

Phylogenetic analyses reveal the shady history of C₄ grasses

Erika J. Edwards^{a,1} and Stephen A. Smith^b

^aDepartment of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912; and ^bNational Evolutionary Synthesis Center, Durham, NC 27705

Edited by Michael J. Donoghue, Yale University, New Haven, CT, and approved December 31, 2009 (received for review August 24, 2009)

Grasslands cover more than 20% of the Earth's terrestrial surface, and their rise to dominance is one of the most dramatic events of biome evolution in Earth history. Grasses possess two main photosynthetic pathways: the C₃ pathway that is typical of most plants and a specialized C₄ pathway that minimizes photorespiration and thus increases photosynthetic performance in high-temperature and/or low-CO₂ environments. C₄ grasses dominate tropical and subtropical grasslands and savannas, and C₃ grasses dominate the world's cooler temperate grassland regions. This striking pattern has been attributed to C₄ physiology, with the implication that the evolution of the pathway enabled C₄ grasses to persist in warmer climates than their C₃ relatives. We combined geospatial and molecular sequence data from two public archives to produce a 1,230-taxon phylogeny of the grasses with accompanying climate data for all species, extracted from more than 1.1 million herbarium specimens. Here we show that grasses are ancestrally a warm-adapted clade and that C₄ evolution was not correlated with shifts between temperate and tropical biomes. Instead, 18 of 20 inferred C₄ origins were correlated with marked reductions in mean annual precipitation. These changes are consistent with a shift out of tropical forest environments and into tropical woodland/savanna systems. We conclude that C₄ evolution in grasses coincided largely with migration out of the understory and into open-canopy environments. Furthermore, we argue that the evolution of cold tolerance in certain C₃ lineages is an overlooked innovation that has profoundly influenced the patterning of grassland communities across the globe.

C₄ photosynthesis | climate niche evolution | cold tolerance | phylogeny

The term “C₄ photosynthesis” refers to a suite of biochemical and anatomical modifications to the standard plant C₃ photosynthetic pathway that work to concentrate CO₂ around the carbon-fixing enzyme Rubisco. The C₄ pathway greatly improves photosynthetic performance in situations that promote photorespiration, typically high-temperature and low-CO₂ environments (1). The pathway also promotes more efficient photosynthetic water use, because the CO₂ concentration mechanism allows C₄ plants to maintain a lower stomatal conductance for a given photosynthetic rate. C₄ photosynthesis is estimated to have evolved at least 50 times in terrestrial plants (2, 3) and is most prominent in grasses, where roughly half the species (~5,000) are C₄ (4), including economically important species such as maize, sugarcane, sorghum, and switchgrass. C₄ grasses currently dominate wide regions of the Earth, and it is estimated that they account for up to 25% of global annual terrestrial primary production (5).

Recent studies suggest that C₄ photosynthesis is a relatively recent innovation in plants, with the earliest appearances coinciding with plummeting atmospheric CO₂ levels during the mid-Oligocene and many origins occurring much later (4, 6). In grasses, the evolutionary history of C₄ photosynthesis is complex, with multiple origins, probable reversals, and a general lag-time between the evolution of the pathway and the formation of C₄-dominated ecosystems (3, 6–8). Although low atmospheric CO₂ certainly was a prerequisite for C₄ evolution, it is thought that multiple stressors worked in concert to promote the pathway. C₄ plants are notably prevalent in arid, high-light, saline, and disturbed environments, and it is largely accepted that water stress

has provided a strong selection pressure for C₄ evolution in eudicots (4). Grasses have long been viewed as an interesting exception to this pattern (9). Significant positive correlations between C₄ grass abundance and growing season temperature have been documented at both continental and regional scales (10–13); C₄ grasses dominate tropical grasslands and savannas but are virtually absent from cool-temperate grasslands and steppes. Furthermore, both experimental measurements of photosynthetic light use efficiency (termed “quantum yield”), and predictions of leaf models of C₃ and C₄ photosynthesis provide strong evidence that C₄ grasses outperform C₃ grasses at higher temperatures (5, 14–16). Most studies concerning C₄ grasses and precipitation have focused on C₃/C₄ mixed temperate grassland systems, where the timing of C₄ growth often is restricted to periods with significant rainfall (12, 13, 17, 18).

The C₄ pathway thus has been largely dismissed as an adaptation to water stress in grasses, with all data indicating that C₄ evolution allowed grasses to invade and diversify successfully into hot climates. However, few studies have compared C₄ grasses with their closest living C₃ relatives. C₄ origins are not distributed uniformly across Poaceae but are clustered in one major grass lineage, informally named the “PACMAD” clade (3, 6). Most of the C₃ grasses that dominate cool-climate grasslands belong to the Pooideae, a lineage that last shared a common ancestor with PACMAD grasses ≈65–50 Mya (19). It thus is possible that differences between Pooideae and PACMAD grasses that have nothing to do with photosynthetic pathway variation are driving the apparently strong sorting of C₃ and C₄ species along temperature gradients.

We employed an explicitly phylogenetic approach to assess the evolutionary history of climate niche space for grasses on a worldwide scale. We used two public archives of data to build the most inclusive phylogeny for Poaceae that permitted analysis of a climate dataset for all taxa (*Materials and Methods*). This analysis resulted in a 1,230 taxon tree with broad coverage of all of the major Poaceae lineages and, importantly, good sampling of known C₃ and C₄ transitions. Our phylogenetic tree includes roughly 10% of all Poaceae species and is the largest grass phylogeny built to date. We identified 21 nodes representing evolutionary transitions between photosynthetic types and used these nodes to generate phylogenetically independent C₃/C₄ pairwise comparisons (*Tables S1 and S2*) (20). Nearly all identified photosynthetic transitions were reconstructed as C₄ origins, with one purported reversal. Fifteen of the 21 transitions occurred within the Panicoideae, a major PACMAD lineage containing more than 3,000 species. To account for topological uncertainty in this area of the phylogeny, we performed Bayesian analyses for a 299-taxon Panicoideae dataset and ran all diver-

Author contributions: E.J.E. and S.A.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: erika_edwards@brown.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0909672107/DCSupplemental.

gence analyses across the Bayesian posterior distribution of trees. We also used this reduced dataset to reconstruct environmental niche evolution in the Panicoideae and tested whether shifts between photosynthetic types corresponded with significant changes in temperature and precipitation niche optima under a stabilizing selection model of evolution.

Results

Climate data extracted from all available geo-referenced herbarium material provided clear evidence that certain grass lineages have specialized in certain habitats (Fig. 1). Importantly, two strictly C₃ grass lineages, the Pooideae and the Danthonioideae, stood apart as inhabiting much cooler environments, measured either by mean annual temperature (MAT) (Fig. 1) or temperature of the wettest, coldest, or warmest month (Fig. S1). All other C₃ lineages were indistinguishable from C₄ lineages with respect to their temperature profiles. Distinctive sorting of precipitation variables was less apparent, although Pooideae occupied the drier end of the spectrum alongside the C₄ lineages Aristidoideae and Chloridoideae (Fig. S2).

Our phylogenetic analyses concurred with these general temperature and precipitation profiles for the different grass clades, particularly the observation that, with the exception of Pooideae and Danthonioideae, grasses are warm-climate specialists (Fig. 2). The evolution of C₄ photosynthesis appears to have had little influence on gross temperature niche: Only 10 of 21 photosynthetic shifts resulted in increases in MAT in the C₄ lineage, with a mean divergence of -0.13 °C. Likewise, there was no significant relationship between photosynthetic pathway and mean temperature of the hottest, coldest, or wettest months (Table 1). In sharp contrast, there were large and predictable shifts in mean annual precipitation (MAP) between C₃ and C₄ sister taxa; 18 of the 21 divergences resulted in C₄ taxa living in areas with lower MAP ($P < 0.01$), with a mean difference of 546 mm year⁻¹. Fifteen of those divergences also resulted in shifts toward increased seasonality of precipitation. Results were statistically similar using a more conservative, reduced dataset generated by removing three focal taxa with questionable placement in the phylogeny and recoding a C₃-C₄ intermediate species as C₃ (Table 1) and also were robust to testing against a wide range of alternative tree topologies within Panicoideae (Table S3). These results are highly consistent with an earlier

analysis of the Hawaiian grass flora that used a similar approach but was based on very limited geographic and phylogenetic sampling (21).

Analyses of niche evolution within Panicoideae provided further support that C₄ evolution was associated with shifts into drier, but not warmer, environments. For most climate variables, two-optimum stabilizing selection Ornstein-Uhlenbeck (OU) models were preferred over Brownian motion or single-optimum models, implying that C₃ and C₄ lineages have experienced divergent selection over time (Table 2). However, the mean C₃ and C₄ temperature optima were not largely different, and in all cases, the model inferred lower temperature optima for C₄ Panicoideae lineages than for C₃ Panicoideae lineages. MAP received the strongest support for a two-optimum model, with the C₄ optimum inferred to be drier than the C₃ optimum by more than 500 mm year⁻¹.

Discussion

These analyses provide clear evidence that C₄ origins in grasses coincided with ecological shifts into drier environments. However, the actual values of precipitation divergences across our datasets largely occurred squarely within what generally is considered a mesophytic moisture niche (22). No reconstructed C₄ ancestral focal node had a MAP of less than 500 mm year⁻¹; across all C₃/C₄ contrasts, the C₃ MAP average was 1,839 mm year⁻¹, and the C₄ average was 1,229 mm year⁻¹. In the tropics, such differences in annual precipitation values often lead to major changes in biome structure, because a MAP of 1,500 mm year⁻¹ generally is considered the minimum amount of rainfall required to support a closed-canopy forest (23). Considering the relative values of MAT, MAP, and precipitation seasonality, one-third of our focal nodes present climate profiles that are highly consistent with a shift from a closed-canopy tropical moist forest to an open tropical woodland/savanna biome (Fig. 3) (24). This finding also is consistent with the unexpected and significant increase in temperature seasonality in C₄ lineages, suggesting geographic movements north and south from the equator and into the savanna belt. This general picture of a shift from closed to open canopies mirrors results of an independent dataset that focused on qualitative ecological assessments for 117 genera of grasses (22) and suggests a direct link between C₄ evolution and the establishment of the savanna biome.

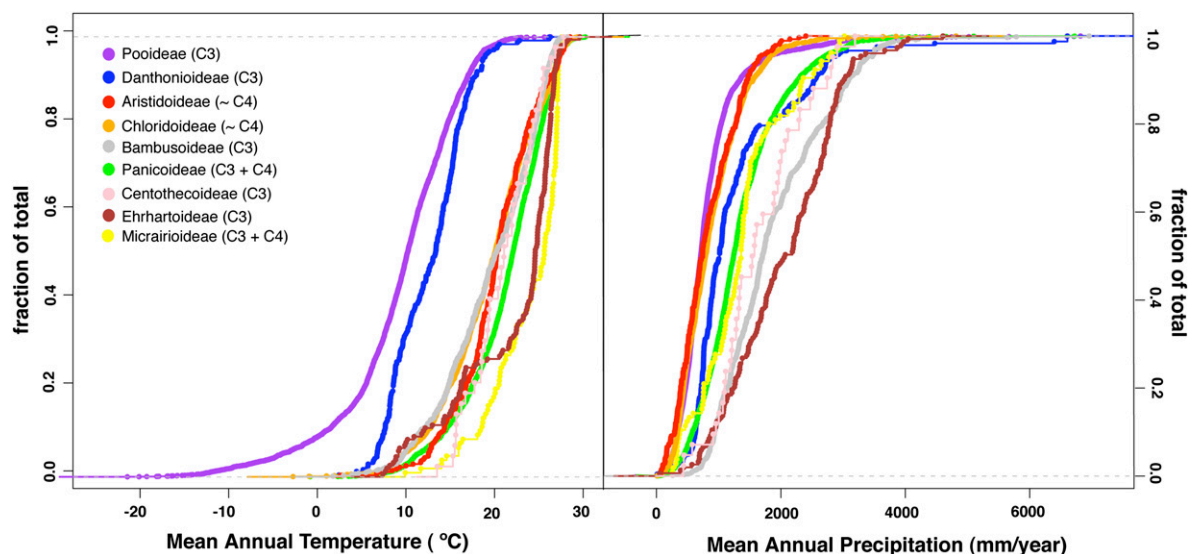


Fig. 1. Species accumulation curves for mean annual temperature and precipitation, sorted by major grass lineage. These data represent 1,584,351 independent collection localities spread across 10,469 taxa. Each point in the curve is a species' mean value.

Table 1. Independent contrast analyses for temperature and precipitation variables

Variable	Contrast (C_4 relative to C_3)	Number of positive contrasts
Full dataset (n = 21)		
MAT (°C)	-0.13	10
T_{max} (°C)	1.24	14
T_{min} (°C)	-1.71	11
$T_{wetestest}$ (°C)	0.90	11
$T_{seasonality}$	1.07	16*
MAP (mm year ⁻¹)	-546**	3**
$P_{seasonality}$	0.20**	15
Reduced dataset (n = 17)		
MAT (°C)	0.60	9
T_{max} (°C)	1.30	11
T_{min} (°C)	-0.202	10
$T_{wetestest}$ (°C)	1.14	9
$T_{seasonality}$	0.53	13*
MAP (mm year ⁻¹)	-541**	2**
$P_{seasonality}$	0.21**	13*

MAP, mean annual precipitation; MAT, mean annual temperature; $P_{seasonality}$, coefficient of variation of monthly average precipitation; T_{max} , average temperature of the warmest month; T_{min} , average temperature of the coldest month; $T_{seasonality}$, SD of monthly average temperature; $T_{wetestest}$, average temperature of the wettest month; *, $P < 0.05$; **, $P < 0.01$.

These factors have been reviewed extensively elsewhere (25, 26) and include traits such as high leaf vein densities, enlarged bundle sheaths, and a propensity for whole-genome duplications. Diversification into cold climates, on the other hand, has happened much more rarely. Thus grasses lend further support to the “tropical conservatism” hypothesis, generated by the observation that relatively few lineages have diversified successfully outside of the tropics, perhaps because low temperature is one of the more difficult environmental barriers for organisms to overcome (27). The Pooideae stand out as an exceptional lineage of grasses in that they occupy both the coldest and the driest climate space in Poaceae. They should not be thought of as typical C_3 grasses nor as a starting point for comparative C_3/C_4 physiology and ecology. Traits that promote chilling and frost tolerance in Pooideae have received attention with regard to understanding the physiology of winter crop cereals (28–30) but have not really been considered in the context of global grassland ecology. It is likely that the evolution of cold tolerance in the Pooideae has been just as relevant as C_4 photosynthesis in shaping current global patterns of grass distribution.

Materials and Methods

Climate Dataset. We extracted all geo-referenced herbarium specimens housed in herbaria and natural history collections that have been made available via the Global Biodiversity Information Facility (GBIF) web portal (<http://www.gbif.org/>). For each set of coordinates we extracted monthly temperature and precipitation values from the Climate Research Unit 10' global gridded climate layers (31). After we purged duplicate records and outliers, our final climate dataset consisted of 1,584,351 independent col-

lection points spanning 10,469 taxa, including subspecies, varieties, and hybrids. From these collection points, we excluded all hybrid taxa as well as taxa that were represented by fewer than 10 independent localities, reducing the dataset to 1,146,612 collection points spanning 4,309 taxa. The mean number of collections per species was 932, although this value was heavily influenced by several taxa (e.g., *Nardus stricta*, *Danthonia decumbens*, *Phragmites australis*, *Poa pratensis*) with very large numbers of collections. After the 5% most heavily collected taxa were removed, the mean number of collections per species for the remaining 95% was 89.

Estimating geographical ranges from herbarium specimens is prone to inherent biases in sampling; for instance, in our dataset, Europe and North America are far more heavily sampled than tropical regions (Fig. S3). However, most of this imbalance results from repeated collecting of a small handful of widespread taxa that represent a very small number of tips on our phylogeny. In general, with the exception of India, the GBIF coverage of the major tropical and subtropical grassland areas (e.g., Eastern Africa, Northern Australia, Northern South America) is quite good.

Phylogenetic Tree Construction. We assembled a 1,230-taxon DNA sequence matrix that consisted of chloroplast regions *atpB* (59 taxa), *matK* (266 taxa), *ndhF* (437 taxa), *rbcl* (251 taxa), *rp16* (176 taxa), and *trnL-trnF* (810 taxa), and nuclear regions *phyB* (93 taxa) and the internal transcribed spacer (ITS; 753 taxa), using the Phylogeny Assembly with Databases (PHLAWD) tool (<http://code.google.com/p/phlawd/>) (32). Before matrix building, we filtered all sequences retrieved from the National Center for Biotechnology Information (NCBI) through our 4,309-taxon list to ensure a complete climate dataset across the tips of the tree. All sequence alignments were conducted using MUSCLE (v. 3.6) (33). The phylogeny was constructed using RAxML 7.1.0 (34) with all genes partitioned, allowing each gene region to have independent parameter estimates for molecular rate matrices. We employed a GTRGAMMA model of nucleotide substitution, with a GTR substitution model and a Γ model of among-site rate heterogeneity. All matrices and trees can be obtained from

Table 2. Mean modeled selection optima for temperature and precipitation variables in Panicoideae, enforcing separate optima for C_3 and C_4 lineages

Variable	OU attraction strength (\pm SD)	C_3 mean (\pm SD)	C_4 mean (\pm SD)	Δ AICc*
MAT (°C)	5.94 (2.29)	21.97 (0.14)	20.12 (0.11)	5.46
T_{max} (°C)	4.66 (0.68)	25.45 (0.14)	25.46 (0.16)	-0.05
T_{min} (°C)	6.08 (2.50)	18.11 (0.15)	14.27 (0.12)	7.91
$T_{wetestest}$ (°C)	5.31 (0.93)	23.27 (0.10)	22.68 (0.10)	1.14
$T_{seasonality}$	6.23 (2.63)	2.60 (0.04)	4.02 (0.03)	11.65
MAP (mm year ⁻¹)	5.33 (1.00)	1781 (25.51)	1269 (16.94)	34.06
$P_{seasonality}$	5.49 (1.18)	0.55 (0.004)	0.58 (0.004)	0.95

*Larger Δ AICc numbers indicate stronger support for a two-optimum model.

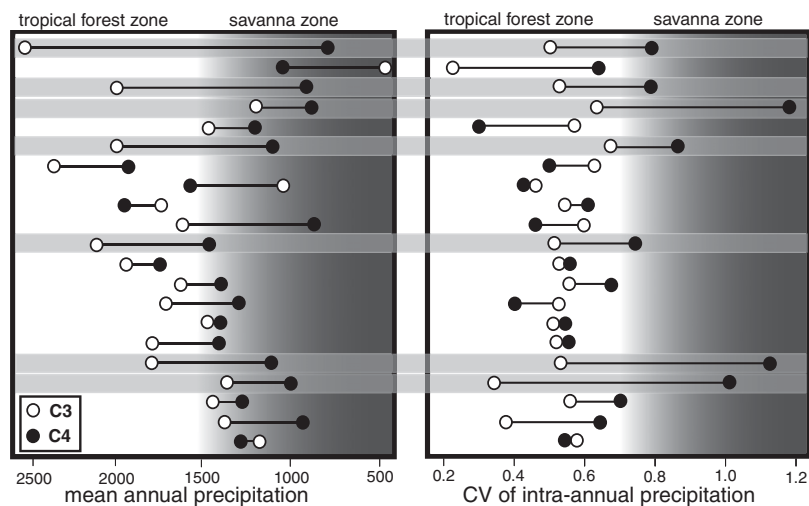


Fig. 3. Maximum likelihood reconstructions of ancestral C_3 and C_4 precipitation niches for 21 C_3/C_4 evolutionary divergences. White dots indicate C_3 value; black dots indicate C_4 value. Background shading indicates gross climate delineations between a closed-canopy tropical forest (white) and open woodland/savanna system (gray); transition occurs around MAP ~ 1500 mm year $^{-1}$, coefficient of variation of precipitation ~ 0.75 . Gray bars highlight C_3/C_4 transitions that are consistent with a movement of the C_4 lineage from tropical forest into open woodland/savanna.

the authors. We rooted Poaceae by including known close relatives *Flagellaria indica*, *Ecdeiocolea monostachya*, and *Joinvillea ascendens*, which subsequently were pruned from the tree for character analyses. Rapid bootstrap analyses were conducted on the dataset before a final maximum likelihood analysis, which used 10 random bootstrap results as starting trees. The Panicoideae were found to be monophyletic in the constructed phylogeny and therefore were used for further analyses. This dataset was excised from the larger dataset and resulted in 299 taxa and 6128 sites. A second phylogenetic analysis on Panicoideae was conducted using MrBayes v. 3.1.2 (35) with a GTR + I + Γ . The MrBayes analysis was conducted with two runs of four chains each run for 10,000,000 generations. Convergence and burn-in was determined by examining time-series plots of likelihood scores and parameter estimates, as well as by examining the effective sample size. Trees were used after removing 3,000,000 generations of burn-in.

Dating. Methods for creating a time-calibrated phylogeny for Poaceae are limited because of the size of the dataset. To overcome this problem, we first constructed a pruned Poaceae tree of 300 tips and ran the nonparametric rate smoothing procedure as implemented in r8s (36). We used the Powell algorithm and restarted three times to verify convergence. We extracted the estimated ages for all internal nodes in the reduced dataset and used these ages as calibrations for the full dataset. To date the remaining nodes in the full dataset, we used the nonparametric dating method PATHd8 (37). For the initial r8s run, we fixed the age of the node subtending crown Poaceae (BEP + PACMAD lineages) to 65 Mya (19). We also set minimum stem ages of 34 Mya for Stipeae and *Chusquea* (7) and set a minimum stem age of *Dicanthelium* to 8 Mya (38). Using fossil phytolith data to date the Poaceae tree is considered controversial by some authors (3, 6); however, it is relative ages, rather than absolute ages, that are important for our analyses. Furthermore, our dates of C_4 origins are largely consistent with those of other recent studies (3, 6), with most occurring between 30 and 18 Mya.

To conduct comparative analyses on the rates of niche evolution and to ensure higher accuracy, we separately obtained time-calibrated phylogenies for the Panicoideae. To accommodate for phylogenetic uncertainty, we dated a sample of 100 trees from the post-burn-in posterior distribution of phylogenetic trees from the MrBayes analysis (see above). Each tree was dated with r8s (36) using the nonparametric rate-smoothing algorithm (39) with the Powell algorithm and restarting three times to verify convergence. Because absolute ages were not necessary for our analyses, we set the root to age 1 to obtain relative ages.

Character Evolution Analyses. We used the Analysis of Traits module in Phylocom (40) to identify 21 phylogenetically independent contrasts between C_3 and C_4 taxa. We also used LASRdisc (v. 1.0) (41) to reconstruct the evolution of photosynthetic shifts using maximum likelihood. Both approaches converged on the same focal nodes. Despite missing several known evolutionary transitions from our dataset (e.g., *Sartidia*, a C_3 lineage in C_4 Aristidoideae;

Merxmuellera rangei, the C_3 sister taxon of C_4 *Centropodia*; and *Neurachne munroi*, which represents a recent C_4 origin within the C_3 *Neurachne*) (3), we have recovered more shifts in pathway than previously had been reported for grasses. Of the 21 nodes, 20 were reconstructed as a C_4 origin rather than as a reversal to C_3 ; the one exception is *Homopholis proluta*, which now stands as a putative reversal to C_3 alongside *Eragrostis walteri* (42) and *Alloteropsis semialata* subsp. *ecksonii* (43) (neither sampled here). Three of our focal nodes result from the addition of taxa that have not been included in previous analyses: *Homopholis proluta*, *Panicum amarum*, and *Panicum decompositum*. These taxa were represented by archived sequence data from NCBI that is not published, and there is no way to confirm that the samples were identified correctly. Another potentially misleading node is the divergence between *Steinchisma laxa* (C_3) and *Steinchisma hians* (a C_3/C_4 intermediate). In the 21-focal node analysis we coded *S. hians* as C_4 . To account for these uncertainties, we ran analyses on a reduced dataset, removing *H. proluta*, *P. amarum*, and *P. decompositum* and recoding *S. hians* as C_3 . This revision resulted in 17 rather than 21 C_3/C_4 divergences. We tested that the magnitude of change in climate variables at focal nodes was significantly different from zero by using a one-sample *t* test and assessed whether the overall directionality of change was significantly positive or negative against the binomial expectation, as recommended in (40).

To determine whether C_3/C_4 species have experienced divergent selection for temperature and precipitation variables, we tested the fit of multiple models of trait evolution using the noncensored approach in BROWNIIE (v. 2.1) (44). The noncensored approach requires ancestral reconstructions for C_3/C_4 at internal nodes, so we used a procedure in BROWNIIE that reconstructs the most likely state at each internal node before testing for model fit. We then reconstructed an OU model (45), allowing the C_4 branches and the C_3 branches to have optimal mean values for climate variables from which values could deviate according to the attraction parameter (OU2). To test for the significance of these results, we compared the Akaike Information Criterion scores corrected for small sample size (AICc), comparing the OU2 model described above with a model with a global mean value (OU1) and a Brownian motion model with a single rate of evolution and with no attraction parameter (BM). We made this comparison for each of the 100 dated trees drawn from the Bayesian posterior distribution and calculated the mean and standard deviation of each estimated parameter. A difference < 2 between the AICc of the OU2 model and the OU1 model was taken as evidence for the OU1 model, whereas a difference > 4 suggested considerable evidence for the OU2 model. In all cases, the difference between AICc of the OU2 model and the BM model was > 200 , suggesting that Brownian motion is a very poor model of trait evolution in this case.

ACKNOWLEDGMENTS. We thank D. Ackerly, J. Beaulieu, A. Bergland, C. Dunn, R. Early, and B. O'Meara for help and advice with analyses and M. Arakaki, M. Donoghue, M. Ogburn, C. Osborne, K. Schmandt, S. Schmerler, E. Spriggs, and A. Williard for helpful discussions and comments on the manuscript.

1. Hatch MD (1971) *Photosynthesis and Photorespiration* eds Hatch MD, Osmond CB, Slayter RO (Academic, San Diego), pp 139–152.
2. Muhaidat R, Sage RF, Dengler NG (2007) Diversity of Kranz anatomy and biochemistry in C₄ eudicots. *Am J Bot* 94:362–381.
3. Christin PA, et al. (2008) Oligocene CO₂ decline promoted C-4 photosynthesis in grasses. *Curr Biol* 18:37–43.
4. Sage RF (2004) The evolution of C-4 photosynthesis. *New Phytol* 161:341–370.
5. Still CJ, Berry JA, Collatz GJ, DeFries RS (2003) Global distribution of C-3 and C-4 vegetation: Carbon cycle implications. *Global Biogeochem Cycles* 17:1006.
6. Vicentini A, Barber JC, Aliscioni SS, Giussani LM, Kellogg EA (2008) The age of the grasses and clusters of origins of C-4 photosynthesis. *Glob Change Biol* 14:2963–2977.
7. Stromberg CAE (2005) Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. *Proc Natl Acad Sci USA* 102: 11980–11984.
8. Cerling TE, et al. (1997) Global vegetation change through the Miocene/Pliocene boundary. *Nature* 389:153–158.
9. Ehleringer JR, Cerling TE, Helliker BR (1997) C-4 photosynthesis, atmospheric CO₂ and climate. *Oecologia* 112:285–299.
10. Teeri JA, Stowe LG (1976) Climatic patterns and distribution of C₄ grasses in North-America. *Oecologia* 23:1–12.
11. Tieszen LL, Senyimba MM, Imbamba SK, Troughton JH (1979) Distribution of C₃-grass and C₄-grass and carbon isotope discrimination along an altitudinal and moisture gradient in Kenya. *Oecologia* 37:337–350.
12. Vogel JC, Fuls A, Danin A (1986) Geographical and environmental distribution of C-3 and C-4 grasses in the Sinai, Negev, and Judean deserts. *Oecologia* 70:258–265.
13. Hattersley PW (1983) The distribution of C₃-grass and C₄-grasses in Australia in relation to climate. *Oecologia* 57:113–128.
14. Ehleringer JR (1978) Implications of quantum yield differences on distributions of C₃ and C₄ grasses. *Oecologia* 31:255–267.
15. Collatz GJ, Berry JA, Clark JS (1998) Effects of climate and atmospheric CO₂ partial pressure on the global distribution of C-4 grasses: Present, past, and future. *Oecologia* 114:441–454.
16. Ehleringer J, Bjorkman O (1977) Quantum yields for CO₂ uptake in C-3 and C-4 plants - dependence on temperature, CO₂, and O₂ concentration. *Plant Physiol* 59:86–90.
17. Epstein HE, et al. (2002) The relative abundance of three plant functional types in temperate grasslands and shrublands of North and South America: Effects of projected climate change. *J Biogeogr* 29:875–888.
18. Paruelo JM, Lauenroth WK (1996) Relative abundance of plant functional types in grasslands and shrublands of North America. *Ecol Appl* 6:1212–1224.
19. Prasad V, Stromberg CAE, Alimohammadian H, Sahn A (2005) Dinosaur coprolites and the early evolution of grasses and grazers. *Science* 310:1177–1180.
20. Maddison WP (2000) Testing character correlation using pairwise comparisons on a phylogeny. *J Theor Biol* 202:195–204.
21. Edwards EJ, Still CJ (2008) Climate, phylogeny, and the ecological distribution of C₄ grasses. *Ecol Lett* 11:266–276.
22. Osborne CP, Freckleton RP (2009) Ecological selection pressures for C₄ photosynthesis in the grasses. *Proc Roy Soc B Biol Sci* 276:1753–1760.
23. Holdridge LR (1967) *Life Zone Ecology* (Tropical Science Center, San Jose, Costa Rica).
24. Nix HA (1983) *Ecosystems of the World 13: Tropical Savannas*, ed Bourliere F (Elsevier Scientific, Amsterdam), pp 37–62.
25. Monson RK (1999) *C₄ Plant Biology*, eds Sage RF, Monson RK (Academic, San Diego), pp 377–410.
26. Sage RF (2001) Environmental and evolutionary preconditions for the origin and diversification of the C-4 photosynthetic syndrome. *Plant Biol* 3:202–213.
27. Donoghue MJ (2008) A phylogenetic perspective on the distribution of plant diversity. *Proc Natl Acad Sci USA* 105:11549–11555.
28. Badawi M, Danyluk J, Boucho B, Houde M, Sarhan F (2007) The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs. *Mol Genet Genomics* 277:533–554.
29. Sowinski P, Rudzinska-Langwald A, Dalbiak A, Sowinska A (2001) Assimilate export from leaves of chilling-treated seedlings of maize. The path to vein. *Plant Physiol Biochem* 39:881–889.
30. del Viso F, et al. (2009) Molecular characterization of a putative sucrose:fructan 6-fructosyltransferase (6-SFT) of the cold-resistant patagonian grass *Bromus pictus* associated with fructan accumulation under low temperatures. *Plant Cell Physiol* 50: 489–503.
31. New M, Lister D, Hulme M, Makin I (2002) A high-resolution data set of surface climate over global land areas. *Clim Res* 21:1–25.
32. Smith SA, Beaulieu JM, Donoghue MJ (2009) Mega-phylogenies for comparative biology: An alternative to supertree and supermatrix approaches. *BMC Evol Biol* 9:37.
33. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
34. Stamatakis A (2006) RaxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
35. Ronquist F, Huelsenbeck JP (2003) MrBayes3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
36. Sanderson MJ (2003) R8s: Inferring absolute rates of evolution and divergence times in the absence of a molecular clock. *Bioinformatics* 19:301–302.
37. Britton T, Anderson CL, Jacquet D, Lundqvist S, Bremer K (2007) Estimating divergence times in large phylogenetic trees. *Syst Biol* 56:741–752.
38. Thomasson JR (1978) Observations on the characteristics of the lemma and palea of the late Cenozoic grass *Panicum elegans*. *Am J Bot* 65:34–39.
39. Sanderson MJ (1997) A nonparametric approach to estimating divergence times in the absence of rate constancy. *Mol Biol Evol* 14:1218–1231.
40. Webb CO, Ackerly DD, Kembel SW (2008) Phylocom: Software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* 24:2098–2100.
41. Jackson VK (2004) *LASRDis: Likelihood Ancestral State Reconstruction for Discrete Characters* (ed. 1.0).
42. Ellis RP (1984) *Eragrostis-Walteri*—a 1st record of non-Kranz leaf anatomy in the sub-family Chloridoideae (Poaceae). *S Afr J Bot* 3:380–386.
43. Ibrahim DG, Burke T, Ripley BS, Osborne CP (2009) A molecular phylogeny of the genus *Alloterospis* (Panicoidae, Poaceae) suggests an evolutionary reversion from C-4 to C-3 photosynthesis. *Ann Bot (Lond)* 103:127–136.
44. O'Meara BC, Ane C, Sanderson MJ, Wainwright PC (2006) Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60:922–933.
45. Butler MA, King AA (2004) Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *Am Nat* 164:683–695.

Supporting Information

Edwards and Smith 10.1073/pnas.0909672107

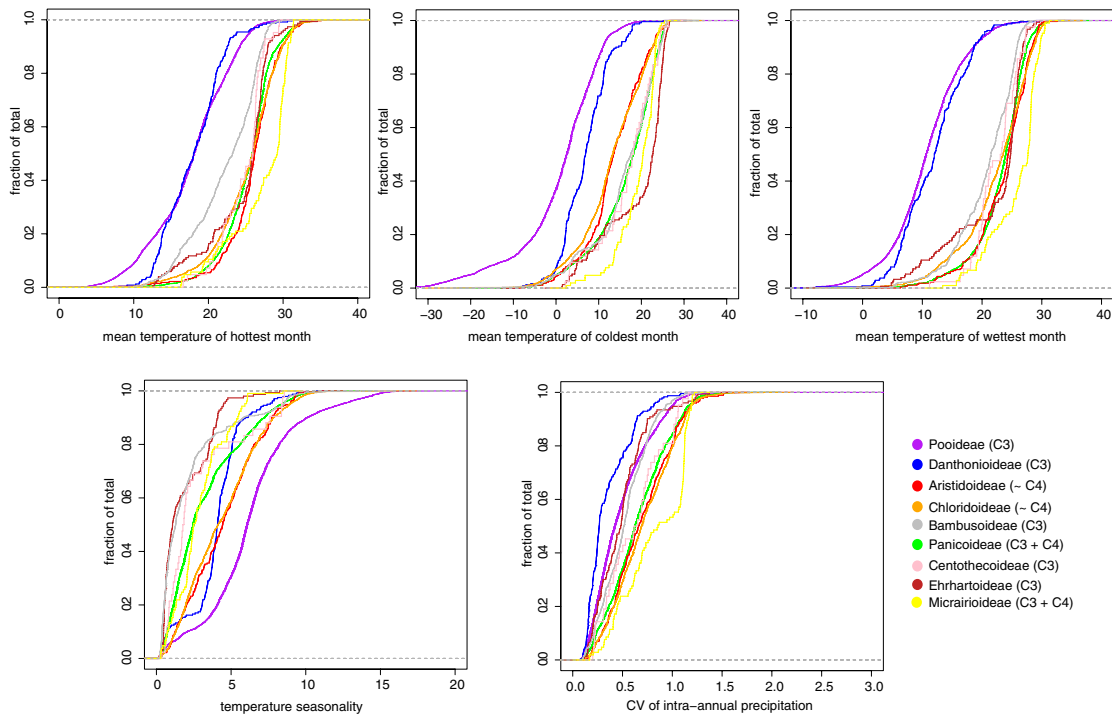


Fig. S1. Species accumulation curves for climate variables, sorted by major grass lineage. These data represent 1,584,351 independent collection localities spread across 10,469 taxa. Each point in the curve is a species' mean value.

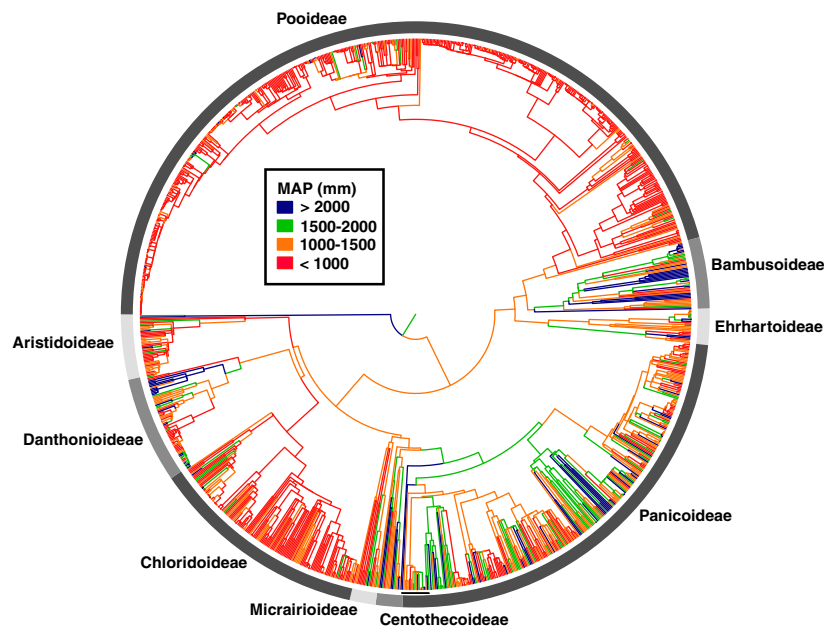


Fig. S2. Maximum likelihood reconstructions of mean annual precipitation (MAP), using species' mean values generated from 1,146,612 geo-referenced herbarium specimens.

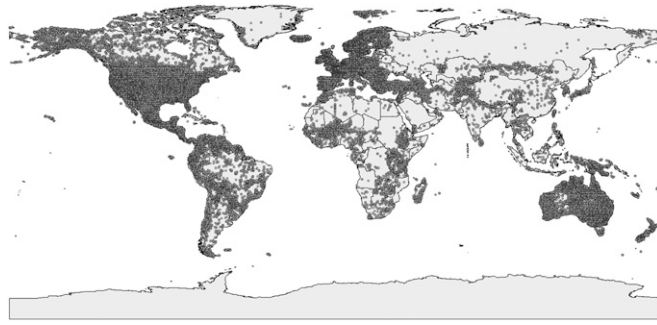


Fig. S3. Global distribution of 1,584,351 geo-referenced grass collections, accessed via the Global Biodiversity Information Facility web portal (<http://www.gbif.org>).

Table S1. Node descriptions, bootstrap support, and climatic ranges for descendant species of C₃/C₄ nodes

Node description	ML bootstrap support values	# C ₃ species	# C ₄ species	C ₃ range MAT	C ₄ range MAT	C ₃ range T _{max}	C ₄ range T _{max}	C ₃ range T _{min}	C ₄ range T _{min}	C ₃ range T _{wettest}	C ₄ range T _{wettest}
Danthoniodeae/ Chloridoideae	61	71	148	4–19	8–28	10–29	16–34	–7–18	–7–25	2–22	5–29
Danth+Chloridoid/ Aristidoideae	83	219	43	4–28	12–28	10–34	21–32	–7–25	–2–24	2–29	12–29
C ₃ Centothecoideae/ Tristachya avenaceae	37	4	1	17–25	20–20	24–26	23–23	11–24	17–17	12–26	22–22
C ₃ Centothecoideae/ Loudetia simplex	39	3	1	15–18	25–25	26–27	28–28	3–7	22–22	18–20	24–24
<i>Streptostachys asperifolia</i> / Axonopus	51*	1	2	26–26	20–25	27–27	25–26	25–25	13–24	26–26	22–24
<i>Ichnanthus</i> / Axonopus +Paspalum	69*	7	28	20–26	16–25	22–27	18–28	13–25	6–23	21–26	17–26
<i>Steinchisma laxa</i> / Panicum decompositum	85*	1	1	21–21	24–24	27–27	26–26	13–13	22–22	25–25	25–25
<i>Steinchisma</i> / Steinchisma hians	96*	3	1	21–24	18–18	25–27	27–27	13–21	9–9	24–25	21–21
<i>Panicum pilosum</i> / Panicum amarum	70*	1	1	16–16	25–25	26–26	26–26	25–25	25–25	23–23	25–25
<i>Steinchisma-Hymanachne</i> / Leptocoryphium lanatum	67*	11	1	16–25	24–24	25–27	26–26	6–23	22–22	21–26	25–25
<i>Panicum mertensii</i> / Panicum caricoides	96*	1	1	26–26	26–26	27–27	28–28	25–25	24–24	26–26	26–26
<i>Panicum mertensii-caricoides</i> / Panicum stenodes	100*	2	1	26–26	26–26	27–28	27–27	25–25	24–24	26–26	26–26
<i>Apochloa chnoodes</i> / Panicum prionits clade	94*	1	4	21–21	15–23	22–22	25–27	20–20	3–18	21–21	19–25
<i>Homolepis</i> etc./ Cyphoanthus discrepans	0*	11	1	15–26	26–26	22–28	28–28	3–25	25–25	19–26	26–26
X = 10 Paniceae/ Andropogoneae	46*	60	103	15–26	9–27	18–28	19–31	3–25	–4–25	17–26	10–30
<i>Homopholis proluta</i> / Panicum repens	100*	1	1	17–17	20–20	24–24	27–27	9–9	13–13	16–16	20–20
<i>Panicum sp.</i> / Echinochloa	0*	4	11	20–22	10–28	23–26	18–32	16–19	1–23	20–22	12–29
<i>Acroceras-Ottochloa</i> / Alloteropsis	0*	2	2	24–24	23–24	26–26	27–27	21–22	18–20	25–26	26–26
Forest shade clade/ Digitaria-Cenchrus-Setaria	0*	47	85	10–28	7–29	18–32	15–34	–2–24	–2–24	12–29	13–29
Centothecoideae +Panicoideae	94	314	1	7–29	23–23	16–34	25–25	–4–25	20–20	10–30	24–24
Loudetiopsis chrysothrix											
<i>Isachne</i> / Eriachne	100	2	5	21–21	22–27	22–23	26–31	17–19	14–23	21–22	26–28

Because of high topological uncertainty within Panicoideae, this region of the tree was subjected to additional Bayesian sensitivity analyses. Bold lettering indicates the C₄ component of the sister taxa. Shaded pairs were removed for the 17-node analyses because of uncertainties either in character coding (*Steinchisma hians*) or in taxon identification from National Center for Biotechnology Information entries (remaining three). Temperature values are °C; precipitation values are mm^{–year}. MAT, mean annual temperature; ML, maximum likelihood; T_{max}, average temperature of the warmest month; T_{min}, average temperature of the coldest month; T_{wettest}, average temperature of the wettest month.

*The node is found within Panicoideae.

Table S2. Node descriptions, bootstrap support, and climatic ranges for descendant species of C₃/C₄ nodes

Node description	ML bootstrap support value	# C ₃ species	# C ₄ species	C ₃ range T _{seasonality}	C ₄ range T _{seasonality}	C ₃ range MAP	C ₄ range MAP	C ₃ range P _{seasonality}	C ₄ range P _{seasonality}
Danthoniodeae/Chloridoideae	61	71	148	0.3–9.4	0.8–10.8	237–3540	70–2601	0.14–0.75	0.20–1.4
Danth+Chloridoid/Aristidoideae	83	219	43	0.3–10.8	0.7–9.5	70–3540	138–2293	0.14–1.42	0.19–1.5
C ₃ Centothecoideae/ <i>Tristachya avenaceae</i>	37	4	1	0.7–4.7	2.3–2.3	425–2906	1140–1140	0.34–0.62	1.13–1.13
C ₃ Centothecoideae/ <i>Loudetia simplex</i>	39	3	1	7.4–8.4	1.9–1.9	1187–1337	1019–1019	0.20–0.22	1.01–1.01
<i>Streptostachys asperifolia</i> /Axonopus	51*	1	2	0.8–0.8	0.6–4.4	2335–2335	1606–2288	0.38–0.65	0.63–0.63
<i>Ichnanthus</i> /Axonopus+ <i>Paspalum</i>	69*	7	28	0.5–4.4	0.7–7.3	1606–2823	499–3054	0.38–0.65	0.24–1.1
<i>Steinchisma laxa</i> / <i>Panicum decompositum</i>	85*	1	1	5.2–5.2	1.6–1.6	936–936	1979–1979	0.79–0.79	0.54–0.54
<i>Steinchisma</i> / <i>Steinchisma hians</i>	96*	3	1	1.6–5.2	6.5–6.5	936–1979	1239–1239	0.46–0.79	0.31–0.31
<i>Panicum pilosum</i> / <i>Panicum amarum</i>	70*	1	1	7.4–7.4	1.1–1.1	813–813	2496–2496	0.79–0.79	0.52–0.52
<i>Steinchisma-Hymanachne</i> / <i>Leptocoryphium lanatum</i>	67*	11	1	1.1–7.4	1.4–1.4	813–2529	1931–1931	0.23–0.95	0.61–0.61
<i>Panicum mertensii</i> / <i>Panicum caricoides</i>	96*	1	1	0.9–0.9	1.0–1.0	1183–1183	914–914	0.65–0.65	1.2–1.2
<i>Panicum mertensii-caricoides</i> / <i>Panicum stenodes</i>	100*	2	1	0.9–1.0	1.0–1.0	914–1183	1844–1844	0.65–1.2	0.69–0.69
<i>Apochloa chnoodes</i> / <i>Panicum prionitis</i> clade	94**	1	4	0.5–0.5	3.6–8.3	1061–1061	1198–1919	0.46–0.46	0.20–0.73
<i>Homolepis</i> etc./ <i>Cyphonanthus discrepans</i>	0*	11	1	0.5–8.3	0.9–0.9	914–2520	1319–1319	0.20–1.2	0.42–0.42
X = 10 Paniceae/ <i>Andropogoneae</i>	46*	60	103	0.5–8.3	0.6–9.4	499–3054	409–2571	0.20–1.2	0.19–1.3
<i>Homopholis proluta</i> / <i>Panicum repens</i>	100*	1	1	5.6–5.6	5.1–5.1	486–486	1017–1017	0.24–0.24	0.64–0.64
<i>Panicum</i> sp./ <i>Echinochloa</i>	0*	4	11	1.8–3.8	3.0–8.9	1500–1803	429–1448	0.33–0.99	0.23–1.3
<i>Acroceras-Ottochloa</i> / <i>Alloteropsis</i>	0*	2	2	1.3–2.1	2.5–3.4	1806–2414	1051–1530	0.47–0.61	0.72–0.82
Forest shade clade/ <i>Digitaria-Cenchrus-Setaria</i>	0*	47	85	0.8–9.1	0.8–9.1	364–2417	258–3180	0.18–1.32	0.11–1.39
Centothecoideae+Panicoideae/ <i>Loudetiopsis chrysothrix</i>	94	314	1	0.5–9.4	1.8–1.8	258–3180	1296–1296	0.11–1.39	0.70–0.70
<i>Isachne</i> / <i>Eriachne</i>	100	2	5	0.8–2.1	1.9–5.5	1615–2527	491–1508	0.61–0.73	0.64–1.17

Because of high topological uncertainty within Panicoideae, this region of the tree was subjected to additional Bayesian sensitivity analyses. Bold lettering indicates the C₄ component of the sister taxa. Shaded pairs were removed for the 17-node analyses because of uncertainties either in character coding (*Steinchisma hians*) or in taxon identification from National Center for Biotechnology Information entries (remaining three). Temperature values are in °C; precipitation values are mm^{-year}. MAP, mean annual precipitation; ML, maximum likelihood; P_{seasonality}, coefficient of variation of monthly average precipitation; T_{seasonality}, standard deviation of monthly average temperature.

*The node is found within Panicoideae.

Table S3. Independent contrast analyses across Bayesian posterior distribution of 7,001 Panicoideae topologies

Variable	Mean value of C ₃ node (±1 SE)	Mean value of C ₄ node (±1 SE)	Mean magnitude of contrast (C ₄ relative to C ₃)	# trees that were positive (P < 0.05)	# trees that were negative (P < 0.05)	# trees that were not different from zero (P < 0.05)	# trees with significant sign tests
MAT (°C)	22.3 (± 0.04)	21.3 (± 0.04)	-1.00	0	2	6,999	29 (-)
T _{max} (°C)	25.4 (±0.02)	26.0 (±0.02)	0.64	120	0	6,881	36 (+)
T _{min} (°C)	18.9 (±0.07)	16.0 (±0.07)	-2.83	0	320	6,681	33 (-)
T _{wettest} (°C)	23.1 (±0.04)	23.4 (±0.03)	0.26	0	0	7,001	0-
T _{seasonality}	2.35 (±0.03)	3.61 (±0.03)	1.27	768	0	6,233	1,256 (+)
MAP (mm year ⁻¹)	1762 (±7.3)	1361 (±4.5)	-400	0	5,442	1,559	1,961 (-)
P _{seasonality}	0.50 (±0.00)	0.62 (±0.00)	0.11	4,441	0	2,590	331 (+)

MAP, mean annual precipitation; MAT, mean annual temperature; P_{seasonality}, coefficient of variation of monthly average precipitation; T_{max}, average temperature of the warmest month; T_{min}, average temperature of the coldest month; T_{seasonality}, SD of monthly average temperature; T_{wettest}, average temperature of the wettest month.