Sensitivity of leaf size and shape to climate: global patterns and paleoclimatic applications

Daniel J. Peppe[1,2], Dana L. Royer[1], Bárbara Cariglino[3,4], Sofia Y. Oliver[1], Sharon Newman[1], Elias Leight[1], Grisha Enikolopov[1], Margo Fernandez-Burgos[1], Fabiany Herrera[5], Jonathan M. Adams[6], Edwin Correa[7], Ellen D. Currano[8], J. Mark Erickson[9], Luis Felipe Hinojosa[10], John W. Hoganson[11], Ari Iglesias[12], Carlos A. Jaramillo[7], Kirk R. Johnson[13], Gregory J. Jordan[14], Nathan J. B. Kraft[15], Elizabeth C. Lovelock[16,17], Christopher H. Lusk[18], Ülo Niinemets[19], Josep Peñuelas[20], Gillian Rapson[21], Scott L. Wing[22] and Ian J. Wright[18]

[1]Department of Earth and Environmental Sciences, Wesleyan University, Middletown, CT 06459, USA; [2]Department of Geology, Baylor University, Waco, TX 76798, USA; [3]Department of Geosciences, Pennsylvania State University, University Park, PA 16802, USA; [4]Museo Argentino de Ciencias Naturales 'B. Rivadavia', Buenos Aires, Argentina; [5]Florida Museum of Natural History and Department of Biology, University of Florida, Gainesville, FL 32611, USA; [6]Department of Biological Sciences, Seoul National University, Seoul 151, Korea; [7]Smithsonian Tropical Research Institute, Apartado Postal 0843-03092, Balboa, Ancon, Panama; [8]Department of Geology, Miami University of Ohio, Oxford, OH 45056, USA; [9]Department of Geology, St Lawrence University, Canton, NY 13617, USA; [10]Faculty of Science, University of Chile, Institute of Ecology and Biodiversity, Santiago, Chile; [11]North Dakota Geological Survey, Bismarck, ND 58505, USA; [12]Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, 1900 La Plata, Argentina; [13]Denver Museum of Nature and Science, Denver, CO 80205, USA; [14]School of Plant Science, University of Tasmania, Private Bag 55, Hobart 7001, Tas., Australia; [15]Biodiversity Research Centre, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; [16]John Day Fossil Beds National Monument, National Parks Service, Kimberly, OR 97848, USA; [17]Department of Earth Science, University of California Santa Barbara, Santa Barbara, CA 93106, USA; [18]Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia; [19]Institute of Agricultural and Environmental Sciences, Estonian University of Life Sciences, Tartu 51014, Estonia; [20]Global Ecology Unit CREAF-CEAB-CSIC, Center for Ecological Research and Forestry Applications, Universitat Autònoma de Barcelona, Barcelona, Spain; [21]Ecology Group, Institute of Natural Resources, Massey University, Private Bag 11222, Palmerston North, New Zealand; [22]Department of Paleobiology, Smithsonian Institution, National Museum of Natural History, Washington, DC 20013, USA

Summary

• Paleobotanists have long used models based on leaf size and shape to reconstruct paleoclimate. However, most models incorporate a single variable or use traits that are not physiologically or functionally linked to climate, limiting their predictive power. Further, they often underestimate paleotemperature relative to other proxies.

• Here we quantify leaf–climate correlations from 92 globally distributed, climatically diverse sites, and explore potential confounding factors. Multiple linear regression models for mean annual temperature (MAT) and mean annual precipitation (MAP) are developed and applied to nine well-studied fossil floras.

• We find that leaves in cold climates typically have larger, more numerous teeth, and are more highly dissected. Leaf habit (deciduous vs evergreen), local water availability, and phylogenetic history all affect these relationships. Leaves in wet climates are larger and have fewer, smaller teeth. Our multivariate MAT and MAP models offer moderate improvements in precision over univariate approaches (± 4.0 vs 4.8°C for MAT) and strong improvements in accuracy. For example, our provisional MAT estimates for most North American fossil floras are considerably warmer and in better agreement with independent paleoclimate evidence.

• Our study demonstrates that the inclusion of additional leaf traits that are functionally linked to climate improves paleoclimate reconstructions. This work also illustrates the need for better understanding of the impact of phylogeny and leaf habit on leaf–climate relationships.

Introduction

The sizes and shapes (physiognomy) of leaves correlate strongly with temperature and moisture from global to local scales, and there are biological bases for these relationships (Bailey & Sinnott, 1915, 1916; Webb, 1968; Lewis, 1972; Givnish, 1979, 1984; Wolfe, 1979, 1993; Hall & Swaine, 1981; Richards, 1996; Wilf, 1997; Wilf *et al.*, 1998; Jacobs, 1999, 2002; Feild *et al.*, 2005; Traiser *et al.*, 2005; Royer & Wilf, 2006). Paleobotanists have long used these leaf–climate correlations to develop proxies for reconstructing paleoclimate (Bailey & Sinnott, 1915, 1916; Dilcher, 1973; Wing & Greenwood, 1993; Wolfe, 1993, 1995; Wilf, 1997; Wilf *et al.*, 1998; Jacobs, 1999, 2002; Kowalski & Dilcher, 2003; Traiser *et al.*, 2005; Adams *et al.*, 2008).

One key leaf–climate association is between leaf teeth and both temperature and local water availability (Baker-Brosh & Peet, 1997; Feild *et al.*, 2005; Royer & Wilf, 2006). The percentage of woody, non-monocotyledonous angiosperms (woody dicots) at a site with toothed leaves (Bailey & Sinnott, 1916; Wolfe, 1979; Wilf, 1997), as well as variables related to tooth count and tooth size (Huff *et al.*, 2003; Royer *et al.*, 2005), all negatively correlate with mean annual temperature (MAT). The prevalence of leaf teeth in cool climates is potentially an adaptation for increased carbon uptake through enhanced sap flow early in the growing season (Billings, 1905; Bailey & Sinnott, 1916; Wolfe, 1993; Baker-Brosh & Peet, 1997; Wilf, 1997; Royer & Wilf, 2006). In cold environments, this early-season pulse in sap flow may allow plants with toothed leaves to maximize the duration of their growing seasons; in warmer climates, the potential benefit is outweighed by the attendant water costs (Wing *et al.*, 2000; Royer & Wilf, 2006). The relationship between leaf teeth and enhanced sap flow may also help explain why, at a given MAT, toothed species are sometimes more abundant in locally wet environments where the water cost associated with teeth may be less important (the 'freshwater-margin effect' in and near swamps, and near lakes and streams; Wolfe, 1993; Burnham *et al.*, 2001; Kowalski & Dilcher, 2003; Greenwood, 2005; Royer *et al.*, 2009a). Teeth may also release excess root pressure through guttation, preventing the flooding of intercellular spaces in the leaf lamina and, in cooler climates, freeze–thaw embolisms (Feild *et al.*, 2005).

Leaf size is also sensitive to climate: site-mean leaf size typically scales with water availability and, to a lesser degree, temperature (Webb, 1968; Dilcher, 1973; Dolph & Dilcher, 1980a,b; Givnish, 1984; Greenwood, 1992; Wilf *et al.*, 1998). Energy balance models predict that for a given level of radiation and wind speed, leaf temperatures are higher in large canopy leaves because of their thicker boundary layers (Vogel, 1968, 1970, 2009; Parkhurst & Loucks, 1972; Givnish, 1979, 1984, 1987; Gates, 1980).

Warmer leaf temperatures promote both photosynthesis and transpiration; thus, plants in drier climates tend to have smaller leaves to reduce evaporative cooling, while in more humid climates larger leaves are common because the attendant water cost is less critical (Givnish, 1984).

Other factors can affect these leaf–climate relationships. It has been commonly claimed, but never rigorously tested, that deciduous species are more likely to be toothed than evergreen species (Bailey & Sinnott, 1916; Givnish, 1979; Wolfe, 1993; Jacobs, 2002). Shared phylogenetic and/or regional histories of floras may also be important. Multiple studies have noted different leaf–climate relationships in the northern and southern hemispheres, with extant southern hemisphere temperate floras typically having a higher percentage of untoothed species than temperature-equivalent northern hemisphere floras (Greenwood, 1992; Jordan, 1997; Jacobs, 1999, 2002; Kennedy *et al.*, 2002; Kowalski, 2002; Greenwood *et al.*, 2004; Aizen & Ezcurra, 2008; Hinojosa *et al.*, 2010; Steart *et al.*, 2010). These differences may be the result of regional differences in environment, such as soil fertility and thermal seasonality, and/or phylogenetic differences (Wolfe and Upchurch, 1987; Jordan, 1997; Greenwood *et al.*, 2004). Other regional differences in leaf–climate relationships exist, although often the differences are not statistically significant (e.g. Gregory-Wodzicki, 2000; Traiser *et al.*, 2005; Miller *et al.*, 2006; Su *et al.*, 2010).

To address these potential problems, regional calibrations have been developed (for example, see Hinojosa *et al.*, 2010; Su *et al.*, 2010) and make the assumption that leaf–climate relationships within a region were the same as they are now. This is a valid assumption in some cases (e.g. late Neogene and Quaternary floras), but not in others (e.g. Cretaceous and early Cenozoic floras), particularly given the uncertainty in the cause for the difference and the major environmental and evolutionary changes since the Cretaceous. If phylogeny is important, then regional calibrations assume that past lineage composition of the fossil flora was similar to the current composition in the region, and that evolution and extinction subsequent to the deposition of the fossils has not changed leaf–climate relationships in those lineages (Jordan, 1997; Hinojosa *et al.*, 2010; Little *et al.*, 2010). If current environment drives regional differences, then regional calibrations must assume that critical environmental features, such as soil fertility and thermal seasonality, were the same in the relevant region at the time of deposition of the fossils, another questionable assumption. Overall, the effects of phylogeny and regional environmental differences on leaf–climate correlations are poorly constrained and have rarely been tested in a proper statistical framework (Hinojosa *et al.*, 2010; Little *et al.*, 2010). As more detailed large-scale assessments of the relationship between phylogeny and leaf traits become available (Little *et al.*, 2010), comparing the leaf–climate correlations

in this and other studies to related methods that incorporate phylogenetic relationships (Felsenstein, 1985; Garland *et al.*, 1992; Westoby *et al.*, 1998) will likely provide additional insights into the ecological and evolutionary forces shaping trait–climate correlations.

The most common leaf physiognomic methods for estimating MAT and mean annual precipitation (MAP), leaf-margin analysis and leaf-area analysis, are each based on a single variable, the percentage of untoothed species at a site and site-mean leaf size, respectively (Wolfe, 1979; Wilf, 1997, 1998; Jacobs, 2002; Miller *et al.*, 2006). Although climate estimates from these methods commonly agree with independent evidence (e.g. Greenwood & Wing, 1995; Wing *et al.*, 2000; Uhl *et al.*, 2003; Wilf *et al.*, 2003a,b; Mosbrugger *et al.*, 2005; Yang *et al.*, 2007; Greenwood *et al.*, 2010), there are many instances where these proxies provide cooler and drier estimates of MAT and MAP than alternative proxy evidence (Utescher *et al.*, 2000; Liang *et al.*, 2003; Fricke & Wing, 2004; Kvacek, 2007; Wing *et al.*, 2009b). Because these are univariate approaches, additional characters may lead to improvements.

To this end, Wolfe (1993, 1995) developed a method called Climate-Leaf Analysis Multivariate Program (CLAMP), which uses 31 categorical leaf states, including leaf-margin and leaf-size categories. The method correlates the characters to climate using canonical correspondence analysis (CCA; Wolfe, 1995). Because CLAMP more thoroughly describes leaf physiognomy, it might be expected to result in more accurate climate estimates than the univariate approaches, but in practice it does not (Jacobs & Deino, 1996; Wilf, 1997; Wiemann *et al.*, 1998; Gregory-Wodzicki, 2000; Kowalski & Dilcher, 2003; Royer *et al.*, 2005; Dilcher *et al.*, 2009; Smith *et al.*, 2009b). This may be caused by errors and biases related to the ambiguity of character definitions, the categorical nature of the character states, weak or non-existent correlations between climate and some character states, and problems related to using CCA in a predictive framework (Jordan, 1997; Wilf, 1997; Wilf *et al.*, 1998, 1999; Green, 2006; Peppe *et al.*, 2010). Thus, although CLAMP is multivariate, it is fraught with systemic problems and does not produce more accurate climate estimates. Other multivariate approaches have been proposed (Wing & Greenwood, 1993; Stranks & England, 1997; Gregory-Wodzicki, 2000), but because they use the CLAMP characters they suffer from many of the same problems.

Recently, Huff *et al.* (2003) and Royer *et al.* (2005) developed a new procedure, called digital leaf physiognomy, which has three major advantages over CLAMP and the univariate approaches. First, it minimizes the ambiguity of CLAMP scoring because computer algorithms process most of the measurements. Second, it uses mostly continuous variables, such as tooth number and size, not categorical characters. Thus, for example, digital leaf physiognomy can discern between a leaf with one and 100 teeth, but CLAMP and leaf-margin analysis do not (Royer *et al.*, 2005, 2008). Third, digital leaf physiognomy incorporates more traits that have a functional and/or physiological connection to climate, such as tooth number, tooth size, leaf area and degree of leaf dissection (see earlier discussion). Importantly, the traits used in digital leaf physiognomy can display some degree of phenotypic plasticity (Royer *et al.*, 2009b), suggesting they can respond quickly to climate change even in the absence of evolutionary responses.

Using digital leaf physiognomy, Huff *et al.* (2003) and Royer *et al.* (2005) observed that leaves from cold climates are more likely to be highly dissected and to have many, large teeth; importantly, these correlations are consistent with the ecophysiological principles outlined earlier. Royer *et al.* (2005) also developed a preliminary, multiple linear regression model for predicting MAT that was considerably more accurate than leaf-margin analysis and CLAMP. A limitation of the study, however, is that it was based on 17 sites from eastern North America and Panama that spanned a limited biogeographic and climatic range (Fig. 1).

Here, we investigate correlations between leaf physiognomy and climate across 92 globally distributed sites from the biomes where fossil leaves are most likely to be preserved (Fig. 1). A major goal of the study was to assess global correlations of MAT and MAP to functionally linked leaf traits using a phylogenetically and climatically diverse data set of extant vegetation (Fig. 1). In addition, we quantitatively tested the importance of two potential confounding factors on these correlations: the evergreen effect (i.e. are woody dicot evergreens less likely to be toothed?) and the freshwater-margin effect (i.e. do freshwater-margin habitats contain a higher percentage of toothed species?). We also compared leaf–climate correlations between extant northern and southern hemisphere floras; however, it is beyond the scope of the present study to employ more formal phylogenetic tests (e.g. Little *et al.*, 2010). Third, we developed multiple linear regression equations derived from the extant vegetation to estimate MAT and MAP. To gauge the accuracy of the equations, we estimated the climate of each extant site using a jack-knife-type approach. We then applied the equations to nine, well-studied fossil floras and compared the climate reconstructions to other climate proxies, including leaf-margin analysis and leaf-area analysis.

Materials and Methods

Calibration sites

We photographed leaves of native, woody dicots from 92 geographically and climatically diverse extant sites (Fig. 1) (*n* = 6525 leaves and 3033 species-site pairs). This data set expands on the 17 calibration sites of Royer *et al.* (2005).
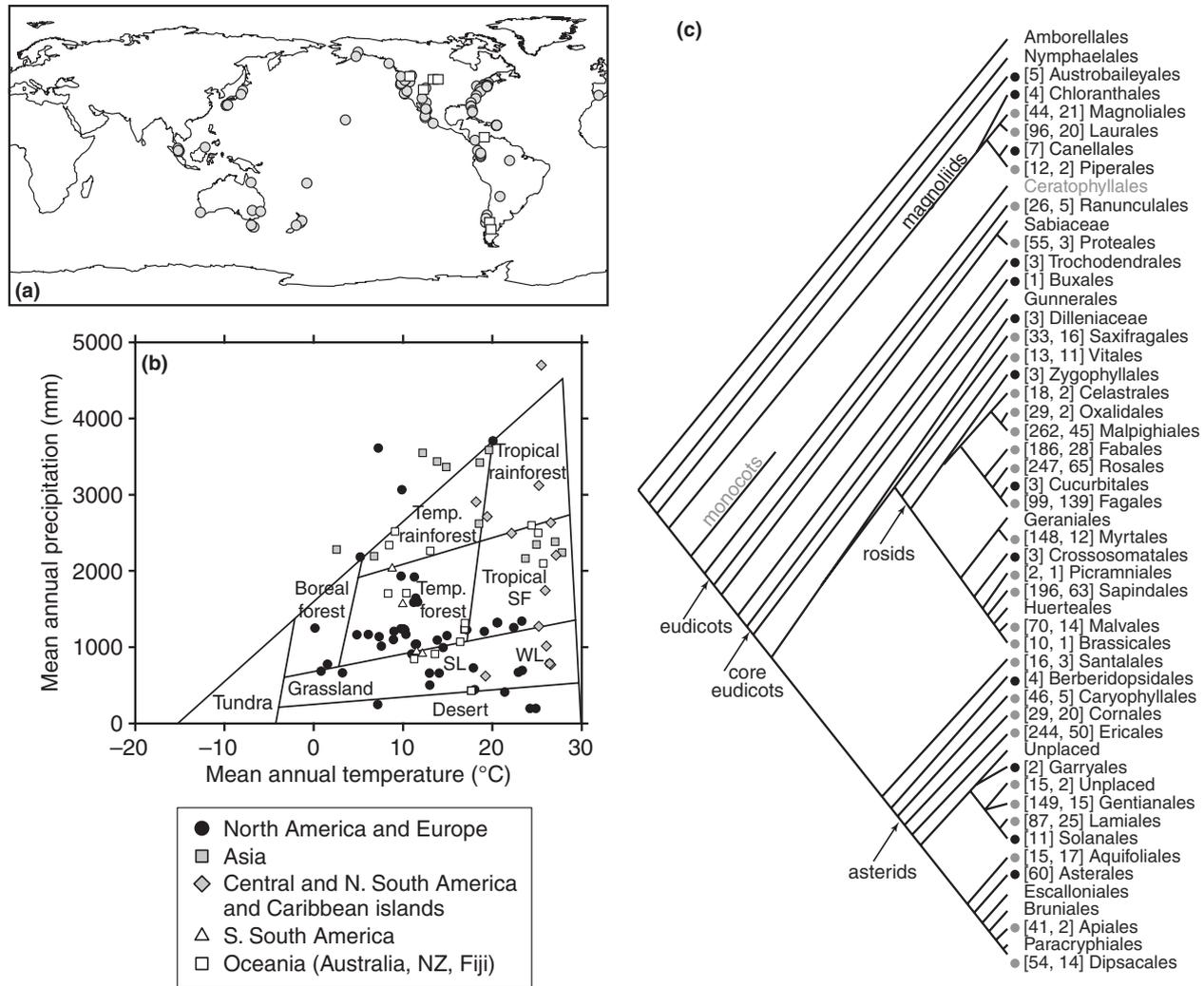
Fig. 1 Geographic, climatic and phylogenetic distribution of data. (a) Geographic distribution of calibration sites (grey circles) and fossil sites (open squares). The paleolatitude of each fossil site is given in Table 1. (b) Climatic distribution of calibration sites. Biomes follow Whittaker (1975) and their boundaries are approximate and do not encompass all samples. SF, seasonal forest; SL, shrubland; WL, woodland. 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. See the Supporting Information, Notes S1 and Table S1, for additional information about sites. (c) Phylogenetic distribution of calibration data. Closed circles represent orders that have been added to the calibration since Royer *et al.* (2005). The first number in brackets is the number of species–site pairs from the 75 new sites; the second number, when present, is the number of pairs from the 17 sites of Royer *et al.* (2005). Ceratophyllales (tinted) is composed solely of herbaceous taxa and is thus not applicable to our study; the monocot clade is also not applicable. Tree follows APG III (Stevens, 2001 onwards; The Angiosperm Phylogeny Group, 2009).

The majority of new sites (*n* = 42) come from the CLAMP collection (Wolfe, 1993; Spicer, 2009), whose voucher specimens are housed in the Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA. Sampling was generally restricted to outer, exposed leaves in the canopy or tree crown (see the Supporting Information for detailed collection protocols). To test the potential of herbs as climate indicators, a collection of 34 herbaceous dicot species was made from north of Reed Gap in Wallingford, Connecticut (see Royer *et al.*, 2010 for sampling details).

Mean annual temperature of our sites ranged from 0.1 to 27.7°C and MAP from 189 to 4694 mm (see the Supporting Information, Table S1). Mean monthly climate data were extracted from a global, interpolated 1 km spatial resolution climate model (WORLDCLIM, Hijmans *et al.*, 2005). Where available, WORLDCLIM matches local climate station data at all but five sites for MAT (± 0.3°C) and three sites for MAP (± 22 mm). For the seven sites where the model deviated strongly from station data (> ± 2.0°C or ± 100 mm), we relied on the latter. We defined the growing season as the period during which the

mean monthly minimum temperature exceeded 0°C and precipitation exceeded 20% of the maximum monthly precipitation, and growing-degree-days as the number of degree-days in a year when the average temperature exceeded 10°C (Table S2) (e.g. Johnson *et al.*, 2000). The functional basis of leaf physiognomy (see the Introduction) may imply that physiognomic traits are more closely linked to growing-season variables such as growing-season precipitation, growing-season mean temperature and mean annual range in temperature (warmest month mean minus coldest month mean); however, we focus here on MAT and MAP because correlations of leaf physiognomy to annual and growing-season climate variables were very similar (Table S3).

Typically, at least two leaves or leaflets per species at each site were used. More than two leaves were used if there was a large variation in leaf form (e.g. compound leaves, species with and without lobes or teeth). Computerized resampling indicates that this level of sampling is sufficient for detecting site-level patterns (Royer *et al.*, 2005). All leaf images used in this study are available from Dryad (http://dx.doi.org/10.5061/dryad.8101) and the personal websites of DJP and DLR. Leaves were manipulated in Adobe Photoshop (Adobe Systems, San Jose, CA, USA) to separate the petiole and teeth (if present) from the blade following the protocols of Royer *et al.* (2005). Most physiognomic characters were calculated using IMAGEJ (http://rsbweb.nih.gov/ij/); presence of teeth and number of teeth were determined visually (see Table S4 for all physiognomic data). Definitions of characters follow Royer *et al.* (2005) (see also Table S2). Site means (Table S1) were calculated from species means. For variables involving teeth, untoothed species were excluded in order to maintain normal distributions (Huff *et al.*, 2003). Because climate impacts leaf physiognomy, we plot climate as the independent variable and leaf traits as the dependent variables. Site-mean data were correlated to climate with single and multiple linear regression (SPSS 17; SPSS Science, Chicago, IL, USA) and with CCA (CANOCO 4.5; Microcomputer Power, Ithaca, NY, USA). Using leaf traits as the independent variables and climate as the dependent variable, we developed predictive multiple linear regression models for MAT and MAP. The variables shape factor (perimeter^2/blade area), compactness ($4\pi \times$ blade area/perimeter^2), number of teeth, tooth area and perimeter/area cannot be calculated in any meaningful way for fragmentary fossils and were excluded from our models (Royer *et al.*, 2005). However, these traits may be useful for studying extant leaf–climate relationships (Royer *et al.*, 2005, 2008; see Table S5 for most significant MAT and MAP models derived using all variables). Models were considered only if: the model and all individual variables in the model were significant at the $\alpha = 0.05$ level, and variables did not show a high degree of co-linearity with the other predictor variables (variance inflation factor < 10; Sokal & Rohlf, 1995). We used the

ordinary least squares regression module in the program SMATR (http://www.bio.mq.edu.au/ecology/SMATR/; Warton *et al.*, 2006) to test for slope and intercept differences between regression lines. We define accuracy as the extent to which a given MAT or MAP estimate agrees with other independent lines of evidence. Precision is defined as uncertainty of an estimate derived from a regression model (i.e. the standard error).

Fossil sites

We applied the digital leaf physiognomy MAT and MAP models, as well as leaf-margin analysis and leaf-area analysis, to 10 fossil floras from the latest Cretaceous and early Paleogene (*c.* 66 to *c.* 47.0 million years ago (Ma)) of North and South America (Fig. 1, Table 1). All floras are well-studied and represent a broad range of interpreted biomes and phylogenetic histories. For each site, we processed 1–48 specimens of each species or morphotype (median = 3; see Tables S6, S7 for all fossil physiognomic data). As fossil specimens are in rock matrix and often fragmentary, additional processing protocols were necessary (Cariglino, 2007; see Methods S1). Because it is possible to determine the margin type (toothed, untoothed) of specimens that cannot be digitally processed, we calculated the percentage of untoothed species based on all species, not just the digitally-processed species.

The Fox Hills flora is from the Linton Member of the Fox Hills Formation and is late Maastrichtian in age (*c.* 66 Ma; Peppe, 2003; Peppe *et al.*, 2007; Table 1). Specimens are stored at the North Dakota Heritage Center in Bismarck, North Dakota, USA, and at St Lawrence University in Canton, New York, USA.

The Fort Union Formation floras (Williston Basin I, II, and III) are from the Fort Union Formation in the Williston Basin of southwestern North Dakota, USA (65.5 to *c.* 58.5 Ma; Peppe, 2009, 2010; Table 1). We grouped these taxa by floral zone following Peppe (2009, 2010). Specimens used in this study are housed at the Yale Peabody Museum in New Haven, Connecticut, USA.

The Palacio de los Loros flora (P. Loros), first described in Berry (1937), is from the westernmost exposures of the Salamanca Formation in southern Chubut Province, Argentina, and is early Paleocene in age (*c.* 61.7 Ma; Iglesias *et al.*, 2007; Table 1). Specimens used in this study come from two outcrops representing the same general depositional environment that are geographically and stratigraphically close to each other (Iglesias *et al.*, 2007). The specimens are reported by Iglesias *et al.* (2007) and are housed at the Museo Paleontológico Egidio Feruglio in Trelew, Argentina.

The Cerrejón flora is from the middle Late Paleocene (*c.* 58 Ma) Cerrejón Formation of Colombia reported by Wing *et al.* (2009b; Table 1). Specimens are housed at INGEOMINAS in Bogotá, Colombia.

Table 1 Age, paleolatitude, number of species, and provisional mean annual temperature and mean annual precipitation estimates for fossil floras

Site	Age (Ma)	Paleolatitude[a]	Number of woody dicotyledonous angiosperm species in flora	Number of woody dicotyledonous angiosperm species processed	Digital leaf physiognomy MAT estimate (°C)[b]	Leaf-margin analysis MAT estimate (°C)[c]	Regional digital leaf physiognomy MAT estimate (°C)[d]	Digital leaf physiognomy MAP estimate (cm)[e]	Leaf-area analysis MAP estimate (cm)[e]
Fox Hills	66.5	49.7	34	25	21.6	14.8	17.1	141 (+116, −64)	152 (+125, −68)
Williston Basin I	65.5–64.0	50.8	26	20	15.7	10.9	12.6	175 (+144, −79)	157 (+129, −71)
Williston Basin II	64.0–63.0	50.8	31	23	15.0	10.2	12.4	148 (+122, −69)	156 (+129, −71)
Williston Basin III	61.0–58.5	50.8	19	18	16.3	9.4	11.6	152 (+125, −68)	157 (+129, −71)
Palacio de los Loros	61.7	−54.7	36	33	12.8	12.9	12.8	125 (+103, −56)	144 (+119, −65)
Cerrejon	58	7.4	48	48	23.6	20.5	14.0	264 (+217, −119)	212 (+174, −96)
Hubble Bubble	55.8	47.6	29	16	20.3	17.9	20.2	147 (+121, −66)	146 (+120, −66)
Laguna del Hunco	51.9	−49.0	132	119	10.9	14.1	16.9	127 (+103, −57)	142 (+117, −64)
Republic	49.4	50.9	45	41	9.0	9.2	8.9	134 (+110, −60)	135 (+111, −61)
Bonanza	47.3	40.4	28	24	[f]	14.8	[f]	[f]	110 (+90, −50)

Ma, million years ago; MAT, mean annual air temperature; MAP, mean annual precipitation.
[a]Paleolatitude reconstruction based on Torsvik *et al.* (2008).
[b]Standard Error (SE) is ± 4.0°C.
[c]Standard Error is ± 4.8°C. Independent proxy evidence suggests that most of these MAT estimates are considerable underestimates (see text).
[d]Regional digital leaf physiognomy models were created for North America and South America. The North American model ($r^2 = 0.81$, SE = ± 3.3°C) used the variables percent untoothed and number of teeth : internal perimeter. The North American model was based on all extant sites in our calibration from North America, Central America, and Asia and was applied to all fossil sites from North America. The South American model ($r^2 = 0.96$, SE = ± 1.7°C) used the variables percent untoothed and Feret's diameter ratio. The model was based on all extant sites in our calibration from South America and was applied to all fossil sites from South America.
[e]Standard errors are asymmetrical because they were converted from logarithmic units.
[f]MAT and MAP for Bonanza were not reconstructed (see discussion in text).

The Hubble Bubble flora (USNM locality 42384) is from the Willwood Formation in the Bighorn Basin, Wyoming, USA, and dates to within the Paleocene–Eocene thermal maximum (PETM, *c.* 55.8 Ma, Currano *et al.*, 2008, 2010; Wing *et al.*, 2009a; Table 1). Specimens are housed in the Department of Paleobiology, National Museum of Natural History, Smithsonian Institution in Washington, DC, USA.

The Early Eocene Laguna del Hunco flora, which was first described by Berry (1925), is 51.91 ± 0.22 Ma and comes from the Tufolitas Laguna del Hunco, a lacustrine unit in the Chubut River volcanoclastic complex in the northwestern Chubut Province in Patagonia, Argentina (Wilf *et al.*, 2003a, 2005a; Table 1). Specimens are stored at the Museo Paleontológico Egidio Feruglio in Trelew, Argentina.

The Bonanza flora, first described by MacGinitie (1969), is from the uppermost Parachute Creek Member of the Green River Formation in northeastern Utah, USA, and is early Middle Eocene in age (*c.* 47.3 Ma, Smith *et al.*, 2008; Table 1). Specimens studied here are a subset of those reported in Wilf *et al.* (2001). The Republic flora (Wolfe & Wehr, 1987; Radtke *et al.*, 2005) is from the Klondike Mountain Formation in northeastern Washington, USA, and is late Early Eocene in age (49.4 ± 0.5 Ma, Radtke *et al.*, 2005; Table 1). Specimens studied here are a subset

of those reported in Wilf *et al.* (2005b). Both collections are housed at the Denver Museum of Nature and Science in Denver, Colorado, USA.

Results and Discussion

Physiognomic correlation with climate

The site means of many leaf physiognomic characters correlate strongly with temperature and precipitation (Figs 2, 3, Table S3). Notably, MAT correlates significantly to tooth-related characters, including percent of untoothed species ($r^2 = 0.58$, $P < 0.001$), number of teeth ($r^2 = 0.23$, $P < 0.001$), tooth area : internal perimeter ($r^2 = 0.11$, $P = 0.001$; internal perimeter is the leaf perimeter after teeth are removed), and number of teeth : internal perimeter ($r^2 = 0.35$, $P < 0.001$), as well as leaf dissection variables such as perimeter ratio ($r^2 = 0.37$, $P < 0.001$; blade perimeter divided by internal perimeter) and shape factor ($r^2 = 0.22$, $P < 0.001$) (Fig. 2). In warmer climates, leaves generally have fewer, smaller teeth and are less dissected, as previously observed by Royer *et al.* (2005).

Leaf-margin analysis models are currently calibrated with woody dicots because the physiognomy of herbaceous angiosperms is considered to be less sensitive to climate (Bailey & Sinnott, 1916). However, we found that the

percentage of untoothed herbaceous dicot species from a central Connecticut site (35.4%) was almost identical to that of woody dicots species from four nearby sites (mean = 34.5%; see Table S1). There may be potential for including herbaceous taxa in leaf-climate proxies, but further work is needed.

Moisture variables also significantly correlate with several physiognomic characters (Fig. 3, Table S3). Correlations are stronger with $\log_e(MAP)$ than with untransformed MAP, probably owing to the non-normal distribution of MAP across sites (Fig. 1) and/or a non-linear relationship between MAP and water stress. As expected, leaf area posi-

tively correlates with $\log_e(MAP)$ ($r^2 = 0.23$, $P < 0.001$; Fig. 3). Tooth area/blade area inversely correlates with $\log_e(MAP)$ ($r^2 = 0.18$, $P < 0.001$), indicating that tooth area normalized to leaf area declines as precipitation increases (Fig. 3b). Although the functional significance of the relationship between precipitation and tooth area/blade area is unclear, it is consistent with a field study of *Acer rubrum* (Royer *et al.*, 2008).

Water availability is a major control on leaf size, but temperature is also important (see the Introduction). In our calibration, MAT weakly correlates with leaf area ($r^2 = 0.09$, $P = 0.003$; Fig. S1). However, the relationship is
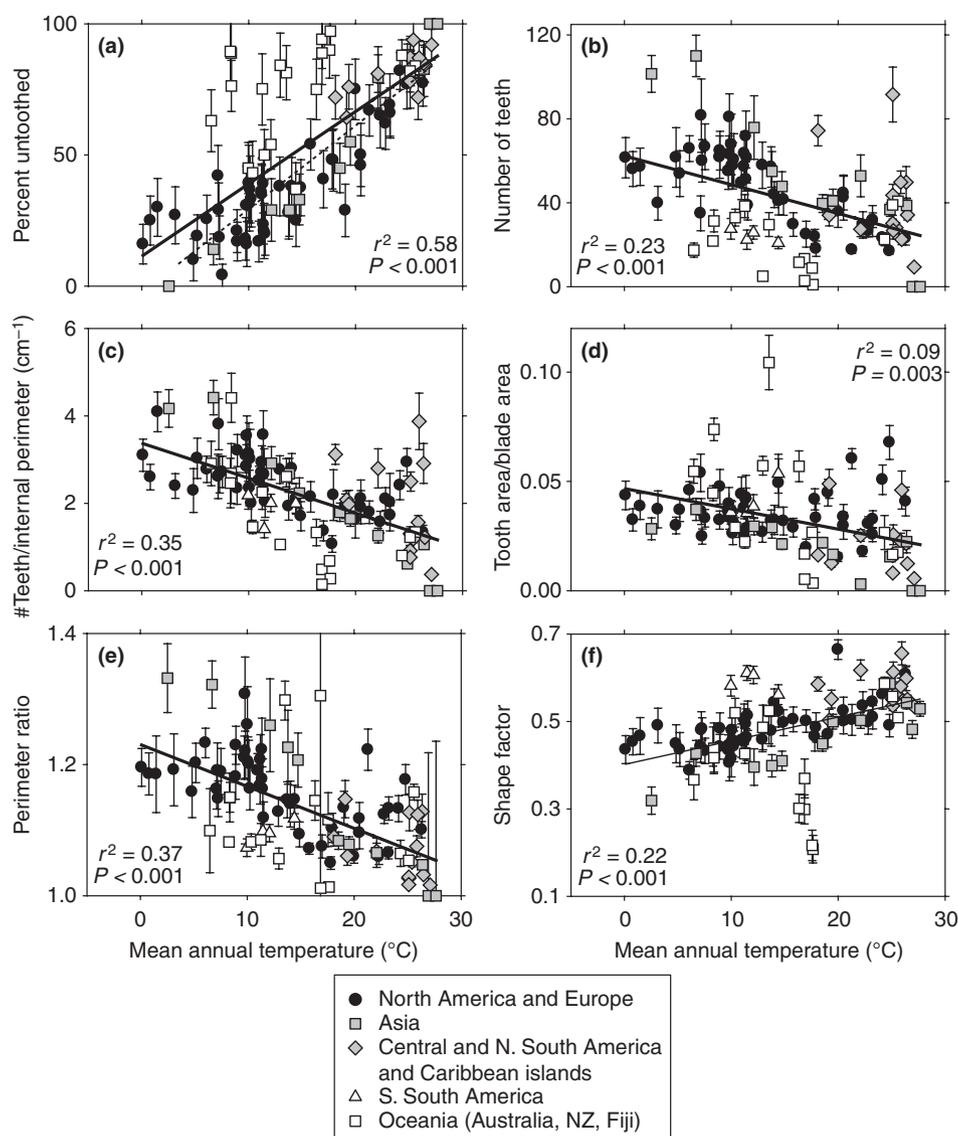

Fig. 2 Relationship between site mean of physiognomic variables and mean annual temperature for the 92 calibration sites. Standard errors of the means for each site are plotted. Linear regression fits and associated r^2 and P values are given in each panel (see also the Supporting Information, Table S3). For comparison, the regression of Wolfe (1979) is plotted in panel (a) (dashed line, $r^2 = 0.98$, $P < 0.001$). Standard errors for percent untoothed character are calculated using Eqn 3 in Miller *et al.* (2006). Internal perimeter is the blade perimeter after teeth are removed, perimeter ratio is blade perimeter divided by internal perimeter, and shape factor is $4\pi \times$ blade area/perimeter^2; see Table S2 for definitions of all variables.
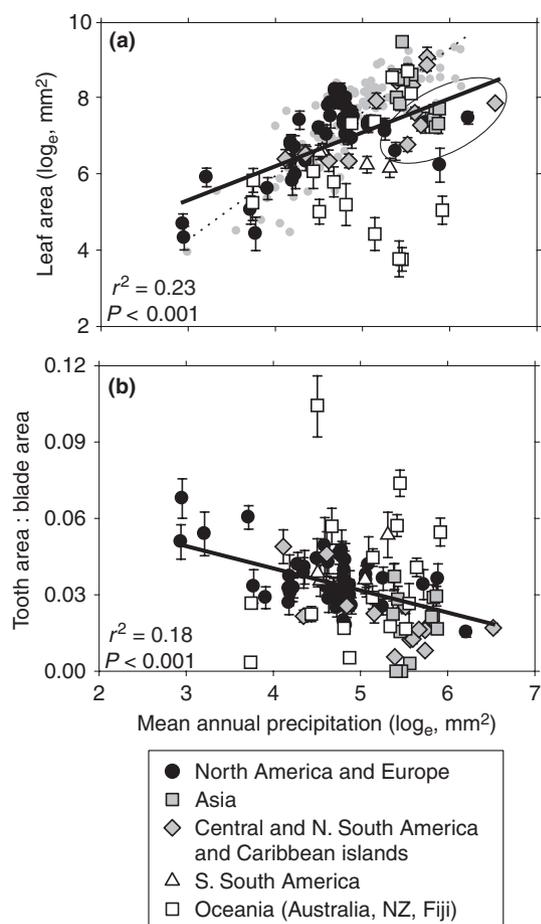
Fig. 3 Relationship between site mean of physiognomic variables and mean annual precipitation for the 92 calibration sites. Standard errors of the site means are plotted. Linear regression fits and associated r^2 and P values are given in each panel (see also the Supporting Information, Table S3). Physiognomic variables are defined in Table S2. For comparison, the leaf area compilation from Jacobs (2002) (grey circles) and associated linear regression (dotted line, $r^2 = 0.70$, $P = <0.001$) is shown in panel (a). Ellipse in panel (a) indicates sites that are warm and wet with relatively small site-mean leaf areas (see Table S1); this climate–physiognomy space is not captured in the Jacobs (2002) compilation. It appears that the correlation between $\log_e$(leaf area) and $\log_e$(MAP) is influenced by sites from Oceania (New Zealand, Australia, Fiji); however, the slope of the regression after these sites are removed is not significantly different ($P = 0.40$) from the full data set. 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.

even weaker after accounting for the covariation between MAT and MAP (Fig. 1b) with partial correlation ($r^2 = 0.07$, $P = 0.01$). As noted by Webb (1968), the leaf size–MAT relationship is strong within the Australia/New Zealand subset (Fig. S1; $r^2 = 0.34$, $P = 0.002$ for partial correlation). These observations raise two points. First, paleoclimate reconstructions based on leaf physiognomy should consider the interactive MAT–MAP control on physiognomy (discussed later). Second, regional differences

in leaf–climate correlations exist (see also Figs 2–3 and the Introduction) and understanding their root causes, whether related to phylogeny, ecology or other factors, will improve paleoclimate reconstructions. Next, we discuss some of these biases.

Potential confounding factors

Freshwater-margin effect Sites with shallow water tables often have a higher percentage of species with teeth (c. 10–15%) than nearby drier sites (e.g. Burnham et al., 2001; Kowalski & Dilcher, 2003; Greenwood, 2005; Royer et al., 2009a). When using leaf-margin analysis to estimate MAT, this freshwater-margin effect could lead to an underestimation of up to 4°C (Burnham et al., 2001; Kowalski & Dilcher, 2003; Greenwood, 2005; Royer et al., 2005, 2009a). Further, the effect may be more severe (up to 10°C) at warmer temperatures (Kowalski & Dilcher, 2003). To test for this bias, we compared the slope of the regression fit between MAT and the percentage of untoothed species in the entire CLAMP data set (Wolfe, 1993) with that of the edaphically wet sites from Kowalski & Dilcher (2003) and Wolfe (1993), and found no statistical difference ($P = 0.12$; Fig. S2). By contrast, the y-intercept of a regression fit for edaphically dry CLAMP sites is shifted towards a higher percentage of untoothed species than that for edaphically wet sites ($P < 0.001$, Fig. S2). Thus, while we detected the freshwater effect, it probably does not strongly affect most paleo-MAT reconstructions because enough calibration sites contain a sufficient proportion of edaphically wet vegetation. The freshwater-margin effect reported by Kowalski & Dilcher (2003) is not representative; instead, a bias of up to 4°C is more plausible (Burnham et al., 2001; Fig. S2). Critically, the additional characters used in digital leaf physiognomy (e.g. number of teeth) generally show less sensitivity to the freshwater-margin effect than does percent of untoothed species (Fig. S3).

Effect of leaf habit and phylogeny Are woody dicots with teeth more likely to be deciduous than evergreen at a given temperature (e.g. Bailey & Sinnott, 1916; Wolfe, 1993; Jacobs, 2002)? We selected sites from our calibration and from the CLAMP calibration that each contained > 15% evergreen and > 15% deciduous species ($n = 29$ sites). At individual sites, deciduous species are more likely to be toothed than evergreen species ($P < 0.001$); at warm temperatures, this discrepancy diminishes such that above 16°C MAT there is no significant effect ($P = 0.18$; Fig. 4). The slope of the relationship between the proportion of toothed deciduous species in a flora and MAT is significantly steeper than that of evergreen species ($P = 0.04$), indicating the presence of a leaf-habit effect. The evergreen effect is also present in many of the digital leaf physiognomy variables (Fig. 4). Evergreen species usually have
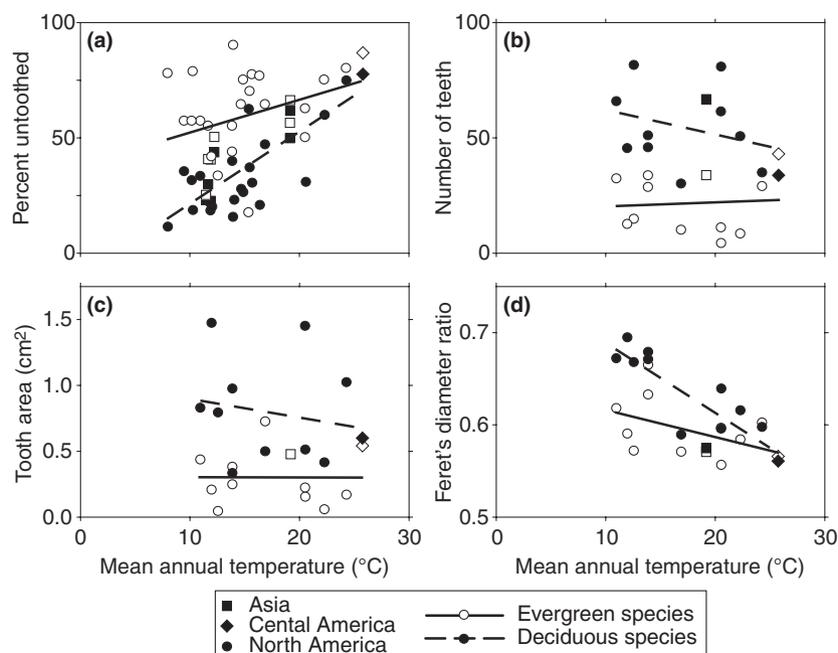
Fig. 4 Relationship between mean physiognomic characters of deciduous and evergreen species in a flora and mean annual temperature (MAT). All sites are from Asia, North America, and Central America. Feret's diameter ratio is the diameter of a circle with same area of a leaf divided by the leaf's longest axis (see the Supporting Information, Table S2 for definitions of variables). (a) Selected sites from our calibration and the CLAMP calibration (see Spicer, 2009) that have > 15% evergreen and > 15% deciduous species ($n = 29$). (b–d) Selected sites from our calibration that have > 15% evergreen and > 15% deciduous species ($n = 12$).

fewer teeth ($P < 0.001$), smaller teeth ($P = 0.005$), and smaller teeth relative to their leaf area ($P = 0.03$) than do deciduous species at the same site. Evergreen species also have lower Feret's diameter ratio (diameter of a circle with the same area as the leaf divided by the leaf's longest axis; $P = 0.009$; Wolfe, 1993; Greenwood and Basinger, 1994). Furthermore, for these traits, either the slope of the regression between the trait and MAT in deciduous taxa is significantly steeper than for evergreen taxa (Feret's diameter ratio: $P = 0.04$), or there is a significant difference in the y-intercept between deciduous and evergreen regressions (number of teeth, $P < 0.001$; tooth area, $P < 0.001$; number of teeth/blade area, $P < 0.001$). As with percent of untoothed species, the effect diminishes at warmer temperatures. We posit that evergreen species are less toothed because leaves in many evergreen taxa flush throughout the growing season, and thus any tooth-driven pulse in sap flow is more muted relative to neighboring deciduous taxa with a more synchronized leaf flush.

The physiognomy of evergreen taxa therefore responds differently to climate than that of deciduous taxa. Across all sites in our calibration and the CLAMP calibration, 11% of the variance in the relationship between MAT and the percentage of untoothed species can be explained by the percentage of evergreen species. This leaf-habit effect can contribute to physiognomic differences both within and across sites (Figs 2, 4, 5). It may even provide a simple explanation for the higher percentage of untoothed species

in southern hemisphere floras compared with northern hemisphere floras (Greenwood et al., 2004; Fig. 5) because southern hemisphere floras are typically dominated by evergreen taxa (mean = 98% vs 25% in our sites). However, as discussed in the Introduction, differing evolutionary or environmental histories of the floras may also contribute to differences.

Estimating climate from leaf physiognomy

A global approach Our models include all 92 calibration sites. The most commonly applied leaf-margin analysis model is based on 34 sites from eastern Asia (Wolfe, 1979; Wing & Greenwood, 1993). Because the correlation between MAT and the percent of untoothed species is remarkably strong in this data set ($r^2 = 0.98$), the standard errors quoted in the paleobotanical literature are typically c. ± 2°C (Wilf, 1997). However, these errors are too low because factors associated with sample size and over-dispersion in the binary data set will inflate them (Miller et al., 2006). Weaker, but similar correlations to those of Wolfe (1979) are found in other regional studies (Wilf, 1997; Jacobs, 1999, 2002; Gregory-Wodzicki, 2000; Kennedy et al., 2002; Kowalski, 2002; Greenwood et al., 2004; Traiser et al., 2005; Miller et al., 2006; Adams et al., 2008; Aizen & Ezcurra, 2008; Hinojosa et al., 2010; Su et al., 2010). The leaf-margin analysis regression using our calibration, which is more climatically, geographically, and
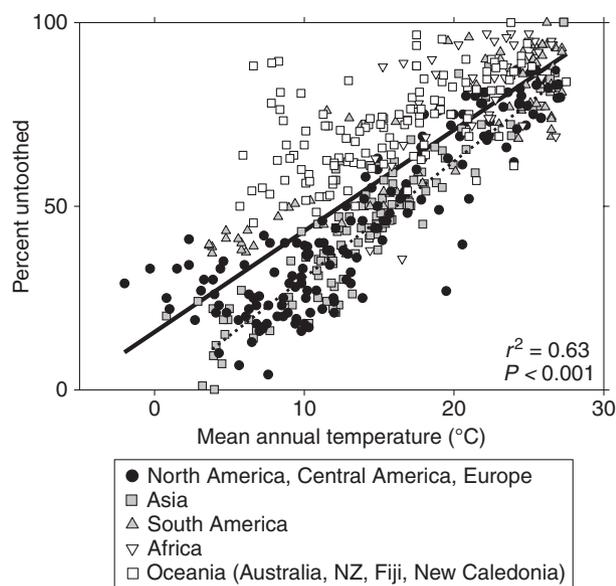
Fig. 5 Relationship between the percentage of untoothed species in a flora and mean annual temperature for 535 globally distributed sites. Linear regression fit, r^2 and P values are given. For comparison, the regression of Wolfe (1979) is plotted (dashed line, $r^2 = 0.98$, $P < 0.001$). Sources include Wolfe (1979, 1993), Midgley *et al.* (1995), Wilf (1997), Burnham *et al.* (2001), Jacobs (1999, 2002), Gregory-Wodzicki (2000), Kennedy (1998), Kowalski (2002), Greenwood *et al.* (2004), Royer *et al.* (2005), Hinojosa *et al.* (2006), Aizen & Ezcurra (2008), Su *et al.* (2010), and this study.

phylogenetically diverse than any single regional data set (Fig. 1), is considerably weaker than most regional equations (Fig. 2a; $r^2 = 0.58$; standard error (SE) = ± 4.8°C). A larger compilation from the literature ($n = 535$ sites) is consistent with this finding ($r^2 = 0.64$; SE = ± 4.1°C; Fig. 5). This suggests that the error associated with a globally derived leaf-margin analysis equation is at least ± 4°C.

The calibration data for leaf-area analysis (Wilf *et al.*, 1998; Jacobs, 1999, 2002; Gregory-Wodzicki, 2000) are primarily from low-latitude in Central America, South America, Asia and Africa. A compilation of these calibration sites suggests a strong univariate correlation between $\log_e$(MAP) and $\log_e$(leaf area) ($r^2 = 0.71$; Jacobs, 2002). Similar to leaf-margin analysis, our more global calibration indicates a much weaker correlation ($r^2 = 0.23$, Fig. 3).

Together, these results raise the obvious question: Why use a global model when regional calibrations are usually more precise (i.e. smaller standard errors)? On one hand, regional models capture the current relationship between leaf physiognomy and climate, which may be appropriate for specific floras. On the other hand, regional models capture a narrower slice of biological and ecological information (see the Introduction), which is not appropriate for fossil floras with a taxonomic composition or environmental setting different from the modern. For example, if the distinct leaf–climate character of Australian vegetation is related to nutrient-poor soils, lack of frost tolerance, ever-

green leaf habit, and/or phylogenetic isolation (Jordan, 1997; Greenwood *et al.*, 2004), any fossils that use an Australia-specific calibration must fit within this relatively narrow phylogenetic and ecological space. We find with our fossil floras that application of regional calibrations typically leads to cooler MAT estimates than the global calibration (Table 1), and that these are more at odds with independent evidence (see 'Application of digital leaf physiognomy to fossil record' section). The regional-based estimates are thus more precise, but may be less accurate.

An advantage of a global calibration for fossil applications is that it increases the likelihood that the appropriate biological and ecological information has been captured, although it may also lead to the incorporation of information not applicable to some fossil floras. For example, the biggest difference between the Jacobs (2002) compilation and our calibration of leaf area is at wet sites. Our data show a much wider range in site-mean leaf area at high MAP, regardless of temperature. That is, some of the warmest, wettest sites have comparatively small leaves (e.g. sites from Colombia, Australia, and Hawaii and Florida, USA, circled in Fig. 3), demonstrating that small leaves at wet sites are not always driven by the confounding influence of cool temperature. There are two possible reasons for the discrepancy between our calibration and the Jacobs (2002) compilation. First, our data contain many sites that are both wetter and drier than the compilation of Jacobs (2002). Second, although the Jacobs (2002) compilation includes sites from Africa, Asia, and Central and South America, many of the sites are from a few discrete areas (e.g. 35% of sites are from Costa Rica and Bolivia). Our calibration includes a greater phylogenetic, geographic and climatic diversity of sites, and probably better reflects the global range of leaf size.

The trade-off with a global calibration is that any single regional signal, which could be important in a fossil application, is diluted through the inclusion of extra-regional sites. Clearly, if sufficient phylogenetic and ecological information is available, approaches that take this information into account would be preferred. We consider our global models to be important, but conservative, first steps for digital leaf physiognomy because a global approach captures the widest range of information and accounts for floras with mixed phylogenetic histories, such as extinct species that are related to extant taxa living in both the northern and southern hemisphere.

Digital leaf physiognomy models The standard error of the best MAT multiple linear regression model that can be applied to fragmentary fossil leaves is ± 4.0°C ($r^2 = 0.70$, $P = 10^{-23}$) (Table 2). Compared with the leaf-margin analysis equation derived from the same 92 sites (± 4.8°C), our model represents a moderate improvement in precision. The multivariate MAT model incorporates the percentage of untoothed species, the number of teeth : internal perimeter,

and Feret's diameter ratio. Both of the tooth variables are probably related functionally to accelerating growth early in the growing season in cooler climates (Royer & Wilf, 2006). Feret's diameter ratio decreases in warmer climates; that is, leaves typically become longer than they are wide as MAT increases. This negative correlation most likely allows leaves to better shed heat in warm climates (e.g. Givnish, 1984). None of these three variables significantly correlate with MAP, even after accounting for the covariation of MAT (Table S3).

The multiple linear regression MAP model ($r^2 = 0.27$, $P = 10^{-6}$; SE = 0.60) is somewhat more precise than the univariate leaf-area analysis MAP model ($r^2 = 0.23$, $P = 10^{-6}$; SE = 0.61; Table 2). For example, the error for the Fox Hills fossil flora using digital leaf physiognomy is +116/−64 cm but with leaf-area analysis is +125/−68 cm (Table 1); the errors are asymmetric because both methods estimate $\log_e$(MAP). The multivariate MAP model incorporates $\log_e$(leaf area, mm^2), $\log_e$(number of teeth : internal perimeter) and $\log_e$(perimeter ratio). Both perimeter ratio and number of teeth : internal perimeter negatively correlate with MAP (i.e. leaves are less toothy at higher rainfalls); the functional basis for this response is not known (see the Introduction). Leaf area increases with MAP, a leaf trait that is functionally related to water loss (Parkhurst & Loucks, 1972). Of the three variables in our MAP model, two also correlate significantly with MAT after controlling for MAP with partial correlation ($\log_e$(perimeter ratio): $r^2 = 0.36$, $P < 0.001$; $\log_e$(number of teeth : internal perimeter): $r^2 = 0.24$, $P < 0.01$; Table S3). This raises the possibility that our paleo-MAP estimates are affected by the confounding influence of MAT.

To gauge the accuracy of our models, MAT and MAP were estimated at each site using the regression based on the other 91 sites (i.e. a jackknife-type approach). For MAT, the standard error of the estimates was smaller for the multi-

variate model than for leaf-margin analysis (4.0 vs 4.8°C). Furthermore, a paired sample t-test indicates that the absolute values of the deleted residuals are significantly smaller in the multivariate model ($P = 0.02$). Our multivariate MAT model is thus more accurate and precise than a similarly-calibrated leaf-margin analysis equation. The patterns for MAP are less convincing. The standard error of the estimates is marginally smaller for the multivariate model than for leaf-area analysis (0.60 vs 0.61), and a paired sample t-test indicates that the absolute value of the deleted residuals are smaller in the multivariate model, but not significantly so ($P = 0.10$). Thus, our MAP model is somewhat more precise, but not significantly more accurate than the univariate leaf-area analysis; further, two of the variables are confounded by the influence of MAT. For these reasons, it is not clear whether our MAP model is worth the additional processing effort relative to leaf-area analysis. In summary, neither our model nor leaf-area analysis are particularly good at estimating MAP.

Application of digital leaf physiognomy to fossil record
We applied our multivariate models to 10, well-studied, latest Cretaceous to Eocene fossil floras (Table 1). We emphasize that the climate estimates presented here are provisional until the potential confounding effects already discussed (especially phylogeny and leaf habit) are more fully accounted for. Nonetheless, we feel an initial application of this new approach is warranted and demonstrates its promise.

First, we used CCA as an initial quality check for our fossils. If a fossil site plotted outside the range of the calibration data, then it occupies uncalibrated physiognomic space; we did not attempt to reconstruct climate from such sites. All fossil sites plotted within our calibrated space except Bonanza (Fig. S4). Bonanza may be an outlier because it mixes two habitats, a lowland lake margin and an upland distal to the lake margin (MacGinitie, 1969). Also,

Table 2 Regression models for predicting mean annual temperature and mean annual precipitation for 92 calibration sites

Regression model	Variables	Coefficient	r^2	SE	F	P
Mean annual temperature						
Leaf-margin analysis	Percent untoothed	0.204	0.58	4.8 (°C)	126.1	10^{-19}
	Constant	4.600				
Digital leaf physiognomy	Percent untoothed	0.210	0.70	4.0 (°C)	69.8	10^{-23}
	Feret's diameter ratio	42.296				
	Number of teeth : internal perimeter	−2.609				
	Constant	−16.004				
Mean annual precipitation						
Leaf-area analysis	Leaf area ($\log_e$, mm^2)	0.283	0.23	0.61 ($\log_e$, cm)	27.0	10^{-6}
	Constant	2.92				
Digital leaf physiognomy	Leaf area ($\log_e$, mm^2)	0.298	0.27	0.60 ($\log_e$, cm)	10.78	10^{-6}
	Perimeter ratio ($\log_e$)	−2.717				
	Number of teeth : internal perimeter ($\log_e$)	0.279				
	Constant	3.033				

Variables defined in the Supporting Information, Table S2. SE, standard error.

among fossil sites, Bonanza has the highest estimated mean leaf mass per area, suggesting a mix of both evergreen and deciduous species (Royer *et al.*, 2007), whereas the other sites were likely composed of a higher percentage of deciduous taxa (Fig. S5). As discussed earlier, leaf habit may influence leaf–climate correlations. For these reasons, we currently do not advocate using leaf physiognomy to reconstruct paleoclimate at Bonanza.

Mean annual temperature estimates made using our leaf-margin analysis equation for the Williston Basin floras are *c.* 10°C (± 4.8°C; Table 1), which is cooler than expected for three reasons. First, high-latitude deep-sea temperatures were *c.* 10°C at this time (Zachos *et al.*, 2001) and are incompatible with low-elevation, mid-latitude MATs of *c.* 10°C. Second, the presence of palm fossils in floral zone Williston Basin I (Peppe, 2009, 2010) suggests a MAT > 10°C (Larcher & Winter, 1981; Sakai & Larcher, 1987; Wing & Greenwood, 1993; Greenwood & Wing, 1995). Third, crocodilian fossils are present throughout the Paleocene sequence in the Williston Basin and across the Western Interior of North America, implying a MAT of ≥ 14°C (Markwick, 1998). The digital leaf physiognomy estimates for the three Williston Basin floral zones are, on average, 5.5°C warmer than leaf-margin analysis estimates (Table 1). These estimates, which are all ≥ 15°C (± 4.0°C), are in better agreement with the independent evidence cited above.

The warmer temperatures with digital leaf physiognomy are mostly driven by the low teeth : internal perimeter values, which negatively correlate with MAT (Fig. S6). The percentage of toothed species at these three sites is quite high (*c.* 75%), which accounts for the cool MAT estimates with leaf-margin analysis, but most of the toothed species have small and few teeth. Thus, these floras demonstrate the usefulness of incorporating climatically meaningful physiognomic variables and provide strong support for the digital leaf physiognomy approach.

The MAT estimate for the Fox Hills flora using digital leaf physiognomy is over 6°C warmer than with leaf-margin analysis (21.6, ± 4.0°C vs 14.8, ± 4.8°C; Table 1), and is more compatible with independent MAT estimates based on oxygen isotopes of shallow-water marine invertebrates from the adjacent, contemporaneous Fox Hills Seaway (18.0°C; Carpenter *et al.*, 2003). As with the Williston Basin floras, the warmer estimate is largely driven by a low teeth : internal perimeter ratio (Fig. S6).

In the case of the Hubble Bubble flora from the PETM in the Bighorn Basin, independent evidence from the basin suggests a warming (Koch *et al.*, 2003; Fricke & Wing, 2004; Wing *et al.*, 2005; Secord *et al.*, 2010) and drying (Kraus & Riggins, 2007; Smith *et al.*, 2009a) during the PETM. Digital leaf physiognomy produces an MAT estimate that is 2.4°C warmer than leaf-margin analysis (Table 1), and thus is in slightly better agreement with the expected temperature increase during the PETM (Fricke &

Wing, 2004). The warmer estimate is again driven primarily by the flora's low teeth : internal perimeter (Fig. S6).

Several lines of evidence are consistent with the Cerrejón flora being a tropical rainforest, including the presence of a large-bodied snake (Head *et al.*, 2009) and soft-shelled turtles (Cadena *et al.*, 2010), as well as the climatic affinities of the nearest living relatives of several Cerrejón plant taxa (Doria *et al.*, 2008; Herrera *et al.*, 2008; Gómez-Navarro *et al.*, 2009; Wing *et al.*, 2009b). The digital leaf physiognomy estimates of MAT and MAP support a tropical rainforest interpretation and are wetter and considerably warmer than estimates from the univariate approaches (Table 1, Fig. S4). We note that our MAP estimate is somewhat drier than the leaf-area analysis estimate of Wing *et al.* (2009b) (264, +217/−119 cm vs 324, +140/−98 cm), but this is because they used the more regional leaf-area analysis regression of Wilf *et al.* (1998).

The digital leaf physiognomy estimates for Republic are similar to the univariate model estimates (Table 1), which broadly agree with some independent evidence (Wolfe & Wehr, 1987) but are cooler than estimates based on the species composition of the flora (*c.* 12–13°C, Greenwood *et al.*, 2005). The MAT and MAP estimates for the two southern hemisphere floras, P. Loros and Laguna del Hunco, are similar to estimates from univariate approaches, but are cooler and drier than expected (Table 1; see also Fig. S4). For example, the presence of a species of *Papuacedrus* in the Laguna del Hunco flora (*P. prechilensis*) suggests that the flora was fairly warm and wet (Wilf *et al.*, 2009). This discrepancy may be due to the phylogenetic histories of the floras (see earlier discussions). Because we have few sites from southern South America in our calibration, we may not have fully characterized the physiognomy–climate space of this region.

Implications and future directions

Our study demonstrates the promise of using leaf–climate correlations in a multivariate context for reconstructing MAT and MAP from fossil floras. Digital leaf physiognomy has three major advantages over the traditional univariate and multivariate methods. First, the physiognomic variables are mostly continuous, highly reproducible, and are functionally linked to climate. Second, digital leaf physiognomy is somewhat more precise than global univariate approaches, offering the potential for more refined climate reconstructions. Third, and perhaps most importantly, climate estimates for fossil floras made using digital leaf physiognomy are typically warmer and wetter, and much closer to independent climate evidence than other leaf–climate approaches. Digital leaf physiognomy thus offers the potential for better understanding ancient greenhouse climates. However, there is room for improvement; in particular, more calibration sites from Europe, Africa, southern South America, Oceania, and the tropics are needed to

increase phylogenetic diversity. Most critically, a quantitative assessment of the impacts of leaf habit and phylogeny (and their interaction) on leaf physiognomy is required so that ecologically and phylogenetically informed calibrations can be developed.

References

Adams JM, Green WA, Zhang Y. 2008. Leaf margins and temperature in the North American flora: recalibrating the paleoclimatic thermometer. *Global and Planetary Change* **60**: 523–534.

Aizen MA, Ezcurra C. 2008. Do leaf margins of the temperate forest flora of Southern South America reflect a warmer past? *Global Ecology and Biogeography* **17**: 164–174.

Bailey IW, Sinnott EW. 1915. A botanical index of Cretaceous and Tertiary climates. *Science* **41**: 831–834.

Bailey IW, Sinnott EW. 1916. The climatic distribution of certain types of angiosperm leaves. *American Journal of Botany* **3**: 24–39.

Baker-Brosh KF, Peet RK. 1997. The ecological significance of lobed and toothed leaves in temperate forest trees. *Ecology* **78**: 1250–1255.

Berry EW. 1925. A Miocene flora from Patagonia: *Johns Hopkins University Studies in Geology* **6**: 183–251.

Berry EW. 1937. A Paleocene flora from Patagonia: Johns Hopkins University Studies. *Geology* **12**: 33–50.

Billings FH. 1905. Precursory leaf serrations of *Ulmus. Botanical Gazette* **40**: 224–225.

Burnham RJ, Pitman NCA, Johnson KR, Wilf P. 2001. Habitat-related error in estimating temperatures from leaf margins in a humid tropical forest. *Johns Hopkins University Studies in Geology* **88**: 1096–1102.

Cadena EA, Bloch JI, Jaramillo CA. 2010. New Podocnemidid turtle (Testudines: Pleurodira) from the Middle-Upper Paleocene of South America. *Journal of Vertebrate Paleontology* **30**: 367–382.

Cariglino B. 2007. *Paleoclimatic analysis of three Eocene lacustrine floras (Laguna del Hunco, Chubut, Argentina; Republic, Washington, USA; and Green River, Utah, USA) using digital leaf physiognomy.* MS thesis, Pennsylvania State University, University Park, PA, USA.

Carpenter SJ, Erickson JM, Holland FD Jr. 2003. Migration of a Late Cretaceous fish. *Nature* **423**: 70–74.

Currano ED, Labandeira CC, Wilf P. 2010. Fossil insect folivory tracks paleotemperature for six million years. *Ecological Monographs* **80**: 547–567.

Currano ED, Wilf P, Wing SL, Labandeira CC, Lovelock EC, Royer DL. 2008. Sharply increased insect herbivory during the Paleocene–Eocene Thermal Maximum. *Proceedings of the National Academy of Sciences, USA* **105**: 1960–1964.

Dilcher D. 1973. A paleoclimatic interpretation of the Eocene floras of southeastern North America. In: Graham A, ed. *Vegetation and vegetational history of northern Latin America.* Amsterdam: Elsevier, 39–59.

Dilcher DL, Kowalski EA, Wiemann MC, Hinojosa LF, Lott TA. 2009. A climatic and taxonomic comparison between leaf litter and standing vegetation from a Florida swamp woodland. *American Journal of Botany* **96**: 1108–1115.

Dolph GE, Dilcher D. 1980a. Variation in leaf size with respect to climate in Costa Rica. *Biotropica* **12**: 91–99.

Dolph GE, Dilcher D. 1980b. Variation in leaf size with respect to climate in the tropics of the Western Hemisphere. *Bulletin of the Torrey Botanical Club* **107**: 154–162.

Doria G, Jaramillo CA, Herrera F. 2008. Menispermaceae from the Cerrejón Formation, Middle to Late Paleocene, Colombia. *American Journal of Botany* **95**: 954–973.

Feild TS, Sage TL, Czerniak C, Iles WJD. 2005. Hydathodal leaf teeth of *Chloranthus japonicus* (Chloranthaceae) prevent guttation-induced flooding of the mesophyll. *Plant, Cell & Environment* **28**: 1179–1190.

Felsenstein J. 1985. Phylogenies and the comparative method. *American Naturalist* **125**: 1–15.

Fricke HC, Wing SL. 2004. Oxygen isotope and paleobotanical estimates of temperature and $\delta^{18}O$-latitude gradients over North America during the Early Eocene. *American Journal of Science* **304**: 612–635.

Garland T, Harvey PH, Ives AR. 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. *Systematic Biology* **41**: 18–32.

Gates DM. 1980. *Biophysical ecology.* New York: Springer-Verlag.

Givnish TJ. 1979. On the adaptive significance of leaf form. In: Solbrig T, Jain S, Johnson GB, Raven PH, eds. *Topics in plant population biology.* New York, NY, USA: Columbia University Press, 375–407.

Givnish TJ. 1984. Leaf and canopy adaptations in tropical forests. In: Medina E, Mooney HA, Vasquez-Yanes C, eds. *Physiological ecology of plants of the wet tropics.* The Hague, the Netherlands: Dr. W. Junk Publishers, 51–84.

Givnish TJ. 1987. Comparative studies of leaf form: assessing the relative roles of selective pressures and phylogenetic constraints. *New Phytologist* **106**: 131–160.

Gómez-Navarro C, Jaramillo C, Herrera F, Wing SL, Callejas R. 2009. Palms (Arecaceae) from a Paleocene rainforest of Northern Colombia. *American Journal of Botany* **96**: 1300–1312.

Green WA. 2006. Loosening the CLAMP: an exploratory graphical approach to the Climate Leaf Analysis Multivariate Program. *Palaeontologia Electronica* 9: 9A.

Greenwood DR. 1992. Taphonomic constraints on foliar physiognomic interpretations of Late Cretaceous and Tertiary paleoclimates. *Review of Palaeobotany and Palynology* 71: 149–190.

Greenwood DR. 2005. Leaf margin analysis: taphonomic constraints. *Palaios* 20: 498–505.

Greenwood DR, Archibald SB, Mathewes RW, Moss PT. 2005. Fossil biotas from the Okanagan Highlands, southern British Columbia and northeastern Washington State: climates and ecosystems across an Eocene landscape. *Canadian Journal of Earth Sciences* 42: 167–185.

Greenwood DR, Basinger JF. 1994. The paleoecology of high latitude Eocene swamp forests from Axel Heiberg Island, Canadian High Arctic. *Review of Palaeobotany and Palynology*, 81: 83–97.

Greenwood DR, Basinger JF, Smith RY. 2010. How wet was the Arctic Eocene rain forest? Estimates of precipitation from Paleogene Arctic macrofloras. *Geology* 38: 15–18.

Greenwood DR, Wilf P, Wing SL, Christophel DC. 2004. Paleotemperature estimation using leaf-margin analysis: is Australia different? *Palaios* 19: 129–142.

Greenwood DR, Wing SL. 1995. Eocene continental climates and latitudinal temperature gradients. *Geology* 23: 1044–1048.

Gregory-Wodzicki KM. 2000. Relationship between leaf morphology and climate, Bolivia: implications for estimating paleoclimate from fossil floras. *Paleobiology* 26: 668–688.

Hall JB, Swaine MD. 1981. *Distribution and ecology of vascular plants in a tropical rain forest: forest vegetation in Ghana*. The Hague, Netherlands: Dr. W. Junk Publishers.

Head JJ, Bloch JI, Hastings AK, Bourque JR, Cadena EA, Herrera FA, Polly PD, Jaramillo CA. 2009. Giant boid snake from the Palaeocene neotropics reveals hotter past equatorial temperatures. *Nature* 457: 715–717.

Herrera FA, Jaramillo CA, Dilcher DL, Wing SL, Gómez-Navarro C. 2008. Fossil Araceae from a Paleocene neotropical rainforest in Colombia. *American Journal of Botany* 95: 1569–1583.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965–1978.

Hinojosa LF, Armesto JJ, Villagran C. 2006. Are Chilean coastal forests Pre-Pleistocene relicts? Evidence from foliar physiognomy, palaeoclimate, and phytogeography. *Journal of Biogeography* 33: 331–341.

Hinojosa LF, Pérez F, Gaxiola A, Sandoval I. 2010. Historical and phylogenetic constraints on the incidence of entire leaf margins: insights from a new South American model. *Global Ecology and Biogeography*, doi: 10.1111/j.1466-8238.2010.00595.x.

Huff PM, Wilf P, Azumah EJ. 2003. Digital future for paleoclimate estimation from fossil leaves? Preliminary results. *Palaios* 18: 266–274.

Iglesias A, Wilf P, Johnson KR, Zamuner AB, Cúneo NR, Matheos SD, Singer BS. 2007. A Paleocene lowland macroflora from Patagonia reveals significantly greater richness than North American analogs. *Geology* 35: 947–950.

Jacobs BF. 1999. Estimation of rainfall variables from leaf characters in tropical Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology* 145: 231–250.

Jacobs BF. 2002. Estimation of low-latitude paleoclimates using fossil angiosperm leaves: examples from the Miocene Tugen Hills, Kenya. *Paleobiology* 28: 399–421.

Jacobs BF, Deino AL. 1996. Test of climate–leaf physiognomy regression models, their application to two Miocene floras from Kenya, and [40]Ar/[39]Ar dating of the Late Miocene Kapturo site. *Palaeogeography, Palaeoclimatology, Palaeoecology* 123: 259–271.

Johnson CM, Zarin DJ, Johnson AH. 2000. Post-disturbance aboveground biomass accumulation in global secondary forests. *Ecology* 81: 1395–1401.

Jordan GJ. 1997. Uncertainty in palaeoclimatic reconstructions based on leaf physiognomy. *Australian Journal of Botany* 45: 527–547.

Kennedy EM. 1998. *Cretaceous and Tertiary megafloras from New Zealand and their climate signals*. PhD Thesis, Open University, Milton Keynes, UK.

Kennedy EM, Spicer RA, Rees PM. 2002. Quantitative palaeoclimate estimates from the Late Cretaceous and Paleocene leaf floras in the northwest of the South Island, New Zealand. *Palaeogeography, Palaeoclimatology, Palaeoecology* 184: 321–345.

Koch PL, Clyde WC, Hepple RP, Fogel ML, Wing SL, Zachos JC. 2003. Carbon and oxygen isotope records from paleosols spanning the Paleocene–-Eocene boundary, Bighorn Basin, Wyoming. In: Wing SL, Gingerich PD, Schmitz B, Thomas E, eds. *Causes and consequences of globally warm climates in the early Paleogene: Boulder, Colorado, Geological Society of America Special Paper 369*: 49–64.

Kowalski EA. 2002. Mean annual temperature estimation based on leaf morphology: a test from tropical South America. *Palaeogeography, Palaeoclimatology, Palaeoecology* 188: 141–165.

Kowalski EA, Dilcher DL. 2003. Warmer paleotemperatures for terrestrial ecosystems. *Proceedings of the National Academy of Sciences, USA* 100: 167–170.

Kraus MJ, Riggins S. 2007. Transient drying during the Paleocene–Eocene Thermal Maximum (PETM): analysis of paleosols in the Bighorn Basin, Wyoming. *Palaeogeography, Palaeoclimatology, Palaeoecology* 245: 444–461.

Kvacek Z. 2007. Do extant nearest relatives of thermophile European Cenozoic plant elements reliably reflect climatic signal? *Palaeogeography, Palaeoclimatology, Palaeoecology* 253: 32–40.

Larcher W, Winter A. 1981. Frost susceptibility of palms: experimental data and their interpretations. *Principes* 25: 143–152.

Lewis MC. 1972. The physiological significance of variation in leaf structure. *Science Progress (Oxford)* 60: 25–51.

Liang M-M, Bruch A, Collinson M, Mosbrugger V, Li C-S, Sun Q-G, Hilton J. 2003. Testing the climatic estimates from different palaeobotanical methods: an example from the Middle Miocene Shanwang flora of China. *Palaeogeography, Palaeoclimatology, Palaeoecology* 198: 279–301.

Little SA, Kembel SW, Wilf P, 2010. Paleotemperature proxies from leaf fossils reinterpreted in light of evolutionary history. *PLoS ONE* 5(12): e15161. doi:10.1371/journal.pone.0015161.

MacGinitie HD. 1969. The Eocene Green River Flora of northwestern Colorado and northeastern Utah. *University of California Publications in Geological Science* 83: 1–202.

Markwick PJ. 1998. Fossil Crocodilians as indicators of Late Cretaceous and Cenozoic climates: implications for using palaeontological data in reconstructing palaeoclimate. *Palaeogeography, Palaeoclimatology, Palaeoecology* 137: 205–271.

Midgley JJ, Van Wyk GR, Everard DA. 1995. Leaf attributes of South African forest species. *African Journal of Ecology* 33: 160–168.

Miller IM, Brandon MT, Hickey LJ. 2006. Using leaf margin analysis to estimate Mid-Cretaceous (Albian) paleolatitude of the Baja BC block. *Earth and Planetary Science Letters* 245: 95–114.

Mosbrugger V, Utescher T, Dilcher DL. 2005. Cenozoic continental climatic evolution of Central Europe. *Proceedings of the National Academy of Sciences, USA* 102: 14964–14969.

Parkhurst DF, Loucks OL. 1972. Optimal leaf size in relation to environment. *Journal of Ecology* 60: 505–537.

Peppe DJ. 2003. *Fox Hills I, a new Upper Maastrichtian megafloral zone within the Williston Basin of North Dakota*. BS thesis, St Lawrence University, Canton, NY, USA.

Peppe DJ. 2009. *A high resolution chronostratigraphic study of the early Paleocene floral record in the northern Great Plains*. PhD thesis, Yale University, New Haven, CT, USA.

Peppe DJ. 2010. Megafloral change in the early and middle Paleocene in the Williston Basin, North Dakota, USA. *Palaeogeography, Palaeoclimatology, Palaeoecology* **298**: 224–234.

Peppe DJ, Erickson JM, Hickey LJ. 2007. Fossil leaf species from the Fox Hills Formation (Upper Cretaceous: Maastrichtian: North Dakota, USA) and their paleogeographic significance. *Journal of Paleontology* **81**: 550–567.

Peppe DJ, Royer DL, Wilf P, Kowalski EA. 2010. Quantification of large uncertainties in fossil leaf paleoaltimetry. *Tectonics* **29**: TC3015, doi:10.1029/2009TC002549.

Radtke MG, Pigg KB, Wehr WC. 2005. Fossil *Corylopsis* and *Fothergilla* leaves (Hamamelidaceae) from the Lower Eocene Flora of Republic, Washington, U.S.A., and their evolutionary and biogeographic significance. *International Journal of Plant Sciences* **166**: 347–356.

Richards PW. 1996. *The tropical rain forest*, 2nd edn. Cambridge, UK: Cambridge University Press.

Royer DL, Kooyman RM, Little SA, Wilf P. 2009a. Ecology of leaf teeth: a multi-site analysis from an Australian subtropical rainforest. *American Journal of Botany* **96**: 738–750.

Royer DL, McElwain JC, Adams JM, Wilf P. 2008. Sensitivity of leaf size and shape to climate within *Acer rubrum* and *Quercus kelloggii*. *New Phytologist* **179**: 808–817.

Royer DL, Meyerson LA, Robertson KM, Adams JM. 2009b. Phenotypic plasticity of leaf shape along a temperature gradient in *Acer rubrum*. *PLoS ONE* **4**: e7653.

Royer DL, Miller IM, Peppe DJ, Hickey LJ. 2010. Leaf economic traits from fossils support a weedy habit for early angiosperms. *American Journal of Botany* **97**: 438–445.

Royer DL, Sack L, Wilf P, Lusk CH, Jordan GJ, Niinemets Ü, Wright IJ, Westoby M, Cariglino B, Coley PD et al. 2007. Fossil leaf economics quantified: calibration, Eocene case study, and implications. *Paleobiology* **33**: 574–589.

Royer DL, Wilf P. 2006. Why do toothed leaves correlate with cold climates? Gas exchange at leaf margins provides new insights into a classic paleo-temperature proxy. *International Journal of Plant Sciences* **167**: 11–18.

Royer DL, Wilf P, Janesko DA, Kowalski EA, Dilcher DL. 2005. Correlations of climate and plant ecology to leaf size and shape: potential proxies for the fossil record. *American Journal of Botany* **92**: 1141–1151.

Sakai A, Larcher W. 1987. *Frost survival of plants*. Berlin, Germany: Springer-Verlag.

Secord R, Gingerich PD, Lohmann KC, MacLeod KG. 2010. Continental warming preceding the Palaeocene–Eocene Thermal Maximum. *Nature* **467**: 955–958.

Smith JJ, Hasiotis ST, Kraus MJ, Woody DT. 2009a. Transient dwarfism of soil fauna during the Paleocene-Eocene Thermal Maximum. *Proceedings of the National Academy of Sciences, USA* **106**: 17655–17660.

Smith ME, Carroll AR, Singer BS. 2008. Synoptic reconstruction of a major ancient lake system: Eocene Green River Formation, western United States. *Geological Society of America Bulletin* **120**: 54–84.

Smith RY, Basinger JF, Greenwood DR. 2009b. Depositional setting, floristics and paleoenvironment of the Early Eocene Falkland Site, a new fossil flora locality from the Okanagan Highlands, British Columbia. *Canadian Journal of Earth Sciences* **46**: 811–822.

Sokal RR, Rohlf FJ. 1995. *Biometry*, 3rd edn. New York, NY, USA: Freeman.

Spicer RA. 2009. *CLAMP, Climate Leaf Analysis Multivariate Program.* The Open University. URL http://www.open.ac.uk/earth-research/spicer/CLAMP/Clampset1.html and http://clamp.ibcas.ac.cn/ [last accessed 1 November 2010].

Steart DC, Spicer RA, Bamford MK. 2010. Is Southern Africa different? An investigation of the relationship between leaf physiognomy and climate in Southern African mesic vegetation. *Review of Palaeobotany and Palynology* **162**: 607–620.

Stevens PF. 2001 onwards. Angiosperm Phylogeny Website. Version 9. URL http://www.mobot.org/MOBOT/research/APweb/ [last accessed 1 November 2010].

Stranks L, England P. 1997. The use of a resemblance function in the measurement of climatic parameters from the physiognomy of woody dicotyledons. *Palaeogeography, Palaeoclimatology, Palaeoecology* **131**: 15–28.

Su T, Xing YW, Liu YS, Jacques FMB, Chen WY, Huang YJ, Zhou ZK. 2010. Leaf margin analysis: a new equation from humid to mesic forests in China. *Palaios* **25**: 234–238.

The Angiosperm Phylogeny Group. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnaean Society* **161**: 105–121.

Torsvik TH, Müller RD, Van der Voo R, Steinberger B, Gaina C. 2008. Global plate motion frames: toward a unified model. *Reviews of Geophysics* **46**: RG3004, doi:10.1029/2007RG000227.

Traiser C, Klotz S, Uhl D, Mosbrugger V. 2005. Environmental signals from leaves – a physiognomic analysis of European vegetation. *New Phytologist* **166**: 465–484.

Uhl D, Mosbrugger V, Bruch A, Utescher T. 2003. Reconstructing palaeotemperatures using leaf floras – case studies for a comparison of Leaf Margin Analysis and the Coexistence Approach. *Review of Palaeobotany and Palynology* **126**: 49–64.

Utescher T, Mosbrugger V, Ashraf AR. 2000. Terrestrial climate evolution in northwest Germany over the last 25 million years. *Palaios* **15**: 430–449.

Vogel S. 1968. Sun leaves and shade leaves: differences in convective heat dissipation. *Ecology* **49**: 1203–1204.

Vogel S. 1970. Convective cooling at low airspeeds and the shapes of broad leaves. *Journal of Experimental Botany* **21**: 91–101.

Vogel S. 2009. Leaves in the lowest and highest winds: temperature, force and shape. *New Phytologist* **183**: 13–26.

Warton DI, Wright IJ, Falster DS, Westoby M. 2006. Bivariate line-fitting methods for allometry. *Biological Reviews of the Cambridge Philosophical Society* **81**: 259–291.

Webb LJ. 1968. Environmental relationships of the structural types of Australian rain forest vegetation. *Ecology* **49**: 296–311.

Westoby M, Cunningham SL, Fonseca CR, Overton J, Wright IJ. 1998. Phylogeny and variation in light capture area deployed per unit investment in leaves: designs for selecting study species with a view to generalising. In: Lambers H, Van Vuuren PHMMI, eds. *Inherent variation in plant growth: physiological mechanisms and ecological consequences*. Leiden, the Netherlands: Backhuys Publishers, 539–566.

Whittaker RH. 1975. *Communities and ecosystems*. New York, NY, USA: Macmillan.

Wiemann MC, Manchester SR, Dilcher DL, Hinojosa LF, Wheeler EA. 1998. Estimation of temperature and precipitation from morphological characters of dicotyledonous leaves. *American Journal of Botany* **85**: 1796–1802.

Wilf P. 1997. When are leaves good thermometers? A new case for leaf margin analysis. *Paleobiology* **23**: 373–390.

Wilf P, Cúneo NR, Johnson KR, Hicks JF, Wing SL, Obradovich JD. 2003a. High plant diversity in Eocene South America: evidence from Patagonia. *Science* **300**: 122–125.

Wilf P, Johnson KR, Cúneo NR, Smith ME, Singer BS, Gandolfo MA. 2005a. Eocene plant diversity at Laguna del Hunco and Río Pichileufú, Patagonia, Argentina. *American Naturalist* **165**: 634–650.

Wilf P, Johnson KR, Huber BT. 2003b. Correlated terrestrial and marine evidence for global climate changes before mass extinction at the Cretaceous–Paleogene boundary. *Proceedings of the National Academy of Sciences, USA* **100**: 599–604.

Wilf P, Labandeira CC, Johnson KR, Coley PD, Cutter AD. 2001. Insect herbivory, plant defense, and early Cenozoic climate change. *Proceedings of the National Academy of Sciences, USA* **98**: 6221–6226.

Wilf P, Labandeira CC, Johnson KR, Cúneo NR. 2005b. Richness of plant–insect associations in Eocene Patagonia: a legacy for South

American biodiversity. *Proceedings of the National Academy of Sciences, USA* **102**: 8944–8948.

Wilf P, Little SA, Iglesias A, Zamaloa MC, Gandolfo MA, Cúneo NR, Johnson KR. 2009. *Papuacedrus* (Cupressaceae) in Eocene Patagonia: a new fossil link to Australasian rainforests. *American Journal of Botany* **96**: 2031–2047.

Wilf P, Wing SL, Greenwood DR, Greenwood CL. 1998. Using fossil leaves as paleoprecipitation indicators: an Eocene example. *Geology* **26**: 203–206.

Wilf P, Wing SL, Greenwood DR, Greenwood CL. 1999. Using fossil leaves as paleoprecipitation indicators: an Eocene example: reply. *Geology* **27**: 92.

Wing S, Greenwood DR. 1993. Fossils and fossil climate: the case for equable continental interiors in the Eocene. *Philosophical Transactions of the Royal Society of London B* **341**: 243–252.

Wing S, Bao H, Koch PL. 2000. An Eocene cool period? Evidence for continental cooling during the warmest part of the Cenozoic. In: Huber BT, MacCleod KG, Wing S, eds. *Warm climates in earth history*. Cambridge, UK: Cambridge University Press, 197–237.

Wing SL, Harrington GJ, Smith FA, Bloch JI, Boyer DM, Freeman KH. 2005. Transient floral change and rapid global warming at the Paleocene–Eocene boundary. *Science* **230**: 993–996.

Wing SL, Bloch JI, Bowen GJ, Boyer DM, Chester S, Diefendorf AF, Harrington GJ, Kraus MJ, Secord R, Smith FA. 2009a. Coordinated sedimentary and biotic change during the Paleocene–Eocene Thermal Maximum in the Bighorn Basin, Wyoming, USA. In: Crouch EM, Strong CP, Hollis CJ, eds. *Climatic and Biotic Events of the Paleogene (CBEP 2009), extended abstracts from an international conference in Wellington, New Zealand, 12–15 January 2009. GNS Science Miscellaneous Series* **18**: 156–162.

Wing SL, Herrera F, Jaramillo CA, Gómez-Navarro C, Wilf P, Labandeira CC. 2009b. Late Paleocene fossils from the Cerréjon Formation, Colombia, are the earliest record of neotropical rainforest. *Proceedings of the National Academy of Sciences, USA* **106**: 18627–18632.

Wolfe JA. 1979. Temperature parameters of humid to mesic forests of eastern Asia and relation to forests of other regions in the Northern Hemisphere and Australasia. *United States Geological Survey Professional Paper* **1106**: 1–37.

Wolfe JA. 1993. A method of obtaining climatic parameters from leaf assemblages. *United States Geological Survey Bulletin* **2040**: 1–71.

Wolfe JA. 1995. Paleoclimatic estimates from Tertiary leaf assemblages. *Annual Review of Earth and Planetary Sciences* **23**: 119–142.

Wolfe JA, Upchurch Jr GR. 1987. North American nonmarine climates and vegetation during the Late Cretaceous. *Palaeogeography, Palaeoclimatology, Palaeoecology* **61**: 33–77.

Wolfe JA, Wehr WC. 1987. Middle Eocene dicotyledonous plants from Republic, northeastern Washington. *United States Geological Survey Bulletin* **1597**: 1–25.

Yang J, Wang Y-F, Spicer RA, Mosbrugger V, Li C-S, Sun Q-G. 2007. Climatic reconstruction at the Miocene Shanwang basin, China, using leaf margin analysis, CLAMP, coexistence approach, and overlapping distribution analysis. *American Journal of Botany* **94**: 599–608.

Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* **292**: 686–693.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Relationship between site-mean leaf area and mean annual temperature for all calibration sites.

Fig. S2 Correlation between the percent of untoothed species in a flora and mean annual temperature for edaphically wet and dry sites.

Fig. S3 Relationship between site-mean leaf physiognomy and mean annual temperature for sites along local water availability gradients.

Fig. S4 Canonical correspondence analysis plot of all calibration and fossil sites based on leaf physiognomy.

Fig. S5 Box-and-whisker plots of leaf mass per area for nine of the 10 fossil floras used in study.

Fig. S6 Box-and-whisker plots of number of teeth to internal perimeter and Feret's diameter ratio for all fossil floras, except Bonanza.

Table S1 Climate variables and site-mean physiognomic variables for all calibration sites

Table S2 Definitions of physiognomic and climatic variables used in study

Table S3 Correlations between site-mean physiognomic variables and climate variables

Table S4 Physiognomic scores for specimens from calibration sites

Table S5 Regression models for predicting mean annual temperature and precipitation for 92 calibration sites using all physiognomic variables

Table S6 Site-mean physiognomic variables for fossil sites

Table S7 Physiognomic scores for specimens from fossil sites

Notes S1 Geography and sampling methods for calibration sites.

Methods S1 Protocol for digitally processing fossil leaves.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.