



Article

Chloroplast Phylogenomics and Evolutionary History of the Alpine Endemic *Eutrema scapiflorum*

Ting Lv ^{1,2,†} , Xiayu Hu ^{3,†}, Lizhi Guo ⁴, Jiasheng Ju ⁵, Yu Zhang ³ and Nan Tang ^{1,2,*}

¹ College of Agriculture and Animal Husbandry, Qinghai University, Xining 810016, China; lvtingiu@163.com

² Key Laboratory of Qinghai Province for Landscape Plants Research, Plateau Flower Research Centre, Qinghai University, Xining 810016, China

³ College of Agronomy, Northwest A&F University, Yangling 712100, China; huxiayu522@163.com (X.H.); yuzhang98@nwfau.edu.cn (Y.Z.)

⁴ Qinghai Institute for Drug Control, Xining 810016, China; paul_567@163.com

⁵ Food and Drug Testing Center, Xianyang 712000, China; starjss0616@163.com

* Correspondence: natasha_tn@163.com

† These authors contributed equally to this work.

Abstract

In this study, we sequenced, assembled, and characterized the first complete chloroplast (cp) genome of *Eutrema scapiflorum*, an alpine species endemic to the Qinghai–Tibet Plateau (QTP). The assembled plastome is 153,041 bp in length and exhibits a typical quadripartite structure, comprising a large single-copy (LSC) region of 83,547 bp and a small single-copy (SSC) region of 17,506 bp, which are separated by two inverted repeats (IRs) of 25,994 bp each. Structurally, the genome encodes 132 unique genes, including 87 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Comparative analysis across eight species revealed that genome size variation is primarily driven by the SSC region. Notably, the IR/SC boundaries in *E. scapiflorum* are highly conserved, which contrasts with the significant IR expansion observed in *Capsella tenella*. Furthermore, simple sequence repeat (SSR) analysis identified 78 loci, predominantly mononucleotide A/T repeats located in intergenic spacers. Nucleotide diversity analysis pinpointed *accD* and *ycf1* as the most variable genes. Selection pressure analysis indicated that most genes are under purifying selection, while seven protein-coding genes (*ycf2*, *nadhE*, *cemA*, *clpP*, *psbH*, *ycf4*, *nadhB*) exhibited signatures of positive selection ($K_a/K_s > 1$). Subsequently, phylogenomic analyses robustly resolved *E. scapiflorum* within the tribe Arabideae, showing its closest relationship to *Alliaria petiolata*. Divergence time estimation dated the split between *E. scapiflorum* and its closest relative to the middle Miocene (~17.57 Ma). Collectively, these findings provide crucial genomic resources and new insights into the structural evolution, phylogenetic placement, and potential adaptive mechanisms of this alpine species within the Brassicaceae family.

Keywords: *Eutrema scapiflorum*; chloroplast genome; high-altitude adaptation; middle Miocene; phylogenomics



Academic Editors: Zulfikar Ali Sahito and Afsheen Zehra

Received: 24 March 2026

Revised: 22 May 2026

Accepted: 26 May 2026

Published: 8 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The Brassicaceae (Cruciferae) family represents a major model system in plant biology, providing unparalleled insights into angiosperm evolution, diversification, and adaptation. The family's ecological amplitude, spanning from tropical lowlands to alpine zones, makes it an ideal group for studying how plants evolve in response to extreme environmental stressors, including high-altitude conditions such as hypoxia, intense ultraviolet radiation,

and temperature fluctuations [1]. Among these diverse lineages, the genus *Eutrema* (encompassing taxa formerly circumscribed as *Pegaeophyton*) comprises multiple high-altitude specialists that serve as exemplary models for investigating alpine adaptation [2]. *Eutrema scapiflorum* (Hook.f. & Thomson) Al-Shehbaz & Warwick (syn. *Pegaeophyton scapiflorum*), a representative taxon endemic to the Qinghai–Tibet Plateau and adjacent Himalayas, predominantly inhabits scree slopes and alpine meadows within an elevational range of 3500 to 5400 m [3–5]. Beyond its significance in evolutionary ecology, *E. scapiflorum* possesses substantial ethnobotanical value, with traditional applications in treating febrile diseases, trauma, and respiratory or digestive disorders. Recent phytochemical studies have further identified a suite of bioactive secondary metabolites including flavonoids, alkaloids, and glucosinolates that likely contribute to its documented anti-inflammatory, antioxidant, and antimicrobial properties [6,7].

Despite the burgeoning interest in high-altitude adaptation and the widespread application of phylogenomic approaches in Brassicaceae systematics, the complete chloroplast genome of *E. scapiflorum* remains uncharacterized [8,9]. This lack of genomic data hinders the precise resolution of its phylogenetic placement and the investigation of potential plastid-mediated adaptations to extreme environments [10]. Chloroplast genomes are valuable in evolutionary studies due to their generally conserved structure, moderate nucleotide substitution rates, and uniparental inheritance, which facilitate comparative genomic analyses and the detection of signatures of selection [11,12]. In high-altitude plants, chloroplast genes involved in photosynthesis, stress responses, and redox homeostasis may be targets of adaptive evolution, making whole-plastome sequencing a powerful tool for uncovering genomic correlates of alpine colonization [13,14].

Previous molecular phylogenetic studies of short-lived Brassicaceae taxa—often based on limited datasets such as nuclear internal transcribed spacer (ITS) and chloroplast loci like *trnL-F*—have revealed complex patterns of relationships among ephemeral and perennial lineages [15,16]. These analyses suggest that short-lived taxa are phylogenetically scattered rather than forming a distinct clade, pointing to possible convergent evolution of life-history traits and/or incomplete phylogenetic resolution due to limited informative characters [9,17]. To clarify the evolutionary history of *E. scapiflorum* and its relatives, a genome-scale approach is needed—one that integrates structural, comparative, and selection analyses to generate robust phylogenetic hypotheses and to identify genomic features associated with alpine life.

In this study, we present the first complete chloroplast genome assembly of *E. scapiflorum* and employ a comprehensive comparative framework to address the following objectives: (1) characterize the structural features, gene content, repetitive elements, and nucleotide diversity of the plastome; (2) examine IR boundary dynamics, sequence divergence, and substitution patterns across related Brassicaceae species; (3) detect signatures of selection in protein-coding genes that may underlie adaptation to high-altitude environments; and (4) reconstruct a robust phylogeny of Brassicaceae using whole chloroplast genomes to clarify the systematic position of *E. scapiflorum* and estimate divergence times within the family.

Our results not only provide a valuable genomic resource for future studies on alpine adaptation and species delimitation in Brassicaceae, but also contribute to a broader understanding of how chloroplast genome evolution may reflect environmental constraints in one of the world's most extreme botanical hotspots.

2. Results

2.1. Structural Characteristics of the Chloroplast Genome

The complete chloroplast genome of *E. scapiflorum* is 153,041 bp in length and exhibits a typical quadripartite structure, comprising a large single-copy (LSC) region of 83,547 bp, a small single-copy (SSC) region of 17,506 bp, and a pair of inverted repeat (IR) regions of 25,994 bp each (Figure 1). The overall GC content of the genome is 36.41%, while the GC contents of the LSC, SSC, and IR regions are 34.10%, 29.49%, and 42.44%, respectively. A total of 132 genes were annotated, including 87 protein-coding genes, 37 transfer RNA (tRNA) genes, and 8 ribosomal RNA (rRNA) genes. Among these, 18 genes are duplicated (Table 1). Additionally, introns were identified in 23 genes, of which 19 harbor a single intron and 4 contain two introns (Table 1).

Table 1. Gene composition in the chloroplast genome of *E. scapiflorum*.

Category	Gene Group	Gene Name
Photosynthesis	Subunits of photosystem I	<i>psaA, psaB, psaC, psaI, psaJ</i>
	Subunits of photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
	Subunits of NADH dehydrogenase	<i>ndhA*, ndhB*(2), ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	Subunits of cytochrome b/f complex	<i>petA, petB*, petD*, petG, petL, petN</i>
	Subunits of ATP synthase	<i>atpA, atpB, atpE, atpF*, atpH, atpI</i>
	Large subunit of rubisco	<i>rbcL</i>
Self-replication	Proteins of large ribosomal subunit	<i>rpl14, rpl16*, rpl2*(2), rpl20, rpl22, rpl23(2), rpl32, rpl33, rpl36</i>
	Proteins of small ribosomal subunit	<i>rps11, rps12**(2), rps14, rps15, rps16*, rps18, rps19, rps2, rps3, rps4, rps7(2), rps8</i>
	Subunits of RNA polymerase	<i>rpoA, rpoB, rpoC1*, rpoC2</i>
	Ribosomal RNAs	<i>rrn16(2), rrn23(2), rrn4.5(2), rrn5(2)</i>
	Transfer RNAs	<i>trnA-UGC*(2), trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG*, trnG-UCC, trnH-GUG, trnI-CAU(2), trnI-GAU*(2), trnK-UUU*, trnL-CAA(2), trnL-UAA*, trnL-UAG, trnM-CAU, trnN-GUU(2), trnP-UGG, trnQ-UUG, trnR-ACG(2), trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC(2), trnV-UAC*, trnW-CCA, trnY-GUA, trnfM-CAU</i>
Other genes	Maturase	<i>matK</i>
	Protease	<i>clpP**</i>
	Envelope membrane protein	<i>cemA</i>
	Acetyl-CoA carboxylase	<i>accD</i>
	c-type cytochrome synthesis gene	<i>ccsA</i>
Genes of unknown function	Conserved hypothetical chloroplast ORF	<i>#ycf1, ycf1, ycf15(2), ycf2(2), ycf3**, ycf4</i>

Notes: Gene*: Gene with one introns; Gene**: Gene with two introns; #Gene: Pseudo gene; Gene(2): Number of copies of multi-copy genes.

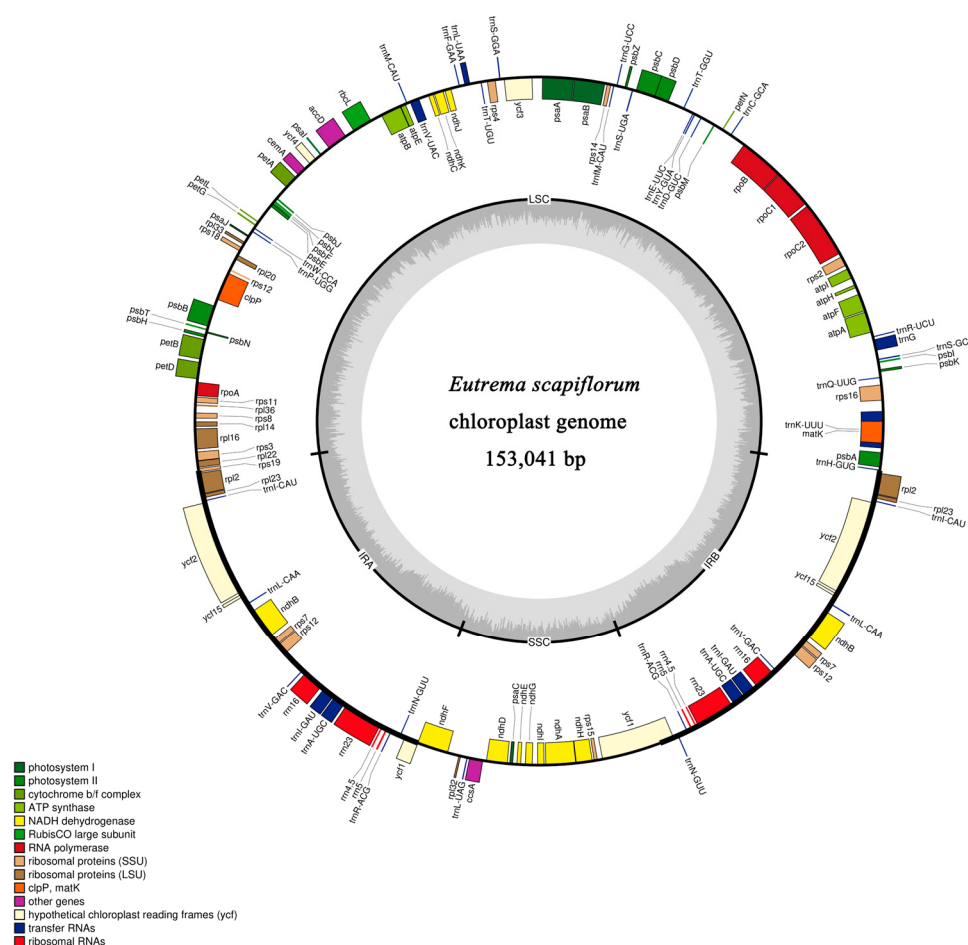


Figure 1. Circularized map of the chloroplast genome of *E. scapiflorum*. The genome shows a typical quadripartite structure: large single-copy (LSC), small single-copy (SSC), and two inverted repeat (IRa and IRb) regions. Genes inside the circle are transcribed clockwise; genes outside are transcribed counterclockwise. Different functional groups are labeled with distinct colors. The inner circle shows the GC content variation across the genome.

2.2. Analysis of SSR Loci

Simple sequence repeats (SSRs) are tandemly repeated sequences composed of 1–6 nucleotide motifs. A total of 78 SSR loci were identified in the *E. scapiflorum* chloroplast genome. These comprised 58 mononucleotide repeats (A, G, and T), 3 dinucleotide repeats (TA), 8 trinucleotide repeats (TTC, AAT, ATT, TAC, and TTG), 7 tetranucleotide repeats (AGAA, ATAG, CAAA, CTTT, GATA, TATT, and TTTC), and 2 pentanucleotide repeats (GGAGA, TCCTC). No hexanucleotide repeats were detected (Figure 2). Analysis of genomic distribution revealed that 54.30% of the SSR loci were situated in intergenic spacers (IGSs), 32.78% in exons, and 12.91% in introns (Table 2). The high abundance of A/T mononucleotide repeats is a conserved feature consistent with most angiosperm plastomes.

A total of 33 dispersed repeats were detected in the *E. scapiflorum* plastome. These included 17 palindromic, 14 forward, 1 reverse, and 1 complementary repeat (Table 3). Palindromic and forward repeats were the predominant types. The lengths of these dispersed repeats ranged from 30 to 65 bp. The predominance of palindromic and forward repeats aligns with expected patterns in related Brassicaceae species.

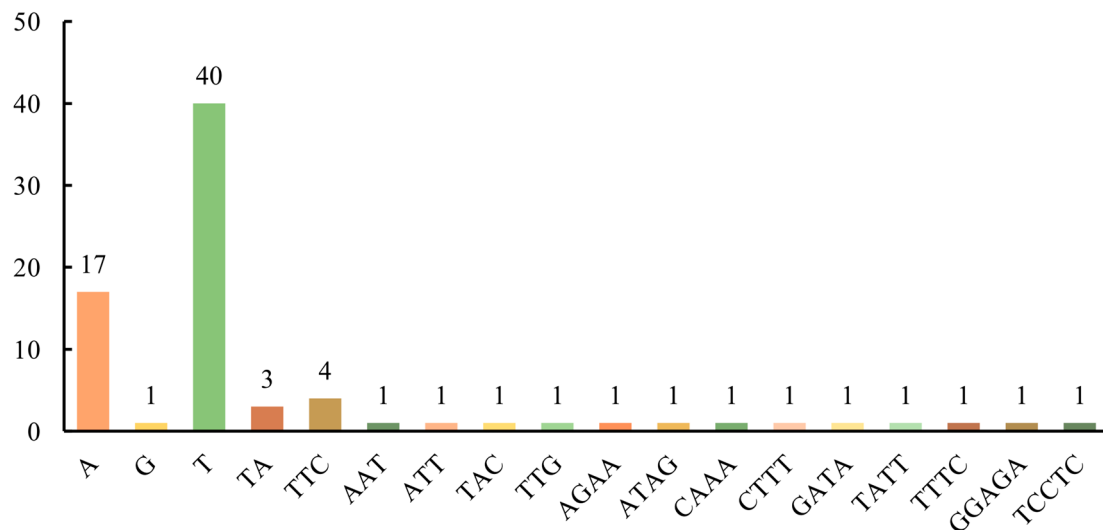


Figure 2. Types and numbers of SSRs in the chloroplast genome of *E. scapiflorum*. Monomers, dimers, trimers, tetramers, and pentamers are shown. No hexanucleotide repeats were detected. Frequencies of each repeat motif are displayed as bar heights.

Table 2. Reactors and numbers of SSRs in the chloroplast genome of *E. scapiflorum*.

Region	Exon	Intron	Intergenic
LSC	37	31	129
SSC	40	4	19
IR	22	4	16
	32.78%	12.91%	54.30%

Table 3. Statistics of the analysis results of scattered repetitive sequences.

Length of Repeat (bp)	Forward	Palindromic	Reverse	Complement	Total
30	5	7	0	1	13
31	2	0	0	0	2
32	2	1	0	0	3
33	2	2	0	0	4
34	0	0	1	0	1
36	0	1	0	0	1
37	1	0	0	0	1
39	0	1	0	0	1
41	0	1	0	0	1
43	1	0	0	0	1
44	0	2	0	0	2
46	1	1	0	0	2
55	0	1	0	0	1

2.3. Genome Collinearity Analysis

Collinearity analysis was performed using the Mauve model to identify potential genome rearrangements and inversions across the chloroplast genomes of eight species. As shown in Figure 3, the alignment revealed distinct locally collinear blocks (LCBs). All LCBs were oriented in the same direction, with no regions appearing below the central axis, indicating a lack of inversions. Consequently, no rearrangements or inversions were observed, suggesting that the chloroplast genome structure is highly conserved within *E. scapiflorum*.

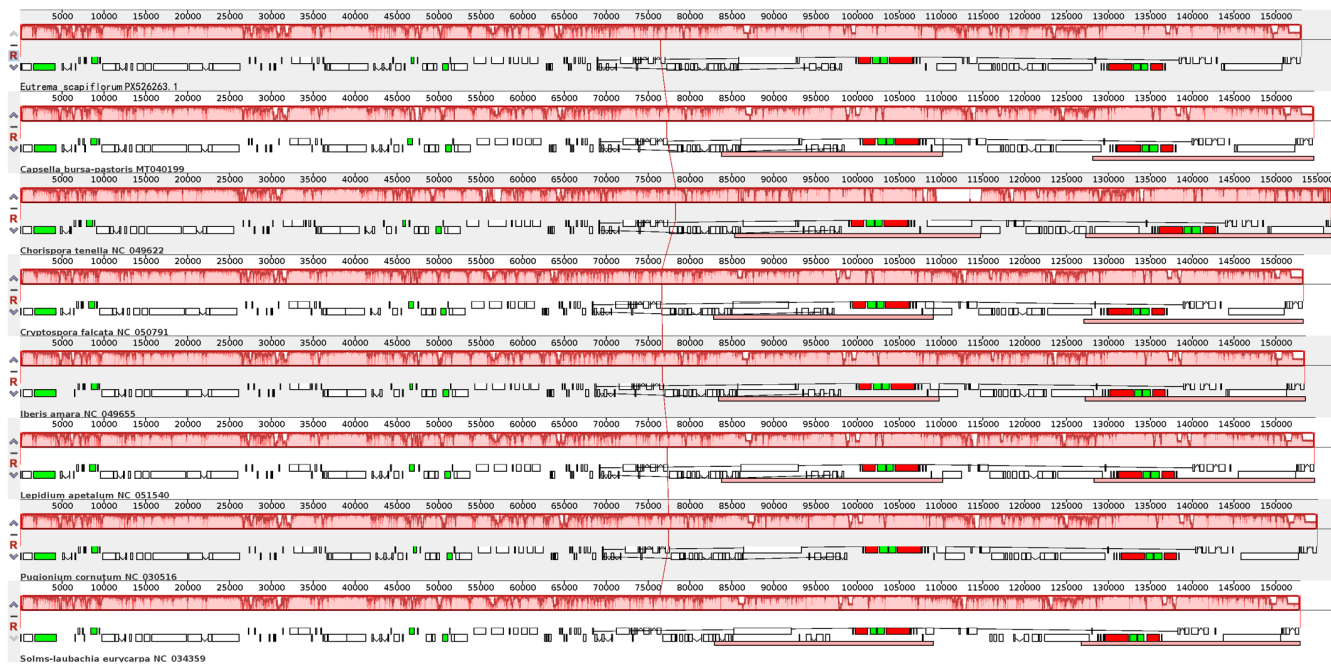


Figure 3. Analysis of chloroplast sequence homology. Locally Collinear Blocks (LCBs) are indicated by colored blocks. Colored blocks represent locally collinear blocks (LCBs) shared among genomes. Blocks above the central are in the forward orientation, while blocks below indicate inversions. Red vertical lines mark LCB boundaries and potential structural variation sites. No large-scale rearrangements were detected among the compared species. All blocks are oriented in the same direction, indicating no large-scale inversions or genome rearrangements. *E. scapiflorum* was used as the reference sequence.

2.4. Analysis of IR Boundary Regions

To investigate the expansion and contraction of the inverted repeat (IR) regions in chloroplast genomes for *E. scapiflorum*, a comparative analysis of IR boundaries was performed on the chloroplast genomes of eight species. Exact boundary locations and region lengths are summarized in Table 4. The results revealed variation in both total genome size and the lengths of individual regions among the analyzed species (Figure 4). Total genome lengths ranged from 152,895 to 156,666 bp. The lengths of the LSC, SSC, and IR regions ranged from 82,927 to 85,364 bp, 12,504 to 18,014 bp, and 25,994 to 29,399 bp, respectively. Notably, *Chorispora tenella* exhibited significant structural divergence. When *C. tenella* was excluded, the length variation decreased substantially: the LSC range narrowed to 82,927–84,104 bp, the SSC to 17,419–18,014 bp, and the IR to 25,994–26,454 bp. Thus, the SSC region was identified as the primary source of genome size variation. In contrast, *E. scapiflorum* showed a highly conserved structure, with regional lengths consistent with those of the other related species.

Table 4. Comparison of IR/SC boundary positions (bp) among eight Brassicaceae chloroplast genomes.

Species	LSC Length	IRa Length	SSC Length	IRb Length	Total Length	JLB	JSB	JSA	JLA
<i>E. scapiflorum</i>	83,547	25,994	17,506	25,994	153,041	<i>ycf2</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf2</i>
<i>A. petiolata</i>	82,927	26,454	18,014	26,454	153,849	<i>ycf2</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf2</i>
<i>C. himalaica</i>	83,421	26,378	17,764	26,378	153,941	<i>ycf2</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf2</i>
<i>S. cheiranthoides</i>	83,672	26,295	17,643	26,295	153,905	<i>ycf2</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf2</i>
<i>T. glabra</i>	83,714	26,183	17,826	26,183	153,906	<i>ycf2</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf2</i>

Table 4. Cont.

Species	LSC Length	IRa Length	SSC Length	IRb Length	Total Length	JLB	JSB	JSA	JLA
<i>C. limprichtiana</i>	84,104	26,077	17,419	26,077	153,677	<i>ycf2</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf2</i>
<i>L. apetalum</i>	84,362	26,245	17,531	26,245	154,383	<i>ycf2</i>	<i>ycf1</i>	<i>ycf1</i>	<i>ycf2</i>
<i>C. tenella</i>	85,364	29,399	12,504	29,399	156,666	<i>rps19</i>	<i>ycf1</i>	<i>ycf1</i>	<i>rps19</i>

Note: JLB: junction between IRb and LSC; JSB: junction between IRb and SSC; JSA: junction between SSC and IRa; JLA: junction between IRa and LSC. Boundary positions are given as gene names located at each junction.

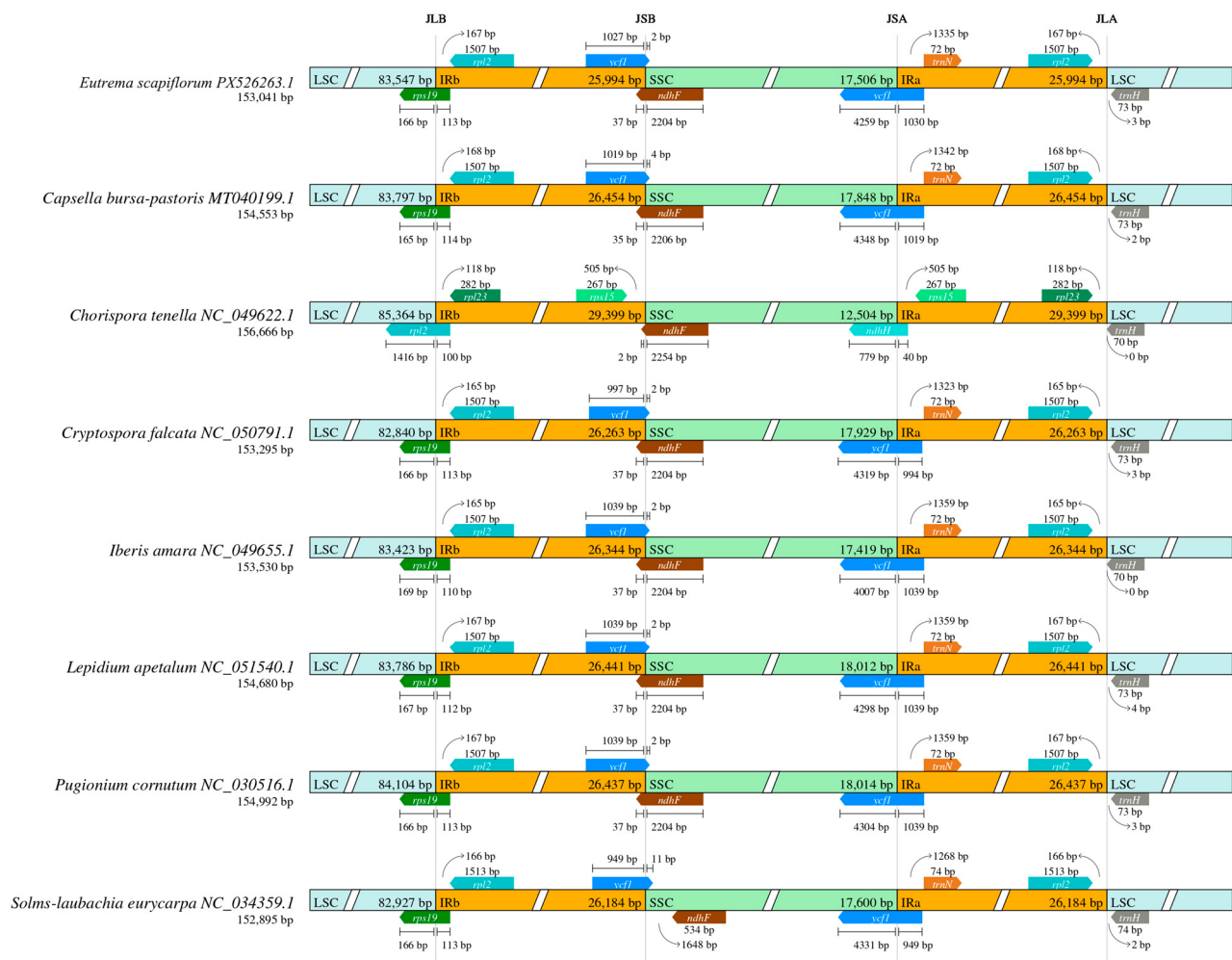


Figure 4. Schematic diagram of chloroplast genome boundary regions of 8 species. JLB: IRb/LSC boundary; JSB: IRb/SSC boundary; JSA: SSC/IRa boundary; JLA: IRa/LSC boundary. Lengths of LSC, SSC, and IR regions are labeled in base pairs (bp). *Chorispora tenella* shows obvious IR expansion compared with the other species.

2.5. Polymorphic Site Analysis

Single nucleotide polymorphisms (SNPs) quantify the extent of nucleic acid sequence variation among species. In this study, nucleotide diversity (P_i) across the *E. scapiflorum* chloroplast genome ranged from 0 to 0.13031. Notably, *accD* and *ycf1* were identified as highly variable hotspots, exhibiting exceptionally high P_i values of 0.13031 and 0.09711, respectively (Table S1, Figure 5). To rule out potential artifacts, these high P_i values were manually verified to ensure they were not driven by poor alignment quality, structural variations, or annotation discrepancies. Given their substantial divergence ($P_i > 0.09$), the observed variations in *accD* and *ycf1* provide sufficient phylogenetic signal, indicating that they can serve as practical and robust molecular markers for future species delimitation.

and population genetic studies in *Eutrema*, rather than merely representing descriptive divergent regions.

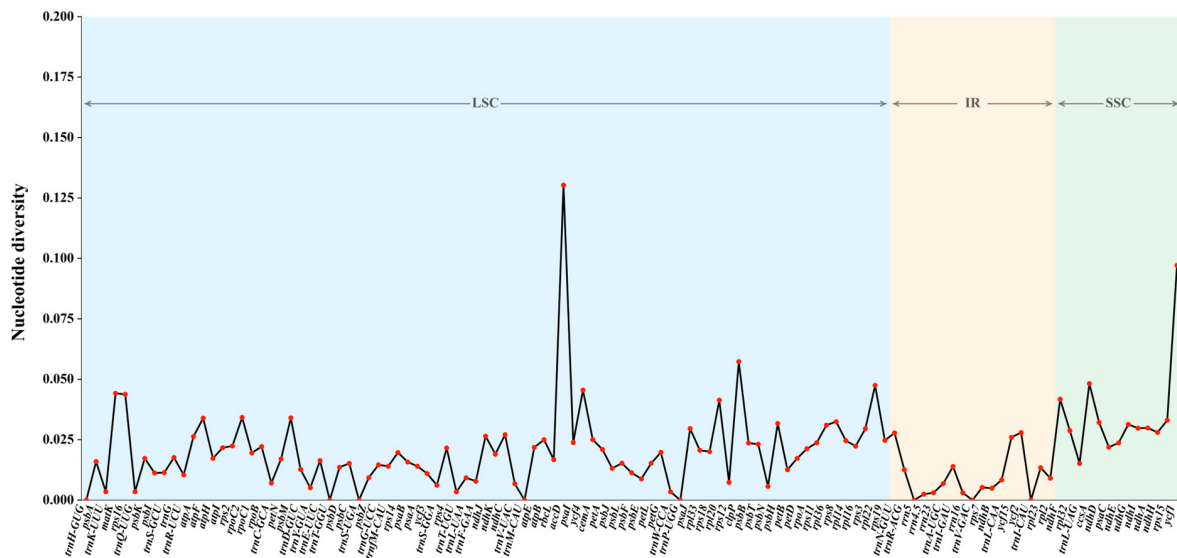


Figure 5. Pi values of chloroplast genome nucleotide polymorphisms in *E. scapiflorum*. Pi values were calculated using DnaSP with a 200-bp window and 600-bp step size. LSC: large single-copy region; IR: inverted repeat region; SSC: small single-copy region. Genes with the highest Pi values are indicated.

2.6. Ka/Ks Analysis

Selective pressure analysis was performed on 76 protein-coding genes shared among *E. scapiflorum* and seven related species, following the removal of invalid data ($Ka/Ks = 0$ or NA). Seven genes exhibited Ka/Ks ratios exceeding 1, including *ycf2*, *ndhE*, *cemA*, *clpP*, *psbH*, *ycf4*, and *ndhB*, with values ranging from 1.001 to 1.594. These ratios indicate that these genes are subject to positive selection. In contrast, the remaining genes showed Ka/Ks ratios below 1, suggesting a predominance of synonymous substitutions and the influence of purifying selection (Figure 6).

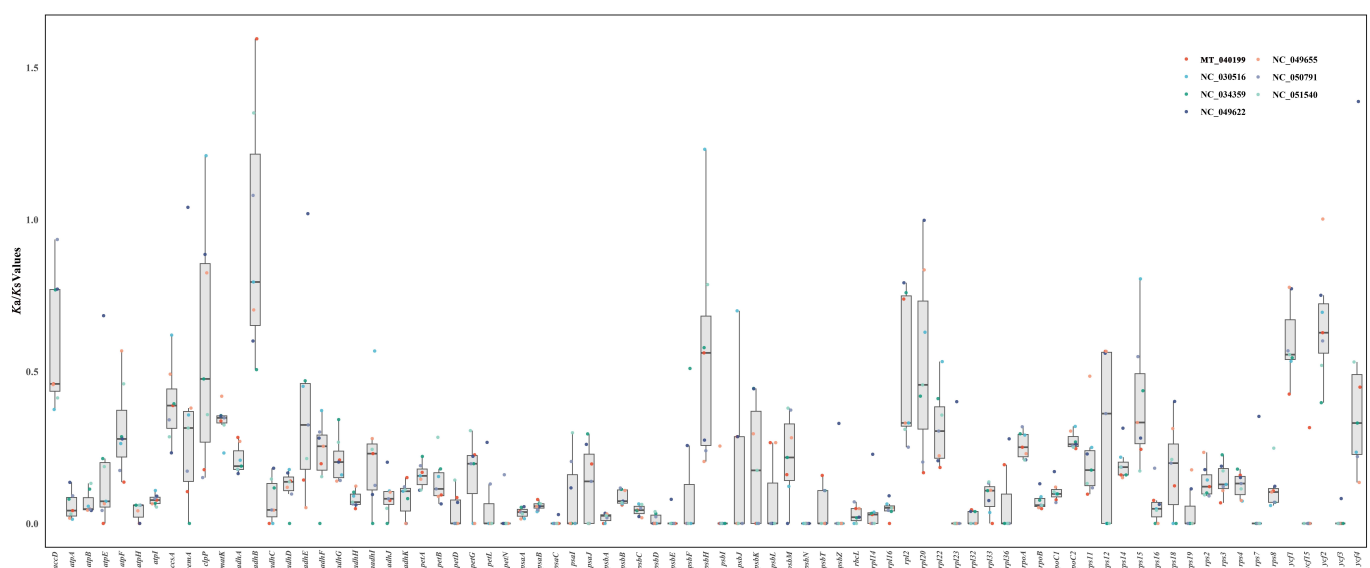


Figure 6. Selection pressure analysis of protein-coding genes in chloroplast genomes of *E. scapiflorum* and eight species. Genes with $Ka/Ks > 1$ are marked as candidates for potential positive selection. Most genes show $Ka/Ks < 1$, indicating purifying selection.

2.7. Phylogenetic and Divergence Time Analysis

To elucidate the phylogenetic relationships within *E. scapiflorum*, phylogenetic trees were reconstructed for 36 Brassicaceae species using both complete chloroplast genomes and CDSs. The phylogenetic analysis based on CDSs resolved the species into four highly supported clades (BS = 100, PP = 1.00). Within this topology, *E. scapiflorum* clustered within a stable clade (BS = 100, PP = 1.00) together with *A. petiolata*, and *S. altissimum* from the tribe Sisymbrieae; *A. flagellosa*, *A. hirsuta*, *A. paniculata*, and *D. fruticulosa* from the tribe Arabideae; *B. oleracea*, *C. limprichtiana*, *O. violaceus*, and *R. sativus* from the tribe Brassiceae; *C. planisiliqua* from the tribe Conringieae; and *I. amara*, *I. gymnocarpa*, and *P. cornutum* from the tribe Lepidieae. *E. scapiflorum* was found to be most closely related to *A. petiolata* (Figure 7). In comparison, the tree based on complete chloroplast genomes yielded a similar topology but identified five clades, with *L. apetalum* forming a separate lineage (BS = 87, PP = 1.00). Despite this minor discrepancy, both datasets consistently supported a sister relationship between *E. scapiflorum* and *A. petiolata* (BS = 98/100, PP = 1.00; Figure 7). Overall, the chloroplast genome data provided strong resolution, establishing a reliable phylogenetic framework for *E. scapiflorum*.

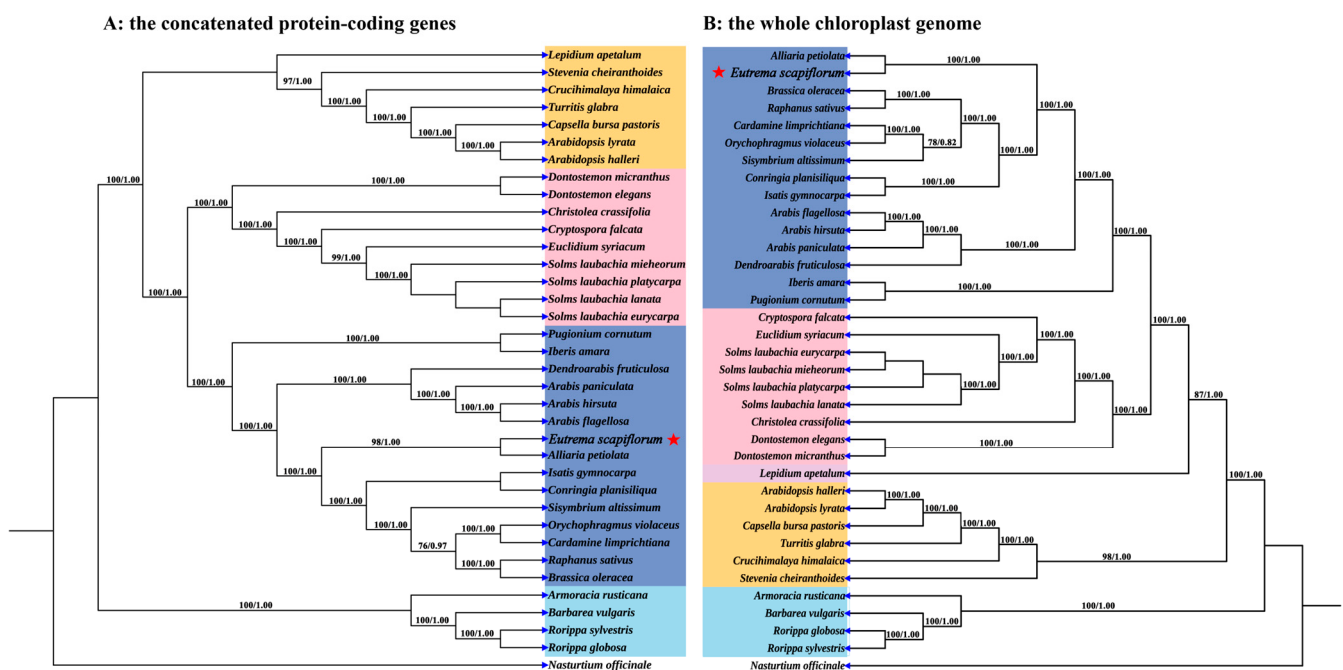


Figure 7. Phylogenetic relationships of 36 Brassicaceae species. The focal species, *E. scapiflorum*, is highlighted with a star (★). (A) is the phylogenetic tree constructed by protein-coding sequences, and (B) is the phylogenetic tree constructed by chloroplast whole genome sequences. Both phylogenies were reconstructed using Maximum Likelihood (ML) and Bayesian Inference (BI) methods. Numbers at the nodes represent ML bootstrap support values and BI posterior probabilities (ML/BI), respectively. Major clades are consistently color-coded across both panels to facilitate visual comparison.

Divergence times for *E. scapiflorum* were estimated using a Bayesian relaxed clock model based on CDSs and secondary calibration points. The resulting time-calibrated phylogeny indicates that the analyzed Brassicaceae lineage diversified during the early Miocene, approximately 22.98 Ma (95% highest posterior density: 17.50–24.56 Ma; Figure 8). Subsequently, the divergence between *E. scapiflorum* and *A. petiolata* was estimated to have occurred around 17.57 Ma (95% HPD: 14.59–20.60 Ma), corresponding to the transition from the early to middle Miocene.

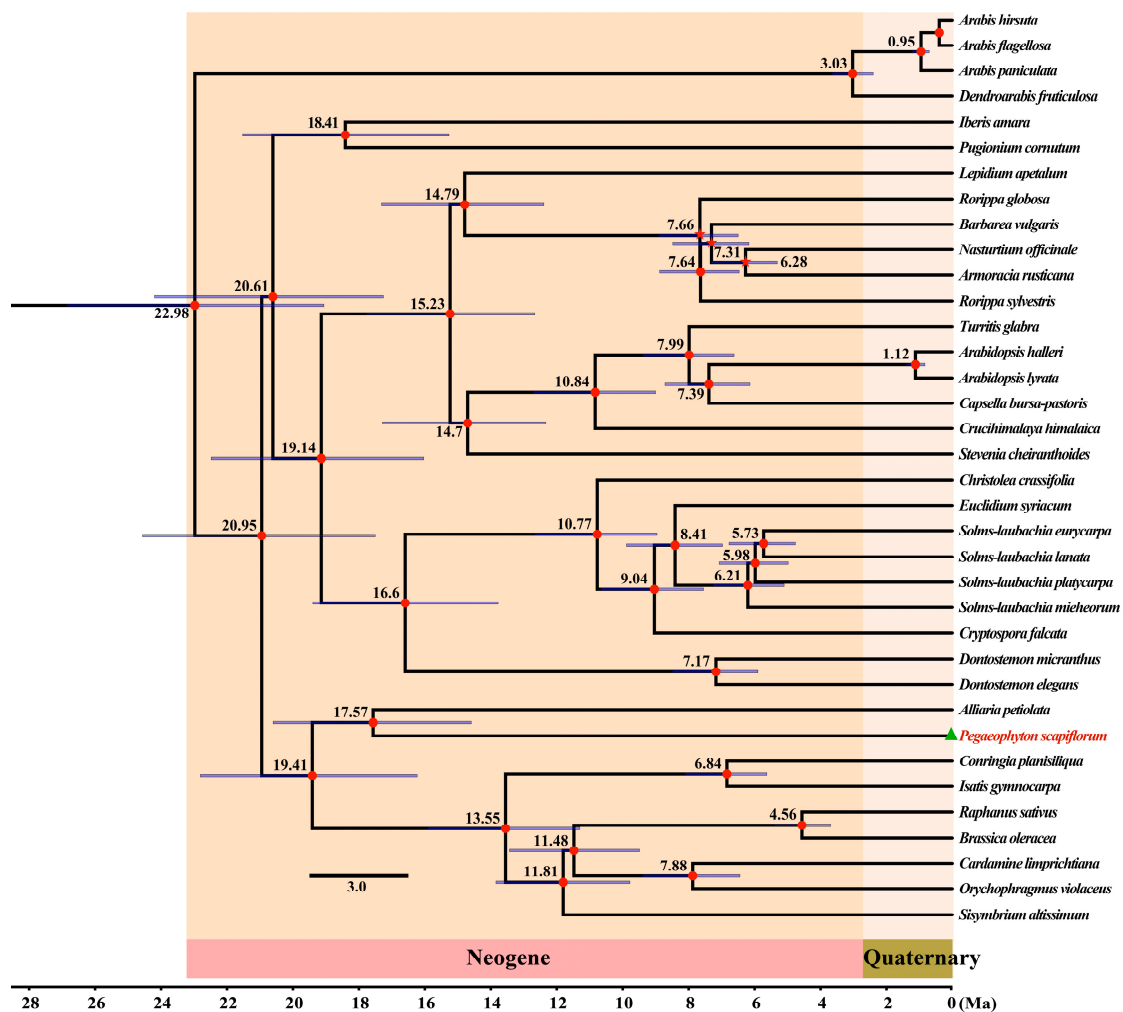


Figure 8. Divergence time estimation based on the phylogenetic tree of chloroplast genomes of 36 species. Numbers at the nodes indicate the median estimated divergence times (Ma). Blue bars represent the 95% highest posterior density (HPD) intervals. The bottom axis displays the geological timescale. The two red stars indicate the calibration points used for divergence time estimation, while the different colors represent distinct geological ages.

3. Discussion

3.1. Comparative Plastome Features and Structural Evolution

Our study presents the first complete chloroplast genome of *E. scapiflorum*, a characteristic alpine species of the QTP. The assembled plastome is 153,041 bp in length and exhibits a characteristic quadripartite structure. Its gene organization, composition, and overall GC content (36.41%) are consistent with those reported for other species [18,19]. Although the observed variation in GC content across structural regions (IRs > LSC > SSC) is consistent with the known pattern driven by GC-rich rRNA genes in the IRs [20], this feature alone does not distinguish *E. scapiflorum* from its close relatives. The annotation identified 132 genes, including 87 protein-coding genes, which represents the conserved gene repertoire typical of autotrophic Brassicaceae plastomes [21].

A more informative finding emerges from comparative analysis of IR boundaries: unlike the pronounced expansion or contraction observed in some alpine or short-lived Brassicaceae species (e.g., significant IR expansion into single-copy regions in *C. tenella*; Figure 4), *E. scapiflorum* exhibits highly conserved IR/SC boundaries relative to closely related taxa. This structural stability is further supported by whole-genome synteny alignments, which reveal no large-scale inversions or rearrangements (Figure 3). Consequently,

macrostructural plastome variation does not appear to be a primary evolutionary driver in this lineage. Nevertheless, the presence of dispersed repeats, particularly palindromic sequences, however, indicates potential recombination hotspots that could facilitate structural variation over extended evolutionary timescales [22,23].

3.2. Genome Divergence, Hotspots, and Implications for Molecular Markers

Our nucleotide diversity analysis identified several highly variable regions in *E. scapiflorum* and related taxa. Consistent with angiosperm-wide patterns [24], intergenic spacers and specific coding regions exhibited elevated Π values. Notably, *accD* and *ycf1* were the most divergent protein-coding genes ($\Pi > 0.09$). The *ycf1* gene, spanning the SSC/IR boundary, is widely recognized as one of the most variable plastid genes. It has proven valuable for resolving shallow phylogenetic relationships [24], a utility that can now be extended to *E. scapiflorum*.

Furthermore, SSR analysis revealed 78 loci, with >50% located in intergenic spacers. The prevalence of A/T-rich repeats in non-coding regions, which is common across flowering plants, contributes to their high polymorphism and makes them excellent candidates for population-specific microsatellite markers. Future studies should leverage these SSRs and hypervariable regions (e.g., *trnH-psbA*, *rpl32-trnL*) to investigate the genetic diversity, population structure, and phylogeography of *E. scapiflorum*. Such efforts, particularly across the topographically complex QTP, will be crucial for understanding the species' response to Quaternary climatic oscillations [25].

3.3. Evolutionary Constraints and Positive Selection

Analysis of K_a/K_s substitution rates reveals that the majority of plastid genes in *E. scapiflorum* exhibit K_a/K_s ratios significantly below 1, consistent with purifying selection maintaining the integrity of genes critical for photosynthesis and metabolism [26]. However, seven genes (*ycf2*, *ndhE*, *cemA*, *clpP*, *psbH*, *ycf4*, *ndhB*) showed $K_a/K_s > 1$. While this pattern is statistically consistent with relaxed selective constraints or positive selection, we caution against overinterpreting these values as definitive evidence of adaptive evolution without additional statistical (e.g., branch-site models) or functional validation [27]. With this caveat, *clpP* (encoding a proteolytic subunit) and *psbH* (involved in photosystem II stabilization) emerge as plausible candidates for further investigation, given their known roles in protein homeostasis and photosynthetic efficiency under environmental stress [27,28]. Additionally, the *accD* gene, which is essential for fatty acid biosynthesis, exhibited the highest nucleotide diversity ($\Pi = 0.13031$). Collectively, these genes warrant targeted study to test whether they contribute to molecular adaptation of *E. scapiflorum* to alpine abiotic stresses (e.g., intense UV radiation, freezing temperatures, oxidative stress). At present, the detected signals are best interpreted as hypothesis-generating rather than conclusive evidence of positive selection.

3.4. Phylogenetic Implications and Taxonomic Insights

Our phylogenomic analyses robustly resolve *A. petiolata* as the sister taxon to *E. scapiflorum*. This relationship illuminates a striking ecological divergence: *A. petiolata* is a biennial herb restricted to forest understories, whereas *E. scapiflorum* is a cushion-forming perennial specialized for high-altitude scree slopes (>3500 m). Given the conserved chloroplast genomic architecture in *E. scapiflorum*, this dramatic niche differentiation suggests that their common ancestor possessed substantial adaptive plasticity, enabling divergent responses to contrasting selective regimes [2,29]. We therefore interpret the cushion-forming perennial habit of *E. scapiflorum* as a derived character state reflecting specialized adaptation to persistent alpine stressors. This finding is consistent with the hypothesis of evolutionary lability and convergent evolution of life-history strategies within alpine Brassicaceae [30].

The estimated divergence time (ca. 17.57 Ma, 95% HPD: 14.59–20.60 Ma) places the evolutionary split of this lineage firmly within the early to middle Miocene. This timeframe, encompassing the estimated HPD interval, coincides temporally with major orogenic uplift and climatic cooling of the Plateau and Himalayas. These Neogene environmental shifts likely promoted diversification and adaptive evolution of the lineage, ultimately leading to the endemic establishment of *E. scapiflorum* [2,16,17,31]. Thus, its evolutionary trajectory appears closely linked to the formation of new alpine habitats and ecological niches during plateau orogenesis. Colonization of scree slope environments may represent a key adaptation that enabled the lineage to exploit these nascent niches, promoting its survival and diversification amidst profound geological heterogeneity.

We acknowledge that evolutionary inferences based solely on the plastid genome are subject to limitations due to its uniparental (maternal) inheritance. Future studies employing nuclear markers or biparentally inherited loci will be necessary to validate the divergence times and phylogenetic relationships presented here.

4. Materials and Methods

4.1. Experimental Materials

Fresh and healthy leaves of *E. scapiflorum* were collected in Maduo County, Qinghai Province (97°51'54" E, 34°24'40" N). The collected leaves were immediately dried and preserved using color-changing silica gel. The DNA voucher has been deposited at the College of Agriculture and Animal Husbandry, Qinghai University (voucher No. 2024-026). Collection was conducted with approval and in compliance with local biodiversity regulations.

4.2. Total DNA Extraction

Genomic DNA was extracted from *E. scapiflorum* leaf tissue using a modified CTAB (cetyltrimethylammonium bromide) method [32]. The integrity of the extracted DNA was subsequently assessed by electrophoresis on a 1.0% agarose gel. Its concentration and purity were then determined by measuring the OD_{260/280} ratio using a Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA).

4.3. Sequencing, Assembly, and Annotation

The quality-controlled DNA samples were randomly fragmented using a Covaris ultrasonicator (Covaris LLC, Woburn, MA, USA). The DNA fragments underwent end repair, poly-A tailing at the 3' end, adapter ligation, and purification. Then, a sequencing library was constructed with PCR amplification. Sequencing was performed on an Illumina NovaSeq 6000 platform (Genepioneer, Nanjing, China) to generate raw data. Raw sequencing reads were filtered using Fastp v0.23.2 [33] with default parameters to remove adapter sequences, low-quality reads (quality score < 20), short reads (<50 bp), and PCR duplicates. The complete chloroplast genome was de novo assembled using GetOrganelle v1.7.7 [34] with the embryophyte plastid database (-F embplant_pt), a k-mer gradient of 21, 45, 65, 85, and 105, and 16 threads.

Furthermore, the integrity of the inverted repeat (IR) boundaries was validated by mapping continuous reads across the junctions in Geneious v2023.0.1. The assembly graph and genome circularity were visually evaluated using Bandage v0.9.0. To validate the assembly and confirm circularization, approximately 5 Gb of raw reads were processed, filtered, and then remapped to the assembled plastome. After filtering and alignment, 3,275,427 paired-end reads were successfully mapped to the chloroplast genome, achieving an average coverage depth of 6797×. Initial annotation was annotated using the online tool GeSeq v2.03 (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>, accessed

on 10 September 2025) [35] (BLAT identity $\geq 85\%$), with the chloroplast genome of *Alliaria petiolata* (NC_049586.1) as a reference.

Subsequently, annotation was performed using the “2544-plastomes” database within the CPGAVAS2 software (<http://47.96.249.172:16019/analyzer/home>, accessed on 10 September 2025) [36]. Conflicting annotations between the two methods were resolved by manual curation based on sequence homology, open reading frame (ORF) integrity, and conserved gene structures in Brassicaceae. Finally, a circular map of the *E. scapiflorum* chloroplast genome was generated using OGDRAW v1.3.1 (OrganellarGenomeDRAW, <https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>, accessed on 10 September 2025) [37].

4.4. Analysis of Chloroplast Genome Features and Repeat Sequences

Geneious Prime 2023.0.1 software [38] was employed to determine the total length of the *E. scapiflorum* chloroplast genome and the lengths of its structural regions, including the large single-copy (LSC), small single-copy (SSC), and inverted repeat (IR) regions. The types, numbers, and base composition of genes were also analyzed and compared to elucidate the sequence characteristics of the *E. scapiflorum* chloroplast genome. Simple sequence repeat (SSR) loci within the *E. scapiflorum* chloroplast genome sequence were identified using the online tool MISA (<https://webblast.ipk-gatersleben.de/misa/index.php?action=1>, accessed on 12 September 2025) [39,40]. The detection parameters for mono- to hexanucleotide repeats were set to 10, 6, 5, 3, 3, and 3, respectively. Dispersed repeat sequences in the *E. scapiflorum* chloroplast genome, including palindromic (P), complementary (C), forward (F), and reverse (R) repeats, were detected using the REPuter (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>, accessed on 12 September 2025) [41]. The search parameters were set to a minimum repeat length of 30 bp, sequence similarity greater than 90%, and a Hamming distance of 3.

4.5. Chloroplast Genome Synteny Comparison

To identify structural rearrangements and inversions, collinearity analysis of the chloroplast genomes from eight species (including *E. scapiflorum*) was conducted using the progressiveMauve algorithm (Mauve v2.4.0) [42] within Geneious v2023.0.1, applying default parameters. The analysis was performed under default parameters, with the *E. scapiflorum* chloroplast genome serving as the reference.

The IRscope online tool (<https://irscope.shinyapps.io/irapp/>, accessed on 14 September 2025) [43] was employed to compare the inverted repeat (IR) region boundaries and analyze IR contraction and expansion among the eight chloroplast genome sequences, thereby detecting boundary shifts in the four genomic regions. The annotated chloroplast genome sequences were sequentially uploaded to automatically generate a visual comparative map. This was followed by manual verification of the positions of each region to ensure the accuracy of the comparison.

4.6. Nucleotide Diversity Calculation

The nucleotide polymorphism (Pi) value was calculated for the *E. scapiflorum* chloroplast genome using DnaSP v6.034 software [44] to identify regions of high nucleotide polymorphism in both coding and non-coding regions. The analysis was performed with a window length of 200 bp and a step size of 600 bp. The generated Pi value box plot was compared against the genome visualization map to determine the corresponding gene regions, and each variable region was annotated accordingly.

4.7. Ka/Ks Analysis

Nucleotide and amino acid sequences of protein-coding genes were extracted from the chloroplast genomes of *E. scapiflorum* and seven related species using PhyloSuite v1.2.2 [45]. Following the identification of homologous genes, nucleotide and amino acid sequences were aligned using ParaAT v2.0 [46] to produce .axt format files. Subsequently, the nonsynonymous (Ka) and synonymous (Ks) substitution rates were calculated using Kaks_Calculator 3.0 [47] for the eight species. The Ka/Ks ratio was used to evaluate selective pressure, where a ratio > 1 indicates positive selection and a ratio < 1 indicates purifying selection [48]. In the study, the genetic code table was set to 11 (Bacterial and Plant Plastid Code), and the GY calculation method (maximum-likelihood) was selected for analysis.

4.8. Phylogenetic Analysis

To further resolve the phylogenetic position of *E. scapiflorum* within the family, a more comprehensive dataset was constructed. Chloroplast genome sequences of 35 species representing various tribes within Brassicaceae were retrieved from the NCBI database (Table 5). Protein-coding sequences (CDS) were extracted from the chloroplast genomes of all 36 analyzed species using Geneious v2023.0.1. Multiple sequence alignments for both whole chloroplast genomes and the concatenated CDSs were performed using MAFFT v7.526 [49] integrated into PhyloSuite platform. The aligned sequences were subsequently trimmed using trimAl v1.5.0 [50]. Phylogenetic analyses were conducted using both Maximum Likelihood (ML) and Bayesian Inference (BI) methods.

Table 5. Information table of tested species.

Tribe	Genus	Species	GenBank ID
Arabideae	<i>Arabidopsis</i>	<i>Arabidopsis halleri</i>	NC_034366.1
		<i>Arabidopsis lyrata</i>	NC_034365.1
		<i>Arabis flagellosa</i>	NC_037475.1
	<i>Arabis</i>	<i>Arabis hirsuta</i>	NC_009268.1
		<i>Arabis paniculata</i>	NC_053754.1
	<i>Barbarea</i>	<i>Barbarea vulgaris</i>	NC_081109.1
	<i>Christolea</i>	<i>Christolea crassifolia</i>	NC_050790.1
	<i>Dendroarabis</i>	<i>Dendroarabis fruticulosa</i>	NC_062040.1
		<i>Dontostemon elegans</i>	NC_081114.1
	<i>Dontostemon</i>	<i>Dontostemon micranthus</i>	NC_049628.1
		<i>Nasturtium officinale</i>	NC_009275.1
	<i>Rorippa</i>	<i>Rorippa globosa</i>	NC_070416.1
		<i>Rorippa sylvestris</i>	NC_069649.1
	<i>Stevenia</i>	<i>Stevenia cheiranthoides</i>	MK_637795.1
Brassicaceae	<i>Turritis</i>	<i>Turritis glabra</i>	NC_061285.1
	<i>Brassica</i>	<i>Brassica oleracea</i>	NC_041167.1
	<i>Cardamine</i>	<i>Cardamine limprichtiana</i>	NC_034287.1
	<i>Raphanus</i>	<i>Raphanus sativus</i>	NC_024469.1
	<i>Orychophragmus</i>	<i>Orychophragmus violaceus</i>	NC_049687.1
		<i>Conringia planisiliqua</i>	NC_049619.1
Conringieae	<i>Conringia</i>	<i>Conringia planisiliqua</i>	NC_049619.1
Drabeae	<i>Armoracia</i>	<i>Armoracia rusticana</i>	NC_060501.1
Euclidieae	<i>Euclidium</i>	<i>Euclidium syriacum</i>	NC_081125.1
Hesperideae	<i>Cryptospora</i>	<i>Cryptospora falcata</i>	NC_050791.1
		<i>Capsella bursa-pastoris</i>	NC_009270.1
	<i>Iberis</i>	<i>Iberis amara</i>	NC_049655.1
		<i>Isatis gymnocarpa</i>	NC_081128.1
Lepidieae	<i>Lepidium</i>	<i>Lepidium apetalum</i>	NC_051540.1
	<i>Pugionium</i>	<i>Pugionium cornutum</i>	NC_030516.1
		<i>Solms-laubachia eurycarpa</i>	NC_034359.1
Matthioleae	<i>Solms-laubachia</i>	<i>Solms-laubachia lanata</i>	NC_050812.1
		<i>Solms-laubachia mieheorum</i>	NC_053342.1
		<i>Solms-laubachia platycarpa</i>	NC_050816.1

Table 5. Cont.

Tribe	Genus	Species	GenBank ID
Sisymbrieae	<i>Alliaria</i>	<i>Alliaria petiolata</i>	NC_062040.1
	<i>Crucihimalaya</i>	<i>Crucihimalaya himalaica</i>	NC_061290.1
	<i>Eutrema</i>	<i>Eutrema scapiflorum</i> *	PX_526263.1
	<i>Sisymbrium</i>	<i>Sisymbrium altissimum</i>	NC_059802.1

Note: * The newly assembled chloroplast genome sequence was used in this study.

The ML tree was constructed using IQ-TREE v2.2.0 [51] with codon partitioning, substitution models selected by ModelFinder v2.2.0, and 5000 ultrafast bootstrap replicates. The BI tree was inferred using MrBayes v3.2.7 [52] with two independent runs of four Markov chains for 10,000,000 generations, sampling every 1000 generations, and discarding the first 25% as burn-in. Convergence was confirmed in Tracer v1.7.1 [53] (effective sample size [ESS] > 200). Support values are presented as ML bootstrap/BI posterior probability at each node.

4.9. Divergence Time Estimation

Divergence times among the analyzed species were estimated based on the concatenated CDS using a Bayesian relaxed-clock model implemented in BEAST v2.7.3. Secondary calibration points were obtained from the TimeTree database (<http://www.timetree.org/>, accessed on 18 September 2025) [54] to constrain well-supported nodes. Specifically, the divergence times between *Nasturtium officinale* and *Armoracia rusticana* (7.4 Ma, 95% HPD: 6.3–8.5 Ma), and between *Barbarea vulgaris* and *Rorippa sylvestris* (6.4 Ma, 95% HPD: 3.3–14.0 Ma) were employed as calibration constraints. A normal prior distribution was applied to these secondary calibrations, as it symmetrically accommodates the mean ages and their empirical 95% confidence intervals derived from the TimeTree database, minimizing subjective parameterization. The analysis utilized a Yule speciation prior, which is appropriate for species-level phylogenies. Given the strong phylogenetic signal provided by the concatenated CDS dataset, the posterior estimates are expected to be robust and primarily driven by the sequence data rather than prior sensitivity. The Markov Chain Monte Carlo (MCMC) was run for 100 million generations, sampling every 1000 generations. Convergence was evaluated in Tracer v1.7.1, ensuring all parameters reached an ESS > 200. The final maximum clade credibility tree was generated using TreeAnnotator v2.7.8 with a 10% burn-in, and node ages were reported with 95% highest posterior density (HPD) intervals to represent uncertainty.

5. Conclusions

This study reported the first complete chloroplast genome of *E. scapiflorum*, a characteristic alpine species endemic to the QTP. Our analyses resolved its precise phylogenetic position, revealed a highly conserved plastome structure, and identified candidate genes potentially under positive selection associated with high-altitude adaptation. The divergence time estimation linked its speciation of *E. scapiflorum* to the dynamic geological history of the QTP. The comprehensive genomic data generated herein provide a foundational resource for future evolutionary and ecological studies. Specifically, the identified SSRs, particularly those located in hypervariable intergenic spacer regions, provided valuable resources for population genetic studies of *E. scapiflorum*, enabling the exploration of its genetic structure, demographic history, and local adaptation. Future research should focus on expanding sampling across the species' distribution to test phylogeographic hypotheses using these markers. Furthermore, integrating nuclear genomic data and conducting functional studies of positively selected genes (e.g., *psbH*, *clpP*) will be crucial to validate

their role in alpine adaptation and to fully unravel the genomic basis of survival in one of Earth's most extreme environments.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms27125195/s1>.

Author Contributions: Conceptualization, T.L. and X.H.; methodology, T.L. and X.H.; software, X.H. and J.J.; validation, T.L. and N.T.; formal analysis, X.H. and J.J.; investigation, Y.Z.; resources, L.G.; data curation, T.L. and L.G.; writing—original draft preparation, X.H. and Y.Z.; writing—review and editing, T.L. and N.T.; project administration, T.L.; funding acquisition, T.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Natural Science Foundation of Qinghai (2024-ZJ-993).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The chloroplast genomes of *E. scapiflorum* under study are deposited in the GenBank database under the following accession numbers: PX526263.1. The other sequences used in this study were downloaded from the NCBI.

Conflicts of Interest: The authors declare no competing interests.

Abbreviations

cp: Chloroplast; QTP: Qinghai–Tibet Plateau; LSC: Large Single-Copy region; SSC: Small Single-Copy region; IR: Inverted repeat; tRNA: Transfer RNA; rRNA: Ribosomal RNA; SSR: Simple sequence repeat; Ka: Nonsynonymous Substitution Rate; Ks: Synonymous Substitution Rate; Ma: Mega-annum; ITS: Internal Transcribed Spacer; CTAB: Cetyltrimethylammonium bromide; OD: Optical Density; PCR: Polymerase Chain Reaction; P: Palindromic; C: Complementary; F: Forward; R: Reverse; Pi: Nucleotide Polymorphism; CDS: Protein-coding Sequence; ML: Maximum Likelihood; BI: Bayesian Inference; 95% HPD: 95% Highest Posterior Density; MCMC: Markov Chain Monte Carlo; ORF: Open Reading Frame; IGS: Intergenic Spacer; LCBs: Locally Collinear Blocks; SNP: Single Nucleotide Polymorphism; BS: Bootstrap Support; PP: Posterior Probability; UV: Ultraviolet Light.

References

- Nowak, M.D.; Birkeland, S.; Mandáková, T.; Choudhury, R.R.; Guo, X.; Gustafsson, A.L.S.; Gizaw, A.; Schröder-Nielsen, A.; Fracassetti, M.; Brysting, A.K.; et al. The genome of *Draba nivalis* shows signatures of adaptation to the extreme environmental stresses of the Arctic. *Mol. Ecol. Resour.* **2021**, *21*, 661–676. [\[CrossRef\]](#)
- Rellstab, C.; Zoller, S.; Sailer, C.; Tedder, A.; Gugerli, F.; Shimizu, K.K.; Holderegger, R.; Widmer, A.; Fischer, M.C. Genomic signatures of convergent adaptation to Alpine environments in three Brassicaceae species. *Mol. Ecol.* **2020**, *29*, 4350–4365. [\[CrossRef\]](#)
- Klimeš, L. Clonal splitters and integrators in harsh environments of the Trans-Himalaya. *Evol. Ecol.* **2008**, *22*, 351–367. [\[CrossRef\]](#)
- Li, R.; Fan, J.; Zhao, Q.; Li, T.; Wang, C.; Ding, R.; Gu, R.; Zhong, S. Estimation of habitat suitability and climatic distribution change of *Pegaeophyton scapiflorum* based on the MaxEnt model. *J. Nanjing For. Univ.* **2024**, *48*, 173–180.
- Yang, C.X.; Suo, N.J.; Longzhu, D.J. Research progress and prospect of *Eutrema scapiflorum*. *Qinghai Pratacult.* **2023**, *32*, 61–63.
- He, Y.J.; Tsering, L.M.; Zhang, Y.C.; Duan, Y.P.; Yuan, R.Y.; Wang, J.L. Study on anti-tussive and anti-inflammatory effects of aqueous extract of *Pegaeophyti Radix Et Rhizoma*. *Chin. Pharm.* **2015**, *26*, 3512–3514.
- Yang, S.Y.; He, S.Q.; Li, Y.; Zhang, J.L.; Chen, Y.; Banma, C.R.; Gu, R. Comparative analysis of biological activities and phytochemical constituents between *Pegaeophyton scapiflorum* and *Solms-laubachia eurycarpa* revealed by UPLC-Q-Orbitrap HRMS-based metabolomics. *Arab. J. Chem.* **2025**, *18*, 4752025. [\[CrossRef\]](#)
- Al-Shehbaz, I.A.; Arai, K.; Ohba, H. A revision of *Pegaeophyton* (Brassicaceae). *Edinb. J. Bot.* **2000**, *57*, 157–170. [\[CrossRef\]](#)
- Hendriks, K.P.; Kiefer, C.; Al-Shehbaz, I.A.; Bailey, C.D.; van Huysduynen, A.H.; Nikolov, L.A.; Nauheimer, L.; Zuntini, A.R.; German, D.A.; Franzke, A.; et al. Global Brassicaceae phylogeny based on filtering of 1,000-gene dataset. *Curr. Biol.* **2023**, *33*, 4052–4068. [\[CrossRef\]](#) [\[PubMed\]](#)

10. Koch, M.; Haubold, B.; Mitchell-Olds, T. Molecular systematics of the Brassicaceae: Evidence from coding plastidic *matK* and nuclear *Chs* sequences. *Am. J. Bot.* **2001**, *88*, 534–544. [\[CrossRef\]](#)
11. Daniell, H.; Lin, C.S.; Yu, M.; Chang, W.J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [\[CrossRef\]](#)
12. Wicke, S.; Müller, K.F.; de Pamphilis, C.W.; Quandt, D.; Wickett, N.J.; Zhang, Y.; Renner, S.S.; Schneeweiss, G.M. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell* **2013**, *25*, 3711–3725. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Streb, P.; Shang, W.; Feierabend, J.; Bligny, R. Divergent strategies of photoprotection in high-mountain plants. *Planta* **1998**, *207*, 313–324. [\[CrossRef\]](#)
14. Zhao, S.; Gao, X.; Yu, X.; Yuan, T.; Zhang, G.; Liu, C.; Li, X.; Wei, P.; Li, X.; Liu, X. Comparative analysis of chloroplast genome of *Meconopsis* (Papaveraceae) provides insights into their genomic evolution and adaptation to high elevation. *Int. J. Mol. Sci.* **2024**, *25*, 2193. [\[CrossRef\]](#)
15. Warwick, S.I.; Al-Shehbaz, I.A.; Price, R.A.; Sauder, C. Phylogeny of *Sisymbrium* (Brassicaceae) based on ITS sequences of nuclear ribosomal DNA. *Can. J. Bot.* **2002**, *80*, 1002–1017. [\[CrossRef\]](#)
16. Zhai, D.; Zhang, L.Y.; Li, L.Z.; Xu, Z.G.; Liu, X.L.; Shang, G.D.; Zhao, B.; Gao, J.; Wang, F.-X.; Wang, J.-W. Reciprocal conversion between annual and polycarpic perennial flowering behavior in the Brassicaceae. *Cell* **2024**, *187*, 3319–3337. [\[CrossRef\]](#)
17. Guo, X.; Liu, J.; Hao, G.; Zhang, L.; Mao, K.; Wang, X.; Zhang, D.; Ma, T.; Hu, Q.; Al-Shehbaz, I.A.; et al. Plastome phylogeny and early diversification of Brassicaceae. *BMC Genom.* **2017**, *18*, 176. [\[CrossRef\]](#)
18. Yang, S.; Wang, Z.Z. The complete chloroplast genome sequence of the medicinal and economic plant woad *Isatis indigotica* (Brassicaceae). *Mitochondrial DNA B* **2017**, *2*, 514–515. [\[CrossRef\]](#)
19. Zhang, Z.; Tao, M.; Shan, X.; Pan, Y.; Sun, C.; Song, L.; Pei, X.; Jing, Z.; Dai, Z. Characterization of the complete chloroplast genome of *Brassica oleracea* var. *italica* and phylogenetic relationships in Brassicaceae. *PLoS ONE* **2022**, *17*, e0263310. [\[CrossRef\]](#)
20. Liu, T.T.; Xu, L.A.; Hu, Z.W.; Xiong, X.P.; An, X.; Cao, J.S. Characteristics of the chloroplast genome and phylogenetic analysis of flowering Chinese cabbage. *J. Zhejiang A&F Univ.* **2025**, *42*, 1–11.
21. Zhou, X.; Ren, H.; Zhang, J.; Xu, D.; Xiao, W.; Huang, H.; Li, G.; Zhang, H.; Zheng, Y. The complete chloroplast genome of *Brassica rapa* var. *purpuraria* (LH Bailey) Kitam 1950 and its phylogenetic analysis. *Mitochondrial DNA B* **2024**, *9*, 143–147. [\[CrossRef\]](#)
22. Li, H.; Guo, Q.; Xu, L.; Gao, H.; Liu, L.; Zhou, X. CPJSDraw: Analysis and visualization of junction sites of chloroplast genomes. *PeerJ* **2023**, *11*, e15326. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Wang, H.; Zhang, Y.; Zhang, L.; Wang, J.; Guo, H.; Zong, J.; Chen, J.; Li, D.; Li, L.; Liu, J.; et al. Molecular characterization and phylogenetic analysis of Centipedegrass [*Eremochloa ophiuroides* (Munro) Hack.] based on the complete chloroplast genome sequence. *Curr. Issues Mol. Biol.* **2024**, *46*, 1635–1650. [\[CrossRef\]](#)
24. Li, J.W.; Li, R.Y.; Chen, Y.M.; Wu, Y.H.; Zou, L.H.; Tang, S.L.; Zhai, J.-W. Comprehensive characterization and phylogenetic analysis of the complete plastomes of two ant-orchids, *Caularthron bicornutum* and *Myrmecophila thomsoniana*. *BMC Plant Biol.* **2024**, *24*, 1146. [\[CrossRef\]](#)
25. Wu, C.S.; Wang, Y.N.; Hsu, C.Y.; Lin, C.P.; Chaw, S.M. Loss of different inverted repeat copies from the chloroplast genomes of Pinaceae and Cupressophytes and influence of heterotachy on the evaluation of Gymnosperm phylogeny. *Genome Biol. Evol.* **2011**, *3*, 1284–1295. [\[CrossRef\]](#)
26. Zhu, B.; Qian, F.; Hou, Y.; Yang, W.; Cai, M.; Wu, X. Complete chloroplast genome features and phylogenetic analysis of *Eruca sativa* (Brassicaceae). *PLoS ONE* **2021**, *16*, e0248556. [\[CrossRef\]](#)
27. Erixon, P.; Oxelman, B. Whole-gene positive selection, elevated synonymous substitution rates, duplication, and indel evolution of the chloroplast *clpP1* gene. *PLoS ONE* **2008**, *3*, e1386. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Zhang, S.Y.; Yan, H.F.; Wei, L.; Liu, T.J.; Chen, L.; Hao, G.; Wu, X.; Zhang, Q.-L. Plastid genome and its phylogenetic implications of Asiatic *Spiraea* (Rosaceae). *BMC Plant Biol.* **2024**, *24*, 23. [\[CrossRef\]](#)
29. Guo, X.Y.; Hu, Q.J.; Hao, G.Q.; Wang, X.J.; Zhang, D.; Ma, T.; Liu, J. The genomes of two *Eutrema* species provide insight into plant adaptation to high altitudes. *DNA Res.* **2018**, *25*, 307–315. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Robert, K.; Koch, M.A. A world-wide perspective on crucifer speciation and evolution: Phylogenetics, biogeography and trait evolution in tribe Arabideae. *Ann. Bot.* **2013**, *6*, 983–1001. [\[CrossRef\]](#)
31. Nikolov, L.A.; Shushkov, P.; Nevado, B.; Gan, X.; Al-Shehbaz, I.A.; Filatov, D.; Bailey, C.D.; Tsiantis, M. Resolving the backbone of the Brassicaceae phylogeny for investigating trait diversity. *New Phytol.* **2019**, *222*, 1638–1651. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
33. Chen, S.F.; Zhou, Y.Q.; Chen, Y.R.; Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **2018**, *34*, i884–i890. [\[CrossRef\]](#)
34. Jin, J.J.; Yu, W.B.; Yang, J.B.; Song, Y.; dePamphilis, C.W.; Yi, T.S.; Li, D.-Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **2020**, *21*, 241. [\[CrossRef\]](#)

35. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq-versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [[CrossRef](#)]
36. Shi, L.C.; Chen, H.M.; Jiang, M.; Wang, L.Q.; Wu, X.; Huang, L.F.; Liu, C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* **2019**, *47*, W65–W73. [[CrossRef](#)]
37. Greiner, S.; Lehwark, P.; Bock, R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **2019**, *47*, W59–W64. [[CrossRef](#)] [[PubMed](#)]
38. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)]
39. Thiel, T.; Michalek, W.; Varshney, R.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [[CrossRef](#)] [[PubMed](#)]
40. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
41. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R.G. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)]
42. Darling, A.C.E.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394–1403. [[CrossRef](#)]
43. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [[CrossRef](#)]
44. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)] [[PubMed](#)]
45. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [[CrossRef](#)]
46. Zhang, Z.; Xiao, J.; Wu, J.; Zhang, H.; Liu, G.; Wang, X.; Dai, L. ParaAT: A parallel tool for constructing multiple protein-coding DNA alignments. *Biochem. Biophys. Res. Commun.* **2012**, *419*, 779–781. [[CrossRef](#)]
47. Zhang, Z. KaKs_Calculator 3.0: Calculating selective pressure on coding and non-coding sequences. *Genom. Proteom. Bioinform.* **2022**, *20*, 536–540. [[CrossRef](#)] [[PubMed](#)]
48. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [[CrossRef](#)]
49. Katoh, K.; Misawa, K.; Kuma, K.I.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)] [[PubMed](#)]
50. Capella, G.S.; Silla, M.M.; Gabaldon, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [[CrossRef](#)]
51. Nguyen, L.T.; Schmidt, H.A.; Haeseler, A.V.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)] [[PubMed](#)]
52. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [[CrossRef](#)] [[PubMed](#)]
53. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior summarisation in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)] [[PubMed](#)]
54. Kumar, S.; Suleski, M.; Craig, J.M.; Kasprowicz, A.E.; Sanderford, M.; Li, M.; Stecher, G.; Hedges, S.B. TimeTree 5: An expanded resource for species divergence times. *Mol. Biol. Evol.* **2022**, *39*, msac174. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.