Anatomical enablers and the evolution of C_4 photosynthesis in grasses

Pascal-Antoine Christin[a], Colin P. Osborne[b], David S. Chatelet[a], J. Travis Columbus[c], Guillaume Besnard[d], Trevor R. Hodkinson[e,f], Laura M. Garrison[a], Maria S. Vorontsova[g], and Erika J. Edwards[a,1]

[a]Department of Ecology and Evolutionary Biology, Brown University, Providence, RI 02912; [b]Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, United Kingdom; [c]Rancho Santa Ana Botanic Garden, Claremont Graduate University, Claremont, CA 91711; [d]Unité Mixte de Recherche 5174, Centre National de la Recherche Scientifique-Université Paul Sabatier-Ecole Nationale de Formation Agronomique, 31062 Toulouse Cedex 9, France; [e]School of Natural Sciences, Trinity College Dublin, Dublin 2, Ireland; [f]Trinity Centre for Biodiversity Research, Trinity College Dublin, Dublin 2, Ireland; and [g]Herbarium, Library, Art and Archives, Royal Botanic Gardens, Kew, Surrey TW9 3AE, United Kingdom

Edited by Elizabeth A. Kellogg, University of Missouri, St. Louis, MO, and accepted by the Editorial Board November 27, 2012 (received for review September 27, 2012)

C_4 photosynthesis is a series of anatomical and biochemical modifications to the typical C_3 pathway that increases the productivity of plants in warm, sunny, and dry conditions. Despite its complexity, it evolved more than 62 times independently in flowering plants. However, C_4 origins are absent from most plant lineages and clustered in others, suggesting that some characteristics increase C_4 evolvability in certain phylogenetic groups. The C_4 trait has evolved 22–24 times in grasses, and all origins occurred within the PACMAD clade, whereas the similarly sized BEP clade contains only C_3 taxa. Here, multiple foliar anatomy traits of 157 species from both BEP and PACMAD clades are quantified and analyzed in a phylogenetic framework. Statistical modeling indicates that C_4 evolvability strongly increases when the proportion of vascular bundle sheath (BS) tissue is higher than 15%, which results from a combination of short distance between BS and large BS cells. A reduction in the distance between BS occurred before the split of the BEP and PACMAD clades, but a decrease in BS cell size later occurred in BEP taxa. Therefore, when environmental changes promoted C_4 evolution, suitable anatomy was present only in members of the PACMAD clade, explaining the clustering of C_4 origins in this lineage. These results show that key alterations of foliar anatomy occurring in a C_3 context and preceding the emergence of the C_4 syndrome by millions of years facilitated the repeated evolution of one of the most successful physiological innovations in angiosperm history.

precursor | preadaptation | phylogeny | Poaceae

The evolvability of some traits can influence the capacity of organisms to adapt to environmental changes, but the factors dictating differences in evolvability among groups are often unknown. Climate changes during the Cenozoic included a marked decline of atmospheric CO_2 levels around 30 million years ago (1, 2), which lowered the efficiency of photosynthesis in plants (3). Several lineages of flowering plants adapted to the resulting CO_2 depletion, particularly in warm, open, and/or dry conditions, by evolving mechanisms that concentrate CO_2 around chloroplasts, where the Calvin cycle uses CO_2 and energy to produce sugars (4, 5). In the so-called C_4 photosynthetic pathway, this increase in internal CO_2 concentration results from a spatial segregation of carbon assimilation and reduction (3). The vast majority of C_4 plants use a dual-cell system, where the assimilation of CO_2 into organic compounds occurs in mesophyll (M) cells of the leaf, whereas carbon reduction by the Calvin cycle, which occurred in the M cells in their C_3 ancestors, happens in the vascular bundle sheath (BS).

The C_4 pathway requires both a specialized foliar anatomy and the addition of a complex biochemical pathway that assimilates and delivers carbon to the Calvin cycle (3). Despite this complexity, it evolved more than 62 times independently in distantly related groups of flowering plants (5). These recurrent C_4 origins were probably facilitated by the existence in most C_3 plants of metabolic modules that are suitable for the C_4 pathway and can be recruited for this function through relatively few mutations (6, 7). In addition, the photorespiratory pump based on glycine decarboxylase is a likely evolutionary stable intermediate phenotype on the road from C_3 to C_4 (3, 8). However, C_4 origins are not randomly distributed across the plant tree of life; although most large (species rich) clades of plants completely lack C_4 taxa, others contain an astonishingly high number of C_3-to-C_4 transitions (5). This is particularly the case with grasses, in which C_4 photosynthesis evolved independently 22–24 times (9). Most grass species belong to either the BEP or PACMAD clade, which contain approximately the same number of species (*ca.* 5,423 and 5,706, respectively); however, the repeated C_4 origins occurred only in the PACMAD clade (9).

The homoplastic tendency of some groups, manifested in the extreme clustering of C_4 origins, suggests that several plant lineages possess traits that increase C_4 evolvability (10, 11). Such enablers may include ecological factors, as well as genomic and anatomical properties (10). For optimal functioning of the C_4 pathway, foliar anatomy must be altered to increase the relative proportion of BS cells and reduce the average path length between M and BS cells (12–15). These modifications guarantee an efficient centripetal increase of CO_2 concentration, but they also imply that C_4 evolution involved many mutations, unless the C_3 ancestors already possessed some C_4-like anatomical characters. Putative C_4 preadaptations have been detected in C_3 taxa closely related to C_4 species in small groups of eudicots with a limited number of C_4 origins (16–19). Recently, a preliminary survey of the proportion of BS tissue and the interveinal distance (IVD) in different C_3 and C_4 grasses suggested that a larger proportion of BS tissue characterizes the C_3 PACMAD, which are closer to C_4 values (20). The same study concluded that short IVD evolved later and was not an important determinant of C_4 evolvability (20). In this work, the different traits that, together, determine the amount of BS tissue or the IVD (e.g., number and size of different types of cells) were not measured separately. The species sampling did not include species outside the BEP and PACMAD clades, preventing any inference about the

EVOLUTION

Author contributions: P.-A.C., C.P.O., D.S.C., and E.J.E. designed research; P.-A.C., D.S.C., J.T.C., G.B., T.R.H., L.M.G., and M.S.V. performed research; P.-A.C. analyzed data; and P.-A.C., C.P.O., and E.J.E. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. E.A.K. is a guest editor invited by the Editorial Board.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. HF558456–HF558529) and Dataset S1, the phylogenetic matrix, and the phylogenetic tree have been deposited in the Dryad repository (http://dx.doi.org/10.5061/dryad.6j9r7).

[1]To whom correspondence should be addressed. E-mail: Erika_Edwards@brown.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1216777110/-/DCSupplemental.

directionality of trait evolution. In addition, the analyses did not use phylogenetic comparative methods, precluding a formal statistical evaluation of the link between foliar characteristics and C_4 evolvability.

In the present work, we measured anatomical traits in 157 grasses representing the BEP and PACMAD clades and several outgroups, as well as the diversity of photosynthetic types known within the group. A phylogenetic framework was used to test statistically for the presence of C_4 precursors and to locate them on the phylogeny, and the evolution of foliar traits was modeled in a phylogenetic context to (*i*) reconstruct the leaf modifications that occurred during the transition from C_3 to C_4 photosynthesis and (*ii*) identify the leaf traits that might explain the propensity of some grass lineages to evolve C_4 photosynthesis.

Results

Phylogenetic Patterns. A total of 172 cross-sections were measured, representing 157 species. Six of these represent outgroup taxa (either grasses sister to the BEP-PACMAD clade or members of families affiliated with grasses), 37 C_3 BEP, 62 C_3 PACMAD, and 52 C_4 PACMAD. The two BS layers that characterize all C_3 and many C_4 grasses [outer sheath (OS) and inner sheath (IS)] were measured separately (Fig. S1). Of the C_4 taxa, 26 use the OS for carbon reduction (C_4-OS) and 26 use the IS (C_4-IS). Only five of the C_4-IS species also possessed an OS layer. A total of 17 C_4 lineages were sampled out of 22–24 described previously (9). This included five C_4-OS lineages, nine C_4-IS, and three that contained both C_4-IS and C_4-OS (Fig. 1). The five C_4 lineages not sampled in this study all belong to Paspaleae and are species-poor groups, putatively using a C_4-IS pathway (21, 22).

The model of C_4 evolution that assumed the existence of precursors (precursor_2) was significantly better than the binary model [Akaike information criterion (AIC), 135.5; AIC difference of 4.7] but only slightly better than the model with equal rates of transition from C_3 to precursor and from precursor to C_4 (precursor_1; difference of AIC, 0.4). In the precursor_2 model, the rate of C_3-to-precursor transition was estimated to be 0.003, and the rate of precursor-to-C_4 transition was estimated to be 0.016, which suggests that some characteristics evolved a small number of times and favored the subsequent repeated origins of C_4 photosynthesis (Fig. 1). A single origin of the precursor was inferred along the branch leading to the PACMAD clade (probability of precursor at the stem, 0.05; at the crown, 1.0) and was followed by multiple transitions to C_4.

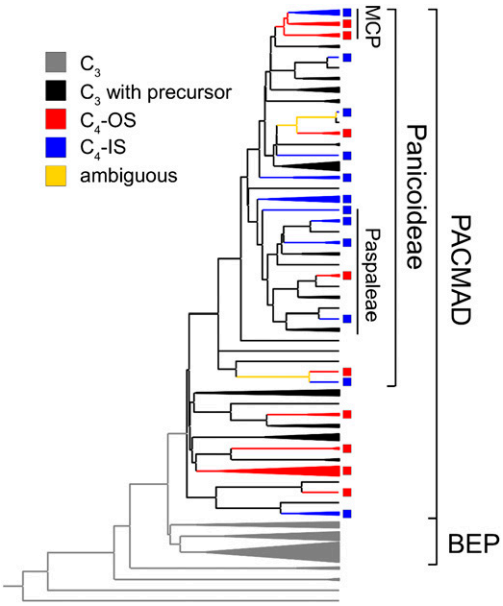

Fig. 1. Sampled species and phylogenetic distribution of photosynthetic types. This time calibrated phylogeny was used in comparative analyses. The C_4-IS and C_4-OS groups are indicated by colored squares. Putative ancestral photosynthetic types are indicated as colors on branches. The evolution of a precursor character was identified with the precursor_2 model as implemented in r8s. The main groups are delimited on the right.

Description of Variables. The individual anatomical trait values all overlapped between C_3 and C_4 taxa, but C_4-OS had higher proportions of OS tissue [%OS = OS area/(OS area + M area)] and C_4-IS higher proportions of IS tissue [%IS = IS area/(IS area + M area)] (Fig. 2). The larger %IS and %OS values result mainly from a decreased amount of M tissue per vein in both C_4-OS and C_4-IS compared with C_3 taxa, which is the consequence of a reduction in BS distance (BSD). The area values of OS tissue per vein and the width of OS cells largely overlapped between C_4-OS species and C_3 taxa. By contrast, the area of IS tissue per vein tends to be larger in C_4-IS than in C_4-OS and C_3 taxa and this is attributable mainly to an increase of the width of IS cells in C_4-IS. The other variables are not consistently different between C_3 and C_4 taxa (Fig. S2).

The C_3 BEP and C_3 PACMAD do not differ in amount of M tissue per vein, which is reduced compared with the outgroups (Fig. 2). However, C_3 PACMAD have greater areas of OS tissue per vein compared with C_3 BEP, which is attributable to larger OS cells. For these two traits, C_3 PACMAD do not differ from C_4-OS species. This results in %OS values in C_3 PACMAD that are intermediate between C_3 BEP and C_4-OS (Fig. 2). The C_3 BEP show smaller IS and OS sizes and, consequently, lower areas of OS and IS per vein compared with the outgroups (Fig. 2).

Comparative Analyses. The evolutionary dynamics of several characters were best explained by a stabilizing selection [Ornstein–Uhlenbeck (OU)] model with a single optimum when considering C_3 taxa only (Table 1). However, M area per vein and BSD evolved around different optima (OU models) in the outgroups compared with the BEP-PACMAD clade (Table 1). Similarly, the optima of OS area per vein, OS width and IS width were reduced in BEP compared with both the C_3 PACMAD and outgroups, and the optima of %OS and leaf thickness were increased in C_3 PACMAD relative to both the BEP clade and the outgroups, whereas the optimum of the ratio of thicknesses was decreased (Table 1). Within PACMAD, the mean vein area was best modeled by an OU model with a single optimum (Table 1). The optima of M area per vein and leaf thickness were decreased in C_4 (both C_4-IS and C_4-OS) compared with C_3 PACMAD. C_4-IS had a lower optimum for OS per vein and OS width but a higher optimum for IS per vein and IS width. Finally, %IS, %OS, and BSD were best modeled with one optimum per photosynthetic type (Table 1).

Variation in the amount of M tissue per vein is explained by a combination of changes in BSD, thickness, ratio of thicknesses, and mean area of veins (R^2 = 0.94), with BSD and thickness alone explaining most of the variation (R^2 = 0.92). Changes in OS area are driven by changes in the width of OS cells, thickness, mean area of veins, and BSD (R^2 = 0.95), with OS cell width and thickness explaining most of the variation (R^2 = 0.94). Finally, modifications of IS area are attributable to a combination of changes in IS cell width, thickness and mean vein area (R^2 = 0.96), with most of the variance explained by IS cell width and vein area (R^2 = 0.96).

Ancestral Reconstructions. The ancestral reconstructions indicate that a higher %OS evolved at the base of the PACMAD clade, as hypothesized previously (20), and was maintained with further increases in multiple clades that led to the emergence of C_4-like

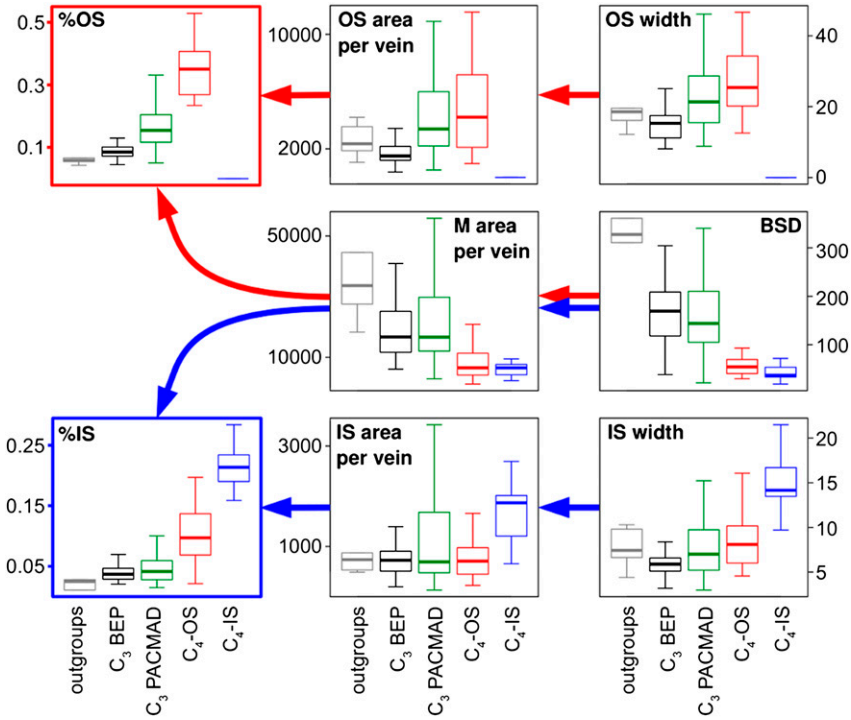

Fig. 2. Distribution of trait values among clades and photosynthetic types. The distribution of eight variables is summarized by boxplots individually for the outgroups (gray; $n = 6$), C_3 BEP (black; $n = 37$), C_3 PACMAD (green; $n = 62$), C_4-OS (red; $n = 26$), and C_4-IS (blue; $n = 26$). Boxes indicate the 25th and 75th percentiles, and whiskers indicate the extreme values still within 1.5 interquartile range of the upper and lower quartiles. Arrows indicate inferred effect, with the changes linked to C_4-OS evolution in red and the changes linked to C_4-IS evolution in blue.

%OS in several parts of the phylogeny (Fig. 3). This increase of %OS was attributable to an increase of OS cell size, which is especially large in several PACMAD lineages, whereas in the BEP clade, the Bambusoideae and Pooideae lineages are generally characterized by very small OS cells (Fig. S3). By contrast, %IS shows smaller differences between BEP and C_3 PACMAD lineages, and C_4-like values evolved only a small number of times: three in the PACMAD and one in the BEP (Fig. S4). However, larger IS cells are more frequent in PACMAD; in BEP, they occurred only in one species (*Lygeum spartum*; Fig. S5). Small BSD values characterize the whole BEP-PACMAD clade, but large values appeared repeatedly within both BEP and PACMAD (Fig. S6).

The inferred transitions between C_3 and C_4-OS involved on average a 1.8-fold increase of %OS (Fig. 4). This was attributable to a strong decrease in the amount of M per vein (except in *Alloteropsis*), which was driven by a strong reduction in BSD. The size of OS cells and, consequently, the amount of OS tissue per vein were not significantly increased, except in the transition that gave rise to the C_4-OS *Alloteropsis*. The evolution of a C_4-IS type from C_3 ancestors required more dramatic changes, with, on average, a 4.0-fold increase in %IS. This resulted from a large reduction of BSD, which strongly reduced the amount of M per vein and, in most cases, an increase of IS width and consequently IS area per vein.

Modeling of Anatomical Enablers. The model that assumed an increase of the probability of being sister to a C_4-OS lineage above a given %OS ratio was significantly better than the null model ($\chi^2 = 12.8$; df = 1; $P < 0.001$). The optimal threshold for a shift of transition rates was 0.15, which appeared at the base of the PACMAD clade and in some BEP (Fig. 3). Above this value, the rate of transition to C_4-OS was 0.010. The evolvability of C_4-OS was also adequately explained by the size of OS cells alone ($\chi^2 = 5.3$;

Table 1. Modeling of evolution among taxonomic groups and photosynthetic types

	C_3 only			PACMAD only (C_3 and C_4)		
Trait	Model*	Groups†	Optima	Model*	Groups†	Optima
%OS	OU2	P [out+B]	0.18 [0.09]	OU3	C_3 OS IS	0.18/0.40/−0.01
%IS	OU	—	0.05	OU3	C_3 OS IS	0.05/0.10/0.22
M per vein	OU2	Out [B+P]	38,260 [17,371]	OU2	C_3 [OS+IS]	18,331 [6,698]
BSD	OU2	Out [B+P]	326 [133]	OU3	C_3 OS IS	140/57/42
OS per vein	OU2	B [out+P]	1,586 [3,534]	OU2	IS [C_3+OS]	−1 [4,774]
OS width	OU2	B [out+P]	14.1 [21.6]	OU2	IS [C_3+OS]	−0.8 [24.2]
IS per vein	OU	—	852	OU2	IS [C_3+OS]	3,080 [943]
IS width	OU2	B [out+P]	6.0 [7.6]	OU2	IS [C_3+OS]	19.4 [7.9]
Mean vein area	OU	—	1,050	OU	—	1,049
Leaf thickness	OU2	P [out+B]	153 [111]	OU2	C_3 [OS+IS]	146 [114]
Ratio of thicknesses	OU2	P [out+B]	0.44 [0.55]	OU2	OS [C_3+IS]	0.27 [0.44]

B, BEP; C_3, C_3 PACMAD; IS, C_4-IS; OS, C_4-OS; OU2, OU with different optima for two groups; OU3, OU with different optima for three groups; out, outgroups; P, PACMAD.

*Best model determined by likelihood ratio tests and Akaike information criteria.

†Groups with significantly different optima. Dashes indicate that a single optimum is favored.

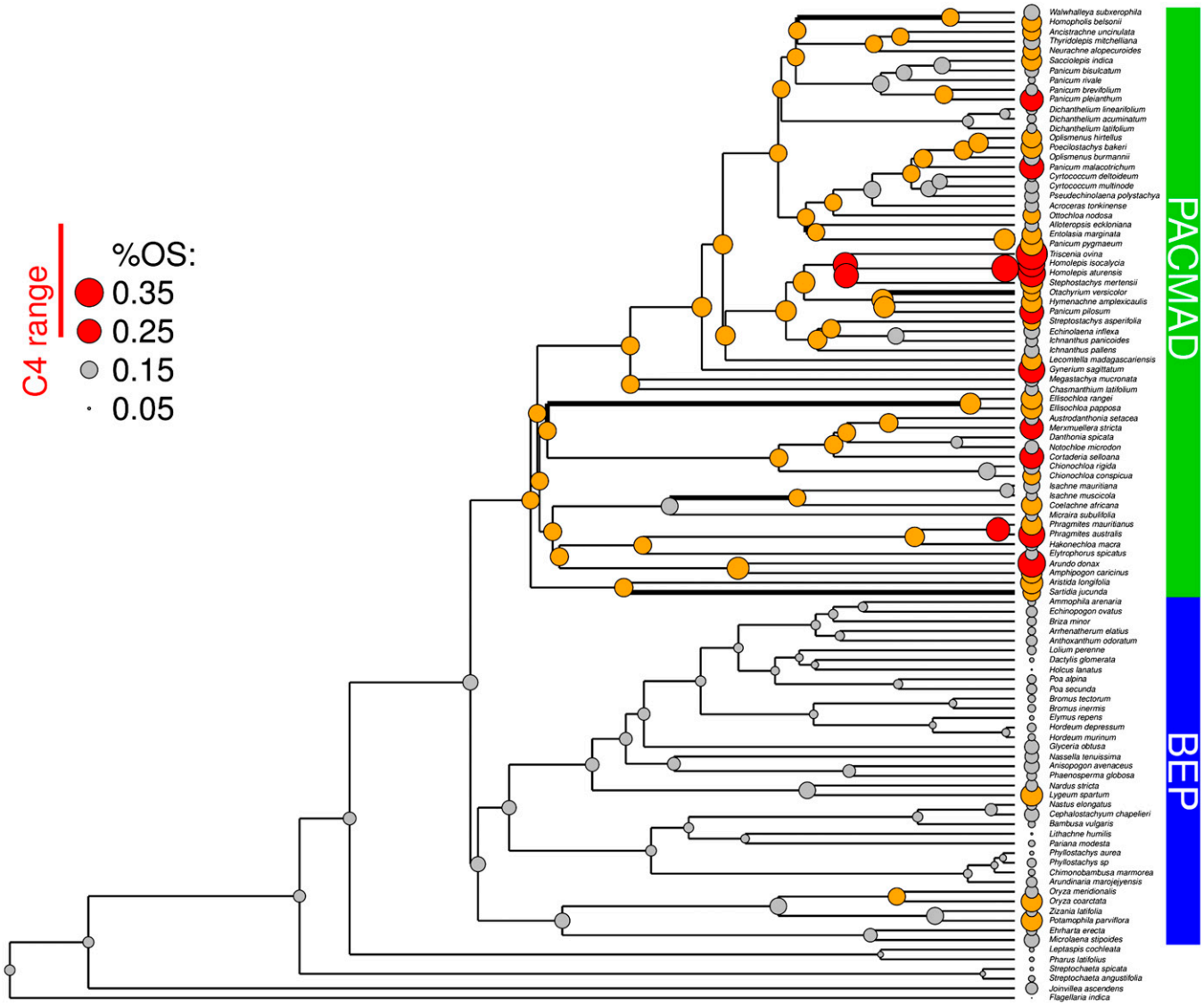

Fig. 3. Evolution of %OS in grasses. The measured and inferred values of %OS are mapped on a calibrated phylogeny of the C_3 grasses included in this study. The diameter of the dots is proportional to %OS values. The values in the C_4 range are indicated in red and those outside the C_4 range but above the threshold that promotes C_4-OS evolution are in orange. Branches where the C_3 to C_4-OS transitions occurred are thicker.

df = 1; P = 0.021), with an OS width above 15.7 leading to a transition rate of 0.005. Finally, BSD also explained the evolvability of C_4-OS (χ^2 = 6.4; df = 1; P = 0.012), with a BSD below 196 leading to a rate of transition of 0.006. Similarly, the evolvability of C_4-IS was explained by %IS, IS width, and BSD (χ^2 = 5.1, χ^2 = 5.9, and χ^2 = 4.1, respectively; df = 1; P = 0.022, P = 0.015, and P = 0.042, respectively), with a %IS above 0.003, an IS width above 6.2, and a BSD below 233 leading to a rate of transition of 0.005.

Discussion

Anatomical Factors Increase C_4 Evolvability in PACMAD. Variation in the evolvability of the C_4-OS type among clades of Poaceae and affiliated taxa is well explained by differences in %OS. Higher %OS in C_3 plants apparently facilitates the evolution of C_4 photosynthesis by reducing the distance between the C_3 and C_4 phenotypes but also potentially by giving access to transitional stages with a selective advantage over the C_3 condition. The exact evolutionary path between C_3 photosynthesis and the C_4-associated biochemistry is not resolved with confidence, but the photorespiratory pump based on glycine decarboxylase is often seen as an intermediate stage between C_3 and C_4 photosynthesis (3, 8, 18). Abundant OS tissue is instrumental in allowing its evolution because it implies that a decrease of glycine decarboxylase in M cells will not be detrimental, as enough of the enzyme will still be expressed in OS cells (3). Once a photorespiratory pump is fixed by natural selection, the biochemical components of the C_4 pathway can be gradually acquired and optimized until a highly efficient C_4 trait emerges. All enzymes of the C_4 pathway already exist in C_3 plants, and C_4 compatible promoter sequences are widespread in angiosperms (6, 7). The evolution of the C_4 trait can consequently require relatively few mutations in plants with a suitable leaf anatomy, which, in grasses, largely includes members of the PACMAD clade (Fig. 3).

The %OS is itself a complex trait, which results from variation in the number and size of cells of different tissues. The main determinants of %OS are BSD and OS width, and both partially explain the evolvability of C_4-OS. A small BSD evolved at the base of the BEP-PACMAD clade, and despite important variation in both clades, it still characterizes most extant members of

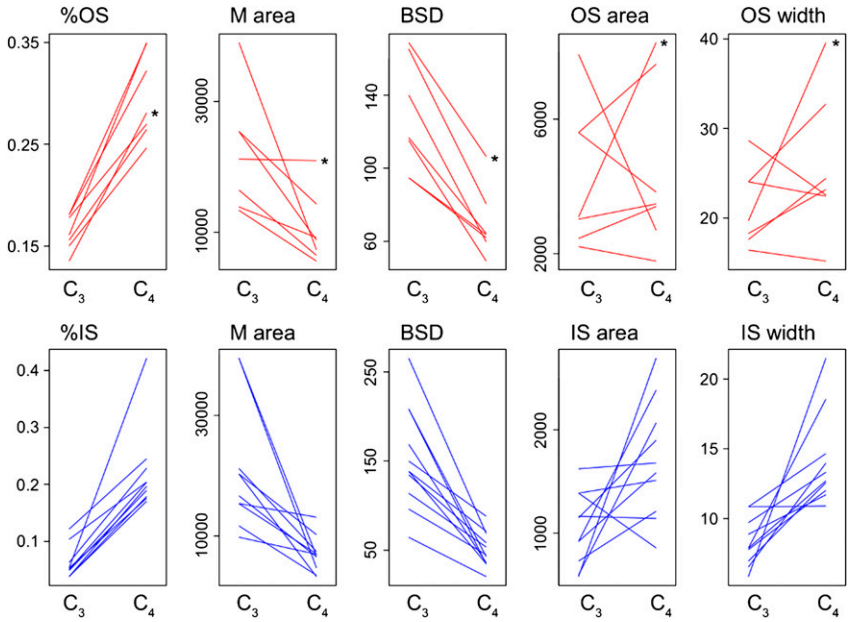

Fig. 4. Changes of anatomical traits inferred during C_4 evolution. The inferred change in five important traits is indicated for all C_3 to C_4-OS and all C_3 to C_4-IS (in blue) transitions. Each line connects the values of the most recent C_3 ancestor (on the left) and the first C_4 descendant (on the right). The asterisks indicate changes that led to the *Alloteropsis cimicina* lineage.

these groups (Fig. S6). This decrease of BSD occurred through a reduction of the number of M cells between OSs without a significant change in the size of these cells (Fig. S7). A small BSD is, however, not sufficient to confer large %OS, which also requires large OS cells. Enlarged OS cells are inferred for the deepest nodes in our phylogeny, which suggests that the ancestor of grasses possessed OS cells that were compatible with C_4 anatomy (Fig. S3). A reduction of OS width occurred in most of the lineages of the BEP clade, with even the partial degradation of OS in several Pooideae. The independent changes in BSD and OS width led to the restriction of large %OS to PACMAD, contributing to the recurrent C_4-OS origins in this clade millions of years later.

Transitions to C_4 Anatomy Involved a Further Reduction of BSD. Compared with C_3 plants, C_4-OS species are characterized by higher %OS, which can be achieved through a decrease of M area or an increase of OS area. A decrease of M area must involve a reduction in the number and/or size of M cells, whereas an increase of OS area must occur through an enlargement of OS cells. In plants, the majority of changes in organ size are attributable to mutations in the control of cell proliferation (23), which suggests that mutations affecting the number of cells are more frequent than those affecting their size. Congruently, in all C_4-OS, a smaller BSD was achieved by a reduction in the average distance between veins present in the C_3 ancestor through a reduction of the number of M cells between veins (Fig. S7). This indicates that the capacity to rapidly reduce BSD is instrumental in the evolution of a C_4-suitable foliar anatomy. The relatively high lability of this trait guarantees a recurrent emergence of lower BSD and a fertile ground for natural selection. In contrast, the width of OS cell was not consistently affected during the transition from C_3 to C_4-OS, and the selective optima for this trait are not different in C_3 and C_4-OS taxa (Table 1). Cell-size variation among plant species is generally a function of genome size (24) and an increased size of a given type of cell is often achieved through the postmitotic duplication of the plant's genome (25). Mutations drastically affecting OS width are, thus, probably infrequent, which limits the capacity of natural selection to fix larger OS cells.

Case of C_4-IS. The evolvability of the C_4-IS type is increased in groups with enlarged IS cells and shorter BSD. Compared with C_4-OS, the C_4-IS phenotype is more distant from the C_3 ancestral condition, and C_4-like values of %IS are extremely rare in both BEP and PACMAD. The only C_4-IS lineage outside Panicoideae (*Aristida* spp.) evolved from a group with a C_4-like %IS (Fig S4), and the independent C_4-IS lineages of Panicoideae all reduced their BSD through the addition of minor veins between the veins present in the C_3 ancestor (Fig. S7) (21), mirroring the evolution of C_4 anatomy in the *Flaveria* (Asteraceae) (26). These minor veins are composed almost exclusively of BS tissue and some species even evolved BS-like cells that punctuate the M without being associated with any vein (21, 27, 28). Their addition resulted in a very dramatic increase of %IS and likely strongly contributed to the recurrent C_4 origins in Panicoideae. However, these additional veins per se did not generate the high %IS values observed in extant taxa, which were possible only through a further increase of IS width.

C_4-IS evolution would have involved fewer changes if it happened from a C_4-OS instead of a C_3 ancestor. In some groups, the transition seems to have occurred directly from C_3 to C_4-IS [e.g., Neurachninae (29)], but a C_3 to C_4-OS transition followed by a change to C_4-IS must have occurred at least once in the MCP clade, which contains a C_4-IS group (Cenchrinae) nested within an otherwise C_4-OS clade (Melinidinae and Panicinae; Fig. 1) and possibly also in *Alloteropsis* (11). In addition, the C_3 *Homolepis* spp. have enlarged OS, filled with centripetal chloroplasts (Fig. S8). Whether these are involved in some kind of photorespiratory pump is unknown, but *Homolepis* is closely related to several C_4-IS taxa, some of which still produce starch in OS cells (9, 21), which might suggest that this C_3 group with a C_4-OS like anatomy is similar to the ancestor that produced C_4-IS descendants. The transition from C_3 to C_4-OS results in an increase of %IS because the decrease of BSD during C_4-OS evolution reduces the M area per vein but also because the enlargement of OS cells is associated with increased IS cell size

(phylogenetic linear model; $R^2 = 0.30$), possibly because the two traits share some genetic determinism. Together with the capacity of Panicoideae to produce additional veins, a passage through C_4-OS might consequently have strongly facilitated the evolution of the C_4-IS type.

Conclusions

Identifying the determinants of relative evolvability among clades is key to understanding the drivers of biological diversification. Most flowering plants possess metabolic modules suitable for C_4 photosynthesis, and yet C_4 origins are tightly clustered only in a handful of lineages. Here, we provide comparative statistical evidence that C_4 evolvability is determined, in large part, by leaf anatomy. Both OS and IS cell sizes determine the probability of transition to C_4-OS and C_4-IS, respectively, with an additional effect of BSD. Previous work has suggested that C_4-like anatomy evolved in PACMAD through adaptation to specific environments (20), but our comparative analyses, instead, emphasize the importance of stochastic historical processes. A reduction of BSD happened at the base of the BEP-PACMAD clade, leading to the coexistence of shorter BSD and large IS and OS cells in the common ancestor of the BEP and PACMAD clades. The foliar anatomy of this common ancestor might, thus, have been compatible with C_4 photosynthesis, but the atmospheric CO_2 levels at this time were high, limiting the advantage of C_4 photosynthesis. Subsequently, OS and IS cells became smaller in most BEP, possibly through selection, and this apparently suppressed the capacity of these plants to later evolve C_4 photosynthesis. When atmospheric CO_2 decreased tens of millions of years after the split of the BEP and PACMAD clades (2, 30–32), a combination of shorter BSD and large OS/IS cells existed only in members of the PACMAD clade, limiting C_4 evolution to this lineage. Changes in the number and size of different cells during the early history of grasses might have had limited importance at the time but allowed photosynthetic diversification when environmental changes opened new ecological niches during the Oligocene, highlighting the importance of historical contingency in evolutionary and ecological adaptation.

Materials and Methods

Species were sampled to maximize the number of independent C_4 origins and their closest C_3 relatives, while maintaining a balanced phylogenetic sampling as established in previous work (9). The species were incorporated in a previously published dataset based on three plastid markers (9), and the phylogenetic relationships between the species sampled for anatomy were extracted. This phylogenetic tree was used for comparative analyses. It was also used to test for the existence of C_4 precursors, following a recently developed approach (33), which compares models that assume the existence of an unidentified precursor (precursor_1 and precursor_2 models) to several binary models. Transitions between C_3 and C_4 photosynthesis are not allowed and can only occur between C_3 photosynthesis and a precursor state and then between the precursor state and C_4 photosynthesis (33).

Cross-sections of the middle part of mature leaves were photographed and used to measure several areas and lengths that were averaged over multiple veins (Fig. S1). C_4 species were classified into those using the OS tissue (C_4-OS) and those using the IS (C_4-IS) for carbon reduction, based on the localization of starch production. The relative proportions of OS and IS tissues were calculated as %OS = OS area/(OS area + M area) and %IS = IS area/(IS area + M area), respectively. The best evolutionary model was determined for each anatomical variable, and the ancestral values of the anatomical traits were reconstructed on the phylogenetic tree. The effect of anatomical traits on C_4-OS and C_4-IS evolvability was tested with models that allow transitions only above (or below) a given value of the anatomical trait. These models were compared with the null models (transition rates independent from anatomical traits) through likelihood ratio tests. The methods are detailed in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Rosa Cerros, Russell Hall, Surrey Jacobs, Simon Malcomber, and Steve Renvoize for their help in acquiring useful samples. This work was funded by Marie Curie International Outgoing Fellowship 252568 (to P.-A.C.), National Science Foundation (NSF) Division of Environmental Biology Grant 0920147 (to J.T.C.), Agence Nationale de la Recherche Grant 10LABX-41 (to G.B.), and NSF Division of Integrative Organismal Systems Grant 0843231 (to E.J.E.).

1. Zachos J, Pagani M, Sloan L, Thomas E, Billups K (2001) Trends, rhythms, and aberrations in global climate 65 Ma to present. *Science* 292(5517):686–693.
2. Beerling DJ, Royer DL (2011) Convergent Cenozoic CO_2 history. *Nat Geosci* 4(7): 418–420.
3. Sage RF, Sage TL, Kocacinar F (2012) Photorespiration and the evolution of C_4 photosynthesis. *Annu Rev Plant Biol* 63:19–47.
4. Christin PA, Osborne CP, Sage RF, Arakaki M, Edwards EJ (2011) C_4 eudicots are not younger than C_4 monocots. *J Exp Bot* 62(9):3171–3181.
5. Sage RF, Christin PA, Edwards EJ (2011) The C_4 plant lineages of planet Earth. *J Exp Bot* 62(9):3155–3169.
6. Hibberd JM, Quick WP (2002) Characteristics of C_4 photosynthesis in stems and petioles of C_3 flowering plants. *Nature* 415(6870):451–454.
7. Brown NJ, et al. (2011) Independent and parallel recruitment of preexisting mechanisms underlying C_4 photosynthesis. *Science* 331(6023):1436–1439.
8. Hylton CM, Rawsthorne S, Smith AM, Jones DA (1988) Glycine decarboxylase is confined to the bundle-sheath cells of leaves of C_3-C_4 intermediate species. *Planta* 175(4): 452–459.
9. Grass Phylogeny Working Group II (2012) New grass phylogeny resolves deep evolutionary relationships and discovers C_4 origins. *New Phytol* 193(2):304–312.
10. Sage RF (2001) Environmental and evolutionary preconditions for the origin and diversification of the C_4 photosynthetic syndrome. *Plant Biol* 3(3):202–213.
11. Christin PA, Freckleton RP, Osborne CP (2010) Can phylogenetics identify C_4 origins and reversals? *Trends Ecol Evol* 25(7):403–409.
12. Hattersley PW (1984) Characterization of C_4 type leaf anatomy in grasses (Poaceae). Mesophyll: Bundle sheath area ratios. *Ann Bot (Lond)* 53(2):163–179.
13. Dengler NG, Dengler RE, Donnelly PM, Hattersley PW (1994) Quantitative leaf anatomy of C_3 and C_4 grasses (Poaceae): Bundle sheath and mesophyll surface area relationships. *Ann Bot (Lond)* 73(3):241–255.
14. Dengler NG, Nelson T (1999) *C_4 Plant Biology*, eds Sage RF, Monson RK (Academic, San Diego), pp 133–172.
15. Muhaidat R, Sage RF, Dengler NG (2007) Diversity of Kranz anatomy and biochemistry in C_4 eudicots. *Am J Bot* 94(3):362–381.
16. Marshall DM, et al. (2007) *Cleome*, a genus closely related to *Arabidopsis*, contains species spanning a developmental progression from C_3 to C_4 photosynthesis. *Plant J* 51(5):886–896.
17. McKown AD, Dengler NG (2007) Key innovations in the evolution of Kranz anatomy and C_4 vein pattern in *Flaveria* (Asteraceae). *Am J Bot* 94(3):382–399.
18. Christin PA, et al. (2011) Complex evolutionary transitions and the significance of C_3-C_4 intermediate forms of photosynthesis in Molluginaceae. *Evolution* 65(3):643–660.
19. Muhaidat R, Sage TL, Frohlich MW, Dengler NG, Sage RF (2011) Characterization of C_3—C_4 intermediate species in the genus *Heliotropium* L. (Boraginaceae): Anatomy, ultrastructure and enzyme activity. *Plant Cell Environ* 34(10):1723–1736.
20. Griffiths H, Weller G, Toy LF, Dennis RJ (2012) You're so vein: Bundle sheath physiology, phylogeny and evolution in C_3 and C_4 plants. *Plant Cell Environ*, 10.1111/j.1365-3040.2012.02585.x.
21. Renvoize SA (1987) A survey of leaf-blade anatomy in grasses XI. *Paniceae*. *Kew Bull* 42(3):739–768.
22. Giussani LM, Cota-Sánchez JH, Zuloaga FO, Kellogg EA (2001) A molecular phylogeny of the grass subfamily Panicoideae (Poaceae) shows multiple origins of C_4 photosynthesis. *Am J Bot* 88(11):1993–2012.
23. Powell AE, Lenhard M (2012) Control of organ size in plants. *Curr Biol* 22(9): R360–R367.
24. Símová I, Herben T (2012) Geometrical constraints in the scaling relationships between genome size, cell size and cell cycle length in herbaceous plants. *Proc Biol Sci* 279(1730):867–875.
25. Sugimoto-Shirasu K, Roberts K (2003) "Big it up": Endoreduplication and cell-size control in plants. *Curr Opin Plant Biol* 6(6):544–553.
26. McKown AD, Dengler NG (2009) Shifts in leaf vein density through accelerated vein formation in C_4 *Flaveria* (Asteraceae). *Ann Bot (Lond)* 104(6):1085–1098.
27. Renvoize SA (1982) A survey of leaf-blade anatomy in grasses III. *Garnotieae*. *Kew Bull* 37(3):497–500.
28. Dengler NG, Donnelly PM, Dengler RE (1996) Differentiation of bundle sheath, mesophyll, and distinctive cells in the C_4 grass *Arundinella hirta* (Poaceae). *Am J Bot* 83(11):1391–1405.
29. Christin PA, et al. (2012) Multiple photosynthetic transitions, polyploidy, and lateral gene transfer in the grass subtribe Neurachninae. *J Exp Bot* 63(17):6297–6308.
30. Pagani M, Zachos JC, Freeman KH, Tipple B, Bohaty S (2005) Marked decline in atmospheric carbon dioxide concentrations during the Paleogene. *Science* 309(5734): 600–603.
31. Christin PA, et al. (2008) Oligocene CO_2 decline promoted C_4 photosynthesis in grasses. *Curr Biol* 18(1):37–43.
32. 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(12):2963–2977.
33. Marazzi B, et al. (2012) Locating evolutionary precursors on a phylogenetic tree. *Evolution* 66(12):3918–3930.

Supporting Information

Christin et al. 10.1073/pnas.1216777110

SI Materials and Methods

Phylogenetic Inference. A previously published 545-taxa dataset of the grasses based on the plastid markers *rbcL*, *ndhF*, and *trnK-matK* (1) was expanded and used for phylogenetic inference. For species sampled for anatomical cross-sections but not included in the published dataset, the markers *ndhF* and/or *trnK-matK* were either retrieved from GenBank when available or were newly sequenced from extracted genomic DNA with the method and primers described previously (1, 2). These new sequences were aligned to the dataset, excluding the regions that were too variable as described previously (1). The final dataset totaled 604 taxa and was used for phylogenetic inference as implemented in the software Bayesian Evolutionary Analysis by Sampling Trees (BEAST) (3).

The phylogenetic tree was inferred under a general time-reversible substitution model with a gamma-shape parameter and a proportion of invariants (GTR+G+I). *Flagellaria indica* (Flagellariaceae) was used to root the tree, following previous results (4). A relaxed molecular clock was implemented with rates that followed a log-normal distribution. No time constraint was used except for the crown of the BEP-PACMAD clade, which was modeled by a normal distribution with a mean of 51.2 and a SD of 6. The speciation pattern was set to follow a Yule process. The starting values and operators were optimized through repeated analyses. Two final analyses were run each for 20,000,000 generations. A tree was sampled every 5,000 generations after a burn-in period of 5,000,000, which was enough for the analysis to converge, as verified with the software TRACER (5). A consensus tree was then computed, and the tree was pruned to include only the species sampled for cross-sections using the package analyses of phylogenetics and evolution APE (6) in R. The anatomical measurements of *Spartina alterniflora* were assigned to its allopolyploid species *Spartina anglica* in the phylogeny, and those of *Stipagrostis obtusa* were assigned to one of the representative of this monophyletic genus. Two species, for which cross-sections have been measured, were not included in the phylogenetic reconstruction because of a lack of good-quality DNA. These species were used for the description of variables but were excluded from phylogeny-based analyses.

The phylogenetic tree was used to test for the existence of precursors, which might have evolved in some clades and allowed the evolution of C_4 photosynthesis in its descendants. This recently developed approach takes a two-state character (C_3 vs. C_4) and models it as a three-state character (C_3, precursor, C_4), where the transition between C_3 and C_4 is done in two steps, from C_3 to precursor and then from precursor to C_4 (7). In the most complex model currently implemented in r8s (precursor_2), the rate is different for C_3/precursor and for precursor/C_4 transitions, but each transition has an equal rate of reversal. This model was compared with the precursor_1 model (identical rate for C_3/precursor and precursor/C_4 transitions) and the two-state model (different rates for C_3-to-C_4 and C_4-to-C_3 transitions) using Akaike information criteria.

Plant Material and Cross-Sections. Mature leaves were either sampled from plants grown in the greenhouse, collected in the wild, or from dried herbarium specimens. When possible, the second or third leaf below the inflorescence was used. Our sampling was supplemented with 30 slides made available from previous studies (8–10). Leaf blade segments from dried herbarium material were rehydrated in water and 10% (wt/wt) aerosol OT for 48 h. Most samples were stored in 100% (vol/vol) ethanol for at least 2 wk. Segments from the center of leaf blades were then embedded in resin (JB-4; Polysciences), following the manufacturer's instructions. Five-micrometer thick cross-sections of the embedded leaf fragments were cut with a microtome and stained with saturated cresyl violet acetate (CVA). Some samples were fixed in formalin-propionic acid-alcohol (FPA), embedded in paraffin, sectioned at 10 μm, and stained with a safranin O-orange G series (11) as described in (12). All slides were made permanent and are available on request.

Anatomical Measurements. All C_3 grasses possess a double BS, with the outer layer derived from ground meristem to form a "parenchyma sheath," and the internal layer derived from the vascular procambium to form a "mestome sheath" (13). Many C_4 grasses also possess these two BS layers, with one of them specialized in carbon reduction and usually referred to as "carbon reduction" or "Kranz" tissue (14). However, this Kranz tissue can correspond either to the OS or the IS of C_3 grasses (13, 15). In addition, several C_4 grasses possess only one sheath, which in all investigated species derives from procambium, and is, therefore, homologous to the IS (= "mestome") of C_3 grasses (15). The terms "mestome" and "Kranz" are, therefore, functionally oriented and are evolutionarily ambiguous. The two sheaths are consequently referred to as IS and OS throughout this manuscript. The definition of all anatomical characters measured in this study aimed at allowing comparison (through homology) among all of the considered taxa.

Photographs of the cross-sections were taken at a magnification of 20× using a Nikon Eclipse E600 compound light microscope (Nikon Instruments) linked to a Nikon DXM1200C digital camera. An area extending from the middle of a vascular bundle to the middle of a distant bundle was measured with ImageJ software (16). When possible, the median vascular bundle was not considered and the measured areas encompassed lateral bundles of different orders. On the selected section, the areas of M, OS, IS, and vascular tissue (total area inside of the IS) were measured together with other characters (Fig. S1). The M was defined as all of the tissue between the epidermis and BS. Structural cells without photosynthetic potential, such as fibers and bulliform cells, were excluded. The OS was defined as the single layer of concentrically arranged cells directly outside the IS. If additional OS cells were present (e.g., BS extension), they were excluded (from both M and OS areas) for comparative purposes. Only cells that were visually differentiated from the surrounding M were considered, and incomplete OS were scored for several taxa. The IS was defined as the single layer of cells that surrounds the vascular tissue (xylem, phloem, and fibers together). In most cases, the IS was complete (i.e., formed a complete ring), but a few exceptions existed, resulting in the measurement of partial IS. When only one BS was present (C_4 taxa only), this was considered as the IS, following previous observations in three independently evolved C_4 taxa (15). It is, however, possible that some C_4 lineages that have not been studied developmentally have a single BS that corresponds to the OS. The measured cross-sectional areas were used to calculate, for each leaf, the proportion of OS compared with M [%OS = OS area/(OS area + M area)] and the proportion of IS compared with M [%IS = IS area/(IS area + M area)]. The center of starch production was identified by the accumulation of chloroplasts, as observed on the cross-sections and described in previous works (8–10), and species were consequently classified as C_4 using the OS (C_4-OS) or C_4 using the IS (C_4-IS) as appropriate. Two species in our sampling are known to use C_2 photosynthesis [also known as C_3-C_4 photosynthesis (17)]

based on the OS (in the genus *Steinchisma*). Because C_2 photosynthesis also represents a derived state over the ancestral C_3 trait, these two species were classified as C_4-OS.

The BSD, defined as the mean distance of M from the outer wall of one BS to the outer wall of an adjacent BS, was also measured on the same sections. In addition, the IVD was measured as the mean distance between centers of veins through the M. In addition, the number of M cells between two adjacent bundles and their longest diameters were measured for a minimum of 15 cells following the shortest path through M. The width of the OS and IS cells was measured for the two most-equatorial cells of each vein along the diameter of the corresponding vein. Leaf thickness was measured from upper epidermis to lower epidermis at the major lateral veins (veins accompanied by fibers and/or separated by bulliform cells). A second leaf thickness (thickness2) was measured from upper epidermis to lower epidermis at the thinnest point between two major veins. A ratio of thicknesses was computed as thickness2 divided by thickness, which estimates whether the leaf is flat (ratio = 1) or thinner between veins than at veins (ratio < 1). The average distance between major veins was also measured (IVDm). All measurements were averaged over the analyzed area and are given in micrometers or square micrometers.

Intraspecific Replicates. To evaluate the intraspecific variability in anatomical traits, measurements were repeated on individuals belonging to the same species but grown in different conditions and/or that originated from different geographic origins (Dataset S1). Replicates were obtained for 15 species, including one outgroup, two C_4, nine C_3 PACMAD, and three C_3 BEP. All measured traits varied within species, which can be attributable to phenotypic plasticity and intraspecific genetic variation. The impact of this variation on the patterns observed in this study was evaluated by linear regressions. For each species, the sample with the largest measured total area was defined as the reference and the sample with the smallest measured area was defined as the duplicate. The R^2 of the linear regression (forcing the intercept to 0) of the duplicates on the references for the different traits ranged from 0.74 (for thickness2) to 0.98 (for OS width), except for the mean vein area ($R^2 = 0.59$). These results indicate that despite intraspecific variation, the trends are consistent among different samples. Consequently, only reference samples were included in comparative analyses.

Comparative Analyses. We used different phylogeny based methods to model the evolution of anatomical characters while correcting for historical factors. All comparative analyses were performed in R using the packages APE (6), GEIGER (18), OUCH (19) and CAPER (20). All variables were log-transformed except for the three ratios (%OS, %IS, and ratio of thicknesses). To differentiate taxonomic and photosynthetic effects, we modeled the variables separately first for the three major clades but excluding C_4 taxa (outgroups, BEP, and PACMAD) and then for all photosynthetic types but in PACMAD only (C_3 PACMAD, C_4-OS, and C_4-IS). The best evolutionary model was selected through hierarchical likelihood ratio tests between a Brownian motion model, an OU model with a single optimum, an OU model with two different optima (best-fit of three possible groupings), and an OU model with three different optima (one per group).

Further analyses were performed on C_3 taxa only to avoid potential biases attributable to different selective regimes and optima in C_3 and C_4 taxa. Correlations between variables were investigated by phylogenetic generalized least squares (GLS) with a lambda correction. Ancestral character states were inferred for all nodes of the tree with a maximum likelihood criterion under a Brownian motion model. The ancestral character states were also estimated for all nodes of the tree that includes both C_3 and C_4 taxa, as well as for the tree that includes all C_3 taxa and all C_4 taxa that possess an OS. To reconstruct the changes that occurred during C_3 to C_4 transitions, the value of the most recent C_3 ancestor of each C_4 lineage was compared with the value at the crown group node of each C_4 lineage. The value of the most recent C_3 ancestor was extracted from the tree that contained C_3 taxa only to avoid values artificially close to C_4 values. Because the node that corresponds to the split of the C_4 lineage was not present in this tree, the average between the nodes that precede and succeed it was considered as representative of the C_3 ancestor. The value for the C_4 crown was extracted from the analysis based on the tree with both C_3 and C_4 taxa. (For %OS, OS area, and OS width, the tree with C_3 taxa and C_4 taxa that possess an OS was used.) For these analyses, the C_4-IS and C_4-OS lineages of *Alloteropsis* were treated as independent C_4 origins, although the most likely scenario involves the existence of unknown intermediate stages (21). The C_4 lineage Tristachyideae was excluded from these analyses because the C_4 type of its ancestor (C_4-IS or C_4-OS) is unknown.

C_3/C_4 transitions were modeled by taking into account a potential effect of anatomical traits. In the case of significant anatomical enablers, the rate of C_4 evolution would be very small below (or above) a given value of the anatomical character and increase once the character becomes larger (or smaller). This was modeled by collapsing to 10^{-5} all branches of the phylogeny for which the reconstructed value of the anatomical trait was below (or above) a given value. The transformed topology was used to model C_3/C_4 with a two-parameter model, and the threshold for the change of probabilities was optimized from the data under maximum likelihood as implemented in the software BayesTraits (22). The model that assumes an increase of transition rates after a given value of an anatomical character can be compared with the model without branch transformation through a likelihood ratio test with one degree of freedom. To circumvent the issue of strong differences between C_3 and C_4 values, the models were optimized on a phylogeny that contained only C_3 taxa and C_3 species were typed as "sister to C_4-OS" or "sister to C_4-IS," which was then used as a proxy for C_4 origins. Because C_3/C_4 transitions happened on branches separating C_4 lineages and their C_3 sister groups, asking whether a given anatomical trait increases the probability of becoming an actual C_4 group or sister to a C_4 group is equivalent. C_3 species not sister to C_4 lineages were given a score of 0, whereas the others were given a score of 1. The C_4-IS groups Andropogoneae and Anthephorinae are both sister to groups that encompass other C_4 lineages, as well as their respective C_3 sister groups (1). All C_3 species in these two groups were consequently given a score of 1 (sister to C_4-IS). The C_3 taxa sister to C_4-IS lineages nested within each of these groups were given a score of 2 to model additional transitions. The trait was modeled as a multistate character where transition rates q01 = q12, q21 = q10, q02 = q20 = 0, forcing the transitions to occur on branches connecting C_3 ancestors and C_4-IS descendants.

1. Grass Phylogeny Working Group II (2012) New grass phylogeny resolves deep evolutionary relationships and discovers C_4 origins. *New Phytol* 193(2):304–312.
2. Taylor SH, et al. (2012) Photosynthetic pathway and ecological adaptation explain stomatal trait diversity amongst grasses. *New Phytol* 193(2):387–396.
3. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
4. Givnish TJ, et al. (2010) Assembling the tree of the Monocotyledons: Plastome sequence phylogeny and evolution of Poales. *Ann Mo Bot Gard* 97(4):584–616.
5. Rambaut A, Drummond AJ (2007) Tracer v1.4. Available at http://beast.bio.ed.ac.uk/Tracer. Accessed June 17, 2010.
6. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20(2):289–290.

7. Marazzi B, et al. (2012) Locating evolutionary precursors on a phylogenetic tree. *Evolution* 66(12):3918–3930.
8. Renvoize SA (1985) A survey of leaf-blade anatomy in grasses V. The bamboo allies. *Kew Bull* 40(3):509–535.
9. Renvoize SA (1986) A survey of leaf-blade anatomy in grasses VIII. *Arundinoideae*. *Kew Bull* 41(2):323–338.
10. Renvoize SA (1987) A survey of leaf-blade anatomy in grasses XI. *Paniceae*. *Kew Bull* 42(3):739–768.
11. Sharman BC (1943) Tannic acid and iron alum with safranin and orange G in studies of the shoot apex. *Stain Technol* 18(3):105–111.
12. Columbus JT (1998) Morphology and leaf blade anatomy suggest a close relationship between *Bouteloua aristidoides* and *B*. (*Chondrosium*) *eriopoda* (Gramineae: Chloridoideae). *Syst Bot* 23(4):467–478.
13. Brown WV (1975) Variations in anatomy, associations, and origins of Kranz tissue. *Am J Bot* 62(4):395–402.
14. Dengler NG, Dengler RE, Donnelly PM, Hattersley PW (1994) Quantitative leaf anatomy of C_3 and C_4 grasses (Poaceae): Bundle sheath and mesophyll surface area relationships. *Ann Bot (Lond)* 73(3):241–255.
15. Dengler NG, Dengler RE, Hatterlsey PW (1985) Differing ontogenetic origins of PCR ("Kranz") sheaths in leaf blades of C_4 grasses (Poaceae). *Am J Bot* 72(2):284–302.
16. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.
17. Sage RF, Sage TL, Kocacinar F (2012) Photorespiration and the evolution of C_4 photosynthesis. *Annu Rev Plant Biol* 63:19–47.
18. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: Investigating evolutionary radiations. *Bioinformatics* 24(1):129–131.
19. Butler MA, King AA (2004) Phylogenetic comparative analysis: A modeling approach for adaptive evolution. *Am Nat* 164(6):683–695.
20. Orme CDL, et al. (2012) Caper: Comparative analyses of phylogenetics and evolution R. Package version 0.5. Available at http://cran.r-project.org/web/packages/caper/index.html. Accessed August 14, 2012.
21. Christin PA, et al. (2012) Adaptive evolution of C_4 photosynthesis through recurrent lateral gene transfer. *Curr Biol* 22(5):445–449.
22. Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53(5):673–684.

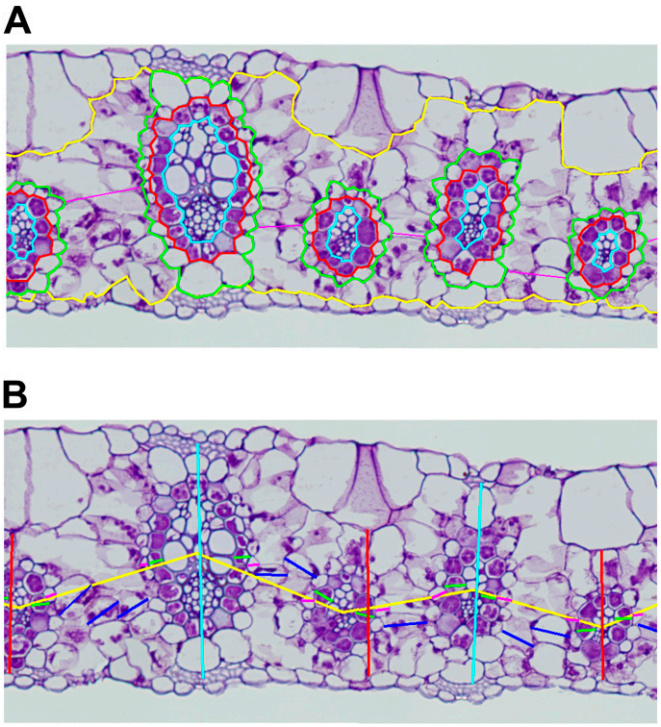

Fig. S1. Representation of the measured areas and lengths. The measured characters are shown on a leaf section of the C_4-IS *Alloteropsis semialata*. (*A*) BSD (purple lines) and main areas (yellow, M; green, OS; red, IS; light blue, veins). (*B*) Other lengths (yellow, IVD; light blue, thickness; red, thickness2; purple, OS size; green, IS size; blue, M size).

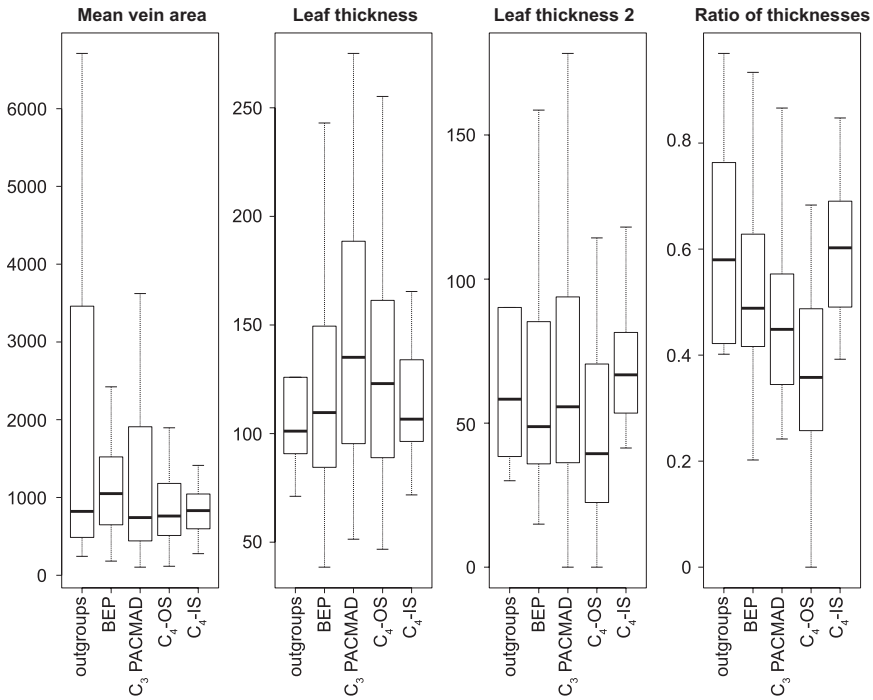

Fig. S2. Distribution of trait values among clades and photosynthetic types. Boxes indicate the 25th and 75th percentiles, and whiskers indicate the extreme values still within 1.5 interquartile range of the upper and lower quartiles, respectively.

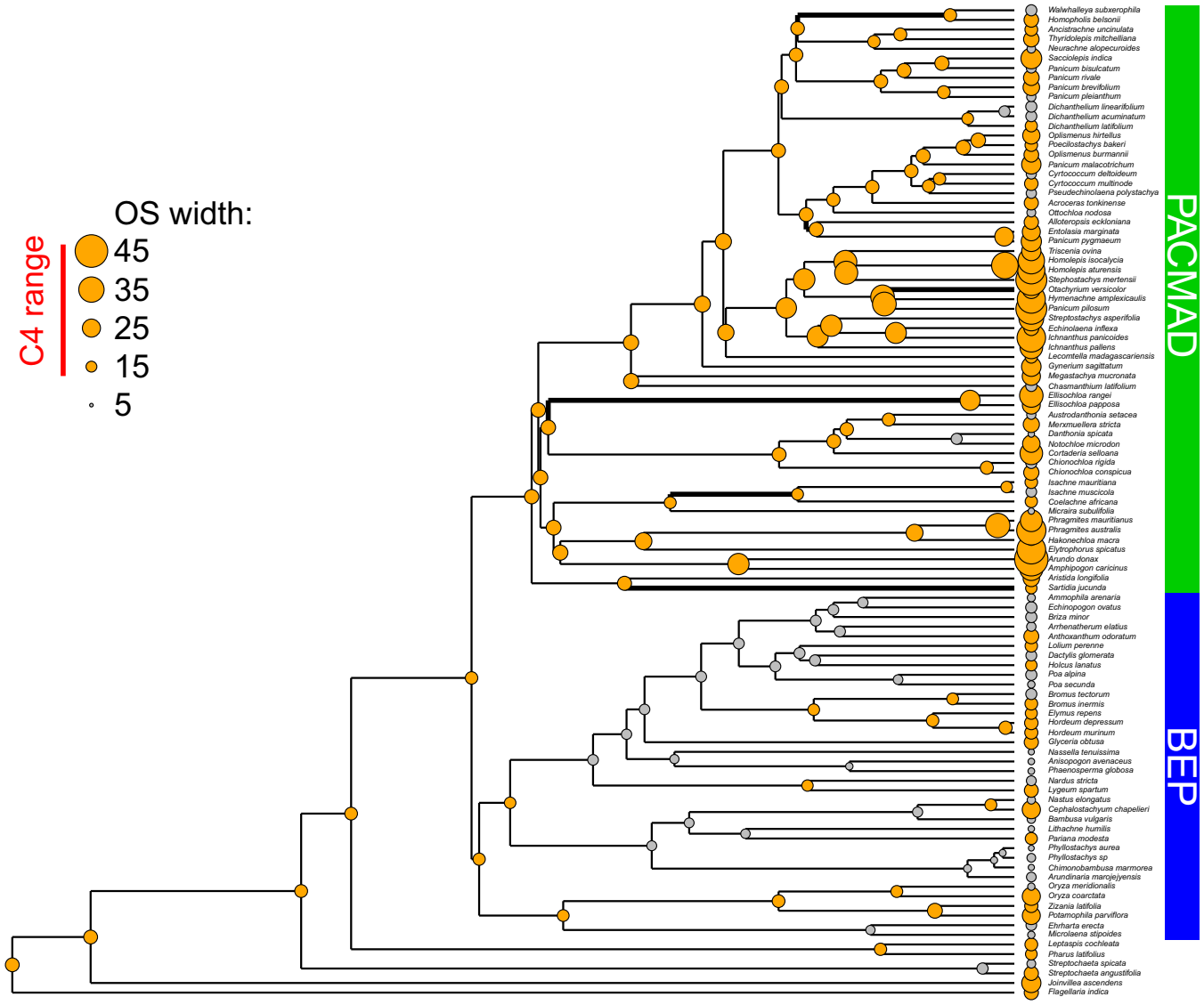

Fig. S3. Mapping of measured and reconstructed values of OS width on the phylogeny. The measured and inferred values of OS width are mapped on a calibrated phylogeny of the C_3 grasses included in this study. The diameter of the dots is proportional to OS width values. The values above the threshold that promotes C_4-OS evolution are in orange. Branches where the C_3 to C_4-OS transitions occurred are thicker.

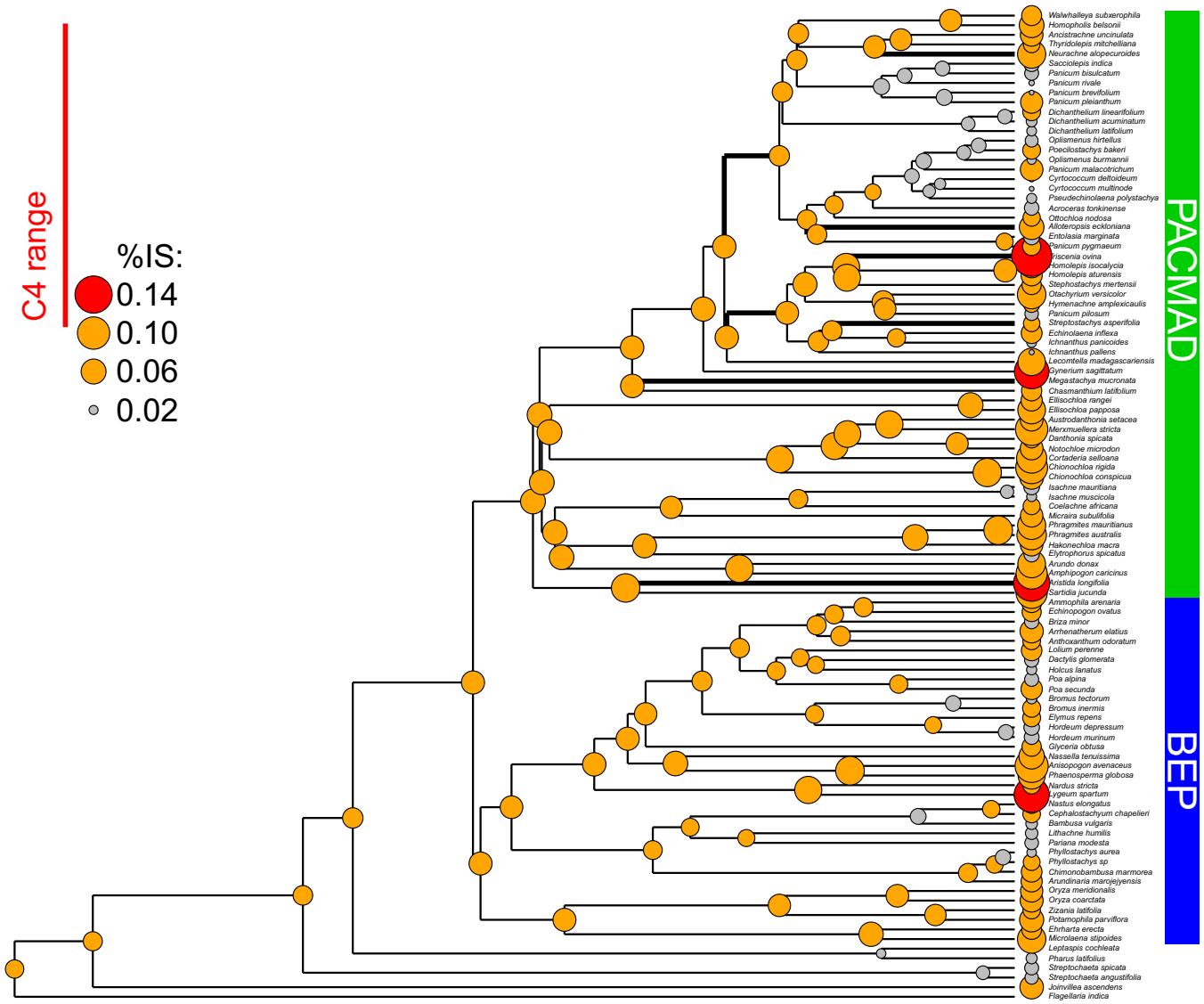

Fig. S4. Mapping of measured and reconstructed values of %IS on the phylogeny. The measured and inferred values of %IS are mapped on a calibrated phylogeny of the C_3 grasses included in this study. The diameter of the dots is proportional to %IS values. The values in the C_4 range are indicated in red, and those outside the C_4 range but above the threshold that promotes C_4-IS evolution are in orange. Branches where the C_3- to C_4-IS transitions occurred are thicker.

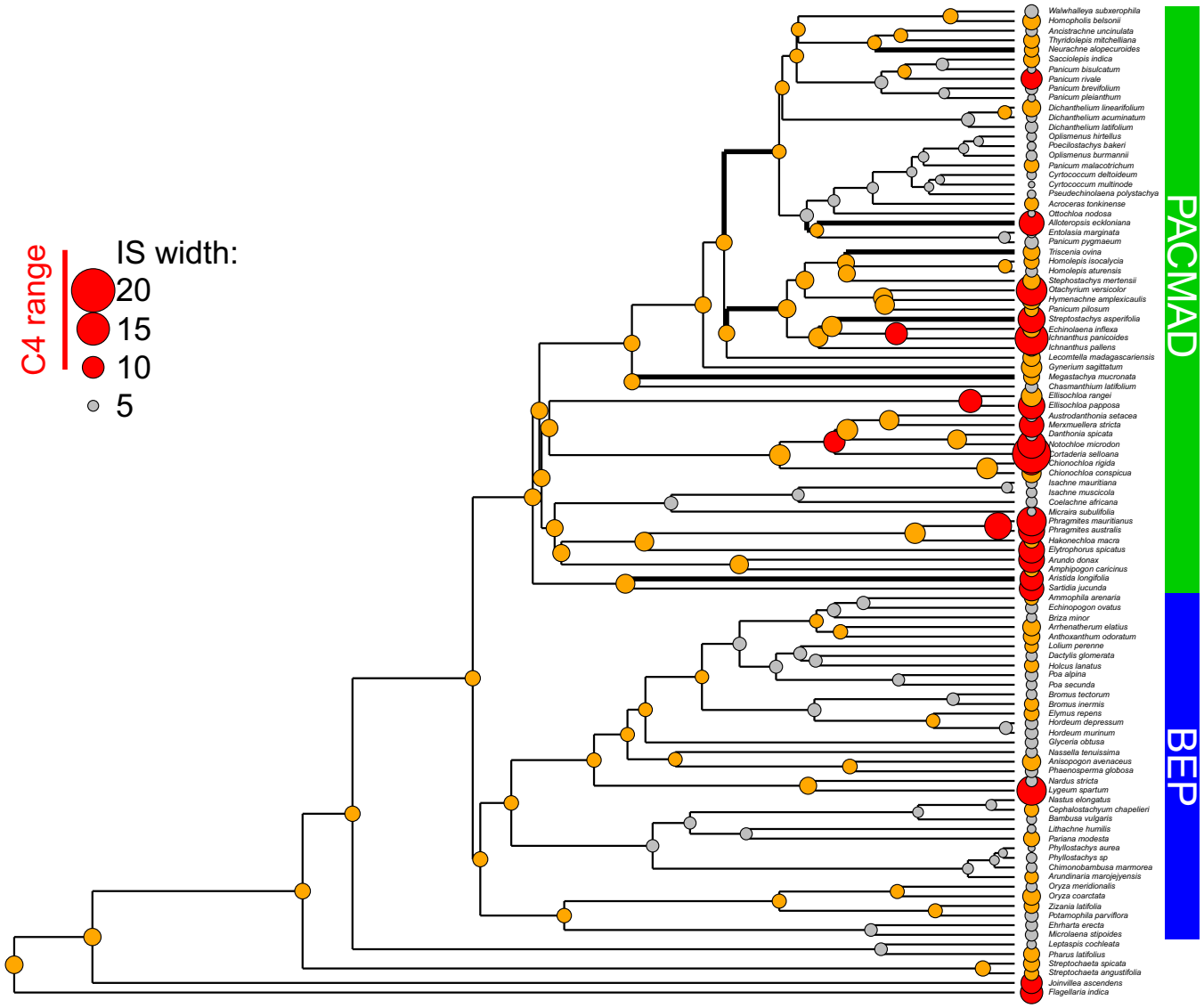

Fig. S5. Mapping of measured and reconstructed values of IS width on the phylogeny. The measured and inferred values of IS width are mapped on a calibrated phylogeny of the C_3 grasses included in this study. The diameter of the dots is proportional to IS width values. The values in the C_4 range are indicated in red, and those outside the C_4 range but above the threshold that promotes C_4-IS evolution are in orange. Branches where the C_3 to C_4-IS transitions occurred are thicker.

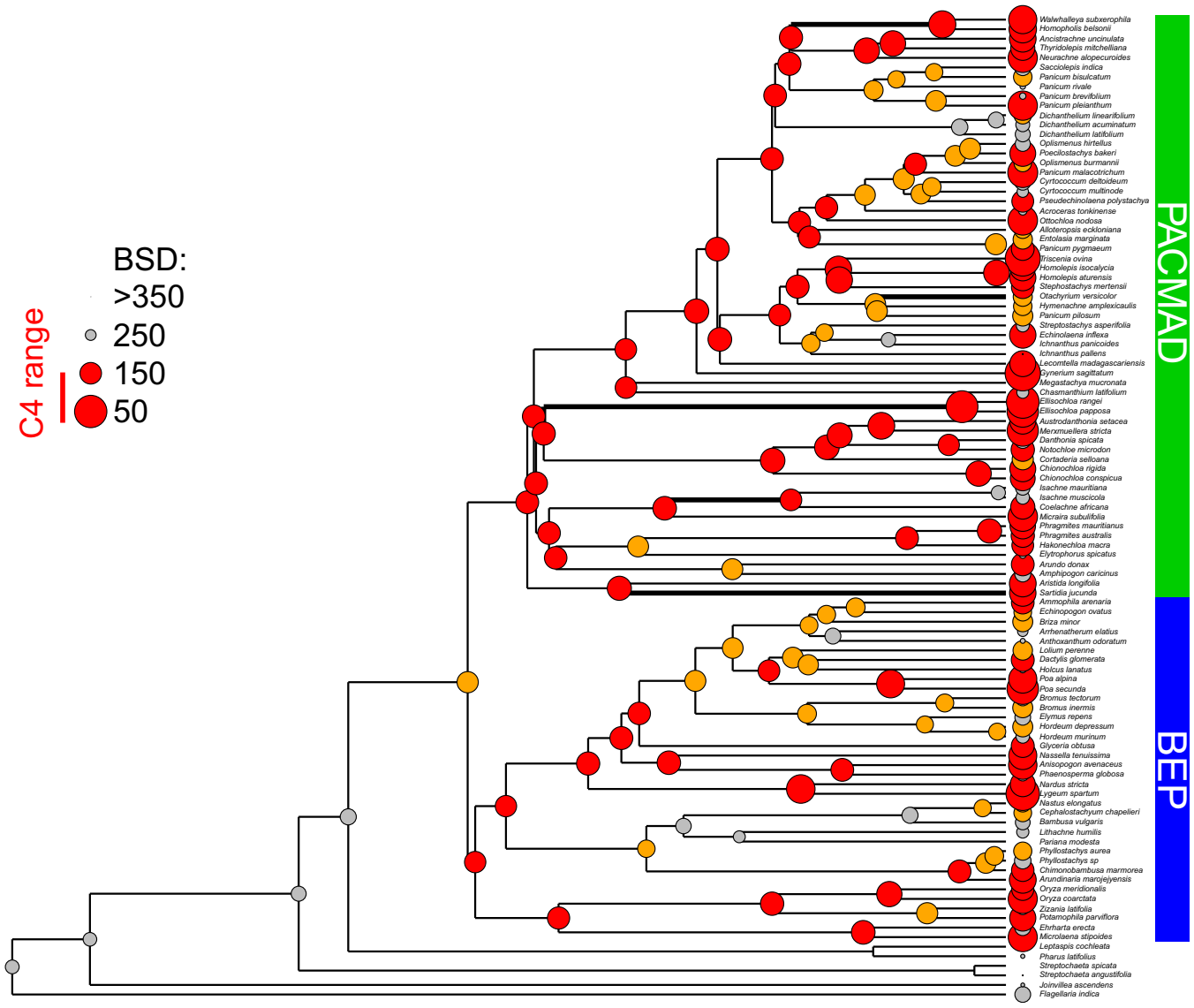

Fig. S6. Mapping of measured and reconstructed values of BSD on the phylogeny. The measured and inferred values of BSD are mapped on a calibrated phylogeny of the C_3 grasses included in this study. The diameter of the dots is inversely proportional to BSD values. The values in the C_4 range are indicated in red, and those outside the C_4 range but below the threshold that promotes C_4-OS evolution are in orange. Branches where the C_3- to C_4-OS transitions occurred are thicker.

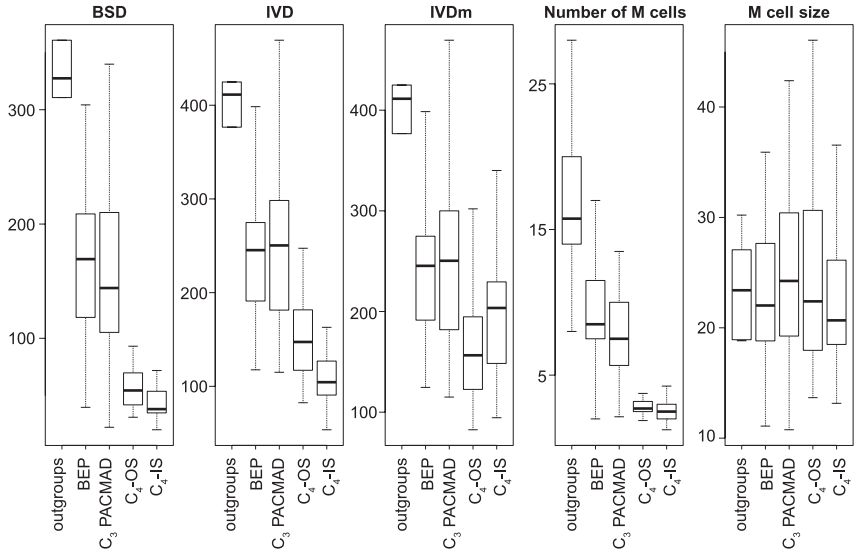

Fig. S7. Distribution of trait linked to BSD among clades and photosynthetic types. The distribution of five variables is summarized by boxplots. Boxes indicate the 25th and 75th percentiles and whiskers the extreme values still within 1.5 interquartile range of the upper and lower quartiles, respectively. BSD, inter-BS distance; IVDm, distance between major veins. Number of M cells is between two adjacent BSs.

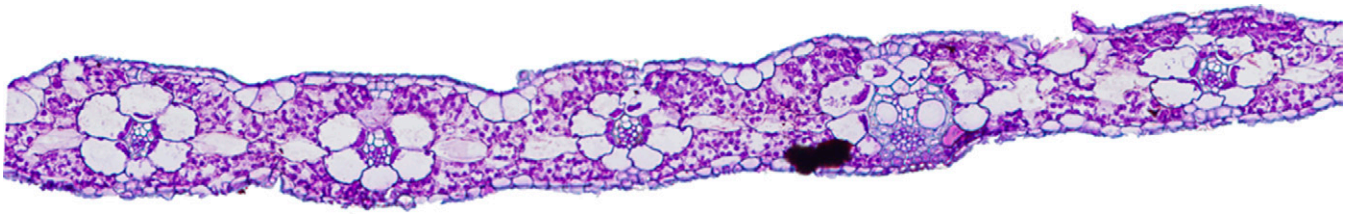

Fig. S8. Detail of a portion of *Homolepis* leaf. This picture shows several vascular bundles of the C_3 *Homolepis aturensis*. The centripetal position of chloroplasts in the outer BS is usually characteristic of C_4 species.

Dataset S1. Sample details and anatomical measurements

Dataset S1

For each species, the measured and extrapolated values are indicated. Phylogeny indicates whether the species was included in the phylogeny. The voucher accession is indicated with the herbarium where the voucher is deposited in parentheses. SisterC4.OS: if the species belongs to a clade sister to a C_4-OS group, its value is 1. SisterC4.IS: if the species belongs to a clade sister to a C_4-IS group, its value is 1 (2 for species sister to a C_4-IS group and nested in a clade sister to another C_4-IS group). Duplicate: the value is 0 if there was only one sample per species, 1 if the sample was considered as the reference, and 2 if it was considered as the replicate. area.bundle, total of OS+IS+vein; area.total, total measured area (M+OS+IS+vein); B&T, seeds acquired from B and T World Seeds (Aigues Vives, France); measured_width, total width (connecting the veins) of the measured area; nb.M, mean number of M cells between two adjacent bundles; nb.veins, number of veins in the section used for measurements; nb.veins.major, number of major veins in the section used for measurements; USDA, seeds acquired from the US Department of Agriculture, Agricultural Research Service (Griffin, GA).