Supporting Figure Captions

Fig. S1 Relationship between site-mean ln(leaf area) and mean annual temperature. Standard errors of the site means are plotted. Linear regression fit and associated r^2 and *P*-values are given (see also Table S3). N. South America = northern South America and includes all sites north of 34 °S latitude; S. South America = southern South America and encompasses all sites south of 34 °S latitude; NZ = New Zealand.

Fig. S2 Correlations between the percent of untoothed species in a flora and mean annual temperature (MAT) for CLAMP sites (Wolfe, 1993) and for sites from Kowalski and Dilcher (2003). Edaphically-wet sites from the CLAMP and Kowalski and Dilcher (2003) calibrations are indicated by blue circles and orange triangles, respectively. The regression between the percent of untoothed species and MAT for the Kowalski and Dilcher (2003) sites is indicated by the orange line. The regression fit of all edaphically-wet sites (CLAMP and Kowalski and Dilcher (2003)) is indicated by dashed line, and the regression fit of all sites in the CLAMP calibration is indicated by the solid black line. Locally-dry sites from the CLAMP calibration are indicated by red circles; the linear regression fit between the percent of untoothed species and MAT for these sites is indicated by the red line. The slopes of the regression of all of the edaphically-wet sites (dashed line) and the Kowalski and Dilcher (2003) sites (orange line) are not statistically distinct from the slope of the regression of all of the sites in the CLAMP calibration ($P = 0.12$). The y-intercept of the regression for locally dry sites (red line) is statistically distinct from the regressions of all CLAMP sites and the edaphically-wet sites ($P < 0.001$).

Fig. S3 Relationship between the site-mean of physiognomic variables and mean annual temperature for the 92 calibration sites. The physiognomic variables shown here are representative of all of the other variables. See Table S2 for definitions of variables. Three sites from Connecticut and two sites from Florida were collected along a local water availability gradient. Sites from Connecticut were collected from a swamp (blue square), along a nearby river margin (tan square), and from a nearby upland site (red square) (see Supporting Notes for details about sites). The sites from Florida were collected from a swamp (blue triangle) and an adjacent upland site (red triangle) (see Kowalski and Dilcher 2003 and Supporting Notes for details about sites). The other 87 calibration sites are indicated by grey circles. For all of these sites, there is no consistent pattern in the tooth characters (i.e., the wet sites are not consistently toothier than the dry sites). Further, except for the percent of untoothed species in the Florida sites, there is no appreciable difference in the physiognomic characters at the wet sites and the rest of the calibration sites. This suggests that the digitally-measured characters are relatively insensitive to effects related to local water availability.

Fig. S4 Canonical correspondence analysis plot of all calibration and fossil sites based on leaf physiognomy. The vectors for mean annual temperature (MAT) and mean annual precipitation (MAP) are also shown. Axis 1 (x-axis) explains 86.1 % of the physiognomy-climate relationships, with axis 2 (y-axis) explaining the remaining 13.9 %. Calibration sites are indicated by circles and fossil sites by squares. The physiognomic variables of the fossil sites are passive and therefore do not impact the ordination. Open symbols indicate sites from North America, Europe, Central America, northern South America, the Caribbean, and Asia; grey symbols indicate sites from Australia, New Zealand, Fiji, and southern South America. Leaf

physiognomy successfully clusters all calibration sites into their correct climate biome (compare with Fig. 1b). The fossil site Bonanza does not plot within the calibrated climate-physiognomy space and is therefore excluded from paleoclimate interpretation. We note that the climate biomes for some fossil sites (e.g., Laguna del Hunco, Palacio de los Loros, and Republic) are probably not reasonable (compare with Table 2) and demonstrate a weakness of the CCA approach (see main text). Thus, we recognize that our protocol for excluding potential fossil sites may not always be appropriate.

Fig. S5 Box-and-whisker plots of leaf mass per area for nine of the ten fossil floras used in study. Line = median, box = ± 1σ, whiskers = 5% and 95% confidence intervals; black dots = species outside the confidence intervals. The numbers inside each box indicate the number of woody dicot species processed for each site. Only species with ≥ 2 measurable specimens are included. The site not plotted, Hubble Bubble, had inadequate sampling (n = 6 species). WBI-III represents a compilation of all the Williston Basin floras (WBI, II, and III). P. Loros is Palacio de los Loros. Republic and Bonanza are from Royer *et al.* (2007) and Cerrejón is from Wing *et al.* (2009). Evergreen taxa typically have a leaf mass per area ≥ 129 g/m^2 (Wright *et al.*, 2004; Royer *et al.*, 2007). The Bonanza flora has the highest estimated site-mean leaf mass per area, suggesting that several of the species were evergreen, which may affect the leaf-climate correlations of the site. All reconstructions are based on the method of Royer *et al.* (2007).

Fig. S6 Box-and-whisker plots of two of the physiognomic variables in the mean annual temperature (MAT) model for the fossil floras, excluding Bonanza. (a) Feret's diameter ratio and

(b) number of teeth/internal perimeter (see Table S2 for definitions of variables). The third variable in the digital leaf physiognomy MAT model is percentage of untoothed species. The position of the sites on the MAT axis is determined by the digital leaf physiognomy estimate of MAT. For comparison, the modern calibration sites are plotted as black dots. Line = median, box = ± 1σ, whiskers = 5% and 95% confidence intervals; squares = species outside the confidence intervals, CR = Cerrejón, HB = Hubble Bubble, FH = Fox Hills, WBI = Williston Basin I, WBII = Williston Basin II, WBIII = Williston Basin III, PL = Palacio de los Loros, LH = Laguna del Hunco, RP = Republic. Estimates of MAT using digital leaf physiognomy are considerably warmer than leaf-margin analysis for six of the floras (Williston Basin I, II, and III, Fox Hills, Hubble Bubble, and Cerrejón; see Table 1). The number of teeth/internal perimeter variable is largely responsible for the warmer estimates because these fossil sites all plot below the calibration regression. The sole influence of Feret's diameter ratio at these sites is to decrease estimated MAT because the sites all plot above the calibration regression.

Figure S1

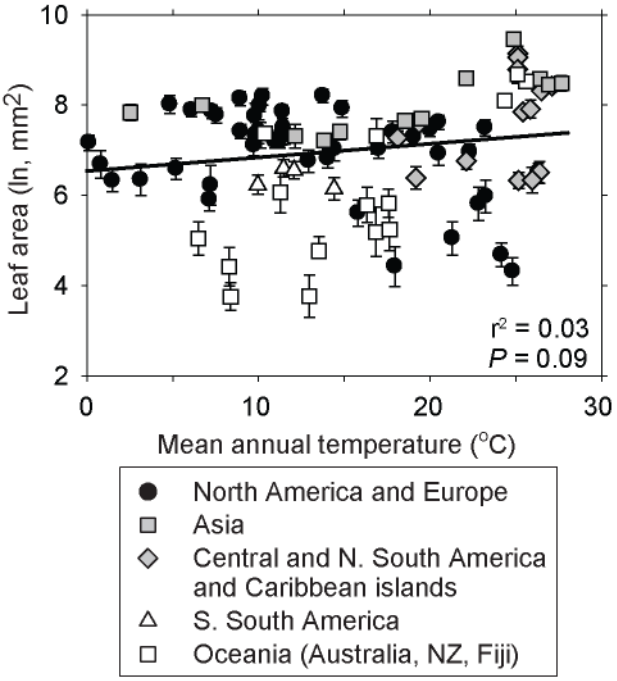

Leaf area (ln, mm2)
10
8
6
4
2
0
10
20
30
Mean annual temperature (°C)
r2 = 0.03
P = 0.09
North America and Europe
Asia
Central and N. South America and Caribbean islands
S. South America
Oceania (Australia, NZ, Fiji)

Figure S2

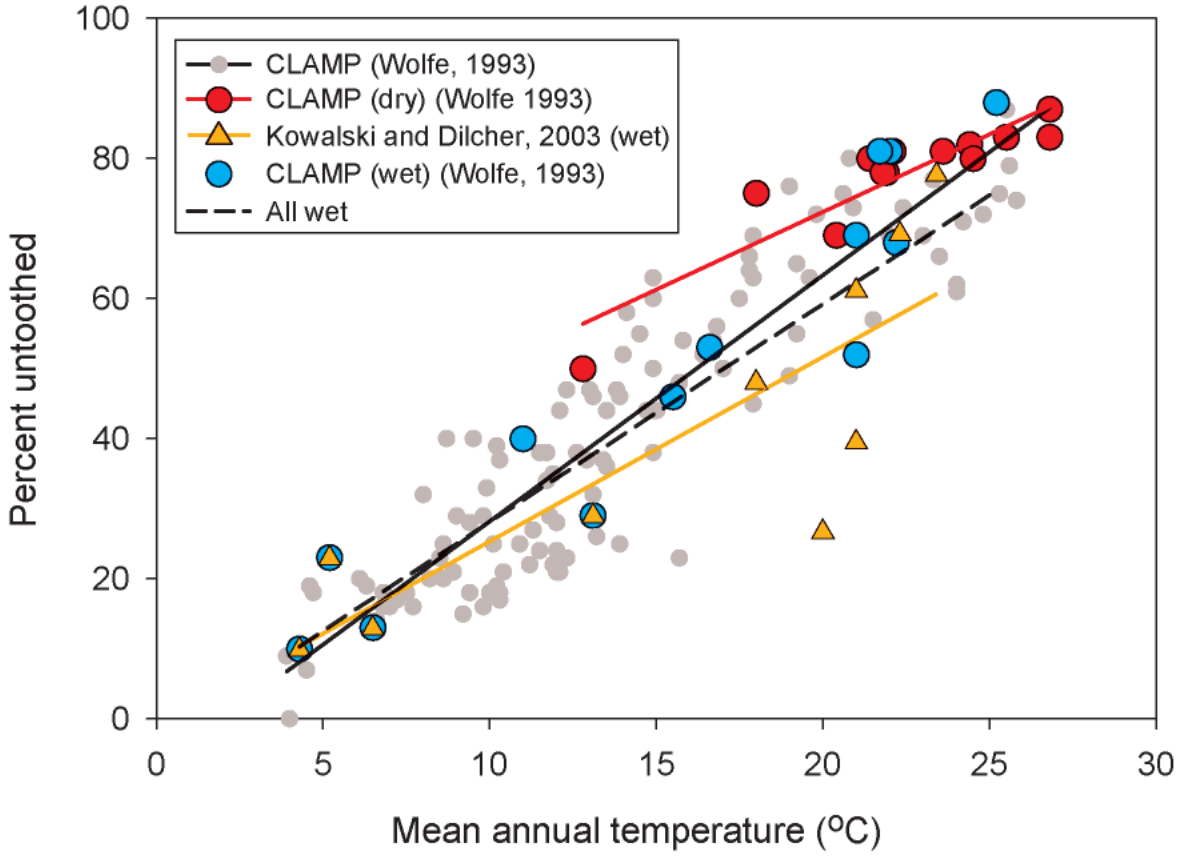

CLAMP (Wolfe, 1993)
CLAMP (dry) (Wolfe 1993)
Kowalski and Dilcher, 2003 (wet)
CLAMP (wet) (Wolfe, 1993)
All wet
Percent untoothed
Mean annual temperature (°C)
0
20
40
60
80
100
0
5
10
15
20
25
30

Figure S3

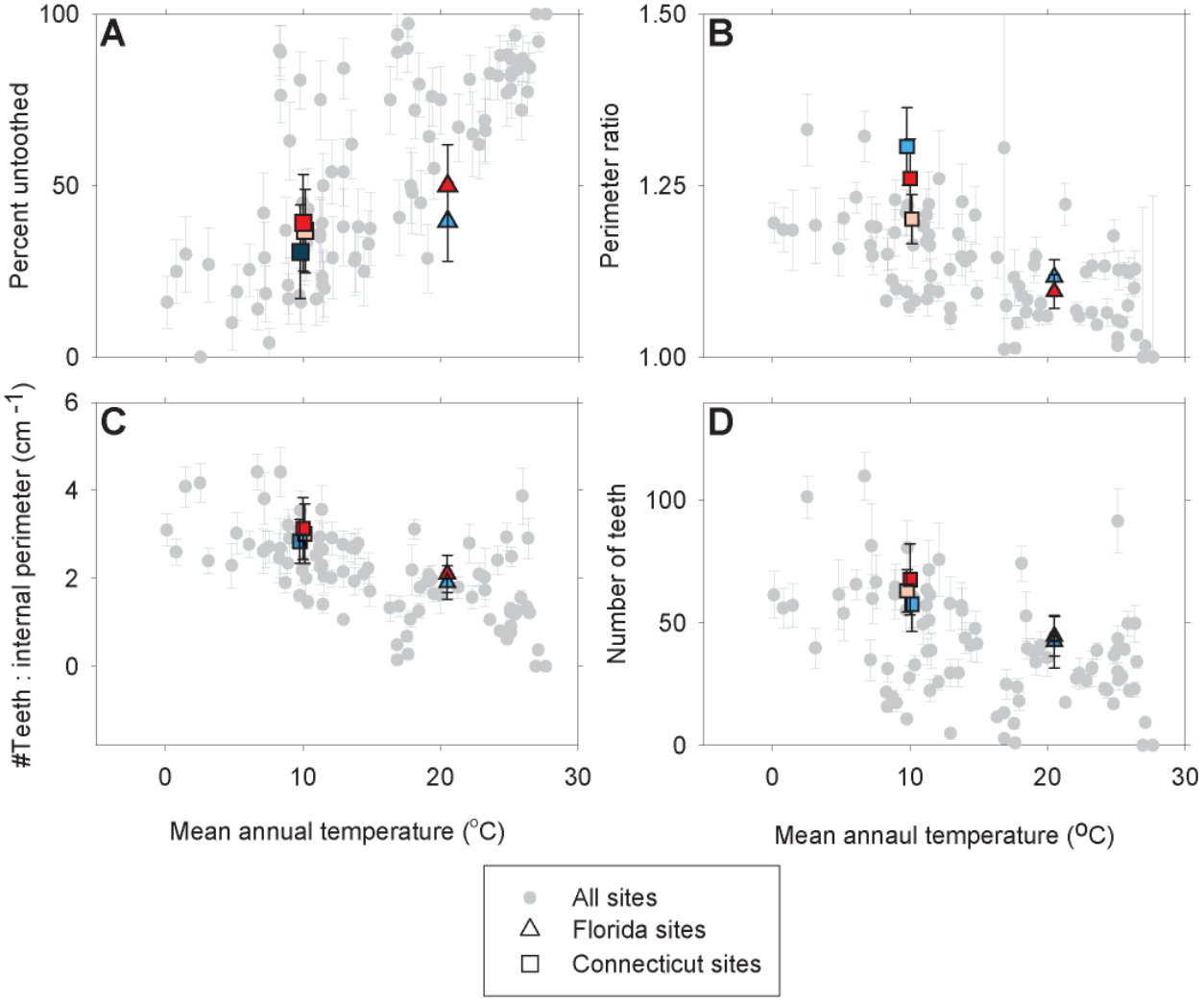

A
B
C
D
Percent untoothed
Perimeter ratio
#Teeth : internal perimeter (cm -1)
Number of teeth
Mean annual temperature (°C)
Mean annaul temperature (°C)
All sites
Florida sites
Connecticut sites

Figure S4

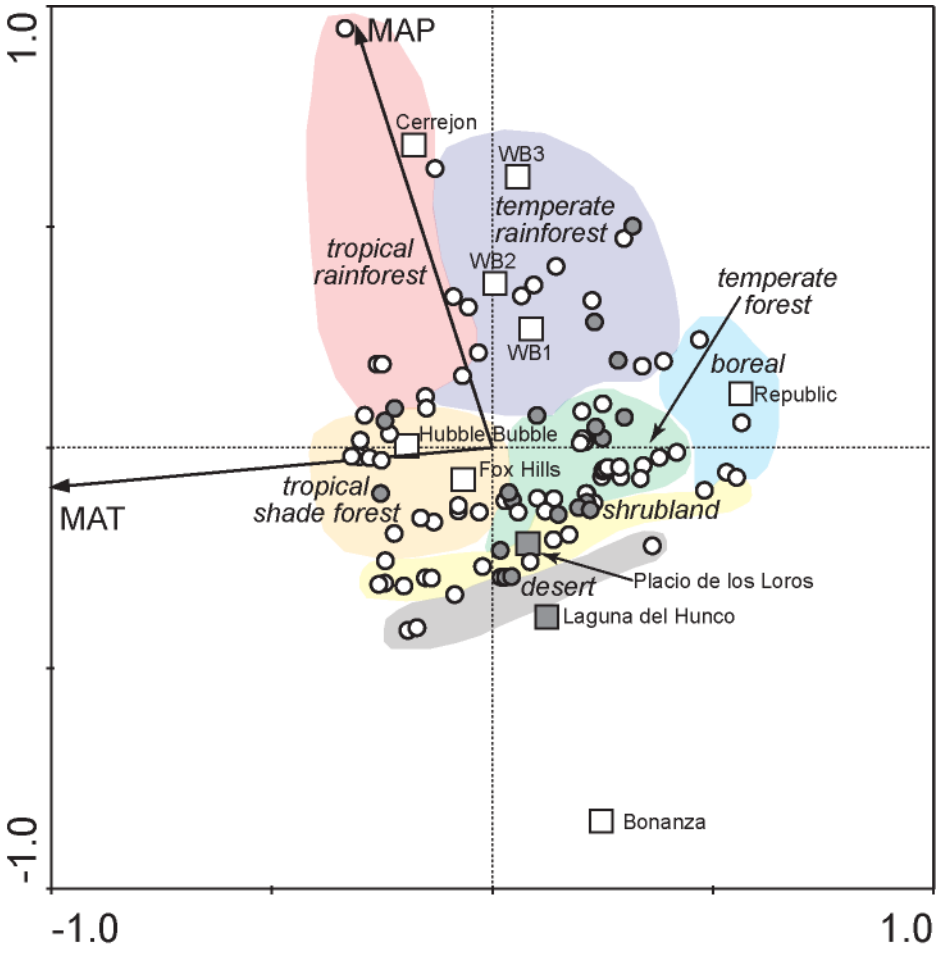

MAP
MAT
1.0
-1.0
-1.0
1.0
Cerrejon
WB3
temperate
rainforest
tropical
rainforest
WB2
WB1
temperate
forest
boreal
Republic
Hubble Bubble
Fox Hills
tropical
shade forest
shrubland
desert
Placio de los Loros
Laguna del Hunco
Bonanza

Figure S5

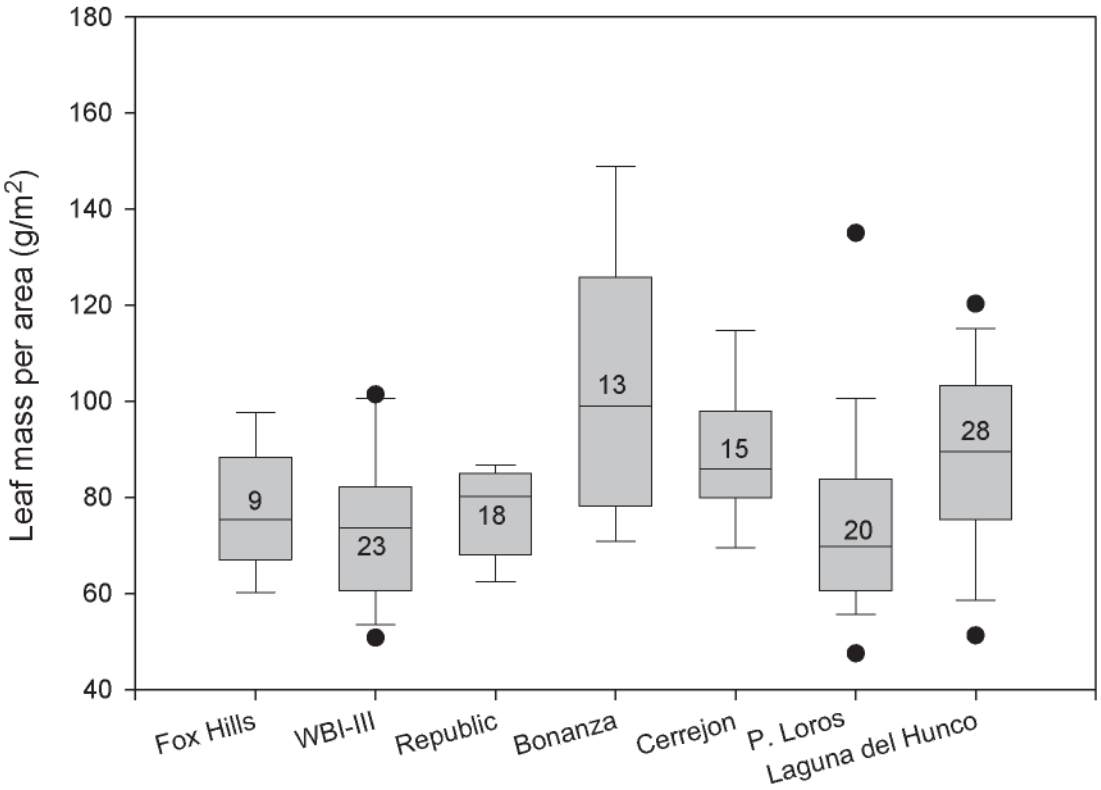
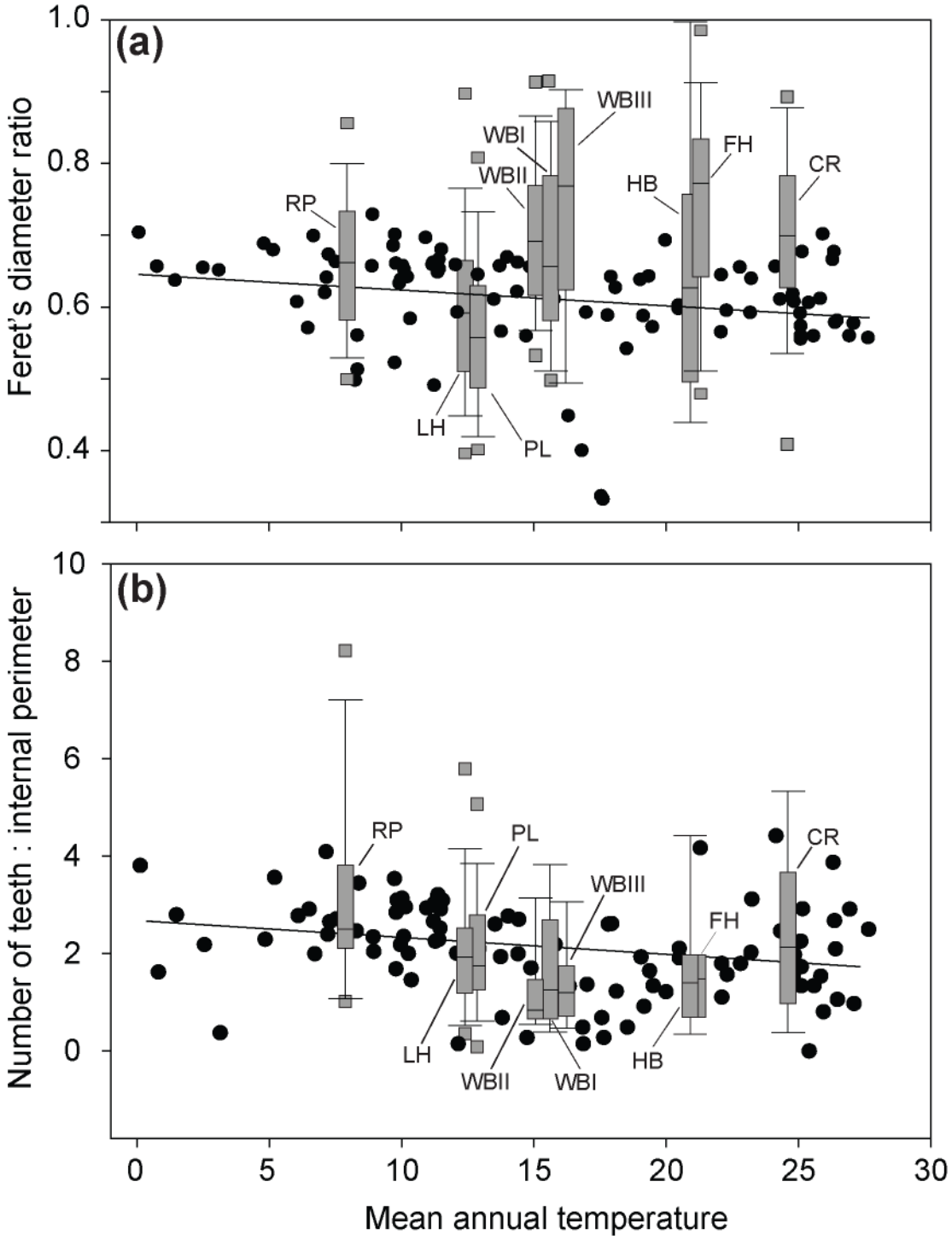

Leaf mass per area (g/m2)
180
160
140
120
100
80
60
40
9
23
18
13
15
20
28
Fox Hills
WBI-III
Republic
Bonanza
Cerrejon
P. Loros
Laguna del Hunco

Figure S6


(a)
Feret's diameter ratio
1.0
0.8
0.6
0.4
RP
LH
WBII
WBI
WBIII
PL
HB
FH
CR
(b)
Number of teeth : internal perimeter
10
8
6
4
2
0
RP
LH
WBII
PL
WBI
WBIII
HB
FH
CR
0
5
10
15
20
25
30
Mean annual temperature