BMC Genomics

RESEARCH Open Access

Mitogenome of *Uncaria rhynchophylla*: genome structure, characterization, and phylogenetic relationships

Lingjian Gui[1,2,3], Zhanjiang Zhang[1,2,3,4], Lisha Song[1,2,3], Changcun Feng[5], Haixia Yu[1,2,3], Limei Pan[1,2,3], Jine Fu[4], Wenjing Liang[1,2,3], Qiulan Huang[6], Ahmed H. El-Sappah[6], Lijun Shi[1,2,3*], Lingyun Wan[1,2,3*] and Shugen Wei[1,2,3*]

Abstract

Background *Uncaria rhynchophylla* is listed in the Chinese pharmacopoeia as one of the five botanical sources of the traditional medicine Gou-Teng, which has been utilized for the treatment of mental and cardiovascular disorders. This particular species is well-known in China for its hook-like structures originating from the leaf axils. Despite available reports on its chloroplast genome, there persists a notable lack of understanding concerning the structural variations and evolution of its mitochondrial genome. This knowledge gap hinders our ability to fully comprehend its genetic attributes.

Results We successfully assembled the mitochondrial genome of *U. rhynchophylla* by seamlessly integrating Illumina short reads with Nanopore long reads, resulting in a non-circular genome comprising 1 circular contig and 2 linear contigs. The total length of this genome is 421,660 bp, encompassing 36 PCGs. The identification of 4 distinct pairs of repeats has unveiled their pivotal role in repeat-mediated recombination. Of the 28 homologous fragments derived from chloroplasts, the majority were observed to have been transferred from the inverted repeat (IR) regions of the chloroplast genome to the mitochondrial genome. The mitochondrial DNA provides a distinctive resolution for the positioning of several species within the Gentianales phylogenetic framework, which remains unresolved by chloroplast DNA.

Conclusion By utilizing a newly assembled, high-quality mitochondrial genome of *U. rhynchophylla*, we have elucidated its intricate genomic structure, distinctive sequence characteristics, and potential for phylogenetic analysis. These findings mark significant strides in advancing our comprehension of the genetics of *Uncaria*.

Keywords *Uncaria*, Mitochondrial genome, Repeat-mediated recombination, Gene transfer, Phylogeny

*Correspondence:
Lijun Shi
slj125@126.com
Lingyun Wan
wanlingyun1984@163.com
Shugen Wei
weishugen2@163.com
[1]Guangxi Key Laboratory of High-Quality Formation and Utilization of Dao-di Herbs, Guangxi Botanical Garden of Medicinal Plants, Nanning 530023, China
[2]Guangxi Traditional Chinese Medicine Breeding Technology Innovation Center, Guangxi Botanical Garden of Medicinal Plants, Nanning 530023, China
[3]National Center for Traditional Chinese Medicine (TCM) Inheritance and Innovation, Guangxi Botanical Garden of Medicinal Plants, Nanning 530023, China
[4]National Engineering Research Center for the Development of Southwestern Endangered Medicinal Materials, Guangxi Botanical Garden of Medicinal Plants, Nanning 530023, China
[5]Guangxi Medical University, Nanning 530021, China
[6]Faculty of Agriculture, Forestry and Food Engineering, Yibin University, Yibin 644000, China

Background

The genus *Uncaria* Schreber, a member of the Rubiaceae family, encompasses approximately 34 species worldwide [1]. Of these, 12 species are endemic to various regions of China, including Guangxi, Guangdong, Guizhou, Yunnan, Hunan, and Sichuan [2]. These evergreen, perennial vines thrive in sparse forests and shrubs adjacent to valleys and streams. To adapt to their environment, *Uncaria* species have evolved hook-like structures originating from the leaf axils, specifically adapted for climbing growth (Fig. 1). During the phenological shift from the vegetative to the reproductive phase, certain stipular hooks undergo a functional metamorphosis, transitioning into inflorescence peduncles. Many *Uncaria* species have been used in traditional Chinese medicines, with *U. rhynchophylla* officially documented in the Chinese Pharmacopoeia as one of five botanical sources for the traditional medicine Gou-Teng [3]. The primary active constituents of *U. rhynchophylla* are indole alkaloids, such as rhynchophylline, isorhynchophylline, corynoxeine, and isocorynoxeine, which are renowned for their efficacy in treating convulsions, hypertension, fever, epilepsy, and eclampsia [4].

Mitochondria are semi-autonomous organelles present in nearly all eukaryotic cells. Their primary function is to produce energy by synthesizing adenosine triphosphate (ATP) [5]. Due to their vital role in converting the potential energy stored in biomolecules into the chemical energy that sustains cellular life, mitochondria are aptly termed the "powerhouses" of cells [6]. In flowering plants, nuclear genetic information exhibit biparental inheritance, whereas mitochondrial and chloroplast (cp.) DNA are typically maternally inherited [7]. This uniparental inheritance pattern facilitates the integration of organellar genomes with nuclear genes, providing a valuable tool for analyzing reticulate evolutionary histories. Despite the relatively slow evolutionary rates of conserved mitochondrial genes, plant mitochondrial genomes possess the capacity to undergo remarkably rapid rearrangements [8–10], and exhibit extensive size variation, ranging from 66 kb [11] to 11.3 Mb [12]. The annotation of plant mitochondrial genomes frequently uncovers novel genes, some of which significantly impact the plant's growth cycle and fertility [13]. Additionally, the evolutionary dynamics of plant mitochondrial genomes are significantly shaped by frequent horizontal gene transfer events, particularly from chloroplast to mitochondrial DNA [14, 15].

Mitochondrial genome rearrangements are believed to originate from recombination events mediated by repetitive sequences [16, 17]. The presence of direct (forward) and inverted (palindromic) repeats generates structural plasticity in mitochondrial genome conformations. Accumulating evidence challenges the canonical circular model of angiosperm mitochondria, revealing instead its existence as dynamic multimeric complexes. For instance, *Viscum scurruloideum* exhibits a bipartite mitochondrial genome comprising circular and linear contigs [11], while *Populus simonii* maintains three distinct circular subgenomes [18]. Medium-length repeats (100-1,000 bp) are considered common in recombination events [17, 19], whereas large repeats (> 1000 bp) are anticipated to yield almost molar or near-molar recombinant molecules in certain taxa [19, 20]. To date, the recombination dynamics and structural organization of the *U. rhynchophylla* mitochondrial genome remain unexplored.

The inherent structural complexity of mitochondrial genomes presents substantial challenges for de novo assembly using short-read sequencing data [21–23]. Recent advances in long-read sequencing technologies enable high-fidelity mitochondrial genome assemblies in plants, significantly improving repeat element resolution while effectively mitigating PCR-derived artifacts [24]. Although cp. genome analyses of *Uncaria* species have been documented [25], a significant knowledge gap remains regarding the structural variations and evolutionary trajectory of their mitochondrial genomes. Here, we present the first complete mitochondrial genome assembly of *U. rhynchophylla*, achieved through hybrid sequencing integrating Illumina short reads and Nanopore long reads. This genomic resource establishes a critical foundation for comparative mitogenomics and evolutionary studies within the *Uncaria* genus.

Materials and methods

Sampling and sequencing

The genetic material for data extraction was sourced from *U. rhynchophylla* plants cultivated at the experimental research base of Guangxi Botanical Garden of Medicinal Plants, located at coordinates 22°51' N, 108°23' E in Nanning, Guangxi, China. Voucher specimens (W20200117) have been deposited in the herbarium of Guangxi Botanical Garden of Medicinal Plants (GXMG). Taxonomic identification of the plant was conducted by Shugen Wei. The mitochondrial genome of *U. rhynchophylla* was obtained as part of our ongoing genome project on this species. Genome sequencing was performed using both the Illumina and Nanopore platforms by Benagen Tech Solutions Company Limited (Wuhan, China).

Assembly and annotation of the mitochondrial genome

We employed two strategies to assemble the mitochondrial genome. Firstly, we utilized Unicycler [26] for hybrid assembly, combining Illumina short reads and Nanopore long reads with the parameters set to --kmers 27, 43, 57, 67, 77, 89. In this process, we initially assembled the mapped Illumina short reads using SPAdes [27]. Subsequently, we employed Nanopore long reads to resolve

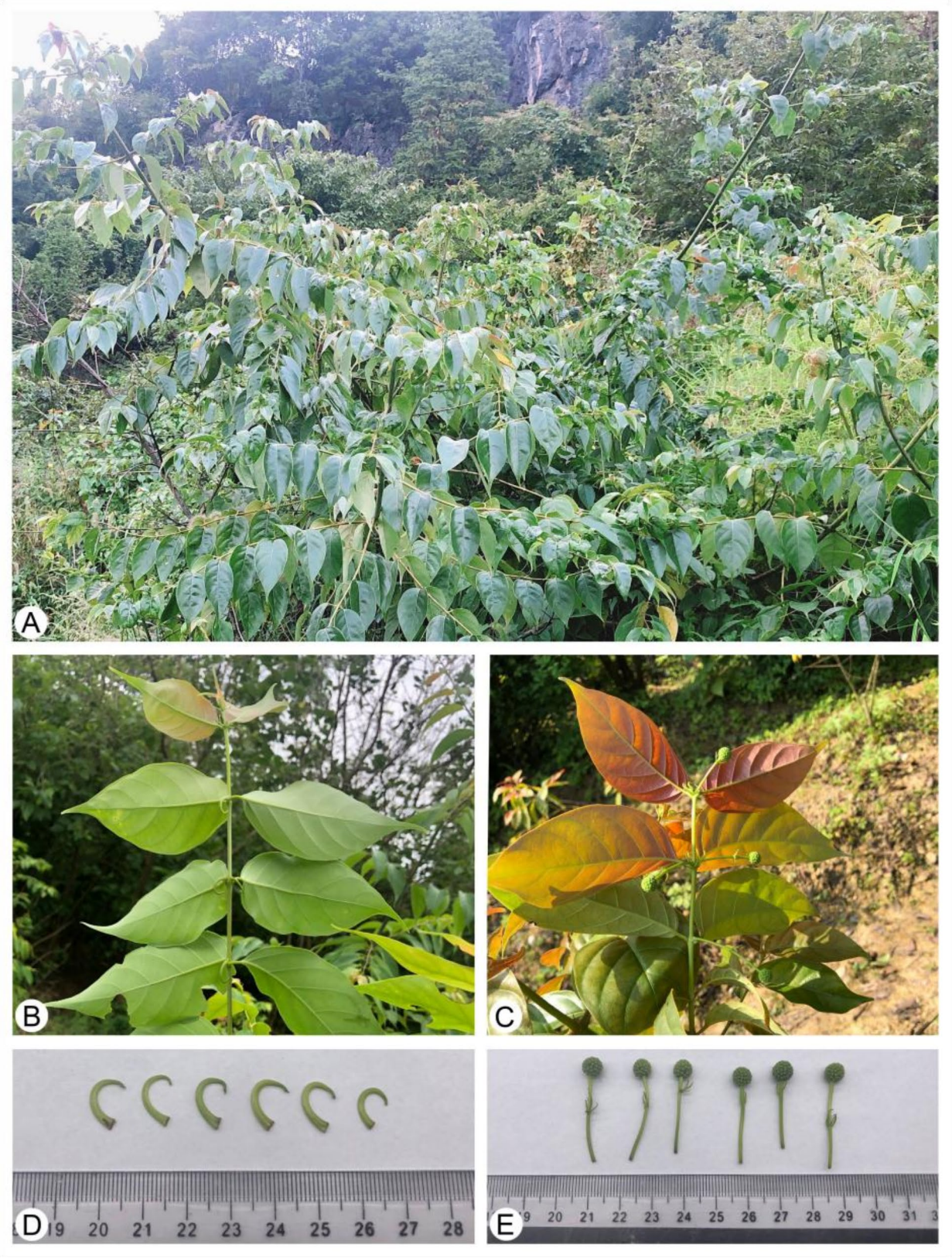

Fig. 1 Ecological habitat and morphological characteristics of *U.rhynchophylla*. (**A**) Habitat in the wild; (**B**) Branches with hooks; (**C**) Branches with inflorescences; (**D**) Detailed morphology of the hooks; (**E**) Detailed morphology of the inflorescences

the repetitive sequence regions of the assembly (Fig. S1) using Minimap2 [28]. The resulting GFA format files generated by Unicycler [26] were then visualized using Bandage [29]. Secondly, we performed de novo assembly of Nanopore long reads using Flye (v.2.9.1-b1780) with the parameter --min-overlap 2000. To identify the draft mitogenome, we created a BLASTn [30] database from the Flye-assembled contigs and queried it with conserved mitochondrial genes from *Arabidopsis thaliana* (NC_037304) and *Liriodendron tulipifera* (NC_021152.1) [31]. Contigs with significant BLAST hits to mitochondrial genes were retained. Finally, we mapped Illumina sequences to the assembled genome and confirmed uniform coverage with a mean depth exceeding 392X, indicating high assembly quality (Fig. S2).

The mitochondrial protein-coding genes (PCGs) were annotated using Geseq [32], referencing the well-annotated mitochondrial genomes of *Arabidopsis thaliana* and *Liriodendron tulipifera*. The tRNA genes were annotated using tRNAscan-SE [33], while the rRNA genes were annotated using BLASTn [30]. For cp. genome assembly and annotation, we employed GetOrganelle v1.7.5 [34], followed by structural and functional annotation using CPGAVAS2 [35]. All automated annotations were manually curated and validated through Apollo v1.6.5 [36]. The newly annotated sequences were submitted to GenBank and assigned the accession numbers OR051899-OR051902.

Codon usage and RNA editing analysis

The PCGs of the mitochondrial genome were extracted using Phylosuite v1.2.3 [37]. Subsequently, a codon usage bias analysis was conducted on these PCGs using Mega7 [38], which facilitated the calculation of relative synonymous codon usage (RSCU) values. For the prediction of RNA editing events, we utilized the online tool PREPACT (http://www.prepact.de/) [39], setting the cutoff value at 0.001.

Repeat sequence and repeat-mediated recombination

We identified repeat sequences using MISA [40], TRF [41], and REPuter [42]. The results were visualized using Circos [43]. To investigate repeat-mediated recombination, we extracted sequences of 1,000 bp upstream and downstream of each repeat unit (labeled as a-d). These flanking sequences, together with the repeat sequences, constituted the reference sequences. Subsequently, we simulated recombination events by swapping the 1,000 bp sequence on one side of the repeat units. Reference sequences representing potential recombination paths were generated. Nanopore long reads that fully covered these reference sequences were searched for. These long reads served as evidence supporting the corresponding recombination paths.

Gene transfer and collinearity analysis

In evolutionary genomics, cp. DNA sequences can undergo interorganellar translocation and subsequent integration into mitochondrial genomes. These transferred fragments often exhibit variability in length and sequence similarity across different species. To gain a deeper understanding of this process, we conducted a comprehensive analysis of homologous fragments between the cp. and mitochondrial genomes using BLASTn [30] with stringent parameters (-evalue 1e-5 -wordsize 7 -outmt 6). We filtered the results to retain only those with an identity greater than 80%. Utilizing MCscanX [44], we generated multiple synteny maps for these species, enabling the visualization and analysis of gene order conservation and collinearity. The species and their corresponding NCBI numbers used in this analysis are listed in Table S1.

Phylogenetic analysis

A total of 27 complete mitochondrial genomes from Gentianales species and outgroups were analyzed. To facilitate a comparative assessment of the phylogenetic relationships among diverse DNA sequences, we additionally downloaded the cp. genomes of the respective species. We used PhyloSuite v1.2.3 [37] to extract homologous and conserved genes from both the mitochondrial and chloroplast genomes. Sequence alignment was then performed using MAFFT [45]. After identifying the most suitable substitution model with ModelFinder [46], we conducted a maximum likelihood (ML) analysis with IQ-TREE [47] and performed Bayesian inference (BI) analysis using MrBayes v3.2 [48]. The resulting phylogenetic trees were visualized and presented using ITOL [49].

Results

Mitochondrial genome features

We identified three contigs, comprising 2 linear contigs and 1 circular contig. These contigs were conventionally designated as Chromosome 1, Chromosome 2, and Chromosome 3. Together, these sequences constitute a non-circular genomic fragment assembly, representing the complete mitochondrial genome of the organism (Fig. 2, Fig. S1, Table S2). The total length of the mitochondrial genome is 421,660 bp, with a GC content of 44.69% (Table 1).

The mitochondrial genome has been fully annotated, revealing a total of 36 PCGs, comprising 24 core genes and 12 non-core genes. The core genes encode essential components of the mitochondrial respiratory chain and related functions, including 5 ATP synthase subunits (*atp1*, *atp4*, *atp6*, *atp8*, *and atp9*), 9 NADH dehydrogenase subunits (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *and nad9*), 4 cytochrome c biogenesis proteins (*ccmB*, *ccmC*, *ccmFC*, *and ccmFN*), 3 cytochrome

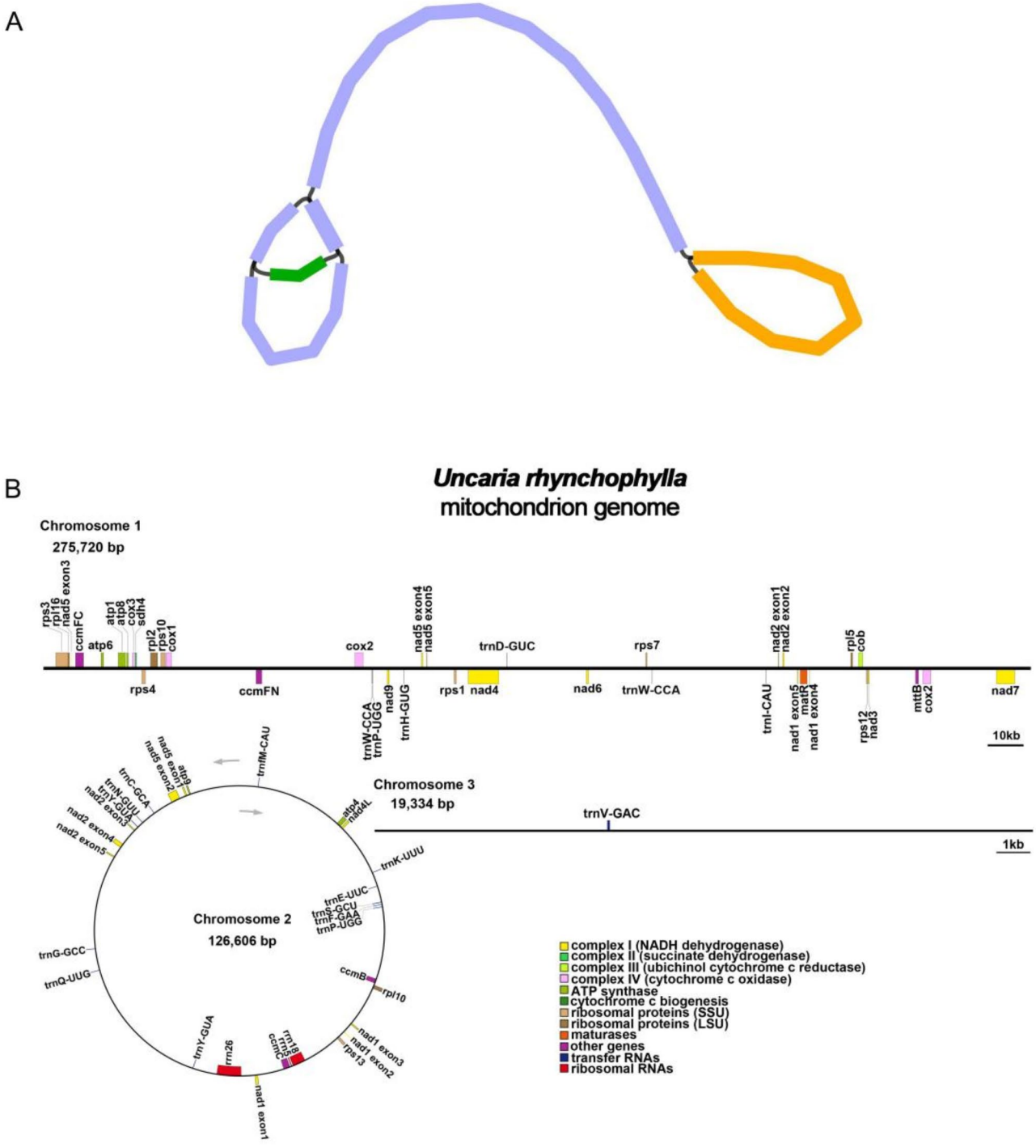

Fig. 2 Mitochondrial genome of *U. rhynchophylla*. (**A**) The graphical fragment assembly of the mitochondrial genome. Chromosome 1, Chromosome 2, and Chromosome 3 are illustrated as distinct contigs, represented by blue, orange, and green lines, respectively; (**B**) Gene annotation map of the mitochondrial genome

c oxidase subunits (*cox1*, *cox2*, *and cox3*), 1 membrane transport protein (*mttB*), 1 maturase-related protein (*matR*), and 1 ubiquinol cytochrome c reductase (*cob*). The non-core genes primarily encode ribosomal proteins, including 4 large subunit proteins (*rpl2*, *rpl5*, *rpl10*, *and rpl16*), 7 small subunit proteins *(rps1*, *rps3*, *rps4*, *rps7*, *rps10*, *rps12*, *and rps13*), and 1 succinate dehydrogenase subunit (*sdh4*) (Table S3). Additionally, the annotation identified 16 tRNA genes (3 of which are duplicated) and 3 rRNA genes. Notably, chromosome 3 of the *U.*

Table 1 Summary of major characteristics of the *U. rhynchophylla* mitochondrial genome

NCBI Accession	Contigs	Type	Length (bp)	GC content (%)
OR051899-OR051901	Chromosome 1–3	branched	421,660	44.69
OR051899	Chromosome 1	linear	275,720	44.76
OR051900	Chromosome 2	circular	126,606	44.77
OR051901	Chromosome 3	linear	19,334	43.25

rhynchophylla mitochondrial genome is devoid of functional PCGs, containing only a single tRNA gene.

Codon usage analysis

Codon usage displays considerable diversity across various species and organisms, reflecting the long-term evolutionary pressures that have ultimately resulted in a balanced state within cells. Codons with an RSCU value greater than 1 are considered to be preferred for their corresponding amino acids. To investigate the codon usage bias within the mitochondrial genome, we analyzed the 36 distinct PCGs (Table S4). Our analysis revealed a general trend towards specific codon usage, with notable exceptions being the start codon AUG and the codon for tryptophan (UGG) (Fig. S3). Among these amino acids, alanine (Ala) exhibits a strong preference for codon GCU, with an RSCU value of 1.62, the highest observed in the mitochondrial PCGs. Similarly, tyrosine (Tyr) prefers the codon UAU, with an RSCU value of 1.54 (Table S4). In contrast, lysine (Lys) and phenylalanine (Phe) exhibit lower RSCU values, below 1.2, suggesting weaker codon usage preferences for these amino acids.

Repetitive sequence analysis

The mitochondrial genome of *U. rhynchophylla* harbors 215 repetitive sequences distributed across its three chromosomes (Table S5). Chromosome 1 contains the highest repeat density, with 87 simple sequence repeats (SSRs) identified, of which monomeric and dimeric SSRs account for 44.83% of the total. Among the monomeric SSRs, thymine (T) repeats predominate, comprising 61.90% (13 out of 21). Hexameric SSRs are absent on Chromosome 1. Ten tandem repeats, ranging from 15 to 24 bp in length with a match percentage exceeding 80%, were annotated. Additionally, 173 dispersed repeat pairs, each at least 30 bp long, were identified, including 88 palindromic and 85 forward repeats. No reverse or complementary repeats were detected (Fig. 3). The longest palindromic repeat spans 3,241 bp, while the longest forward repeat covers 3,489 bp.

Chromosome 2 exhibited reduced repeat complexity, with a total of 26 SSRs identified. Monomeric and dimeric SSRs accounted for 42.31% of the total SSRs, with adenine (A) monomeric repeats representing 66.67% (2 out of 3) of all monomeric SSRs. Six tandem repeats, ranging from 17 to 25 bp in length and showing a match percentage greater than 66%, were also detected. Additionally, 40 dispersed repeat pairs were identified, including 12 palindromic and 28 forward repeats (Fig. 3). The maximum lengths of the palindromic and forward repeats were 95 bp and 124 bp, respectively.

Chromosome 3 displayed minimal repeat content, containing only 7 SSRs, with 57.14% being monomeric or dimeric forms. Additionally, a single tandem repeat of 23 bp with an 86% match was identified. Dispersed repeat analysis revealed two pairs of repeats, each at least 30 bp, with one pair being palindromic and the other reverse (Fig. 3).

Repeat-mediated recombination

Apart from enlarging the mitochondrial genome, the presence of repeats fosters genomic recombination. Typically, recombination is mediated solely by forward and palindromic repeats. Our analysis identified 214 pairs of short repeats that have the potential to promote isomer formation (Fig. 4A). These repeats generate distinct genomic configurations: Path A1 and A2 represent the primary configurations, while Path B1 and B2 correspond to alternative configurations resulting from repeat-mediated recombinations (Fig. 4B).

Under stringent criteria, 4 pairs of repeats exhibited significant evidence of recombination activity and were designated as R1 to R4 (Table 2, Table S6). R1 comprises a pair of 3,241 bp palindromic repeats supported by 18 long reads for the primary conformation (56.25%) and 14 long reads for the alternative conformation (43.75%), while R3 exhibits a peak recombination rate of 49.46%, with 47 and 46 long read segments favoring the primary and alternative conformations, respectively. In contrast, the remaining two repeats, R2 and R4, show significantly lower recombination frequencies (< 4%), with their structural dynamics primarily favoring the maintenance of the primary conformation.

RNA editing events

We predicted RNA editing in 36 PCGs within the mitochondrial genome, identifying a total of 433 potential RNA editing sites. All these sites exclusively involve the substitution of cytosine (C) for uracil (U). The *mttB* gene exhibited the highest number of RNA editing events, with 37 sites, followed closely by the *ccmFN* gene, which had 34 events. In contrast, the *rpl2* gene showed only 1 predicted RNA editing event, while no RNA editing was predicted for the *rps3* gene (Fig. 5).

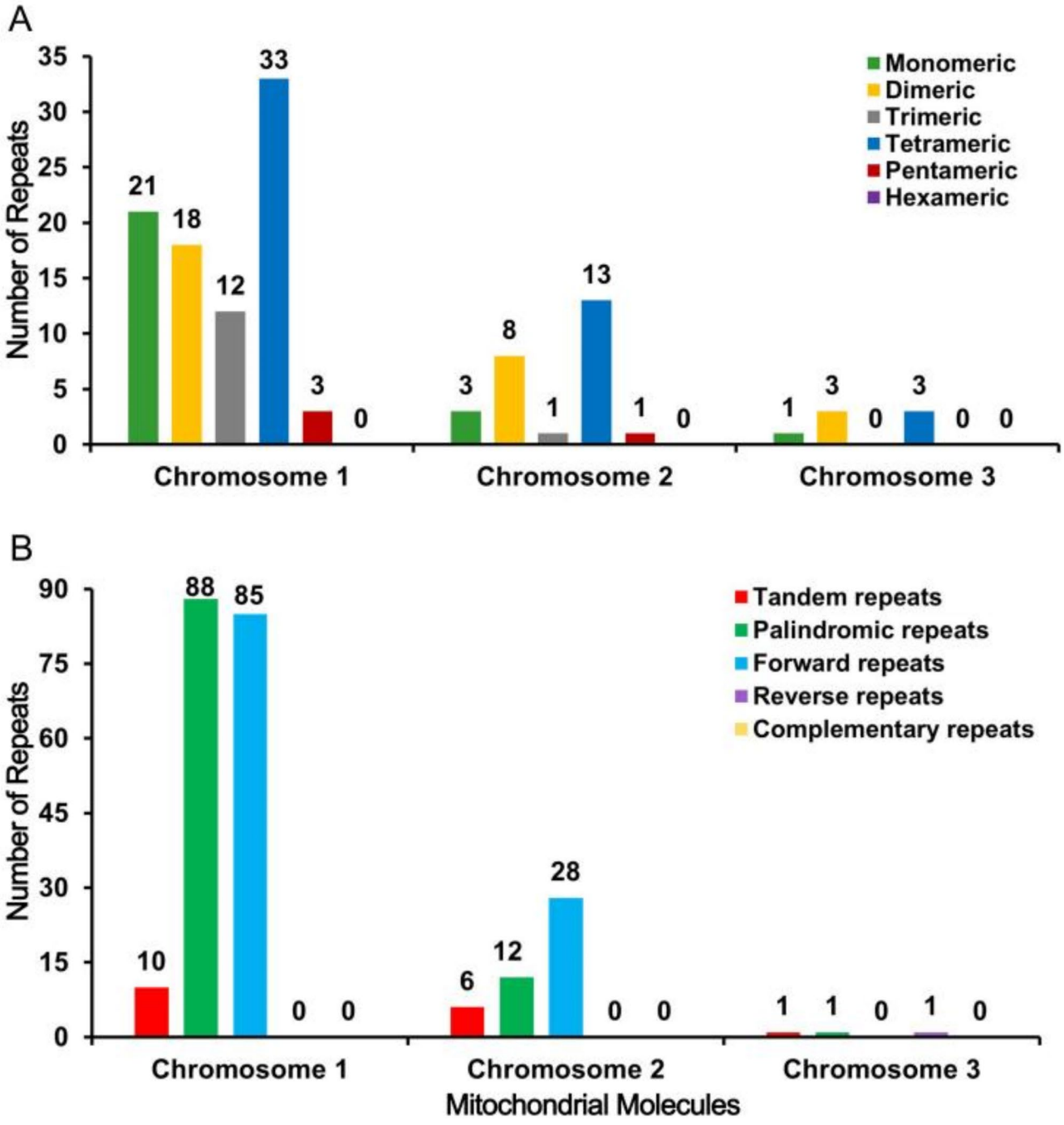

Fig. 3 Repeat sequence analysis of three mitochondrial chromosomes. (**A**) Statistics on SSRs in the mitochondrial genome; (**B**) Statistics on tandem repeats and dispersed repeats

Sequence transfer analysis

The sequence similarity analysis revealed 28 homologous mitochondrial plastid DNA (MTPT) fragments (Fig. 6A), totaling 21,961 bp and accounting for 5.21% of the mitochondrial genome. Intriguingly, the majority of MTPTs originate from the inverted repeat (IR) regions of the cp. genome, with 18 MTPTs constituting 85.12% of the entire transferred sequence length. While MTPT17 and MTPT18 represented the largest transferred segments (5,074 bp each), the remaining MTPTs ranged from 41 to 1,732 bp. Functional annotation revealed 8 intact transferred genes, including 2 PCGs (*petG* and *rpl23*) and 6 tRNA genes (*trnD-GUC*, *trnI-CAU*, *trnNGUU*, *trnP-UGG*, *trnW-CCA*, and *trnY-GUA*) (Table S7).

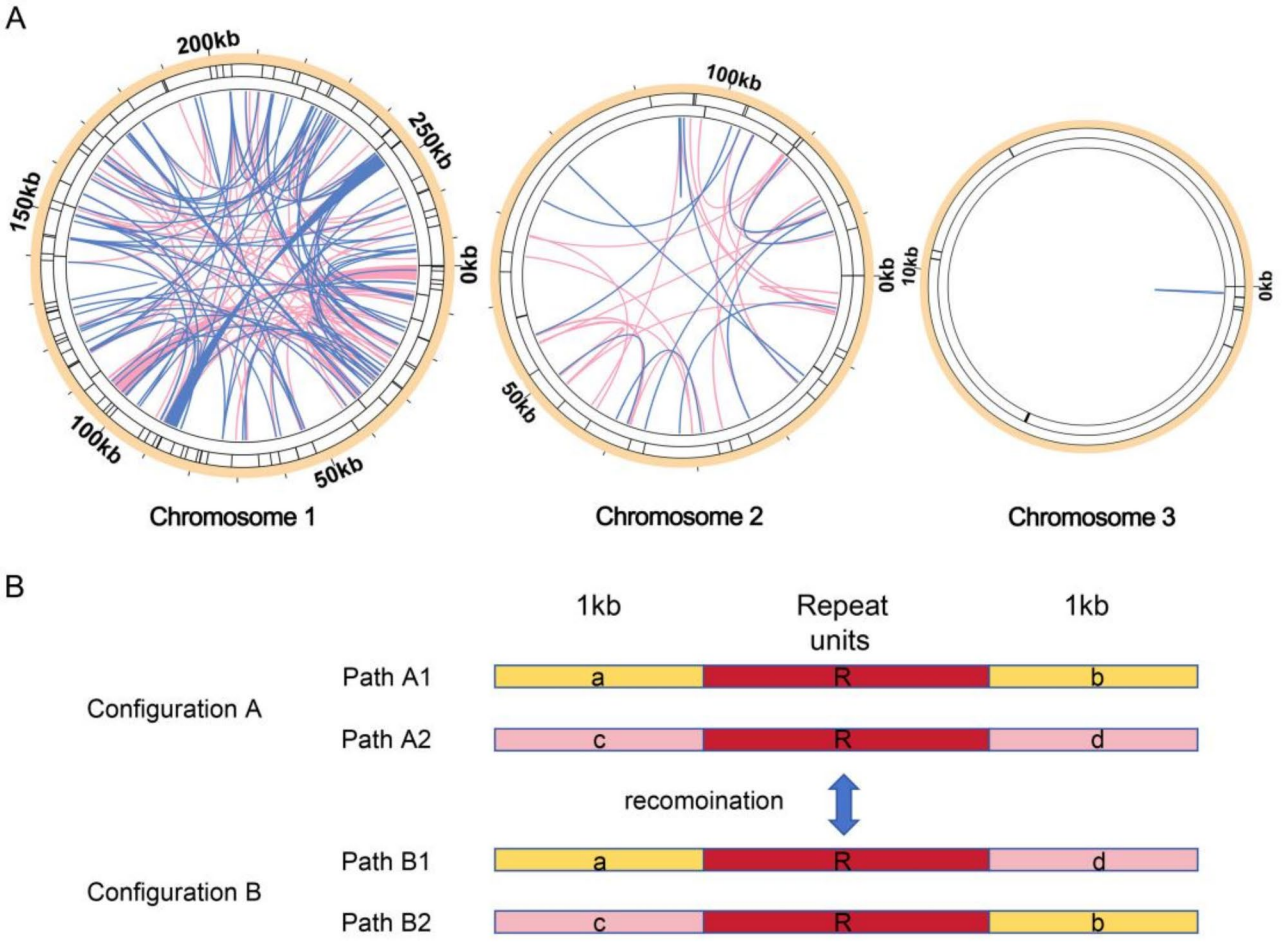

Fig. 4 The palindromic and forward repeats on mitochondria and repeat-mediated recombination. (**A**) Distribution of 101 palindromic repeats (blue line) and 113 forward repeats (red line); (**B**) Recombination mediated by the identified repeats. Repeat units (R) and their 1 kb flanking regions (a-d) were analyzed for potential recombination events. Primary configuration A shows native sequence path A1 and path A2, while alternative configuration B illustrates recombinant path B1 and path B2

Table 2 Long-read support for repeat-mediated recombination configurations

	Path A1	Path A2	Path B1	Path B2	Configuration A	Configuration B
R1	14	4	4	10	18 (56.25%)	14 (43.75%)
R2	14	33	0	1	47 (97.92%)	1 (2.08%)
R3	23	24	22	24	47 (50.54%)	46 (49.46%)
R4	11	14	1	0	25 (96.15%)	1 (3.85%)

Collinearity analysis

The collinearity analysis of the cp. genomes of *U. rhynchophylla*, *Amborella trichopoda* (an early angiosperm representative), *Vitis vinifera* (a dicotyledonous representative), and *Coffea canephora* (a closely related species) revealed a high degree of genomic collinearity, with virtually no large blocks of recombination or significant insertions observed (Fig. 6B). This finding suggests a remarkable conservation of gene order across the cp. genomes of these diverse taxa.

A collinearity analysis was conducted on the mitochondrial genome of *U. rhynchophylla*, compared with those of closely related species within Gentianales. The results identified numerous homologous collinear blocks, although these segments were relatively short in length. Additionally, unique regions specific to *U. rhynchophylla* were detected, showing no detectable homology with the other analyzed species. These findings suggest that the mitochondrial genome of *U. rhynchophylla* has undergone substantial genomic rearrangements relative to its closely related species (Fig. 6C).

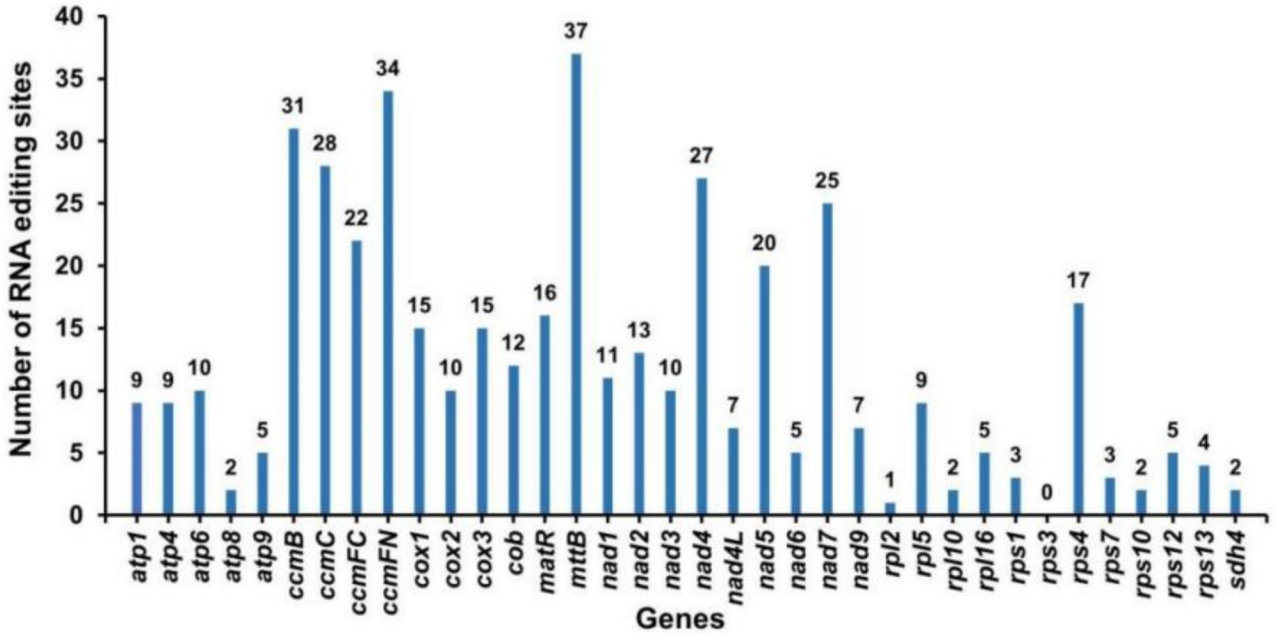

Fig. 5 Number of RNA editing sites predicted for each PCG in the mitochondrial genome

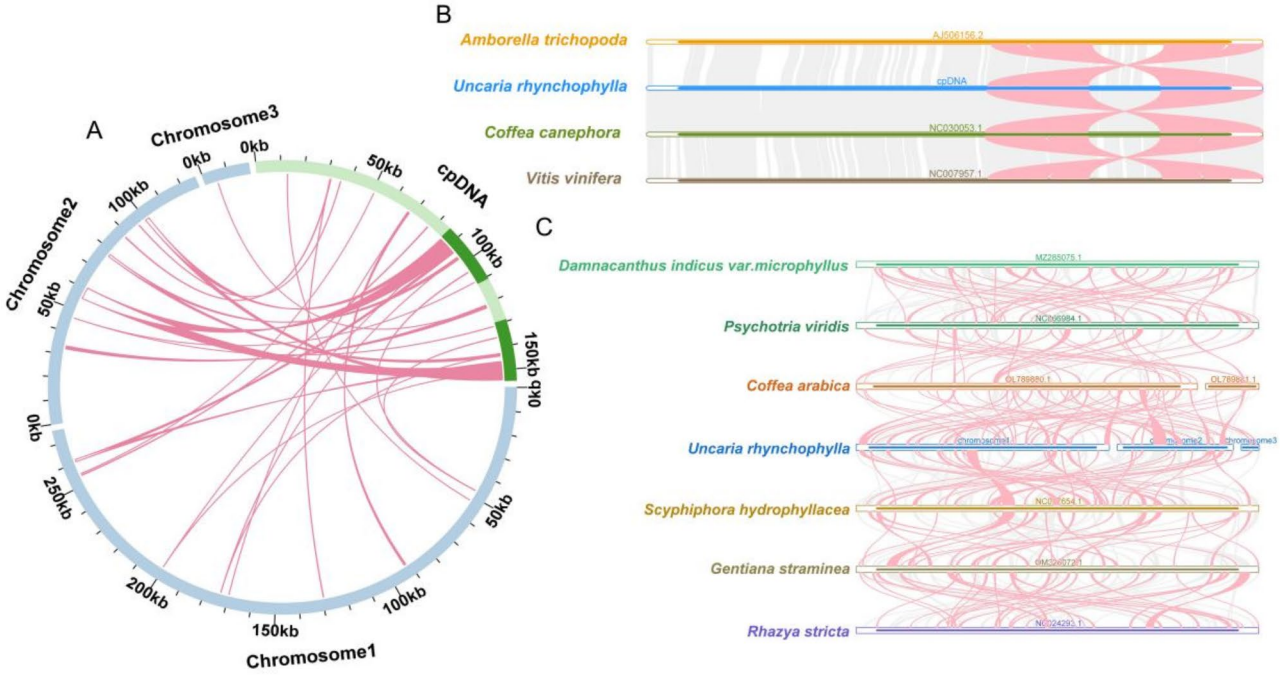

Fig. 6 Comparative genomic architecture. (**A**) MTPT transfer events in *U. rhynchophylla*. Dark green regions represent the IR regions of the cp. genome; (**B**) Collinearity of cp. genomes among representative angiosperms; (**C**) Structural rearrangements in the mitochondrial genome among species of the Gentianales order. Gray areas denote regions with significant sequence homology, while red arcs highlight loci where inversions have been identified

Phylogenetic analysis

A phylogenetic analysis was conducted on 28 species within the order Gentianales, utilizing DNA sequences from 26 conserved mitochondrial PCGs (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFC*, *ccmFN*, *cob*, *cox2*, *cox3*, *matR*, *mttB*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *rpl2*, *rps3*, *rps4*, *rps12*, *and rps13*). Two species of *Rhododendron* were designated as outgroups. The analysis robustly placed *U. rhynchophylla* within the Rubiaceae family, showing closest relationship to *Coffea arabica* and *Scyphiphora hydrophyllacea*. When compared with the phylogeny derived from cp. genomes, a significant

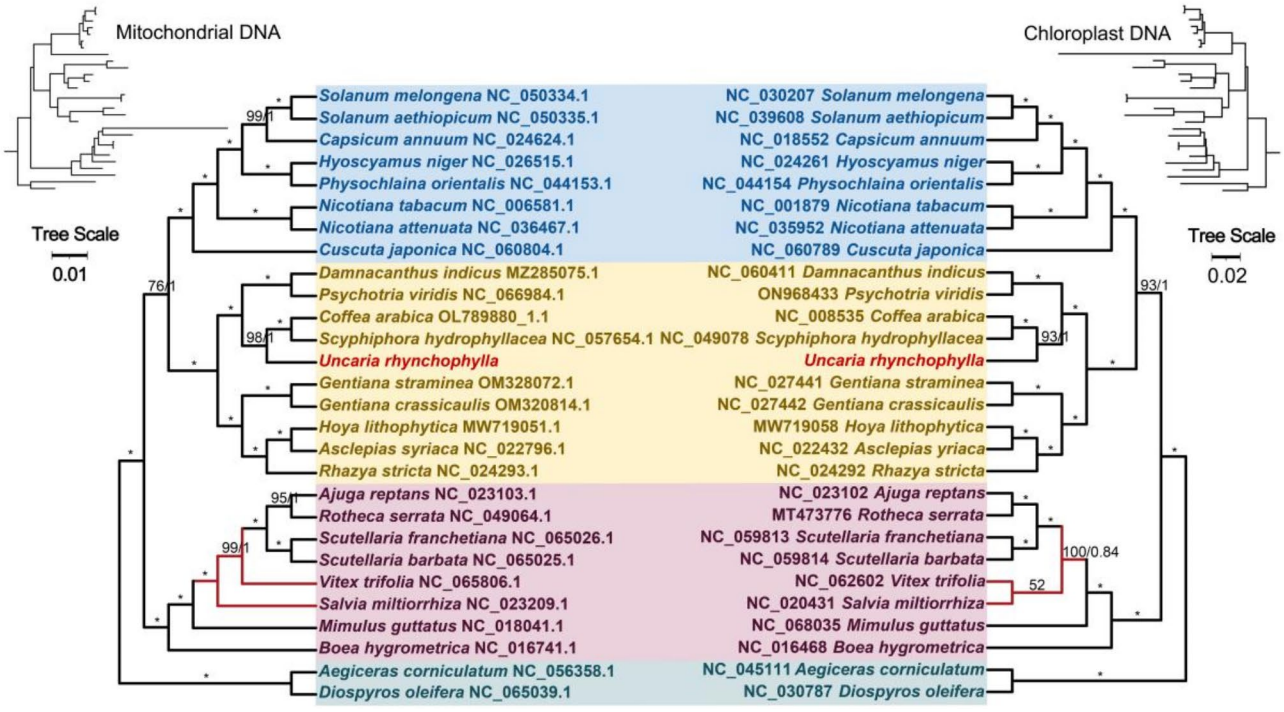

Fig. 7 Comparative phylogenetic analysis of Gentianales based on mitochondrial DNA (**A**) and cp. DNA (**B**). Red branches indicate phylogenetic relationships resolved in the mitochondrial DNA but not in the cp. DNA. ML and BI tree values are shown on branches, with asterisks (*) denoting maximum support

topological congruence was observed. However, the relationship between *Vitex trifolia* and *Salvia miltiorrhiza* in the cpDNA tree appeared unresolved (Fig. 7), while these taxa's positions were distinctly clarified in the mitochondrial DNA tree, highlighting the complementary phylogenetic signal of mitochondrial genomes.

Discussion

Plant mitochondria possess several distinctive characteristics that set them apart from their animal and fungal counterparts. These features include high copy number variation, dynamic genome reorganization, and a critical role in cytoplasmic male sterility [50]. Such specialized traits not only highlight the unique biological significance of plant mitochondria but also open new avenues for significant scientific inquiry. Understanding these differences can provide deeper insights into mitochondrial function and evolution across different kingdoms of life.

Genomic plasticity and evolution

Plant mitochondrial genomes are traditionally depicted as circular molecules [17]. However, advancements in sequencing technology and research have revealed a remarkable diversity in plant mitochondrial genomes, encompassing variations in structure, size, and gene content [51]. For instance, the mitochondrial genome of *Psilotum nudum* has been assembled into 2 circular chromosomes [52], while that of the early angiosperm *Amborella trichopoda* comprises 5 circular chromosomes ranging from 118.7 to 3179.3 kb [53]. In various *Silene* species, the chromosome count varies significantly, with *Silene conica* exhibiting the highest known number of mitochondrial chromosomes (>128), whereas other polyploid *Silene* species display chromosome numbers between 2 and 5 [54]. This study reports the presence of 1 circular contig and 2 linear contigs in the mitochondrial genome of *U. rhynchophylla*. The deduced final structure of the mitochondrial genome reveals a multi-branched, closed configuration. Although not circular, this structure is highly reliable. Collinearity analysis indicates that the mitochondrial genome sequences of *U. rhynchophylla* and its closely related species exhibit a highly non-conservative arrangement, characterized by frequent genomic recombination events.

While repetitive sequences are hypothesized to facilitate recombination, empirical evidence suggests limited recombination activity. For instance, only 74 recombination events were detected among 886,982 mitochondrial repeats in a water lily [55]. Similarly, in *Picea abies*, the majority of detected repeated sequences exhibited little or no evidence of repeat-mediated recombination [56]. In the water lily, the recombination frequencies of the two largest repeats were 0.2% and 8.2%, respectively [55], while in *Lactuca sativa* and *L. serriola*, the proportion of short repeat-mediated recombination ranged from approximately 1–10% [17]. In our study

of *U. rhynchophylla*, we observed a similar pattern: only 4 pairs of repeats displayed significant evidence of recombination activity, with 2 exhibiting extremely low recombination rates below 4%, while the other 2 demonstrated repeat-mediated recombination rates exceeding 40%. Overall, repeat-mediated recombination is typically much less common than the numerous scattered repeats in the mitochondrial genome, with many of these recombination probabilities being low, even though some recombinations may be highly active. Notably, due to the significant copy number of organelles, the occurrence of mitochondrial genome isomers in plants remains notably prevalent, adding another layer of complexity to the understanding of repeat-mediated recombination in plants.

The structure and size of plant mitochondrial genomes are influenced not only by complex repetitive sequences but also by the frequent transfer of exogenous DNA into mitochondria, a factor that can induce variations within the mitochondrial genome [57]. We identified 28 cp-derived fragments in the mitochondrial genome, including 8 functional genes (2 PCGs and 6 tRNAs). This gene transfer can cause disruptions in the mitochondrial genome and may contribute to the high level of mitochondrial genome rearrangement [58]. Interestingly, we observed that DNA transfer primarily occurred within the IR regions of the *U. rhynchophylla* cp. genome (Fig. 4), which may be attributed to the frequency of replication and gene conversion in the IR region of the cp. genome [59]. Collectively, genome-wide rearrangements, recombination events, and the integration of exogenous DNA constitute substantial driving forces shaping the structural development and evolution of *U. rhynchophylla* mitochondria.

Mitogenomic phylogenetic analysis

It is crucial to recognize that despite the complex and dynamic nature of mitochondrial structures, their coding sequences remain conserved. Previous studies have established phylogenetic relationships among different species using various methods, including morphological analysis, ITS sequence analysis, and cp. DNA analysis [60–62]. However, phylogenetic relationships derived from analyses of cp. and mitochondrial genomes occasionally exhibit divergences and inconsistencies [63]. These discrepancies may arise due to differences in evolutionary rates, gene transfer events, or selective pressures. Research has indicated occasional paternal leakage of either the mitochondrial or cp. genome during speciation and differentiation events [64]. Therefore, mitochondrial DNA analysis holds significant potential for providing novel insights into phylogenetic relationships.

In this study, we aimed to explore the phylogeny of Gentianales by constructing a sequence-based phylogenetic tree using PCGs from the mitochondrial genome. The findings indicated that mitochondrial DNA provides high-resolution insights into branches that remained unresolved by cpDNA analyses (Fig. 7), highlighting the immense potential of mitochondrial genomes in establishing plant phylogeny frameworks. However, it is important to acknowledge that the phylogenetic insights derived from this study have been constrained by several factors, including inadequate sampling, incomplete lineage sorting, and variations in evolutionary rates. To attain more robust and comprehensive phylogenetic resolutions, future research endeavors must prioritize the enhancement of mitochondrial data collection and the refinement of analytical methodologies.

Conclusion

Using Illumina and Nanopore sequencing reads, we precisely assembled and annotated the complete mitochondrial genome of *U. rhynchophylla*. Unlike the traditional view of plant mitochondrial genomes as single circular structures, the mitochondrial genome of *U. rhynchophylla* was assembled into two linear contigs and one circular contig, ultimately forming a complete mitochondrial genome with a closed structure. Through rigorous analyses, we gained profound insights into the complex-type structure, codon usage patterns, repeat-mediated recombination, and DNA transfer mechanisms within the mitochondrial genome of *U. rhynchophylla*. Compared to the cp. genome, the mitochondrial genome demonstrates remarkable potential in elucidating the phylogenetic framework of Gentianales. These findings provide invaluable genetic resources that significantly contribute to a deeper understanding of the biological characteristics of mitochondria in *Uncaria*.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11372-9.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We would like to express our gratitude to Shixin Feng and Kun Zhang for their invaluable assistance in sample collection.

Author contributions

LJG, ZJZ, LYW designed the project, JEF, CCF, WJL, QLH contributed to plant sample collection; HXY, LSS, LMP performed genome assembly, annotation and data analyses; LJG wrote the manuscript; AHE, LJS, SGW revised the manuscript. All authors read and approved the final manuscript.

Funding

Guangxi Major Science and Technology Project of China (GuikeAA22096021), National Natural Science Foundation of China (82360753), Innovative Team for Traditional Chinese Medicinal Materials Quality of Guangxi (GZKJ2305), Key Laboratory Construction Program of Guangxi Health commission (ZJC2020003), Key Techniques Research and Promotion of Guangxi Medicinal

Materials Varieties (GZKJ2314), Scientific Research Funding Project of Guangxi Botanical Garden of Medicinal Plants (GXYYXGD202202).

Data availability

The mitochondrial and cp. genomes of ***U. rhynchophylla*** have been deposited in the GenBank database (https://www.ncbi.nlm.nih.gov) with the GenBank accessions OR051899-OR051902. Additionally, the data is stored on figshare as well (https://figshare.com/articles/dataset/Uncaria_rhynchophylla_mitochondrion_complete_genome_/23578635).

Declarations

Ethics approval and consent to participate

The *U. rhynchophylla* plants were cultivated at an experimental research base of the Guangxi Botanical Garden of Medicinal Plants in Nanning, Guangxi, China (22°51 N, 108°23 E). The voucher specimens (W20200117) have been preserved in the herbarium of Guangxi Botanical Garden of Medicinal Plants (GXMG). Taxonomic identification of the plant was conducted by Shugen Wei. The collection, experimental research, and field studies involving these plants were conducted in accordance with the IUCN Policy Statement on Research Involving Species at Risk of Extinction and the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 16 June 2023 / Accepted: 16 February 2025
Published online: 26 February 2025

References

1. Liang JH, Wang C, Huo XK, Tian XG, Zhao WY, Wang X, Sun CP, Ma XC. The genus *Uncaria*: A review on phytochemical metabolites and biological aspects. Fitoterapia. 2020;147:104772.
2. Ndagijimana A, Wang X, Pan G, Zhang F, Feng H, Olaleye O. A review on Indole alkaloids isolated from *Uncaria rhynchophylla* and their Pharmacological studies. Fitoterapia. 2013;86:35–47.
3. Kuramochi T, Chu J, Suga T. Gou-teng (from *Uncaria rhynchophylla* Miquel)-induced endothelium-dependent and-independent relaxations in the isolated rat aorta. Life Sci. 1994;54(26):2061–9.
4. Lim HB, Lee HR. Safety and biological activity evaluation of *Uncaria rhynchophylla* ethanolic extract. Drug Chem Toxicol. 2022;45(2):907–18.
5. Bi C, Lu N, Xu Y, He C, Lu Z. Characterization and analysis of the mitochondrial genome of common bean (*Phaseolus vulgaris*) by comparative genomic approaches. Int J Mol Sci. 2020;21(11):3778.
6. Van der Bliek AM, Sedensky MM, Morgan PG. Cell biology of the mitochondrion. Genetics. 2017;207(3):843–71.
7. Smith DR, Keeling PJ. Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. *Proceedings of the National Academy of Sciences* 2015, 112(33):10177–10184.
8. Drouin G, Daoud H, Xia J. Relative rates of synonymous substitutions in the mitochondrial, Chloroplast and nuclear genomes of seed plants. Mol Phylogenet Evol. 2008;49(3):827–31.
9. Gualberto JM, Newton KJ. Plant mitochondrial genomes: dynamics and mechanisms of mutation. Annu Rev Plant Biol. 2017;68:225–52.
10. Sloan DB, Havird JC, Sharbrough J. The on-again, off-again relationship between mitochondrial genomes and species boundaries. Mol Ecol. 2017;26(8):2212–36.
11. 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. 2015;112(27):E3515–24.
12. Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 2012;10(1):e1001241.
13. Jiang H, Lu Q, Qiu S, Yu H, Wang Z, Yu Z, Lu Y, Wang L, Xia F, Wu Y. Fujian cytoplasmic male sterility and the fertility restorer gene OsRf19 provide a promising breeding system for hybrid rice. Proc Natl Acad Sci. 2022;119(34):e2208759119.
14. Turmel M, Otis C, Lemieux C. Mitochondrion-to-chloroplast DNA transfers and intragenomic proliferation of Chloroplast group II introns in Gloeotilopsis green algae (Ulotrichales, Ulvophyceae). Genome Biol Evol. 2016;8(9):2789–805.
15. Sloan DB, Wu Z. History of plastid DNA insertions reveals weak deletion and AT mutation biases in angiosperm mitochondrial genomes. Genome Biol Evol. 2014;6(12):3210–21.
16. 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):2773–85.
17. Kozik A, Rowan BA, Lavelle D, Berke L, Schranz ME, Michelmore RW, Christensen AC. The alternative reality of plant mitochondrial DNA: one ring does not rule them all. PLoS Genet. 2019;15(8):e1008373.
18. Bi C, Qu Y, Hou J, Wu K, Ye N, Yin T. Deciphering the multi-chromosomal mitochondrial genome of *Populus simonii*. Front Plant Sci 2022, 13.
19. Maréchal A, Brisson N. Recombination and the maintenance of plant organelle genome stability. New Phytol. 2010;186(2):299–317.
20. Guo W, Grewe F, Fan W, Young GJ, Knoop V, Palmer JD, Mower JP. *Ginkgo* and *Welwitschia* mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. Mol Biol Evol. 2016;33(6):1448–60.
21. Choi K-S, 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.
22. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A. The Norway Spruce genome sequence and conifer genome evolution. Nature. 2013;497(7451):579–84.
23. Jackman SD, Warren RL, Gibb EA, Vandervalk BP, Mohamadi H, Chu J, Raymond A, Pleasance S, Coope R, Wildung MR. Organellar genomes of white Spruce (*Picea glauca*): assembly and annotation. Genome Biol Evol. 2016;8(1):29–41.
24. Shearman JR, Sonthirod C, Naktang C, Pootakham W, Yoocha T, Sangsrakru D, Jomchai N, Tragoonrung S, Tangphatsornruang S. The two chromosomes of the mitochondrial genome of a sugarcane cultivar: assembly and recombination analysis using long PacBio reads. Sci Rep. 2016;6(1):31533.
25. Ling LZ, Zhang SD. The complete Chloroplast genome of the traditional Chinese herb, *Uncaria rhynchophylla* (Rubiaceae). Mitochondrial DNA Part B. 2020;5(1):424–5.
26. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(6):e1005595.
27. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.
28. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100.
29. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de Novo genome assemblies. Bioinformatics. 2015;31(20):3350–2.
30. Chen Y, Ye W, Zhang Y, Xu Y. High speed BLASTN: an accelerated megablast search tool. Nucleic Acids Res. 2015;43(16):7762–8.
31. Richardson AO, Rice DW, Young GJ, Alverson AJ, Palmer JD. The fossilized mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. BMC Biol. 2013;11(1):1–17.
32. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. GeSeq–versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45(W1):W6–11.
33. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25(5):955–64.
34. Jin J-J, Yu W-B, Yang J-B, Song Y, DePamphilis CW, Yi T-S, Li D-Z. GetOrganelle: a fast and versatile toolkit for accurate de Novo assembly of organelle genomes. Genome Biol. 2020;21:1–31.
35. Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 2019;47(W1):W65–73.
36. Lewis SE, Searle S, Harris N, Gibson M, Iyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby M. Apollo: a sequence annotation editor. Genome Biol. 2002;3:1–14.
37. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT. 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.
38. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4.
39. Lenz H, Hein A, Knoop V. Plant organelle RNA editing and its specificity factors: enhancements of analyses and new database features in PREPACT 3.0. BMC Bioinformatics. 2018;19(1):1–18.
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 JV, 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 H, Meltzer P, Davis S. RCircos: an R package for circos 2D track plots. BMC Bioinformatics. 2013;14:1–5.
44. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49–49.
45. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80.
46. Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587–9.
47. Nguyen L-T, 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.
48. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–42.
49. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256–9.
50. Arimura S-i. Fission and fusion of plant mitochondria, and genome maintenance. Plant Physiol. 2018;176(1):152–61.
51. Petersen G, Anderson B, Braun H-P, Meyer EH, Møller IM. Mitochondria in parasitic plants. Mitochondrion. 2020;52:173–82.
52. Guo W, Zhu A, Fan W, Mower JP. Complete mitochondrial genomes from the ferns *Ophioglossum californicum* and *Psilotum nudum* are highly repetitive with the largest organellar introns. New Phytol. 2017;213(1):391–403.
53. Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, Barry K, Boore JL, Zhang Y, DePamphilis CW. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm Amborella. Science. 2013;342(6165):1468–73.
54. Wu ZQ, Liao XZ, Zhang XN, Tembrock LR, Broz A. Genomic architectural variation of plant mitochondria—A review of multichromosomal structuring. J Syst Evol. 2022;60(1):160–8.
55. Dong S, Zhao C, Chen F, Liu Y, Zhang S, Wu H, Zhang L, Liu Y. The complete mitochondrial genome of the early flowering plant *Nymphaea colorata* is highly repetitive with low recombination. BMC Genomics. 2018;19:1–12.
56. Sullivan AR, Eldfjell Y, Schiffthaler B, Delhomme N, Asp T, Hebelstrup KH, Keech O, Öberg L, Møller IM, Arvestad L. The mitogenome of Norway Spruce and a reappraisal of mitochondrial recombination in plants. Genome Biol Evol. 2020;12(1):3586–98.
57. Bergthorsson U, Adams KL, Thomason B, Palmer JD. Widespread horizontal transfer of mitochondrial genes in flowering plants. Nature. 2003;424(6945):197–201.
58. Shidhi PR, Biju VC, Anu S, Vipin CL, Deelip KR, Achuthsankar SN. Genome characterization, comparison and phylogenetic analysis of complete mitochondrial genome of evolvulus alsinoides reveals highly rearranged gene order in solanales. Life. 2021;11(8):769.
59. 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.
60. Turner IM. A revised conspectus of *Uncaria* (Rubiaceae). Webbia. 2018;73(1):9–21.
61. Zhu S, Li Q, Chen S, Wang Y, Zhou L, Zeng C, Dong J. Phylogenetic analysis of *Uncaria* species based on internal transcribed spacer (ITS) region and ITS2 secondary structure. Pharm Biol. 2018;56(1):548–58.
62. Chen M-M, Zhang M, Liang Z-S, He Q-L. Characterization and comparative analysis of Chloroplast genomes in five *Uncaria* species endemic to China. Int J Mol Sci. 2022;23(19):11617.
63. Zhang L, Chen F, Zhang X, Li Z, Zhao Y, Lohaus R, Chang X, Dong W, Ho SYW, Liu X, et al. The water Lily genome and the early evolution of flowering plants. Nature. 2020;577(7788):79–84.
64. Tsujimura M, Mori N, Yamagishi H, Terachi T. A possible breakage of linkage disequilibrium between mitochondrial and Chloroplast genomes during emmer and Dinkel wheat evolution. Genome. 2013;56(4):187–93.

Publisher's note