Phylogenomics of C_4 Photosynthesis in Sedges (Cyperaceae): Multiple Appearances and Genetic Convergence

Guillaume Besnard,[1] A. Muthama Muasya,† Flavien Russier,* Eric H. Roalson,‡ Nicolas Salamin,*§ and Pascal-Antoine Christin**

*Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland; †Botany Department, University of Cape Town, Private Bag X3, Rondebosch, South Africa; ‡School of Biological Sciences and Center for Integrated Biotechnology, Washington State University, Pullman, USA; and §Swiss Institute of Bioinformatics, Genopode, Lausanne, Switzerland

C_4 photosynthesis is an adaptive trait conferring an advantage in warm and open habitats. It originated multiple times and is currently reported in 18 plant families. It has been recently shown that phospho*enol*pyruvate carboxylase (PEPC), a key enzyme of the C_4 pathway, evolved through numerous independent but convergent genetic changes in grasses (Poaceae). To compare the genetics of multiple C_4 origins on a broader scale, we reconstructed the evolutionary history of the C_4 pathway in sedges (Cyperaceae), the second most species-rich C_4 family. A sedge phylogeny based on two plastome genes (*rbcL* and *ndhF*) has previously identified six fully C_4 clades. Here, a relaxed molecular clock was used to calibrate this tree and showed that the first C_4 acquisition occurred in this family between 19.6 and 10.1 Ma. According to analyses of PEPC-encoding genes (*ppc*), at least five distinct C_4 origins are present in sedges. Two C_4 *Eleocharis* species, which were unrelated in the plastid phylogeny, acquired their C_4-specific PEPC genes from a single source, probably through reticulate evolution or a horizontal transfer event. Acquisitions of C_4 PEPC in sedges have been driven by positive selection on at least 16 codons (3.5% of the studied gene segment). These sites underwent parallel genetic changes across the five sedge C_4 origins. Five of these sites underwent identical changes also in grass and eudicot C_4 lineages, indicating that genetic convergence is most important within families but that identical genetic changes occurred even among distantly related taxa. These lines of evidence give new insights into the constraints that govern molecular evolution.

Introduction

C_4 photosynthesis is a complex adaptation over the classical C_3 pathway (von Caemmerer and Furbank 2003; Sage 2004). It consists of a set of morphological and biochemical modifications that together create a CO_2 pump, which concentrates CO_2 around Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). This suppresses photorespiration, a phenomenon resulting from the oxygenase activity of Rubisco and leading to CO_2 release and energy loss (Foyer et al. 2009). The C_4 pathway thereby confers an increased photosynthetic efficiency in all conditions promoting high rates of photorespiration, and especially in open and warm habitats (Ehleringer et al. 1997; Sage 2004). The leaf anatomy generally associated with the C_4 syndrome, called Kranz anatomy, is characterized by two rings of cells surrounding the vascular bundles, the mesophyll (M) and bundle sheath (BS) cells. Atmospheric CO_2 reaches M cells via the plant stomata and is converted into HCO_3^-. This molecule is then fixed on phospho*enol*pyruvate by phospho*enol*pyruvate carboxylase (PEPC) into oxaloacetate. This four-carbon acid is then transported to the BS cells via a biochemical path that varies among C_4 plants (Kanai and Edwards 1999). Therein, decarboxylating enzymes release CO_2 from the four-carbon compound, which can then enter the Calvin cycle as in C_3 plants through its fixation by Rubisco. Three biochemical C_4 subtypes are recognized based on the decarboxylating enzymes: nicotinamide adenine dinucleotide phosphate–malic enzyme (NADP–ME), nicotinamide adenine dinucleotide–malic enzyme (NAD–ME), and phospho*enol*pyruvate carboxykinase (PCK).

Despite its complexity, C_4 photosynthesis is one of the best examples of convergent evolution, having evolved more than 50 times in at least 18 plant families (Sage 2004; Conway Morris 2006). Only 3% of all angiosperms use the C_4 pathway, but they account for one-fifth of global primary production (Lloyd and Farquhar 1994; Ehleringer et al. 1997), which highlights the ecological importance of C_4 plants (Sage et al. 1999). In addition, this adaptive trait is used by major crops, such as maize, sorghum, and sugarcane, conferring it a large economical interest. These have led to extensive research on C_4 photosynthesis, especially in the grass family, which encompasses about 60% of all C_4 species (Sage 2004).

Using comparative phylogenetic analyses, recent investigations showed that the C_4 pathway evolved in grasses about 30 million years ago (Ma), probably as a response to past decline of atmospheric CO_2 concentration (Christin, Besnard, et al. 2008; Vicentini et al. 2008). Phylogenetic tools also showed that the recurrent C_4 appearances in grasses had been caused by important parallel adaptive changes on genes encoding the different enzymes involved in the C_4 pathway, such as PEPC (Christin et al. 2007) or PCK (Christin et al. 2009). In the C_4 cycle, PEPC is responsible for the initial fixation of atmospheric CO_2 into organic compounds. However, PEPC-encoding genes (*ppc* genes) form a ubiquitous multigene family not restricted to C_4 plants. These genes encode different enzymes responsible for various functions, mainly anaplerotic (Latzko and Kelly 1983). One of these isoforms evolved the C_4 function at least eight times independently in grasses. In this family, similar or identical changes were identified on 21 codons and occurred in up to eight of these C_4 PEPC origins (Christin et al. 2007). Such modifications probably resulted in adaptive convergence in kinetic properties of the C_4 PEPCs (Dong et al. 1998; Bläsing et al. 2000; Gowik et al. 2006; Rao et al.

[1] Present address: Imperial College London, Silwood Park Campus, Ascot, Berkshire, United Kingdom.

Key words: C_4 photosynthesis, convergence, Cyperaceae, molecular dating, phospho*enol*pyruvate carboxylase, positive selection.

E-mail: g.besnard@imperial.ac.uk.

Mol. Biol. Evol. 26(8):1909–1919. 2009
doi:10.1093/molbev/msp103
Advance Access publication May 21, 2009

© The Author 2009. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

2008). These first insights into the genetics of convergence shed light on the constraints shaping molecular evolution. These were however limited to one of the 18 plant families in which C_4 emerged and drawing general patterns about molecular evolution of convergence would benefit from studying more C_4 systems by enabling multi-scale comparisons (e.g., intra and interfamilies).

Besides grasses, eudicot C_4 systems have been the subject of numerous and fruitful investigations on C_4 molecular evolution. The genus *Flaveria* (Asteraceae) in particular has received considerable attention. It contains several C_3 and C_4 relatives, as well as photosynthetic intermediates representing a gradual increase from C_3 to C_4 photosynthesis (McKown et al. 2005). The *ppc* genes and their encoded enzymes have been particularly well characterized in this genus (Bläsing et al. 2000; Svensson et al. 2003; Westhoff and Gowik 2004; Akyildiz et al. 2007; Jacobs et al. 2008). A study was also performed on *ppc* genes of the genus *Alternanthera* (Amaranthaceae), which contains C_3, C_3–C_4, and C_4 species (Gowik et al. 2006). Genetic information for these C_4 eudicots has already allowed a first and rough interfamilial comparison of C_4 genetic mutations (Christin et al. 2007). Recently, *ppc* genes were also analyzed in a facultative C_4 Hydrocharitaceae, *Hydrilla verticillata* (Rao et al. 2008). Unfortunately, the limited number of C_4 species and independent lineages in these groups hampered efficient statistical testing of adaptation and other C_4-rich systems have to be investigated.

The sedge family (Cyperaceae) is the second most important C_4 family, with approximately 1,500 C_4 species (more than 20% of all C_4 plants; Sage 2004). However, C_4 sedges have been subject to limited investigations compared with other C_4 systems. The different C_4 species of Cyperaceae present anatomical and biochemical variations suggesting that their C_4 traits are not always homologous (Bruhl et al. 1987; Ueno and Koyama 1987; Ueno et al. 1989; Bruhl and Perry 1995; Soros and Dengler 2001). Four anatomical types have been recognized, which are generally accompanied by a given C_4 biochemical subtype (Soros and Dengler 2001). The rhynchosporoid (*Rhynchospora*), chlorocyperoid (Cypereae), and fimbristyloid (Abildgaardieae) anatomical types are associated with the NADP–ME subtype, whereas species presenting the eleocharoid (*Eleocharis*) anatomical type (and one anomalous fimbristyloid species, *Eleocharis vivipara*; Bruhl et al. 1987; Ueno et al. 1988) use the NAD–ME biochemical type. Phylogenetic investigations suggest at least four independent C_4 origins in Cyperaceae (Soros and Bruhl 2000; Bruhl and Wilson 2007), but further considerations are needed, especially for clades containing C_3 and C_4 species and whose phylogeny is not completely resolved (such as genera *Eleocharis* and *Rhynchospora*, or Tribe Abildgaardieae; Roalson and Friar 2000; Ghamkhar et al. 2007; Thomas et al. 2009). Investigations of C_4 genetics in sedges are even more sparse and concerned exclusively *E. vivipara* (Agarie et al. 2002), one of a small group of species capable of employing either C_3 or C_4 photosynthesis depending on environmental conditions (Ueno et al. 1988; Ueno 2001, 2004; Edwards et al. 2004; Murphy et al. 2007).

The present study used species relationships deduced from plastid markers (Christin, Salamin, et al. 2008) to estimate the dates of the C_4 origins in this family. Analyses of nuclear *ppc* genes enabled testing for independent acquisitions of the C_4 pathway in the different C_4 lineages deduced from plastid markers. This phylogenetic framework was used here to shed light on the genetics of convergence, and specifically to 1) test for a role of adaptive genetic evolution in the acquisition of C_4-specific PEPCs, 2) detect the C_4-adaptive mutations and compare them among the different C_4 lineages of sedges, and 3) compare these changes with those that occurred in C_4 grasses and C_4 eudicots. These multi-scale comparisons revealed strong intrafamilial genetic convergence and even angiosperm-wide convergence of some genetic adaptations.

Materials and Methods

Plant Material and Molecular Dating of the Plastid DNA Phylogenetic Tree

The 104 species of Cyperaceae (including 31 C_4 species) analyzed with plastid DNA markers by Christin, Salamin, et al. (2008; supplementary table 1) were chosen to represent main tribes of Cyperaceae and all known C_4 lineages and their C_3 relatives (Bruhl and Wilson 2007; Muasya et al. 2009). A subsample of 63 species (including 18 C_4 species) was characterized in the present study for *ppc* gene variation (see below). This subsample represents all of the main tribes of Cyperaceae, and all C_4 lineages detected with plastid markers as well their C_3 relatives (supplementary table 1, Supplementary Material online).

The phylogenetic tree of Poales previously inferred from two plastid markers, *rbcL* and *ndhF*, using Bayesian methods (Christin, Salamin, et al. 2008) contains 334 commelinid species and four Asparagales that were used as outgroup. In this tree, C_4 sedges cluster in six fully C_4 lineages. This topology was used for Bayesian molecular dating, using the multidivtime software (Kishino et al. 2001; Thorne and Kishino 2002). The analysis was run as described by Christin, Besnard, et al. (2008). Calibration points also followed Christin, Besnard, et al. (2008), except that the upper bound set at 60 My on the stem node (divergence of the clade from its sister group) on the BEP–PACMAD clade of grasses was removed as suggested by Vicentini et al. (2008). This constraint had no significant effect on the age estimates (supplementary table 2, Supplementary Material online).

Isolation of PEPC-Encoding Genes

Only one *ppc* gene sequence of Cyperaceae (a cDNA of C_4 *ppc* from *E. vivipara*; EMBL no AB085948) was available in public databases. Polymerase chain reaction (PCR) isolation of the whole *ppc* genes was not easily feasible on genomic DNA due to its relatively large number of base pairs (bp). We firstly isolated sequences of *ppc* genes transcribed in leaves of a C_3 species, *Cyperus eragrostis*. Fresh leaves from a plant grown in a greenhouse were sampled at noon and total mRNAs were isolated using the NucleoSpin RNA Plant kit (Macherey-Nagel). Double-stranded complementary DNAs were then obtained with the ProtoScript First Strand cDNA Synthesis Kit

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

(BioLabs). Two primers able to PCR amplify a long *ppc* cDNA segment in monocots (~2,000 bp) were developed: *ppc*-850F (^{5}CAG TTC TCY TCY TGG ATG GG^3) and *ppc*-2872R (^{5}GCR GCR ATR CCC TTC ATG GT^3). The PCR reaction mixture contained ~100 ng of cDNA template, 5 μl of 10× AccuPrime PCR Buffer, 0.2 μM of each primer, 3 mM of $MgSO_4$, and 1 unit of proof-reading *Taq* polymerase (AccuPrime *Taq* DNA Polymerase High Fidelity, Invitrogen) in a total volume of 50 μl. This mixture was incubated at 94 °C for 2 min, followed by 36 cycles consisting of 45 s at 94 °C, 45 s at 53 °C, and 2 min 30 s at 68 °C. The last cycle was followed by 20 min at 68 °C. PCR products were purified with the QIAquick Gel Extraction Kit (Qiagen). The purified fragments were cloned into the pTZ57R/T vector with the InsT/Aclone PCR Product Cloning Kit (Fermentas). Positive clones were identified after PCR amplification using the M13 primers. PCR products were then digested with *Taq*I (Invitrogen), and clones with a distinct restriction pattern were purified and sequenced as in Christin et al. (2007). In addition, to assess the diversity of *ppc* genes present in the sedges genomes, we also screened genomic DNAs of six species (*Carex pendula* [C_3], *Chrysitrix dodii* [C_3], *Cyperus nipponicus* [C_4], *Fimbristylis littoralis* (C_4], *Pycreus sanguinolentus* [C_4], and *Scirpoides holoschoenus* [C_3]; see supplementary table 1, Supplementary Material online, for herbarium vouchers) using two primer combinations (supplementary table 3, Supplementary Material online). PCR, purification, cloning, and sequencing were conducted as described above. Genes encoding C_4 PEPC were identified by the presence of a serine at position 780 (according to maize C_4 PEPC; CAA33317), which characterizes all C_4 PEPCs and is necessary for the C_4 function (Bläsing et al. 2000; Christin et al. 2007).

According to the preliminary assessment of sedge *ppc* diversity, all genes encoding C_4 PEPC in this family belong to the same gene lineage (*ppc-1*; see Results). We thus focused on this gene lineage for the analysis of the subsample of 63 Cyperaceae species (supplementary table 1, Supplementary Material online). A gene segment covering exons 8–10 and carrying major C_4 amino acids (Bläsing et al. 2000; Christin et al. 2007) was characterized. Two introns are generally present in this *ppc* gene segment (Christin et al. 2007; supplementary fig. 1, Supplementary Material online). Based on *ppc* sequences isolated from *Cy. eragrostis* (EMBL no FM208065 and FM208066) and *E. vivipara* (AB085948), one first primer pair was designed: *ppc*-1336F ($^{5'}$TTT GGT CTC TCT YTT GTG CGT C$^{3'}$) and *ppc*-2727R ($^{5'}$GTT SGG GTC CCT GAT YCT TTT G$^{3'}$). These primers allow amplification of a *ppc* segment with 1,369 bp of coding sequence (supplementary fig. 1, Supplementary Material online). In a 50-μl volume, the PCR mixture contained ~100 ng of genomic DNA, 5 μl of 10× AccuPrime PCR Buffer II, 0.2 μM of each primer, 3 mmol of $MgSO_4$, 2.5 μl of dimethyl sulfoxide and 1 unit of *Taq* DNA polymerase (AccuPrime, Invitrogen). This mixture was incubated at 94 °C for 2 min, followed by 36 cycles consisting of 45 s at 94 °C, 45 s at 50 °C, and 2 min at 68 °C. The last cycle was followed by 20 min at 68 °C. PCR product purifications were done as described in the previous section. Twenty-nine species were successfully characterized with these primers (supplementary table 1, Supplementary Material online). In other cases, PCR failed (probably because DNAs were not of sufficiently high quality for amplification of DNA segments up to 2,000 bp) and thus a second primer pair was designed to amplify a shorter *ppc* segment (with 1,201 bp of coding sequence; supplementary fig. 1, Supplementary Material online): *ppc*-1426F ($^{5'}$GGG TCM TAC CGT GAG TGG TC$^{3'}$) and *ppc*-2647R ($^{5'}$CTT TGY TTT AGG TAG GGA TCT CC$^{3'}$). The PCR conditions used were exactly the same than those previously described. For sequencing, five different primers were used to directly sequence the *ppc* segments: *ppc*-2032F ($^{5'}$GAG CAG TCR TTT GGT GAG GAG C$^{3'}$), *ppc*-2209R ($^{5'}$GRC GGA AAT ACT CAA CAA AGC G$^{3'}$), *ppc*-2246F ($^{5'}$TGG AGT ATG GCC GYA TGA ACA T$^{3'}$), *ppc*-2411R ($^{5'}$CAT STG RAT GTT GCG CAC ATC CT$^{3'}$), and *ppc*-2600F ($^{5'}$CMA AGA AKC TYC TTC TTC AGG T$^{3'}$). Sequences obtained with these different primers overlap with one another. The sequencing protocol previously described was used (Christin et al. 2007). Every DNA sequencing chromatogram was checked and edited. In some species, double-peak or unreadable chromatograms were obtained due to the presence of multiple alleles or gene paralogs. To separate the different sequences, the purified fragments were cloned following the protocol described for cDNA. The distinct clones (identified by their *Taq*I restriction profile) were then sequenced.

Analyses of PEPC-Encoding Genes

Sequences obtained from genomic DNA were aligned with cDNAs, and introns were identified following the GT–AG rule. Phylogenetic trees were then reconstructed from coding sequences using Bayesian inference as implemented in MrBayes 3.1 (Ronquist and Huelsenbeck 2003). The best-fit substitution model, determined through hierarchical likelihood tests, was the general time reversible (GTR) model with a gamma shape parameter and a proportion of invariant sites (GTR + G + I). All model parameters were optimized independently for first, second, and third positions of codons. Two analyses, each of four parallel chains, were run for 10,000,000 generations. A tree was sampled each 1,000 generations after a burn-in period of 3,000,000 generations. A first phylogeny was inferred from coding sequences of sedge genes isolated with monocot *ppc* primers and PEPC-encoding genes from other plant families (supplementary table 4, Supplementary Material online). A second phylogenetic tree was then inferred only from coding sequences of gene lineage *ppc-1* (supplementary table 1, Supplementary Material online).

To confirm the phylogenetic pattern observed for *ppc-1* genes of the *Eleocharis* clade (see Results), their two introns were manually aligned and used to infer a phylogenetic tree, using the analysis parameters described above. The best fit-model for these introns was a GTR + G.

Positive Selection Tests

Three codon models were optimized on coding sequences of *ppc-1* from sedges, using the topology previously obtained. The software codeml, implemented in PAML 4

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

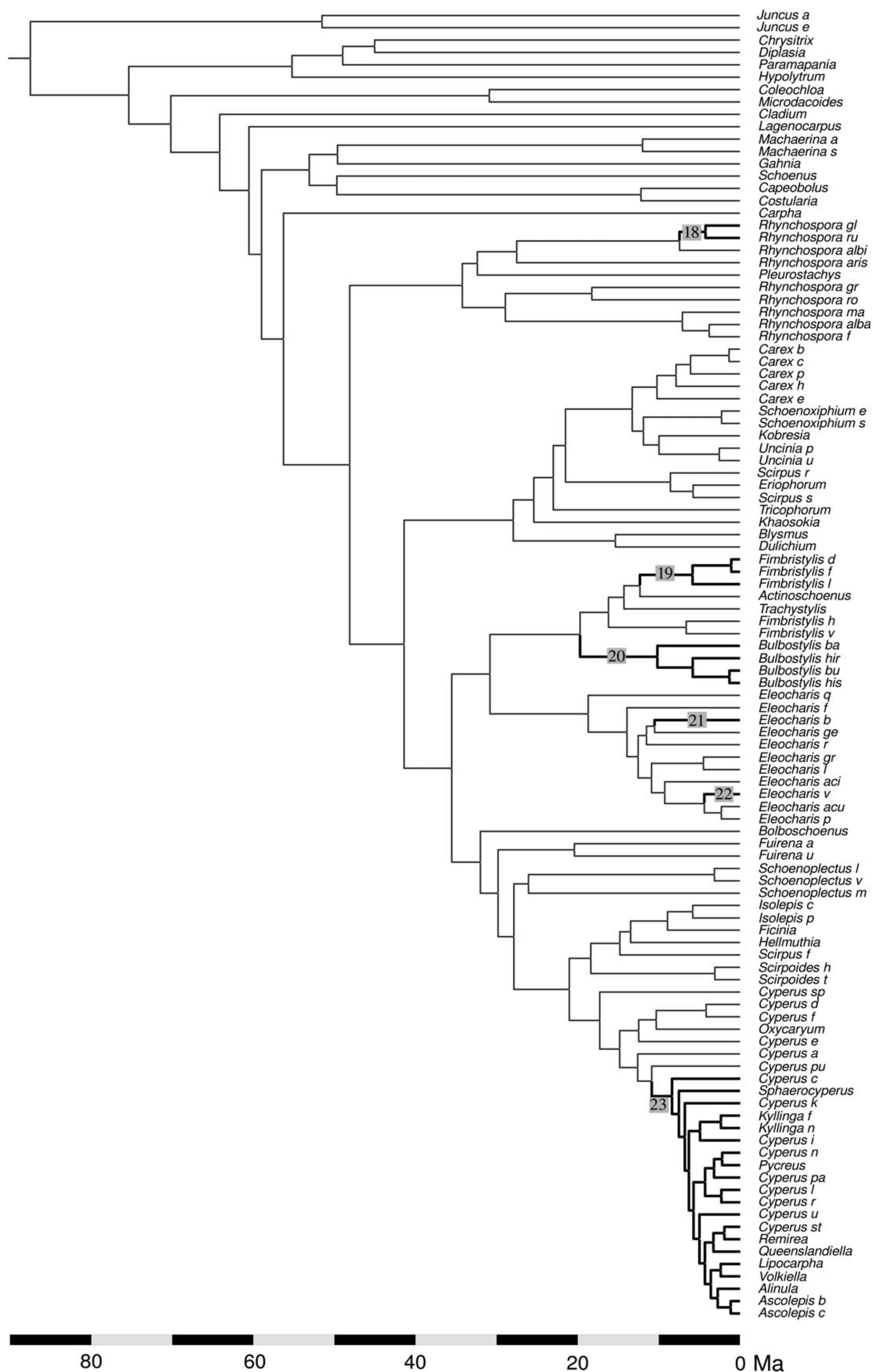

FIG. 1.—Calibrated Cyperaceae phylogenetic tree. The part of the phylogenetic tree corresponding to sedges and Juncaceae is presented. Branch lengths are proportional to estimated times, in million years before present (Ma). Branches leading to C_4 clades are in bold. C_4 lineages are numbered according to Christin, Salamin, et al. (2008). Estimates of C_4 evolutionary tempo can be found in table 1. Taxon names were abbreviated to the genus plus the first letter(s) of the species (when several species of the same genus were analyzed).

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

(Yang 2007), was used. These models evaluate the selective pressures driving the evolution of genes via the nonsynonymous versus synonymous mutations rate ratio (d_N/d_S; omega). An omega smaller than 1 indicates purifying selection, whereas an omega equal to 1 suggests relaxed selection. An omega value significantly greater than 1 demonstrates a proportion of fixed nonsynonymous mutations that is greater than expected by chance, which is usually interpreted as evidence for positive selection. The first model M1a is a site model allowing omega to vary among codons, being either smaller than 1 or equal to 1 (Yang et al. 2000). The alternative model A is a branch-site model that allows the omega ratio to vary both among codons and among branches of the phylogenetic tree (Yang and Nielsen 2002). In this model, sites are attributed to four different classes; either purifying selection in the whole tree, relaxed selection in the whole tree, purifying selection but positive selection in foreground branches, or relaxed selection but positive selection in foreground branches. The last model A′ is identical to model A except that positive selection is replaced by relaxed selection (Zhang et al. 2005). Both models M1a and A′ are nested in model A, enabling comparisons through likelihood ratio tests. The first test compares models M1a and A, whereas the second test, which compares models A′ and A, is more conservative (Zhang et al. 2005). In models A and A′, the foreground branches on which positive selection is expected have to be defined a priori. In this study, they were set simultaneously to all branches leading to a group in which all members encode C_4 PEPC with a serine at position 780.

The Bayes empirical Bayes procedure (Yang et al. 2005) implemented in codeml calculates the posterior probability of each codon to have evolved under positive selection in foreground branches. Codons for which this probability was greater than 0.999 were considered as having evolved under positive selection.

Results

Molecular Dating

The calibrated Cyperaceae phylogenetic tree is presented in figure 1. For taxa already present in our previous calibrated tree (Christin, Besnard, et al. 2008), divergence time values were fully congruent with the prior estimates (supplementary table 2, Supplementary Material online). Consequently, only the part of the tree corresponding to Cyperaceae is presented here. The divergence of sedges from Juncaceae (stem node) was estimated at 87.3 (± 7.9) Ma. The first divergence of two sedges (crown node) was at 75.1 (± 7.7) Ma. The age estimates for the sedge C_4 lineages are presented in table 1. Following our analyses, the first C_4 appearance occurred between 19.6 (± 4.9) and 10.1 (± 3.6) ago, in the *Bulbostylis* clade (table 1; fig. 1).

PEPC Evolutionary History

Using monocot *ppc* primers, 18 gene sequences were initially isolated from seven species (*Ca. pendula*, *Ch. dodii*, *Cy. eragrostis*, *Cy. nipponicus*, *F. littoralis*, *P. sanguinolentus*, and *S. holoschoenus*). Phylogenetic reconstructions showed that these genes belonged to five different gene lineages (named *ppc-1*, *ppc-2*, *ppc-3*, *ppc-4*, and *ppc-5*; fig. 2). Lineage *ppc-1* (isolated from six species) is apparently sister to *ppc-aL1* lineages of grasses. Closely related lineages *ppc-2* (isolated only from *Cy. nipponicus*), *ppc-3* (isolated only from *S. holoschoenus*), and *ppc-4* (isolated from six different species) are positioned as sister to grass lineages *ppc-aR*, *ppc-B1*, and *ppc-B2* (fig. 2). The fifth lineage, named *ppc-5*, was isolated only from *F. littoralis*, and is related to grass *ppc-aL2* (fig. 2). In the phylogenetic tree, this gene was positioned inside grass *ppc-aL2*, which likely results from a bias due to the absence of other sedge *ppc-5* and the small length of the isolated fragment (854 bp of coding sequence). Genes encoding C_4-PEPCs (with Ser^{780}) were identified in *E. vivipara* (AB085948; Agarie et al. 2002), *F. littoralis* and *P. sanguinolentus* and were only detected in lineage *ppc-1*.

Table 1
Age Estimates of Sedge C_4 Lineages

Lineage No	Clade Name	Stem Group Age	Crown Group Age
18	C_4 *Rhynchospora*	7.4 (± 2.8)	4.2 (± 2.2)
19	C_4 *Fimbristylis*	12.3 (± 3.8)	5.8 (± 2.6)
20	*Bulbostylis* clade	19.6 (± 4.9)	10.1 (± 3.6)
21	*Eleocharis baldwinii*	10.5 (± 3.2)	NA
22	*Eleocharis vivipara*	4.4 (± 2.1)	NA
23	C_4 Cypereae	10.9 (± 3.4)	8.4 (± 2.9)

We then focused on the *ppc-1* gene lineage for which 78 sequences were isolated from 63 sedge species (supplementary table 1, Supplementary Material online). The species relationships deduced from the phylogenetic tree inferred from coding sequences of this gene (fig. 3) were almost perfectly congruent with those deduced from plastid markers (fig. 1). The presence of two distinct *ppc-1* gene clusters in C_4 *Fimbristylis* indicates that gene duplication has occurred before the divergence of this clade. In *Eleocharis*, up to three different genes were isolated from the same individual (in *E. limosa*). Inside this genus, species relationships deduced from *ppc-1* genes do not concord with those inferred from plastid markers or ribosomal DNA internal transcribed spacer (ITS) sequences (Roalson and Friar 2000). The topology retrieved from *ppc-1* introns was identical to that inferred from coding sequences (supplementary fig. 2, Supplementary Material online). The fine-scale phylogenetic incongruence between plastid, ITS, and *ppc-1* genes points to a complex evolutionary history in *Eleocharis*.

At least one C_4 *ppc-1* sequence (i.e., encoding PEPC with Ser^{780}) was isolated from all C_4 species analyzed. Five distinct C_4 *ppc* gene lineages were identified in the phylogenetic tree (fig. 3). They display a faster rate of nonsynonymous substitutions compared to non-C_4 *ppc-1*, as attested by the branch lengths inferred from amino acid sequences (supplementary fig. 3, Supplementary Material online). The *ppc-1* gene isolated from *Kyllinga* underwent a large number of synonymous mutations, which led to a very long branch in the phylogeny inferred from nucleotides (fig. 3) but not in the amino acid tree (supplementary fig. 3, Supplementary Material online) indicating an increased evolutionary rate in this

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

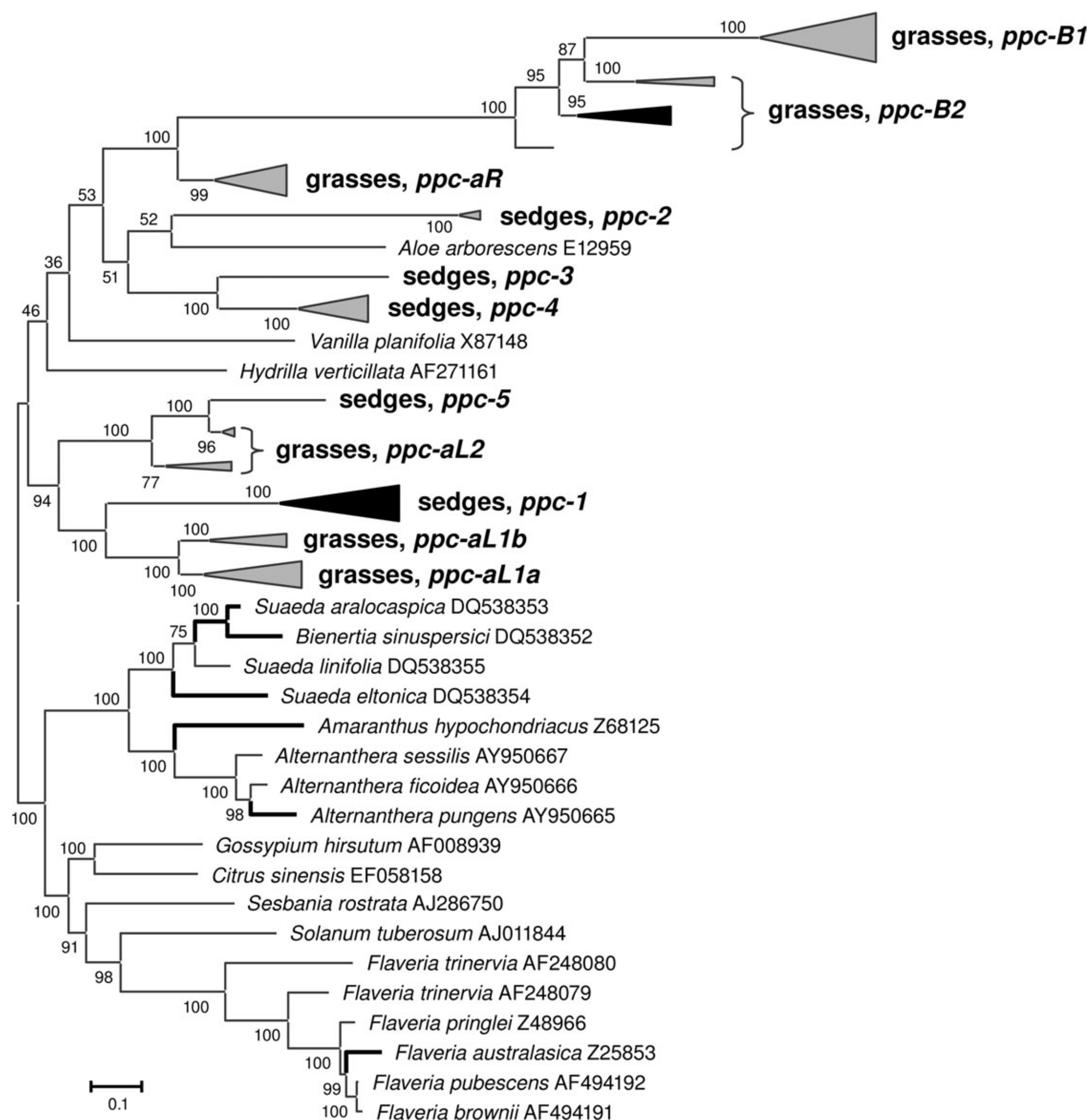

FIG. 2.—Phylogenetic tree of *ppc* genes in angiosperms. This phylogenetic tree was obtained through Bayesian inference on coding sequences of *ppc* genes of sedges and other angiosperms (see supplementary table 3, Supplementary Material online). Grass *ppc* clades, as previously defined (Christin et al. 2007), are compressed, as are the *ppc* gene lineages from sedges. Bayesian support values are indicated near branches. Branches leading to genes encoding C_4 PEPC (with a serine at position 780) are in bold. The two black compressed lineages contain both non-C_4 and C_4 genes.

species. The five C_4 *ppc* groups correspond to the C_4 lineages inferred from plastid markers, except for *Eleocharis baldwinii* and *E. vivipara*, which formed two distinct lineages in the plastid tree (fig. 1) but strongly clustered together in the *ppc-1* phylogeny (fig. 3).

Positive Selection

The model A, implementing positive selection in branches leading to sedge C_4 *ppc*, was significantly better than both null models (models M1a vs. A; chi-squared = 379.6, df = 2, P value < 0.0001; models A′ vs. A; chi-squared = 64.18, df = 1, P value < 0.0001). In model A, 31 sites were assigned to positive selection with a posterior probability greater than 0.95. By setting the threshold to 0.999, the number of codons under positive selection fell to 16. The amino acids encoded by these codons are shown for both C_4 and non-C_4 *ppc-1* groups (fig. 3). Many parallel genetic changes are observed between the five C_4 *ppc* lineages on these codons. In addition to the alanine to serine transition at position 780 (numbered according to *Zea mays* sequence, CAA33317), the amino acid at position 540 changed five times independently from a proline to a threonine, the one at position 665 from a histidine to a asparagine and other positions underwent similar or identical changes between two and four times independently (fig. 3). Only five positions (572, 665, 733, 761, and 780) were significantly detected as evolving under positive selection during the evolution of C_4-specific PEPC in both sedges (fig. 3) and grasses (Christin et al. 2007). The amino acid changes at these positions were the same in both families and even in some eudicots (table 2; Christin et al. 2007). In addition, some C_4-adaptive changes detected in PEPC genes of either grasses or sedges were also observed in some lineages of

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

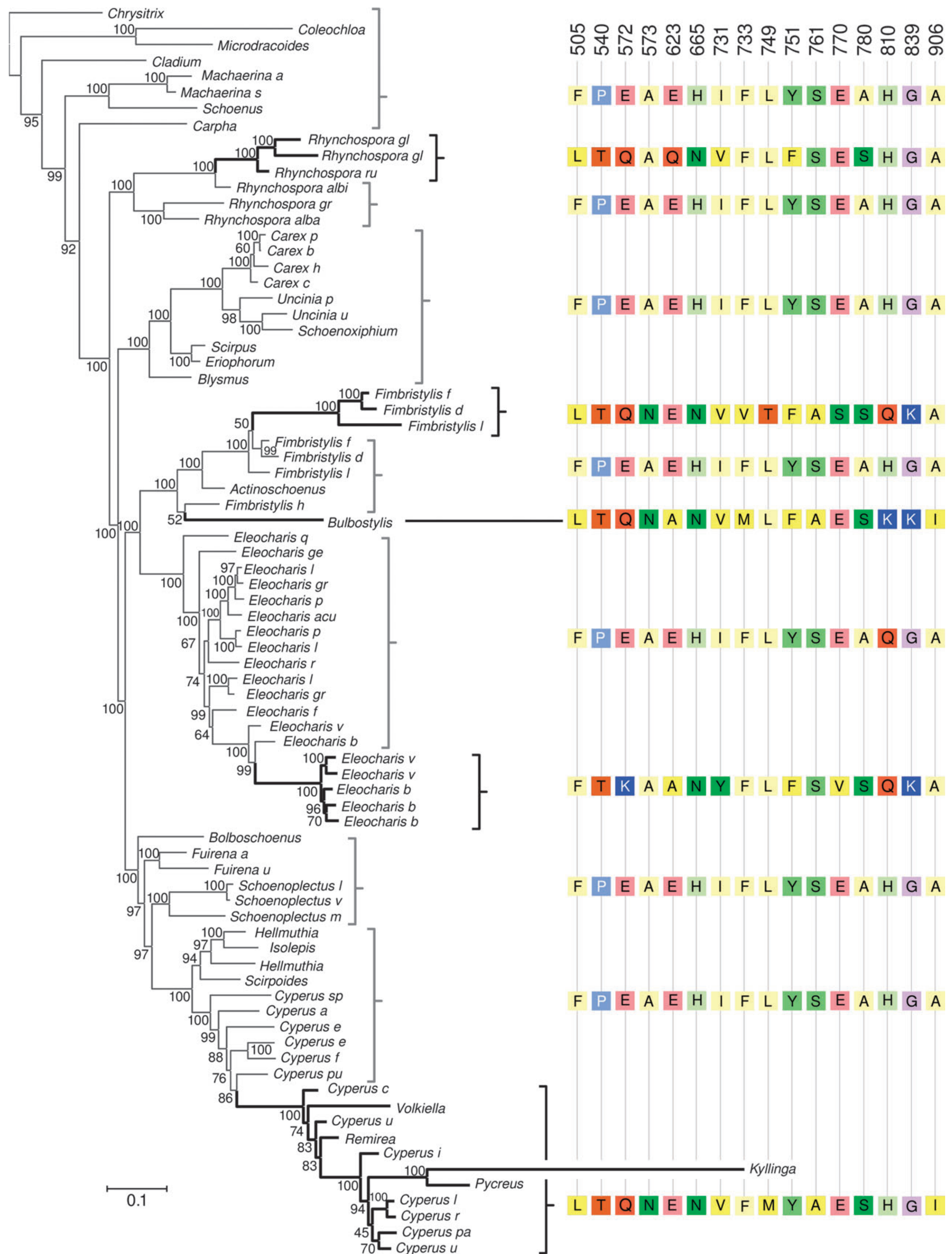

FIG. 3—Evolutionary history of sedge PEPC encoding genes. This phylogenetic tree was obtained through Bayesian inference on coding sequences of *ppc-1* genes from sedges. It was rooted on *Chrysitrix dodii* sequence (EMBL no FM208000). Bayesian support values are indicated near branches. Branches leading to C_4 *ppc* are in bold. For each C_4 and non-C_4 *ppc* clade, the most abundant amino acid for each codon under positive selection (numbered according to *Zea mays* sequence, CAA33317) is indicated on the right. C_4-specific amino acids are brightened and in bold. Taxon names were abbreviated to the genus plus the first letter(s) of the species (when several species of the same genus were analyzed).

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

Table 2
Adaptive Changes Shared by Sedge and Grass C_4 *ppc*

Group	*N*	Codon (No) 572	665	733	761	780
Grasses	8	E → Q: 7	H → N: 6 H → R: 1	F → V: 3 F → M: 2	S → A: 6	A → S: 8
Sedges	5	E → Q: 4 E → K: 1	H → N: 5	F → V: 1 F → M: 1	S → A: 3	A → S: 5
Eudicots	5	E → Q: 1 E → K: 1	H → N: 3		S → A: 3	A → S: 5

Note.—For the five positions (numbered according to *Zea mays* accession CAA33317) that were detected as evolving under positive selection in both sedges (this study) and grasses (Christin et al. 2007), the number of times C_4-linked amino acid mutations occurred is indicated, for sedges, grasses, and eudicots. The number of independent lineages (*N*), for which C_4 PEPC is available, is indicated for each group.

the other family, although they were not significantly attributed to positive selection in this second family (supplementary table 5, Supplementary Material online).

Discussion

Number and Tempo of C_4 Origins in Sedges

Six distinct C_4 lineages were present in the phylogenetic tree inferred from plastid markers (fig. 1). Analyses of *ppc-1*, which encodes the key enzyme of the C_4 pathway, unequivocally showed that this gene acquired C_4 attributes at least five times independently (fig. 3), confirming the multiple origins of the C_4 pathway in Cyperaceae. Intriguingly, *E. baldwinii* and *E. vivipara*, two C_4 species that are clearly distinguished based on plastid DNA (fig. 1), ITS (Roalson and Friar 2000), as well as phenotypic C_4 characteristics (Ueno et al. 1988; Soros and Dengler 2001), were strongly supported as monophyletic considering their C_4 *ppc* genes, which are very similar (fig. 3). This incongruence between *ppc* genes and other phylogenetic markers has been suggested to be due to phylogenetic bias caused by adaptive evolution (Roalson 2007) as seen for grass *ppc* genes (Christin et al. 2007). However, even when considering only introns, the C_4 *ppc* genes of *E. vivipara* and *E. baldwinii* clustered together with very high support (supplementary fig. 2, Supplementary Material online). These two unrelated taxa seem to have acquired their C_4 *ppc* from the same source. The very low divergence between these *ppc* genes (fig. 3; supplementary fig. 2, Supplementary Material online) rules out an acquisition of C_4 *ppc* during the early diversification of the *Eleocharis* genus followed by recurrent losses in C_3 species. It is thus likely that one of the two lineages evolved a C_4-specific *ppc* gene that was recently transmitted to the other taxa (probably *E. vivipara*; Roalson forthcoming), either through horizontal gene transfer or hybridization. C_4 evolution in these two *Eleocharis* lineages is thus not completely independent. Polyploidization is apparently very frequent in the genus *Eleocharis* (Yano et al. 2004; da Silva et al. 2008; Roalson 2008a) and an acquisition of C_4 genes through allopolyploidization between C_3 and C_4 parents is also likely. Further investigations are required to solve this issue. Species sampling of *Eleocharis* and number of molecular markers should especially be increased.

In the family Cyperaceae, the first C_4 evolution (*Bulbostylis*) has likely occurred between 19.6 (± 4.9) and 10.1 (± 3.6) Ma, but the other C_4 appearances all arose during the last 12 My (table 1), making C_4 sedges generally younger than their grass counterparts (Christin, Besnard, et al. 2008). This raises questions about the drivers of these evolutionary events (Roalson 2008b). Indeed, if the Oligocene CO_2 decline created the general condition for C_4 photosynthesis to be advantageous in some environments (Christin, Besnard, et al. 2008; Roalson 2008b), other factors, such as heat, drought, fire, and other disturbances, could have locally selected for the C_4 pathway (Osborne 2008). Increasing perturbations of the midlatitude ecosystems throughout the Miocene have thus probably contributed to the success of C_4 sedges, as for C_4 grasses (Keeley and Rundel 2005; Beerling and Osborne 2006; Osborne 2008). However, C_4 sedges generally occupy wetter habitats than C_4 grasses within the fire-adapted vegetation (Linder and Rudall 2005), and they probably evolved their C_4 systems in a wetland context (Li et al. 1999; Stock et al. 2004). Photorespiration is generally reduced in such conditions because water availability allows stomatal aperture, which helps in maintaining significant intracellular CO_2 concentrations. It is thus possible that ecosystem-scale feedbacks between fire and vegetation have lately affected wetter grasslands, partially explaining why C_4 sedges have generally appeared later than C_4 grasses.

In Cyperaceae, however, a few C_3 and C_4 taxa occupy seasonally dry habitats, such as shallow soils of inselbergs in arid tropical savannah (Porembski and Barthlott 2000; Proctor and Pence 2002; Proctor and Tuba 2002). Some taxa are desiccation tolerant (e.g., poikilohydry) having the ability to evade drought, by rapid equilibration of the plant's water content to that of the surrounding environment, and yet retain the ability to resume active metabolism when conditions are favorable. Desiccation tolerance has evolved independently in sedge tribes Trilepideae (C_3; at least 5 species of 17), Cariceae (C_3; 1 species), Abildgaardieae (C_4; 2 species), and Cypereae (C_4; 3 species). Nevertheless, this adaptive trait is a common characteristic only in the Trilepideae (e.g., *Coleochloa* and *Microdracoides*), which have a crown age of 30.8 (±7.5) My in our phylogenetic analysis (fig. 1). The few C_4 sedge taxa that have colonized inselbergs (e.g., *Kyllinga alba*, *Bulbostylis leucostachya*) may have acquired the desiccation-tolerance trait relatively recently. The C_4 pathway, which raises the water-use efficiency compared with the C_3 type, has also probably contributed to such evolutionary transitions to arid habitats in these lineages as recently suggested in grasses (Osborne and Freckleton 2009).

Constraints in C_4 PEPC Evolution

In sedges, the five C_4 *ppc-1* lineages have clearly undergone a very high number of nonsynonymous substitutions (supplementary fig. 3, Supplementary Material online). Positive selection analyses showed that this was due to important adaptive changes along the protein sequence, as in grasses (Christin et al. 2007). The evolution of C_4-specific enzymes likely involves not only changes of regulatory sequences to acquire the light-dependent expression pattern in M cells (Akyildiz et al. 2007), but also kinetic modifications (Dong et al. 1998; Bläsing et al. 2000; Gowik et al. 2006; Rao

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

et al. 2008) that were likely achieved through the adaptive amino acid changes identified on C_4 *ppc-1*.

The *ppc* lineages that evolved the C_4 function are clearly not randomly distributed in the phylogeny because only one *ppc* lineage of five in sedges became recurrently involved for the C_4 pathway. Similarly, of the six *ppc* gene lineages that exist in grasses, the *ppc-B2* lineage developed a C_4 function at least eight times independently (Christin et al. 2007). Interestingly, gene lineages *ppc-1* and *ppc-B2* are clearly not orthologous (fig. 2). Recurrent recruitments of the same gene lineages in Cyperaceae or Poaceae suggest that they have predispositions for the C_4 function. For instance, a non-C_4 gene allowing the acquisition of a light-induced and cell-specific expression through simple genetic changes (as shown in *Flaveria*; Akyildiz et al. 2007) is likely to be preferentially selected for the C_4 cycle. The possibility of easily altering the enzyme expression patterns likely constrained the originations of the C_4 PEPC to some *ppc* lineages. Once the C_4 transcription was acquired, kinetic properties had to be adapted to the new substrate and product concentrations present in the C_4 photosynthetic cells (Dong et al. 1998; Bläsing et al. 2000). The different C_4 sedge groups evolved a C_4 PEPC starting from very similar genes, and the high similarity of amino acid sequences of non-C_4 *ppc-1* (supplementary fig. 3, Supplementary Material online) suggests that, before C_4 evolution, genes belonging to this lineage had conserved functions and kinetic properties. This probably accounts for the recurrence of many changes on *ppc-1* in the five C_4 sedge lineages. The number of amino acid changes possible along a sequence is extremely large, but many of them would be detrimental because they would reduce or modify enzymatic activity. Out of the small proportion of amino acid changes that would be advantageous after the recruitment of a *ppc* gene in the C_4 pathway, many will be out of reach. A new optimum is accessible only through step mutations that have to be all advantageous or at least neutral, which limits the number of possible evolutionary paths (Weinreich et al. 2006). The molecular convergence during sedge C_4-PEPC evolution probably results from both a limited number of new optima and the genetic background constraining the coverage of the protein space (Hodin 2000).

Intra and Interfamilial Genetic Convergence

Of the 16 codons with a high probability of having evolved under positive selection in branches leading to C_4 *ppc* of sedges, only five were also detected as C_4-adaptive in grasses (Christin et al. 2007). Despite showing a strong genetic convergence within sedges and grasses, respectively, C_4-specificities of PEPC strongly vary among families. This is probably due to the strong divergence of the non-C_4 genes that were recruited for the C_4 pathway in each of the two families (fig. 2). Grass *ppc* genes diverged from each other between 10 and 30 My before acquiring a C_4 function (Christin, Besnard, et al. 2008) and sedge *ppc* genes less than 40 My (fig. 1). On the other side, C_4 recruitments of *ppc* occurred more than 90 My after the divergence of grasses and sedges (Christin, Besnard, et al. 2008). Non-C_4 *ppc* genes of grasses and sedges are only distantly related and are putatively responsible for different functions, with other associated kinetics. Therefore, it is likely that the path to the C_4 optimum varies depending on the starting point. The same explanation probably accounts for the nontransferability of most grass C_4-adaptive amino acid changes to eudicot systems. A common starting point strongly constrained the evolutionary path to C_4-optimized enzymes in closely related plant lineages, whereas the important divergence of non-C_4 genes among major C_4 clades opened the road toward other C_4 optima.

Despite family-specific adaptive changes to fulfill the C_4 function, it is worth noting that five codons underwent similar changes in sedges, grasses, and eudicots (table 2). In addition to the alanine at position 780 that mutated to a serine at least five times in sedges, eight times in grasses, and five times in eudicots, the histidine at position 665 mutated to an asparagine at least five, six, and three times in the three groups, respectively. Recurrent changes in many C_4 groups are observed at these five positions (table 2). Whatever the gene recruited in the C_4 pathway, these five amino acids were often mutated to an identical residue. They are thus indicative of the very strongly convergent genetic evolution linked to the acquisition of C_4-specific PEPCs, and of the repeatability of some evolutionary processes, at the genetic level and across very broad taxonomic scales. These five sites are also potential universal determinants of the C_4 function and could be used to transform *ppc* genes of any plant to a C_4-specific enzyme, opening promising opportunities for the engineering of the C_4 pathway in C_3 plants (Hibberd et al. 2008). In the next decade, phylogenomic analyses of C_4 evolution should be extended to other families and additional enzymes toward a comprehensive understanding of the genetic mechanisms linked to the numerous C_4 origins. In the coming genomic era, multi-level C_4 comparative studies could position C_4 photosynthesis as a case study for molecular evolution.

Supplementary Material

Supplementary tables 1– 5 and supplementary figures 1– 3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

This work was funded by Swiss NSF grant 3100AO-105886 to G.B. It partially results from a Switzerland–South Africa joint meeting, funded by the Swiss NSF grant IB32A0-112804. P. Danton, F. Dessimoz, N. Fumeaux, B. Gehrke, F. Jacquemoud, and O. Ueno provided very useful samples.

Literature Cited

Agarie S, Kai M, Takatsuji H, Ueno O. 2002. Environmental and hormonal regulation of gene expression of C_4 photosynthetic enzymes in the amphibious sedge *Eleocharis vivipara*. Plant Sci. 163:571–580.

Akyildiz M, Gowik U, Engelmann S, Koczor M, Streubel M, Westhoff P. 2007. Evolution and function of cis-regulatory module for mesophyll-specific gene expression in the C_4 dicot *Flaveria trinervia*. Plant Cell. 19:3391–3402.

Beerling DJ, Osborne CP. 2006. The origin of the savanna biome. Glob Change Biol. 12:2023–2031.

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

Bläsing OE, Westhoff P, Svensson P. 2000. Evolution of C_4 phospho*enol*pyruvate carboxylase in *Flaveria*, a conserved serine residue in the carboxyl-terminal part of the enzyme is a major determinant for C_4-specific characteristics. J Biol Chem. 275:27917–27923.

Bruhl JJ, Perry S. 1995. Photosynthetic pathway-related ultrastructure of C_3, C_4 and C_3-like C_3-C_4 intermediate sedges (Cyperaceae), with special reference to *Eleocharis*. Aust J Plant Physiol. 22:521–530.

Bruhl JJ, Stone NE, Hattersley PW. 1987. C_4 acid decarboxylation enzymes and anatomy in sedges (Cyperaceae): first record of NAD–malic enzyme species. Aust J Plant Physiol. 14:719–728.

Bruhl JJ, Wilson KL. 2007. Towards a comprehensive survey of C_3 and C_4 photosynthetic pathway in Cyperaceae. Aliso. 23:99–148.

Christin PA, Besnard G, Samaritani E, Duvall MR, Hodkinson TR, Savolainen V, Salamin N. 2008. Oligocene CO_2 decline promoted C_4 photosynthesis in grasses. Curr Biol. 18:37–43.

Christin PA, Petitpierre B, Salamin N, Büchi L, Besnard G. 2009. Evolution of C_4-related phospho*enol*pyruvate carboxykinase in grasses, from genotype to phenotype. Mol Biol Evol. 26:357–365.

Christin PA, Salamin N, Muasya AM, Roalson EH, Russier F, Besnard G. 2008. Evolutionary switch and genetic convergence on *rbcL* following the evolution of C_4 photosynthesis. Mol Biol Evol. 25:2361–2368.

Christin PA, Salamin N, Savolainen V, Duvall MR, Besnard G. 2007. C_4 photosynthesis evolved in grasses via parallel adaptive genetic changes. Curr Biol. 17:1241–1247.

Conway Morris S. 2006. Evolutionary convergence. Curr Biol. 16:R826–R827.

da Silva CRM, González-Elizondo MS, do Nascimento Andrade de Alemeida Rego L, Torezan JMD, Vanzela ALL. 2008. Cytogenetical and cytotaxonomical analysis of some Brazilian species of *Eleocharis* (Cyperaceae). Aust J Bot. 56:82–90.

Dong LY, Masuda T, Kawamura T, Hata S, Izui K. 1998. Cloning, expression, and characterization of a root-form phospho*enol*pyruvate carboxylase from *Zea mays*: Comparison with the C_4-form enzyme. Plant Cell Physiol. 39: 865–873.

Edwards GE, Franceschi VR, Voznesenskaya EV. 2004. Single-cell C_4 photosynthesis versus the dual-cell (Kranz) paradigm. Annu Rev Plant Biol. 55:173–196.

Ehleringer JR, Cerling TE, Helliker BR. 1997. C_4 photosynthesis, atmospheric CO_2, and climate. Oecologia. 112:285–299.

Foyer CH, Bloom A, Queval G, Noctor G. 2009. Photorespiratory metabolism: genes, mutants, energetics, and redox signaling. Annu Rev Plant Biol. 60:455–484.

Ghamkhar K, Marchant AD, Wilson KL, Bruhl JJ. 2007. Phylogeny of Abildgaardieae (Cyperaceae) inferred from ITS and *trn*L-F data. Aliso. 23:149–164.

Gowik U, Engelmann S, Bläsing OE, Raghavendra AS, Westhoff P. 2006. Evolution of C_4 phospho*enol*pyruvate carboxylase in the genus *Alternanthera:* gene families and the enzymatic characteristics of the C_4 isozyme and its orthologues in C_3 and C_3/C_4 Alternantheras. Planta. 223:359–368.

Hibberd JM, Sheehy JE, Langdale JA. 2008. Using C_4 photosynthesis to increase the yield of rice: rationale and feasibility. Curr Opin Plant Biol. 11:228–231.

Hodin J. 2000. Plasticity and constraints in development and evolution. J Exp Zool (Mol Dev Evol). 288:1–20.

Jacobs B, Engelmann S, Westhoff P, Gowik U. 2008. Evolution of C_4 phosphoenolpyruvate carboxylase in *Flaveria*: determinants for high tolerance towards the inhibitor L-malate. Plant Cell Environ. 31:793–803.

Kanai R, Edwards GE. 1999. The biochemistry of C_4 photosynthesis. In: Sage RF, Monson RK, editors. C_4 plant biology. San Diego (CA): Academic Press. p. 49–87.

Keeley JE, Rundel PW. 2005. Fire and the Miocene expansion of C_4 grasslands. Ecol Lett. 8:683–690.

Kishino H, Thorne JL, Bruno WJ. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol Biol Evol. 18:352–361.

Latzko E, Kelly GJ. 1983. The many-faced function of phospho*enol*pyruvate carboxylase in C_3 plants. Physiol Veg. 21:805–815.

Li MR, Wedin DA, Tieszen LL. 1999. C_3 and C_4 photosynthesis in *Cyperus* (Cyperaceae) in temperate eastern North America. Can J Bot. 77:209–218.

Linder HP, Rudall PJ. 2005. Evolutionary history of Poales. Annu Rev Ecol Evol Syst. 36:107–124.

Lloyd J, Farquhar GD. 1994. ^{13}C discrimination during CO_2 assimilation by the terrestrial biosphere. Oecologia. 99:201–215.

McKown AD, Moncalvo JM, Dengler NG. 2005. Phylogeny of *Flaveria* (Asteraceae) and inference of C_4 photosynthesis evolution. Am J Bot. 92:1911–1928.

Muasya AM, Simpson DA, Verboom GA, Goetghebeur P, Naczi RFC, Chase MW, Smets E. 2009. Phylogeny of Cyperaceae based on DNA sequence data: current progress and future prospects. Bot Rev. 75:2–21.

Murphy LR, Barroca J, Franceschi VR, Lee R, Roalson EH, Edwards GE, Ku MSB. 2007. Diversity and plasticity of C_4 photosynthesis in *Eleocharis* (Cyperaceae). Funct Plant Biol. 34:571–580.

Osborne CP. 2008. Atmosphere, ecology and evolution: what drove the Miocene expansion of C_4 grasslands? J Ecol. 96:35–45.

Osborne CP, Freckleton RP. 2009. Ecological selection pressures for C_4 photosynthesis in the grasses. Proc Biol Sci. 276:1753–1760.

Porembski S, Barthlott W. 2000. Granitic and gneissic outcrops (inselbergs) as centers of diversity for desiccation-tolerant vascular plants. Plant Ecol. 151:19–28.

Proctor MCF, Pence VC. 2002. Vegetative tissues: bryophytes, vascular 'resurrection plants' and vegetative propagules. In: Black M, Pritchard H, editors. Desiccation and survival in plants: drying without dying. Wallingford (United Kingdom): CAB International. p. 207–237.

Proctor MCF, Tuba Z. 2002. Poikilohydry and homoihydry: antithesis or spectrum of possibilities? New Phytol. 156: 327–349.

Rao SK, Reiskind JB, Bowes G. 2008. Kinetic analyses of recombinant isoforms of phospho*enol*pyruvate carboxylase from *Hydrilla verticillata* leaves and the impact of substituting a C_4-signature serine. Plant Sci. 174:475–483.

Roalson EH. 2007. C_4 photosynthesis: convergence upon convergence upon. Curr Biol. 17:R776–R778.

Roalson EH. 2008a. A synopsis of chromosome number variation in the Cyperaceae. Bot Rev. 74:209–393.

Roalson EH. 2008b. C_4 photosynthesis: differentiating causation and coincidence. Curr Biol. 18:R167–R168.

Roalson EH. Forthcoming. Origins and transitions in photosynthetic pathway types in the monocots: a review and reanalysis. In: Raghavendra AS, Sage RF, editors. C_4 photosynthesis and related CO_2 concentrating mechanisms. New York: Springer.

Roalson EH, Friar EA. 2000. Infrageneric classification of *Eleocharis* (Cyperaceae) revisited: evidence from the internal transcribed spacer (ITS) region of nuclear ribosomal DNA. Syst Bot. 25:323–336.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19:1572–1574.

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024

Sage RF. 2004. The evolution of C_4 photosynthesis. New Phytol. 161:341–370.
Sage RF, Wedin DA, Li MR. 1999. The biogeography of C_4 photosynthesis: patterns and controlling factors. In: Sage RF, Monson RK, editors. C_4 plant biology. San Diego (CA): Academic Press. p. 313–373.
Soros CL, Bruhl JJ. 2000. Multiple evolutionary origins of C_4 photosynthesis in the Cyperaceae. In: Wilson KL, Morrison DA, editors. Monocots: systematics and evolution. Collingwood (Australia): CSIRO Publishing. Vol. 1, p. 629–636.
Soros CL, Dengler NG. 2001. Ontogenic deviation and cell differentiation in photosynthetic tissues of C_3 and C_4 Cyperaceae. Am J Bot. 88:992–1005.
Stock WD, Chuba DK, Verboom GA. 2004. Distribution of South African C_3 and C_4 species of Cyperaceae in relation to climate and phylogeny. Austral Ecol. 29:313–319.
Svensson P, Bläsing OE, Westhoff P. 2003. Evolution of C_4 phospho*enol*pyruvate carboxylase. Arch Biochem Biophys. 414:180–188.
Thomas WW, Araújo AC, Alves MV. 2009. A preliminary molecular phylogeny of the Rhynchosporeae (Cyperaceae). Bot Rev. 75:22–29.
Thorne JL, Kishino H. 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst Biol. 51:689–702.
Ueno O. 2001. Environmental regulation of C_3 and C_4 differentiation in the amphibious sedge *Eleocharis vivipara*. Plant Physiol. 127:1524–1532.
Ueno O. 2004. Environmental regulation of photosynthetic metabolism in the amphibious sedge *Eleocharis baldwinii* and comparisons with related species. Plant Cell Environ. 27:627–639.
Ueno O, Koyama T. 1987. Distribution and evolution of C_4 syndrome in *Rhynchospora* (Rhynchosporeae–Cyperaceae). Bot Mag Tokyo. 100:63–85.
Ueno O, Samejima M, Koyama T. 1989. Distribution and evolution of C_4 syndrome in *Eleocharis*, a sedge group inhabiting wet and aquatic environments, based on culm anatomy and carbon isotope ratios. Ann Bot. 64: 425–438.
Ueno O, Samejima M, Muto S, Miyachi S. 1988. Photosynthetic characteristics of an amphibian plant, *Eleocharis vivipara*: expression of C_4 and C_3 modes in contrasting environments. Proc Natl Acad Sci USA. 85:6733–6737.
Vicentini A, Barber JC, Aliscioni SA, Giussiani LM, Kellogg EA. 2008. The age of the grasses and clusters of origins of C_4 photosynthesis. Glob Change Biol. 14:2963–2977.
von Caemmerer S, Furbank RT. 2003. The C_4 pathway: an efficient CO_2 pump. Photosynth Res. 77:191–207.
Weinreich DM, Delaney NF, DePristo MA, Hartl DL. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science. 312:111–114.
Westhoff P, Gowik U. 2004. Evolution of C_4 phospho*enol*pyruvate carboxylase. genes and proteins: a case study with the genus *Flaveria*. Ann Bot. 93:13–23.
Yang ZH. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591.
Yang ZH, Nielsen R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. Mol Biol Evol. 19:908–917.
Yang ZH, Nielsen R, Goldman N, Krabbe Pedersen AM. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics. 155:431–449.
Yang ZH, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acids sites under positive selection. Mol Biol Evol. 22:1107–1118.
Yano O, Katsuyama T, Tsubota H, Hoshino T. 2004. Molecular phylogeny of Japanese *Eleocharis* (Cyperaceae) based on ITS sequence data, and chromosomal evolution. J Plant Res. 117:409–419.
Zhang JZ, Nielsen R, Yang ZH. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 22:2472–2479.

David Irwin, Associate Editor

Accepted May 5, 2009

Downloaded from https://academic.oup.com/mbe/article/26/8/1909/981544 by guest on 20 April 2024