The Origins of C₄ Grasslands: Integrating Evolutionary and Ecosystem Science

Erika J. Edwards, Colin P. Osborne, Caroline A. E. Strömberg, Stephen A. Smith, C₄ Grasses Consortium

The evolution of grasses using C₄ photosynthesis and their sudden rise to ecological dominance 3 to 8 million years ago is among the most dramatic examples of biome assembly in the geological record. A growing body of work suggests that the patterns and drivers of C₄ grassland expansion were considerably more complex than originally assumed. Previous research has benefited substantially from dialog between geologists and ecologists, but current research must now integrate fully with phylogenetics. A synthesis of grass evolutionary biology with grassland ecosystem science will further our knowledge of the evolution of traits that promote dominance in grassland systems and will provide a new context in which to evaluate the relative importance of C₄ photosynthesis in transforming ecosystems across large regions of Earth.

Photosynthesis is the fundamental biological process that transforms solar energy into the chemical fuel for life by generating sugars from water and CO₂. The ancestral pathway (C₃ photosynthesis) evolved in a CO₂-rich atmosphere more than 2800 million years ago (Ma), but depletion of atmospheric CO₂ about 30 Ma has reduced the efficiency and rate of carbon uptake in many terrestrial plants, especially under high temperatures and water deficits (1). This limitation has been alleviated through the convergent evolution of C₄ photosynthesis in more than 45 independent flowering plant lineages (1). C₄ photosynthesis is a coordinated system of anatomical and physiological traits that concentrate CO₂ around the C₃ photosynthetic machinery, through the use of a solar-powered biochemical cycle. The emergence of ecosystems dominated by C₄ species has transformed the biosphere; although comprising only 3% of vascular plant species (1), they account for some 25% of terrestrial photosynthesis (2).

Sixty percent of C₄ species are grasses, dominating warm-climate grasslands and savannas (Fig. 1A), where their high rates of foliage production sustain Earth’s highest levels of herbivore consumption (3). Stable carbon isotopic data (δ¹³C) collected over the past 20 years document a worldwide expansion of C₄ grasslands through the displacement of C₃ vegetation during the Late Miocene and Pliocene (3 to 8 Ma) (4). This was a dramatic event of biome evolution in Earth’s history, outpacing the rise to dominance of flowering plants during the Cretaceous by one order of magnitude (5), but its drivers are still debated.

The last decade has seen much progress in our understanding of C₄ grass ecophysiology, C₄ grassland ecosystem ecology and geologic history, and the evolutionary history of the C₄ pathway within the grass lineage. However, these independent strands of research are not well integrated. Here, we outline a framework that explicitly links the evolutionary biology of grasses with the ecology and history of grasslands. We review the current state of knowledge in the field and argue that a shift in emphasis from photosynthetic pathway to broader assemblages of plant traits may be essential for understanding the rise of C₄ grasslands. This boils down to one question: Just how responsible is C₄ photosynthesis for the distribution of C₄ grasslands?

Crossing Environmental Thresholds

Today’s C₄ grasses are mostly confined to low latitudes and altitudes, whereas C₃ species dominate at higher latitudes and elevations (Fig. 1A). These patterns correlate best with temperature, with several classic studies (6) showing the relationship on every continent. Explanations of these gradients have traditionally focused on fundamental physiological differences between C₃ and C₄ photosynthesis. At high temperatures and low atmospheric CO₂, the key C₃ photosynthetic enzyme rubisco fails to completely distinguish CO₂ and O₂. The process of O₂ uptake leads to photorespiration in C₃ plants, resulting in net losses of ≤40% of photosynthetic carbon in today’s low-CO₂ atmosphere (7). C₄ photosynthesis suppresses photorespiration by concentrating CO₂ internally, but this comes with an energetic cost, which exceeds the photorespiratory costs of C₃ photosynthesis at high temperatures.

Fig. 1. (A) Global distribution of forests (green) and woodlands, savannas, and grasslands with a ground cover dominated by either C₄ (orange) or C₃ (yellow) grasses. Cropland (red) and shrubs, desert, bare ground, and ice (beige/brown) are also shown. See the supporting online material (SOM) (49) for a full description of land-cover data and its categorization. White circles with black dots in the center indicate the regions in which the geological history of C₄ grasslands is best described (see Fig. 3). (B) The predicted atmospheric CO₂ and growing-season temperature conditions that favor the growth of C₃ or C₄ grasses, based on the quantum yield of photosynthesis, a measure of the inefficiency caused by photorespiration [adapted from (10)].
high CO2 and low temperatures (7, 8). All else being equal, C4 grasses will therefore outperform C3 grasses below a critical threshold in CO2, the level of which depends on growing-season temperature (Fig. 1B) (7, 8). By saturating rubisco with CO2, the C4 pathway also allows the enzyme to achieve maximum catalytic rates under high-light conditions (9); conversely, the overall C4 advantage is often lost in shaded forest understories, where cool conditions improve the quantum efficiency in C3 species (7).

This functional model forms the central basis for understanding the current distribution of C4 grasses and grasslands (Fig. 1A) and the general absence of C4 grasses from forest understory habitats (10). It explains glacial-interglacial cycles of C4 grassland expansion and contraction (10) and underpins forecasts of future impacts of global change on Earth’s C3-C4 balance (11). The extension of this model to the geological past generates the hypothesis that declining atmospheric CO2 drove the displacement of C3 plants by C4 grasses (4, 8, 10). Because lower temperatures reduce the crucial CO2 threshold for a C4 photosynthetic advantage (Fig. 1B), C4 grasslands should have appeared first in the tropics at 350 to 550 parts per million (ppm) CO2 and then spread to higher latitudes as CO2 declined further (10).

**Reconstructing a Botanical Revolution**

It is probable that tectonic events ultimately drove a major decline in CO2 during the Early Oligocene (12, 13), but the subsequent history of CO2 is less certain. Most CO2 proxy records and model calculations indicate CO2 levels substantially lower than 550 ppm for the past 28 million years (My) (Fig. 2), thus also indicating an uncoupling of C4 grassland expansion from atmospheric CO2 during the Miocene (12). In contrast, two recent studies show CO2 variation close to the C4 crossover threshold during the Middle or Late Miocene (Fig. 2) (14, 15). The assumptions, uncertainties, and imprecision inherent to each CO2 proxy, as well as the range of uncertainty in the CO2 crossover threshold itself, make these alternatives difficult to evaluate. Nevertheless, CO2 does seem to have fallen below the upper bound of the threshold 2 My before the origin of C4 grasslands, provoking a reappraisal of the ecological context and environmental drivers of C4 grassland expansion.

Recent phylogenetic reconstructions show that C4 photosynthesis has evolved multiple times in grasses (16, 17). Time-calibration of these phylogenies using fossilized grass pollen and inflorescences places the earliest probable origin in the Early Oligocene (~30 to 32 Ma) (Fig. 2) and suggests that subsequent origins arose in clusters (for example, in the Middle Miocene). This timing has led researchers to hypothesize that the Early Oligocene drop in CO2 triggered evolution of the C4 pathway (16, 17). However, the proposal is challenged by the discovery of Late Cretaceous microscopic plant silica (phytoliths) diagnostic of grasses (Fig. 2), suggesting that this lineage may be much older than previously thought (18). A recalibration with these fossils would date the earliest C4 grasses to the Middle Eocene (17), a time of warm equable climates and probably high CO2 (Fig. 2). Even more controversial are δ13C records from leaf-wax molecules (n-alkanes) in marine sediments, indicating that C4 photosynthesis existed in Cretaceous land plants (19), albeit not necessarily in grasses.

New palaeontological evidence also reveals crucial information about the Miocene environments that preceded C4 grasslands. Rather than being forested, as initially thought (20), it now appears that landscapes were relatively open. The evolution of ungulate grazers or mixed feeders (feeding on grasses and broad-leaved plants) and pollen data (21) supplemented by new, phytolith-based reconstructions of vegetation (22) document the emergence of savannas or woodlands with predominantly C3 grasses in the Early-Middle Miocene (11 to 24 Ma), several million years before C4 grasslands spread (Fig. 3). This vegetation shift is evident in all of the studied cases, although its timing and pace seem to have varied among regions (Fig. 3).

C4 grasses occurred in the landscape soon after this transition. Phytoliths show that C4 Chloridoideae species were represented in North American grassland communities 19 Ma (Fig. 2). Similarly, δ13C records from fossil soils suggest that C4 grasses contributed 20 to 40% of local vegetation in several regions for many million years before C4 species completely dominated communities (Fig. 3) (23, 24). Spatially detailed sedimentological and isotopic reconstructions of the paleolandscape (25, 26) indicate substantial heterogeneity in vegetation structure, with tree-grass mosaics before and during the C3-C4 shift. C4 grasses seemed to have first invaded drier parts of floodplains, whereas C3 plants preferred moister habitats in topographic lows (25, 26).

The explosive, broadly synchronous Late Miocene–to-Pliocene spread of C4 grasses, originally diagnosed by the δ13C of fossil soil carbonates (20), has since been abundantly documented by δ13C records of ungulate teeth (4, 27) and n-alkanes in soils and marine sediments (24, 28) across many low- to mid-latitude regions (Fig. 3). However, with more C4 proxy data available for each region, it has become clear that the δ13C
records of tooth enamel and soil carbonate record different aspects of C₄ expansion. Evaluating both proxies concurrently allows us to distinguish overall vegetation change, recorded by soil δ¹³C, from changes in herbivore food source, mirrored in tooth enamel δ¹³C. Specifically, data from North America, Siwaliks (Himalayan foreland in Pakistan), Argentina, and Kenya show that particular mammals started feeding largely on C₄ grasses ≥1 My before these grasses became abundant (>50%) in this region originated only 3 Ma (29). The rapid adoption of C₄ grazing in these herbivores may mark the emergence of C₄ dietary specialists in faunal communities or else expansion of C₄ grasses at the expense of C₃ grasses, but not necessarily at the expense of C₃ woody vegetation.

In contrast, the overall C₃-C₄ transition in vegetation (inferred from soil carbonate δ¹³C) seems to have typically been much slower, with little evidence of the latitudinal gradient in C₄ grassland appearance expected from the crossover threshold model. It occurred ~3 My earlier in the Siwaliks and China, compared with temperate North America and tropical Kenya (Fig. 3). The Eastern Mediterranean, North America, and China were located at roughly equivalent paleolatitudes, yet only the latter two became C₄-dominated in the Late Miocene, and at slightly different times and rates. The Eastern Mediterranean remained dominated by C₃ plants throughout the Neogene, although the summer droughts that are now thought to exclude C₄ grasses from this region originated only 3 Ma (30).

The wealth of high-resolution data now available point to a complex Late Miocene–to-Pliocene ecological transition, where the roughly correlated C₃-C₄ shifts across continents differ in many details. The decoupling in time of the transition from forest to C₃ grassland and the later expansion of C₄ grasses indicate that different drivers were probably involved. Similarly, the lag between the origins of C₄ grasses and their subsequent rise to dominance suggest separate triggers. These observations necessitate a reassessment of the factors driving C₄ grassland origins and suggest that the problem is best framed as two related questions: (i) What drove the forest-to-grassland transition? (ii) Why did these grasslands later become C₄-dominated in warm-climate regions?

**What Drove the Forest-to-Grassland Transition?**

The abundance of woody vegetation (i.e., trees and shrubs) versus grasses in modern grasslands is maintained by two primary mechanisms. First, C₃ and C₄ grasses may tolerate soil factors or climatic extremes that limit the establishment or survival of woody plants. Second, grasslands are sustained by fires and herbivory, which limit the recruitment and growth of woody species (31). Improved understanding of these different drivers and their interactions has inspired novel hypotheses about the origins of C₄ grasslands and has revived some old ones. All invoke regional or local factors and focus on increased aridity and/or shifts in disturbance regime. Stable oxygen isotope ratios (δ¹⁸O), sedimentology, and floral records point to the development of seasonal climates with warm-season precipitation in South Asia (20, 32). The onset of a dry season causing intensified fire cycles in this region is therefore favored as a driver of the forest-to-grassland transition and is supported by charcoal records indicating increased occurrence of fire on several continents (29, 33). Based on δ¹⁸O and leaf wax hydrogen isotope ratios (δD), an overall increase in aridity (as opposed to seasonality) has been proposed for both South Asia and East Africa (28, 34, 35), with no explicit role for disturbance. In a third scenario using fossil-soil data from South Asia, East Africa, and North America, dry-adapted C₄ grasses evolved herbivore resistance and/or traits leading to fuel accumulation, allowing them to expand their ecological niche into more mesic habitats, at the expense of trees (36, 37).

A major problem with these Late Miocene and Pliocene scenarios is that they rely on mechanisms that today maintain the grass-to-tree/shrub balance in C₄-dominated ecosystems, whereas the fossil record shows that the (much earlier) forest-to-grassland transition involved mainly C₃ grass species (Fig. 3). Whether similar processes drove the spread of C₃ grasslands is an important question that remains to be tackled (22). Nevertheless, modern ecology illustrates that the abundance of grass versus woody vegetation is controlled by complex sets of factors playing out over different spatial scales. At the regional scale, grasslands are sustained by climate/soil/disturbance interactions, mediated by the traits of the grass flora and herbivore fauna. Locally, topographic effects generate microhabitat variation that supports a mosaic of distinct grass communities with very different trait combinations (38, 39). The general message is clear: Traits that are, at best, only indirectly related to C₄ photosynthesis currently play important roles in allowing grasses to dominate vast areas of natural grassland.

**Why Did Grasslands Become C₄-Dominated?**

Questions of this scale have historically fallen into the realm of ecosystem ecologists and plant physiologists and have rarely received attention from the field of evolutionary biology (40). Our understanding of grass phylogenetics and patterns of C₄ evolution in grasses has improved considerably over the past decade (16, 17, 41). Though obviously im-

---

**Fig. 3.** Neogene record of changes in vegetation structure (from forest vegetation to an environment with substantial amount of grass cover), and of photosynthetic pathway (C₃ versus C₄), in vegetation and in diet of ungulates from various regions (49). Ol., Oligocene.
important for understanding C₄ origins, phylogeny should play an equally large role in evaluating ecological and physiological consequences of the pathway.

Most previous studies comparing C₃ and C₄ grasses have chosen taxa that span the deepest divergences within the grass lineage, comparing members of the Pooidae, a strictly C₃ lineage, with C₄ members of “PACMAD” (an acronym for Panicoideae, Aristidoideae, Chloridoideae, Micrairioideae, Arundinoideae, and Danthonioideae lineages), a large clade of C₃ and C₄ grasses containing ~18 independent origins of the C₄ pathway (Fig. 4). However, Pooidae and PACMAD last shared a common ancestor >50 Ma, allowing both lineages to evolve many differences that will confound any potential C₃-C₄ signal (40). Recent work indicates that PACMAD species tend to be warm-adapted whether or not they are C₄, suggesting that the evolution of cold tolerance in Pooidae may be as important as C₁-C₄ differences in establishing the ecological sorting of grass species along temperature gradients (42).

Therefore, isolating the effects of C₄ photosynthesis on any aspect of grass biology (from biochemistry to ecology) requires comparisons between closely related C₃ and C₄ taxa within the PACMAD clade. For example, studies show that responses to climatic extremes of cold and drought are as important as differing photosynthetic performance in determining the ecological characteristics of C₃ and C₄ subspecies of the grass *Allotetroper semilandata* (43).

Similarly, grouping the multiple, independently derived C₄ lineages within PACMAD into a single C₄ functional type probably masks underlying variation in other traits that could be important at the community and/or ecosystem level. Key differences among C₃ plants have been acknowledged for many years because of the presence of C₄ photosynthetic subtypes distinguished by variation in anatomy, biochemistry, and physiology (44–46). The recurrent evolution of the C₄ pathway in grasses presents another opportunity for variation, as each origin arose within a unique internal (and probably external) environment. Researchers have already documented that independent C₄ lineages exhibit different growth responses to elevated CO₂ (47), and that previously recognized sorting of subtypes along climate gradients is better explained by the sorting of different C₃ lineages (48).

The above studies indicate that seemingly self-evident relationships between C₄ photosynthesis and physiological properties or ecological tolerances become decidedly less certain when viewed within a phylogenetic context. Phylogenies emphasize the rich organismal diversity that is contained within the C₄ grass functional type and also provide the perfect framework for evaluating how different grassland systems have been assembled from these independent lineages. The major challenge in this area is to understand how the C₄ pathway has been functionally integrated into each of these diverse organismal backgrounds, and in turn, how the resulting ecological characteristics of these lineages influence such large-scale processes as biome development and ecosystem function.

**Most Grass Species Do Not Dominate Grasslands**

The complexity of many ecosystems makes it difficult to understand how individual species influence ecosystem function. However, just a handful of species make up the majority of the standing biomass in most grasslands, and species richness is largely contributed by taxa occurring in very low densities. Furthermore, a minority of grasses (~600 out of ~11,000 species) is documented as being ecologically dominant in grasslands (49). These dominant species are phylogenetically clustered, suggesting that certain clades of grasses are more prone than others to evolve traits that promote ecological dominance (Fig. 4). The pattern implies that important characteristics other than the C₄ pathway enable these particular species to become abundant in grasslands. If we assume that palegrasslands had a similar diversity structure as their modern analogs, then understanding the Miocene expansion of C₄-dominated ecosystems hinges on the question of “How did particular grassland dominants (as opposed to C₄ grasses in general) come to occupy such a large fraction of the land surface?” A phylogenetic approach can immediately help to organize new research questions about the evolution of grassland dominants. Below, we highlight key examples of how knowledge from other fields can be integrated within this framework.

**How do environmental niche preferences differ between dominant and nondominant grasses?**

A fundamental question is whether the Late Miocene–Pliocene spread of grasslands resulted from the evolution of new environmental niche tolerances in a handful of C₄ grasses, or whether climatological changes promoted the geographical expansion of particular niches already inhabited by those grasses. The truth is likely to be some combination of both scenarios. We currently have little quantitative information on how environmental niches are distributed within the PACMAD clade. By reconstructing the evolution of environmental niche space across the grass phylogeny (42), we will be able to address when and how key changes occurred. Is the environmental niche relatively conserved across certain clades of PACMAD, or is it evolutionarily labile? Have dominants from different lineages converged on similar niche space when living in the same biogeographical region? Do dominant species always exhibit broader ecological ranges than their nondominant sister taxa?

**How do grass species dominate grasslands?**

Understanding the timing and phylogenetic distribution of C₄ origins in grasses is obviously fundamental to understanding the rise of C₄ grasslands. The same is true for other traits that promote species’ dominance, highlighted in the rich ecological literature on competitive interactions in grasslands (50). Multiple comparisons of dominant and subdominant sister taxa will help to identify traits associated with shifts to dominance, and broader surveys of how those traits are distributed across PACMAD will improve our understanding of their evolutionary history. In fire-driven grasslands of African monsoonal climate regions, for instance, dominant grasses typically have well-protected buds and storage reserves, and they respout quickly after defoliation (31). Other traits act to induce fire: High wet-season growth rates and slow rates of leaf decomposition result in large fuel accumulation (31). When and where
in grasses did these traits evolve, and in what order? Dating the appearance of such important ecological traits, in addition to shifts in photosynthetic pathway, may bridge the time gap between the origins of C3 grasses and the origins of C4 grasslands.

**What role does C4 photosynthesis play in the evolution of dominance?** Clearly, the C4 pathway must fundamentally influence the ability of PACMAD grasses to establish dominance, as evidenced by the dearth of C3 PACMAD species with high abundances in grasslands (16 of 145 species sampled for Fig. 4). However, C4 photosynthesis will also probably confer different advantages in different circumstances. For instance, in the fire-adapted example above, the C4 pathway would facilitate biomass accumulation by supporting high photosynthetic rates and nitrogen-use efficiencies, especially in the high-light environment after a fire (37, 46). In contrast, the North American short-grass prairies are water-limited and far less productive than tropical savannas; in this case, the higher water-use efficiency afforded by C4 metabolism might provide the competitive edge (37). These two ecosystem types are dominated by different C4 PACMAD lineages. In African Andropogonaeae, C4 works within an organismal context of a large, fire-adapted plant that quickly accumulates biomass between fires. In North American Chloridoideae, C4 operates within a short-statured, drought-resistant plant. C4 photosynthesis is a well-integrated component of each of these strategies, resulting in two grassland systems that are dominated by C4 grasses, but for very different reasons.

Establishing these patterns of trait assembly for modern grasses could permit specific ecological inferences to be made about C3/C4 lineages that are recognized in the fossil record using phytoliths and other plant fossils. Such inferences will allow detailed reconstruction of the grass communities present in the Miocene and, by inference, the type of environment that may have promoted C4 grass expansion in different regions (37).

### Research Priorities

The current acceleration of computing power and molecular-sequence accumulation makes a completely sampled grass phylogeny a realistic near-term goal, which will greatly facilitate a lineage-centered focus on the C4 grassland problem. Linking the grass phylogeny to databases detailing climatic, ecological, morphological, and physiological information for individual grass taxa is already possible, permitting the application of tools developed for ecological bioinformatics and spatial ecology. Similarly, reconstructions of turnover in paleocommunities are becoming much more refined, using phytolith data to track major grass lineages through time. This growing body of paleobotanical evidence also promises the ongoing improvement in dating of key events in grass evolutionary history. To keep pace with these developments, comparative biological investigations of grass species are urgently required to understand trait evolution in grassland dominants. A stronger synthesis between evolutionary processes, plant function, and ecosystem composition will provide new opportunities for sharpening our hypotheses about the factors that drove the Late Miocene–to-Pliocene emergence of C4 grassland biomes.

### References and Notes

5. R. Lupia, S. Lupidg, P. R. Cran, Paleobiology 25, 305 (1999).
22. E. A. Kellogg,15 Alan K. Knapp,16 Andrew D. B. Leakey,17 Elizabeth A. Kellogg,1,6 John D. Kingston, L. L. Jacobs,16 Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland. 21The Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland. 22School of Biological Sciences, Washington State University, Pullman, WA 99164, USA. 23Department of Biological and Ecology Evolution, University of Sheffield, Sheffield S10 2TN, UK. 24Department of Biology and Burke Museum, University of Washington, Seattle, WA 98195–1800, USA. 25Department of Geology and Geophysics, University of California, Santa Cruz, CA 95064, USA. 26Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA. 27Department of Biology, University of Missouri, St. Louis, MO 63121, USA. 28Department of Biology, Colorado State University, Fort Collins, CO 80523, USA. 29Department of Biology, Strathclyde University, 2455 S.55 (2003). 30Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095, USA. 31Department of Biology, University of Missouri, St. Louis, MO 63121, USA. 32Department of Biology, Colorado State University, Fort Collins, CO 80523, USA. 33Institute for Genomic Research, University of Maryland, Center for Environmental Science Appalachian Laboratory, Frostburg, MD 21532, USA. 34Canadian Museum of Nature, Research Division, Ottawa, Ontario K1P 6P4, Canada. 35Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario M5S 3B2, Canada. 36The Swiss Institute of Bioinformatics, Quartier Sorge, 1015 Lausanne, Switzerland. 37Department of Geography and Institute for Computational Earth System Science, University of California, Santa Barbara, CA 93106, USA. 38Department of Biology, University of Utah, Salt Lake City, UT 84112, USA. 39Supporting Online Material

www.sciencemag.org/cgi/content/full/318/5798/587/DC1 Materials and Methods

SDM Table

Table S1

References

10.1126/science.1177216