JSE Journal of Systematics and Evolution

doi: 10.1111/jse.70034

Research Article

Integrative phylogenomics and morphology reveal the evolution and biogeography of *Encephalartos* (Zamiaceae)

Sadaf Habib[1,2], Anders Lindstrom[3], James A.R. Clugston[4,5], Yiqing Gong[2], Shanshan Dong[2], Yunhua Wang[2], Dennis Stevenson[6], Chen Feng[1]*, and Shouzhou Zhang[2]*

[1]Jiangxi Provincial Key Laboratory of Ex Situ Plant Conservation and Utilization, Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang 332900, China

[2]Key Laboratory of Southern Subtropical Plant Diversity, Fairy Lake Botanical Garden, Shenzhen & Chinese Academy of Sciences, Shenzhen 518004, China

[3]Global Biodiversity Conservancy 144/124 Moo 3, Soi Bua Thong, Bangsalae, Sattahip, Chonburi 20250, Thailand

[4]Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales 2751, Australia

[5]Montgomery Botanical Center, Coral Gables, Florida 33156, USA

[6]New York Botanical Garden, Bronx, New York 10458, USA

*Authors for correspondence: Shouzhou Zhang. E-mail: shouzhouz@szbg.ac.cn; Chen Feng. E-mail: fengc@lsbg.cn

Received 5 January 2025; Accepted 23 October 2025

Abstract *Encephalartos*, an African endemic genus within the Zamiaceae, comprises 65 extant species whose phylogenetic relationships have remained unresolved due to limited genetic differentiation observed in previous studies. This research reconstructs the evolutionary history of *Encephalartos* utilizing 3545 single-copy nuclear genes derived from transcriptomes of 64 species. The study estimates divergence times and reconstructs ancestral states for 12 key morphological traits. Phylogenetic analyses definitively resolve eight major clades, supported by both molecular and morphological evidence. Although these clades partially align with previous morphology and geography based classifications, the genomic data provide novel insights, necessitating a revised infrageneric system. Biogeographic reconstructions indicate that *Encephalartos* originated in southern Africa during the Oligocene (~26.3 Ma), subsequently dispersing into eastern and northern Africa through the Zimbabwe–Mozambique corridor during the Miocene, followed by expansion into Central Africa. Speciation rates decreased markedly during the Pliocene and Pleistocene, potentially due to intensified climatic drying and cooling. Morphological character mapping identified ancestral traits, including aerial stems, green leaves, and red sarcotesta. Specific transitions such as subterranean stems in clade IV and bluish-green leaves in clades II and V, further substantiate clade differentiation. These findings resolve long-standing taxonomic uncertainties and emphasize the Oligocene-Miocene as a crucial period for *Encephalartos* diversification, influenced by Cenozoic climate change. This research establishes a robust framework for future systematic and conservation studies while demonstrating the effectiveness of transcriptome data in resolving phylogenies of slowly evolving lineages.

Key words: biogeography, cycads, *Encephalartos*, phylogenetics, transcriptome data, Zamiaceae.

1 Introduction

Cycads represent one of the most ancient extant seed plant lineages, having persisted through major climatic fluctuations for over 330 million years. Modern cycad genera underwent their principal radiations during the Neogene (beginning ~23 Ma; Coiro et al., 2023). Although broadly distributed across tropical and subtropical regions of the Americas, Africa, Asia, and Australia (Donaldson et al., 2003), more than 70% of cycad species are currently classified as threatened (IUCN, 2025). The primary threats include habitat loss, anthropogenic activities, climate change, and illegal collection for horticultural trade (Donaldson et al., 2003).

Encephalartos is the largest cycad genus in Africa, comprising 65 extant species that span tropical and subtropical regions of central and eastern Africa (Calonje et al., 2025). The genus has the highest diversity in southern Africa, where over 70% of species occur. *Encephalartos* inhabits diverse environments, including forests, grasslands, woodlands, savannahs, shrublands, granite outcrops, quartzite hills, cliffs, and gorges. However, most species inhabit woodland–savannah ecosystems (Mabunda, 2007). Most species display disjunct distributions across southern,

central, and eastern Africa. *Encephalartos barteri* Carruth. ex Miq. remains the sole member of the genus found in West African coastal regions.

Encephalartos is also among the most threatened cycad genera, with over 35 species facing threats from habitat destruction, land conversion, and illegal harvesting (Bezuidenhout, 2020). Scarcity and high horticultural demand have driven extensive overexploitation, indicating the importance of integrated conservation strategies that combine in situ and ex situ measures, including the establishment of conservation populations in botanical gardens (Donaldson et al., 2003; Griffith et al., 2020). Effective conservation requires an understanding of the population genetic structure including diversity, heterozygosity, and differentiation, to maintain wild genetic variation and inform management priorities (Clugston et al., 2022). However, classification within cycads presents inherent challenges due to high morphological variability and frequent hybridization, which obscure species boundaries (Keppel, 2009; Tao et al., 2021). Resolving these taxonomic ambiguities and defining biologically meaningful conservation units require a landscape level genomics approach targeting specific species groups (Balkenhol et al., 2019). This approach integrates detailed phylogenetic reconstruction with targeted genomic sequencing that, in turn, enhances the precision of sampling and informs adaptive management (Walters & Osborne, 2004; Lemmon & Lemmon, 2013; McCormack et al., 2013).

Substantial research has aimed to reconstruct phylogenetic relationships and elucidate the biogeographic and temporal distribution patterns across major cycad genera. However, traditional genetic markers have often lacked the resolving power necessary to clarify evolutionary relationships within and among cycad genera, primarily due to the low sequence divergence inherent to these markers (Treutlein et al., 2005; Clugston et al., 2016; Gutierrez-Ortega et al., 2018; Liu et al., 2018; Calonje et al., 2019; Medina-Villarreal et al., 2019; Mankga et al., 2020). Even extensive data sets of plastid protein-coding genes have shown limited utility in resolving relationships among closely related species, as exemplified by studies within *Macrozamia* (Habib et al., 2022). In contrast, transcriptome-derived single-copy nuclear genes (SCGs), which evolve at comparatively higher rates, have demonstrated strong efficacy in resolving infrageneric relationships within several cycad genera (Habib et al., 2022, 2023; Lindstrom et al., 2024; Liu et al., 2024). These findings demonstrate the critical value of large-scale genomic data sets in providing unprecedented resolution for phylogenetic studies within cycad genera.

Understanding the evolutionary relationships and morphological diversification within the genus *Encephalartos* has advanced significantly through extensive morphological observations and molecular data analyses (Coetzee, 1993; Osborne et al., 1993; Vorster, 1993, 2004; van der Bank et al., 2001; Grobbelaar, 2004; Vorster et al., 2004; Treutlein et al., 2005; Konings, 2006; Chaiprasongsuk et al., 2007; Mabunda, 2007; Rousseau, 2012; Mankga et al., 2020; Stewart et al., 2023). Species classification within the genus has been established using integrated molecular, morphological, and spatial distribution data. Vorster (2004) developed a significant infrageneric classification system, organizing *Encephalartos* into 18 groups based on morphological features and geographic relationships. Rousseau (2012) enhanced this classification by incorporating additional recognized species and utilizing various DNA barcoding regions, though genetic variation at the species level remained low. Mankga et al. (2020) analyzed three plastid markers (*rbcLa, matK, trnH-psbA*) and two nuclear markers (ITS, PHYP) to reconstruct the phylogeny and spatio-temporal diversification of *Encephalartos*. Their study estimated the genus's divergence at approximately 9 Ma, with recent diversification occurring within the last 2.5 Ma. However, their biogeographical analysis was constrained by limited statistical support for numerous nodes (Mankga et al., 2020). More recently, Stewart et al. (2023) proposed dividing the genus into ten major clades using combined plastid (*rbcL, matK-trnK, trnH-psbA, cab*), mitochondrial (*nad1*), and nuclear (26S, AGAMOUS, NEEDLY, nrITS1) markers, achieving improved resolution for several nodes. Nevertheless, substantial gaps remain in our understanding of relationships among closely related species.

This study extends previous research by analyzing transcriptome data from Liu et al. (2022) to investigate phylogenetic relationships within *Encephalartos*, Africa's most species rich cycad lineage. Through examination of this data set, this research aims to elucidate infrageneric relationships within the genus, with major clades supported by morphological character analysis of ancestral state evolution. Furthermore, this phylogenetic framework enables investigation of the historical biogeography and diversification patterns of *Encephalartos* across Africa.

2 Materials and Methods

2.1 Plant sampling and gene data retrieval

Transcriptome data were obtained from Liu et al. (2022), derived from fresh leaf samples of 339 cycad specimens. The data set included 67 samples from the genus *Encephalartos*, collected from living specimens of known wild origin in two botanical gardens: 63 accessions from Nong Nooch Tropical Botanical Garden (NNTBG) in Thailand and four accessions from Fairy Lake Botanical Garden in China (SZG). These collections represent 64 of 65 accepted *Encephalartos* species (excluding 6 subspecies) and two species currently classified as synonyms (*E. kanga* Pócs & Q. Luke = *E. kisambo* Faden & Beentje; *E. flavistrobilus* I. S. Turner & Sclavo = *E. schaijesii* Malaisse, Sclavo & Crosiers) in the World List of Cycads (Calonje et al., 2025). An undescribed species (*Encephalartos* sp. nov.) from Mozambique was also included to determine its phylogenetic position within the genus. For outgroup comparison, two species from related Australian genera were selected: *Macrozamia johnsonii* D. L. Jones & K. D. Hill (SRR18094522) and *Lepidozamia hopei* (W. Hill) Regel (SRR18094519). A detailed list of all accessions, distribution areas, and accession numbers is shown in Table 1. Raw transcriptome reads are accessible through the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under BioProject number PRJNA929092.

Illumina reads were subjected to trimming using Trimmomatic v0.39 (Bolger et al., 2014) to remove adapters, low-quality sequences, short inserts, and duplicate reads,

Table 1 List of taxa examined in the current study, along with their collection No. at NNTBG or SZG, herbarium and Sequence Read Archive (SRA) data accessions, and distribution ranges

No.	Specimen collection No.	Taxon names	Herbarium accession number	SRA accession	Distribution range
1	165	*E. aemulans*	SZG00120108	SRR23279675	South Africa: Natal
2	166	*E. altensteinii*	SZG00120101	SRR23279674	South Africa: Eastern Cape
3	169	*E. aplanatus*	SZG00120146	SRR23279627	South Africa: Swaziland
4	170	*E. arenarius*	SZG00120121	SRR23279614	South Africa: Eastern Cape
5	172	*E. barteri*	W 2012-0010268	SRR23279497	Northern Africa: Nigeria, Benin, Ghana, Togo
6	10 898	*E. brevifoliolatus*	SZG00120148	SRR23279486	South Africa: Transvaal
7	173	*E. bubalinus*	W 2012-0010143	SRR23279475	Northern Africa: Tanzania
8	364	*E. afer*	SZG00120134	SRR23279570	South Africa: Eastern Cape
9	10 916	*E. cerinus*	SZG00120096	SRR23279559	South Africa: Natal, Mozambique
10	16 063	*E. chimanimaniensis*	SZG00120145	SRR23279548	Mozambique-Zimbabwe: Chimanimani Mountain Range
11	7471	*E. concinnus*	SZG00120144	SRR23279673	Mozambique-Zimbabwe: South Central Zimbabwe
12	229	*E. cupidus*	SZG00120118	SRR23279662	South Africa: Transvaal
13	16 148	*E. delucanus*	W 2012-0009990	SRR23279651	Central Africa: Tanzania
14	368	*E. dolomiticus*	SZG00120128	SRR23279608	South Africa: Transvaal
15	370	*E. dyerianus*	SZG00120130	SRR23279597	South Africa: Transvaal, Swaziland
16	371	*E. equatorialis*	SZG00120103	SRR23279586	Northern Africa: Lake Victoria, Uganda
17	179	*E. eugene-maraisii*	SZG00120122	SRR23279543	South Africa: Transvaal
18	11 531	*E. flavistrobilus*	SZG00120097	SRR23279532	Central Africa: Zaire
19	18 755	*E. ferox*	SZG00120141	SRR23279521	South Africa: Natal, Mozambique
20	181	*E. friderici-guilielmi*	SZG00120152	SRR23279638	South Africa: Eastern Cape
21	10 904	*E. gratus*	SZG00120129	SRR23279626	Eastern Africa: Malawi, Mozambique
22	11 535	*E. heenanii*	SZG00120107	SRR23279623	South Africa: Transvaal, Swaziland
23	E23-FLG	*E. hildebrandtii*	W 2012-0010039	SRR23279622	Eastern Africa: Tanzania, Kenya
24	10 907	*E. hirsutus*	SZG00120120	SRR23279621	South Africa: Transvaal
25	184	*E. horridus*	SZG00120115	SRR23279620	South Africa: Eastern Cape
26	185	*E. humilis*	SZG00120150	SRR23279619	South Africa: Transvaal
27	228	*E. inopinus*	SZG00120127	SRR23279618	South Africa: Transvaal
28	208	*E. ituriensis*	SZG00120105	SRR23279617	Northern Africa: Zaire, Uganda
29	11 539	*E. kanga*	SZG00120102	SRR23279616	Eastern Africa: Tanzania
30	186	*E. kisambo*	SZG00120125	SRR23279615	Eastern Africa: Kenya
31	187	*E. laevifolius*	SZG00120151	SRR23279613	South Africa: Transvaal, Swaziland
32	188	*E. lanatus*	SZG00120149	SRR23279612	South Africa: Transvaal
33	192	*E. laurentianus*	SZG00120099	SRR23279505	Central Africa: Angola, Zaire
34	193	*E. lebomboensis*	SZG00120114	SRR23279504	South Africa: Natal, Transvaal, Swaziland
35	194	*E. lehmannii*	SZG00120119	SRR23279503	South Africa: Eastern Cape
36	222	*E. mackenziei*	SZG00120106	SRR23279502	Northern Africa: Didinga Hills (S. Sudan)
37	10 911	*E. macrostrobilus*	SZG00120139	SRR23279501	Northern Africa: Uganda
38	10 606	*E. manikensis*	SZG00120140	SRR23279500	Mozambique-Zimbabwe
39	16 067	*E. marunguensis*	SZG00120135	SRR23279499	Central Africa: Zaire, Uganda
40	E40-FLG	*E. middelburgensis*	SZG00120090	SRR23279498	South Africa: Transvaal, Swaziland
41	199	*E. msinganus*	SZG00120110	SRR23279496	South Africa: Natal
42	10 608	*E. munchii*	SZG00120147	SRR23279495	Mozambique-Zimbabwe: Zembe Mountain

Continued

Table 1 Continued

No.	Specimen collection No.	Taxon names	Herbarium accession number	SRA accession	Distribution range
43	201	*E. natalensis*	SZG00120109	SRR23279494	South Africa: Natal
44	376	*E. ngoyanus*	SZG00120133	SRR23279493	South Africa: Natal, Mozambique
45	377	*E. nubimontanus*	SZG00120123	SRR23279492	South Africa
46	211	*E. paucidentatus*	SZG00120124	SRR23279491	South Africa: Transvaal, Swaziland
47	11 530	*E. poggei*	W 2012-0010036	SRR23279490	Central Africa: Zaire, Angola
48	213	*E. princeps*	SZG00120113	SRR23279489	South Africa: Eastern Cape
49	10 609	*E. pterogonus*	SZG00120143	SRR23279488	Mozambique-Zimbabwe: Mt. Mruwere (Mozambique)
50	10 604	*Encephalartos* sp. nov.	SZG00120142	SRR23279487	Mozambique
51	19 506	*E. schaijesii*	SZG00120137	SRR2279485	Central Africa: Zaire
52	10 918	*E. schmitzii*	W 2012-0010269	SRR23279484	Central Africa: Zaire, Zambia
53	17 068	*E. sclavoi*	SZG00120126	SRR23279483	Eastern Africa: Usambara Mountains (Tanzania)
54	174	*E. senticosus*	SZG00120112	SRR23279482	South Africa: Swaziland
55	216	*E. septentrionalis*	SZG00120136	SRR23279481	Northern Africa: Northern Uganda, Southern Sudan
56	217	*E. tegulaneus*	SZG00120094	SRR23279480	Northern Africa: Mt. Lolokwe (Kenya)
57	10 934	*E. transvenosus*	SZG00120100	SRR23279479	South Africa: Transvaal
58	219	*E. trispinosus*	SZG00120117	SRR23279478	South Africa: Eastern Cape
59	226	*E. turneri*	SZG00120132	SRR23279477	Mozambique: Nampula
60	381	*E. umbeluziensis*	SZG00120138	SRR23279476	South Africa: Swaziland
61	221	*E. villosus*	SZG00120098	SRR23279474	South Africa: Cape, Natal, Transvaal, Swaziland
62	17 069	*E. whitelockii*	SZG00120104	SRR23279579	Northern Africa: Mpanga River Falls (Uganda)
63	224	*E. woodii*	SZG00120111	SRR23279578	South Africa: Natal
64	195	*E. longifolius*	SZG00120116	SRR23279577	South Africa: Eastern Cape
65	E65-FLG	*E. cycadifolius*	SZG00120091	SRR23279576	South Africa: Eastern Cape
66	E66-FLG	*E. ghellinckii*	SZG00120092	SRR23279575	South Africa: Cape, Natal
67	E67-FLG	*E. latifrons*	SZG00120093	SRR23279574	South Africa: Eastern Cape
68	254	*M. johnsonii*	BK 084009	SRR18094522	Australia
69	350	*L. hopei*	BK 084013	SRR18094519	Australia

BK, W, and SZG are herbarium codes for Bangkok Herbarium (Department of Agriculture, Bangkok, Thailand), Naturhistorisches Museum Wien (Austria), and Fairy Lake Botanical Garden (Shenzhen, China), respectively.

following default parameters (LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:36). The read statistics of transcriptome data are presented in Table S1. The cleaned reads were utilized to retrieve SCGs. Initially, de novo assembly was conducted using the Trinity v2.15.1 pipeline with default settings (Grabherr et al., 2011). The longest assembled transcripts were annotated using TransDecoder v5.7.1 (https://github.com/TransDecoder), and single-copy ortholog detection of annotated transcripts was performed using OrthoFinder v2.4.0 (Emms & Kelly, 2019). SCGs were selected for phylogenetic reconstruction using KinFin v1.1.1 (Laetsch & Blaxter, 2017). After removing paralogs and excluding genes with species coverage below 70% using the "rmdup" command in SeqKit v2.5.0 (Shen et al., 2016), 3545 retrieved genes were aligned using MAFFT v5.0 (Katoh & Standley, 2013) via TranslatorX v14.0 (Abascal et al., 2010). Ambiguous portions in the alignment were eliminated with Gblocks v0.91b (Talavera & Castresana, 2007) using default settings. Alignments were subjected to manual inspection in Geneious v.8.0.2 (https://www.geneious.com) before subsequent phylogenetic analyses.

2.2 Phylogenetic reconstruction

To reconstruct the evolutionary history of *Encephalartos*, species trees were inferred using two approaches: (1) a maximum likelihood (ML) approach based on a concatenated data set of all genes (hereafter ML-CONCAT) and (2) a multispecies coalescent approach (MSC-ASTRAL) using Weighted ASTRAL (Zhang & Mirarab, 2022), which infers a species tree from gene trees. Two data sets were analyzed: one including all codon positions (CD-NT) and another excluding the fast-evolving third codon position (CD-NT12).

For the ML-CONCAT approach, individual gene alignments were concatenated into a supermatrix using the Perl script FASconCAT-G_v1.04.pl (Kück & Longo, 2014). Partitioning schemes were selected based on codon positions (the 1st, 2nd, and 3rd codon positions), and substitution models for each partition were determined in Mode lFinder

(Kalyaanamoorthy et al., 2017), specifically GTR + F + R2 for codon positions 1 and 2 and GTR + F + R10 for codon position 3. The ML tree was reconstructed in IQ-TREE v2.3.6 (Minh et al., 2020), implementing ultra-fast bootstrap analysis of 1000 replicates. Branches with poor support (<85%) were collapsed using TreeCollapserCL4 (http://emmahodcroft.com/TreeCollapseCL.html) to minimize incongruence among trees from different data sets.

For the MSC-ASTRAL framework, ML gene trees for each ortholog were reconstructed using IQ-TREE v2.3.6 (Minh et al., 2020) with identical parameter settings. The optimal ML gene tree and 100 associated bootstrap replicate trees were used to infer the species tree topology and node support, including local posterior probabilities (LPPs) and quartet support values (q), using the hybrid weighting method (wASTRAL-h) implemented in Weighted ASTRAL (Zhang & Mirarab, 2022). To examine site concordance, the "-scfl" option was used in IQ-TREE 2 (Minh et al., 2020; Mo et al., 2022), using the final tree as input.

To evaluate reticulation events such as hybridization and recombination within *Encephalartos*, two approaches were implemented. Single-nucleotide polymorphism site (SNP) data from the complete data set were used to generate a Neighbor-Net network with uncorrelated p-distances in SplitsTree v.6.1.10 (Huson & Bryant, 2006). Hybridization events among the most recent common ancestors (MRCAs) were inferred using PhyloNet v2.4 (Than et al., 2008) via the "InferNetworks ML" command. For computational efficiency, the analysis was restricted to 45 species and 1999 gene trees, focusing on samples representing major clades, broad geographic ranges, signals of introgression, and evidence of natural hybridization from published literature or personal observations. The selection included only genes with high species coverage and sufficient sequence variability. *Encephalartos tegulaneus* Melville, despite having the highest proportion of missing data (537 of 3545 possible genes), was retained in the PhyloNet analyses due to conflicting phylogenetic placements across data sets and methods.

2.3 Divergence time and ancestral area reconstruction

Based on computational methods from previous studies (Habib et al., 2022, 2023; Lindstrom et al., 2024), the divergence time for the genus *Encephalartos* was analyzed using a subset of 50 genes from the total 3545 SCGs. Gene selection utilized the SortaDate package (Smith et al., 2018), considering clock-likeness, minimal topological conflict, and substantial branch length variations. These 50 genes were concatenated and analyzed in RAxML v8 (Stamatakis, 2014) to generate 1000 trees, using the optimal phylogenetic tree (CD-NT) as a constraint. Secondary calibrations (root min = 51.7, max = 102.1; crown min = 14.20, max = 35.45) derived from Coiro et al. (2023) provided recent total-evidence molecular dating analysis for Cycadales. Divergence times were calculated using penalized likelihood in treePL (Sanderson, 2002; Smith & O'Meara, 2012), utilizing trees from 50-gene ML analysis. Maximum clade credibility (MCC) trees were constructed using TreeAnnotator v2.6.2 (Helfrich et al., 2018), displaying median ages and 95% highest posterior density (HPD 95%) intervals on nodes, with the initial 25% (250/1000) trees removed as burn-in. The chronogram was visualized in FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

The TreePL chronogram served as input for historical biogeographic analysis, with species distributions based on the World List of Cycads (Calonje et al., 2025). Following Stewart et al. (2023), five areas were defined: (A) South Africa, (B) Mozambique-Zimbabwe (MZB), (C) Central Africa, (D) Eastern Africa, and (E) Northern Africa. These regions represent the current distribution of the genus in Africa and facilitate detection of major historical biogeographical patterns. The R package "BioGeoBEARS" (Matzke, 2018) was used to conduct likelihood ancestral range estimation in RASP v4.2 (Yu et al., 2020) using three models: (i) DEC (dispersal extinction cladogenesis (Ree & Smith, 2008)), (ii) DIVALIKE (likelihood-based dispersal vicariance analysis, Ronquist (1997)), and (iii) BAYAREALIKE (likelihood implementation of BayArea (Landis et al., 2013)). Models with and without founder-event speciation ('j') were evaluated in the package.

2.4 Ancestral character state reconstructions

Twelve morphological characters were analyzed, including growth habit, stem, leaf, leaflet, strobili, and seed features, using data from systematic and floristic literature (Vorster, 1993; van der Bank et al., 2001; Vorster, 2004; Vorster et al., 2004; Treutlein et al., 2005; Konings, 2006; Mabunda, 2007) and specimen observations as detailed in Table 2. All traits were reconstructed using the coalescent CD-NT data set's species tree. Character evolution analysis was performed in Mesquite (Maddison & Maddison, 2023) using likelihood methods under the Markov *k*-state one-parameter (Mk1) model, assuming symmetrical rates of character gain and loss.

3 Results

3.1 Data statistics

Transcriptome sequencing of 69 samples produced between 20 463 947 and 27 714 413 reads, averaging 22 984 288 reads per sample (Table S1). The analysis identified 3545 single-copy orthologous groups for phylogenetic analyses. The CD-NT data set showed a mean alignment length of 3210 bp per gene, with the concatenated supermatrix totaling 4 198 320 bp, containing 296 062 variable and 89 727 parsimony-informative sites. The CD-NT12 data set showed a mean gene alignment of 2462 bp, with a concatenated matrix of 2 799 076 bp, comprising 179 741 variable and 41 178 parsimony-informative sites. All in-group species contained over 70% of genes, except for *E. tegulaneus* (15%) and *E. longifolius* (Jacq.) Lehm. (63%). The remaining in-group species averaged 15% gap or ambiguous sites, ranging from 7.1% in *E. kisambo* to 29% in *E. senticosus* Vorster.

3.2 Phylogenetic analyses

Phylogenetic analyses for *Encephalartos* were conducted using two methodological approaches (Figs. 1, S1). The CD-NT data set yielded better supported topologies compared to the CD-NT12 data set. For interpretation purposes, topological support is categorized as strong or full support when local posterior probability (LPP) = 1.0 or maximum likelihood

Table 2 Twelve morphological characters optimized on the phylogenetic tree, number of species where the character state appeared, and the well-supported (>80%) ancestral states for the genus *Encephalartos* and its major clades

				Ancestral state	
No.	Characters	States	No. of species	Genus	Clades
1	Stem	0) subterranean	8		IV
		1) aerial trunk	60	✓	I–III, V–VIII
2	Basal petiole collar	0) absent	57	✓	I–VIII
		1) present	10		
3	Leaf pinnacanth	0) absent	54	✓	I–VIII
		1) reduced	8		
		2) present	5		
4	Pinnae color	0) green	52	✓	I, III, IV, VI, VII
		1) bluish green	13		II, V
		2) yellowish-green to grey	2		
5	Pinnae margin	0) toothed	40	*	VII
		1) entire	27	*	I, II
6	Glaucous leaves	0) absent	55	✓	I–VIII
		1) present	12		
7	Leaf texture	0) coriaceous	39	*	III, V
		1) papyraceous	28	*	I, IV, VII, VIII
8	Strobilus emergence	0) single	59	✓	I–VIII
		1) multiple	8		
9	Strobilus pubescence	0) absent	57	✓	II–VIII
		1) present	8		I
10	Megasporangiate strobilus color	0) green	21		V, VII
		1) yellowish-green to yellow	33	✓	I, III, IV, VIII
		2) bluish-green	5		
		3) olive	2		
		4) orange	3		
		5) red	1		
11	Microsporangiate strobilus color	0) green	14		V
		1) yellowish-green to yellow	37		I, VI, VIII
		2) bluish-green	8		VII
		3) yellowish-green to yellow	5		
		4) orange	2		
		5) red	1		
12	Sarcotesta color	0) red	37	✓	III, IV, VI–VIII
		1) orange	15		II
		2) yellow	11		I
		3) amber brown	3		

*indicate the equivocal states.

bootstrap (MLBS) = 100%; high support when $1.0 > \text{LPP} \geq 0.95$ or $100\% > \text{MLBS} \geq 95\%$; moderate support when $0.95 > \text{LPP} \geq 0.85$ or $95\% > \text{MLBS} \geq 90\%$; and weak or low support when $\text{LPP} < 0.85$ or $\text{MLBS} < 90\%$. The analyses revealed eight major clades (I–VIII) within the genus, with strongly supported and largely congruent relationships across data sets. Primary discrepancies involved the phylogenetic placement of *E. tegulaneus* and certain intra-clade relationships within clades III, V, and VIII (Figs. 1, S1). Clade I, representing the earliest diverging lineage, contains seven species with consistently strong support for their relationships (Figs. 1, S1). A newly identified, well-supported monophyletic group (clade II) comprising *E. inopinus* R. A. Dyer and *E. hirsutus* P. J. H. Hurter appears sister to clade I. Clade III, the most species-rich group, contains 17 species with moderately supported relationships in the CD-NT12 data set and mostly well-resolved relationships in the CD-NT data set. Within this clade, *E. woodii* Sander represents the earliest diverging taxon (subclade III-A), sister to two subclades containing seven (III-B) and nine species (III-C), respectively. In subclade III-C, the concatenated CD-NT data set indicates *E. transvenosus* Stapf & Burtt Davy as sister to *E. heenanii* R. A. Dyer and *E. paucidentatus* Stapf & Burtt Davy (MLBS = 93). Clade IV comprises six species with generally congruent relationships, except for *E. aplanatus* Vorster, whose phylogenetic position showed weak support in comparative analyses. Clade V also includes six species but demonstrates substantial conflict at some internal nodes. The MSC-ASTRAL analysis recovered a strongly supported (LPP = 1.0) monophyletic group of *E. eugene-maraisii* I. Verd. and *E. dolomiticus*

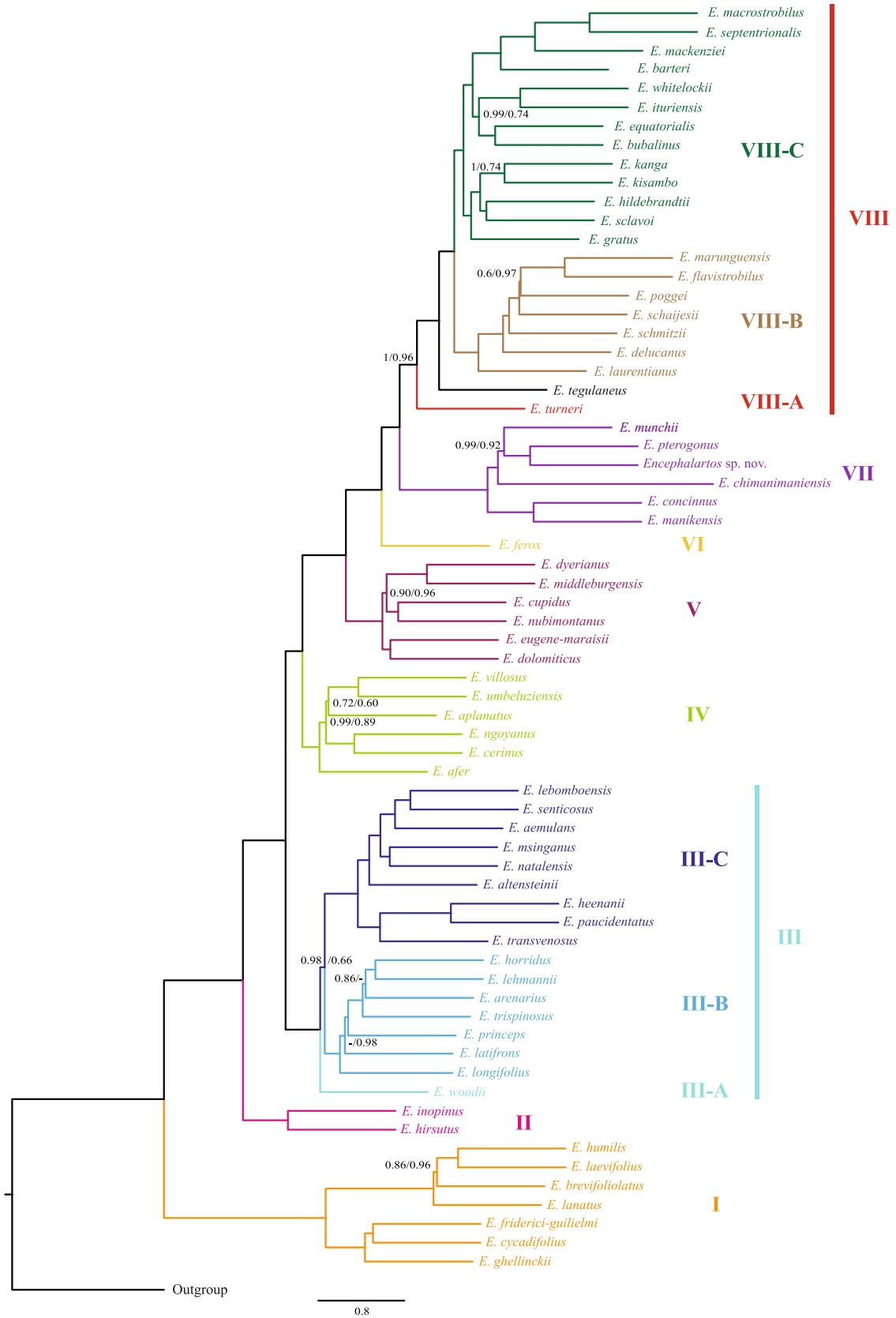

Fig. 1. *Encephalartos*' Phylogeny shows all major clades based on coalescent analyses of 3545 single-copy nuclear genes. All the branches are maximally supported (LPP = 1) for both "all codon positions (NT) data set" and "1st and 2nd codon position only (NT12) data set" unless otherwise noted along the branches if LPP < 1 in either case. "-" indicate LPP < 80%.

Lavranos & D. L. Goode as sister to the remaining species in the clade (Fig. 1). Conversely, the ML-CONCAT analysis (CD-NT data set) positioned these two species as sisters to *E. cupidus* R. A. Dyer and *E. nubimontanus*, with MLBS = 99% (Fig. S1). However, neither relationship received robust support (MLBS < 85%) in the CD-NT12 data set (Fig. S1), indicating conflicting signals between analytical approaches. *Encephalartos ferox* G. Bertol. constitutes a monospecific clade (clade VI) with maximum support across all analyses. Clade VII encompasses six species with generally consistent phylogenetic support and relationships (Figs. 1, S1). Clade VIII, containing 22 species, shows *E. turneri* Lavranos & D. L. Goode (subclade VIII-A) and *E. tegulaneus* as successive sister taxa to two major monophyletic groups (subclades VIII-B and VIII-C). Subclade VIII-B includes seven species with either *E. poggei* (MSC-ASTRAL CD-NT12 data set, MLBS = 99.7%) or *E. schaijesii* (ML-CONCAT CD-NT12 data set, MLBS = 90%) as sister to *E. marunguensis* Devred and *E. flavistrobilus*. Neither relationship received significant support in the CD-NT data set under either analytical method (Figs. 1, S1). Subclade VIII-C contains 13 or 14 species, with varying node support across methods; while ML-CONCAT analyses lacked robust support for several relationships, MSC-ASTRAL trees provided well-supported placements in the CD-NT data set. The main inconsistency concerns *E. tegulaneus* placement: ML-CONCAT analyses robustly position this species within subclade VIII-C (MLBS = 100; Fig. S1), whereas MSC-ASTRAL analyses place it either as sister to *E. turneri* (adjacent to combined subclades VIII-B and VIII-A) in the CD-NT data set or within clade VI (sister to the monophyletic clade VII–VIII) in the CD-NT12 data set, both with full statistical support (LPP = 1.0; Fig. 1). These conflicting placements highlight methodological differences in resolving the species' phylogenetic position.

Quartet support values for the primary topology (q1) demonstrated high values (> 50%) at several deep nodes, including clades I, II, III, V, VI, and VII, compared to alternative topologies (q2 and q3). However, q1 support for clades IV and VIII showed moderate values (35%–50%) (Fig. S2). The phylogenetic position of *E. tegulaneus* received support from 43% of the analyzed gene trees. Sister relationships within major clades showed clear differentiation through significant quartet support for q1 relative to q2 and q3, indicating robust phylogenetic relationships in the CD-NT data set of MSC-ASTRAL, with conflicting nodes showing q1 values of 35%–40%. The MSC-ASTRAL CD-NT tree analysis was supplemented with site concordance factor (sCF) values to further validate phylogenetic relationships. The sCF values demonstrated enhanced resolution (sCF > 35%) for major clades and shallower relationships (Fig. S2).

The Neighbor-Net network graph revealed distinct clustering of species within their major clades, with certain closely related groups corresponding to nodes with low bootstrap support. *Encephalartos tegulaneus* showed deep edges with the group comprising clades VIII-B and VIII-C (aligning with MSC-ASTRAL-based trees), while also displaying shallow edges with species of VIII-C (consistent with ML-CONCAT analyses). Additionally, *E. woodii* demonstrated closer relationships to species in clade III-C (Fig. S3) and showed a weak association as sister to clades III-B and III-C in the CD-NT12 MSC-ASTRAL analysis (Fig. 1). The PhyloNet network analyses indicated that *E. tegulaneus* shares edges with the most recent common ancestors (MRCA) of clade III B + C and clade VIII-C, or clade IV and clade VIII-C, suggesting hybridization among ancestral lineages of these groups (Fig. S3).

3.3 Divergence time estimation and spatial distribution

The crown age of *Encephalartos* was estimated at 26.34 Ma (95% HPD: 25.54–26.84), corresponding to the Oligocene epoch (Fig. 2). The subsequent divergence of clades II, III, and IV from other major clades occurred in the early Miocene at 19.57, 17.15, and 16.12 Ma, respectively. Divergences within clades III, IV, and V emerged during the middle Miocene (approximately 13–14 Ma), while divergence within clade VI was estimated at 11.69 Ma (Fig. 2). The majority of cladogenetic events within the genus occurred during the late Miocene from 11 to 5.3 Ma, including major divergence events within clades VII and VIII. Subsequently, species divergence decreased notably, with limited speciation events within clade VIII occurring during the Pliocene and Pleistocene epochs (~5–2 Ma) (Fig. 2).

Spatial patterns of divergence were analyzed across five areas of endemism using BioGeoBEARS analyses. The DIVALIKE + J model emerged as the best-fit biogeographical model for the genus. South Africa was identified as the ancestral region for *Encephalartos* (Fig. 2), with the highest probability (1.0). The species in clades I–VI diversified within this region, primarily during the mid and late Miocene. Diversification toward eastern and northern Africa (clade VIII, excluding subclade VIII-A) proceeded via the Zimbabwe–Mozambique region (clade VII), with subsequent expansion into central Africa (subclade VIII-B) during the late Miocene.

3.4 Reconstruction of ancestral states of *Encephalartos* morphological characters

Using the resolved *Encephalartos* phylogeny with branch lengths, we reconstructed ancestral states for 12 morphological characters (Table 2; Fig. 3). Four characters, i.e., leaf margin type and the coloration of the megasporangiate strobilus, microsporangiate strobilus, and sarcotesta, demonstrated no transitions along the phylogenetic backbone (Table 2; Fig. 3). Aerial stems likely represented the ancestral state in early *Encephalartos*, while subterranean stems evolved independently in the crown ancestor of clade IV and separately within clades I and V. The absence of a basal petiole collar characterized all major clades, except for clades II, III-B, and V, where this feature's presence remained equivocal (i.e., overlapping probabilities for 'present' and 'absent') (Table 2; Fig. 3). Reduced basal pinnae (pinnacanths) were ancestrally absent but evolved five times as complete reductions and eight times as partial reductions, predominantly among South African species. The ancestral leaf color of *Encephalartos* was green, with bluish-green leaves evolving independently in clades II and V, and in clades I, III-B, IV (twice in each), VII, and VIII-B (three times in each) (Table 2; Fig. 3). Pinna margins were ambiguously toothed or entire in the ancestor, with entire margins retained in clades I and II and toothed margins arising in clade VIII. Patterns for leaf margin evolution showed complexity across other clades. Glaucous leaves were likely ancestral in *Encephalartos* and have generally persisted throughout the genus.

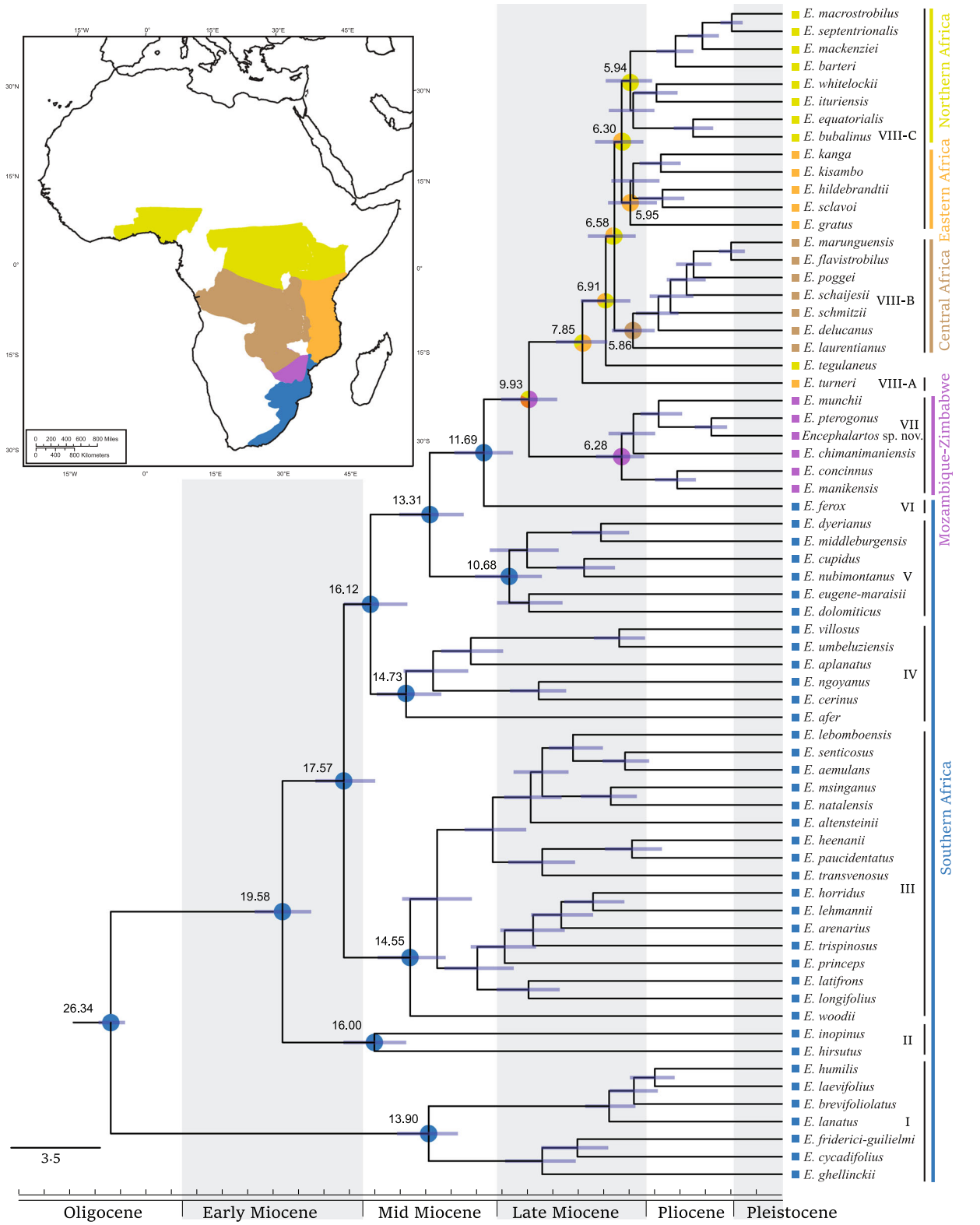

Fig. 2. Ancestral area reconstruction inferred in BioGeoBEARS by DIVALIKE analyses on the chronograms of *Encephalartos*. 50-gene ML tree with topologies constrained to the NT tree was used with three secondary calibration points derived from Coiro et al. (2023). Numbers at selected nodes indicate the median divergence time estimates of major clades and blue bars show the 95% HPD intervals of each node age. Pie charts at nodes represent the most likely ancestral areas reconstructed in BioGeoBEARS by DIVALIKE analyses. The map shows the geographical distribution of major clades in Africa.

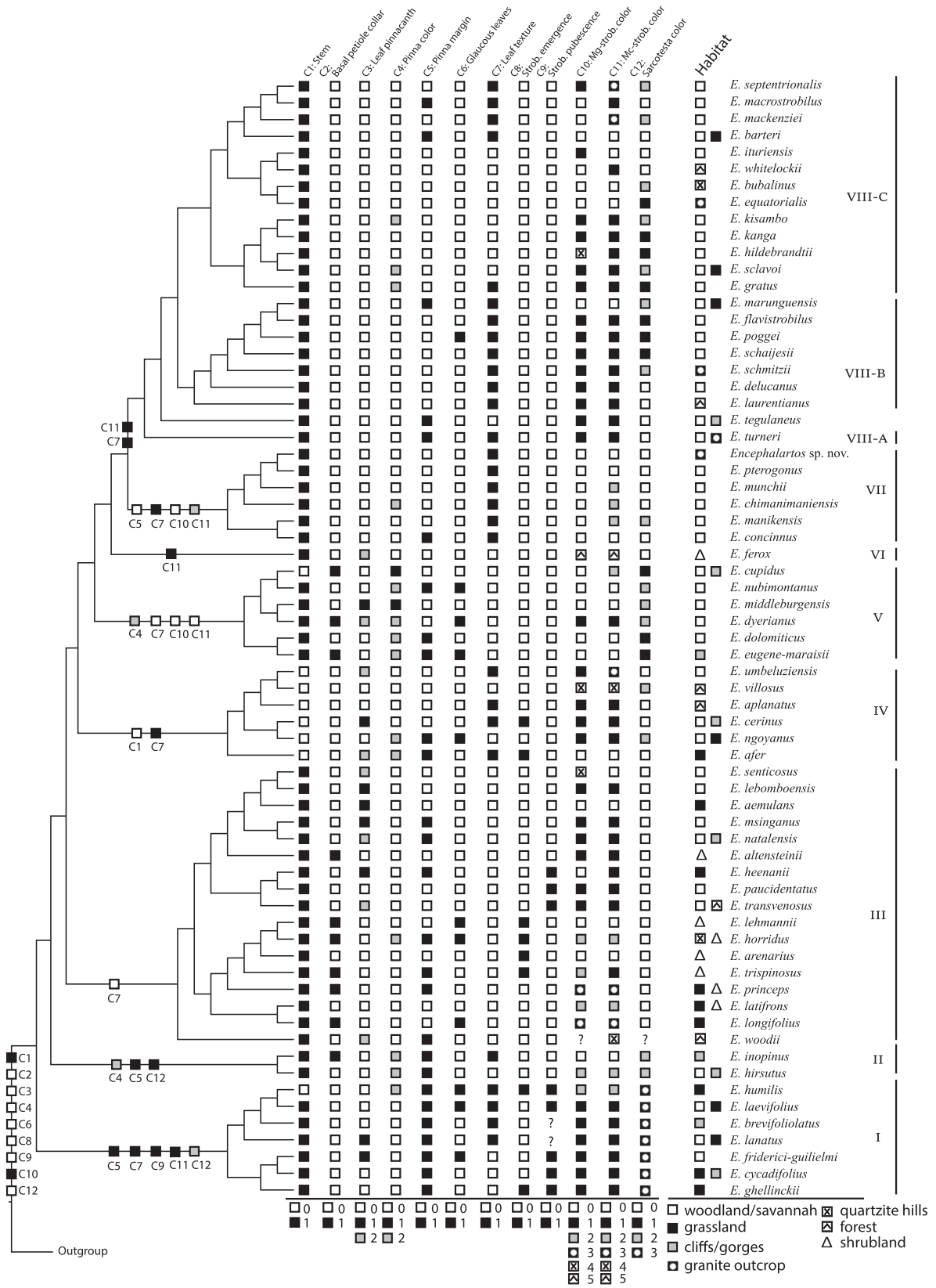

Fig. 3. Distribution of twelve morphological characters of *Encephalartos* (C1–C12) along with the habitat, optimized onto the maximum likelihood tree in coalescent analyses (NT data set). "?" indicate the missing data for character states. Ancestral states of the genus and major clades are indicated on the nodes (probability > 80% only). Mc-strob, mircosporangiate strobilus; Mg-strob, megasporangiate strobilus; Strob, Strobilus.

The absence of glaucous leaves evolved three times (each) within clades I, III-B, and V, and once (each) in clades IV and VIII-B. Leaf texture, though ancestrally ambiguous (coriaceous vs. papyraceous), distinguishes clades: coriaceous leaves predominate in clades III, V, and VI, while papyraceous leaves characterize clades IV and VII, with additional subgroup differentiation in clades I and VIII (Table 2; Fig. 3). A single strobilus per growing point was ancestral for *Encephalartos* and all major clades. The "multiple" strobili state evolved mainly in clade III-B and twice independently in each of clades I and IV. Strobilus pubescence was absent ancestrally and throughout the genus, except at the crown of clade I and in a subgroup of clade III-C. Ancestral megasporangiate strobili were yellowish-green to yellow, with green coloration appearing in clades V, VII, and a subgroup of VIII. Rare colors, such as red-salmon in *E. ferox*, evolved sporadically. Microsporangiate strobili maintained ancestral yellowish-green/yellow coloration in clades I, III-C, IV, and VIII. The ancestral seed sarcotesta was red, with transitions to yellow (clade I) and orange (clades II and V), while clade VIII displayed substantial color variability (Table 2; Fig. 3).

4 Discussion

4.1 A well-resolved *Encephalartos* phylogeny

Previous phylogenetic studies of *Encephalartos* resolved two major lineages subdivided into geographically and morphologically congruent clades (Mankga et al., 2020; Stewart et al., 2023), and yet, significant challenges persisted in reconciling incongruent clade placements and achieving robust nodal support across the genus. This study advances understanding by establishing a well-supported phylogenetic framework for *Encephalartos*, utilizing transcriptome-derived SCGs and incorporating both concatenation (ML-CONCAT) and coalescent (MSC-ASTRAL) approaches. This methodology achieves comprehensive resolution of species relationships, addressing long-standing ambiguities and identifying two deeply supported major lineages: the Cycadifolius Lineage (Clade I), restricted to southern Africa, and the Inopinus Lineage (Clades II–VIII), which spans southern Africa and extends into central, eastern, and northern Africa (Fig. 1).

For MSC-ASTRAL trees, the analysis revealed higher quartet support and site concordance factors for the main topology, an uncommon finding compared to other recent studies on cycad genera such as *Ceratozamia*, *Macrozamia*, and *Zamia* (Habib et al., 2022, 2023; Lindstrom et al., 2024). In *Encephalartos*, data set incongruence may arise from statistical limitations such as data insufficiency or biological phenomena including hybridization, incomplete lineage sorting (ILS), or minimal sequence divergence from rapid radiation. Notably, the phylogenetic position of *E. tegulaneus* in either data set is supported by only 15% of the total investigated genes (537 out of 3545), potentially explaining its variable placement across trees. The Neighbor-Net Algorithm and PhyloNet indicated that this species shares edges at both deep and shallow levels with species showing close association in alternative methods (Fig. S3). This pattern suggests that both recent and ancient hybridization may have contributed to topological uncertainty for *E. tegulaneus*, warranting additional investigation. Hybridization has significantly influenced phylogenetic relationships in other cycad genera, introducing complexity to their evolutionary histories (Habib et al., 2022, 2023; Lindstrom et al., 2024; Liu et al., 2024). Limited genetic diversity, frequent conflicts between molecular and morphological data, and ILS create challenges in establishing robust phylogenies. ILS particularly affects phylogenetic placement during rapid successive divergence events over time, the likelihood of alternative alleles occurring in different lineages increases, producing gene trees with conflicting signals (Whitfield & Lockhart, 2007; Oliver, 2013). The CD-NT tree from the MSC-ASTRAL method provided well-resolved positions for most nodes, aligning with geographical distribution patterns. This tree therefore serves as the basis for divergence time estimation, character evolution analyses, and subsequent discussion.

4.2 Origin and diversification of *Encephalartos*

This study uses three divergence time estimates from Coiro et al. (2023) as calibration points. In contrast to previous hypotheses that suggested a late Miocene origin (~9 Ma) for *Encephalartos* (Mankga et al., 2020), these results indicate that the genus originated in the Oligocene, approximately 26.34 Ma. This timeline establishes an earlier origin for *Encephalartos* and corresponds more closely with other cycad lineages such as *Ceratozamia* (22.2 Ma) and *Zamia* (26.3 Ma) (Coiro et al., 2023). The analyses demonstrate that major species diversification occurred during the late Miocene (11–5.3 Ma), characterized by significant climatic and ecological changes. These finding challenges earlier assumptions of simultaneous origin and diversification, revealing a temporal separation between *Encephalartos* origin and subsequent radiation.

The Miocene (~23–5 Ma) represented a period of substantial geological and climatological change in Africa, likely influencing *Encephalartos* speciation. While the early Miocene was warmer and more humid than the Oligocene, substantial global cooling in the middle Miocene attributed to the closure of the Tethys Sea, resulted in increased aridification (Zachos et al., 2001; Plana, 2004). This climate shift promoted grassland expansion and forest contraction (Jacobs, 2004), potentially limiting *E. woodii*'s distribution range. These environmental changes transformed Africa into a mosaic of isolated refugial forests separated by savannah (Jacobs, 2004; Plana, 2004), facilitating dispersal and diversification in *Encephalartos*. These findings differ from patterns observed in African rainforest trees, which demonstrate reduced diversification during the middle Miocene. This trend is not surprising because most extant *Encephalartos* species do prefer arid or open grassland habitats.

Consistent with previous research (Mankga et al., 2020; Stewart et al., 2023), our results strongly support a southern African origin for *Encephalartos*. During this period, lowland rainforests in regions of South Africa were gradually replaced by savannah woodland, while the rainforest belt shifted southward, potentially extending from the western to the eastern African coast (Coetzee, 1993). These environmental changes likely facilitated early diversification within the

genus. Initial range expansion remained confined to southern Africa during the early to middle Miocene (Fig. 3), with gradual northward expansion reaching central Africa only by the middle-late Miocene. Geological uplift and volcanism began in the late Oligocene–early Miocene (Wichura et al., 2010). Africa's current topography was largely established by the late Miocene, characterized by alternating moist and dry periods that led to repeated forest expansions and contractions (Morley, 2001; Plana, 2004). Evidence suggests rapid early divergences within the genus from the fragmentation of ancestral populations during the early to middle Miocene. This hypothesis is supported by paleo environmental data demonstrating the dominance of tropical and subtropical forests in early Miocene Africa (Zachos et al., 2008). Although aridification reduced rainforest flora diversity (Maley, 1996), it appears to have enhanced diversification in *Encephalartos*. Similar patterns are observed in *Acridocarpus* (Malpighiaceae), a lineage that adapted to drought-prone habitats during the middle Tertiary. The Plio-Pleistocene represented a critical period for species diversification in both rainforests and savannahs (Plana, 2004), with numerous contemporary African species lineages traceable to the late Miocene. Increasing humidity in the early Pliocene (5–3.5 Ma) prompted rainforest expansion and savannah contraction (Plana, 2004), likely isolating arid-adapted *Encephalartos* in higher and drier mountainous regions. Our findings indicate that *Encephalartos* diversification reached its peak before the end of the Miocene and in the early Pliocene, followed by a transition toward significantly drier climate around 3.48 Ma (late Pliocene), with drying and cooling cycles (Coetzee, 1993), potentially influencing the genus's current diversity.

4.3 Systematics considerations for infrageneric classification of the genus *Encephalartos*

This investigation offers novel perspectives on *Encephalartos* phylogeny. The results largely congruent with recent phylogenetic frameworks (Stewart et al., 2023), with exceptions regarding the placement of clade II and sister group relationships within clade VIII. Furthermore, several clades received more robust nodal support in our analysis. The phylogenetic tree generally aligns with previous taxonomic proposals (Vorster, 2004; Treutlein et al., 2005; Stewart et al., 2023), though some discrepancies appear with the expanded molecular sampling. Fig. 4 presents the revised infrageneric classification scheme, integrating molecular, morphological, and biogeographic evidence, and illustrates the diversity of vegetative and reproductive characters across all major clades. *Encephalartos* comprises two major lineages: the Cycadifolius Lineage and the Inopinus Lineage. The Cycadifolius Lineage (clade I) includes species with pubescent strobili typically found in grasslands. The Inopinus Lineage (clades II–VIII) contains species with minimal or no strobilus pubescence and occurring across diverse habitats, including forests, savannahs, shrublands, granite outcrops, and cliffs, with species being less common in grasslands (Mabunda, 2007).

4.3.1 Cycadifolius Clade (Clade I)

This clade includes *E. brevifoliolatus* Vorster, *E. cycadifolius* (Jacq.) Lehm., *E. friderici-guilielmi* Lehm., *E. ghellinckii* Lem., *E. humilis* I. Verd., *E. laevifolius* Stapf & Burtt Davy, and *E. lanatus* Stapf & Burtt Davy (Fig. 4). This group is consistently supported in molecular and morphological studies (Osborne et al., 1993; Vorster, 2004; Treutlein et al., 2005; Mabunda, 2007; Rousseau, 2012; Yessoufou et al., 2014; Mankga et al., 2020; Stewart et al., 2023). These species are native to eastern South Africa, typically occurring in grasslands at elevations of 700–2000 m and frequently experiencing frost or snow. They have mostly entire, lanceolate pinnae at maturity (~8–10 mm wide) that show a slight reduction toward the base and have a distinct petiole. Their strobili are distinctly pubescent and emerge in spring, with a brief cone disintegration period (about three months) relative to other *Encephalartos* species (six months or longer). These phenological characteristics prevent hybridization with other species outside their clade. The seed sarcotesta is yellow to amber and the sclerotesta is ribbed with prominent venation patterns (Vorster, 2004; Treutlein et al., 2005; Rousseau, 2012).

4.3.2 Inopinus Clade (Clade II)

This clade comprising *E. hirsutus* and *E. inopinus* is newly recognized as a monophyletic group across all data sets and inference methods (Fig. 4). Both species are endemic to Limpopo Province, South Africa; *E. hirsutus* is found in only two populations within quartzite hills (Hurter & Glen, 1996), while *E. inopinus* has a wider distribution from woodland/savannah to cliffs and gorges (Mabunda, 2007). Previously, *E. inopinus* was classified as a distinct group (Vorster, 2004) or positioned as sister to three subterranean species (*E. ngoyanus* I. Verd., *E. cerinus* Lavranos & D. L. Goode, and *E. poggei* Asch.) (Osborne et al., 1993). *Encephalartos hirsutus* is distinguishable from similar species (e.g., *E. eugene-maraisii, E. dolomiticus, E. dyerianus* Lavranos & D. L. Goode, and *E. middelburgensis* Vorster, Robbertse & S. van der Westh) by its decurrent pinnae and glabrous, waxy sporophylls (Hurter & Glen, 1996). Rousseau (2012) discussed *E. inopinus* and *E. hirsutus* as a distinct species group, but noted that *E. inopinus* demonstrates reproductive incompatibility with other species (Vorster, 2004). The cones of *E. hirsutus* share characteristics with those of *E. inopinus*, both displaying a glabrous and glaucous blue-green color. *Encephalartos hirsutus* remains unique among *Encephalartos* species, with its persistent rough indumentum on young leaves (Rousseau, 2012).

4.3.3 Altensteinii Clade (Clade III)

This clade comprises three subclades (Fig. 4), each characterized by specific distribution patterns in South Africa.

Woodii subclade (III-A): Contains only *E. woodii*, known from a single wild male plant discovered in KwaZulu-Natal in 1895 (Prakash et al., 2008). Currently extinct in the wild, it persists only through cultivated vegetative offsets. Historical assumptions suggested close relations or a hybridization event with *E. natalensis* R. A. Dyer & I. Verd. (Viljoen & van Staden, 2006). However, Inter-Specific Sequence Repeats (ISSR's) analysis demonstrated that *E. woodii* and *E. natalensis* are only distantly related (Prakash et al., 2008). This is corroborated by the current study that also confirms them as distinct species. Morphologically, *Encephalartos*

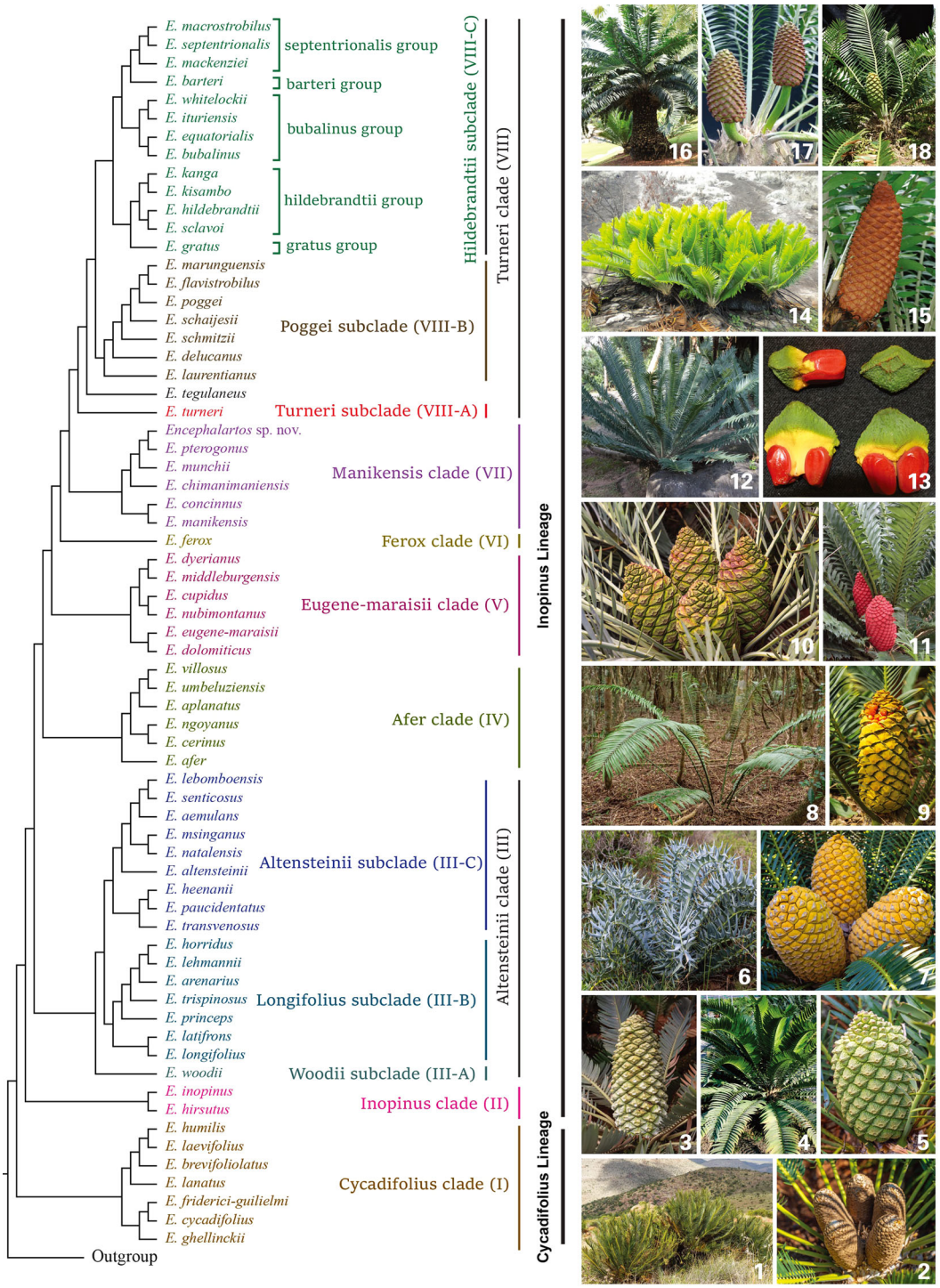

Fig. 4. Phylogenetic relationships and proposed infrageneric classification of *Encephalartos*. Major clades identified in this study are shown on the left, with representative taxa illustrated on the right. Clades and their representatives are as follows: Cycadifolius Clade: (1) *E. cycadifolius* (Jacq.) Lehm. and (2) *E. laevifolius* Stapf & Burtt Davy; Inopinus Clade: (3) *E. hirsutus* P. J. H. Hurter; Altensteinii Clade: (4) *E. woodii* Sander of woodii subclade, (5) *E. trispinosus* (Hook. f.) R. A. Dyer and (6) *E. horridus* (Jacq.) Lehm. of longifolius subclade, (7) *E. lebomboensis* I. Verd. of altensteinii subclade; Afer Clade: (8) *E. villosus* Lem. and (9) *E. cerinus* Lavranos & D. L. Goode; Eugene-maraisii Clade: (10) *E. middelburgensis* Vorster, Robbertse & S. van der Westh; Ferox Clade: (11) *E. ferox* G. Bertol; Manikensis Clade: (12) *E. concinnus* R. A. Dyer & I. Verd. and (13) *E. chimanimaniensis* R. A. Dyer & I. Verd.; Turneri Clade: (14) *E. turneri* Lavranos & D. L. Goode, (15) and (16) *E. gratus* Prain of turneri subclade, (17) *E. barteri* Carruth. ex Miq. of hildebrandtii subclade, (18) *E. delucanus* R. A. Dyer & I. Verd. of poggei subclade. Photographs by A. Lindstrom (4, 11–18) and J. A. R. Clugston (1–3, 5–10).

woodii shows the unique, distinctive characters of 3–4 teeth at the base of the upper margin of the pinna in juvenile plants and lower leaflets in mature plants (Rousseau, 2012).

Longifolius subclade (III-B): Includes *E. arenarius* R. A. Dyer, *E. horridus* (Jacq.) Lehm., *E. latifrons* Lehm., *E. lehmannii* Lehm., *E. longifolius*, *E. princeps* R. A. Dyer, and *E. trispinosus* (Hook. f.) R. A. Dyer (Fig. 4). These species occur in a relatively limited area in the Eastern Cape region of South Africa, typically growing in arid environments (Vorster, 2004). The species have glaucous leaves and/or prominent spiny-lobed pinnae with recurved apices terminating in a spine. Smaller plants commonly produce single strobili that are brown to yellow-green and largely covered with russet indumentum. Additional shared traits include a high proportion of sterile sporophylls, red sarcotesta, and an elevated seed sarcotesta index (Vorster, 2004; Rousseau, 2012).

Altensteinii subclade (III-C): Comprises two groups: (1) *E. transvenosus*, *E. heenanii*, and *E. paucidentatus*, native to the northern Swaziland, Mpumalanga, and Limpopo provinces in South Africa, and (2) *E. aemulans* Vorster, *E. altensteinii* Lehm., *E. lebomboensis* I. Verd., *E. msinganus* Vorster, *E. natalensis* R. A. Dyer & I. Verd., and *E. senticosus*, primarily occurring in the northern KwaZulu-Natal province in South Africa. This subclade is distinguished both morphologically, through its robust stature and bright green leaves in all species and warty terminal bullae on megasporangiate strobili in some species, and with molecular data (Treutlein et al., 2005; Rousseau, 2012; Mankga et al., 2020; Stewart et al., 2023).

4.3.4 Afer Clade (Clade IV)

This clade includes *E. afer* (Thunb.) Lehm., *E. aplanatus, E. cerinus, E. ngoyanus, E. umbeluziensis* R. A. Dyer, and *E. villosus* Lem. (Fig. 4) and shows a wide distribution along eastern South Africa's coast. Their range extends from the Eastern Cape (*E. afer*) north through KwaZulu-Natal (*E. cerinus, E. ngoyanus, E. villosus*) to Swaziland and Mozambique (*E. aplanatus, E. umbeluziensis*). This group, identified here for the first time as a distinct phylogenetic clade, has previously been considered to have close relationships (Treutlein et al., 2005; Rousseau, 2012; Mankga et al., 2020). Morphologically, these species share a non-emergent or slightly emergent subterranean stem, sparse and predominantly erect leaves, and similar strobili with drooping and/or fringed sporophylls and rather indistinct facets (Vorster, 2004).

4.3.5 Eugene-maraisii Clade (Clade V)

This clade comprises *E. cupidus* R. A. Dyer, *E. dolomiticus, E. dyerianus, E. eugene-maraisii, E. middelburgensis*, and *E. nubimontanus* (Fig. 4). These species inhabit South Africa's northern escarpment in cool, high-elevation woodland/savannahs. Their close phylogenetic relationship has received support in previous studies (Treutlein et al., 2005; Rousseau, 2012; Mankga et al., 2020; Stewart et al., 2023). Morphologically, these species share glaucous foliage, amphistomatic pinnae, megasporophyll extensions over the strobilus surface, and large seeds (Vorster, 2004; Stewart et al., 2023).

4.3.6 Ferox Clade (Clade VI)

This clade contains only *E. ferox* (Fig. 4) and is endemic to the shrubland of Mozambique and KwaZulu-Natal, South Africa, within 100 m of the ocean (Mabunda, 2007). This species is distinguished by its shortly emergent soft stem and coriaceous, dull green leaves with densely dentate and broad leaflets, and distinctively salmon-red-colored strobili. The species demonstrates reproductive isolation, as it does not hybridize with other *Encephalartos* species in cultivation (Vorster, 2004).

4.3.7 Manikensis Clade (Clade VII)

This clade includes *E. chimanimaniensis* R. A. Dyer & I. Verd., *E. concinnus* R. A. Dyer & I. Verd., *E. manikensis* (Gilliland) Gilliland, *E. munchii* R. A. Dyer & I. Verd., *E. pterogonus* R. A. Dyer & I. Verd., and an undescribed species (*Encephalartos* sp. nov.) (Fig. 4). The Manikensis Clade is endemic to the Zimbabwe–Mozambique border region in woodland/savannahs (Vorster, 2004; Mabunda, 2007). The members have papyraceous foliage, basal pinnae reduced to pinnacanths, emerging leaves with whitish indumentum, stout erect trunks, green glabrous strobili, and red sarcotesta (Vorster, 2004; Treutlein et al., 2005).

4.3.8 Turneri Clade (Clade VIII)

This clade extends across eastern, central, northern, and northwestern Africa (Fig. 2) and is subdivided into three subclades:

Turneri subclade (VIII-A): This subclade contains only *E. turneri*, which was previously considered to be related to *E. gratus* Prain (Treutlein et al., 2005; Rousseau, 2012; Mankga et al., 2020; Stewart et al., 2023) or considered to be morphologically similar to species in the Manikensis subclade (Vorster, 2004). Current evidence supports its classification as a distinct species (Fig. 4), occurring in the Nampula and Niasa provinces of Mozambique regions, characterized by an extremely hot and arid climate (Rousseau, 2012). Its megasporangiate strobili are olive green, with a distinctive pinkish discoloration around the edges of the strobilus megasporophylls. This unique coloration serves as an autapomorphy of this species (Rousseau, 2012).

Poggei subclade (VIII-B): This subclade includes *Encephalartos delucanus* Malaisse, Sclavo & Crosiers, *E. flavistrobilus, E. poggei, E. marunguensis, E. laurentianus* De Wild., *E. schaijesii*, and *E. schmitzii* Malaisse (Fig. 4), native to Central African countries (Fig. 2). These species are primarily found in the Democratic Republic of Congo, with some species extending into Zambia (*E. flavistrobilus* and *E. schmitzii*), Tanzania (*E. delucanus* and *E. marunguensis*), and the Angola–Congo border region (*E. laurentianus*). Previous studies have supported the phylogenetic relationships within this subclade (Treutlein et al., 2005; Rousseau, 2012; Mankga et al., 2020; Stewart et al., 2023). Most species are small plants, except for *E. laurentianus,* which produces some of the largest leaves in *Encephalartos*. The leaves, deciduous in the dry season, show glaucous green coloration. The pinnae are typically entire or, less commonly, with a few small teeth, and gradually reduced to pinnacanths toward the leaf base. The strobili emerge with a glaucous green color that turns yellowish at maturity, with a glabrous or very short russet-brown indumentum (Rousseau, 2012).

Hildebrandtii subclade (VIII-C): This subclade includes 13 species divided into five distinct groups (Fig. 4): gratus, hildebrandtii, bubalinus, barteri, and septentrionalis. The species in these subclades show broad geographical distributions. They primarily inhabit woodland/savannah ecosystems, with some occurring in forests (*E. whitelockii* P. J. H. Hurter), granite outcrops (*E. barteri*), and quartzite hills (*E. bubalinus* Melville) (Fig. 3). The gratus group inhabits southern Malawi and Mozamsbique (Eastern Africa sensu Stewart et al. (2023)); the hildebrandtii group occurs in northwestern Tanzania and southern Kenya (Eastern Africa sensu Stewart et al., (2023)); the bubalinus group extends across the northeastern Democratic Republic of Congo, Uganda, and Kenya (Northern Africa sensu Stewart et al. (2023)); the barteri group spans tropical West African Ghana, Benin, and Nigeria (Northern Africa sensu Stewart et al. (2023)); and the septentrionalis group occupies north-eastern Uganda (*E. macrostrobilus* Scott Jones & Wynants), southern and south-eastern Sudan (*E. mackenziei* L. E. Newton, *E. septentrionalis* Schweinf. ex Eichler) (Northern Africa sensu Stewart et al. (2023)). While these groups have subtle morphological differences, they share the characteristic of microsporangiate strobili emerging in succession. The gratus group is characterized by papyraceous leaves; glabrous, salmon-pink, cylindrical strobili; and red sarcotesta (Vorster, 2004). The hildebrandtii group has coriaceous leaves; glabrous, green to yellow, cylindrical/ovoid strobili; and yellow sarcotesta (Rousseau, 2012). The bubalinus group's key characteristics include coriaceous leaves; glabrous, yellow-green strobili; and yellow sarcotesta (Rousseau, 2012). Barteri group species have papyraceous leaves; glabrous, olive green, ovoid strobili; and red sarcotesta (Vorster, 2004). The septentrionalis group has papyraceous leaves; a slightly russet-brown indumentum on green, cylindrical strobili; and a red sarcotesta (Vorster, 2004).

5 Conclusions

This study elucidates the complex phylogeny of *Encephalartos*, Africa's most diverse cycad genus, through analysis of 3545 nuclear genes from 64 species. The integration of transcriptome-scale nuclear data with comprehensive morphological and biogeographic evidence represents the most thorough phylogenomic analysis of *Encephalartos* to date. This research resolves critical taxonomic uncertainties, identifies eight major clades, and reconstructs the evolutionary, temporal, and spatial dynamics underlying contemporary diversity. The findings show that *Encephalartos* originated in southern Africa during the Oligocene and experienced significant Miocene radiations, with climatic and geological changes driving continental range expansions and contractions. The marked reduction in speciation during the Pliocene and Pleistocene demonstrates the genus's sensitivity to environmental changes, which is a critical consideration, given current global climate change. The integration of molecular and morphological data facilitates a revised infrageneric classification that combines scientific rigor with practical utility. This framework enhances species identification and conservation efforts for *Encephalartos* taxa, many facing extinction risks. Additionally, it establishes a model for systematic studies of other ancient and taxonomically challenging plant groups. The results highlight the significance of combining genomic, morphological, and paleoclimatic approaches to understand relict lineage evolution. These insights should inform coordinated international conservation strategies, emphasizing collaborative research, policy development, habitat protection, and *ex situ* conservation. As ancient seed plant representatives, *Encephalartos* and other cycads provide invaluable insights into plant diversification, resilience, and adaptation across geological time. Their preservation maintains unique evolutionary lineages and ensures that critical knowledge about plant diversity origins and persistence remains available for future study.

Acknowledgements

The authors thank Dr. Liu Jian for valuable comments during manuscript writing. We are also grateful to Mr. Kampon Tansacha, the President of Nong Nooch Tropical Botanical Garden, for giving us permission to collect and use samples from his extensive cycad ex situ collection. This study was supported by the Scientific Foundation of Urban Management Bureau of Shenzhen (Nos. 202019 and 202409 to Shouzhou Zhang) and, in part, by the Special Startup Project Grant from Lushan Botanical Garden, Jiujiang (2024ZWZX04 to Sadaf Habib). Yiqing Gong has benefited from Propagation and Wild Reintroduction of *Cycas diannanensis* and *C. segmentifida* (2024–2025) from the National Key Wildlife and Plant Conservation Expenditure under the Central Forestry and Grassland Ecological Protection and Restoration Fund.

Conflicts of Interest

The authors declare no conflict of interest.

Data Availability Statement

The raw sequencing data are accessible through the NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) under BioProject accession number PRJNA929092. Sequence alignments underlying analyses and phylogenetic trees are available from the Science Data Bank (https://doi.org/10.57760/sciencedb.j00148.00016).

References

Abascal F, Zardoya R, Telford MJ. 2010. TranslatorX: Multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* 38: W7–W13.

Balkenhol N, Dudaniec RY, Krutovsky KV, Johnson JS, Cairns DM, Segelbacher G, Selkoe KA, von der Heyden S, Wang IJ, Selmoni O, Joost S. 2019. Landscape genomics: Understanding relationships between environmental heterogeneity and genomic characteristics of populations. In: Rajora OP ed. *Population genomics: Concepts, approaches and applications*. Cham: Springer International Publishing. 261–322.

Bezuidenhout H. 2020. Assessment results (2015–2018) of re-established poached cycad trees in Addo Elephant National Park, Eastern Cape, South Africa. *Koedoe* 62: a1626.

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

Calonje M, Meerow AW, Griffith MP, Salas-Leiva D, Vovides AP, Coiro M, Francisco-Ortega J. 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, Stevenson DW, Osborne R. 2025. The World List of Cycads (Version 2025.09.04-r1). Coral Gables, FL: Montgomery Botanical Center. https://doi.org/10.5281/zenodo.17059408

Chaiprasongsuk M, Mingmuang M, Thongpan A, Namwongprom K. 2007. Molecular identification of *Encephalartos* (Zamiaceae) species and their relationships to morphological characters. *Agriculture and Natural Resources* 41: 43–60.

Clugston JAR, Griffith MP, Kenicer GJ, Husby CE, Calonje MA, Stevenson DW, Little DP. 2016. *Zamia* (Zamiaceae) phenology in a phylogenetic context: Does in situ reproductive timing correlate with ancestry? *Edinburgh Journal of Botany* 73: 345–370.

Clugston JAR, Ruhsam M, Kernicher GJ, Henwood M, Milne RI, Nagalingum N. 2022. Conservation genomics of an Australian cycad *Cycas calcicola*, and the absence of key genotypes in botanic gardens. *Conservation Genetics* 23: 449–465.

Coetzee J. 1993. African flora since the terminal Jurassic. *Biological relationships between Africa and South America* 37: 37–61.

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.

Donaldson JS, Hill KD, Stevenson DW. 2003. Cycads of the World: An Overview. In: Donaldson JS ed. *Cycads: Status survey and conservation action plan*. Gland, Switzerland and Cambridge, UK: IUCN. 3–8.

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

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

Griffith MP, Clase T, Toribio P, Encarnación Piñeyro Y, Jiménez F, Gratacos X, Sanchez V, Meerow A, Meyer A, Kramer A, Fant J, Havens K, Magellan TM, Dosmann M, Hoban S. 2020. Can a botanic garden metacollection better conserve wild plant diversity? A case study comparing pooled collections with an ideal sampling model. *International Journal of Plant Sciences* 181: 485–496.

Grobbelaar N. 2004. *Cycads with special reference to the Southern African species*. South Africa: N. Grobbelaar, Pretoria.

Gutiérrez-Ortega JS, Salinas-Rodríguez MM, Martínez JF, Molina-Freaner F, Pérez-Farrera MA, Vovides AP, Matsuki Y, Suyama Y, Ohsawa TA, Watano Y, Kajita T. 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.

Habib S, Gong Y, Dong S, Lindstrom A, Stevenson DW, Wu H, Zhang S. 2023. Phylotranscriptomics shed light on intra-generic relationships and historical biogeography of *Ceratozamia* (Cycadales). *Plants* 12: 478.

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

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, Isahara H, Maegaard B, Mariani J, Mazo H, Moreno A, Odijk J, Piperidis S, Tokunaga T eds. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. Japan: European Language Resources Association (ELRA).

Hurter P, Glen H. 1996. *Encephalartos hirsutus* (Zamiaceae): A newly described species from South Africa. *South African Journal of Botany* 62: 46–48.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.

IUCN. 2025. The IUCN red list of threatened species. Version 2025-1 [online]. Available from https://www.iucnredlist.org. [accessed 05 May 2025].

Jacobs BF. 2004. Palaeobotanical studies from Tropical Africa: Relevance to the evolution of forest, woodland and savannah biomes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359: 1573–1583.

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.

Keppel G. 2009. Morphological variation, an expanded description and ethnobotanical evaluation of *Cycas seemannii* A.Braun (Cycadaceae). *The South Pacific Journal of Natural Science* 27: 20–27.

Konings KM. 2006. *Life history traits of South African* Encephalartos *spp*. (Zamiaceae) *and their implications for understanding population structure, responses to threats and effective conservation action*. Masters Dissertation. South Africa: University of Cape Town.

Kück P, Longo GC. 2014. FASconCAT-G: Extensive functions for multiple sequence alignment preparations concerning phylogenetic studies. *Frontiers in Zoology* 11: 81.

Laetsch DR, Blaxter ML. 2017. Kinfin: Software for taxon-aware analysis of clustered protein sequences. *G3: Genes, Genomes, Genetics* 7: 3349–3357.

Landis MJ, Matzke NJ, Moore BR, Huelsenbeck JP. 2013. Bayesian analysis of biogeography when the number of areas is large. *Systematic Biology* 62: 789–804.

Lemmon EM, Lemmon AR. 2013. High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics* 44: 99–121.

Lindstrom A, Habib S, Dong S, Gong Y, Liu J, Calonje M, Stevenson D, Zhang S. 2024. Transcriptome sequencing data provide a solid base to understand the phylogenetic relationships, biogeography and reticulated evolution of the genus *Zamia* L. (Cycadales: Zamiaceae). *Annals of Botany* 134: 747–768.

Liu J, Lindstrom A, Gong Y, Dong S, Liu Y, Zhang S, Xun G. 2024. Eco-evolutionary evidence for the global diversity pattern of *Cycas* (Cycadaceae). *Journal of Integrative Plant Biology* 66: 1170–1191.

Liu J, Zhang S, Nagalingum N, Chiang YC, Lindstrom A, Xun G. 2018. Phylogeny of the gymnosperm genus *Cycas* L. (Cycadaceae) as

inferred from plastid and nuclear loci based on a large-scale sampling: Evolutionary relationships and taxonomical implications. *Molecular Phylogenetics and Evolution* 127: 87–97.

Liu Y, Wang S, Li L, Yang T, Dong S, Wei T, Wu S, Liu Y, Gong Y, Feng X, Ma J, Chang G, Huang J, Yang Y, Wang H, Liu M, Xu Y, Liang H, Yu J, Cai Y, Zhang Z, Fan Y, Mu W, Sahu SK, Liu S, Lang X, Yang L, Li N, Habib S, Yang Y, Lindstrom AJ, Liang P, Goffinet B, Zaman S, Wegrzyn JL, Li D, Liu J, Cui J, Sonnenschein EC, Wang X, Ruan J, Xue J-Y, Shao Z-Q, Song C, Fan G, Li Z, Zhang L, Liu J, Liu Z-J, Jiao Y, Wang X-Q, Wu H, Wang E, Lisby M, Yang H, Wang J, Liu X, Xu X, Li N, Soltis PS, Van de Peer Y, Soltis DE, Gong X, Liu H, Zhang S. 2022. The *Cycas* genome and the early evolution of seed plants. *Nature Plants* 8: 389–401.

Mabunda MA. 2007. *Species-level phylogenetic reconstruction of the African cycad genus Encephalartos* (Zamiaceae). Masters Dissertation. South Africa: University of the Western Cape.

Maddison WP, Maddison DR. 2023. Mesquite: A modular system for evolutionary analysis. Version 3.81 [online]. Available from http://www.mesquiteproject.org [accessed 15 September 2024].

Maley J. 1996. The African rain forest–main characteristics of changes in vegetation and climate from the upper cretaceous to the quaternary. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences* 104: 31–73.

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

Matzke N. 2018. BioGeoBEARS: BioGeography with bayesian (and likelihood) evolutionary analysis with R scripts. version 1.1. 1. San Francisco, CA: GitHub. Available online at: https://github.com/nmatzke/BioGeoBEARS.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution* 66: 526–538.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* 37: 1530–1534.

Medina-Villarreal A, González-Astorga J, de Los Monteros AE. 2019. Evolution of *Ceratozamia* cycads: A proximate-ultimate approach. *Molecular Phylogenetics and Evolution* 139: 106530.

Mo YK, Lanfear R, Hahn MW, Minh BQ. 2022. Updated site concordance factors minimize effects of homoplasy and taxon sampling. *Bioinformatics* 39: btac741.

Morley RJ. 2001. Why are there so many primitive angiosperms in the rain forests of the far East? In: Metcalfe I, Smith JMB, Morwood M, Davidson I eds. *Faunal and floral migrations and evolution in SE Asia-Australia.* The Netherlands: A. A. Balkema, Rotterdam. 185–199.

Oliver JC. 2013. Microevolutionary processes generate phylogenomic discordance at ancient divergences. *Evolution* 67: 1823–1830.

Osborne R, Grobbelaar N, Vincent P. 1993. A numerical phenetic study of the genus Encephalartos Lehm. The biology, structure and systematics of the Cycadales. In: Stevenson DW, Norstog KJ eds. *Proceedings of CYCAD 90, The Second International Conference on Cycad Biology*. Milton: Palm and Cycad Societies of Australia. 279–280.

Plana V. 2004. Mechanisms and tempo of evolution in the African Guineo–Congolian rainforest. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359: 1585–1594.

Prakash S, Grobbelaar N, Van Staden J. 2008. Diversity in *Encephalartos woodii* collections based on Random Amplified DNA markers (RAPD's) and Inter-Specific Sequence Repeats (ISSR's). *South African Journal of Botany* 74: 341–344.

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

Ronquist F. 1997. Dispersal-vicariance analysis: A new approach to the quantification of historical biogeography. *Systematic Biology* 46: 195–203.

Rousseau P. 2012. *A molecular systematic study of the African endemic cycads*. Masters Dissertation. South Africa University of Johannesburg.

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

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, 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.

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

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

Stewart R, Clugston JAR, Williamson J, Niemann HJ, Little DP, van der Bank M. 2023. Species relationships and phylogenetic diversity of the African genus *Encephalartos* Lehm. (Zamiaceae). *South African Journal of Botany* 152: 165–173.

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

Tao Y, Chen B, Kang M, Liu Y, Wang J. 2021. Genome-wide evidence for complex hybridization and demographic history in a group of *Cycas* from China. *Frontiers in Genetics* 12: 717200.

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

Treutlein J, Vorster P, Wink M. 2005. Molecular relationships in *Encephalartos* (Zamiaceae, Cycadales) based on nucleotide sequences of nuclear ITS 1&2, *rbcl*, and genomic ISSR fingerprinting. *Plant Biology* 7: 79–90.

Van der Bank H, Wink M, Vorster P, Treutlein J, Brand L, van der Bank M, Hurter J. 2001. Allozyme and DNA sequence comparisons of nine species of *Encephalartos* (Zamiaceae). *Biochemical Systematics and Ecology* 29: 241–266.

Viljoen CD, van Staden J. 2006. The genetic relationship between *Encephalartos natalensis* and *E. woodii* determined using RAPD fingerprinting. *South African Journal of Botany* 72: 642–645.

Vorster P. 1993. Taxonomy of Encephalartos (Zamiaceae): Taxonomically useful external characteristics. In: Stevenson DW, Norstog KJ eds. *Proceedings of CYCAD 90, The Second International Conference on Cycad Biology*. Queensland: Palm and Cycad Societies of Australia. 294–299.

Vorster P. 2004. Classification concepts in *Encephalartos* (Zamiaceae). In: Walters T, Osborne R eds. *Cycad classification: Concepts and recommendations*. UK: CABI Publishing Wallingford. 69–83.

Vorster P, Van Der Bank F, Van der Bank M, Wink M. 2004. Phylogeny of *Encephalartos*: Some Eastern Cape species. *The Botanical Review* 70: 250–259.

Walters T, Osborne R. 2004. *Cycad classification: Concepts and recommendations*. UK: CABI Publishing, Wallingford.

Whitfield JB, Lockhart PJ. 2007. Deciphering ancient rapid radiations. *Trends in Ecology & Evolution* 22: 258–265.

Wichura H, Bousquet R, Oberhänsli R, Strecker MR, Trauth MH. 2010. Evidence for Middle Miocene uplift of the East African Plateau. *Geology* 38: 543–546.

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 J, Pagani M, Sloan L, Thomas E, Billups K. 2001. Trends, rhythms, and aberrations in global climate 65 MA to present. *Science* 292: 686–693.

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.

Supplementary Material

The following supplementary material is available online for this article at http://onlinelibrary.wiley.com/doi/10.1111/jse.70034/suppinfo:

Fig. S1. Phylogenetic tree of *Encephalartos* based on a concatenated data set of NT (left) and NT12 (right). The branches with MLBS < 100 are mentioned, and those with MLBS < 80 are split into polytomies.

Fig. S2. Support values from the site concordance factor (sCFl) and quartet analysis (q1) are displayed (scfl/q1) along the branches of species tree generated as for the coalescent CD-NT data set.

Fig. S3. (A) A phylogenetic network of the concatenated data set of *Encephalartos* species generated using SplitsTree6. The colors of branches are consistent with those in Fig. 1, representing major clades within the genus. (B) Gene flow among most recent common ancestors (MRCA) inferred with PHYLONET, with the best species tree (reticulations = 2) for the genus *Encephalartos*. Red lines indicate reticulation edges connecting reticulation nodes.

Table S1. The read statistics of 69 taxa investigated in this study.