BMC Genomics

RESEARCH

Open Access

Complete mitochondrial genomes of the hemiparasitic genus *Cymbaria* (Orobanchaceae): insights into repeat-mediated recombination, phylogenetic relationships, and horizontal gene transfer

Yang Ma[1], Jordi López-Pujol[2,3], Dongqing Yan[1], Zekun Deng[1], Zhen Zhou[1] and Jianming Niu[1,4,5*]

Abstract

Background The Orobanchaceae family is widely recognized as an exemplary model system for examining the evolutionary dynamics of parasitic plants. However, reports on the mitochondrial genome (mitogenome) of the hemiparasitic tribe Cymbarieae are currently lacking. Here, we sequenced, assembled and characterized the complete mitogenome of the genus *Cymbaria* L. sensu stricto (*C. mongolica* and *C. daurica*).

Results A total of 51 unique mitochondrial genes, including 33 protein-coding genes, three rRNA genes, and 15 tRNA genes, are shared by the mitogenomes of the two hemiparasitic plants, exhibiting the gene content characteristic of autotrophic plants. The mitogenomes of *C. mongolica* and *C. daurica* are characterized by a pentacyclic chromosome structure (their major conformation), with lengths of 1,576,465 bp and 1,539,836 bp, respectively. Moreover, we identified and validated the presence of four minor conformations mediated by four pairs of large repeats (> 1000 bp in size) in *C. mongolica* and eight minor conformations mediated by six large repeats in *C. daurica*. We further explored codon usage, RNA editing sites, selective pressure, and nucleotide diversity in two *Cymbaria* mitogenomes. Phylogenetic analyses of 26 species of Lamiales revealed that the two *Cymbaria* species form a sister clade to the other lineages of Orobanchaceae. Extensive mitogenomic rearrangements are also observed between *Cymbaria* and five closely related species. Although we identified mitochondrial plastid sequences in the *Cymbaria* mitogenomes, The mitochondrial plastid sequences (MTPTs) in their mitogenomes represent only 2.37% and 1.74%, respectively. Additionally, there is minimal evidence of intracellular and horizontal gene transfer, with only a few genes (*rpl22*, *rps3*, and *ycf2*) showing low bootstrap support (BS ≤ 70%) for the relationships with the potential host plants *Allium mongolicum*, *Leymus chinensis*, and *Saposhnikovia divaricata,* respectively.

Conclusions We reported the mitochondrial genome in hemiparasitic *Cymbaria* species for the first time, which are characterized by multiple repeat-mediated recombination and little to no intracellular and horizontal gene transfer. Our findings provide valuable genetic insights for further studies on the mitogenome evolution of hemiparasitic plants.

Keywords Orobanchaceae, Hemiparasite, Polycyclic molecules, Repeat-mediated recombination, Gene transfer

*Correspondence:
Jianming Niu
111964401@imu.edu.cn
Full list of author information is available at the end of the article

Background

Mitochondria serve as semi-autonomous organelles, playing a pivotal role in various cellular physiological processes related to energy synthesis and conversion. Mitochondrial genome is akin to chloroplast genome, both exhibiting maternal (uniparental) inheritance, which distinguishes them from the biparentally inherited nuclear genome [1]. The assembly of mitogenomes presents significant challenges, as reflected in the relatively few published mitogenomes compared to chloroplast and plastid genomes. As of February 2025, the NCBI database listed 14,533 chloroplast genomes, 2564 plastid genomes, and 988 mitogenomes. The angiosperm mitogenome shows notable variation in size (ranging from ~ 66 Kb to ~ 12 Mb [2, 3]), which is attributed to factors such as repeat-mediated recombination, intracellular gene transfer (IGT), and horizontal gene transfer (HGT) [4–6]. Recent evidence suggests that the in vivo conformation of angiosperm mitogenomes is highly variable and diverse, despite the common perception of them as monocyclic structures; for instance, they can be found as linear, branched, or polycyclic structures [7–9]. Therefore, the complexity and variability of angiosperm mitogenomes underscore the need for more comprehensive studies, especially when compared to plastomes. Approximately 4530 of the 354,000 angiosperm species (i.e. 1.3%), are classified as parasitic plants, which are those that get their essential nutrients from host plants through haustoria [10]. So far, only a small portion of the mitochondrial genomes of parasitic plants has been reported, with a primary focus on holoparasitic species, including *Cuscuta* L. (Convolvulaceae) [11, 12], *Aeginetia* L. (Orobanchaceae) [13, 14], *Boschniakia* C.A. Mey. ex Bong. (Orobanchaceae) [15], and *Viscum* L. (Viscaceae) [16]. Previous studies have revealed some features that characterize the mitogenomes, such as significant size reduction and gene loss [2], the presence of minicircular molecules and pronounced heteroplasmy [17], and widespread HGT [18]. Despite these promising findings, detailed knowledge of mitogenomes in parasitic plants remains limited, particularly in hemiparasites.

The Orobanchaceae family is the largest among parasitic angiosperms and displays diverse lifestyles including autotrophy, hemiparasitism, and holoparasitism [10]. This family provides an excellent model system for studying the organelle genome evolution of parasitic plants, due to its various parasitic strategies and phylogenetic relationships [19]. Currently, Orobanchaceae is classified into nine tribes [20]. Nevertheless, the phylogenetic relationships among these tribes are still controversial, particularly regarding the clade sister to the remaining parasitic lineages. The hemiparasitic tribe Cymbarieae, comprising about five genera and 20 species, is considered to have played a key role in the transition from autotrophism to heterotrophism because of its phylogenetic position [21]. This tribe has been invariably resolved as sister to other parasitic lineages within Orobanchaceae [22–24], but a recent study on the holoparasitic tribe Orobancheae challenged this traditional classification [25]. Unfortunately, only 12 mitogenomes from four of the Orobanchaceae tribes (which do not include tribe Cymbarieae) are available to date, in sharp contrast to the over 80 plastomes across all nine tribes that are deposited in GenBank. This scarcity of genetic resources impedes our understanding of the evolutionary dynamics of the Orobanchaceae mitogenome.

The hemiparasitic genus *Cymbaria* L., belonging to the tribe Cymbarieae, comprises about two or three species of perennial herbs [26], as it is still unclear whether *Cymbaria borysthenica* Pall. ex Schltdl. (≡ *Cymbochasma borysthenicum* (Pall. ex Schltdl.) Klokov & Zoz) belongs in this genus. *Cymbaria mongolica* Maxim. and *C. daurica* L. are widely distributed in eastern areas of the Eurasian steppe [27]. Both species are facultative hemiparasites that remain photosynthetic and obtain part of their water and nutrients from the diverse host plants through haustoria [28]. They usually parasitize on roots of steppe plants, including members of *Stipa* L. (Poaceae), *Allium* L. (Amaryllidaceae), and *Caragana* Lam. (Fabaceae). Therefore, genetic materials can be exchanged between *Cymbaria* and host plants via haustorial connection. Moreover, our previous research revealed that eight NAD(P)H-dehydrogenase complex genes are either pseudogenized or lost in the *Cymbaria* chloroplast genomes [29]. However, the fate of these lost plastid genes is still largely unknown. Overall, these two *Cymbaria* species provide an ideal system for investigating the IGT and HGT assumption that are common in hemiparasitic plants.

In this study, we assembled the high-quality mitogenomes of *Cymbaria mongolica* and *C. daurica* by the combination of Illumina and Nanopore strategies. Our study aimed to (1) perform comprehensive comparisons between thees two mitogenomes, (2) explore evolutionary relationships within Orobanchaceae, and (3) assess the assumption for the occurrence of IGT and HGT events in the hemiparasites. Our findings will contribute to a deeper understanding of mitogenome evolution in the hemiparasitic Cymbarieae and the conservation and utilization of *Cymbaria* germplasm resources.

Materials and methods

Sampling and DNA extraction

Fresh samples of *C. mongolica* and *C. daurica* were collected from Linfen (35°14′38.63"N, 108°13′06.11"E) and Ejin Horo Banner (39°33′08.35"N, 110°11′42.41"E) in China. Young leaves were immediately frozen in liquid nitrogen at −80 °C. Voucher specimens were deposited in the Herbarium of Inner Mongolia University, with accession numbers MYCM22052206 and MYCD22062608 for *C. mongolica* and *C. daurica*, respectively. Genomic DNA was extracted using the modified CTAB method [30]. DNA concentration was determined using NanoDrop spectrophotometry (NanoDrop Technologies, Inc, Wilmington, DE, USA).

Sequencing, assembly, and annotation

High-quality mitogenomes were obtained by the combination of Illumina sequencing and Nanopore single-molecule sequencing strategies. DNA libraries (350 bp) were constructed with the NEBNext® library building kit according to standard Illumina and Oxford Nanopore protocols. The Illumina HiSeq 2500 platform produced ~ 5.64 Gb and ~ 5.80 Gb of clean data (Illumina Inc., San Diego, CA, USA), while the Oxford Nanopore Technologies GridION platform generated ~ 13.09 Gb and ~ 11.10 Gb of clean data (Oxford Nanopore Technologies Ltd., Oxford, UK), for *C. mongolica* and *C. daurica*, respectively. Total RNA was extracted using the RNeasy Plus Mini Kit (Qiagen Inc., Hilden, Germany). RNA-seq libraries (350 bp) were constructed with the Truseq RNA Library Prep kit following the standard Illumina protocols and sequenced on the Illumina HiSeq 2500 platform (Illumina Inc., San Diego, CA, USA). The de novo assembly of the two *Cymbaria* mitogenomes was performed using Wtdbg2 v2.5 [31]. Candidate contigs were identified by BLASTN v2.2.25 [32]. These contigs were polished using a hybrid error correction method comprising Pilon v1.23 [33] for Nanopore long reads and Racon v1.4.20 [34] for Illumina short reads. Visualization of the assembly results was performed using Bandage v0.8.1 [35]. Protein-coding genes (PCGs) and ribosomal RNAs (rRNAs) were annotated using Geseq v1.43 [36] and BLASTN [32], respectively, with the mitogenomes of *Castilleja paramensis* F. González & Pabón-Mora (NC_031806.1) and *Aeginetia indica* L. (NC_069194.1) as references. Transfer RNAs (tRNAs) were identified by tRNAscan-SE v2.0.7 [37]. The circular maps of chromosomes were produced by OGDRAW v1.3.1 [38] and the heatmaps of PCGs were plotted using Tbtools-II v1.120 [39].

Comparative mitogenome analyses

For repeats element analysis, simple sequence repeats (SSRs), tandem repeats, and dispersed repeats were identified using the MISA, Tandem Repeats Finder (TRF), and REPuter web servers, respectively [40–42]. The distribution of repeats within and across chromosomes was visualized using Tbtools-II [39]. Relative synonymous codon usage (RSCU) was calculated using PhyloSuite v1.2.2 [43]. For identification of RNA editing event, RNA-seq data were mapped to *C. mongolica* and *C. daurica*, respectively. was. After mapping to mitogenomes using BWA v0.7.10-r789 [44], putative RNA editing sites were predicted with Samtools v1.7 with depth > 10 × [45]. For selective pressure analysis, Nonsynonymous (*Ka*) and synonymous substitution rates (*Ks*) were calculated using the yn00 program in PAML v4.9 [46].

Identification and validation of the homologous recombination mediated by large repeats

Previous studies have proven that large repeats (> 1000 bp in size and > 85% in identity) are more prone to mediate homologous recombination [7]. We first detected the high-scoring sequence pairs within and across chromosomes using BLASTN (E-value < 1e-5) [32]. Then, we extracted each pair of these repeats and their flanking sequences of 1000 bp as references. Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast) web server was used for primer design. The PCR amplification was 50 μl in total, consisting of 1 μl template DNA, 2 μl forward primer, 2 μl reverse primer, 25 μl 2 × Phanta Max Master Mix, and 20 μl ddH_2O. The PCR reactions were conducted under the following conditions: 95 °C for 3 min; 35 cycles of 95 °C for 15 s, 55 °C for 15 s, and 72 °C for 15 s; and 72 °C for 5 min. The PCR products were analyzed using agarose (1.5%) gel electrophoresis and further sequenced using the Sanger method.

Phylogenetic and synteny analysis

To determine the phylogenetic placement of the hemiparasitic genus *Cymbaria*, mitogenomes of 24 closely related species of order Lamiales were downloaded from NCBI. *Nicotiana tabacum* L. (NC_006581.1) and *Ginkgo biloba* L. (NC_027976.1) were served as outgroups. PhyloSuite v1.2.2 [43] was employed to extract shared PCGs among the species, and their sequences were aligned using MAFFT v7.450 [47]. Phylogenetic relationships were inferred using IQ–TREE v2.1.4 [48] with GTR + F + I + G4 model and MrBayes v3.2.6 [49] with GTR + I + G mode, respectively. The resulting phylogenetic trees through Maximum Likelihood (ML) and Bayesian Inference (BI) methods were visualized with iTOL v6 [50]. To assess the collinear relationships between two *Cymbaria* species

and five related species (*Aeginetia indica*, *Castilleja paramensis*, *Pedicularis kansuensis* Maxim., *Pedicularis chinensis* Maxim., and *Rehmannia glutinosa* (Gaertn.) Libosch. ex Fisch. & C.A. Mey.), collinear blocks longer than 500 bp were identified using BLASTN [32]. The synteny analysis was visualized with the MCscanX program of Tbtools-II [39]. Furthermore, sequence alignment of the seven species was performed using Mauve v2.4.0 [51], and pairwise comparisons of dot plots between *C. mongolica*, *C. daurica*, and *C. paramensis* were generated by the MAFFT online service [52].

Identification of intracellular and horizontal gene transfer

To detect mitochondrial plastid sequences (MTPTs), the mitogenomes of *C. mongolica* and *C. daurica* were compared against their respective chloroplast genomes (NC_064388.1 and NC_064104.1) using BLASTN [32]. Due to the observed pseudogenization/loss of PCGs in the hemiparasitic *Cymbaria* chloroplast genomes, we also extended this BLASTN analysis to compare each *Cymbaria* mitogenome with the reference plastome of three autotrophic species (*Lindenbergia philippensis* (Cham. & Schltdl.) Benth., *Rehmannia henryi* N.E. Br., and *Triaenophora shennongjiaensis* Xiao D. Li, Y.Y. Zan & J.Q. Li). The results were visualized using Tbtools-II [39]. Phylogenetic analyses were conducted to validate the IGTs and HGTs. Multiple sequence alignments of each mitochondrial and plastid-derived PCGs or fragments were performed using MAFFT [47]. Single-gene Maximum Likelihood (ML) trees were constructed using RAxML with the GTRGAMMA model and 1000 bootstrap replicates [53]. According to Zhong's method [13], a gene was considered an IGT if, in the phylogenetic trees, Lamiales was resolved as monophyletic and *Cymbaria* nested within Orobanchaceae species with bootstrap support (BS) > 80%. Conversely, a gene was categorized as an HGT if, despite Lamiales being monophyletic, *Cymbaria* clustered with non-Lamiales species with BS > 80%.

Results

Features of the *Cymbaria* mitogenomes

Five circular chromosomes were resolved in both *C. mongolica* and *C. daurica* mitogenomes, measuring in total 1,576,465 bp with 45.4% GC content and 1,539,836 bp with 45.5% GC content, respectively (Figs. 1 and S1). The nucleotide composition for *C. mongolica* was 27.3% of A, 22.7% of C, 22.7% of G, and 27.3% of T, and for *C. daurica*, 27.2% of A, 22.8% of C, 22.8% of G, and 27.2% of T. Additional details on the size, GC content, and nucleotide composition of each molecule are provided in Table 1.

The *C. mongolica* mitogenome comprised 51 unique genes, including 33 protein-coding genes (PCGs), three ribosomal RNA (rRNA) genes, and 15 transfer RNA (tRNA) genes. The *C. daurica* mitogenome contained 53 unique genes, consisting of 34 PCGs, three rRNA genes, and 16 tRNA genes (Fig. 2 and Table S1). Each mitogenome had two copies of *trnM-CAU*, *trnP-UGG* , and *trnQ-UUG*. The *Cymbaria mongolica* mitogenome exclusively possessed the duplicated genes *trnC-GCA*, *trnG-GCC*, and *trnW-CCA*, whereas an additional copy of *trnF-GAA* was only present in the mitogenome of *C. daurica*. Interestingly, the exons of genes *nad1*, *nad2*, and *nad5* were distributed across different chromosomes in both mitogenomes. The two mitogenomes had eight introns in common. The genes *cox1* and *rps10* each contained one intron, while *nad4* and *nad7* each held three introns. Additionally, the *ccmFC* gene in *C. daurica* included one intron. Six (*trnC-GCA*, *trnD-GUC*, *trnG-GCC*, *trnM-CAU*, *trnP-UGG*, and *trnW-CCA*) and seven (*trnC-GCA*, *trnD-GUC*, *trnG-GCC*, *trnH-GUG*, *trnL-UAA*, *trnM-CAU*, and *trnW-CCA*) plastid-derived tRNA genes were identified in the mitogenomes of *C. mongolica* and *C. daurica*, respectively, with the remaining tRNA genes being mitochondria-native.

Comparative mitogenome analyses

In the mitogenomes of *C. mongolica* and *C. daurica*, 278 and 279 simple sequence repeats (SSRs) were identified, respectively, with tetra-nucleotide repeats being the most prevalent type (Fig. 3 and Tables S2). The most frequent mono- to tetra-nucleotide repeats in both species were A/T (31, 34), AG/CT (41, 39), AAG/CTT (24, 24), and AAAG/CTTT (57, 40). The *C. mongolica* mitogenome contained 16 tandem repeats with periods ranging from 12 to 29 bp and a matching degree greater than 72%. In contrast, *C. daurica* had 13 tandem repeats with matching degrees over 75% and period sizes between 12 to 71 bp (Fig. 3 and Tables S3). There were 4924 and 5000 pairs of dispersed repeats (≥ 30 bp) in *C. mongolica* and *C. daurica*, respectively, accounting for 19.45% and 19.22% of the total mitogenome (Tables S4). The majority of these repeats were shorter than 70 bp (99.05% and 99.08%), while 47 forward and 46 palindromic repeats exceeded 70 bp (Fig. 3).

The PCGs of the mitogenomes (33 in *C. mongolica* and 34 in *C. daurica*) were analyzed for relative synonymous codon usage (RSCU) and RNA editing site prediction. A total of 9659 and 9448 codons were identified in the mitogenomes of *C. mongolica* and *C. daurica*, respectively (Table S5). A broad preference for the RSCU was observed in both mitogenomes, except for the methionine (AUG) and tryptophan (UGG) start codons. Furthermore, Leucine (Leu) was the most frequent amino acid, while cysteine (Cys) was the least one (Fig. S2). The mitogenomes of *C. mongolica* and

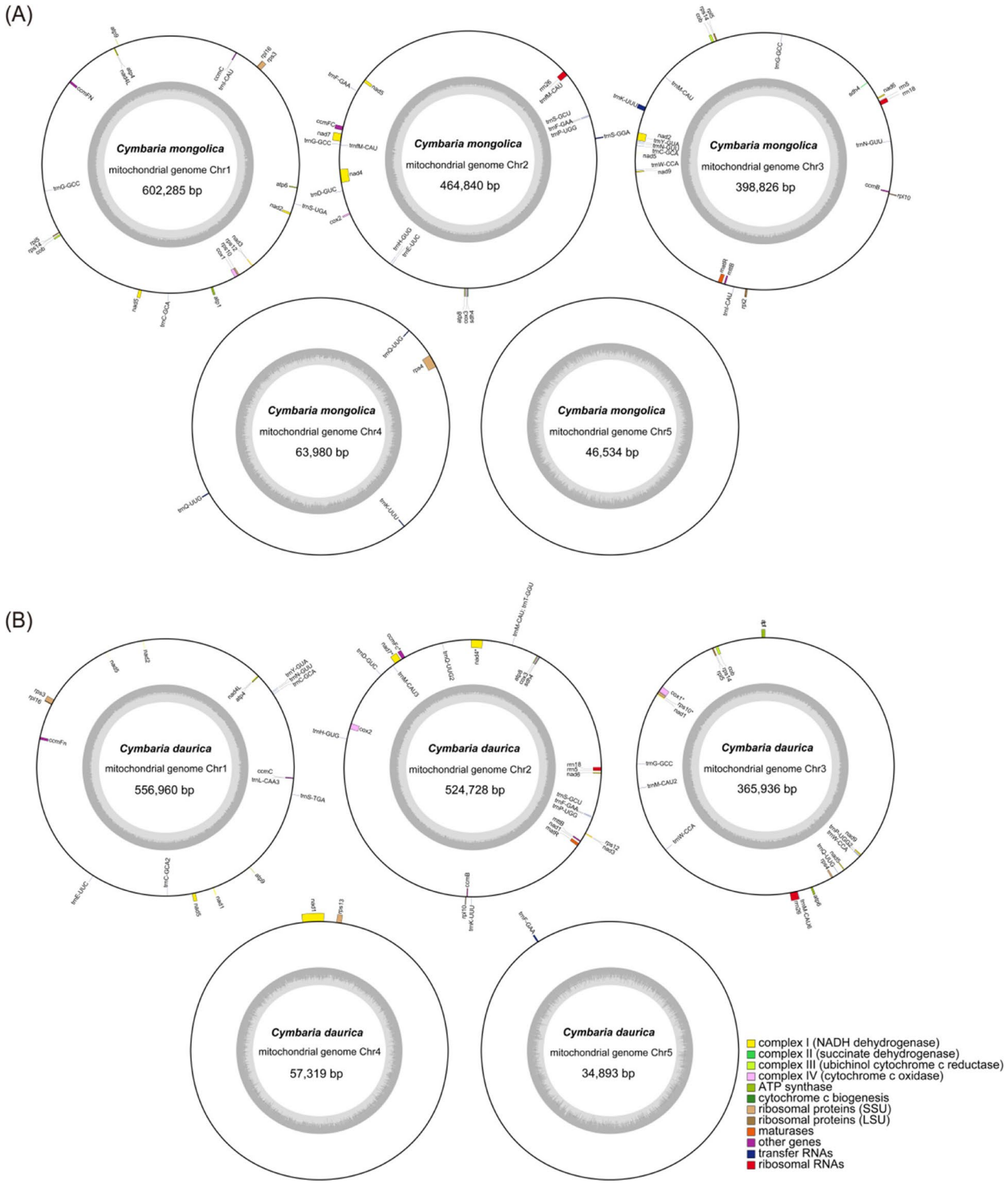

Fig. 1 Circular maps of the C*ymbaria mongolica* (**A**) and *C. daurica* (**B**) mitogenomes. 'Chr' means 'chromosome'. Gene transcript clockwise or counter-clockwise strands are drawn on the upper and on the lower part of the circles, respectively. Genes belonging to different functional groups are color-coded. Asterisks (*) indicate genes containing intron(s)

Table 1 Genomic features of the *Cymbaria mongolica* and *C. daurica* mitogenomes. 'Chr' means 'chromosome'

Species	Feature	A (%)	C (%)	G (%)	T (%)	GC (%)	Size (bp)
C. mongolica	Whole genome	27.3	22.7	22.7	27.3	45.4	1,576,465
	Chr 1	27.3	22.7	22.6	27.4	45.3	602,285
	Chr 2	27.4	22.7	22.7	27.2	45.4	464,840
	Chr 3	27.1	22.9	22.7	27.3	45.6	398,826
	Chr 4	26.6	22.6	22.8	28	45.4	63,980
	Chr 5	27.5	22.6	22.6	27.4	45.1	46,534
C. daurica	Whole genome	27.2	22.8	22.8	27.2	45.5	1,539,836
	Chr 1	27.3	22.7	22.7	27.3	45.4	556,960
	Chr 2	27	22.9	22.8	27.3	45.7	524,728
	Chr 3	27.3	22.6	22.9	27.2	45.5	365,936
	Chr 4	27.3	22.6	22.9	27.3	45.5	57,319
	Chr 5	27.2	23	22.7	27	45.7	34,893

C. daurica contained 110 and 188 RNA editing sites, respectively (Table S6). The C-to-U editing changes were predominant, accounting for 95.45% in *C. mongolica* and 97.87% in *C. daurica*, and they mostly occurred at the first two positions of the triplet codon (94.54% and 92.55%, respectively). Three genes (*atp6*, *nad3*, and *cox1*) were edited more than 10 times in both species. Additionally, *nad4*, *nad7*, *atp9*, and *nad4L* had more than 10 editing sites in *C. daurica*. The ratio of synonymous (*Ka*) and nonsynonymous (*Ks*) substitution rates as well as nucleotide diversity (Pi) of 33 common PCGs in both mitogenomes were also calculated (Table S7). Most *Ka*/*Ks* ratios were less than 1.0, indicating predominant negative selection. However, three genes (*atp1*, *atp8*, and *rps3*) had *Ka*/*Ks* ratios above 1.0, indicating positive selection. Additionally, while most genes (31 of 33) had Pi values below 0.1, the *nad1* and *nad5* genes showed higher Pi values above this value (0.17492 and 0.14127, respectively) (Table S7).

Large repeat-mediated recombination analysis

Interestingly, we found four and six pairs of large repeats with a total length of 248,864 and 220,143 bp in the mitogenomes of *C. mongolica* and *C. daurica*, respectively (Table S8). Their length ranged from 2572 to 92,624 bp, which were too long to amplify the entire regions. Thus, specific primers were designed for boundary validation (Fig. 4A and B; Table S9). All junction sequences were validated by PCR amplification (Fig. 4C and D) and Sanger sequencing (Supplementary Additional file 3). The sequencing results were consistent with those expected, thus confirming the occurrence of large repeat-mediated recombination.

Based on experimental verification, we proposed a potential homologous recombination process for the two *Cymbaria* mitogenomes (Fig. 5). The major conformation of *C. mongolica* mitogenome (five molecules) could form minor conformations Mic01, Mic03 and Mic04) by means of the forward repeats R1, R3 and R4 mediated recombination, respectively. Of these, Mic01 could shift into Mic02 through the reserve repeat R2. Similarly, the major conformation of *C. daurica* mitogenome (five molecules) could shift into the three minor conformations Mic02, Mic04, and Mic07 by the three forward repeats R1, R3, and R6, and further into five minor conformations Mic01, Mic03, Mic05, Mic06, and Mic08 by the three reserve repeats R2, R4, and R5, respectively. These minor conformations form four circular structures, except for Mic04 of *C. mongolica* mitogenome with six circular molecules. In summary, four and six pairs of large repeats were identified that could mediated recombination in the *C. mongolica* and *C. daurica* mitogenomes, making their major conformation to form four and eight minor conformations, respectively.

Phylogenetic relationships and synteny analysis

To elucidate the evolutionary status of the genus *Cymbaria*, phylogenetic trees of 26 Lamiales (Table S10) were constructed using 24 shared mitochondrial PCGs. Both Maximum Likelihood (ML) and Bayesian Inference (BI) methods yielded mutually consistent topologies (Fig. 6), aligning with the Angiosperm Phylogeny Group IV (APG IV) classification system. These findings robustly supported the placement of *C. mongolica* and *C. daurica* within the tribe Cymbarieae, forming a sister clade to the other parasitic lineages in Orobanchaceae.

Synteny analysis revealed numerous small-sized homologous collinear blocks between the two *Cymbaria* and the five closely related species (Fig. 7), indicating low

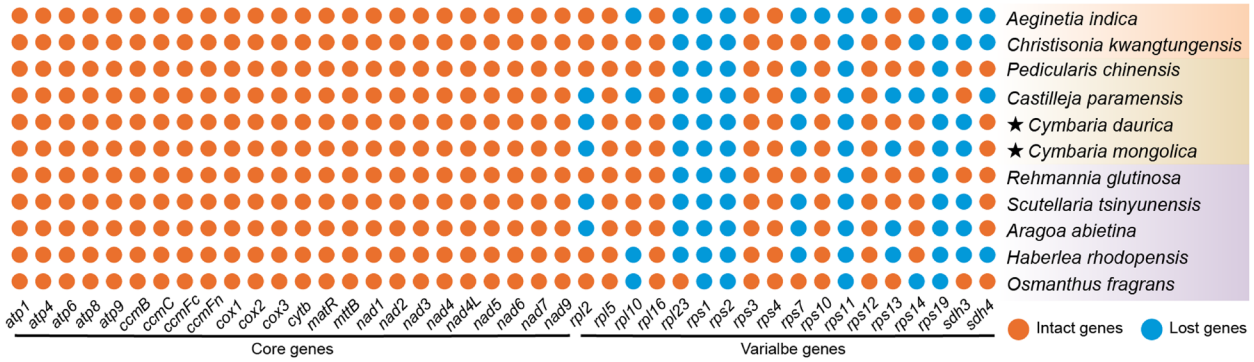

Fig. 2 Protein-coding genes annotation of the mitogenomes of the 11 species of Lamiales. Background colors in purple, brown, and orange indicate autotrophs, hemiparasites, and holoparasites, respectively

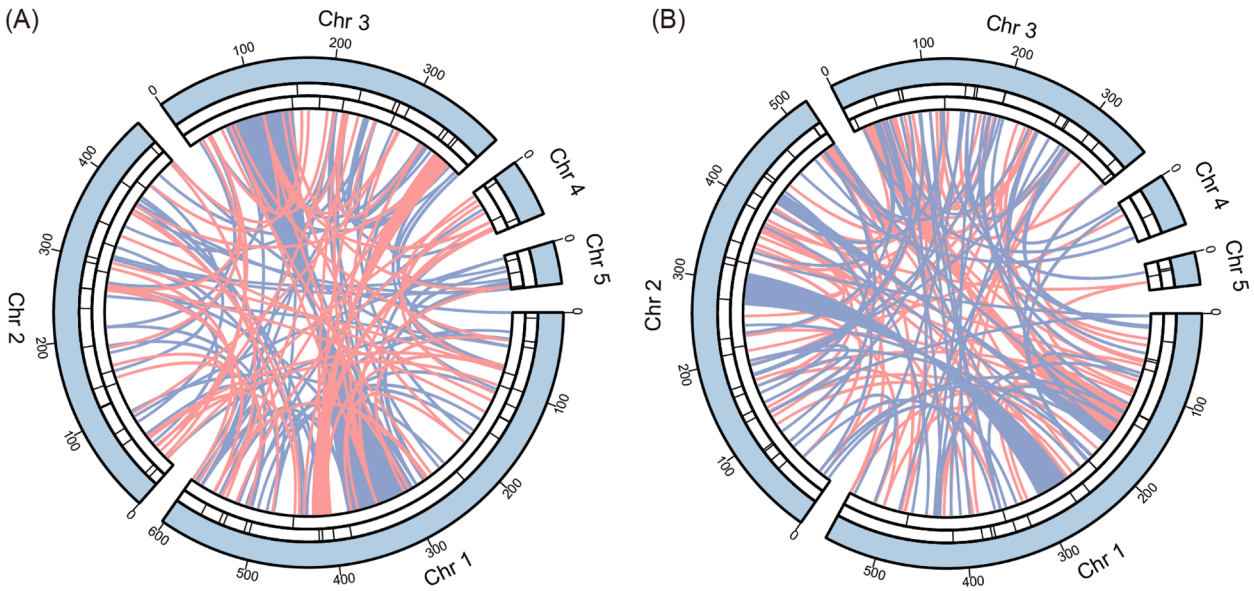

Fig. 3 Schematic diagram of the repetitive sequences identified in the mitogenomes of *Cymbaria mongolica* (left) and *C. daurica* (right). 'Chr' means 'chromosome'. The colored lines on the innermost circle connect pairs of dispersed repeats (≥ 70 bp in size), of which the blue and pink lines represent forward and palindromic repeats, respectively. The black segments on the middle circle and on the outermost circle represent tandem repeats and SSRs, respectively

collinearity across these seven Orobanchaceae mitogenomes. Furthermore, the arrangement of these blocks varied, suggesting an extremely non-conserved structure. These findings indicated that the *Cymbaria* mitogenomes had undergone extensive genomic rearrangements compared to mitogenomes of related species. Similar observations were made using the Mauve program (Fig. S3A). Dot-plot analysis revealed longer collinear sequences with higher homology between *C. mongolica* and *C. daurica* (Fig. S3B, S3C, and S3D). In addition, the lack of homology in some regions of these mitogenomes suggested that the segments were not rearranged.

Identification of intracellular and horizontal gene transfer

A total of 103 and 94 homologous fragments were identified in the mitogenomes of *C. mongolica* and *C. daurica*, spanning 37,292 bp and 26,782 bp, respectively (Fig. 8; Table S11). These MTPTs account for 24.96% (2.37%) and 17.67% (1.74%) of their chl- (mito-) genomes, respectively. The identified fragments included complete PCGs such as *rpl23* in *C. mongolica* and *rps14* in *C. daurica*, partial PCGs (e.g., *ycf2*, *psbB*, and *rpl22*), and intergenic spacer (IGS) regions. Notably, a 462 bp sequence of the gene region in *C. daurica* with similarities to the pseudogene *ndhH* was detected from the BLASTN search

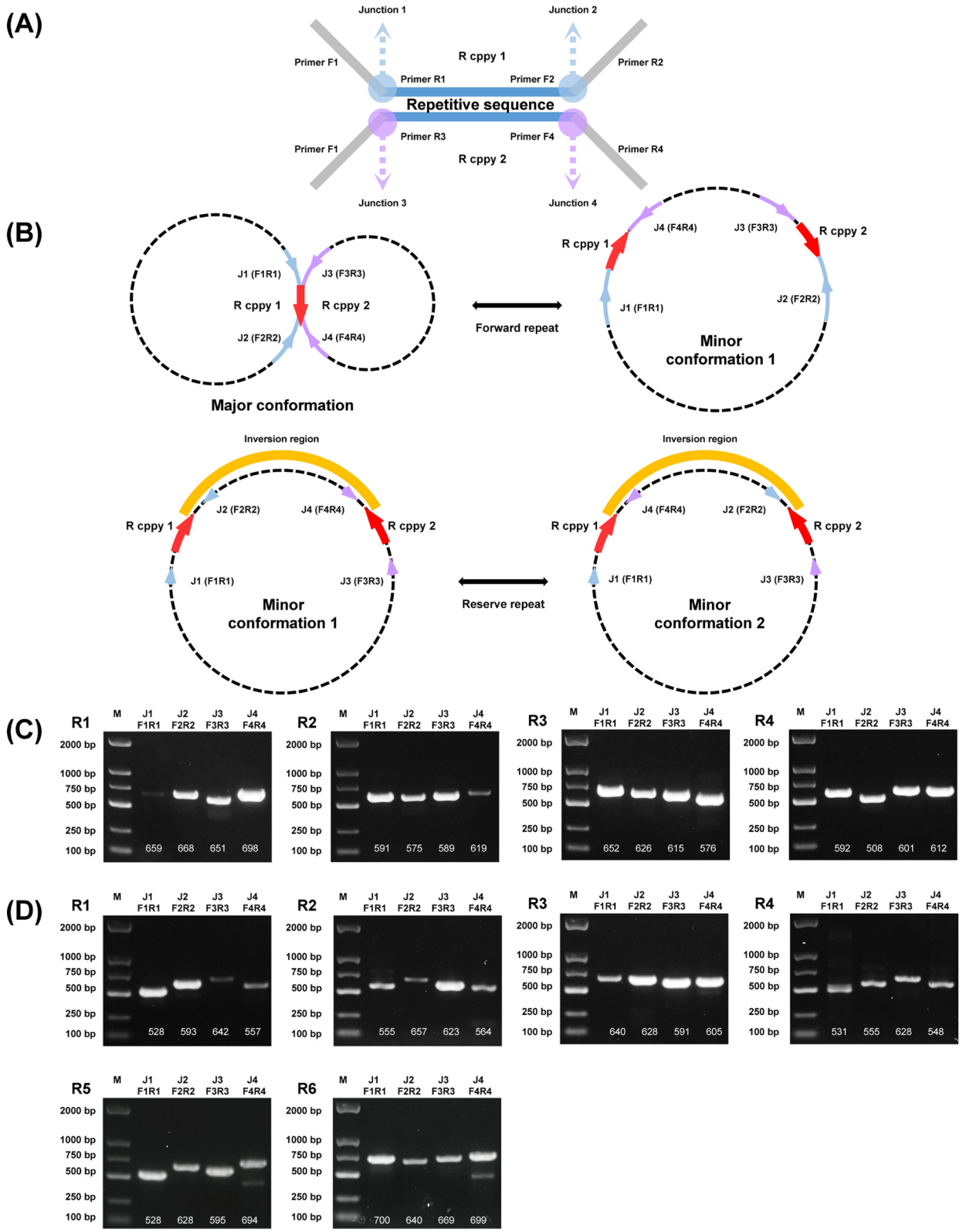

Fig. 4 Validation of the homologous recombination mediated by large repeats. Schematic diagram of primer design (**A**) and experimental design (**B**) for the junction site validation. Agarose gel electrophoresis results of the PCR products for the junction sequences of large repeats in the mitogenome of *Cymbaria mongolica* (**C**) and *C. daurica* (**D**). The junction sites, PCR primer names, and PCR product sizes are shown above each panel

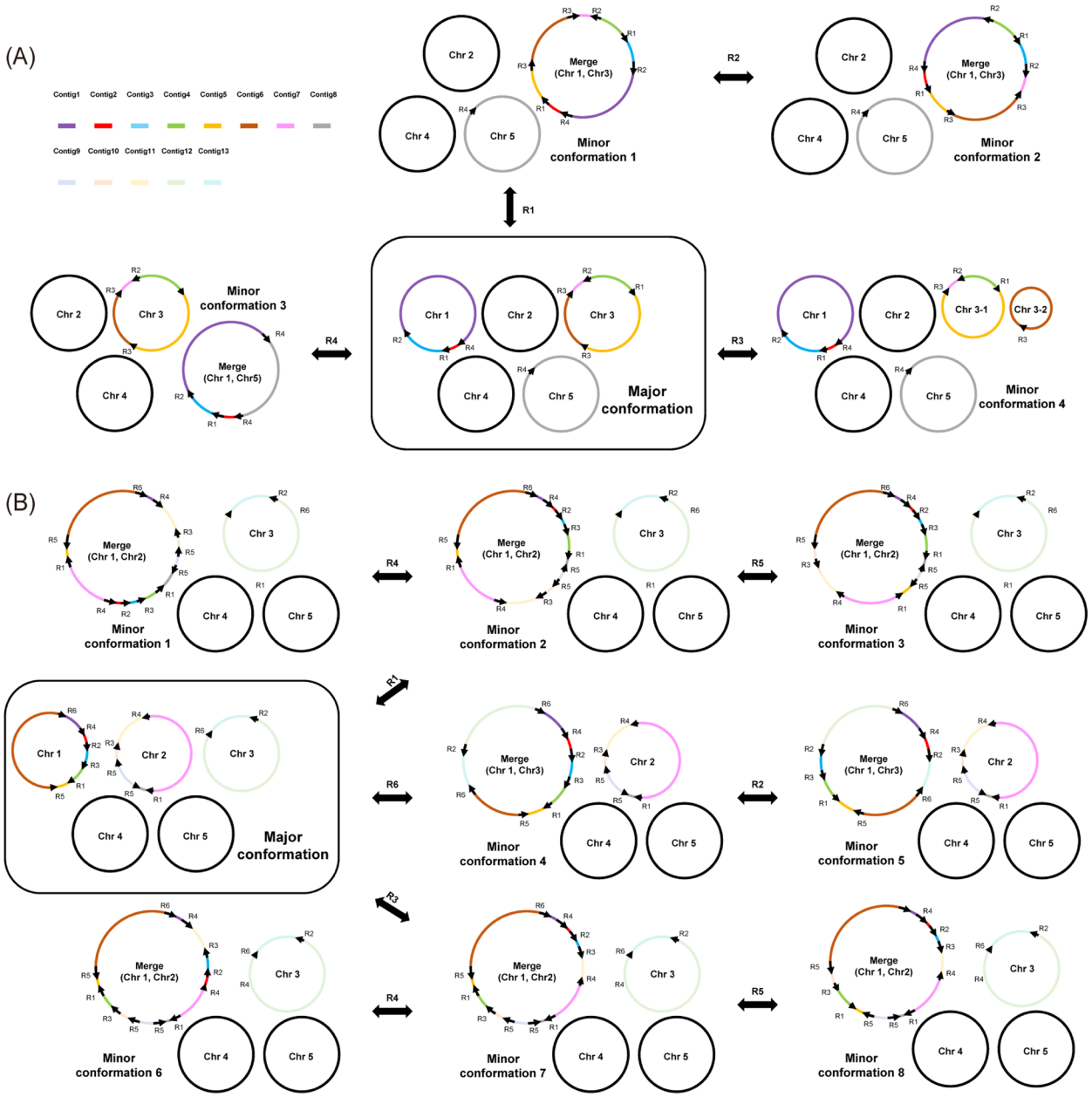

Fig. 5 Hypothetical products generated by recombination mediated by large repeats in the mitogenome of *Cymbaria mongolica* (**A**) and *C. daurica* (**B**). The circles represent the mitogenome conformations. The black arrows represent the repeats, and the colored lines represent the fragments between the repeats

against reference autotrophic plastomes (Tables S12 and S13).

To identify IGTs and HGTs in the two *Cymbaria* species, we selected reference species from the same family Orobanchaceae and order Lamiales, as well as their host families(e.g. Poaceae, Fabaceae, and Asteraceae) (Tables S14 and S15). The most homologous fragments, such as *nad2*, were grouped with non-Lamiales species (bootstrap support, BS < 80%) or nested within Orobanchaceae species (BS < 80%), suggesting a lack of IGTs or HGTs (Figs. 9A and S4). Additionally, most of individual mitochondrial PCGs showed that *Cymbaria* species clustered with other Lamiales species rather than with their potential hosts (Fig. S5). Interestingly, only a few genes (*rpl22*, *rps3*, and *ycf2*) exhibited low bootstrap

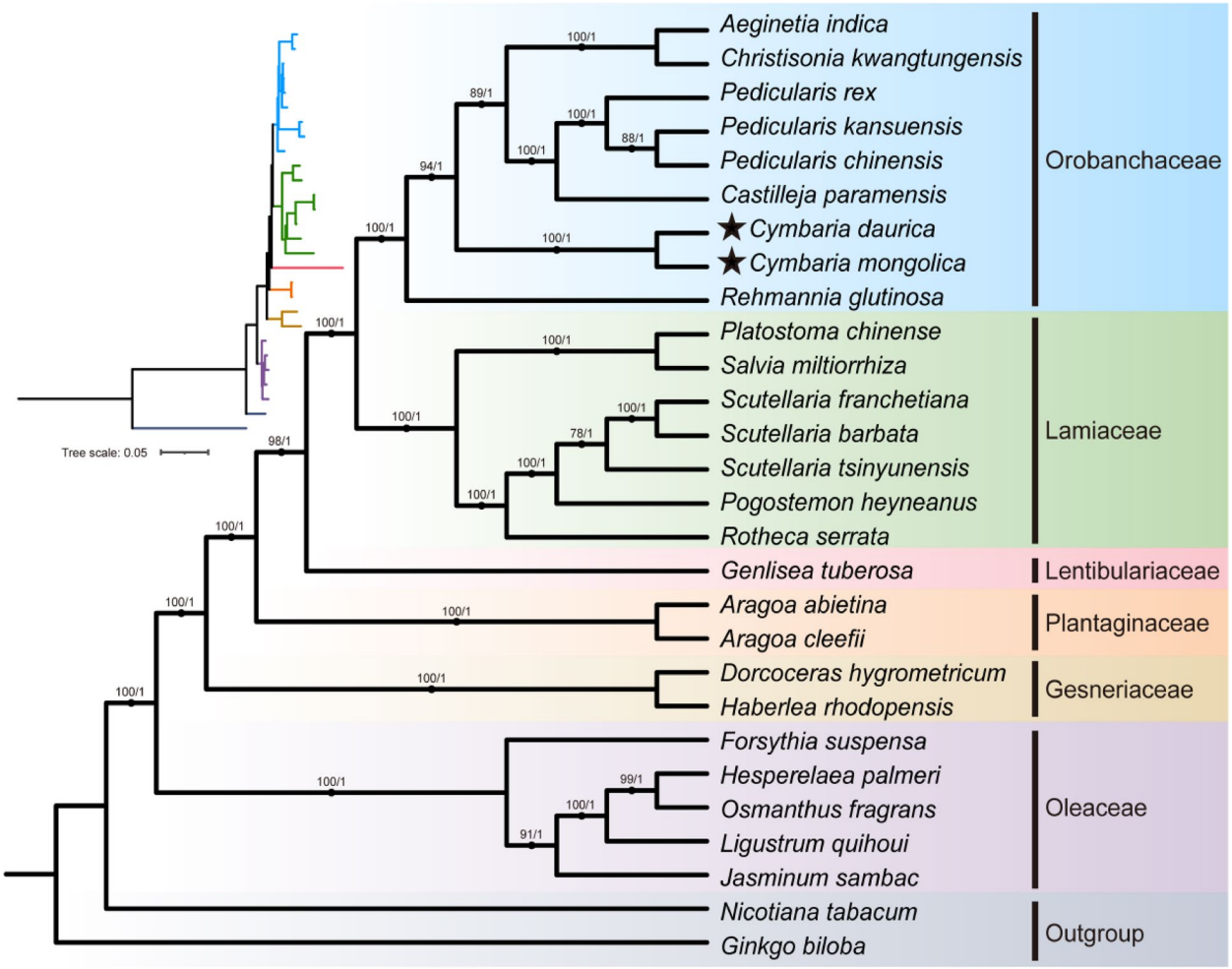

Fig. 6 Phylogenetic relationship of genus *Cymbaria* with 24 additional species of Lamiales. The tree was inferred from ML and BI methods based on 30 shared mitochondrial genes. Numbers above the branches indicate bootstrap values (BS, left) and posterior probabilities (PP, right). The stars (★) indicate the two newly sequenced *Cymbaria* species

support for their relationships with their putative host plants: *Allium mongolicum* (BS = 70%, Fig. 9B), *Leymus chinensis* (BS = 62%), and *Saposhnikovia divaricata* (Turcz.) Schischk. (BS = 48%), respectively.

Discussion

Hemiparasitic *Cymbaria* exhibits a mitochondrial gene content characteristic of autotrophic plants

Angiosperm mitogenomes tend to be more complex than those of animals and fungi due to size and structural variations, increased gene content and duplications, presence of introns, and the influence of various evolutionary processes [54]. Third-generation sequencing technologies have significantly facilitated the assembly and analysis of angiosperm mitogenomes [55]. In this study, a hybrid sequencing approach combining short reads and long reads was utilized to assemble the first two complete mitogenomes within the hemiparasitic tribe Cymbarieae of Orobanchaceae.

While researchers have unveiled a substantial depletion of mitochondrial genes in the holoparasitic mistletoe *Viscum* [2, 56], the broad significance of this finding for parasitic plants is becoming more disputed. The two *Cymbaria* mitogenomes shared a total of 51 unique genes, including 33 PCGs, three rRNA genes, and 15 tRNA genes, of which all core genes were intact with only several variable genes (*rpl2*, *rpl23*, *rps1*, *rps2*, *rps7*, *rps11*, *rps19*, and *sdh3*) being lost. It should be noted that these lost variable genes were also observed in five autotrophic Lamiales species in this study and other angiosperms [57, 58]. As mitochondrial genomes of parasitic plants such as *Rhopalocnemis phalloides* Jungh. [17], *Aeginetia indica* [13], and *Cuscuta* [59] are being documented, they also have intact core genes and do not show considerable gene reductions. One might hypothesize that these

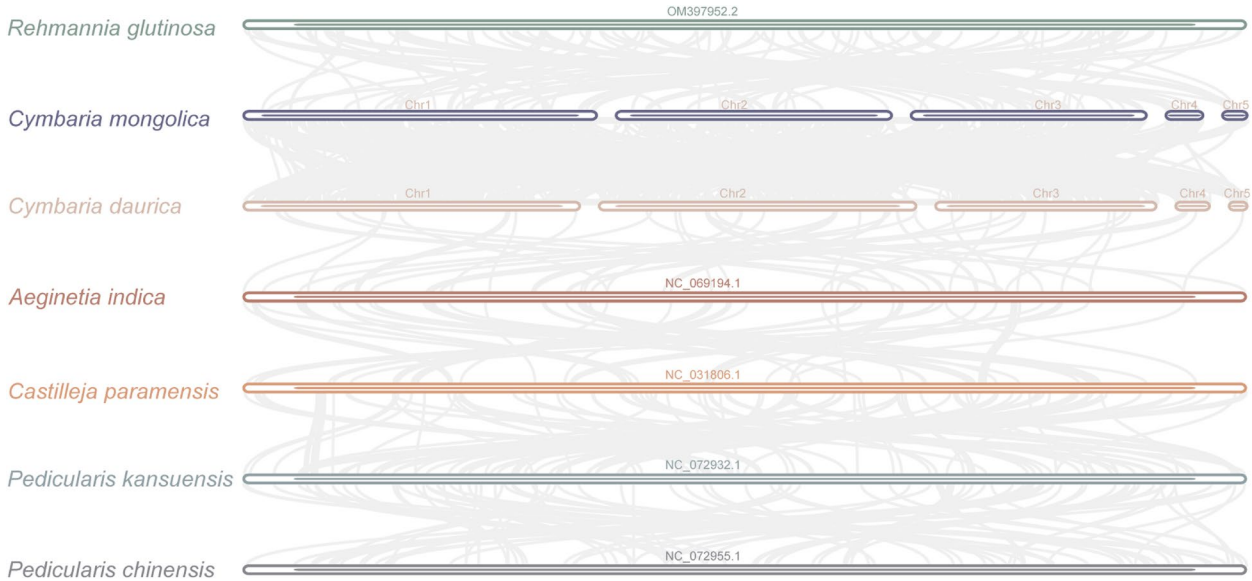

Fig. 7 Synteny analysis of the mitogenomes of two *Cymbaria* species and five closely related species

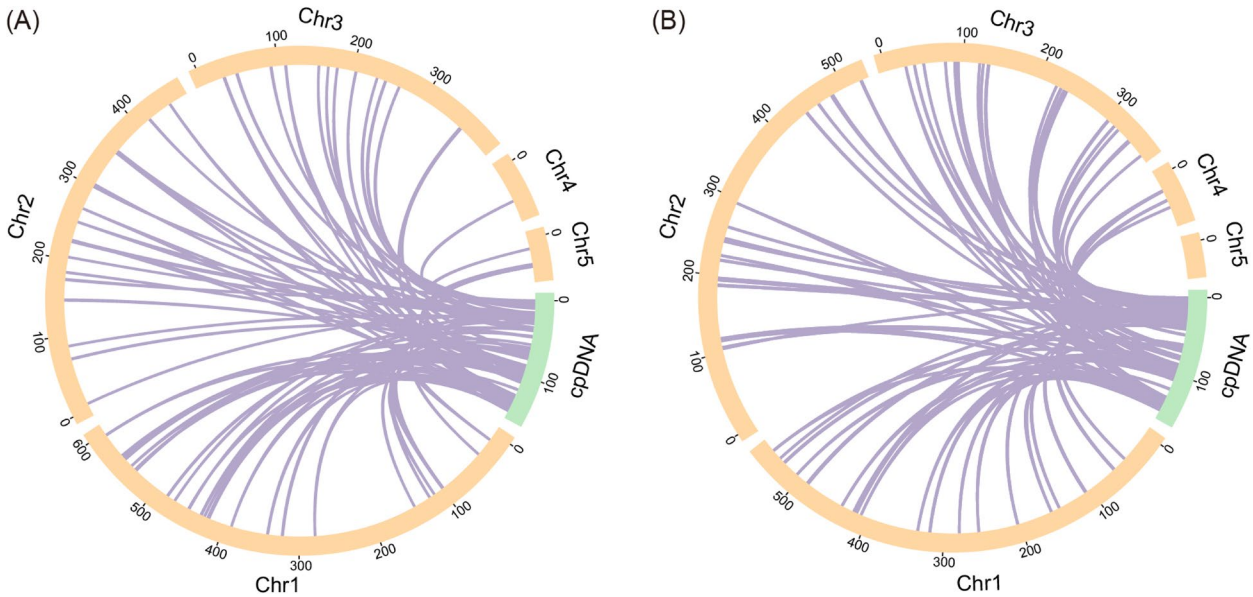

Fig. 8 Schematic diagram of the mitochondrial plastid sequences (MTPTs) in C*ymbaria mongolica* (**A**) and *C. daurica* (**B**). The green and yellow arcs correspond to the chloroplast genomes and mitogenomes, respectively. The purple lines represent homologous DNA fragments

losses would be less likely to affect mitochondrial activity detrimentally, as each loss is typically compensated by nuclear-encoded homologs that ensures functional maintenance [60], though validation through sequencing of the *Cymbaria* nuclear genome will be necessary to confirm this hypothesis. In summary, the hemiparasitic genus *Cymbaria* exhibits mitochondrial gene content comparable to that of autotrophic species, whereas its chloroplast gene content is slightly reduced relative to autotrophic plants [29].

To further understand functional implications of mitochondrial gene variation in *C. mongolica* and *C. daurica* mitogenomes, various other aspects have been investigated, including RSCU, RNA editing sites, selective pressure, and nucleotide diversity. Most PCGs of two *Cymbaria* mitogenomes initiate with the typical ATG

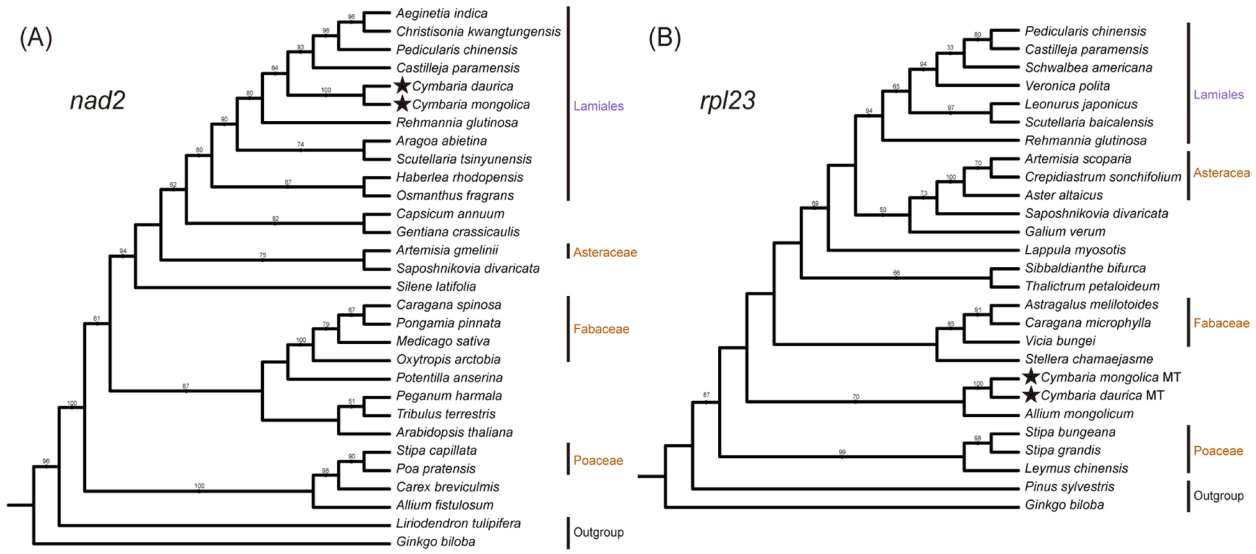

Fig. 9 The single-gene ML trees constructed using plastid-derived *nad2* gene (**A**) and mitochondrion-native *rpl23* gene (**B**). Bootstrap values > 50% were shown on branches

codon, and the distribution of amino acid compositions exhibits high similarity to that of other angiosperms [61]. As described in previous studies [62], we have detected a strong A or T bias in the third position of the codon in the PCGs. Consistent with the mitogenomes of most land plants [9], most of RNA editing sites in *Cymbaria* are located at the first or the second position of the triplet codon. Our results have shown that 30 out of 33 shared mitochondrial PCGs undergo negative selection, as it is commonly observed in angiosperms [63]. In contrast, the remaining three genes (*atp*1, *atp*8, and *rps*3) are under positive selection, which is of significance for further studies on the evolutionary dynamics of *Cymbaria* species.

Homologous recombination mediated by large repeats contributes to the larger size and pentacyclic structure of the *Cymbaria* mitogenomes

Angiosperm mitogenomes demonstrate extensive diversity in size and structural configuration, typically characterized by a single circular chromosome ranging between 200 and 750 Kb [2]. The length of the two *Cymbaria* mitogenomes exceed 1500 Kb, surpassing most published ones in Orobanchaceae species, except for *Cistanche salsa* (1,708,661 bp) and *C. tubulosa* (1,860,774 bp) [64]. Moreover, the *Cymbaria* mitogenome consist of five circular chromosomes, mirroring the polycyclic molecule pattern observed in some Orobanchaceae mitogenomes like *Christisonia kwangtungensis* (Hu) G.D. Tang, J.F. Liu & W.B. Yu [65] with three circular chromosomes, *Boschniakia rossica* (Cham. & Schltdl.) B. Fedtsch. with three circular chromosomes [15], and *B. himalaica* Hook. f. & Thomson with 15 circular chromosomes [15]. This pentacyclic structure with a relatively large genome size in *C. mongolica* and *C. daurica* merits further investigation into its underlying causes.

Repetitive sequences have been shown to mediate frequent homologous recombination in angiosperm mitogenomes [63]. Large repeats often play a crucial role in inter- or intra-molecular recombination, leading to structural variations and increased genome sizes [66]. Moreover, the high-frequency recombination tends to be mediated by large repeats [67]. Our results revealed that large repeats make up a relatively substantial proportion of the total mitogenome in *C. mongolica* and *C. daurica*, 15.79% and 14.30% respectively. This implies their significant role in size expansion. Except for R2 in *C. mongolica*, these repeats are located on different circular molecules. We further confirmed that they can mediate homologous recombination through PCR amplification and Sanger sequencing. These large repeats act as homologous regions, promoting the exchange of genetic material between circular chromosomes, leading to the formation of their main configuration of the pentacyclic chromosome structure through recombination. Furthermore, we speculated that four and six pairs of large repeats could mediate the formation of four and eight minor conformations in the *C. mongolica* and *C. daurica* mitogenomes, respectively. In *C. mongolica*, the minor conformation Mic04 contains six circular molecules, the major conformation contains five circular molecules, and the three remaining minor conformations (Mic01, Mic02, and

Mic03) contain four circular molecules. In *C. daurica*, the major conformation of its mitogenome contains five circular molecules, while the eight minor conformations (Mic01 to Mic08) contain four molecules. Similar results have been reported in *Gelsemium elegans* (Gardner & Champ.) Benth. [68], *Indigofera amblyantha* Craib [69], and *Agrostis stolonifera* L. [70]. Interesting, our prior investigations have demonstrated that the distinctive *rbcL-matK* inversion observed in *Cymbaria* chloroplast genomes originates from a repeat-mediated recombination [29]. These observations may may originate from a specialized DNA repair mechanism, playing a crucial role in maintaining genome stability and influencing species evolution [71]. Overall, homologous recombination mediated by large repeats contributes to the large size and polycyclic structure in the two *Cymbaria* mitogenomes.

Cymbaria species form an independent sister clade to all other parasitic Orobanchaceae

The Orobanchaceae family is highly valued as a model system for investigating the evolutionary dynamics of parasitic plants due to its wide range of lifestyles, including autotrophic, hemiparasitic, and holoparasitic ones [72]. Considerable debate has arisen recently on the sister clade to all other parasitic Orobanchaceae lineages [24, 25]. Historically, the hemiparasitic tribe Cymbarieae was considered the most basal group within Orobanchaceae [22, 23], while more recent phylogenetic studies have placed the holoparasitic tribe Orobancheae as sister to all other lineages [25]. Phylogenetic analysis supported a closer relationship between *C. mongolica* and *C. daurica* with *Cymbaria* diverging the earliest within parasitic Orobanchaceae. This is consistent with the synteny analysis, which shows that more collinear blocks exist between the two *Cymbaria* species rather than between these two species and their five closely related species. Our findings support recent conclusions drawn from chloroplast genomes [29] and 907 low-copy orthogroups [24] and challenge the earlier results of [25] based on a limited number of nuclear genes. Although currently only 14 Orobanchaceae mitogenomes are available for phylogenetic analysis, our study would be viewed as a valuable contribution to the ongoing exploration of plant parasitism evolution.

Little to no intracellular and horizontal gene transfer into the hemiparasitic *Cymbaria* mitogenomes

Typically, the proportion of plastid-derived sequences in mitogenomes ranges from 0.56% to 10.85% [73]. In the organelle genomes of *C. mongolica* and *C. daurica*, homologous fragments totaling 37,292 bp and 26,782 bp, respectively, were identified, corresponding to 2.37% and 1.74% of their mitogenomes. Thus, the relatively low proportion of MTPTs in the *Cymbaria* mitogenome suggested that their contribution in the size expansion of the *Cymbaria* mitogenome was relatively insignificant. Previous research indicates that the most common direction of DNA fragment migration within organelle genomes is from the plastome to the mitogenome [74]. Intracellular transfer of genes from the chloroplast to the mitogenome, particularly those related to photosynthesis, has been observed in holoparasitic *Orobanche* L. species [75], *Aeginetia indica* [13, 14, 76], and *Boschniakia rossica* [15]. However, no evidence of IGT events was found in the holoparasitic *Christisonia kwangtungensis* [65], in *Boschniakia himalaica* [15], and in the two hemiparasitic *Cymbaria* species as revealed here. A few fragments homologous to six partial genes (*ndhJ*, *psbB*, *rpl22*, *rpoB*, *rpoC2*, and *ycf4*) were identified within Orobanchaceae species (BS ≥ 80%), but these fragments are incomplete and nonfunctional. Furthermore, these genes are still present in their respective chloroplast genomes. It should be noted that 8 *ndh* genes that have been lost or pseudogenized in the two *Cymbaria* chloroplast genomes have not been transferred to their mitochondrial genomes, indicating that these genes may have been transferred to the nuclear genome. Therefore, it is necessary to carry out sequencing of the *Cymbaria* nuclear genome in the future.

Reports of horizontal gene transfer events involving mitochondrial genomes are scarce and predominantly focus on parasitic plants. However, there is no conclusive determination regarding the uniformity and variability of HGT between parasites and their hosts [77]. In previous studies, it has been reported that the Holoparasitic Orobanchaceae species, including *Christisonia kwangtungensis* [65], *Boschniakia rossica* and *B. himalaica* [15] have undergone transfer of chloroplast and/or mitochondrial genes, but while the hemiparasitic *C. paramensis* [78], the holoparasitic *Cuscuta* species [12], and *Aeginetia indica* [13] have not undergone such transfer. Our phylogenetic analysis indicates that *Cymbaria* consistently clusters with other species of Lamiales, except for a few genes (*rpl22*, *rps3*, and *ycf2*) that exhibit low bootstrap support for their affiliations with potential host species, namely *Allium mongolicum* (BS = 70%), *Leymus chinensis* (BS = 62%), and *Saposhnikovia divaricata* (BS = 48%), respectively. The minimal HGT observed in *Cymbaria* likely stems from its facultative root hemiparasitic nature and recently-evolved parasitic association with hosts. As a facultative parasite, spatial (e.g., apoplastic DNA degradation) and temporal (asynchronous life stages) barriers restrict horizontal transfer. On the other hand, phylogenetic divergence from hosts may render transferred genes non-functional, while evolutionary time frames (recent

parasitism or ancient barriers) further limit detectable HGT events. In other words, if the parasitic association is recent, insufficient time may have elapsed for HGT to occur and become fixed. Conversely, ancient associations might indicate long-term transfers. It has been found that there exists HGT between *Christisonia kwangtungensis* and its probable historical hosts [65]. Intriguingly, the nuclear genome of *A. indica* [79] and *Cuscuta* species [80] have acquired several genes from their hosts, indicating that host-to-parasite HGTs might vary between mitogenomes and nuclear genomes. This underscores the need for further investigation into HGT patterns in hemiparasitic *Cymbaria* nuclear genomes.

Conclusions

In this study, we have characterized the first complete mitogenomes of the hemiparasitic tribe Cymbarieae within the Orobanchaceae family. The mitogenomes of the hemiparasitic *C. mongolica* and *C. daurica* exhibit gene compositions that are typical of autotrophic flowering plants. Notably, large repeats have been validated by PCR amplification and Sanger sequencing that confirm the mediate homologous recombination, thereby contributing to the larger size and pentacyclic structure of these two mitogenomes. Phylogenetic analyses revealed that the two *Cymbaria* species form an independent sister clade to all other parasitic Orobanchaceae lineages. Additionally, the *Cymbaria* mitogenome demonstrated little to no intracellular and horizontal gene transfer, which is probably due to its facultative root hemiparasitic nature and recently-evolved parasitic association with hosts. These findings provide critical insights for studying the evolution of hemiparasitic plants within Orobanchaceae.

Abbreviations

PCGs	Protein-coding genes
rRNA	Ribosomal RNA genes
tRNA	Transfer RNA genes
SSRs	Simple sequence repeats
RSCU	Relative synonymous codon usage
ML	Maximum likelihood
BI	Bayesian Inference
MTPTs	Mitochondrial plastid sequences
IGT	Intracellular gene transfer
HGT	Horizontal gene transfer

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11474-4.

Additional file 1: Figure S1. Graphical mitogenome assembly of Cymbaria mongolica (A) and C. daurica (B) with possible contigs (color lines) and connections (black lines). Figure S2. Codon usage bias in mitochondrial PCGs of Cymbaria mongolica (A) and C. daurica (B). Figure S3. Colinear and dot plot analysis of the mitogenomes of two Cymbaria species and five closely related species. Figure S4. The single-gene ML trees constructed using plastid-derived PCGs. Figure S5. The single-gene ML trees constructed using mitochondrial PCGs.

Additional file 2: Table S1. Gene annotation of the two Cymbaria mitogenomes. Table S2. SSRs identified in the two Cymbaria mitogenomes. Table S3. Tandem repeats identified in the two Cymbaria mitogenomes. Table S4. Dispersed repeats identified in the two Cymbaria mitogenomes. Table S5. Relative synonymous codon usage (RSCU) in the two Cymbaria mitogenomes. Table S6. RNA editing sites predicted in mitochondrial PCGs of the two Cymbaria species. Table S7. Selective pressure and nucleotide diversity calculated in the shared PCGs between the two Cymbaria mitogenomes. Table S8. Large repeats identified in the two Cymbaria mitogenomes. Table S9. Primers used to detect potential minor conformations. Table S10. Species used in Lamiales phylogenetic analysis. Table S11. The homologous DNA fragments identified between the mitogenome and chloroplast genome of two Cymbaria species. Table S12. The lost/Pesudogenes identified between the two Cymbaria mitogenomes and three reference chloroplast genomes. Table S13. The Pesudogene ndhH identified between the Cymbaria daurica mitogenome and three reference chloroplast genomes. Table S14. Genbank accession numbers for analysis of intracellular gene transfer. Table S15. Genbank accession numbers for analysis of horizontal gene transfer.

Additional file 3: Validation of repeat mediated recombination in the two Cymbaria mitogenomes, including original agarose gel electrophoresis, reference sequence, sanger sequencing results, and sequence alignment.

Acknowledgements

Not applicable.

Authors' contributions

YM collected samples, analyzed data, conducted experiments, and drafted the manuscript; JLP and DQY reviewed the manuscript; ZZ and ZKD collected samples; JMN conceived the study. All authors read and approved the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (grant no. 31860106, 32260304), the Major Science and Technology Projects of Inner Mongolia Autonomous Region (grant no. 2019ZD008), and the Chinese Scholarship Council.

Data availability

The raw sequencing data for the Illumina and Nanopore platforms and the annotated mitogenome sequences have been deposited in NCBI (https://www.ncbi.nlm.nih.gov/) with accession numbers PRJNA1174065, SAMN44326656–SAMN44326657, SRR31045243–SRR31045246, OR698879–OR698888, respectively.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]School of Ecology and Environment, Inner Mongolia University, Hohhot 010020, People's Republic of China. [2]Botanic Institute of Barcelona (IBB), CSIC-CMCNB, Barcelona 08038, Spain. [3]Escuela de Ciencias Ambientales, Universidad Espíritu Santo (UEES), Samborondón 091650, Ecuador. [4]Ministry of Education Key Laboratory of Ecology and Resource Use of the Mongolian Plateau, Hohhot 010020, People's Republic of China. [5]Inner Mongolia Key Laboratory of Grassland Ecology and the Candidate State Key Laboratory of Ministry of Science and Technology, Hohhot 010020, People's Republic of China.

Received: 15 November 2024 Accepted: 12 March 2025
Published online: 31 March 2025

References

1. Birky CW. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc Natl Acad Sci USA. 1995;92(25):11331–8.
2. Skippington E, Barkman TJ, Rice DW, Palmer JD. Miniaturized mitogenome of the parasitic plant *Viscum scurruloideum* is extremely divergent and dynamic and has lost all *nad* genes. Proc Natl Acad Sci USA. 2015;112(27):E3515–24.
3. Putintseva YA, Bondar EI, Simonov EP, Sharov VV, Oreshkova NV, Kuzmin DA, et al. Siberian larch (Larix sibirica Ledeb.) mitochondrial genome assembled using both short and long nucleotide sequence reads is currently the largest known mitogenome. BMC Genomics. 2020;21:654.
4. Wynn EL, Christensen AC. Repeats of unusual size in plant mitochondrial genomes: identification, incidence and evolution. G3 Genes|Genomes|Genetics. 2019;9(2):549–59.
5. Park S, Grewe F, Zhu A, Ruhlman TA, Sabir J, Mower JP, et al. Dynamic evolution of *Geranium* mitochondrial genomes through multiple horizontal and intracellular gene transfers. New Phytol. 2015;208(2):570–83.
6. Bergthorsson U, Adams KL, Thomason B, Palmer JD. Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature. 2003;424(6945):197–201.
7. Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, Michelmore RW, et al. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. PLoS Genet. 2019;15(8): e1008373.
8. Li J, Li J, Ma Y, Kou L, Wei J, Wang W. The complete mitochondrial genome of okra (*Abelmoschus esculentus*): Using nanopore long reads to investigate gene transfer from chloroplast genomes and rearrangements of mitochondrial DNA molecules. BMC Genomics. 2022;23:481.
9. Shan Y, Li J, Zhang X, Yu J. The complete mitochondrial genome of *Amorphophallus albus* and development of molecular markers for five *Amorphophallus* species based on mitochondrial DNA. Front Plant Sci. 2023;14: 1180417.
10. Qian H, Zhang J, Zhao J. How many known vascular plant species are there in the world? An integration of multiple global plant databases. Biodiv Sci. 2022;30(7):22254.
11. Lin Y, Li P, Zhang Y, Akhter D, Pan R, Fu Z, et al. Unprecedented organelle genomic variations in morning glories reveal independent evolutionary scenarios of parasitic plants and the diversification of plant mitochondrial complexes. BMC Biol. 2022;20:49.
12. Anderson BM, Krause K, Petersen G. Mitochondrial genomes of two parasitic *Cuscuta* species lack clear evidence of horizontal gene transfer and retain unusually fragmented *ccmFC* genes. BMC Genomics. 2021;22:816.
13. Zhong Y, Yu R, Chen J, Liu Y, Zhou R. Highly active repeat–mediated recombination in the mitogenome of the holoparasitic plant *Aeginetia indica*. Front Plant Sci. 2022;13: 988368.
14. Choi KS, Park S. Complete plastid and mitochondrial genomes of *Aeginetia indica* reveal intracellular gene transfer (IGT), horizontal gene transfer (HGT), and cytoplasmic male sterility (CMS). Int J Mol Sci. 2021;22(11): 6143.
15. Zhang J, Huang Z, Fu W, Zhang C, Zan T, Nan P, et al. Host shift promotes divergent evolution between closely related holoparasitic species. Mol Phylogenet Evol. 2023;186: 107842.
16. Skippington E, Barkman TJ, Rice DW, Palmer JD. Comparative mitogenomics indicates respiratory competence in parasitic *Viscum* despite loss of complex I and extreme sequence divergence, and reveals horizontal gene transfer and remarkable variation in genome size. BMC Plant Biol. 2017;17:49.
17. Yu R, Sun C, Zhong Y, Liu Y, Sanchez-Puerta MV, Mower JP, et al. The minicircular and extremely heteroplasmic mitogenome of the holoparasitic plant Rhopalocnemis phalloides. Curr Biol. 2022;32(2):470–9.
18. Petersen G, Anderson B, Braun HP, Meyer EH, Møller IM. Mitochondria in parasitic plants. Mitochondrion. 2020;52:173–82.
19. Westwood J, dePamphilis CW, Das M, Fernández-Aparicio M, Honaas L, Timko M, et al. The parasitic plant genome project: New tools for understanding the biology of *Orobanche* and *Striga*. Weed Sci. 2012;60(2):295–306.
20. Jiang N, Dong L, Yang J, Tan Y, Wang H, Randle C, et al. Herbarium phylogenomics: resolving the generic status of the enigmatic *Pseudobartsia* (Orobanchaceae). J Syst Evol. 2022;60(5):1218–28.
21. Schneeweiss GM. Phylogenetic relationships and evolutionary trends in Orobanchaceae. Parasitic Orobanchaceae. Berlin, Heidelberg: Springer; 2013. p. 243–65.
22. Bennett JR, Mathews S. Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A. Am J Bot. 2006;93(7):1039–51.
23. McNeal JR, Bennett JR, Wolfe AD, Mathews S. Phylogeny and origins of holoparasitism in Orobanchaceae. Am J Bot. 2013;100(5):971–83.
24. Xu Y, Zhang J, Ma C, Lei Y, Shen G, Jin J, et al. Comparative genomics of orobanchaceous species with different parasitic lifestyles reveals the origin and stepwise evolution of plant parasitism. Mol Plant. 2022;15(8):1384–99.
25. Li X, Feng T, Randle C, Schneeweiss GM. Phylogenetic relationships in Orobanchaceae inferred from low–copy nuclear genes: Consolidation of major clades and identification of a novel position of the non–photosynthetic *Orobanche* clade sister to all other parasitic Orobanchaceae. Front Plant Sci. 2019;10: 902.
26. Hong DY, Yang HB, Jin CL, Holmgren NH. Scrophulariaceae. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China vol 18: Scrophulariaceae through Gesneriaceae: Science Press. Beijing, China and Missouri Botanical Garden Press, St. Lours, Missouri: USA; 1998. p. 1–212.
27. Zhao Y. The species research and its floristic analysis of plants of *Cymbaria*. Journal of Inner Mongolia University (Natural Science Edition). 1999;30(3):351–3.
28. Fischer E. Scrophulariaceae. In: Flowering plants· dicotyledons: Lamiales (except Acanthaceae including Avicenniaceae). Springer; 2004. p. 333–432.
29. Ma Y, López-Pujol J, Yan D, Zhou Z, Deng Z, Niu J. Complete chloroplast genomes of the hemiparasitic genus *Cymbaria*: Insights into comparative analysis, development of molecular markers, and phylogenetic relationships. Ecol Evol. 2024;14(7):e11677.
30. Allen GC, Flores-Vergara M, Krasynanski S, Kumar S, Thompson W. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nat Protoc. 2006;1(5):2320–5.
31. Ruan J, Li H. Fast and accurate long-read assembly with wtdbg2. Nat Methods. 2020;17(2):155–8.
32. Chen Y, Ye W, Zhang Y, Xu Y. High speed BLASTN: an accelerated MegaBLAST search tool. Nucleic Acids Res. 2015;43(16):7762–8.
33. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 2014;9(11): e112963.
34. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 2017;27(5):737–46.
35. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: Interactive visualization of de novo genome assemblies. Bioinformatics. 2015;31(20):3350–2.
36. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq–versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45(W1):W6–11.
37. Chan PP, Lin BY, Mak AJ, Lowe TM. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Research. 2021;49(16):9077–96.
38. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Research. 2019;47(W1):W59–64.
39. Chen C, Wu Y, Li J, Wang X, Zeng Z, Xu J, et al. TBtools–II: A "one for all, all for one" bioinformatics platform for biological big-data mining. Mol Plant. 2023;16(11):1733–42.
40. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: A web server for microsatellite prediction. Bioinformatics. 2017;33(16):2583–5.
41. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80.
42. Kurtz S, Choudhuri J, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: The manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29(22):4633–42.
43. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li W, et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 2020;20(1):348–55.

44. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England). 2010;26(5):589–95.
45. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078–879.
46. Yang Z. PAML 4: Phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24(8):1586–91.
47. Katoh K, Standley D. MAFFT Multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.
48. Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268–74.
49. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 2012;29(8):1969–73.
50. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6.
51. Darling A, Mau B, Perna N. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 2010;5: e11147.
52. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019;20(4):1160–6.
53. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–3.
54. Mower JP. Variation in protein gene and intron content among land plant mitogenomes. Mitochondrion. 2020;53:203–13.
55. Kumar KR, Cowley MJ, Davis RL: Next-generation sequencing and emerging technologies. In: Seminars in thrombosis and hemostasis: 2024. Thieme Medical Publishers: 1026–1038. p. 1026-1038.
56. Petersen G, Cuenca A, Møller I, Seberg O. Massive gene loss in mistletoe (*Viscum*, Viscaceae) mitochondria. Sci Rep. 2015;5(1):17588.
57. Asaf S, Khan AL, Khan AR, Waqas M, Kang SM, Khan MA, et al. Mitochondrial genome analysis of wild rice (*Oryza minuta*) and its comparison with other related species. PLoS One. 2016;11(4): e0152937.
58. Unseld M, Marienfeld JR, Brandt P, Brennicke A. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. Nat Genet. 1997;15(1):57–61.
59. Lin Q, Banerjee A, Stefanović S. Mitochondrial phylogenomics of Cuscuta (Convolvulaceae) reveals a potentially functional horizontal gene transfer from the host. Genome Biology and Evolution. 2022;14(6):evac091.
60. Kleine T, Maier UG, Leister D. DNA Transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Annu Rev Plant Biol. 2009;60:115–38.
61. Zhou P, Zhang Q, Li F, Huang J, Zhang M. Assembly and comparative analysis of the complete mitochondrial genome of Ilex metabaptista (Aquifoliaceae), a Chinese endemic species with a narrow distribution. BMC Plant Biol. 2023;23(1):1–23.
62. Liu D, Qu K, Yuan Y, Zhao Z, Chen Y, Han B, Li W, El-Kassaby YA, Yin Y, Xie X. Complete sequence and comparative analysis of the mitochondrial genome of the rare and endangered Clematis acerifolia, the first clematis mitogenome to provide new insights into the phylogenetic evolutionary status of the genus. Front Genet. 2023;13: 1050040.
63. Sloan DB. One ring to rule them all? Genome sequencing provides new insights into the 'master circle' model of plant mitochondrial DNA structure. New Phytol. 2013;200(4):978–85.
64. Miao Y, Chen H, Xu W, Liu C, Huang L. *Cistanche* species mitogenomes suggest diversity and complexity in lamiales–order mitogenomes. Genes. 2022;13(10): 1791.
65. Zhang C, Lin Q, Zhang J, Huang Z, Nan P, Li L, et al. Comparing complete organelle genomes of holoparasitic *Christisonia kwangtungensis* (Orobanchaceae) with its close relatives: How different are they? BMC Plant Biol. 2022;22:444.
66. Ruhlman TA, Zhang J, Blazier JC, Sabir JS, Jansen RK. Recombination-dependent replication and gene conversion homogenize repeat sequences and diversify plastid genome structure. Am J Bot. 2017;104(4):559–72.
67. Varré JS, d'Agostino N, Touzet P, Gallina S, Tamburino R, Cantarella C, et al. Complete sequence, multichromosomal architecture and transcriptome analysis of the *Solanum tuberosum* mitochondrial genome. Int J Mol Sci. 2019;20(19): 4788.
68. You C, Cui T, Zhang C, Zang S, Su Y, Que Y. Assembly of the complete mitochondrial genome of *Gelsemium elegans* revealed the existence of homologous conformations generated by a repeat mediated recombination. Int J Mol Sci. 2022;24(1): 527.
69. Zhao J, Chen J, Xiong Y, He W, Xiong Y, Xu Y, et al. Organelle genomes of *Indigofera amblyantha* and *Indigofera pseudotinctoria*: comparative genome analysis, and intracellular gene transfer. Ind Crops Prod. 2023;198: 116674.
70. 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;24:466.
71. Christensen AC. Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. Genome Biol Evol. 2013;5(6):1079–86.
72. Wicke S, Müller KF, de Pamphilis CW, Quandt D, Wickett NJ, Zhang Y, et al. Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. Plant Cell. 2013;25(10):3711–25.
73. Zhao N, Wang Y, Hua J. The roles of mitochondrion in intergenomic gene transfer in plants: A source and a pool. Int J Mol Sci. 2018;19(2): 547.
74. Filip E, Skuza L. Horizontal gene transfer involving chloroplasts. Int J Mol Sci. 2021;22(9): 4484.
75. Cusimano N, Wicke S. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. New Phytol. 2015;210(2):680–93.
76. Chen J, Yu R, Dai J, Liu Y, Zhou R. The loss of photosynthesis pathway and genomic locations of the lost plastid genes in a holoparasitic plant *Aeginetia indica*. BMC Plant Biol. 2020;20:199.
77. Wang J, Kan S, Liao X, Zhou J, Tembrock LR, Daniell H, Jin S, Wu Z. Plant organellar genomes: much done, much more to do. Trends Plant Sci. 2024;29(7):754–69.
78. Fan W, Zhu A, Kozaczek M, Shah N, Pabón-Mora N, González F, et al. Limited mitogenomic degradation in response to a parasitic lifestyle in Orobanchaceae. Sci Rep. 2016;6(1): 36285.
79. Kado T, Innan H. Horizontal gene transfer in five parasite plant species in Orobanchaceae. Genome Biol Evol. 2018;10(12):3196–210.
80. Yang Z, Wafula EK, Kim G, Shahid S, McNeal JR, Ralph PE, et al. Convergent horizontal gene transfer and cross-talk of mobile nucleic acids in parasitic plants. Nature Plants. 2019;5(9):991–1001.

Publisher's Note