Reconciling fossils with phylogenies reveals the origin and macroevolutionary processes explaining the global cycad biodiversity

Mario Coiro, Rémi Allio, Nathan Mazet, Leyla J. Seyfullah, Fabien L Condamine

To cite this version:

Mario Coiro, Rémi Allio, Nathan Mazet, Leyla J. Seyfullah, Fabien L Condamine. Reconciling fossils with phylogenies reveals the origin and macroevolutionary processes explaining the global cycad biodiversity. New Phytologist, 2023, 240 (4), pp.1616-1635. 10.1111/nph.19010 . hal-04144335

Reconciling fossils with phylogenies reveals the origin and macroevolutionary processes explaining the global cycad biodiversity

Mario Coiro[1,2], Rémi Allio[3], Nathan Mazet[4], Leyla J. Seyfullah[1] and Fabien L. Condamine[4]

[1]Department of Palaeontology, University of Vienna, 1090 Vienna, Austria; [2]Ronin Institute for Independent Scholarship, Montclair, NJ 07043, USA; [3]Centre de Biologie pour la Gestion des Populations, INRAE, CIRAD, IRD, Montpellier SupAgro, Université de Montpellier, 34988 Montpellier, France; [4]CNRS, Institut des Sciences de l'Evolution de Montpellier, Université de Montpellier, Place Eugène Bataillon, 34095 Montpellier, France

Authors for correspondence:
Mario Coiro
Email: mar.coiro@gmail.com

Fabien L. Condamine
Email: fabien.condamine@gmail.com

Received: *25 November 2022*
Accepted: *1 May 2023*

New Phytologist (2023)
doi: 10.1111/nph.19010

Key words: Antarctica, Cycadales, fossil leaves, Greenland, historical biogeography, total-evidence analysis.

Summary

- The determinants of biodiversity patterns can be understood using macroevolutionary analyses. The integration of fossils into phylogenies offers a deeper understanding of processes underlying biodiversity patterns in deep time. Cycadales are considered a relict of a once more diverse and globally distributed group but are restricted to low latitudes today. We still know little about their origin and geographic range evolution.
- Combining molecular data for extant species and leaf morphological data for extant and fossil species, we study the origin of cycad global biodiversity patterns through Bayesian total-evidence dating analyses. We assess the ancestral geographic origin and trace the historical biogeography of cycads with a time-stratified process-based model.
- Cycads originated in the Carboniferous on the Laurasian landmass and expanded in Gondwana in the Jurassic. Through now-vanished continental connections, Antarctica and Greenland were crucial biogeographic crossroads for cycad biogeography. Vicariance is an essential speciation mode in the deep and recent past. Their latitudinal span increased in the Jurassic and restrained toward subtropical latitudes in the Neogene in line with biogeographic inferences of high-latitude extirpations.
- We show the benefits of integrating fossils into phylogenies to estimate ancestral areas of origin and to study evolutionary processes explaining the global distribution of present-day relict groups.

Introduction

Extinction is an important process shaping biodiversity. Given the pervasiveness of extinction in deep time, comparative studies of living organisms offer a limited window on evolutionary patterns and processes, especially when looking at geological timescales (Crisp *et al.*, 2011; Marshall, 2017). In the past three decades, an extended toolbox has been developed for macroevolutionary analyses of dated phylogenies (Nee *et al.*, 1992, 1994; Morlon, 2014). However, even if some neontological approaches seem to be robust in the absence of extinct lineages (Morlon *et al.*, 2011; Beaulieu & O'Meara, 2015; Condamine *et al.*, 2020), studies have shown that integrating fossil information can greatly improve macroevolutionary inferences (Slater *et al.*, 2012; Fritz *et al.*, 2013; Hunt & Slater, 2016; Oliveros *et al.*, 2020). Recognition of the fundamental importance of fossils in macroevolutionary studies has led to the development of new methods that use the fossil record to infer diversification dynamics (Silvestro *et al.*, 2014, 2018a; Mitchell *et al.*, 2019), as well as methods to intertwine fossils in molecular phylogenies (Zhang *et al.*, 2016; Gavryushkina *et al.*, 2017). This has led to a modest fossil renaissance in macroevolution (King *et al.*, 2017; Slater *et al.*, 2017), which may have important implications in our understanding of the geographic patterns and processes underlying global diversity and distribution, especially for ancient groups like plants (Crisp *et al.*, 2011; Mao *et al.*, 2012; May *et al.*, 2021).

Despite the expansion of statistical methods, the plant fossil record remains particularly problematic and yet central to macroevolutionary analyses. Different plant organs tend to fossilize separately (Bateman & Hilton, 2009), and the phylogenetic informativeness of plant fossils varies between organ types (Bateman & Simpson, 1998; Coiro & Barone Lumaga, 2018; Coiro *et al.*, 2020a) and conservation modes, as well as within and between different clades. This complicates both the taxonomic assignment of plant fossils and the estimate of stratigraphic ranges. However, this has not discouraged people from using generic-level occurrences to infer macroevolutionary dynamics at different scales (vascular plants: Silvestro *et al.*, 2015; ferns: Lehtonen *et al.*, 2017; conifers: Condamine *et al.*, 2020). Even if such investigations are probably robust, the possibility of systematic error introduced by misinterpretation of the fossil record cannot be excluded. Integrating plant fossils during phylogenetic reconstruction represents an alternative. Total-evidence dating (Ronquist *et al*., 2012) with the fossilized birth-death (FBD) model (Heath *et al*., 2014; Zhang *et al*., 2016) has been used to estimate both phylogenetic relationships and divergence times in many fossil-rich plant groups (Grimm *et al*., 2015; Larson-Johnson, 2016; Renner *et al*., 2016; May *et al*., 2021; Zhang *et al*., 2022). These analyses take advantage of well-preserved fossils or well-understood relationships between fossil and extant taxa to improve phylogenetic and divergence time inferences.

Among plants, cycads (order Cycadales) include > 360 extant species (Govaerts *et al*., 2021) of which 68% are threatened with extinction (IUCN, 2022). Originating in the Paleozoic (Hermsen *et al*., 2006; Condamine *et al*., 2015), cycads are widely considered quintessential 'living fossils', due to their supposed long-term morphological stasis and their apparent dominance in the fossil record followed by a decline toward the present. Indeed, cycad evolution was thought to have reached its pinnacle in the Middle Jurassic, with a subsequent decline in morphological and taxonomic diversity (Niklas *et al*., 1983), possibly driven by competition with the angiosperms (Norstog & Nichols, 1997) or the decline of nonavian dinosaurs (Mustoe, 2007; Butler *et al*., 2009). This decline is thought to have led to a depauperate modern cycad flora (Harris, 1961). The idea of cycad diversity as a relic from the Mesozoic has been challenged on molecular grounds, since the extant species diversity is inferred to be the result of Miocene-Pleistocene radiations in most genera (Treutlein & Wink, 2002; Nagalingum *et al*., 2011; Salas-Leiva *et al*., 2013; Condamine *et al*., 2015; Liu *et al*., 2022). Despite this, the old 'living fossil' viewpoint still continues to influence some views on the evolution of the group (Zhang *et al*., 2015; Nackey *et al*., 2018), its biotic interactions in deep time (Cai *et al*., 2018; Salzman *et al*., 2020), and its geographic origin and range expansion (Salas-Leiva *et al*., 2013).

Our poor understanding of the cycad fossil record hinders testing hypotheses about their macroevolutionary trajectories and especially their historical biogeography. The relationships of the main Mesozoic taxa are still poorly understood (Hermsen *et al*., 2006; Coiro & Pott, 2017), and it remains unclear whether many genera represent biological rather than purely taxonomic units (Pott *et al*., 2007). Fossil cycads representing reproductive structures are rare and have been often overlooked or have little or unclear phylogenetic value, though better-preserved specimens have led to fruitful insights (Spencer *et al*., 2017; Rothwell *et al*., 2022; Elgorriaga & Aktinson, 2023). Although the leaf record is much more abundant, it is also poorly understood. Recently, the cycadalean nature of one of the most common leaf taxa from the Mesozoic, *Nilssonia* Brongniart, has been questioned on chemical grounds (Vajda *et al*., 2017). Similarly, the Early Cretaceous leaf *Mesodescolea* S.Archang., thought to be the closest relative of extant *Stangeria* Hook. ex Hook.f., has been reinterpreted as an angiosperm leaf (Coiro *et al*., 2020b). Even if some of the extant genera start to be recognized in sediments as old as the Eocene (Hill, 1978; Carpenter, 1991; Kvaček, 2002; Su *et al*., 2014; Erdei *et al*., 2018), many Cenozoic leaf fossils defy any attempt of classification into the extant cycad groups, such as *Eostangeria* Barthel from the Paleocene of North America and Europe, Eocene of Germany, and Miocene of Bulgaria (Barthel, 1976; Kvaček & Manchester, 1999; Uzunova *et al*., 2001), and *Pseudodioon* Erdei, Agkun et Barone Lumaga from the Miocene of Turkey (Erdei *et al*., 2009). Moreover, the Mesozoic genus *Ctenis* Lindl. & Hutt. is present in the Eocene of Oregon (Erdei & Manchester, 2015), a locality where cycads are absent today. Likewise, *Dioonopsis* Horiuchi & Kimura, with three species from the Paleocene of Japan and Eocene of North America (Horiuchi & Kimura, 1987; Erdei *et al*., 2012), has been linked with the extant genus *Dioon* Lindl. based on general leaf morphology (Moretti *et al*., 1993) as well as phylogenetic analyses of anatomy and morphology (Hermsen *et al*., 2006; Martínez *et al*., 2012), but this link has been recently questioned (Barone Lumaga *et al*., 2015; Erdei & Manchester, 2015; Erdei *et al*., 2019). Given the age of cycads and long branches subtending radiations of extant genera, the uncertainties in fossil placements hinder any macroevolutionary inferences that would be based solely on dated phylogenies (Crisp *et al*., 2011).

Few phylogenetic analyses have included fossil cycads (Hermsen *et al*., 2006; Martínez *et al*., 2012). Although these pioneering works have merits, they still present many issues. First, the inferred topology is invariably incompatible with the topology of extant cycads obtained using molecular data, a conflict that calls into question the relationship retrieved between fossil and extant taxa (Coiro & Pott, 2017). Second, these analyses were conducted using generic-level taxa, even though paleobotanical practice may lead to the creation of heterogeneous, nonmonophyletic genera for practical use (Harris, 1961). Third, they included genera with uncertain cycadalean affinities, such as the seed fern-like *Ticoa* S.Archang. and *Kurtziana* Frenguelli, or the angiosperm *Mesodescolea* S.Archang. (Coiro *et al*., 2020b). Fourth, they included both leaf and stem taxa, each scored for only a few non-overlapping and limited characters, which can create conflicting phylogenetic signals that cannot be disentangled.

Here, we study the macroevolutionary processes underlying the assembly of global diversity and distribution of cycads by reconciling fossils and phylogeny through Bayesian total-evidence analysis. Based on morphological analyses of leaves, 60 fossil cycad species were integrated in a single phylogenetic inference performed with *c*. 87% of all extant species for which both morphological and molecular data were collected. The extant cycad distribution is highly disjunct today with sister genera located on different continents (Fig. 1), such as *Encephalartos* Lehm. (Africa) and *Lepidozamia* Regel (Australia) or *Zamia* L. + *Microcycas* A.DC. (Americas), and *Stangeria* (Africa). We assess the roles of dispersal, extinction, and vicariance in establishing the current cycad distribution by inferring the ancestral geographic origin and tracing the historical biogeography with a time-stratified and tectonic-informed process-based model. We also tested the effect of treating fossil ranges as fundamentally ambiguous by scoring absences as unknown. Given their appearance in the fossil record in the Paleozoic and limited dispersal ability, a Pangean origin and vicariance events during the breakup of ancient supercontinents are expected. Moreover, the latitudinal span of extant cycads is currently

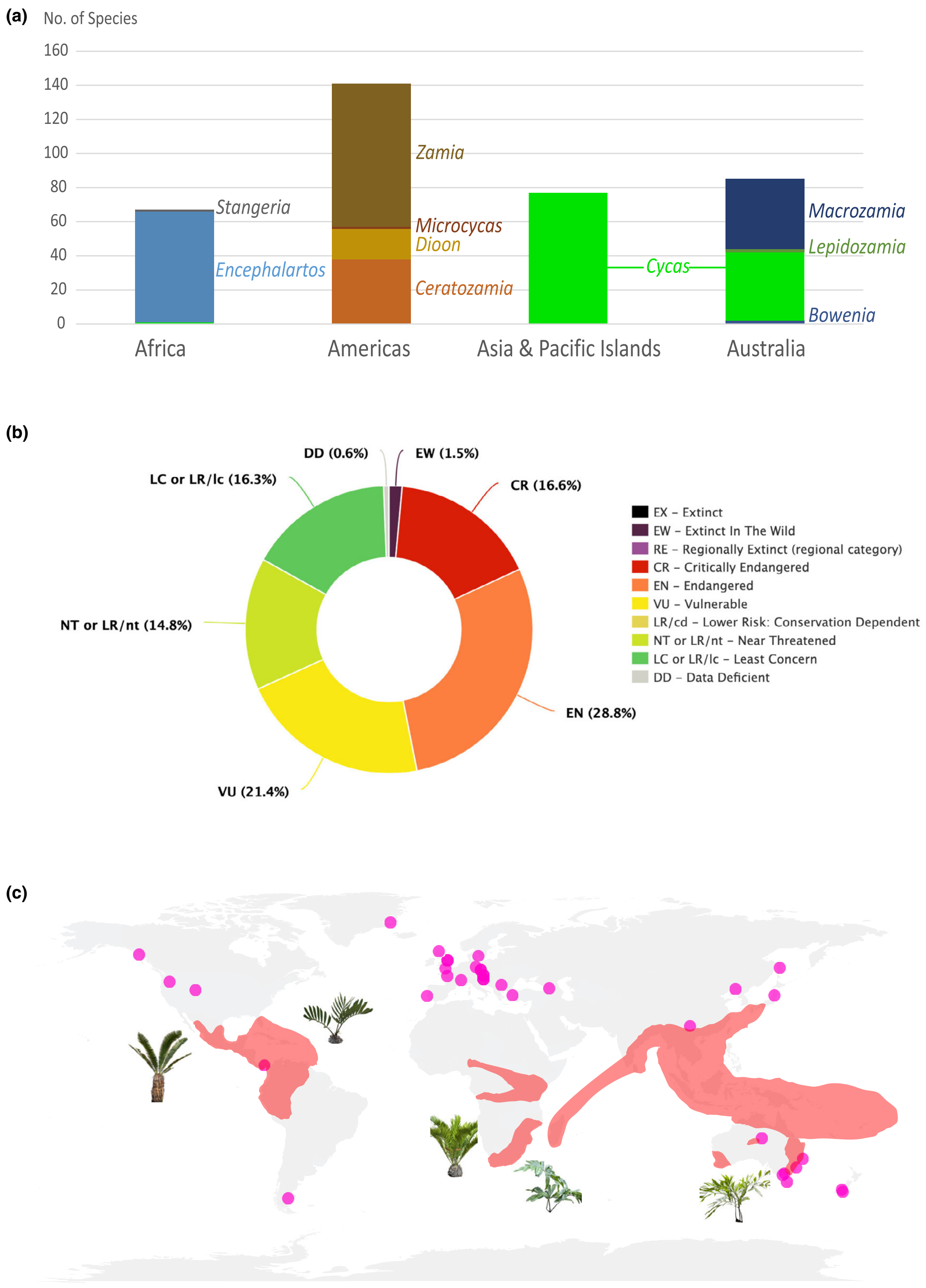

Fig. 1 Global distribution of Cycadales. (a) Species richness of the 10 extant cycad genera across continents. (b) IUCN threat categories for 337 extant cycad species (redrawn from IUCN, 2022). (c) Map showing the extant distribution of cycad diversity (red areas) and sampled fossil localities (purple dots), indicating broader geographic distribution in the past.

restricted to subtropical latitudes, but fossils indicate wider latitudinal distribution in the Mesozoic (Harris, 1926; Smoot *et al*., 1985; Hermsen *et al*., 2006; Fig. 1), suggesting high-latitude extirpations (Meseguer & Condamine, 2020). We finally infer the latitudinal span of cycads through time to test the hypothesis of poleward extinctions in the Cenozoic.

Materials and Methods

Global cycad diversity and distribution

The World Checklist of Vascular Plants (WCVP; Govaerts *et al.*, 2021) lists 362 recognized species of Cycadales. However, new cycad species are regularly described each year (Calonje *et al.*, 2021; Pérez-Farrera *et al.*, 2021a,b, 2022; Martínez-Domínguez *et al.*, 2022), so that there are now 370 cycad species (The World List of Cycads: http://cycadlist.org/, Calonje *et al.*, 2022). From the WCVP data, we extracted the distribution range of each species ('geographic_area') to categorize their continental distribution for biogeographic analyses (to be described later).

Molecular phylogenetics

We sampled 321 extant species out of 370 recognized species, including all genera and a species sampling across all genera as follows: *Bowenia* Hook. ex Hook.f. 2/2, *Ceratozamia* Brongn. 33/38, *Cycas* L. 104/118, *Dioon* 13/18, *Encephalartos* 65/65, *Lepidozamia* 2/2, *Macrozamia* Miq. 27/41, *Microcycas* 1/1, *Stangeria* 1/1, and *Zamia* 74/84. This results in 86.8% (321/370) of the total species richness. *Ginkgo biloba* L. was added as an outgroup, which is recognized as the sister lineage of cycads (Liu *et al.*, 2022; Yang *et al.*, 2022). The molecular matrix was assembled from sequences available on GenBank. We included 18 loci including plastidial (matK, rbcL), mitochondrial (26S), and nuclear genome sequences (AC3, F3H, GroES, RPB1, SAMS, CyAG, HTS, WRKY4, 40S, 5.8S, LiSH, GTP, PHYP, HZP, and PEX4).

We aligned the noncoding genes using MAFFT 7.110 (Katoh & Standley, 2013) with the E-INS-i algorithm, while we aligned the coding genes using the OMM MACSE pipeline (omm_macse_v11.05.sif; Ranwez *et al.*, 2018), which includes nucleotides and amino acid alignment steps combined with several cleaning steps (including HMMCLEANER). All the resulting alignments were checked for codon stops and eventually refined by eye with MESQUITE 3.7 (Maddison & Maddison, 2021). All gene alignments were concatenated into a nucleotide supermatrix. The final molecular matrix contained 11 895 nucleotides.

We performed maximum likelihood (ML) to reconstruct phylogenetic relationships of extant species. ML inference was implemented with IQ-TREE 2.2.0 (Minh *et al.*, 2020) using MODELFINDER to select the best-fit partition scheme and the best-fitting substitution model for each partition (*-m MFP+MERGE*, Chernomor *et al.*, 2016; Kalyaanamoorthy *et al.*, 2017). For IQ-TREE analyses, we estimated the most likely tree with 100 separate ML searches, as well as 100 searches, which after initial model optimization on a parsimony tree used 100 random tree topologies as starting trees for each search. As recommended, we optimized ML searches to avoid local optima by (1) increasing the number of unsuccessful iterations before stopping tree optimization to 500 (*-nstop 500*), and (2) decreasing the perturbation strength for randomized NNI to 0.2 (*-pers 0.2*). Branch supports were evaluated with 1000 ultrafast bootstraps (UFBS; Hoang *et al.*, 2018), with strong UFBS values ≥95% considered as strong support.

Sampling fossils and morphological data

We generated a matrix of 31 morphological characters covering leaf morphology and cuticular anatomy, based on literature and direct observations (Supporting Information Methods S1; Table S1; Notes S1). We then scored characters for 60 species of fossil cycad leaves, including macrofossils and dispersed cuticles, as well as for the extant taxa of the molecular matrix. Coding was based on observations of the specimens or descriptions and photographs of the specimens available in the literature (Greguss, 1968; Mickle *et al.*, 2011; Barone Lumaga *et al.*, 2015; Coiro & Pott, 2017; Vovides *et al.*, 2018; Erdei *et al.*, 2019; Coiro *et al.*, 2020a, 2021; Glos *et al.*, 2022).

Bayesian total-evidence analyses

Estimates of divergence times were carried out under a Bayesian total-evidence dating approach with the FBD model, which explicitly models the speciation, extinction, fossilization, and sampling processes (Heath *et al.*, 2014; Zhang *et al.*, 2016; Gavryushkina *et al.*, 2017). For this analysis, the outgroup (*Ginkgo biloba*) was removed, to avoid issues with the long unsampled fossil history of this lineage. Bayesian inferences were performed with MRBAYES 3.2.7a, with the six molecular partitions as estimated by ModelFinder in IQ-TREE and set to have their own evolutionary model. Models of sequence evolution were set with the reversible-jump Markov Chain Monte Carlo (MCMC) with the gamma rate for site heterogeneity and the proportion of invariable sites (Huelsenbeck *et al.*, 2004). We also set one partition for all the morphological characters. Morphological evolution was computed with the Markov-k model (Lewis, 2001) with the correction for variable characters and a gamma-distributed rate variation across characters. Variable gamma rates were chosen as the preferred model to analyze the different partitions. Substitution model parameters (rates, gamma shape, and invariable sites) were unlinked between partitions. Two runs of eight incrementally heated MCMC starting from a random tree were performed.

The dating was conducted using the independent gamma rates (IGR) model in which tree branches have their own evolutionary rates (equivalent to the uncorrelated lognormal relaxed clock; Drummond *et al.*, 2006). The IGR clock with an exponential prior for the variance parameter was used for the molecular partitions (*prset clockvarpr= igr*) with an exponential prior on the variance of the gamma distribution from which the branch lengths are drawn in the IGR model (*prset igrvarpr= exp(10)*), while a strict clock was used for morphology. The mean clock rate (mean substitution rate per site per million years (Myr)) is assigned a lognormal prior (*prset clockratepr= lognorm(-6,0.5)*); giving a mean *c.* 0.001 substitution rate per site per Myr.

Since preliminary analyses failed to converge without topological constraints, we enforced the best IQ-TREE topology for extant species with *partial* constraints, meaning that the placement of fossil taxa is inferred while the topology of extant species remained fixed. The 60 fossil taxa were added as tips based on single specimens or multiple specimens from the same locality.

We assign priors for the fossil ages based on the geological time scale. This typical step in total-evidence dating aims at calibrating the fossil taxa instead of the internal nodes of the tree. Uncertainty about the dating of the fossiliferous localities was implemented as a uniform distribution for the age of the fossil tips, for example, *Bowenia johnsonii* R.S.Hill, K.E.Hill, Carpenter et Jordan is an early Eocene fossil that translates into *uniform (47.8,56)*. Sources for the fossil ages are listed in Table S2. The speciation, extinction, fossilization, and sampling processes are explicitly modelled using the FBD process: *prset brlenspr = clock: fossilization* (Heath *et al.*, 2014; Zhang *et al.*, 2016). The three FBD parameters were set as follows: (1) the fossilization prior with a beta distribution: *prset fossilizationpr = beta(1,1)*; (2) the speciation rate with an exponential prior: *prset speciationpr = exp (10)*; and (3) the relative extinction rate with a beta prior: *prset extinctionpr = beta(1,1)*. The sampling strategy of extant taxa was set to diversity (*prset samplestrat = diversity*) with sampling fraction set to 0.87 (*prset sampleprob = 0.87*), wherein fossils are sampled randomly and can be tips or ancestors (Zhang *et al.*, 2016). A uniform prior was set on the tree age, bounded by the lowest age of the Lopingian at 259.1 million years ago (Ma) (that correspond to the older fossils assignable with confidence to the cycads) and the upper age of the Famennian at 358.9 million years ago (Ma) (that correspond to some of the oldest stem seed plants): *prset treeagepr = uniform(259.1,358.9)*.

The total-evidence dating was finally run for 50 million MCMC generations with trees and associated model parameters sampled every 50 000 generations. All the analyses were carried out on the CIPRES science gateway (Miller *et al.*, 2010). We performed the analyses three times to ensure repeatability of the results. Convergence diagnostics were checked for each analysis (i.e. average standard deviation of split frequencies (ASDSF) < 0.05, potential scale reduction factor (PSRF) close to 1.0) as well as the effective sample size (ESS) > 200 in TRACER 1.7.1 (Rambaut *et al.*, 2018). A consensus tree was obtained after discarding 25% of the generations as burn-in to compute posterior probability (PP), median age, and 95% highest posterior density (HPD) for each node.

Representing uncertainty in fossil placement

We used RoguePlots (Klopfstein & Spasojevic, 2019) to investigate the uncertainty in the fossil placement of some taxa, namely the members of the two genera *Dioonopsis* Horiuchi & Kimura and *Eostangeria* Barthel. These were selected for their importance as calibrations (*Dioonopsis*) and their uncertain relationship in the literature (*Eostangeria*). From the posterior distribution of trees from the constrained analysis, we generated tree summaries to show the placement of these taxa on the consensus phylogeny.

Estimating ancestral latitudes

The latitudinal span of the Cycadales throughout their evolutionary history was reconstructed using the directional Brownian–Motion method of Silvestro *et al.* (2018b) with the variable trend implemented in Zhang *et al.* (2022). Mean latitude for the extant genera was obtained using occurrence data obtained from the Global Biodiversity Information Facility (GBIF). These occurrences were obtained by querying the different genera using the function *occ_search* from the R package RGBIF (Chamberlain *et al.*, 2019), and then cleaned using the function *clean_coordinates* from the R package COORDINATECLEANER (Zizka *et al.*, 2019). Further cleaning of occurrences was done manually. This resulted in mean latitude estimates for 207 extant taxa. Paleolatitudes for the fossil taxa were obtained using the paleolatitude calculator (van Hinsbergen *et al.*, 2015). Approximate position of the fossil localities was queried, and the average age of the localities was used to obtain paleolatitudes from the model. This resulted in 48 fossils with constrained paleolatitudes. The MCMC was run for 100 000 generations using 100 trees randomly sampled from the posterior sample. Traits were rescaled by 10 to help convergence.

Estimating historical biogeography

We estimated the ancestral areas of origin and geographic range evolution for Cycadales using the ML approach of dispersal–extinction–cladogenesis (DEC, Ree & Smith, 2008) as implemented in the C++ version (Beeravolu & Condamine, 2016), as well as the R package BIOGEOBEARS (Matzke, 2018). To infer the biogeographic history of a clade, DEC requires a time-calibrated tree (i.e. the consensus tree obtained from the Bayesian TED analysis), the current distribution of each species, a set of geographic areas, and a time-stratified geographic model that is represented by connectivity matrices for specified time intervals spanning the entire evolutionary history of the group.

We first defined a set of geographic areas based on paleogeographic knowledge as follows: (1) West Palearctic, defined as Western Europe to the Urals, (2) East Palearctic, defined as east of the Urals, above 3000 m in the Himalayas and north of Sichuan in China, (3) West Nearctic, defined as Western North America including the Rocky Mountains, (4) East Nearctic, defined as North America east of the Rocky Mountains, (5) Central America, going from the northern border of Mexico southward to the border between Panama and Colombia, (6) Caribbean Islands, excluding Trinidad and Tobago, (7) South America, defined as all countries from Colombia to Argentina and including Trinidad and Tobago, (8) Africa, defined as the whole African continent and Arabian Peninsula but excluding the islands in the Indian Ocean, (9) Madagascar, defined as the island of Madagascar and all other Indian Ocean islands in the vicinity, (10) India, defined as the area below 3000 m from NW Pakistan to the border with Myanmar, (11) Indonesia and Wallacea, defined as Myanmar, SE Asia, southern China, western Indonesia to Lydekker's Line; including the Lesser Sunda Islands but excluding Timor, Wetar and associated islands, which are Australasian in origin, and (12) Australasia, defined as everywhere east of Lydekker's Line but including Timor, Wetar and small nearby islands. Furthermore, there is evidence for the ancient (Triassic) presence of cycads in Antarctica (e.g. *Antarcticycas schopfii* Smoot, Taylor, et Delevoryas emend. Hermsen, T. N. Taylor, E. L. Taylor, et Stevenson, Smoot *et al.*, 1985; Hermsen *et al.*, 2006), and Greenland (e.g. *Anthrophyopsis crassinervis* Nathorst (Harris, 1926, 1932)). To reflect their increasingly recognized role in global plant biogeography (e.g. Estrella *et al.*, 2019), we added Antarctica and Greenland to the geographic areas to consider the possibility that cycads colonized these continents. We choose these 14 areas to test biogeographic hypotheses and estimate implications from the breakup of ancient Mesozoic supercontinents, whereas Wallace's bioregions do not capture tectonic and geological changes.

Relying on the WCVP (Govaerts *et al.*, 2021), the geographic distribution for all extant cycad species (column 'geographic_area' in the database) was categorized by coding the presence or absence of each species in each of the above-defined areas. We also used data available in the literature (e.g. http://cycadlist.org/, Calonje *et al.*, 2022). Occurrences of introduced species or marginally entering an area were not considered. Similarly, we coded the 60 fossil species of our dataset. However, fossil data provide evidence on the past geographic presence of taxa, but not about geographic absence due to the incompleteness of the fossil record. In BioGeoBEARS, but not yet in DECX, one can include this information by coding the geographic range of fossils with missing data '?' instead of true absence (*useAmbiguities* = *TRUE* option). Coding presence in region A and unknown presence or absence in other regions mean that any geographic range including A will have tip likelihood of 1, and any geographic range excluding A will have a tip likelihood of 0. Fossil lineages thus represent a positive constraint on ancestral range estimates, which is conservative because it gives more weight to large ranges than to single-area ranges. However, in most empirical analyses like in cycads, many species occur only in a single area, and a few are widespread. Hence, we performed the analyses with the positive constraint strategy, using '?' for fossil species, and ran it again with the assumption that fossil ranges are the true ranges. We then compared the effects of these assumptions on ancestral range estimates. An overview of the global geographic distribution of extant and extinct Cycadales is presented in Fig. 1.

A time-stratified geographic model was built using connectivity matrices that consider paleogeographic changes through time (Beeravolu & Condamine, 2016). Connectivity matrices specify constraints on area connectivity by coding 0 if any two areas are not connected or 1 if they are connected at a given period based on paleogeographic reconstructions (e.g. Blakey, 2008; Seton *et al.*, 2012; Kocsis & Scotese, 2021). We created connectivity matrices to represent major changes in tectonic conditions that may have affected cycad distribution and to define biological plausibility of ranges over time. For instance, there are wide disjunctions between Australasia and North America such that no species is found on both. We did not add dispersal constraints because setting the values for dispersal rates between regions through time is subjective, and it has been shown that dispersal probability categories had minor effects on ancestral state estimates (Chacón & Renner, 2014). We assumed a dispersal matrix with equal rates between areas. The time-slicing protocol introduced by Upchurch *et al.* (2002) is followed here. Four time slices were selected to construct the time-stratified model and correspond to the major geological periods. The first covers the Carboniferous to the Late Triassic (358.9–201.3 Ma), corresponding to the assembly of Pangea (Kocsis & Scotese, 2021). The second ranges from the Early Jurassic to the Late Cretaceous (201.3–66 Ma), which represents the breakup of Pangea into Gondwana and Laurasia (Blakey, 2008; Seton *et al.*, 2012). There are two time slices for the Cenozoic: one covers the Paleogene (66–23 Ma), and the other encompasses the Neogene to the present, corresponding to the tectonic plate motion to current positions. These time slices were designated because: (1) they are based on the major periods with important geological changes (Seton *et al.*, 2012; Müller *et al.*, 2016; Kocsis & Scotese, 2021); (2) subdivision of the Mesozoic into two time slices allows testing of biogeographic hypothesis (e.g. supercontinent breakups, Ezcurra & Agnolín, 2012); and (3) subdivision of the Cenozoic into Paleogene and Neogene with the latter witnessing the establishment of the present terrestrial biodiversity (Fine & Ree, 2006) and marked by a modern plate configuration (Kocsis & Scotese, 2021) and the establishment of the present climate regime during a long cooling process (Westerhold *et al.*, 2020).

We used the most likely ancestral range estimates to count the biogeographic events such as dispersals 'into' and 'out of' a region as well as local extinctions (extirpations from a region). Biogeographic events can occur at nodes (cladogenesis) and along branches (anagenesis). We made a custom R script to retrieve the biogeographic events and calculate their timing across the phylogeny (see Data availability). Following previous studies (e.g. Antonelli *et al.*, 2018; Meseguer & Condamine, 2020), anagenetic events were dated at middle points of branches and constrained by the time-stratified geographic model to account for appearances and disappearances of regions and connectivities between them through time.

Results

Total-evidence phylogeny of cycads

The ML phylogeny of cycads recovers generic relationships in agreement with the literature (Fig. S1) and is generally robust (UFBS ≥ 95, Fig. S2). However, many nodes within the extant genera remain unresolved. This topology served as backbone topological constraints to improve convergence of Bayesian inferences. After checking for convergence of the Bayesian total-evidence analyses (ASDSF = 0.013; on average PSRF = 0.9999 and ESS = 623 for all parameters; $n = 94$), the consensus tree shows an extant-genus topology in agreement with previous studies and that most cycad fossils are not closely related to the extant groups (Figs 2, S3, S4). Indeed, only five fossil taxa are strongly associated with Cycadaceae (i.e. are included in a clade with *Cycas* with PP = 1) including the extinct genus *Paracycas* and the two *Cycas* fossil species, and 18 are strongly associated with Zamiaceae (Fig. 3). Many relationships based on morphological comparative analyses are retrieved, that is, *Ceratozamia hoffmannii* Ettinghausen and *Ceratozamia floersheimensis* (Engelhardt) Kvaček form a clade with *Ceratozamia* (PP = 1), *Macrozamia australis* Carpenter is sister to *Macrozamia* (PP = 1), and *Lepidozamia hopeites* (Cookson) L. Johnson and *Lepidozamia foveolata* are

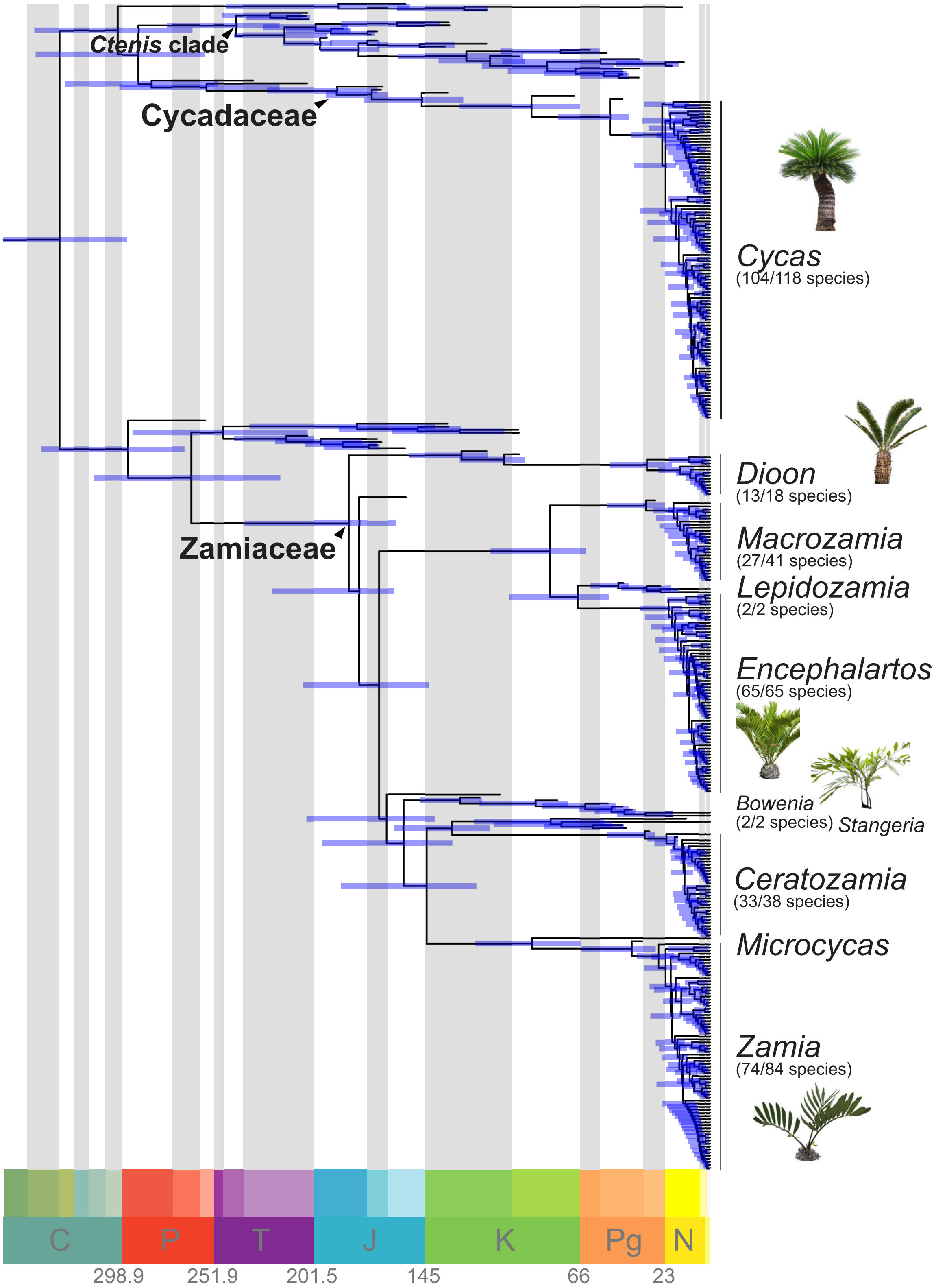

Fig. 2 Bayesian total-evidence dated phylogeny of Cycadales. This chronogram is the resulting consensus tree from the MrBayes analyses performed with the fossilized-birth-death model and an uncorrelated relaxed molecular clock. The tree includes 321 extant species and 60 extinct species with median divergence times along with 95% Highest Posterior Density (blue bars) for each node. C, Carboniferous; J, Jurassic; K, Cretaceous; N, Neogene; P, Permian; Pg, Paleogene; T, Triassic.

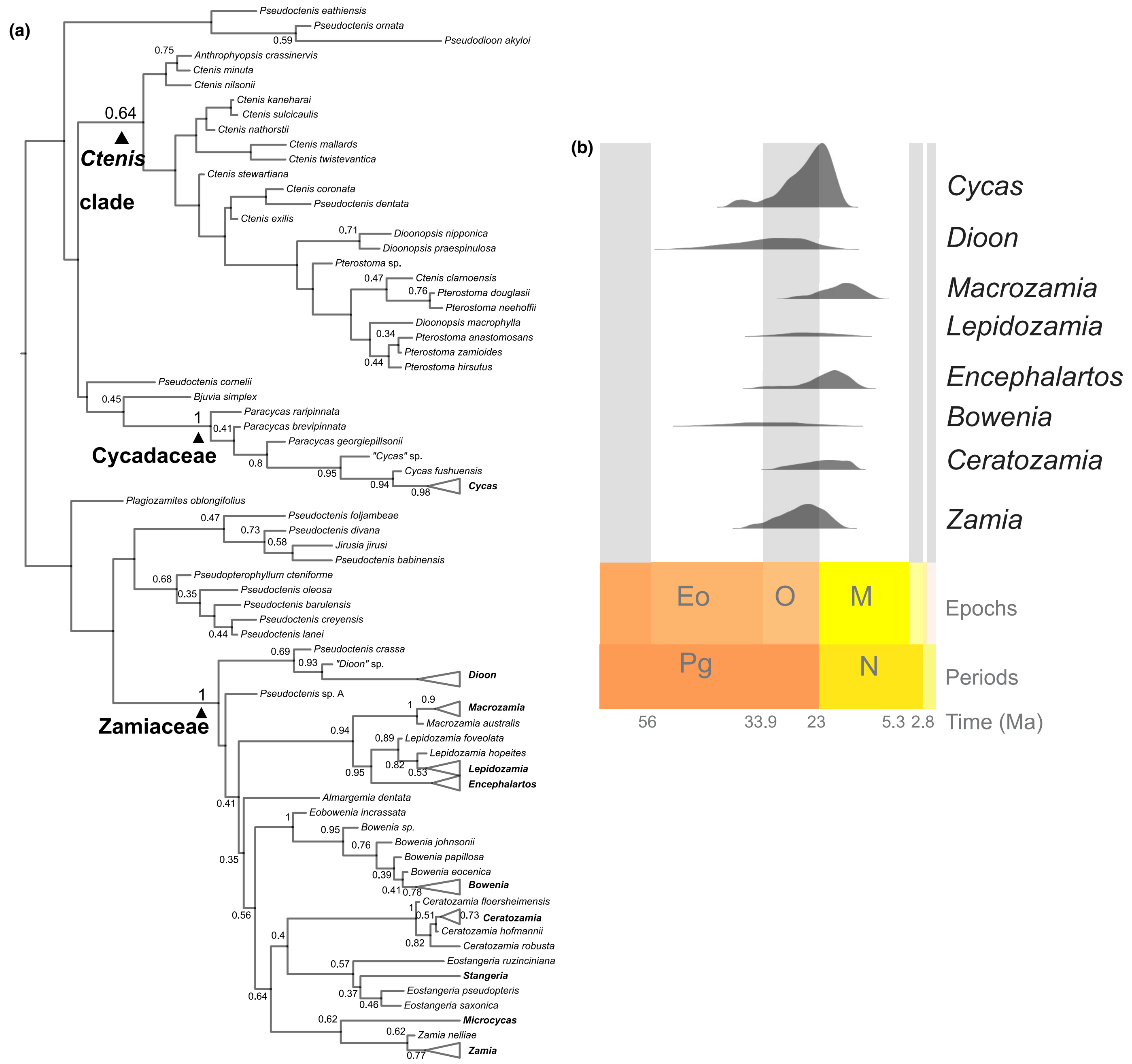

Fig. 3 Ages of extant genera and fossil placements. (a) Phylogenetic relationships for extant (genera are collapsed) and extinct cycads with support values (Posterior Probability indicated when > 0.3). Node support for extant genera is only indicated when < 1. (b) Posterior distributions for the crown ages of the extant genera. Eo, Eocene; M, Miocene; N, Neogene; O, Oligocene; Pg, Paleogene. Ma, million years ago.

sister to *Lepidozamia* (PP = 0.82 and 0.89, respectively). Species of the genus *Eostangeria* Barthel are retrieved in a clade with extant *Stangeria*, though with lower levels of support (PP = 0.57; Fig. S5). Likewise, extinct species of the genus *Bowenia* are retrieved in a clade with extant *Bowenia* with high support (PP = 0.95), and sister to genus *Eobowenia* with maximal support (PP = 1). The relationships between fossil taxa are less strongly supported than the relationships between fossil and extant taxa. There are 27 fossil species along the stem Cycadaceae and 10 fossil species along the stem Zamiaceae; most of them have low node support (PP < 0.5). Nonetheless, clades' PP > 0.5 are retrieved for several groups of fossil taxa, indicating that even a small morphological matrix can contain some phylogenetic informativeness (Fig. 3).

The fossil leaves *Dioonopsis praespinulosa* (Hollick) Erdei, Manchester et Kvaček and *Dioonopsis macrophylla* (Potbury) Erdei, Manchester et Kvaček, which have been previously associated with extant *Dioon* and even used as calibrations for molecular dating (Nagalingum *et al.*, 2011; Condamine *et al.*, 2015; Gutiérrez-Ortega *et al.*, 2018), are not related to *Dioon* nor

Zamiaceae, but are nested in an extinct clade including *Ctenis* species along the stem branch of Cycadaceae (Fig. S5).

Our analyses do not support the monophyly of any fossil cycad genus (Fig. 3). The genus *Pseudoctenis* Seward, defined as having broadly attached multiveined leaflets like *Ctenis* but lacking anastomoses, is wildly polyphyletic. It includes taxa related to the Zamiaceae (i.e. *Pseudoctenis* species A from the Late Jurassic of France and *Pseudoctenis crassa* S.Archang. & Baldoni from the Early Cretaceous of Argentina (Archangelsky & Baldoni, 1972)), taxa nested among species assigned to *Ctenis* (*Pseudoctenis dentata* S.Archang. & Baldoni from the Early Cretaceous of Argentina), as well as two clades including respectively *Pseudoctenis* species from the Jurassic as well as the genus *Pseudopterophyllum* Florin, and *Pseudoctenis* species from the Cretaceous and the genus *Jirusia* Bayer. A relationship between the Early Cretaceous *Pseudoctenis ornata* A.Archang., R.Andreis, S.Archang. & A.Artabe, and the Miocene *Pseudodioon akioly* Erdei, Agkun & Barone Lumaga is also retrieved (0.59). Interestingly, the type species of the genus (*Pseudoctenis eathiensis* (Richards) Seward) does not receive appreciable support for any placement.

The phylogenetic analysis finds some evidence (although not strong) of a long-lasting lineage of cycads that did not leave any extant representatives. This clade includes the mostly Mesozoic genus *Ctenis* together with the Cenozoic genera *Pterostoma* Hill and *Dioonopsis* Horiuchi & Kimura (supported with PP = 0.64). Such a lineage was postulated on comparative grounds because of the presence of peculiar H-anastomoses in the venation of these leaves, a character absent in all extant cycads (Erdei & Manchester, 2015). However, not all members of the lineage retrieved here possess anastomoses (i.e. *Pseudoctenis dentata*).

Total-evidence dating of cycads

Using the FBD model, the Bayesian dating analyses estimate the divergence between Zamiaceae and Cycadaceae *c.* 330 Ma at the boundary between the Early and Late Carboniferous (Fig. 2; 95% HPD = 296.2–358.9 Ma). These results place the split at an older date than the previous node-dating analyses (Table 1). The origin of the crown-group Zamiaceae is estimated at 183.5 Ma between the Late Triassic and the Late Jurassic (95% HPD = 159.8–236.3 Ma). *Bowenia* is inferred to have split from its sister clade *c.* 155.6 Ma between the Early Jurassic and the Early Cretaceous (95% HPD = 130.9–197 Ma). Zaminae (genera *Ceratozamia*, *Stangeria*, *Zamia*, and *Microcycas*) appears to have an older crown age than Encephalartinae (genera *Encephalartos*, *Lepidozamia*, *Macrozamia*), with the former clade originating at 143.7 Ma between the Jurassic and Early Cretaceous (95% HPD = 118.8–187.3 Ma), and the latter originating at 81.3 Ma between the Early Cretaceous and the Paleocene (95% HPD = 63.2–111.5 Ma). A Cretaceous age is inferred for the divergence between *Microcycas* and *Zamia* (95% HPD = 65.7–119.3 Ma).

Genus crown ages inferred in our analyses are summarized in Table 2. Even though we do retrieve a relatively young origin of extant species diversity (Fig. 3), the hypothesis of a synchronous radiation of the extant genera is weakened in our analysis. Based on a sample of 100 random trees from the posterior, a repeated measurement ANOVA retrieves significant differences between the crown ages of the genera ($P \leq 0.0001$). *Post hoc* testing finds significant differences between the mean ages of most genera (Table 2).

Table 1 Ages for the main cycad clades retrieved in our analysis and compared with ages reported from the node-calibration analysis of Condamine *et al.* (2015).

	Ages (Ma) from Condamine *et al.* (2015)			Ages (Ma) from the total-evidence dating		
Cycad clades	Median	Min	Max	Median	Min	Max
Cycadales	274.5	235.0	332.4	330.4	296.2	358.9
Zamiaceae	156.1	107.0	207.9	183.5	159.8	236.3
Zaminae	107.8	74.3	147.7	143.7	118.8	187.3
Encephalartinae	56.2	39.3	82.1	81.3	63.2	111.5
Ceratozamia-Stangeria	84.9	55.0	118.9	131.0	112.1	160.4
Encephalartos-Lepidozamia	39.1	33.9	55.0	67.2	51.7	102.1
Zamia-Microcycas	57.0	34.2	84.5	90.5	65.5	119.3
Cycas	17.6	10.1	29.1	24.4	18.1	40
Bowenia	5.5	0.8	14.6	32.8	18.7	47.2
Dioon	15.4	7.5	24.6	32.1	19.4	50.9
Ceratozamia	19.2	9.5	33.2	22.2	16.5	31.6
Encephalartos	10.5	6.2	16.3	21.1	14.2	35.45
Macrozamia	9.1	5.1	15.3	18.6	11.7	28.8
Lepidozamia	10.9	3.2	23.1	25.3	15.7	34.3
Zamia	14.6	9	22.1	26.3	18.4	37.2

For the latter study, only ages from the analysis with the full fossil dataset and the birth-death tree prior are reported. Min and max ages represent the upper and lower boundary of the 95% Highest Posterior Density. Ma, million years ago.

Historical biogeography and latitudinal distribution through time

When including both extant and extinct cycads (Fig. S6), the DEC analyses estimated a most likely ancestral origin for Cycadales in Western and Eastern Palearctic (Fig. 4, relative probability = 0.192). The second-best inference included only Western Palearctic (relative probability = 0.166), and the third best encompassed Western-Eastern Palearctic and Greenland (relative probability = 0.140). Only geographic areas of Laurasia are recovered as part of the estimated ancestral ranges for the cycad origin (cumulative relative probability = 0.788). When including only extant cycads (Fig. S7), the DEC analyses found a widely Pangean origin with the most likely range comprising Western–Eastern Palearctic, Eastern Nearctic, Africa, and Antarctica. Without fossils, there was higher uncertainty at the root ancestral range. There were 22 and 10 ‘equally likely’ ancestral ranges (i.e. ranges having a log-likelihood lower than 2 units compared with the most likely range) for the root when excluding and including fossils, respectively. Generally, the discrepancies of biogeographic inferences between the two analyses are found in the deep nodes (e.g. crowns of Zamiaceae, Zaminae, and Encephalartinae; Fig. 4; Table 2). These results were robust when

Table 2 Ancestral ranges for main cycad clades as estimated with the biogeographic analyses under the Dispersal–Extinction–Cladogenesis model (14 areas and four time slices) applied to the phylogeny including extant species only or extant and extinct species.

		Ancestral ranges estimated with extant and extinct cycads				Ancestral ranges estimated with extant cycads only			
Cycad clades	Median age (Ma)	Best	Second	Third	Interpretation	Best	Second	Third	Interpretation
Cycadales	330.4	WP + EP	WP	WP + EP + GR	Laurasia	WP + EP + EN + AF + AN	WP + EN + AF + AU + AN	WP + EP + EN + AF	Pangea
Zamiaceae	183.5	WP + EN + CA + SA + GR	WP + EN + WI + SA + GR	WP + EN + AF + GR + AN	Pangea	CA + AF + IN+AU + AN	CA + WI + AF + AU + AN	EN + CA + AF + AU + AN	Gondwana
Zaminae	143.7	WP + EN + CA + GR	WP + EN + WI + GR	WP + EN + CA + WI + GR	Laurasia	CA + AF	CA + WI + AF	CA + AF + AN	Western Gondwana
Encephalartinae	81.3	EN + WI + SA + AU + AN	CA + WI + SA + AU + AN	WI + SA + AU + AN	Western Gondwana	AU	–	–	Australasia
Ceratozamia-Stangeria	131.0	WP	WP + GR	WP + EN + GR	Laurasia	CA + AF	–	–	Western Gondwana
Encephalartos-Lepidozamia	67.2	EN + WI + SA + AU + AN	CA + WI + SA + AU + AN	–	Western Gondwana	AU	–	–	Australasia
Zamia-Microcycas	90.5	CA + WI	CA	WI	Central America	CA + WI	CA	WI	Central America
Cycas	24.4	EP + WA	WA	–	Eastern Asia	WA	EP + WA	–	Eastern Asia
Bowenia	32.8	AU	–	–	Australasia	AU	–	–	Australasia
Dioon	32.1	CA	–	–	Central America	CA	–	–	Central America
Ceratozamia	22.2	WP + EP + WN + CA	–	–	Holarctic	CA	–	–	Central America
Encephalartos	21.1	EP + AF + WA + AU	EP + AF + WA	–	Paleotropics	EP + AF + WA + AU	EP + AF + WA	–	Paleotropics
Macrozamia	18.6	AU	–	–	Australasia	AU	–	–	Australasia
Lepidozamia	25.3	AU	–	–	Australasia	AU	–	–	Australasia
Zamia	26.3	CA	–	–	Central America	CA	CA + WI	–	Central America

For each clade, the most likely ancestral range (Best) as well as the second and third most likely range are reported. A range can be composed of several areas (denoted with +). Area abbreviations defined as follows: WP, West Palearctic; EP, East Palearctic; WN, West Nearctic; EN, East Nearctic; CA, Central America; WI, Caribbean Islands; SA, South America; AF, Africa; IN, India; WA, Southeast Asia; AU, Australasia; GR, Greenland; AN, Antarctica. The Interpretation column translates the estimated ancestral ranges into major biogeographical and geological entities. Ma, million years ago.

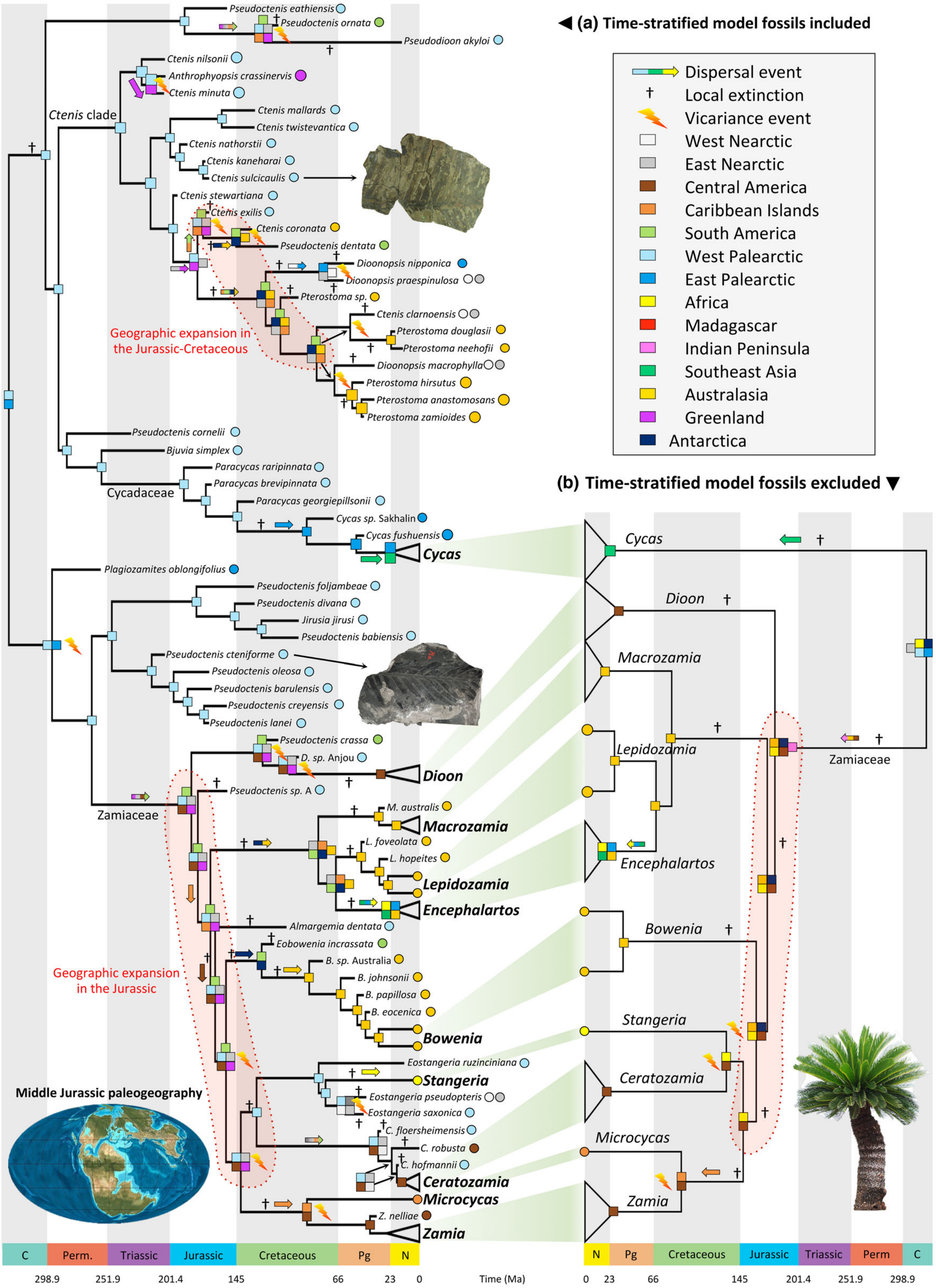

Fig. 4 Historical biogeography of cycads. Estimates of ancestral areas were performed with a time-stratified model in Dispersal–Extinction–Cladogenesis (DEC) with fossils included (a) and with fossils excluded (b). The extant genera have been collapsed to focus on deep-time biogeography (for details within each genus, see Supporting Information Fig. S6 when fossils are included and Fig. S7 when fossils are excluded). The bottom-right corner legend indicates colored areas used in this study corresponding to colored squares for each node, representing inferred ancestral area(s) with the DEC model, and colored circles for fossil species representing known distributions (except for extant genera with one or two species only). The red-highlighted shades show the cycad expansion into Gondwana during the Jurassic and Cretaceous. The bottom-left corner map represents the global paleogeography in the Jurassic (180 million years ago (Ma)). Paleomap used with permission © 2020 Colorado Plateau Geosystems Inc. Arrows indicate fossil species illustrated. Pictures from Mario Coiro. C, Carboniferous; N, Neogene; Perm, Permian; Pg, Paleogene.

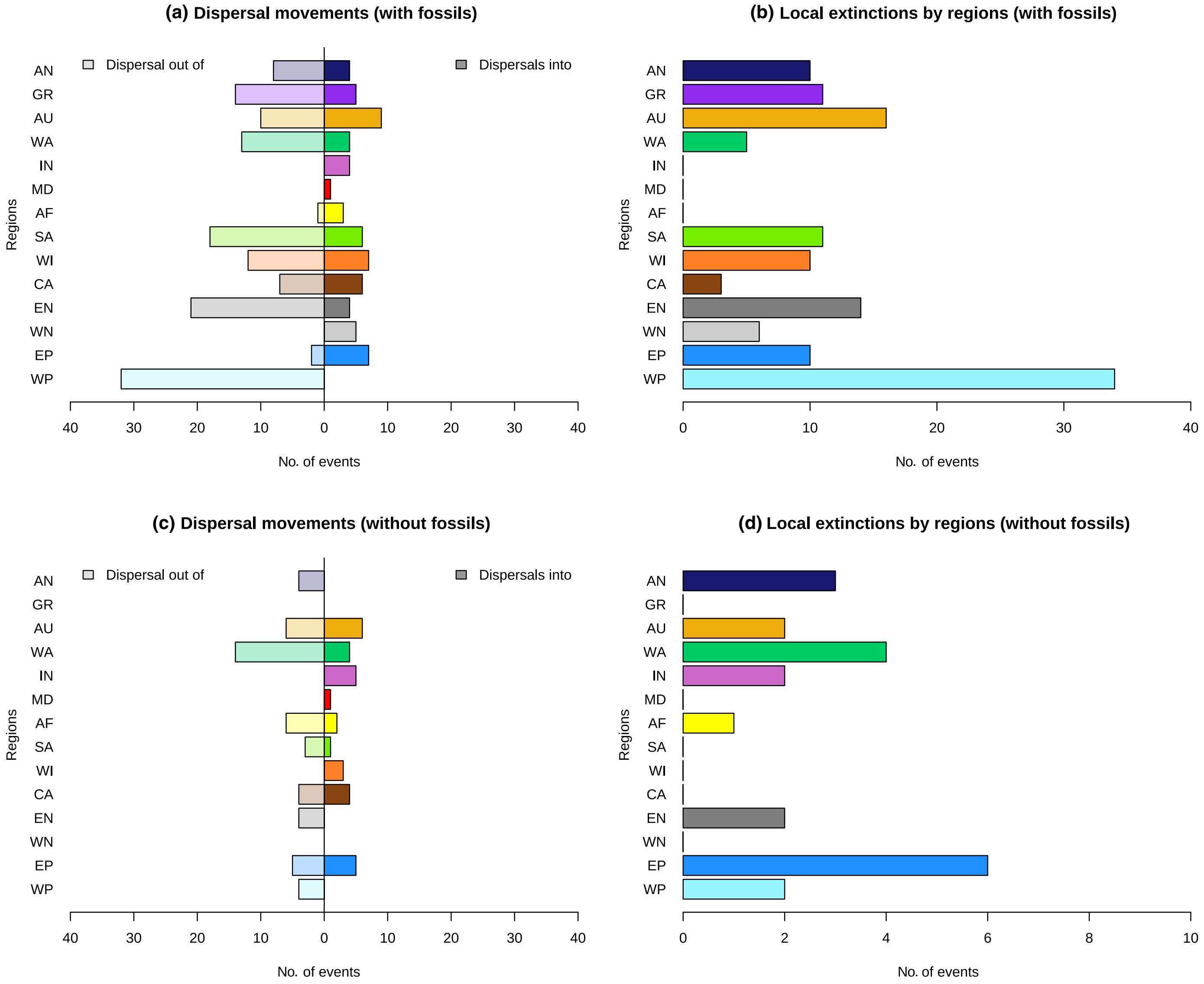

Fig. 5 Biogeographic processes explaining the global distribution pattern of cycads. The number of dispersal events into a region and out of a region as well as the number of local extinctions (extirpations) are compared between analyses including fossils (a, b) and analyses excluding fossils (c, d). Area names: WP, West Palearctic; EP, East Palearctic; WN, West Nearctic; EN, East Nearctic; CA, Central America; WI, Caribbean Islands; SA, South America; AF, Africa; IN, India; WA, Southeast Asia; AU, Australasia; GR, Greenland; AN, Antarctica.

coding fossil geographic ranges with missing data instead of true absence (Figs S8, S9) and with uncertainties in fossil placements and age estimates (Fig. S10).

Using ancestral estimates, we extracted biogeographic processes (dispersals, local extinctions (= extirpations), and vicariance) explaining the geographic range evolution in cycads. Since determining which area the dispersal event came from with an ancestral range including two or more areas is challenging, we counted one dispersal event per area of the ancestral range that likely leads to an overestimation of dispersal events. Comparing biogeographic processes between the phylogenies including or excluding fossils, we found differences in the extent and number of events (Fig. 5). Dispersal events were separated into two categories: dispersal into a region and dispersal out of a region (Fig. 5a,c). We estimated 25 dispersals into high-latitude regions with fossils included, but only 5 when fossils are excluded. We further found 40 dispersals into low-latitude regions compared with 26 when fossils are excluded. Moreover, we estimated 77 dispersals out of high-latitude regions with fossils included, while only 17 when fossils are excluded. We also recovered 61 dispersals out of low-latitude regions compared with 33 when fossils are excluded. Despite these discrepancies, there are also similarities such as the role of the Indomalayan region as a source of dispersal as well as Central and South America to a lesser extent, or East Palearctic and Australasia as sinks of diversity.

The biogeographic analyses revealed numerous vicariance events (Fig. 4), with 32 events inferred with fossils and 20 events without fossils. Local extinctions, or extirpations (Figs 5b,d, S11), are fifth as often inferred when fossils are included (130) than when they are excluded (22). Among extirpations, we found

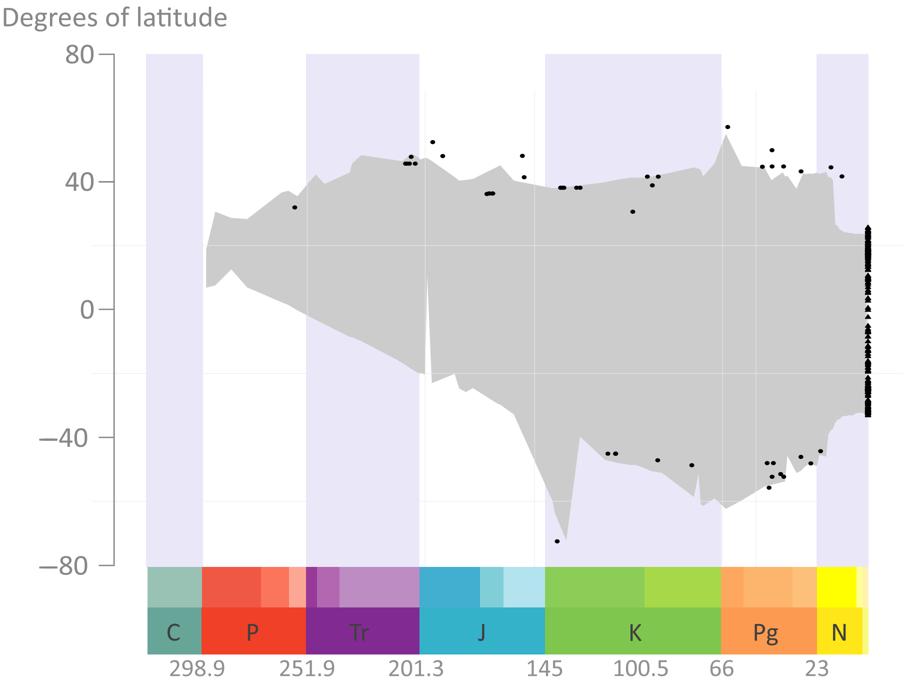

Fig. 6 Reconstruction of the latitudinal span of cycads during geological times. This estimate has been obtained using the method of Silvestro *et al*. (2018b) as implemented in Zhang *et al*. (2022). Gray polygon represents the reconstructed span, black dots represent the paleolatitude and age of fossil tips, while black triangles show the latitude of extant species. C, Carboniferous; J, Jurassic; K, Cretaceous; N, Neogene; P, Permian; Pg, Paleogene; Tr, Triassic.

85 high-latitude and 45 low-latitude extirpations when including fossils, while there are only 13 and 9 when fossils are excluded.

The analysis of latitudinal distribution indicates a mid-to-low northern latitude origin of the Cycadales (Fig. 6). A similar latitudinal range is maintained until the Late Jurassic, when the group spreads to high latitudes in the Southern Hemisphere. During the Cretaceous and Paleogene, the group persists at mid-to-high latitudes in both Hemispheres. A contraction between the end of the Paleogene (Oligocene) and the mid-Miocene leads to the current subtropical to tropical distribution.

Discussion

Our results show that the integration of fossils as tips in a total-evidence phylogeny of cycads provides a more complete view of their macroevolutionary history. Here, we discuss the consequence of our results on cycad phylogeny and how this new phylogenetic hypothesis informs us on the biogeographic history processes explaining global cycad biodiversity.

Phylogeny of cycads

Our study shows the feasibility and advantages of the total-evidence approach for cycads, and it allows us to reach insights that would be impossible using other approaches. Our analyses provide new insights into long-standing issues with the cycad phylogeny. Most Triassic and Early to Middle Jurassic cycad leaves are not particularly closely related to extant cycad families, contrary to previous suggestions based on morphology (Hermsen *et al.*, 2006; Martínez *et al.*, 2012) or morphology + molecules (Coiro & Pott, 2017). Previous studies may have been swayed by the coding of polyphyletic genera as single taxa, and the inclusion of only a few characters relating to leaf morphology and anatomy. The polyphyletic nature of fossil leaf genera in Cycadales is revealed in our analyses, particularly for the genus *Pseudoctenis* (Pott *et al.*, 2007). This suggests that the use of generic-level analyses as a proxy for species-level dynamics (Roy *et al.*, 1996) should be viewed with caution. Coding of taxa at the level of single species and the implementation of the FBD prior could represent a more promising avenue in groups with a poorly understood fossil record and without many whole-plant reconstructions, a widespread situation in land plants.

Our results validate some previous suggestions of relatedness between fossil taxa and extant genera (Hill, 1978; Carpenter, 1991; Kvaček, 2002; Su *et al.*, 2014; Erdei *et al.*, 2018; Hill *et al.*, 2019a). Previously puzzling forms, such as *Eostangeria* (Barthel, 1976; Kvaček & Manchester, 1999; Uzunova *et al.*, 2001), appear to represent transitional forms between the *Zamia-Microcycas* clade and the extremely derived genus *Stangeria*, combining the epidermal anatomy of the former with the unique macromorphology of the latter (i.e. leaflets with pinnate venation). This contrasts with the idea of *Eostangeria* as a separate lineage in the Zamiaceae that converged with *Stangeria*. The tree also confirms the relationship between the Mesozoic genus *Paracycas* Harris and extant *Cycas*. The genus *Eobowenia* is confirmed as sister to extant *Bowenia*, strengthening inferences on its understory habit based on comparison with other fossil and extant species of this lineage (Hill *et al.*, 2019b).

Another important result is the emergence of the hypothesis of an entirely extinct clade spanning the Triassic to the Miocene, including both common Mesozoic forms assigned to *Ctenis* and controversial Cenozoic fossils such as *Pterostoma* (Florin, 1933; Harris, 1964). The morphology of the leaves in this clade differs from that of the two extant families mostly in the presence of H or N anastomoses. This clade includes *Dioonopsis* validating results based on comparative investigations (Barone Lumaga *et al.*, 2015; Erdei & Manchester, 2015), and weighing strongly against its use as a calibration for the age of *Dioon* (e.g. Li *et al.*, 2019).

The phylogenetic distance between most fossil cycads and the extant clades argues against the use of extant cycadalean ecophysiology to infer past preferences of fossil cycads (McElwain, 1998), especially for members of the *Ctenis* clade. Indeed, the only *Ctenis* species for which we have a good paleoecological understanding (Wing *et al.*, 1993) occupied a fern-dominated open habitat on a peaty substrate, unlike any extant cycad. More studies are needed to better understand the ecology of fossil cycads and provide clues into the causes of their extinction.

Origin and early historical biogeography of cycads

Our dating analyses first show a more ancient Paleozoic age for the origin of Cycadales (Carboniferous vs Permian or even Triassic in molecular-only studies; Nagalingum *et al.*, 2011; Salas-Leiva *et al.*, 2013; Condamine *et al.*, 2015). Furthermore, our results indicate a much wider range in the ages of crown-group genera (Fig. 3b). Despite clear limits in our study (to be described later), this new phylogenetic framework allows the estimation of the historical biogeography and latitudinal range evolution of Cycadales. Nonetheless, we found that the ancestral state estimates were robust when coding fossil geographic ranges with missing data instead of true absence and when considering uncertainties in fossil placements and divergence times.

To our knowledge, only one study attempted to reconstruct the geographic origin of cycads with phylogenetic approaches, showing an origin in Australia, China, and Mexico (Salas-Leiva *et al.*, 2013). However, such a geographic range is unlikely given the nonadjacency of the three areas, an issue probably due to the lack of a time-stratified geographic model. Here, by including Antarctica and Greenland with now-vanished continental connections, we relied on more accurate information from a paleogeographic perspective. We also compared the effect of including the fossils into the ancestral estimates. We show important discrepancies between the analyses including or not the fossils as tips, particularly in deep nodes. This could be expected given the long branches subtending radiations of extant genera bearing little phylogenetic information in deep times, supporting the view of Crisp *et al.* (2011) that fossil lineages are crucial to study historical biogeography. However, treating fossil distributions as unknown did not impact the analysis dramatically. This could be due to the high endemism of cycad species (only 11 species out of 381 have more than one area for their range), thus limiting the possibility of unobserved presence of widespread fossil species.

Cycadales likely originated in the northern part of Pangea (Laurasia). This result agrees with the most ancient cycad fossil lineage, *Crossozamia* Pomel, described from the Permian of China (Gao & Thomas, 1989). The Carboniferous was a time of active mountain-building as the supercontinent Pangea coalesced (von Raumer *et al.*, 2003). The southern continents were united as Gondwana, which collided with the West Palearctic and East Nearctic. This collision resulted in the Hercynian orogeny in Palearctic, and the Alleghenian orogeny in Nearctic (von Raumer *et al.*, 2003). These mountain ranges could have limited cycad distribution in northern Pangea. Contemporaneously, much of present eastern Eurasian plate welded itself to the West Palearctic along the Ural Mountains, which makes plausible an eastern Laurasian origin for cycads. During the Carboniferous, global climate progressively cooled and dried, eventually culminating with an extinction period known as the Carboniferous rainforest collapse that drastically affected terrestrial biodiversity in Laurasia (Sahney *et al.*, 2010) of which cycads survived. Cycads remained in Laurasia (including Greenland), potentially constrained by the Central Pangean Mountains in the Permian. By the Middle Triassic, the Central Pangean mountains had been substantially reduced in size, and by the earliest Jurassic *c.* 200 Ma the Pangean range in West Palearctic was reduced to upland areas surrounded by marine basins (Scotese & Schettino, 2017). In the Early Jurassic, our inferences suggest Zamiaceae and the *Ctenis* clade independently colonized Gondwana through East Nearctic and Central America or Caribbean going to South America.

The role of Antarctica and Greenland

From the Jurassic until the end-Cretaceous, cycads expanded to all continents through active dispersals into low-latitude regions of Gondwana. Four ancient independent colonizations of Antarctica were inferred between the Late Jurassic and Early Cretaceous (two in the *Ctenis* clade between 174 and 147 Ma and 179 and 123 Ma; *Eobowenia* and *Bowenia* between 155 and 130 Ma; Encephalartinae between 170 and 81 Ma). In line with the dense fossil record of Patagonia (Artabe & Stevenson, 1999; Cúneo *et al.*, 2010), our analyses indicate West Gondwana is a biogeographic crossroad for cycads with numerous (18) dispersals out of South America. At that time, South America was connected to Antarctica, itself connected to Australia (Blakey, 2008; Seton *et al.*, 2012; Kocsis & Scotese, 2021), thus creating a large land bridge for terrestrial biodiversity.

The case of Encephalartinae illustrates well the impact of fossils and how they help clarify their origins. The African genus *Encephalartos* is sister to *Lepidozamia* endemic to Australia, separated at the K/Pg boundary (67 Ma), and both are sister to *Macrozamia*, which is also restricted to Australia and diverged in the Late Cretaceous (81 Ma). Both *Lepidozamia* and *Macrozamia* have fossil taxa distributed in Australia, and the Patagonian *Austrozamia stockeyi* Wilf, D.Stevenson et Cuneo is thought to be related to the *Encephalartos*-*Lepidozamia* clade (Wilf *et al.*, 2016). However, no fossil taxon is known to be closely related to *Encephalartos*. Our analyses suggest a widespread ancestor in Gondwana (including Antarctica) for Encephalartinae and the common ancestor of *Encephalartos* and *Lepidozamia*. However, this wide range does not include Africa, which was not connected to other Gondwanan continents in the Late Cretaceous. We estimated the colonization of Africa along the stem of *Encephalartos* likely started from Australia and extended northward to Southeast Asia and passing through the East Palearctic, followed by numerous Paleogene extirpations. If correct, we can expect to discover fossil taxa possibly related to *Encephalartos* in Cenozoic deposits of Eurasia. However, the presence of fossils with potential affinities with Encephalartinae in South America could also indicate a Western Gondwanan route, with dispersal to Africa via Antarctica supported by the South African origin of *Encephalartos* (Mankga *et al*., 2020a). The lack of stomatal morphology in these South American fossils unfortunately does not allow testing of this hypothesis. Interestingly, the analyses without fossils recover the same range for *Encephalartos*, but an Australasian origin for Encephalartinae.

Likewise, Greenland was key in cycad biogeography. We recovered five independent colonizations from the Late Triassic (two in the *Ctenis* clade) to the Early Jurassic (in Zamiaceae). Greenland was particularly prominent in ancestral range estimates for Zamiaceae and acted as biogeographic crossroads for all lineages of the backbone (14 dispersals out of Greenland). Our study highlights the role of Antarctica and Greenland in biogeographic estimates as proposed previously (de la Estrella *et al*., 2019). Indeed, Antarctica and Greenland were, separately, sources of common ancestors until the Paleocene–Eocene transition, after which the local climate became unfavorable for plant growth (Pross *et al*., 2012; Suan *et al*., 2017; Klages *et al*., 2020). Importantly, we estimated that nine vicariance events out of a total of 31 involve Antarctica or Greenland, when added together, with subsequent disjointed descending lineages. The recognition of the role of Antarctica and Greenland in historical biogeography has rarely been implemented in macroevolutionary analyses. We argue that the incorporation of fossils allows recovering the role of these continents, which is otherwise challenging without them. However, our analyses excluding fossils did recover the role of Antarctica in the early nodes of Zamiaceae, but not of Greenland.

Our results including or excluding the fossils indicate similar geographic origins for the crown of extant genera, except for *Cycas* and *Ceratozamia*. We find different ancestral ranges because of the inclusion of one East Palearctic and two West Palearctic fossils, respectively. For *Ceratozamia*, the discrepancy is stronger because *C. hofmannii* (early Miocene of Austria) is found within the extant genus, unlike a stem lineage for *C. floersheimensis* (early Oligocene of Germany). Hence, we estimated a broad Laurasian range from the early Oligocene to early Miocene followed by geographic extirpations in the Neogene. Although the impact of fossils is less obvious toward the present, extinct taxa bear direct evidence of the presence of related extant representatives that can occur in different regions. Including them can alter the inference of evolutionary processes explaining their current distribution pattern.

Within the most species-rich cycad genera, *Cycas* and *Zamia* show a dynamic biogeographic history with multiple area colonizations by range expansions and several vicariance events (11 for *Cycas* and 6 for *Zamia*). *Zamia* colonized the Caribbean Islands during the second emersion of the Aves Ridge 16 Ma (Garrocq *et al*., 2021). Our analyses underestimate the number of allopatric events because of the broad-scale analyses with continents as biogeographic units. Finer-scale studies have already unveiled evolutionary processes explaining extant species distribution within genera and showing the role of allopatry (Calonje *et al*., 2019; Mankga *et al*., 2020b; Habib *et al*., 2022).

Evolution of the latitudinal gradient of cycad biodiversity

We complemented historical biogeography with ancestral latitude estimates. Although the origin of cycads at high latitude might be due to the poor record in tropical and equatorial latitudes, our analysis clearly shows that cycads were heavily impacted at the end of the Paleogene and in the mid-Miocene (40 high-latitude extinction events vs 22 low-latitude extinction events, Fig. S10). This period corresponds with the extinction of the *Ctenis* clade, as well as the disappearance of many relatives of extant genera from Europe and North America. Interestingly, many Cenozoic species of the *Ctenis* clade are distributed at high latitudes (*Pterostoma* in Australia and New Zealand, *Dioonopsis* in Alaska). This pattern is reminiscent of the high-latitude refugium hypothesis advanced for many fossil and extant groups (Bomfleur *et al*., 2018), including ferns that cooccurred with *Ctenis* in the past (Wing *et al*., 1993). Testing such a hypothesis on mechanistic grounds could represent an interesting avenue for further investigations.

Elevated extinction of gymnosperms during the Cenozoic has been proposed as a partial explanation for the imbalance in species richness between gymnosperms and angiosperms (Crisp & Cook, 2011; Condamine *et al*., 2020), and are retrieved in analyses of both phylogenies (May *et al*., 2016) and the fossil record (Crepet & Niklas, 2009). Whether such an elevated extinction was driven by one or more events remains hard to infer, but phylogeny-based diversification models unveil an increasing extinction rate in cycads, with diversity decline starting 125 Ma (Mazet *et al*., 2022). The succession of many cooling events, in the Maastrichtian (Linnert *et al*., 2014), at the Eocene–Oligocene transition (Houben *et al*., 2012), the Oligocene–Miocene transition (Beddow *et al*., 2015), and the mid-Miocene transition, leading to an Icehouse Earth, could have led high-latitude cycads to extinction, and shifted the distribution of the group to more subtropical and tropical latitudes (Meseguer & Condamine, 2020). The maximum latitudinal expansion of the group broadly corresponds to a period of Greenhouse Earth with terrestrial ecosystems extending to Antarctica and Greenland (Suan *et al*., 2017; Klages *et al*., 2020).

Such periods of elevated extinction offer a better explanation for the patterns of diversification of the extant genera. Even if the origins of the crown groups of the extant genera appear not to be entirely synchronous, the ages retrieved in the tree are still relatively young. If this would go against the idea of a global 'trigger' for the radiation of the genera, it would be compatible with a period of elevated turnover. This would indicate that cycad evolution has been mainly influenced by abiotic factors (Barnosky, 2001; Benton, 2009). However, the influence of biotic factors cannot be fully excluded: The distribution of the members of the *Ctenis* clade in high-latitude environments might be a consequence of competitive exclusion from lower-latitude environments, which might have predisposed the clade to eradication by Neogene climatic change.

Conclusion

Although our dated phylogeny includes a fraction of the known fossil cycad diversity, this study represents a step forward in our understanding of the global cycad biodiversity. Cycads originated in the Carboniferous on the Laurasian landmass and colonized Gondwana in the Jurassic. We revealed the crucial role of Antarctica and Greenland as biogeographic crossroads and found numerous vicariance events in the deep and recent past. Cycad latitudinal range increased in the Jurassic but retreated toward subtropical latitudes in the Neogene due to important high-latitude extirpations. However, we remain cautious about our biogeographic inferences. For instance, we were unable to incorporate taxa like *Antarcticycas* (Triassic of Antarctica) or *Crossozamia* (Permian of China) because of the lack of preserved leaf traits. We think that including more fossil cycad taxa can alter phylogenetic relationships, which in turn can affect divergence times and biogeographic inferences. Further studies are needed to incorporate new cycad taxa to deepen our understanding of cycad origin and range evolution. This species-based analysis represents a starting point for building a more species-rich and character-rich matrix that avoids the pitfall of previous analyses, namely the use of polyphyletic operational taxonomic units. Such an effort would require the reinvestigation of the fossil record of other cycad organs, and the establishment of more whole-plant reconstructions for this group.

Acknowledgements

We thank Michael R. May, Ixchel González-Ramírez, Yaowu Xing, and Jason Hilton for providing constructive comments that improved the study. MC has been supported by the Austrian Science Fund (FWF), Lise Meitner project no. M3168. The visit of MC to the NRM and the NHM was funded by the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 226506 (SYNTHESYS project SE-TAF-5774 and project GB-TAF-5763). The computational results presented have been achieved in part using the Vienna Scientific Cluster (VSC). This project benefited from an ‘Investissements d’Avenir’ program managed by the Agence Nationale de la Recherche (CEBA, ref. ANR-10-LABX-25-01) and the ANR GAARAnti project (ANR-17-CE31-0009). MC acknowledges Peta Hayes for the assistance with the collections at the NHM, and Christian Pott for assistance with the collections at the NRM. MC also wishes to thank Maria Rosaria Barone Lumaga, Christian Pott, and Boglarka Erdei, for discussion about the evolutionary history of cycads. We also thank the Montgomery Botanical Center for the use of some images used as silhouettes. James Doyle is thanked for comments on a previous version of the manuscript. FLC and NM thank Iago Bonnici for help in developing DECX. We dedicate this study to the memory of our esteemed colleague Nathalie S. Nagalingum.

Competing interests

None declared.

Author contributions

MC and FLC conceived the project. MC, RA, LJS and FLC extracted data for the analyses. MC, RA, NM, LJS and FLC analyzed data. MC and FLC wrote the first draft of the manuscript and prepared figures. All authors edited the manuscript.

ORCID

Rémi Allio https://orcid.org/0000-0003-3885-5410
Mario Coiro https://orcid.org/0000-0002-0113-0320
Fabien L. Condamine https://orcid.org/0000-0003-1673-9910
Nathan Mazet https://orcid.org/0000-0002-6611-0312
Leyla J. Seyfullah https://orcid.org/0000-0002-0199-9923

Data availability

The data that support the findings of this study and code used in this study are openly available in FigShare at doi: 10.6084/m9.figshare.21399576.v1.

References

Antonelli A, Zizka A, Carvalho FA, Scharn R, Bacon CD, Silvestro D, Condamine FL. 2018. Amazonia is the primary source of Neotropical biodiversity. *Proceedings of the National Academy of Sciences, USA* **115**: 6034–6039.

Archangelsky S, Baldoni A. 1972. Notas sobre la flora de la zona de Tico, provincia de Santa Cruz. X. Dos nuevasespecies de *Pseudoctenis* (Cycadales). *Ameghiniana* **9**: 241–257.

Artabe AE, Stevenson DWM. 1999. Fossil Cycadales of Argentina. *The Botanical Review* **65**: 219–238.

Barnosky AD. 2001. Distinguishing the effects of the Red Queen and Court Jester on Miocene mammal evolution in the northern Rocky Mountains. *Journal of Vertebrate Paleontology* **21**: 172–185.

Barone Lumaga MR, Coiro M, Truernit E, Erdei B, De Luca P. 2015. Epidermal micromorphology in *Dioon*: did volcanism constrain *Dioon* evolution? *Botanical Journal of the Linnean Society* **179**: 236–254.

Barthel M. 1976. Eozane Floren des Geiseltales: Farne und Cycadeen. Abhandlungen des Zentralen Geologischen Institutes. *Paläontologische Abhandlungen* **26**: 439–498.

Bateman RM, Hilton J. 2009. Palaeobotanical systematics for the phylogenetic age: applying organ-species, form-species and phylogenetic species concepts in a framework of reconstructed fossil and extantwhole-plants. *Taxon* **58**: 1254–1280.

Bateman RM, Simpson NJ. 1998. Comparing phylogenetic signals from reproductive and vegetative organs. In: Owens SJ, Rudall PJ, eds. *Reproductive biology*. Kew, UK: Royal Botanic Gardens, 231–253.

Beaulieu JM, O’Meara BC. 2015. Extinction can be estimated from moderately sized molecular phylogenies. *Evolution* **69**: 1036–1043.

Beddow HM, Liebrand D, Sluijs A, Wade BS, Lourens LJ. 2015. Global change across the Oligocene–Miocene transition: high-resolution stable isotope records from IODP Site U1334 (equatorial Pacific Ocean). *Paleoceanography* **31**: 81–97.

Beeravolu CR, Condamine FL. 2016. An extended maximum likelihood inference of geographic range evolution by dispersal, local extinction and cladogenesis. *bioRxiv*. doi: 10.1101/038695.

Benton MJ. 2009. The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. *Science* **323**: 728–732.

Blakey RC. 2008. Gondwana paleogeography from assembly to breakup – a 500 my odyssey. *Geological Society of America Special Papers* **441**: 1–28.

Bomfleur B, Blomenkemper P, Kerp H, McLoughlin S. 2018. Polar regions of the Mesozoic–Paleogene greenhouse world as refugia for relict plant groups. In: Krings M, Harper C, Cuneo NR, Rothwell GW, eds. *Transformative paleobotany.* Cambridge, MA, USA: Academic Press.

Butler RJ, Barrett PM, Kenrick P, Penn MG. 2009. Testing co-evolutionary hypotheses over geological timescales: interactions between Mesozoic non-avian dinosaurs and cycads. *Biological Reviews of the Cambridge Philosophical Society* **84**: 73–89.

Cai C, Escalona HE, Li L, Yin Z, Huang D, Engel MS. 2018. Beetle pollination of cycads in the Mesozoic. *Current Biology* **28**: 2806–2812.

Calonje M, Hernandez JC, Coca LF, Jaramillo D, Aristizabal A. 2021. Two new species of *Zamia* (Zamiaceae, Cycadales) from the Magdalena-Urabá moist forests ecoregion of northern Colombia. *Phytotaxa* **497**: 1–19.

Calonje M, Meerow AW, Griffith MP, Salas-Leiva DE, Vovides AP, Coiro M, Francisco-Ortega J. 2019. A time-calibrated species tree phylogeny of the New World cycad genus *Zamia* L. *International Journal of Plant Sciences* **180**: 286–314.

Calonje M, Stevenson DW, Osborne R. 2022. *The World list of Cycads, online edition [Internet].* [WWW document] URL www.cycadlist.org [accessed 01 November 2022].

Carpenter R. 1991. *Macrozamia* from the early Tertiary of Tasmania and a study of the cuticles of extant species. *Australian Systematic Botany* **4**: 433–444.

Chacón J, Renner SS. 2014. Assessing model sensitivity in ancestral area reconstruction using Lagrange: a case study using the Colchicaceae family. *Journal of Biogeography* **41**: 1414–1427.

Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L, Ram K. 2019. *RGBIF: interface to the global biodiversity information facility API.* R package v.1.2.0. [WWW document] URL https://CRAN.R-project.org/package=rgbif [accessed 10 May 2019].

Chernomor O, Von Haeseler A, Minh BQ. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology* **65**: 997–1008.

Coiro M, Barone Lumaga MR. 2018. Disentangling historical signal and pollinator selection on the micromorphology of flowers: an example from the floral epidermis of the Nymphaeaceae. *Plant Biology* **45**: 669–676.

Coiro M, Barone Lumaga MR, Rudall PJ. 2021. Stomatal development in the cycad family Zamiaceae. *Annals of Botany* **128**: 577–588.

Coiro M, Jelmini N, Neuenschwander H, Calonje MA, Vovides AP, Mickle JE, Lumaga MRB. 2020a. Evolutionary signal of leaflet anatomy in the Zamiaceae. *International Journal of Plant Sciences* **52**: 192–201.

Coiro M, Martínez LCA, Upchurch GR, Doyle JA. 2020b. Evidence for an extinct lineage of angiosperms from the Early Cretaceous of Patagonia and implications for the early radiation of flowering plants. *New Phytologist* **228**: 344–360.

Coiro M, Pott C. 2017. *Eobowenia* gen. nov. from the Early Cretaceous of Patagonia: indication for an early divergence of *Bowenia*? *BMC Evolutionary Biology* **17**: 97.

Condamine FL, Nagalingum NS, Marshall CR, Morlon H. 2015. Origin and diversification of living cycads: a cautionary tale on the impact of the branching process prior in Bayesian molecular dating. *BMC Evolutionary Biology* **15**: 65.

Condamine FL, Silvestro D, Koppelhus EB, Antonelli A. 2020. The rise of angiosperms pushed conifers to decline during global cooling. *Proceedings of the National Academy of Sciences, USA* **117**: 28867–28875.

Crepet WL, Niklas KJ. 2009. Darwin's second 'abominable mystery': why are there so many angiosperm species? *American Journal of Botany* **96**: 366–381.

Crisp MD, Cook LG. 2011. Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. *New Phytologist* **192**: 997–1009.

Crisp MD, Trewick SA, Cook LG. 2011. Hypothesis testing in biogeography. *Trends in Ecology & Evolution* **26**: 66–72.

Cúneo NR, Escapa IH, Villar de Seoane L, Artabe AE, Gnaedinger S. 2010. Review of the Cycads and Bennettitaleans from the Mesozoic of Argentina. In: Gee C, ed. *Plants in Mesozoic time: innovations, phylogeny, ecosystems (Ted Delevoryas festschrift).* Bloomington, IN, USA: Indiana University Press, 187–212.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**: e88.

Elgorriaga A, Aktinson BA. 2023. Cretaceous pollen cone with three-dimensional preservation sheds light on the morphological evolution of cycads in deep time. *New Phytologist* **238**: 1695–1710.

Erdei B, Akgün F, Barone Lumaga MR. 2009. *Pseudodioon akyoli* gen. et sp. nov., an extinct member of Cycadales from the Turkish Miocene. *Plant Systematics and Evolution* **285**: 33–49.

Erdei B, Calonje M, Hendy A, Espinosa N. 2018. A review of the Cenozoic fossil record of the genus *Zamia* L. (Zamiaceae, Cycadales) with recognition of a new species from the late Eocene of Panama – evolution and biogeographic inferences. *Bulletin of Geosciences* **93**: 185–204.

Erdei B, Coiro M, Miller I, Johnson KR, Griffith MP, Murphy V. 2019. First cycad seedling foliage from the fossil record and inferences for the Cenozoic evolution of cycads. *Biology Letters* **15**: 20190114.

Erdei B, Manchester SR. 2015. *Ctenis clarnoensis* sp. n., an unusual cycadalean foliage from the Eocene Clarno formation, Oregon. *International Journal of Plant Sciences* **176**: 31–43.

Erdei B, Manchester SR, Kvaček Z. 2012. *Dioonopsis* Horiuchi et Kimura Leaves from the Eocene of Western North America: a cycad shared with the Paleogene of Japan. *International Journal of Plant Sciences* **173**: 81–95.

de la Estrella M, Buerki S, Vasconcelos T, Lucas EJ, Forest F. 2019. The role of Antarctica in biogeographical reconstruction: a point of view. *International Journal of Plant Sciences* **180**: 63–71.

Ezcurra MD, Agnolín FL. 2012. A new global palaeobiogeographical model for the late Mesozoic and early Tertiary. *Systematic Biology* **61**: 553–566.

Fine PV, Ree RH. 2006. Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *American Naturalist* **168**: 796–804.

Florin R. 1933. *Studienüber die Cycadales des Mesozoikums, nebst Erörterungen über die Spaltöffnungsapparate der Bennettitales.* Stockholm, Sweden: Almqvist & Wiksells boktryckeri.

Fritz SA, Schnitzler J, Eronen JT, Hof C, Böhning-Gaese K, Graham CH. 2013. Diversity in time and space: wanted dead and alive. *Trends in Ecology & Evolution* **28**: 509–516.

Gao Z, Thomas BA. 1989. A review of fossil cycad megasporophylls, with new evidence of *Crossozamia* Pomel and its associated leaves from the lower Permian of Taiyuan, China. *Review of Palaeobotany and Palynology* **60**: 205–223.

Garrocq C, Lallemand S, Marcaillou B, Lebrun J, Padron C, Klingelhoefer F, Laigle M, Münch P, Gay A, Schenini L *et al.* 2021. Genetic relations between the Aves Ridge and the Grenada back-arc Basin, East Caribbean Sea. *Journal of Geophysical Research: Solid Earth* **126**: e2020JB020466.

Gavryushkina A, Heath TA, Ksepka DT, Stadler T, Welch D, Drummond AJ. 2017. Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Systematic Biology* **66**: 57–73.

Glos RA, Salzman S, Calonje M, Vovides AP, Coiro M, Gandolfo MA, Specht CD. 2022. Leaflet anatomical diversity in *Zamia* (Cycadales: Zamiaceae) shows little correlation with phylogeny and climate. *The Botanical Review* **88**: 1–16.

Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A. 2021. The World Checklist of vascular plants, a continuously updated resource for exploring global plant diversity. *Scientific Data* **8**: 215.

Greguss P. 1968. *Xylotomy of the living Cycads, with a description of their leaves and epidermis.* Budapest, Hungary: Akademiai Kiado.

Grimm GW, Kapli P, Bomfleur B, McLoughlin S, Renner SS. 2015. Using more than the oldest fossils: dating Osmundaceae with three Bayesian clock approaches. *Systematic Biology* **64**: 396–405.

Gutiérrez-Ortega J, Yamamoto T, Vovides AP, Angel Pérez-Farrera M, Martínez JF, Molina-Freaner F, Watano Y, Kajita T. 2018. Aridification as a driver of biodiversity: a case study for the cycad genus *Dioon* (Zamiaceae). *Annals of Botany* **121**: 47–60.

Habib S, Gong Y, Dong S, Lindstrom A, William Stevenson D, Liu Y, Wu H, Zhang S. 2022. Phylotranscriptomics reveal the spatio-temporal distribution and morphological evolution of *Macrozamia*, an Australian endemic genus of Cycadales. *Annals of Botany* **130**: 671–685.

Harris T. 1961. The fossil cycads. *Palaeontology* **4**: 313–323.

Harris TM. 1926. The Rhaetic flora of Scoresby Sound, east Greenland. *Meddelelser om Grønland* **68**: 43–147.

Harris TM. 1932. The Fossil Flora of Scoresby Sound, East Greenland: Part 2: Description of seed plants *incertae sedis* together with a discussion of certain cycadophyte cuticles. *Meddelelser om Grønland* **85**: 1–114.

Harris TM. 1964. *The Yorkshire Jurassic Flora. II. Caytoniales, Cycadales & Pteridosperms.* London, UK: British Museum (Natural History).

Heath TA, Huelsenbeck JP, Stadler T. 2014. The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences, USA* **111**: E2957–E2966.

Hermsen EJ, Taylor TN, Taylor EL, Stevenson DW. 2006. Cataphylls of the Middle Triassic cycad *Antarcticycas schopfii* and new insights into cycad evolution. *American Journal of Botany* **93**: 724–738.

Hill RS. 1978. Two new species of *Bowenia* Hook, ex Hook, f. from the Eocene of eastern Australia. *Australian Journal of Botany* **26**: 837–846.

Hill RS, Hill KE, Carpenter RJ, Jordan GJ. 2019a. New macrofossils of the Australian cycad *Bowenia* and their significance in reconstructing the past morphological range of the genus. *International Journal of Plant Sciences* **180**: 128–140.

Hill KE, Hill RS, Watling JR. 2019b. Pinnule and stomatal size and stomatal density of living and fossil *Bowenia* and *Eobowenia* specimens give insight into physiology during Cretaceous and Eocene paleoclimates. *International Journal of Plant Sciences* **180**.

van Hinsbergen DJJ, de Groot LV, van Schaik SJ, Spakman W, Bijl PK, Sluijs A, Langereis CG, Brinkhuis H. 2015. A paleolatitude calculator for paleoclimate studies (model v.2.1). *PLoS ONE* **10**: e0126946.

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* **35**: 518–522.

Horiuchi J, Kimura T. 1987. *Dioonopsis nipponica* gen. et sp. nov., a new cycad from the Palaeogene of Japan. *Review of Palaeobotany and Palynology* **51**: 213–225.

Houben AJP, van Mourik CA, MTri A, Coccioni R, Brinkhuis H. 2012. The Eocene–Oligocene transition: changes in sea level, temperature or both? *Palaeogeography, Palaeoclimatology, Palaeoecology* **335–336**: 75–83.

Huelsenbeck JP, Larget B, Alfaro ME. 2004. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular Biology and Evolution* **21**: 1123–1133.

Hunt G, Slater G. 2016. Integrating paleontological and phylogenetic approaches to macroevolution. *Annual Review of Ecology, Evolution, and Systematics* **47**: 189–213.

IUCN. 2022. The IUCN Red List of Threatened Species. v.2022-1.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods* **14**: 587–589.

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

King B, Qiao T, Lee MSY, Zhu M, Long JA. 2017. Bayesian morphological clock methods resurrect placoderm monophyly and reveal rapid early evolution in jawed vertebrates. *Systematic Biology* **66**: 499–516.

Klages JP, Salzmann U, Bickert T, Hillenbrand C-D, Gohl K, Kuhn G, Bohaty SM, Titschack J, Müller J, Frederichs T *et al.* 2020. Temperate rainforests near the South Pole during peak Cretaceous warmth. *Nature* **580**: 81–86.

Klopfstein S, Spasojevic T. 2019. Illustrating phylogenetic placement of fossils using RoguePlots: an example from ichneumonid parasitoid wasps (Hymenoptera, Ichneumonidae) and an extensive morphological matrix. *PLoS ONE* **14**: e0212942.

Kocsis ÁT, Scotese CR. 2021. Mapping paleocoastlines and continental flooding during the Phanerozoic. *Earth-Science Reviews* **213**: 103463.

Kvaček Z. 2002. A new Tertiary *Ceratozamia* (Zamiaceae, Cycadopsida) from the European Oligocene. *Flora* **197**: 303–316.

Kvaček Z, Manchester S. 1999. *Eostangeria* Barthel (extinct Cycadales) from the Paleogene of western North America and Europe. *International Journal of Plant Sciences* **160**: 621–629.

Larson-Johnson K. 2016. Phylogenetic investigation of the complex evolutionary history of dispersal mode and diversification rates across living and fossil Fagales. *New Phytologist* **209**: 418–435.

Lehtonen S, Silvestro D, Karger DN, Scotese CR, Tuomisto H, Kessler M, Pena C, Wahlberg N, Antonelli A. 2017. Environmentally driven extinction and opportunistic origination explain fern diversification patterns. *Scientific Reports* **7**: 4831.

Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50**: 913–925.

Li H-T, Yi T-S, Gao L-M, Ma P-F, Zhang T, Yang J-B, Gitzendanner MA, Fritsch PW, Cai J, Luo Y *et al.* 2019. Origin of angiosperms and the puzzle of the Jurassic gap. *Nature Plants* **5**: 461.

Linnert C, Robinson SA, Lees JA, Bown PR, Pérez-Rodríguez I, Petrizzo MR, Falzoni F, Littler K, Arz JA, Russell EE. 2014. Evidence for global cooling in the Late Cretaceous. *Nature Communications* **5**: 1–7.

Liu Y, Wang S, Li L, Yang T, Dong S, Wei T, Wu S, Liu Y, Gong Y, Feng X *et al.* 2022. The *Cycas* genome and the early evolution of seed plants. *Nature Plants* **8**: 389–401.

Maddison WP, Maddison D. 2021. *Mesquite: a modular system for evolutionary analysis.* v.3.70. [WWW document] URL www.mesquiteproject.org [accessed 15 May 2022].

Mankga LT, Yessoufou K, Chitakira M. 2020a. On the origin and diversification history of the African genus *Encephalartos. South African Journal of Botany* **130**: 231–239.

Mankga LT, Yessoufou K, Mugwena T, Chitakira M. 2020b. The cycad genus *Cycas* may have diversified from Indochina and occupied its current ranges through vicariance and dispersal events. *Frontiers in Ecology and Evolution* **8**: 44.

Mao K, Milne RI, Zhang L, Peng Y, Liu J, Thomas P, Mill RR, Renner SS. 2012. Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National Academy of Sciences, USA* **109**: 7793–7798.

Marshall CR. 2017. Five palaeobiological laws needed to understand the evolution of the living biota. *Nature Ecology and Evolution* **1**: 165.

Martínez L, Artabe AE, Bodnar J. 2012. A new cycad stem from the Cretaceous in Argentina and its phylogenetic relationships with other Cycadales. *Botanical Journal of the Linnean Society* **170**: 436–458.

Martínez-Domínguez L, Nicolalde-Morejón F, Vergara-Silva F, Stevenson DW. 2022. *Ceratozamia oliversacksii* (Zamiaceae), a new species of gymnosperm from western Oaxaca, Mexico. *Kew Bulletin* **77**: 211–219.

Matzke NJ. 2018. *BioGeoBEARS: biogeography with Bayesian (and Likelihood) evolutionary analysis with R scripts.* v.1.1.1. doi: 10.5281/zenodo.1478250

May MR, Contreras DL, Sundue MA, Nagalingum NS, Looy CV, Rothfels CJ. 2021. Inferring the total-evidence timescale of Marattialean fern evolution in the face of model sensitivity. *Systematic Biology* **70**: 1232–1255.

May MR, Höhna S, Moore BR. 2016. A Bayesian approach for detecting the impact of mass-extinction events on molecular phylogenies when rates of lineage diversification may vary. *Methods in Ecology and Evolution* **7**: 947–959.

Mazet N, Fabre P-H, Morlon H, Condamine FL. 2022. Estimating clade-specific diversification rates and palaeodiversity dynamics from reconstructed phylogenies. *bioRxiv*. doi: 10.1101/2022.05.10.490920.

McElwain JC. 1998. Do fossil plants signal the paleoatmospheric CO_2 concentration in the geological past? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **353**: 83–96.

Meseguer S, Condamine FL. 2020. Ancient tropical extinctions at high latitudes contributed to the latitudinal diversity gradient. *Evolution* **74**: 1966–1987.

Mickle JE, Barone Lumaga MRB, Moretti A, De Luca P. 2011. Scanning electron microscopy studies of cuticle micromorphology in *Cycas* L. (Cycadaceae). *Plant Biosystems* **145**: 191–201.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: *Proceedings of the gateway computing environments workshop*. IEEE, 1–8.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von 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.

Mitchell JS, Etienne RS, Rabosky DL. 2019. Inferring diversification rate variation from phylogenies with fossils. *Systematic Biology* **68**: 1–18.

Moretti A, Caputo P, Cozzolino S, de Luca P, Gaudio L, Gigliano GS, Stevenson DW. 1993. A phylogenetic analysis of *Dioon* (Zamiaceae). *American Journal of Botany* **80**: 204–214.

Morlon H. 2014. Phylogenetic approaches for studying diversification. *Ecology Letters* **17**: 508–525.

Morlon H, Parsons TL, Plotkin JB. 2011. Reconciling molecular phylogenies with the fossil record. *Proceedings of the National Academy of Sciences, USA* **108**: 16327–16332.

Müller RD, Qin X, Sandwell DT, Dutkiewicz A, Williams SE, Flament N, Maus S, Seton M. 2016. The GPlatesportal: cloud-based interactive 3D visualization of global geophysical and geological data in a web browser. *PLoS ONE* **11**: e0150883.

Mustoe G. 2007. Coevolution of cycads and dinosaurs. *Cycad Newsletter* **30**: 6–9.

Nackey LL, Midgley GF, de Bosenber JW, Donaldson JS. 2018. A cycad's non-saturating response to carbon dioxide enrichment indicates Cenozoic carbon limitation in pre-historic plants. *Austral Ecology* **43**: 447–455.

Nagalingum NS, Marshall CR, Quental TB, Rai HS, Little DP, Mathews S. 2011. Recent synchronous radiation of a living fossil. *Science* **334**: 796–799.

Nee S, May RM, Harvey PH. 1994. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **344**: 305–311.

Nee S, Mooers AO, Harvey PH. 1992. Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences, USA* **89**: 8322–8326.

Niklas KJ, Tiffney BH, Knoll AH. 1983. Patterns in vascular land plant diversification. *Nature* **303**: 614–616.

Norstog KJ, Nichols TJ. 1997. *The biology of the Cycads*. New York, NY, USA: Cornell University Press.

Oliveros CH, Andersen MJ, Hosner PA, Mauck WM III, Sheldon FH, Cracraft J, Moyle RG. 2020. Rapid Laurasian diversification of a pantropical bird family during the Oligocene–Miocene transition. *Ibis* **162**: 137–152.

Pérez-Farrera MA, Gutiérrez-Ortega JS, Gregory TJ, Chemnick J, Salas-Morales S, Calonje M, Díaz-Jiménez P. 2022. *Ceratozamia schiblii* (Zamiaceae): a new cycad species from the eastern mountains of Oaxaca, Mexico. *Taxonomy* **2**: 324–338.

Pérez-Farrera MA, Gutiérrez-Ortega JS, Haynes JL, Chemnick J, Salas-Morales SH, Calonje M, Vovides AP. 2021a. *Ceratozamia aurantiaca* (Zamiaceae): a new cycad species from the northern rainforests of Oaxaca, Mexico. *Taxonomy* **1**: 243–255.

Pérez-Farrera MA, Gutiérrez-Ortega JS, Vovides AP, Calonje M, Díaz-Jiménez P. 2021b. *Ceratozamia dominguezii* (Zamiaceae): A new cycad species from southeastern Mexico. *Taxonomy* **1**: 345–359.

Pott C, Kerp JHF, Krings M. 2007. *Pseudoctenis cornelii* nov. spec. (cycadalean foliage) from the Carnian (Upper Triassic) of Lunz, Lower Austria. *Annalen des Naturhistorischen Museums Wien* **109A**: 1–17.

Pross J, Contreras L, Bijl PK, Greenwood DR, Bohaty SM, Schouten S, Bendle JA, Röhl U, Tauxe L, Raine JI *et al.* 2012. Persistent near-tropical warmth on the Antarctic continent during the early Eocene epoch. *Nature* **488**: 73–77.

Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**: 901–904.

Ranwez V, Douzery EJ, Cambon C, Chantret N, Delsuc F. 2018. Macse v.2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. *Molecular Biology and Evolution* **35**: 2582–2584.

von Raumer JF, Stampfli GM, Bussy F. 2003. Gondwana-derived microcontinents—the constituents of the Variscan and Alpine collisional orogens. *Tectonophysics* **365**: 7–22.

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

Renner SS, Grimm GW, Kapli P, Denk T. 2016. Species relationships and divergence times in beeches: new insights from the inclusion of 53 young and old fossils in a birth–death clock model. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **371**: 20150135.

Ronquist F, Klopfstein S, Vilhelmsen L, Schulmeister S, Murray DL, Rasnitsyn AP. 2012. A total-evidence approach to dating with fossils, applied to the early radiation of the Hymenoptera. *Systematic Biology* **61**: 973–999.

Rothwell GW, Stockey RA, Stevenson DW, Zumajo-Cardona C. 2022. Large permineralized seeds in the Jurassic of Haida Gwaii, Western Canada: exploring the mode and tempo of cycad evolution. *International Journal of Plant Sciences* **183**: 674–690.

Roy K, Jablonski D, Valentine JW. 1996. Higher taxa in biodiversity studies: patterns from eastern Pacific marine molluscs. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **351**: 1605–1613.

Sahney S, Benton MJ, Falcon-Lang HJ. 2010. Rainforest collapse triggered Carboniferous tetrapod diversification in Euramerica. *Geology* **38**: 1079–1082.

Salas-Leiva DE, Meerow AW, Calonje M, Griffith MP, Francisco-Ortega J, Nakamura K, Stevenson DW, Lewis CE, Namoff S. 2013. Phylogeny of the cycads based on multiple single-copy nuclear genes: congruence of concatenated parsimony, likelihood and species tree inference methods. *Annals of Botany* **112**: 1263–1278.

Salzman S, Crook D, Crall JD, Hopkins R, Pierce NE. 2020. An ancient push-pull pollination mechanism in cycads. *Science Advances* **6**: eaay6169.

Scotese C, Schettino A. 2017. Late Permian-Early Jurassic paleogeography of western Tethys and the world. In: Soto JI, Flinch JF, Tari G, eds. *Permo-Triassic salt provinces of Europe, North Africa and the Atlantic margins*. Amsterdam, the Netherlands: Elsevier, 57–95.

Seton M, Müller RD, Zahirovic S, Gaina C, Torsvik T, Shephard G, Talsma A, Gurnis M, Turner M, Maus S *et al.* 2012. Global continental and ocean basin reconstructions since 200 Ma. *Earth-Science Reviews* **113**: 212–270.

Silvestro D, Cascales-Miñana B, Bacon CD, Antonelli A. 2015. Revisiting the origin and diversification of vascular plants through a comprehensive Bayesian analysis of the fossil record. *New Phytologist* **207**: 425–436.

Silvestro D, Schnitzler J, Liow LH, Antonelli A, Salamin N. 2014. Bayesian estimation of speciation and extinction from incomplete fossil occurrence data. *Systematic Biology* **63**: 349–367.

Silvestro D, Warnock RC, Gavryushkina A, Stadler T. 2018a. Closing the gap between palaeontological and neontological speciation and extinction rate estimates. *Nature Communications* **9**: 5237.

Silvestro D, Tejedor MF, Serrano-Serrano ML, Loiseau O, Rossier V, Rolland J, Zizka A, Höhna S, Antonelli A, Salamin N. 2018b. Early arrival and climatically-linked geographic expansion of New World monkeys from tiny African ancestors. *Systematic Biology* **68**: 78–92.

Slater G, Harmon L, Alfaro M. 2012. Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* **66**: 3931–3944.

Slater GJ, Goldbogen JA, Pyenson ND. 2017. Independent evolution of baleen whale gigantism linked to Plio-Pleistocene ocean dynamics. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **284**: 20170546.

Smoot E, Taylor T, Delevoryas T. 1985. Structurally preserved fossil plants from Antarctica. I. *Antarcticycas*, gen. nov., a Triassic cycad stem from the Beardmore Glacier area. *American Journal of Botany* **72**: 1410–1423.

Spencer AR, Garwood RJ, Rees AR, Raine RJ, Rothwell GW, Hollingworth NT, Hilton J. 2017. New insights into Mesozoic cycad evolution: an exploration of anatomically preserved Cycadaceae seeds from the Jurassic Oxford Clay biota. *PeerJ* **5**: e3723.

Su K, Quan C, Liu Y-S. 2014. *Cycas fushunensis* sp. nov. (Cycadaceae) from the Eocene of northeast China. *Review of Palaeobotany and Palynology* **204**: 43–49.

Suan G, Popescu S-M, Suc J-P, Schnyder J, Fauquette S, Baudin F, Yoon D, Piepjohn K, Sobolev NN, Labrousse L. 2017. Subtropical climate conditions and mangrove growth in Arctic Siberia during the early Eocene. *Geology* **45**: 539–542.

Treutlein J, Wink M. 2002. Molecular phylogeny of cycads inferred from rbcL sequences. *Naturwissenschaften* **89**: 221–225.

Upchurch P, Hunn CA, Norman DB. 2002. An analysis of dinosaurian biogeography: evidence for the existence of vicariance and dispersal patterns caused by geological events. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **269**: 613–621.

Uzunova K, Palamarev E, Kvaček Z. 2001. *Eostangeria ruzinciniana* (Zamiaceae) from the Middle Miocene of Bulgaria and its relationship to similar taxa of fossil *Eostangeria*, and extant *Chigua* and *Stangeria* (Cycadales). *Acta Palaeobotanica* **41**: 177–193.

Vajda V, Pucetaite M, McLoughlin S, Engdahl A, Heimdal J, Uvdal P. 2017. Molecular signatures of fossil leaves provide unexpected new evidence for extinct plant relationships. *Nature Ecology and Evolution* **1**: 1093–1099.

Vovides AP, Clugston JAR, Gutiérrez-Ortega JS, Pérez-Farrera MA, Sánchez-Tinoco MY, Galicia S. 2018. Epidermal morphology and leaflet anatomy of Dioon (Zamiaceae) with comments on climate and environment. *Flora: Morphology, Distribution, Functional Ecology of Plants* **239**: 20–44.

Westerhold T, Marwan N, Drury AJ, Liebrand D, Agnini C, Anagnostou E, Barnet JS, Bohaty SM, De Vleeschouwer D, Florindo F *et al.* 2020. An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science* **369**: 1383–1387.

Wilf P, Stevenson DW, Cúneo NR. 2016. The last Patagonian cycad, *Austrozamia stockeyi* gen. et sp. nov., early Eocene of Laguna del Hunco, Chubut, Argentina. *Botany* **94**: 817–829.

Wing SL, Hickey LJ, Swisher CL. 1993. Implications of an exceptional fossil flora for Late Cretaceous vegetation. *Nature* **363**: 342–344.

Yang Y, Ferguson DK, Liu B, Mao K-S, Gao L-M, Zhang S-Z, Wan T, Rushforth K, Zhang Z-X. 2022. Recent advances on phylogenomics of gymnosperms and an updated classification. *Plant Diversity* **44**: 340–350.

Zhang C, Stadler T, Klopfstein S, Heath TA, Ronquist F. 2016. Total-evidence dating under the fossilized birth-death process. *Systematic Biology* **65**: 228–249.

Zhang Q, Ree RH, Salamin N, Xing Y, Silvestro D. 2022. Fossil-informed models reveal a boreotropical origin and divergent evolutionary trajectories in the walnut family (Juglandaceae). *Systematic Biology* **71**: 242–258.

Zhang YJ, Cao KF, Sack L, Li N, Wei XM, Goldstein G. 2015. Extending the generality of leaf economic design principles in the cycads, an ancient lineage. *New Phytologist* **206**: 817–829.

Zizka A, Silvestro D, Andermann T, Azevedo J, Ritter CD, Edler D, Farooq H, Herdean A, Ariza M, Scharn R *et al.* 2019. COORDINATECLEANER: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* **10**: 744–751.

Supporting Information

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

Fig. S1 Maximum-likelihood phylogeny of Cycadales.

Fig. S2 Node support for phylogenetic analyses of Cycadales.

Fig. S3 Bayesian chronogram from the total-evidence dating analysis including extant and extinct taxa.

Fig. S4 Bayesian chronogram from the total-evidence dating analysis excluding the fossil taxa.

Fig. S5 Phylogenetic placements of fossil cycads illustrated using RoguePlots.

Fig. S6 Estimate of the historical biogeography for Cycadales with extant and extinct species.

Fig. S7 Estimate of the historical biogeography for Cycadales with extant species only.

Fig. S8 Estimate of the historical biogeography for Cycadales using BioGeoBEARS (DEC model) with extant and extinct species by coding fossil geographic ranges with missing data instead of true absences (option *useAmbiguities* = *TRUE*).

Fig. S9 Estimate of the historical biogeography for Cycadales using BioGeoBEARS (DEC model) with extant and extinct species by coding fossil geographic ranges as true absences (option *useAmbiguities* = *FALSE*).

Fig. S10 Estimate of the historical biogeography for Cycadales using BioGeoBEARS (DEC model) with extant and extinct species by taking into account the uncertainties in fossil placements and divergence times, and coding fossil geographic ranges with missing data (option *useAmbiguities* = *TRUE*).

Fig. S11 Number of local extinctions (extirpations) per time bin (Cenozoic vs Mesozoic and Paleozoic) compared between analyses excluding fossils and analyses including fossils.

Methods S1 Examination of fossil specimens.

Notes S1 Morphological characters used in the total-evidence dating analyses.

Table S1 Specimens re-examined during the coding of the matrix in this study.

Table S2 Fossil species used in this study, and references for ages and morphology.