Annals of Botany **XX**: 1–22, 2024
https://doi.org/10.1093/aob/mcae065, available online at www.academic.oup.com/aob

ANNALS OF BOTANY
Founded 1887

ORIGINAL ARTICLE

Transcriptome sequencing data provide a solid base to understand the phylogenetic relationships, biogeography and reticulated evolution of the genus *Zamia* L. (Cycadales: Zamiaceae)

Anders Lindstrom[1,*,#], Sadaf Habib[2,3,#], Shanshan Dong[3], Yiqing Gong[3], Jian Liu[4], Michael Calonje[5], Dennis Stevenson[6] and Shouzhou Zhang[3,*]

[1]Global Biodiversity Conservancy 144/124 Moo 3, Soi Bua Thong, Bangsalae, Sattahip, Chonburi, 20250, Thailand, [2]Jiangxi Provincial Key Laboratory of Ex Situ Plant Conservation and Utilization, Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang 332900, China, [3]Key Laboratory of Southern Subtropical Plant Diversity, Fairy Lake Botanical Garden Shenzhen & Chinese Academy of Sciences, Shenzhen, 518004, China, [4]Department of Economic Plants and Biotechnology, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, China, [5]Montgomery Botanical Center, Coral Gables, FL 33156, USA and [6]New York Botanical Garden, Bronx, NY 10458, USA
[*]For correspondence: shouzhouz@szbg.ac.cn, or ajlindstrom71@gmail.com
[#]Contributed equally to this work.

Received: 5 January 2024 Returned for revision: 26 March 2024 Editorial decision: 25 April 2024 Accepted: 16 May 2024

• **Background and Aims:** Cycads are a key lineage to understand the early evolution of seed plants and their response to past environmental changes. However, tracing the evolutionary trajectory of cycad species is challenging when the robust relationships at inter- or infrageneric level are not well resolved.
• **Methods:** Here, using 2901 single-copy nuclear genes, we explored the species relationships and gene flow within the second largest genus of cycads, i.e. *Zamia*, based on phylotranscriptomic analyses of 90 % extant *Zamia* species. Based on a well-resolved phylogenetic framework, we performed gene flow analyses, molecular dating and biogeographical reconstruction to examine the spatiotemporal evolution of *Zamia*. We also performed ancestral state reconstruction of a total of 62 traits of the genus to comprehensively investigate its morphological evolution.
• **Key Results:** *Zamia* comprises seven major clades corresponding to seven distinct distribution areas in the Americas, with at least three reticulation nodes revealed in this genus. Extant lineages of *Zamia* initially diversified around 18.4–32.6 (29.14) million years ago in Mega-Mexico, and then expanded eastward into the Caribbean and southward into Central and South America. Ancestral state reconstruction revealed homoplasy in most of the morphological characters.
• **Conclusions:** This study revealed congruent phylogenetic relationships from comparative methods/datasets, with some conflicts being the result of incomplete lineage sorting and ancient/recent hybridization events. The strong association between the clades and the biogeographic areas suggested that ancient dispersal events shaped the modern distribution pattern, and regional climatic factors may have resulted in the following *in situ* diversification. Climate cooling starting during the mid-Miocene is associated with the global expansion of *Zamia* to tropical South America that has dramatically driven lineage diversification in the New World flora, as well as the extinction of cycad species in the nowadays cooler regions of both hemispheres, as indicated by the fossil records.

Key words: Cycadales, *Zamia*, biogeography, transcriptomics, character evolution, gymnosperms.

INTRODUCTION

Cycads represent one of the earliest lineages of seed plants with an extensive fossil record. The divergence between Zamiaceae and Cycadaceae was estimated to have occurred ~330 million years ago (MA) in Carboniferous period (Coiro *et al.,* 2023). However, most of the extant genera of cycads diversified during the Neogene (Y. Liu *et al.*, 2022; Coiro *et al.,* 2023), implying a relatively recent radiation of extant cycads. Within all ten extant cycad genera, four are endemically distributed within the Neotropical region, i.e. *Zamia* (86 species), *Ceratozamia* (40 species), *Dioon* (18 species) and *Microcycas* (1 species)

(Calonje *et al.*, 2024). Among them, *Zamia* represents the most diverse genus and is broadly distributed from southern USA (Florida) through Central America and the Caribbean to South America as far south as Bolivia. It occurs primarily in subtropical and tropical regions and is particularly diverse in tropical Panama and Colombia (Pérez-Farrera *et al.*, 2023; Calonje *et al.*, 2024). Besides, *Zamia* also displays the most extensive morphological diversity and adaptation to varied habitats in the Neotropics (e.g. sea-sides, cloud forests, mangroves, grasslands, white sand forest, and semideserts) (Fig. 1), and from near sea level to >2000 m in elevation (Norstog and Nicholls, 1997; Calonje *et al.*, 2019). The ecological niche diversity of this genus is also exemplified by the only true epiphytic cycad species, *Z. pseudoparasitica*, and several specialist lithophytes that grow on steep cliff faces (e.g. *Z. cremnophila*, *Z. meermanii*, *Z. magnifica*, *Z. sandovalii*). Some species (e.g. *Z. herrerae*, *Z. boliviana*, *Z. standleyi*) are adapted to seasonally dry habitats and may become deciduous during periods of extreme drought (A. Lindstrom, pers. observ.).

Despite their wide distribution ranges, a total of 62 % of extant *Zamia* species are considered as threatened, and another 24 % are vulnerable or conservation-dependent low-risk taxa (Calonje *et al.*, 2024). In this context, a well-resolved phylogenetic tree is crucial to enhancing and organizing the conservation strategies. Nevertheless, due to their intensive ecological adaptations and complex morphological characters and the inaccessibility of some narrowly distributed rare species for research, the evolutionary relationships of *Zamia* are controversial and difficult to resolve (Calonje *et al.*, 2019). Morphology-based studies of the traditional classification of *Zamia* have emphasized characters such as leaflet shape and superficial strobilus characters (Schutzman, 1984, 1998; Newell, 1985; Schutzman *et al.*, 1988; Stevenson, 1993, 2001*a*, 2004; Taylor *et al.*, 2008; Nicolalde-Morejón *et al.*, 2009). However, morphological stasis of some characters (e.g. petiole prickles) coupled with ample variation and parallel evolution of others (e.g. plicate leaflets) within a conserved framework have made it problematic to identify any ingroup morphological synapomorphies (Caputo *et al.*, 2004; Calonje *et al.*, 2019) or even to utilize anatomical characters for clade diagnoses (Glos *et al.*, 2022). Therefore, whether extensive morphological homoplasy exists in *Zamia* remains to be tested based on a robust molecular phylogeny.

Earlier molecular phylogenetic work in *Zamia* laid a foundation to investigate the evolutionary relationships within the genus. Caputo *et al.* (2004) conducted parsimony-based phylogenetic analysis for 23 currently accepted species using nuclear ribosomal DNA (internal transcribed spacer 2, ITS2) and additional morphological data. This study indicated that phylogenetic relationships within *Zamia* species correlate more closely to its geographical distribution rather than to gross morphological similarities, and that convergent evolution of morphological characters is pervasive in the genus. Nagalingum *et al.* (2011) produced a time-calibrated phylogeny of cycads, and investigated 199 cycad accessions including 33 *Zamia* species, using nuclear phytochrome P (PHYP) and two plastid genes (*rbcL* and *matK*). The phylogeny exhibited low node support between the species within the genus *Zamia*, which might have occurred due to the presence of fewer informative sites in the selected

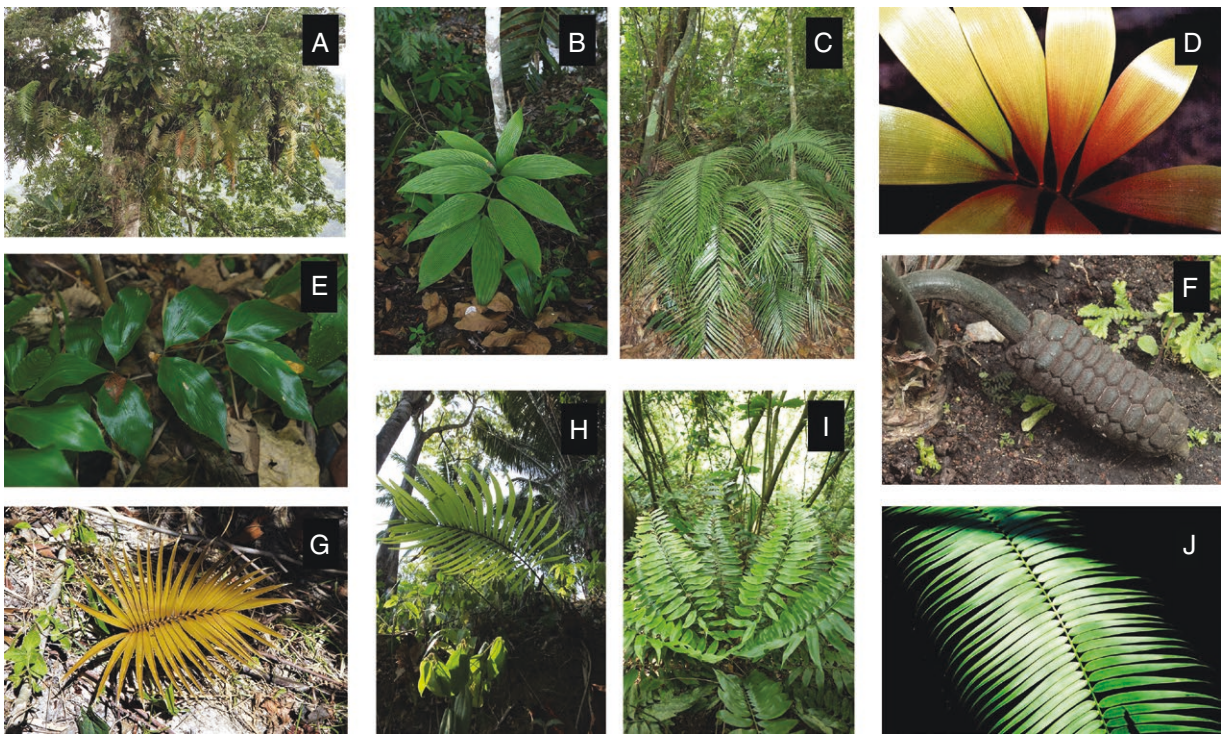

FIG. 1. Morphology and habitat variations in selected taxa of *Zamia*. (A) *Zamia pseudoparasitica*, epiphytic in habitat, Panama; (B) *Z. dressleri*, corrugated leaves in habitat, Panama; (C) *Z. soconuscensis,* in montane forest, Chiapas (Mexico); (D) *Z. pyrophylla*, emergent leaf, Colombia; (E) *Z. restrepoi*, raised leaf venation details, Colombia; (F) *Z. splendens*, decumbent female strobili, Mexico; (G) *Z. oreillyi*, emergent leaf, Honduras; (H) *Z. sandovallii*, decumbent leaf, Honduras; (I) *Z. onan-reyesii*, rainforest habitat, Honduras; (J) *Z. chigua*, linear parallel leaf details, Colombia. Photographs by A. Lindstrom.

markers. Using a concatenated dataset of nine single-copy nuclear genes (SCNGs) and one plastid gene with extensive taxon sampling (90 % of species), Calonje *et al.* (2019) recovered a phylogenetic tree of *Zamia* with a strong geographical delimitation of clades. The study resolved the division of the genus into six major clades: the Caribbean clade, the Fischeri clade, the Mesoamerican clade, the Mega-Mexico clade, the Isthmus clade and the Central Meridional clade. It was a major step forward towards understanding the phylogeny of *Zamia* in terms of sampling and the number of markers used. Nonetheless, their *Zamia* phylogeny was still not well resolved at shallower nodes, suggesting that incomplete lineage sorting and hybridization played roles in blurring the species relationships, and leaving unsettled specific questions like spatiotemporal diversification and character evolution of *Zamia*.

According to Calonje *et al.* (2019), the diversification of extant *Zamia* species started ~9.54 MA, during the middle Miocene, and species richness peaked during the Pliocene and Pleistocene periods. Moreover, their study, along with that of Caputo *et al.* (2004), suggested the Caribbean or Meso-America as a possible point of radiation for *Zamia.* Comparative age estimates provided in several previous studies were either older (Nagalingum *et al.*, 2011; Condamine *et al.*, 2015) or younger (Salas-Leiva *et al.*, 2013) than those reported in Calonje *et al.* (2019). More recently, the diversification of cycads re-examined by Coiro *et al.* (2023) with a combination of extant and fossil taxa (i.e. total-evidence method) revealed an earlier diversification of *Zamia* in the Oligocene (26.3 MA). This age is comparable with the estimated crown ages of two other Neotropical cycad genera, *Ceratozamia* (22.2 MA) and *Dioon* (32.2 MA) (Coiro *et al.*, 2023), suggesting a synchronized diversification of the cycads in the New World. Combining a robust phylogenetic framework with reliable time calibrations would allow the exploration of implications from a well-resolved phylogeny of the genus *Zamia*. This would contribute to a better understanding of the biogeographic distribution pattern of the genus, allow the testing of previously hypothesized relationships, and provide a basis for comparisons with other cycad genera in the Neotropics.

Organellar and nuclear genomic region assemblies of cycads highlighted the extremely slow evolution and mutation rates, as well as highly similar genomic profiles within each genus (Rai *et al.*, 2003; Chang *et al.*, 2020; Habib *et al.*, 2021, 2022; J. Liu *et al.*, 2022). Thus, extensive mining of a gene dataset for an in-depth exploration of molecular evolution remained necessary to better reveal the infrageneric relationships within the genus. Recent phylogenomic studies on cycads utilized comprehensive datasets of thousands of SCNGs retrieved from transcriptome data to understand the infrageneric relationships within different genera (Habib *et al.*, 2022, 2023; J. Liu *et al.*, 2024). These studies systematically compared and summarized concatenated and coalescent species tree inference methods as well as different datasets, and demonstrated the advantages of phylotranscriptomics in resolving species-level relationship in cycads.

In this study, we assembled the transcriptome data of 83 *Zamia* accessions representing 77 of the known species, to investigate the molecular phylogeny of the genus as well as its implications for biogeography and morphological evolution. The main objectives of this study are (1) to provide insight into phylogenetic relationships and reticulate evolution within *Zamia* based on an extensive SCNG dataset, and compare our result with previous phylogenetic inferences (Caputo *et al.*, 2004; Calonje *et al.*, 2019); (2) to investigate the spatiotemporal diversification of *Zamia* and link this process to other cycad genera in this region; and (3) to examine the evolutionary pattern of morphological characters in *Zamia* and test if homoplasy exists extensively.

MATERIALS AND METHODS

Plant sampling and gene data retrieval

Our sampling comprises 83 accessions of *Zamia* representing 77 of the currently recognized *Zamia* species, including two species currently considered synonyms (*Z. kickxii*, *Z. ottonis*). Three species were sampled with duplicate accessions: *Z. integrifolia*, *Z. paucijuga* and *Z. variegata*. Nine other currently accepted species that are not sampled in the study are *Z. brasiliensis*, *Z. gomeziana*, *Z. imbricata*, *Z. lindosensis*, *Z. macrochiera*, *Z. magnifica*, *Z. melanorrhachis*, *Z. multidentata* and *Z. stenophyllidia*. Additionally, three representative species of closely related genera, i.e. *Microcycas calocoma*, *Stangeria eriopus* and *Ceratozamia hildae*, were included as the outgroups (86 accessions in total).

Fresh leaf samples of all the investigated taxa included in our study were obtained from the Nong Nooch Tropical Botanical Garden (NNTBG), Thailand. These plants belong to original wild collections that are now cultivated at NNTBG. A list of investigated taxa is provided in Table 1 along with voucher specimens and origin details. The transcriptome data for 75 ingroup *Zamia* accessions and three outgroup taxa were sampled for the cycad genome project to investigate the early evolution of seed plants (Y. Liu *et al.*, 2022). Eight samples were newly added here.

Fresh samples were stored in liquid nitrogen and then preserved in an ultra-low temperature (−80 °C) freezer. RNA extraction of each sample was done using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). Each sample's RNA quality and quantity were examined using 1 % agarose gel electrophoresis and a Qubit fluorometer (Invitrogen). Approximately 1 μg high-quality RNA was used to generate paired-end sequencing libraries with insert fragments of 200–300 bp of the corresponding cDNA, according to the manufacturer's instructions (Illumina, CA, USA). The libraries were then sequenced on an Illumina HiSeq 2000 at NextonOmics Biosciences (Wuhan, China).

The raw NGS data were trimmed and filtered for adapters, low-quality reads, undersized inserts and duplicate reads using Trimmomatic v0.39 (Bolger *et al.*, 2014) with custom parameters (LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:36). The newly generated clean read data (transcriptomic samples) are available at the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under BioProject number PRJNA930880 or in CNGB Sequence Archive (CNSA) (https://db.cngb.org) of China National GeneBank DataBase (CNGBdb) with project accession CNP0005845 (Table 1). The read statistics for 86 investigated taxa are given in Supplementary Data Table S1. The cleaned transcriptome reads were processed in the following four

TABLE 1. *Summary of datasets of 83* Zamia *and three outgroup accessions included for phylogenetic, spatiotemporal diversification and ASR analyses in this study.*

No.	Taxon	Origin	Herbarium code	NCBI/CNGBdb accessions	Raw/clean reads (Gb)
1	*Z. acuminata*	Costa Rica (San José)	SING0291185	SRR23444884	3.5/3.4
2	*Z. amazonum*	Brazil (Amazonas)	SING3048468	SRR23444883	3.5/3.3
3	*Z. amplifolia*	Colombia (Valle del Cauca)	W 2012-0010122	SRR23444872	3.3/3.1
4	*Z. angustifolia*	Bahamas (Eleuthera)	W 2012-0010215	SRR23444861	3.2/3.1
5	*Z. boliviana*	Bolivia (Cochabamba)	SING0383080	SRR23444850	3.5/3.5
6	*Z. chigua*	Colombia (Chocó)	SING0291175	SRR23444839	3.4/3.3
7	*Z. cremnophila*	Mexico (Tabasco)	SING0383075	SRR23444828	3.4/3.2
8	*Z. cunaria*	Panama (Kuna Yala)	W 2012-0010171	SRR23444817	3.3/3.2
9	*Z. decumbens*	Belize (Cayo)	W 2012-0010197	SRR23444808	3.6/3.5
10	*Z. disodon*	Colombia (Antioquia)	W 2012-0010009	SRR23444805	3.3/3.2
11	*Z. dressleri*	Panama (Colón)	W 2012-0010282	SRR23444882	3.7/3.6
12	*Z. elegantissima*	Panama (Colón)	W 2012-0009999	SRR23444881	3.4/3.2
13	*Z. encephalartoides*	Colombia (Santander)	W 2012-0010295	SRR23444880	3.2/3.1
14	*Z. erosa*	Puerto Rico	W 2012-0010236	SRR23444879	3.3/3.2
15	*Z. fairchildiana*	Costa Rica (Puntarenas)	W 2012-0010286	SRR23444878	3.7/3.5
16	*Z. fischeri*	Mexico (Tamaulipas)	W 2012-0009992	CNS1121654	3.23/2.78
17	*Z. furfuracea*	Mexico (Veracruz)	W 2012-0010152	SRR23044174	12/7.79
18	*Z. gentryi*	Ecuador (Carchi)	W 2012-0010289	SRR23444877	3.6/3.4
19	*Z. grijalvensis*	Mexico (Chiapas)	BK 084314	SRR23444876	4.1/4.0
20	*Z. hamannii*	Panama (Bocas del Toro)	SING0291183	SRR23444875	3.6/3.5
21	*Z. herrerae*	El Salvador (Sonsonate)	W 2012-0009996	SRR23444874	3.1/3.0
22	*Z. huilensis*	Colombia (Huila)	BK 084515	SRR23444873	3.5/3.4
23	*Z. hymenophyllidia*	Peru (Loreto)	W 2012-0010078	SRR23444871	3.4/3.3
24	*Z. imperialis*	Panama (Veraguas)	W 2012-0010187	SRR23444870	3.5/3.3
25	*Z. incognita*	Colombia (Santander)	SING0291184	SRR23444869	3.4/3.3
26	*Z. inermis*	Mexico (Veracruz)	W 0044918	SRR23444868	3.5/3.4
27	*Z. integrifolia* (B)	Bahamas (Andros)	W 2012-0010233	SRR23444867	3.6/3.5
28	*Z. integrifolia* (F)	U.S.A. (Florida)	SING383086	SRR23444806	3.4/3.3
29	*Z. ipetiensis*	Panama (Cuna Yala)	W 2012-0010173	SRR23444866	3.4/3.3
30	*Z. katzeriana*	Mexico (Chiapas)	W 2012-0010205	SRR23444865	3.9/3.8
31	*Z. kickxii*[*]	Cuba (Pinar del Rio)	W 2012-0010227	SRR23444815	3.4/3.3
32	*Z. lacandona*	Mexico (Chiapas)	W 2012-0010121	SRR23444864	3.8/3.7
33	*Z. lecointei*	Venezuela (Amazonas)	W 2012-0010207	SRR23444863	3.6/3.5
34	*Z. lindenii*	Ecuador (Los Ríos)	W 2012-0010125	SRR23444862	3.5/3.4
35	*Z. lindleyi*	Panama (Chiriquí)	BK 084316	SRR23444860	3.5/3.5
36	*Z. loddigesii*	Mexico (Veracruz)	W 2012-0010135	SRR23444810	3.6/3.5
37	*Z. lucayana*	Bahamas (Long Island)	W 2012-0010212	SRR23444859	3.6/3.5
38	*Z. manicata*	Colombia (Antioquia)	W 2012-0010175	SRR23444858	3.3/3.2
39	*Z. meermanii*	Belize (Belize)	BK 084317	SRR23444857	3.5/3.4
40	*Z. montana*	Colombia (Antioquia)	BK 084318	SRR23444855	3.4/3.3
41	*Z. monticola*	Guatemala (Alta Verapaz)	SING0383083	SRR23444854	3.4/3.3
42	*Z.* aff. *monticola*	Guatemala	SING0383084	SRR23444812	3.3/3.2
43	*Z. muricata*	Venezuela (Carabobo)	W 2012-0010063	SRR23444853	3.3/3.3
44	*Z. nana*	Panama (Coclé)	SING0383074	SRR23444852	3.2/3.1
45	*Z. nesophila*	Panama (Bocas del Toro)	W 2012-0010185	SRR23444851	2.9/2.9

TABLE 1. *Continued*

No.	Taxon	Origin	Herbarium code	NCBI/CNGBdb accessions	Raw/clean reads (Gb)
46	*Z. neurophyllidia*	Costa Rica	W 2012-0010274	SRR23444849	3.4/3.4
47	*Z. obliqua*	Panama (Darién)	W 2012-0010279	SRR23444848	3.4/3.3
48	*Z. oligodonta*	Colombia (Risaralda)	BK 084319	SRR23444847	3.3/3.2
49	*Z. onan-reyesii*	Honduras (Cortés)	SING0291179	SRR23444846	3.4/3.3
50	*Z. oreillyi*	Honduras (Atlántida)	SING0383081	SRR23444845	3.3/3.2
51	*Z. orinoquiensis*	Colombia (Meta)	SING0291178	SRR23444813	3.4/3.3
52	*Z. ottonis*[*]	Cuba (Sancti spiritus)	W 2012-0010221	SRR23444814	3.3/3.3
53	*Z. paucifoliolata*	Colombia (Valle del Cauca)	SING0383088	SRR23444844	3.5/3.4
54	*Z. paucijuga* (O)	Mexico (Oaxaca)	SING0383082	SRR23444843	3.5/3.3
55	*Z. paucijuga* (N)	Mexico (Nayarit)	SING0383085	CNS1121653	3.1/2.66
56	*Z. poeppigiana*	Puerto Rico	W 2012-0010222	SRR23444842	3.4/3.3
57	*Z. portoricensis*	Peru (Huánuco)	W 2012-0010155	SRR23444841	3.4/3.3
58	*Z. prasina*	Belize (Cayo)	W 2012-0009993	SRR23444840	3.0/2.9
59	*Z. pseudomonticola*	Costa Rica (Puntarenas)	SING0291180	SRR23444809	3.5/3.4
60	*Z. pseudoparasitica*	Panama (Cocle)	SING0291176	SRR23444807	3.6/3.5
61	*Z. pumila*	Dominican Republic	W 2012-0010140	SRR23444838	3.4/3.3
62	*Z. purpurea*	Mexico (Oaxaca)	W 2012-0009994	SRR23444837	3.4/3.3
63	*Z. pygmaea*	Cuba (W Cuba)	W 2012-0010230	SRR23444836	3.4/3.3
64	*Z. pyrophylla*	Colombia (Chocó)	SING0383078	SRR23444835	3.6/3.6
65	*Z. restrepoi*	Colombia (Córdoba)	W 2012-0010065	SRR23444834	3.5/3.5
66	*Z. roezlii*	Colombia (Nariño)	W 2012-0010117	SRR23444833	3.3/3.2
67	*Z. sandovalii*	Honduras (Atlántida)	SING0291181	SRR23444832	3.4/3.3
68	*Z. sinuensis*	Colombia (Antioquia)	W 2012-0010060	SRR23444856	3.5/3.3
69	*Z. skinneri*	Panama (Bocas del Toro)	W 2012-0010277	SRR23444831	3.6/3.5
70	*Z. soconuscensis*	Mexico (Chiapas)	W 2012-0010128	SRR23444830	3.2/3.1
71	*Z. spartea*	Mexico (Oaxaca)	W 2012-0010181	SRR23444829	3.3/3.3
72	*Z. splendens*	Mexico (Chiapas)	W 2012-0010114	SRR23444827	3.3/3.3
73	*Z. standleyi*	Honduras (Atlántida)	W 2012-0010112	SRR23444826	3.4/3.3
74	*Z. stevensonii*	Panama (Panamá)	SING0291186	SRR23444825	4.2/4.0
75	*Z. stricta*	Cuba (Santiago de Cuba)	SING0291177	SRR23444824	3.5/3.5
76	*Z. tolimensis*	Colombia (Tolima)	BK 084320	SRR23444823	2.9/2.8
77	*Z. tuerckheimii*	Guatemala (Alta Verapaz)	SING0291187	SRR23444822	3.5/3.4
78	*Z. ulei*	Brazil (Acre)	SING0383087	SRR23444821	3.6/3.5
79	*Z. urep*	Peru (Huánuco)	W 2012-0010012	SRR23444820	3.3/3.3
80	*Z. variegata* (B)	Belize (Toledo)	W 2012-0010284	SRR23444819	3.8/3.6
81	*Z. variegata* (G)	Guatemala (Alta Verapaz)	W 2012-0010195	SRR23444811	3.5/3.4
82	*Z. vazquezii*	Mexico (Veracruz)	SING0291182	SRR23444818	3.0/3.0
83	*Z. wallisii*	Colombia (Antioquia)	SING0383079	SRR23444816	3.9/3.8
84	*M. calocoma*	Cuba (Pinar del Rio)	W 2012-0010150	SRR23044176	12/6.6
85	*S. eriopus*	South Africa	W 2012-0010304	SRR23044175	12/6.85
86	*C. hildae*	Mexico	W 2012-0010028	SRR23044169	12/6.49

[*]Considered as synonym by The World List of Cycads, Calonje, Stev. and Osbourne (10 May, 2023).

Herbarium codes: BK, Bangkok Herbarium, Bangkok, Thailand; SING, Singapore Botanical Garden, Singapore; W, Naturhistorisches Museum Wien, Austria.

steps: (1) *de novo* assembly of clean raw reads using the Trinity v2.8.4 pipeline (Grabherr *et al.*, 2011); (2) annotation of longest assembled transcripts with TransDecoder v5.7.0 (https://github.com/TransDecoder); (3) orthologue detection of annotated transcripts using OrthoFinder v2.4.0 (Emms and Kelly, 2019); and (4) selection of single-copy genes for phylogenetic reconstruction using KinFin software v1.0.3 (Laetsch and Blaxter, 2017). All of the software packages mentioned above were used with default settings. Paralogues were filtered and gene families with only single-copy genes having species coverage >70 % were retained using the 'rmdup' command in SeqKit v2.5.0 (Shen *et al.*, 2016). The alignment of each single-copy gene dataset was performed using a local version of program TranslatorX v14.0(2) (Abascal *et al.*, 2010). This program first translated the nucleotide sequence into an amino acid sequence and then produced an amino acid alignment with MAFFT v5.0 (Katoh and Standley, 2013). Ambiguous portions in the alignment were trimmed by Gblocks v0.91b (Talavera and Castresana, 2007) with the default settings. In the final step, these cleaned amino acid alignments were used as a guide to generate the alignment for nucleotide sequences for subsequent phylogenetic analyses.

Phylogenetic tree reconstruction

A total of 2901 coding sequences (CDSs) of SCNGs were used for subsequent analyses, with two types of nucleotide dataset, i.e. one with all codon positions (CDS-nt) and another with only the first and second codon positions (CDS-nt12). We carried out multiple analyses to find and compare the phylogenetic inferences.

For coalescence analyses, a coalescence approach using unrooted gene trees was performed for the single-copy nuclear genes. Single-gene trees were inferred using IQ-TREE v2.2.0 (Minh *et al.*, 2020) by implementing the best-fitting model in ModelFinder (Kalyaanamoorthy *et al.*, 2017), 100 randomized tree searches, a rapid bootstrap analysis and a search for the best-scoring maximum likelihood estimation (MLE) tree. The resulting unrooted gene trees were then used as input to construct a species tree using the weighted ASTRAL (wASTRAL) optimization algorithm (Zhang and Mirarab, 2022) with the hybrid weighting (wASTRAL-h) option. Support values of the gene trees were also calculated using quartet scores (q1, q2, q3) in ASTRAL-III (Zhang *et al.*, 2018). For the final coalescent tree, the gene and site concordance factors (gCF and sCF, respectively) were calculated using IQ-TREE 2 with '–gcf' and '–scf' options (Minh *et al.*, 2020), to quantify genealogical concordance, which provides additional insight into agreement and disagreement among gene and species trees.

For concatenated analyses, to construct the phylogenetic tree based on a supermatrix approach, individual coding sequences were concatenated (CDS-nt_concat and CDS-nt12_concat) using SeqKit v2.5.0 (Shen *et al.*, 2016). We used partitioning scheme selection based on codon positions (the first, second and third codon positions), and the best-fitting model for each partition was selected in ModelFinder (Kalyaanamoorthy *et al.*, 2017). Then, the ML tree was reconstructed using IQ-TREE v2.2.0 (Minh *et al.*, 2020), with ultra-fast bootstrap analysis of 1000 replicates using the codon-based partitioning schemes and nucleotide substitution models inferred (i.e. TIM3 + F + R10 for codon position 1, TVM + F + R10 for codon position 2, and GTR + F + R10 for codon position 3), as specified by ModelFinder. Output trees were visualized and further processed using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Phylogenetic network analyses were performed to investigate the reticulate evolutionary history of *Zamia*. To reduce the computational burden, we first discarded redundant or unavailable samples from all sampled *Zamia* accessions, including (1) those with replicated accessions, (2) those showing no signals of introgression based on coalescent analysis (the nodal supports leading to these species pairs are decisive and show no evidence of uncertainty), or (3) those having low gene occupancies (<70 %) that might affect the hybridization analyses afterwards. Within the retained accessions, we further screened 40 representative species based on three criteria: (1) representative species must cover all major clades in the phylogeny; (2) species show overlapped geographical distribution or have debated taxonomical history; and (3) species may have undergone natural hybridization based on previous reports or our own field observations. We used PhyloNet v2.4 (Than *et al.*, 2008) to infer the possible hybrid events using the 'InferNetworks_ML' command.

Divergence time, net diversification analysis and ancestral area reconstruction

To explore the divergence times and diversification patterns for *Zamia*, we screened 50 loci from the complete dataset of 2901 genes, using the SortaDate package (Smith *et al.*, 2018) due to computational time constraints. The software involves filtering of loci for clock-likeness, and selects the genes with least topological conflict and enough informative sites. These 50 genes covered >95 % of the ingroup species with concordant evolutionary histories. A phylogenetic tree was constructed for the 50-gene concatenated dataset in RAxML v8 to generate 1000 trees, with tree topologies constrained to an output tree from multi-species coalescent analyses (CDS-nt12). The resulting trees were then used as the input data for TreePL to calculate the divergence times using the penalized likelihood method (Sanderson, 2002; Smith and O'Meara, 2012). Secondary calibrations for the root (minimum 65.5, maximum 119.3) and crown (minimum 18.4, maximum 37.2) of the genus were derived from the most recent total-evidence molecular dating analysis of the order Cycadales (Coiro *et al.*, 2023). The maximum clade credibility (MCC) trees were then constructed using TreeAnnotator v2.6.2 (Helfrich *et al.*, 2018), with median ages and 95 % highest posterior density (HPD) intervals on nodes, with the initial 25 % (250 out of 1000) trees discarded as burn-in. The chronogram was visualized in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Diversification rates of *Zamia* were estimated with Bayesian Analysis of Macroevolutionary Mixtures (BAMM) v2.5.0 (Rabosky, 2014), which uses a reversible-jump Markov chain Monte Carlo (MCMC) method to sample numerous possible diversification regimes based on the dated phylogeny. The time-calibrated tree without outgroups was obtained from the TreePL analysis. The prior assumption (i.e. one shift event) and

Table 2. *List of 10/62 morphological characters selected in ASR analysis for their phylogenetic significance.*

No.	Morphological	States			
	character	**0**	**1**	**2**	**3**
1	Stem habit	Hypogeous	Epigeous	Arborescent	
2	Petiole prickles	Present	Absent		
3	Leaflets number	Few 2–5	Medium 6– 20	Plenty 21–50	Numerous 51–105
4	Pinnae plication	Present	Absent		
5	Pinnae venation	Prominent	Inconspicuous	Furrows	
6	Eophyll	Bifoliolate	Tetrafoliolate	Pinnate	
7	Male peduncle length	Shorter than strobili	Equal to strobili	Longer than strobili	
8	Female peduncle length	Shorter than strobili	Equal to strobili	Longer than strobili	
9	Sarcotesta texture (ripe)	Coriaceous (leathery)	Membranous (thin)	Mucilaginous (gelatinous)	Sticky
10	Sclerotesta length	Small (8–11 mm)	Medium (12–20 mm)	Large (21–30 mm)	Extreme (31–42 mm)

options were selected using the 'setBAMMpriors' function of the R package BAMMtools v2.1.10 (Rabosky *et al.*, 2014). The MCMC was run for 2 000 000 generations and sampled every 1000 generations. Post-run analysis and visualization were also implemented using the R package BAMMtools. The initial 20 % of samples of the MCMC run were discarded as burn-in, and the remaining data were assessed for convergence using R package coda v0.19 (Plummer *et al.*, 2006) to ensure that the ESS values were >200. We conducted three separate BAMM runs to ensure convergence on comparable distributions of branch-specific marginal rate shift probabilities. To compare the diversification dynamics of *Zamia* with other cycad genera, we also extracted the fossil-calibrated phylogenies of *Ceratozamia*, *Dioon*, *Encephalartos* and *Macrozamia* from a recent work (Coiro *et al.*, 2023). We then performed BAMM analyses to infer speciation rate through time plots of each genus based on the methods described above. The speciation rate of *Zamia* was not quantified from the above study, as their published topology deviated from our high-resolution phylogenetic tree.

To obtain the distribution ranges of *Zamia* for biogeographical analyses, we first carefully extracted available occurrence data from extensive herbarium records and the online Global Biodiversity Information Facility (GBIF, http://www.gbif.org/), which were further confirmed by the World List of Cycads database (Calonje *et al.*, 2024; http://www.cycadlist.org) and with our own observations during field visits. Ancestral area reconstruction was conducted using the BioGeoBEARS package (Matzke, 2018) implemented in RASP v4.2 (Yu *et al.*, 2020). Seven biogeographical regions were defined based on the investigated distribution of *Zamia*: (A) Mega-Mexico (including Mexico, Belize, Guatemala, El Salvador and Honduras); (B) Caribbean; (C) Florida; (D) Isthmus (Costa Rica and Panama); (E) West Andes; (F) East Andes; and (G) Northern Colombia. The input tree for BioGeoBEARS analysis was obtained from TreePL analysis. We compared all alternative models of possible transition (simple analysis and with the +J parameter) and determined the best-fit biogeographical model according to the Akaike information criterion cumulative weight (AICc_wt).

Morphological character evolution

Data for morphological characters and alternating character states were produced from *in situ* field data associated with the documented collections for each sample used, from extensive data mining of herbarium specimens held at numerous worldwide herbaria, and from published descriptions (Stevenson, 1990*a*, 1991, 1993, 2001*a*, *b*, 2004; Lindstrom, 2009; Nicolalde-Morejón *et al.*, 2009; Lindstrom *et al.*, 2013; Zonneveld and Lindstrom, 2016; Calonje *et al.*, 2019; Segalla *et al.*, 2019). Additional morphological character states were recorded from the living cycad collection at NNTBG. Character states were obtained either by being labelled statistically or by being absent or present. Characters with morphometric ranges were also recorded in order to cover a larger range of characters. The relationships between morphological and ecological characters are commonly predicted to reflect the association between form and function, with this hypothesis being well supported in restricted taxonomic and geographical contexts (Kennedy *et al.*, 2020). Thus, we divided the morphological characters into two groups: (1) characters related to plant ecology and habitat adaptation based on all non-reproductive characters; and (2) characters related to reproductive biology, which included reproductive characters related to cone and seed characters. The detailed character matrix was completed initially for 62 morphological characters, with several characters coded here for the first time (Supplementary Data Table S2). However, due to the complex morphological states to be traced on the phylogenetic tree, we only selected ten morphological characters (four reproductive and six vegetative characters) for the final data matrix (Table 2). The character states were optimized onto the tree generated from the ML analysis of the coalescent dataset (CDS-nt12) in Mesquite v3.81 (Maddison and Maddison, 2023), using the ML criterion with the Markov *k*-state one-parameter (Mk1) model (Lewis, 2001). Species belonging to two outgroup genera, *Ceratozamia hildae* and *Stangeria eriopus*, were excluded from our character state analyses as they differ significantly in several morphological states not comparable with *Zamia*, and only *Microcycas*, the closest sister to *Zamia*, was included in such analyses.

RESULTS

Data statistics

The filtered, clean transcriptome data generated for each of the 83 *Zamia* accessions range from 2.84 Gb (*Z. nesophila*) to 4.03 Gb (*Z. stevensonii*). The average alignment length of 2901 CDSs used in phylogenetic analyses of *Zamia* was 3643 bp for CDS-nt and 2394 bp for CDS-nt12. The concatenated alignment length was 3 218 922 bp, with 480 193 variable and 198 203 parsimony-informative sites for CDS-nt and 195 044 477 bp with 270 178 variable and 80 160 parsimony-informative sites for the CDS-nt12 dataset. The average gap/ambiguous sites among the ingroup species was 14 %, with the lowest value reported in *Z. pseudoparasitica* (3.8 %) and the maximum in *Z. disodon* (48 %).

Phylogenetic analyses

We retrieved highly resolved phylogenetic relationships (>90 %) for CDS-nt and CDS-nt12 datasets, with a single major conflict in the case of the concatenated analyses (i.e. the phylogenetic position of clade II; Supplementary Data Fig. S1) or fewer shallow conflicts in cases of coalescent analyses (Supplementary Data Fig. S2), such as paraphyly of clade VI species and phylogenetic position of *Z. encephalartoides*. The genus is divided into seven major clades, and the placement of these major clades is mostly consistent between the two methods, except for clades II (the Fischeri clade), III and VI (Fig. 2). Additionally, some discrepancies occurred between the trees generated from coalescent and concatenated methods (Fig. 2; see details below).

Clade II is placed as sister to clade I in both coalescent analyses based on the CDS-nt and CDS-nt12 datasets (100 % support), and in concatenated analysis based on the CDS-nt12 dataset (99 % support; Fig. 2; Supplementary Data Figs S1 and S2). In contrast, clade II is sister to the monophyletic group of clades III–VII in concatenated analyses based on the CDS-nt dataset (BS = 92; Supplementary Data Fig. S1).

Clade III is divided into two subclades, III-A and III-B. Species trees supported the monophyly of subclade III-B with eight species within it (CDS-nt PP, 0.89; CDS-nt12 PP, 0.99). In contrast, concatenation trees did not support this relationship, and three species of subclade III-B (i.e. *Z. oreillyi*, *Z. standleyi*, *Z. sandovalii*) along with *Z. purpurea* of subclade III-A formed a small clade sister to subclade III-A with high support (Fig. 2).

Clade VI appeared to be paraphyletic in concatenated and coalescent CDS-nt analyses. However, the CDS-nt12 tree supported this group as monophyletic with maximum support (Fig. 2; Supplementary Data Fig. 2). Other conflicts occurred between the species at a shallow level within the major clades. Higher quartet support (>40 %) values were observed for the main topology (q1) compared with their first and second alternatives (q2 and q3) of clades I, II and V, whereas quartet support for the remaining clades was <40 % (Supplementary Data Fig. S3). The sister relationships within major clades I–III were less differentiated with equal quartet support for q1 and q2 or q3 (33–39 %; Supplementary Data Fig. S3). On the other hand, the sister relationships within clades IV–VI have significantly higher q1 values compared with q2 or q3 (Supplementary Data Fig. S3). In our phylogenomic tree, most of the nodes received 100 % bootstrap support, but only moderate quartet support. Thus, we calculated gene and site concordance factors (gCF and sCF) as complementary information. Overall, the sCF values were higher (better resolved) than gCF values for all the nodes (clades) (Supplementary Data Fig. S3). The sCF values for the major clades range from 38 % (clade III) to 97 % (clade I). The gCF showed low values for most of the major clades as well as deeper nodes.

Detection of gene flow

Phylogenetic relationships within *Zamia* were further investigated for potential evidence of hybridization (Supplementary Data Fig. S4). Putative gene flow events were detected among the clades using 40 *Zamia* species as representatives. The species tree network of $n = 3$ had a higher probability in likelihood than any other reticulations. The infrageneric relationships of the *Zamia* network tree were largely consistent with our presented phylogeny and were found to have three reticulation nodes with probability of 0.5 from each of their ancestors. The first reticulation node (node 1) revealed that ancestors of clade II (Fischeri clade) were inherited from the ancestor of clade I and clade III (without *Z. meermanii*, *Z. standleyi* and *Z. tuerckheimii*). The second node suggested an event where species from clades VI and VII likely shared an ancestor from a most recent common ancestor (MRCA) of clades I–IV and *Z. soconuscensis* of clade IV. Reticulation on node 3 denoted that the ancestors of clade V were inherited from *Z. soconuscensis* of clade IV and the MRCA of clades VI + VII.

Spatiotemporal distribution and relationships among major clades

DEC + j was found to be the best-fit biogeographical model for *Zamia* in BioGeoBEARS analyses. Among the investigated areas, Mega-Mexico (area A) was found to be the ancestral area for the whole genus (Fig. 3) with maximum probability of 0.75, followed by Mega-Mexico + Caribbean (area AB) with the probability of 0.21. Our analyses of spatiotemporal distribution (Fig. 3; Table 3) indicated that after the split of the genus *Zamia* from *Microcycas* around 99.23 MA (HPD 95 %, 85.38–103.84; node 1) during the Cretaceous epoch, the genetic divergence of clades I and II from rest of the genus occurred at 29.14 MA (HPD 95 %, 18.4–32.63; node 2) in the Oligocene. Species of clade I (Caribbean species with inclusion of *Zamia integrifolia* from Florida) split from clade II (Fischeri clade from Veracruz, Mega-Mexico) around 26.32 MA (HPD 95 %, 15.87–30.2; node 3). Speciation within the Fischeri clade occurred in the early Miocene at 18.95 MA (HPD 95 %, 10.63–22.57; node 5), followed by speciation occurring within clade I around 8.17 MA (HPD 95 %, 4.51–10.33; node 4) during the late Miocene.

The second event of independent divergence of species within Mega-Mexico occurred around 24.83 MA (HPD 95 %, 14.51–28.58; node 6). The clade diverged into two independent lineages (clade III and clade IV) within approximately the same time frame. Species of clade III diversified in Mega-Mexico at 22.9 MA (HPD 95 %, 12.74–26.83; node 7) during the Oligocene. This clade also showed a distinct diversification

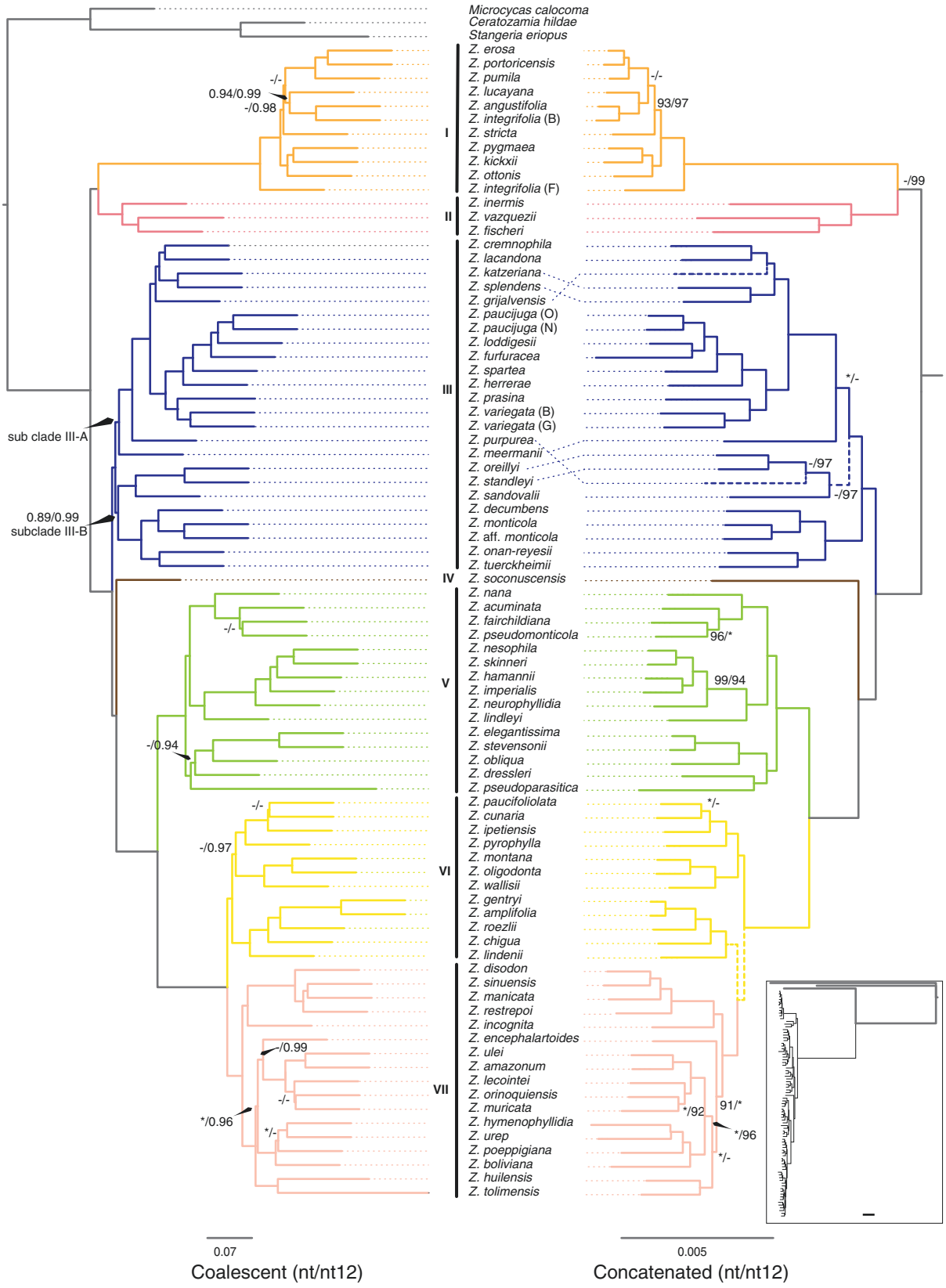

FIG. 2. Phylogenetic tree of *Zamia* using coalescent (left) and concatenated (right) analyses, based on nucleotide dataset of 2901 SCNGs using all codon positions (nt) and first and second codon positions only (nt12). Support values are not mentioned if branches are maximally supported in both datasets. Otherwise, * and – indicate maximum support and branch support <80 %, respectively. Dotted branches highlight incongruence among two analyses. Major phylogenetic clades in the genus are indicated with black vertical bars with species names.

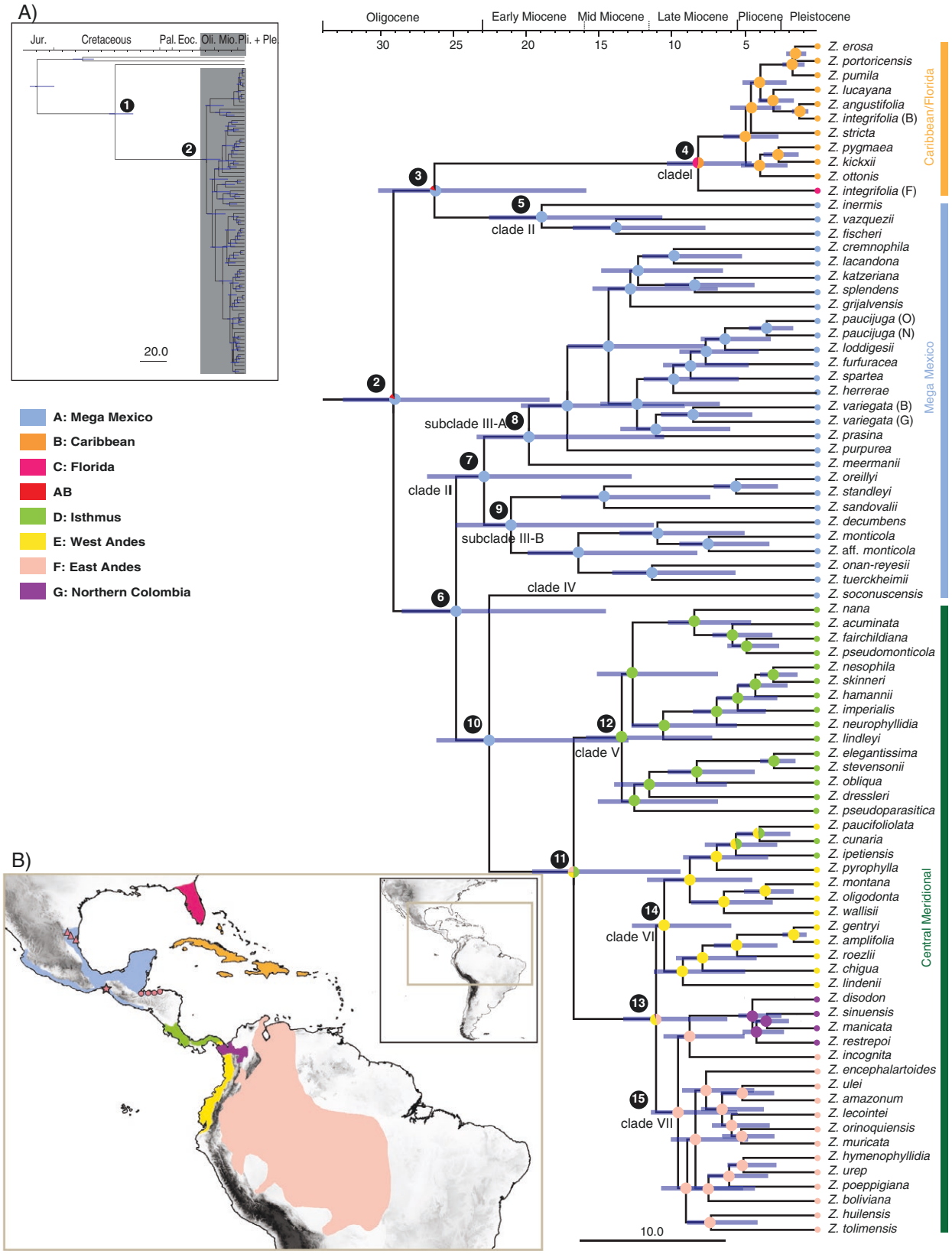

FIG. 3. Spatiotemporal evolution of *Zamia*. (A) Chronograms and ancestral area reconstruction of *Zamia* derived from the 50-gene ML tree with topologies constrained to the coalescent tree (nt12), using three secondary calibration points derived from Coiro *et al.* (2023). Numbers 1–15 represent major divergence events in *Zamia*. Nodes 1 and 2 are mentioned in the inset, and the grey-shaded area of the inset is enlarged to indicate the position of nodes 3–15. Blue bars represent the 95 % HPD. The mean divergence times and 95 % HPSs are provided for each node of interest in Table 3. The pie charts indicate the ancestral area of each node (>5 % only). (B) The map of the Americas in the inset is enlarged to show the geographical distribution of major clades among the seven distinct areas of the genus distribution in Mega-Mexico and adjacent regions.

Table 3. *Mean divergence times, 95 % HPDs and ancestral area probability for each node of interest in accordance with Fig. 3.*

Node	Diversification event	Time (MA): 95 % HPD (mean)	Ancestral area probability
1	*Microcycas* > *Zamia*	85.38–103.84 (99.23)	A: 0.50, A: 0.50
2	Mega-Mexico > diversification	18.4–32.63 (29.14)	A: 0.89
3	Mega-Mexico (Veracruz + Caribbean)	15.87–30.2 (26.32)	A: 0.98
4	Caribbean > inter-island dispersion	4.51–10.33 (8.17)	B: 0.50, C: 0.50
5	Veracruz > diversification	10.63–22.57 (18.95)	A: 1.00
6	Mega-Mexico + Isthmus/South America	14.51–28.58 (24.83)	A: 0.99
7	Mega-Mexico + Northern Central America	12.74–26.83 (22.9)	A: 1.00
8	Mega-Mexico > inland diversification	10.51–23.42 (22.13)	A: 1.00
9	Honduras/Guatemala > diversification	11.22–24.83 (23.49)	A: 1.00
10	Mega-Mexico (Chiapas) + Central Meridional	12.96–26.19 (22.5)	A: 0.99
11	Central Meridional (Isthmus + South America)	9.39–19.6 (16.74)	D: 0.5, E: 0.25, F:0.25
12	Isthmus > diversification	7.2–18.89 (13.74)	D: 1.00
13	East Andes + West Andes	6.16–13.38 (11.07)	E: 0.49, F: 0.49
14	West Andes > diversification	5.89–12.73 (10.51)	E: 0.99
15	East Andes + Northern Colombia	5.46–11.42 (9.54)	F: 0.99

pattern in a biogeographical framework, with species of subclade III-A distributed through most of Mexico and south to the extreme west of Guatemala and Belize. In contrast, species of subclade III-B are restricted to Southern Mega-Mexico (present-day Honduras and bordering Eastern Guatemala and Western Belize). The only geographically overlapping exceptions are of *Z. meermanii* at the first divergence point and *Z. variegata/Z. prasina* from Southeastern Mexico and extending their distribution into Belize and Guatemala, as seen in subclade III-A rather than with the Southern Mega-Mexico clade. The next independent divergence in Mega-Mexico occurred at around 22.5 MA (HPD 95 %, 12.96–26.19; node 10), i.e. the individual placement of *Z. soconuscensis* (clade IV) from Chiapas (Mexico).

Based on our biogeographical analyses, species subsequently dispersed from ancestral Mega-Mexico towards the Central Meridional region (Isthmus of Costa Rica and Panama, West Andes, East Andes, and Northern Colombia) in clades V–VII. These clades showed a distinct distribution pattern with few range expansion events between the investigated areas. Species of clade V are separated from other Central Meridional taxa (i.e. South American species, or clades VI and VII) around 16.74 MA (HPD 95 %, 9.39–19.6; node 11) during the mid-Miocene. Then this clade diversified within a relatively short time in the present-day Isthmus of Costa Rica and Panama at 13.74 MA (HPD 95 %, 7.2–18.89; node 12) within the same epoch. Furthermore, *Zamia* in South America (clades VI + VII) started to diversify around 11.07 MA (HPD 95 %, 6.16–13.38; node 13), with the species displaying a distinct distribution pattern in the West Andes and a single dispersal event towards Isthmus (clade VI). This pattern contrasts with species restricted to the East Andes/Northern Colombia (clade VII), where dispersal between these two regions is uncertain. The diversification of two respective clades occurred quickly at 10.51 MA (HPD 95 %, 5.89–12.73; node 14) and 9.54 MA (HPD 95 %, 5.46–11.42; node 15). Overall, the cladogenesis events within the extant *Zamia* was initiated during the Oligocene (~29.14 MA), reached its maximum during the Miocene, and started decreasing later during the Pliocene and Pleistocene epochs (Fig. 3).

Diversification analyses

BAMM analyses detected no shift in the diversification rate across *Zamia* (Fig. 4). Furthermore, a slow decrease in diversification rate through time was found, which is largely congruent with the overall decreasing global paleotemperature since the evolutionary origin of the genus (Fig. 4). The per-clade rates of net diversification estimated by BAMM varied between 0.28 and 0.13 lineages per MA (Supplementary Data Fig. S5). Despite some differences among rate estimates, all estimates concurred in finding low diversification rates, with little evidence for solid rate heterogeneity across branches. The results also indicate relatively constant extinction rates.

The lineage including the Caribbean and Fischeri clades showed the steepest net diversification loss, from 0.28 lineages per MA at the root decreasing to ~0.10 lineages per MA at the tip. The major Mega-Mexico clade, i.e. clade III, was similar, with a decreasing net diversification rate from 0.27 to 0.09 lineages per MA. The two South American clades, VI and VII, showed less of an inclined loss, with 0.13 lineages per MA at the root and 0.08 lineages per MA at the tip (Supplementary Data Fig. S5).

Our analysis on the speciation rates of different cycad genera showed that the genus *Encephalartos* had a loss of speciation very similar to that found in *Zamia* (Supplementary Data Fig. S6). *Macrozamia* and *Ceratozamia* were found to have a slow, gradual and uniform decrease in speciation rate. However, the genus *Dioon*, which shares its distribution with *Ceratozamia* and partly with the genus *Zamia*, showed a different prediction (Supplementary Data Fig. S6) – a slow and minor but distinct increase in speciation from an estimated 7 MA, followed by a more stable continuation until the present time.

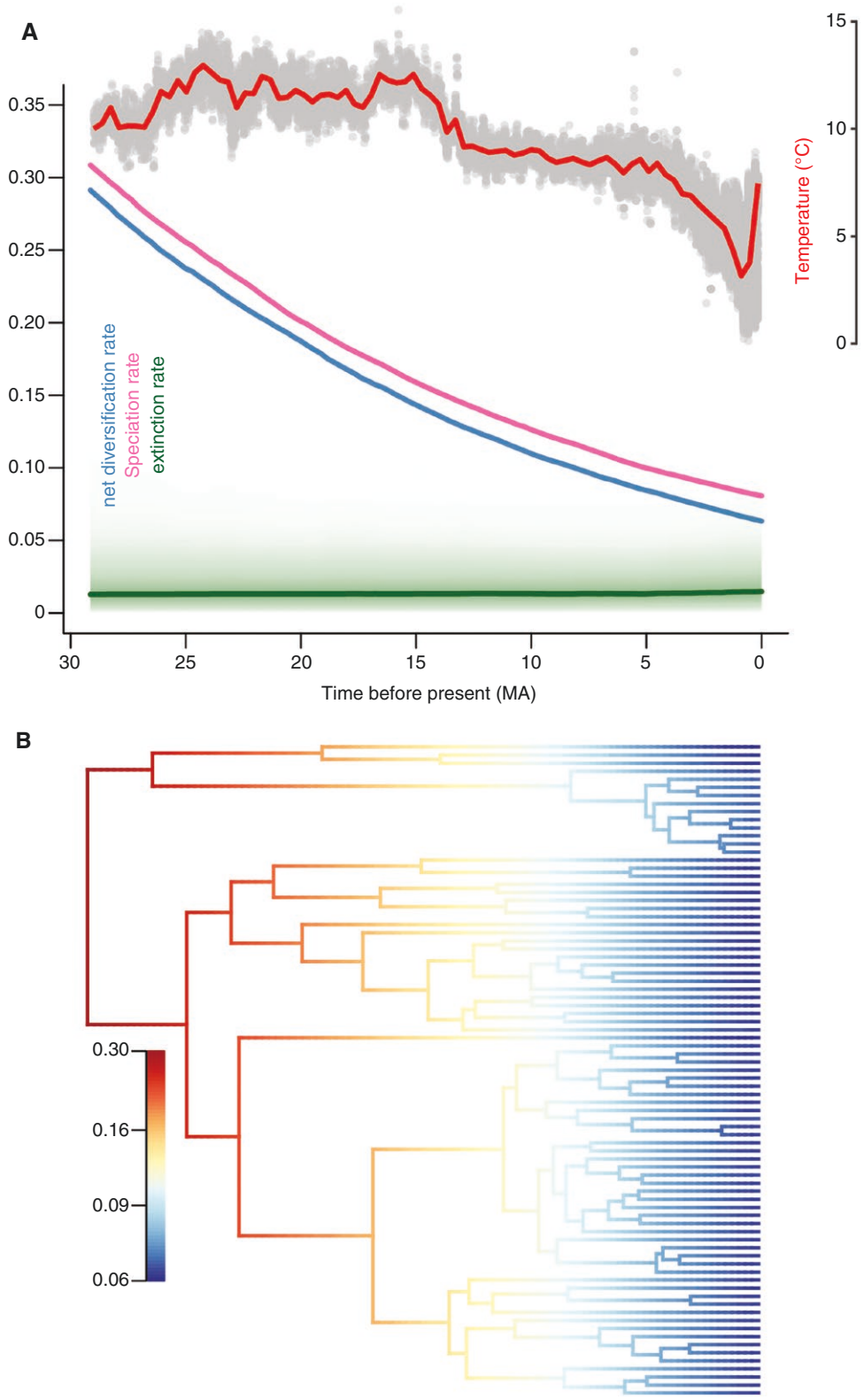

FIG. 4. Macroevolution of *Zamia*. (A) Median rates of speciation (pink), net diversification (blue) and extinction (green) estimated by BAMM, with colour density shading to denote confidence in evolutionary rates through time. The unit of speciation, extinction and diversification is speciation events per million years and the rates of niche traits are unitless. Global surface temperature changes are denoted by the red curve using the dataset of Zachos *et al.* (2008). (B) Phylogenetic distribution of net diversification rate shifts on the dated phylogeny from TreePL. The net diversification rates are displayed on branches with colours from low (blue) to high (red) values.

Character state evolution

Our character selections included 62 morphological characters having several characters that were not previously studied. Among these, ancestral states of 21 characters were found to be equivocal and 41 characters were found to have a strongly supported ancestral state in *Zamia* (Supplementary Data Table S2). Most of the investigated characters were homoplastic and experienced convergent evolution within the genus. Only a few characters showed somewhat consistent evolutionary trends in the phylogenetic tree. Thus, we selected ten (the four most significant characters in the main text as Figs 5 and 6, and the remaining six with a somewhat consistent pattern in Supplementary Data Figs S7–S9) out of 62 morphological characters to discuss their significance for major clades within the genus *Zamia*.

The ancestral state of stem habit (Fig. 5A) was found to be equivocal for the genus. The hypogeous stem habit appeared to be the most likely ancestral state, but with weak support (61 %). This character state is prominent in clades I, II and VII and subclade III-A and evolved once or twice in clades V and VI and subclade III-B. An arborescent stem was found as the most likely ancestral state (79 %) for clade V and equivocally for clade VI along with the other alternative character states. It appeared homoplasious once in clade II and subclade III-B, and twice in clade VII. An epigeous stem was prominent in clade IV and subclade III-B and appeared to have convergently evolved in several different clades (i.e. clades I, III-A, V, VI and VII). 'Tetrafoliate' eophylls (Fig. 5B) was found to be the ancestral state for the genus and all the major clades except subclade III-B. No uniformity of the other two states was found within any other clade, with the exception of the North Colombian species of clade VII, which showed uniformity in having eophylls with four or more leaflets (pinnate).

Pinna venation was treated as two segregated characters (Fig. 6), one with distinct plicate and deeply folded venation (e.g. in *Z. skinneri*) (Fig. 6A) and the other with pinnae with visible venation or raised veins (not plicate) (e.g. *Z. disodon*) (Fig. 6B). Both of these characters showed that non-plicate and inconspicuous venation of the leaflets is ancestral in the genus, as shown in Fig. 6A and B, respectively. Significant occurrence of the derived traits was reported only within clade V (Isthmus, 99 %) and clade VI (West Andes, 97 %). Our results indicate that these characters are homoplastic and have convergently evolved in seven separate lineages.

The character state of petiole prickles being uniformly 'absent' was only found in species of clades I and II (Supplementary Data Fig. S7A) and the state of petiole prickles being present was ancestral in all other clades. The absence of petiole prickles has evolved independently four times in clade V (Isthmus) and three times in clade VII (East Andes/North Colombia). The leaflet number was found for most of the species within the genus *Zamia* to be large (21–50 leaflets), followed by medium (6–20 leaflets). Moreover, the ancestral state for this character is uncertain for the whole genus as well as the major clades (Supplementary Data Fig. S7B).

Male peduncle length was found to be equivocal for the genus, with no pattern of occurrence in analyses. Only the North Colombian species of clade VII showed a uniform characteristic, with male peduncle length 'longer than strobili' (Supplementary Data Fig. S8A). In contrast, a female peduncle length character state of 'shorter than strobili' was found to be ancestral to the genus and for all major clades. It shows reticulate evolution of peduncle length as shorter, equal to or longer than strobili within clades III, VI and VII (Supplementary Data Fig. S8B).

Sarcotesta texture (Supplementary Data Fig. S9A) had equivocal character states for the genus. Only the Caribbean clade (clade I) with the mucilaginous texture state and the Northern Colombian species of clade VII with membranous texture state had complete conformity within the character. Sclerotesta of 'medium (12–20 mm)' length was supported as an ancestral state for the genus, with 83 % support; however, this character showed a complex evolutionary history among the major clades (Supplementary Data Fig. S9B).

DISCUSSION

Distinguishing *Zamia* species based on transcriptome data

The results presented here, based on a dataset of 2901 SCNGs with extensive taxon sampling, is by far the largest dataset used to date for the genus of *Zamia*. Furthermore, the utility of the coalescent method is analysed for the first time within this genus, and has also proven advantageous to reveal phylogenetic relationships within other cycad genera (Habib *et al.*, 2022, 2023; Liu *et al.*, 2024). With the suite of single-copy nuclear genes utilized here, seven major clades within the genus *Zamia* were successfully identified based on utilizing both concatenation and coalescent approaches. Consistent with previous studies, we found that the highly conserved individual gene sequences with fewer parsimoniously informative sites led to a large number of alternative trees with respect to the main consensus tree. Therefore, we retrieved the lower quartet support and concordance factor values for almost all the representative nodes of phylogenetic trees.

Phylogenetic relationships among the major *Zamia* clades found in our study were in general agreement with previous studies (Calonje *et al.*, 2019). Major topological conflicts were found with respect to the position of the Fischeri clade (clade II) and monophyly versus paraphyly of subclade III-B and clade VI in the different datasets and their analyses. Additionally, there were few conflicts regarding the phylogenetic placement of species at the shallow internal branches within the major clades. Generally, removal of the fast-evolving third codon is considered to yield better phylogenetic inferences because it reduces the impact of substitution saturation in gene tree estimation (Breinholt and Kawahara, 2013; Liu *et al.*, 2014). Several other factors, such as horizontal gene transfer, hybridization or incomplete lineage sorting, are also likely to cause conflicts. Nearly identical quartet frequencies for major (q1) and alternative topologies (q2, q3) for conflicting positions and lower concordance factor values indicate these conflicts may be as a consequence of rapid diversification and incomplete lineage sorting. PhyloNet analyses detected gene flow in the Fischeri clade (a clade with conflicting phylogenetic position) and Central Meridional clades (Supplementary Data Fig. S4), suggesting that ancient hybridization might have played essential roles in the evolution of *Zamia*. Moreover, concatenated

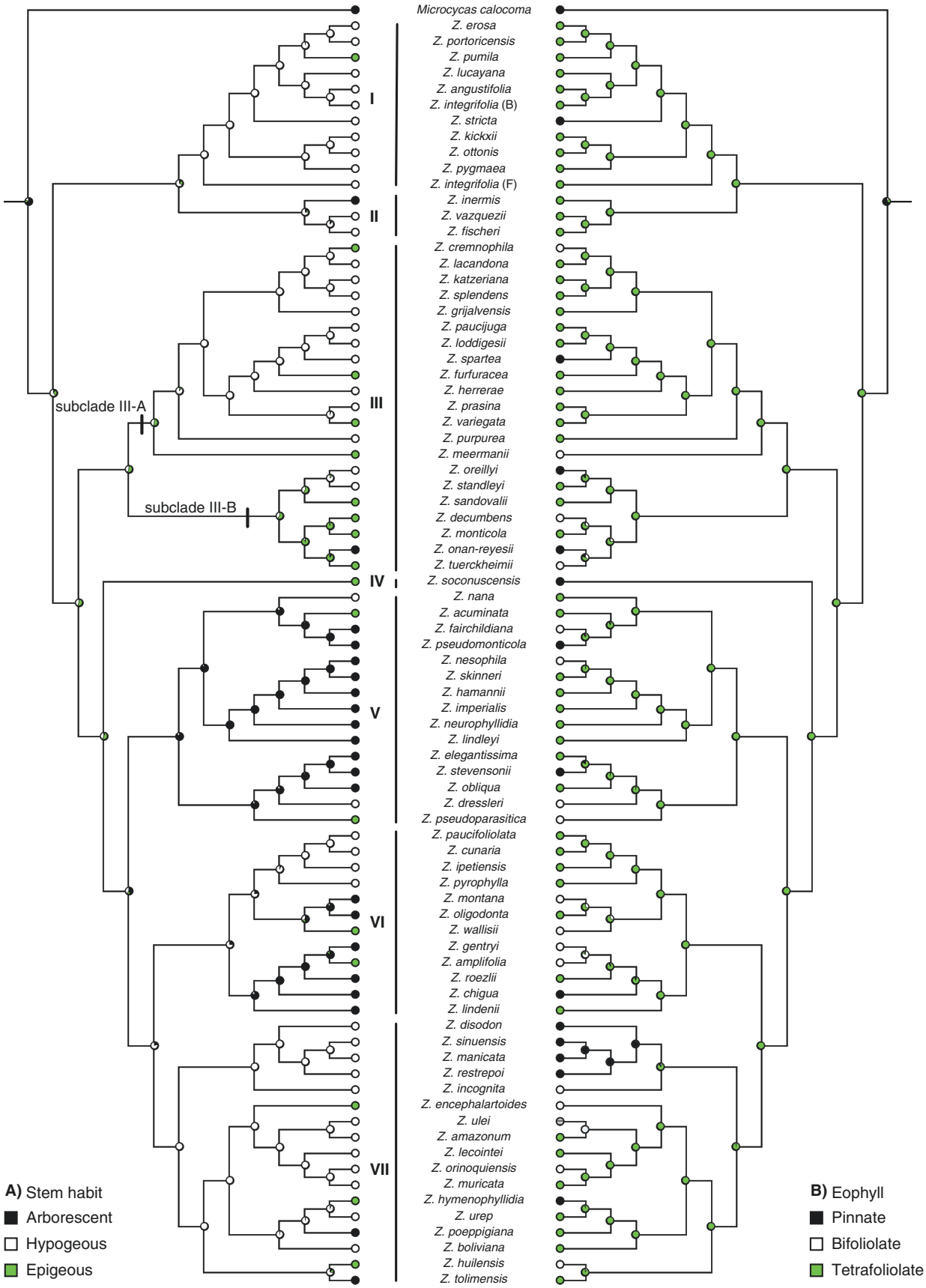

FIG. 5. Ancestral state reconstruction of (A) stem habit and (B) eophyll in Mesquite under the Mk1 model, on the phylogenetic tree of *Zamia* derived from coalescent (nt12) analyses. The pie chart at each node represents the probability of each character state for that morphological character.

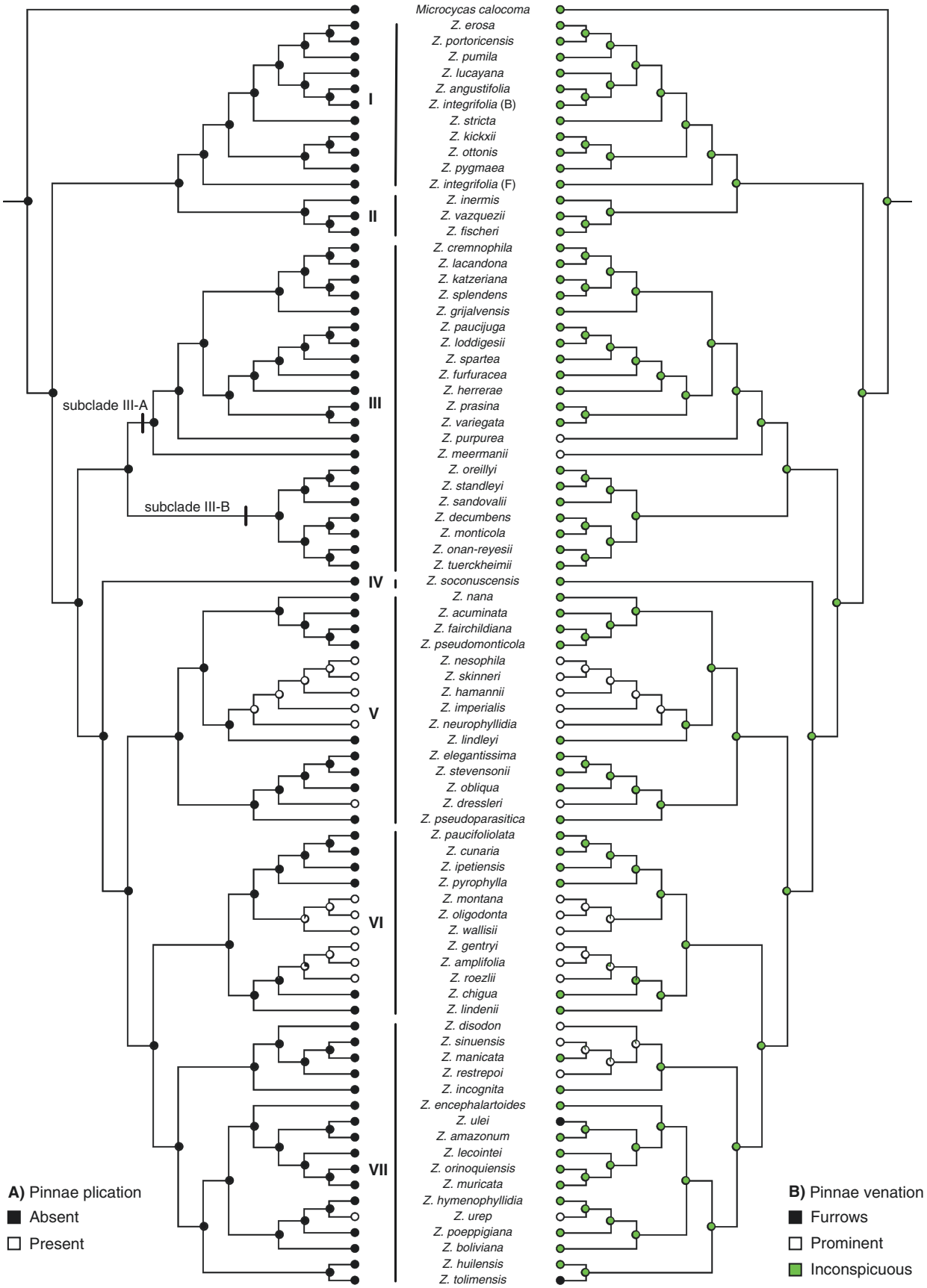

FIG. 6. Ancestral state reconstruction of (A) pinna plication and (B) pinna venation in Mesquite under the Mk1 model, on the phylogenetic tree of *Zamia* derived from coalescent (nt12) analyses. The pie chart at each node represents the probability of each character state for that morphological character.

alignment methods cannot handle heterogeneity among gene trees with a large dataset and this might produce false trees with high confidence values. Therefore, the phylogenetic positions recovered by the CDS-nt12 coalescent method are favoured (Supplementary Data Fig. S2). Further evidence from geographical distribution supports these phylogenetic placements. Hence, the phylogenetic tree that resulted from the coalescent CDS-nt12 analysis was used for later dating and character evolution analyses.

Phylogenetic placement of clades

Infrageneric distinction between major clades within *Zamia* is consistent with its regional diversity and biogeographical distribution, as suggested in previous studies (Caputo *et al.*, 2004; Zonneveld and Lindstrom, 2016; Calonje *et al.*, 2019). The ancestral origin of *Zamia* has been speculated previously to be from two widely separated areas: (1) South America (Passalia *et al.*, 2010), based partly on apparently misidentified plant fossil material and the assumed distinctiveness of the genus *Chigua* that is now known to be nested in *Zamia* (Caputo *et al.*, 2004; Erdei *et al.*, 2012); and (2) North America, which was based on a strong phylogenetic pattern of genome size (Zonneveld and Lindstrom, 2016) and multilocus sequence data with ten independent loci (Calonje *et al.*, 2019). Our result, based on a well-supported phylogenetic tree derived from transcriptome data, clearly settled the dispute on the origin of extant *Zamia* and confirmed it to be of North American origin rather than South American. These results are in congruence with the fossil origin of North American Zamiaceae as related to fossils, such as *Eostangeria* (Uzunova *et al.*, 2001). By comparing leaf anatomical characters with extant cycad genera, the fossil genus *Eostangeria* is confirmed to be part of Zamiaceae with closest relationship to *Zamia*, rather than any other genera within the family (Kvaček and Manchester, 1999; Uzunova *et al.*, 2001).

The Caribbean clade (clade I), including species from mainland USA (Florida), which consists of a debatable number of species (Eckenwalder, 1980; Stevenson, 1987; González-Géigel, 2003; Meerow *et al.*, 2018), was recovered here as a strongly supported monophyletic clade that corresponds with the findings of Calonje *et al.* (2019). The latest checklist shows a total of eight species with one variety from the Caribbean (Calonje *et al.*, 2024). We include a total of nine nomenclatural valid species and added *Z. kickxii* and *Z. ottonis* (Fig. 2), which are considered synonymous to *Z. pygmaea* in the World List of Cycads (Calonje *et al.*, 2024). The Caribbean taxa plus Florida (clade I) have long been recognized to be a uniform clade (Eckenwalder, 1980; Stevenson, 1987) in overall morphology, with subterranean stems, petioles lacking prickles, and a high leaflet width diversity from needle-like to 5 cm wide (Newell, 1985). Recent detailed molecular work on this clade has shown an intricate gene flow not only between populations but also between islands (Meerow *et al.*, 2007, 2012, 2018; Griffith *et al.*, 2022).

The Fischeri clade (clade II) consists of three morphologically distinct species distributed in a rather restricted area (Veracruz, Hidalgo, Querétaro, San Luis Potosí and Tamaulipas) within Mega-Mexico (Figs 2 and 3). *Zamia fischeri* was considered to be related to Caribbean species because of its unarmed petioles (Miquel, 1861). Caputo *et al.* (2004) also showed this species to be part of the Caribbean clade. Our results strongly support a distinct Fischeri clade, first mentioned by Calonje *et al.* (2019), comprising *Z. vazquezii*, *Z. fischeri* and *Z. inermis.* However, most of our results resolved this clade as sister to the Caribbean clade rather than the rest of mainland American species mentioned in Calonje *et al.* (2019). The genome size of the species within this clade was one of the smallest genome sizes (Zonneveld and Lindstrom, 2016). Recent findings suggest that small genome size is often found in species or species complexes that show particular adaptations to local conditions (Pyšek *et al.*, 2023). The Fischeri clade appears to have diverged along with the Caribbean clade and not as part of the Mega-Mexico clade at 26.32 MA (Fig. 3). This connection is also evident from the shared morphological characters such as the non-prickly petiole and similar reproductive characters. The persistence of the Fischeri clade in mainland Mega-Mexico is also interesting in the biogeographical context and supports our findings that the ancestral area was, in fact, part of Mega-Mexico.

Clade III is divided into two subclades based on their distinct regional distribution within Mega-Mexico. Subclade III-A consists of 15 species described from Mega-Mexico (14 sampled here), all hypogeous with rather uniform morphological characters (Fig. 5A), usually in semideciduous habitats, but some in more mesic habitats. This subclade has been shown to be a natural group based on genome size (Zonneveld and Lindstrom, 2016). The constituent species of this subclade, excluding *Z. meermanii*, also formed a strongly supported clade ($P = 1$; 'Mexico clade') in Calonje *et al.* (2019)'s phylogeny. The species of subclade III-A included both eastern and western Mexican species as sister taxa. Both samples of *Z. paucijuga*, representing the northern and southern extremes of the species distribution, are clearly shown to be part of the *loddigesii–furfuracea* complex. These findings are in contrast to those of Calonje *et al.* (2019), who found the Western Mega-Mexico species to be clearly separated in their own subclade although still within their Furfuracea subclade.

Species from Honduras, Guatemala and Belize within Mega-Mexico (Fig. 2, subclade III-B) form a distinct group that according to our results originated from a southern Mexican region. The subclade consists of seven described and one possibly undescribed (aff. *monticola*) species. Species within this subclade show high morphological and ecological diversity. These seem to be adapted to wet cloud forest or to drier deciduous forest habitats, or some are lithophytes with pendent leaves (Nelson Sutherland, 2006; Nelson Sutherland and Sandoval González, 2008).

Clade IV is a monospecific clade consisting of a single species, *Z. soconuscensis*, that grows in isolated pockets in cloud forest on steep mountain slopes in Chiapas, Mexico (Figs 2 and 3). Morphologically, it is a very distinct species. *Zamia soconuscensis* is sister to the Central Meridional species rather than its biogeographically closer Mega-Mexican species of clades II and III. This unexpected placement was first recovered by Calonje *et al.* (2019) and is fully supported by our data. Multiple divergence events within Mega-Mexico can be seen in a considerably short time frame and fast range extension at

node 7 (22.9 MA; Table 3) with the divergence between Mega-Mexico and ancestral Southern Central America, and Chiapas taxa at node 10 (22.5 MA; Table 3).

Clade V, the Isthmus clade, with a total of 15 described species, is the most morphologically diverse and ecologically adapted clade within the genus (Figs 2 and 3). Ecological adaptations range from the only known true epiphytic cycad to species adapted to beach-side habitats with high salinity (Taylor *et al.*, 2008). The morphology in this clade ranges from plicate to papyraceous leaflets and tall arborescent to small subterranean growth habits. The placement of *Z. dressleri* with *Z. pseudoparasitica* is peculiar as the morphological and habitat differences between these two species are extreme. *Zamia dressleri* is a subterranean small terrestrial species that has only one or two plicate leaves, whereas *Z. pseudoparasitica* is the only epiphytic cycad and has several inconspicuously veined leaves (Stevenson, 1993; Bell-Doyon and Villarreal, 2020). The sister subclade to this consists of *Z. obliqua*, *Z. elegantissima* and *Z. stevensonii*. These species are morphologically more similar to *Z. pseudoparasitica*. Calonje *et al.* (2019) resolved *Z. pseudoparasitica* (with PP = 0.70 support) to be in its own subclade diverged from *Z. dressleri*, although still within the Obliqua subclade, which corresponds to our findings. Diversification of species in South America at 11.07 MA follows a West versus East Andes clades, including Northern Colombia. This high *Zamia* species diversity within a small geographical area might be a result of the recent geological origin of the Isthmus of Panama (O'Dea *et al.*, 2016).

The West Andes (clade VI) clade consists of 12 described species and was indicated by Caputo *et al.* (2004) to be divergent from the East Andes clade (Figs 2 and 3). Species within this clade commonly inhabit wet rainforest and are mostly arborescent, with some species inhabiting coastal swamps (e.g. *Z. roezlii*) while others live at the highest altitude recorded within the genus in cloud forest (e.g. *Z. gentryi*). The West Andes clade in our results was resolved with strong support, a resolution lacking in the internal linages recovered in the analyses presented by Calonje *et al.* (2019). *Zamia wallisii*, *Z. oligodonta* and *Z. montana* formed a morphologically uniform subclade supported by morphological data in Calonje *et al.* (2019). Our study confirms that *Z. montana*, not sampled in the previous study, is part of this subclade. A subgroup within clade VI, consisting of four species, was also recorded as a distinct clade (Cunaria clade) by Calonje *et al.* (2019) with a uniform set of combining morphological characters. This subgroup has two species, *Z. cunaria* and *Z. ipetiensis*, that occur geographically within the Panamanian Isthmus clade and two species, *Z. pyrophylla* and *Z. paucifoliolata*, that are within the geographical boundaries of the West Andes, South America.

The East Andes/Northern Colombia (clade VII) clade consists of 17 species with the Northern Colombian species within it as a monophyletic group (Figs 2 and 3). Our study further corroborates the finding of Caputo *et al.* (2004) and Lindstrom (2009) that the genus *Chigua* is undoubtedly part of the genus *Zamia* together with a small clade of related taxa. Species of *Zamia* from North Colombia were considered as a separate clade sister to East Andes species (Calonje *et al.*, 2019) or nested within species from West Andes (Nagalingum *et al.*, 2011). The results from all datasets and comparative analyses in our study supported the concept that the North Colombian *Zamia* species were nested within the East Andes clade and that *Z. incognita* (East Andes) is sister to all the species in Northern Colombia. All the remaining taxa were inhabiting the East Andes region (Fig. 2, clade VII) and are highly morphologically diverse, from small subterranean to tall arborescent stems, and from ecologically adapted to extremely wet rainforest to dry lowland to high-altitude semi-desert. Most species included in this clade are very poorly documented and rarely collected, and the clade is expected to yield an increasing number of species in the future.

Diversification of *Zamia*

Results from our diversification analyses based on different lineages are coherent, and consistently indicate that the rate of diversification in *Zamia* was high at its origin and underwent a constant decrease throughout its evolutionary history (Fig. 4; Supplementary Data Fig. S5). Our time tree generated for *Zamia* was based on the most recent cycad fossil calibration (Coiro *et al.*, 2023), and consequently showed an older age than in a previous publication (Calonje *et al.*, 2019), which found a high diversification rate in the Caribbean clade and a rapid, early net diversification in the South American clade followed by a decrease in diversification rate over time. They found that the diversification rates only differed significantly in the South America clade using Monte Carlo Constant Rates (MCCR) tests. In accordance with Calonje *et al.* (2019), there was no shift in diversification rates reported across the tree.

Our results for diversification rates are in contrast to recent research on the genus *Cycas*, where these estimates showed an early burst of speciation, followed by a slowly decreasing rate from ~48 MA to ~20 MA, and a relatively stable to slightly increasing trend since then with a decline during the Pleistocene, including one significant rate shift identified within the section *Cycas* (Liu *et al.*, 2024). Besides the section *Cycas*, other major clades in the genus *Cycas* may have experienced a constantly decreasing rate (ranging from 0.3 to 0.15 lineages per MA) over time. This pattern is similar to *Zamia* (Supplementary Data Fig. S5), thus implying a slow descending trend of net diversification assumed for different cycad genera. For *Zamia*, the estimated initial speciation rates varied from 0.13 to 0.28 lineages per MA for different clades, which is comparable to that in *Cycas* (~0.15 to 0.30), but higher than pines (~0.08; Jin *et al.*, 2021). *Zamia* overall showed significantly higher historical speciation rates than other Neotropical cycad genera (i.e. *Dioon* and *Ceratozamia*; Supplementary Data Fig. S6), which can be attributed to its higher diversity and wider ranges in the New World.

Considering the similar speciation pattern of Neotropical cycads (Supplementary Data Fig. S6), the distribution pattern of *Zamia* is thus particularly interesting, as it has been with only minor exceptions one-directional towards the south, originating from Northern Central America, heading south through the Central American land bridge and into South America, all the way south to present-day Bolivia. This contrasts with the other two cycad genera in the Neotropics, which are currently restricted to habitat niches in montane deserts or rainforest habitats in the Mexican transition zone. Specifically, *Dioon* is

a relict cycad group that evolved during the formation of the current biogeographic provinces in the Mexican transition zone but did not disperse further south than present-day Honduras. Meanwhile, *Ceratozamia* are considerably younger in evolutionary terms but are restricted in their southern dispersal to present-day Honduras. It has been speculated that a possible factor that might have promoted the initial divergence of New World cycads is related to orogeny (Gutiérrez-Ortega *et al*., 2018). Orogeny, a significant geological process, has played a crucial role in Mexico's biogeography. It formed natural barriers that decreased the levels of precipitation in southern Mexico. This, in turn, allowed the origin and expansion of seasonally dry tropical forests and deserts, influencing the distribution and diversification of cycads. It is noteworthy that the genus *Dioon* has produced distinct speciation patterns in several geographical directions, along the west coast heading north and the east coast heading north, with only a single lineage dispersing south to present-day Honduras (Gutiérrez-Ortega *et al*., 2018). It has been suggested that non-random evolution with shifts towards more seasonal environments at high latitudes or shifts towards humid or dry environments at low latitudes explains the diversification of *Ceratozamia* (Gutiérrez-Ortega *et al.*, 2023*b*). For *Zamia*, it can be inferred that the dispersal through a more diverse range of habitats during the Neogene climate change and ongoing orogeny might both have facilitated geographical niche isolation, leading to higher speciation rates in *Zamia* than in other cycad genera in this region.

Our results also showed that the extinction rate was low in *Zamia*, and this is in contradiction to the hypothesized assumption that long stems in phylogenetic trees such as ours (Fig. 2A) are the result of mass extinction events. Moreover, if multiple co-existent lineages have long stems with crowns of similar age, this could indicate an extinction event driven by a major environmental change (Crisp and Cook, 2009, 2011). It should be noted that during the Cenozoic era at ~29 MA a large extinction of gymnosperms occurred at a rate approximately seven times higher than in angiosperms and was followed by similarly differential extinctions, but with less alterations (lower peaks) around 16 and 7–5 MA (Niklas, 1997; Crepet and Niklas, 2009; Crisp and Cook, 2011).

Morphological diversity in *Zamia*

Ancestral state reconstructions (ASRs) (Jermann *et al.*, 1995) are used to map morphological and/or ecological character traits onto a molecular phylogeny (Holland *et al.*, 2020). The selected characters can be examined and can sometimes be useful to re-evaluate past classifications. There are, however, cautions to be made, as voiced by Holland *et al.* (2020), that have shown some states in ASR are not consistent in many evolutionary systems, specifically for directional selection traits. The increasing error rates are linked to node depth, true number of state transitions, and rates of state transition as well as preferential extinction of species with the ancestral character state (Holland *et al.*, 2020). The genus *Zamia* has very short branches consistent with other cycad genera (Nagalingum *et al.*, 2011), and the true number of the rates of state transitions (Holland *et al.*, 2020) seems to have been less prone to extinction (Calonje *et al.*, 2019) than what has been suggested for other cycad genera, such as *Encephalartos* (Yessoufou *et al.*, 2014). Very limited information is available regarding major evolutionary transitions within *Zamia*, such as the ancestral origin of certain characters. This is, in part, a result of a poor fossil record.

In the genus *Zamia*, several character states were found to be homoplastic. Morphological characters in the genus seem to be related to either ecology and/or habitat adaptation or reproductive biology. We assume that our results might reflect a slightly different ASR if an expanded dataset of related genera as outgroups were used. Recent studies (Corio *et al.*, 2020; Glos *et al.*, 2022) also have recovered many characters as being homoplastic.

The morphological states of the outgroups could impact the overwhelming number of homoplastic characters. This phenomenon can be highlighted using stem habit as an example. If we had included *Stangeria*, which has a hypogeous stem, and *Ceratozamia*, which is polymorphic, then with a hypogeous stem being present in the first two clades, the arborescent *Microcycas* would be homoplasious.

The arborescent stem habit seems to be a result of convergent evolution and an adaptation to mesic habitat preference, rather than an ancestral trait in the genus *Zamia* (Fig. 5A), which is supported by our findings as being equivocal. However, several, if not most, known fossil cycads are described from stems that were arborescent in one way or another (Martínez *et al.*, 2012), as fleshy subterranean stem are much less likely to be fossilized. One of earliest cycad fossil stems is interpreted as slender and often branched (Delevoryas and Hope, 1976), as in the extant *Z. obliqua*. The subterranean (hypogeous) stem habit (Fig. 5A) as being an ancestral state is rejected as well, in contrast to Calonje *et al.* (2019), who found it to be ancestral. We included the eophyll character (Fig. 5B) because it was first speculated to be of taxonomic value (Schutzman, 2004), but it seems to be of limited taxonomic use for the genus as a whole. Calonje *et al.* (2019) investigated four morphological characters for ASR and found all of them to be homoplastic over the entire phylogeny. Our current results resolved subterranean stems, petiole lacking prickles, leaflets entire and prominent veins to be the ancestral state characters for the genus as confirmed by ASR.

All characters related to male and female strobili, peduncles and seeds are part of the reproductive biology. It is unknown at present which, if any, characters are directly linked to the pollinators and/or the dispersal agents and their ecology, but nevertheless they can all be considered as part of the reproductive biology of the plants themselves. The peduncle length of the strobili in both sexes has been used for taxonomic implications (Vovides *et al.*, 1983; Schutzman, 1989; González-Géigel, 2003) and even considered part of a set of characters to separate the genus *Chigua* from other related genera (Stevenson, 1990*b*, 2001*a*). The extended peduncle length in ovulate strobili found in all species within the North Colombian clade (clade VII) used to distinguish the genus *Chigua* but is now confirmed to be a combining character of this clade.

The texture of the sarcotesta when the seeds are ripe (Supplementary Data Fig. S8A) indicates usefulness in taxonomic circumscriptions of certain clades. The implication of sarcotesta texture within the Caribbean clade is most likely an adaptation to a non-specialist dispersal agent and observational data indicate that seeds in this clade attract birds and mammals as potential dispersal agents (Tang, 1989). The function of the

membranous texture of the sarcotesta in the North Colombian clade is unknown, but birds were observed swallowing seeds of species within this clade (Taborda-López and López-Gallego, 2018). The thin membranous texture means that the ripe seeds are exposed to faster desiccation, perhaps an adaptation to the constant moist and high-humidity habitat the species in this clade share.

Sclerotesta (seed) length (Supplementary Data Fig. S8B) has been suggested to be of dispersal importance (Symes, 2018). There is evidence of a positive relationship between seed size and animal seed dispersal (Eriksson, 2008; Monteza-Moreno *et al.*, 2022). No clear indication of any relationship was found between the phylogeny and this morphological character, as species belonging to different clades showed a variety of sclerotesta sizes. Notably, within the Caribbean clade (clade I) there is a subclade consisting of the smallest species within the genus, *Z. pygmaea*, *Z. kickxii* and *Z. ottonis*, and it is not surprising that this subclade shows a clear species versus seed size relationship. However, within the same Caribbean clade the only non-insular species, i.e. *Z. integrifolia* (Florida), has a larger than average (extreme) seed size.

Phylogenetic relationships affirmed in context of robust phylogeny

The relationships between major lineages in *Zamia*, as shown in the well supported phylogenetic tree, clarify the phylogenetic position of several species. The results are coherent with the genome size studied by Zonneveld and Lindstrom (2016), which correlates with most clades having the same geographical distribution. The phylogenetic position of *Z. ottonis* (clade I) supported it as a valid species (Fig. 2) rather than a synonym of *Z. pygmaea* (Calonje *et al.*, 2024). This concept was already suggested by González-Géigel (2003), but never widely accepted. *Zamia kickxii* was resolved as sister to *Z. pygmaea* in our analyses, but it is noteworthy that the type locality of these two species is not known as they were described from cultivated specimens. Nevertheless, our results clearly showed that there are at least two distinct separate species in Cuba, with *Z. ottonis* having diverged earlier than the latter two taxa. Previous studies placed *Z. purpurea* along with *Z. splendens*, *Z. katzeriana*, *Z. lacandona*, *Z. cremnophila* and *Z. grijalvensis* in a 'purpurea' clade (Calonje *et al.*, 2019); however, our results show that *Z. purpurea* is more genetically distinct than previously realized and appears to be nearer to *Z. meermanii* and sister to all remaining taxa of clade III-A (Fig. 2). *Zamia katzeriana* has been suggested to be of natural hybrid origin (Perez-Farrera *et al.*, 2016). A recent study has confirmed it not to be of hybrid origin and as a distinct species (Gutiérrez-Ortega *et al.*, 2023*a*). Our data indicate that this species is not phylogenetically close to one of the purported parents, *Z. loddigesii*. *Zamia katzeriana* is sister to the morphologically similar *Z. splendens*, but it has enough morphological distinction to be considered a distinct species based on the erect female strobilus peduncle and the narrow oblong-lanceolate leaflets (Perez-Farrera *et al.*, 2016). *Zamia restrepoi* was once considered to be part of a distinct genus, *Chigua* (as *Chigua restrepoi*), based on the apparent midrib on the leaflets (Stevenson, 1990*b*, 2001*a*), but was later synonymized based on being nested within *Zamia* with molecular data (Caputo *et al.*, 2004) and morphological differences (Lindstrom, 2009). Thus, the leaflet midrib is an autapomorphy within *Zamia*. Our study used extensive data and taxon sampling, and confirmed the inclusion and phylogenetic position of *Z. restrepoi* within the Northern Colombia subgroup along with *Z. disodon*, *Z. manicata* and *Z. sinuensis*. We did not sample *Z. melanorrhachis* and *Z. imbricata*, but we believe they are likely part of this subgroup based on morphological synapomorphy. This assemblage of species, referred to as the Manicata clade by Calonje *et al.* (2019, 2021), shares a distinct set of morphological characters such as toothed leaflets, megasporophylls with relatively flat shields, and microsporophylls with a much-reduced fertile section of the lamina. The incomplete sampling is most likely what resulted in less resolution within this group. In addition to these, several poorly supported relationships in previous studies (Nagalingum *et al.*, 2011; Calonje *et al.*, 2019) within the West Andes (clade VI) and the East Andes/North Colombia clade (clade VII) showed a strong support in our phylogenetic tree with a slight inconsistency compared with previous studies. The subgroup within clade VII differs from Calonje *et al.* (2019), who had weak support (PP = 0.64) for *Z. amazonum* placed with *Z. urep/Z. hymenophyllidia*. Our results resolved *Z. amazonum* and *Z. ulei* together as sister to a clade comprising *Z. muricata* and *Z. lecointei* as a species pair sister to the newly described *Z. orinoquiensis*. *Zamia orinoquiensis* as a newly described species was not morphologically compared to *Z. lecointei* but rather to *Z. muricata* because populations of *Z. orinoquiensis* had been included in Stevenson's (2001*a*) circumscription of *Z. muricata* and thus had to be segregated from that species rather than from *Z. lecointei*. Overall, the common observation among these studies that the phylogenetic position of species is consistent with the geographical distribution clearly shows species diversification along a geographical gradient in *Zamia*, which is also reported in the other genera from this region, i.e. *Dioon* and *Ceratozamia* (Gutiérrez-Ortega *et al.*, 2023*a*, 2024; Habib *et al.*, 2023).

CONCLUSIONS

Based on an extensive transcriptome dataset, the robust phylogeny reconstructed in this study provides insights for better understanding of the deep relationships within the genus *Zamia*. The incongruences between phylogenetic trees from coalescent and concatenated analyses suggests that hybridization or incomplete lineage sorting may have occurred in the early diversification of the genus, which is evident from quartet and PhyloNet analyses. Furthermore, because the concatenation alignment of a large dataset cannot handle heterogeneity among gene trees, this may have led to the incongruence between the two analyses. The morphological characters showed a complex evolutionary history and infrageneric relationships that are shaped based on the biogeographical distribution of the species in *Zamia*. To critically explore the evolutionary complexity of cycad genera and association of infrageneric relationships with the regional distribution, a comprehensive study with a sampling strategy covering all currently recognized cycad species utilizing nuclear genomic data is needed. This in turn will provide insight

into global and regional diversity of this evolutionarily significant plant group.

SUPPLEMENTARY DATA

Supplementary data are available at *Annals of Botany* online and consist of the following. Figure S1: phylogenetic backbone of major clades of *Zamia* based on concatenated datasets of nt12 (left) and nt (right). BS >100 is noted above and the numbers of species per clade are indicated below the branches. Figure S2: phylogenetic tree of the genus *Zamia* generated with wASTRAL analyses, based on a nucleotide dataset of 2901 SCNGs using all codon positions (nt) on the right, and first and second codon positions only (nt12) on the left. Branch support is not mentioned if branches are maximally supported in both datasets. Dotted lines indicate conflicting phylogenetic positions of taxa among two datasets. Major phylogenetic groups, i.e. clades and subclades, in the genus are recognized and indicated with black vertical bars with species names and on branches, respectively. Figure S3: species trees of the genus *Zamia* using coalescent analyses of the first and second codon positions only dataset (nt12). The site concordance factor (sCF) and gene concordance factor (gCF) are displayed above the branches. Support values from quartet analysis are shown below the branches. '–' indicates quartet support <0.3. Figure S4: the best species tree network inferred with PHYLONET (reticulations = 3) for hybridization simulation. Blue lines indicate reticulation edges connecting reticulation nodes. Figure S5: median rates of speciation in time before present for the genus *Zamia* and all of its major clades estimated by BAMM, with colour-density shading to denote confidence in evolutionary rates through time. Figure S6: plot of speciation rates through time for different cycad genera. The rate trends of these genera were estimated by BAMM based on the fossil-calibrated phylogenies in Coiro *et al.* (2023), except *Zamia* (in red), for which the dated phylogenetic tree was inferred in this study. Figure S7: ancestral state reconstruction of (A) petiole prickles and (B) leaflet numbers in Mesquite under the Mk1 model, on the phylogenetic tree of *Zamia* derived from coalescent (nt12) analyses. Figure S8: ancestral state reconstruction of (A) male peduncle length and (B) female peduncle length in Mesquite under the Mk1 model, on the phylogenetic tree of *Zamia* derived from coalescent (nt12) analyses. Figure S9: ancestral state reconstruction of (A) sarcotesta texture and (B) sclerotesta length in Mesquite under the Mk1 model, on the phylogenetic tree of *Zamia* derived from coalescent (nt12) analyses. Table S1: read statistics of 86 taxa investigated in this study. Table S2: list of morphological characters investigated for ancestral state reconstructions in Mesquite of the genus *Zamia*. Ancestral states of characters are highlighted in red, if present.

FUNDING

This study was supported by the Scientific Foundation of Urban Management Bureau of Shenzhen (No. 202019 to S.Z., No. 202105 to Y.G., No. 202208 to S.Z.) and the Biodiversity Survey and Assessment Project of the Ministry of Ecology and Environment, China (No. 2019HJ2096001006 to S.Z.) and, in part, by NSF Grants BSR-8607049, EF-0629817 and DEB-2140319 to D.S.

ACKNOWLEDGEMENTS

S.Z. and A.L. conceived the idea for experiment; A.L. collected the plant materials; S.H. and Y.G. performed the experiments; S.H., S.D., J.L. and Y.G. performed data analyses; A.L. and S.H. drafted the manuscript. J.L., D.S. and M.C. provided extensive scientific and grammatical revision. The manuscript was finalized with input from all the authors. We would like to thank Nong Nooch Tropical Botanical Garden, Thailand, especially the President Kampon Tansacha for allowing us to collect plant samples for this study. We also thank Dr Juan Carlos of Université Laval, Québec, Canada, Dr Liu Yang of Fairy Lake Botanical Garden, Shenzhen, China, and Dr Yong Yang from Nanjing Forest University, Nanjing, China, for their helpful comments.

LITERATURE CITED

Abascal F, **Zardoya R**, **Telford MJ. 2010**. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* **38**: W7–13.

Bell-Doyon P, **Villarreal JC. 2020**. New notes on the ecology of the epiphytic gymnosperm and Panamanian endemic *Zamia pseudoparasitica*. *Neotropical Naturalist* **2**: 1–7.

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

Breinholt JW, **Kawahara AY. 2013**. Phylotranscriptomics: saturated third codon positions radically influence the estimation of trees based on next-gen data. *Genome Biology and Evolution* **5**: 2082–2092.

Calonje M, **Meerow AW**, **Griffith MP**, ***et al***. **2019**. A time-calibrated species tree phylogeny of the New World cycad genus *Zamia* L. (Zamiaceae, Cycadales). *International Journal of Plant Sciences* **180**: 286–314.

Calonje M, **Hernández JC**, **Coca LF**, **Jaramillo D**, **Aristizábal 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, **Stevenson DW**, **Osborne R. 2024**. *The world list of cycads*. Online edition. https://www.cycadlist.org/ (29 Feburary 2024).

Caputo P, **Cozzolino S**, **de Luca P**, **Moretti A**, **Stevenson DW. 2004**. Molecular phylogeny of *Zamia* (Zamiaceae). In: **Walters T**, **Osborne R**, eds. *Cycad classification: concepts and recommendations*. Wallingford: CABI Publishing, 149–157.

Chang ACG, **Lai Q**, **Chen T**, ***et al***. **2020**. The complete chloroplast genome of *Microcycas calocoma* (Miq.) A. DC. (Zamiaceae, Cycadales) and evolution in Cycadales. *PeerJ* **8**: e8305.

Coiro M, **Jelmini N**, **Neuenschwander H**, ***et al***. **2020**. Evolutionary signal of leaflet anatomy in the Zamiaceae. *International Journal of Plant Sciences* **181**: 697–715.

Coiro M, **Allio R**, **Mazet N**, **Seyfullah LJ**, **Condamine FL. 2023**. Reconciling fossils with phylogenies reveals the origin and macroevolutionary processes explaining the global cycad biodiversity. *New Phytologist* **240**: 1616–1635.

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**: 1–18.

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. 2009**. Explosive radiation or cryptic mass extinction? Interpreting signatures in molecular phylogenies. *Evolution* **63**: 2257–2265.

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

Delevoryas T, **Hope RC. 1976**. More evidence for a slender growth habit in Mesozoic cycadophytes. *Review of Palaeobotany and Palynology* **21**: 93–100.

Eckenwalder JE. 1980. Taxonomy of the West Indian cycads. *Journal of the Arnold Arboretum* **61**: 701–722.

Emms DM, **Kelly S. 2019**. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**: 1–14.

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.

Eriksson O. 2008. Evolution of seed size and biotic seed dispersal in angiosperms: paleoecological and neoecological evidence. *International Journal of Plant Sciences* **169**: 863–870.

Glos RAE, **Salzman S**, **Calonje M**, ***et al***. **2022**. Leaflet anatomical diversity in *Zamia* (Cycadales: Zamiaceae) shows little correlation with phylogeny and climate. *Botanical Review* **88**: 437–452.

González-Géigel L. 2003. Zamiaceae. In: **Greuter W**, ed. *Flora de la República de Cuba. Serie A. Plantas Vasculares*. Liechtenstein: Gantner, 3–22.

Grabherr MG, **Haas BJ**, **Yassour M**, ***et al***. **2011**. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**: 644–652.

Griffith MP, **Meerow AW**, **Calonje M**, **Gonzalez E**, **Nakamura K**, **Francisco-Ortega J. 2022**. Genetic patterns of *Zamia* in Florida are consistent with ancient human influence and recent near extirpation. *International Journal of Plant Sciences* **183**: 169–185.

Gutiérrez-Ortega JS, **Salinas-Rodríguez MM**, **Martínez JF**, ***et al***. **2018**. The phylogeography of the cycad genus *Dioon* (Zamiaceae) clarifies its Cenozoic expansion and diversification in the Mexican transition zone. *Annals of Botany* **121**: 535–548.

Gutiérrez-Ortega JS, **Pérez-Farrera MA**, **Lopez S**, **Vovides AP. 2023***a*. Demographic history and species delimitation of three *Zamia* species (Zamiaceae) in south-eastern Mexico: *Z. katzeriana* is not a product of hybridization. *Botanical Journal of the Linnean Society* **202**: 110–133.

Gutiérrez-Ortega JS, **Pérez-Farrera MA**, **Matsuo A**, ***et al***. **2023***b*. The phylogenetic reconstruction of the Neotropical cycad genus *Ceratozamia* (Zamiaceae) reveals disparate patterns of niche evolution. *Molecular Phylogenetics and Evolution* **190**: 107960.

Gutiérrez-Ortega JS, **Pérez-Farrera MA**, **Sato M**, ***et al***. **2024**. Evolutionary and ecological trends in the Neotropical cycad genus *Dioon* (Zamiaceae): an example of success of evolutionary stasis. *Ecological Research* **39**. doi:10.1111/1440-1703.12442.

Habib S, **Dong S**, **Liu Y**, **Liao W**, **Zhang S. 2021**. The complete mitochondrial genome of *Cycas debaoensis* revealed unexpected static evolution in gymnosperm species. *PLoS One* **16**: e0255091.

Habib S, **Gong Y**, **Dong S**, ***et al***. **2022**. Phylotranscriptomics reveal the spatio-temporal distribution and morphological evolution of *Macrozamia*, an Australian endemic genus of Cycadales. *Annals of Botany* **130**: 671–685.

Habib S, **Gong Y**, **Dong S**, ***et al***. **2023**. Phylotranscriptomics shed light on intrageneric relationships and historical biogeography of *Ceratozamia* (Cycadales). *Plants (Basel)* **12**: 478.

Helfrich P, **Rieb E**, **Abrami G**, **Lücking A**, **Mehler A. 2018**. TreeAnnotator: versatile visual annotation of hierarchical text relations. In: **Calzolari N**, **Choukri K**, **Cieri C**, **Declerck T**, **Goggi S**, **Hasida K** *et al*., eds. Proceedings of the Eleventh International Conference *on* Language Resources and Evaluation (LREC *2018*). European Language Resources Association (ELRA). https://aclanthology.org/L18-1308.

Holland BR, **Ketelaar-Jones S**, **O'Mara AR**, **Woodhams MD**, **Jordan GJ. 2020**. Accuracy of ancestral state reconstruction for non-neutral traits. *Scientific Reports* **10**: 7644.

Jermann TM, **Opitz JG**, **Stackhouse J**, **Benner SA. 1995**. Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily. *Nature* **374**: 57–59.

Jin WT, **Gernandt DS**, **Wehenkel C**, **Xia XM**, **Wei XX**, **Wang XQ. 2021**. Phylogenomic and ecological analyses reveal the spatiotemporal evolution of global pines. *Proceedings of the National Academy of Sciences of the USA* **11**: e2022302118.

Kalyaanamoorthy S, **Minh BQ**, **Wong TK**, **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 version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kennedy JD, **Marki PZ**, **Fjeldsa J**, **Rahbek C. 2020**. The association between morphological and ecological characters across a global passerine radiation. *Journal of Animal Ecology* **89**: 1094–1108.

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

Laetsch DR, **Blaxter ML. 2017**. KinFin: software for taxon-aware analysis of clustered protein sequences. *G3 (Bethesda)* **7**: 3349–3357.

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

Lindstrom AJ. 2009. Typification of some species names in *Zamia* L. (Zamiaceae), with an assessment of the status of *Chigua* D. Stev. *Taxon* **58**: 265–270.

Lindstrom AJ, **Calonje M**, **Stevenson DW**, **Husby C**, **Taylor A. 2013**. Clarification of *Zamia acuminata* and a new *Zamia* species from Coclé Province, Panama. *Phytotaxa* **98**: 27–42.

Liu Y, **Cox CJ**, **Wang W**, **Goffinet B. 2014**. Mitochondrial phylogenomics of early land plants: mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias. *Systematic Biology* **63**: 862–878.

Liu J, **Lindstrom AJ**, **Gong X. 2022**. Towards the plastome evolution and phylogeny of *Cycas* L. (Cycadaceae): molecular-morphology discordance and gene tree space analysis. *BMC Plant Biology* **22**: 116.

Liu Y, **Wang S**, **Li L**, ***et al***. **2022**. The *Cycas* genome and the early evolution of seed plants. *Nature Plants* **8**: 389–401.

Liu J, **Lindstrom A**, **Gong Y**, ***et al***. **2024**. Eco-evolutionary evidence for the global diversity pattern of *Cycas* (Cycadaceae). *Journal of Integrative Plant Biology* **66**: 1170–1191. https://doi.org/10.1111/jipb.13638.

Maddison WP, **Maddison DR. 2023**. *Mesquite: A MODULAR SYSTEM FOR EVOLUTIONARY ANALYSIS*. Version 3.81. http://www.mesquiteproject.org (15 November 2023, date last accessed).

Martínez LCA, **Artabe AEE**, **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.

Matzke N. 2018. *BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis with R scripts. Version 1.1. 1*. San Francisco: GitHub. https://github.com/nmatzke/BioGeoBEARS.

Meerow AW, **Stevenson DW**, **Moynihan J**, **Francisco-Ortega J. 2007**. Unlocking the coontie conundrum: the potential of microsatellite DNA studies in the Caribbean *Zamia pumila* complex (Zamiaceae). *Memoirs of the New York Botanical Garden* **98**: 484–518.

Meerow AW, **Francisco-Ortega J**, **Calonje M**, ***et al***. **2012**. *Zamia* (Cycadales: Zamiaceae) on Puerto Rico: asymmetric genetic differentiation and the hypothesis of multiple introductions. *American Journal of Botany* **99**: 1828–1839.

Meerow AW, **Salas-Leiva DE**, **Calonje M**, ***et al***. **2018**. Contrasting demographic history and population structure of *Zamia* (Cycadales: Zamiaceae) on six islands of the Greater Antilles suggests a model for population diversification in the Caribbean clade of the genus. *International Journal of Plant Sciences* **179**: 730–757.

Minh BQ, **Schmidt HA**, **Chernomor O**, ***et al***. **2020**. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**: 1530–1534.

Miquel FAW. 1861. *Prodromus Systematis Cycadearum: In Honorem Festi Diei Xv Kal. M. Julii Mdccclxi Quo Academia Rheno-Trajectina Exacta Xlv Lustra Celebrat*. Utrecht: C. Van der Post Jr.

Monteza-Moreno CM, **Rodriguez-Castro L**, **Castillo-Caballero PL**, **Toribio E**, **Saltonstall K. 2022**. Arboreal camera trapping sheds light on seed dispersal of the world's only epiphytic gymnosperm: *Zamia pseudoparasitica*. *Ecology and Evolution* **12**: e8769.

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

Nelson Sutherland CH. 2006. Dos plantas del genero *Zamia* (Gimnosperma) nuevas de Honduras. *Ceiba* **46**: 41–44.

Nelson Sutherland CH, **Sandoval González GG. 2008**. Una especie nueva de *Zamia* (Zamiaceae) de Honduras. *Ceiba* **49**: 135–135.

Newell SJ. 1985. Intrapopulational variation in leaflet morphology of *Zamia pumila* L. in relation to microenvironment and sex. *American Journal of Botany* **72**: 217–221.

Nicolalde-Morejón F, **Vovides AP**, **Stevenson DW. 2009**. Taxonomic revision of *Zamia* in Mega-Mexico. *Brittonia* **61**: 301–335.

Niklas KJ. 1997. *The Evolutionary Biology of Plants*. Chicago: University of Chicago Press.

Norstog K, **Nicholls T. 1997**. *Biology of the Cycads*. Ithaca: Cornell University Press.

O'Dea A, **Lessios HA**, **Coates AG**, ***et al*. 2016**. Formation of the Isthmus of Panama. *Science Advances* **2**: e1600883.

Passalia MG, **Del Fueyo G**, **Archangelsky S. 2010**. An Early Cretaceous zamiaceous cycad of south west Gondwana: *Restrepophyllum* nov. gen. from Patagonia, Argentina. *Review of Palaeobotany and Palynology* **161**: 137–150.

Perez-Farrera MA, **Vovides AP**, **Ruiz-Castillejos C**, **Galicia S. 2016**. Anatomy and morphology suggest a hybrid origin of *Zamia katzeriana* (Zamiaceae). *Phytotaxa* **270**: 161–181.

Pérez-Farrera MA, **Gutiérrez-Ortega JS**, **Martínez-Martínez MG**, **Calonje M. 2023**. *Zamia magnifica* (Zamiaceae, Cycadales): a new rupicolous cycad species from Sierra Norte, Oaxaca, Mexico. *Taxonomy* **3**: 232–249.

Plummer M, **Best N**, **Cowles K**, ***et al*. 2006**. CODA: convergence diagnosis and output analysis for MCMC. *R News* **6**: 7–11.

Pyšek P, **Lučanová M**, **Dawson W**, ***et al*. 2023**. Small genome size and variation in ploidy levels support the naturalization of vascular plants but constrain their invasive spread. *New Phytologist* **239**: 2389–2403.

Rabosky DL. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PLoS One* **9**: e89543.

Rabosky DL, **Grundler M**, **Anderson C**, ***et al*. 2014**. BAMM tools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. *Methods in Ecology and Evolution* **5**: 701–707.

Rai HS, **O'Brien HE**, **Reeves PA**, **Olmstead RG**, **Graham SW. 2003**. Inference of higher order relationships in the cycads from a large chloroplast data set. *Molecular Phylogenetics and Evolution* **29**: 350–359.

Salas-Leiva DE, **Meerow AW**, **Calonje M**, ***et al*. 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.

Sanderson MJ. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* **19**: 101–109.

Schutzman B. 1984. A new species of *Zamia* L. (Zamiaceae, Cycadales) from Chiapas, Mexico. *Phytologia* **55**: 299–304.

Schutzman B. 1989. A new species of *Zamia* from Honduras. *Systematic Botany* **14**: 214–219.

Schutzman B. 1998. *Revisionary studies of Mesoamerican Zamia L. (Zamiaceae, Cycadales)*, PhD Thesis, University of Florida, USA.

Schutzman B. 2004. Systematics of Meso-American *Zamia* (Zamiaceae). In: **Walters T**, **Osborne R**, eds. *Cycad Classification: Concepts and Recommendations*. Wallingford: CABI Publishing, 159–172.

Schutzman B, **Vovides AP**, **Dehgan B. 1988**. Two new species of *Zamia* (Zamiaceae, Cycadales) from southern Mexico. *Botanical Gazette* **149**: 347–360.

Segalla R, **Telles FJ**, **Pinheiro F**, **Morellato P. 2019**. A review of current knowledge of Zamiaceae, with emphasis on *Zamia* from South America. *Tropical Conservation Science* **12**: 1–21.

Shen W, **Le S**, **Li Y**, **Hu F. 2016**. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**: e0163962.

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

Smith SA, **Brown JW**, **Walker JF. 2018**. So many genes, so little time: a practical approach to divergence-time estimation in the genomic era. *PLoS One* **13**: e0197433.

Stevenson DW. 1987. Again the West Indian *Zamias*. *Fairchild Tropical Garden Bulletin (USA)* **42**: 23–27.

Stevenson DW. 1990*a*. Morphology and systematics of the Cycadales. *Memoirs of the New York Botanical Garden* **57**: 8–55.

Stevenson DW. 1990*b*. *Chigua*, a new genus in the Zamiaceae with comments on its biogeographic significance. *Memoirs of the New York Botanical Garden* **57**: 169–172.

Stevenson DW. 1991. Zamiaceae. In: **Görts Van Rijn ARA**, ed. *Flora of the Guianas. Series A: Fascicle 9*. Koenigstein: Koeltz Scientific Books, 7–11.

Stevenson DW. 1993. The Zamiaceae in Panama with comments on phytogeography and species relationships. *Brittonia* **45**: 1–16.

Stevenson DW. 2001*a*. Orden Cycadales. In: **Bernal R**, **Forero E**, eds. *Flora de Colombia*. Bogotá: UNIBIBLOS, 1–92.

Stevenson DW. 2001*b*. Zamiaceae. In: **Stevens W**, **Ulloa C**, **Pool A**, **Montiel O**, eds. *Flora de Nicaragua*. St Louis: Missouri Botanical Garden Press, 6–7.

Stevenson DW. 2004. Cycads of Colombia. *Botanical Review* **70**: 194–234.

Symes CT. 2018. Cycad seed dispersal: the importance of large frugivorous birds. *Biodiversity Observations* **9**: 1–9.

Taborda-López D, **López-Gallego C. 2018**. *Patrones de exito de la polinizacion y dispersion secundaria de semillas en Zamia manicata en un bosque humedo tropical del Darien Colombia*. Thesis, University of Antioquia, Colombia.

Talavera G, **Castresana J. 2007**. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* **56**: 564–577.

Tang W. 1989. Seed dispersal in the cycad *Zamia pumila* in Florida. *Canadian Journal of Botany* **67**: 2066–2070.

Taylor AS, **Haynes JL**, **Holzman G. 2008**. Taxonomical, nomenclatural and biogeographical revelations in the *Zamia skinneri* complex of Central America (Cycadales: Zamiaceae). *Botanical Journal of the Linnean Society* **158**: 399–429.

Than C, **Ruths D**, **Nakhleh L. 2008**. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics* **9**: 1–16.

Uzunova K, **Palamarev E**, **Kvacek Z. 2001**. *Eostangeria ruzinciniа*na (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.

Vovides A, **Rees J**, **Vázquez-Torres M. 1983**. *Flora de Veracruz. Zamiaceae. Fascículo 26*. Xalapa: Instituto Nacional de Investigaciones sobre Recursos Bióticos.

Yessoufou K, **Bamigboye SO**, **Daru BH**, **van der Bank M. 2014**. Evidence of constant diversification punctuated by a mass extinction in the African cycads. *Ecology and Evolution* **4**: 50–58.

Yu Y, **Blair C**, **He X. 2020**. RASP 4: ancestral state reconstruction tool for multiple genes and characters. *Molecular Biology and Evolution* **37**: 604–606.

Zachos JC, **Dickens GR**, **Zeebe RE. 2008**. An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature* **451**: 279–283.

Zhang C, **Mirarab S. 2022**. Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution* **39**: msac215.

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

Zonneveld B, **Lindstrom AJ. 2016**. Genome sizes for 71 species of *Zamia* (Cycadales: Zamiaceae) correspond with three different biogeographic regions. *Nordic Journal of Botany* **34**: 744–751.