**C₄ eudicots are not younger than C₄ monocots**

Pascal-Antoine Christin¹*, Colin P. Osborne², Rowan F. Sage³, Mónica Arakaki¹ and Erika J. Edwards¹

¹ Department of Ecology and Evolutionary Biology, Brown University, 80 Waterman St, Box G-W, Providence, RI 02912, USA
² Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK
³ Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON M5S3B2, Canada

* To whom correspondence should be addressed. E-mail: Pascal-Antoine_Christin@brown.edu

Received 17 November 2010; Revised 27 January 2011; Accepted 28 January 2011

**Abstract**

C₄ photosynthesis is a plant adaptation to high levels of photorespiration. Physiological models predict that atmospheric CO₂ concentration selected for C₄ grasses only after it dropped below a critical threshold during the Oligocene (~30 Ma), a hypothesis supported by phylogenetic and molecular dating analyses. However the same models predict that CO₂ should have reached much lower levels before selecting for C₄ eudicots, making C₄ eudicots younger than C₄ grasses. In this study, different phylogenetic datasets were combined in order to conduct the first comparative analysis of the age of C₄ origins in eudicots. Our results suggested that all lineages of C₄ eudicots arose during the last 30 million years, with the earliest before 22 Ma in Chenopodiaceae and Aizoaceae, and the latest probably after 2 Ma in Flaveria. C₄ eudicots are thus not globally younger than C₄ monocots. All lineages of C₄ plants evolved in a similar low CO₂ atmosphere that predominated during the last 30 million years. Independent C₄ origins were probably driven by different combinations of specific factors, including local ecological characteristics such as habitat openness, aridity, and salinity, as well as the speciation and dispersal history of each clade. Neither the lower number of C₄ species nor the frequency of C₃–C₄ intermediates in eudicots can be attributed to a more recent origin, but probably result from variation in diversification and evolutionary rates among the different groups that evolved the C₄ pathway.

**Key words:** C₄ photosynthesis, eudicots, evolution, molecular dating, multiple origins, phylogeny.

**Introduction**

C₄ photosynthesis is a highly convergent trait that has evolved more than 60 times in at least 18 families of flowering plants (Sage et al., 2011a). It consists of anatomical and biochemical innovations that increase the internal CO₂ concentration around the carboxylating enzyme Rubisco, thereby suppressing oxygenation of RuBP and photorespiration. C₄ photosynthesis thus provides an advantage in all conditions where photorespiration levels are deleteriously high, especially in warm, dry, and/or saline habitats (Sage, 2004). However, the energetic requirements of the C₄ pathway can represent a net cost in conditions where photorespiratory rates are low, most notably at low temperature or in atmospheres of elevated CO₂. One index that effectively describes the relationship between the efficiency of C₃ and C₄ photosynthesis and the environment is the maximum quantum yield of photosynthesis (QY) measured at low light intensities (Ehleringer et al., 1997). Maximum QY (hereafter referred to as simply QY) is a measure of the maximum light use efficiency of photosynthesis, calculated as the number of CO₂ molecules fixed per absorbed photon. As an index of photorespiratory inhibition, QY is useful in C₃ and C₄ comparisons both at low light where variation in light use efficiency is directly proportional to carbon gain, and at high light, where QY is inversely proportional to the ratio of photorespiration to photosynthesis (Sage and Kubien, 2003). Due to the increase in photorespiration, QY in C₃ plants declines with decreasing CO₂ levels and increasing temperature. By contrast, C₄ QY is little affected by variation in CO₂ and temperature, but is below that of C₃ plants at high CO₂ and/or low temperature due to the two ATP equivalents
needed for each turn of the C₄ metabolic cycle (Ehleringer et al., 1997).

Because of its simplicity and close relationship with the ratio of photorespiration to photosynthesis, QY has been widely used to compare C₃ and C₄ species across a range of environments, and to model conditions, both present and past, when the C₄ pathway would be advantageous over the C₃ pathway. Using a QY model, Ehleringer and co-workers (Ehleringer et al., 1991, 1997; Cerling et al., 1997) predicted that C₄ plants would have greater evolutionary fitness under the low CO₂ levels of recent geological time than C₃ species, leading to the hypothesis that low CO₂ was a selection factor for the rise of the C₄ pathway. With improved estimates of paleo atmospheric CO₂ and evidence of C₄ plants occurring in the Mid-Miocene, the date proposed for the earliest C₄ plants has been placed in the late Oligocene period, when global climate underwent important changes and atmospheric CO₂ levels fell to near or below current levels (Sage, 2001, 2004; Pagani et al., 2005). This hypothesis has been empirically supported in C₄ monocots (grasses and sedges), where phylogenetic modelling indicates a marked increase in the probability of C₄ evolution after the Oligocene CO₂ decline (Christin et al., 2008; Besnard et al., 2009). It has not been tested in the numerous eudicot lineages with C₄ species, however. QY comparisons generally demonstrated lower values of light use efficiency in C₄ eudicots than C₄ monocots (Ehleringer and Pearcy, 1983). This observation led to the hypothesis that C₄ evolution in eudicots would have only occurred at lower atmospheric CO₂ levels than in monocots, constraining C₄ evolution in eudicots to the Pleistocene (Ehleringer et al., 1997). This prediction can now be evaluated using phylogenetic information accumulated in the past 15 years for eudicot lineages where C₄ photosynthesis appears.

Phylogenetic data have increased exponentially in recent years, leading to a better understanding of the relationships between C₄ taxa and affiliated C₃ taxa. The grasses (Poaceae) have received the lion’s share of phylogenetic attention and the number and timing of C₄ origins in this group are now relatively well understood, with a minimum of 17 C₄ origins, starting around 30 million years ago (Ma) and continuing to more recent geological time (Giussani et al., 2001; Christin et al., 2008; Vicentini et al., 2008; Edwards et al., 2010; Roalson, 2011). This pattern, along with considerations of the ecological setting for C₄ evolution, has led to the current view that Oligocene CO₂ decline met an essential environmental precondition for the evolution of C₄ photosynthesis, while individual C₄ origins were probably driven by additional local factors, such as warmth, seasonality, and habitat openness (Sage, 2001, 2004; Roalson, 2008; Osborne and Freckleton, 2009; Edwards and Smith, 2010). Besides grasses, five C₄ origins in Cyperaceae (sedges) were estimated to have occurred between 20 and 4 Ma (Besnard et al., 2009; Roalson, 2011). The C₄ origins in sedges thus also occurred in a low CO₂ world, perhaps as a response to increasing disturbance and fire frequencies in the wetter parts of the warm biomes (Linder and Rudall, 2005). Together, sedges and grasses encompass about 80% of all C₄ species, but represent only a minority of C₄ origins, the majority of which (about 60%) occurred in eudicots (Sage et al., 2011a). Some of these eudicot groups serve as model systems for the study of C₄ genetics and evolution, including members of the Cleomaceae, Amaranthaceae, and, especially, Flaveria in the Asteraceae (Kadereit et al., 2003; Svensson et al., 2003; Brown et al., 2005; McKown et al., 2005; McKown and Dengler, 2007). It has long been postulated that C₄ eudicots were much younger than C₄ monocots, based mainly on the physiological models described above, but also on the presence of C₃–C₄ intermediates and the lower number of C₄ species in eudicot groups (Ehleringer et al., 1997; Sage, 2004). Some well-resolved and relatively well-sampled phylogenies are now available for several groups of C₄ eudicots (McKown et al., 2005; Kapralov et al., 2006; Sage et al., 2007; Feodorova et al., 2010; Kadereit et al., 2010; Ocampo and Columbus, 2010; Christin et al., 2011). However, accurate time calibration of these phylogenies has been hampered by the lack of fossil records for the studied groups.

The first attempt to date C₄ eudicots used a phylogeny based on rbcL for Amaranthaceae/Chenopodiaceae, which suggested that the first C₄ origins in this group could have occurred more than 14.5 Ma (Kadereit et al., 2003), challenging the hypothesis of a Pleistocene origin of C₄ eudicots. This potential early origin of C₄ Amaranthaceae/Chenopodiaceae has been confirmed by more densely sampled phylogenetic analyses (Kadereit et al., 2010; Kadereit and Freitag, 2011). Very recently, two more studies have used time estimates obtained for the major angiosperm lineages to calibrate phylogenies for Molluginaceae and Cleomaceae (Feodorova et al., 2010; Christin et al., 2011). Surprisingly, these C₄ groups with very few species (two, and less than five, respectively) were found to be up to 10 million years (My) old. These new lines of evidence demonstrate the need for a re-evaluation of the timing of C₄ evolution in eudicots and their relationship with past fluctuation of atmospheric CO₂ levels.

The goal of the present study was to estimate the potential ages of the different eudicot C₄ lineages described in the literature. First, a phylogenetetic tree containing the major eudicot clades was inferred and dated and then these time estimates were used to calibrate a series of smaller and more detailed phylogenies inferred from fast-evolving markers. The resultant time-calibrated phylogenetic framework contains many of the postulated C₄ origins in eudicots and is used to address variation in the probability of C₄ evolution through geological time. This study also identified groups for which phylogenetic information is limited or completely absent, but are critical for understanding the timing, evolvability, and reversibility of transitions between C₃ and C₄ photosynthesis in angiosperms.

Materials and methods

General methodology and eudicot phylogeny

The GenBank database was screened for genes available for different C₄ lineages (according to Sage et al., 2011a). Non-coding
genes, such as nuclear internal transcribed spacers (ITS) and intergenic spacers from the plastid genome, are available for numerous C₄ groups. Unfortunately, these fast-evolving markers cannot be unambiguously aligned among distantly related plant groups. Therefore, slower-evolving coding markers were selected that allowed the largest number of C₄ taxa to be incorporated. The two markers selected upon were genes from the plastid genome encoding matK and rbcL. They have been used in various attempts to reconstruct the angiosperm phylogeny (Soltis et al., 2000; Hilu et al., 2003) and are relatively well sampled for several groups that contain C₃/C₄ transitions (Cuénot et al., 2002; Kadereit et al., 2003; Sage et al., 2007; Christin et al., 2011). These two markers were retrieved for species spanning the main eudicot lineages, the purported sister group of eudicots (Ceratophyllum) and one monocot, which was used to root the tree (Acorus). Clades previously reported to contain C₃ taxa were densely sampled and, where possible, the closest C₃ relatives of each C₄ group was included in order to improve the estimates of divergence times. This sampling was completed by de novo sequencing of matK and rbcL for several Caryophyllales species in order to improve the resolution of the phylogeny and, in particular, the relationships between C₃ and C₄ taxa (see Supplementary Data S1 at JXB online). These markers were amplified and sequenced with the primers developed in Christin et al. (2011), following the procedure described therein. Sequences were aligned with ClustalW (Thompson et al., 1994) and the alignment was then manually edited. A phylogenetic tree was inferred simultaneously from rbcL and matK through a Bayesian procedure implemented in MrBayes 3.2 (Ronquist and Huelsenbeck, 2003). The substitution model used for the analyses was a general time-reversible model with a gamma shape parameter and a proportion of invariant sites (GTR+G+I). Two analyses, each consisting of four parallel chains, were run for 7 000 000 generations after a burn-in period of 3 000 000, and a tree was sampled each 1000 generations. The consensus phylogeny resulting from the 14 000 sampled trees was used for molecular dating with Model parameters were estimated with baseasl (Yang, 2007) for the two genes separately. Branch lengths and the variance–covariance matrix were then optimized using estbranches (Thorne et al., 1998). A Bayesian MCMC procedure implemented in multidivtime (Kishino et al., 2001; Thorne and Kishino, 2002) approximated the posterior distributions of substitution rates and divergence times, given a set of time constraints. The MCMC procedure run for 1 000 000 generations after a burn-in of 100 000 generations, with a sampling frequency of 100 generations. The outgroup (Acorus americanus) was removed during the analysis. Calibration points, based on reported fossils, were set as described in Christin et al. (2011), with lower bounds on the stems of Buxales at 102.2 Ma, of Malpighiales at 91.2 Ma, of Fabales at 59.9 Ma, of Malvales at 69.7 Ma, of Myrtales at 88.2 Ma, of Ericales at 91.2 Ma, a lower bound on the divergence of Polycarpon from the higher Caryophyllaceae at 34 Ma, a lower bound on the stem of eudicots at 120 Ma and an upper bound on the crown of eudicots at 130 Ma. This eudicot phylogeny gave relatively good estimates for several C₄ origins, but many groups were poorly represented. For this reason, additional phylogenies were reconstructed with fast-evolving markers for several groups containing C₄ taxa. Since good fossils are not available for most groups at this taxonomic scale, time estimates for nodes of the eudicot phylogeny, together with their associated confidence interval, were used to calibrate these lower scale phylogenies. With this approach, the ages estimated for the different groups are not independent, increasing the probability of accurate estimates of the relative ages, even in the presence of misleading calibration points. For constrained nodes, the bound was set as the mean time minus the standard deviation and the upper bound to the estimated divergence time plus the standard deviation. All trees were inferred and calibrated with the method described above. Specific details for each group are detailed below.

To account for recent suggestions that angiosperms could be older than previously thought (Smith et al., 2010), dating of the eudicot tree was repeated by removing the upper constraint on the crown of eudicots and setting the maximal age of the root (eudicots+Ceratophyllium) to 200 Ma. Divergence times estimated in this analysis were then used to calibrate lower taxonomic scale phylogenies under the hypothesis of an earlier origin of angiosperms.

Detailed analyses of selected clades

The phylogeny for Flaveria was reconstructed using the nuclear ITS and the plastid intergenic spacer trnL-trnF. Sequences for Flaveria were extracted from the dataset of McKown et al. (2005) and other Asteraceae were added to allow for more calibration points. Only one accession per Flaveria species was considered. Scaevola aemula (Goodeniaceae) was used as the outgroup (removed during the dating analysis). The inferred phylogeny was calibrated with the estimated divergence times between Cichorioideae (represented by Lactuca in the eudicot phylogeny) and Asteroideae (i.e. Helianthus and Flaveria) and between Helianthus and Oryzea. The photosynthetic types were determined by McKown et al. (2005) and McKown and Dengler (2007).

A phylogeny for the Chenopodiaceae was inferred from nuclear ITS and the non-coding plastid marker psbB-psbH, from data originating from previous studies (Kapralov et al., 2006; Akhani et al., 2007; Wen et al., 2010). The phylogeny was rooted with Chenopodiacoideae (Atriplex and Chenopodiaceae, removed during the dating analysis), according to previous phylogenies (Kadereit et al., 2003; Kapralov et al., 2006). It was calibrated with the estimated divergence time between Salsoloideae (Salsola, Camphorosma, and Bassia) and Suaedoeideae/Salicornioideae (Bienertia, Allenrollea, and Haloedum). Photosynthetic types were reported from the literature (Jacobs, 2001; Kadereit et al., 2003, and references therein, 2010; Kapralov et al., 2006; Akhani et al., 2007).

A phylogenetic tree for C₄ and related C₃ Nyctaginaceae was inferred from four markers, the nuclear ITS, the plastid coding gene ndhF, and the plastid non-coding genes rpl16 and rps16. These data come from a densely sampled phylogenetic dataset by Douglas and Manos (2007). The phylogenetic tree was rooted with Colignonia scandens, according to Douglas and Manos (2007). It was calibrated with the estimated divergence times between Bougainvillea glabra and the other species and between the Mirabilis genus and the group composed of Allionia/Okenia. Photosynthetic types were deduced from carbon isotope ratio measurements (RF Sage, unpublished data).

A phylogeny was inferred using ITS for Sesuvioideae (data generated by Hassan et al., 2005), other Aizoaceae, and Montiaceae (used as the outgroup, removed in the dating analysis). The resulting phylogeny was calibrated using the estimated divergence time between Gelania+Delosperma+Trichodiadema and Sesuvioideae and the estimated divergence time between Galenia and Delosperma+Trichodiadema. Photosynthetic types were reported from previous studies (Sankhla et al., 1975; Kocacinar and Sage, 2003; RF Sage, unpublished data).

GenBank was screened for markers representing Euphorbia species previously identified as C₄. ITS was selected as it was available for a large number of C₄ as well as C₃ Euphorbia. Euphorbia species spread across the phylogeny were selected together with Manihot esculenta and Jatropha curcas (used as the outgroup, removed during the dating analysis). The resulting phylogeny was calibrated with the estimated divergence time between Manihot esculenta and the Euphorbia genus. Photosynthetic types were reported from the literature (Pearcy and Troughton, 1975; Webster et al., 1975; Batanouny et al., 1991; Sage et al., 2011b).
Modelling of photosynthetic transition probabilities

Two models of transition between C₃ and C₄ photosynthesis were optimized on the eudicot phylogeny inferred from rbcL and matK, with branch lengths estimated from molecular markers and incorporating the estimated divergence times. The null model allows different transition rates from C₃ to C₂ and from C₄ to C₃, but these rates are constant through time. The alternative model allows for one change in the transition rates, independently optimizing transition probabilities between C₁ and C₄ before and after a given time threshold (Christin et al., 2008). The two models are nested and can be compared through likelihood ratio tests. Species were coded as C₄ or C₃ (including C₃–C₄ intermediates) and all parameters were estimated from the data, using MLtree software (Christin et al., 2008).

Results

Eudicot phylogeny

The phylogeny inferred from matK and rbcL (see Supplementary Fig. S1A, B at JXB online) is congruent with the known relationships among angiosperms (APG III, 2009; Brockington et al., 2009), and our age estimates for major clades are consistent with recent estimates (see Supplementary Fig. 1A at JXB online; Bell et al., 2010; Moore et al., 2010). However, to account for recent suggestions that angiosperms could be older than previously thought (Smith et al., 2010), the dating analysis was repeated with relaxed constraints on the maximal age of the eudicots and the root. Removing these constraints affected the estimated ages of the major clades, but did not significantly change those for nodes closer to the tips, which included all potential C₄ origins (see Supplementary Table S1 at JXB online).

Our sampling of rbcL and matK incorporated C₄ taxa forming a total of 31 C₄-like or C₄ groups (Table 1). Of the 36 C₃ lineages of eudicots hypothesized by Sage et al. (2011a), eight are missing because data in GenBank were insufficient (Blepharis, Anticharis, Heliotropium, Pectis, Chrysanthellum/Isostigma, Polycarpaea, and two C₄ groups of Cleome). The two C₄ Atriplex, for which only rbcL was available, were not monophyletic, congruent with previous analysis of this marker (Kadereit et al., 2003). However, more recent analyses with other markers and a denser sampling have shown the monophony of the C₄ Atriplex (Kadereit et al., 2010). All other relationships are perfectly congruent with previous phylogenies for C₄ eudicots (Kadereit et al., 2003; Kapralov et al., 2006; Akhani et al., 2007; Sage et al., 2007; Christin et al., 2011).

For each C₄ group, C₄ photosynthesis could have originated at any time between the divergence of the C₄ group and its C₃ sister group (stem group node), and the earliest evidence of diversification within the C₄ group (crown group node; Table 1). Thus, the limited species sampling in our rbcLmatK tree produces large uncertainties about the timing of some C₄ origins. In addition, many groups are only represented by a single species, which makes it impossible to define a minimum age bracket with a crown group. The upper estimate for the C₄ origin in Portulaca is at 9.7 Ma; Table 1). Dates for the Portulacineae suborder (which contains Portulaca) are older than in a previous study that calibrated the phylogeny with the appearance of the Hawaiian islands (Ocampo and Columbus, 2010). C₄ photosynthesis could be 63.5 My old in the Tribulus group (with a minimum age of 13.8 Ma), but the only markers available for these species were rbcL or a small fragment of matK, decreasing the accuracy of the age estimation and leading to very wide confidence intervals (Table 1). All other C₄ origins are estimated to have occurred during the last 30 My (Table 1).

Detailed analyses of selected clades

According to our time estimates, the appearance of C₃–C₄ photosynthesis in Flaveria occurred either once between 3.6 and 3.1 Ma (with a reversal in F. robusta) or twice, between 2.8 Ma and the present in F. sonorense, and between 3.1 and 2.9 in the common ancestor of clades A and B (see Supplementary Fig. S1C at JXB online). This C₃-C₄ type was co-opted twice to evolve a C₄-like trait, between 0.4 Ma and the present in F. brownii and between 1.8 and 1.3 Ma in clade A (Fig. 1). The number of C₄ origins from the C₄-like type are difficult to infer, since different scenarios are equally probable; in all cases, they occurred in the last 2 My (Table 1).

The reconstruction of C₃/C₄ transitions is highly ambiguous in the inferred Chenopodiaceae phylogeny and the most parsimonious scenarios would imply C₄ to C₃ reversals (see Supplementary Fig. S1D at JXB online). Such reversions have been hypothesized (Pyankov et al., 2001), but other studies concluded that the family contains mainly C₄ origins based on anatomical variation (Kadereit et al., 2003, 2010; Kadereit and Freitag, 2011). According to our estimate, the evolution of single-celled C₄ photosynthesis (Edwards et al., 2004) occurred between 20.8 and 2.6 Ma in Bienertia and between 7.7 Ma and the present in Suaeda aralocaspica (Table 1). If two transitions from C₃ to C₄ are assumed in Suaeda, they would have occurred between 9.9 and 7.0 Ma and between 5.6 and 4.5 Ma, respectively. In Salsoloideae, a single origin followed by reversals would be estimated between 28.3 and 26.2 Ma. In the case of multiple C₄ origins in Salsoloideae, these would be spread across the last 25 My (Table 1). The oldest lower estimate for a C₄ origin in the eudicots is in the Caroxyloneae at 22.1 Ma. Dates for C₄ Chenopodiaceae are slightly younger than previous estimates based on phylogenetic trees encompassing only Amaranthaceae/Chenopodiaceae (Kadereit et al., 2003, 2010; Kadereit and Freitag, 2011).

In Nyctaginaceae, C₄ taxa form two groups (see Supplementary Fig. S1E at JXB online). If they correspond to two C₄ origins, the Allionia lineage evolved the C₄ pathway between 6.1 Ma and the present, while C₄ appeared between 4.7 and 2.2 Ma in the Boerhavia/Okenia group (Table 1). The evolutionary history of photosynthetic types in Sesuvioideae is difficult to reconstruct (see Supplementary Fig. S1F at JXB online). The observed distribution of C₄ taxa is compatible with a single C₄ origin, which would
Table 1. Estimated ages (in millions of years, with standard deviation in brackets) of stem and crown nodes for the different photosynthetic transitions, assuming no reversal

<table>
<thead>
<tr>
<th>Clade</th>
<th>Transition</th>
<th>Stem age</th>
<th>Crown age</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portulacaceae</td>
<td>Portulaca</td>
<td>C₃ to C₄</td>
<td>28.8 (5.0)</td>
<td>9.7 (3.3) For details, see Ocampo et al. (2010)</td>
</tr>
<tr>
<td>Molluginaceae</td>
<td>Mollugo cerviana/ Hypeptelis spargulacea</td>
<td>C₃ to C₃–C₄</td>
<td>17.1 (3.5)</td>
<td>7.0 (2.0) For details, see Christin et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>Mollugo cerviana group</td>
<td>C₂–C₂ to C₄</td>
<td>7.0 (2.0)</td>
<td>0.5 (0.4) For details, see Christin et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>Mollugo fragilis group</td>
<td>C₂–C₂ to C₄</td>
<td>5.3 (1.6)</td>
<td>1.4 (0.6) For details, see Christin et al. (2011)</td>
</tr>
<tr>
<td>Gisekiaceae</td>
<td>Gisekiia</td>
<td>C₃ to C₄</td>
<td>4.8 (3.6)</td>
<td>0</td>
</tr>
<tr>
<td>Amaranthaceae</td>
<td>Gomphreneae</td>
<td>C₃ to C₄</td>
<td>8.6 (2.5)</td>
<td>7.0 (2.2)</td>
</tr>
<tr>
<td></td>
<td>Tidestromia</td>
<td>C₃ to C₄</td>
<td>13.0 (3.3)</td>
<td>3.2 (1.7)</td>
</tr>
<tr>
<td></td>
<td>Alternanthera</td>
<td>C₃ to C₃–C₄</td>
<td>10.3 (2.9)</td>
<td>7.5 (2.4)</td>
</tr>
<tr>
<td></td>
<td>Alternanthera</td>
<td>C₂–C₄ to C₄</td>
<td>7.5 (2.4)</td>
<td>5.9 (2.1)</td>
</tr>
<tr>
<td></td>
<td>Aerva</td>
<td>C₃ to C₄</td>
<td>14.5 (3.7)</td>
<td>2.5 (1.5)</td>
</tr>
<tr>
<td></td>
<td>Amaranthus</td>
<td>C₃ to C₄</td>
<td>15.4 (4.1)</td>
<td>1.8 (1.2)</td>
</tr>
<tr>
<td>Chenopodiaceae</td>
<td>Atriplex</td>
<td>C₃ to C₄</td>
<td>12.9 (2.4)</td>
<td>4.7 (2.3) Recent evidence suggests one C₄ origin in Atriplex 14.1-10.9 Mya (Kadereit et al., 2010)</td>
</tr>
<tr>
<td></td>
<td>Suaeda sect. Salsina</td>
<td>C₃ to C₄</td>
<td>9.9 (3.1)</td>
<td>7.0 (2.4)</td>
</tr>
<tr>
<td></td>
<td>Suaeda sect. Schoberia</td>
<td>C₃ to C₄</td>
<td>5.6 (2.1)</td>
<td>4.5 (1.8)</td>
</tr>
<tr>
<td></td>
<td>Bienertia</td>
<td>C₃ to single-celled C₄</td>
<td>20.8 (3.9)</td>
<td>2.6 (2.0)</td>
</tr>
<tr>
<td></td>
<td>Suaeda aralocaspica</td>
<td>C₃ to single-celled C₄</td>
<td>7.7 (3.2)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Halosarcia indica</td>
<td>C₃ to C₄</td>
<td>6.5 (3.9)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Camphorosmeae</td>
<td>C₃ to C₄</td>
<td>16.7 (4.6)</td>
<td>13.1 (4.1) Recent evidence suggests two C₄ origins, 21.3-8.3 and 13.8-9.8 Mya (Kadereit and Freitag, 2011)</td>
</tr>
<tr>
<td></td>
<td>Caroxyloneae</td>
<td>C₃ to C₄</td>
<td>24.6 (3.4)</td>
<td>22.1 (3.6)</td>
</tr>
<tr>
<td></td>
<td>Salsola kali group</td>
<td>C₃ to C₄</td>
<td>22.9 (3.5)</td>
<td>17.9 (4.0)</td>
</tr>
<tr>
<td></td>
<td>Halothamnus</td>
<td>C₃ to C₄</td>
<td>21.2 (3.5)</td>
<td>7.2 (3.3)</td>
</tr>
<tr>
<td></td>
<td>Haloxylon/Anabis</td>
<td>C₃ to C₄</td>
<td>13.7 (2.9)</td>
<td>12.7 (2.8)</td>
</tr>
<tr>
<td>Nyctaginaceae</td>
<td>Boerhavia</td>
<td>C₃ to C₄</td>
<td>4.7 (0.9)</td>
<td>2.2 (0.5)</td>
</tr>
<tr>
<td></td>
<td>Allonia</td>
<td>C₃ to C₄</td>
<td>6.1 (1.0)</td>
<td>0</td>
</tr>
<tr>
<td>Alizoaceae</td>
<td>Trianthema</td>
<td>C₃ to C₄</td>
<td>22.1 (4.9)</td>
<td>20.2 (4.7) Postulating all Trianthema are C₄ origins</td>
</tr>
<tr>
<td></td>
<td>Zaleya</td>
<td>C₃ to C₄</td>
<td>20.1 (4.8)</td>
<td>4.2 (3.4) Postulating all Zaleya are C₄</td>
</tr>
<tr>
<td></td>
<td>Sesuvium</td>
<td>C₃ to C₄</td>
<td>4.8 (3.3)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Cypselia</td>
<td>C₃ to C₄</td>
<td>11.3 (4.1)</td>
<td>0</td>
</tr>
<tr>
<td>Polygonaceae</td>
<td>Calligonum</td>
<td>C₃ to C₄</td>
<td>19.9 (5.2)</td>
<td>1.2 (1.1)</td>
</tr>
<tr>
<td>Asteraceae</td>
<td>Flaviera</td>
<td>C₃ to C₃–C₄</td>
<td>3.1 (0.9)</td>
<td>2.9 (0.8)</td>
</tr>
<tr>
<td></td>
<td>Flaviera sonorensis</td>
<td>C₃ to C₃–C₄</td>
<td>2.8 (0.8)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Flaviera brownii</td>
<td>C₃–C₄ to C₄-like</td>
<td>0.4 (0.2)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Flaviera clade A1</td>
<td>C₃–C₄ to C₄-like</td>
<td>1.8 (0.6)</td>
<td>1.3 (0.5)</td>
</tr>
<tr>
<td></td>
<td>F. campestris</td>
<td>C₄-like to C₄</td>
<td>0.5 (0.3)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>F. trinervia group</td>
<td>C₄-like to C₄</td>
<td>1.0 (0.4)</td>
<td>0.2 (0.2)</td>
</tr>
<tr>
<td></td>
<td>F. bidentis</td>
<td>C₄-like to C₄</td>
<td>0.6 (0.3)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>F. kochiana</td>
<td>C₄-like to C₄</td>
<td>0.3 (0.2)</td>
<td>0</td>
</tr>
<tr>
<td>Cleomeaceae</td>
<td>Cleome gynandra</td>
<td>C₃ to C₄</td>
<td>17.7 (4.1)</td>
<td>0</td>
</tr>
</tbody>
</table>
The model allowing variation of transition rates through time was marginally better than the null model (log-likelihood=−86.65; χ²=7.49; df=3; P value=0.058). The significance of this alternative model should be re-evaluated when more data are available, but the present result, with limited taxon sampling, suggests that transition probabilities were not constant through time. The optimal threshold for the change of transition probabilities was at 28.8 Ma (Fig. 2), essentially identical to what was estimated for grasses (27.6 Ma; Christin et al., 2008). According to this model, the probability of C₄ evolution was almost zero before 28.8 Ma, but strongly increased after this threshold. After 28.8 Ma, the probability of C₄ to C₃ reversal was positive in eudicots while it was null in grasses (Christin et al., 2008).

### Discussion

**Timing of C₄ origins in the eudicots**

C₄ eudicots have been proposed to be of recent origin compared to C₄ monocots (grasses and sedges), based on arguments such as a low number of C₄ species and the existence of numerous C₃–C₄ intermediates (Ehleringer et al., 1997; Kellogg, 1999; Sage, 2004). A combination of physiological models and estimations of past climatic conditions even led to suggestions of a Pleistocene origin of C₄ eudicots (Ehleringer et al., 1997). While our phylogenetic sampling was limited, which hampered an accurate estimate of the timing of C₄ origins in several eudicot lineages, narrow confidence intervals were still obtained for others. Our analysis provided little evidence for Pleistocene origins of C₄ photosynthesis in the eudicots. The only C₄ origin for which the maximal bound was inferred in the Pleistocene is Flaveria (Table 1). Several origins, within Portulaca, Euphorbia, and most C₄ groups of Aizoaceae, Amaranthaceae, and Chenopodiaceae, are estimated to have occurred in the Miocene or even the late Oligocene, supporting previous estimates (Kadereit et al., 2003, 2010; Ocampo and Columbus, 2010). Excluding the Chloridoideae subfamily of grasses (first estimated C₄ origin 32–25 Ma; Christin et al., 2008), C₄ origins in monocots are also spread throughout the last 25 My and largely overlap with C₄ origins in eudicots (Fig. 3). C₄ eudicots should thus no longer be considered young compared to monocots.

The contemporaneous nature of C₄ monocots and eudicots emphasizes that neither the number of species nor the presence of C₃–C₄ intermediates accurately predicts the age of the different C₄ groups. A recent study suggests that C₃–C₄ intermediacy has been evolutionarily stable for more
than 20 My in some Molluginaceae (Christin et al., 2011) and, according to our present results, the C3–C4 type evolved at least 7.7 Ma in Alternanthera (Table 1). With regard to the species richness of various C4 clades, the Trianthema genus contains fewer than 20 species, but could have appeared more than 20 Ma. Conversely, Cyperus and affiliated taxa in sedges acquired the C4 pathway at less than 11 Ma but encompass more than 550 C4 species (Bruhl and Wilson, 2007; Besnard et al., 2009). Thus, there appears to be no relationship between the age of a C4 clade and the number of species it contains. The larger number of C4 species in monocots may instead be the consequence of grass and sedge families being prone to high diversification rates, since both lineages also contain highly diversified and widespread groups of C3 species (Bruhl and Wilson, 2007; Edwards and Smith, 2010).

Ecological drivers of C4 photosynthesis in eudicots

Past variations of atmospheric CO2 levels have long been viewed as instrumental in driving the evolution of C4 photosynthesis (Ehleringer et al., 1991, 1997; Sage, 2004). Our results provide further evidence in support of this hypothesis; however, the proposal that C4 eudicots evolved in the Pleistocene, only after CO2 fell to lower levels than promoted C4 evolution in the grasses, is not supported by our data. The atmospheric CO2 concentration is estimated to have drastically decreased around 30 Ma to below current levels (Pagani et al., 2005). Further CO2 declines during the last 10 My are difficult to reconstruct with certainty (Edwards et al., 2010), although recent evidence suggests that atmospheric CO2 fell below 350 ppm between 15 and 8 Ma and reached its lowest levels during glacial episodes in the Pleistocene (Tripati et al., 2009). Several C4 eudicot lineages evolved more than 20 Ma (Table 1), and our modelling of C3/C4 transitions suggested that the probability of C4 evolution increased at the same time in monocots and eudicots, around 28 Ma (Fig. 3). These time estimates are consistent with a global effect of the decline in atmospheric CO2 during the Oligocene, and could explain the widespread C4 origins in many different geographic regions beginning around 30 Ma (Sage et al., 2011a).

The similarity in the timing of C4 origins in eudicots and monocots is inconsistent with the physiologically-based predictions on the estimated lower QY of eudicots. A number of possibilities could explain this discrepancy. First, most C4 species occur in high light environments where QY is not directly limiting. At moderate to high light levels, photoprotective processes such as zeaxanthin quenching of excess light energy become engaged, reducing QY below the maximum values used in interspecies comparisons (Sage and Kubien, 2003). Hence, QY differences between C4 eudicots and monocots would not be directly related to...
carbon gain and fitness. Second, past CO\textsubscript{2} levels cannot be reconstructed with certainty when lower than 500 ppm (Edwards et al., 2010), and post-Oligocene levels could have been below the threshold proposed for both the evolution of C\textsubscript{4} monocot and eudicots, based on QY models (Ehleringer et al., 1997). Third, QY differences between C\textsubscript{4} monocots and eudicots are small, and are not consistent between all groups (Skillman, 2008). While QY of C\textsubscript{4} Amaranthaceae/Chenopodiaceae and Flaveria are clearly lower than those of C\textsubscript{4} grasses analysed by the same authors, some C\textsubscript{4} Euphorbia had higher QY than many C\textsubscript{4} grasses (Ehleringer and Bjorkman, 1977; Ehleringer and Pearcy 1983; Monson et al., 1986). In this light, QY differences among taxonomic groups could have been of limited importance for C\textsubscript{4} evolution. Rather, ecological factors such as heat, aridity, and salinity might have been key determinants of C\textsubscript{4} evolvability once the low CO\textsubscript{2} precondition was met.

Length of the transition from C\textsubscript{3} to C\textsubscript{4}

It has been proposed that the multiple transformations in anatomy and biochemistry required for C\textsubscript{4} photosynthesis have been acquired gradually (Griffiths, 1989; Sage, 2004). The time required for the whole transition from a C\textsubscript{3} ancestor to completely C\textsubscript{4} descendants is difficult to estimate, since the precise moment at which the different C\textsubscript{4} characteristics appeared cannot be easily mapped onto a phylogenetic tree (Christin et al., 2010). One approach is to consider the time that separates the first appearance of C\textsubscript{4} characters from the first fully C\textsubscript{4} node of the same lineage. In Flaveria, the first C\textsubscript{3}-C\textsubscript{4} ancestor emerged between 3.1 and 2.9 Ma, but anatomical preconditions appeared earlier, between 3.6 and 3.1 Ma. The first C\textsubscript{4} plants of Flaveria (if the possibility of reversals is excluded) evolved between 1 and 0.2 Ma (Table 1). This indicates that the transition from a C\textsubscript{3} ancestor to C\textsubscript{4} plants took at least 2 My in this clade (Fig. 1). In Molluginaceae, enlarged bundle-sheath cells evolved more than 20 Ma, probably in a C\textsubscript{3} context, and were co-opted to evolve a C\textsubscript{4} trait that was optimized in the last 1 My (Christin et al., 2011). This suggests that the complete transition from the appearance of C\textsubscript{4} preconditions in a C\textsubscript{3} context to completely developed C\textsubscript{4} plants took at least 15 My (Fig. 4). The presence of anatomical C\textsubscript{4} preconditions in taxa related to C\textsubscript{4} lineages has also been suggested for Heliotropium and Cleome (Marshall et al., 2007; Vogan et al., 2007). Unfortunately, phylogenies are not available for Heliotropium and the taxonomic distribution of these preconditions is uncertain for Cleome (Feodorova et al., 2010).

Conclusions

By adopting a multi-faceted phylogenetic approach, it has been shown that the numerous C\textsubscript{4} origins in eudicots have been spread across the last 30 My, and were contemporaneous with those in monocots. So far, no conclusive evidence of pre-Oligocene C\textsubscript{4} plants have been found by molecular dating or analysis of fossilized organic remains (Edwards et al., 2010, Urban et al., 2010). The period extending from the Oligocene to the present has thus seen an exceptional burst of C\textsubscript{4} origins in numerous distantly related groups of flowering plants (Sage et al., 2011a), suggesting the action of global environmental triggers. The decline of atmospheric CO\textsubscript{2} probably increased the probability of C\textsubscript{4} evolution by establishing photorespiration that inhibited C\textsubscript{3} photosynthesis on the one hand, while on the other creating an internal pool of photorespired CO\textsubscript{2} that could serve as resource for driving evolutionary innovation (Hylton et al., 1988; Sage, 2004). Low CO\textsubscript{2} alone was probably insufficient, as shown by the relative lack of C\textsubscript{4} species in habitats such as moist, shaded forests. Other environmental changes, notably increased seasonality, fire, and aridification possibly enhanced the probability of C\textsubscript{4} evolution by opening up landscapes, reducing competition, and restricting stomatal aperture. Together with anatomical and genetic factors that enhanced C\textsubscript{4} evolvability in some clades, these environmental factors might have been instrumental in driving the repeated origins of C\textsubscript{4} photosynthesis in a variety of angiosperms, and profoundly changing the plant communities of many biomes across the planet.

Supplementary data

Supplementary data can be found at JXB online

Supplementary Document S1. List of sequences used for phylogenetic reconstructions.

Supplementary Fig. S1. Calibrated phylogenetic trees.

Supplementary Table S1. Comparison of estimates of stem and crown ages (in millions of years, with standard deviation in brackets) with those obtained with less calibration constraints.

Acknowledgements

This study was funded by the Swiss National Science Foundation grant PBLAP3-129423 and the Marie Curie
IOF 252568 fellowship to PAC, and the National Science Foundation grant DEB-1026611 to EJE.

References


