Acta* **75**, 6225-6238.  
530

531 Peterse, F., van der Meer, M.T.J., Schouten, S., Jia, G., Ossebaar, J., Blokker, J.,  
532 Sinninghe Damsté, J.S., 2009. Assessment of soil *n*-alkane  $\delta D$  and branched tetraether  
533 membrane lipid distributions as tools for paleoelevation reconstruction.  
534 *Biogeosciences* **6**, 2799-2807.  
535

536 Peterse, F., van der Meer, J., Schouten, S., Weijers, J.W.H., Fierer, N., Jackson, R.B.,  
537 Kim, J.-H., et al., 2012. Revised calibration of the MBT–CBT paleotemperature proxy

538 based on branched tetraether membrane lipids in surface soils. *Geochimica et*  
539 *Cosmochimica Acta* **96**, 215-229.

540

541 Peterse, F., Martínez-García, A., Zhou, B., Beets, C.J., Prins, M.A., Zheng, H.,  
542 Eglinton, T.I., 2014. Molecular records of continental air temperature and monsoon  
543 precipitation variability in East Asia spanning the past 130,000 years. *Quaternary*  
544 *Science Reviews* **83**, 76-82.

545

546 RStudio Team, 2015. RStudio: Integrated Development for R. RStudio, Inc., Boston,  
547 MA (USA).

548

549 Russell, N.J., 1984. Mechanisms of thermal adaptation in bacteria: blueprints for  
550 survival. *Trends in Biochemical Sciences* **9**, 108-112.

551

552 Schlitzer, R., 2015. Ocean Data View. <http://odv.awi.de>.

553

554 Schoon, P.L., de Kluijver, A., Middelburg, J.J., Downing, J.A., Sinninghe Damsté,  
555 J.S., Schouten, S., 2013. Influence of lake water pH and alkalinity on the distribution  
556 of core and intact polar branched glycerol dialkyl glycerol tetraethers (GDGTs) in  
557 lakes. *Organic Geochemistry* **60**, 72-82.

558

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

562

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

566

567 Sinninghe Damsté, J.S., Rijpstra, W.I.C., Hopmans, E.C., Weijers, J.W.H., Foesel,  
568 B.U., Overmann, J., Dedysh, S.N., 2011. 13,16-Dimethyl Octacosanedioic Acid (*iso*-  
569 Diabolic Acid), a Common Membrane-Spanning Lipid of *Acidobacteria* Subdivisions  
570 1 and 3. *Applied and Environmental Microbiology* **77**, 4147-4154.

571

572 Weijers, J.W.H., Schouten, S., Hopmans, E.C., Geenevasen, J.A.J., David, O.R.P.,  
573 Coleman, J.M., Pancost, R.D., et al., 2006a. Membrane lipids of mesophilic anaerobic  
574 bacteria thriving in peats have typical archaeal traits. *Environmental Microbiology* **8**,  
575 648-657.

576

577 Weijers, J.W.H., Schouten, S., Spaargaren, O.C., Sinninghe Damsté, J.S., 2006b.  
578 Occurrence and distribution of tetraether membrane lipids in soils: Implications for  
579 the use of the TEX<sub>86</sub> proxy and the BIT index. *Organic Geochemistry* **37**, 1680-1693.

580

581 Weijers, J.W.H., Schefuß, E., Schouten, S., Damsté, J.S.S., 2007a. Coupled Thermal  
582 and Hydrological Evolution of Tropical Africa over the Last Deglaciation. *Science*  
583 **315**, 1701-1704.

584

585 Weijers, J.W.H., Schouten, S., van den Donker, J.C., Hopmans, E.C., Sinninghe  
586 Damsté, J.S., 2007b. Environmental controls on bacterial tetraether membrane lipid  
587 distribution in soils. *Geochimica et Cosmochimica Acta* **71**, 703-713.

588

589 Weijers, J.W.H., Panoto, E., van Bleijswijk, J., Schouten, S., Rijpstra, W.I.C., Balk,  
590 M., Stams, A.J.M., et al., 2009. Constraints on the Biological Source(s) of the Orphan  
591 Branched Tetraether Membrane Lipids. *Geomicrobiology Journal* **26**, 402-414.

592

593 Weijers, J.W.H., Bernhardt, B., Peterse, F., Werne, J.P., Dungait, J.A.J., Schouten, S.,  
594 Sinninghe Damsté, J.S., 2011. Absence of seasonal patterns in MBT–CBT indices in  
595 mid-latitude soils. *Geochimica et Cosmochimica Acta* **75**, 3179-3190.

596

597 Xiao, W., Xu, Y., Ding, S., Wang, Y., Zhang, X., Yang, H., Wang, G., et al., 2015.  
598 Global calibration of a novel, branched GDGT-based soil pH proxy. *Organic*  
599 *Geochemistry* **89–90**, 56-60.

600

601 Yang, H., Lü, X., Ding, W., Lei, Y., Dang, X., Xie, S., 2015. The 6-methyl branched  
602 tetraethers significantly affect the performance of the methylation index (MBT<sup>1</sup>) in  
603 soils from an altitudinal transect at Mount Shennongjia. *Organic Geochemistry* **82**,  
604 42-53.

605

606 Zell, C., Kim, J.H., Balsinha, M., Dorhout, D., Fernandes, C., Baas, M., Sinninghe  
607 Damsté, J.S., 2014. Transport of branched tetraether lipids from the Tagus River basin  
608 to the coastal ocean of the Portuguese margin: consequences for the interpretation of  
609 the MBT/CBT paleothermometer. *Biogeosciences* **11**, 5637-5655.

610

611 Zheng, Y., Li, Q., Wang, Z., Naafs, B.D.A., Yu, X., Pancost, R.D., 2015. Peatland  
612 GDGT records of Holocene climatic and biogeochemical responses to the Asian  
613 Monsoon. *Organic Geochemistry* **87**, 86-95.

614

615

## 616 **Figure captions**

617 Figure 1; World map with topography and location of soils used in this study, created  
618 using Ocean Data View (Schlitzer, 2015).

619

620 Figure 2; Fractional abundance of the brGDGT-Ia, -IIa, -IIIa, -IIa', and -IIIa' in the  
621 global soil sample data set versus mean annual air temperature (MAAT, top row),  
622 mean annual precipitation (MAP, middle row), and moisture index (MI, bottom row).  
623 Climatic parameters are obtained using PeatStash. Linear regressions are shown for  
624 those brGDGTs which relative abundance has a linear correlation coefficient ( $R^2$ ) of  
625 at least 0.2. Zero values (below detection limit) are not included. Correlation of other  
626 brGDGTs to these parameters is not significant ( $R^2 < 0.2$ ) and not shown.

627

628 Figure 3; Same as figure 2, but now for mean warmest month temperature (MWMT)  
629 and growing degree days above zero (GDD<sub>0</sub>). Samples from soils with MAAT < 5 °C  
630 are highlighted in red in the MAAT plots.

631

632 Figure 4;  $MBT_{5me}'$  plotted versus a) mean annual air temperature (MAAT) and d)  
633 growing degree days above 0 °C ( $GDD_0$ ) together with Deming regression (purple  
634 line) and simple linear regression (black dotted line). Also shown are the residuals for  
635 both the Deming (b and e) and multiple linear regressions (c and f). Gray area in the  
636 residual plots indicated missed variation because  $MBT_{5me}'$  reached 1.

637

638 Fig. 5; Histograms of pH values from mineral soils, showing that samples with  $IR_{6me}$   
639  $< 0.5$  are predominantly from acidic soils.

640

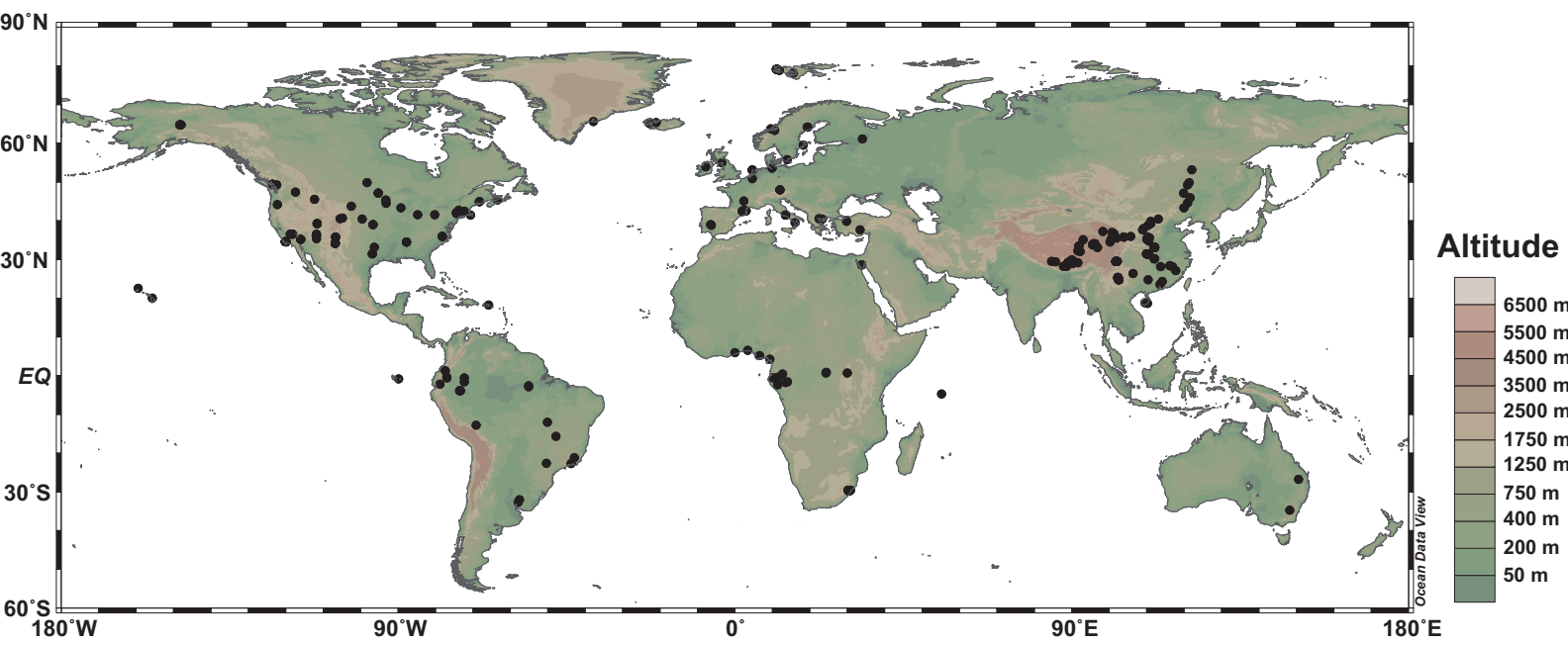
641 Fig. 6: The correlation coefficient ( $R^2$ ) between MAAT and  $MBT_{5me}'$  versus the  $IR_{6me}$   
642 cut-off value as well as number of soils in each dataset. The total dataset ( $IR_{6me}$  cut-  
643 off = 1) has a  $R^2$  of 0.6 and consists of 350 samples.

644

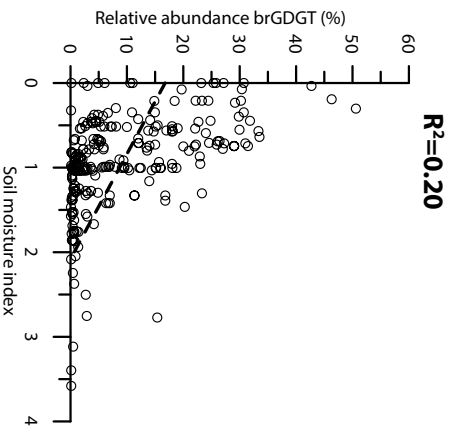
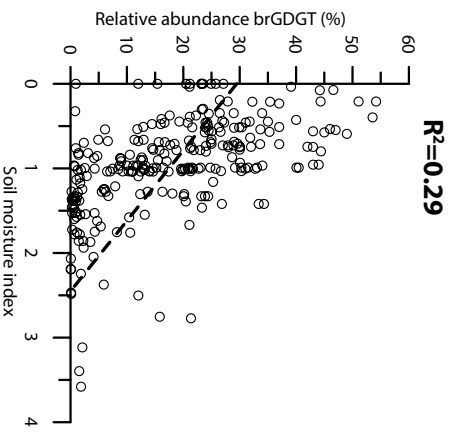
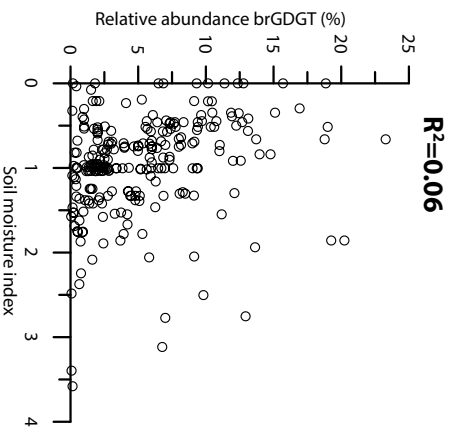
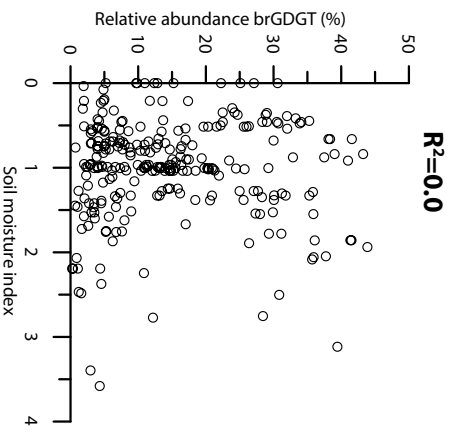
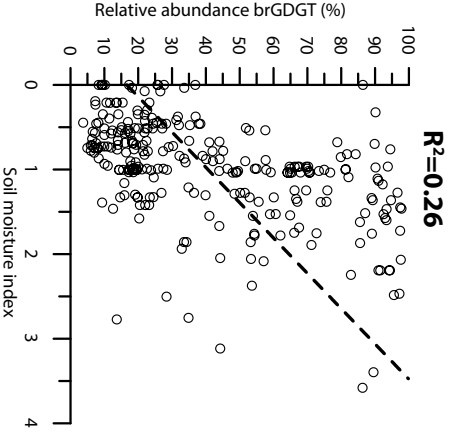
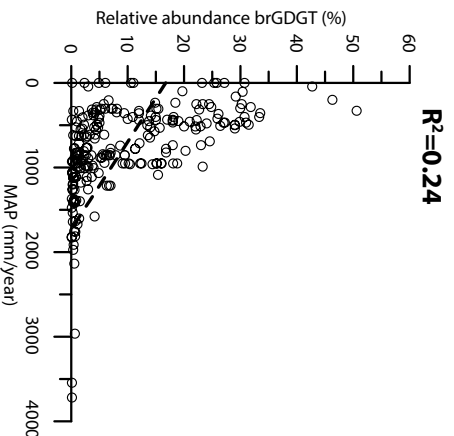
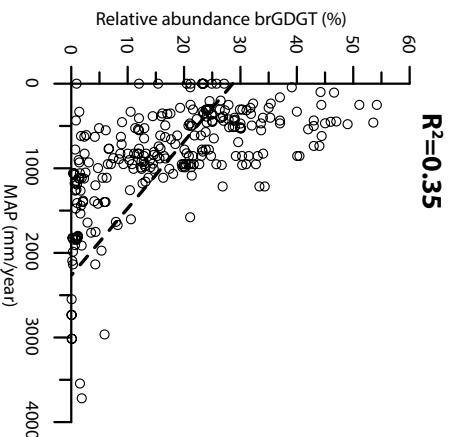
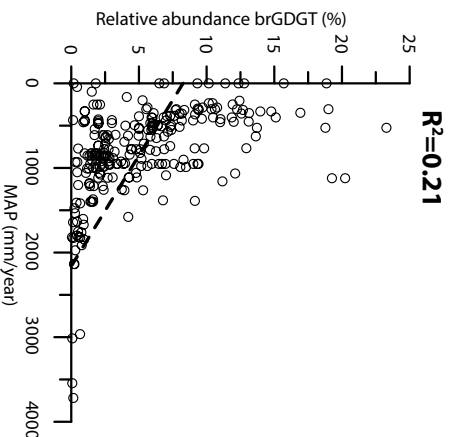
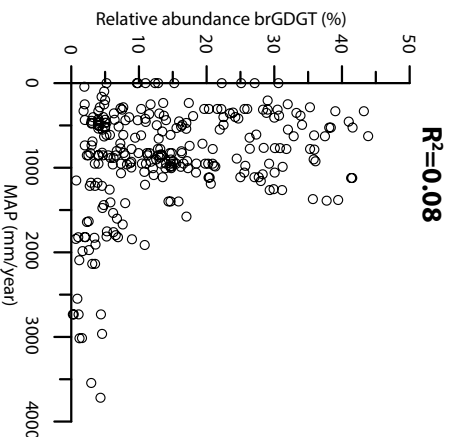
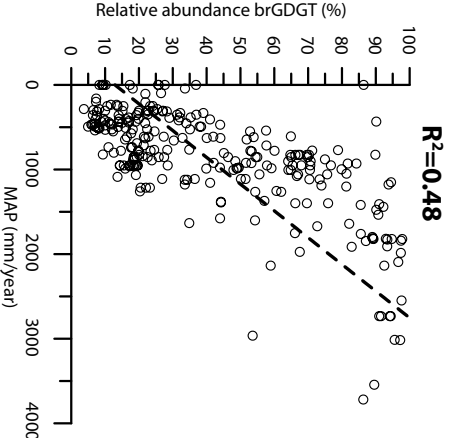
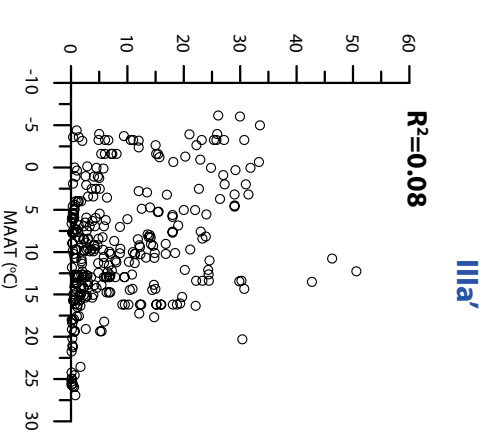
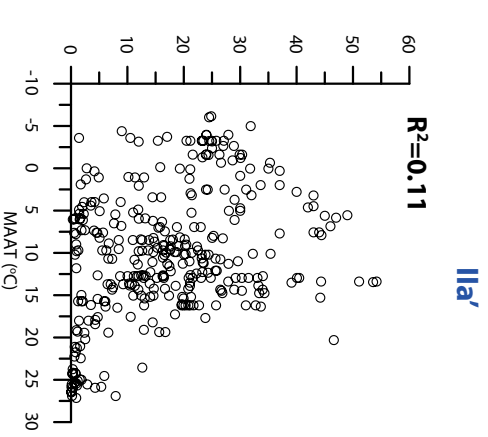
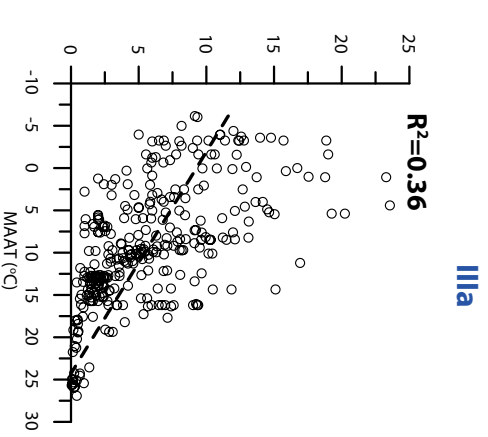
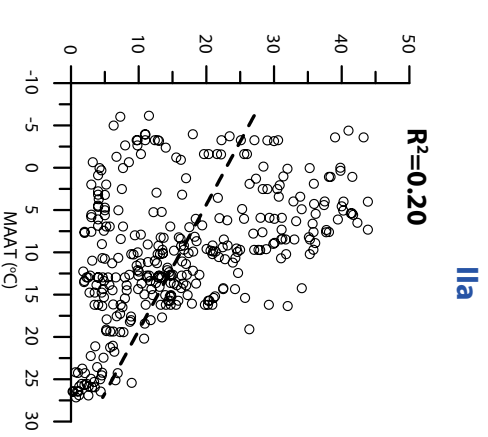
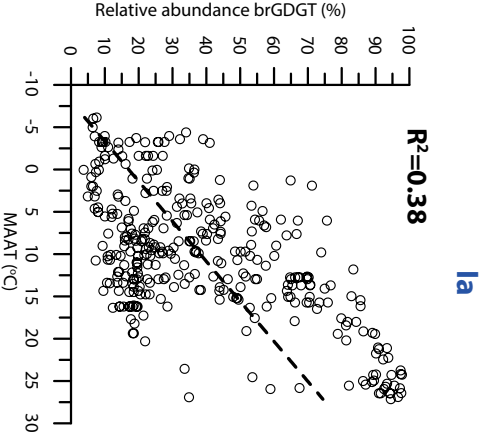
645 Fig. 7; Relative abundance of brGDGT-Ia, -IIa, and -IIIa versus MAAT and  $GDD_0$  for  
646 the complete soil data set (black,  $n=350$ ) and soil samples with  $IR_{6me} < 0.5$  (pink,  
647  $n=177$ ).

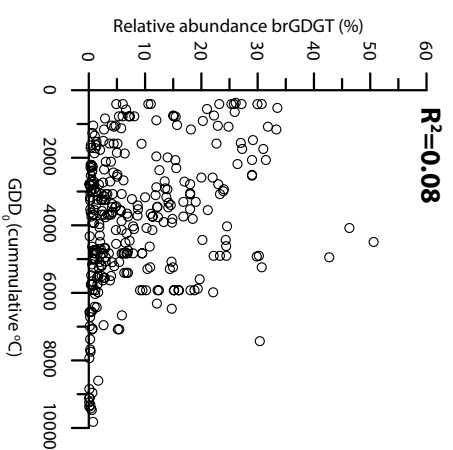
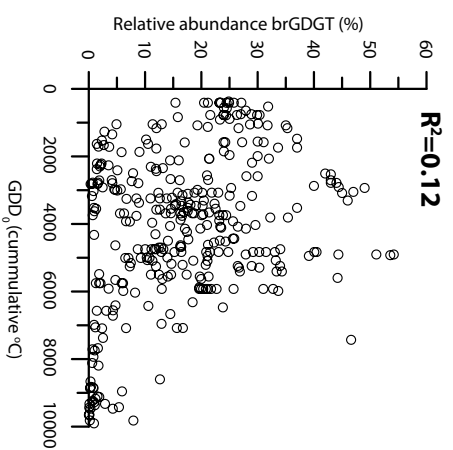
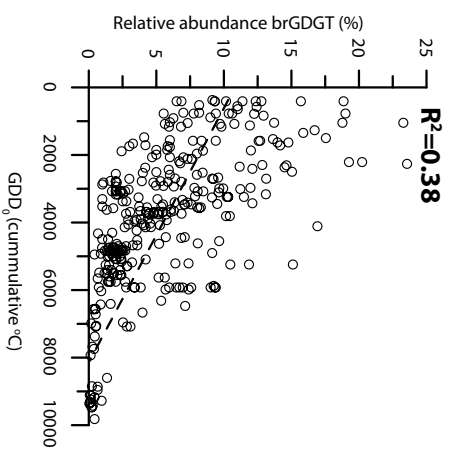
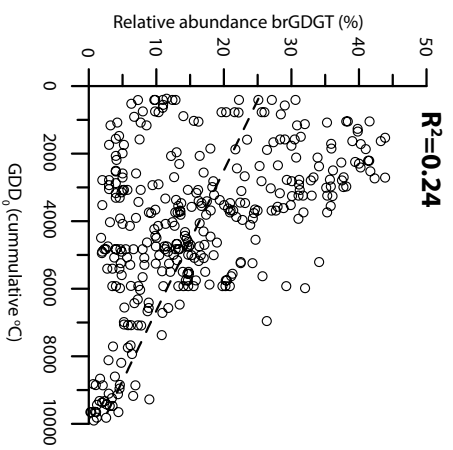
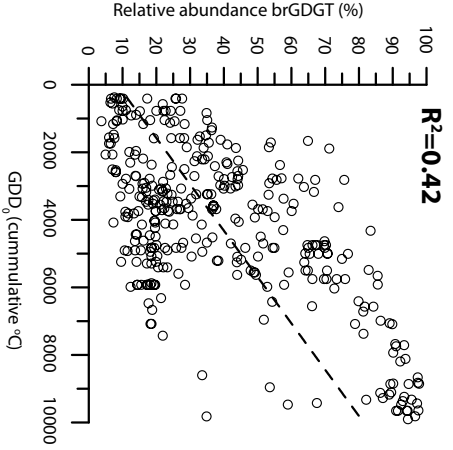
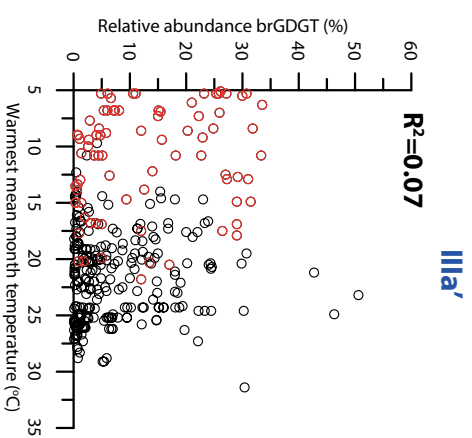
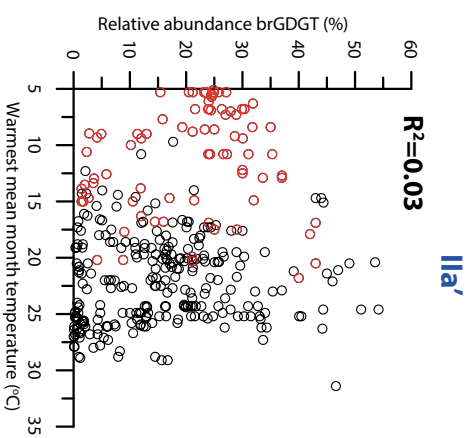
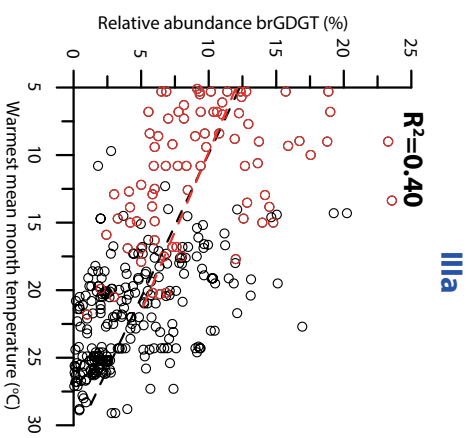
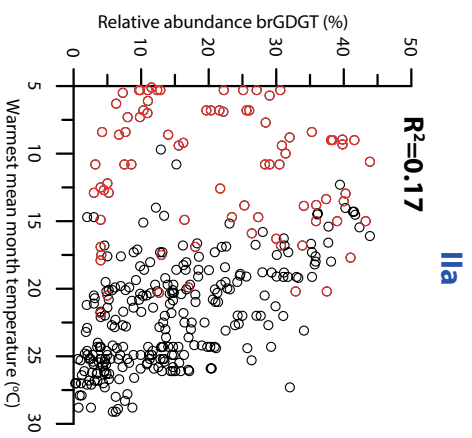
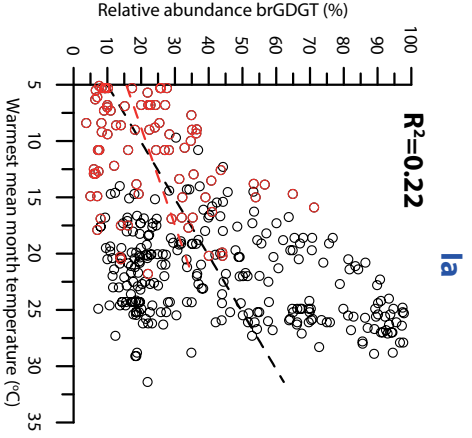
648

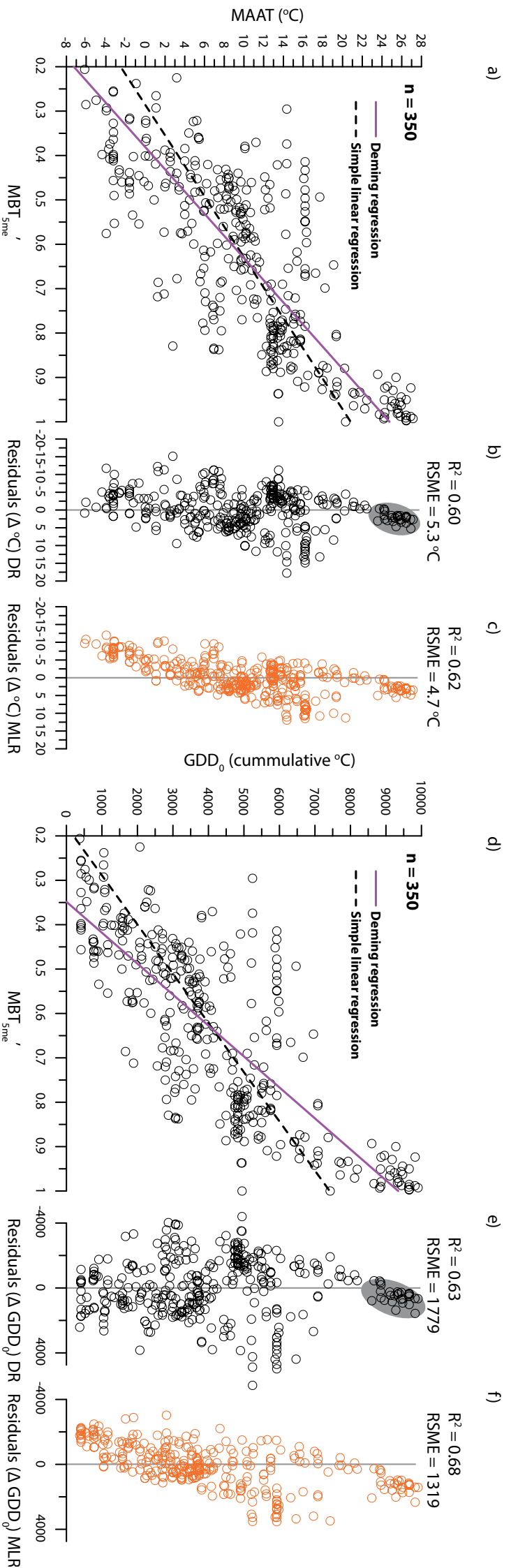
649 Figure 8;  $MBT_{5me}'$  of samples with  $IR_{6me} < 0.5$  plotted versus a) mean annual air  
650 temperature (MAAT) and d) growing degree days above 0 °C ( $GDD_0$ ) together with  
651 Deming regression (purple line) and simple linear regression (black dotted line). Also  
652 shown are the residuals for both the Deming (b&e) and multiple linear regressions  
653 (e&f). Gray area in the residual plots indicated missed variation because  $MBT_{5me}'$   
654 reached 1.



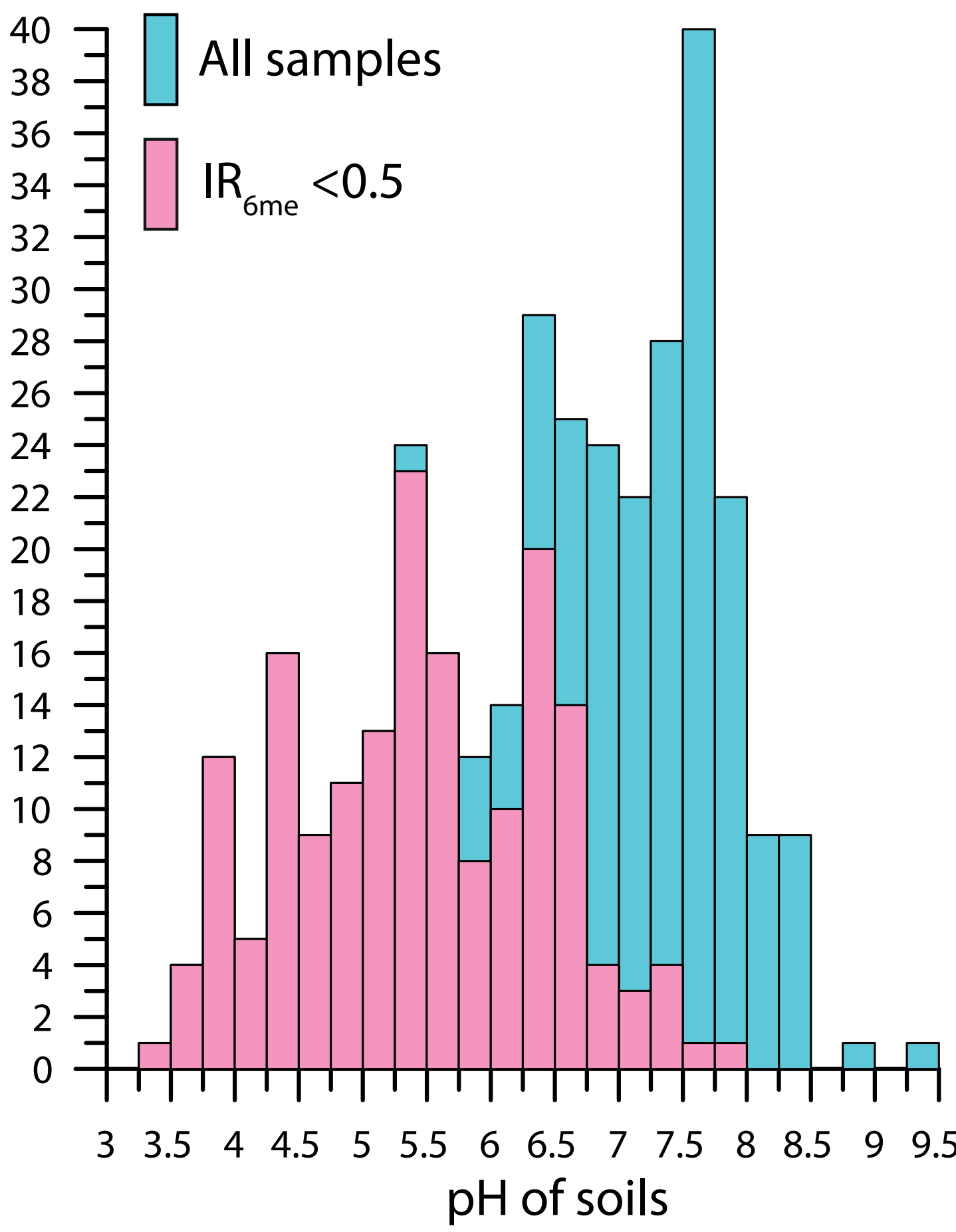


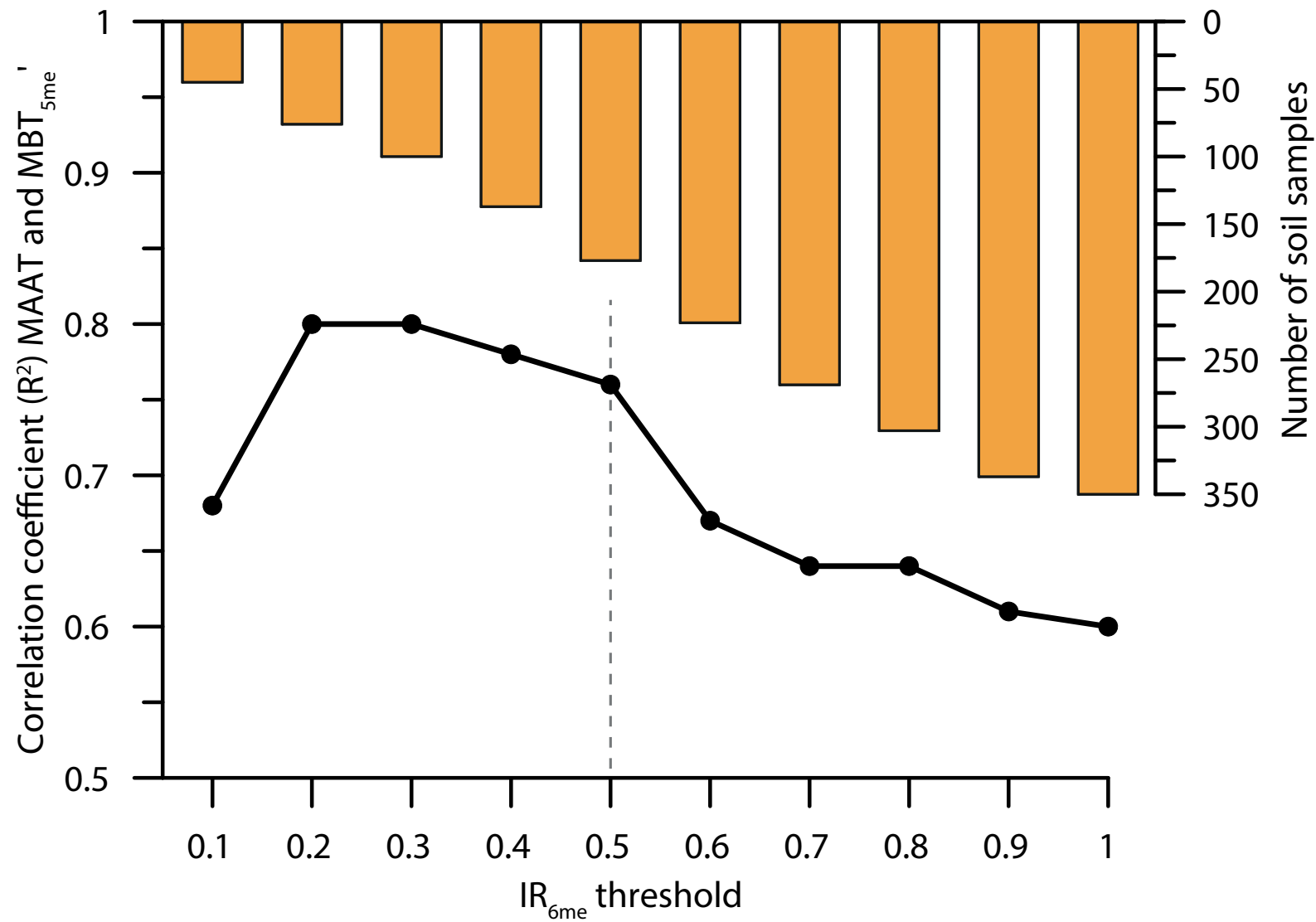


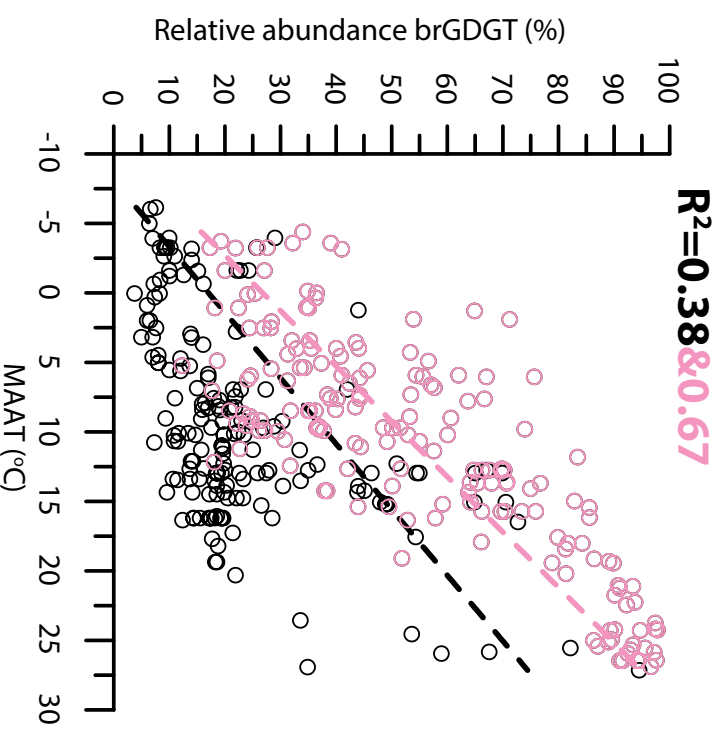
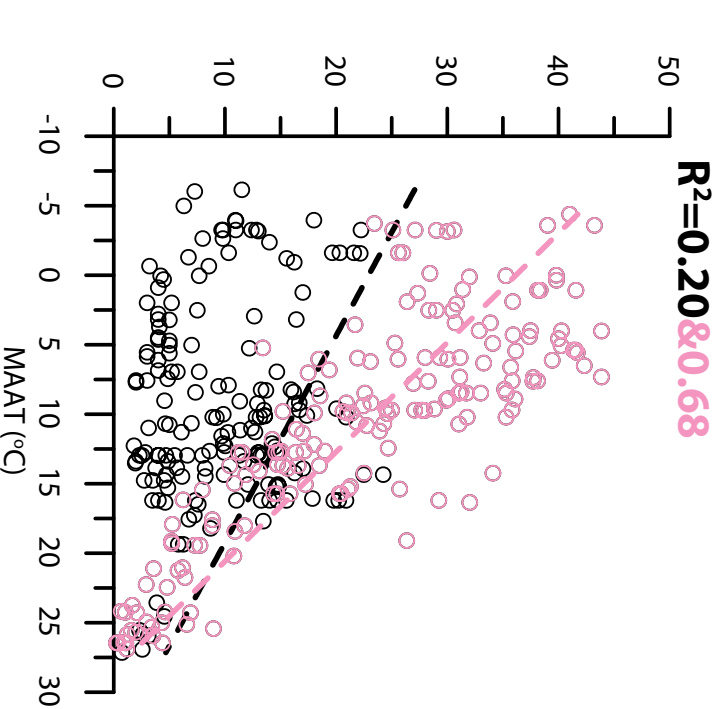
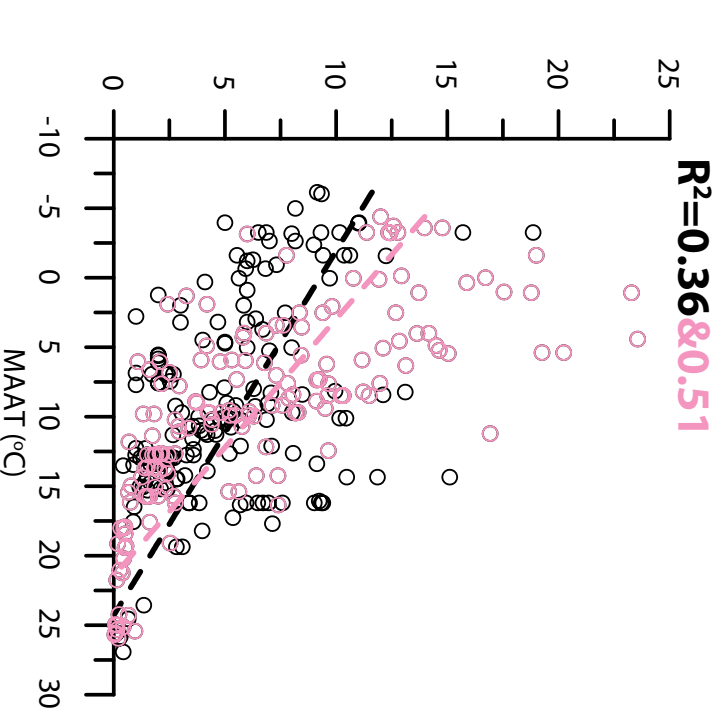
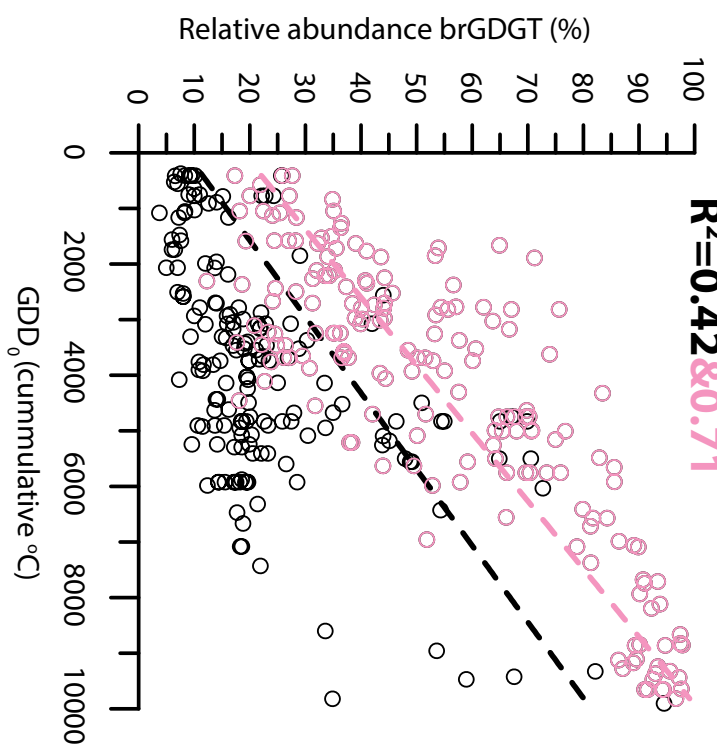
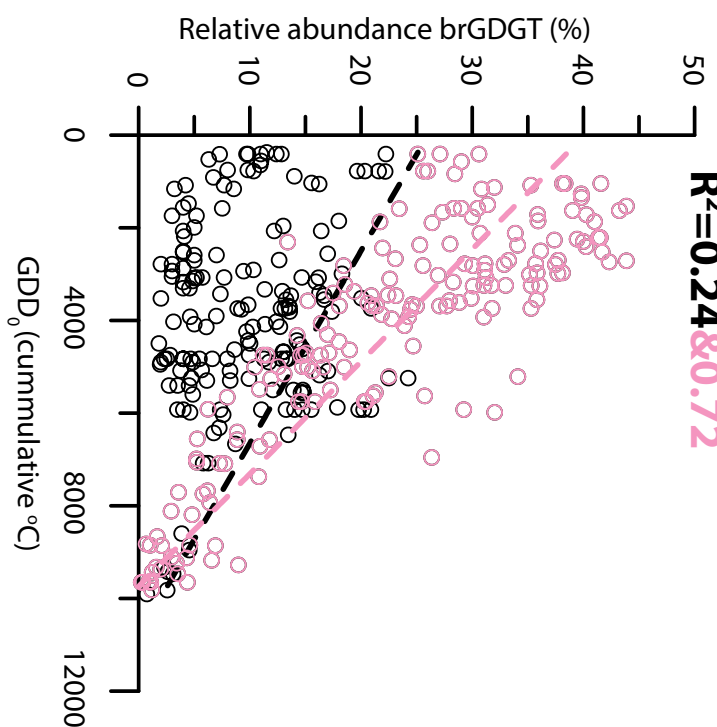




Number of soil samples





**Ia****IIa****IIIa****R<sup>2</sup>=0.42&0.71****R<sup>2</sup>=0.24&0.72****R<sup>2</sup>=0.38&0.51**