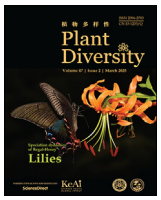
Plant Diversity 47 (2025) 229–243

Contents lists available at ScienceDirect

Plant Diversity

journal homepage: http://www.keaipublishing.com/en/journals/plant-diversity/
http://journal.kib.ac.cn



Research paper

An integrative approach clarifies species delimitation and biogeographic history of *Debregeasia* (Urticaceae)

Amos Kipkoech [a, c], Ke Li [a, d], Richard I. Milne [e], Oyetola Olusegun Oyebanji [f, g], Moses C. Wambulwa [h], Xiao-Gang Fu [a, c], Dennis A. Wakhungu [b, c], Zeng-Yuan Wu [a, **], Jie Liu [a, b, *]

[a] *Germplasm Bank of Wild Species & Yunnan Key Laboratory of Crop Wild Relatives Omics, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, Yunnan, China*
[b] *CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, Yunnan, China*
[c] *University of Chinese Academy of Sciences, Beijing 100093, China*
[d] *School of Ecology and Environment Science, Yunnan University, Kunming, China*
[e] *Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh, EH9 3JH, UK*
[f] *Department of Biology, University of Louisiana, Lafayette, LA, USA*
[g] *Department of Biological Sciences, University of Arkansas, Fayetteville, AR, USA*
[h] *Department of Life Sciences, School of Science and Computing, South Eastern Kenya University, Kitui, Kenya*

ARTICLE INFO

Article history:
Received 3 April 2024
Received in revised form
19 November 2024
Accepted 21 November 2024
Available online 26 November 2024

Keywords:
Debregeasia
Plastome
Phylogenomics
Reticulate evolution
Species delimitation
Ultra-barcodes

ABSTRACT

Integrative data from plastid and nuclear loci are increasingly utilized to resolve species boundaries and phylogenetic relationships within major angiosperm clades. *Debregeasia* (Urticaceae), an economically important genus, presents challenges in species delimitation due to its overlapping morphological traits and unstable taxonomic assignments. Here, we analyzed 14 morphological traits and generated 12 data matrices from the plastomes and nrDNA using genome skimming from the nine recognized morpho-species to clarify species boundaries and assess barcode performance in *Debregeasia*. We also used a universal set of 353 nuclear genes to explore reticulate evolution and biogeographic history of *Debregeasia*. Plastomes of *Debregeasia* exhibited the typical quadripartite structure with conserved gene content and marginal independent variations in the SC/IR boundary at inter- and intra-specific levels. Three *Debregeasia* species were non-monophyletic and could not be discerned by any barcode; however, ultra-barcodes identified the remaining six (67%), outperforming standard barcodes (56%). Our phylogenetic analyses placed *Debregeasia wallichiana* outside the genus and suggested six monophyletic clades in *Debregeasia*, although the placement between *Debregeasia hekouensis* and *Debregeasia libera* varied. There was extensive trait overlap in key morphologically diagnostic characters, with reticulation analysis showing potentially pervasive hybridization, likely influenced by speciation patterns and overlaps between species ranges. We inferred that *Debregeasia* crown diversification began at *ca.* 12.82 Ma (95% HPD: 11.54–14.63 Ma) in the mid-Miocene within Australia, followed by vicariance and later long-distance dispersal, mainly out of southern China. Our findings highlight the utility of genomic data with integrative lines of evidence to refine species delimitation and explore evolutionary relationships in complex plant lineages.

* Corresponding author. Germplasm of Bank of Wild Species & CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, Yunnan, China
** Corresponding author.
E-mail addresses: pkoech160@gmail.com (A. Kipkoech), like@mail.kib.ac.cn (K. Li), R.Milne@ed.ac.uk (R.I. Milne), oyebanjioyetola@gmail.com (O.O. Oyebanji), mcwambulwa@gmail.com (M.C. Wambulwa), fuxiaogang@mail.kib.ac.cn (X.-G. Fu), dennis@mail.kib.ac.cn (D.A. Wakhungu), wuzengyuan@mail.kib.ac.cn (Z.-Y. Wu), liujie@mail.kib.ac.cn (J. Liu).

Peer review under the responsibility of Editorial Office of Plant Diversity.

https://doi.org/10.1016/j.pld.2024.11.004

1. Introduction

Species delimitation is a fundamental cornerstone for conservation, management, and sustainable utilization of biodiversity (Hong, 2016; Wu et al., 2023). However, both conceptually and in practice, defining what constitutes a species remains highly challenging. Traditionally, species identification relied on morphological traits (Mishra et al., 2016; Shen et al., 2019). However, morphological analysis is often complicated by phenotypic plasticity—including both divergence within the same population and convergence across different species—resulting in morphological overlap (Fu et al., 2023). The advent of DNA barcoding using short-standardized sequences has significantly helped to overcome these constraints, offering a more reliable method for delineating species boundaries in complex groups (Hebert et al., 2003). Commonly used universal or standard plant barcodes include three plastid regions (*matK, rbcL*, and *trnH-psbA*) and the nuclear ITS (CBOL, 2009; Li et al., 2011). However, the discriminatory power of these fragments is rather limited for certain closely related taxa, particularly within recently diverged plant groups, mainly due to the lack of adequate genetic variation, potentially pervasive hybridization (Coissac et al., 2016; Zhang et al., 2023) and/or rapid radiation (Zan et al., 2023). Consequently, there is a pressing need for a more robust dataset of plant DNA barcodes to enhance deeper taxonomic resolution (Hollingsworth et al., 2016).

Urticaceae comprises over 2000 species and approximately 54 genera with a pantropical distribution (Plants of the World Online, 2024). Despite several phylogenetic analyses (Hadiah et al., 2008; Wu et al., 2013, 2015, 2018), some clades within the family remain poorly resolved due to the lower discriminatory power of limited DNA loci (Kim et al., 2015; Tseng et al., 2019). Studies have indicated that *Debregeasia* Gaudich (Fig. 1) is a monophyletic genus of the tribe Boehmerieae endemic to the Old World (Wang et al., 2020b; Wu et al., 2013, 2018) and distributed mainly in the Himalaya, East Asia, and Southeast Asia. Several *Debregeasia* species are valued for their fruits and stem fibers (Chen et al., 2003), as well as medicinal properties (Almubayedh and Ahmad, 2019). Available evidence indicates that *Debregeasia* originated from Eurasia and spread to Africa and Southeast Asia (Wu et al., 2018), with the current pantropical and rapid radiation likely occurring in Pliocene (Wang et al., 2020). However, these studies involved limited genome coverage and sampled few individuals per taxon. Hence, a deeper examination of the historical biogeography of *Debregeasia* is required. Since *Debregeasia* was described (Weddell, 1869), the number of accepted *Debregeasia* species has varied (see Table S1) among studies (Wilmot-Dear, 1988, 1994; Chen et al., 2003; Wilmot-Dear and Friis, 2012; Wang, 2016). Furthermore, the convergent evolution of morphological traits evident in Urticaceae (Wu et al., 2015) and likely in *Debregeasia* (Fig. S1) calls into question taxonomic assignments based on morphology alone. Hence DNA based analysis is particularly relevant in this genus, but analyses using few gene loci have met limited success in resolving relationships within the Boehmerieae tribe (Wu et al., 2013, 2018). This might be because the evolutionary trajectory of *Debregeasia* was markedly influenced by multiple reticulation events. Moreover, the role of biogeographical radiation in the speciation of the genus remains largely unexplored. Therefore, much greater discriminatory power marker is needed compared to previous work.

Next-generation sequencing (NGS) technologies have made it possible to generate low-coverage genomes of three organelles (plastome, mitochondrial, and nuclear) at a relatively low cost through genome skimming (Zhang et al., 2017). Plastomes and entire nuclear ribosomal DNA (nrDNA) have been proposed as suitable targets for next-generation barcodes in plants (Tong et al., 2022; Wang et al., 2022) and are often referred to as ultra-barcodes (Kane et al., 2012), organelle-scale barcodes (Yang et al., 2013), or super-barcodes (Ji et al., 2019). With more informative sites (Hollingsworth et al., 2016; Ji et al., 2019), these next-generation barcodes have improved species discrimination (Zhou et al., 2019; Guo et al., 2021; Jiang et al., 2024) and phylogenetic resolution in various plant taxa, such as the Millettioid/Phaseoloid clade from Fabaceae (Oyebanji et al., 2020), Rubiaceae (Amenu et al., 2022), the tribe Urticeae (Ogoma et al., 2022), and genus *Triplostegia* (Fu et al., 2023). In addition, plastome sequencing allows the detection and exploitation of hypervariable loci and lineage-specific indels, which has proven crucial in discriminating *Taxus* species (Liu et al., 2011, 2018). Furthermore, high-throughput sequences also offer a pathway for deciphering the organellar discordance in plant lineages, such as cyto-nuclear conflicts, by exploring 353 single-copy nuclear genes (Johnson et al., 2019). Unlike plastomes, which are maternally inherited, nuclear genomes are biparentally inherited, allowing for the detection of complex biological phenomena such as introgression, hybridization, and incomplete lineage sorting (ILS) (Ravinet et al., 2017). This capacity makes nuclear data particularly valuable for resolving relationships where plastome data may fall short. Despite the increasing application of ultra-barcodes (Ślipiko et al., 2020; Yu et al., 2022), their discriminatory power remains disputed. Also, the biological events leading to topological discordances are worth exploring. It is essential to continuously evaluate the utility of ultra-barcodes and biological complexities driving the evolutionary relationships for different plant groups (Li et al., 2021).

Here, standard DNA barcoding sampling was conducted for the nine recognized *Debregeasia* morphospecies, permitting comparative and integrative species delimitation and evolutionary analyses using plastome and nuclear genetic data. Our goals were to (i) characterize *Debregeasia* plastomes for potential informative structural variations, ii) compare the discriminatory power of the three categories of barcode data using multiple species delimitation algorithms, iii) explore trait evolution and identify key taxonomic diagnostic traits for monophyletic clades in *Debregeasia* populations and species based on the newly reconstructed phylogeny, and iv) investigate the reticulate evolution and biogeographic history of *Debregeasia*. The findings from this study will clarify the diversity and evolution of *Debregeasia* and contribute to a better understanding of evolutionary processes in Urticaceae lineages.

2. Material and methods

2.1. Sample collection

A total of 1963 individuals of *Debregeasia* were collected between 2018 and 2023. In addition, 60 newly sampled individuals representing the nine recognized species of *Debregeasia* (Fig. 1) were included for genomic analysis. Leaves from newly sampled individuals were dried and stored in silica gel for preservation. To provide a comparative framework for phylogenetic and evolutionary analysis, we also sampled 11 outgroup species (Table S2), representing six tribes within the Urticaceae family(Conn and Hadiah, 2009; Wu et al., 2013). Three datasets (I, II, and III) were generated based on our research goals. Dataset I, consisting of 51 individuals of *Debregeasia* and 11 outgroups, was used to characterize *Debregeasia* plastomes, assess barcode performance, and evaluate species boundaries in *Debregeasia*. For each species, 2 to 10 individuals were sampled from different locations, ensuring comprehensive coverage of their entire distribution range (Fig. 2 and Table S2). Dataset II, comprised 270 *Debregeasia* individuals from 91 populations, was designed to study evolution of diagnostic morphological traits in *Debregeasia* (Table S3). Finally, dataset III,

Fig. 1. Morphological diversity of *Debregeasia* species. (A) *D. longifolia* (Burm.f.) Wedd.; (B) *D. libera* J.J. Chien and C.J. Chen; (C) *D. elliptica* C.J. Chen; (D) *D. orientalis* C.J. Chen (E) *D. hekouensis* W.T. Wang; (F) *D. saeneb* (Forssk.) Hepper and J.R.I. Wood; (G) *D. squamata* King ex Hook.f.; (H) *D. wallichiana* Wedd.; (I) *D. australis* Friis, Wilmot-Dear and C.J. Chen. Photo credits: A–G by Zeng-Yuan Wu and Jie Liu; H by Debabrata Maity; I was downloaded from atlas of living Australia (https://www.ala.org.au/).

which included all individuals from dataset I, with nine additional *Debregeasia* individuals, was used to investigate the evolutionary relationships and biogeographic history of *Debregeasia* (Table S6).

2.2. DNA extraction, library preparation, and sequencing

A modified cetyltrimethylammonium ammonium bromide (CTAB) protocol (Doyle and Doyle, 1987) was used to extract total DNA from ~20 mg of the silica gel-dried leaves. Genomic DNA from each sample was then assessed for quality using a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA), and 1% agarose gel electrophoresis was run before library preparation. The library was built using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs, Ipswich, Massachusetts, USA) according to the manufacturer's instructions. Sequencing was done on an Illumina DNBSEQ-T7 (Illumina, San Diego, CA, United States) platform, yielding 150 bp paired-end reads and *ca.* 8 Gb of clean data generated for each individual.

2.3. Data assembly and annotation

The GetOrganelle v.1.7.5.3 toolkit (Jin et al., 2020) with k-mer values 21, 45, 65, 85, 105, and 127 was used for *de novo* assembly of the newly generated plastome and nrDNA (18S-ITS1-5.8S-ITS2-26S) data. The complete plastome (MN189948) and nrDNA (OM892755) from *Debregeasia elliptica* accession De19 were used as reference genomes for both assembly and annotation. The assembled plastomes were visualized and filtered to generate a complete circular genome in Bandage v.0.8 (Wick et al., 2015). The inverted repeat (IR) regions were identified by automatically flanking the IR using the function "find repeat" in Geneious v.8.1 (Kearse et al., 2012). Further curation of the annotated plastomes was done to confirm the start and stop codon, and tRNAscan-SE v.1.21 (Schattner et al., 2005) was used in default settings to verify the tRNA genes. The newly assembled and annotated genomes were deposited at the China National GenBank for accession numbers (Tables S2 and S6).

2.4. Plastome structural variation, divergence, and mutational hotspot analyses

We characterized the features of 51 plastomes, including genome size and structural elements, gene content [protein-coding (PCG) genes, tRNA, rRNA], and GC content. Comparative analyses of IR expansion and contraction at the four junctions of the sampled plastomes were conducted in Geneious and plotted in IRscope (Amiryousefi et al., 2018). Further, gene rearrangements were investigated in Mauve v.2.4.0 (Darling et al., 2010) with default settings, while sequence divergence analysis was performed in mVISTA in Shuffle-LAGAN mode (Frazer et al., 2004) using

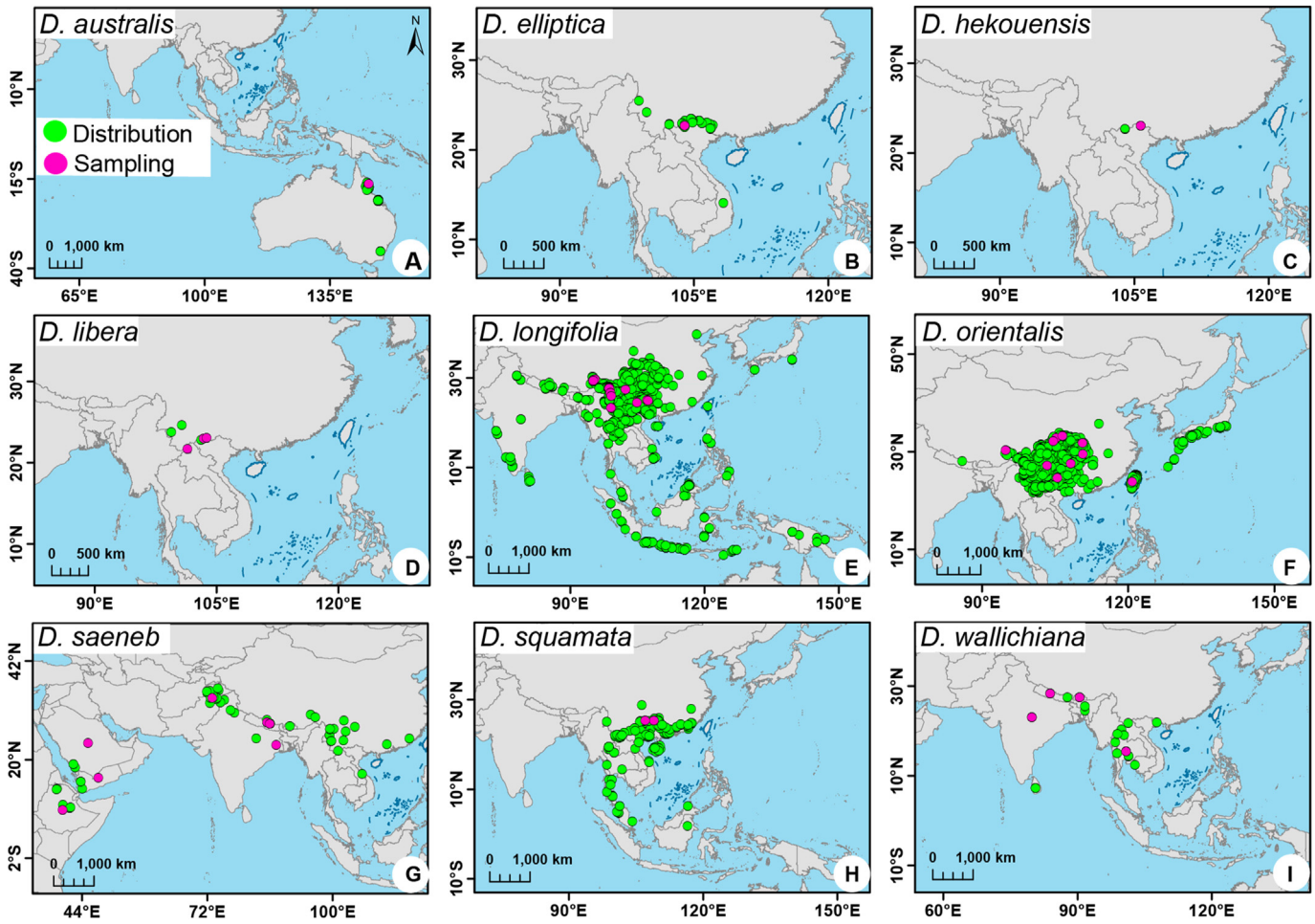


Fig. 2. Geographic distribution of the occurrence points and sampled localities of nine species of *Debregeasia*. Green dots represent the occurrence points obtained from online databases and literature, while pink dots indicate sampling sites in this study. The north arrow and point legend of Fig. 2B–I are consistent with those in Fig. 2A.

D. elliptica De19 as the reference plastome. We further assessed the hypervariable regions based on nucleotide diversity and the variation of parsimony informative sites and nucleotide diversity of all *Debregeasia* plastomes as an additional tool for species barcode assessment. The coding and non-coding regions of the sampled plastomes were extracted and separately analyzed. Parsimony informative sites were estimated in PAUP* v.4.0b10 (Swofford, 2002) while the nucleotide diversity analysis was conducted using a sliding window analysis in DnaSP v.5 (Librado and Rozas, 2009), with a step size of 200 bp and a window length of 600 bp. The top ten sequences based on nucleotide diversity were selected as hypervariable regions. We generated a physical circular map of the plastome in Chloroplot (Zheng et al., 2020).

2.5. Species discrimination analysis

We aligned the newly generated plastomes and nrDNA data separately using MAFFT v.7.22 (Katoh and Standley, 2013) with the LINSI algorithm; from this, multiple data matrices were generated to explore species delimitation among the sampled individuals. Six regions were extracted from the plastome using Geneious: coding sequences (CDS), non-coding sequences (nCDS), the hypervariable *ycf1* region, and three universal standard DNA barcode sequences—*rbcL, matK*, and *trnH-psbA*. Two nuclear datasets were extracted: ITS (ITS1-5.8S-ITS2), and nrDNA (18S-ITS1-5.8S-ITS2-26S). These datasets were combined to create 12 data matrices that were then classified into three barcode categories: ultra-barcodes (whole plastome, whole plastome + nrDNA, CDS, and nCDS), specific (nrDNA, *ycf1, matK* + *rbcL* + *trnH-psbA,* and *matK* + *rbcL* + *trnH-psbA* + ITS), and standard barcodes (*matK, rbcL, trnH-psbA*, and ITS). Each data matrix was aligned separately using MAFFT and manually adjusted in Geneious, while the most suitable nucleotide substitution model was determined in jModelTest 2 (Darriba et al., 2012). Sites that were parsimony-informative, constant, and variable were determined using PAUP* v.4.0b10 (Swofford, 2002), with gaps treated as missing data (Table 1).

Two types of broadly used species discrimination methods, i.e. tree-based and distance-based (see details in supplementary file 1), were employed in each of the three barcode data categories. Tree-based methods use a phylogenetic tree approach to calculate the variation along the branches to delimit species, and four were used here: Maximum likelihood (ML), Bayesian Inference (BI), Generalized Mixed Yule Coalescent (GMYC: Pons et al., 2006), and Poisson Tree Processes (PTP: Kapli et al., 2017). Conversely, distance-based methods mainly rely on the level of differences between sequences to calculate intraspecific and interspecific thresholds; of these, we employed three: Barcoding gap (BG), Automatic barcode gap discovery (ABGD; Puillandre et al., 2012) and assembly species by automatic partitioning (ASAP; Puillandre et al., 2021). All these methods use assembled data, and so a third type of method, skmer analysis (Sarmashghi et al., 2019), was performed on the unassembled reads so that its power could be compared to tree and

Table 1
Summary of sequence characteristics for 12 data matrices derived from 51 *Debregeasia* individuals. PIS (percentage of parsimony-informative sites) shows the proportion of informative characters contributing to resolving the evolutionary relationships.

Sequence	Aligned length (bp)	Variable sites (%)	PIS sites (%)	No. of constant sites (%)	Evolutionary models
Single-copy nuclear genes	277,835	72,158 (25.97)	66,336 (23.88)	153,277 (55.17)	GTRGAMMAI
Plastome + nrDNA	170,286	7555 (4.44)	6434 (3.78)	162,731 (95.56)	GTRGAMMA
Plastome	164,191	7093 (4.32)	6096 (3.71)	157,098 (95.68)	GTRGAMMA
CDS	80,557	2720 (3.38)	2332 (2.89)	77,837 (96.62)	GTRGAMMA
nCDS	92,050	4778 (5.19)	3938 (4.28)	87,272 (94.81)	GTRGAMMA
nrDNA	5922	458 (7.73)	300 (5.07)	5464 (92.27)	GTRGAMMAI
ycf1	5809	577 (9.91)	550 (9.45)	5243 (90.09)	GTRGAMMA
matK + *rbcL* + *trnH-psbA*	4790	293 (6.12)	276 (5.76)	4497 (93.88)	GTRGAMMAI
matK + *rbcL* + *trnH-psbA* + ITS	5468	483 (8.83)	444 (8.12)	4985 (91.17)	GTRGAMMAI
matK	1542	206 (6.68)	190 (6.16)	2878 (93.32)	GTRGAMMA
rbcL	1449	38 (2.62)	38 (2.62)	1411 (97.38)	GTRGAMMAI
trnH-psbA	257	59 (22.96)	49 (19.07)	198 (77.04)	GTRGAMMA
ITS	741	96 (40.85)	90 (38.30)	139 (59.15)	GTRGAMMAI

Note: (%) indicates the percentage of aligned length.

distance-based methods. The species discrimination power of ML, BI, and skmer was determined by whether all representatives of a given species clustered into a monophyletic group, whereas in BG analysis, a species was successfully identified when the minimum interspecific distance was higher than its maximum intraspecific distance. Other analyses relied on matches, defined as occurring when morphospecies correspond to operational taxonomic units (OTUs, defined as comprising all individuals of one species but no others). Hence, GMYC, bPTP (Bayesian Poisson Tree Processes), mPTP (multi-rate Poisson Tree Processes), ABGD, and ASAP were evaluated based on the number of matches obtained, following Magoga et al. (2021).

2.6. Morphological trait analysis

We used dataset II to comprehensively analyze 14 morphological traits that vary within *Debregeasia*, measuring two to three specimens in each sampled population (Table S3). All morphological terms used for both the qualitative (leaf base, leaf margin, leaf tip, leaf blade abaxial colour, leaf shape, perianth type, and pubescence of branch and petiole) and quantitative (number of secondary veins, inflorescence length, peduncle length, leaf area, leaf length, leaf width, and leaf weight) traits followed Harris and Harris (1994) and *Flora of China* (Chen et al., 2003). Quantitative traits were measured manually using a vernier calliper (SHAHE caliper, Wenzhou, Zhejiang, China). The leaf area was calculated using ImageJ software (Abràmoff et al., 2004), and leaf weight using ME-T analytical balance (Mettler-Toledo, Columbus, Ohio, USA). To account for within-individual variation, we measured leaf traits from three leaves of the middle stem, and the longest spreading inflorescences of the lower, middle, and upper stem parts. We calculated Spearman's rank correlation coefficients to define the relationship between characters and performed Principal Coordinates Analysis (PCoA) using 'ggfortify,' 'ggrepel,' and 'ggplot2' packages on the Gower distances from character trait data generated using the 'gowdis' function in 'FD' package to visualize the variations and clustering in the morphological traits of the *Debregeasia* species. All analyses were done in R (R Core Team, 2024).

Additionally, we performed ancestral state reconstruction (ASR) based on seven qualitative characters, with *Debregeasia australis* scored following Wilmot-Dear and Friis (2012). We used maximum parsimony (MP) in Mesquite v.2.75 (Maddison, 2007) on the newly reconstructed phylogeny based on the concatenated plastome and nrDNA data while treating character states as unordered. The character coding and data matrices are presented in Tables S4 and S5, respectively.

2.7. Single-copy nuclear genes assembly and alignment

We extracted the universal set of 353 single-copy nuclear genes from genome skimming data using Captus v.0.9.83 (Ortiz et al., 2023) following the recently published pipeline (https://github.com/edgardomortiz/Captus) based on dataset III (Table S6). MEGAHIT v.1.2.9 (Li et al., 2015) was used with the "assemble" function to combine paired reads into contigs. We then extracted the target genes using the "extract" function and aligned the extracted markers in MAFFT v.7.22.

2.8. Phylogenetic inference and reticulate evolution analysis

Phylogenetic relationships were inferred using ML and the species tree for the single-copy nuclear genes matrix. For whole plastome data, we relied solely on the ML tree. This was based on the performance of this method and the barcode in species delimitation in prior analysis. ML followed the procedure described in section 2.5 (see details in supplementary file 1), while the species tree was inferred using the multispecies coalescence (MSC) model in ASTRAL-III (Zhang et al., 2018). The MSC approach can detect biological events such as incomplete lineage sorting and hybridization causing topological conflicts or discordance. Thus, we compared the plastid and nuclear trees from ML and ASTRAL-III for gene and species trees, respectively. These alternative tree topology frameworks were used to provide insights into the reticulate evolutionary histories of the genus *Debregeasia*. We quantified the gene concordance using PhyParts v.0.0.1 (Smith et al., 2015). Next, the quartets (SNaQ) method implemented in the Julia package PhyloNetworks v.0.6.0 (Bezanson et al., 2017; Solís-Lemus et al., 2017) was employed to infer phylogenetic networks from the species networks, for five separate runs with the maximum number of hybridization events varying from 0 to 4. The best phylogenetic network with the smallest pseudolikelihood value was identified as showing the most likely pattern of past hybridization events. A coalescent simulation analysis was performed to measure the goodness-of-fit of the coalescent model and whether ILS explains the gene tree discordance. If the simulated gene trees correspond well to the empirical tree, ILS may explain discordance (Yang et al., 2020; Morales-Briones et al., 2021).

2.9. Molecular dating and ancestral area reconstruction

We conducted molecular dating on the newly reconstructed ML tree of 352 single-copy nuclear genes based on a penalized likelihood (PL) in treePL (Smith and O'Meara, 2012). We utilized the ML tree from the single-copy nuclear genes matrix due to its higher

support values and resolved relationships. The nodes were calibrated using two selected fossil points from the outgroups, Forsskaoleeae (33.9 Ma) and *Girardinia* (11.6 Ma). Additionally, one secondary calibration (crown divergence) for Urticaceae 68.7 Ma (56.2–81.7 Ma) was applied, as inferred from previous study based on the fossil (Wu et al., 2018). Following the empirical guide described by Maurin (2020), the smoothing parameter was determined using the cross-validation option, and priming was used to determine the best optimization scores. We used TreeAnnotator v.2.6.7 (Bouckaert et al., 2014) to summarize the dated bootstrapped trees and produce the confidence intervals (CI) for a consensus tree. The results were visualized in FigTree v.1.4.0 (Rambaut, 2009).

We conducted the biogeographical analyses using Bayesian Binary MCMC Analysis (BBM) implemented in RASP (Yu et al., 2020) based on the dated ML tree from single-copy nuclear genes data with the outgroup pruned. The number of maximum areas for ancestral nodes was set at two, representing the maximum number of areas observed in extant *Debregeasia* species. The distribution area was divided into seven regions: A, Southern China; B, Third Pole (Tibetan Plateau and adjacent mountain regions); C, Southeast Asian archipelago; D, Indian subcontinent; E, Australian archipelago; F, Yemen, Saudi Arabia, and Ethiopia; G, Taiwan (China) and Japan. These areas were extrapolated from the literature in *Flora of China* and plants of the world based on the geological and climatic history, distribution, and endemism of *Debregeasia* (Chen et al., 2003).

3. Results

3.1. Plastome structural variation, sequence divergence, and hypervariable regions

Plastomes of all 51 individuals of the nine *Debregeasia* species examined had a quadripartite structure (Fig. 3A) and exhibited limited intraspecific variation in size (Table S7). The lengths of the large single-copy (LSC), small single-copy (SSC), and IR regions had ranges of 84,904–90,893 bp, 18,532–19,217 bp, and 22,536–25,689 bp, respectively (Table S7). There was marginal variation in total GC content, ranging from 36.30 to 36.80%. However, the GC content for IR regions was higher (42.6–42.7%) than that of LSC (34.0–34.5%) and SSC (29.4–30.2%). *Debregeasia* plastomes are highly conserved, and encode between 125 and 130 genes, of which 80–85 are protein-coding, eight encode rRNA, and 37 tRNA genes.

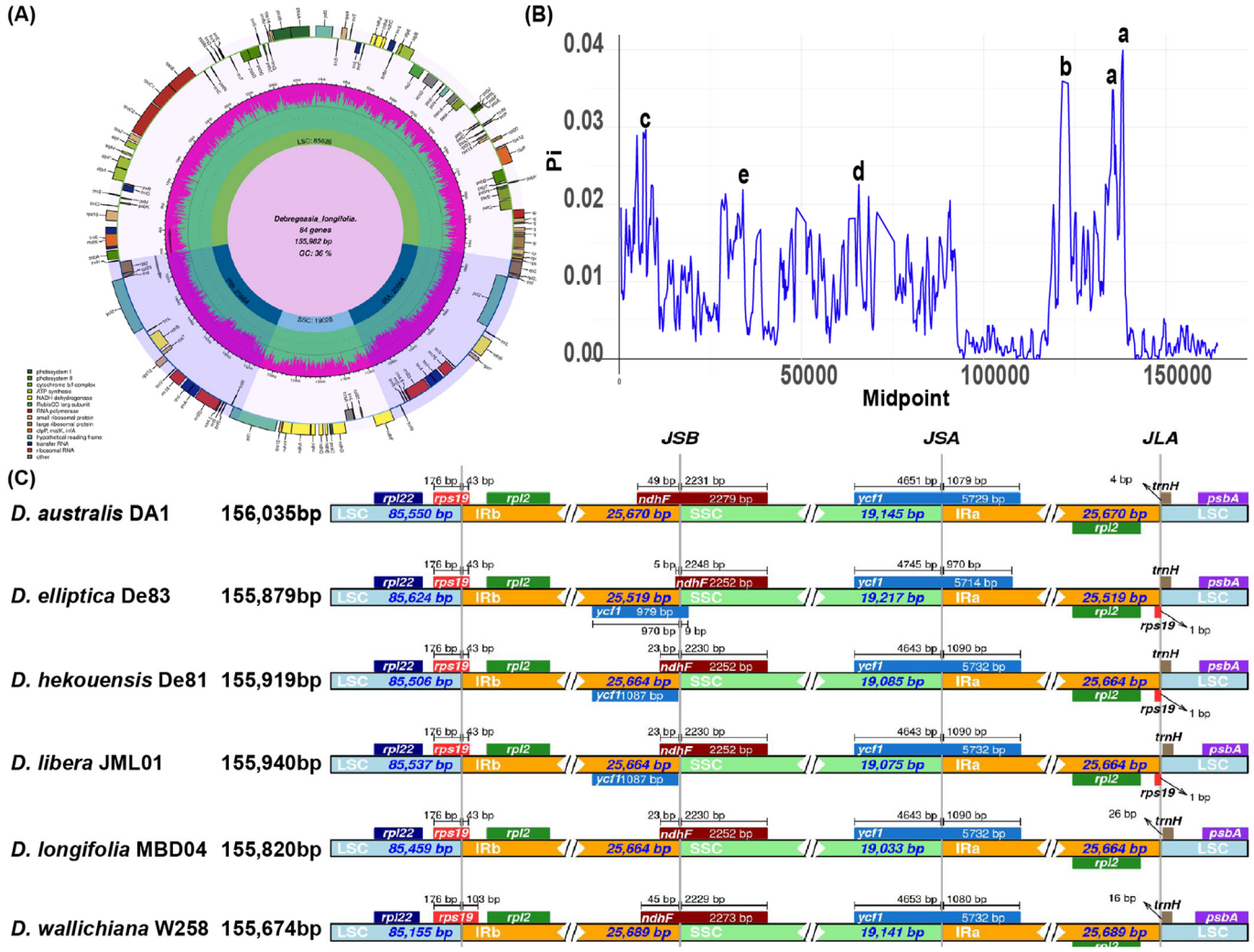


Fig. 3. Plastome structure of *Debregeasia*. (A) Gene map of the complete plastome of *D. longifolia* showing the typical plastome gene organization in *Debregeasia*. (B) Nucleotide diversity (π) in the homologous regions of 51 *Debregeasia* plastomes. The a–e represent the top five hypervariable regions in the order shown in Table S8, from highest to lowest. (C) Comparison of SC and IR borders among six representatives *Debregeasia* plastomes showing typical gene organization and interspecific boundary variation in *Debregeasia*.

The plastomes of *Debregeasia* species studied have experienced independent boundary shifts in the IR both between (Fig. 3C) and within species (Fig. S2). The LSC/IRb border is embedded in the *rps*19 gene (of which 27–43 bp is located within IRb), causing an independent boundary shift (Fig. S2). The IRb/SSC boundary is located within the *ndhF* gene (of which 21–36 bp is located within IRb), while the IRa/LSC boundary is within the *rpl2–trnH-GUG* region (Fig. S2). The boundary SSC/IRa is conserved throughout the studied plastomes and is located within the *ycf*1 coding gene (of which 970–1370 bp is located within IRa).

There were no significant evolutionary gene rearrangements or inversions detected in the studied plastomes (Fig. S3). Similarly, sequence variation analysis of *Debregeasia* plastomes revealed low sequence divergence within and between species (Fig. S4). Nucleotide diversity (π) ranged from 0 to 0.0422 across the plastomes, and the most hypervariable regions included *rps16–trnQ-UUG, ycf4–cemA, trnT-GGU–psbD, rpl32–trnL-UAG, and ycf1* (Table S8 and Fig. 3B). Parsimony-informative sites for all regions ranged from 0 to 516 sites (Figs. S5 and S6; Tables S9 and S10). Because *ycf1* exhibited both the highest π value and parsimony-informative sites, its species discriminatory power was further investigated.

3.2. Characteristics of data matrix

Of the 12 data matrices, an ultra-barcode (plastome + nrDNA) and standard barcode (ITS) had the longest (170,286 bp) and shortest (741 bp) alignment length, respectively. However, one standard barcode (ITS) had the highest percentage of variable (40.85%) and parsimony-informative (38.30%) sites, in contrast to another standard barcode (*rbcL*), which had the smallest percentage (2.62%) of both variable and parsimony-informative sites (Table 1).

3.3. Species discrimination

3.3.1. Tree-based species discrimination

BI and ML topologies derived from all 12 data matrices for the nine species were congruent regarding which species were monophyletic; in addition, these methods outperformed other tree-based methods used in this study (Table 2 and Figs. S7–S18). Six species were monophyletic in all analyses, whereas *Debregeasia longifolia, D. hekouensis,* and *D. libera* were non-monophyletic, and hence could not be discriminated by any barcode. Three data matrices from the ultra-barcodes (plastome + nrDNA, plastome, and CDS) and two specific barcodes (nrDNA and *matK* + *rbcL* + *trnH-psbA* + ITS) showed the highest discriminatory power (67%; six of nine species discriminated). The standard barcodes *matK, rbcL,* and ITS gave an average performance of 56% (five of nine species discriminated), while *trnH-psbA* had the least resolution, with only two out of nine species discriminated (22%; Table 2). Additionally, we noted significant topological incongruence between the plastome vs. nrDNA results (Fig. 4A and B).

We observed contrasting discriminatory powers of the data matrices with the different algorithm thresholds. By delimiting five species, the *ycf1* specific barcode and CDS ultra-barcodes had the highest discriminatory power in bPTP and mPTP, respectively. In contrast, the nrDNA specific barcode and all the four standard barcodes (ITS, *matK, rbcL, trnH-psbA*) had the least resolution, delimiting no species in mPTP. In mGMYC, the plastome, plastome + nrDNA, and ITS performed slightly below-average, delimiting four out of nine species, whereas the CDS ultra-barcodes could not delimitate any species (0/9). In sGMYC, the standard barcode *matK* gave the highest resolution (5/9), whereas both the ultra-barcodes (plastome and plastome + nrDNA) and ITS (standard barcode) delimited 4/9, and the others even fewer. In putative OTU analysis (Tables S11–S14), plastome + nrDNA yielded similar ML OTUs in both sGMYC and mGMYC, whereas the plastome ultra-barcodes alone gave more in mGMYC compared to sGMYC (Table S11). The CDS ultra-barcodes stand out with the highest OTUs (38) in sGYMC, whereas ITS exhibited the lowest (1). Both *matK* and *rbcL* resulted in more ML clusters in sGMYC than mGMYC (Table S11). The bPTP analysis showed a high acceptance rate and maximum split for the specific barcode (*matK* + *rbcL* + *trnH-psbA*) (Table S12), and most OTUs in nrDNA specific barcode (22) and the fewest (2) in ITS standard and CDS ultra-barcodes. In contrast, mPTP analysis gave the most OTUs in the CDS ultra-barcodes (10) and the fewest in the nrDNA-specific barcode (1), exhibiting a significant discrepancy between the two methods (Table S13). Plastome ultra-barcodes gave the same number of OTUs (8) in both mPTP and bPTP, whereas the four standard barcodes (*matK, rbcL, trnH-psbA*, and ITS) gave fewer (between 2 and 6) in both mPTP and bPTP.

3.3.2. Distance-based species discrimination

The discriminatory power of the distance-based analyses varied considerably, identifying 1–6 species of the nine morphospecies. In barcode gap analysis, a specific barcode (*matK* + *rbcL* + *trnH-psbA* + ITS) had the highest discriminatory power (6/9 species;

Table 2
Species discrimination success rate of twelve data matrices based on tree-based and distance-based methods.

Data matrices	Tree-based method						Distance-based method		
	ML	BI	bPTP	mPTP	sGMYC	mGYMC	BG	ABGD	ASAP
	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
Plastome + nrDNA	6/9 (67)	6/9 (67)	3/9 (33)	0/9 (0)	2/9 (22)	3/9 (44)	5/9 (56)	5/9 (56)	4/9 (44)
Plastome	6/9 (67)	6/9 (67)	4/9 (44)	4/9 (44)	4/9 (44)	4/9 (44)	5/9 (56)	4/9 (44)	4/9 (44)
CDS	6/9 (67)	6/9 (67)	1/9 (11)	5/9 (56)	1/99 (11)	0/9 (0)	5/9 (56)	2/9 (22)	3/9 (33)
nCDS	5/9 (56)	5/9 (56)	4/9 (44)	2/9 (22)	1/9 (11)	3/9 (33)	5/9 (56)	2/9 (22)	4/9 (44)
nrDNA	6/9 (67)	6/9 (67)	1/9 (11)	0/9 (0)	2/9 (22)	3/9 (33)	3/9 (33)	1/9 (11)	1/9 (11)
matK + *rbcL* + *trnH-psbA* + ITS	6/9 (67)	6/9 (67)	2/9 (22)	1/9 (11)	4/9 (44)	3/9 (33)	6/9 (67)	4/9 (44)	2/9 (22)
matK + *rbcL* + *trnH-psbA*	5/9 (56)	5/9 (56)	2/9 (22)	1/9 (11)	2/9 (22)	1/9 (11)	5/9 (56)	5/9 (56)	5/9 (56)
ycf1	5/9 (56)	5/9 (56)	5/9 (56)	1/9 (11)	4/9 (44)	3/9 (33)	3/9 (33)	2/9 (22)	3/9 (33)
ITS	5/9 (56)	5/9 (56)	1/9 (11)	0/9 (0)	4/9 (44)	4/9 (44)	5/9 (56)	5/9 (56)	4/9 (44)
matK	5/9 (56)	5/9 (56)	1/9 (11)	0/9 (0)	5/9 (56)	2/9 (22)	4/9 (44)	5/9 (56)	3/9 (33)
rbcL	5/9 (56)	5/9 (56)	1/9 (11)	0/9 (0)	0/9 (0)	2/9 (22)	5/9 (56)	5/9 (56)	5/9 (56)
trnH-psbA	2/9 (22)	2/9 (22)	0/9 (0)	0/9 (0)	4/9 (44)	1/9 (11)	5/9 (56)	3/9 (33)	3/9 (33)

Note: Percentage of discriminated species presented in (); ML, maximum likelihood; BI, Bayesian Inference; sGMYC, single-threshold generalized mixed yule coalescent; mGMYC, multiple-threshold generalized mixed yule coalescent; bPTP, Bayesian Poisson Tree Processes; mPTP, multi-rate Poisson Tree Processes; BG, barcoding gap; ABGD, automatic barcode gap discovery; ASAP, assembly species by automatic partitioning.

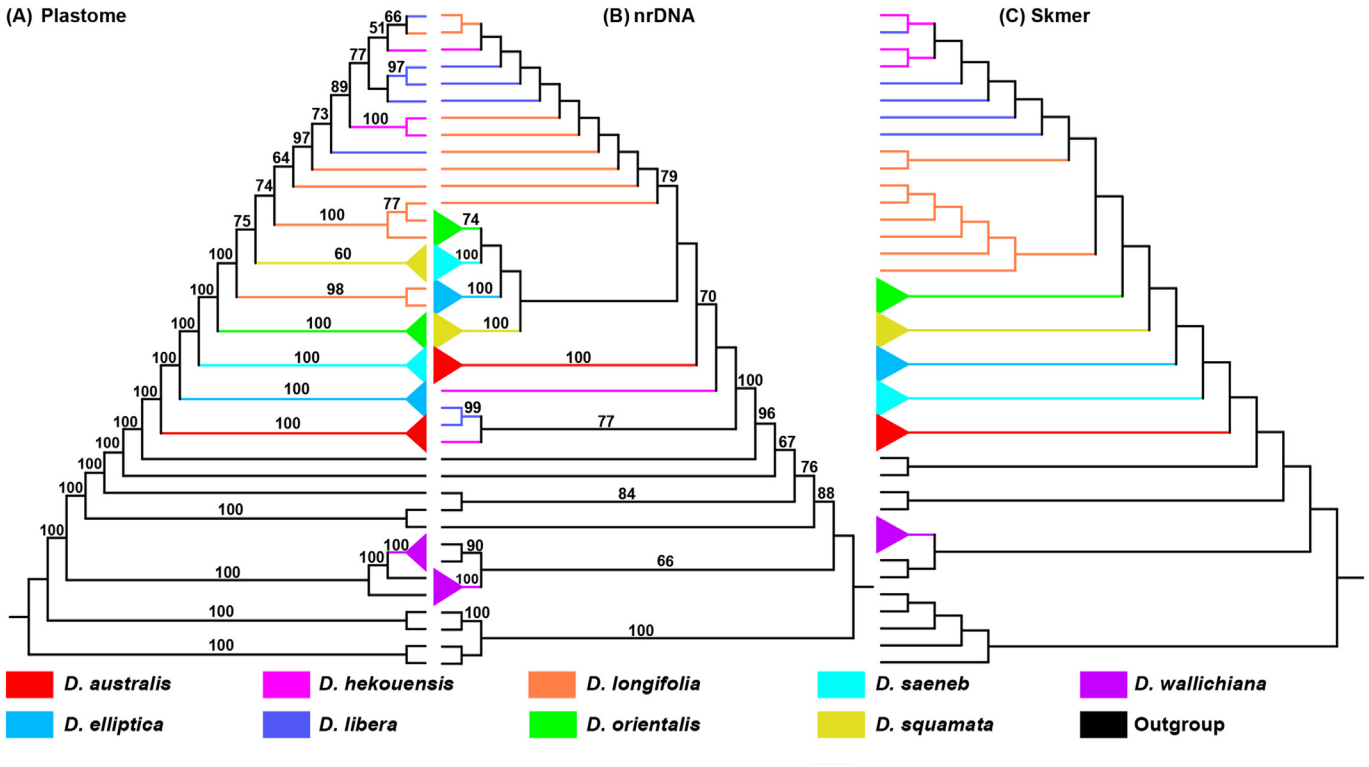


Fig. 4. Tree-based clustering relationships for 51 individuals of *Debregeasia*. (A and B) Species discrimination based on tree topology, comparing the plastome (left) and nrDNA (right) barcodes of *Debregeasia* using maximum likelihood analysis. Only bootstrap support values above 50% are shown. (C) Skmer tree derived from genomic distances of unassembled reads. Triangles represent monophyletic clades corresponding to individual species.

67%); next best were the four ultra-barcodes (plastome + nrDNA, plastome, CDS, and nCDS) and a standard barcode (ITS), each of which identified five of the nine species (Fig. S19A). In ABGD analysis, five barcodes, including three standard barcodes (*matK, rbcL*, and ITS), one ultra-barcodes (plastome + nrDNA) and one specific barcode (*matK* + *rbcL* + *trnH-psbA*), performed best (6/9; 67%) whereas nrDNA performed the worst (1/9; 11%). Using the ASAP algorithm, the specific barcode (*matK* + *rbcL* + *trnH-psbA*) and standard barcode (*rbcL*) performed best (5/9; 56%), while other barcodes discriminated up to 4/9 taxa (Table 2). The number of generated OTUs varied across ABGD analysis with the different prior intraspecific divergence both in initial and recursive partitions (Tables S14 and S15), with the standard barcode (ITS) generating the highest number of OTUs (25) in recursive partition and the fewest in the *ycf1* specific barcode (2) in initial partition, both at 0.001 prior intraspecific divergence. The three models, K80 Kimura, Jukes-Cantor (JC69), and simple distance, produced comparable OTUs for each dataset. The number of discriminated OTUs similarly varied across different ASAP score values (Table S16), with the nCDS specific barcode generating the most OTUs (33), while *matK* + *rbcL* + *trnH-psbA*, *ycf1*, and ITS generated the fewest (2). The skmer method equally discriminated up to six out of nine species (67%; Fig. 4C).

3.4. Comparison of species resolution for different analytical methods

Species discrimination rates varied across analysis method and barcodes (Fig. S19B). ML and BI outperformed other methods, with better resolution in six out of twelve barcodes and a maximum resolution of 67% (6 out of 9 species discriminated). In contrast, mPTP performed the worst, with no species discriminated by any of the standard barcodes (0% resolution). Among the distance-based analyses, the barcode gap method showed a better resolution power (> 56% in nine of twelve barcodes) than ASAP and ABGD. Equally, skmer analyses resulted in the best resolution (67%), similar to the maximum performance achieved in BI and ML methods.

3.5. Morphological evolution

Spearman's rank test showed varying relationships among the examined morphological characters, with eight characters recovered as uncorrelated (Table S17). PCoA of all 14 morphological traits showed a better clustering than when the qualitative and quantitative traits were separately analyzed, with four morphospecies (*Debregeasia elliptica, D. libera, D. squamata,* and *D. wallichiana*) displaying distinct clusters. A similar pattern was resolved when only the eight selected uncorrelated traits were analyzed (Fig. 5).

ASR revealed multiple cases of homoplasy involving key diagnostic morphological characters (Figs. S20A–B). Leaf shape barely defined half of the species, with the rest of the characters exhibiting rampant homoplasy. The leaf tip, base, perianth type, and margin showed high homoplasy with shared character states across species (Figs. S21–S25). Three synapomorphies were found in *Debregeasia elliptica* (rounded leaf base, finely denticulate leaf margin, and elliptic leaf shape), two in *D. saeneb* (thickly snow-white tomentose and sparsely spreading hirtellous), and *D. squamata* (cordate leaf base, fleshy spreading protuberances, and mixed appressed hair), and a single trait in *D. wallichiana* (shortly acuminate leaf tip).

3.6. Sequence characteristics of single-copy nuclear genes

We recovered 352 exons of the universal set of 353 targeted loci. The length of each locus after trimming ranged from 195 to 3657

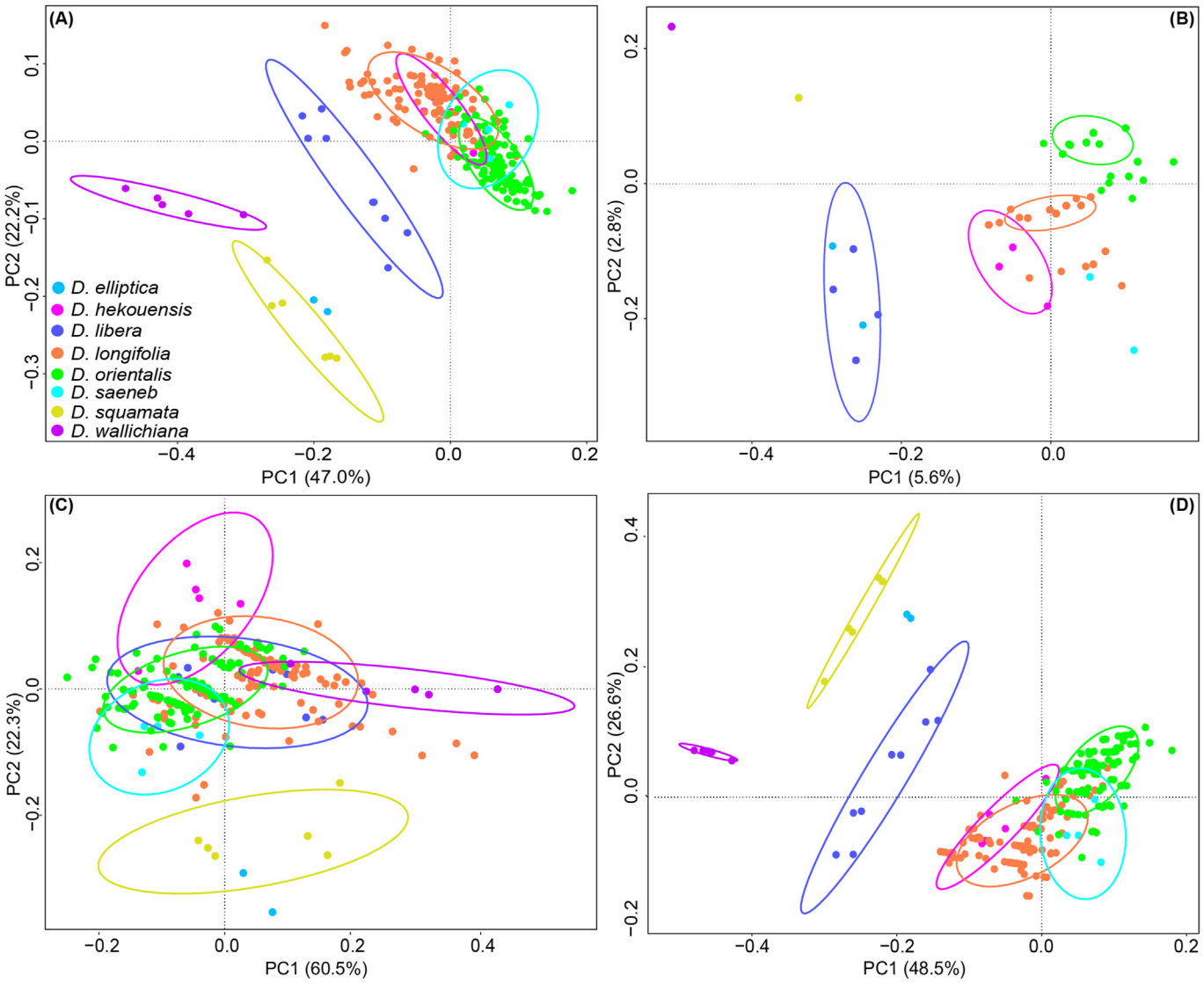


Fig. 5. Principal Coordinates Analysis (PCoA) analysis of 14 morphological traits of *Debregeasia*. Individual data points are shown with the proportion of variance explained by PC1 and PC2 displayed for: (A) all 14 traits, (B) seven qualitative traits, (C) seven quantitative traits, and (D) the eight uncorrelated traits selected based on Spearman's rank correlation coefficient. The color coding for each species in Fig. 5B–D is consistent with that of Fig. 5A.

bp, with GC content ranging from 39.0% to 68.7%. These genes were assembled into a concatenated matrix of aligned length 277,835 bp, with 153,277 constant sites, 5822 parsimony-uninformative, and 66,336 parsimony-informative sites. GC content was 46.8%.

3.6.1. Phylogenetic resolution and reticulation patterns

The nuclear tree was largely consistent with the plastome tree but with conflicting phylogenetic positions detected in the concatenated *Debregeasia* tree (Fig. S26), where *D. hekouensis* and *D. libera* placements varied significantly. In both data matrices (Figs. S27–S29), *Debregeasia* is supported as monophyletic (BS = 100%; PP = 1.00), excluding *D. wallichiana*, which clustered with *Coussapoa contorta* and *Maoutia setosa* with strong support (BS = 100%; PP = 1.00). The single-nuclear copy data matrix recovered six *Debregeasia* species as fully supported monophyletic clades (*D. elliptica, D. saeneb, D. orientalis, D. squamata, D. longifolia*, and *D. australis*), whereas the plastome recovered only five (with *D. longifolia* in two clades). The conflict analysis based on the nuclear gene trees showed that most gene trees (144/118) supported the monophyly of *Debregeasia*, suggesting potential evolutionary processes such as ILS and hybridization (Fig. 6A). In line with this, phylonet recovered two phylogenetic networks of possible hybridizations from the four runs with (Likelihood: −5463.62) (Fig. 6C) and (Likelihood: −5286.41) (Fig. 6D) where one hybridization event was identified as the best-fit model based on the smallest pseudolikelihood value (Fig. 6C). This network suggests that 68.3% of ancestral lineage in *Debregeasia* are from *D. australis* and 31.7% from another ancestral sister Urticaceae species (possibly the close relatives used as outgroups). Coalescent simulation analysis between the empirical and simulated gene-to-gene tree at the species level showed much overlap between the plastid and simulated gene tree (Fig. 6B). Although the simulations did not rule out ILS, SNaQ analysis suggests that other alternatives, including gene flow and plastid capture, might explain the conflict.

3.6.2. Molecular dating and ancestral area reconstruction

Our dating analysis suggests that *Debregeasia* started to diversify at *ca.* 12.82 Ma (crown-age; 95%HPD: 11.54–14.63Ma) (Fig. 7D) in

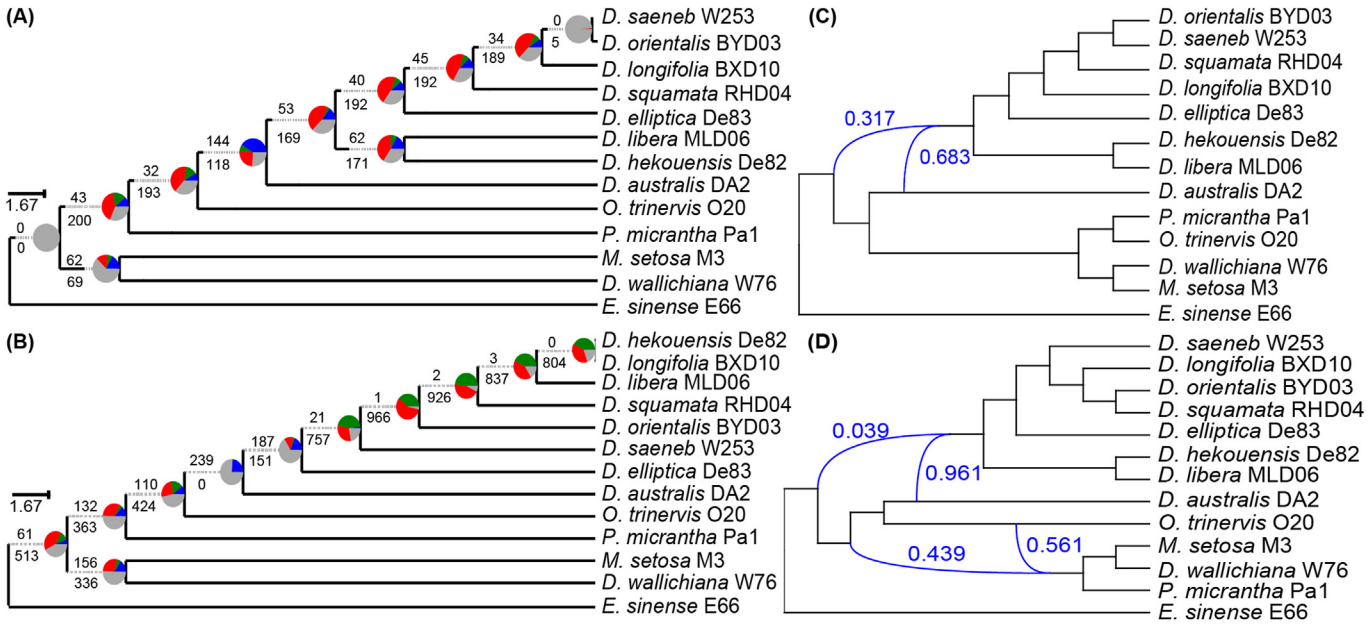


Fig. 6. Coalescent simulation and phylogenetic network analysis. The analysis involving nine *Debregeasia* species and four outgroups. (A) Relationships within *Debregeasia* inferred using ASTRAL-III, based on 352 nuclear genes, showing gene-tree concordance and conflict. Pie charts at the nodes represent the proportion of gene trees in concordance (blue), in conflict (green, the most common alternative; red, other alternatives), and uninformative (gray) with the bipartition. Numbers above and below the branches indicate the number of concordant and conflicting genes, for each bipartition. (B) A summary of 1000 chloroplast trees simulated under the coalescent using the ASTRAL-III tree, based on 352 nuclear genes, as the guide tree. (C and D) Phylogenetic networks inferred from SNaQ network analysis, allowing for a maximum of one (C) and two (D) reticulations. Blue curved branches indicate possible hybridization events, with inheritance probabilities labeled at the nodes.

the mid-Miocene era of the Neogene period. Our analyses further indicated that the biogeographical ancestral area of *Debregeasia* may be the Australian archipelago (Prob: 51.75%), and that the genus originated during the mid-Miocene. A variance event subsequently split *Debregeasia* into two lineages. The first lineage consists of the Australian species, *D. australis,* whereas the second lineage consists of the remaining species from Asia and Africa (Fig. 7A and B). At least six dispersals and one variance event occurred during the Miocene, which influenced the present-day distribution pattern of the genus (Fig. 7C).

4. Discussion

4.1. Plastome characteristics and implications for DNA barcoding

Plastomes of *Debregeasia* were found to be highly conserved and exhibited typical quadripartite structure, similar to other Urticaceae plastomes such as *Oreocnide* (Wu et al., 2022), *Urtica* (Li et al., 2022), members of Urticeae (Ogoma et al., 2022) and other photosynthetic angiosperms (Wang et al., 2020b; Amenu et al., 2022). No inversions were detected, but several species had independent IR-SC boundary shifts, as detected for other taxa (Park et al., 2018; Ogoma et al., 2022). In *Debregeasia*, the maximum possible species discrimination based on plastomes was 6/9, with *D. libera, D. hekouensis,* and *D. longifolia* similarly non-monophyletic in nrDNA (Fig. 4). Possible reasons for this are discussed below, but in practice, barcode performance could therefore not exceed six discerned species, or 67%. Hence, plastome barcodes achieved the best possible performance alone or combined with nrDNA and also as coding sequences only (CDS) (Fig. S19B). This confirms the reported utility of complete plastomes as barcodes (Fu et al., 2022; Xu et al., 2022; Zhang et al., 2023), especially where species are difficult to resolve, as also shown in *Ficus* (Bruun-Lund et al., 2017), *Diospyros* (Li et al., 2018), *Panax* (Ji et al., 2019), and *Taxus* (Fu et al., 2019). Although standard barcodes have performed well for *Taxus* (Liu et al., 2018), *Alnus* (Ren et al., 2010), *Dendrobium* (Singh et al., 2012), and *Pedicularis* (Yu et al., 2015), their performance may be suboptimal in recently diverged species or taxonomically complex plant taxa (Liu et al., 2022; Yu et al., 2022). Here, most standard individual barcodes performed relatively poorly, resolving a maximum of five *Debregeasia* species, indicating a need for specific or ultra-barcodes. However, combining the standard barcodes with others (*matK* + *rbcL* + *trnH-psbA* + ITS) offered an improved resolution, discriminating all six monophyletic species in barcoding gap analysis and both ML and BI. This indicates that discrimination success is improved by increasing the number of variable characters examined.

Taxon-specific markers offer a potential compromise between the expense involved in ultra-barcodes, such as the whole plastome, and the potential problem of limited genetic variation for standard barcodes. We identified five mutational hotspots for *Debregeasia*: *rps16–trnQ-UUG, ycf4–cemA, trnT-GGU–psbD, rpl32–trnL-UAG*, and *ycf*1. These mutational hotspots were similar to hotspots for other Urticaceae (Li et al., 2021; Ogoma et al., 2022). These regions had π values ranging from 0.0223 to 0.0421. The *ycf1* hotspot offered similar discrimination power (56%) to three of the standard barcodes (*matK, rbcL*, and ITS) and was better compared to *trnH-psbA* (22%). Therefore, it can serve as an alternative where these standard barcodes fail. Considering the cost and time constraints of plastome sequencing, these gene regions are cost-effective and efficient for future population genetic studies on *Debregeasia*, contributing to the expanding dataset for taxon-specific barcodes.

4.2. Species delimitation of Debregeasia

Species delimitation success varied depending on the data and methods used (see details in the supplementary file 1, section 4.2.1), similar to previous studies (Liu et al., 2018; Magoga et al., 2021). Complete plastome data indicated that five of nine

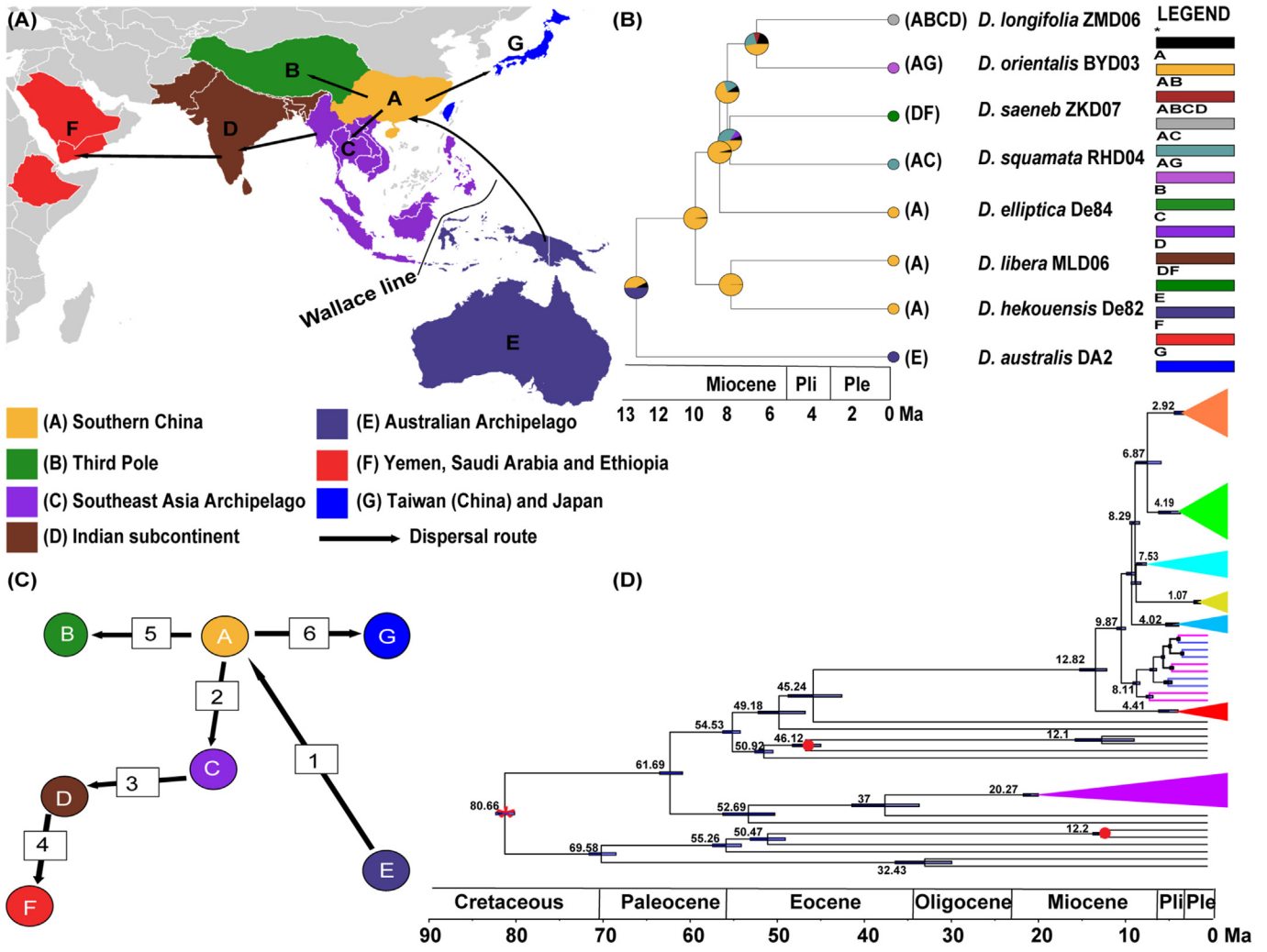


Fig. 7. Biogeographical history of *Debregeasia*. (A–B) Ancestral area reconstruction for *Debregeasia* based on Bayesian Binary MCMC (BBM) implemented in RASP. Ancestral ranges are mapped as follows: A, Southern China; B, Third Pole (Tibetan Plateau and adjacent mountain regions); C, Southeast Asian archipelago; D, Indian subcontinent; E, Australian archipelago; F, Yemen, Saudi Arabia, and Ethiopia; G, Taiwan (China) and Japan. (C) The major dispersal events during the Miocene that contributed to the current distribution of *Debregeasia*. (D) A time-calibrated phylogenetic tree of *Debregeasia*. Median age estimates are derived from treePL analysis, incorporating secondary calibration points (red asterisks) and two fossil records (red circles). Blue bars at nodes represent 95% highest posterior densities (HPD) intervals. Triangles denote monophyletic clades comprising a single species, with color codes corresponding to those in Fig. 4.

Debregeasia species are monophyletic, with three species, *D. hekouensis, D. libera*, and *D. longifolia*, polyphyletic; the remaining species, *D. wallichiana* should be removed from the genus. However, phylogenies based on the universal set of 353 nuclear genes proved more powerful in recovering six monophyletic species, including *D. longifolia*. In phylogenetic trees based on both plastome and single-copy nuclear genes, *D. wallichiana* formed a clade with *Coussapoa contorta* and *Maoutia setosa* with maximum support (Figs. S2–S4) and was hence separated from all other *Debregeasia* species. Therefore, *Debregeasia*, as currently circumscribed, is polyphyletic; this is a new finding, as previous studies did not include *D. wallichiana* (Wu et al., 2013, 2018; Wang et al., 2020a). The taxonomic treatment of *D. wallichiana* has been unstable: first described as *Missiessya wallichiana* (Weddell, 1854), it was later placed in both *Debregeasia* and *Morocarpus,* the latter of which was later synonymized into *Debregeasia* (Wilmot-Dear, 1988). *Debregeasia wallichiana* has very similar morphology to other *Debregeasia* species, such as the cordate leaf base and the greenish-gray color of the abaxial leaf blade, as noted in the *Flora of China* (Chen et al., 2003). However, it differs in leaf shape, leaf margin, and leaf tip traits, which are key diagnostic characters in *Debregeasia*. Although we did not sample the full *D. wallichiana* species range due to sampling challenges around the southern Himalaya, the seven individuals examined did cover a wide geographical range (Fig. 2). Therefore, we strongly recommend that *D. wallichiana* be removed from *Debregeasia,* and restored to *Morocarpus* as *M. wallichiana.*

Further, ASR revealed key synapomorphies for four monophyletic species, including three traits in *Debregeasiaelliptica* (rounded leaf base, finely denticulate leaf margin, and elliptic leaf shape), two each in *D. saeneb* (thickly snow-white tomentose and sparsely spreading hirtellous), and in *D. squamata* (cordate leaf base and fleshy spreading protuberances mixed with appressed hair) and one trait in *D. wallichiana* (shortly acuminate leaf tip). In contrast, the complex of *D. hekouensis,* and *D. libera* showed rampant overlapping morphological traits with similar leaf base, tip, margin, and abaxial blade color. These two species were already known to share many similarities in general growth, habit, and reproductive structures (Chen et al., 2003). *D.ebregeasia libera* and *D. hekouensis* have similar leaf shape (broadly lanceolate) (Fig. 1)

and can only be distinguished based on inflorescence length, with *D. hekouensis* being longer. Therefore, discrimination of these two species is morphologically complex, and they might not be truly distinct, in which case no amount of barcode data will separate them. Case in point, plastome, nrDNA sequences, and the universal set of 353 nuclear genes were unable to distinguish between these species. To fully test their status might require the additional power offered by sequencing more of the nuclear genome and samples, or perhaps all of it e.g. (Wu et al., 2024) , to better reflect their evolutionary history. For instance, taxon-specific datasets, such as the whole genome (Gu et al., 2023), have been helpful in resolving critical species complexes) and could be tried here. Phenotypic clustering and ancestral state reconstruction indicate that no single character of *Debregeasia's* main traits is a useful tool for classifying this species complex. Hence, the relationship and boundaries of these two species should be one of the priorities for future research in *Debregeasia*.

Our data indicated that *Debregeasia* began diversifying in the mid-Miocene (12.82 Ma) (Fig. 7), with present distributions arising from at least six dispersal events. Southern China is identified as a key source of dispersal events into Southeast Asia, the Indian subcontinent, parts of Taiwan (China) and Japan, and Africa. Our dates for *Debregeasia* are generally older than those of previous studies that did not include *D. australis* (Wu et al., 2018). Insufficient sampling within a group can result in an underestimation of the crown age of the plant group (Lu et al., 2020). The convergence of land masses after the mid-Miocene is likely to have facilitated the dispersal of *Debregeasia* throughout Southeast Asia, with the Wallacean islands potentially serving as stepping-stones among Southeast Asian islands (Wu et al., 2019; Low et al., 2021). Additionally, the geological stability of Southeast Asia during this period suggests that the dispersal may have been facilitated by the intensification of the East Asian summer monsoon between 15 and 10 Ma (An et al., 2001). Asian monsoons have played a significant role in the diversification patterns of several plant groups, including *Primulina* (Gesneriaceae) (Kong et al., 2017), Melocanninae (Zhou et al., 2022), *Oreocnide* (Urticaceae) (Wu et al., 2022), and *Cymbidium* (Orchidaceae) (Chen et al., 2024). Further investigation into the mechanisms behind the long-distance dispersal of *Debregeasia* from Asia through Southeast Asia to Australia is necessary to understand the potential asymmetry in this exchange (Peng et al., 2021). It is, therefore, highly likely that long-distance dispersal is the main factor influencing the present distribution patterns of *Debregeasia*, with seeds possibly being dispersed across oceans over considerable distances (Wu et al., 2018). Also, analysis of reticulate evolution showed that hybridization might have been pervasive in *Debregeasia,* influenced perhaps by speciation patterns and overlaps between species ranges, contributing to the complex evolutionary history of the genus.

4.3. Phylogenetic resolution and reticulate evolution in Debregeasia

Phylogenomic analyses based on single-copy nuclear genes have provided valuable insights into the phylogenetic relationships of many plant groups where plastid genomes alone failed (Zhang et al., 2021; Gu et al., 2024; Yan et al., 2024). In this study, we utilized 352 single-copy nuclear genes to resolve phylogenetic relationships in *Debregeasia*. Our results recovered six monophyletic clades with stronger support than plastome and nrDNA sequences. Notably, neither the plastome nor the nrDNA sequences resolved any of *D. hekouensis, D. libera,* or *D. longifolia* as monophyletic, which was partially resolved in the universal set of 353 nuclear genes. These instead formed a mixed clade together but with differing relationships among individual accessions (Fig. 4). Such cyto-nuclear discordance can be associated with hybridization, introgression, and incomplete lineage sorting (Xiang et al., 2015; Folk et al., 2017; Loiseau et al., 2021). Coalescent simulation analysis revealed that almost none of the simulated trees supported the plastid topology Fig. 6B. This implies that ILS alone cannot account for the cyto-nuclear discordance observed in *Debregeasia*, which, therefore, must be driven by gene flow and plastid capture.

The overlapping geographical distributions and similar flowering and fruiting times of *Debregeasia* species suggest that species within the genus have adequate opportunity for interspecific gene flow. For example, *D. longifolia, D. hekouensis*, and *D. libera* are sympatric in southern Yunnan and show high conflict and gene tree discordance, as noted above, which might reflect introgression. Hybridization is common among angiosperms (Mallet, 2007) and can blur species boundaries (Petit and Excoffier, 2009; Hollingsworth et al., 2016). Introgression is plausible because *D. hekouensis, D. libera,* and *D. longifolia* occur sympatrically (Fig. 2) with overlapping flowering time (August–December), and where this occurs, no single barcode may be effective at species delimitation. The skmer analysis resolved *D. libera* and *D. hekouensis* as a single intermixed clade, sister to a clade of *D. longifolia* material from southern Yunnan. Sister to these was a clade comprising all other *D. longifolia* material, which was all from northwestern Yunnan. These two regions differ substantially in climate, with the southern region much warmer and wetter, which might explain the evolutionary patterns behind the two clades. However, clade membership varies according to analysis and region (Fig. 4), so further comprehensive analysis with wide sampling and new genomic data is needed to investigate plastome capture and reticulation patterns among these three species.

5. Conclusions

We examined plastome structural variations and investigated the efficacy of plastome, nrDNA, and single-copy nuclear genes for resolving species boundaries and evolutionary relationships based on within and between species variation in *Debregeasia*. Our study identified varying genetic and morphology-based species boundaries. We revealed multiple cases of overlapping diagnostic characters in *Debregeasia* and uncovered key synapomorphic traits for *D. elliptica*, *D. wallichiana*, *D. saeneb*, and *D. squamata*. Compared to the standard barcodes, the complete plastome proved to be an effective plant ultra-barcode for distinguishing closely related plant species. Full discrimination was not achieved, almost certainly because three species were genuinely non-monophyletic, with accessions intermixed for both plastome and nrDNA. However, single-copy nuclear genes resolved this complex and suggested six monophyletic clades in *Debregeasia*, excluding *D. wallichiana,* which our data indicates should be removed from *Debregeasia.* Our results indicate that hybridization might have been pervasive in *Debregeasia,* influenced perhaps by speciation patterns and overlaps between species ranges, providing an exciting opportunity for future studies to investigate further reticulation events in this group. In addition, estimated divergence times of *Debregeasia* suggest that long-distance dispersal has played an important role in forming current intercontinental distributions, with seeds dispersed across oceans. This study serves as an exemplary case for the utility of recently proposed deep genome skimming methods to retrieve nuclear and organelle genes, enabling robust phylogenetic and biogeographic reconstructions. Additionally, our study demonstrates that extensive exploration of nuclear genes can circumvent the challenges faced by ultra-barcodes and can pave the way for clarifying species relationships in evolutionarily intractable taxa.

CRediT authorship contribution statement

Amos Kipkoech: Writing – review & editing, Writing – original draft, Methodology, Formal analysis. **Ke Li:** Writing – review & editing, Methodology, Formal analysis. **Richard I. Milne:** Writing – review & editing. **Oyetola Olusegun Oyebanji:** Writing – review & editing, Formal analysis. **Moses C. Wambulwa:** Writing – review & editing. **Xiao-Gang Fu:** Writing – review & editing. **Dennis A. Wakhungu:** Writing – review & editing. **Zeng-Yuan Wu:** Writing – review & editing, Supervision, Funding acquisition, Data curation, Conceptualization. **Jie Liu:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Data availability

All DNA sequences generated in this study have been submitted to China National GeneBank for accession numbers presented in Tables S2 and S6.

Authors statement

All co-authors have reviewed and approved the final version of the manuscript, and those entitled to authorship have been named accordingly. This paper contains original content, and no other studies employing this database have been submitted elsewhere. The document includes all of the article's contents, raw data information, and protocols.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We extend our gratitude to Mr. Tao Liu and Mr. Xuewen Liu for their effort in field sampling. We are sincerely thankful to Prof. Chia-Jui Chen, Prof. De-Zhu Li, Prof. Ran Nathan, Associate Prof. Ya-Huang Luo and Prof. Debabrata Maity for their valuable advice and insightful comments during this study. Special thanks to the Royal Botanic Garden Edinburgh (E) and the Royal Botanic Gardens, Kew (K), for providing DNA material. This research was funded by the National Natural Science Foundation of China (42171071, 42211540718), Key Research Program of Frontier Sciences, CAS (ZDBS-LY-7001), Yunnan Fundamental Research Projects (202401AT070190, 202201BC070001), Top-notch Young Talents Project of Yunnan Provincial "Ten Thousand Talents Program" (YNWR-QNBJ-2018-146, YNWR-QNBJ-2020-293), and CAS "Light of West China" Program. Jie Liu and Zeng-Yuan Wu were also supported by the China Scholarship Council (202304910135 and 202304910138) for a one-year study at the University of Toronto, Canada. Molecular experiments were conducted at the Laboratory of Molecular Biology and data analysis was partially facilitated by the iFlora HPC Center (iFlora High-Performance Computing Center), Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pld.2024.11.004.

References

Abràmoff, M.D., Magalhães, P.J., Ram, S.J., 2004. Image processing with ImageJ. Biophot. Int. 11, 36–42.
Almubayedh, H., Ahmad, R., 2019. Ethnopharmacological uses, phytochemistry, biological activities of *Debregeasia salicifolia*: a review. J. Ethnopharmacol. 231, 179–186. https://doi.org/10.1016/j.jep.2018.11.023.
Amenu, S.G., Wei, N., Wu, L., et al., 2022. Phylogenomic and comparative analyses of Coffeeae alliance (Rubiaceae): deep insights into phylogenetic relationships and plastome evolution. BMC Plant Biol. 22, 88. https://doi.org/10.1186/s12870-022-03480-5.
Amiryousefi, A., Hyvönen, J., Poczai, P., 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics 34, 3030–3031. https://doi.org/10.1093/bioinformatics/bty220.
An, Z., Kutzbach, J.E., Prell, W.L., et al., 2001. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. Nature 411, 62–66. https://doi.org/10.1038/35075035.
Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B., 2017. Julia: a fresh approach to numerical computing. SIAM Rev. 59, 65–98. https://doi.org/10.1137/141000671.
Bouckaert, R., Heled, J., Kühnert, D., et al., 2014. Beast 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537.
Bruun-Lund, S., Clement, W.L., Kjellberg, F., et al., 2017. First plastid phylogenomic study reveals potential cyto-nuclear discordance in the evolutionary history of *Ficus* L. (Moraceae). Mol. Phylogenet. Evol. 109, 93–104. https://doi.org/10.1016/j.ympev.2016.12.031.
CBOL Plant Working Group, 2009. A DNA barcode for land plants. Proc. Natl. Acad. Sci. U.S.A. 106, 12794–12797. https://doi.org/10.1073/pnas.0905845106.
Chen, C.J., Lin, Q., Friis, I., et al., 2003. Urticaceae. In: Wu, Z.Y., Raven, P.H. (Eds.), Flora of China, vols. 76–1. Science Press, St. Petersburg. St. Louis: Missouri Botanical Garden Press, Beijing, pp. 107–111.
Chen, H.Y., Zhang, Z.R., Yao, X., et al., 2024. Plastid phylogenomics provides new insights into the systematics, diversification, and biogeography of *Cymbidium* (Orchidaceae). Plant Divers. 46, 448–461. https://doi.org/10.1016/j.pld.2024.03.001.
Coissac, E., Hollingsworth, P.M., Lavergne, S., et al., 2016. From barcodes to genomes: extending the concept of DNA barcoding. Mol. Ecol. 25, 1423–1428. https://doi.org/10.1111/mec.13549.
Conn, B.J., Hadiah, J.T., 2009. Nomenclature of tribes within the Urticaceae. Kew Bull. 64, 349–352. https://doi.org/10.1007/s12225-009-9108-4.
Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss, and rearrangement. PLoS One 5, e11147. https://doi.org/10.1371/journal.pone.0011147.
Darriba, D., Taboada, G.L., Doallo, R., et al., 2012. jModelTest 2: more models, new heuristics, and high-performance computing. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109. PMID:22847109;PMCID:PMC4594756.
Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf material. Phytochem. Bull. 19, 11–15.
Folk, R.A., Mandel, J.R., Freudenstein, J.V., 2017. Ancestral gene flow and parallel organellar genome capture result in extreme phylogenomic discord in a lineage of angiosperms. Syst. Biol. 66, 320–337. https://doi.org/10.1093/sysbio/syw083.
Frazer, K.A., Pachter, L., Poliakov, A., et al., 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32, W273–W279. https://doi.org/10.1093/nar/gkh458.
Fu, C.-N., Wu, C.-S., Ye, L.-J., et al., 2019. Prevalence of isomeric plastomes and effectiveness of plastome super-barcodes in yews (*Taxus*) worldwide. Sci. Rep. 9, 2773. https://doi.org/10.1038/s41598-019-39161-x.
Fu, C.-N., Mo, Z.Q., Yang, J.B., et al., 2022. Testing genome skimming for species discrimination in the large and taxonomically difficult genus *Rhododendron*. Mol. Ecol. Resour. 22, 404–414. https://doi.org/10.1111/1755-0998.13479.
Fu, Q.-L., Mo, Z.-Q., Xiang, X.-G., et al., 2023. Plastome phylogenomics and morphological traits analyses provide new insights into the phylogenetic position, species delimitation and speciation of *Triplostegia* (Caprifoliaceae). BMC Plant Biol. 23, 645. https://doi.org/10.1186/s12870-023-04663-4.
Gu, T.T., Wu, H., Yang, F., et al., 2023. Genomic analysis reveals a cryptic pangolin species. Proc. Natl. Acad. Sci. U.S.A. 120, 40. https://doi.org/10.1073/pnas.2304096120.
Gu, W., Zhang, T., Liu, S.Y., et al., 2024. Phylogenomics, reticulation, and biogeographical history of Elaeagnaceae. Plant Divers. 46, 683–697. https://doi.org/10.1016/j.pld.2024.07.001.
Guo, C., Ma, P.-F., Yang, G.-Q., et al., 2021. Parallel ddRAD and genome skimming analyses reveal a radiative and reticulate evolutionary history of the temperate bamboos. Syst. Biol. 70, 756–773. https://doi.org/10.1093/sysbio/syaa076.
Hadiah, J.T., Conn, B.J., Quinn, C.J., 2008. Infra-familial phylogeny of Urticaceae, using chloroplast sequence data. Aust. Syst. Bot. 21, 375–385. https://doi.org/10.1071/SB08041.
Harris, J.G., Harris, M.W., 1994. Plant Identification Terminology: an Illustrated Glossary. Spring lake publishing, Utah, p. 198.
Hebert, P.D., Cywinska, A., Ball, S.L., et al., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B-Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218.
Hollingsworth, P.M., Li, D.-Z., van der Bank, M., et al., 2016. Telling plant species apart with DNA: from barcodes to genomes. Phil. Trans. Biol. Sci. 371, 20150338. https://doi.org/10.1098/rstb.2015.0338.

Hong, D.-Y., 2016. Biodiversity pursuits need a scientific and operative species concept. Biodivers. Sci. 24, 979. https://doi.org/10.17520/biods.2016203.
Ji, Y., Liu, C., Yang, Z., et al., 2019. Testing and using complete plastomes and ribosomal DNA sequences as the next generation DNA barcodes in *Panax* (Araliaceae). Mol. Ecol. Resour. 19, 1333–1345. https://doi.org/10.1111/1755-0998.13050.
Jiang, Y., Yang, J., Folk, R.A., et al., 2024. Species delimitation of tea plants (*Camellia* sect. *Thea*) based on super-barcodes. BMC Plant Biol. 24, 181. https://doi.org/10.1186/s12870-024-04882-3.
Jin, J.-J., Yu, W.-B., Yang, J.-B., et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21, 1–31. https://doi.org/10.1186/s13059-020-02154-5.
Johnson, M.G., Pokorny, L., Dodsworth, S., et al., 2019. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. Syst. Biol. 68, 594–606. https://doi.org/10.1093/sysbio/syy086.
Kane, N., Sveinsson, S., Dempewolf, H., et al., 2012. Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. Am. J. Bot. 99, 320–329. https://doi.org/10.3732/ajb.1100570.
Kapli, P., Lutteropp, S., Zhang, J., et al., 2017. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. Bioinformatics 33, 1630–1638. https://doi.org/10.1093/bioinformatics/btx025.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780. https://doi.org/10.1093/molbev/mst010.
Kearse, M., Moir, R., Wilson, A., et al., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.
Kim, C., Deng, T., Chase, M., et al., 2015. Generic phylogeny and character evolution in Urticeae (Urticaceae) inferred from nuclear and plastid DNA regions. Taxon 64, 65–78. https://doi.org/10.12705/641.20.
Kong, H.H., Condamine, F.L., Harris, A.J., et al., 2017. Both temperature fluctuations and East Asian monsoons have driven plant diversification in the karst ecosystems from southern China. Mol. Ecol. 26, 6414–6429. https://doi.org/10.1111/mec.14367.
Li, D.Z., Gao, L.M., Li, H.T., et al., 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. Proc. Natl. Acad. Sci. U.S.A. 108, 19641–19646. https://doi.org/10.1073/pnas.1104551108.
Li, D., Liu, C.M., Luo, R., et al., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. Bioinformatics 31, 1674–1676. https://doi.org/10.1093/bioinformatics/btv033.
Li, J., Tang, J., Zeng, S., et al., 2021. Comparative plastid genomics of four *Pilea* (Urticaceae) species: insight into interspecific plastid genome diversity in *Pilea*. BMC Plant Biol. 21, 1–13. https://doi.org/10.1186/s12870-020-02793-7.
Li, K., Zhang, H., Shi, M., et al., 2022. The complete chloroplast genome sequence of *Urtica fissa*. Mitochondrial DNA B 7, 1005–1007. https://doi.org/10.1080/23802359.2022.2080017.
Li, W., Liu, Y., Yang, Y., et al., 2018. Interspecific chloroplast genome sequence diversity and genomic resources in *Diospyros*. BMC Plant Biol. 18, 1–11. https://doi.org/10.1186/s12870-018-1421-3.
Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452. https://doi.org/10.1093/bioinformatics/btp187.
Liu, B.B., Ren, C., Kwak, M., et al., 2022. Phylogenomic conflict analyses in the apple genus *Malus* s.l. reveal widespread hybridization and allopolyploidy driving diversification, with insights into the complex biogeographic history in the Northern Hemisphere. J. Integr. Plant Biol. 64, 1020–1043. https://doi.org/10.1111/jipb.13246.
Liu, J., Milne, R.I., Möller, M., et al., 2018. Integrating a comprehensive DNA barcode reference library with a global map of yews (*Taxus* L.) for forensic identification. Mol. Ecol. Resour. 18, 1115–1131. https://doi.org/10.1111/1755-0998.12903.
Liu, J., Moeller, M., Gao, L.M., et al., 2011. DNA barcoding for the discrimination of Eurasian yews (*Taxus* L., Taxaceae) and the discovery of cryptic species. Mol. Ecol. Resour. 11, 89–100. https://doi.org/10.1111/j.1755-0998.2010.02907.x.
Loiseau, O., Mota Machado, T., Paris, M., et al., 2021. Genome skimming reveals widespread hybridization in a Neotropical flowering plant radiation. Front. Ecol. Evol. 9, 668281. https://doi.org/10.3389/fevo.2021.668281.
Low, S.L., Yu, C.C., Ooi, I.H., et al., 2021. Extensive Miocene speciation in and out of Indochina: the biogeographic history of *Typhonium sensu stricto* (Araceae) and its implication for the assembly of Indochina flora. J. Syst. Evol. 59, 419–428. https://doi.org/10.1111/jse.12689.
Lu, L.M., Hu, H.H., Peng, D.X., et al., 2020. Noise does not equal bias in assessing the evolutionary history of the angiosperm flora of China: a response to Qian (2019). J. Biogeogr. 47, 2286–2291. https://doi.org/10.1111/jbi.13947.
Maddison, W.P., 2007. Mesquite: a modular system for evolutionary analysis. Version 2.0. https://cir.nii.ac.jp/all?q=http://mesquiteproject.org.
Magoga, G., Fontaneto, D., Montagna, M., 2021. Factors affecting the efficiency of molecular species delimitation in a species-rich insect family. Mol. Ecol. Resour. 21, 1475–1489. https://doi.org/10.1111/1755-0998.13352.
Mallet, J., 2007. Hybrid speciation. Nature 446, 279–283. https://doi.org/10.1038/nature05706.
Maurin, K.J.L., 2020. An empirical guide for producing a dated phylogeny with treePL in a maximum likelihood framework. arXiv, 2008.07054v2. https://arxiv.org/pdf/2008.07054.
Mishra, P., Kumar, A., Nagireddy, A., et al., 2016. DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant Biotechnol. J. 14, 8–21. https://doi.org/10.1111/pbi.12419.
Morales-Briones, D.F., Kadereit, G., Tefarikis, D.T., et al., 2021. Disentangling sources of gene tree discordance in phylogenomic data sets: testing ancient hybridizations in Amaranthaceae s.l. Syst. Biol. 70, 219–235. https://doi.org/10.1093/sysbio/syaa066.
Ogoma, C.A., Liu, J., Stull, G.W., et al., 2022. Deep insights into the plastome evolution and phylogenetic relationships of the tribe Urticeae (Family Urticaceae). Front. Plant Sci. 13, 870949. https://doi.org/10.3389/fpls.2022.870949.
Ortiz, E.M., Hoewener, A., Shigita, G., et al., 2023. A novel phylogenomics pipeline reveals complex pattern of reticulate evolution in Cucurbitales. bioRxiv. https://doi.org/10.1101/2023.10.27.564367, 2023-10.
Oyebanji, O., Zhang, R., Chen, S.-Y., et al., 2020. New insights into the plastome evolution of the Millettioid/Phaseoloid clade (Papilionoideae, Leguminosae). Front. Plant Sci. 11, 151. https://doi.org/10.3389/fpls.2020.00151.
Park, S., An, B., Park, S., 2018. Reconfiguration of the plastid genome in *Lamprocapnos spectabilis*: IR boundary shifting, inversion, and intraspecific variation. Sci. Rep. 8, 13568. https://doi.org/10.1038/s41598-018-31938-w.
Peng, D.X., Dang, V.C., Habib, S., et al., 2021. Historical biogeography of *Tetrastigma* (Vitaceae): insights into floristic exchange patterns between Asia and Australia. Cladistics 37, 803–815. https://doi.org/10.1111/cla.12462.
Petit, R.J., Excoffier, L., 2009. Gene flow and species delimitation. Trends Ecol. Evol. 24, 386–393. https://doi.org/10.1016/j.tree.2009.02.011.
Plants of the World Online, 2024. Royal Botanic Gardens. Kew. Available at: https://powo.science.kew.org/. (Accessed 13 March 2024).
Pons, J., Barraclough, T.G., Gomez-Zurita, J., et al., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Syst. Biol. 55, 595–609. https://doi.org/10.1080/10635150600852011.
Puillandre, N., Brouillet, S., Achaz, G., 2021. ASAP: assemble species by automatic partitioning. Mol. Ecol. Resour. 21, 609–620. https://doi.org/10.1111/1755-0998.13281.
Puillandre, N., Lambert, A., Brouillet, S., et al., 2012. ABGD, automatic barcode gap discovery for primary species delimitation. Mol. Ecol. 21, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x.
R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. http://www.r-project.org.
Rambaut, A., 2009. FigTree. Inst. Evol. Biol. Univ. Edinburgh, Edinburgh, UK. http://tree.bio.ed.ac.uk/software/figtree/.
Ravinet, M., Faria, R., Butlin, R.K., et al., 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. J. Evol. Biol. 30, 1450–1477. https://doi.org/10.1111/jeb.13047.
Ren, B.Q., Xiang, X.G., Chen, Z.D., 2010. Species identification of *Alnus* (Betulaceae) using nrDNA and cpDNA genetic markers. Mol. Ecol. Resour. 10, 594–605. https://doi.org/10.1111/j.1755-0998.2009.02815.x.
Sarmashghi, S., Bohmann, K., Gilbert, P., et al., 2019. Skmer: Assembly-free and alignment-free sample identification using genome skims. Genome Biol. 20, 1–20. https://doi.org/10.1186/s13059-019-1632-4.
Schattner, P., Brooks, A.N., Lowe, T.M., 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. 33, W686–W689. https://doi.org/10.1093/nar/gki366.
Shen, Z., Lu, T., Zhang, Z., et al., 2019. Authentication of traditional Chinese medicinal herb "Gusuibu" by DNA-based molecular methods. Ind. Crops Prod. 141, 111756. https://doi.org/10.1016/j.indcrop.2019.111756.
Singh, H.K., Parveen, I., Raghuvanshi, S., et al., 2012. The loci recommended as universal barcodes for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA barcoding of *Dendrobium* species. BMC Res. Notes 5, 1–11. https://doi.org/10.1186/1756-0500-5-42.
Ślipiko, M., Myszczyński, K., Buczkowska, K., et al., 2020. Molecular delimitation of European leafy liverworts of the genus *Calypogeia* based on plastid super-barcodes. BMC Plant Biol. 20, 1–15. https://doi.org/10.1186/s12870-020-02435–y.
Smith, S.A., Moore, M.J., Brown, J.W., Yang, Y., 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. BMC Evol. Biol. 15, 1–15. https://doi.org/10.1186/s12862-015-0423-0.
Smith, S.A., O'Meara, B.C., 2012. treePL: divergence time estimation using penalized likelihood for large phylogenies. Bioinformatics 28, 2689–2690. https://doi.org/10.1093/bioinformatics/bts492.
Solís-Lemus, C., Bastide, P., Ané, C., 2017. PhyloNetworks: a package for phylogenetic networks. Mol. Biol. Evol. 34, 3292–3298. https://doi.org/10.1093/molbev/msx235.
Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0b10. Sinauer Associates, Sunderland.
Tong, R., Gui, C., Zhang, Y., et al., 2022. Phylogenomics, plastome structure and species identification in *Mahonia* (Berberidaceae). BMC Genomics 23, 766. https://doi.org/10.1186/s12864-022-08964-0.
Tseng, Y.-H., Monro, A.K., Wei, Y.-G., et al., 2019. Molecular phylogeny and morphology of *Elatostema s.l.* (Urticaceae): implications for inter-and infrageneric classifications. Mol. Phylogenet. Evol. 132, 251–264. https://doi.org/10.1016/j.ympev.2018.11.016.

Wang, H.-X., Liu, H., Moore, M.J., et al., 2020a. Plastid phylogenomic insights into the evolution of the Caprifoliaceae *s.l.* (Dipsacales). Mol. Phylogenet. Evol. 142, 106641. https://doi.org/10.1016/j.ympev.2019.106641.
Wang, J., Fu, C.-N., Yang, J.-B., et al., 2022. Testing the complete plastome for species discrimination, cryptic species discovery, and phylogenetic resolution in *Cephalotaxus* (Cephalotaxaceae). Front. Genet. 13, 768810. https://doi.org/10.3389/fpls.2022.768810.
Wang, R.-N., Milne, R.I., Du, X.-Y., et al., 2020b. Characteristics and mutational hotspots of plastomes in *Debregeasia* (Urticaceae). Front. Genet. 11, 527033. https://doi.org/10.3389/fgene.2020.00729.
Wang, W., 2016. Two new species of Urticaceae from China. Bull. Bot. Res. 36, 481–483. https://doi.org/10.7525/j.issn.1673-5102.2016.04.001.
Weddell, H.A., 1854. Revue de la famille des Urticacées. Ann. Sci. Nat. Bot. Sér. 4, 173–212.
Weddell, H.A., 1869. Urticaceae. In: Candolle, A. de (Ed.), Prodromus Systematis Naturalis Regnis Vegetabilis. V. Masson, Paris, pp. 32–235.
Wick, R.R., Schultz, M.B., Zobel, J., et al., 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31, 3350–3352. https://doi.org/10.1093/bioinformatics/btv383.
Wilmot-Dear, C., 1988. An account of the genus *Debregeasia* (Urticaceae-Boehmerieae). Kew Bull. 673–692. https://doi.org/10.2307/4129966.
Wilmot-Dear, C., 1994. *Debregeasia ceylanica* (Urticaceae): a change of status. Kew Bull. 49, 468. https://doi.org/10.2307/4114470.
Wilmot-Dear, C., Friis, I., 2012. *Debregeasia australis* sp. nov. (Urticaceae), with a new synopsis of and a new key to the genus. Edinb. J. Bot. 69, 301–311. https://doi.org/10.1017/S096042861200011X.
Wu, X., Liu, X., Kodrul, T., et al., 2019. *Dacrycarpus* pattern shedding new light on the early floristic exchange between Asia and Australia. Natl. Sci. Rev. 6, 1086–1090. https://doi.org/10.1093/nsr/nwz060.
Wu, Z.Y., Chapman, M.A., Liu, J., et al., 2024. Genomic variation, environmental adaptation, and feralization in ramie, an ancient fiber crop. Plant Commun. 5, 100942. https://doi.org/10.1016/j.xplc.2024.100942.
Wu, Y., Hipp, A.L., Fargo, G., et al., 2023. Improving species delimitation for effective conservation: a case study in the endemic maple-leaf oak (*Quercus acerifolia*). New Phytol. 238, 1278–1293.
Wu, Z.-Y., Liu, J., Provan, J., et al., 2018. Testing Darwin's transoceanic dispersal hypothesis for the inland nettle family (Urticaceae). Ecol. Lett. 21, 1515–1529. https://doi.org/10.1111/ele.13132.
Wu, Z.-Y., Milne, R.I., Chen, C.-J., et al., 2015. Ancestral state reconstruction reveals rampant homoplasy of diagnostic morphological characters in Urticaceae, conflicting with current classification schemes. PLoS One 10, e0141821. https://doi.org/10.1371/journal.pone.0141821.
Wu, Z.-Y., Milne, R.I., Liu, J., et al., 2022. Phylogenomics and evolutionary history of *Oreocnide* (Urticaceae) shed light on recent geological and climatic events in SE Asia. Mol. Phylogenet. Evol. 175, 107555. https://doi.org/10.1016/j.ympev.2022.107555.
Wu, Z.-Y., Monro, A.K., Milne, R.I., et al., 2013. Molecular phylogeny of the nettle family (Urticaceae) inferred from multiple loci of three genomes and extensive generic sampling. Mol. Phylogenet. Evol. 69, 814–827. https://doi.org/10.1016/j.ympev.2013.06.022.
Xiang, Q.-P., Wei, R., Shao, Y.-Z., et al., 2015. Phylogenetic relationships, possible ancient hybridization, and biogeographic history of *Abies* (Pinaceae) based on data from nuclear, plastid, and mitochondrial genomes. Mol. Phylogenet. Evol. 82, 1–14. https://doi.org/10.1016/j.ympev.2014.10.008.
Xu, Y.-L., Shen, H.-H., Du, X.-Y., et al., 2022. Plastome characteristics and species identification of Chinese medicinal wintergreens (*Gaultheria*, Ericaceae). Plant Divers. 44, 519–529. https://doi.org/10.1016/j.pld.2022.06.002.
Yan, Y., da Fonseca, R.R., Rahbek, C., et al., 2024. A new nuclear phylogeny of the tea family (Theaceae) unravels rapid radiations in genus *Camellia*. Mol. Phylogenet. Evol. 196, 108089. https://doi.org/10.1016/j.ympev.2024.108089.
Yang, J.-B., Tang, M., Li, H.-T., et al., 2013. Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. BMC Evol. Biol. 13, 1–12. https://doi.org/10.1186/1471-2148-13-84.
Yang, Y., Sun, P., Lv, L., et al., 2020. Prickly waterlily and rigid hornwort genomes shed light on early angiosperm evolution. Nat. Plants 6, 215–222. https://doi.org/10.1038/s41477-020-0594-6.
Yu, W.-B., Liu, M.-L., Wang, H., et al., 2015. Towards a comprehensive phylogeny of the large temperate genus *Pedicularis* (Orobanchaceae), with an emphasis on species from the Himalaya-Hengduan Mountains. BMC Plant Biol. 15, 1–14. https://doi.org/10.1186/s12870-015-0547-9.
Yu, X.Q., Jiang, Y.Z., Folk, R.A., et al., 2022. Species discrimination in *Schima* (Theaceae): next-generation super-barcodes meet evolutionary complexity. Mol. Ecol. Resour. 22, 3161–3175. https://doi.org/10.1111/1755-0998.13683.
Yu, Y., Blair, C., He, X., 2020. Rasp 4: ancestral state reconstruction tool for multiple genes and characters. Mol. Biol. Evol. 37, 604–606. https://doi.org/10.1093/molbev/msz257.
Zan, T., He, Y.-T., Zhang, M., et al., 2023. Phylogenomic analyses of *Camellia* support reticulate evolution among major clades. Mol. Phylogenet. Evol. 182, 107744. https://doi.org/10.1016/j.ympev.2023.107744.
Zhang, C., Huang, C.H., Liu, M., et al., 2021. Phylotranscriptomic insights into Asteraceae diversity, polyploidy, and morphological innovation. J. Integr. Plant Biol. 63, 1273–1293. https://doi.org/10.1111/jipb.13078.
Zhang, C., Rabiee, M., Sayyari, E., et al., 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19, 15–30. https://doi.org/10.1186/s12859-018-2129-y.
Zhang, L., Huang, Y.W., Huang, J.L., et al., 2023. DNA barcoding of *Cymbidium* by genome skimming: Call for next-generation nuclear barcodes. Mol. Ecol. Resour. 23, 424–439. https://doi.org/10.1111/1755-0998.13719.
Zhang, N., Erickson, D.L., Ramachandran, P., et al., 2017. An analysis of *Echinacea* chloroplast genomes: implications for future botanical identification. Sci. Rep. 7, 216. https://doi.org/10.1038/s41598-017-00321-6.
Zheng, S., Poczai, P., Hyvönen, J., et al., 2020. Chloroplot: an online program for the versatile plotting of organelle genomes. Front. Genet. 11, 576124. https://doi.org/10.3389/fgene.2020.576124.
Zhou, M.Y., Liu, J.X., Ma, P.F., et al., 2022. Plastid phylogenomics shed light on intergeneric relationships and spatiotemporal evolutionary history of Melocanninae (Poaceae: Bambusoideae). J. Syst. Evol. 60, 640–652. https://doi.org/10.1111/jse.12843.
Zhou, Q., Lin, C.-W., Ng, W.L., et al., 2019. Analyses of plastome sequences improve phylogenetic resolution and provide new insight into the evolutionary history of Asian Sonerileae/Dissochaeteae. Front. Plant Sci. 10, 485977. https://doi.org/10.3389/fpls.2019.01477.