



OPEN ACCESS

EDITED BY
Zhiqiang Wu,
Chinese Academy of Agricultural
Sciences, China

REVIEWED BY
Rongpeng Liu,
Beijing University of Chinese Medicine,
China
Jia Minlong,
Shanxi Agricultural University, China

*CORRESPONDENCE
Fang Wen
✉ wenfang760608@139.com
Xinxiang Bai
✉ xxbai@gzu.edu.cn
Longfei Fu
✉ longfeifu@126.com

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 12 December 2025
REVISED 01 March 2026
ACCEPTED 04 March 2026
PUBLISHED 23 March 2026

CITATION
Xin Z, Qian X, Li X, Fu L, Bai X and Wen F
(2026) Comparative and phylogenetic
analysis of chloroplast genomes in the
subtribe Leptoboeinae (Gesneriaceae).
Front. Plant Sci. 17:1766257.
doi: 10.3389/fpls.2026.1766257

COPYRIGHT
© 2026 Xin, Qian, Li, Fu, Bai and Wen.
This is an open-access article distributed
under the terms of the [Creative
Commons Attribution License \(CC BY\)](#).
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.

Comparative and phylogenetic analysis of chloroplast genomes in the subtribe Leptoboeinae (Gesneriaceae)

Zibing Xin^{1,2†}, Xiaomao Qian^{1,2,3†}, Xiaojuan Li^{1,2}, Longfei Fu^{1,2*},
Xinxiang Bai^{3*} and Fang Wen^{1,2*}

¹Guangxi Key Laboratory of Plant Conservation and Restoration Ecology in Karst Terrain, Guangxi
Institute of Botany, Guangxi Zhuang Autonomous Region and Chinese Academy of Sciences, Guilin,
Guangxi, China, ²Gesneriad Committee of China Wild Plant Conservation Association, National
Gesneriaceae Germplasm Resources Bank of GXIB, The Gesneriad Conservation Center of China,
Guilin Botanical Garden, Chinese Academy of Sciences, Guilin, Guangxi, China, ³College of Forestry,
Guizhou University, Guiyang, Guizhou, China

The subtribe Leptoboeinae (Gesneriaceae: Trichosporeae) comprises seven genera—*Rhynchosyche* Blume, *Boeica* C.B. Clarke, *Beccarinda* Kuntze, *Leptoboea* Benth., *Platystemma* Wall., *Crassicaulis* Lei Cai & Mich. Moller, and *Championia* Gardner. It is highly valuable in horticulture and medicine. However, long-standing taxonomic issues, including the polyphyly of *Boeica* and the uncertain placement of *Championia*, remain unresolved. To address these problems, we sequenced and assembled complete chloroplast genomes for 17 species representing all seven genera, and conducted comparative genomic and phylogenomic analyses. The genomes ranged from 153,114 to 154,647 bp and exhibited a typical quadripartite structure consistent with other Gesneriaceae. A total of 1,861 SSRs, 1,179 dispersed repeats, and 29 high-frequency codons with similar usage patterns were identified. We also detected 10 highly variable regions that may serve as potential molecular markers. The plastid tree and coalescent tree constructed based on CDS sequences in this study showed highly congruent topologies at the generic level, supporting the division of *Boeica* into five clades, two of which can be distinguished morphologically. *Championia* formed a sister relationship with all sampled genera of Trichosporeae and Epithemateae, confirming that it does not belong to Leptoboeinae. Moreover, structural variation in the plastome helped clarify the polyphyly of *Boeica*: the five clades exhibited significant differences in total GC content, LSC GC content, and SSC GC content. Given the overlapping geographic distributions of these clades, we infer that plastome divergence may reflect independent lineage evolution. Overall, this study offers new insights into the phylogenetic relationships within Leptoboeinae and provides a foundation for future taxonomic revision and evolutionary research.

KEYWORDS

comparative chloroplast genomics, Leptoboeinae, molecular markers, phylogenetic relationships, polyphyly

1 Introduction

The subtribe Leptoboeinae was established in 1884 by Clarke, based on the characteristics that its capsules dehisce into 2 valves septicidally and its seeds have no appendages (Clarke, 1884). Later, Fritsch (Fritsch, 1894) merged the genera *Championia*, *Platystemma*, *Boeica*, and *Leptoboea* into the Championieae on the basis of their narrowly elongate calyx, 4 fertile stamens, and dehiscent fruits, and later transferred them to the Trichosporeae. At present, phylogenetic relationships within *Primulina* Hance and *Petrocodon* Hance (Trichosporeae) are relatively clear (Gu et al., 2024; Hsieh et al., 2022). However, members of Leptoboeinae are scattered across multiple countries along the southern slopes of the Himalayas, southern China, the Indochinese Peninsula, and two countries in the Indian subcontinent (Wang et al., 1990; POWO, 2025). The difficulty of collecting representative samples has long hindered research on this subtribe, to the extent that it has been largely overlooked, whether deliberately or unintentionally. In modern classification of Gesneriaceae, Leptoboeinae is treated as a subtribe of Trichosporeae and comprises seven genera—*Radiaticorollarius*, *Rhynchotechum*, *Boeica*, *Beccarinda*, *Leptoboea*, *Platystemma*, *Championia*, and *Crassicaulis* (Weber et al., 2013; Yang et al., 2025), encompassing approximately 48 species, distributed mainly in southern and southwestern China (Wang et al., 1990, 1998; Weber et al., 2013; Wen et al., 2022), and elsewhere in the above-mentioned regions. Members of this subtribe show striking morphological diversification and conspicuous differentiation, but they share relatively consistent traits such as an inconspicuous or capitate stigma, capsules dehiscing loculicidally into 2 valves, and 4 stamens (Weber et al., 2013), which distinguish them from other subtribes. Some genera of Leptoboeinae are traditionally used in folk medicine. For instance, many species of *Rhynchotechum* and *Boeica fulva* C. B. Clarke have medicinal or even edible value. For example, *Rhynchotechum vestitum* Wall. ex C. B. Clarke can be used in the treatment of hepatitis A and hepatitis B (Lu et al., 1998). At the generic level, the placement of *Championia* has remained controversial. Such as in subtribe Didymocarpeae of tribe Cyrtandreae by Bentham (1876); subtribe Championiinae of tribe Championieae by Fritsch 1894; later it was reassigned by different researchers to the tribe Didymocarpeae of subfamily Cyrtandroideae (Burt, 1963; Ivanina, 1965; Burt and Wiehler, 1995). Weber et al. (2013) on the basis of the tetramerous corolla, suggested its possible affinity with Leptoboeinae, but molecular phylogenetic evidence was lacking. Ranasinghe et al. (2024) using phylogenies based on *ndhF*, *matK*, *rps16*, and *trnL-F* sequences, found *Championia* to fall outside the Leptoboeinae clade. Moreover, plastome phylogenies based on protein-coding genes and coding regions indicated that *Radiaticorollarius* forms a sister group with some species of *Boeica* and *Rhynchotechum* (Wen et al., 2022), whereas whole plastome data (excluding one inverted repeat) supported *Radiaticorollarius* as

sister to *Leptoboea* (Cui et al., 2023). However, neither relationship is well supported, resulting in continued uncertainty over its circumscription. Accordingly, Weber et al. (2025) after re-evaluating the morphological characteristics and phylogenetic relationships of *Radiaticorollarius*, treated it as part of the genus *Boeica*. *Boeica* itself was established by Clarke in 1874 on the basis of floral traits, including an almost actinomorphic corolla, four free stamens, and cylindrical capsules dehiscing dorsally along both locules. Several studies have shown that *Boeica* is not monophyletic, with some of its species forming a sister group to *Rhynchotechum* (Yang et al., 2020; Wen et al., 2022; Yang et al., 2023a), but Weber et al. (2025) argued that the sampling of *Boeica* and *Rhynchotechum* in the studies by Yang et al. (2020; 2023a) and Wen et al. (2022) was insufficient, and thus their results require further verification (Middleton et al., 2015), using ITS and *trnL-F*, also found *Boeica* and *Rhynchotechum* to be closely related and noted that several species of both genera exhibit alternating phyllotaxy with opposite leaves. Furthermore, Yang et al. (2025) reconstructed phylogenetic trees using plastid genome and low-copy nuclear (LCN) gene datasets, which revealed a complex reticulate evolutionary history within the subtribe and confirmed the polyphyly of *Boeica*, although no further analysis of this polyphyly was conducted.

With the rapid rise of next-generation DNA sequencing (NGS) technologies, the cost of sequencing has dropped dramatically, driving the swift advancement of genomics and enabling the application of large-scale molecular datasets in phylogenetic research, thereby giving rise to the field of phylogenomics (Sun et al., 2022; Wang et al., 2024a; Torruella et al., 2025). Chloroplasts are ubiquitous organelles in plants, and chloroplast genomes, due to their structural stability, moderate evolutionary rate, sequence conservation, good collinearity, and predominantly maternal inheritance, have become highly valuable in phylogenetic studies. They not only enhance the support and resolution of phylogenetic trees but also play an important role in elucidating species evolution and interspecific relationships (Zhai et al., 2021; Wu et al., 2024; Cheng et al., 2025; Dong et al., 2025; Xia et al., 2025).

Therefore, in response to the above issues and current situation, the main objectives of this study are: (1) to construct a phylogenetic framework of Leptoboeinae based on different chloroplast genome datasets, investigate conflicts between plastid trees and coalescent trees, and elucidate the phylogenetic relationships among genera within the subtribe; (2) to identify structural variations and potential phylogenetic signals through comparative analyses of chloroplast genomes, in order to clarify the polyphyly of *Boeica*; and (3) to detect highly variable regions of the chloroplast genome in Leptoboeinae and screen for potential DNA barcode markers.

2 Materials and methods

2.1 Plant materials and DNA extraction

In this study, we conducted species-level sampling of all known genera within the subtribe Leptoboeinae, collecting 17 species representing 7 genera (Supplementary Table 2), thereby covering

Abbreviations: BI, Bayesian inference; IR, Inverted repeat; JLB, LSC and IRb boundary; JSB, IRb and SSC boundary; JSA, SSC and IRa boundary; JLA, IRa and LSC boundary; LSC, Larger single copy; ML, Maximum likelihood; nrDNA, nuclear ribosomal DNA; PP, Posterior probability; RSCU, Relative Synonymous Codon Usage; SSC, Small single copy.

all genera of the subtribe. To enhance dataset completeness, the 17 newly sequenced chloroplast genomes were integrated with previously published chloroplast genome data from 69 accessions of Leptoboeinae (Wen et al., 2022; Yang et al., 2025). This resulted in a final dataset of 86 chloroplast genomes representing 37 species (77.08% of the subtribe). The sampling included 1 species of *Championia* (100% of the genus), 1 species of *Platystemma* (100% of the genus), 1 species of *Crassicaulis* (100% of the genus), 8 species of *Beccarinda* (75% of the genus), 13 species of *Boeica* (76.47% of the genus), 2 species of *Leptoboea* (50% of the genus), 11 species of *Rhynchoetechum* (61.11% of the genus). Among these, 37 representative samples were selected for chloroplast genomes structural analyses. Detailed information is provided in [Supplementary Table 2](#).

Total genomic DNA of each sample was extracted using the modified CTAB method (Doyle and Doyle, 1987). A 1% agarose gel was prepared, and electrophoresis was performed to assess the quality of the extracted DNA, while DNA concentration and purity were measured using a Nanodrop spectrophotometer. Qualified samples were then sent to Sangon Biotech (Shanghai) Co., Ltd. for genomic library construction and high-throughput sequencing.

2.2 Sequencing, assembly and annotation

DNA libraries were sequenced on the Illumina high-throughput platform using paired-end 150 bp (PE) sequencing, generating 4 GB of raw data for each sample. The quality of the raw reads was assessed using FASTQ 0.36 (Chen et al., 2018), and adapters as well as low-quality reads were filtered with Trimmomatic v0.39 (Bolger et al., 2024) to obtain high-quality data (clean reads). Clean reads were assembled with GetOrganelle v1.7.5 (Jin et al., 2020) using the parameters “-R 20 -k 21, 45, 65, 85, 105, 127 -F embplant_pt.” Redundant contigs were manually removed with Bandage to obtain complete chloroplast genomes. Using *Beccarinda tonkinensis* (Pellegr.) Burt (GenBank: NC_072338) as the reference, 30 chloroplast genome sequences were annotated with the GeSeq online annotation tool (Tillich et al., 2017). Coding sequence (CDS) start and stop codons were manually checked and adjusted in Geneious 9.0.2 (Kearse et al., 2012), while the type, length, and anticodon of tRNA genes were verified and corrected with tRNAscan SE (Chan and Lowe, 2019). The annotated information was visualized with OGDRAW v1.3.163 (Greiner et al., 2019). To assemble the nrDNA repeat sequences from Illumina reads, GetOrganelle v1.7.5 (Jin et al., 2020) was also used with default settings and the parameters “-R 20 -k 21, 45, 65, 85, 105, 127 -F embplant_nr,” resulting in contigs for each sample. These contigs were then imported into Geneious 9.0.2 and aligned against the nrDNA reference sequence of *Beccarinda tonkinensis* (GenBank: KJ475423) for validation.

2.3 Chloroplast genome IR/SC boundary analysis, similarity comparison, and identification of highly variable regions

The boundaries of the 37 chloroplast genomes were compared and visualized using CPJSDraw (Li et al., 2023). To identify and visually detect interspecific variation regions in the chloroplast

genomes, the online software mVISTA (<https://genome.lbl.gov/vista/mvista/submit.shtml>) (Frazer et al., 2004) was employed, with the annotation of *Beccarinda argentea* XZB2310 as the reference and LAGAN selected as the alignment program, to generate visual similarity comparisons of the 37 chloroplast genomes. Furthermore, DnaSP v.6.12.03 (Rozas et al., 2017) was used to analyze the aligned sequence matrix of the 37 chloroplast genomes in order to further quantify nucleotide polymorphism levels and detect highly variable regions within the subtribe Leptoboeinae. The sliding window analysis was performed with a window length of 800 bp and a step size of 200 bp, and the top 5% of windows with the highest *Pi* values were considered mutational hotspot regions (Wang et al., 2022; Gu et al., 2024).

2.4 Codon usage bias analysis

To quantify codon usage bias, CodonW 1.4.2 (Peden, 1999) was used to analyze and filter 37 CDS sequences. Since short sequences cannot accurately calculate the effective number of codons (Liu, 2013), only CDS longer than 300 bp were selected for Relative Synonymous Codon Usage (RSCU) analysis in this study. RSCU represents the ratio between the observed frequency of a codon and its expected frequency under equal usage (Sharp and Li, 1987). When $RSCU > 1$, the codon is used more frequently and considered a high-frequency codon; when $RSCU < 1$, the codon is used less frequently and considered a low-frequency codon; and when $RSCU = 1$, the codon shows no usage bias (Mazumdar et al., 2017).

2.5 Repeat sequence analysis

Simple sequence repeats (SSRs) were detected using Krait (Du et al., 2018) with the following parameters: mononucleotide repeat units ≥ 10 , dinucleotide repeat units ≥ 5 , trinucleotide repeat units ≥ 4 , and tetra-, penta-, and hexanucleotide repeat units ≥ 3 . Dispersed repeats, which are distributed throughout the genome, include 4 types: forward (F), reverse (R), palindromic (P), and complement (C) repeats (Jelinek et al., 1980). Dispersed repeats were identified using the online REPuter software (Kurtz et al., 2001), with a minimum repeat size of 30 bp and a sequence identity threshold of 90% (Hamming distance = 3).

2.6 Phylogenetic analysis

To investigate the systematic issues within the subtribe Leptoboeinae, including intergeneric relationships and generic delimitation, we conducted phylogenetic analyses based on 103 chloroplast genomes and 107 nrDNA sequences. The chloroplast dataset included 37 species (86 samples) of Leptoboeinae as the ingroup and 17 species from the closely related tribes Trichosporeae, Epithemateae, and Gesnerieae, as well as Scrophulariaceae and Lamiaceae, as outgroups. The nrDNA dataset comprised 37 species (86 samples) of Leptoboeinae and 21 outgroup species corresponding to those used in the chloroplast dataset. These outgroups were selected because they are known to share close phylogenetic affinities with Leptoboeinae, providing a robust framework for the analyses (Ranasinghe et al., 2024), while species from Scrophulariaceae and Lamiaceae were included to root

the phylogenetic trees (Li et al., 2021). Four datasets were employed for phylogenetic analyses: (a) the nrDNA region 18S–ITS1–5.8S–ITS2–26S, (b) 81 concatenated CDS derived from 37 chloroplast genomes, (c) 81 CDS analyzed using a coalescent-based approach. The concatenation method was used based on the traditional view that chloroplast genomes evolve as a single unit, which was thought to avoid conflicts between gene trees and species trees (Doyle, 1992; Wicke et al., 2011), and (d) complete chloroplast genome sequences (with one IR region removed). However, recent studies have shown inconsistencies between gene trees and species trees constructed from chloroplast genomes (Walker et al., 2019; Xue et al., 2024). Therefore, this study used concatenation and coalescent methods based on chloroplast genome data to detect any structural inconsistencies. Phylogenetic trees were reconstructed using Bayesian inference (BI) and maximum likelihood (ML) methods. Multiple sequence alignments of the chloroplast genomes and nrDNA data were performed with MAFFT v.7.5.1.1 (Katoh and Standley, 2013) using default parameters. The alignments were manually checked and corrected in MEGA 11.0.13 (Tamura et al., 2021) to obtain final alignment matrices. DAMBE v5.3.19 (Xia, 2013) was used to evaluate the substitution saturation index (Iss) of the data matrices, with results indicating that the observed index of substitution saturation (Iss = 0.0800) was significantly lower than the critical Iss.c value (0.8610; $P = 0.0000 < 0.05$), suggesting that the data were not saturated and were suitable for phylogenetic tree reconstruction.

The optimal substitution model for both maximum likelihood (ML) analysis and Bayesian inference (BI) was determined using PhyloSuite v.1.2.3 (Zhang et al., 2020) based on the AICc criterion, with GTR+I+G identified as the best-fit model. ML and BI phylogenetic trees were constructed using IQ-TREE 1.6.12 and MrBayes 3.2.7, respectively, via PhyloSuite v.1.2.3. Branch support in the ML analysis was assessed using standard 1,000 bootstrap replicates (BS). BI inference was run for 1,000,000 generations of MCMC, sampling every 1,000 generations, and convergence was verified with ESS > 200. After discarding the first 25% of samples as burn-in, posterior probabilities (PP) were used to evaluate branch reliability. Single-gene trees were inferred in IQ-TREE 1.6.12 with 1,000 bootstrap replicates and automatic model selection. Coalescent tree inference was performed in ASTRAL-III v5.7.8 (Zhang et al., 2018), with branch support evaluated using local posterior probabilities (LPP; Sayyari and Mirarab, 2016) and quartet scores. Quartet frequencies were used to detect gene tree conflicts, reflecting the degree of incongruence among gene trees. Bootstrap values from gene trees were mapped onto the ASTRAL topology, and phypartspiecharts.py was used to generate pie charts visualizing the proportion of gene trees supporting each topology. Phylogenetic trees were visualized using Figtree v.1.4.4 (<https://github.com/rambaut/figtree/releases>, accessed 18 January 2025) and iTOL (Letunic and Bork, 2021).

2.7 Comparative analysis of chloroplast genomes

Samples used for chloroplast comparative genomic analyses were grouped according to the clades defined in Figure 1: E1 (2 species, 5 samples), E2 (2 species, 8 samples), E3 (1 species,

2 samples), E4 (6 species, 12 samples), and E5 (2 species, 3 samples). The chloroplast genome size, LSC, SSC, IR regions, and their corresponding GC contents were summarized for each clade and visualized using EcoAmp v0.13. We further compiled these genomic characteristics for all clades (A–G) of Leptoboeinae and generated boxplots in R v3.6.1. One-way ANOVA was performed in IBM SPSS Statistics v27, followed by Tamhane's T2 *post-hoc* test without assuming equal variances, to assess significant differences among genera and subclades. Details and results are provided in Supplementary Figure 3 and Supplementary Table 6.

3 Result

3.1 Structure and characteristics of chloroplast genomes

In this study, the chloroplast genomes of 37 species of Leptoboeinae were characterized, of which 17 were newly generated (Supplementary Tables 1, 2). The total length of the chloroplast genomes ranges from 153,114 bp (*Platystemma violoides* Wall., B8) to 154,647 bp (*Rhynchotechum nirijuliense* Taram & D. Borah, B9), exhibiting a typical quadripartite circular structure (Figure 2). Each genome consisted of a large single-copy region (LSC) of 84,457–85,796 bp, a small single-copy region (SSC) of 17,691–19,569 bp, and two inverted repeats (IRs) of 24,635–25,534 bp. The overall GC content of the 37 chloroplast genomes was similar (Supplementary Table 1), ranging from 37.28% to 37.9%. The GC contents of the LSC and SSC regions were 35.17%–35.94% and 31.06%–31.94%, respectively, whereas that of the IR regions was higher, at 43.22%–43.29%. For protein-coding sequences (CDSs), the total GC content ranged from 37.28% to 37.9%, with codon positions 1, 2, and 3 (GC1, GC2, and GC3) showing GC contents of 34.16%–45.68%, 31.9%–42.34%, and 29.84%–42.28%, respectively (Supplementary Table 1).

Each of the 37 chloroplast genomes was annotated with 130–139 functional genes, including 84–89 protein-coding genes (PCGs), 37 transfer RNA (tRNA) genes, and 8 ribosomal RNA (rRNA) genes (Supplementary Table 1). Based on the different functions of the genes, these genes were classified into protein synthesis and DNA replication genes (71–78), photosynthesis-related genes (44–46), other protein-coding genes (6), and genes of unknown function (6–9) (Supplementary Table 3). Among the 130–139 functional genes, 15–17 genes contained a single intron, while 3 genes (*clpP1*, *ycf3*, and *rps12*) contained 2 introns.

3.2 Analysis of simple sequence repeats and dispersed repeats

This study analyzed the simple sequence repeats (SSRs) and dispersed repeat sequences in the chloroplast genomes of 37 species from 7 genera of the subtribe Leptoboeinae.

A total of 1861 SSRs were detected (Supplementary Table 4), with each species containing 37 *Boeica multinervia* K. Y. Pan (G39) to 69 *Championia reticulata* Gardner (G32) SSRs. These included 18–44 mononucleotide (Mono) repeats, 5–14 dinucleotide (Di) repeats,

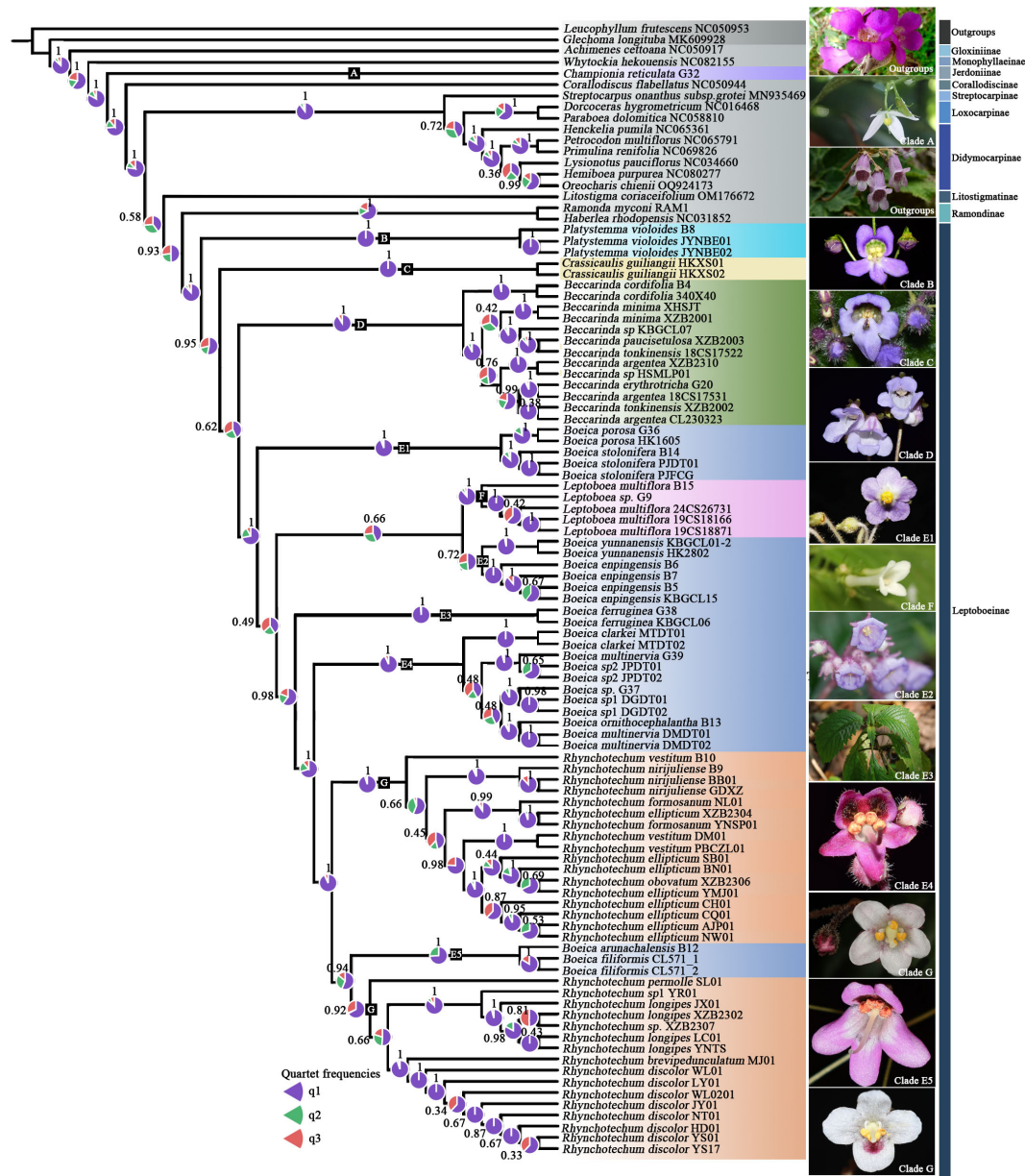


FIGURE 1

Coalescent tree: phylogenetic relationships of the subtribe Leptoboeinae inferred from 81 CDSs, with branch values indicating local posterior probability (LPP) support; pie charts around branches show the relative frequencies of three alternative topologies (purple, consistent with the coalescent tree; orange, first alternative topology; green, second alternative topology).

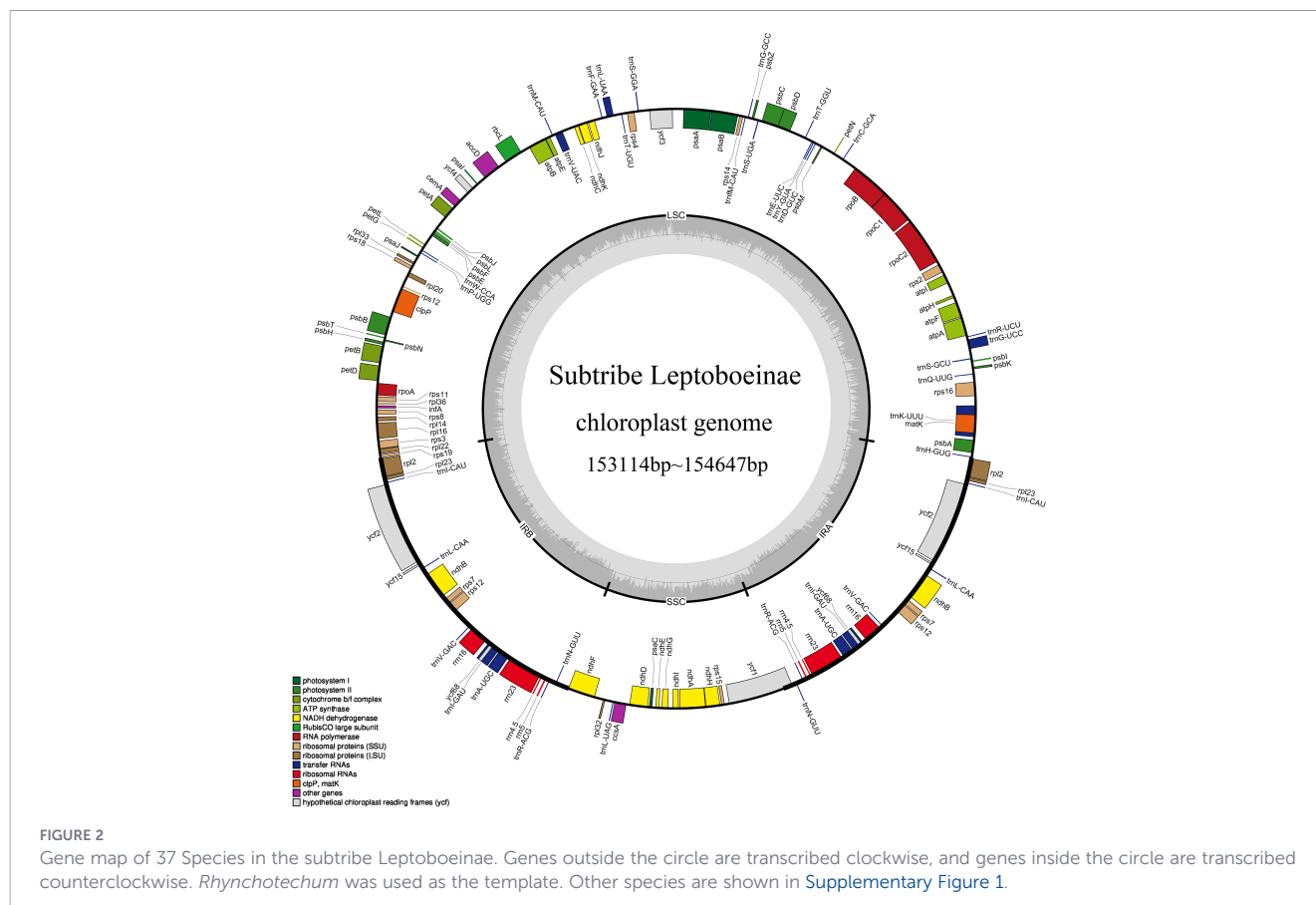
1–9 trinucleotide (Tri) repeats, 0–13 tetranucleotide (Tetra) repeats, 0–3 pentanucleotide (Penta) repeats, and 0–1 hexanucleotide (Hexa) repeats (Figure 3A). The repeat units were mainly composed of adenine (A) and thymine (T), with mononucleotide A/T repeats being the most abundant, followed by dinucleotide AT/AT repeats (Figure 3B). Among the chloroplast genomes of the 37 species, only *Leptoboea* sp. G9 and *Rhynchochotum* sp. XZB2307 contained all 6 types of SSRs (Figure 3A).

Dispersed repeats are distributed throughout the genome and include four types: forward (F), reverse (R), palindromic (P), and complement (C) repeats (Jelinek et al., 1980). In this study, a total of 1179 dispersed repeats were detected across the 37 chloroplast genomes (Supplementary Table 4). The number of dispersed

repeats per species ranged from 23 in *Boeica* sp. G37 to 44 in *Beccarinda erythrorhiza* G20 and *Beccarinda minima* XZB2001. These included 11–20 forward repeats, 12–23 palindromic repeats, 0–6 reverse repeats, and 0–2 complement repeats. Forward and palindromic repeats were the most abundant types, and reverse and complement repeats were absent in some species (Figure 3C).

3.3 Expansion and contraction of SC/IR boundaries

The 37 chloroplast genomes of the subtribe Leptoboeinae exhibit a typical quadripartite structure, with 4 junctions between the single-copy (SC) regions and the inverted repeat (IR) regions.



The results showed that the SC/IR boundaries of the 37 chloroplast genomes are relatively conserved. The LSC/IRB boundary (JLB line) is located in the *rps19* gene, the IRB/SSC boundary (JSB line) lies within the overlapping region of the *ndhF* gene and the pseudogene *ycf1*, the SSC/IRA boundary (JSA line) is located in the *ycf1* gene, and the IRA/LSC boundary (JLA line) falls in the intergenic region between *rpl2* and *trnH* (Figure 4).

At the LSC/IRb boundary, the *rps19* gene spans the JLB line, with a length of 202–250 bp in the LSC region and 24–76 bp in the IRb region. At the IRb/SSC boundary, the *ndhF* gene crosses the JSB line, with a length of 7–111 bp in the IRb region and 2,161–2,219 bp in the SSC region. In some species, the *ndhF* gene shows partial extension into the IRb region, such as in *Boeica arunachalensis* D. Borah, R. Kr. Singh, Taram & A. P. Das (B12), *B. ornithocephalantha* F.Wen, T.V.Do & Y.G.Wei (B13), and *B. stolonifera* K. Y. Pan (B14), which exhibit additional expansions of 25 bp, 18 bp, and 14 bp, respectively. At the SSC/IRA boundary, the *ycf1* gene spans the JSA line, with a length of 4,549–4,716 bp in the SSC region and 102–806 bp in the IRA region. At the IRA/LSC boundary, the *rpl2* gene is located to the left of the JLA line, 84–1,607 bp away, while *trnH* is positioned to the right of the JLA line, at a distance of 0–14 bp.

3.4 Comparative analysis of whole chloroplast genome alignment and nucleotide polymorphism

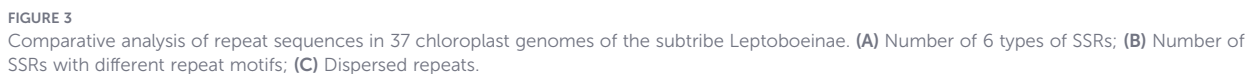
Using *Beccarinda argentea* (Anthony) Burt (XZB2310) as the reference sequence, multiple whole-genome alignments of the

chloroplast genomes from 37 species of the subtribe Leptoboeinae were performed and visualized (Figure 5). The results showed that the SC regions exhibited higher nucleotide variability than the IR regions, and the non-coding regions were more variable than the coding regions. Within the coding regions, all genes were relatively conserved except for *ycf1*.

Comparative analysis of nucleotide polymorphism (Pi values) in 800 bp sliding windows across the alignment of the 37 chloroplast genomes showed that Pi values ranged from 0 to 0.03574, with the lowest Pi value among the top 5% windows being 0.02136. A total of ten mutation hotspot regions (Pi > 0.02136) were identified as potential molecular markers: *trnK^{UUU}*, *trnK^{UUU}-rps16*, *rps16*, *trnC^{GCA}*, *ndhF*, *ndhF-rpl32*, *rpl32-trnL^{UAG}*, *ccsA*, *ndhG*, *ndhH-rps15-ycf1* and *ycf1*. All ten mutation hotspots were located in the SC regions, whereas Pi values in the IR regions were generally lower and more conserved (Figure 6).

3.5 Codon usage bias analysis

In this study, CDS longer than 300 bp were selected, resulting in a total of 58 CDS used for codon usage bias analysis. A total of 64 codons were detected, with the total number of codons ranging from 19,558 *Beccarinda cordifolia* (Anthony) B. L. Burt (B4) to 21,391 *Championia reticulata* G32. UGA (stop codon) and AUU (encoding isoleucine) were the least and most frequently used codons among the 37 chloroplast genomes, respectively, while cysteine (Cys, C) and leucine (Leu, L) were the least and most abundant amino acids, respectively (Supplementary Table 5).



All chloroplast genomes of Leptoboeinae exhibit the conserved quadripartite structure. The genome sizes, the lengths of the LSC, SSC, and IR regions, and their corresponding GC contents for each

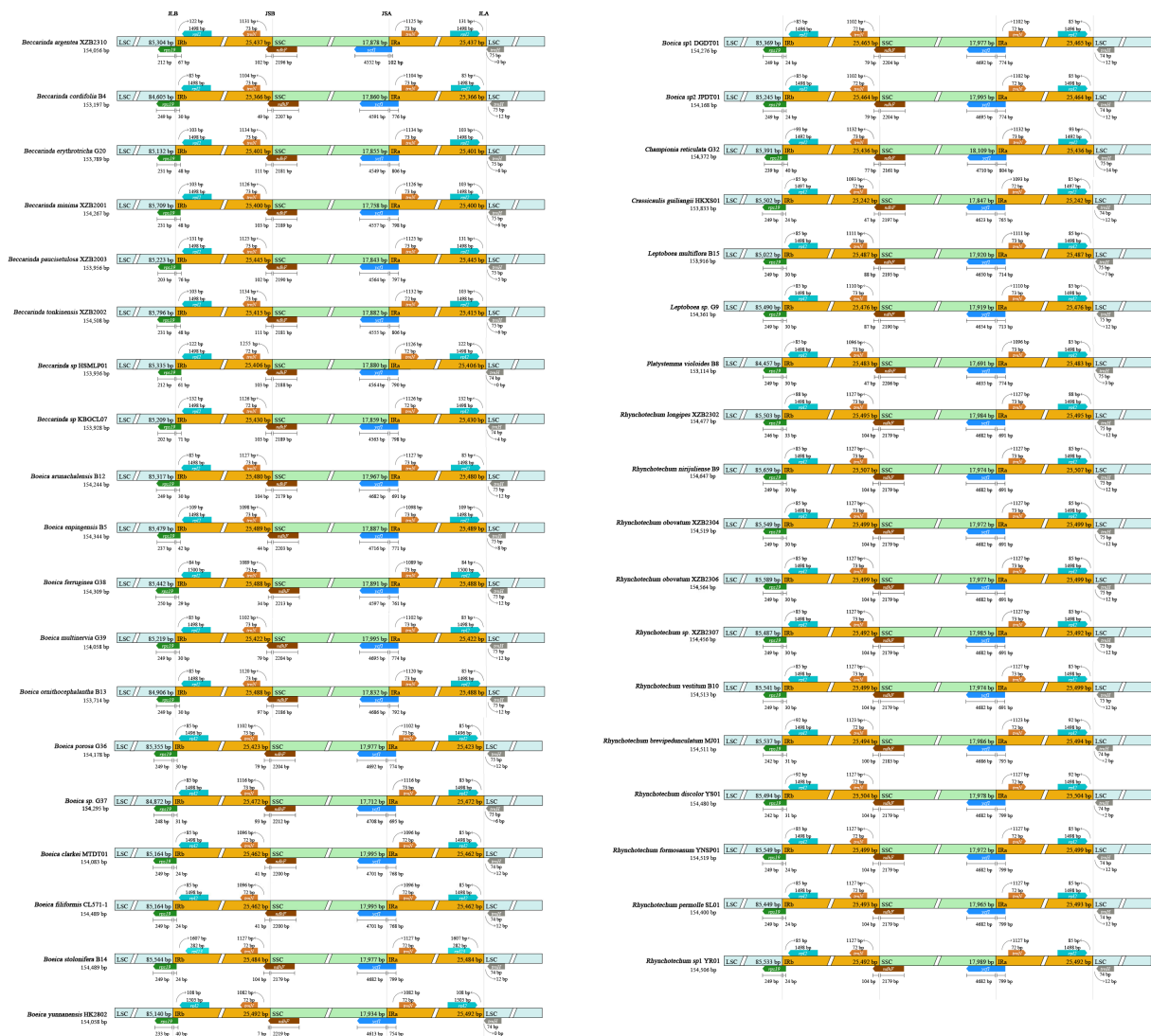


FIGURE 4
Comparison of SC/IR boundaries among 37 chloroplast genomes of the subtribe Leptoboeinae.

genus are shown in [Supplementary Figure 3](#) and [Supplementary Table 1](#), and the one-way ANOVA results are provided in [Supplementary Table 6](#). The statistical analyses indicate that the ten clades show significant differences in chloroplast genome size and GC content, and the subclades of *Boeica* also exhibit significant variation in GC content ([Figure 8](#)).

4 Discussion

4.1 Conservatism of the chloroplast genomic structure in the subtribe Leptoboeinae

In this study, the chloroplast genomes of 37 species of the subtribe Leptoboeinae were all found to possess the typical quadripartite structure. Their chloroplast genome organization, gene types, gene content, SC/IR boundaries, and codon usage bias

are relatively conserved. The high degree of conservation of chloroplast genomes has also been observed in the tribe Trichosporeae, as well as in the genera *Primulina* and *Paraboea* (Clarke) Ridley within Gesneriaceae ([Cui et al., 2023](#); [Gu et al., 2024](#); [Wang et al., 2022](#)), indicating that chloroplast genome structure across Gesneriaceae is subject to strong evolutionary and functional constraints. The sizes of the chloroplast genomes ranged from 153,114 bp to 154,647 bp, which falls within the general size range (115 kb–165 kb) of photosynthetic plant chloroplast genomes ([Jansen et al., 2005](#); [Zhang et al., 2021](#)). This narrow size range further supports the hypothesis that plastome evolution in Leptoboeinae has been dominated by purifying selection rather than by large-scale structural rearrangements. Given that most species of this subtribe inhabit shaded understory environments with relatively stable microclimatic conditions, strong functional constraints associated with photosynthetic efficiency may have limited structural divergence of chloroplast genomes ([Kan et al., 2024](#)). Similar ecological filtering effects have been proposed to maintain plastome stability in other shade-adapted angiosperm

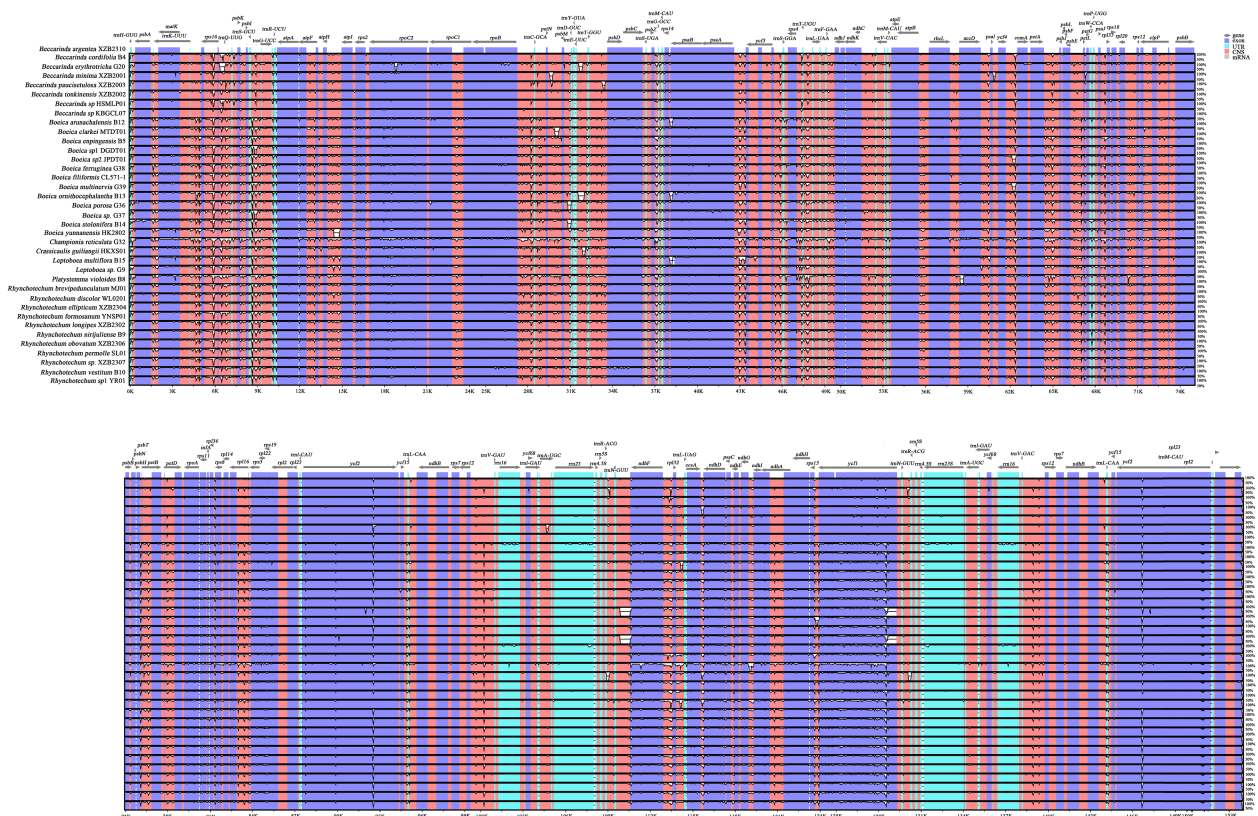


FIGURE 5
Global alignment analysis of 37 chloroplast genomes of the subtribe Leptoboeinae using the LAGAN algorithm on the mVISTA platform, with *Beccarinda argentea* XZB2310 as the reference sequence.

lineages (Magota et al., 2024). The GC content of the chloroplast genome is unevenly distributed, with the highest GC content in the IR regions and the lowest in the SSC region, which may be related to the presence of 4 rRNA genes (*rrn4.5*, *rrn5*, *rrn16*, *rrn23*) in the IR regions (Lan et al., 2022; Peng et al., 2022; Yan et al., 2022). In addition, we found that the GC content range of Leptoboeinae species is consistent with that previously reported for other chloroplast genomes of Gesneriaceae, such as Trichosporeae

(37.2%–37.8%; Cui et al., 2023) and *Primulina* (37.71%–38.80%; Gu et al., 2024). This consistency suggests that base composition evolution of the chloroplast genome has remained relatively conserved during the diversification of Leptoboeinae, providing a stable molecular background for phylogenetic inference.

Contraction, expansion, or gene loss of the IR regions often leads to changes in the length of the genome sequence (Goulding et al., 1996; Raubeson et al., 2007; Wang et al., 2008). In this study,

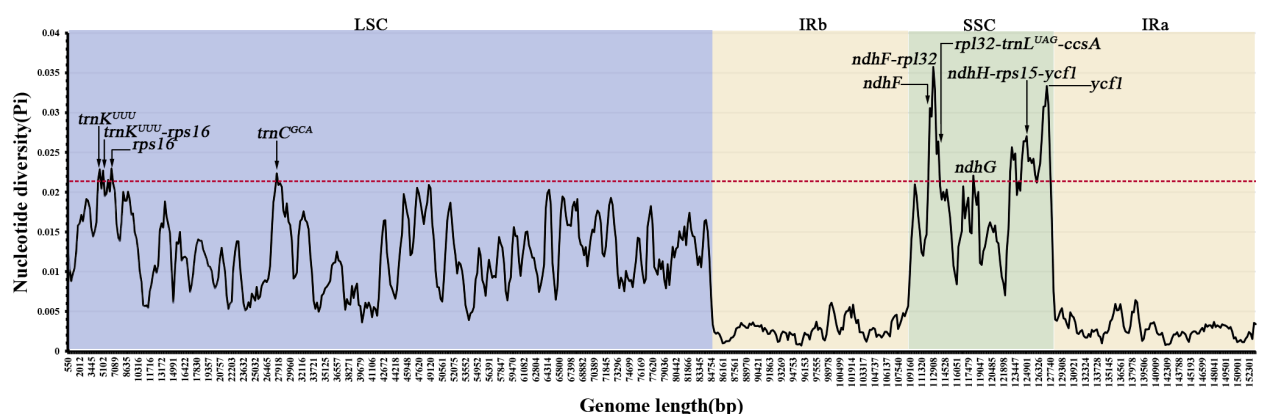


FIGURE 6
Comparison of nucleotide diversity (Pi) among 37 chloroplast genomes. The red dashed line represents the lowest Pi value within the top 5% of windows, and regions with Pi values above this line are identified as highly variable regions.

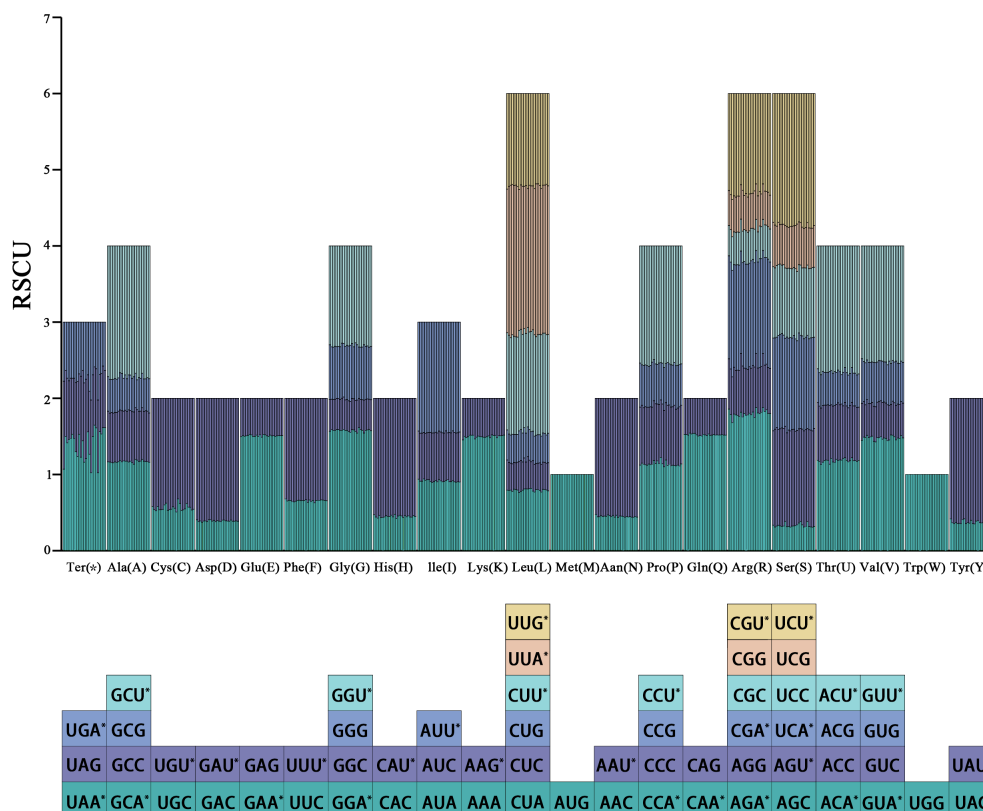


FIGURE 7

Relative synonymous codon usage (RSCU) values of codons encoding 20 amino acids and stop codons in the 37 chloroplast genomes. Each column in the bar chart represents one species, and color in the column graph corresponds to the codons depicted below the figure. "*" indicates codons with RSCU > 1.

the SC/IR boundaries of the 37 chloroplast genomes were relatively conserved, with no obvious IR expansion or contraction observed, which is largely consistent with the SC/IR boundary patterns reported in other species of Gesneriaceae (Cui et al., 2023; Gu et al., 2024). Although the boundary regions of chloroplast genomes are relatively stable, expansions and contractions of the inverted repeat (IR) regions can lead to the extension of the *ndhF* gene located at the boundary, a phenomenon that has also been observed in other plant lineages (Asaf et al., 2021; Liang et al., 2022), and are more likely attributable to minor boundary shifts rather than directional evolution of the inverted repeat regions. This further supports the overall structural stability of chloroplast genomes in the subtribe Leptoboeinae.

Codon usage bias is crucial for understanding species evolution, predicting gene function, and estimating expression levels (Uddin et al., 2015). Analysis of codon usage bias showed that the frequencies of 64 codons were generally similar across the 37 chloroplast genomes. A total of 31 high-frequency codons, 33 low-frequency codons, and 2 unbiased codons were identified. Among the high-frequency codons, most ended with A/U, except for UUG (encoding leucine). In contrast, the low-frequency codons tended to end with C/G. This pattern is consistent with the relatively low GC3 content (Supplementary Table 1) and is also in line with previous findings (Hsieh et al., 2022; Cui et al., 2023; Gu et al., 2024).

4.2 Repeats in the chloroplast genome

In the chloroplast genomes of the 37 species, the detected SSRs were mainly composed of mono-, di-, tri-, and tetranucleotide repeats. The SSR repeat units were predominantly composed of A or T, with A/T and AT/AT repeats being the most common, while other types were relatively rare. This pattern is consistent with the major SSR repeat units identified in the chloroplast genomes of other Gesneriaceae species (Hsieh et al., 2022; Cui et al., 2023; Gu et al., 2024). Previous studies have shown that among all SSR types, the contents of A and T bases are much higher than those of G and C, which corresponds to the general SSR characteristics of angiosperm chloroplast genomes (Asaf et al., 2020). This pronounced AT bias is consistent with the overall low GC content of chloroplast genomes and further reflects the presence of biased mutational pressure that favors replication slippage in AT-rich regions (Massouh et al., 2016; Kan et al., 2024). In addition, forward and palindromic repeats were the main types of dispersed repeats, which is consistent with reports in other chloroplast genomes (Yang et al., 2023b; Qin et al., 2025). The dominance of forward and palindromic dispersed repeats further suggests that even in the absence of large-scale genome rearrangements, localized recombination events or slipped-strand mispairing may drive small-scale variations in chloroplast genomes (Wang and Lanfear, 2019; Cauz-Santos, 2025).

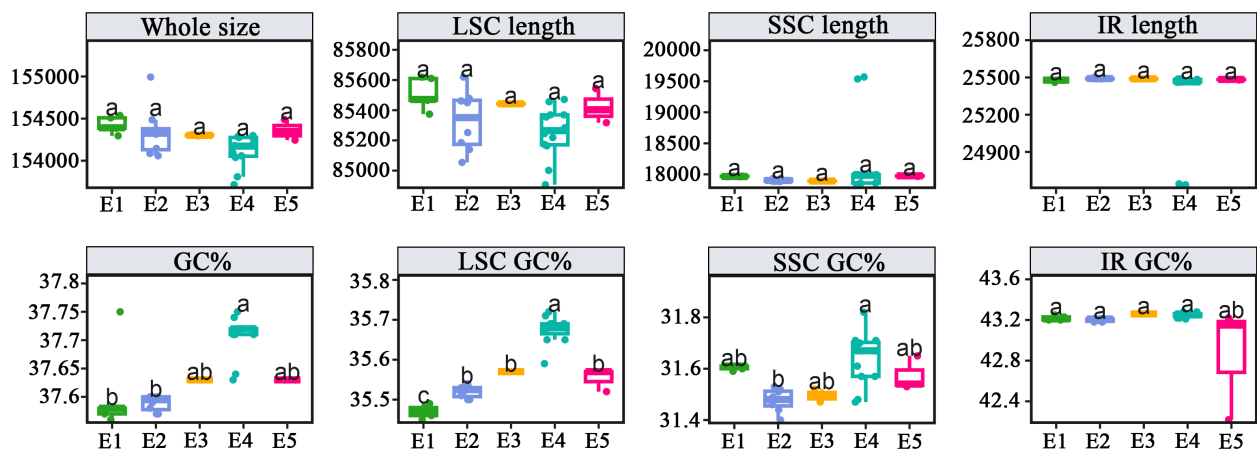


FIGURE 8

Comparative analysis of genomic characteristics among clades within *Boeica*. The x-axis represents the clade numbers, and different lowercase letters indicate significant differences among the five clades at the $P < 0.05$ level.

4.3 Identification of highly variable regions as DNA barcodes

In this study, results from multiple alignments of the 37 chloroplast genomes and mVista similarity analyses showed that the IR regions are more conserved than the SC regions, with greater nucleotide divergence observed in the non-coding regions of the SC and in the *ycf1* gene. This pattern is largely attributable to the homogenizing effect of concerted evolution between inverted repeat sequences, which reduces sequence divergence within the IR regions (Ruhlman et al., 2017). Further quantification of nucleotide polymorphism levels identified a total of ten mutational hotspot regions, providing important insights into the evolutionary dynamics of chloroplast genomes in the subtribe Leptoboeinae. Among these regions, *ycf1*, *trnK^{UUU}-rps16*, and *ndhF* are considered potential efficient molecular markers, which have also been reported in other angiosperms (Wang et al., 2024b), indicating that these loci represent evolutionarily unstable regions. Their elevated variability is likely attributable to relaxed functional constraints or higher nucleotide substitution rates (Dong et al., 2015; Dugas et al., 2015). Moreover, the candidate DNA barcodes identified in this study establish a practical bridge between species identification and deeper phylogenetic reconstruction, and thus have direct implications for taxonomic studies, conservation genetics, and evolutionary biology within Gesneriaceae.

4.4 Phylogenetic relationships

4.4.1 Phylogenetic relationships of genera within the subtribe Leptoboeinae

Based on CDS sequences, this study reconstructed the most comprehensive phylogeny to date for the subtribe Leptoboeinae, including all 37 sampled species representing all genera, using both coalescent and concatenation approaches. The two methods produced largely congruent topologies at the generic level (Figure 1; Supplementary Figure 2), indicating that chloroplast genome data provide reliable phylogenetic signals for resolving deep relationships

within the subtribe. However, differences in the placement of several species are likely attributable to chloroplast-specific features and methodological biases (Wu et al., 2015). Recent work has demonstrated the presence of cytonuclear conflict within the subtribe (Yang et al., 2025). Compared with Yang et al., our coalescent tree differs in several respects: *Platystemma* is recovered as basal; *Boeica yunnanensis* group with *B. enpingensis*; *B. ferruginea* forms an independent lineage; and *Boeica filiformis* is nested within *Rhynchotechum*. In contrast, Yang et al. recovered *Crassicaulis* as basal; *Boeica yunnanensis* as a separate lineage; *B. ferruginea* and *B. enpingensis* as sister taxa; and *Boeica filiformis* as sister to *Rhynchotechum*. Such topological incongruence may result from hybridization (Robertson et al., 2010), chloroplast capture (Yang et al., 2021), introgression (Pelser et al., 2010; Cai et al., 2021), or incomplete lineage sorting (ILS) (Feng et al., 2022). Future work integrating broader nuclear genomic data will be necessary to resolve these reticulate evolutionary patterns. Overall, confirming the monophyly of Leptoboeinae supporting the exclusion of *Championia* from Leptoboeinae, consistent with earlier hypotheses (Middleton et al., 2015; Ranasinghe et al., 2024). In this study, species of *Boeica* were scattered across five clades and showed varying degrees of cytonuclear conflict. However, the type species of *Boeica* (*Boeica fulva*) was not sampled here or in any previous phylogenetic work, leaving its placement among the recovered clades (E1–E5) unresolved. Traditional morphological characters used to delimit genera within Leptoboeinae show partial correspondence with the phylogenetic relationships reconstructed here. For example, *Platystemma* and *Beccarinda* are closely related, sharing the traits of free sepals and ovoid ovaries; however, *Platystemma* is distinguished from other Leptoboeinae genera by having only 1–2 terminal leaves, while *Beccarinda* is set apart by nearly transverse capsule insertion (Möller et al., 2017). The relationships among *Boeica*, *Leptoboea*, and *Rhynchotechum* have long been established (Fritsch, 1894; Wei et al., 2010), likely due to their shared morphological traits of being suffrutescent with stems and having four stamens (Weber et al., 2013). In contrast, *Leptoboea* and *Boeica* bear capsules, whereas *Rhynchotechum* produces indehiscent berries; additionally, *Leptoboea* has opposite leaves and anthers that do not fuse apically. *Crassicaulis* is

morphologically distinct from other genera by its stout stems and the presence of glandular hairs on the pedicel, calyx, and corolla (Yang et al., 2025). The polyphyly of *Boeica* has been repeatedly recognized in previous studies, but neither morphological nor molecular phylogenetic analyses have fully resolved it (Yang et al., 2020; 2023a; Yang et al., 2025). Morphologically, leaf arrangement and other characters are often unstable and overlap with those of related genera; molecularly, pervasive cyto-nuclear conflict and signals of reticulate evolution further complicate phylogenetic inference. Within the clades defined here (E1–E5), among which E1 and E3 can be distinguished from the other *Boeica* lineages by their nearly actinomorphic corollas and opposite leaves (Wang et al., 1990, 1998), whereas no distinct morphological characters were found for E2, E4 and E5. This pattern suggests that a combination of morphological stasis or convergence and reticulate evolution has blurred lineage boundaries within *Boeica*.

4.4.2 The polyphyly of *Boeica* may be partially congruent with its chloroplast genome structure

Structural variation in chloroplast genomes often contributes to the formation of genetic diversity among lineages, which is not only associated with phylogenetic relationships but also linked to environmental adaptation (Šmarda et al., 2014). The polyphyly of *Boeica* has been repeatedly demonstrated in previous studies, yet its taxonomic controversy remains unresolved (Ranasinghe et al., 2024; Yang et al., 2025). In this study, we found that clades E1 and E3 of *Boeica* can be distinguished by morphological characters, whereas clades E2, E4, and E5 are difficult to separate. Structural variation in chloroplast genome has proven valuable for reconstructing evolutionary relationships (Li et al., 2015; Zhan et al., 2024), and this is also supported by our findings (Figure 8). Analyses of chloroplast genome structure and ANOVA indicate that genome size and GC content vary among clades, largely driven by significant differences in LSC length. Although clade E2 is closely related to E3 and E4, their GC contents differ markedly, particularly in the LSC, likely due to the region's high gene density, functional diversity, and rapid evolutionary rate (Li et al., 2020; Qin et al., 2025; Yu et al., 2025). Notably, clade E5 differs significantly from clades E2, E3, and E4. Habitat analyses show that E5 occurs mainly in shaded tropical forests, and elevated GC content has been associated with adaptation to extreme temperatures (Carels et al., 1998; Šmarda et al., 2014), suggesting that environmental stress may drive GC-content evolution in this clade. Although E5 is morphologically distinct from *Rhynchoetechum*, its nested position within that genus and its GC-content differences relative to the G clade of *Rhynchoetechum* further highlight incongruence, a pattern also observed in other plant groups (Keating et al., 2023; Oyston et al., 2022). Because multiple *Boeica* clades have overlapping distributions, plastid genomic divergence may reflect independent lineage evolution. Overall, chloroplast genome characteristics provide new evidence for resolving the polyphyly of *Boeica*, a framework successfully applied to other polyphyletic groups (Zhang et al., 2023; Zhan et al., 2024). While taxonomic problems remain unresolved, this study offers new directions for future evolutionary and systematic research, and integrating morphological character evolution may further clarify generic boundaries.

5 Conclusion

This study presents the first comprehensive chloroplast genome analysis of the subtribe Leptoboecinae and successfully assembled the chloroplast genomes of 17 species within the group. Comparative analyses revealed that the overall genome structure, gene categories, gene content, and codon usage bias of these chloroplast genomes are relatively conserved. Across the 37 chloroplast genomes, with a total of 1,861 SSRs and 1,179 dispersed repeats identified. Ten hypervariable regions were detected, which represent potential molecular markers for genus-level delimitation. Codon usage patterns were highly similar, with the same 31 codons most frequently used. Phylogenetic analyses based on chloroplast genomes recovered well-resolved phylogenetic trees. Although conflicts were detected between the plastid tree and the coalescent tree regarding the phylogenetic placement of some taxa, both trees consistently supported the polyphyly of *Boeica*, which was resolved into five distinct lineages, and revealed that *Rhynchoetechum* is paraphyletic as a consequence. The monophyly of *Beccarinda* was confirmed, whereas the boundaries of *Rhynchoetechum* and *Boeica* require redefinition. In addition, structural variation within chloroplast genomes partially explains the polyphyly of *Boeica*. Significant differences were observed in GC content, LSC GC content, and SSC GC content among closely related lineages. Overall, these findings offer new insights for future evolutionary studies and systematic classification, and highlight the need for further integration of nuclear genomic data with morphological evidence to resolve the complex polyphyly of *Boeica*.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Author contributions

ZX: Data curation, Writing – original draft, Visualization. XQ: Data curation, Writing – original draft, Visualization. XL: Supervision, Writing – review & editing. LF: Supervision, Writing – review & editing. XB: Supervision, Writing – review & editing. FW: Supervision, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was supported by the Guangxi Natural Science Foundation (2025GXNSFAA069869, 2024GXNSFAA010462), the National Natural Science Foundation of China (32560392, 32460422, 32260782).

Acknowledgments

We thank Dr. Lei Cai and Dr. Fengmao Yang from the Kunming Institute of Botany, Chinese Academy of Sciences for providing data and assistance in the conduct of this study; we also thank Xiaoying Fu from the College of Life Sciences, Guangxi Normal University for her contributions to this study.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

References

- Asaf, S., Jan, R., Khan, A. L., and Lee, I. J. (2020). Complete chloroplast genome characterization of *Oxalis corniculata* and its comparison with related species from family Oxalidaceae. *Plants*. 9, 928. doi: 10.3390/plants9080928
- Asaf, S., Khan, A. L., Numan, M., and Al-Harrasi, A. (2021). Mangrove tree (*Avicennia marina*): insight into chloroplast genome evolutionary divergence and its comparison with related species from family Acanthaceae. *Sci. Rep.* 11, 3586. doi: 10.1038/s41598-021-83060-z
- Bentham, G. (1876). "Gesneriaceae," in *Genera plantarum*, vol. 2. Eds. G. Bentham and J. D. Hooker (Lovell Reeve, Co, London), 990–1025.
- Bolger, A. M., Lohse, M., and Usadel, B. (2024). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*. 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Burtt, B. L. (1963). Studies in the Gesneriaceae of the Old World. XXIV. Tentative keys to the tribes and genera. *Edinburgh: Notes R. Bot. Gard.* 24, 205–220.
- Burtt, B. L., and Wiehler, H. (1995). Classification of the family gesneriaceae. *Gesneriana*. 1, 1–4.
- Cai, L., Xi, Z., Lemmon, E. M., Lemmon, A. R., Mast, A., Buddenhagen, C. E., et al. (2021). The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Syst. Biol.* 70, 491–507. doi: 10.1093/sysbio/syaa083
- Carels, N., Hatey, P., Jabbari, K., and Bernardi, G. (1998). Compositional properties of homologous coding sequences from plants. *J. Mol. Evol.* 46, 45–53. doi: 10.1007/PL00006282
- Cauz-Santos, L. A. (2025). Beyond conservation: the landscape of chloroplast genome rearrangements in angiosperms. *New Phytol.* 247. doi: 10.1111/nph.70364
- Chan, P. P., and Lowe, T. M. (2019). tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* 1962, 1–14. doi: 10.1007/978-1-4939-9173-0_1
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Cheng, X., Shi, C., Yang, T., Ge, Q., Kress, W. J., and Liu, X. (2025). Unveiling the conserved nature of *Heliconia* chloroplast genomes: insights from the assembly and analysis of four complete chloroplast genomes. *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1535549
- Clarke, C. B. (1884). "Gesneriaceae," in *Flora of british India*, vol. . p. Ed. J. D. Hooker (Reeve, London), 336–375.
- Cui, Y. F., Zhou, P., Xiang, K. L., Zhang, Q., Yan, H., Zhang, L. G., et al. (2023). Plastome evolution and phylogenomics of Trichosporeae (Gesneriaceae) with its morphological characters appraisal. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1160535
- Dong, W., Xu, C., Li, C., Sun, J., Zuo, Y., Shi, S., et al. (2015). *ycf1*, the most promising plastid DNA barcode of land plants. *Sci. Rep.* 5, 8348. doi: 10.1038/srep08348
- Dong, Y., Cao, Q., Yu, K., Wang, Z., Chen, S., Chen, F., et al. (2025). Chloroplast phylogenomics reveals the maternal ancestry of cultivated chrysanthemums. *Genomics Commun.* 2, e019. doi: 10.1186/s12864-021-07807-8
- Doyle, J. J. (1992). Gene trees and species trees: molecular systematics as one-character taxonomy. *Syst. Bot.* 17, 144–163. doi: 10.2307/2419070
- Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.
- Du, L. M., Zhang, C., Liu, Q., Zhang, X. Y., and Yue, B. S. (2018). Krait: an ultrafast tool for genome-wide survey of microsatellites and primer design. *Bioinformatics*. 34, 681–683. doi: 10.1093/bioinformatics/btx665
- Dugas, D. V., Hernandez, D., Koenen, E. J. M., Schwarz, E., Straub, S., Hughes, C. E., et al. (2015). Mimosoid legume plastome evolution: IR expansion, tandem repeat expansions, and accelerated rate of evolution in *clpP*. *Sci. Rep.* 5, 16958. doi: 10.1038/srep16958
- Feng, S., Bai, M., Rivas-González, I., Li, C., Liu, S., Tong, Y., et al. (2022). Incomplete lineage sorting and phenotypic evolution in marsupials. *Cell*. 185, 1646–1660. doi: 10.1016/j.cell.2022.03.034
- Frazer, K. A., Pachte, R. L., Poliakov, A., Rubin, E. M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res.* 32, W273–W279. doi: 10.1093/nar/gkh458
- Fritsch, K. (1894). "Gesneriaceae," in *Die natürlichen Pflanzenfamilien IV. 3b*. Prantl. Eds. A. Engler and K. Prantl (Leipzig: Wilhelm Engelmann Press) 133–185.
- Goulding, S. E., Olmstead, R. G., Morden, C. W., and Wolfe, K. H. (1996). Ebb and flow of the chloroplast inverted repeat. *Mol. Gen. Genet.* 252, 195–206. doi: 10.1007/BF02173220

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.2026.1766257/full#supplementary-material>.

SUPPLEMENTARY TABLE 1

Chloroplast genome features of 37 species of the tribe Leptoboeinae.

SUPPLEMENTARY TABLE 2

Sampling information of 37 species of the tribe Leptoboeinae.

SUPPLEMENTARY TABLE 4

Number of SSRs and INEs in 37 chloroplast genomes.

SUPPLEMENTARY TABLE 6

Results of one-way ANOVA for chloroplast genome features among different clades of the tribe Leptoboeinae.

- Greiner, S., Lehwark, P., and Bock, R. (2019). OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 47, W59–W64. doi: 10.1093/nar/gkz238
- Gu, J. M., Li, M. J., He, S. T., Li, Z., Wen, F., Tan, K., et al. (2024). Comparative chloroplast genomes analysis of nine *Primulina* (Gesneriaceae) rare species, from karst region of southwest China. *Sci. Rep.* 14, 1–16. doi: 10.1038/s41598-024-81524-6
- Hsieh, C. L., Xu, W. B., and Chung, K. F. (2022). Plastomes of limestone karst gesneriad genera *Petrocodon* and *Primulina*, and the comparative plastid phylogenomics of *Gesneriaceae*. *Sci. Rep.* 12, 15800. doi: 10.1038/s41598-022-19812-2
- Ivanina, L. I. (1965). Application of the carpological method to the taxonomy of Gesneriaceae. *Notes R. Gov Bhutan* 26, 383–403.
- Jansen, R. K., Raubeson, L. A., Boore, J. L., Pamphilis, C. W., Chumley, T. W., Haberle, R. C., et al. (2005). Methods for obtaining and analyzing whole chloroplast genome sequences. *Meth. Enzym.* 395, 348–384. doi: 10.1016/S0076-6879(05)95020-9
- Jelinek, W. R., Toomey, T. P., Leinwand, L., Duncan, H., Biro, P. A., Choudary, P. V., et al. (1980). Ubiquitous, interspersed repeated sequences in mammalian genomes. *Proc. Natl. Acad. Sci.* 77, 1398–1402. doi: 10.1073/pnas.77.3.1398
- Jin, 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
- Kan, S., Su, X., Yang, L., Zhou, H., Qian, M., Zhang, W., et al. (2024). From light into shadow: comparative plastomes in *Petrocosmea* and implications for low light adaptation. *BMC Plant Biol.* 24, 949. doi: 10.1186/s12870-024-05669-2
- 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
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199
- Keating, J. N., Garwood, R. J., and Sansom, R. S. (2023). Phylogenetic congruence, conflict and consistency between molecular and morphological data. *BMC Ecol. Evol.* 23, 30. doi: 10.1186/s12862-023-02131-z
- Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633
- Lan, Z. H., Shi, Y. H., Yin, Q. G., Gao, R. R., Liu, C. L., Wang, W. T., et al. (2022). Comparative and phylogenetic analysis of complete chloroplast genomes from five *Artemisia* species. *Front. Plant Sci.* 13, 1049209. doi: 10.3389/fpls.2022.1049209
- Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301
- Li, H., Guo, Q. Q., Xu, L., Gao, H. D., Liu, L., and Zhou, X. G. (2023). CPJSDraw: analysis and visualization of junction sites of chloroplast genomes. *PeerJ.* 11, e15326. doi: 10.7717/peerj.15326
- Li, L., Jiang, Y., Liu, Y., Niu, Z., Xue, Q., Liu, W., et al. (2020). The large single-copy (LSC) region functions as a highly effective and efficient molecular marker for accurate authentication of medicinal *Dendrobium* species. *Acta Pharm. Sin. B* 10, 1989–2001. doi: 10.1016/j.apsb.2020.01.012
- Li, H. T., Luo, Y., Gan, L., Ma, P. F., Gao, L. M., Yang, J. B., et al. (2021). Plastid phylogenomic insights into relationships of all flowering plant families. *BMC Biol.* 19, 232. doi: 10.1186/s12915-021-01166-2
- Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., and Chen, S. (2015). Plant DNA barcoding: from gene to genome. *Biol. Rev.* 90, 157–166. doi: 10.1111/brv.12104
- Liang, D., Wang, H., Zhang, J., Zhao, Y., and Wu, F. (2022). Complete chloroplast genome sequence of *Fagus longipetiolata* Seemen (Fagaceae): Genome structure, adaptive evolution, and phylogenetic relationships. *Life.* 12, 92. doi: 10.3390/life12010092
- Liu, X. E. (2013). A more accurate relationship between ‘effective number of codons’ and GC3s under assumptions of no selection. *Comput. Biol. Chem.* 42, 35–39. doi: 10.1016/j.compbiolchem.2012.11.003
- Lu, Y., Xu, P. J., Chen, Z. N., and Liu, G. M. (1998). Anthraquinone glycosides from *Rhynchotechum vestitum*. *Phytochemistry.* 49, 1135–1137. doi: 10.1016/S0031-9422(98)00079-X
- Magota, K., Gotoh, E., Sakaguchi, S., Ikeda, H., and Setoguchi, H. (2024). High-intensity light promotes adaptive divergence of photosynthetic traits between sun-exposed and shaded populations in *Saxifraga fortunei*. *Am. J. Bot.* 111, e16371. doi: 10.1002/ajb2.16371
- Massouh, A., Schubert, J., Yaneva-Roder, L., Ulbricht-Jones, E. S., Zupok, A., Johnson, M. T., et al. (2016). Spontaneous chloroplast mutants mostly occur by replication slippage and show a biased pattern in the plastome of *Oenothera*. *Plant Cell.* 28, 911–929. doi: 10.1105/tpc.15.00879
- Mazumdar, P., Binti, O. R., Mebus, K., Ramakrishnan, N., and Ann, H. J. (2017). Codon usage and codon pair patterns in non-grass monocot genomes. *Ann. Bot.* 120, 893–909. doi: 10.1093/aob/mcx112
- Middleton, D. J., Nishii, K., Puglisi, C., Forrest, L. L., and Möller, M. (2015). *Chayamaritia* (Gesneriaceae: Didymocarpoideae), a new genus from southeast Asia. *Pl. Syst. Evol.* 301, 1947–1966. doi: 10.1007/s00606-015-1213-2
- Möller, M., Nampy, S., Janeesha, A. P., and Weber, A. (2017). The Gesneriaceae of India: Consequences of updated generic concepts and new family classification. *Rheedea.* 27, 23–41. doi: 10.22244/rheedea.2017.27.1.5
- Oyston, J. W., Wilkinson, M., Ruta, M., and Wills, M. A. (2022). Molecular phylogenies map to biogeography better than morphological ones. *Commun. Biol.* 5, 521. doi: 10.1038/s42003-022-03482-x
- Peden, J. F. (1999). Analysis of codon usage (dissertation). University of Nottingham, Nottingham.
- Pelzer, P. B., Kennedy, A. H., Tepe, E. J., Shidler, J. B., Nordenstam, B., Kadereit, J. W., et al. (2010). Patterns and causes of incongruence between plastid and nuclear *Senecioneae* (Asteraceae) phylogenies. *Am. J. Bot.* 97, 856–873. doi: 10.3732/ajb.0900287
- Peng, J. Y., Zhang, X. S., Zhang, D. G., Wang, Y., Deng, T., Huang, X. H., et al. (2022). Newly reported chloroplast genome of *Sinosenecio albonervius* Y Liu & QE Yang and comparative analyses with other *Sinosenecio* species. *BMC Genom.* 23, 639. doi: 10.1186/s12864-022-08872-3
- POWO (2025). Plants of the world online. Facilitated by the royal botanic gardens, kew. Available online at: <https://powo.science.kew.org/> (Accessed November 27, 2025).
- Qin, Q., Dong, Y. J., Chen, J. L., Wang, B., Peng, Y. X., Zhang, X. P., et al. (2025). Comparative analysis of chloroplast genomes reveals molecular evolution and phylogenetic relationships within the *Papilionoideae* of *Fabaceae*. *BMC Plant Biol.* 25, 1–20. doi: 10.1186/s12870-025-06138-0
- Ranasinghe, S. W., Nishii, K., Möller, M., Atkins, H. J., Clark, J. L., Perret, M., et al. (2024). Biogeographic history of the pantropical family Gesneriaceae with a focus on the Indian plate and diversification through the Old World. *Front. Biogeogr.* 16. doi: 10.21425/F5FBG61637
- Raubeson, L. A., Peery, R., Chumley, T. W., Dziubek, C., Fourcade, H. M., Boore, J. L., et al. (2007). Comparative chloroplast genomes: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* 8, 1–27. doi: 10.1186/1471-2164-8-174
- Robertson, A., Rich, T. C., Allen, A. M., Houston, L., Roberts, C. A. T., Bridle, J. R., et al. (2010). Hybridization and polyploidy as drivers of continuing evolution and speciation in *Sorbus*. *Mol. Ecol.* 19, 1675–1690. doi: 10.1111/j.1365-294X.2010.04585.x
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34, 3299–3302. doi: 10.1093/molbev/msx248
- Ruhlman, T. A., Zhang, J., Blazier, J. C., Sabir, J. S., and Jansen, R. K. (2017). Recombination-dependent replication and gene conversion homogenize repeat sequences and diversify plastid genome structure. *Am. J. Bot.* 104, 559–572. doi: 10.3732/ajb.1600453
- Sayyari, E., and Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Mol. Biol. Evol.* 33, 1654–1668. doi: 10.1093/molbev/msw079
- Sharp, P. M., and Li, W. H. (1987). The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15, 1281–1295. doi: 10.1093/nar/15.3.1281
- Šmarda, P., Bureš, P., Horová, L., Leitch, I. J., Mucina, L., Pacini, E., et al. (2014). Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proc. Natl. Acad. Sci.* 111, E4096–E4102. doi: 10.1073/pnas.1321152111
- Sun, Y., Shang, L., Zhu, Q. H., Fan, L., and Guo, L. (2022). Twenty years of plant genome sequencing: achievements and challenges. *Trends Plant Sci.* 27, 391–401. doi: 10.1016/j.tplants.2021.10.006
- Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120
- 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
- Torruella, G., Galindo, L. J., Moreira, D., and López-García, P. (2025). Phylogenomics of flagellated protists supports a revised eukaryotic tree of life. *Curr. Biol.* 35, 198–207. doi: 10.1016/j.cub.2024.10.075
- Uddin, A., Mazumdar, T. H., Choudhury, M. N., and Chakraborty, S. (2015). Codon bias and gene expression of mitochondrial ND2 gene in chordates. *Bioinformatics.* 11, 407–412. doi: 10.6026/97320630011407
- Walker, J. F., Walker-Hale, N., Vargas, O. M., Larson, D. A., and Stull, G. W. (2019). Characterizing gene tree conflict in plastome-inferred phylogenies. *PeerJ.* 7, e7747. doi: 10.7717/peerj.7747

- Wang, R. J., Cheng, C. L., Chang, C. C., Wu, C. L., Su, T. M., and Chaw, S. M. (2008). Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol. Biol.* 8, 36. doi: 10.1186/1471-2148-8-36
- Wang, J., Kan, S., Liao, X., Zhou, J., Tembrock, L. R., Daniell, H., et al. (2024a). Plant organellar genomes: much done, much more to do. *Trends Plant Sci.* 29, 754–769. doi: 10.1016/j.tplants.2023.12.014
- Wang, W., and Lanfear, R. (2019). Long-reads reveal that the chloroplast genome exists in two distinct versions in most plants. *Genome Biol. Evol.* 11, 3372–3381. doi: 10.1093/gbe/evz256
- Wang, W. T., Pan, K. Y., and Li, Z. Y. (1990). “Gesneriaceae,” in *Flora republicae popularis sinicae*. Ed. W. T. Wang (Science Press, Beijing), 125–581.
- Wang, W. T., Pan, K. Y., Li, Z. Y., Weitzman, A. L., and Skog, L. E. (1998). “Gesneriaceae,” in *Flora of China*, vol. 18. Eds. Z. Y. Wu and P. H. Raven (Beijing: Science Press, St. Louis: Missouri Botanical Garden Press), 244–401.
- Wang, Y. F., Wen, F., Hong, X., Li, Z. L., Mi, Y. L., and Zhao, B. (2022). Comparative chloroplast genome analyses of *Paraboea* (Gesneriaceae): insights into adaptive evolution and phylogenetic analysis. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1019831
- Wang, Y. H., Xu, C., Guo, X., Wang, Y., Chen, Y. Y., Shen, J., et al. (2024b). Phylogenomics analysis of *Scutellaria* (Lamiaceae) of the world. *BMC Biol.* 22, 185. doi: 10.1186/s12915-024-01982-2
- Weber, A., Clark, J. L., and Möller, M. (2013). A new formal classification of Gesneriaceae. *Selbyana* 31, 68–94.
- Weber, A., Middleton, D., and Moeller, M. (2025). *Skogea* is a synonym of *boeica* (Gesneriaceae). *Edinburgh J. Bot.* 82, 1–8. doi: 10.1002/fedr.202400003
- Wei, Y. G., Wen, F., Chen, W. H., Shui, Y. M., and Möller, M. (2010). *Litostigma*, a new genus from China: a morphological link between basal and derived didymocarpoid Gesneriaceae. *Edinburgh J. Bot.* 67, 161–184. doi: 10.1017/S0960428609990291
- Wen, F., Xin, Z. B., Hong, X., Cai, L., Chen, X. Y., Liang, J. J., et al. (2022). *Actinostephanus* (Gesneriaceae), a new genus and species from Guangdong, South China. *PhytoKeys* 193, 89–106. doi: 10.3897/phytokeys.193.80715
- Wicke, S., Schneeweiss, G. M., Depamphilis, C. W., Müller, K. F., and Quandt, D. (2011). The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Mol. Biol.* 76, 273–297. doi: 10.1007/s11103-011-9762-4
- Wu, L., Fan, P., Cai, J., Zang, C., Lin, Y., Xu, Z., et al. (2024). Comparative genomics and phylogenomics of the genus *Glycyrrhiza* (Fabaceae) based on chloroplast genomes. *Front. Pharmacol.* 15. doi: 10.3389/fphar.2024.1371390
- Wu, Z., Tembrock, L. R., and Ge, S. (2015). Are differences in genomic data sets due to true biological variants or errors in genome assembly: an example from two chloroplast genomes. *PLoS One* 10, e0118019. doi: 10.1371/journal.pone.0118019
- Xia, X. H. (2013). DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* 0, 1720–1728. doi: 10.1093/molbev/mst064
- Xia, P., Huang, Y., and Zhu, J. (2025). The complete chloroplast genome sequences of 11 *Panax* species: Providing insights for evolution and species identification. *Ind. Crops Prod.* 223, 120160. doi: 10.1016/j.indcrop.2024.120160
- Xue, T. T., Janssens, S. B., Liu, B. B., and Yu, S. X. (2024). Phylogenomic conflict analyses of the plastid and mitochondrial genomes via deep genome skimming highlight their independent evolutionary histories: A case study in the cinquefoil genus *Potentilla* sensu lato (Potentilleae, Rosaceae). *Mol. Phylogenet. Evol.* 190, 107956. doi: 10.1016/j.ympev.2025.108325
- Yan, K., Ran, J., Bao, S., Li, Y., Islam, R., Zhang, N., et al. (2022). The complete chloroplast genome sequence of *eupatorium fortunei*: genome organization and comparison with related species. *Genes* 14, 64. doi: 10.3390/genes14010064
- Yang, L. H., Harris, A. J., Wen, F., Li, Z., Feng, C., Kong, H., et al. (2023a). Phylogenomic analyses reveal an allopolyploid origin of core Didymocarpaceae (Gesneriaceae) followed by rapid radiation. *Syst. Biol.* 72, 1064–1083. doi: 10.1093/sysbio/syad029
- Yang, F., Ling, S., Yang, R., Wen, F., Gu, W., Hu, X., et al. (2025). Phylogenomic reconstruction reveals deep reticulate evolutionary history and polyphyletic genus (*Boeica*) in the subtribe Leptoboeinae (Gesneriaceae), with description of a novel genus. *Cladistics* 41, 1–16. doi: 10.1111/cla.70010
- 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, L. H., Wen, F., Kong, H. H., Sun, Z. X., Su, L. Y., and Kang, M. (2020). Two new combinations in *Oreocharis* (Gesneriaceae) based on morphological, molecular and cytological evidence. *PhytoKeys* 157, 43. doi: 10.3897/phytokeys.157.32609
- Yang, T. G., Wu, Z. H., Tie, J., Qin, R., Wang, J. Q., and Liu, H. (2023b). A comprehensive analysis of chloroplast genome provides new insights into the evolution of the genus *Chrysosplenium*. *Int. J. Mol. Sci.* 24, 14735. doi: 10.1186/s12864-020-07045-4
- Yu, F. M., He, W. C., Erst, A. S., Peng, H. W., Ling, Y. Y., Erst, T. V., et al. (2025). Phylogenomics and structural variations of plastid genomes of the columbine genus (*Ranunculaceae*). *BMC Plant Biol.* 25. doi: 10.3389/fpls.2024.1415253
- Zhai, Y. F., Yu, X. Q., Zhou, J. G., Li, J., Tian, Z., Wang, P. Q., et al. (2021). Complete chloroplast genome sequencing and comparative analysis reveals changes to the chloroplast genome after allopolyploidization in *Cucumis*. *Genome* 64, 627–638. doi: 10.1139/gen-2020-0134
- Zhan, M., Xue, L., Zhou, J. J., Zhang, Q., Qin, X. M., Liao, X. W., et al. (2024). Polyphyly of *Boehmeria* (Urticaceae) congruent with plastome structural variation. *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1297499
- Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., et al. (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* 20, 348–355. doi: 10.1111/1755-0998.13096
- Zhang, X. F., Landis, J. B., Wang, H. X., Zhu, Z. X., and Wang, H. F. (2021). Comparative analysis of chloroplast genome structure and molecular dating in Myrtales. *BMC Plant Biol.* 21, 219. doi: 10.1186/s12870-021-02985-9
- Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinf* 19, 153. doi: 10.1186/s12859-018-2129-y
- Zhang, D., Ren, J., Jiang, H., Wang, V. O., Dong, X., and Hu, G. (2023). Comparative and phylogenetic analysis of the complete chloroplast genomes of six *Polygonatum* species (Asparagaceae). *Sci. Rep.* 13, 7237. doi: 10.1038/s41598-023-34083-1