The First Complete mitochondrial genome of Neopicrorhiza scrophulariiflora: insights into structure, Codon usage, repeats, and RNA editing

Feiya Zhao
Lijiang Culture and Tourism College
Jing Li
Kunming University
Chunfan Xiang
Lijiang Culture and Tourism College
Yunjuan Li
Lijiang Agricultural Science Research Institute
Canxing Wang
Lijiang Agricultural Science Research Institute
Jialong Song
Lijiang Agricultural Science Research Institute
Shi Peng
Lijiang Agricultural Science Research Institute
Aien Tao
2515073996@qq.com
Lijiang Culture and Tourism College
Yan Zhao
Yunnan Agricultural University

Research Article

Keywords: Neopicrorhiza scrophulariiflora, Mitochondrial genome, RNA editing, Homologous fragments, Phylogeny

Posted Date: August 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10253361/v1

Additional Declarations: No competing interests reported.

The First Complete mitochondrial genome of *Neopicrorhiza scrophulariiflora*: insights into structure, Codon usage, repeats, and RNA editing

Feiya Zhao[1], Jing Li[1,2], Chunfan Xiang[1], Yunjuan Li[3], Canxing Wang[3], Jialong Song[3], Shi Peng[3*], Aien Tao[1*], Yan Zhao[4*]

1 College of Medicine, Lijiang Culture and Tourism College, Lijiang 674100, China

2 School of Medicine, Kunming University, Kunming, 650214, China

3 Lijiang Agricultural Science Research Institute, Lijiang 674100, China

4 Yunnan Agricultural University，Kunming 650201, China

*Correspondence:

Shi Peng: pengshinks@163.com

Aien Tao: 2515073996@qq.com

Yan Zhao: zhaoyankm@126.com

Abstract

Background *Neopicrorhiza scrophulariiflora* (Pennell) D. Y. Hong, a traditional medicinal plant adapted to extreme alpine environments, has a long history of use in Tibetan and Chinese medicine. The species was recognized as an endangered (EN) species in the IUCN Red List of Threatened Species in 1987. Despite its ecological and medicinal importance, its mitochondrial genome remains uncharacterized, limiting insights into its evolutionary adaptations and genomic architecture.

Results In this study, we assembled and comprehensively analyzed the complete mitochondrial genome of *N. scrophulariiflora*. The genome spans 606,216 bp with a

GC content of 44.31% and exhibits a multi-chromosomal architecture comprising four circular chromosomes, encompassing 56 genes, including 34 protein-coding genes (PCGs), 3 rRNA genes, and 17 tRNA genes. Analysis of codon usage bias revealed a preference for A/U-ending synonymous codons, with 23 of the 30 preferentially used codons terminating in A or U, aligning with trends observed in other angiosperms. Repetitive sequence analysis identified 66 SSRs, 11 tandem repeats, and 557 dispersed repeat pairs (284 forward and 273 palindromic repeats), suggesting their involvement in genomic rearrangement and structural diversification. A total of 492 high-confidence RNA editing sites (score ≥ 0.9) were predicted across all PCGs, all of which were C-to-U transitions, with nad5 harboring the highest number of editing sites (102) and rps14 the fewest (2). Comparative genomic analysis uncovered 49 chloroplast-derived homologous fragments (MTPTs) totaling 37,452 bp, accounting for approximately 6.18% of the mitogenome, providing evidence of frequent inter-organellar DNA transfer. Phylogenetic reconstruction based on 32 conserved mitochondrial PCGs across 28 taxa placed *N. scrophulariiflora* within Plantaginaceae, clustering with Digitalis ferruginea with moderate bootstrap support (65%). Synteny analysis with 8 closely related species revealed extensive structural rearrangements, with homologous regions covering 34.57% of the mitogenome with PX523758.1. Additionally, Ka/Ks analysis of 33 shared PCGs indicated that most mitochondrial genes (24 genes) are under strong purifying selection, with 1 gene (*sdh3*) exhibiting a median Ka/Ks value ≥ 1, suggesting possible positive selection linked to high-altitude adaptation, while *rps14* showed particularly strong purifying selection reflecting its essential role in mitochondrial translation.

Conclusions This study presents the first complete mitochondrial genome of *N. scrophulariiflora*, an endangered medicinal plant endemic to the Himalayan region, and offers a thorough characterization of its genomic architecture. Our findings expand the mitochondrial genomic resources for Plantaginaceae and elucidate the structural features, codon usage patterns, repetitive elements, RNA editing landscape, and inter-organellar gene transfer events of *N. scrophulariiflora*. Additionally, we establish its phylogenetic position within the family. The results confirm the placement of *N. scrophulariiflora* in Plantaginaceae, consistent with the APG IV system, and identify candidate genes under positive selection that may be linked to high-altitude adaptation. This work provides a valuable genomic foundation for future conservation genetics, evolutionary studies, and the sustainable utilization of this endangered medicinal species.

Keywords: *Neopicrorhiza scrophulariiflora*, Mitochondrial genome, RNA editing, Homologous fragments, Phylogeny

Background

Neopicrorhiza scrophulariiflora (Pennell) D.Y.Hong *(N. scrophulariiflora)* is a perennial medicinal plant within the genus *Neopicrorhiza* of the family Plantaginaceae. Its distribution encompasses southern Tibet, northwestern Yunnan, western Sichuan in China, and Nepal. This species is congeneric with *Neopicrorhiza kurroa* Royle ex. Benth. (*N. kurroa*), representing the only two species within this genus. Due to overharvesting and the absence of organized cultivation in recent years, the habitat range of *N. scrophulariiflora* has become increasingly restricted, leading to a significant

decline in its population. Consequently, this species has been classified as endangered by the IUCN[1]. *N. scrophulariiflora* has a medicinal history that spans thousands of years, with its therapeutic components increasing as the plant matures[2]. The rhizome of *N. scrophulariiflora* is utilized in medicine, demonstrating hepatoprotective, antioxidant, anti-allergic, and anti-asthmatic effects. Traditionally, the rhizome is boiled in water to alleviate colds and fevers. Recognized as a legitimate Chinese herbal medicine, it is included in the Pharmacopoeia of China (2020). The pharmaceutical industry incorporates this plant into 22 types of Chinese patent medicines[3].

Mitochondria are pivotal for oxidative metabolism and ATP production, while also mediating signaling, cell division, and apoptosis. Derived from endosymbiotic prokaryotes, plant mitochondrial genomes are maternally inherited, a feature that simplifies genetic analyses and makes them valuable for phylogenetics[4]. Since the first land plant mtDNA was sequenced, mitogenomes from numerous angiosperms have been characterized[5]. Unlike chloroplast genomes, plant mtDNAs exhibit intricate structures (linear, cyclic, or branched) and extensive size variation (200 kb to >3 Mb) even within genera, driven by foreign DNA insertions and repetitive sequences. Despite this variability, gene content is relatively conserved, encompassing core respiratory genes. However, plant mtDNAs are marked by sparse gene distribution, abundant noncoding DNA, pervasive repeats, and extensive RNA editing, which collectively cause structural instability and pose substantial assembly challenges[6]. Consequently, mtDNA sequencing has lagged behind that of chloroplasts, but recent advances in long-read technologies and assembly tools are rapidly expanding the repertoire of plant

mitogenomes, promising deeper insights into their evolution and functional biology[5].

Understanding the genetic architecture of *N. scrophulariiflora* is crucial for its conservation. Previous studies utilizing ISSR markers have demonstrated that this species exhibits significant genetic diversity at the species level, with 100% polymorphic bands[7]. However, genetic diversity within populations is notably low, with a mean percentage of polymorphic bands (PPB) of 30.56%. This is coupled with high genetic differentiation among populations (Gst = 0.6955) and extremely limited gene flow (Nm = 0.2198). These genetic patterns, influenced by clonal propagation, genetic drift, and geographical isolation in alpine habitats, suggest that *N. scrophulariiflora* populations are highly fragmented and susceptible to local extinction[8]. Furthermore, overharvesting, driven by high trade demand and insufficient cultivation efforts, has positioned this species in the highest threat category among Himalayan medicinal plants[9]. Populations in unprotected areas are experiencing significant declines in both density and reproductive individuals. Mitochondrial genomes, characterized by maternal inheritance and relatively slow evolutionary rates, have been extensively used to trace population history, assess effective population sizes, and infer phylogeographic patterns in endangered species, thereby providing a valuable tool for informing conservation strategies[6, 10].

To date, no mitochondrial genome has been reported for the genus *Neopicrorhiza* or any species within the tribe Veroniceae, highlighting a significant gap in the genomic resources available for conservation-oriented research on this endangered lineage. Although the complete chloroplast genome of *N. scrophulariiflora* has been sequenced

and characterized, its mitochondrial genome remains unexplored[11]. Within Plantaginaceae, only the mitogenomes of two *Aragoa* species have been characterized; both possess a typical suite of genes comprising 34 proteins, 17 tRNAs, and three rRNAs, while no complete mitogenome sequence has been published for the tribe Veroniceae[12]. In this study, we assembled and annotated the first complete mitochondrial genome of *N. scrophulariiflora* by employing a combination of Illumina short-read and Nanopore long-read sequencing technologies. Our analyses included genome structure, gene content, codon usage bias, identification of repetitive elements, prediction of RNA editing sites, and comparative genomics with other Lamiales species. The results provide essential genomic data to support the conservation and sustainable management of this endangered medicinal plant and contribute to a broader understanding of mitogenome evolution within Plantaginaceae.

Materials and methods

Plant material sampling, DNA extraction, and sequencing

Fresh leaf material of *N. scrophulariiflora* was collected from Haihongyuan Traditional Chinese Medicine Co., Ltd., Ludian Township, Yulong County, Lijiang City, Yunnan Province, China (27° 11′ N, 99° 27′ E; altitude 2,392 m) (Fig.1). The plant sample was identified by Professor Tao Aien, and a voucher specimen has been deposited in the Herbarium of Medical College, Lijiang Culture and Tourism College under the accession number: 2401-LJCTC-HHL-0001. The samples were promptly frozen in liquid nitrogen for genomic DNA extraction using the CTAB method. Long-read sequencing libraries were prepared with the SQK-LSK109 kit and sequenced on

the Nanopore PromethION platform (Nanopore Technologies, Wilmington, DE, USA). Data processing employed NanoFilt and NanoPlot from NanoPack. Simultaneously, high-quality DNA libraries were prepared using the Nextera XT DNA Library Prep Kit (Illumina, San Diego, CA, USA), and sequenced on the Illumina NovaSeq 6000 system, yielding **19.75** GB of raw data. Subsequent quality control analysis with the NGS QCToolkit v2.3.3 produced **19.64** GB of clean data with a GC content of **39.65**%, achieving Q20 and Q30 scores of **99.32**% and **97.24**%, respectively.

Genome assembly and annotation

The Nanopore long reads were assembled *de novo* using miniasm (v0.3-r179) following an all-vs-all alignment conducted with minimap2 (v2.15-r905). The resulting assembly underwent further polishing with nextPolish (v1.3.1) to enhance accuracy. Subsequently, second-generation sequencing reads were aligned to the polished assembly using Bowtie2 (v2.3.5.1). The Illumina short reads were integrated into the draft contigs via Unicycler (v0.4.8), and the assembly graph was visualized using Bandage. To finalize the assembly, Nanopore long reads were realigned to the contig sequences with minimap2, which facilitated the incorporation of connectivity information from Bandage, validation of inter-contig linkages, and manual curation to resolve ambiguities and complete the mitogenome structure. This iterative process enabled the successful reconstruction of the complete mitogenome of *N. scrophulariiflora*. Annotation of the mitogenome was conducted using MFANNOT and MITOFY[13]. The circular map of the mitogenome was generated with OGDRAW[14].

Codon usage analysis and RNA editing analysis

Codon usage bias is a prevalent characteristic of genomes, reflecting long-term evolutionary pressures that balance mutation biases with translational efficiency. To assess this bias, the protein-coding sequences were extracted using Phylosuite v1.1.16[15], and relative synonymous codon usage (RSCU) values were calculated with MEGA (v7.0 https://www.megasoftware.net/) software. Furthermore, Deepred-Mt software, integrated within the MGA workflow, was utilized to predict C-to-U RNA editing sites in the protein-coding genes of the mitogenome, retaining predictions with a score greater than 0.9 for subsequent analysis.

Ka/Ks analysis

The selective pressure on mitochondrial genes was assessed by calculating the ratio of non-synonymous (Ka) to synonymous (Ks) substitution rates for 20 protein-coding genes shared across representative species of Lamiales. Ka/Ks values were computed using KaKs_Calculator (v2.0) and visualized as box plots with Origin 2021 software[16].

Repeat sequence analysis

The MISA (https://webblast.ipk-gatersleben.de/misa/) system identified simple repetitive sequences (SSRs) in the *N. scrophulariiflora* mitogenome. The minimum repeat lengths were defined as 10 for mononucleotides, 5 for dinucleotides, 4 for trinucleotides, 3 for tetranucleotides, and 3 for pentanucleotides. Tandem repeat analysis was conducted using the Tandem Repeats Finder (TRF) v4.09 (https://tandem.bu.edu/trf/trf.advanced.submit.html) with the parameters set as follows: match = 2, mismatch = 7, insertion/deletion (indel) = 7, minimum alignment

score = 50, and maximum cycle length = 500. The REPuter web server was utilized for the analysis of random element (RE) sequences.

Homologous sequence analysis between chloroplast and mitochondria

The chloroplast genome of *N. scrophulariiflora* was analyzed using the BLASTN program[17] with default parameters. Furthermore, the amino acid sequences obtained from the plant's mitogenome were compared to those of the chloroplast genome to identify potential sequence similarities. The resulting homologous regions were visualized using Circos software[18]. To explore evolutionary relationships, Mauve (version 2.3.1) was utilized to align sequencing reads against reference mitogenomes from related species. Collinearity block lengths were determined based on BLAST-derived alignments, facilitating a more detailed examination of genomic conservation and divergence among species. Additionally, the analysis of gene transfer events from chloroplasts to mitochondria was performed using TBtools (v2.056)[19].

Phylogenetic and collinearity analysis

Pairwise BLASTN comparisons were performed as a preliminary step to identify homologous regions among mitochondrial genomes, followed by MCscanX-based synteny analysis to assess structural conservation across related species[20]. Syntenic blocks between *N. scrophulariiflora* and other Plantaginaceae (or related Lamiales) species were identified using MCscanX software[21], and a comparative syntenic map was generated. The identified syntenic blocks were visualized using the Circos package in R v4.2. For phylogenetic reconstruction, amino acid sequences of conserved mitochondrial PCGs shared by Lamiales species (including Solanaceae, Gesneriaceae,

Phrymaceae, Lamiaceae, Oleaceae, Orobanchaceae, Lentibulariaceae, Plantaginaceae, and Linderniaceae) were aligned using MAFFT v7.505[22], and poorly aligned regions were removed using TrimAl (v1.2rev57)[23]. A maximum likelihood (ML) tree was constructed using IQ-TREE v2.2.0[24], with ModelFinder selecting the best-fit substitution model and 1,000 ultrafast bootstrap replicates. The resulting phylogenetic tree was visualized using iTOL (https://itol.embl.de/). The analysis included the newly assembled mitogenome of *N. scrophulariiflora* (designated as HHL) and publicly available mitogenomes from representative species across the aforementioned families, with *Nicotiana tabacum* (NC_006581) as the outgroup to resolve evolutionary relationships within Lamiales.

Results

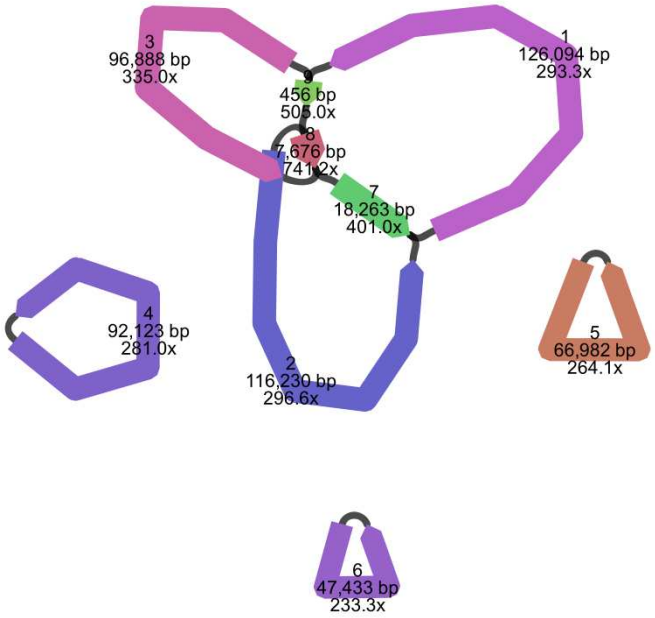
Structural features and gene composition of the *N. scrophulariiflora* mitogenome

We assembled the complete mitochondrial genome of *N. scrophulariiflora*, which exhibits a multi-chromosomal architecture comprising four circular chromosomes (Fig.2). The genome spans 520,966 bp with a GC content of 44.40%. The genome spans 606,216 bp with a GC content of 44.31%. The four chromosomes are shown as gene maps in Fig.3. Genome annotation revealed 56 genes, including 34 protein-coding genes (PCGs), three rRNA genes, and 17 tRNA genes (Table 1).

Based on functional classification, the 34 PCGs comprise 5 ATP synthase genes (*atp1, atp4, atp6, atp8, atp9*), 9 NADH dehydrogenase genes (*nad1, nad2, nad3, nad4, nad4L, nad5, nad6, nad7, nad9*), 8 ribosomal protein genes (*rpl10, rpl5, rps10, rps12, rps13, rps14, rps3, rps4*), 2 succinate dehydrogenase genes (*sdh3, sdh4*), 1 cytochrome

bc1 complex gene (*cob*), 3 cytochrome c oxidase genes (*cox1, cox2, cox3*), 4 cytochrome c biogenesis genes (*ccmB, ccmC, ccmFc, ccmFn*), 1 maturase gene (*matR*), and 1 membrane transport protein gene (*mttB*). No unclassified hypothetical ORFs were identified in the annotation.

Fig.1 Morphology of *N.scrophulariiflora*. (Photo by the authors.)



228

229 **Fig.2** Assembly of the mitogenome of the *N. scrophulariiflora*, visualized using

230 Bandage software. The mt contigs are represented by different color segments.

231

232 **Table 1** Gene annotation of the *N. scrophulariiflora* mitochondrial genome

Group of genes	**Name of genes**
ATP synthase, Complex V	*atp1, atp4, atp6, atp8, atp9*
Cytochrome bc1 complex, Complex III	*cob*
Cytochrome *c* oxidase	*cox1, cox2, cox3*
NADH dehydrogenase, Complex I	*nad1, nad2, nad3, nad4, nad4L, nad5, nad6, nad7, nad9*
Ribosomal proteins	*rpl10, rpl5, rps10, rps12, rps13, rps14, rps3, rps4*
Succinate dehydrogenase, Complex II	*sdh3, sdh4*
RNA maturase	*matR*
Cytochrome c biogenesis	*ccmB, ccmC, ccmFc, ccmFn*
Transport membrane protein	*mttB*

233

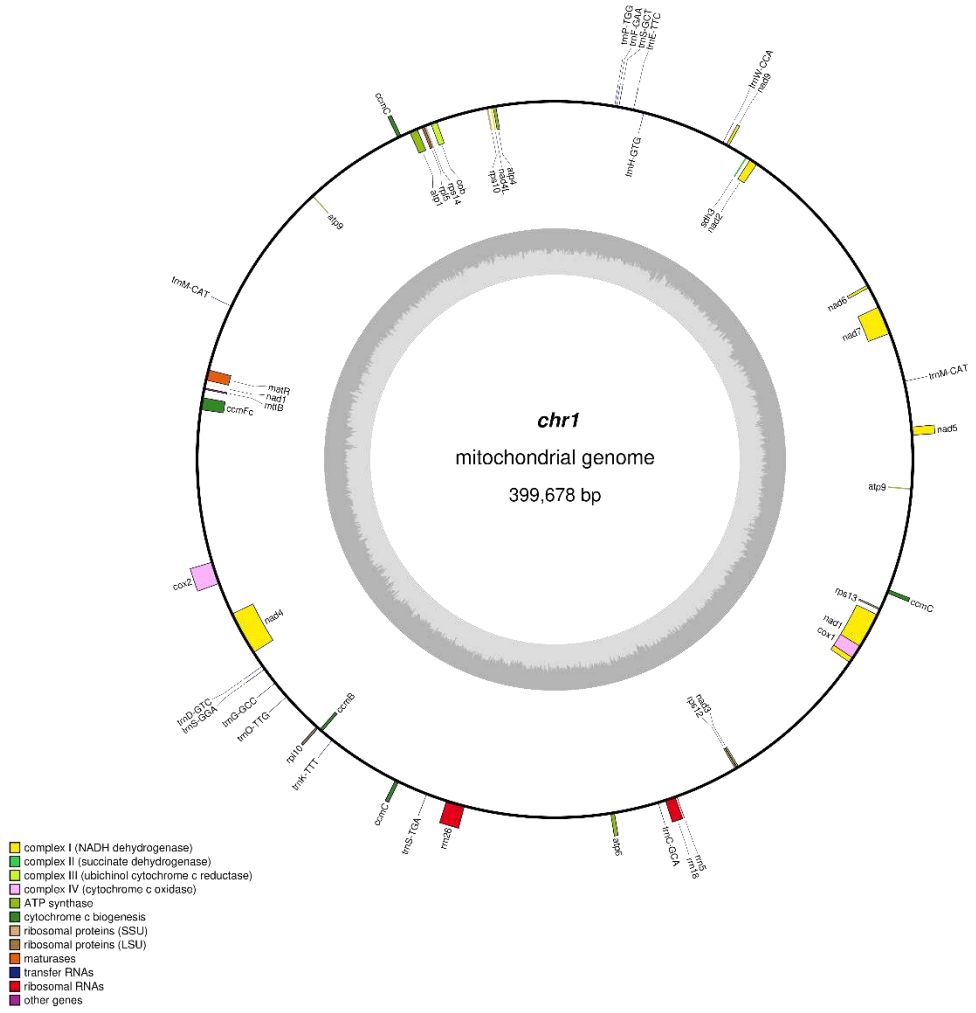


Fig.3 Circular gene map of Chromosome 1 (Ch1) of the *N. scrophulariiflora* mitogenome.

Codon usage bias analysis

Codon usage bias was evaluated for all 34 PCGs of the *N. scrophulariiflora* mitogenome. A total of 11,917 codons were identified, encoding 20 amino acids and termination signals. The codon TTT (Phe) was the most frequently used, occurring 509 times, while leucine (Leu) was the most abundant amino acid (1,321 occurrences), and cysteine (Cys) the least frequent (178 occurrences). RSCU analysis revealed that 30

codons exhibited RSCU values greater than 1, of which 27 terminate with A or U, indicating a strong preference for A/U-rich codons (Fig.4). Conversely, 32 codons displayed RSCU values less than 1, reflecting non-preferred usage. The codons ATG (Met) and TGG (Trp) showed no usage bias (RSCU ≈ 1).

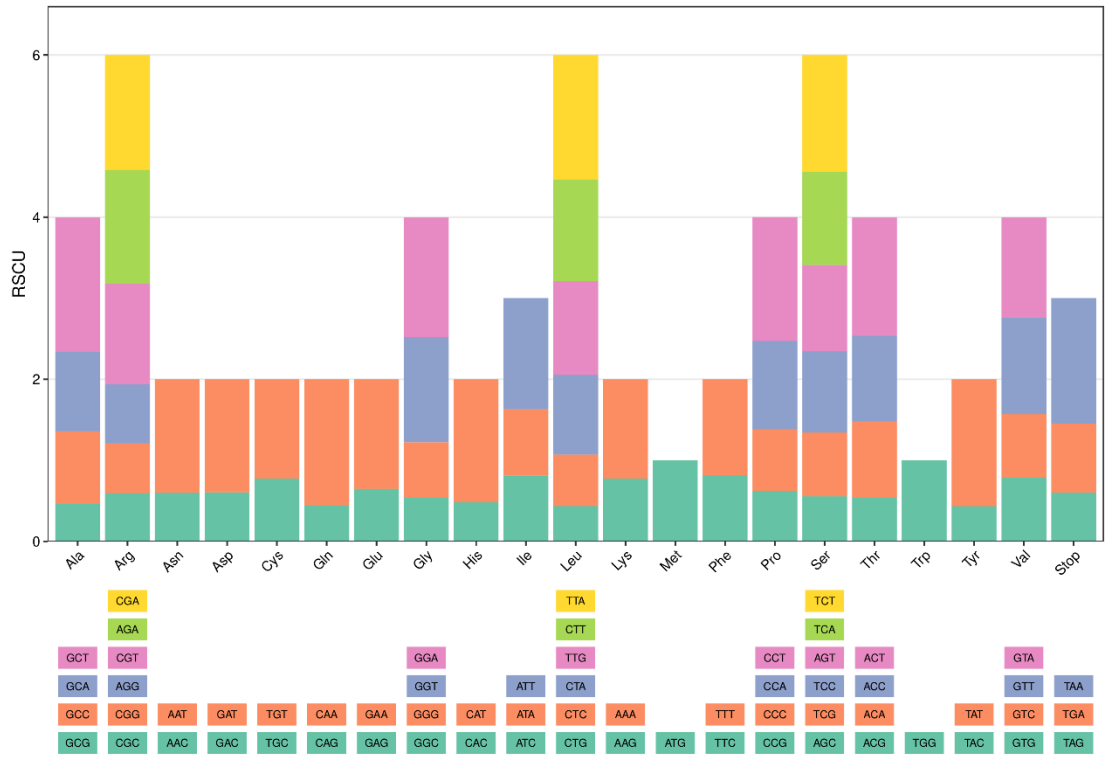


Fig.4 Relative synonymous codon usage (RSCU) analysis of PCGs in the *N. scrophulariiflora* mitogenome. Codon families are shown on the x-axis, and RSCU values on the y-axis. Codons encoding the same amino acid are distinguished by different colours. RSCU values > 1 indicate preferential codon usage.

RNA editing site prediction

RNA editing is a crucial post-transcriptional modification that ensures proper gene expression in plant mitochondria. We predicted RNA editing sites across all protein-coding genes of the *N. scrophulariiflora* mitogenome using Deepred-Mt and

retained high-confidence predictions with a score ≥ 0.9. A total of 492 high-confidence potential RNA editing sites were identified, with C-to-U as the predominant editing type, consistent with the pattern observed in most angiosperm mitogenomes (Fig.5). These editing events predominantly occurred at the first or second codon positions and are primarily involved in restoring protein sequence conservation to ensure proper protein function. Among all genes analysed, nad5 harboured the highest number of editing sites (102), while rps14 contained the fewest (2).

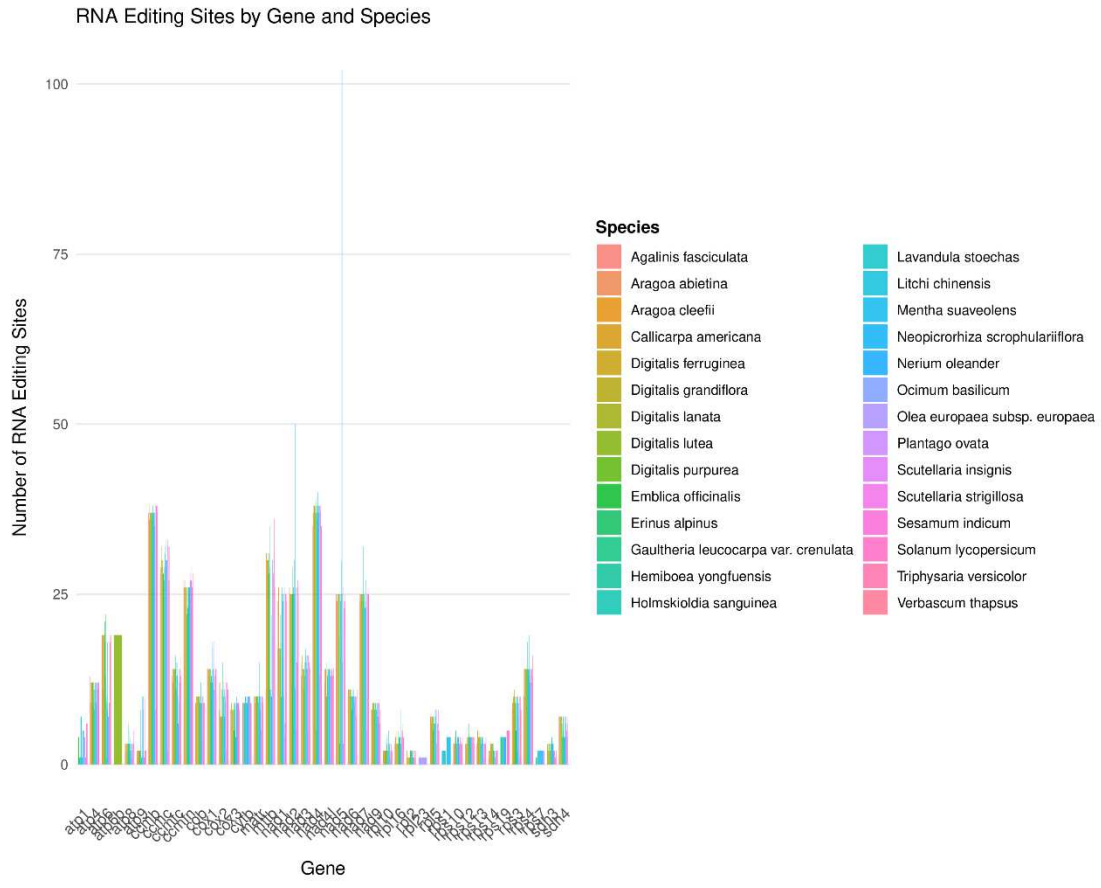


Fig.5 Distribution of RNA editing sites across mitochondrial PCGs in *N. scrophulariiflora* in comparison with related species. The x-axis represents different genes, and the y-axis represents the number of RNA editing sites. Different colours denote different species.

Ka/Ks analysis

To investigate the evolutionary dynamics of the mitochondrial genome in the *N. scrophulariiflora* lineage, we analysed the ratio of non-synonymous (Ka) to synonymous (Ks) substitution rates (ω = Ka/Ks) for 33 shared protein-coding genes for which valid estimates could be obtained (Fig.6). Overall, most genes exhibited strong evolutionary constraints. Based on the distribution of Ka/Ks values, the genes were classified into three categories. 24 genes (including *atp9, nad1, rps13, and rps12*) showed median Ka/Ks values considerably less than 1, indicating strong purifying selection and high functional conservation. These genes primarily encode core subunits of the respiratory chain complex and ATP synthase, and their sequence stability is critical for maintaining fundamental mitochondrial functions. 1 gene (*sdh3*) exhibited a median Ka/Ks value greater than or equal to 1, suggesting that this gene may have experienced positive selection during evolution, potentially associated with adaptive evolution to specific environmental conditions. The remaining 8 genes (including *atp4, ccmFc, ccmFn* and *mttB*) had Ka/Ks values close to the neutral threshold ($0.5 \leq \omega < 1$), indicating that they may be transitioning from purifying selection to adaptive evolution, or are subject to weaker selective pressures.

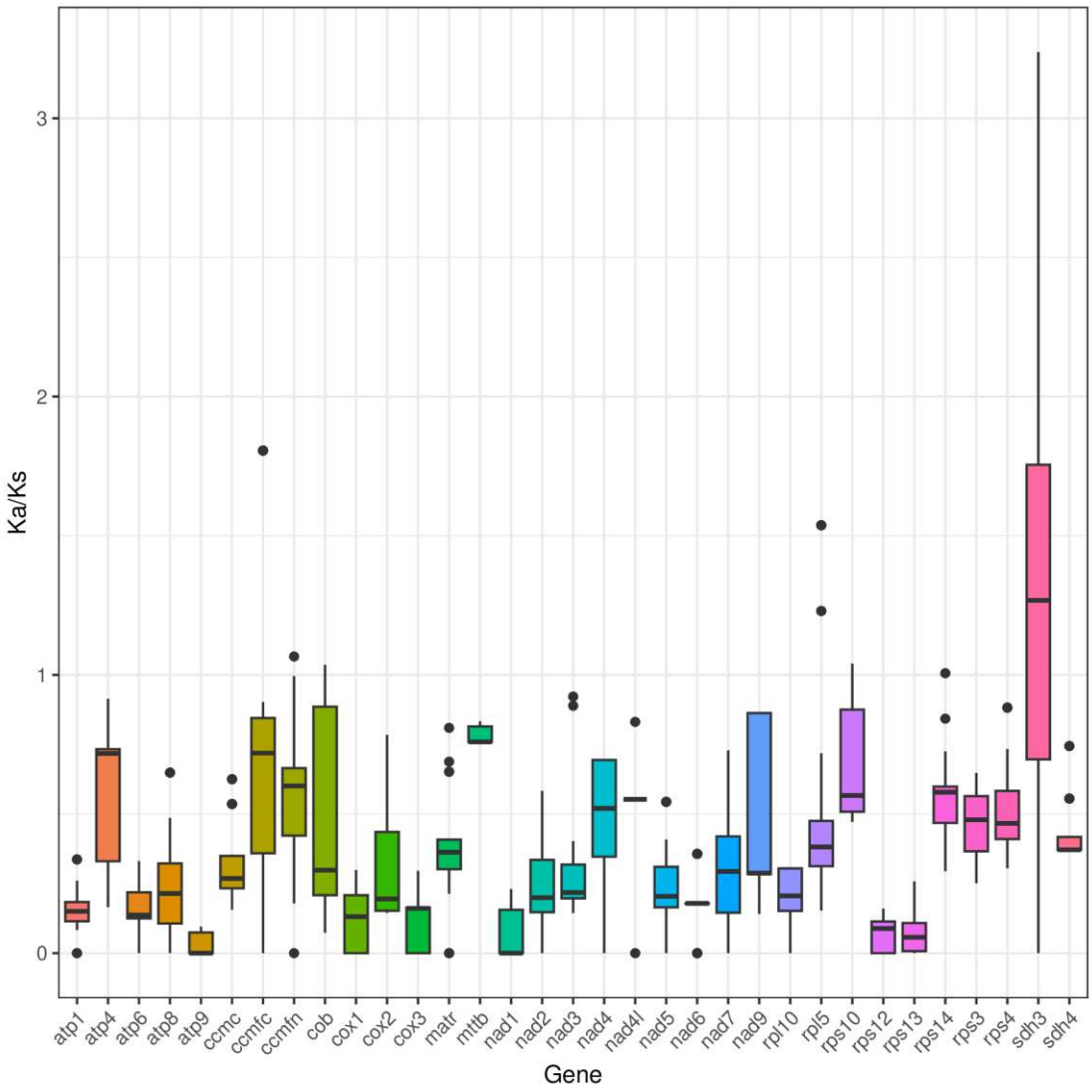


Fig.6 Box plots showing Ka/Ks ratios of 33 shared mitochondrial PCGs across representative species. The x-axis represents different genes, and the y-axis represents Ka/Ks values. The box plots display the distribution of Ka/Ks ratios, with median values indicated by horizontal lines. Genes with median Ka/Ks values considerably less than 1 indicate strong purifying selection, those with values greater than or equal to 1 suggest positive selection, and those between 0.5 and 1 indicate near-neutral evolution.

Repetitive sequence analysis

Repetitive sequences are key drivers of structural variation and genome evolution

297 in plant mitochondria. We therefore performed a comprehensive repeat analysis using
298 multiple complementary tools.

299 For SSRs, MISA detected 66 repeats with a total length of 761 bp. Mononucleotide
300 repeats were the most abundant (56 occurrences, 84.85%), followed by dinucleotides
301 (nine occurrences, 13.64%). A single trinucleotide repeat was identified, whereas no
302 tetra-, penta-, or hexanucleotide repeats were found (Fig.7). Among mononucleotide
303 SSRs, A/T motifs were predominant. Tandem repeats were identified using TRF,
304 revealing 11 loci with a combined length of 555 bp (Fig.8). For dispersed repeats,
305 Vmatch-based self-alignment identified 557 repeat pairs with a combined length of
306 118,111 bp, consisting of 284 forward repeats and 273 palindromic repeats, with lengths
307 ranging from 30 to 26,399 bp (Fig.9)

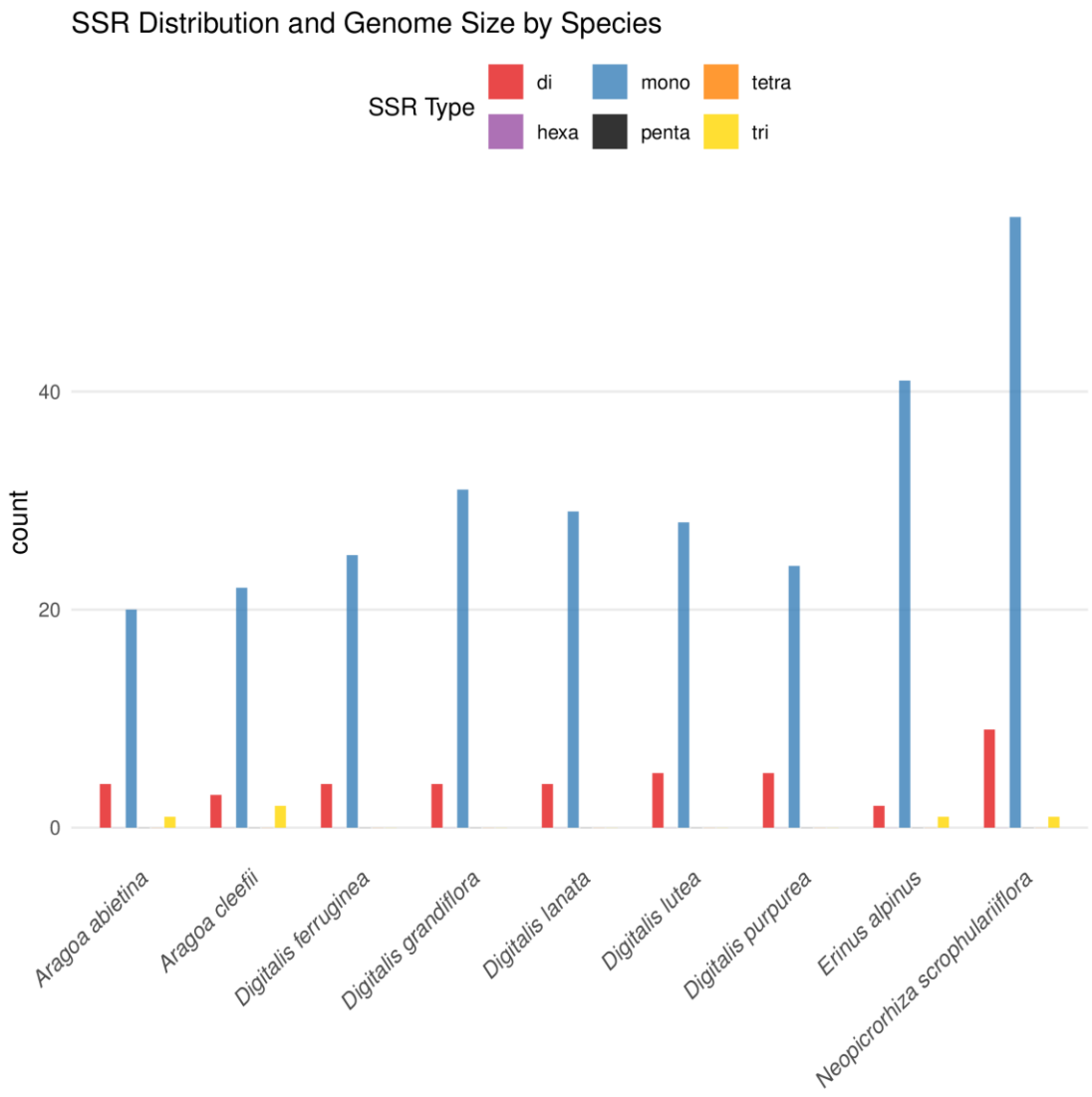


Fig.7 Types and frequencies of SSRs in the *N. scrophulariiflora* mitogenome in comparison with related species. The x-axis represents different species, and the y-axis represents the number of SSRs. Different colours indicate distinct SSR types (mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats).

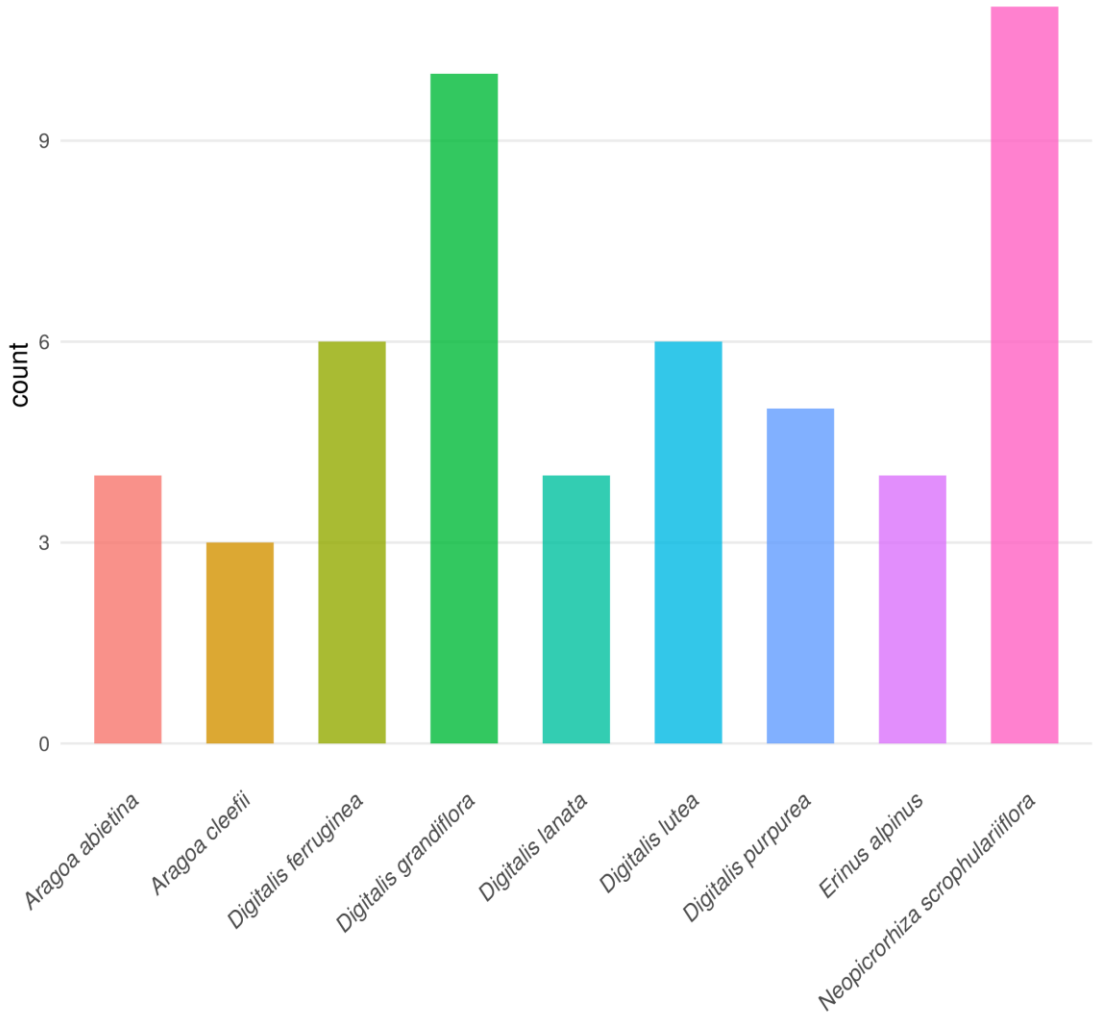


Fig.8 Tandem repeat counts in the *N. scrophulariiflora* mitogenome in comparison with related species. The x-axis represents different species, and the y-axis represents the number of tandem repeats.

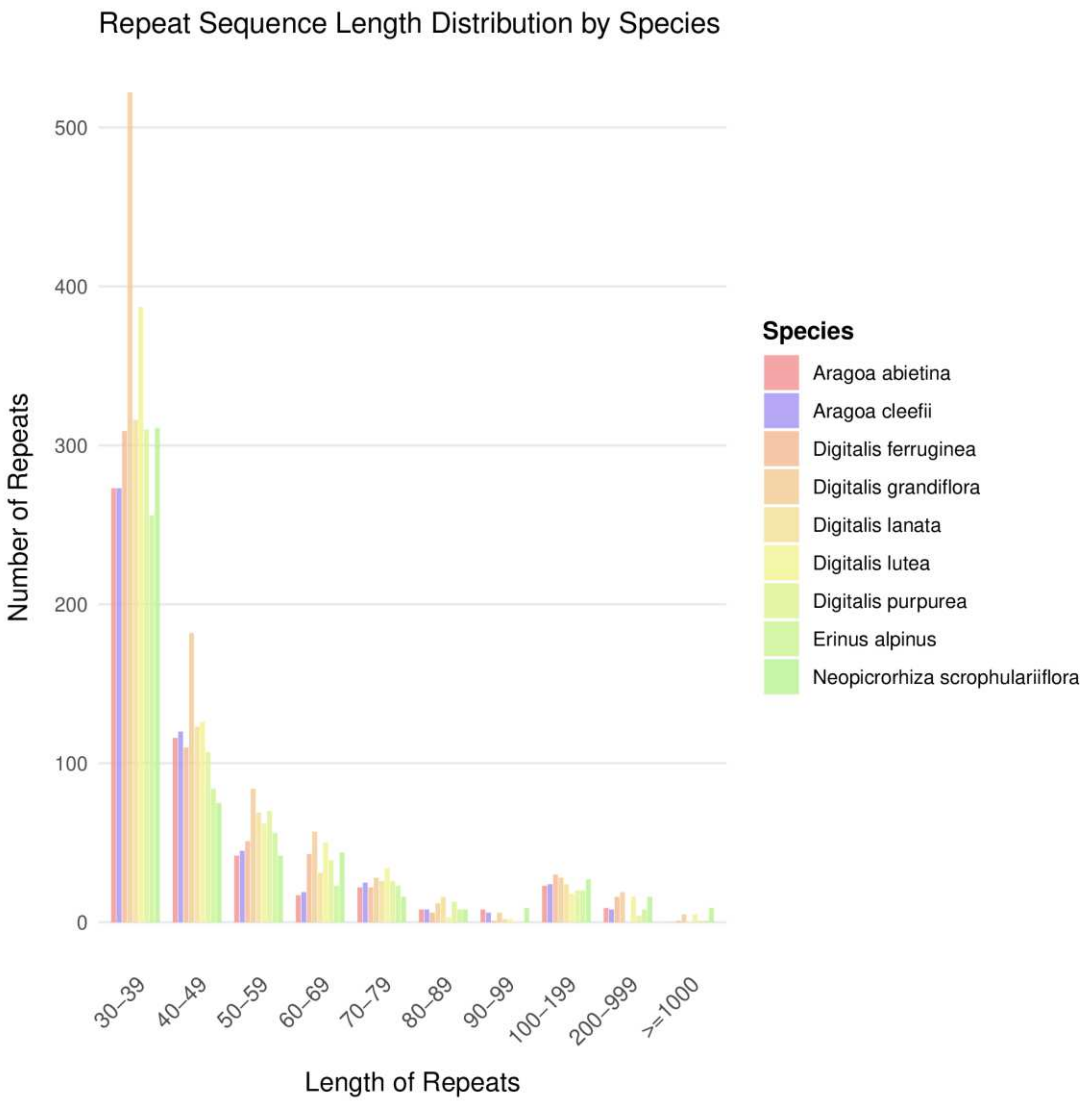


Fig.9 Length distribution of dispersed repeats in the *N. scrophulariiflora* mitogenome in comparison with related species. The x-axis represents repeat length ranges (bp), and the y-axis represents the number of repeats. Different colours denote different species.

MTPT analysis

Inter-organellar DNA transfer is a hallmark of plant genome evolution. To investigate mitochondrial plastid DNAs (MTPTs) in the *N. scrophulariiflora* mitogenome, we performed a detailed analysis of homologous fragments between the mitochondrial and chloroplast genomes. A total of 49 MTPT

fragments were identified, with a combined length of 37,452 bp, accounting for approximately 6.18% of the total mitogenome length (606,216 bp) (Fig.10). These fragments ranged from 42 bp to 3,600 bp in length, with sequence identity ranging from 73.82% to 100.0%. Annotation of these fragments revealed that a substantial number of chloroplast genes had been transferred—either completely or partially—to the mitogenome. In total, 29 distinct genes were involved, comprising 18 protein-coding genes, 9 tRNA genes, and 2 rRNA genes. These results indicate that extensive and frequent chloroplast-to-mitochondrion DNA transfer events have occurred during the evolution of this lineage.

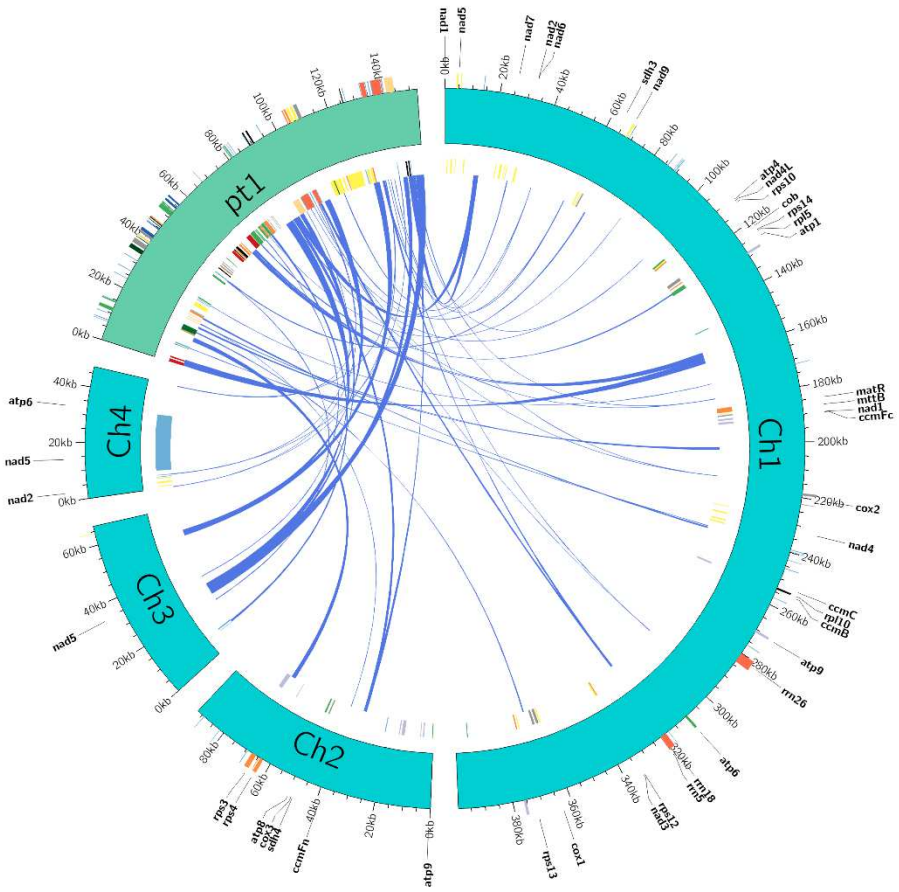


Fig.10 Homology analysis between the mitochondrial and chloroplast genomes of *N. scrophulariiflora*. The mitochondrial and chloroplast genomes are represented by blue and green arcs, respectively. Orange lines connecting the arcs represent homologous fragments (MTPTs) transferred from the chloroplast to the mitochondrion.

Synteny analysis

To further explore the structural evolution of the *N. scrophulariiflora* mitogenome and evaluate the integrity of our assembly, we performed comparative synteny analysis with 8 closely related species (OK514181.1, OK514182.1, PX523756.1, PX523757.1,

PX523758.1, PX523759.1, PX523760.1, PX523761.1) (Fig.11). Among the species compared, the mitogenome of *N. scrophulariiflora* exhibited the highest homology with PX523758.1 (Digitalis lanata). The homologous regions covered 34.57% of the *N. scrophulariiflora* mitogenome, with a non-overlapping total length of 209,594 bp (out of a total mitogenome size of 606,216 bp). This high coverage rate supports the completeness of our assembly and indicates a close phylogenetic relationship with NC_086689.1. Despite this close relationship, the synteny map revealed differences in collinear block order, reflecting structural rearrangements and highlighting the highly non-conservative nature of plant mitogenome structure.

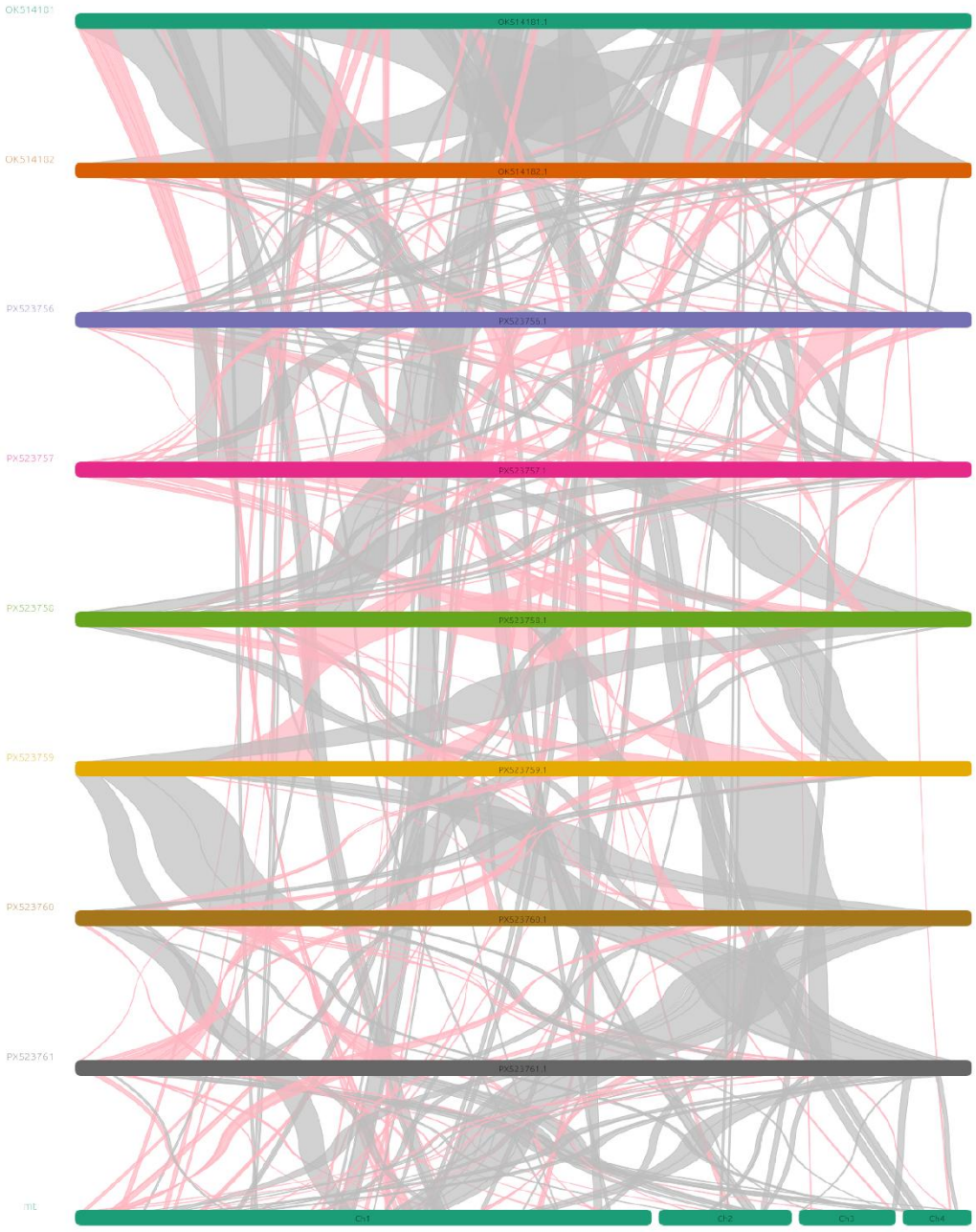


Fig.11 Mitochondrial genome synteny analysis of *N. scrophulariiflora* and 8 closely related species. Each coloured bar represents a mitochondrial genome, with the species name indicated on the left. Ribbons connecting adjacent bars represent homologous syntenic blocks. The y-axis represents the percentage of sequence coverage, and different colours denote different species.

Phylogenetic analysis

To determine the phylogenetic position of *N. scrophulariiflora*, we constructed a maximum likelihood (ML) tree using concatenated sequences of 32 core protein-coding genes shared by 28 taxa (Fig.12). The resulting topology showed that *N. scrophulariiflora* (labelled as HHL) clustered with Digitalis ferruginea (PX523756) within the family Plantaginaceae, with moderate bootstrap support (65%). The overall tree topology was highly congruent with taxonomic classifications, with species clustering into their respective families: Apocynaceae, Ericaceae, Gesneriaceae, Lamiaceae, Oleaceae, Orobanchaceae, Pedaliaceae, Phyllanthaceae, Plantaginaceae, Sapindaceae, and Scrophulariaceae. The placement of *N. scrophulariiflora* within Plantaginaceae is consistent with its taxonomic classification, and the overall topology agrees with the APG IV system.

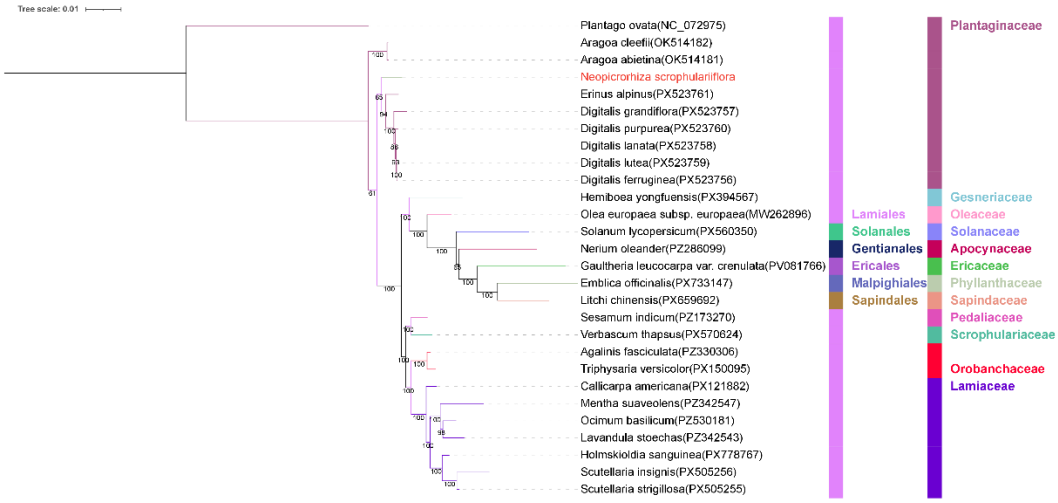


Fig. 12 Maximum likelihood phylogenetic tree based on mitogenomes of 28 species. Bootstrap support values are shown at the corresponding nodes. The red font indicates the species of the *N. scrophulariiflora*. Different colours represent different families.

Discussion

Plant mitochondrial genomes display substantial variation in size and structural organization across angiosperms, with sizes ranging from approximately 200 kb to over 3 Mb[25]. The *N. scrophulariiflora* mitogenome, at 606,216 bp, falls within the typical range observed in Lamiales, falls within the typical range observed in Lamiales, which generally spans from approximately 350 kb to over 600 kb[26]. The genome exhibits a multi-chromosomal architecture comprising four circular chromosomes, adding to the growing body of evidence that plant mitochondrial genomes exist in diverse conformational states beyond the canonical single circular molecule. This view has been increasingly supported by long-read sequencing technologies, which have revealed complex multi-chromosomal or branched structures in multiple plant lineages[27]. The multi-chromosomal architecture observed in *N. scrophulariiflora* aligns with this emerging paradigm and underscores the structural plasticity that characterizes plant mitogenomes. Despite this structural complexity, the number of protein-coding genes (34) is comparable to that reported in other angiosperm mitogenomes, and the presence of 17 tRNA genes and three rRNA genes reflects the functional conservation of core mitochondrial genes, reflecting the functional conservation of core mitochondrial genes across flowering plants[28]. The gene complement includes a complete set of core respiratory genes, consistent with the essential role of mitochondria in energy metabolism[29]. The GC content (44.31%) is similar to that of other Plantaginaceae species, reinforcing the notion that this genomic feature is evolutionarily conserved within the family, even as genome size and structure

exhibit considerable variation[30].

Analysis of codon usage bias in the 34 PCGs revealed a pronounced preference for A/U-terminating codons, with 27 of the 30 preferentially used codons ending in A or U. This bias is consistent with observations across diverse angiosperm lineages, including other Lamiales species, and is thought to reflect the combined effects of mutational pressure and translational selection[31]. The predominance of A/U-rich codons is a recurring theme in plant mitochondrial genomes, where the overall AT-rich base composition of non-coding regions exerts a strong influence on synonymous codon usage. Leucine was the most abundant amino acid, while cysteine was the least frequent, a pattern that has been consistently reported in plant mitogenomes and likely reflects the combined effects of codon availability and functional constraints on amino acid composition. The A/U bias observed in the preferentially used codons may have functional implications for translation efficiency, particularly under the low-temperature and hypoxic conditions characteristic of high-altitude environments, as A/U-rich codons tend to reduce mRNA secondary structure stability, potentially facilitating ribosome progression.[31]

Repetitive sequences are key drivers of structural variation and genome evolution in plant mitochondria[32], and the *N. scrophulariiflora* mitogenome harbors a substantial number of repetitive elements, including 66 SSRs, 11 tandem repeats, and 557 dispersed repeat pairs. Mononucleotide SSRs predominated, with A/T motifs being the most frequent, a pattern consistent with the overall AT-rich nature of plant mitochondrial genomes[33]. The dispersed repeats, comprising 284 forward repeats and

273 palindromic repeats with lengths up to 26,399 bp, suggest that recombination mediated by these repeats may have contributed to the structural complexity and multi-chromosomal architecture observed in this species. The absence of reverse and complementary repeats is notable and has been documented in other Lamiales mitogenomes, suggesting that forward and palindromic repeats are the primary drivers of recombination events in this lineage[34]. The abundance and distribution of repetitive elements in *N. scrophulariiflora* align with the broader understanding that repeat-mediated recombination is a major driver of mitogenome structural diversification, generating the genomic plasticity that characterizes plant mitochondrial genomes[35].

RNA editing is a hallmark of plant mitochondrial gene expression, predominantly involving C-to-U conversions that restore conserved amino acid sequences and ensure proper protein function[36]. In the *N. scrophulariiflora* mitogenome, we predicted 492 high-confidence potential RNA editing sites (score ≥ 0.9) across all PCGs. This number is within the range reported in other Lamiales species, although the editing frequency may be underestimated due to the stringent filtering threshold. This figure substantially exceeds the numbers reported in most other Lamiales species, suggesting that RNA editing may play a particularly prominent role in this species. The high editing frequency may reflect the extreme environmental pressures of its high-altitude habitat, where optimization of respiratory chain function is critical for survival[37]. All identified editing events were C-to-U transitions, consistent with the pattern observed across angiosperms, and the editing sites were predominantly located at the first or

second codon positions[38]. The gene *nad5* harbored the highest number of editing sites, while *rps14* contained the fewest. The substantial editing burden on respiratory genes such as *matR* and *ccmFn* suggests that RNA editing serves to optimize the function of proteins involved in energy metabolism, potentially as an adaptive mechanism to high-altitude stress[39].

Inter-organellar DNA transfer is a pervasive feature of plant genomic evolution, and the *N. scrophulariiflora* mitogenome contains 49 MTPT fragments with a combined length of 37,452 bp, accounting for approximately 6.18% of the total mitogenome length. This proportion falls within the range commonly observed among angiosperms[26]. The transferred fragments included 18 protein-coding genes, 9 tRNA genes, and 2 rRNA genes, indicating extensive and recurrent chloroplast-to-mitochondrion gene transfer. The enrichment of tRNA genes among the transferred sequences is noteworthy, as chloroplast-derived tRNAs may contribute to mitochondrial translation efficiency by expanding the tRNA repertoire available for protein synthesis[40]. The presence of these MTPT fragments highlights the dynamic interplay between organellar genomes and underscores the role of inter-organellar gene transfer in shaping mitochondrial genome complexity[41]. This phenomenon, driven by homologous recombination-mediated integration events, represents a continuous source of genomic novelty and structural expansion in plant mitochondrial genomes[42].

Phylogenetic analysis based on 32 conserved mitochondrial PCGs across 28 taxa placed *N. scrophulariiflora* within Plantaginaceae, consistent with its taxonomic

466 classification and the APG IV system[43]. The species clustered with Digitalis
467 ferruginea, with moderate bootstrap support (65%), providing greater resolution than
468 previous analyses based on fewer markers, suggesting that additional mitochondrial
469 markers or broader taxon sampling may be required to fully resolve the phylogenetic
470 position of *Neopicrorhiza* within Plantaginaceae[44]. The overall tree topology was
471 highly congruent with established family-level classifications, confirming the utility of
472 mitochondrial PCGs for resolving phylogenetic relationships at the family level within
473 Lamiales. Synteny analysis with 8 closely related species revealed that while *N.*
474 *scrophulariiflora* shares extensive homologous regions with its relatives (covering
475 34.57% of the mitogenome with PX523758.1), the arrangement of collinear blocks
476 differs substantially among species. This observation reflects the frequent genomic
477 rearrangements—including inversions and recombination events—that characterize
478 plant mitochondrial genome evolution and highlights the highly dynamic nature of
479 mitogenome structure even among closely related lineages[45]. The presence of
480 species-specific regions lacking homology further underscores the structural plasticity
481 of plant mitochondrial genomes and their capacity for rapid evolutionary change[46].

482 Analysis of Ka/Ks ratios for 33 shared PCGs revealed that the majority of genes
483 (24 genes) are evolving under strong purifying selection, consistent with the functional
484 conservation of core mitochondrial genes across angiosperms[47]. Genes such as *atp9,*
485 *nad1, rps13,* and *rps12* exhibited particularly strong purifying selection, reflecting its
486 essential role in mitochondrial translation[48]. 1 gene (*sdh3*) showed a median Ka/Ks
487 value ≥ 1, suggesting possible positive selection. These genes are primarily involved in

energy metabolism and cytochrome *c* biogenesis, and their accelerated evolution may be associated with adaptive responses to the extreme environmental conditions of the Qinghai-Tibet Plateau[49]. Genes exhibiting elevated Ka/Ks ratios are of particular interest for understanding species-specific adaptation and molecular evolution. The remaining 8 genes (including *atp4, ccmFc, ccmFn,* and *mttB*) exhibited near-neutral Ka/Ks values, indicating relaxed selective constraints or a transition towards functional divergence[50]. This differential pattern of selective pressure provides valuable insights into the adaptive evolution of mitochondrial genes in alpine plant species.

Conclusions

In this study, we assembled and annotated the first complete mitochondrial genome of *N. scrophulariiflora*, a medicinal plant species endemic to the Himalayan region. The genome spans 606,216 bp with a GC content of 44.31% and exhibits a multi-chromosomal architecture comprising four circular chromosomes, containing 56 genes including 34 protein-coding genes, 3 rRNA genes, and 17 tRNA genes. Comprehensive analyses revealed a preference for A/U-ending codons, 492 high-confidence C-to-U RNA editing sites (score $\geq$ 0.9) predominantly affecting respiratory genes, extensive repetitive sequences (66 SSRs, 11 tandem repeats, and 557 dispersed repeats) that likely contribute to the structural complexity of this mitogenome, and 49 MTPT fragments (37,452 bp, 6.18% of the mitogenome) providing evidence of frequent chloroplast-to-mitochondrion gene transfer. Phylogenetic reconstruction confirmed the placement of *N. scrophulariiflora* within Plantaginaceae consistent with the APG IV classification, while synteny analysis revealed extensive structural

rearrangements compared to closely related species. Selection pressure analysis indicated that most mitochondrial genes are under purifying selection, Phylogenetic reconstruction confirmed the placement of *N. scrophulariiflora* within Plantaginaceae, clustering with Digitalis ferruginea with moderate bootstrap support (65%). Selection pressure analysis indicated that most mitochondrial genes are under purifying selection, with 1 gene (*sdh3*) showing signals of positive selection The results of this study provide essential genomic resources for conservation genetics and evolutionary research on *N. scrophulariiflora*, while also contributing to a broader understanding of mitogenome evolution within Plantaginaceae and the adaptive strategies of alpine medicinal plants. The characterized mitochondrial markers and adaptive genomic features offer a foundation for future investigations into stress tolerance, population genetics, and sustainable utilization of this endangered species.

Declarations

Ethics approval and consent to participate

This study does not involve human subjects, animal experiments, or clinical trials. The plant material (Neopicrorhiza scrophulariiflora) was collected from cultivated sources at Haihongyuan Traditional Chinese Medicine Co., Ltd. in Lijiang, Yunnan, China, with the permission of the cultivation base. The collection and experimental procedures complied with relevant institutional, national, and international guidelines for the use of plant genetic resources. Therefore, formal ethics approval and consent to participate are not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The complete mitochondrial genome sequence of *Neopicrorhiza scrophulariiflora* has been deposited in the GenBank database under the accession number **[PZ729975-PZ729978]** (Neopicrorhiza scrophulariiflora isolate HHLL01 chromosome 4 mitochondr - Nucleotide - NCBI).The datasets generated and analysed during the current study are also available from the corresponding author on reasonable request.

Competing interests

The authors declare no competing interests.

Funding

Fourth batch of university-level key discipline development project 1 – Traditional Chinese Medicine Resources (2025xk01), Xingli Talent Support Program for

Innovative Science and Technology Professionals (2024XLRC03), Lijiang City's Project for Promoting Young and Middle-Aged Academic and Technical Leaders (2023ljrc02). Yunnan Provincial Department of Education Scientific Research Fund Young Talent Project (2026J2411).

Authors' contributions

F.Z. and J.L. contributed equally to this work and share first authorship. They jointly conceived and designed the study, performed the experiments, and drafted the manuscript. C.X. contributed to data collection and formal analysis. Y.L., C.W., and J.S. provided resources and assisted with sample preparation. Y.Z. and S.P. supervised the research, provided critical reagents, and revised the manuscript. A.T. participated in data interpretation, manuscript editing, and project administration. All authors read and approved the final manuscript.

Acknowledgements

We thank the Editor and the anonymous reviewers for their insightful comments and suggestions on the manuscript. The authors thank Beijing Novogene Biotechnology Technology Co., Ltd. for help in genome sequencing and the analysis.

.

References:

1. Rao K, Liu S, Tang X, Jia G, Yang S, Ren C, Pei J: **Integrative analysis of metabolite and transcriptome reveals the biosynthetic pathway and candidate genes for iridoid glycoside biosynthesis in Neopicrorhiza scrophulariiflora (Pennell) D.Y.Hong**. *FRONT PLANT SCI* 2025, **16**:1527477.
2. Zhou F, Wang Y, Liu Y, Zhao X, Wu K, Feng C, Li W, Zhang W, Fang W, Zhan Z: **Textual Research on Picrorhizae Rhizoma in Classical Prescriptions**. *Chinese Journal of Experimental Traditional Medical Formulae* 2025, **31**(09):228-239.
3. Zhou P, Wang J, Li C, Li L, Duan L, Wang W, Liu X, Arshad KT, Liang Y, Zhao Y: **Multi-Omics and Functional Insights into Triterpenoid Biosynthesis Pathways in Neopicrorhiza scrophulariiflora (Pennell) D.Y.Hong**. *PLANTS-BASEL* 2025, **14**(10).
4. Duminil JRM: **Mitochondrial genome and plant taxonomy**. *Methods Mol Biol* 2014, **1115**:121-140.
5. Møller IM, Rasmusson AG, Van Aken O: **Plant mitochondria - past, present and future**. *PLANT J* 2021, **108**(4):912-959.
6. Chevigny N, Schatz-Daas DB, Lotfi FDRQ, Gualberto JM: **DNA Repair and the Stability of the Plant Mitochondrial Genome**. *INT J MOL SCI* 2020, **21**(1).
7. Mower JP, Wang J, Ahmed R, Wu Z: **Rapid turnover of mitogenomic repeats and congruent phylogenomic signal among Digitalideae genomes**. *Genomics Communications* 2026, **3**(1):e004.
8. Liu XL, Qian ZG, Liu FH, Yang YW, Pu CX: **Genetic diversity within and among populations of Neopicrorhiza scrophulariiflora (Scrophulariaceae) in China, an endangered medicinal plant**. *BIOCHEM SYST ECOL* 2011, **39**(4):297-301.
9. Zeng Z, Zhang Z, Norbu N, Bonjor N, Tan X, Zhang S, Tso N, Wang J, Qiong L: **Comprehensive Analysis of the Complete Mitochondrial Genome of Paeonia ludlowii Reveals a Dual-Circular Structure and Extensive Inter-Organellar Gene Transfer**. *BIOLOGY-BASEL* 2025, **14**(7).
10. Liu YJ, Sulc J, Auwerx J: **Mitochondrial genetics, signalling and stress responses**. *NAT CELL BIOL* 2025, **27**(3):393-407.
11. Zhang YM, Qian ZG, Zhang AL, Yang CW, Li GD, Liu XL: **The complete plastid genome sequence of Neopicrorhiza scrophulariiflora (Plantaginaceae): an endangered species endemic to The Himalayas regions**. *MITOCHONDRIAL DNA B* 2019, **4**(2):2504-2505.
12. Mower JP, Hanley L, Wolff K, Pabón-Mora N, González F: **Complete Mitogenomes of Two Aragoa Species and Phylogeny of Plantagineae (Plantaginaceae, Lamiales) Using Mitochondrial Genes and the Nuclear Ribosomal RNA Repeat**. *PLANTS-BASEL* 2021, **10**(12).
13. Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD: **Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae)**. *MOL BIOL EVOL* 2010, **27**(6):1436-1448.
14. 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**(W1):W59-W64.
15. Xiang CY, Gao F, Jakovlić I, Lei HP, Hu Y, Zhang H, Zou H, Wang GT, Zhang D: **Using PhyloSuite for molecular phylogeny and tree-based analyses**. *iMeta* 2023, **2**(1):e87.
16. Zhang Z, Li J, Zhao XQ, Wang J, Wong GKS, Yu J: **KaKs_Calculator: Calculating Ka and Ks Through Model Selection and Model Averaging**. *GENOM PROTEOM BIOINF* 2006, **4**(4):259-263.

17. Gautheret D, Lambert A: **Direct RNA motif definition and identification from multiple sequence alignments using secondary structure profiles11Edited by J. Doudna**. *J MOL BIOL* 2001, **313**(5):1003-1011.
18. Zhang H, Meltzer P, Davis S: **RCircos: an R package for Circos 2D track plots**. *BMC BIOINFORMATICS* 2013, **14**(1):244.
19. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R: **TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data**. *MOL PLANT* 2020, **13**(8):1194-1202.
20. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones S, Marra M: **CIRCOS: an information aesthetic for comparative genomics**. *GENOME RES* 2009, **19**:1639-1645.
21. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TEH, Jin H, Marler B, Guo H *et al*: **MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity**. *NUCLEIC ACIDS RES* 2012, **40**(7):e49.
22. Katoh K, Misawa K, Kuma KI, Miyata T: **MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform**. *NUCLEIC ACIDS RES* 2002, **30**(14):3059-3066.
23. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T: **trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses**. *BIOINFORMATICS* 2009, **25**(15):1972-1973.
24. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R: **IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era**. *MOL BIOL EVOL* 2020, **37**(5):1530-1534.
25. Huang R, Deng H, Wang R, Wei R, Yan S, Wu G, Zheng H: **The first complete multichromosomal mitogenome of Cunninghamia lanceolata (Chinese fir): assembly, characterization and comparative analysis**. *GENE* 2026, **1001**:150210.
26. Wang J, Liu X, Zhang M, Liu R: **The mitochondrial genome of Lavandula angustifolia Mill. (Lamiaceae) sheds light on its genome structure and gene transfer between organelles**. *BMC GENOMICS* 2024, **25**(1):929.
27. Yang H, Chen H, Ni Y, Li J, Cai Y, Ma B, Yu J, Wang J, Liu C: **De Novo Hybrid Assembly of the Salvia miltiorrhiza Mitochondrial Genome Provides the First Evidence of the Multi-Chromosomal Mitochondrial DNA Structure of Salvia Species**. *INT J MOL SCI* 2022, **23**(22).
28. Hu H, Sun P, Yang Y, Ma J, Liu J: **Genome-scale angiosperm phylogenies based on nuclear, plastome, and mitochondrial datasets**. *J INTEGR PLANT BIOL* 2023, **65**(6):1479-1489.
29. Miranda M, Bonekamp NA, Hl IKU: **Starting the engine of the powerhouse: mitochondrial transcription and beyond**. *BIOL CHEM* 2022, **403**:779-805.
30. Gui L, Zhang Z, Song L, Feng C, Yu H, Pan L, Fu J, Liang W, Huang Q, El-Sappah AH *et al*: **Mitogenome of Uncaria rhynchophylla: genome structure, characterization, and phylogenetic relationships**. *BMC GENOMICS* 2025, **26**(1):199.
31. Li J, Chen Y, Liu Y, Wang C, Li L, Chao Y: **Complete mitochondrial genome of Agrostis stolonifera: insights into structure, Codon usage, repeats, and RNA editing**. *BMC GENOMICS* 2023, **24**(1):466.
32. Cho KS, Lee HO, Lee SC, Park HJ, Seo JHE, Cho JH, Park YE, Choi JG, Yang TEJ: **Mitochondrial genome recombination in somatic hybrids of Solanum commersonii and S. tuberosum**. *SCI REP-UK* 2022, **12**(1):8659.
33. Yang H, Ni Y, Zhang X, Li J, Chen H, Liu C: **The mitochondrial genomes of Panax notoginseng reveal recombination mediated by repeats associated with DNA replication**. *INT J BIOL MACROMOL* 2023, **252**:126359.

34. Wynn EL, Christensen AC: **Repeats of Unusual Size in Plant Mitochondrial Genomes: Identification, Incidence and Evolution**. *G3-GENES GENOM GENET* 2019, **9**(2):549-559.
35. Yang Z, Ni Y, Lin Z, Yang L, Chen G, Nijiati N, Hu Y, Chen X: **De novo assembly of the complete mitochondrial genome of sweet potato (Ipomoea batatas [L.] Lam) revealed the existence of homologous conformations generated by the repeat-mediated recombination**. *BMC PLANT BIOL* 2022, **22**(1):285.
36. Small ID, Schallenberg-Rüdinger M, Takenaka M, Mireau H, Ostersetzer-Biran O: **Plant organellar RNA editing: what 30 years of research has revealed**. *PLANT J* 2020, **101**(5):1040-1056.
37. Zhang A, Jiang X, Zhang F, Wang T, Zhang X: **Dynamic response of RNA editing to temperature in grape by RNA deep sequencing**. *FUNCT INTEGR GENOMIC* 2020, **20**(3):421-432.
38. Hu YX, Huang A, Li Y, Molloy DP, Huang C: **Emerging roles of the C-to-U RNA editing in plant stress responses**. *PLANT SCI* 2024, **349**:112263.
39. Liu H, Tian Z, Danzin T, Tan X, Wang J, La Q, Li W: **Characterization and comparative analysis of the complete mitochondrial genome of Phlomoides rotata, a traditional Tibetan medicinal plant**. *BMC GENOMICS* 2025, **26**(1):727.
40. Stadler P, Florentz C, Middendorf M, Donath A, Bernt M, Pütz J, Jühling F: **Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements**. *NUCLEIC ACIDS RES* 2012, **40**(7):2833-2845.
41. Lyu ZY, Yang GM, Zhou XL, Wang SQ, Zhang R, Shen SK: **Deciphering the complex organelle genomes of two Rhododendron species and insights into adaptive evolution patterns in high-altitude**. *BMC PLANT BIOL* 2024, **24**(1):1054.
42. Bai X, Zhu T, Chen H, Wang X, Liu J, Feng Y, Huang Y, Lee J, Kokubugata G, Qi Z *et al*: **Unraveling the mitochondrial genome of the medicinal Chinese motherwort (Leonurus japonicus, Lamiaceae): structural dynamics, organelle-to-nuclear gene transfer, and evolutionary implications**. *FRONT PLANT SCI* 2025, **16**:1546449.
43. The APG, Chase MW, Christenhusz MJM, Fay MF, Byng JW, Judd WS, Soltis DE, Mabberley DJ, Sennikov AN, Soltis PS *et al*: **An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV**. *BOT J LINN SOC* 2016, **181**(1):1-20.
44. Yang H, Chen H, Ni Y, Li J, Cai Y, Ma B, Yu J, Wang J, Liu C: **De Novo Hybrid Assembly of the Salvia miltiorrhiza Mitochondrial Genome Provides the First Evidence of the Multi-Chromosomal Mitochondrial DNA Structure of Salvia Species**. *INT J MOL SCI* 2022, **23**(22).
45. Feng S, Pozzi A, Stejskal V, Opit G, Yang Q, Shao R, Dowling DK, Li Z: **Fragmentation in mitochondrial genomes in relation to elevated sequence divergence and extreme rearrangements**. *BMC BIOL* 2022, **20**(1):7.
46. Cole LW, Guo W, Mower JP, Palmer JD: **High and Variable Rates of Repeat-Mediated Mitochondrial Genome Rearrangement in a Genus of Plants**. *MOL BIOL EVOL* 2018, **35**(11):2773-2785.
47. Han F, Peng Y, Xu L, Xiao P: **Identification, characterization, and utilization of single copy genes in 29 angiosperm genomes**. *BMC GENOMICS* 2014, **15**(1):504.
48. Cadorna CAE, Pahayo DG, Rey JD: **The first mitochondrial genome of Calophyllum soulattri Burm.f**. *SCI REP-UK* 2024, **14**(1):5112.
49. Huang JL, Shi YJ, Tian FF, Mi JX, Zhang F, Zhou SX, Wan XQ, Huang CD: **Structural diversity and phylogenomic insights from the mitochondrial genomes of two populus species from the**

Qinghai-Tibet Plateau. *FRONT PLANT SCI* 2025, **Volume 16 - 2025**.
50. Chen Z, Zhou W, Wang Z, Chen Z, You X, Gong Y: **Analysis of the mitochondrial genome of the Camellia sinensis cv. 'Zhuyeqi': multichromosomal structure, RNA editing sites, and evolutionary characterization**. *FRONT PLANT SCI* 2025, **Volume 16 - 2025**.