BMC Plant Biology

RESEARCH Open Access

Comparative analysis of mitochondrial genomes of *Stemona tuberosa* lour. reveals heterogeneity in structure, synteny, intercellular gene transfer, and RNA editing

De Xu[1], Tao Wang[1], Juan Huang[1], Qiang Wang[1], Zhide Wang[1], Zhou Xie[1], Dequan Zeng[1], Xue Liu[2,3*] and Liang Fu[1*]

Abstract

Background *Stemona tuberosa*, a vital species in traditional Chinese medicine, has been extensively cultivated and utilized within its natural distribution over the past decades. While the chloroplast genome of *S. tuberosa* has been characterized, its mitochondrial genome (mitogenome) remains unexplored.

Results This paper details the assembly of the complete *S. tuberosa* mitogenome, achieved through the integration of Illumina and Nanopore sequencing technologies. The assembled mitogenome is 605,873 bp in size with a GC content of 45.63%. It comprises 66 genes, including 38 protein-coding genes, 25 tRNA genes, and 3 rRNA genes. Our analysis delved into codon usage, sequence repeats, and RNA editing within the mitogenome. Additionally, we conducted a phylogenetic analysis involving *S. tuberosa* and 17 other taxa to clarify its evolutionary and taxonomic status. This study provides a crucial genetic resource for evolutionary research within the genus *Stemona* and other related genera in the Stemonaceae family.

Conclusion Our study provides the inaugural comprehensive analysis of the mitochondrial genome of *S. tuberosa*, revealing its unique multi-branched structure. Through our investigation of codon usage, sequence repeats, and RNA editing within the mitogenome, coupled with a phylogenetic analysis involving *S. tuberosa* and 17 other taxa, we have elucidated its evolutionary and taxonomic status. These investigations provide a crucial genetic resource for evolutionary research within the genus Stemona and other related genera in the Stemonaceae family.

Keywords *Stemona tuberosa*, Mitochondrial genome, Repeated sequences, Phylogenetic relationship, RNA editing

*Correspondence:
Xue Liu
liu0906xue@163.com
Liang Fu
1426236936@qq.com
[1]Dazhou Academy of Agricultural Sciences, Dazhou 635000, China
[2]Chongqing Key Laboratory of Traditional Chinese Medicine Resource, Endangered Medicinal Breeding National Engineering Laboratory, Chongqing Academy of Chinese Materia Medica, Chongqing 400065, China
[3]College of Pharmacy, Chongqing Medical University, Chongqing 400016, China

Introduction

The genus *Stemona* (Stemonaceae) encompasses approximately 27 species globally that are predominantly found across Southeastern Asia. Among these, *Stemona tuberosa* stands out as a significant medicinal plant. It is recognized as one of three protospecies officially listed in the 2020 Chinese Pharmacopoeia of the People's Republic of China for its properties of tonifying Qi, moistening the lungs, and exterminating insects [1–5]. Due to its broad distribution and noteworthy therapeutic efficacy, *S. tuberosa* is preeminent among the Bai Bu medicinal materials. Challenges such as low yield and growth rates hamper the development of industries based on *S. tuberosa*. Currently, the majority of *S. tuberosa* resources are sourced from the wild, with limited artificial cultivation. Overexploitation, especially in easily accessible areas, has severely damaged wild resources. This has led to isolated populations facing significant threats, making sustainable use and conservation of this plant crucial. The plight of *S. tuberosa* has garnered considerable attention from government agencies and researchers. While studies on this species have extensively covered areas like chemistry, pharmacology, breeding, and quality assessment, there remains a significant gap in our understanding of its molecular genetics. Furthermore, taxonomic disputes within the genus *Stemona* complicate the classification and status determination of *S. tuberosa* [6]. This underscores the critical need for further exploration of its molecular and genetic information through genomic studies. Molecular phylogenetic studies within the Stemonaceae have provided a complete chloroplast genome assembly and detailed investigations into the chromosomal-scale genome of *S. tuberosa* [7, 8]. but no mitochondrial genome (mitogenome) for *S. tuberosa* has been reported, which significantly restricts further research in this area.

Apart from the nucleus, chloroplast and mitochondria are the only two organelles in a plant cell to possess genetic material and have evolved independently of the nuclear genome. The chloroplast genome is unique to plants compared to animal organelles, and the mitogenome is much larger and structurally variable [9]. Whole chloroplast genomes contain numerous variations, such as single-nucleotide polymorphisms (SNPs), single-sequence repeats (SSRs), insertion or deletion polymorphisms (indels), of which have been instrumental in characterizing genetic diversity and divergence in medicinal species [10], discerning population structure

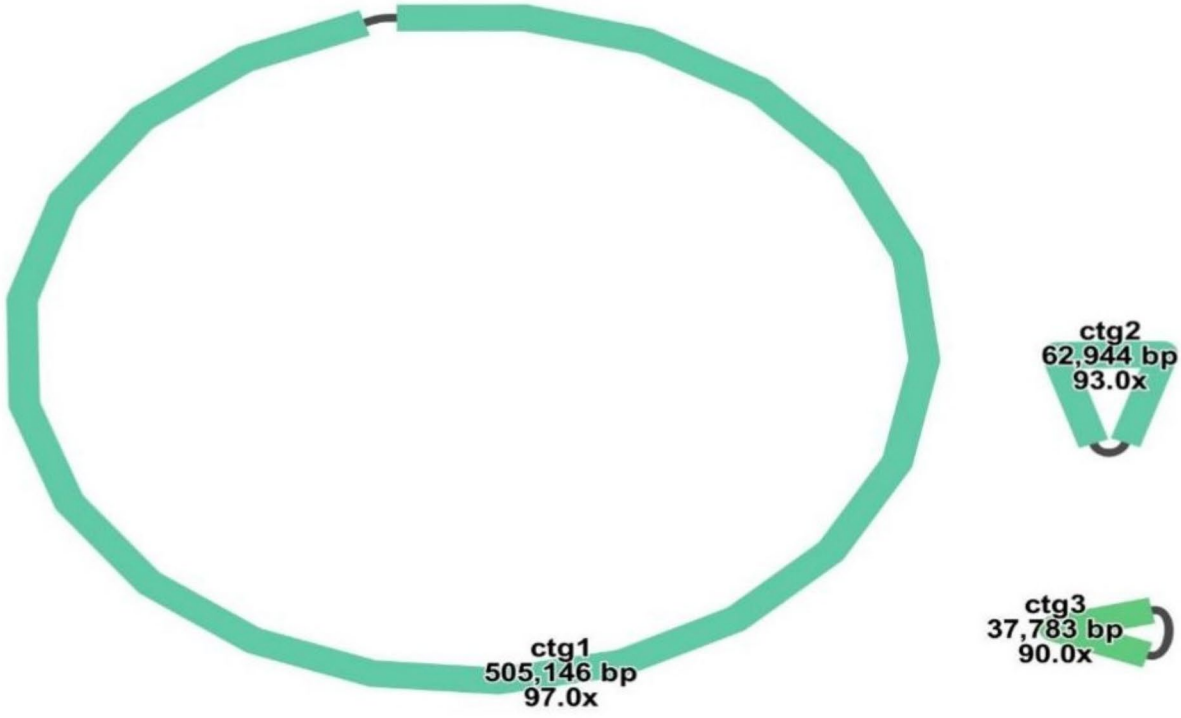

Fig. 1 Circular representation of the mitochondrial genome assembly of*S. tuberosa*. The figure shows the assembly result visualized in Bandage, displaying three circular chromosomes. Chromosome 1 (ctg1) has a length of 505,146 bp with 97x coverage, Chromosome 2 (ctg2) is 62,944 bp long with 93x coverage, and Chromosome 3 (ctg3) measures 37,783 bp with 90x coverage. Each circular chromosome is indicated by a closed loop, representing the structure of the *S. tuberosa* mitochondrial genome

Table 1 Summary of Mitochondrial Genome Assembly for *S. tuberosa*

NCBI Accession number	Contigs	Type	Length (bp)	GC content (%)
	Contig 1–3	NA	605,873	45.63
PQ374236	Contig 1	circular	505,146	45.67
PQ374237	Contig 2	circular	62,944	44.80
PQ374238	Contig 3	circular	37,783	46.52

[11] and evaluating gene flow [12]. The mitogenome contains a large number of exogenous sequences and repetitive sequences from the nuclear and chloroplast genomes [13], and involves in numerous metabolic processes and plays a critical role in energy metabolism, gene expression, stress response, and plant growth in many seed plants. Plant mitogenomes generally exhibit a circular genome structure; however, their physical organization is highly complex, varying in size and structure due to homologous recombination between repeats [14, 15]. This results in a mix of linear [16, 17], circular [18] and branched structures [19]. For instance, the mitogenome of *Arabidopsis thaliana*is typically organized as a single circular structure [20], whereas in *Silene conica*, it presents complex multichromosomal configurations [21]. Similarly, the cucumber (*Cucumis sativus*) mitogenome consists of three circular chromosomes [22], and the onion (*Allium cepa*) comprises two circular chromosomes [23].

This study represents the first successful sequencing and assembly of the *S. tuberosa* mitogenome, achieved through the integration of next-generation sequencing (NGS) and third-generation sequencing technologies (TGS). A comprehensive investigation into the genome's multichromosomal structure was conducted. Additionally, analyses of repeat sequences, codon usage bias, phylogenetic relationships, RNA editing, and intergenomic sequence transfer revealed key insights into potential genomic recombination and dynamic evolutionary changes in *S. tuberosa*. These results could provide a solid theoretical foundation and valuable resources for the structural and functional characterization of the *S. tuberosa* mitogenome, while also offering important insights for further research into its genetic mechanisms and evolutionary history.

Materials and methods

Plant material and mitogenomic sequencing

The sample material for this study was provided by the Dazhou Academy of Agricultural Sciences in Dazhou, China. Total DNA was isolated from fresh leaves of *S. tuberosa* and purified by the cationic detergent cetyltrimethylammonium bromide (CTAB) method [24]. The mitogenome of *S. tuberosa* was sequenced utilizing both Illumina and Nanopore technologies. We constructed paired-end libraries with an insert size of 300 bp, which were sequenced on the Illumina HiSeq 2500 platform. To ensure data quality, low-quality reads were removed using the SOAP-nuke (version 2.1.4) tool (available at https://github.com/BGI-ffexlab/SOAPnuke). For Nanopore sequencing, the SQK-LSK109 ligation kit was employed following the manufacturer's guidelines. The prepared library was loaded onto primed R9.4 Spot-on Flow Cells and sequenced using a PromethION sequencer (Oxford Nanopore Technologies, Oxford, UK) over a 48-hour period. Base calling of the raw data was performed using Oxford Nanopore's GuPPy v1.2.0.

Table 2 Gene composition in the mitogenome of *S. tuberosa*

Group of genes	Name of genes
ATP synthase	*atp1*, *atp4*, *atp6*, *atp8*, *atp9*
NADH dehydrogenase	*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, *nad9*
Cytochrome b	*cob*
Cytochrome c biogenesis	*ccmB*, *ccmC*, *ccmFC*, *ccmFN*
Cytochrome c oxidase	*cox1*,*cox2*, *cox3*
Maturases	*matR*
Protein transport subunit	*mttB*
Ribosomal protein large subunit	*rpl2*, *rpl5*,*rpl16*
Ribosomal protein small subunit	*rps1*, *rps2*, *rps3*, *rps4*, *rps7*, *rps10*, *rps11*, *rps12*, *rps13*, *rps14*, *rps19*
Ribosome RNA	*rrn5*, *rrn18*, *rrn26*
Transfer RNA	*trnA-UGC*, *trnC-GCA*, *trnD-GUC*, *trnE-UUC*, *trnF-GAA* (×2), *trnfM-CAU*, *trnH-GUG*, *trnI-CAU* (×2), *trnI-GAU*, *trnK-UUU*, *trnL-CAA*, *trnL-UAA*, *trnM-CAU* (×2), *trnN-GUU* (×2), *trnP-UGG*, *trnQ-UUG*, *trnR-ACG*, *trnR-CCU*, *trnS-GCU*, *trnS-GGA*, *trnS-UGA*, *trnT-UGU*, *trnV-GAC*, *trnW-CCA*, *trnY-GUA*

Note: "x2" represents the number of copies. For example, *trnfM-CAU* had two copies

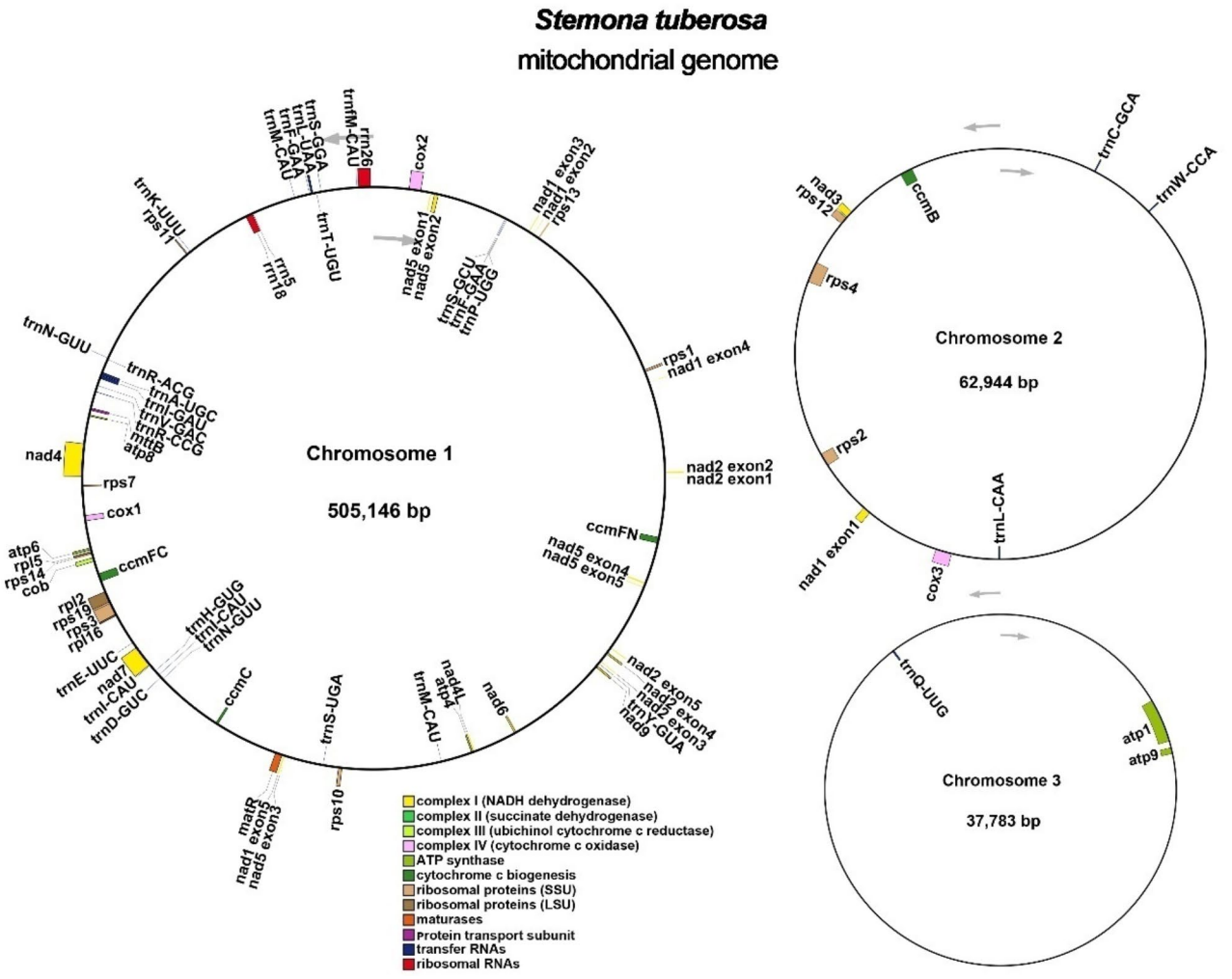

Fig. 2 The map of the mitogenome of *S. tuberosa*. The arrows shown transcriptional direction of the mitogenome. Genes with different functions were depicted using different colors

Mitogenome assembly and annotation

The assembly of the *S. tuberosa* mitogenome was accomplished using GetOrganelle software (version 1.7.5) with specific parameters: -R 20 -k 21,45,65,85,105 -P 1,000,000 -F embplant-mt [25]. Visualization of the assembled mitogenome was facilitated by Bandage software (version 0.8.1), which also enabled the manual removal of extended fragments from the chloroplast and nuclear genomes [26]. The alignment of the Nanopore data with the circular mitogenome was conducted using BWA software (version 0.7.17) [27]. For annotating the protein-coding genes (PCGs) in the *S. tuberosa* mitogenome, we referred to two mitogenome sequences from *Arabidopsis thaliana* (NC_037304) and *Liriodendron tulipifera* (NC_021152.1). Annotation was performed using Geseq (version 2.03) [28] and IPMGA (available at http://www.1kmpg.cn/ipmga/). Additionally, tRNA and rRNA within the mitogenome were annotated using tRNAscan-SE (version 2.0.11) [29] and BLASTN software (version 2.13.0) [30], respectively. Any errors in the annotation were meticulously corrected through a manual process using Apollo software (version 1.11.8) [31]. The final assembly and annotated files were subsequently deposited in the NCBI database (https://www.ncbi.nlm.nih.gov/).

Analysis of codon usage bias, repeat fragments, and prediction of RNA editing sites

Protein-coding sequences were extracted using PhyloSuite software (version 1.1.16) [32] with default settings. The analysis of codon usage bias and the calculation of relative synonymous codon usage (RSCU) were performed using MEGA software (version 7.0) [33] based on the protein-coding genes from the mitogenome. Analyses of Short Tandem Repeats (STR), tandem repeats, and dispersed repeats were conducted using various tools: MISA (version 2.1), accessible online at [https://webblast.ipk-gatersleben.de/misa/] [34], the Tandem Repeats

Table 3 Relative synonymous codon usage for each amino acid in the mitogenome of *S. tuberosa*

Amino	Codon 1	Codon 2	Codon 3	Codon 4	Codon 5	Codon 6
	RSCU	RSCU	RSCU	RSCU	RSCU	RSCU
Ala	GCU	GCA	GCC	GCG		
	1.61	1.01	0.88	0.5		
Arg	AGA	CGA	CGU	CGG	AGG	CGC
	1.43	1.28	1.24	0.73	0.72	0.6
Asn	AAU	AAC				
	1.32	0.68				
Asp	GAU	GAC				
	1.41	0.59				
Cys	UGU	UGC				
	1.12	0.88				
End	UAA	UGA	UAG			
	1.41	0.88	0.71			
Gln	CAA	CAG				
	1.52	0.48				
Glu	GAA	GAG				
	1.37	0.63				
Gly	GGA	GGU	GGG	GGC		
	1.44	1.33	0.69	0.54		
His	CAU	CAC				
	1.54	0.46				
Ile	AUU	AUA	AUC			
	1.31	0.85	0.85			
Leu	UUA	CUU	UUG	CUA	CUG	CUC
	1.39	1.26	1.18	0.9	0.63	0.63
Lys	AAA	AAG				
	1.12	0.88				
Met	AUG					
	1.0					
Phe	UUU	UUC				
	1.14	0.86				
Pro	CCU	CCA	CCC	CCG		
	1.44	1.15	0.8	0.61		
Ser	UCU	UCA	UCC	AGU	UCG	AGC
	1.41	1.14	1.0	0.99	0.84	0.62
Thr	ACU	ACC	ACA	ACG		
	1.36	1.0	1.0	0.64		
Trp	UGG					
	1.0					
Tyr	UAU	UAC				
	1.47	0.53				
Val	GUU	GUA	GUG	GUC		
	1.17	1.12	0.92	0.79		

Finder (TRF, version 4.09) available at [https://tandem.bu.edu/trf/trf.unix.help.html] [35], and the REPuter server at [https://bibiserv.cebitec.uni-bielefeld.de/reputer/] [36], respectively. Visualization of these genomic elements was achieved using the Circos package (version 0.69.9) [37] and Excel 2021. Additionally, RNA editing events were predicted using the online tool PREPACT3 (available at http://www.prepact.de/) [38], with a cutoff value set at 0.001.

Identification of homologous fragment and collinear analysis

The chloroplast genome of *S. tuberosa* was assembled using GetOrganelle software. Annotation of this genome was performed using CPGAVAS2 software (version 2.0) [39]. Homologous sequences between the mitochondrial and chloroplast genomes were analyzed using BLASTN software (version 2.13.0) with default settings, and the resulting homologous fragments were visualized using

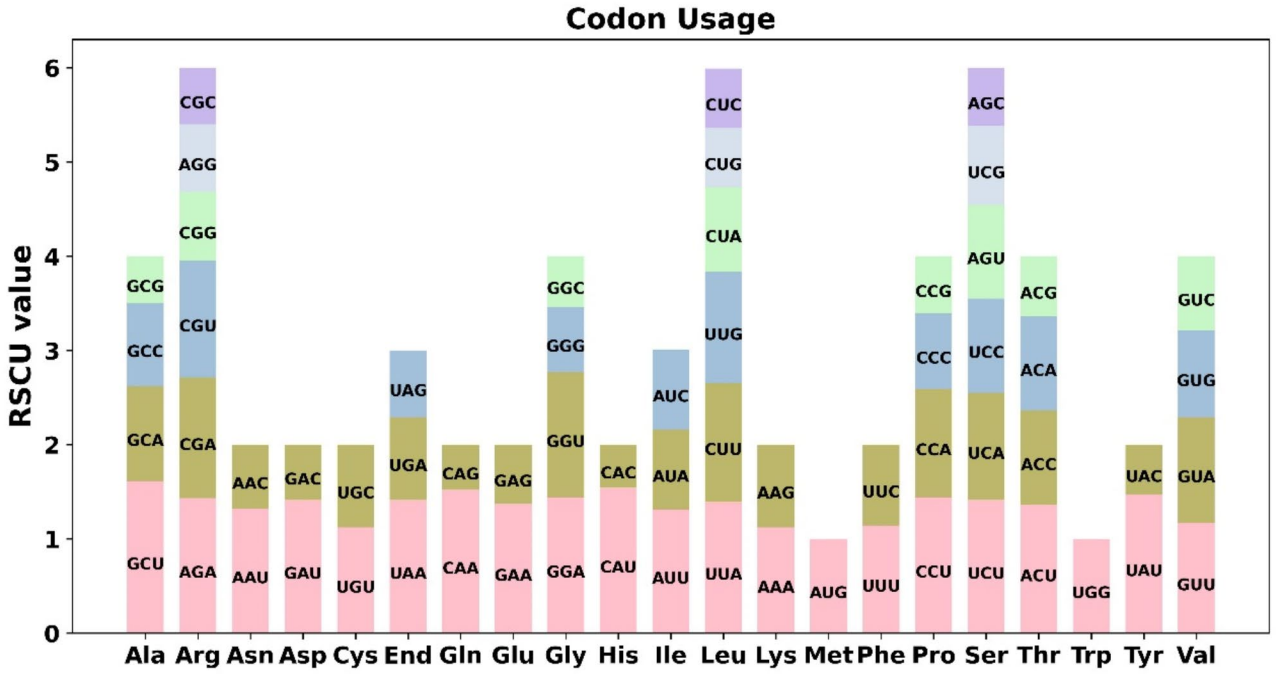

Fig. 3 Relative synonymous codon usage (RSCU) in the mitochondrial protein-coding genes of*S. tuberosa*. The figure displays the RSCU values for the 38 unique protein-coding genes in the *S. tuberosa* mitochondrial genome. The codon usage patterns are represented for 20 amino acids and stop codons (End), showing the preference for certain codons over others. Codons with higher RSCU values indicate a greater frequency of usage relative to other synonymous codons

the Circos package (version 0.69.9). Further evolutionary analysis was conducted using the BLAST program to examine species evolution. Additionally, MCscanX [40] software was utilized to generate a Multiple Synteny Plot, mapping synteny between *S. tuberosa* and closely related species. This integrated approach provides a comprehensive view of the genomic architecture and evolutionary relationships of *S. tuberosa*.

Construction of maximum likelihood tree based on the PCGs

Seventeen complete mitogenomes from five different orders (Asparagales, Arecales, Pandanales, Alismatales, and Ranunculales) were retrieved from the National Center for Biotechnology Information (NCBI) database. These genomes include species such as *Chlorophytum comosum* (MW411187.1), *Asparagus officinalis* (NC_053642.1), *Allium cepa* (NC_030100.1), *Hemerocallis citrina* (MZ726801_3.1), *Crocus sativus* (OL804177.1), and others up to *Aconitum kusnezoffii* (NC_053920.1). For phylogenetic analysis, these mitogenomes were used, with *Pulsatilla dahurica* and *Aconitum kusnezoffii* serving as outgroups. Using PhyloSuite, 24 conserved protein-coding genes (PCGs) such as *atp1*, *atp4*, *atp6*, and others up to *nad9* were extracted. These multiple sequences were aligned using MAFFT software (v7.505, parameter "–auto") [41]. Phylogenetic analysis was conducted using IQ-TREE software (version 1.6.12) with specific parameters:--alrt 1000 -B 1000 [42], and the resulting maximum likelihood tree was visualized with ITOL software (version 4.0) [43]. This robust methodology provides insights into the evolutionary relationships among these diverse plant species.

Results

Characteristics of the mitogenomes of *S. tuberosa*

The mitogenome of *S. tuberosa* exhibits a branched structure, comprising three circular contigs as depicted in Fig. 1. These contigs vary in size and GC content: contig 1 measures 505,146 bp with a GC content of 45.67%, contig 2 is 62,944 bp with a GC content of 44.80%, and contig 3 spans 37,783 bp with a GC content of 46.52%. Collectively, the total size of the *S. tuberosa* mitogenome is 605,873 bp, with an overall GC content of 45.63%. The GenBank accession number for this mitogenome is detailed in Table 1. A total of 66 genes were identified within the mitogenome, comprising 38 unique protein-coding genes (PCGs), 25 tRNA genes, and 3 rRNA genes, as listed in Table 2. Among the 38 unique PCGs, 24 are considered core genes, which include five ATP synthase genes (*atp1*, *atp4*, *atp6*, *atp8*, and *atp9*), nine NADH dehydrogenase genes (*nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, and *nad9*), four cytochrome c biogenesis genes (*ccmB*, *ccmC*, *ccmFC*, and *ccmFN*), three

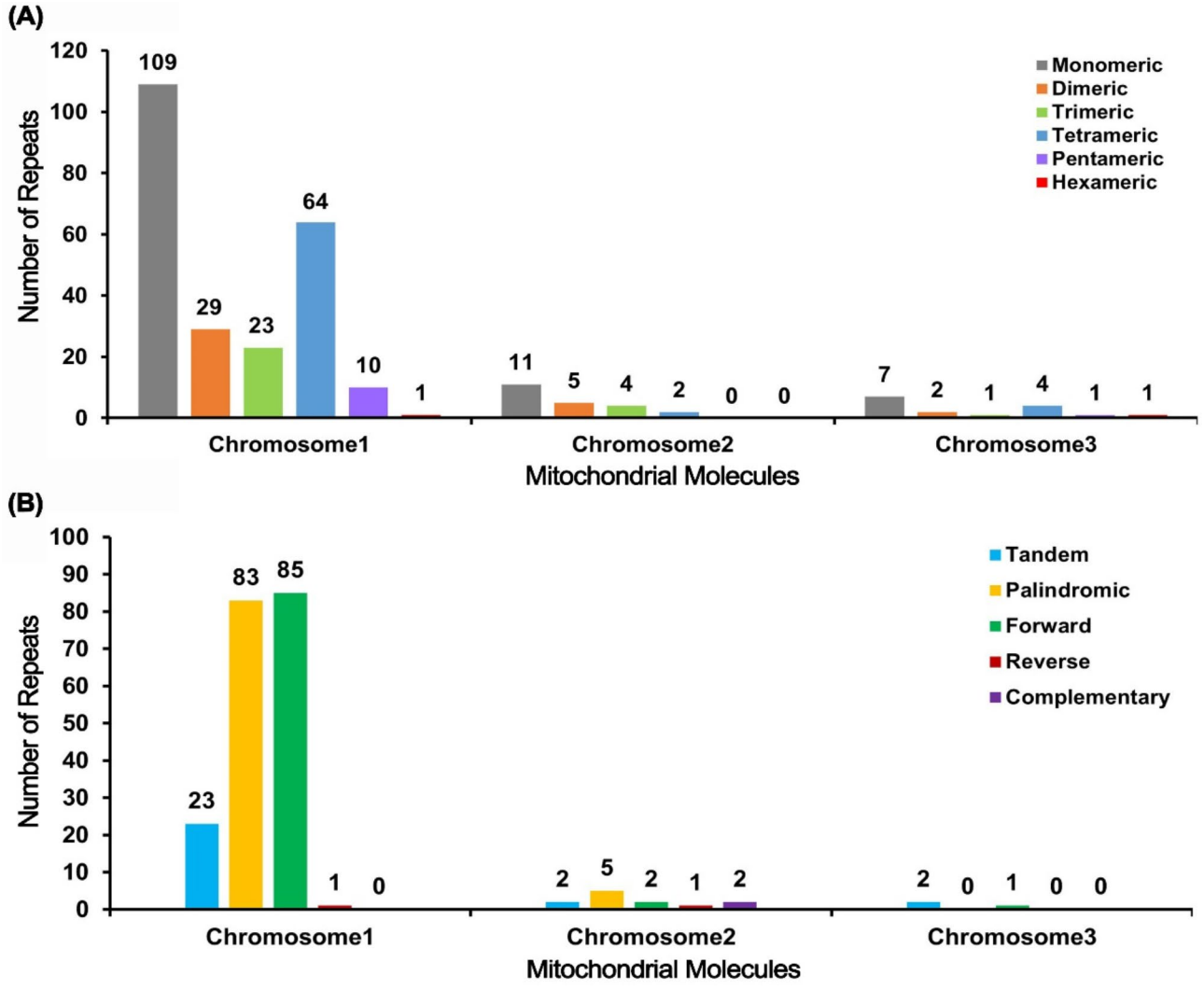

Fig. 4 Analysis of repeat elements in the mitochondrial genome of*S. tuberosa*.(**A**) Distribution of repeat motifs classified by repeat unit length (monomeric, dimeric, trimeric, tetrameric, pentameric, and hexameric) across the three mitochondrial chromosomes of *S. tuberosa*. (**B**) Classification of repeats based on structural types, including tandem, palindromic, forward, reverse, and complementary repeats

cytochrome c oxidase genes (*cox1*, *cox2*, and *cox3*), one protein transport subunit gene (*mttB*), one maturase gene (*matR*), and one cytochrome b gene (*cob*). The non-core genes are represented by three ribosomal large subunit genes (*rpl2*, *rpl5*, and *rpl16*) and eleven ribosomal small subunit genes (*rps1*, *rps2*, *rps3*, *rps4*, *rps7*, *rps10*, *rps11*, *rps12*, *rps13*, *rps14*, and *rps19*), as shown in Fig. 2.

Analysis of relative synonymous codon usage

In this study, we analyzed the codon usage patterns of the 38 unique protein-coding genes (PCGs) in the *S. tuberosa* mitogenome. Relative synonymous codon usage (RSCU) values greater than 1 indicate a preference for specific codons, suggesting bias towards certain amino acids, while values less than 1 suggest the opposite. Detailed codon usage for each amino acid is presented in Table 3. Within the mitogenome PCGs, a distinct preference for specific codons was observed beyond the standard AUG (Met), UCC(Ser), UGG (Trp), ACC, and ACA(Thr). For example, alanine (Ala) showed the highest preference for the codon GCU, with an RSCU value of 1.61. Additionally, most amino acids are represented by at least two different codons, whereas arginine, leucine, and serine each have six associated codons, as depicted in Fig. 3. These patterns align with findings from Xie's study [44], which reported no significant codon usage differences within the *Stemona* genus. Furthermore, among the 28 codons with RSCU values exceeding 1, 27 codons—representing 96.43%—showed a consistent preference for U/A-ending codons at the third position in the *S. tuberosa* mitogenome. This observation underscores a strong bias towards specific nucleotide endings in this species.

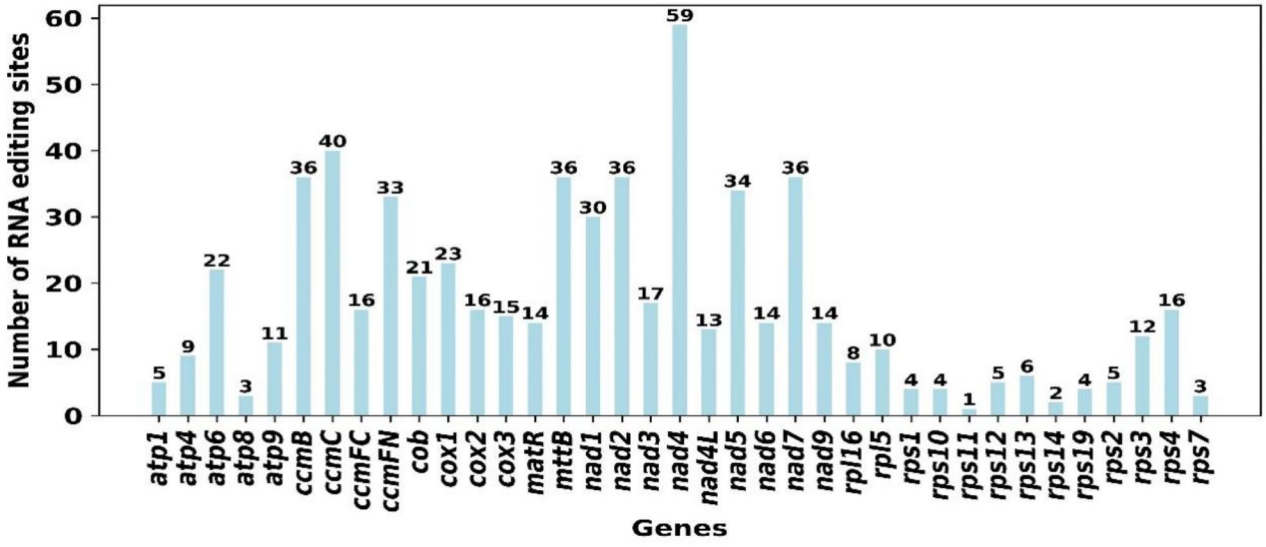

Fig. 5 Predicted RNA Editing Sites Based on Protein-Coding Genes. This bar chart displays the number of predicted RNA editing sites in various protein-coding genes. The x-axis represents different genes, while the y-axis indicates the number of RNA editing sites for each gene. Each bar corresponds to the number of editing sites in a gene, visually representing the distribution of editing sites across the genes

Repeat sequences and prediction of RNA editing events

In the *S. tuberosa* mitogenome, we identified a total of 274 simple sequence repeats (SSRs) distributed across three chromosomes: 236 in chromosome 1, 22 in chromosome 2, and 16 in chromosome 3. Monomeric repeats constituted the largest proportion of SSRs, accounting for 46.17%, 50.00%, and 43.75% in chromosome 1, 2, and 3, respectively (Fig. 4A). Notably, no pentameric or hexameric repeats were found in chromosome 2. Furthermore, we detected 27 tandem repeats within the mitogenome, ranging from 2 to 41 base pairs (bp). Of these, 23 were located on contig 1, while chromosome 2 and 3 each contained 2 tandem repeats. A detailed analysis revealed that over 70% of the 23 tandem repeats, ranging from 3 to 41 bp, were found on chromosome 1. Additionally, more than 77% of the tandem repeats, ranging from 2 to 21 bp, and over 89% of the tandem repeats, ranging from 4 to 5 bp, matched on chromosome 2 and 3, respectively.

Moreover, 180 dispersed repeats were identified across the three chromosomes, and each repeat being at least 30 bp in length, of which chromosome1 contained 169 of these dispersed repeats, predominantly in the form of forward (85) and palindromic (83) repeats, which comprised 50.31% and 49.11% of the repeats, respectively. In contrast, chromosome 2 contained 10 dispersed repeats, including palindromic (5), complementary (2), reverse (1), and forward (2) repeats. Only one dispersed repeat, a forward repeat, was identified in chromosome 3 (Fig. 4B). This comprehensive analysis highlights significant variability in repeat types and distributions across the contigs of the *S. tuberosa* mitogenome.

RNA editing events are pivotal in plant growth and development. In this study, we identified 633 RNA editing sites within the *S. tuberosa* mitogenome, across 38 unique protein-coding genes (Fig. 5), all involving cytidine to uridine (C to U) transitions. Supplementary Table S1 lists these 633 C to U editing sites. The *nad4* gene exhibited the highest number of editing sites, with 59 occurrences, followed by the *ccmC* gene with 40 sites. Our analysis revealed substantial variability in RNA editing site distribution across different mitochondrial genes. For instance, a significant concentration of RNA editing sites was found in the NADH dehydrogenase (*nad2*, *nad4*, and *nad7*), cytochrome c biogenesis (*ccmB* and *ccmC*), and protein transport subunit (*mttB*) genes. In contrast, no RNA editing sites were detected in the *rpl2* gene.

Further examination showed that most RNA editing sites were nonsynonymous, leading to changes in 19 types of amino acids. Conversely, synonymous editing, affecting codon usage without altering the encoded amino acid, was responsible for 10 types of amino acid conversions. These conversions included cysteine (1), valine (3), serine (4), leucine (5), isoleucine (5), phenylalanine (7), tyrosine (1), proline (3), arginine (1), and glycine (2). Interestingly, these synonymous changes primarily occurred at the third positions of codons, underscoring their role in amino acid variation.

Intracellular gene transfer (IGT)between chloroplast and mitochondrial organelles

Sequence alignment revealed 29 homologous fragments between chloroplast and mitochondrial organelles

Table 4 The homologous DNA fragment in the mitochondrial genome of *S. tuberosa*

Number	Identity(%)	Alignment Length(bp)	Chloroplast Genome		Mitochondrial Genome		MTPT Annotation
			Start	End	Start	End	
MTPT1	100	1668	33,504	31,837	328,618	330,285	partial *psbD*; partial *psbC*
MTPT2	100	41	119,515	119,555	51,682	51,722	IGS(*ndhA -ndhA*)
MTPT3	100	40	113,035	113,074	59,095	59,134	IGS(*rpl32 -trnL-UAG*)
MTPT4	100	28	98,674	98,647	418,506	418,533	IGS(*rps12 -trnV-GAC*)
	100	28	138,011	138,038	418,506	418,533	IGS(*trnV-GAC-rps12*)
MTPT5	99.922	6415	91,500	97,909	45,308	51,722	partial *ycf2*;complete *trnL-CAA*; complete *ndhB* ;complete *rps7* ;partial *rps12*
MTPT6	99.907	6418	145,185	138,773	45,308	51,725	partial *ycf2*
MTPT7	99.879	2484	2677	194	330,275	332,758	partial *psbA*; partial *trnK-UUU*; partial *matK*
MTPT8	99.749	2789	85,342	82,554	308,608	311,391	partial *rpl22*
	99.749	2789	151,343	154,131	308,608	311,391	partial trn*I-CAU*; complete *rpl23* ;complete *rpl2* ;complete trn*H-GUG*; complete *rps19* ;partial *rpl22*
MTPT9	99.698	331	13,697	14,027	62,614	62,944	partial *atpI*
MTPT10	99.694	4907	90,884	85,993	408,840	413,746	partial *ycf2*
	99.694	4907	145,801	150,692	408,840	413,746	partial *ycf2*
MTPT11	99.666	898	68,947	68,052	1	898	partial *clpP*
MTPT12	99.647	7082	50,321	57,398	147,318	154,384	complete *trnV-UAC*; complete *trnM-CAU*; complete *atpE* ;complete *atpB* ;complete *rbcL* ;partial *accD*
MTPT13	99.419	10,507	109,379	98,894	216,823	227,300	partial *ndhF*
MTPT14	99.415	2736	17,805	20,538	345,570	348,305	partial *rpoC2* ;partial *rpoC1*
MTPT15	99.18	122	14,012	14,133	38,630	38,751	partial *atpI*
MTPT16	98.963	482	63,317	62,836	317,321	317,801	partial *psbE*
MTPT17	97.966	934	69,146	70,075	407,907	408,840	partial *clpP*
MTPT18	97.952	14,798	35,610	50,328	132,434	147,178	partial *trnfM-CAU*; complete *rps14* ;complete *psaB* ;complete *psaA* ;complete *ycf3*;complete *trnS-GGA*; complete *rps4* ;complete *trnT-UGU*; complete *trnL-UAA*; complete *trnF-GAA*; complete *ndhJ* ;complete *ndhK* ;complete *ndhC*
MTPT19	97.619	84	107,768	107,851	313,736	313,819	complete *trnN-GUU*
	97.619	84	128,917	128,834	313,736	313,819	IGS(*trnN-GUU-trnN-GUU*)
MTPT20	94.382	267	56,832	57,097	53,625	53,889	partial *accD*
MTPT21	94.118	102	57,576	57,477	8821	8922	partial *accD*
MTPT22	93.671	79	51,355	51,277	397,675	397,753	IGS(*trnM-CAU-trnM-CAU*)
MTPT23	93.537	851	23,374	22,534	457,780	458,617	partial *rpoB*
MTPT24	90.551	127	107,482	107,356	15,405	15,527	partial *trnR-ACG*
	90.551	127	129,203	129,329	15,405	15,527	partial *trnR-ACG*
MTPT25	89.831	59	23,374	23,316	272,795	272,853	partial *rpoB*
MTPT26	88.667	150	33,748	33,894	400,691	400,840	partial *psbC*
MTPT27	87.665	1289	9307	10,547	462,156	463,420	partial *atpA*
MTPT28	86.498	237	64,365	64,600	401,280	401,507	IGS(*psbE -petL*)
MTPT29	81.366	483	65,665	65,201	7316	7779	IGS(*trnW-CCA-trnW-CCA*)

(MTPTs), as detailed in Table 4. Collectively, these transfer fragments span 66,408 bp, comprising 10.96% of the *S. tuberosa* mitogenome (Fig. 6). Notably, 11 of these 29 fragments exceed 1,000 bp in size. The largest of these, MTPT18, measures 14,798 bp, making it the most substantial fragment among the identified homologous sequences. Further annotation of these sequences revealed the presence of 25 complete genes, including 16 protein-coding genes (PCGs) and 9 tRNA genes. The PCGs identified are *atpB*, *atpE*, *ndhB*, *ndhC*, *ndhJ*, *ndhK*, *psaA*, *psaB*, *rbcL*, *rpl2*, *rpl23*, *rps14*, *rps19*, *rps4*, *rps7*, and *ycf3*. The tRNA genes include *trnF-GAA*, *trnH-GUG*, *trnL-CAA*, *trnL-UAA*, *trnM-CAU*, *trnN-GUU*, *trnS-GGA*, *trnT-UGU*, and *trnV-UAC*.

Analysis of the mitochondrial genome collinearity among *S. tuberosa* and other species

To better elucidate the conservatism of mitogenome evolution among *S. tuberosa* and five other species (*Crocus sativus*, *Phoenix dactylifera*, *Pandanus odoratifer*, *Zantedeschia aethiopica*, *Pinelliaternata*), MCscanX was employed to generate multiple synteny plots

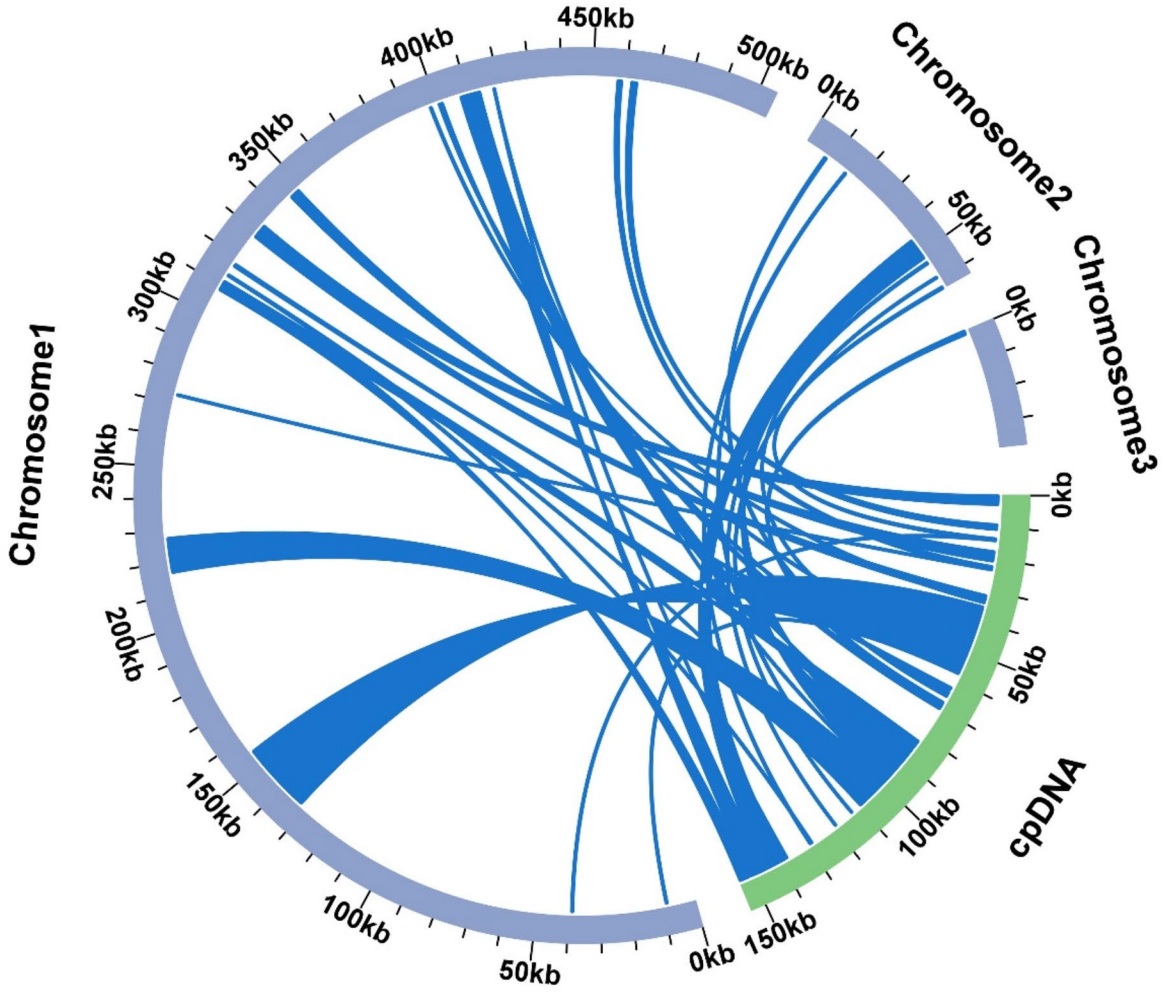

Fig. 6 Homologous analysis between two organelles. The blue arc represents mtDNA. The green arc represents chloroplast genome. The homologous fragments are indicated using the yellow lines between blue and green arcs

based on sequence similarity. Figure 7 illustrates varying arrangements of co-linear blocks across the mitogenomes of these species. The analysis revealed numerous homologous co-linear blocks, which were notably short in length. Additionally, some blocks were absent in the compared genomes, indicating sequences unique to the mitochondrial genome of *S. tuberosa*. Furthermore, the arrangement of these co-linear blocks varied among the six species, suggesting that their mitogenomes have undergone extensive gene rearrangements.

Phylogenetic analysis

Understanding the evolutionary status of plants is crucial. In present study, PhyloSuite software was utilized to extract 24 conserved protein-coding genes (PCGs) from the mitogenomes of 18 species across five orders—Asparagales, Pandanales, Arecales, Alismatales, and Ranunculales—with *Pulsatilla dahurica* (NC_071219.1) and *Aconitum kusnezoffii* (NC_053920.1) serving as outgroups (Fig. 8). These 24 PCGs included *atp1*, *atp4*, *atp6*, *atp8*, *atp9*, *ccmB*, *ccmC*, *ccmFC*, *ccmFN*, *cob*, *cox1*, *cox2*, *cox3*, *matR*, *mttB*, *nad1*, *nad2*, *nad3*, *nad4*, *nad4L*, *nad5*, *nad6*, *nad7*, and *nad9*. Phylogenetic analysis revealed that *S. tuberosa* and *Pandanus odorifer* within the Pandanales order clustered together with a 100% bootstrap support rate. This mitochondrial DNA-based phylogeny aligns with the most recent classification by the Angiosperm Phylogeny Group (APG), confirming the reliability of using plant mitochondrial protein-coding genes to construct a maximum likelihood (ML) phylogenetic tree.

Discussion

The mitochondrial and chloroplast genomes are both crucial for energy production and cellular metabolism in plants, yet they function in different aspects of cellular activity. The mitogenome of *S. tuberosa* is involved in

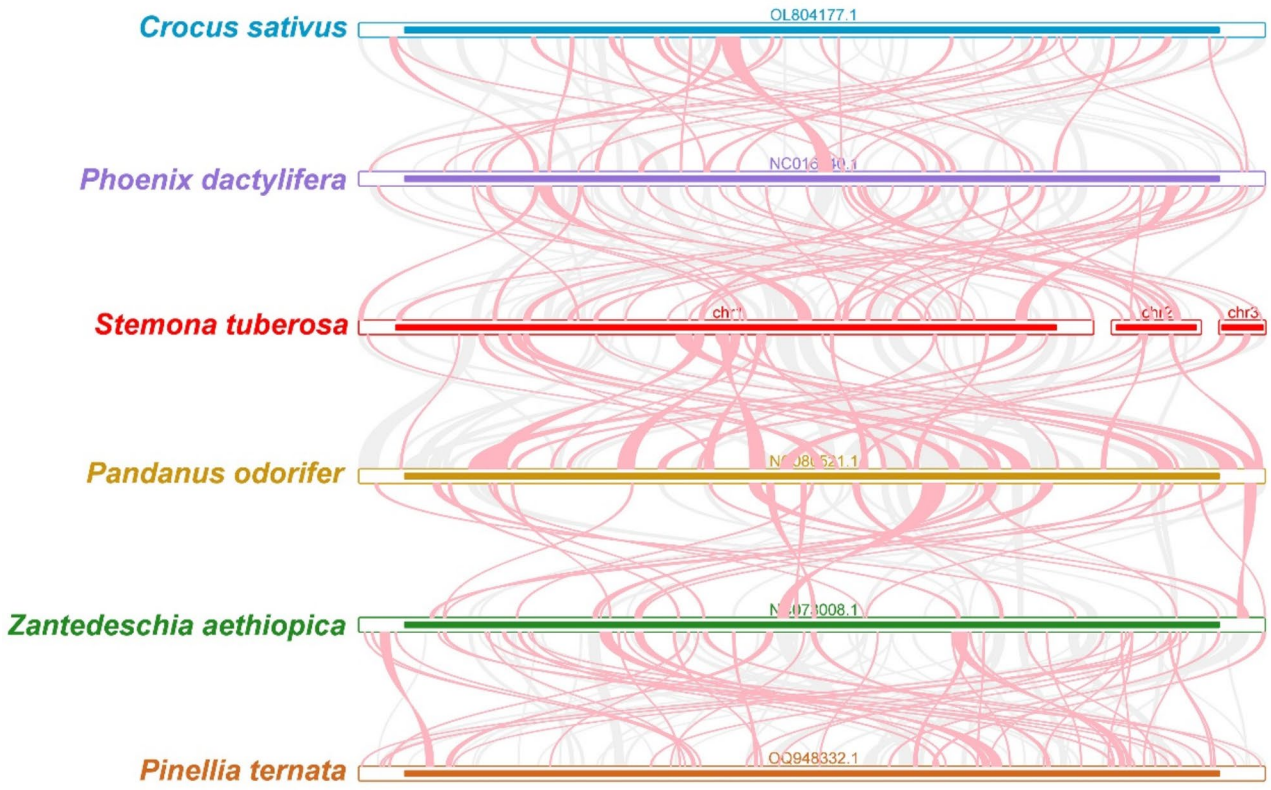

Fig. 7 Collinear analysis of sixspecies. The pink arcs indicated inverted regions. The gray arcs indicated better homologous regions. The regions with no colinear blocks are indicated as unique in the species

energy metabolism, supporting processes like respiration and ATP production, while the chloroplast genome plays a key role in photosynthesis. Despite their different functions, both genomes are involved in similar evolutionary processes, such as gene transfer and genome rearrangements. Compared to plant chloroplasts and animal mitogenomes, plant mitogenomes exhibit more complex and variable features. These include intricate structures and size differences, multipartite arrangements, low gene density, extensive post-transcriptional RNA editing, gene sequence transfer or loss, and foreign sequence capture [45]. To date, numerous plant mitogenomes have been characterized, revealing diverse structural variations such as multiple circular replicons, branched, linear, or mixed genomic structures [46]. Recent advancements in Illumina and Nanopore sequencing technologies have further highlighted the complexity of plant mitochondrial genomes. For instance, the mitogenomes of *Paphiopedilum micranthum*, *Salvia officinalis*, and *A. biserrata* consist of twenty-six, two, and six circular chromosomes, respectively [47, 48]. In this article, we sequenced the first complete mitogenome of *S. tuberosa*. The chloroplast genome of *S. tuberosa* is typical, presenting as a circular structure with tetrad features and a total length of 154,379 bp [8]. In contrast, the mitogenome of *S. tuberosa* consists of three circular chromosomes totaling 605,873 bp. This configuration differs markedly from that of *S. sessilifolia*—another member of the *Stemona* genus—which exhibits one linear and six circular chromosomes totaling 724,751 bp [44]. Besides, its close relative *P. odorifer*'s mitogenome exhibits one circular chromosome totaling 330,962 bp [49]. These results suggest that the presence of multiple molecular forms may be more common within the Stemona genus than previously anticipated.

The guanine-cytosine (GC) content plays a crucial role in determining the amino acid composition within protein groups during the evolutionary process among land plants [50]. The GC content of the *S. tuberosa* mitogenome is 45.63%, aligning closely with the GC content observed in the mitogenomes of other plant species such as *S. sessilifolia*, *(A) leptophyllum*, *(B) chinense*, and *S. divaricate* [44, 51, 52]. Although many studies have highlighted similarities in GC content across various plant mitogenomes, significant variations do exist among seed plants, underscoring the evolutionary diversity within this group.

Codon usage is significant in the context of genetic mutations, with a preference for specific synonymous codons playing a vital role in defining the genetic makeup of organisms. In this paper, an analysis of codon usage among the mitochondrial protein-coding genes (PCGs)

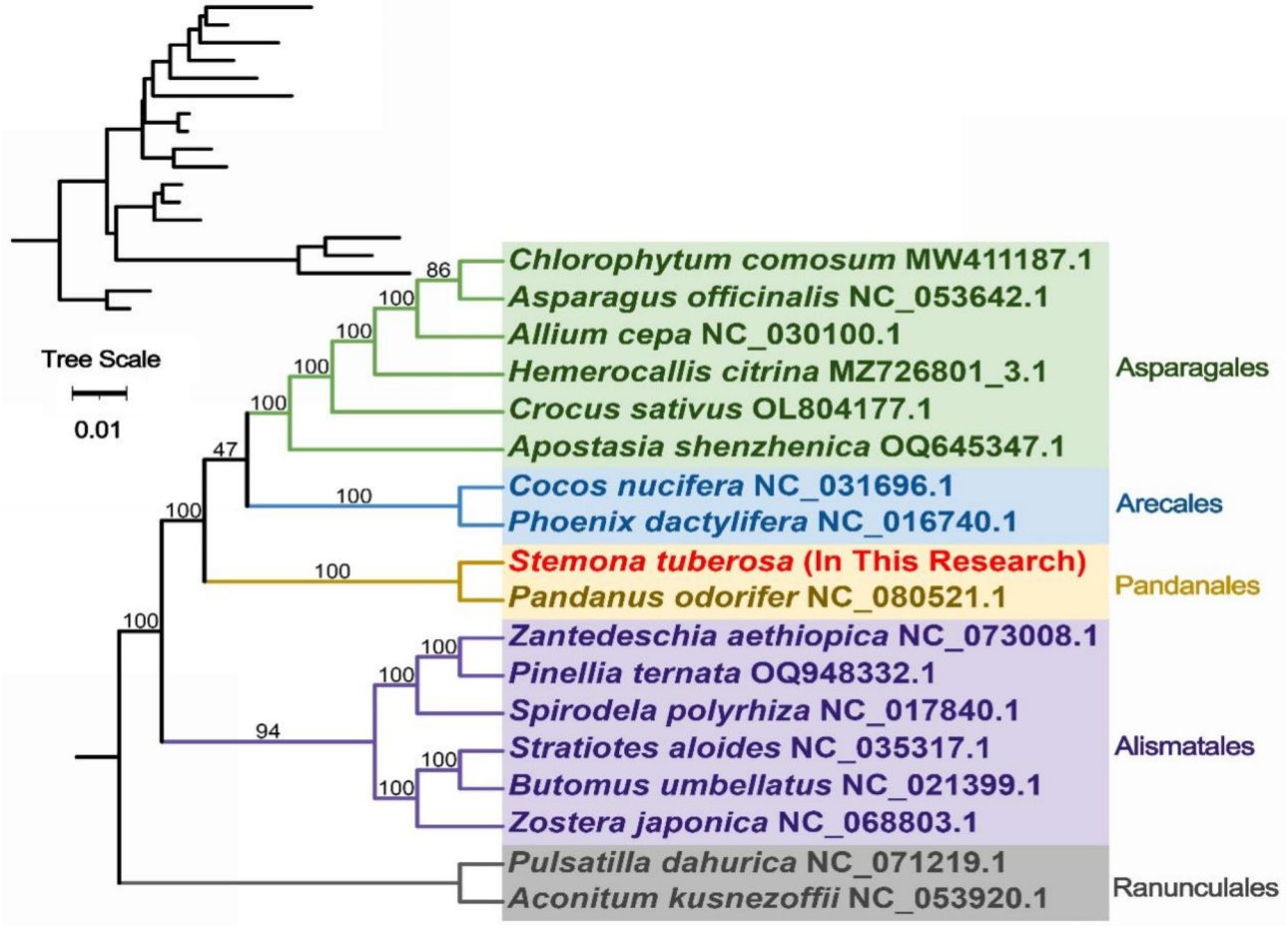

Fig. 8 Construction of the maximum likelihood tree based on the 18 species

in *S. tuberosa* revealed preferential codon usage for certain amino acids. For example, alanine (Ala) showed a marked preference for the codon GCU, while histidine (His) favored the codon CAU. Additionally, our findings indicated a tendency for U/A-ending codons at the third positions within the *S. tuberosa* mitogenome. This pattern contrasts with findings from other species such as *A. biserrata*, *Mangifera longipes*, *Mangifera persiciformis*, and *Mangifera sylvatica*, which tend to favor A/T bases and A/T-ending codons in the third positions [48, 53]. Understanding these codon usage patterns deepens our insight into the molecular evolution and functional constraints of mitochondrial genes, highlighting the nuanced differences that influence mitochondrial DNA evolution across species.

RNA editing events are highly frequent in plant mitogenomes and result in amino acid changes through insertions, deletions, and substitutions, thereby contributing to substantial genetic diversity [54]. Predicting potential RNA editing sites is essential for understanding the expression of plant mitochondrial genes. In present study, a total of 633 RNA editing sites across 38 unique mitochondrial protein-coding genes (PCGs) were identified. Predominantly, these edits were from cytosine to uridine (C to U), although guanine to uridine (G to U) and adenine to uridine (A to U) edits were also observed. These variations may be influenced by RNA structure or genetic differences between individuals, indicating a degree of diversity among species. Additionally, our results indicated that RNA editing sites predominantly affect amino acid changes at the first or second base positions of codons, with the second position experiencing more frequent alterations. This observation aligns with findings from previous studies, highlighting the significant impact of RNA editing on the functional dynamics of mitochondrial genes.

In our study, we conducted a homologous sequence analysis that revealed 29 homologous fragments (MTPTs), totaling 66,408 bp and constituting 10.96% of the *S. tuberosa* mitogenome, between the chloroplast and mitochondria. These mitochondrial plastid sequences (MTPTs) include complete and partial sequences of plastid protein-coding genes (PCGs), transfer RNA (tRNA), and ribosomal RNA (rRNA). The partial loss

of these plastid sequences suggests that they may have become nonfunctional pseudogenes in the mitogenome, although some tRNA genes might retain functionality [55]. This aligns with the hypothesis that DNA fragments from plastomes typically become nonfunctional upon transfer, underscoring the complexity of inter-organelle genetic exchange in plants. These fragments encompass 16 protein-coding genes and 9 tRNA genes, which are likely crucial for fundamental functions such as energy metabolism and translation. Prior studies on Amborella trichopoda and Liriodendron tulipifera supported the predominant direction of gene transfer from chloroplasts to mitochondria [56, 57]. For instance, the mitochondrial genome of *Salvia miltiorrhiza* contains gene fragments of chloroplast origin, providing direct evidence for the transfer of DNA segments from chloroplasts to mitochondria [58]. Additionally, research on *Saposhnikovia divaricata* has indicated the potential transfer of chloroplast repeat regions to mitochondria, further endorsing gene flow from chloroplasts to mitochondria in plants [59]. These findings not only enhance our understanding of the dynamics of plant mitochondrial genomes but also have significant implications for comprehending plant evolution and adaptability.

Repeated sequences are critical in shaping mitogenome structures through genome rearrangements, duplications, and recombination events. Previous studies have identified that three pairs of repetitive sequences mediated genome recombination into eight and seven different conformations in the mitogenomes of *Prunus salicina* and *I. batatas*, respectively. In the current study, a total of 274 simple sequence repeats (SSRs) were identified across all chromosomes, with monomeric polymers being the most prevalent. Dispersed and tandem repeats also showed variations in their distribution across different chromosomes. The findings are consistent with the findings in mitochondria of Stemonaceae species, including *S. mairei* [60], *S. sessilifolia* [44], and *S. parviflora* [60]. While we have confirmed the existence of these genomic structures, the specific functions they perform in the mitochondrial context remain to be further investigated.

Conclusions

This study provides the first detailed analysis of the mitogenome of *S. tuberosa*, revealing its unique multi-branched structure. The *S. tuberosa* mitogenome consists of three circular contigs with a total length of 605,873 bp, and 66 genes were annotated, including 38 protein-coding genes, 25 tRNA genes, and 3 rRNA genes. Our findings on codon usage patterns, RNA editing sites, and repeat sequences significantly enhance our understanding of the genetic characteristics and evolutionary dynamics of *S. tuberosa*. Notably, the mitogenome exhibits a preference for U/A-ending codons at the third positions, differing from previous studies and indicating diversity in mitochondrial codon usage bias across species. Additionally, RNA editing events are predominantly C-to-U, with some G-to-U and A-to-U edits, which may be influenced by RNA structure or genetic variations. Future studies should focus on the impact of RNA editing on mitochondrial gene expression in *S. tuberosa* to further elucidate its population genetics and evolutionary processes. Including more species from the *Stemona* genus will also enrich future analyses and offer broader insights into their evolutionary patterns.

Abbreviations

PCGs Protein	coding genes
RSCU	Relative synonymous codon usage
MTPT	Mitochondrial plastid DNA sequence
tRNA	Transfer RNA
rRNA	Ribosomal RNA
APG	Angiosperm phylogeny group
SSRs	Simple sequence repeats
IGT	Intracellular gene transfer
STR	Short tandem repeats

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-06034-z.

Supplementary Material 1

Acknowledgements

We thank the Editor and the anonymous reviewers for their insightful comments and suggestions on the manuscript. The authors thank Wuhan Benagen Technology Co., Ltd. for help in genome sequencing and the analysis of RNA editing.

Author contributions

D.X. and T.W. collaborated on the analysis and writing of this manuscript. J.H. Q.W. provided the material. Z.D.W. D.Q.Z and Z.X. undertook the formal identification of the plant material. X.L and L.F contributed to the design and editing of this manuscript. All authors reviewed and approved the final manuscript.

Funding

This work was supported by the Sichuan innovation team of national modern agricultural industry technology system, Foundation [SCCXTD-2024-19] and Dazhou Technological Innovation Special Project [24ZDYF0016].

Data availability

The datasets presented in this study can be found in onlinerepositories. The names of the repository/repositories and accessionnumber(s) can be found below: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1164997,https://www.ncbi.nlm.nih.gov/biosample/SAMN43911811,https://www.ncbi.nlm.nih.gov/sra/SRR30802486,https://www.ncbi.nlm.nih.gov/sra/SRR30802487,https://www.ncbi.nlm.nih.gov/nuccore/PQ374236,https://www.ncbi.nlm.nih.gov/nuccore/PQ374237,https://www.ncbi.nlm.nih.gov/nuccore/PQ374238.

Declarations

Ethical approval and consent to participate

We collected fresh leaf materials of *Stemona tuberosa* for this study. The study, including plant samples, complies with relevant institutional, national, and international guidelines and legislation. No specifc permits were required for plant collection.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 8 October 2024 / Accepted: 30 December 2024
Published online: 07 January 2025

References

1. Zhang RR, Tian HY, Wu Y, Sun XH, Zhang JL, Ma ZG, Jiang RW. Isolation and chemotaxonomic significance of stenine- and stemoninine-type alkaloids from the roots of *Stemona tuberosa*. Chin Chem Lett. 2014;25(9):1252–5.
2. Chung HS, Hon PM, Lin G, But PP, Dong H. Antitussive activity of *Stemona* alkaloids from *Stemona tuberosa*. Planta Med. 2003;69(10):914–20.
3. Xu YT, Hon PM, Jiang RW, Cheng L, Li SH, Chan YP, Xu HX, Shaw PC, But PP. Antitussive effects of *Stemona tuberosa* with different chemical profiles. J Ethnopharmacol. 2006;108(1):46–53.
4. Xu YT, Shaw PC, Jiang RW, Hon PM, Chan YM, But PP. Antitussive and central respiratory depressant effects of *Stemona tuberosa*. J Ethnopharmacol. 2010;128(3):679–84.
5. Huang SZ, Kong FD, Ma QY, Guo ZK, Zhou LM, Wang Q, Dai HF, Zhao YX. Nematicidal *Stemona* alkaloids from *Stemona parviflora*. J Nat Prod. 2016;79(10):2599–605.
6. Zhang X, Ge J, Yang J, Dunn B, Chen G. Genetic diversity of *Stemona parviflora*: a threatened myrmecochorous medicinal plant in China. Biochem Syst Ecol. 2017;71:193–169.
7. Liu J, Jiang M, Chen H, Liu Y, Liu C, Wu W. Comparative genome analysis revealed gene inversions, boundary expansions and contractions, and gene loss in the Stemona sessilifolia (miq.) Miq. Chloroplast genome. PLoS ONE. 2021;16(6):e0247736.
8. Lian Y, Huang F, Zhu WT, Liu XF, Wu H, Jiang GH, Yin XM. Chloroplast genome structure of *Stemona tuberosa* and phylogenetic analysis based on PacBio sequencing. Chin J Experimental Traditional Med Formulae. 2023;29(14):123–32.
9. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA. 1987;84(24):9054–8.
10. Torre S, Sebastiani F, Burbui G, Pecori F, Pepori AL, Passeri I, Ghelardini L, Selvaggi A, Santini A. Novel insights into refugia at the southern margin of the distribution range of the endangered species *Ulmus Laevis*. Front Plant Sci. 2022;13:826158.
11. Xiao S, Xu P, Deng Y, Dai X, Zhao L, Heider B, Zhang A, Zhou Z, Cao Q. Comparative analysis of chloroplast genomes of cultivars and wild species of sweetpotato (*Ipomoea batatas* [L.] Lam). BMC Genomics. 2021;22(1):262.
12. Huang DI, Hefer CA, Kolosova N, Douglas CJ, Cronk QCB. Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely related balsam poplars, Populus balsamifera and *P. trichocarpa* (Salicaceae). New Phytol. 2014;204(3):693–703.
13. Palmer JD, Herbon LA. Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol. 1988;28(1–2):87–97.
14. Gualberto JM, Mileshina D, Wallet C, Niazi AK, Weber-Lotfi F, Dietrich A. The plant mitochondrial genome: dynamics and maintenance. Biochimie. 2014;100:107–20.
15. Yurina NP, Odintsova MS. Mitochondrial genome structure of photosynthetic eukaryotes. Biochem (Mosc). 2016;81(2):101–13.
16. Allen JO, Fauron CM, Minx P, Roark L, Oddiraju S, Lin GN, Meyer L, Sun H, Kim K, Wang C, Du F, Xu D, Gibson M, Cifrese J, Clifton SW, Newton KJ. Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. Genetics. 2007;177(2):1173–92.
17. He T, Ding X, Zhang H, Li Y, Chen L, Wang T, Yang L, Nie Z, Song Q, Gai J, Yang S. Comparative analysis of mitochondrial genomes of soybean cytoplasmic male-sterile lines and their maintainer lines. Funct Integr Genomics. 2021;21(1):43–57.
18. Wang Y, Chen S, Chen J, Chen C, Lin X, Peng H, Zhao Q, Wang X. Characterization and phylogenetic analysis of the complete mitochondrial genome sequence of *Photinia serratifolia*. Sci Rep. 2023;13(1):770.
19. Jackman SD, Coombe L, Warren RL, Kirk H, Trinh E, MacLeod T, Pleasance S, Pandoh P, Zhao Y, Coope RJ, Bousquet J, Bohlmann J, Jones SJM, Birol I. Complete mitochondrial genome of a gymnosperm, sitka spruce (*Picea sitchensis*), indicates a complex physical structure. Genome Biol Evol. 2020;12(7):1174–9.
20. Sloan DB, Wu Z, Sharbrough J. Correction of persistent errors in *Arabidopsis* reference mitochondrial genomes. Plant Cell. 2018;30(3):525–7.
21. 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.
22. Alverson AJ, Rice DW, Dickinson S, Barry K, Palmer JD. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. Plant Cell. 2011;23(7):2499–513.
23. Tsujimura M, Kaneko T, Sakamoto T, Kimura S, Shigyo M, Yamagishi H, Terachi T. Multichromosomal structure of the onion mitochondrial genome and a transcript analysis. Mitochondrion. 2019;46:179–86.
24. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.
25. Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 2020;21(1):241.
26. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 2015;31(20):3350–2.
27. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–60.
28. 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):6–11.
29. 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.
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. Lewis SE, Searle SM, Harris N, Gibson M, Lyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby MA, Kaminker JS, Matthews BB, Prochnik SE, Smithy CD, Tupy JL, Rubin GM, Misra S, Mungall CJ, Clamp ME. Apollo: a sequence annotation editor. Genome Biol. 2002. https://doi.org/10.1186/gb-2002-3-12-research0082
32. 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.
33. 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.
34. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33(16):2583–5.
35. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80.
36. 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.
37. Zhang H, Meltzer P, Davis S. RCircos: an R package for Circos 2D track plots. BMC Bioinformatics. 2013;14:244.
38. 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):255.
39. 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):65–73.
40. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40(7):e49.
41. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66.
42. 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):268–24.
43. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256–259.
44. Xie Y, Liu W, Guo L, Zhang X. Mitochondrial genome complexity in *Stemona sessilifolia*: nanopore sequencing reveals chloroplast gene transfer and DNA rearrangements. Front Genet. 2024;15:1395805.

45. Hong Z, Liao X, Ye Y, Zhang N, Yang Z, Zhu W, Gao W, Sharbrough J, Tembrock LR, Xu D, Wu Z. A complete mitochondrial genome for fragrant Chinese rosewood (*Dalbergia Odorifera*, Fabaceae) with comparative analyses of genome structure and intergenomic sequence transfers. BMC Genomics. 2021;22(1):672.
46. Doré G, Barloy D, Barloy-Hubler F. De novo hybrid assembly unveils multi-chromosomal mitochondrial genomes in *Ludwigia* species, highlighting genomic recombination, gene transfer, and RNA editing events. Int J Mol Sci. 2024;25(13):7283.
47. Yang JX, Dierckxsens N, Bai MZ, Guo YY. Multichromosomal mitochondrial genome of paphiopedilum micranthum: compact and fragmented genome, and rampant intracellular gene transfer. Int J Mol Sci. 2023;24(4):3976.
48. Wang L, Liu X, Xu Y, Zhang Z, Wei Y, Hu Y, Zheng C, Qu X. Assembly and comparative analysis of the first complete mitochondrial genome of a traditional Chinese medicine *Angelica Biserrata* (Shan Et Yuan) Yuan Et Shan. Int J Biol Macromol. 2024;257(Pt 1):128571. https://doi.org/10.1016/j.ijbiomac.2023.128571
49. Darshetkar AM, Patil SS, Pable AA, Nadaf AB. Loss of mitogenome-encoded respiratory genes complemented by nuclear transcripts in halophyte *Pandanus odorifer* (Forssk.) Kuntze. Plant Biotechnol Rep. 2024;18:91–104.
50. Wang YR. Mitochondrial genome of *Selaginella Sinensis* (Desv.) Spring (Lycophyte). Master's thesis. Hainan University. 2020; https://doi.org/10.27073/d.cnki.ghadu.2020.001092
51. Ma Q, Wang Y, Li S, Wen J, Zhu L, Yan K, Du Y, Ren J, Li S, Chen Z, Bi C, Li Q. Assembly and comparative analysis of the first complete mitochondrial genome of *Acer Truncatum* Bunge: a woody oil-tree species producing nervonic acid. BMC Plant Biol. 2022;22(1):29.
52. Wicke S, Schneeweiss GM, dePamphilis CW, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 2011;76(3–5):273–97.
53. Niu Y, Gao C, Liu J. Complete mitochondrial genomes of three *Mangifera* species, their genomic structure and gene transfer from chloroplast genomes. BMC Genomics. 2022;23(1):147.
54. Sun T, Bentolila S, Hanson MR. The unexpected diversity of plant organelle RNA editosomes. Trends Plant Sci. 2016;21(11):962–73.
55. Kitazaki K, Kubo T, Kagami H, Matsumoto T, Fujita A, Matsuhira H, Matsunaga M, Mikami T. A horizontally transferred tRNA(Cys) gene in the sugar beet mitochondrial genome: evidence that the gene is present in diverse angiosperms and its transcript is aminoacylated. Plant J. 2011;68(2):262–72.
56. Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, Barry K, Boore JL, Zhang Y, dePamphilis CW, Knox EB, Palmer JD. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. Science. 2013;342(6165):1468–73.
57. 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;11:29.
58. Yang H, Chen H, Ni Y, Li J, Cai Y, Ma B, Yu J, Wang J, Liu C. De novo hybrid assembly of the *Salvia miltiorrhiza* mitochondrial genome provides the first evidence of the multi-chromosomal mitochondrial DNA structure of *Salvia* species. Int J Mol Sci. 2022;23(22):14267.
59. Ni Y, Li J, Chen H, Yue J, Chen P, Liu C. Comparative analysis of the chloroplast and mitochondrial genomes of *Saposhnikovia divaricata* revealed the possible transfer of plastome repeat regions into the mitogenome. BMC Genomics. 2022;23(1):570.
60. 60 Lu Q, Ye W, Lu R, Xu W, Qiu Y: Phylogenomic and comparative analyses of complete plastomes of Croomia and Stemona (Stemonaceae). Int J Mol Sci. 2018;19(8):2383.
61. Wei R, Li Q. The complete chloroplast genome of endangered species *Stemona parviflora*: insight into the phylogenetic relationship and conservation implications. Genes (Basel). 2022;13(8):1361.

Publisher's note