

Global analysis of Poales diversification – parallel evolution in space and time into open and closed habitats

Tammy L Elliott*, Department of Botany and Zoology, Faculty of Science, Masaryk University, Kotlarska 2, Brno 611 37, Czech Republic; Department of Biological Sciences, University of Cape Town, Cape Town 7700, South Africa, 0000-0003-2162-8537,

Daniel Spalink*, Department of Ecology and Conservation Biology, Texas A&M University, College Station, Texas, USA,

Isabel Larridon, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK; Systematic and Evolutionary Botany Lab, Department of Biology, Ghent University, K.L. Ledeganckstraat 35, 9000 Gent, Belgium, 0000-0003-0285-722X,

Alexandre Rizzo Zuntini, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK, 0000-0003-0705-8902,

Marcial Escudero, Department of Plant Biology and Ecology, Faculty of Biology, University of Seville, Reina Mercedes 6, Seville 41012, Spain,

Jan Hackel; Department of Biology, University of Marburg, Karl-von-Frisch-Str. 8, 35043 Marburg, Germany; Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK,

Russell L Barrett, National Herbarium of New South Wales, Botanic Gardens of Sydney, Australian Botanic Garden, Locked Bag 6002, Mount Annan, New South Wales 2567, Australia, 0000-0003-0360-8321,

Santiago Martín-Bravo, Botany area, Department of Molecular Biology and Biochemical Engineering, Universidad Pablo de Olavide, ctra. de Utrera km 1, 41013, Seville, Spain, 0000-0003-0626-0770,

José Ignacio Márquez-Corro, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK; Botany area, Department of Molecular Biology and Biochemical Engineering, Universidad Pablo de Olavide, ctra. de Utrera km 1, 41013, Seville, Spain, 0000-0003-4277-2933,

Carolina Granados Mendoza, Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico, 0000-0003-4001-619X,

Aluoneswi C. Mashau, Foundational Research and Services, South African National Biodiversity Institute (SANBI), Private Bag X101, Pretoria, 0184, South Africa, 0000-0001-6028-7064,

Katya J Romero-Soler, Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico, 0000-0002-5075-9862,

Daniel A. Zhigila, Department of Botany, Gombe State University, Tudun Wada, Gombe, Nigeria, 0000-0002-5529-5964,

Berit Gehrke, Universitetet i Bergen, Universitetsmuseet, Postboks 7800, NO-5020 Bergen, Norway, 0000-0001-5866-4430,

Caroline Oliveira Andrino, Departamento de Botânica, Universidade de Brasília, Brasília, Distrito Federal, Brazil, 0000-0003-1107-5692,

Darren M Crayn, Sir Robert Norman Building (E2), James Cook University, PO Box 6811, Cairns QLD 4870, Australia, 0000-0001-6614-4216,

Maria S Vorontsova, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK,

Félix Forest, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK, 0000-0002-2004-433X,

William J Baker, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK, 0000-0001-6727-1831,

Karen L Wilson, National Herbarium of New South Wales, Botanic Gardens of Sydney, Australian Botanic Garden, Locked Bag 6002, Mount Annan, New South Wales 2567, Australia, 0000-0001-7419-8222,

David A Simpson, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AE, UK; Botany Department, School of Natural Sciences, Trinity College, The University of Dublin, Dublin 2, Ireland,

A Muthama Muasya, Department of Biological Sciences, University of Cape Town, Cape Town 7700, South Africa, 0000-0002-0763-0780

correspondence: muthama.muasya@uct.ac.za

*equal contribution as lead authors

Summary

- Poales are one of the most species-rich, ecologically and economically important orders of plants and often characterise open habitats, enabled by unique suites of traits. We test the hypotheses that Poales species are assembled into distinct phyloregions, with centres of high phylogenetic diversity and endemism clustered in tropical regions, and that cosmopolitan families show parallel transitions into open and closed habitats at different times.
- We sampled 42% of Poales species and obtained taxonomic and biogeographic data from the World Checklist of Vascular Plants database, which was combined with open/closed habitat data scored by taxonomic experts. A dated supertree of Poales was constructed. We integrated spatial phylogenetics with regionalization analyses, historical biogeography, ancestral state estimations, and models of contingent evolution.
- Diversification in Poales and assembly of open and closed habitats result from dynamic evolutionary processes that vary across lineages, time, space, and traits, most prominently in tropical and southern latitudes. Our results reveal parallel and recurrent patterns of habitat and trait transitions in the species-rich families Poaceae and Cyperaceae, yet other smaller families display unique evolutionary trajectories.
- The Poales have achieved global dominance via parallel evolution in open habitats, with notable, spatially and phylogenetically restricted divergences into strictly closed habitats.

5–8 key words: biogeography, evolution, evolutionary transitions, grass-like plants, grasslands, savannas

Introduction

Open and closed habitats are functionally distinct categories of terrestrial environments (Bond, 2019), each with unique ecology and differing in ground layer vegetation strongly constrained by available light (Ratnam *et al.*, 2011; Bond, 2022). Open habitats, comprising nearly 60% of land area (Dinerstein *et al.*, 2017), include treetops that provide suitable habitats for epiphytes, areas too extreme to support trees (e.g., deserts, high altitudes, tundra, rock walls), and grasslands – regions that could potentially support trees but are dominated by grass-like plants because of disturbances (Strömberg, 2011; Buisson *et al.*, 2022). Grasslands occupy a wide biogeographic range and cover nearly 40% of the terrestrial biosphere, including tropical and subtropical savannas, boreal and temperate prairies, and Eurasian steppes (Bond, 2019; Buisson *et al.*, 2022). They provide habitat for a wide diversity of animals and plants, and support the livelihood of over one billion people worldwide (Buisson *et al.*, 2022). Until now, biogeographical and evolutionary studies of open habitats have focussed on regional scales (e.g., Solofondranohatra *et al.*, 2020) or on the most diverse family occupying these habitats, Poaceae (e.g., Edwards *et al.*, 2010; Linder *et al.*, 2018; Gallaher *et al.*, 2022), although other closely related families can be equally important components of their respective habitats (e.g., Cyperaceae; Barrett, 2013).

The ability of plants to occur in either open or closed habitat has evolved repetitively across the tree of life, in some cases possibly through parallel evolution – the repeated evolution of similar traits in closely-related lineages (Bailey *et al.*, 2017; Durán-Castillo *et al.*, 2022) – or through divergent evolution involving large evolutionary distances between populations (or species) in genotypic or phenotypic space (Bolnick *et al.*, 2018). As the morphological and physiological traits of plants are linked to their habitat preferences, the repetitive evolution of traits suggests parallel adaptive evolution (Bolnick *et al.*, 2018). To identify macroevolutionary processes, such as whether evolution is parallel or divergent, it is important to determine the appropriate phylogenetic scope to avoid excluding key lineages that unveil important patterns (Folk *et al.*, 2018). Large comparative studies including thousands of species across a single higher-level lineage present opportunity to distinguish which traits might be exceptional at smaller compared to larger phylogenetic scales, as well as incorporate all main clades into the study, including those where the evolution of a trait is typical compared to most of the other lineages (Beaulieu & O’Meara, 2018).



Figure 1. Representatives of twelve Poales families. **(a)** *Tillandsia tovarensis* (Bromeliaceae), epiphytic in cloud forest, Kuelap, N Peru. **(b)** Insect-pollinated *Rhynchospora colorata* (Cyperaceae), in forest gaps, S Ecuador. **(c)** *Ecdeiocolea rigens* (Ecdeiocoleaceae), arid heath, SW Australia. **(d)** *Paepalanthus ensifolius* (Eriocaulaceae), in cloud forest, Podocarpus National Park, S Ecuador. **(e)** *Flagellaria indica* (Flagellariaceae), rocky savannah, NW Australia. **(f)** *Mayaca fluviatilis* (Mayacaceae), wetlands, Singapore. **(g)** *Micraira* sp. Purnululu (Poaceae), a rapid-resurrection species from sandstone pavements in NW Australia. **(h)** *Guacamaya superba* (Rapataceae) in a sedge and grass swamp in E Colombia. **(i)** *Lepidobolus preissii* (Restionaceae) from sandy heath, SW Australia. **(j)** *Prionium serratum* (Thurniaceae) in fynbos, Cape Province, South Africa. **(k)** *Sparganium japonicum* (Typhaceae) from wetlands in E Russia. **(l)** *Xyris complanata* (Xyridaceae) from an ephemeral wetland in savannah, NW Australia. Photos by Russell Barrett; except for **f**, **h**, **j** & **k**, all posted on iNaturalist as CC-BY-NC: **f** by CheongWeei Gan; **h** by Carlos Eduardo; **j** by Linda Hibbin; **k** by Sergei Prokopenko.

Dominant in open habitats, Poales are among the most species-rich orders of angiosperms, comprising 12 families and 24,302 species (Govaerts, 2022; APG IV, 2016; Fig. 1; Table 1), and including two of the three largest monocot families (Poaceae and Cyperaceae). The 12 Poales families differ in their species richness and distribution, yet plant families exhibiting high species richness tend to be cosmopolitan and varying in life form, such as having both woody and herbaceous species (Ricklefs & Renner, 1994). Although the two largest families are predominately herbaceous, they are cosmopolitan and among the ten most species-rich plant families (Govaerts, 2022). The mid-sized families show latitudinal bias (e.g., Restionaceae in austral temperate regions; Bromeliaceae in the Neotropics) or are restricted to a single continent (e.g., Rapateaceae in South America). The ancestor of Poales is hypothesised to have occupied open and seasonally dry habitats during the Cretaceous (Bouchenak-Khelladi *et al.*, 2014) and would have been herbaceous, with typical monocot features including rhizomes, monocarpic tillers, C₃ photosynthesis, and wind pollination (Dahlgren *et al.*, 1985). Several lineages of Poales have entered closed habitats since the Late Cretaceous, but such lineages are predominantly species-poor (e.g., Flagellariaceae), with the exception of Bromeliaceae which has rapidly diversified during the Miocene (Givnish *et al.*, 2011, 2014). Conversely, several species-rich Poales lineages have diversified in open habitats, apparently in conjunction with Crassulacean Acid Metabolism (CAM) and C₄ photosynthetic systems (Bouchenak-Khelladi *et al.*, 2014). This generalisation, however, is based on incomplete sampling (below 2% of species; Bouchenak-Khelladi *et al.*, 2014), which is likely to obscure underlying patterns and processes, and thus remains to be tested with more comprehensive sampling. From an ancestor occurring in open habitats, multiple lineages of Poales have evolved traits possibly in parallel or divergently to occupy distinct niches.

In recent decades, spatial phylogenetic analyses (Mishler *et al.*, 2014) have been increasingly used to evaluate the geographic structure of lineage assembly (Holt *et al.*, 2013; Thornhill *et al.*, 2016; Zhang *et al.*, 2022), to link evolutionary processes to the manifestation of biodiversity through time and space (Carter *et al.*, 2022; Nitta *et al.*, 2022; Sanbonmatsu & Spalink, 2022), and for the development of conservation policies (Sechrest *et al.*, 2002; Gonzalez-Orozco *et al.*, 2014; Zhang *et al.*, 2021). While many spatial phylogenetic metrics exist, emerging methods can identify hotspots of remarkable phylogenetic diversity and distinguish between areas of significant palaeo- and neo-endemism (Mishler *et al.*, 2014). Such analyses provide insight into the spatial patterns of evolutionary processes, integrating

phylogenetic information (i.e., branch lengths and species relationships) with the ranges of species and clades. When coupled with functional traits, this inclusion of evolutionary information provides a more complete understanding of macroecological processes (Tucker *et al.*, 2017; Spalink *et al.*, 2018).

To understand how Poales diversity has accumulated in space and time in the context of open and closed habitats, we focus on the six largest families (Bromeliaceae, Cyperaceae, Eriocaulaceae, Juncaceae, Poaceae, and Restionaceae) to link spatial phylogenetics with models of historical biogeography, phylogenetic regionalization, and ancestral state estimations. We investigate where and when each family diversified across continents, how these processes contribute to phylogenetic regionalization, the spatial and temporal patterns of diversification into open and closed habitats, and whether these patterns demonstrate parallel or divergent evolution. We expect patterns of regionalization to exhibit strong partitioning between tropical and temperate zones, reflecting a tropical origin and select temperate diversification of clades. In addition, we predict that evolutionary transitions between open and closed habitats have occurred frequently, rapidly, and in parallel in Poales, as three of the largest families (Poaceae, Cyperaceae, and Bromeliaceae) have diversified in both habitats. We also expect that evolutionary transitions to, and diversification within, open habitats has occurred in synchrony around the world and across lineages, with the rapid accumulation of species beginning in the Eocene as open habitats became more widespread (Strömberg, 2011). The manifestation of these processes on spatial patterns of diversity should result in significantly high phylogenetic diversity and endemism across the tropics, where most Poalean families have a hypothesised origin, and significantly low endemism in northern temperate habitats. Across Poales, we hypothesise that centres of neo-endemism occur on young islands and mountainous regions, whereas palaeo-endemism should occur in areas with high bioclimatic stability since the Eocene. Finally, we expect family-level spatial phylogenetic patterns to reflect the historical biogeographic context of each clade, with partitioning of neo- and palaeo-endemism dependent on when and where each family diversified.

Materials and Methods

Phylogenomic backbone reconstruction

We produced a family-level phylogenomic backbone using nuclear data from 353 loci (Angiosperms353; Johnson *et al.*, 2019). The sampling for the backbone aimed towards 50% of the currently accepted genera and involved new data produced and samples mined from public repositories. The genomic data production was conducted following Baker *et al.* (2022), with DNA extractions (mostly from herbarium materials) using CTAB (Doyle & Doyle, 1987). We used the NEBNext Ultra II DNA Library Prep kit (New England Biolabs) for standard pair-ended library preparation and libraries were hybridised with myBaits Angiosperms353 v1 probe kit (Arbor Biosciences).

The sequence recovery from raw data (target enrichment and mined reads) started with reads being trimmed for short and/or low-quality sequences using Trimmomatic (Bolger *et al.*, 2014) and then assembled with a *de novo* approach implemented in HybPiper v.1.3.1 (Johnson *et al.*, 2016). In HybPiper, trimmed reads were initially binned into genes using BLASTN, which were assembled into scaffolds using SPADes (Bankevich *et al.*, 2012), and the coding regions later extracted with Exonerate (Slater & Birney, 2005). For assembled datasets (i.e., whole genomes and transcriptomes) sequence recovery followed Baker *et al.* (2022). Only coding regions were used for the analysis.

We inferred the phylogenomic backbone using a multi-species coalescent framework (MSC). For that, we computed individual gene trees, each computed as follows. The sequences were aligned in MAFFT (Katoh & Standley, 2013) in *insi*-mode, with gappy sites (> 90% missing data) removed using Phyutility (Smith & Dunn, 2008). The gene tree was inferred using IQ-TREE 2 (Minh *et al.*, 2020), with support assessed via UltraFast bootstrap (UFBS; Hoang *et al.*, 2018). Once all gene trees were computed, we used TreeShrink (Mai & Mirarab, 2018) to identify outliers that significantly increased the tree space. For the gene trees with outliers, a new iteration of sequence alignment, trimming and tree inference was performed as described above for the remaining sequences. Once the second iteration of gene tree inference was completed, all gene trees were trimmed for poorly supported branches (UFBS < 30%) and used as input for the MSC analysis in ASTRAL-III (Zhang *et al.*, 2018). In order to obtain a species tree with branch lengths proportional to the genetic distance, we first ranked the genes according to the congruence of their resulting trees to the species tree using SortaDate (Smith *et al.*, 2018) and then concatenated the alignments of the 25 most

congruent genes. Using the MSC species tree as topological constraint and this concatenated alignment, a new phylogram was inferred in IQ-Tree 2. For more details on library preparation and data analyses, please refer to Baker *et al.* (2022).

Reconstruction of species-level phylogeny

We inferred separate trees for five groups of families identified with the backbone phylogeny: 1) Bromeliaceae + Typhaceae; 2) Rapateaceae + Thurniaceae + Juncaceae + Cyperaceae + Mayacaceae (the cyperid clade); 3) Xyridaceae + Eriocaulaceae (the xyrid clade); 4) Restionaceae; and 5) Joinvilleaceae + Flagellariaceae + Ecdeiocoleaceae + Poaceae. Available sequences for different marker sets (Table S1) were retrieved from GenBank (Benson *et al.*, 2010). In general, we downloaded as many sequences as possible per marker while disregarding those corresponding to hybrids and not determined to species level, with the exception of the cyperid clade where we preferentially used sequences included in the phylogenetic reconstruction of Elliott *et al.* (2022). Reconciliation of the GenBank names for sequences with those in the ‘Plants of the World Online’ (POWO, 2022) database was done using the R package *TAXIZE* (Chamberlain & Szocs, 2013). We selected one sequence per marker per taxa, while considering synonymy. Sequence matrices for the five clades were aligned using MAFFT v.7.453 (Kato & Standley, 2013), edited manually in AliView v.1.26 (Larsson, 2014) and concatenated. Trees were then inferred under a GTRCAT model in RAxML v.8.2.12 (Stamatakis, 2014), using rapid bootstrapping with 100 replicates followed by a thorough maximum likelihood search on the Czech National Grid Infrastructure. The preliminary phylogenetic trees were then manually checked for obviously spurious species placements, as well as nomenclatural abnormalities. At this point, we verified that species nomenclature aligned with the ‘World Checklist of Vascular Plants’ (WCVP; Govaerts, 2022). Corrected matrices were then realigned with outgroup taxa (Table S1) and manually edited. A second set of phylogenetic constructions were then conducted using constraint trees derived from the backbone tree.

The backbone and clade trees were time-calibrated using penalized likelihood as implemented in treePL, with the smoothing value identified through cross-validation (Smith & O’Meara, 2012). We used secondary calibration points for the backbone tree, setting fixed ages for the family crown nodes (Table S2) obtained from the plastome tree of Givnish *et al.* (2018). Separate analyses were conducted for Poaceae and Cyperaceae, incorporating

additional fossil and secondary priors within the clades using best available information (Spalink *et al.*, 2016; Gallaher *et al.*, 2022; see Table S2). For the remaining clade trees, the root ages were set to one. All subtrees were then grafted into the backbone tree after scaling their age to that of the corresponding nodes in the dated backbone tree.

Species distributions and habitats

We used the WCVF dataset to create a species presence/absence matrix in each botanical country (Level 3) as specified by the International Working Group on Taxonomic Databases for Plant Sciences (TDWG; Brummitt *et al.*, 2001). Species not present in the phylogeny were removed from the matrix prior to downstream analyses. Occurrence in open, closed or both habitats was scored from the literature, especially regional floras, in addition to taxonomic expertise. Habitats noted as ‘forests’ in floristic treatments, interpreted as closed canopy at least during the active growing season, were scored as ‘closed’. We scored species occurring in open and closed habitats as ‘both’. For Bromeliaceae, species occurring on sun-exposed rock-walls or treetops were scored as ‘open’. A full list of taxa included in the study and their habitat is available in Table S3.

Regionalization

To determine phylogenetic patterns of regionalization in Poales, we first calculated phylogenetic beta diversity across all botanical countries. We then used three metrics to identify the optimal number of phyloregions: K-means and silhouette scores (Hartigan & Wong, 1979; Rousseeuw, 1987), the elbow method (Salvador & Chan, 2004; Zhao *et al.*, 2011), and the gap statistic (Tibshirani *et al.*, 2001). For each analysis, we set the maximum number of phyloregions to 20, allowing for the possibility of high levels of regionalization. The elbow and K-means approach each identified 13 phyloregions as optimal, whereas the gap statistic identified 20 phyloregions. For simplicity, we used 13 phyloregions for downstream comparative analyses, with regions partitioned to the optimal number identified above using non-metric multidimensional scaling (NMDS) and hierarchical dendrogram clustering based on a beta-diversity similarity matrix of botanical countries. To determine nestedness of the 13 regions, we completed a final analysis that limited the number of phyloregions to three. Regionalization analyses were conducted using the R packages

phyloregion v.1.0.6 (Daru *et al.*, 2017, 2020a, 2020b), *cluster* v.2.1.3 (Maechler *et al.*, 2022) and *stats* v.4.1.2 (R Core Team, 2021). To summarise the spatial and temporal patterns of lineage assembly, we used the R package *phytools* v.1.2.0 (Revell, 2012) to construct lineage through time plots for the open and closed species of each family in each phyloregion.

Spatial phylogenetics

We used the R packages *picante* v.1.8.2 (Kembel *et al.*, 2010) and *canaper* v.0.0.2 (Nitta & Iwasaki, 2021) to calculate phylogenetic diversity (PD; Faith, 1992) and phylogenetic endemism (PE; Rosauer *et al.*, 2009), as well as to perform categorical analysis of neo- and palaeo-endemism (CANAPE; Mishler *et al.*, 2014) across botanical countries. To calculate the standardised effect size of PD and PE, we used two different sequential randomization algorithms – swap and curveball. The swap algorithm shuffles species presence/absence across the community matrix, while maintaining species ranges sizes and species richness within communities (Gotelli & Entsminger, 2003). The curveball algorithm identifies the species that occur in only one of two randomly selected communities, then distributes those species across communities, while maintaining species richness in communities (Strona *et al.*, 2014). For both randomizations, we performed 1,000 replications of 100,000 iterations each. These were followed by two-tailed tests to determine if PD and PE were significantly over-dispersed or clustered. CANAPE methods follow those outlined in Mishler *et al.* (2014) and Nitta & Iwasaki (2021). For PD, PE and CANAPE, analyses were conducted for all Poales combined, as well as separately for the six largest families (Bromeliaceae, Cyperaceae, Eriocaulaceae, Juncaceae, Poaceae and Restionaceae). In addition, PD was calculated for both open and closed habitat species.

Ancestral state estimations

The ancestral area estimation provides biogeographical context for lineage shifts between open and closed habitats. Our geographical unit of comparison is the Poales phyloregions identified above, but simplified from 13 regions to 7 to reduce the state-space explored by the model (see inset map in Fig. 4). We *a priori* selected the dispersal-extinction-cladogenesis (DEC) model of ancestral estimation as the model with the most appropriate parameters for our study system (Ree *et al.*, 2005; Ree & Smith, 2008; see Notes S1 for justification), which

we implemented with stochastic mapping using the R package *BioGeoBEARS* v1.1.2 (Matzke 2013, 2014). For comparison, we also ran the full suite of DEC, DIVA-like, and BAYAREA-like models available in *BioGeoBEARS*, with and without the additional J parameter.

Stochastic maps were summarised to estimate the frequency of dispersals between bioregions.

To estimate ancestral states of open/closed habitat and the transition rates between them, we used Generalised Hidden Markov models, as implemented in the function *corHMM* in the R package *corHMM* v.2.8 (Boyko & Beaulieu, 2021). We ran "symmetric rates" and "all rates differ" Markov models separately with one transition regime and with two transition regimes associated with two hidden states fitted during the analyses. Akaike information criterion (AIC) was used to select the best fitting model, and ancestral states in the phylogeny were inferred with maximum likelihood and stochastic mapping. We then linked the output of the best-fitting *corHMM* models to the DEC analysis to summarise spatial patterns of open/closed habitat transitions, identifying the nodes at which character states shifted, and the geographical areas that those ancestors were inferred to have occupied based on the DEC analysis.

Results

We inferred a phylogenetic supertree encompassing 10,100 species of Poales (41.6%; Table 1). All families and 88% of genera in the order were represented, but the percentage of species sampled varied from highest in Restionaceae (82%, 453 species), medium for the largest families Poaceae (45%, 5341 species) and Cyperaceae (46%, 2690 species), to lowest in Xyridaceae (3%, 13 species). The majority of species studied occur in open habitats (vs. 14% in closed). Open/closed habitat scoring was carried out for all species, but with some gaps where information was not available (Table S3; > 80% of taxa scored per family).

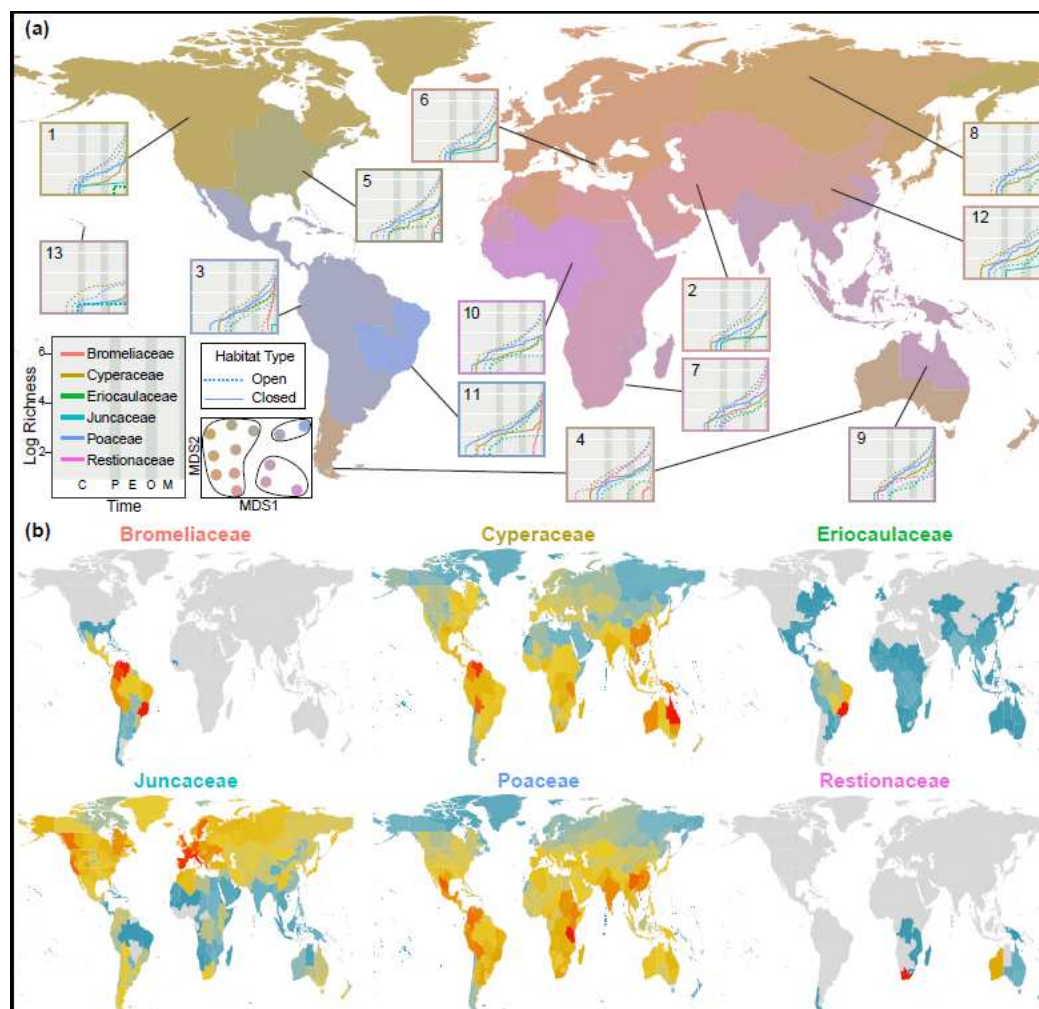


Figure 2. (a) Phylogenetic regionalization of the six largest families of Poales. Graphs represent the logarithmic richness of the six families through time within each of the 13 phyloregions identified using the elbow and K-means approach, where dotted lines represent open habitat lineages and solid lines indicate closed habitat lineages. A geological timeline is located in the inset in the far left-hand corner, with the following abbreviations: Cretaceous (C: 66–56 Mya); Paleocene (P: 66–56 Mya); Eocene (E: 56–33.9 Mya); Oligocene (O: 33.9–23 Mya) and Miocene (M: 23–5.3 Mya). Inset to the right of the geological timeline shows nestedness of the 13 regions based on non-metric multidimensional scaling (NMDS), where each colour dot represent a phyloregion and the three clusters each with similar colour dots represent the three major botanical kingdoms (Temperate, Neotropics, Palaeotropics). **(b)** Phylogenetic diversity (PD) of the six most speciose families of Poales, with red and dark blue indicating botanical countries with relatively high and low PD values, respectively. Botanical regions indicated by grey in **(b)** lack species for the respective family.

Regionalization and lineages through time

Phylogenetic regionalization identified 13 regions (Fig. 2a), which were nested into three major zones (SI Fig. S1): 1) **Temperate** (1: temperate North America + Russian far east; 2: northern Africa [excluding Algeria] + Arabian peninsula [excluding Yemen] + southwestern Asia; 4: Patagonia + Antarctica + western and southern Australia + New Zealand; 5: central and eastern North America; 6: Europe + mediterranean Eurasia + northern Africa [Algeria]; 8: northeastern Asia; 12: central and eastern China; 13: north central Pacific), 2) **Neotropics** (3: Central + South America [excluding central and eastern Brazil]; 11: central and eastern Brazil), and 3) **Palaeotropics** (7: Sub-Saharan Africa [excluding central and western Africa; Madagascar + Mozambique] + southeastern Arabian Peninsula [Yemen] + Mauritania; 9: Indomalayan, comprising India + southeastern Asia + northeastern Australia + Madagascar + Mozambique; 10: central and western tropical Africa [excluding Mauritania]). A NMDS plot shows these 13 regions separated latitudinally, falling into three groups, where northern and southern hemisphere temperate regions form a single cluster separated from the Holotropics, and the latter separates into Neotropics and Palaeotropics (Figs. 2a, S1).

Major families differ in their stem age and lineage accumulation over time among the 13 regions (Fig. 2a). When considering the six largest families, either the Poaceae or Cyperaceae are the oldest in all regions, except for the Restionaceae which are oldest in temperate Australia. Bromeliaceae is the youngest family, whereas the small families (Eriocaulaceae, Juncaceae) are younger than Poaceae and Cyperaceae in regions where they occur. Poaceae and Cyperaceae occur in all regions, Juncaceae is nearly cosmopolitan but is poorly represented in tropical regions, while the other nine families are restricted to one or few regions. Considering the six largest families, three are represented in Asia and northern Europe (Poaceae, Cyperaceae, Juncaceae); five occur in North, Central, and South America (Poaceae, Cyperaceae, Bromeliaceae, Eriocaulaceae, Juncaceae); five are found in the Austral, tropical and subtropical Africa regions (Poaceae, Cyperaceae, Eriocaulaceae, Juncaceae, Restionaceae; Fig. 2a,b).

In all tropical regions, Poales – and specifically Poaceae – diversified first in closed habitats before lineages began accumulating in open habitats (Fig. 2a). By contrast, in northern temperate regions open habitat lineages - specifically, Cyperaceae – diversified first (Fig. 2a). The exception to this latter pattern is in the eastern Nearctic, where closed habitat Poaceae were the first to diversify. However, this exception is likely driven by the inclusion

of some subtropical Caribbean islands and southern Florida in this bioregion. The initial diversification of Poaceae in most temperate regions included both open and closed habitat lineages. Likewise, but in tropical regions only, the initial diversification of open and closed habitat Cyperaceae was concurrent. Open habitat Poaceae displayed the greatest accumulation of lineages (especially during the Oligocene and Miocene) followed by open habitat Cyperaceae in all but the north central Pacific phyloregion, where lineage diversification was more gradual and where there was a big lapse in time until Cyperaceae diversified in closed habitats (Fig. 2a). The most rapid accumulation of lineages occurred in the closed habitat Bromeliaceae. In all bioregions, open habitat species richness is higher than closed habitat species richness in all families except Bromeliaceae.

Spatial Patterns of Diversity

The species in our phylogeny were generally representative of the cosmopolitan distribution of Poales, except for the underrepresentation of tropical areas in South America, India and parts of China (SI Fig. S2). Across Poales and in all families except Bromeliaceae, phylogenetic diversity is substantially higher in open habitats than closed habitats in all bioregions (Figs. 3a, SI Fig. S3), with the highest values attributed to open habitat species in southern Africa, western Australia, and northern South America. High open habitat PD is driven by Cyperaceae, Poaceae, and Restionaceae in southern Africa; Cyperaceae, Poaceae, and Restionaceae in western Australia; and by Bromeliaceae, Cyperaceae, Eriocaulaceae, and Poaceae in northern South America (Figs. 2b, 3; SI Fig. S3). Although closed habitat PD is lower overall than open habitat PD, northern South America is a high centre of diversity of these species (Fig. 3b), with high representation of Bromeliaceae, Cyperaceae, and Poaceae. Despite the widespread availability of forested habitats throughout the Holarctic, these regions have low PD of closed habitat Poales. Those species that do occur in closed habitats in the Holarctic represent only Cyperaceae, Juncaceae, and Poaceae (SI Fig. S3). Apart from Juncaceae, closed habitat PD is higher in all major families south of the Tropic of Cancer (SI Fig. S3).

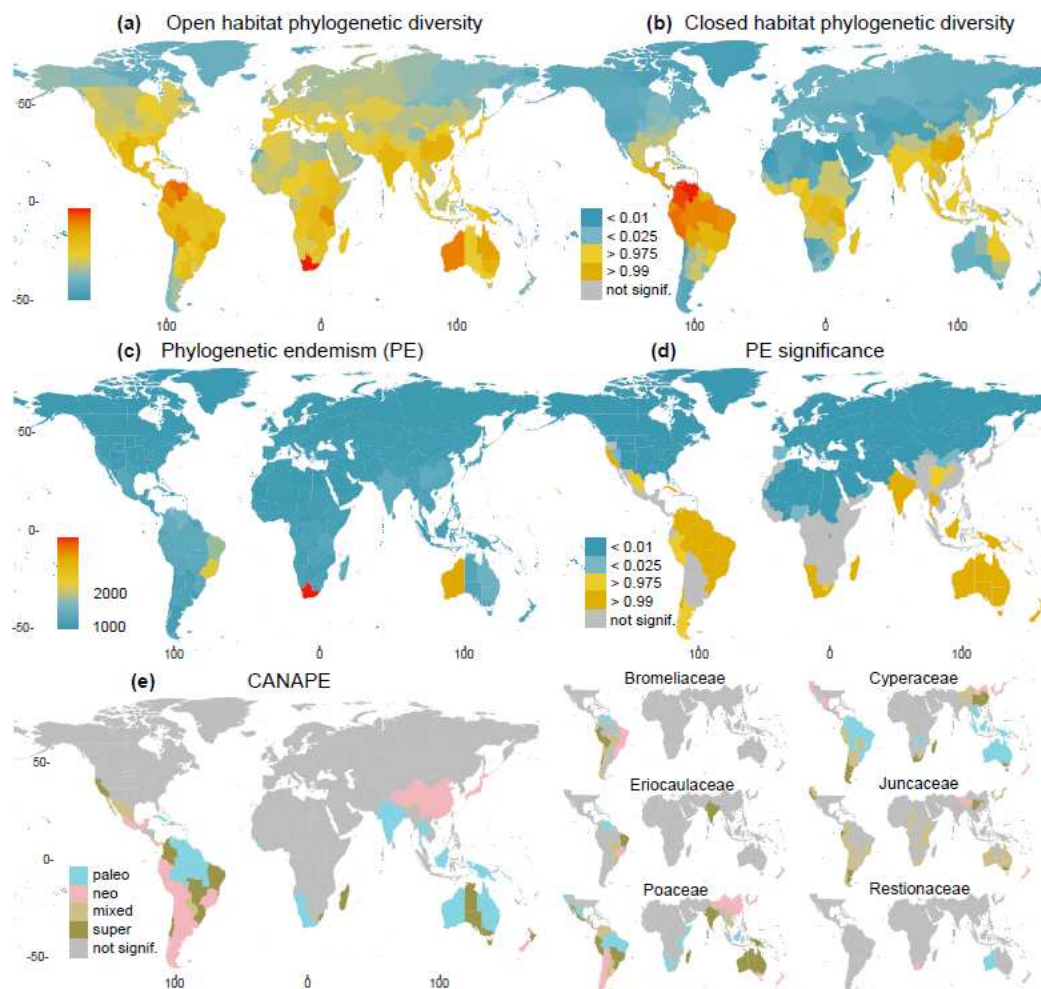


Figure 3. Biodiversity patterns in Poales represented by Faith’s phylogenetic diversity (PD) for open habitat species **(a)** PD for closed habitat species; **(c)** Rosauer’s phylogenetic endemism (PE) and **(d)** PE significance; and **(e)** centres of palaeo- and neo-endemism determined using CANAPE for all Poales (left) and the six most speciose families (right). High PE significance indicates that the region has an overrepresentation of short, rare branches. Low PE significance indicates that short, rare branches are underrepresented.

By far, southern Africa displays the highest PE in Poales (Fig. 3c), driven by high endemism of Cyperaceae, Poaceae, and Restionaceae (SI Fig. S4). PE is also high in Western Australia, again because of restricted lineages from the same three families, and southeastern Brazil, largely due to Bromeliaceae and Eriocaulaceae (Figs. 3c, SI Fig. S4). The Tropic of Cancer clearly demarcates centres of significant PE, with nearly all Holarctic regions exhibiting significantly low PE and many botanical countries in the tropics and subtropics exhibiting significantly high PE. California is the only north temperate zone with

significantly high PE, driven by Poaceae and Juncaceae (Figs. 3c, SI Fig. S4). Overall, family level patterns of significant PE match those of Poales, with low PE in the North and high PE in the South, except for Typhaceae which tends to have low PE in most of its range.

Categorical analysis of endemism (CANAPE; Fig. 3e) shows a general contrasting pattern for Poales between Holarctic (mostly non-significant) and Holotropic + Austral (significant) regions. Palaeo-endemism in Poales was supported in northern South America, southern Africa and Australia – a pattern also shared with Poaceae (including Tanzania, California) and Cyperaceae, whereas palaeo-endemism is only evident in Australia for the Cyperaceae and Restionaceae. Neo-endemism is observed in western South America (Andean mountains) in the Poales and Poaceae, but in the eastern South American region (Brazil) for Bromeliaceae and Eriocaulaceae; in most of East Asia for the order and families Poaceae and Cyperaceae; and in southern Africa for the Restionaceae. A mixture of both palaeo- and neo-endemism (i.e., supra-endemism) is observed in Australia and South America for both the Poales and Poaceae. Though areas of endemism in Poaceae and Cyperaceae often co-occur, frequently centres of palaeo- or neo-endemism in one family are centres of mixed/super endemism in the other (e.g., Andean Mountains, Australia, southeast Brazil, eastern China, southern Africa).

Historical Biogeography and Ancestral Habitat Estimations

The BioGeoBEARS model with the highest AICw was the BayAreaLIKE +J model, followed by DEC without the additional J parameter (Table S4). Given that the BayArea model does not parameterize vicariant speciation but instead allows widespread sympatric speciation, we do not consider it a reasonable model for the global analysis of a clade that spans 120 million years of evolution. Indeed, the BayAreaLIKE model places the Poales in a nearly cosmopolitan range for the first 40 million years of its evolution, which is neither supported by fossil data nor biologically plausible. Thus, subsequent analyses are based on the DEC model. Poales are inferred to have originated in the Neotropics in Western Gondwana (current South America) (Fig. 4a). The crown nodes of most families occur in the Cretaceous (e.g., Cyperaceae, Eriocaulaceae, Juncaceae, Poaceae, and Restionaceae) to the Neogene (e.g., Bromeliaceae) (Fig. 4a). The family crown nodes are distributed in all the major zones (Fig. 4a; Table 1): Bromeliaceae, Cyperaceae, Eriocaulaceae, Rapataceae, and Xyridaceae are assigned a Neotropic origin, while Flagellariaceae and Poaceae are placed in the

Palaeotropics. The Holarctic region is ancestral for the Juncaceae and Typhaceae, whereas the Austral region is ancestral for the Restionaceae. Below family rank, Schoeneae is shown to have originated in the Austral region, *Carex* and Pooideae in Eurasia, Cypereae and the PACMAD clade in Sub-Saharan Africa, Abildgaardieae in Indomalayan region (including NE Australia, Mozambique, and Madagascar) and Bambusoideae in the Neotropics (Fig. 4a). Of these clades, Schoeneae is the oldest originating c. 70 Mya, while *Carex* is the youngest with an origin dated at c. 34 Mya.

Poaceae and Cyperaceae experienced a several million-year lag before co-occurring in any bioregion (Figs. 2, 4). This pattern of concurrent but spatially separated diversification is repeated for main lineages within these families as well. In the Paleocene, the grass clades Bambusoideae, PACMAD, and Pooideae originate allopatrically in the Neotropics, Afrotropics, and Eurasia, respectively. In the Eocene, the sedge clades Abildgaardieae, Cypereae, *Carex* and Schoeneae originate in the Indomalayan tropics, Afrotropics, Palearctic, and Austral region, respectively.

Dispersal across all bioregions has occurred many times throughout the evolution of Poales, most frequently between the Nearctic and Neotropics and between the Palearctic and Paleotropics (Fig. 4b). Eurasia is a strong one-way source to North America, tropical Africa and tropical Asia, while equatorial Africa is a notable recipient of lineages from Sub-Saharan Africa (Fig. 4a,b). A relatively high number of transitions between open and closed habitats have occurred in the Neotropics and Eurasia, while numbers are low for equatorial Africa and the Austral region (Fig. 4a,c). Transitions from open to closed are higher than the opposite for the Neoarctic and Sub-Saharan Africa, whereas closed to open transitions are higher for the Indomalayan region (Fig. 4a,c).

The asymmetrical ("all-rates-different") with two rate regimes is the best fit corHMM model of evolution for open/closed habitats (AIC = 8814.2; Fig. 5; Table S5). An open habitat (Fig. 5) is inferred at the root of Poales and most families, with a slow rate shift to closed habitat (0.000000001 transitions per million years [Mya] in R2 [Regime rate 2]). However, there are rapid rate transitions between open and closed (0.31 transitions / Mya in R1 [Regime rate 1]) and closed to open (0.51 transitions / Mya in R1) habitats among Bromeliaceae, Cyperaceae (*Carex* clade) and Poaceae (Bambusoideae clade), occurring predominantly during the Miocene.

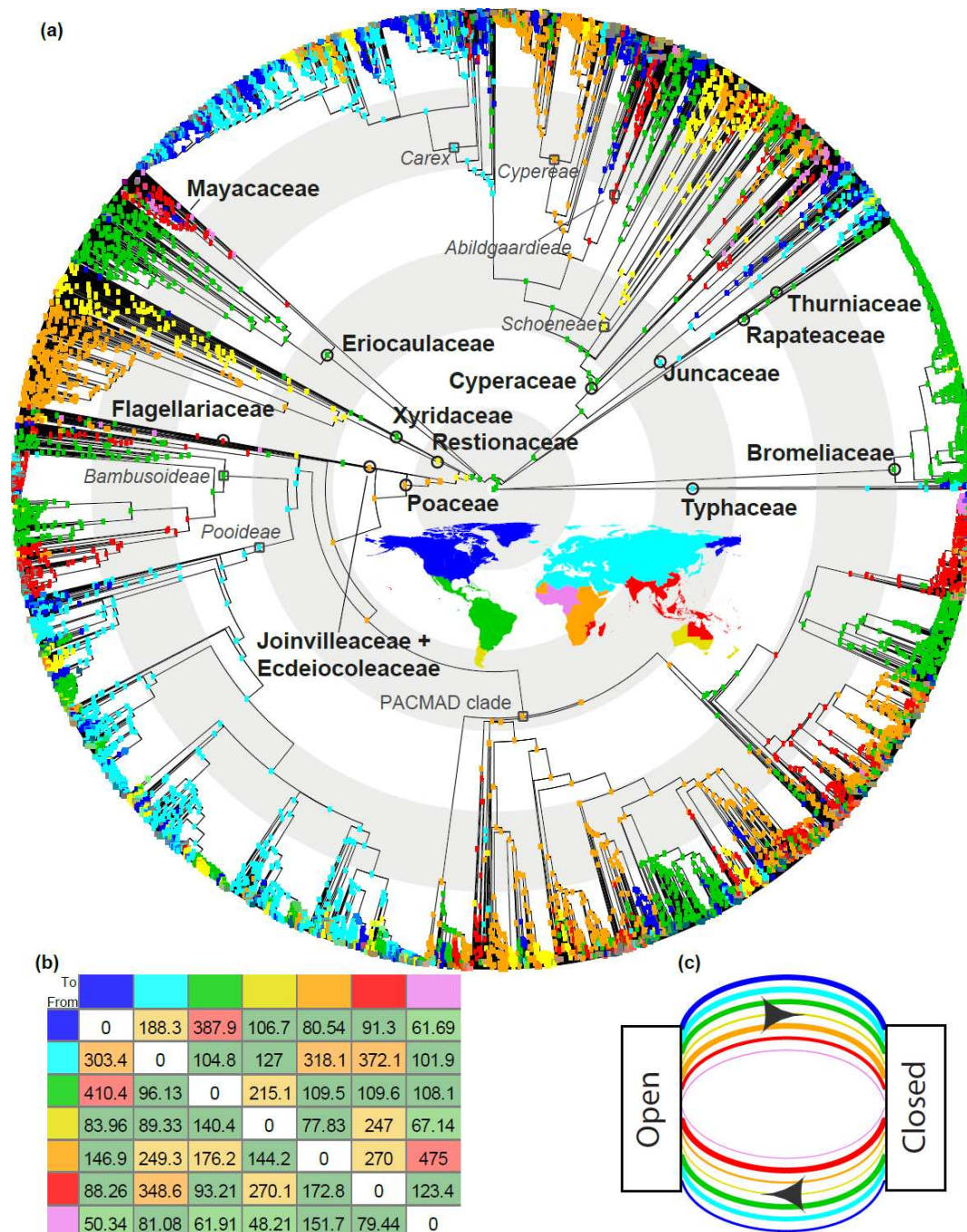


Figure 4. Ancestral area reconstruction within Poales based on seven regions, obtained using the DEC model in BioGeoBEARS (a). A global map showing colours corresponding to the seven defined areas for the BioGeoBEARS analysis is in the centre of the phylogenetic reconstruction. The crown nodes of the families within Poales are shown with black circles, whereas dark grey squares are used to depict lineages important for the study's interpretations. Concentric light grey/white rings underlying the phylogeny indicate time slots of 20 million years intervals. Note that Joinvilleaceae and Ecdeiocoleaceae are depicted together for visual purposes. Number of dispersal events between the seven regions inferred

using Biogeographical Stochastic Mapping on the DEC model **(b)**, and **(c)** number of transitions between open and closed habitats inferred to have occurred within each of the seven areas, calculated through comparison of best fitting corHMM models and historical biogeographical estimates (line thickness indicates relative frequency). The colours of row and column labels in **(b)** and arrows in **(c)** correspond to those on inset map in **(a)**.

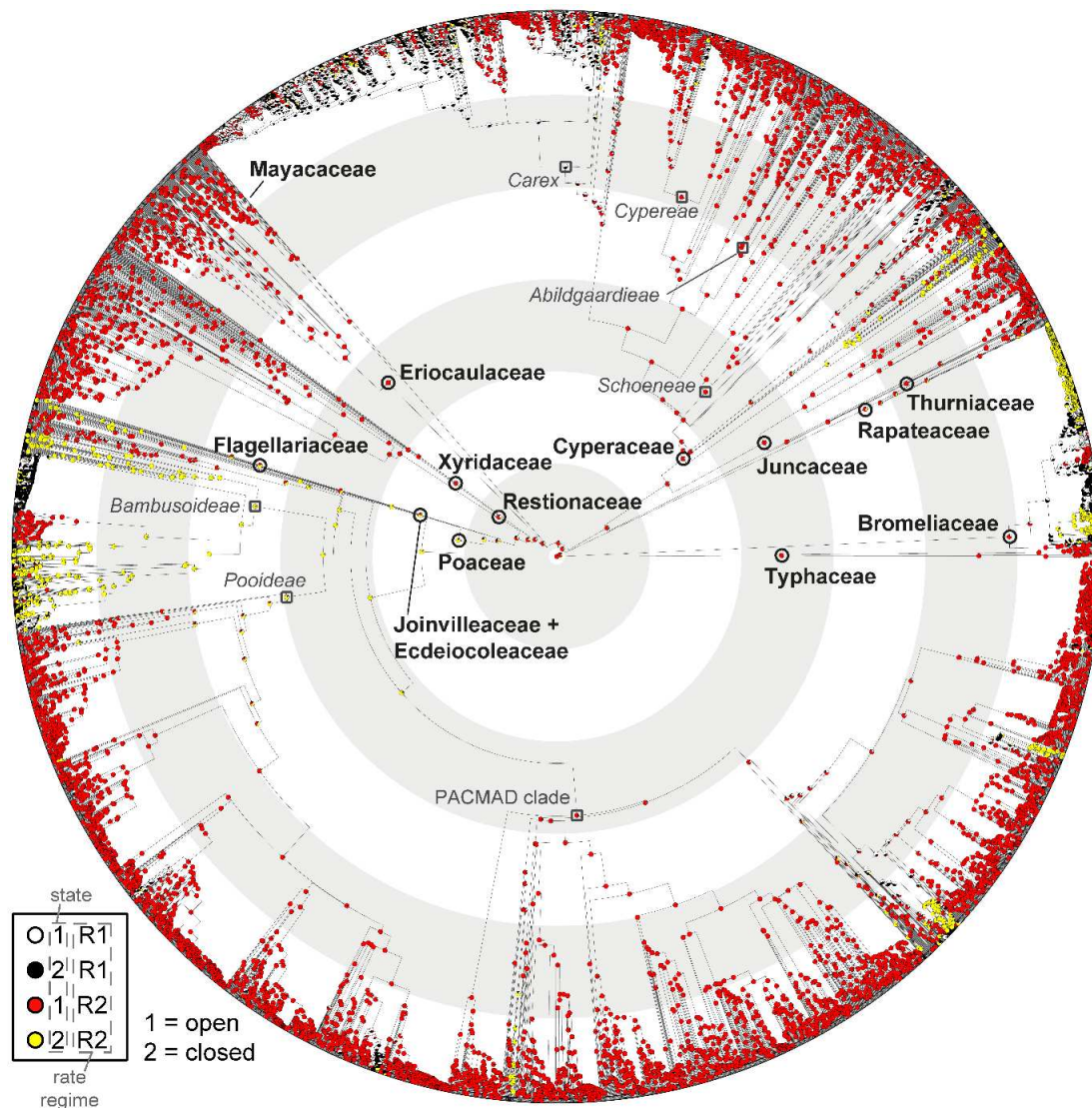


Figure 5. Ancestral states of open/closed habitats based on Generalized Hidden Markov models, with two rate regimes and asymmetric rates, implemented with the corHMM package in R. The crown nodes of the families within Poales are shown with black circles, whereas dark grey squares are used to depict lineages important for the study's interpretations. Concentric light grey/white rings underlying the phylogeny indicate time slots of 20 million years intervals. Detailed transition rates between the states and rate are given in Table S5.

Discussion

This is the most comprehensive evolutionary study of Poales to date, with the goal of understanding the assembly of open habitats around the world. The key outcome is an emerging picture of parallel evolution – in space, through time, across lineages – resulting in the global assembly of open habitats, with notable, spatially and phylogenetically restricted divergences into strictly closed habitats. Our analyses show that each family exhibits a contrasting historical biogeography and pattern of habitat transition, which manifest in unique spatial patterns of assembly. And yet, patterns of evolution and assembly are repeated in most families.

The biogeography and regionalization of Poales

The Poales represent over 120 million years of diversification, with an estimated Cretaceous origin in Western Gondwana (Fig. 4a). Consistent with previous phylogenetic studies, our model places the Poaceae origin in the Afrotropics with eventual migration to the Neotropics in the Eocene (Bouchenak-Khelladi *et al.*, 2010; Gallaher *et al.*, 2022). Meanwhile Cyperaceae originates in the Neotropics and quickly shifts to Patagonia-Australia before landing in Africa in the Paleocene (Spalink *et al.*, 2016; Larridon *et al.*, 2021a). The earliest Cyperaceae fossils also place extinct Cyperaceae lineages in southeast Asia by the Paleogene (Smith *et al.*, 2009). The following global dominance of these Poaceae and Cyperaceae clades is a tale of parallel diversification with repeated spatial convergence of distantly related lineages, which is also observed in transitions between open and closed habitats and has observable impacts on spatial patterns of phylogenetic endemism.

In contrast to Poaceae and Cyperaceae, the other families have experienced more limited diversification, either spatially or in terms of species richness. The oldest of these is Restionaceae, which originated in Western Australia in the Late Cretaceous and dispersed to southern Africa in the Palaeocene (Fig. 4a; Linder *et al.*, 2003). The remarkable diversity of this family of C₃ perennials remains restricted to open habitats in these two regions, where they have evolved strategies to withstand frequent fires (reseeder/resprouter; Litsios *et al.*, 2014) and thrive on oligotrophic soils via enhanced nutrient acquisition strategies (cluster roots; Lambers *et al.*, 2006), traits only shared by the Schoeneae (Cyperaceae; Barrett, 2013)

and lacking in the Poaceae. Typhaceae and Juncaceae, both of Palearctic origin near the KP boundary, are now cosmopolitan but remain relatively species poor. Eriocaulaceae originated in the Neotropics at about the same time, and only a few lineages in the family have diversified outside of this ancestral zone. Bromeliaceae are perhaps the most unique among the Poales, diversifying nearly exclusively in Central and South America and with suites of traits not found in other families, such as epiphytism and lithophytism, which are uncommon traits in the Poales encountered outside the Bromeliaceae among the poikilohydric clade of Cyperaceae (Trilepideae; Muasya *et al.*, 2010; Porembski *et al.*, 2021).

These historical processes have resulted in substantial spatial structure to phylogenetic regionalization in Poales, with lineages forming thirteen distinct phyloregions (Fig. 2a) clustered into three floristic kingdoms (*sensu* Carta *et al.*, 2022; SI Fig. S1). Despite the propensity for long distance dispersal in many Poales groups (Linder *et al.*, 2018; Martín-Bravo *et al.*, 2019; Benítez-Benítez *et al.*, 2021; Larridon *et al.*, 2021a; Gallaher *et al.*, 2022), regionalization in this clade largely mirrors that of all vascular plants, where the tectonic biogeographic legacy of Gondwana and Laurasia persists in determining, at least in part, the distribution of lineages (Takhtajan *et al.*, 1986; Carta *et al.*, 2022). In Poales, this is evident in northern temperate regions being distinct from those in the tropical south, and a united Austral-Patagonian (holantarctic, *sensu* Takhtajan *et al.*, 1986) region. Unlike other vascular plants, temperate zones in both hemispheres share more common lineages than they do with tropical zones, and the Neotropics are phylogenetically distinct from the Paleotropics (SI Fig. S1). The unified North-South Temperate floristic kingdom may be partly due to bipolar disjunctions within species (at least 14) and genera (at least eight) in Poales (Villaverde *et al.*, 2017). Likely, a more important driver of this pattern is the ecological filtering of lineages with respect to traits that evolved earlier in the diversification of Poales. These would include traits like frost hardiness and seasonality tolerance in Juncaceae, Typhaceae, and select Poaceae and Cyperaceae clades (Vigeland *et al.*, 2013; Ambroise *et al.*, 2019; Martín-Bravo *et al.*, 2019; Schubert *et al.*, 2020), and edaphic specificity coupled with functional traits associated with adaptations to temperature and water extremes, as well as disturbance regimes (e.g., C₄, CAM, silica, tannins, epiphytism; Givnish *et al.*, 2011, 2014; Linder *et al.*, 2018) across Poales. The relatively few lineages that have evolved tolerance to frost and strong climatic seasonality have diversified extensively, virtually in all habitable places where these conditions persist (Fig. 2). The phylogenetic differentiation of the Neo- and Palaeo- tropical kingdoms has roots deep in the Poales phylogeny. The Neotropical origin of

Cyperaceae and Afrotropical origin of Poaceae, the near restriction of Eriocaulaceae, Bromeliaceae, and Rapateaceae in the Neotropics, and the restriction of Restionaceae to austral regions (Australia, New Zealand, southern Africa and Patagonia) drive strong phylogenetic beta diversity across the tropics.

Evolutionary parallelism in open and closed habitats

The remarkable parallel evolution of open and closed lineages in Poales is most evident in Poaceae and Cyperaceae, which collectively represent ~74% of species richness in the order and the majority of open-habitat PD (SI Fig. S3). Both families encompass two species-rich lineages (Cyperaceae: *Carex* and Cypereae; Poaceae: PACMAD and Pooideae), each of which (except Pooideae) is inferred to have an open-habitat ancestor (Fig. 5). The exclusively C₃ *Carex* and Pooideae are predominately temperate lineages, likely originating in eastern Asia (Fig. 4; Gallaher *et al.*, 2022; Spalink *et al.*, 2016; Martín-Bravo *et al.*, 2019), while tropical Africa is inferred as the origin for the heavily C₄ PACMAD and Cypereae clades (Figs. 4a, 5). These species-rich lineages evolved from C₃ ancestors, acquiring cold tolerance (*Carex*, Pooideae; Vigeland *et al.*, 2013; Martín-Bravo *et al.*, 2019; Schubert *et al.*, 2020) and C₄ photosynthesis (Cypereae, Abildgaardieae, PACMAD; Fig. 5) in parallel. The tempo of habitat transition has shifted from slow to fast multiple times (Fig. 5), most notably in *Carex* (> 2000 species; Larridon *et al.*, 2021b), Bambusoideae (c. 1700 species; Soreng *et al.*, 2022), and Bromeliaceae (c. 3700 species, Gouda & Butcher, 2023). Bambusoideae uniquely grow tall enough to reach the canopy (Soreng *et al.*, 2015; Attigala *et al.*, 2016). Bromeliaceae likely originated in open habitats, but has rapidly and repeatedly shifted between both habitats (Fig. 5).

Ultimately, our analyses indicate that there was likely a high diversity in all phyloregional species pools from multiple families with beneficial adaptations for survival in open habitats prior to their expansion in the Miocene (Edwards *et al.*, 2010; Veldman *et al.*, 2015). Consistent with the fossil record (Prasad *et al.*, 2005; Vicentini *et al.*, 2008), the origin of PACMAD (encompassing all known origins of C₄ in Poaceae; Edwards, 2019) and Abildgaardieae and Cypereae (majority of C₄ Cyperaceae) is placed in the Palaeocene and Eocene (Fig. 4; Prasad *et al.*, 2011; Gallaher *et al.*, 2022), but most of the diversification in these clades coincides and is possibly linked with the low CO₂ concentrations, cooler, drier conditions, and increased fire activity and herbivory during the Miocene (Osborne &

Freckleton, 2009; Edwards & Smith, 2010; Strömberg, 2011; Maurin *et al.*, 2014; Sage *et al.*, 2018; Peppe *et al.*, 2023). Diversification of most major Poales families increased in all phyloregions during this time, with open-habitat species accumulating most rapidly in the Miocene (Fig. 2).

Spatial phylogenetics of Poales

The spatial structure of Poales phylogenetic diversity at the resolution presented here (Fig. 3, SI Fig. S3–S4) is ultimately the result of the biogeographical and ecological processes rather than local contemporary ecological processes (Meynard *et al.*, 2013; Ross *et al.*, 2021). The resulting diversity patterns are complex (Fig. 3). Each family exhibits peak PD (Fig. 2) in a different botanical country – usually, but not always in their phyloregion of origin (Fig. 4a). The PD of the order as a whole, and each family independently, defies a traditional latitudinal diversity gradient (Kreft & Jetz, 2007). Though this is the dominant trend in open habitat lineages, the PD of closed habitat lineages peaks in equatorial forests and decreases poleward (Fig. 3a,b). These patterns reflect the tremendous diversification of Poales in open habitats around the world (Figs. 2, 4, SI Fig. S3), with much more limited diversification of closed habitat lineages outside of ancestral ranges.

Patterns of diversity in Poales generally align with our stated hypotheses. PD peaks in the tropics, where most families originated. Hotspots of PE (Fig. 3c-e) are almost strictly a southern hemispheric phenomenon, and with the exception of California in North America, never occur north of 33°N at this spatial resolution. This pattern is likely driven both by the presence of spatially restricted families in the Southern Hemisphere (e.g., Bromeliaceae, Restionaceae, and Rapateaceae; Fig. 2b) and the propensity of lineages to migrate across the Holarctic (Donoghue & Smith, 2004).

To our knowledge, this is the first study coupling CANAPE with historical biogeographical estimations on a global scale. We find contrasting patterns of palaeo- and neo-endemism clearly reflect historical biogeographical processes at the family level (Fig. 3e). In Bromeliaceae, centres of palaeoendemism occur in the Guayana Shield – the site of putative origin of the family – whereas centres of neoendemism are restricted to the Brazilian Shield, where Bromelioideae subsequently radiated (Givnish *et al.*, 2011). Eriocaulaceae mirror this pattern (Andrino *et al.*, 2023), with the added area of super-endemism in India,

where nearly 20% of the family occurs in the Western Ghats region alone (Sunil *et al.*, 2015). In Restionaceae, Western Australia, the ancestral home of the family, hosts palaeoendemics, while neoendemics are restricted to South Africa (Linder *et al.*, 2003). And finally, Poaceae and Cyperaceae show high levels of both endemism types around the world, reflecting their rapid migration early in their evolutionary history (Fig. 4) and *in situ* diversification across phyloregions (Fig. 2). In both families, tropical savannas and forested regions primarily harbour palaeoendemics, whereas more recent diversification is found frequently on islands – both young and old – and mountains. Significant palaeoendemism is observed within old landscapes with buffered climates (South America, southern Africa, India, and Australia; Hopper *et al.*, 2016), while significant neoendemism is predominantly within regions that have experienced Miocene geomorphological evolution – especially in Andean (Gregory-Wodzicki, 2000) and Himalayan mountains (Wang *et al.*, 2022). Areas containing a mixture of both neo- and palaeoendemism are observed in regions where old and young landscapes are interspersed (Hopper *et al.*, 2016; Barros-Souza & Borges, 2022; Vasconcelos *et al.*, 2020).

Conclusions

The biogeography of individual families in the Poales is distinct from early in their evolution, such that few major clades originating in approximate synchrony do so in the same places. This is particularly evident in Poaceae and Cyperaceae, whose evolution is like two dancers on a global stage - a biogeo(choreo)graphy - characterised by parallel movement through a shared temporality but with varying tempos, spatial staging, and ecological rhythms, ultimately resulting in a final tableau of global dominance. Both families originate in the late Cretaceous, but on either side of the widening Atlantic Ocean, dispersing and evolving in parallel to achieve cosmopolitan distribution. In contrast, the other families have experienced more limited, spatially restricted diversification. We identify latitudinal and longitudinal patterns of diversification, most evident in the phylogenetic distinctiveness of the temperate regions and also in the neotropics separating from the palaeotropics, with highest endemism found south of the Tropic of Cancer. Parallel evolution is observed in habitat transition, whose tempo has shifted from slow to fast multiple times especially since the Eocene, driven by repeated evolution of traits enabling colonisation of open and closed habitats in different latitudes.

Data availability

The datasets analysed in this study are available as Supporting Information. The new Angiosperms353 raw reads are deposited in the European Nucleotide Archive (PRJEB35285).

Acknowledgements

The authors acknowledge the team who delivered the WCVF on which many of the analyses performed in this study are based. This work was partly funded by grants from the Calleva Foundation to the Plant and Fungal Trees of Life Project (PAFTOL) at the Royal Botanic Gardens, Kew. The authors are thankful for the Research/Scientific Computing teams at The James Hutton Institute and NIAB for providing computational resources used for the phylogenomic analyses performed in the “UK’s Crop Diversity Bioinformatics HPC” (BBSRC grant BB/S019669/1). Computational resources were also supplied by the project ‘e-Infrastruktura CZ’ (e-INFRA CZ ID:90140) supported by the Ministry of Education, Youth and Sports of the Czech Republic. Lizzy Wenk extracted data from AusTraits. This work was supported in part by USDA National Institute of Food and Agriculture (McIntire Stennis Project 1018692) and NSF #1902064 to DS; UNAM–DGAPA–PAPIIT (No. IA202319), “Investigación Científica Básica” CONACYT (No. 286249) to CGM; MICINN-AEI to ME (PID2021-122715NB-I00) and SM-B (PID2020-113897GB-I00). KR thanks DGAPA-UNAM for a postdoctoral grant (2022). ARZ and JH were each funded by a “Future Leader in Plant and Fungal Science” fellowship from the Royal Botanic Gardens, Kew.

Conflict of interest

The authors have no conflicts of interest to declare.

Author contributions

AMM, TLE, DS and IL conceived the study. AMM, TLE, ME, DS, IL, JH, RLB, SM-B, JIM-C, CGM, ACM, KJR-S, DAZ, COA, DC, MSV, KLW and DAS collected data, while

TLE, DS, JH, ARZ, and ME conducted the analyses. AMM, TLE and DS wrote the manuscript with input from all co-authors. JIM-C, DS, ME and RLB created the figures.

References

- Ambroise V, Legay S, Guerriero G, Hausman JF, Cuypers A, Sergeant K. 2020.** The roots of plant frost hardiness and tolerance. *Plant and Cell Physiology* **61**: 3–20.
- Andrino CO, Costa FN, Simon MF, Missagia RV, Sano PT. 2023.** Eriocaulaceae: a new classification system based on morphological evolution and molecular evidence. *Taxon*. <https://doi.org/10.1002/tax.12915>
- APG IV. 2016.** An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* **181**: 1–20.
- Attigala L, Wysocki WP, Duvall MR, Clarke LG. 2016.** Phylogenetic estimation and morphological evolution of Arundinarieae (Bambusoideae: Poaceae) based on plastome phylogenomic analysis. *Molecular Phylogenetics and Evolution* **101**: 111–121.
- Bailey SF, Blanquart F, Bataillon T, Rees K. 2017.** What drives parallel evolution? How population size and mutational variation contribute to repeated evolution. *BioEssays* **39**: 1-9.
- Baker WJ, Bailey P, Barber V, Barker A, Bellot S, Bishop D, Botigué LR, Brewer G, Carruthers T, Clarkson JJ, et al. 2022.** A comprehensive phylogenomic platform for exploring the Angiosperm tree of life. *Systematic Biology* **71**: 301–319.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012.** SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**: 455–477.
- Barrett RL. 2013.** Ecological importance of sedges: a survey of the Australasian Cyperaceae genus *Lepidosperma*. *Annals of Botany* **111**: 499–529.
- Barros-Souza Y, Borges LM. 2022.** Spatial and lineage dependent processes underpin floristic assembly in the megadiverse Eastern South American mountains. *bioRxiv* 2022.05.26.493493; doi:10.1101/2022.05.26.493493

Beaulieu JM, O'Meara BC. 2018. Can we build it? Yes we can, but should we use it? Assessing the quality and value of a very large phylogeny of campanulid angiosperms. *American Journal of Botany* **105**: 417–432.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2010. GenBank. *Nucleic Acids Research* **38**: D46–D51.

Benítez-Benítez C, Martín-Bravo S, Bjorå CS, Gebauer S, Hipp AL, Hoffmann MH, Luceño M, Pedersen TM, Reznicek A, Roalson E, Volkova P. 2021. Geographical vs. ecological diversification in *Carex* section *Phacocystis* (Cyperaceae): patterns hidden behind a twisted taxonomy. *Journal of Systematics and Evolution* **59**: 642–667.

Bolnick DI, Barrett RDH, Oke KB, Rennison DJ, Stuart YE. 2018. (Non) parallel evolution. *Annual Review of Ecology, Evolution, and Systematics* **49**: 303–330.

Bouchenak-Khelladi Y, Verboom GA, Savolainen V, Hodkinson TR. 2010. Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Botanical Journal of the Linnean Society* **162**: 543–557.

Bouchenak-Khelladi Y, Muasya AM, Linder P. 2014. A revised evolutionary history of Poales: Origins and diversification. *Botanical Journal of the Linnaean Society* **175**: 4–16.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bond WJ. 2019. *Open ecosystems: Ecology and evolution beyond the forest edge*. Oxford: Oxford University Press.

Bond WJ. 2022. Out of the shadows: Ecology of open ecosystems. *Plant Ecology & Diversity* **14**: 205–222.

Boyko JD, Beaulieu JM. 2021. Generalized hidden Markov models for phylogenetic comparative datasets. *Methods in Ecology and Evolution* **12**: 468–478.

Brummitt RK, Pando F, Hollis S, Brummitt N. 2001. *World geographical scheme for recording plant distributions*. 2nd ed. Pittsburgh, PA: Hunt Institute for Botanical Documentation, Carnegie-Mellon University.

- Buisson E, Archibald S, Fidelis A, Suding KN. 2022.** Ancient grasslands guide ambitious goals in grassland restoration. *Science* **377**: 594–598.
- Carta A, Peruzzi L, Ramírez-Barahona S. 2022.** A global phylogenetic regionalization of vascular plants reveals a deep split between Gondwanan and Laurasian biotas. *New Phytologist* **233**: 1494–1504.
- Carter BE, Misiewicz TM, Mishler BD. 2022.** Spatial phylogenetic patterns in the North American moss flora are shaped by history and climate. *Journal of Biogeography* **49**:1327–38.
- Chamberlain S, Szocs E. 2013.** taxize - taxonomic search and retrieval in R. F1000Research. <https://f1000research.com/articles/2-191/v2>
- Dahlgren RM, Clifford HT, Yeo PF 1985.** *The families of the monocotyledons*. Berlin: Springer-Verlag.
- Daru BH, Elliott TL, Park DS, Davies TJ. 2017.** Understanding the processes underpinning patterns of phylogenetic regionalization. *Trends in Ecology and Evolution* **32**: 845–860.
- Daru BH, Farooq H, Antonelli A, Faurby S. 2020a.** Endemism patterns are scale dependent. *Nature Communications* **11**: 2115.
- Daru BH, Karunarathne P, Schliep K. 2020b.** phyloregion: R package for biogeographic regionalization and macroecology. *Methods in Ecology and Evolution* **11**: 1483–1491.
- Dinerstein E, Olson D, Joshi A, Vynne C, Burgess ND, Wikramanayake E, Hahn N, Palminteri S. et al. 2017.** An ecoregion-based approach to protecting half the terrestrial realm. *BioScience* **67**: 534–545.
- Donoghue MJ, Smith SA. 2004.** Patterns in the assembly of temperate forests around the Northern Hemisphere. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**: 1633–1644.
- Doyle JJ, Doyle JL. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**: 11–15.

Durán-Castillo M, Hudson A, Wilson Y, Field DL, Twyford AD. 2022. A phylogeny of *Antirrhinum* reveals parallel evolution of alpine morphology. *New Phytologist* **233**: 1426–1439.

Edwards EJ. 2019. Evolutionary trajectories, accessibility and other metaphors: the case of C₄ and CAM photosynthesis. *New Phytologist* **223**: 1742–1755.

Edwards EJ, Smith SA. 2010. Phylogenetic analyses reveal the shady history of C₄ grasses. *Proceedings of the National Academy of Sciences* **107**: 2532–2537.

Edwards EJ, Osborne CP, Strömberg CAE, Smith SA, C₄ Grasses Consortium, Bond WJ, Christin P-A, Cousins AB, Duvall MR, Fox DL, et al. 2010. The origins of C₄ grasslands: integrating evolutionary and ecosystem science. *Science* **328**: 587–591.

Elliott TL, Zedek F, Barrett RL, Bruhl JJ, Escudero M, Hroudová Z, Joly S, Larridon I, Luceño M, Márquez-Corro JI, Martín-Bravo S, et al. 2022. Chromosome size matters: genome evolution in the cyperid clade. *Annals of Botany* **130**: 999–1013.
<https://doi.org/10.1093/aob/mcac136>.

Faith DP. 1992. Conservation evaluation and phylogenetic diversity. *Biological Conservation* **61**: 1–10.

Folk RA, Sun M, Soltis PS, Smith SA, Soltis DE, Guralnick RP. 2018. Challenges of comprehensive taxon sampling in comparative biology: Wrestling with rosids. *American Journal of Botany* **105**: 433–445.

Gallaher TJ, Peterson PM, Soreng RJ, Zuloaga FO, Li DZ, Clark LG, Tyrrell CD, Welker AD, Kellogg EA, Teisher JK. 2022. Grasses through space and time: An overview of the biogeographical and macroevolutionary history of Poaceae. *Journal of Systematics and Evolution* **60**: 522–569.

Givnish TJ, Barfuss MHJ, van Ee B, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily R, Crayn DM, Smith JAC, et al. 2011. Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: Insights from an eight-locus plastid phylogeny. *American Journal of Botany* **98**: 872–895.

Givnish TJ, Barfuss MHJ, Van Ee B, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily RS, Crayn DM, Smith JAC, *et al.* 2014. Adaptive radiation, correlated and contingent evolution, and net species diversification in Bromeliaceae. *Molecular Phylogenetics and Evolution* **71**: 55–78.

Givnish TJ, Zuluaga A, Spalink D, Soto Gomez M, Lam VK, Saarela JM, Sass C, Iles WJ, De Sousa DJ, Leebens-Mack J, *et al.* 2018. Monocot plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of monocots. *American Journal of Botany* **105**: 1888–1910.

Gonzalez-Orozco CE, Pollock LJ, Thornhill AH, Mishler BD, Knerr N, Laffan SW, Miller JT, Rosauer DF, Faith DP, Nipperess DA, Kujala H, *et al.* 2016. Phylogenetic approaches reveal biodiversity threats under climate change. *Nature Climate Change* **6**: 1110–1114.

Gotelli NJ, Entsminger NJ. 2003. Swap algorithms in null model analysis. *Ecology* **84**: 532–535.

Gouda EJ, Butcher D. 2023. The new bromeliad taxon list, version 4. Utrecht: The Netherlands: University Botanic Gardens. Retrieved from <https://bromeliad.nl/taxonlist/>

Govaerts R (ed.). 2022. WCVF: World Checklist of Vascular Plants. Facilitated by the Royal Botanic Gardens, Kew. [WWW document] URL http://sftp.kew.org/pub/data-repositories/WCVF/Special_Issue_28_Feb_2022/ [accessed 28 February 2022].

Gregory-Wodzicki KM. 2000. Uplift history of the Central and Northern Andes: a review. *Geological Society of America Bulletin* **112**: 1091–1105.

Hartigan JA, Wong MA. 1979. Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics* **28**: 100–108.

Holt BG, Lessard JP, Borregaard MK, Fritz SA, Araújo MB, Dimitrov D, Fabre PH, Graham CH, Graves GR, Jönsson KA *et al.* 2013. An update of Wallace’s zoogeographic regions of the world. *Science* **339**: 74–78. <http://dx.doi.org/10.1126/science.1228282>

Hoang DT, Chernomor O, Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution* **35**: 518–522.

Hopper SD, Silveira FA, Fiedler PL. 2016. Biodiversity hotspots and Ocbil theory. *Plant and Soil* **403**: 167–216.

Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* **4**: 1600016.

Johnson MG, Pokorny L, Dodsworth S, Botigue LR, Cowan RS, Devault A, Eiserhardt WL, Epitawalage N, Forest F, Kim JT, et al. 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology* **68**: 594–606.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kembel SW, Cowan PD, Helmus MR, Cornwell WL, Morlon H, Ackerly DD, Blomberg SP, Webb CO. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**: 1463–1464.

Kreft H, Jetz W. 2007. Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences* **104**: 5925–5930.

Larridon I, Spalink D, Jiménez-Mejías P, Márquez-Corro JI, Martín-Bravo S, Muasya M, Escudero M. 2021a. The evolutionary history of sedges (Cyperaceae) in Madagascar. *Journal of Biogeography* **48**: 917–932.

Larridon I, Zuntini AR, Lévillé-Bourret E, Barrett RL, Starr JR, Muasya AM, Villaverde T, Bauters K, Brewer GE, Bruhl JJ, et al. 2021b. A new classification of Cyperaceae (Poales) supported by phylogenomic data. *Journal of Systematics and Evolution* **59**: 852–895.

Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**: 3276–3278.

Linder HP, Eldenas P, Briggs BG. 2003. Contrasting patterns of radiation in African and Australian Restionaceae. *Evolution* **57**: 2688–2702.

Linder HP, Lehmann CER, Archibald S, Osborne CP, Richardson DM. 2018. Global grass (Poaceae) success underpinned by traits facilitating colonization, persistence and habitat transformation. *Biological Reviews* **93**: 1125–1144.

Litsios G, Wüest RO, Kostikova A, Forest F, Lexer C, Linder HP, Pearman PB, Zimmermann NE, Salamin N. 2014. Effects of a fire response trait on diversification in replicated radiations. *Evolution* **68**: 453–465.

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. 2022. cluster: Cluster Analysis Basics and Extensions. R package version 2.1.3. <https://CRAN.R-project.org/package=cluster>

Mai U, Mirarab S. 2018. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics* **19**: 272.

Martín-Bravo S, Jiménez-Mejías P, Villaverde T, Escudero M, Hahn M, Spalink D, Roalson EH, Hipp AL, Global *Carex* Group. 2019. A tale of worldwide success: Behind the scenes of *Carex* (Cyperaceae) biogeography and diversification. *Journal of Systematics and Evolution* **57**: 695–718.

Matzke, NJ. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography* **5**: 242–248.

Matzke, NJ. 2014. Model Selection in Historical Biogeography Reveals that Founder-event Speciation is a Crucial Process in Island Clades. *Systematic Biology* **63**: 951–970.

Maurin O, Davies TJ, Burrows JE, Daru BH, Yessoufou K, Muasya M, van der Bank M, Bond WJ. 2014. Savanna fire and the origins of the ‘underground forests’ of Africa. *New Phytologist* **204**: 201–214.

Meynard CN, Lavergne S, Boulangeat I, Garraud L, Van Es J, Mouquet N, Thuiller W. 2013. Disentangling the drivers of metacommunity structure across spatial scales. *Journal of Biogeography* **40**: 1560–1571.

- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Haeseler A, Lanfear R. 2020.** IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**: 1530–1534.
- Mishler BD, Knerr N, González-Orozco CE, Thornhill AH, Laffan SW, Miller JT. 2014.** Phylogenetic measures of biodiversity and neo- and paleo-endemism in Australian *Acacia*. *Nature Communications* **5**: 4473.
- Muasya AM, Harvey Y, Cheek M, Tah K, Simpson DA. 2010.** *Coleochloa domensis* (Cyperaceae), a new epiphytic species from Cameroon. *Kew Bulletin* **65**: 323–325.
- Nitta JH, Iwasaki W. 2021.** canaper: Categorical analysis of neo- and paleo-endemism in R. doi:10.5281/zenodo.5094032
- Nitta JH, Mishler BD, Iwasaki W, Ebihara A. 2022.** Spatial phylogenetics of Japanese ferns: Patterns, processes, and implications for conservation. *American Journal of Botany* **109**: 727–745.
- Osborne CP, Freckleton RP. 2009.** Ecological selection pressures for C₄ photosynthesis in the grasses. *Proceedings of the Royal Society B: Biological Sciences* **276**: 1753–1760.
- Peppe DJ, Cote SM, Deino AL, Fox DL, Kingston JD, Kinyanjui RN, Lukens WE, MacLachy LM, Novello A, Strömberg CA, Driese SG, et al. 2023.** Oldest evidence of abundant C₄ grasses and habitat heterogeneity in eastern Africa. *Science* **380**: 173–177.
- Porembski S, Rexroth J, Weising K, Bondi L, Mello-Silva R, Centeno DC, Datar MN, Watve A, Thiombano A, Tindano E, Rabarimanarivo MN. 2021.** An overview on desiccation-tolerant mat-forming monocotyledons on tropical inselbergs. *Flora* **285**: 151953.
- POWO. 2022.** Plants of the world online. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; <http://www.plantsoftheworldonline.org/>
- Prasad V, Stromberg CA, Alimohammadian H, Sahni A. 2005.** Dinosaur coprolites and the early evolution of grasses and grazers. *Science* **310**: 1177–1180.
- Prasad V, Strömberg CAE, Leaché AD, Samant B, Patnaik R, Tang L, Mohabey DM, Ge S, Sahni A. 2011.** Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae. *Nature Communications* **2**: 480.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

Ratnam J, Bond WJ, Fensham RJ, Hoffmann WA, Archibald S, Lehmann CER, Anderson MT, Higgins SI, Sankaran M. 2011. When is a ‘forest’ a savanna, and why does it matter? *Global Ecology and Biogeography* **20**: 653–660.

Ree RH, Moore BR, Webb CO, Donoghue MJ. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* **59**: 2299–2311.

Ree RH, Smith SA. 2008. Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systematic Biology* **57**: 4–14.

Revell LJ. 2012. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**: 217–223.

Ricklefs RE, Renner SS. 1994. Species richness within families of flowering plants. *Evolution* **48**: 1619–1636.

Rosauer D, Laffan SW, Crisp MD, Donnellan SC, Cook LG. 2009. Phylogenetic endemism: A new approach for identifying geographical concentrations of evolutionary history. *Molecular Ecology* **18**: 4061–4072.

Ross S-J, Suzuki Y, Kondoh M, Suzuki K, Villa Martín P, Dornelas M. 2021. Illuminating the intrinsic and extrinsic drivers of ecological stability across scales. *Ecological Research* **36**: 364– 378.

Rousseeuw PJ. 1987. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**: 53–65.

Sage RF, Monson RK, Ehleringer JR, Adachi S, Pearcy RW. 2018. Some like it hot: the physiological ecology of C₄ plant evolution. *Oecologia* **187**: 941–966.

Salvador S, Chan P. 2004. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In: *Proceedings of the Sixteenth IEEE International Conference on Tools with Artificial Intelligence*. Piscataway, New Jersey, USA: Institute of Electrical and Electronics Engineers. pp 576–584.

Sanbonmatsu KK, Spalink D. 2022. A global analysis of mosses reveals low phylogenetic endemism and highlights the importance of long-distance dispersal. *Journal of Biogeography* **49**: 654–667.

Sechrest W, Brooks TM, da Fonseca GAB, Konstant WR, Mittermeier RA, Purvis A, Rylands AB, Gittleman JL. 2002. Hotspots and the conservation of evolutionary history. *Proceedings of the National Academy of Sciences USA* **99**: 2067–2071.

Schubert M, Humphreys AM, Lindberg CL, Preston JC, Fjellheim S. 2020. To coldly go where no grass has gone before: a multidisciplinary review of cold adaptation in Poaceae. *Annual Plant Reviews* **3**: 523–562.

Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**: 31.

Smith SA, Dunn CW. 2008. Phyutility: a phyloinformatics tool for trees, alignments and molecular data. *Bioinformatics* **24**: 715–716.

Smith SA, O'Meara BC. 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. *Bioinformatics* **28**: 2689–2690.

Smith SA, Brown JW, Walker JF. 2018. So many genes, so little time: A practical approach to divergence-time estimation in the genomic era. *PLOS ONE* **13**: e0197433. <https://doi.org/10.1371/journal.pone.0197433>

Smith SY, Collinson ME, Simpson DA, Rudall PJ, Marone F, Stampanoni M. 2009. Elucidating the affinities and habitat of ancient, widespread Cyperaceae: *Volkeria messelensis* gen. et sp. nov., a fossil mapanioid sedge from the Eocene of Europe. *American Journal of Botany* **96**: 1506–1518.

Solofondranohatra CL, Vorontsova MS, Hempson GP, Hackel J, Cable S, Vololoniaina J, Lehmann CER. 2020. Fire and grazing determined grasslands of central Madagascar represent ancient assemblages. *Proceedings of the Royal Society B: Biological Sciences* **287**: 20200598.

Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, Filgueiras TS, Davis JI, Morrone O. 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae). *Journal of Systematics and Evolution* **53**: 117–137.

Soreng RJ, Peterson PM, Zuloaga RO, Romaschenko K, Clarke LG, Teisher JK, Gillespie LJ, Barberá P, Welker CAD, Kellogg EA. 2022. A worldwide phylogenetic classification of the Poaceae (Gramineae) III: An update. *Journal of Systematics and Evolution* **60**: 476–521.

Spalink D, Drew BT, Pace MC, Zaborsky JG, Starr JR, Cameron KM, Givnish TJ, Sytsma KJ. 2016. Biogeography of the cosmopolitan sedges (Cyperaceae) and the area- richness correlation in plants. *Journal of Biogeography* **43**: 1893–1904.

Spalink D, Pender J, Escudero M, Hipp AL, Roalson EH, Starr JR, Waterway MJ, Bohs L, Sytsma KJ. 2018. The spatial structure of phylogenetic and functional diversity in the United States and Canada: An example using the sedge family (Cyperaceae). *Journal of Systematics and Evolution* **56**: 449–465.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Strömberg CAE. 2011. Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences* **39**: 517–544.

Strona G, Nappo D, Boccacci F, Fattorini S, San-Miguel-Ayanz J. 2014. A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications* **5**: 4114.

Sunil CN, Ratheesh Narayanan MK, Sivadasan M, Alfarhan AH, Abdul Jaleel V. 2015. *Eriocaulon vandaanamense* sp. nov. (Eriocaulaceae) from Kerala, India. *Nordic Journal of Botany* **33**: 155-158.

Takhtajan A. 1986. Floristic regions of the world. University of California Press, California.

Thornhill AH, Mishler BD, Knerr NJ, González-Orozco CE, Costion CM, Crayn DM, Laffan SW, Miller JT. 2016. Continental-scale spatial phylogenetics of Australian angiosperms provides insights into ecology, evolution and conservation. *Journal of Biogeography* **43**: 2085-2098.

Tibshirani R, Walther G, Hastie T. 2001. Estimating the number of data clusters via the Gap statistic. *Journal of the Royal Statistical Society B* **63**: 411–423.

Tucker CM, Cadotte MW, Carvalho SB, Davies TJ. 2017. A guide to phylogenetic metrics for conservation, community ecology and macroecology. *Biological Review* **92**: 698–715.

Vasconcelos TN, Alcantara S, Andrino CO, Forest F, Reginato M, Simon MF, Pirani JR. 2020. Fast diversification through a mosaic of evolutionary histories characterizes the endemic flora of ancient Neotropical mountains. *Proceedings of the Royal Society B*. **287**: 20192933.

Veldman JW, Buisson E, Durigan G, Fernandes GW, Le Stradic S, Mahy G, Negreiros D, Overbeck GE, Veldman RG, Zaloumis NP. 2015. Toward an old-growth concept for grasslands, savannas, and woodlands. *Frontiers in Ecology and the Environment* **13**: 154–162.

Vigeland MD, Spannagl M, Asp T, Paina C, Rudi H, Rognli O-A, Fjellheim S, Sandve SR. 2013. Evidence for adaptive evolution of low-temperature stress response genes in a Pooideae grass ancestor. *New Phytologist* **199**: 1060–1068.

Villaverde T, Escudero M, Martín-Bravo S, Jiménez-Mejías P, Sanmartín I, Vargas P, Luceño M. 2017. Bipolar distributions in vascular plants: A review. *American Journal of Botany* **104**: 1680-1694.

Vicentini A, Barber JC, Aliscioni SS, Giussani LM, Kellogg EA. 2008. The age of the grasses and clusters of origins of C₄ photosynthesis. *Global Change Biology* **14**: 2963–2977.

Wang Q, Huang J, Zang R, Li Z, El-Kassaby YA. 2022. Centres of neo-and paleo-endemism for Chinese woody flora and their environmental features. *Biological Conservation* **276**: 109817.

Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**: 153.

Zhang Y, Qian L, Spalink D, Sun L, Chen J, Sun H. 2021. Spatial phylogenetics of two topographic extremes of the Hengduan Mountains in southwestern China and its implications for biodiversity conservation. *Plant Diversity* **43**: 181-191

Zhang XX, Ye JF, Laffan SW, Mishler BD, Thornhill AH, Lu LM, Mao LF, Liu B, Chen YH, Lu AM, Miller JT. 2022. Spatial phylogenetics of the Chinese angiosperm flora provides insights into endemism and conservation. *Journal of Integrative Plant Biology* **64**: 105-117.

Zhao X, Valen E, Parker BJ, Sandelin A. 2011. Systematic clustering of transcription start site landscapes. *PLoS ONE* **6**: e23409.

Table 1. Summary of number of species, habitat and phyloregions corresponding to crown nodes for Poales.

Family	Total	Studied	Studied %	Closed	Phyloregion	Habitat
Poaceae	11994	5341	45	449	Afrotropical	Closed
Cyperaceae	5870	2690	46	398	Neotropic	Open
Bromeliaceae	3564	1066	30	551	Neotropic	Open
Eriocaulaceae	1206	348	29	6	Neotropic	Open
Restionaceae	552	453	82	3	Austral	Open
Juncaceae	514	209	41	5	Holarctic	Open
Xyridaceae	406	13	3	1	Neotropic	Open
Rapateaceae	97	24	25	2	Neotropic	Open
Typhaceae	74	26	35	0	Holarctic	Open
Thurniaceae	8	4	50	0	Neotropic	Open
Flagellariaceae	5	4	80	2	Indomalayan	Closed
Mayacaceae	5	1	20	0	Neotropic	Open
Joinvilleaceae	4	2	50	2	Indomalayan	Closed
Ecdeiocolaceae	3	2	67	0	Austral	Open
Total	24302	10183		1419		
%		42		14		

Total number of species per family were derived from the ‘World Checklist of Vascular Plants’ (WCVP; Govaerts, 2022). The number of species found in closed habitats was calculated from Supporting Information Table S3, while Phyloregions corresponding to the crown nodes were inferred from Fig. 4.

Supporting Information (brief legends)

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Figure S1. Poales botanical regions grouped into three ‘floristic kingdoms’ based on phylogenetic beta diversity, indicated by different colours and numbers.

Figure S2. The number of species of Poales missing from the phylogenetic dataset compared to the number listed in the World Checklist of Vascular Plants (WCVP) as of 28 February 2022, mapped per botanical region.

Figure S3. Phylogenetic diversity (PD) of the six largest Poales families categorised into open and closed habitats.

Figure S4. Phylogenetic endemism (PE) mapped per botanical region for Poales and eight families with the highest number of species in the dataset.

Fig. S5. Ancestral area reconstruction within Poales based on seven regions, obtained using BayAREA-like model in BioGeoBEARS.

Table S1. Sequences for the phylogenetic reconstruction.

Table S2. Calibrations used in the treePL (Smith & O’Meara, 2012) configuration file.

Table S3. Taxa included in this study, showing habitat scoring across the Poales.

Table S4. Comparison of six ancestral area reconstruction models based on BioGeoBEARS analyses for Poales.

Table S5. Results from corHMM ancestral state reconstructions.

Notes S1. Justification for selecting dispersal-extinction-cladogenesis (DEC) model of ancestral estimation.