



OPEN ACCESS

EDITED BY

Wenpan Dong,
Beijing Forestry University, China

REVIEWED BY

Hengchang Wang,
Chinese Academy of Sciences (CAS), China
Yong-Hua Zhang,
Wenzhou University, China

*CORRESPONDENCE

Hong Yu

✉ hongyu@ynu.edu.cn

Li-Bing Zhang

✉ Libing.Zhang@mobot.org

Xin-Mao Zhou

✉ xinmao.zhou@ynu.edu.cn

†These authors share first authorship

RECEIVED 22 March 2024

ACCEPTED 17 June 2024

PUBLISHED 16 July 2024

CITATION

Zhao J, He Z-R, Fang S-L, Han X-K, Jiang L-Y, Hu Y-P, Yu H, Zhang L-B and Zhou X-M (2024) Phylogenomic data resolved the deep relationships of Gymnogynoideae (Selaginellaceae). *Front. Plant Sci.* 15:1405253. doi: 10.3389/fpls.2024.1405253

COPYRIGHT

© 2024 Zhao, He, Fang, Han, Jiang, Hu, Yu, Zhang and Zhou. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Phylogenomic data resolved the deep relationships of Gymnogynoideae (Selaginellaceae)

Jing Zhao^{1†}, Zhao-Rong He^{2†}, Shao-Li Fang^{1†}, Xu-Ke Han¹, Lu-Yao Jiang¹, Yu-Ping Hu¹, Hong Yu^{1*}, Li-Bing Zhang^{3,4*} and Xin-Mao Zhou^{1*}

¹School of Ecology and Environmental Science, Yunnan University, Kunming, Yunnan, China,

²School of Life Sciences, Yunnan University, Kunming, Yunnan, China, ³Missouri Botanical Garden, St. Louis, MO, United States, ⁴Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, Sichuan, China

The unresolved phylogenetic framework within the Selaginellaceae subfamily Gymnogynoideae (ca. 130 species) has hindered our comprehension of the diversification and evolution of Selaginellaceae, one of the most important lineages in land plant evolution. Here, based on plastid and nuclear data extracted from genomic sequencing of more than 90% species of all genera except two in Gymnogynoideae, a phylogenomic study focusing on the contentious relationships among the genera in Gymnogynoideae was conducted. Our major results included the following: (1) Only single-copy region (named NR) and only one ribosomal operon was firstly found in *Afroselaginella* among vascular plants, the plastome structure of Gymnogynoideae is diverse among the six genera, and the direct repeats (DR) type is inferred as the ancestral state in the subfamily; (2) The first strong evidence was found to support *Afroselaginella* as a sister to *Megaloselaginella*. Alternative placements of *Ericetorum* and *Gymnogynum* were detected, and their relationships were investigated by analyzing the variation of phylogenetic signals; and (3) The most likely genus-level relationships in Gymnogynoideae might be: ((*Bryodesma*, *Lepidoselaginella*), (((*Megaloselaginella*, *Afroselaginella*), *Ericetorum*), *Gymnogynum*)), which was supported by maximum likelihood phylogeny based on plastid datasets, maximum likelihood, and Bayesian inference based on SCG dataset and concatenated nuclear and plastid datasets and the highest proportion of phylogenetic signals of plastid genes.

KEYWORDS

conformation, lycophyte, phylogenetic conflict, phylogenetic signal, plastomes

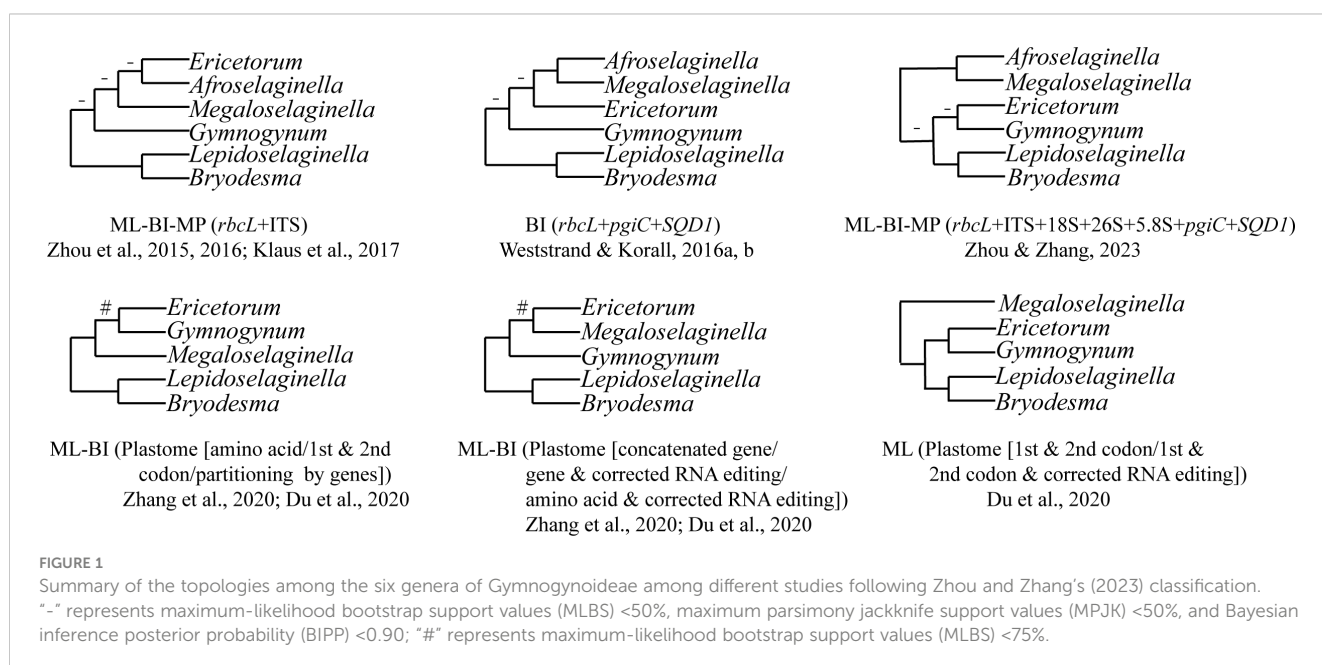
1 Introduction

Selaginellaceae, the largest family in lycophytes, are nearly cosmopolitan and are estimated to contain 700–800 species (Jermy, 1986, 1990; Tryon and Lugardon, 1991; Zhou and Zhang, 2015; PPG I, 2016; Zhou et al., 2016; Weststrand and Korall, 2016a, b; Zhou and Zhang, 2023) but may contain up to 1000 species. Diverged from their closest living relatives ca. 383 million years ago in the Devonian, they were hyper-diverse already in the Cretaceous based on fossil records and molecular dating (Thomas, 1992, 1997; Kenrick and Crane, 1997; Korall et al., 1999; Taylor et al., 2009; Arrigo et al., 2013; Klaus et al., 2017; Morris et al., 2018; Schmidt et al., 2020, 2022). Selaginellaceae are an ideal example of morphological stasis in plants, and both extant species and fossil materials have morphological stability over hundreds of millions of years (Schmidt et al., 2022). Because of the high intraspecific and low interspecific variability, Selaginellaceae are difficult to identify, and how to subdivide the family has frequently been contentious (e.g., Jermy, 1986, 1990; Zhou and Zhang, 2015; Zhou et al., 2016; Weststrand and Korall, 2016a, b; Zhou and Zhang, 2023).

Most recently, using the largest taxon and character sampling and integrating morphological characters, geographical distribution, Zhou and Zhang (2023) recognized seven subfamilies and 19 genera in Selaginellaceae. Here we adopt their new classification and focus on the phylogeny of the subfamily Gymnogynoideae (= “S. subg. *Ericetorum*” sensu Zhou and Zhang, 2015). Following the new classification, Gymnogynoideae include six genera: *Afroselaginella* (= S. sect. *Myosurus* Li Bing Zhang & X.M.Zhou sensu Zhou and Zhang, 2015; = S. subg. *Exaltatae* Weststrand & Korall p.p. sensu Weststrand and Korall, 2016b), *Bryodesma* (= S. sect. *Homoeophyllae* Spring sensu Zhou and Zhang, 2015; = S. subg. *Rupestreae* Weststrand & Korall sensu Weststrand and Korall, 2016b), *Ericetorum* (= S. sect. *Lyallia* (Rothm.) Li Bing Zhang & X.M.Zhou sensu Zhou and Zhang, 2015; = S. subg. *Ericetorum* Jermy sensu Weststrand and Korall, 2016b), *Gymnogynum*

(= S. sect. *Articulatae* (Spring) Li Bing Zhang & X.M.Zhou sensu Zhou and Zhang, 2015; = S. subg. *Gymnogynum* (P.Beauv.) Weststrand & Korall sensu Weststrand and Korall, 2016b), *Lepidoselaginella* (= S. sect. *Lepidophyllae* Li Bing Zhang & X.M.Zhou sensu Zhou and Zhang, 2015; = S. subg. *Lepidophyllae* (Li Bing Zhang & X.M.Zhou) Weststrand & Korall sensu Weststrand and Korall, 2016b), and *Megaloselaginella* (= S. sect. *Megalosporarum* Li Bing Zhang & X.M.Zhou sensu Zhou and Zhang, 2015; = S. subg. *Exaltatae* p.p. sensu Weststrand and Korall, 2016b).

Gymnogynoideae are well characterized by the often dorsal rhizophores and reticulate megaspores (Zhou and Zhang, 2015; Zhou et al., 2016; Zhou and Zhang, 2023). Species of Gymnogynoideae show high morphological diversity in megaspores, microspores, habit, leaves, strobili, etc (Zhou and Zhang, 2015; Zhou et al., 2016; Weststrand and Korall, 2016a; Zhou and Zhang, 2023). Although the monophyly of the six genera in Gymnogynoideae was each well supported phylogenetically and morphologically (Zhou and Zhang, 2015; Zhou et al., 2016; Weststrand and Korall, 2016a; Zhou and Zhang, 2023), the inter-generic relationships have never been fully resolved (Zhou and Zhang, 2015; Zhou et al., 2016; Weststrand and Korall, 2016a; Zhou and Zhang, 2023). The sister relationship between *Lepidoselaginella* and *Bryodesma* has been confirmed in previous phylogenetic studies using different analyses or datasets (Zhou et al., 2016; Weststrand and Korall, 2016a; Klaus et al., 2017; Du et al., 2020; Zhang et al., 2020a; Zhou et al., 2022; Zhou and Zhang, 2023). However, the relationships among the remaining genera were strongly volatile (Figure 1). Zhou et al. (2016) resolved *Gymnogynum* as sister to the remaining genera (except *Lepidoselaginella* and *Bryodesma*) with weak support, and the phylogenetic relationships among *Megaloselaginella*, *Ericetorum*, and *Afroselaginella* were unresolved in maximum likelihood (ML), Bayesian inference (BI), and maximum parsimony (MP) analyses based on the concatenated dataset (*rbcL* and ITS)



(Figure 1). Weststrand and Korall (2016b) also resolved *Gymnogynum* as the second earliest-diverging lineage in BI analysis, but with low support based on the combined three-region dataset (*rbcl* + *pgiC* + *SQD1*), and resolved *Megaloselaginella* as sister to *Afroselaginella* clade with strong support (Figure 1). Recently, the phylogenetic analysis based on plastid *rbcl* and five nuclear markers (18S, 26S, ITS, *SQD1*, and *pgiC*) showed that Gymnogynoideae were comprised by four well-supported clades: *Ericetorum*, *Gymnogynum*, the *Megaloselaginella* + *Afroselaginella* clade, and the *Lepidoselaginella* + *Bryodesma* clade. However, phylogenetic relationships among them were entirely unresolved in maximum likelihood (ML), Bayesian inference (BI), and maximum parsimony (MP) analyses (MPJK and MLBS <50% and BIPP <90%) (Figure 1; Zhou and Zhang, 2023).

In recent years, with the boosting of next-generation sequencing, complete or nearly complete plastomes have been largely used to infer the phylogeny of Selaginellaceae (Du et al., 2020; Zhang et al., 2020a; Zhou et al., 2022). Gymnogynoideae included about 130 species with worldwide distribution (Zhou and Zhang, 2015; Weststrand and Korall, 2016b; Zhou and Zhang, 2023). In previous phylogenomic studies, only fewer than eight species in Gymnogynoideae were investigated (Du et al., 2020; Zhang et al., 2020a; Zhou et al., 2022). It is well known that appropriate and extensive taxon sampling is one of the most important determinants to improve the accuracy of phylogeny inferences (Heath et al., 2008). *Afroselaginella* contains about five species (Zhou and Zhang, 2015; Weststrand and Korall, 2016a, b; Zhou and Zhang, 2023), but this genus was never included in previous phylogenomic studies (Du et al., 2020; Zhang et al., 2020a). The inter-generic relationships in Gymnogynoideae were still unsolved or not well supported even by studies based on plastome data. Analyses of the phylogeny of Selaginellaceae with different derived datasets (e.g., amino acid dataset, the first and second codon sites, concatenated unpartitioned dataset, concatenated datasets partitioned by genes or codon positions, and dataset corrected by RNA-edited sites) and analytical methods (ML/BI) were conducted (Figure 1; e.g., Du et al., 2020; Zhang et al., 2020a; Zhou et al., 2022; Zhou and Zhang, 2023). However, except *Lepidoselaginella* being constantly resolved as sister to *Bryodesma*, the remaining genera usually showed alternative relationships largely caused by the uncertain placement of *Megaloselaginella*. *Megaloselaginella* was resolved as either sister to the *Gymnogynum* + *Ericetorum* clade (Du et al., 2020: CDS, AA) or sister to the *Lepidoselaginella* + *Bryodesma* + *Gymnogynum* + *Ericetorum* clade (Du et al., 2020: codons 1 and 2) or sister to *Ericetorum* (Du et al., 2020: CDSr, AAr). Using the same data processing of codons 1 and 2, Zhang et al. (2020) resolved *Megaloselaginella* as sister to the *Gymnogynum* + *Ericetorum* clade, which was different from the results of Du et al. (2020). In those analyses, the inter-generic relationships received weak to strong support (ML bootstrap support—MLBS = 50% – 100%/Bayesian inference posterior probability—BIPP = 0.73–1) (Du et al., 2020; Zhang et al., 2020a). Different inference methods, samplings, and strategies of data processing strongly affected the resolutions in Gymnogynoideae (Figure 1).

In this study, based on the largest taxon sampling so far, we used plastome data and 140 single-copy nuclear genes to further explore

the evolutionary relationships in Gymnogynoideae. Multi-datasets were applied to gain insights into the deeper relationships in this subfamily. Our results provide the important insight into plastome-based and nuclear-based phylogenetic relationships in the subfamily and outline the significance for future phylogenomic studies among genera or subfamilies employing the plastomes in Selaginellaceae.

2 Materials and methods

2.1 Taxon sampling and sequencing

Nine samples were added and sequenced here, plus 24 accessions available in GenBank and our previous work (Zhou et al., 2022) (Supplementary Table S1). The total genomic DNA of nine samples was extracted from herbarium specimens. Single-stranded circular DNA libraries were constructed and sequenced on Illumina HiSeq 2500 platform using 150-bp pair-end sequencing by Beijing Novogene Technology Co., Ltd. (Tianjin, China). In total, there were 33 samples representing ca. 25 taxa of all six genera in Gymnogynoideae *sensu* Zhou and Zhang (2023), and all six sections of *Selaginella* subg. *Gymnogynum* *sensu* Zhou and Zhang (2015) correspond to all subgenera in Weststrand and Korall's (2016b) classification. These samples represent over 90% species of all genera in the subfamily, except *Bryodesma* and *Gymnogynum*. Four other Selaginellaceae outside of Gymnogynoideae were included as part of the ingroup based on the previous works of Zhou et al. (2022) and Zhao et al. (unpubl. data) (Supplementary Table S1). Two species of *Isoetes* were selected as outgroups (Wood et al., 2020) (Supplementary Table S1).

2.2 Plastome assembling and annotation

Adapters, containing Ns, and low-quality bases from the raw data were first filtered by Fastp v0.12.4 (Chen et al., 2018). *De novo* assemblies were constructed using GetOrganelle v1.7.5 (Jin et al., 2020). Bandage v0.8.1 (Wick et al., 2015) was also used to visualize *de novo* assembly graphs. Complete plastome annotation was performed through the online program Geseq (Tillich et al., 2017). The initial annotation was subsequently inspected and adjusted manually by comparing it with the complete plastome, which was published in our previous study, confirming the start and stop codons and the exon–intron boundaries of genes in Geneious Prime 2019.2.1. A Blastn search was conducted with default parameter settings for any uncertain genes. In addition, all tRNA genes were validated using the online tRNAscan-SE v2.0 service (Chan et al., 2021).

2.3 Alignment and phylogenetic inference

The complete plastomes of Gymnogynoideae were acquired in our study. However, previous studies have shown that the plastome of Selaginellaceae has different degrees of gene loss and inversion (Tsuji et al., 2007; Smith, 2009; Xu et al., 2018; Mower et al., 2019;

Zhang et al., 2019a, b; Kang et al., 2020; Zhang et al., 2020a; Xiang et al., 2022; Zhou et al., 2022), and it could not be aligned with confidence. In addition, to alleviate the impact of missing data, the Python scripts “get_annotated_regions_from_gb.py” (<https://github.com/Kinggerm/PersonalUtilities/>) were used to automatically extract 51 protein-coding gene loci and 23 intergenic loci which was shared by all samples. All loci were aligned individually in Mafft v7.450 (Katoh and Standley, 2013) with setting “E-INS-i” and trimmed poorly aligned regions using trimAI v1.3 (Capella-Gutierrez et al., 2009) with setting “automated1”.

We generated four basic datasets, namely: (1) 51 concatenated gene (gene), (2) 23 concatenated intergenic regions (intergene), (3) 74 concatenated gene and intergenic regions (gene_intergene), and (4) first and second codon sites of 51 concatenated gene (codon12). ModelFinder (Kalyaanamoorthy et al., 2017) was used to select the best-fitting likelihood model (Supplementary Table S2) for maximum likelihood (ML) and Bayesian inference (BI) under the Corrected Akaike Information Criterion (AICc). ML tree searches were conducted using IQ-tree v2.1.3 (Nguyen et al., 2015) with 5,000 ultrafast bootstraps (Hoang et al., 2018) analyses in a single run. BI was conducted using MrBayes 3.2.2 (Ronquist et al., 2012). Four Markov chain Monte Carlo chains were run, each beginning with a random tree and sampling one tree every 100 generations of 2,000,000 generations, and the first 25% of samples were discarded as burn-in. Each tree was visualized with their maximum-likelihood bootstrap support values (MLBS) and Bayesian inference posterior probability (BIPP) in Figtree v1.4.3 (Rambaut, 2017).

2.4 Quantification of phylogenetic signal for alternative tree topologies

To assess the variation of the phylogenetic signal of each alternative topology of the conflicting nodes, both delta site-wise log-likelihood scores (Δ SLS) and delta gene-wise log-likelihood support (Δ GLS) were used to quantify the distribution of phylogenetic signal for three alternative topologies of *Ericetorum* and four alternative topologies of *Gymnogynum*. For each of the above-mentioned four basic datasets, the analytical methods of Shen et al. (2017) were followed. The average Δ SLS for each locus was separately calculated to avoid the influence of length in the four basic datasets, and standard deviation was used to identify outliers. Boxplot method was used to identify these outlier loci, in which the average Δ SLS values were more significant than the upper bound or smaller than the lower bound of the boxplot, and plotted in OriginPro v2022b (OriginLab Corporation, Northampton, MA, USA). Then, those involved outlier genes were pruned from the above-mentioned four basic datasets in an effort to reduce the conflict at the position of *Gymnogynum* and *Ericetorum* (Supplementary Table S3). For the placement of *Gymnogynum*, we generated another four reduced datasets (codon12_RG, gene_RG, intergene_RG, and gene_intergene_RG) after removing the outlier loci of *Gymnogynum* from the basic datasets. We also generated another four reduced datasets of codon12_RE, gene_RE, intergene_RE, and gene_intergene_RE after removing the outlier

loci of *Ericetorum* from basic datasets. Finally, another four datasets (codon12_REG, gene_REG, intergene_REG, and gene_intergene_REG) were generated for *Gymnogynum* and *Ericetorum* after removing the outlier genes. Phylogenetic trees were reconstructed using IQ-tree and MrBayes as described previously. The phylogenetic signal was also recalculated, and plots were made using OriginPro.

2.5 Assembly of single-copy orthologues and phylogeny reconstruction

HybPiper v1.3.1 (Johnson et al., 2016) was used to assemble single-copy genes (SCGs) from sequenced quality-filtered reads. The sequences of our previous study (Zhao et al., unpublished data) had identified 140 SCGs that were used as target input file for HybPiper. Each gene was aligned individually in Mafft with setting “E-INS-i”, and poorly aligned regions were trimmed using trimAI with setting “automated1”. ModelFinder was used to select the best-fitting likelihood model (Supplementary Table S2) for ML and BI under the Corrected Akaike Information Criterion (AICc) based on the concatenated single-copy genes matrix (SCGs). ML tree searches were conducted using IQ-tree with 5,000 ultrafast bootstrap analyses in a single run. BI was conducted using MrBayes. Four Markov chain Monte Carlo chains were run, each beginning with a random tree and sampling one tree in every 100 generations of 2,000,000 generations, and the first 25% of the samples were discarded as burn-in. In addition, the four basic datasets of plastome were concatenated with single-copy genes matrix (SCGs), respectively. For the four derived datasets (SCGs_codon12, SCGs_gene, SCGs_intergene, and SCGs_gene_intergene) that combined, nuclear and plastid loci datasets were used in the same tree searches (ML and BI) as described previously.

2.6 Concordance, character evolution, and phylogenetic networks

The program PhyParts (Smith et al., 2015) was used to identify both plastid loci tree and nuclear loci tree concordance and discordance for each node in the concatenated trees. Each locus tree was reconstructed using ML as implemented by IQ-tree, with the substitution model selected using ModelFinder (set -AICc) and nodal support assessed using 1,000 ultra-fast bootstrap replicates. The results were visualized using the Python script “phypartspiecharts.py” (<https://github.com/mossmatters/MJPythonNotebooks>). Two discrete characters were reconstructed in their marginal ancestral states in R/phytools (Revell, 2011) with our tip coding states (Supplementary Table S4) (model = “ER”). Habit (hydrophytes, xerophytes, mesophytes) and plastome master structures [DR (the repeats region of plastome was direct repeats), IR (the repeats region of plastome was two inverted repeats), NR (the plastome without repeat regions), and DR-IR coexisting (direct repeats and inverted repeats co-existing in plastome)] were scored. They were selected and coded following previous studies (Tryon, 1955; Chu, 2006; Zhang et al., 2013; Zhou and Zhang, 2015; Zhou

et al., 2016; Weststrand and Korall, 2016a, b; Zhang et al., 2019a; Zhou and Zhang, 2023, b; Zhang et al., 2020a; Zhou et al., 2022) and our investigations for Selaginellaceae. To explore specific reticulate evolution events and gene flow, a maximum of five reticulation events was set with the command “inferNetworks_MPL” in PhyloNet (Yu and Nakhleh, 2015). An individual SCG tree used as input was generated by IQ-tree with 1,000 rapid bootstrap replicates and then visualized in Dendroscope v3.8.1 (Huson and Scornavacca, 2012).

3 Results

3.1 Characteristics of plastomes in Gymnogynoidae

All sequenced plastomes of nine samples representing four genera (*Lepidoselaginella*, *Ericetorum*, *Megaloselaginella*, and *Afroselaginella*) in Gymnogynoidae were completely assembled. The plastomes of *L. lepidophylla* and *B. peruvianum* had a typical quadripartite structure composed of a large single-copy (LSC), a small single-copy (SSC), and two inverted repeats (IR), but the remaining species had diverse and unique plastome structures which were divided into four types: DR, IR, NR, and DR–IR. *L. lepidophylla* and *B. peruvianum* were IR type with two isomeric forms via a homologous recombination between the two IR copies. The plastomes of all species of *Ericetorum* sequenced were DR type with only one conformation and had a pair of direct repeats (DR) without other short or medium repeats in SC. In *Gymnogynum*, except that the clade A comprised *G. arthriticum* and *G. lingulatum* were DR–IR type with three conformations for both IR and DR in single-copy region (SC), the other clades were DR type with only one conformation. All samples of *Afroselaginella* were NR type and with only one conformation, which was first found in this study. Same as the clade A in *Gymnogynum*, *Megaloselaginella* also has three conformations for both IR and DR in SC. In *Bryodesma*, except that the plastome of *B. peruvianum* was IR type with two conformations for a pair of inverted repeats in SC, the remaining species were DR type.

The plastome features of Gymnogynoidae were strongly different among the six genera (Supplementary Table S5). The sizes of the 33 plastomes ranged from 100,119 bp in *Afroselaginella myosurus* (voucher: C.J. Rothfels et al. 08–183) to 131,938 bp in *Gymnogynum remotifolium* (voucher: X.-M. Zhou 722). The plastome GC content ranged from 50.70% in *Ericetorum pectinatum* (voucher: F. Rakotondrainibe 6494) to 56.70% in *G. arthriticum*. *Afroselaginella* has NR type of plastome with only SC. The length of LSC was from 45,276 bp in *E. lyallii* (Zhang et al., 2019a) to 91,273 bp in *Bryodesma peruvianum*. The length of SSC was from 16,732 bp in *G. lingulatum* to 46,373 bp in *G. remotifolium* (voucher: X.-M. Zhou 722). The shortest repeat (RC) was 364 bp in *B. peruvianum* (lost in *Afroselaginella*), and the longest was 18,194 bp in *B. rupestre*. For the gene contents, *Afroselaginella* has minimal genes (72), tRNAs (10), and rRNAs (4). Other genera have genes ranging from 80 to 98, tRNAs from 10 to 17, and rRNAs from four to eight, respectively.

3.2 Phylogenetic relationships of Gymnogynoidae based on plastid loci

All six genera in Gymnogynoidae were recovered as monophyletic and fully supported in four basic datasets (codon12, gene, intergene, and gene_intergene) in both ML and BI analyses (Figure 2; Supplementary Figures S1–S4). However, four alternative phylogenetic relationships among these six genera were recovered. In all ML analyses based on four basic datasets, *Ericetorum* was sister to the *Megaloselaginella* + *Afroselaginella* clade with medium support (MLBS = 61–87), and *Gymnogynum* was recovered as sister to a clade composed of *Ericetorum* + *Megaloselaginella* + *Afroselaginella* with weak to strong support based on four basic datasets (gene: MLBS = 99, intergene: MLBS = 38, gene_intergene: MLBS = 99, and codon12 dataset: MLBS = 98; Figure 2A). In BI analysis, the relationship of the codon12 dataset revealed that *Ericetorum* was sister to other genera of Gymnogynoidae with maximum support (BIPP = 1; Figure 2B), and *Gymnogynum* was resolved as sister to a clade containing *Megaloselaginella* + *Afroselaginella* with strong support (BIPP = 0.98; Figure 2B). Based on both the gene and gene_intergene datasets, BI analysis revealed that the *Ericetorum* + *Gymnogynum* clade (BIPP = 1.0; Figure 2C) was consistently recovered as sister to a clade containing *Megaloselaginella* + *Afroselaginella* (BIPP = 1.0; Figure 2C). Comparing the resolutions based on the intergene dataset with BI tree, the sister relationship between *Gymnogynum* and a clade comprised by *Bryodesma* + *Lepidoselaginella* was recovered with strong support (BIPP = 0.98; Figure 2D), and *Ericetorum* was sister to a clade composed of *Megaloselaginella* + *Afroselaginella* with maximum support (BIPP = 1.0; Figure 2D). Overall, *Ericetorum* had three alternative placements in all ML and BI analyses based on four basic datasets of plastid: T1 ((*Ericetorum*, (*Megaloselaginella*, *Afroselaginella*)), *Gymnogynum*); T2 (*Ericetorum*, other genera of Subfamily Gymnogynoidae); and T3 ((*Ericetorum*, *Gymnogynum*), (*Megaloselaginella*, *Afroselaginella*)) (Supplementary Figure S5); *Gymnogynum* had four alternative placements in all ML and BI analyses based on four basic datasets of plastid: T1 (*Gymnogynum*, (*Ericetorum*, (*Megaloselaginella*, *Afroselaginella*))); T2 ((*Gymnogynum*, *Ericetorum*), (*Megaloselaginella*, *Afroselaginella*)); T3 ((*Gymnogynum*, (*Bryodesma*, *Lepidoselaginella*)), other genera of subfamily Gymnogynoidae); and T4 ((*Gymnogynum*, (*Megaloselaginella*, *Afroselaginella*)), *Ericetorum*) (Supplementary Figure S6).

After the outlier loci from basic datasets of plastid were removed, both ML and BI analyses were reconducted for the alternative topologies of *Gymnogynum* and *Ericetorum*. For the alternative topologies of *Gymnogynum*, the genus involved both codon12_RG and gene_RG removing two outlier loci (*ycf1* and *ycf2*) from codon12 and gene dataset, intergene_RG dataset removing four outlier loci (*psaA-psaB*, *rpl22-rps19*, *psbH-psbN*, and *psbF-psbL*) from gene_intergene dataset, and gene_intergene_RG dataset removing one outlier loci (*atpA-ycf12*) from gene_intergene dataset (Figure 3; Supplementary Table S3). For the alternative topologies of *Ericetorum*, the genus involved codon12_RE dataset removing three outlier loci (*clpP*, *ycf1*, and *ycf2*) from the codon12 dataset, the gene_RE dataset removing two outlier loci (*ycf1* and *ycf2*) from the gene dataset, the intergene_RE dataset removing two outlier loci

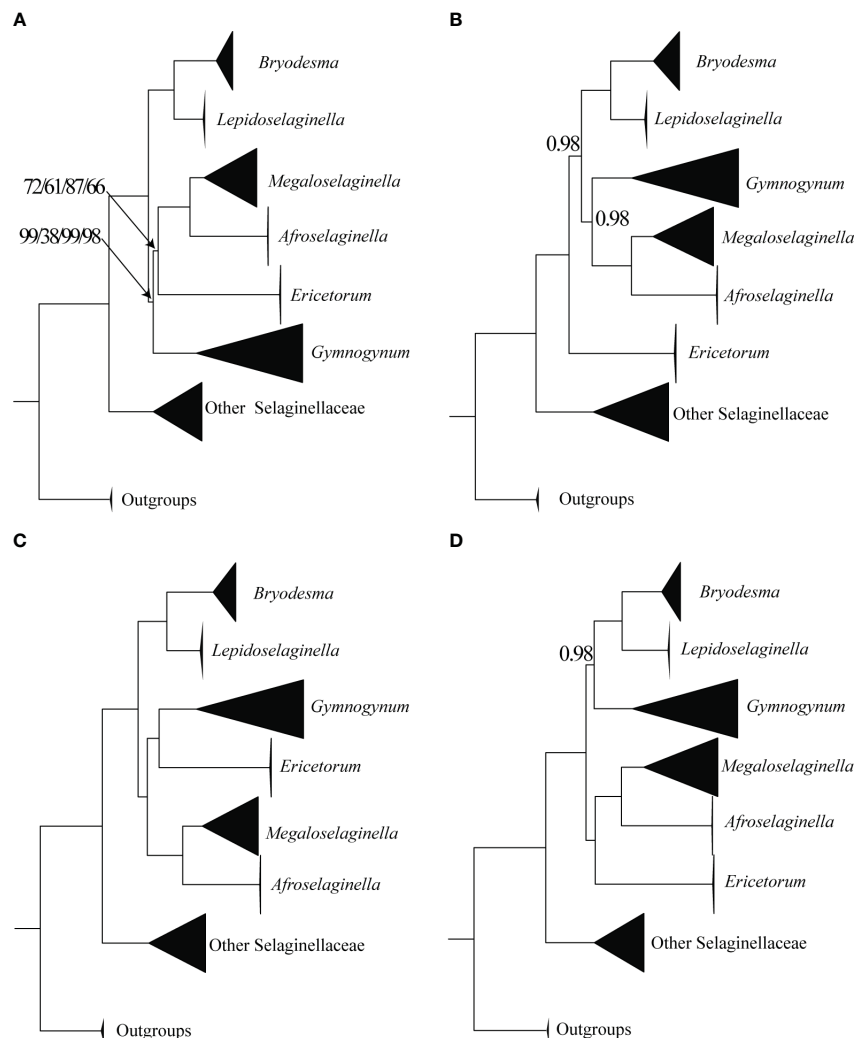


FIGURE 2

Simplified phylogeny of Gymnogoideae based on different datasets and methods. (A) Maximum likelihood phylogeny based on gene/intergene/gene_intergene/codon12/dataset. (B) Bayesian inference phylogeny based on codon12 dataset. (C) Bayesian inference phylogeny based on gene and gene_intergene dataset. (D) Bayesian inference phylogeny based on intergene dataset. The size of the black triangles was in proportion to the sampled size of individual clades. Maximum-likelihood bootstrap support values (MLBS) and Bayesian inference posterior probability (BIPP) are shown above the branches and are 100/1.0 unless otherwise indicated.

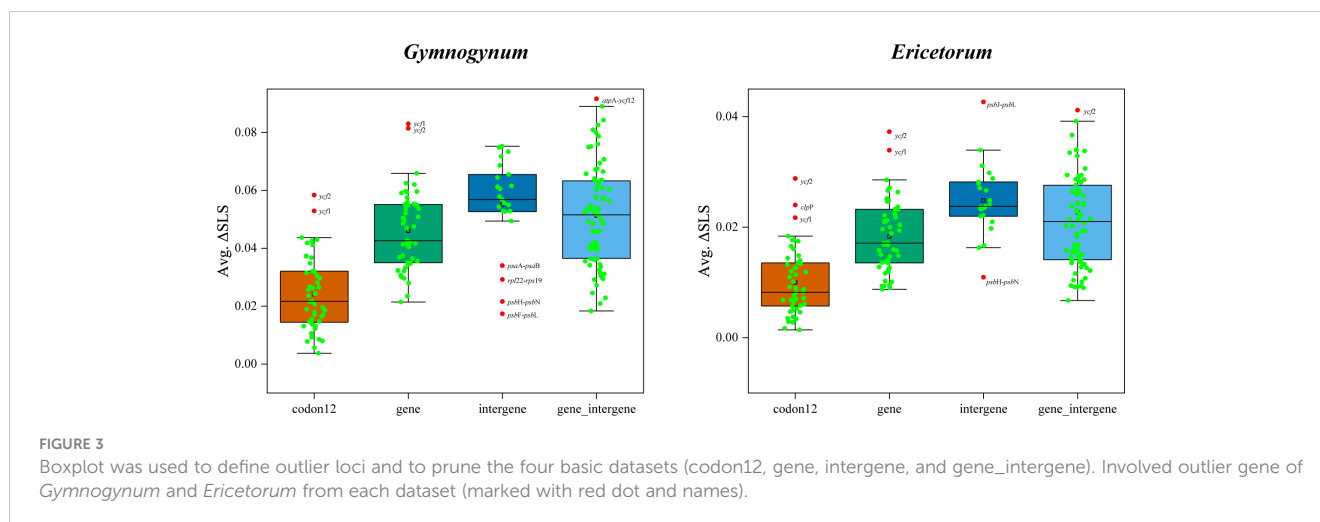
(*psbH-psbN* and *psbJ-psbL*) from intergene dataset, the *gene_intergene_RE* dataset removing one outlier loci (*ycf2*) from *gene_intergene* dataset to generate (Figure 3; Supplementary Table S3). All analyses showed that the topologies of ML and BI were each identical to those based on the same datasets with the outlier loci included, but their support values were variable among different datasets (Supplementary Figures S1-S4).

3.3 Conflicting phylogenetic signals of plastid loci

The phylogenetic signals for *Gymnogynum* and *Ericetorum* with conflicting topologies as noted above were examined. Both *Gymnogynum* and *Ericetorum* were plotted in the tree, marking the nodes with the alternative positions (Figures 4A, 5A; Supplementary Figures S5, S6). For *Ericetorum*, intergene dataset

supported T2 with the highest proportion of Δ GLS (43%), and 41%–45% loci supported T1 with the highest proportion of Δ GLS based on the remaining three basic datasets (codon12, gene, and *gene_intergene* dataset) (Figure 4D). After the outlier loci were removed, T1 (38%) and T2 (39%) had a similar proportion of Δ GLS support based on *intergene_RE* and *gene_intergene_REG* datasets. T1 was still supported with the strongest Δ GLS (41%–45%) with all derived datasets (Figure 4D). The examination of Δ SLS values showed the same trend proportions of sites supporting T1 to T3. Both fundamental and derived datasets showed the highest proportion of Δ SLS supporting T1 (42%–54%; Figure 4E). In summary, for the 12 datasets of *Ericetorum*, T1 showed a higher support with strongest Δ GLS (46.4%) and strongest Δ SLS (42.1%; Figures 4B, C).

Within *Gymnogynum*, the examination of Δ GLS and Δ SLS values showed a greater variation of proportions of loci and sites supporting from T1 to T4 based on different datasets. The codon12



dataset showed the highest proportion of Δ GLS supporting T1 (33%) and supported T1 (37% and 33%, respectively) based on codon12_RG and codon12_REG datasets after the outlier loci were removed (Figure 5D). The gene dataset showed a similarly high proportion of Δ GLS supporting T1 (31%) and T2 (31%), but with the highest proportion of Δ GLS supporting T2 (31%) with both gene_RG and gene_REG datasets (Figure 5D). However, the intergene dataset supported T4 with the highest proportion Δ GLS (Figure 5D) and with a similarly high proportion of Δ GLS supporting T3 and T4 (37% and 33%, respectively) with intergene_RG and intergene_REG datasets (Figure 5D). With gene_intergene dataset, T1 had the highest proportion of Δ GLS (31%). T1 and T3 had a similarly high proportion (30%) of Δ GLS with gene_intergene_RG dataset (Figure 5D). T3 and T4 had an equally high proportion (33%) of Δ GLS with gene_intergene_REG dataset (Figure 5D). Within the qualification of Δ SLS (Figure 5E), T2 had the highest proportion of Δ SLS with gene (39%), gene_RG (39%), and gene_REG (39%) datasets (Figure 5E). T3 had the highest proportion of Δ SLS with intergene (35%), intergene_RG (34%), and intergene_REG (35%) datasets (Figure 5E). When the outlier sites were included, T3 (45%) had the highest proportion of Δ SLS with the codon12 dataset (Figure 5E), but both codon12_RG and codon12_REG datasets showed the highest proportion of Δ SLS supporting T1 (49%; Figure 5E). T1 had the highest proportion of Δ SLS supporting T1 with gene_intergene and gene_intergene_RG datasets, and T3 (35%) had the highest proportion of Δ SLS with gene_intergene_REG, which was generated by the further removal of outlier loci in *Ericetorum* (Figure 5E). In summary, for the 12 datasets of *Gymnogynum*, T1 showed higher support with the strongest Δ GLS (34.3%) and the strongest Δ SLS (30.5%; Figures 5B, C).

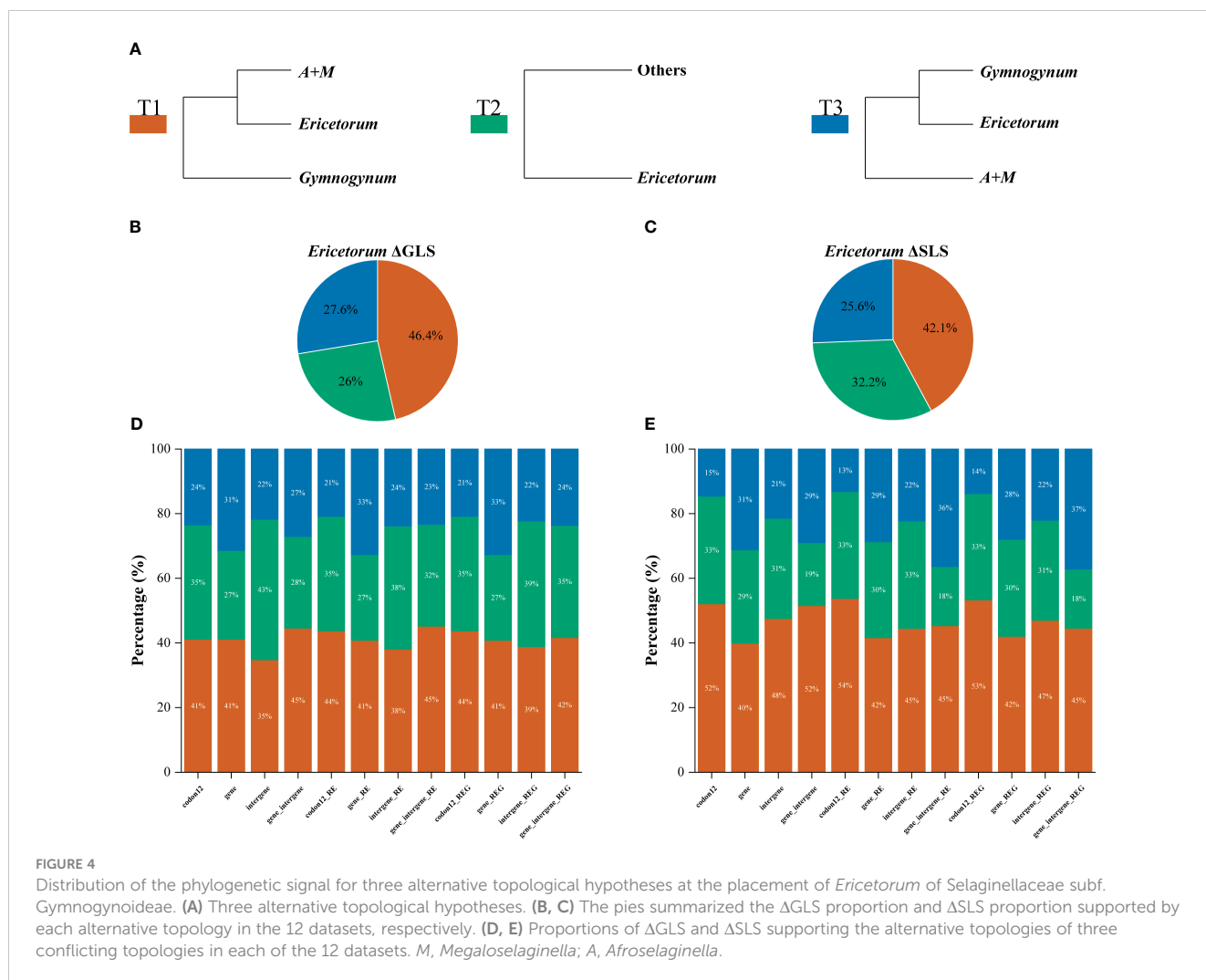
3.4 Phylogenetic inference based on single-copy nuclear genes and combined plastid and nuclear datasets

In total, 140 single-copy nuclear genes were used as baits to extract from our high-throughput sequencing, but only 139 genes were

used for further analyses because one gene failed to be assembled from reads (Supplementary Figure S7; Supplementary Table S6). In addition, fewer loci were successfully assembled for some old herbarium specimens with degraded DNA—for example, only 10 loci of the specimen *Ericetorum lyallii*-2, 24 loci of the specimen *E. lyallii*-1, 11 loci of the specimen *E. pectinatum*-1, and 63 loci of the specimen *E. pectinatum*-2 were assembled (Supplementary Figure S7; Supplementary Table S6). Both BI and ML methods were used to reconstruct tree species based on a concatenated single-copy genes (SCGs) supermatrix. In all analyses, all six genera of Gymnogynoideae were recovered as monophyletic and fully supported (MLBS = 100; BIPP = 1.0; Figure 6). The topology was the same as the ML topology based on all plastid datasets: *Bryodesma* was sister to *Lepidoselaginella* (MLBS = 100; BIPP = 1.0), and they together were sister to all other genera of Gymnogynoideae (Figure 6); *Gymnogynum* was sister to *Megaloselaginella* + *Afroselaginella* + *Ericetorum* with strong support (MLBS = 98; BIPP = 0.84), and *Ericetorum* was sister to *Megaloselaginella* + *Afroselaginella* with strong support (MLBS = 100; BIPP = 1.0; Figure 6). The combined data of the single-copy nuclear genes and each one of the plastid datasets resolved intergeneric relationships in Gymnogynoideae as follows: ((*Bryodesma*, *Lepidoselaginella*), ((*Megaloselaginella*, *Afroselaginella*), *Ericetorum*), *Gymnogynum*), all relationships being fully supported except the sister relationship between *Ericetorum* and *Megaloselaginella* + *Afroselaginella* which received medium support (MLBS = 68–86; BIPP = 0.78–0.90; Supplementary Figure S8).

3.5 Discordance, ancestral state reconstruction, and network evolution

The conflict analysis from Phyparts showed that less gene tree discordance was detected among nuclear genes and plastid genes regarding the placement of the six major clades (Figures 6, 7). Phyparts suggested the monophyly of Gymnogynoideae with 138 out of 139 informative nuclear gene trees and all 74 plastid locus trees (Figures 6, 7). The ICA values varied among the trees for the monophyly of the six genera each and for the relationships among



them. The monophyly of *Bryodesma*, *Lepidoselaginella*, *Megaloselaginella*, *Afroselaginella*, *Ericetorum*, and *Gymnogynum* each was supported by 67 (ICA = 0.55), 73 (ICA = 0.90), 56 (ICA = 0.50), 74 (ICA = 1), 74 (ICA = 1), and 63 (ICA = 0.45) out of the 74 single-locus plastid trees (Figure 7), respectively. The clade comprising *Bryodesma* + *Lepidoselaginella* was supported by 105 nuclear gene trees (out of 139; ICA = 0.40; Figure 6) and 60 plastid loci trees (out of 74; ICA = 0.45; Figure 7), respectively. The clade comprising *Megaloselaginella* + *Afroselaginella* was supported by 111 nuclear gene trees (out of 139; ICA = 0.52; Figure 6) and 57 plastid loci trees (out of 74; ICA = 0.44; Figure 7), respectively. The sister relationship between the clade composed of *Bryodesma* + *Lepidoselaginella* and the clade composed of *Megaloselaginella* + *Afroselaginella* + *Ericetorum* + *Gymnogynum* was supported by 88 nuclear gene trees (out of 139; ICA = 0.28; Figure 6) and 57 plastid loci trees (out of 74; ICA = 0.32; Figure 7). However, the phylogenetic placement of *Ericetorum* and *Gymnogynum* showed a high proportion of discordance. The sister relationship between *Ericetorum* and the clade comprising *Megaloselaginella* + *Afroselaginella* was supported by only 15 nuclear gene trees (out

of 139; ICA = 0.18; Figure 6) and only 18 plastid loci trees (out of 74; ICA = 0.13; Figure 7), while the sister relationship between *Gymnogynum* and the clade comprising *Ericetorum* + *Megaloselaginella* + *Afroselaginella* was supported by only 49 nuclear gene trees (out of 139; ICA = 0.14; Figure 6) and only 19 single-locus trees (out of 74; ICA = 0.12; Figure 7).

The character evolution of habits (hydrophytes, xerophytes, and mesophytes) and plastome master structures (DR, IR, NR, and DR-IR coexisting) is shown in Figures 8A, B. Both were not much homoplasious and thus useful to characterize the clades morphologically (Figures 8A, B). To analyze the potential causes of nuclear gene trees' conflict, we employed SCGs datasets to perform a phylogenetic network analysis in PhyloNet while accounting for both hybridization and incomplete lineage sorting (ILS). One reticulation event from ancestral Selaginellaceae and *Bryodesma* lineage was inferred in all five examinations (Figure 8C), with an inheritance probability of 0.017 from the ancestral Selaginellaceae lineage and 0.983 from the *Bryodesma* lineage (Figure 8C). Four inferred hybridization events between *Bryodesma* and *Ericetorum* were also detected (Figure 8C: network 2–network 5).

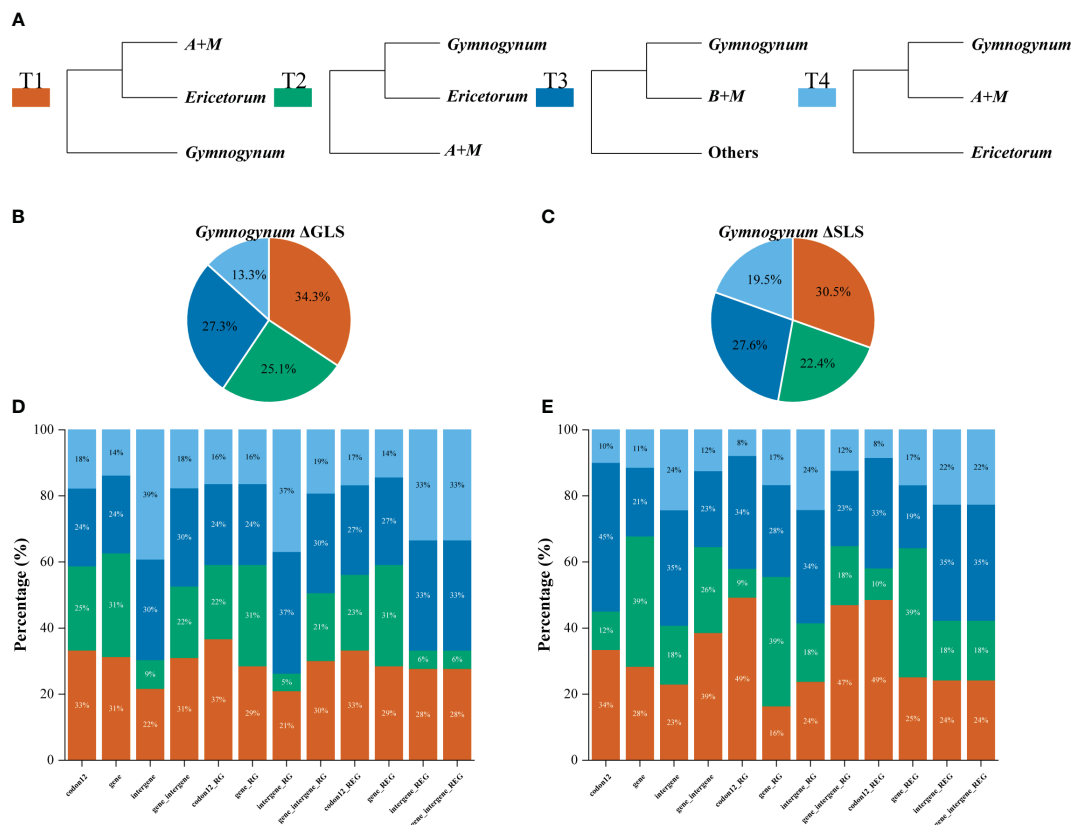


FIGURE 5 Distribution of phylogenetic signal for four alternative topological hypotheses at the placement of *Gymnogynum* of Selaginellaceae subf. Gymnogynoideae. (A) Four alternative topological hypotheses. (B, C) The pies summarized the Δ GLS proportion and Δ SLS proportion supported by each alternative topology in the 12 datasets, respectively. (D, E) Proportions of Δ GLS and Δ SLS supporting the alternative topologies of three conflicting topologies in each of the 12 datasets. M, *Megaloselaginella*; A, *Afroselaginella*; B, *Bryodesma*; L, *Lepidoselaginella*.

4 Discussion

4.1 The relationship between *Afroselaginella* and *Megaloselaginella*

Although the ML trees are somewhat discordant with the BI trees based on plastid datasets, our phylogenetic inference strongly supported the sister relationship between *Afroselaginella* and *Megaloselaginella* in all analyses based on both plastid and nuclear loci (Figures 2–7; Supplementary Figures S1–S5). Based on two molecular makers (*rbcL* + ITS; Zhou et al., 2016) and morphological difference, *Megaloselaginella* (including *M. exaltata*) and *Afroselaginella* (including *A. myosurus*) were recognized by Zhou and Zhang (2015), albeit as sections of their S. subg. *Gymnogynum*. Subsequently, Weststrand and Korall (2016b) combined the two taxa and coined S. subg. *Exaltatae* based on the Bayesian analysis of three genes (*rbcL* + *pgiC* + *SQD1*; Weststrand and Korall, 2016a, b). Several previous phylogenomic studies of Selaginellaceae involved Gymnogynoideae (e.g., Du et al., 2020; Zhang et al., 2020a), but *Afroselaginella* was still not sampled. Here we provide the first strong phylogenomic evidence that *Afroselaginella* is indeed sister to *Megaloselaginella* with full support (Figures 2–7; Supplementary Figures S1–S4, S8).

Despite the confirmed sister relationship between *Afroselaginella* and *Megaloselaginella*, considering the distinction between them, here we support the recognition of both at generic rank. Firstly, species of *Megaloselaginella* have large and erect plants, articulate stems, and pillared to baculite microspore surfaces (Zhou and Zhang, 2015; Weststrand and Korall, 2016a, b; Zhou and Zhang, 2023), whereas those of *Afroselaginella* have medium-sized and creeping plants, inarticulate stems and branches, and microspores with an equatorial ring and verrucate surfaces (Zhou and Zhang, 2015; Weststrand and Korall, 2016a, b; Zhou and Zhang, 2023). In previous studies, *Megaloselaginella* was generally thought to be only distributed in the Central and South America, whereas those of *Afroselaginella* only in western Africa (Zhou and Zhang, 2015, 2023). However, this phylogenetic analysis showed that the western Africa species *Afroselaginella grillipes* was clustered into *Megaloselaginella*. We further studied the specimen of *Afroselaginella grillipes* and confirmed that it is indeed a member of *Megaloselaginella*. This species with large plants and articulate stems is highly consistent with those members of *Megaloselaginella*. *Afroselaginella grillipes* should be transferred to *Megaloselaginella*. *Afroselaginella* has the smallest plastome size (ca. 100 kb) in Gymnogynoideae (Supplementary Table S5) and a unique plastome structure with only SC region but without a repeat region and lacking one ribosomal operon in Selaginellaceae

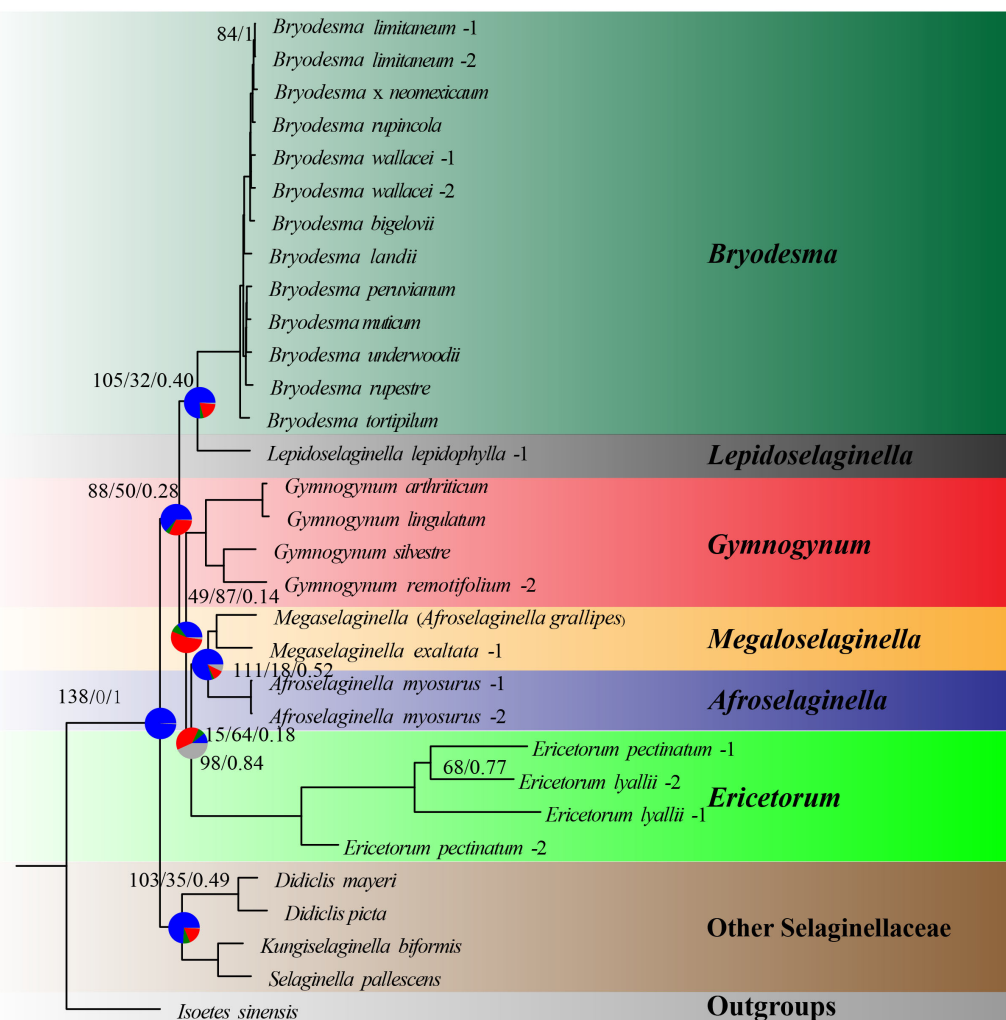


FIGURE 6

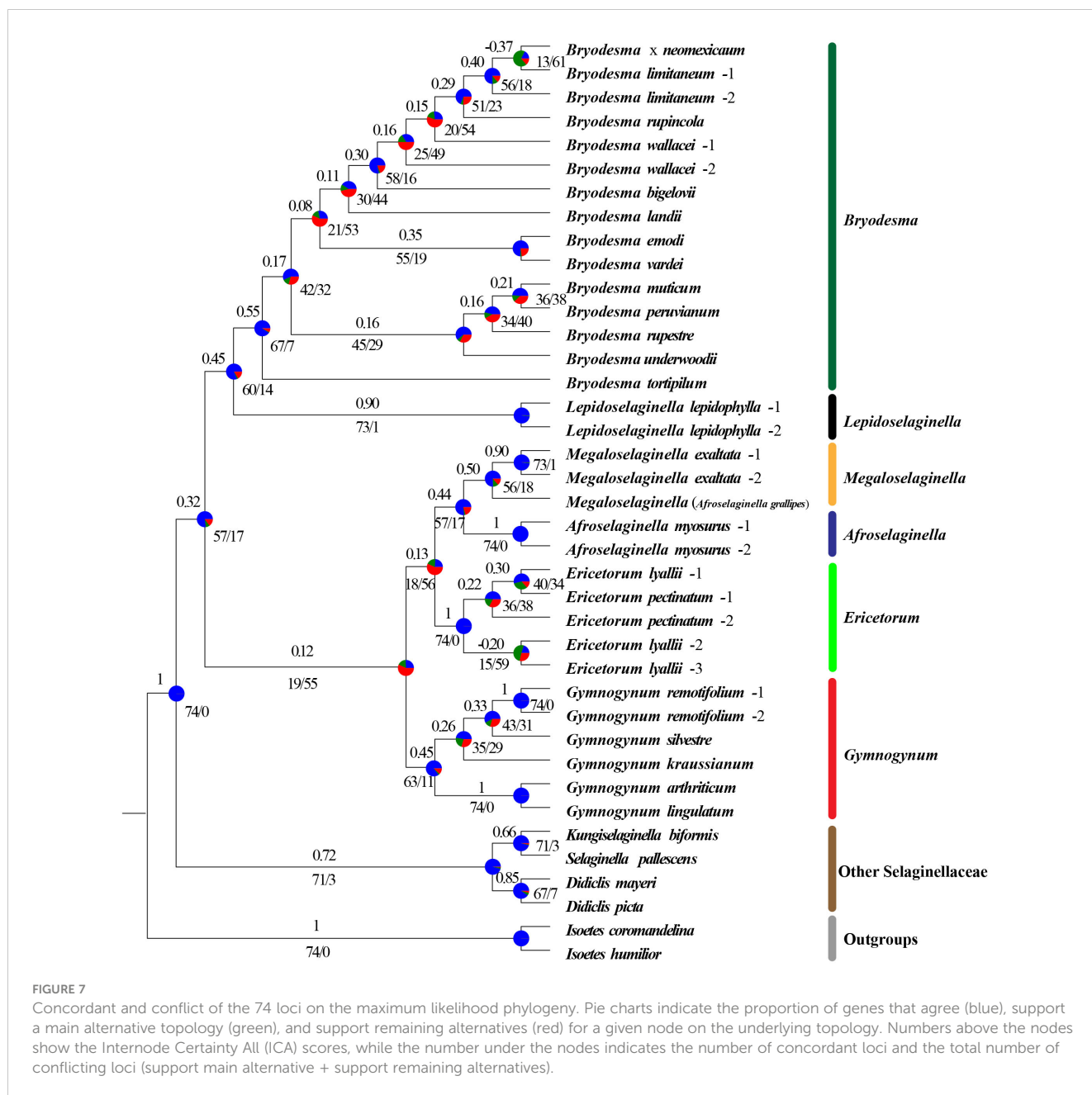
Maximum likelihood phylogeny of Selaginellaceae subf. Gymnogynoideae based on single-copy nuclear genes. Pie charts on the focused seven nodes, indicating the proportion of genes that agree (blue), support a main alternative topology (green), and support remaining alternatives (red) for a given node on the underlying topology. Numbers above the nodes indicate the number of concordant loci, the total number of conflicting loci, and the Internode Certainty All (ICA) scores. Maximum-likelihood bootstrap support values (MLBS) and Bayesian inference posterior probability (BIPP) are shown under the branches and are 100/1.0 unless otherwise indicated.

(Supplementary Table S5), whereas *Megaloselaginella* has DR–IR coexisting plastome structure with three conformations *in vivo*. In addition, molecular dating showed that *Afroselaginella* and *Megaloselaginella* diverged from each other ca. 210 Ma (Klaus et al., 2017), an age older than the stem ages of most of the fern families recognized by PPG I (2016) following the molecular datings by Schuettpelz and Pryer (2009); Rothfels et al. (2015), and Testo and Sundue (2016). Similar arguments also apply to the recognition of other four genera in Gymnogynoideae.

4.2 Phylogenetic signals of plastid loci for incongruent nodes

Regarding *Ericetorum* and *Gymnogynum*, our concordance analysis detected an obvious discordance among trees based on different plastid loci, and ML and BI trees based on all plastid loci

datasets demonstrated a high proportion of conflict (Figures 2–5, 7; Supplementary Figures S1–S4). The topologies of BI trees are identical to one another based on all datasets after removing the outlier loci (Supplementary Figures S2–S4). The topologies of BI trees generally did not change obviously based on any datasets, although BIPP values varied, even after the outlier loci were removed (Supplementary Figures S2–S5). To assess the alternative placements of *Ericetorum* and *Gymnogynum*, we used the approaches of Shen et al. (2017) to measure the phylogenetic signal of each plastid locus. As previous studies had shown, conflicts among different sequence types were also detected [coding (gene) vs. non-coding regions (intergene)] (Zhang et al., 2020b; Yang et al., 2021; Zhang et al., 2022). For the alternative resolutions of *Ericetorum*, except T2 with a higher proportion of ΔGLS (35%) in the intergene dataset, the remaining two fundamental datasets (codon12 and gene datasets) showed that T1 had higher ΔGLS (41%) (Figure 4D). Removing the outlier loci,



all the derived four datasets (codon12_RE, gene_RE, codon12_REG, and gene_REG) showed that T1 had a relatively higher proportion of ΔGLS (41%–44%), but T1 and T2 had the same high proportion in intergene_RE dataset (38%) and intergene_REG dataset (39%), respectively (Figure 4D). For *Gymnogynum*, T4 had a higher proportion of ΔGLS (39%) in the intergene dataset, T1 had a higher proportion of ΔGLS (33%–37%) in codon12, codon12_RG, and codon12_REG datasets, and T2 had a relatively higher proportion of ΔGLS (31%) in gene, gene_RG, and gene_REG datasets (Figure 5D). For the T1 of *Gymnogynum*, maybe the actual positions had higher summarized ΔGLS and ΔSLS proportions (Figures 5B, C). The different support values (T1 vs. T2) of *Gymnogynum* (codon12/codon12_RG/codon12_REG vs. gene/gene_RG/gene_REG) also demonstrated that the selection of codon positions might affect the accuracy and precision of

phylogeny (Simmons et al., 2006b). Moreover, ΔGLS had a decrease in different degrees based on gene_intergene datasets-concatenated coding and non-coding regions, but the T1 of *Ericetorum* and *Gymnogynum* had a higher proportion of ΔGLS and ΔSLS with gene datasets (Figures 4, 5). It could be caused by the difference in evolutionary force in the coding and non-coding regions (Gielly and Taberlet, 1994).

The ΔGLS is correlated with the gene length of plastid genes, and the topology supported by the largest number of sites (ΔSLS) was further used to assess the variation of the phylogenetic signal. All datasets showed that T1 of *Ericetorum* had a higher proportion of ΔSLS (Figure 4E), but T2 of *Ericetorum* had a higher proportion of ΔGLS based on intergene datasets (Figure 4D). Because noises are randomly distributed in sequences (Delsuc et al., 2005), those data with strong signals rather than simply with more loci should be

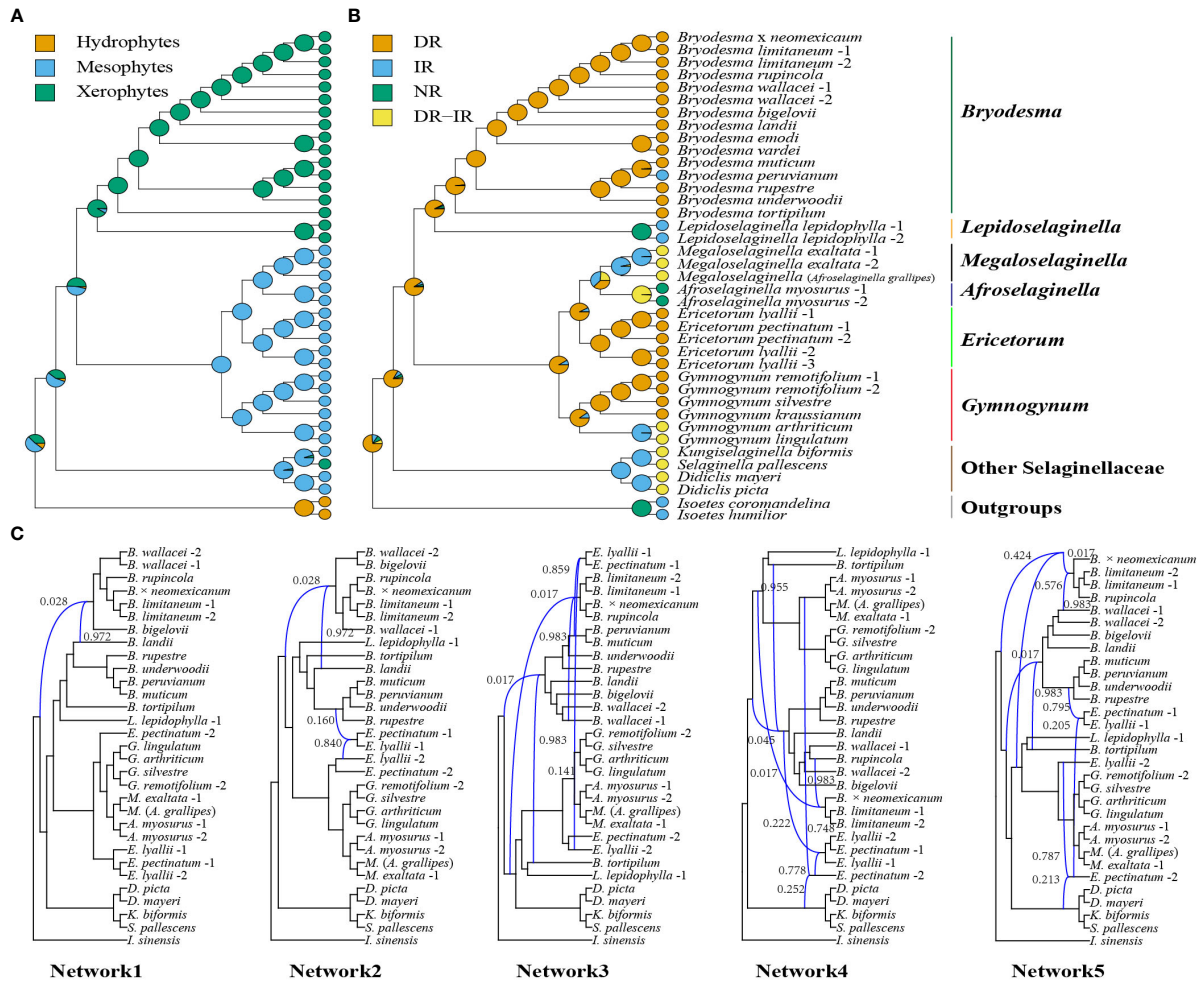


FIGURE 8

Ancestral state reconstruction of habit (A) and plastome master structures (B) in Selaginellaceae subfamily Gymnogoideae. The pie charts on nodes summarize the results of stochastic character mapping. The genus-level divisions are indicated behind the tip names. (C) Species network inferred from PhyloNet pseudolikelihood analyses with one to five hybridization events based on SCGs dataset. The curved branches indicate the minor and major edges of hybrid nodes, respectively. Numbers next to the curved branches indicate the inheritance probabilities for each hybrid node.

selected and used in phylogenetic inference. Recent studies have discovered that plastid genes were largely uninformative in rosids, Fagales, and Leguminosae (Gonçalves et al., 2019; Koenen et al., 2020; Zhang et al., 2020b; Yang et al., 2021). However, our concordance analyses showed that, except for each node of *Eriectorum* and *Gymnogynum*, almost all plastid loci provided a high proportion of trees that are congruent with one another for the remaining nodes (Figure 7). Nevertheless, after removing the outliers, the phylogenetic signals of most possible topologies in each derived dataset generally increased in various proportions (Figures 3–5). Our discordance and phylogenetic signal analysis show that both T1 of *Eriectorum* and that of *Gymnogynum* are most strongly supported by plastid data. Although the sources of conflict in plastome phylogenies are poorly understood, in this study they can be a combination of the unique plastome traits, high substitution rate, high level of rate heterogeneity, missing data, high GC content, unusual RNA editing, long evolutionary history, the analysis methods (ML, BI) themselves, etc. The extremely high

substitution rates and high level of rate heterogeneity have been reported for the plastid genes of Selaginellaceae, which could cause issues in phylogenetic inference (Korall and Kenrick, 2004; Zhang et al., 2020a). Plastomes of Selaginellaceae generally have a high GC content (>50%) (Supplementary Table S5) (Tsuji et al., 2007; Smith, 2009; Xu et al., 2018; Du et al., 2020; Zhang et al., 2020a, 2021; Zhang and Zhang, 2022, 2022; Xiang et al., 2022; Yang et al., 2022; Zhou et al., 2022; Tang et al., 2023), which was thought to be correlated with a high level of RNA editing (Du et al., 2020; Xiang et al., 2022; Yang et al., 2022). The plastome of *Didiclis uncinata* (Desv. ex Poir.) Li Bing Zhang & X.M. Zhou, a member of the same family, was reported to have 3,415 RNA editing sites (Oldenkott et al., 2014). Extreme plastid RNA editing may confound the phylogenetic reconstruction (Du et al., 2020; Zhang et al., 2022). Because old herbarium materials were used in sequencing, a few critical plastomes have quite a bit of missing data—for example, the four samples of *Eriectorum* have 10–63 missing loci (Supplementary Figure S7; Supplementary Table S6).

This added the uncertainty to the phylogenetic estimates (Philippe et al., 2000; Goremykin et al., 2010). A single search and holding only a single tree in ML analysis have been known for spurious inferences, and BI has been known for inflated support values under various circumstances (Simmons et al., 2004, 2006a, Simmons and Freudenstein, 2011). In addition, our analyses showed that there was a very high level of heterogeneity among single-copy nuclear gene trees (Figure 6). While these inconsistencies have never been well explained, it is often believed that the phylogenetic incongruence in plants is mainly due to hybridization (Liu et al., 2022; Zhao et al., 2023). We also detected inter-genus gene flow among almost all lineages of Gymnogynoideae (Figure 8C), so that topology conflict in Gymnogynoideae at the genus level was potentially caused by hybridization. All of these factors might have acted simultaneously and/or cumulatively contributing to the observed conflicts in phylogenetic inference in this study.

4.3 The overall phylogeny in Gymnogynoideae

In this study, the monophyly of six genera in Gymnogynoideae is confirmed with strong support based on plastid data (MLBS = 100; BIPP = 1.0; ICA = 0.44–1.0; Figures 2, 7; Supplementary Figures S1–S4). In previous studies based on complete or nearly complete plastomes (Du et al., 2020; Zhang et al., 2020a; Zhou et al., 2022), the relationships among the genera of Gymnogynoideae were debatable depending on datasets (Figure 1). Part of the reasons may be attributed to the limited taxon samplings (fewer than eight accessions), particularly, *Afroselaginella* was never sampled before in any phylogenomic studies (Du et al., 2020; Zhang et al., 2020a; Zhou et al., 2022).

Using expanded taxon and character sampling and comprehensive analytical methods for dissecting signal and conflict among loci, despite some existing conflicting plastid trees, the most plausible overall phylogeny in Gymnogynoideae appears to be ((*Bryodesma*, *Lepidoselaginella*), (((*Megaloselaginella*, *Afroselaginella*), *Ericetorum*), *Gymnogynum*)) (Figure 2A). This topology is not only supported by most plastid evidence/tree topologies but also is consistent with the nuclear tree based on single-copy nuclear genes (Figure 6). *Bryodesma* is resolved as sister to *Lepidoselaginella* (Figures 1, 2, 7; Supplementary Figures S1–S4), consistent with the results of most previous studies (Arrigo et al., 2013; Zhou et al., 2016; Weststrand and Korall, 2016a; Klaus et al., 2017; Du et al., 2020; Zhang et al., 2020a). *Bryodesma* + *Lepidoselaginella* together are sister to the rest of the subfamily, in accordance with the results of Weststrand and Korall (2016a) based on *rbcL*, *pgiC*, and *SQD1* data and those of Zhou et al. (2016) based on *rbcL* and ITS data, although both latter two studies provided weak support for this resolution. *Gymnogynum* is sister to the remaining three genera together, and *Ericetorum* is sister to *Megaloselaginella* + *Afroselaginella*, in agreement with the results of Weststrand and Korall (2016a) and Zhou et al. (2016), too; again, the latter two provided weak support. Overall, our large plastid and nuclear datasets independently well resolved the relationships in the

subfamily, consistent with the two most recent large phylogenies (Zhou et al., 2016; Weststrand and Korall, 2016a) based on Sanger sequence data, and we here provide the strongest support for these relationships for the first time. Notably, based on six plastid and nuclear markers, Zhou and Zhang (2023) resolved *Ericetorum* as sister to *Gymnogynum*, but the support values were low (<50%).

4.4 Evolution of habit and plastome master structures in Gymnogynoideae

With the relationships well resolved in the subfamily, we can understand how habit and plastome master structures evolved in Gymnogynoideae. The ancestral state of the habit of clade A (*Bryodesma* + *Lepidoselaginella*) was xerophytes (or “resurrection”) and that of clade B (*Megaloselaginella*, *Afroselaginella*, *Ericetorum*, and *Gymnogynum*) was mesophytes (Figure 8A), but the ancestral state of the habit of Gymnogynoideae had nearly the same proportion of xerophytes and mesophytes (Figure 8A). Reconstruction of plastome master structures showed that DR was the ancestral state of Gymnogynoideae, and the other three (IR, NR, and DR–IR) were merged with the evolution of plastomes (Figure 8B).

Data availability statement

The new sequenced chloroplast genome of this study has been submitted to GenBank under the accession number of PP693476 to PP693484, respectively. Multiple sequence alignments, phylogenetic trees, trait ancestral character state reconstructions, are available on the Figshare repository: [10.6084/m9.figshare.25750449.v1](https://doi.org/10.6084/m9.figshare.25750449.v1).

Author contributions

JZ: Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. ZRH: Writing – review & editing, Conceptualization. SLF: Writing – original draft, Software, Methodology. XKH: Writing – original draft. LYJ: Writing – original draft. YPH: Writing – original draft. HY: Writing – review & editing, Conceptualization. LBZ: Writing – review & editing, Conceptualization. XMZ: Writing – review & editing, Conceptualization, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by National Natural Science Foundation of China (NSFC) to X-MZ (#32260050, #31900186), Yunnan Fundamental Research Projects X-MZ (Grant NO. 202301BF07001-016) and the Glory Light International Fellowship for Chinese Botanists at Missouri Botanical Garden (MO) to X-MZ.

Acknowledgments

We thank MO for providing some materials of Selaginellaceae; Rong Zhang, Jian-Jun Jing, and Jian-Jun Yang for their help with data analysis; and Xia Wan for helping to check specimens at MO. The authors also thank the two reviewers and editors for their critiques and suggestions which greatly improved our manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2024.1405253/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Maximum likelihood phylogeny of Selaginellaceae subf. Gymnogynoideae based on different datasets. Maximum likelihood bootstrap support values (MLBS) are shown above the branches and are 100/1.0 unless otherwise indicated. The MLBS from left to right were based on the dataset of codon12, gene, intergene, gene_intergene, codon12_RE, gene_RE, intergene_RE, gene_intergene_RE, codon12_RG, gene_RG, intergene_RG, gene_intergene_RG, codon12_REG, gene_REG, intergene_REG, and gene_intergene_REG, respectively.

SUPPLEMENTARY FIGURE 2

Phylogenetic tree reconstruction using Bayesian inference (BI) based on the dataset of codon12, codon12_RE, codon12_RG, and codon12_REG, respectively. Numbers above the branches represent Bayesian inference posterior probability (BIPP) and are 1.0/* unless otherwise indicated at the nodes.

SUPPLEMENTARY FIGURE 3

Phylogenetic tree reconstruction using Bayesian inference (BI) based on the dataset of intergene, intergene_RE, intergene_RG, and intergene_REG, respectively. Numbers above the branches represent Bayesian inference posterior probability (BIPP) and are 1.0/* unless otherwise indicated at the nodes.

SUPPLEMENTARY FIGURE 4

Phylogenetic tree reconstruction using Bayesian inference (BI) based on the dataset of gene, gene_intergene, gene_RE, gene_intergene_RE, gene_RG, gene_intergene_RG, gene_REG, and gene_intergene_REG, respectively. Numbers above the branches represent Bayesian inference posterior probability (BIPP) and are 1.0/* unless otherwise indicated at the nodes.

SUPPLEMENTARY FIGURE 5

Three alternative topologies for the *Ericetorum* (red) placement in the Selaginellaceae subf. Gymnogynoideae.

SUPPLEMENTARY FIGURE 6

Four alternative topologies for the *Gymnogynum* (red) placement in the Selaginellaceae subf. Gymnogynoideae.

SUPPLEMENTARY FIGURE 7

HybPiper recovery heatmap to visualize the recovery efficiency. Each row shows a sample, and each column is a gene. The amount of shading in each box corresponds to the length of the gene recovered for that sample, relative to the length of the reference.

SUPPLEMENTARY FIGURE 8

Maximum likelihood phylogeny of Selaginellaceae subf. Gymnogynoideae based on the datasets of SCGs_codon12, SCGs_gene, SCGs_intergene, and SCGs_gene_intergene. Maximum-likelihood bootstrap support values (MLBS) and Bayesian inference posterior probability (BIPP) are shown above the branches and are 100/1.0 unless otherwise indicated.

SUPPLEMENTARY TABLE 1

List of plastomes used in this study. Subgenus *Gymnogynum* in the row of "Subgenus (Zhou and Zhang, 2015)" is identical to *S. subg. Ericetorium* Jerm *sensu* Zhou and Zhang (2015).

SUPPLEMENTARY TABLE 2

Best-fitting likelihood model for maximum likelihood (ML) and Bayesian inference analyses (BI) for each dataset in this study.

SUPPLEMENTARY TABLE 3

List of the loci in each dataset.

SUPPLEMENTARY TABLE 4

Character coding of the taxa used in this study.

SUPPLEMENTARY TABLE 5

Features of plastomes used in this study.

SUPPLEMENTARY TABLE 6

Length of sequence recovered by HybPiper for each single-copy gene and sample in this study.

References

- Arrigo, N., Therrien, J., Anderson, C. L., Windham, M. D., Haufler, C. H., and Barker, M. S. (2013). A total evidence approach to understanding phylogenetic relationships and ecological diversity in *Selaginella* subg. *Tetragonostachys*. *Am. J. Bot.* 100, 1672–1682. doi: 10.3732/ajb.1200426
- Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Chan, P.-P., Lin, B.-Y., Mak, A. J., and Lowe, T. M. (2021). tRNAseScan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* 49, 9077–9096. doi: 10.1093/nar/gkab688
- Chen, S.-F., Zhou, Y.-Q., Chen, Y.-R., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Chu, W.-M. (2006). "Selaginellaceae," in *Flora Yunnanica*, vol. 20. Ed. C. Y. Wu (Beijing: Science Press), 35–93.
- Delsuc, F., Brinkmann, H., and Philippe, H. (2005). Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–375. doi: 10.1038/nrg1603
- Du, X.-Y., Lu, J.-M., and Li, D.-Z. (2020). Extreme plastid RNA editing may confound phylogenetic reconstruction: a case study of *Selaginella* (lycophytes). *Plant Divers.* 42, 356–361. doi: 10.1016/j.pld.2020.06.009

- Gielly, L., and Taberlet, P. (1994). The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbcL* sequences. *Mol. Biol. Evol.* 11, 769–777. doi: 10.1093/oxfordjournals.molbev.a040157
- Gonçalves, D. J. P., Simpson, B. B., Ortiz, E. M., Shimizu, G. H., and Jansen, R. K. (2019). Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. *Mol. Phylogenet. Evol.* 138, 219–232. doi: 10.1016/j.ympev.2019.05.022
- Goremykin, V. V., Nikiforova, S. V., and Bininda-Emonds, O. R. P. (2010). Automated removal of noisy data in phylogenomic analyses. *J. Mol. Evol.* 71, 319–331. doi: 10.1007/s00239-010-9398-z
- Heath, T. A., Hedtke, S. M., and Hillis, D. M. (2008). Taxon sampling and the accuracy of phylogenetic analyses. *J. Syst. Evol.* 46, 239–257. doi: 10.3724/SP.J.1002.2008.08016
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* 35, 518–522. doi: 10.1093/molbev/msx281
- Huson, D. H., and Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* 61, 1061–1067. doi: 10.1093/sysbio/syso62
- Jermey, A. C. (1986). Subgeneric names in *Selaginella*. *Fern Gazette* 13, 117–118.
- Jermey, A. C. (1990). “Selaginellaceae,” in *The families and genera of vascular plants, vol. 1, Pteridophytes and gymnosperms*. Eds. K. U. Kramer and P. S. Green (Springer, Berlin, Heidelberg & New York), 39–45.
- Bin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C. W., Yi, T.-S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. *Genome Biol.* 21, 241. doi: 10.1186/s13059-020-02154-5
- Johnson, M. G., Gardner, E. M., Liu, Y., Medina, R., Goffinet, B., Shaw, A. J., et al. (2016). HybPiper: extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Appl. Plant Sci.* 4, 1600016. doi: 10.3732/apps.1600016
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Kang, J.-S., Zhang, H.-R., Wang, Y.-R., Liang, S.-Q., Mao, Z.-Y., Zhang, X.-C., et al. (2020). Distinctive evolutionary pattern of organelle genomes linked to the nuclear genome in Selaginellaceae. *Plant J.* 104, 1657–1672. doi: 10.1111/tpj.15028
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kenrick, P., and Crane, P. R. (1997). *The origin and early diversification of land plants: a cladistic study* (Washington, DC, USA: Smithsonian Institution Press). doi: 10.1080/106351501753328875
- Klaus, K. V., Schulz, C., Bauer, D. S., and Stützel, T. (2017). Historical biogeography of the ancient lycophyte genus *Selaginella*: early adaptation to xeric habitats on Pangea. *Cladistics* 33, 469–480. doi: 10.1111/cla.12184
- Koenen, E. J. M., Ojeda, D. I., Steeves, R., Migliore, J., Bakker, F. T., Wieringa, J. J., et al. (2020). Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near simultaneous evolutionary origin of all six subfamilies. *New Phytol.* 225, 1355–1369. doi: 10.1111/nph.16290
- Korall, P., and Kenrick, P. (2004). The phylogenetic history of Selaginellaceae based on DNA sequences from the plastid and nucleus: extreme substitution rates and rate heterogeneity. *Mol. Phylogenet. Evol.* 31, 852–864. doi: 10.1016/j.ympev.2003.10.014
- Korall, P., Kenrick, P., and Therrien, J. P. (1999). Phylogeny of Selaginellaceae: evaluation of generic/subgeneric relationships based on *rbcL* gene sequences. *Int. J. Plant Sci.* 160, 585–594. doi: 10.1086/314137
- Liu, B.-B., Ren, C., Kwak, M., Hodel, R. G., Xu, C., He, J., 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. doi: 10.1111/jipb.13246
- Morris, J. L., Puttick, M. N., Clark, J. W., Edwards, D., Kenrick, P., Pressel, S., et al. (2018). The timescale of early land plant evolution. *Proc. Natl. Acad. Sci. U. S. A.* 115, E2274–E2283. doi: 10.1073/pnas.1719588115
- Mower, J. P., Ma, P.-F., Grewe, F., Taylor, A., Michael, T. P., VanBuren, R., et al. (2019). Lycophyte plastid genomics: extreme variation in GC, gene and intron content and multiple inversions between a direct and inverted orientation of the rRNA repeat. *New Phytol.* 222, 1061–1075. doi: 10.1111/nph.15650
- Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Oldenkott, B., Yamaguchi, K., Tsuji-Tsukinoki, S., Knie, N., and Knoop, V. (2014). Chloroplast RNA editing going extreme: more than 3400 events of C-to-U editing in the chloroplast transcriptome of the lycophyte *Selaginella uncinata*. *RNA* 20, 1499–1506. doi: 10.1261/rna.045575.114
- Philippe, H., Lopez, P., Brinkmann, H., Budin, K., Germot, A., Laurent, J., et al. (2000). Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. *Proc. Biol. Sci.* 267, 1213–1221. doi: 10.1098/rspb.2000.1130
- PPG I (2016). A community-derived classification for extant lycophytes and ferns. *J. Syst. Evol.* 54, 563–603. doi: 10.1111/jse.12229
- Rambaut, A. (2017). *FigTree-version 1.4.3*. Available at: <http://tree.bio.ed.ac.uk/software/figtree/>.
- Revell, L. J. (2011). Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* 3, 217–223. doi: 10.1111/j.2041-210X.2011.00169.x
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542. doi: 10.1093/sysbio/syso29
- Rothfels, C. J., Li, F.-W., Sigel, E. M., Huiet, L., Larsson, A., Burge, D. O., et al. (2015). The evolutionary history of ferns inferred from 25 low-copy nuclear genes. *Am. J. Bot.* 102, 1089–1107. doi: 10.3732/ajb.1500089
- Schmidt, A. R., Korall, P., Krings, M., Weststrand, S., Bergschneider, L., Sadowski, E. M., et al. (2022). *Selaginella* in cretaceous amber from Myanmar. *Willdenowia* 52, 179–245. doi: 10.3372/wi.52.52203
- Schmidt, A. R., Regalado, L., Weststrand, S., Korall, P., Sadowski, E. M., Schneider, H., et al. (2020). *Selaginella* was hyperdiverse already in the Cretaceous. *New Phytol.* 228, 1176–1182. doi: 10.1111/nph.16600
- Schuettpelz, E., and Pryer, K. M. (2009). Evidence for a Cenozoic radiation of ferns in an angiosperm-dominated canopy. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11200–11205. doi: 10.1073/pnas.0811136106
- Shen, X.-X., Hittinger, C. T., and Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nat. Ecol. Evol.* 1, 1–10. doi: 10.1038/s41559-017-0126
- Simmons, M. P., and Freudenstein, J. V. (2011). Spurious 99% bootstrap and jackknife support for unsupported clades. *Mol. Phylogenet. Evol.* 61, 177–191. doi: 10.1016/j.ympev.2011.06.003
- Simmons, M. P., Pickett, K. M., and Miya, M. (2004). How meaningful are Bayesian support values? *Mol. Biol. Evol.* 21, 188–199. doi: 10.1093/molbev/msh014
- Simmons, M. P., Zhang, L.-B., Webb, C. T., and Reeves, A. (2006b). How can third codon positions outperform first and second codon positions in phylogenetic inference? An empirical example from the seed plants. *Syst. Biol.* 55, 245–258. doi: 10.1080/10635150500481473
- Simmons, M. P., Zhang, L.-B., Webb, C. T., Reeves, A., and Miller, J. A. (2006a). The relative performance of Bayesian and parsimony approaches when sampling characters evolving under homogeneous and heterogeneous sets of parameters. *Cladistics* 22, 171–185. doi: 10.1111/j.1096-0031.2006.00098.x
- Smith, D. R. (2009). Unparalleled GC content in the plastid DNA of *Selaginella*. *Plant Mol. Biol.* 71, 627–639. doi: 10.1007/s11103-009-9545-3
- Smith, S. A., Moore, M. J., Brown, J. W., and Yang, Y. (2015). Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evol. Biol.* 15, 150. doi: 10.1186/s12862-015-0423-0
- Tang, J.-Y., Wei, R., Zhang, X.-C., and Xiang, Q.-P. (2023). Mitogenome-based phylogenomics provides insights into the positions of the enigmatic *sinensis* group and the *sanguinolenta* group in Selaginellaceae (Lycophyte). *Mol. Phylogenet. Evol.* 179, 107673. doi: 10.1016/j.ympev.2022.107673
- Taylor, T. N., Taylor, E. L., and Krings, M. (2009). “Paleobotany,” in *The biology and evolution of fossil plants, 2nd edn* (Elsevier/Academic Press, New York NY, USA).
- Testo, W., and Sundue, M. (2016). A 4000-species dataset provides new insight into the evolution of ferns. *Mol. Phylogenet. Evol.* 105, 200–211. doi: 10.1016/j.ympev.2016.09.003
- Thomas, B. A. (1992). Paleozoic herbaceous lycopsids and the beginnings of extant *Lycopodium* sens. lat. and *Selaginella* sens. lat. *Ann. Mo. Bot. Gard.* 79, 623–631. doi: 10.2307/2399756
- Thomas, B. A. (1997). Upper Carboniferous herbaceous lycopsids. *Rev. Palae. Pale.* 95, 129–153. doi: 10.1016/S0034-6667(96)00032-2
- Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., et al. (2017). GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 45, W6–W11. doi: 10.1093/nar/gkx391
- Tryon, R. M. (1955). *Selaginella rupestris* and its allies. *Ann. Mo. Bot. Gard.* 42, 1–99. doi: 10.2307/2477710
- Tryon, A. F., and Lugardon, B. (1991). *Spores of the pteridophyta: surface, wall structure, and diversity based on electron microscope studies* (New York: Springer), 606–621. doi: 10.1007/978-1-4613-8991-0
- Tsuiji, S., Ueda, K., Nishiyama, T., Hasebe, M., Yoshikawa, S., Konagaya, A., et al. (2007). The chloroplast genome from a lycophyte (microphylophyte), *Selaginella uncinata*, has a unique inversion, transpositions and many gene losses. *J. Plant Res.* 120, 281–290. doi: 10.1007/s10265-006-0055-y
- Weststrand, S., and Korall, P. (2016a). Phylogeny of Selaginellaceae: there is value in morphology after all! *Am. J. Bot.* 103, 2136–2159. doi: 10.3732/ajb.1600156
- Weststrand, S., and Korall, P. (2016b). Subgeneric classification of *Selaginella* (Selaginellaceae). *Am. J. Bot.* 103, 2160–2169. doi: 10.3732/ajb.1600288
- Wick, R. R., Schultz, M. B., Zobel, J., and Holt, K. E. (2015). Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31, 3350–3352. doi: 10.1093/bioinformatics/btv383

- Wood, D., Besnard, G., Beerling, D. J., Osborne, C. P., and Christin, P. A. (2020). Phylogenomics indicates the "living fossil" *Isoetes* diversified in the Cenozoic. *PLoS One* 15, e0227525. doi: 10.1371/journal.pone.0227525
- Xiang, Q.-P., Tang, J.-Y., Yu, J.-G., Smith, D. R., Zhu, Y.-M., Wang, Y.-R., et al. (2022). The evolution of extremely diverged plastomes in Selaginellaceae (lycophyte) is driven by repeat patterns and the underlying DNA maintenance machinery. *Plant J.* 111, 768–784. doi: 10.1111/tj.15851
- Xu, Z.-C., Xin, T.-Y., Bartels, D., Li, Y., Gu, W., Yao, H., et al. (2018). Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of desiccation tolerance. *Mol. Plant* 11, 983–994. doi: 10.1016/j.molp.2018.05.003
- Yang, Y.-Y., Qu, X.-J., Zhang, R., Stull, G. W., and Yi, T.-S. (2021). Plastid phylogenomic analyses of Fagales reveal signatures of conflict and ancient chloroplast capture. *Mol. Phylogenet. Evol.* 163, 107232. doi: 10.1016/j.ympev.2021.107232
- Yang, J., Xiang, Q.-P., and Zhang, X.-C. (2022). Uncovering the hidden diversity of the rosette-forming *Selaginella tamariscina* group based on morphological and molecular data. *Taxon* 72, 8–19. doi: 10.1002/tax.12817
- Yu, Y., and Nakhleh, L. (2015). A maximum pseudo-likelihood approach for phylogenetic networks. *BMC Genomics* 16, S10. doi: 10.1186/1471-2164-16-S10-S10
- Zhang, X.-C., Kato, M., and Nootboom, H. P. (2013). "Selaginellaceae," in *Flora of China*, vol. 2–3. Eds. Z. Y. Wu, P. H. Raven and D. Y. Hong (Science Press; St. Louis: Missouri Botanical Garden Press, Beijing), 37–66.
- Zhang, R., Wang, Y.-H., Jin, J.-J., Stull, G. W., Bruneau, A., Cardoso, D., et al. (2020b). Exploration of plastid phylogenomic conflict yields new insights into the deep relationships of Leguminosae. *Syst. Biol.* 69, 613–622. doi: 10.1093/sysbio/syaa013
- Zhang, M.-H., Wei, R., Xiang, Q.-P., Ebihara, A., and Zhang, X.-C. (2021). Integrative taxonomy of the *Selaginella helvetica* group based on morphological, molecular and ecological data. *Taxon* 71, 1163–1187. doi: 10.1002/tax.12565
- Zhang, H.-R., Wei, R., Xiang, Q.-P., and Zhang, X.-C. (2020a). Plastome-based phylogenomics resolves the placement of the *sanguinolenta* group in the spikemoss of lycophyte (Selaginellaceae). *Mol. Phylogenet. Evol.* 147, 106788. doi: 10.1016/j.ympev.2020.106788
- Zhang, H.-R., Xiang, Q.-P., and Zhang, X.-C. (2019b). The unique evolutionary trajectory and dynamic conformations of DR and IR/DR-coexisting plastomes of the early vascular plant Selaginellaceae (Lycophyte). *Genome Biol. Evol.* 11, 1258–1274. doi: 10.1093/gbe/evz073
- Zhang, M.-H., Xiang, Q.-P., and Zhang, X.-C. (2022). Plastid phylogenomic analyses of the *Selaginella sanguinolenta* group (Selaginellaceae) reveal conflict signatures resulting from sequence types, outlier genes, and pervasive RNA editing. *Mol. Phylogenet. Evol.* 173, 107507. doi: 10.1016/j.ympev.2022.107507
- Zhang, M.-H., and Zhang, X.-C. (2022). Integrative species delimitation of *Selaginella labordei* and closely related species: uncovering the mysterious identity of *S. jugorum* and *S. tibetica*, and description of a new species. *Taxon* 71, 1155–1169. doi: 10.1002/tax.12800
- Zhang, H.-R., Zhang, X.-C., and Xiang, Q.-P. (2019a). Direct repeats co-occur with few short-dispersed repeats in plastid genome of a spikemoss, *Selaginella vardei* (Selaginellaceae, Lycopodiopsida). *BMC Genomics* 20, 484. doi: 10.1186/s12864-019-5843-6
- Zhao, J., Zhou, X.-M., Fang, S.-L., Zhu, Z.-M., Li, Y.-X., Yu, H., et al. (2023). Transcriptome-based study on the phylogeny and hybridization of Marattiacean ferns (Marattiaceae). *Plants* 12, 2237. doi: 10.3390/plants12122237
- Zhou, X.-M., Rothfels, C. J., Zhang, L., He, Z.-R., Le Pêcheon, T., He, H., et al. (2016). A large-scale phylogeny of the lycophyte genus *Selaginella* (Selaginellaceae: Lycopodiopsida) based on plastid and nuclear loci. *Cladistics* 32, 360–389. doi: 10.1111/cla.12136
- Zhou, X.-M., and Zhang, L.-B. (2015). A classification of *Selaginella* (Selaginellaceae) based on molecular (chloroplast and nuclear), macro-morphological, and spore features. *Taxon* 64, 1117–1140. doi: 10.12705/646.2
- Zhou, X.-M., and Zhang, L.-B. (2023). Phylogeny, character evolution, and classification of Selaginellaceae (lycophytes). *Plant Divers.* 45, 630–684. doi: 10.1016/j.pld.2023.07.003
- Zhou, X.-M., Zhao, J., Yang, J.-J., Le Pêcheon, T., Zhang, L., He, Z.-R., et al. (2022). Plastome structure, evolution, and phylogeny of *Selaginella*. *Mol. Phylogenet. Evol.* 169, 107410. doi: 10.1016/j.ympev.2022.107410