BMC Genomics

RESEARCH

Open Access

The first complete mitochondrial genome of *Curcuma amarissima* (Zingiberaceae): insights into multi-branch structure, codon usage, and phylogenetic evolution

Heng Liang[1,4†], Jiabin Deng[3,4†], Yidan Wang[5], Gang Gao[2,4*] and Ruiwu Yang[4]

Abstract

Background As a key genus in Zingiberaceae, *Curcuma* is widely studied for its taxonomic diversity, the presence of bioactive curcuminoids and volatile oils, and its extensive applications in traditional medicine and economic products such as spices and cosmetics. Although chloroplast genomes have been assembled and published for over 20 *Curcuma* species, mitochondrial genomic data remain limited.

Results We successfully sequenced, assembled, and annotated the mitogenome of *Curcuma amarissima* (*C. amarissima*) using both Illumina short reads and Nanopore long reads, achieving the first complete mitogenome characterization in the Zingiberaceae family. The *C. amarissima* mitogenome features a unique multi-branched structure, spanning 6,505,655 bp and consisting of 39 distinct segments. It contains a total of 43 protein-coding genes, 63 tRNA genes, and 4 rRNA genes, with a GC content of 44.04%. Codon usage analysis indicated a weak bias, with neutrality plot analysis suggesting natural selection as a key factor shaping mitochondrial codon usage in *C. amarissima*. The mitogenome provides valuable insights into genome size, coding genes, structural features, RNA editing, repetitive sequences, and sequence migration, enhancing our understanding of the evolution and molecular biology of multi-branched mitochondria in Zingiberaceae. The high frequency of repeat sequences may contribute to the structural stability of the mitochondria. Comparing chloroplast genome, phylogenetic analysis based on the mitochondrial genome establishes a foundation for further exploration of evolutionary relationships within Zingiberaceae.

Conclusions In short, the mitochondrial genome characterized here advances our understanding of multi-branched mitogenome organization in Zingiberaceae and offers useful genomic resources that may support future breeding, germplasm conservation, and phylogenetic studies, though further research is necessary.

Keywords Mitogenome, *Curcuma amarissima*, Comparative genomics, Phylogeny

†Heng Liang and Jiabin Deng contributed equally to this work.

*Correspondence:
Gang Gao
gaogang870522@163.com
Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Background

Plant mitochondria, organelles measuring about 0.5 to 1.0 μm in diameter and enclosed by a double membrane, contain their own genetic material and system [1], is critical for plant development and growth [2]. As the main site for aerobic respiration and energy production, they are widely present in eukaryotic cells [3]. Mitochondria are essential for energy production and management, supplying the energy required for diverse cellular functions, and they influence cell growth and the cell cycle [4]. Additionally, mitochondria play a vital role in essential cellular processes, including differentiation, apoptosis, division, and the synthesis and breakdown of specific molecules [5]. Mitochondria originated from ancient bacteria capable of respiration, with their cell membranes folding inward to form a primitive membrane structure with oxidative respiration and genetic material, eventually evolving into modern mitochondria over a long evolutionary history [6].

The mitogenome in plants is more structurally complex than the chloroplast genome [7]. Recent advancements in sequencing and assembly techniques have facilitated the assembly of numerous plant mitogenomes, enhancing our understanding of their complex structures [8]. Plant mitogenomes display diverse complex structures, such as circular master molecules, circular subgenomic forms, linear molecules, and highly branched variants [9]. For example, both "master circles" and multiple "subgenomic circles" have been identified in the mitogenome of *Panax notoginseng* [10]. Additionally, recent studies have revealed that some plant mitogenomes also display multi-branch structures [11]. Mitochondria are specialized organelles with unique functions, playing a crucial role in long-term symbiosis. They participate in various metabolic activities, including energy generation and the synthesis and breakdown of different compounds. The mitogenome contains a amount of repetitive sequences, RNA editing sites, and nucleotide polymorphisms, all of which are linked to mitogenome variation and recombination [12]. The exchange of similar genetic fragments among different genomes plays a key role in the evolution of plant mitogenomes. During prolonged symbiosis, mitochondria have lost some of their original DNA, with only the transferred portions responsible for encoding certain functions remaining [6]. Exploring mitogenome variations through complete genome analysis offers certain advantages in phylogenetic reconstruction [13]. For most plants, both chloroplast and mitogenomes are maternally inherited [20], a genetic mechanism that simplifies genetic studies and has recently been widely used for deeper phylogenetic inference in taxonomy [14, 15].

Curcuma, a genus within the Zingiberaceae family with over 100 species, consists of perennial herbaceous plants with fleshy, aromatic rhizomes. The Plant List (http://www.theplantlist.org/) indicates that these plants are typically found in tropical and subtropical areas. Most species are mainly located in East and Southeast Asia, with some also present in northern Australia and the South Pacific. Research indicates that *Curcuma* species contain over 40 chemical compounds, with their rhizomes and tubers being particularly rich in curcuminoids and volatile oils, and also possess resins, sugars, fatty acids, peptides, alkaloids, curcumin acid, and trace elements [16]. Beyond their medicinal USES, *Curcuma* are economically valuable in the production of pigments, spices, dyes, cosmetics, and insecticides [17, 18]. As a key genus in the Zingiberaceae family, *Curcuma* holds importance in plant taxonomy, phytochemistry, and medicinal botany.

Although research on the biological characteristics of *Curcuma* continues to expand, there are still relatively few studies specifically focused on mitochondrial genomes within this genus, as well as within Zingiberaceae and Zingiberales, limiting our comprehensive understanding of mitochondrial biology in these taxa. Our study provides the first complete mitochondrial genome of *C. amarissima*, offering valuable preliminary insights into mitochondrial genomic structure and biology. The results presented here can serve as an important initial resource to support further mitogenome research and comparative analyses within the Zingiberaceae family and beyond.

Materials and methods

Sequencing, assembly, and annotation of the mitogenome

Samples of *C. amarissima* leaves were collected at MengA Farm in Xishuangbanna, Yunnan, China. They were immediately frozen in liquid nitrogen and subsequently stored at − 80 °C. High-quality genomic DNA was obtained using the TIANamp Genomic DNA Kit from Tiangen, known for its efficient and high-purity extraction method. We utilized a combination of third-generation sequencing via Oxford Nanopore Technologies (ONT) and second-generation sequencing using Illumina to construct the genomic library. The second-generation sequencing experiment was carried out according to the standard protocol provided by Illumina, including sample quality assessment, library construction, library quality control, and sequencing. After confirming the quality of the genomic DNA samples, the DNA was fragmented using ultrasonic sonication. The fragmented DNA then underwent purification, end repair, addition of a 3' A tail, and ligation of sequencing adapters. Size selection was performed using agarose gel electrophoresis, followed by PCR amplification to create the sequencing library. The constructed library was quality checked, and the qualified library was sequenced using the Illumina Novaseq6000

platform. After confirming the quality of the samples, genomic DNA is randomly fragmented. Large fragments of DNA are then enriched and purified using magnetic beads, followed by gel extraction. The fragmented DNA undergoes damage repair. After purification, end repair and 3' A addition are performed on the DNA fragments, and the adapters from the SQK-LSK109 kit are ligated. Finally, the constructed DNA library is quantified. Once the library is complete, a specific concentration and volume of the DNA library is added to the flow cell, which is then transferred to the Oxford Nanopore PromethION sequencer for real-time single-molecule sequencing. The third-generation sequencing data is filtered using Filtlong (v0.2.1, https://github.com/rrwick/Filtlong) with the parameters: –min_length 1000 –min_mean_q 7.

The mitogenome was assembled using the following methods. We utilized minimap2 (v.2.24) [19] to align sequences from closely related species. The aligned data were then processed with Flye (v.2.9.3) [20] and corrected using Racon [21]. Subsequently, second-generation data were aligned to the corrected results with Bowtie2 [22]. Finally, Unicycler (v.0.5.1) [23] was employed for hybrid assembly, and the genome was circularized based on coverage from third-generation data to obtain the final result.

For mitogenome annotation, BlastN [24] was initially employed to identify mitochondrial genes against the NCBI database. Next, tRNA genes were identified using tRNA scan-SE software (v.2.0.12) (http://lowelab.ucsc.edu/tRNAscan-SE/). The boundaries of protein-coding genes (PCGs) and rRNA genes were established by comparing them with homologous genes in other species while considering tRNA gene locations. To further validate the predicted positions, we used R script to assess the completeness of the nucleotide sequences of the PCGs and to translate them into amino acids. The mitogenome map was created using OGDRAW [25].

Analysis of relative synonymous codon usage (RSCU) and Effective Number of Codons (ENC)

Codon usage bias, the variation in synonymous codon frequency, commonly influences speciation and evolutionary divergence among species [26]. MEGA7 [27] was used to examine codon usage patterns, calculating the RSCU for the mitogenome's PCGs. Additionally, the online tool CUSP (https://www.bioinformatics.nl/cgi-bin/emboss/cusp) was utilized to assess the GC content of these coding genes. This combined analysis provides a comprehensive view of codon bias and nucleotide composition in *C. amarissima*'s mitogenome. The effective number of codons (ENC) was assessed with CodonW [28] to reflect deviations in codon usage.

RNA editing sites of *C. amarissima*

A previous study demonstrated that mitochondrial genes undergo numerous RNA editing events [29]. RNA editing sites for the protein-coding genes in each species were identified using PREPACT [30], combined with BLASTx (e-value cutoff =0.001, filter threshold =30%). A Perl script was then used to generate the coding sequences resulting from RNA editing for the species.

Analysis of repeat sequences

Repetitive sequences are features of genomes, influencing their evolution, inheritance, and variability. They are crucial for gene expression, transcription regulation, chromosomal structure, and metabolic processes. Simple sequence repeats (SSRs) were detected using MISA software (v1.0) with parameters set to 1–10, 2–5, 3–4, 4–3, 5–3, and 6–3. Tandem repeats were identified using TRF software (trf409.linux64) with the following settings: 2, 7, 7, 80, 10, 50, 2000, -f, -d, -m. For dispersed repeats, BLASTN (v2.10.1) was employed with a word size of 7 and an e-value of 1e- 5, ensuring that redundancy and tandem repeats were excluded. The final results were visualized using Circos (v0.69–5).

Migration analyses of chloroplast-to-mitochondrial DNA

The assembly of the chloroplast genome of *C. amarissima* was primarily based on NGS data. We used GetOrganelle for the assembly, employing default parameters. Annotation was carried out using CPGAVAS2, and all annotation results were subsequently edited and corrected manually to ensure accuracy. Homologous sequences between the chloroplast and mitogenomes were identified using BLAST (version 2.6, https://blast.ncbi.nlm.nih.gov/Blast.cgi), applying an E-value threshold of 1e- 5 and requiring at least 70% similarity.

Comparative genomics analysis

For lack of published mitochondrial Zingiberales genomes, we selected the closest relative and representative species, *Pontederia crassipes* (PP905391), *Aechmea fasciata* (PP436815), and *Phoenix dactylifera* (JN375330) from sister order of the Commelinanae for comparative analysis. To evaluate the relationship between the mitogenome of *C. amarissima* and other species within Commelinanae, we performed a pairwise alignment of mitogenomes with BLASTN (v2.10.1) [31], identifying homologous sequences with a minimum collinear block length cutoff of 500 bp. Visualization and collinearity analysis for the four species were conducted using Mauve software [32].

We used Phylosuite software to extract PCGs and identified homologous sequences through BLAST.

These sequences were aligned using MAFFT (v7.149b) [33]. Following alignment, Ka and Ks values for each gene pair were calculated with KaKs_Calculator v2.0 (available at https://sourceforge.net/projects/kakscalculator2/), applying the MLWL method. We compiled the Ka/Ks ratios for all gene pairs and created box plots to visualize the results. Nucleotide diversity (Pi) for PCGs genes was calculated by using DnaSP6 [34].

Phylogenetic tree construction

Phylogenetic trees were constructed using both maximum likelihood (ML) and Bayesian inference (BI) methods. The phylogenetic analysis involved a total of 9 species, *Pontederia crassipes*, *Avena longiglumis*, *Saccharum officinarum*, *Aechmea fasciata*, *Poa pratensis*, *Oryza sativa*, *Phoenix dactylifera*, *Apostasia fujianica* and *Dracaena cochinchinensis*, with PCGs from both mitochondrial and chloroplast genomes. We determined the best evolutionary model for the PCGs using ModelTest-NG based on the AIC criterion. Maximum likelihood (ML) analysis was performed on both datasets with RAxML-NG [35], utilizing 1,000 rapid bootstrap replicates. For Bayesian inference (BI), we ran MrBayes (v3.2.7) with a Markov Chain Monte Carlo (MCMC) analysis over 1,000,000 generations, sampling every 100 generations and discarding 25% of the phylogenetic trees as burn-in [36].

Results

Multibranched structure assembly and annotation of the *C. amarissima* Mitogenome

After sequencing and assembly, and excluding nodes originating from nuclear and chloroplast genomes, we successfully generated the complete mitogenome of *C. amarissima* using the '3 + 2' method strategy with Bandage (Fig. 1). The mitogenome of *C. amarissima,* with 6,505,655 bp in total size, displayed a complex multibranched conformation, assembled into 39 segments, and contains a GC content of 44.04%. Only contig 15 showed a single circular molecule, other 38 segments organized into a linear molecule. The segments varied in size, with contig 39 measuring 2,651 bp and contig 1 measuring 675,091 bp. Contig 35 exhibited the highest GC content at 48.72%, while contig 38 had the lowest at 37.48%.

In annotating the 39 segments of *C. amarissima*, we identified a total of 43 unique PCGs, which included 24 core mitochondrial genes and 12 non-core genes, along with 63 tRNA genes (two of which were multicopy), four rRNA genes (including one multicopy), and 11 pseudogenes (Table S1). The 24 core genes consisted of 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 cytochrome c oxidase genes (*cox1*, *cox2*, and *cox3*), a transport membrane protein (*mttB*), maturases (*matR*),

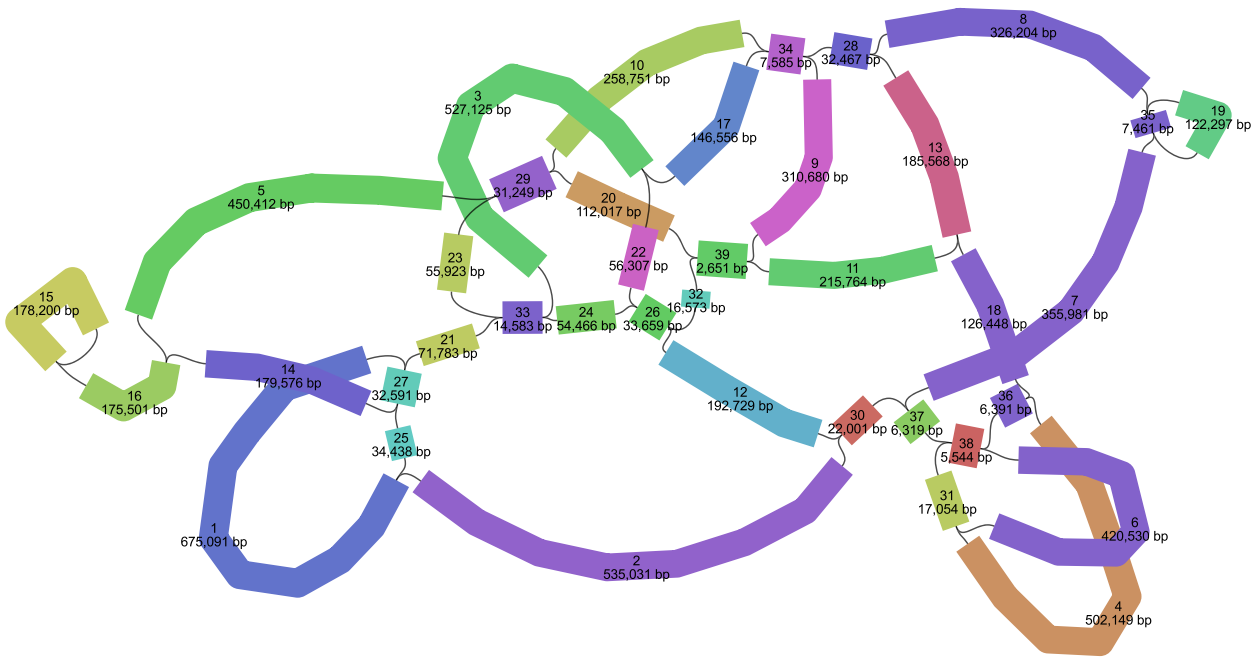

Fig. 1 The assembly result of the mitochondrial genome of *C. amarissima*

and a ubichinol cytochrome c reductase gene (*cob*). The non-core genes included four large ribosomal subunits (*rpl10*, *rpl16*, *rp2*, and *rpl5*), eight small ribosomal subunits (*rps1*, *rps12*, *rps13*, *rps14*, *rps19*, *rps3*, *rps4*, and *rps7*), along with two succinate dehydrogenase genes (*sdh3* and *sdh4*). The arrangement and orientation of these genes are depicted in Fig. S1. No complete genes were found in contigs 20, 21, 25, 27, 30, 32, 33, 35, 36, 37, 38, and 39.

Codon usage patterns of PCGs and RNA editing site predictions

The total length of PCGs in the *C. amarissima* mitogenome is 34,771 bases, with the majority starting with the standard ATG codon (Table S1). Notably, the genes *nad1* and *rps10* utilize ACG as their start codon, likely due to C-to-U RNA editing at the second position. Three termination codons (TAA, TGA, and TAG) were identified, but no RNA editing was detected in these codons.

The analysis of codon usage in the mitogenome of *C. amarissima* revealed distinct patterns in codon preference (Fig. 2A and Table S2). Methionine (Met) was predominantly encoded by AUG, which exhibited the highest relative synonymous codon usage (RSCU) value of 3.00 among the PCGs. Alanine (Ala) was mainly represented by GCU, showing the second highest RSCU value of 1.61. Additionally, histidine (His), glutamine (Gln), and tyrosine (Tyr) all demonstrated strong preferences, with maximum RSCU values exceeding 1.45. In plant mitogenomes, ATG serves as the sole initiation codon, resulting in the lowest RSCU values of 0 for CUG and UUG, both encoding Met. A notable preference for A or T(U) at the third position was observed across the mitogenome. The effective number of codons (ENC) for the PCGs ranged from 35.64% to 61%, with an average ENC exceeding 50, indicating relatively weak codon usage bias (Table S3). Neutrality plot analysis revealed a low correlation between GC12 and GC3 (Fig. 2B), indicating that natural selection significantly influences mitochondrial codon usage. To explore factors influencing codon patterns, we plotted ENC values against GC3 for the PCGs (Fig. 2C and Table S3). Most genes were positioned below the standard curve, except for three genes (*rps11*, *rps14*, and *atp4*), indicating that selection pressure likely influences codon preferences.

We predicted and analyzed RNA editing events in the mitogenome (Fig. 2D). A total of 35 unique PCGs were assessed, resulting in the identification of 492 potential RNA editing sites, all of which involved cytidine to uridine (C–T/U) conversions. The *ccmFn* gene exhibited the highest number of RNA editing sites, with 38 identified, followed closely by *ccmB*, which had 37 editing events. Conversely, *rps7* and *cox1* were the least edited genes, each with only one predicted editing site. Following RNA editing (Table S4), we found that 12.20% and 29.47% of amino acids remained unchanged in terms of hydrophobicity. Additionally, 48.37% of amino acids changed from hydrophilic to hydrophobic, 8.74% changed from hydrophobic to hydrophilic, and only 1.22% transitioned from hydrophilic to stop codons.

Repeat sequence analysis

The mitochondrial genome of *C. amarissima* contains numerous repetitive elements, including simple sequence repeats (SSRs), tandem repeats, and dispersed repetitive sequences, indicating structural complexity and frequent genomic rearrangements. We identified three distinct types of repeat sequences in this genome (Fig. 3A). Dispersed repeats are sequences that recur at different locations throughout the genome [7]. In the mitogenome of *C. amarissima*, we identified 3,901 dispersed repeats that are 500 bp or longer, including 1,783 forward and 2,118 palindromic repeats (Table S5). The longest forward repeat was 2,103 bp, while the longest palindromic repeat measured 2,141 bp, both located in chr6. Most of these repeats fell within the 500–2,141 bp range.

Simple sequence repeats (SSRs), consist of 1–6 bp DNA fragments, are valuable for phylogenetic analysis, population genetics, and species identification due to their polymorphism and widespread distribution in the genome. In the mitogenome of *C. amarissima*, we identified a total of 1,957 SSRs, including monomers, dimers, trimers, tetramers, pentamers, and hexamers (Table S6). Tetramers were the most common, representing 39.99% of the total, followed by dimers at 27.39% and trimers at 14.67%. Monomers and hexamers were less frequent. Notably, more than 90% of monomer SSRs comprised A/T bases, while 64.55% of dimer repeats contained AG/CT bases.

Tandem repeats, also known as satellite DNAs, consist of repeated core units typically ranging from 1 to 200 bases. In the mitogenome of *C. amarissima*, we identified a total of 4,863 tandem repeats. Among these, 3,915 tandem repeats fell within the length range of 9 to 39 bp, while the overall lengths varied from 2 to 524 bp (Table S7).

Homologous Fragment analysis between mitochondria and chloroplasts

The transfer of genetic material between cellular organelles, particularly within mitogenomes, is a common occurrence in higher plants [37]. In contrast, chloroplast-derived sequence fragments exhibit relatively lower conservation [38]. To examine this in *C. amarissima*, we performed a sequence similarity analysis to detect potential instances of sequence migration from the chloroplast to the mitochondrion (Fig. 3B).

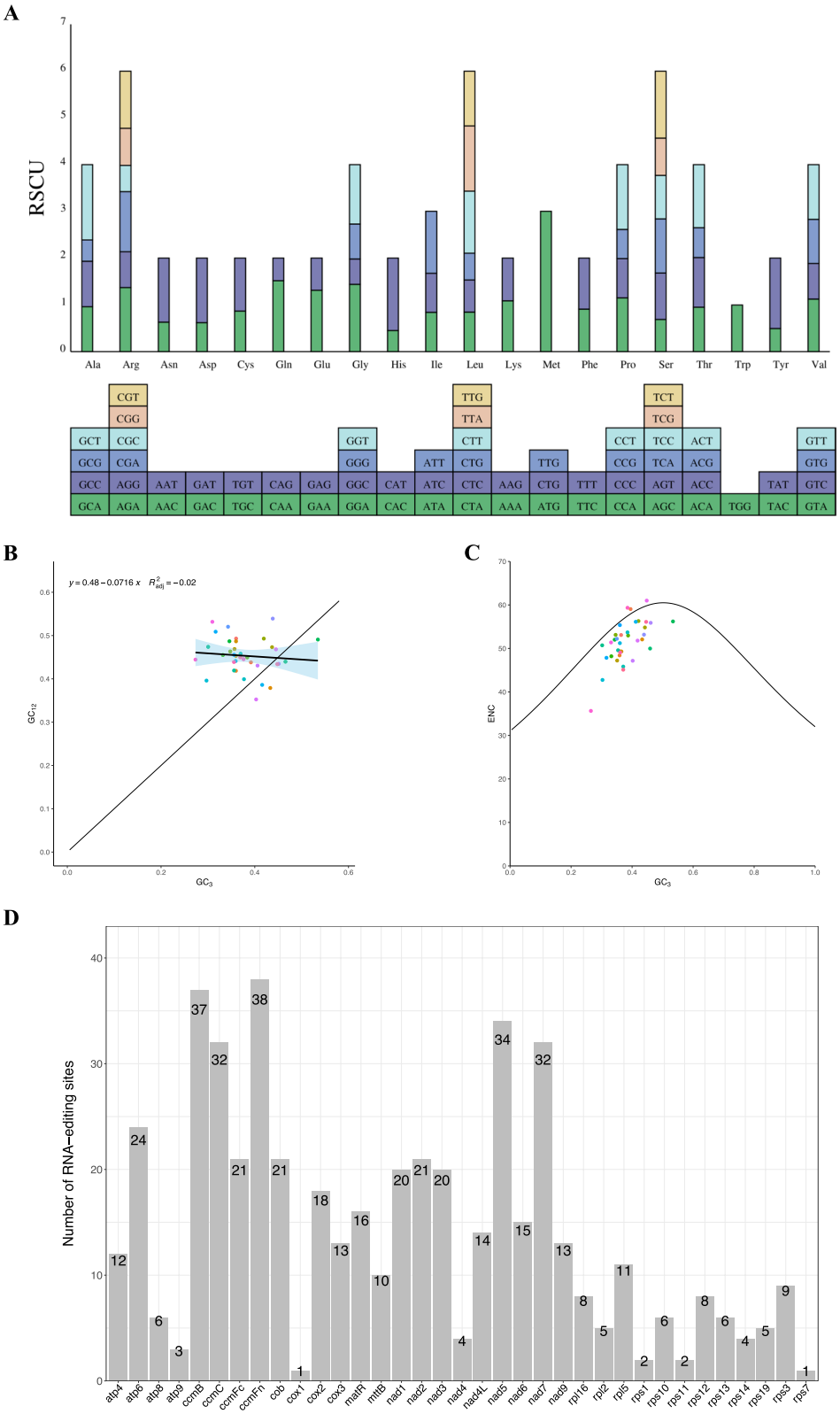

Fig. 2 The statistics of codon in *C. amarissima* mitogenome. **A** The relative synonymous codon usage (RSCU). Codon families displayed on the x-axis. RSCU values represent the frequency of a specific codon relative to the expected frequency of that codon under uniform synonymous codon usage. **B** GC content of different positions from PCGs. **C** ENC-plot against GC3 of mitochondrial genome of *C. amarissima*. **D** Number of RNA editing sites predicted by individual PCGs

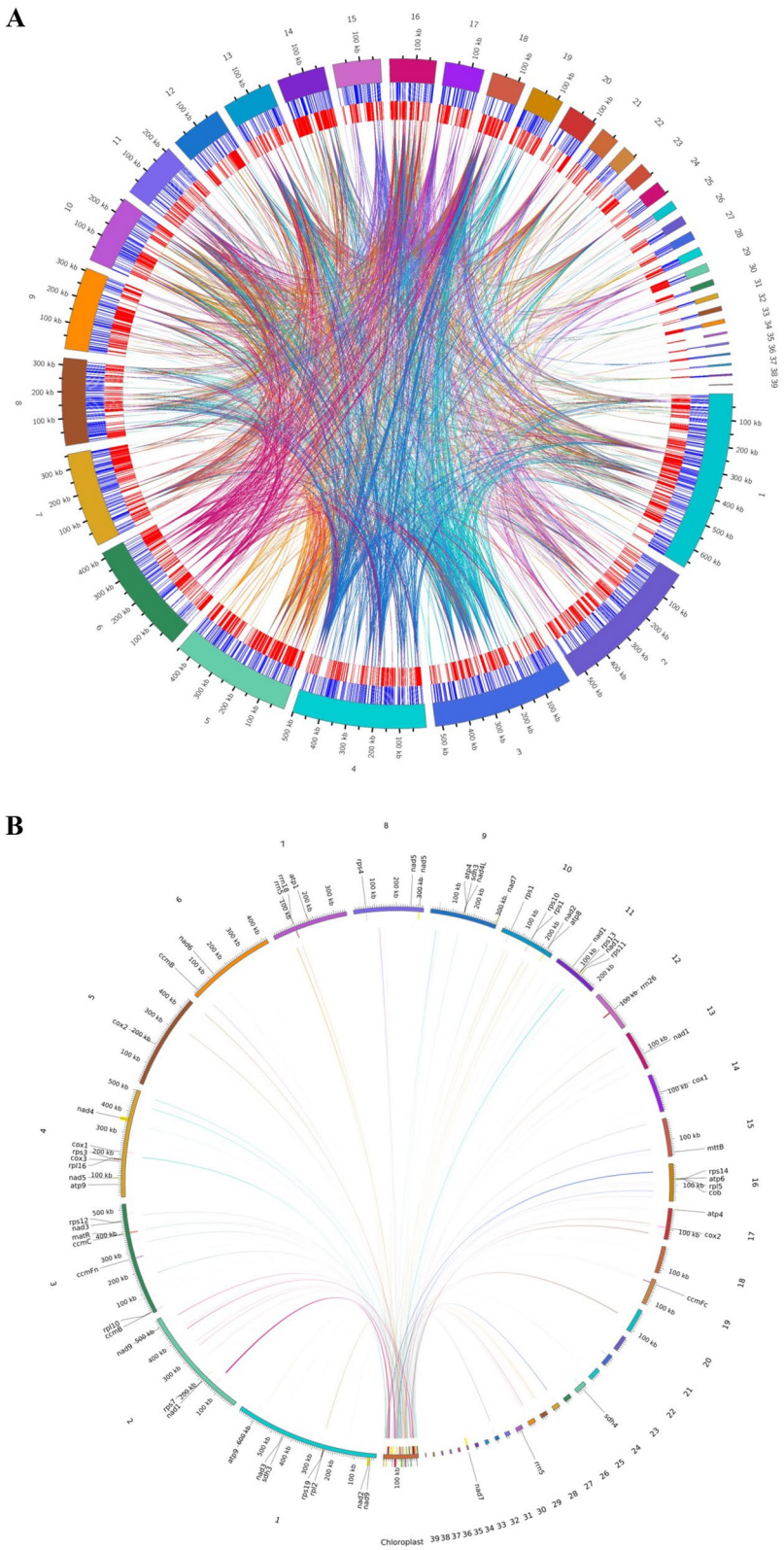

Fig. 3 The distribution of repetitive sequences and homologous in *C. amarissima* mitogenome. **A** Distribution of repetitive sequences. The outermost circle represents the mitochondrial genome; the inner circle is SSR, tandem repeat (red), and dispersed repeat (turquoise). **B** Homologous analysis based on different organelles shows the arc representing mitogenome DNA

We identified over 200 homologous fragments sharing between the chloroplast and mitogenomes of *C. amarissima*. These fragments ranged in alignment lengths from 10 to 511 bp, with mismatches ranging from 0 to 210. These fragments, 69,219 bp in total, account for approximately 1.06% of the *C. amarissima* mitochondrial DNA (Table S8). Upon annotating these sequences, we identified 16 complete genes: 18 tRNA genes (*trnA-TGC, trnD-GTC, trnE-TTC, trnF-GAA, trnH-GTG, trnK-CTT, trnL-CAA, trnL-TAG, trnM-CAT, trnN-GTT, trnP-TGG, trnR-ACG, trnR-TCT, trnS-GGA, trnS-TGA, trnT-TGT, trnV-GAC* and *trnW-CCA*). These tRNA genes in the mitogenome may have lost their functionality or undergone changes to become pseudogenes.

Comparative mitogenome analysis

To investigate the evolutionary characteristics of the mitogenome of *C. amarissima*, we compared it with three species from Commelinanae. The GC content of their mitogenomes varied from 44.04% in *C. amarissima* to 47.01% in *Pontederia crassipes*. The mitogenome sizes of four species varied from 397,847 bp (*Pontederia crassipes*) to 6,505,655 bp (*C. amarissima*). Additionally, the number of rRNAs, tRNAs, introns, and PCGs ranged from 4 to 3, 63 to 18, 31 to 19, and 32 to 43, respectively (Table 1).

To assess variation in PCGs and clarify evolutionary patterns in *C. amarissima*, we calculated the nonsynonymous/synonymous mutation ratio (Ka/Ks) and nucleotide diversity (Pi). Our study focused on how environmental stress affects mitogenome evolution. The average Pi for genes ranged from 0.01392 in *nad7* to 0.09955 in *atp9* (Fig. 3A and Table S9). *Atp9* had the highest Pi at 0.09955, whereas *nad7* had the lowest at 0.01392. Fast-evolving genes included *atp9* (0.09955), *rps10* (0.07576), *atp8* (0.05379), *atp4* (0.05353), and *rps19* (0.05243). In contrast, slow-evolving genes were *nad7* (0.01392), *cob* (0.01549), and *rrn5* (0.01709) (Fig. 4A). The analysis revealed that median Ka/Ks values exceeding 1 for *ccmB*, *mttB*, *rps1*, and *rps13* indicated positive selection, while other genes showed purifying selection (Fig. 4B).

Table 1 General characteristics of four species mtDNAs

	C. amarissima	*P. dactylifera*	*A. fasciata*	*P crassipes*
Size(bp)	6,505,655	715,001	1,166,331	880,048
GC%	44.04	45.14	44.27	45.57
rRNAs	4	3	3	3
tRNAs	63	18	37	18
introns	31	24	28	19
PCGs	38	43	42	32

We focused on conserved collinearity blocks of at least 500 bp, retaining those over 0.5 kb to effectively illustrate the collinear results (Fig. 4C). Our findings identified several co-linear genomic blocks exhibiting diverse structural arrangements, suggesting notable genomic rearrangements within the mitochondrial genome of *C. amarissima* compared to Commelinanae species.

Phylogenetic analysis

We extracted protein-coding gene sequences from seven Commelinanae species, and two outgroup species (*Apostasia fujianica* and *Dracaena cochinchinensis*). In the phylogenetic tree based on mitochondrial PCGs, seven out of eight nodes showed strong bootstrap support (BS > 90%) and posterior probabilities (PP = 1). Only one node had lower BS values (BS = 77) in Poaceae clade (Fig. 5A). We also obtained a tree with a topology identical almost universal to that derived from chloroplast PCGs, but only six out of eight nodes showed higher support (BS > 80 and PP = 1). Among Commelinanae species, *Phoenix dactylifera* formed the basal clade (though BS < 90 and PP = 1). *C. amarissima* clustered with *Pontederia crassipes* (BS = 100 and PP = 1), while *Saccharum oficinarum* and *Oryza sativa* were clustered in one branch in Poaceae clade (Fig. 5A). In contrast, *Oryza sativa* was nested within the three samples of Poaceae (Fig. 5B). The reasons for this discrepancy require further investigation.

Discussion

Sequencing and analyzing plant mitogenomes presents challenges due to their complex structure, characterized by numerous repetitive elements, integration of chloroplast genetic material, and extensive rearrangements, which complicate the assembly process [8, 39]. Plant mitogenomes, characterized by elevated recombination rates, display a flexible structure with various forms, including major loops, subloops, and linear configurations within the mitochondria. [40, 41]. In here, the *C. amarissima* mitogenome features a complex, multi-branched conformation with a length of 6,505,655 bp and a GC content of 44.04%, differing from the conventional circular arrangement found in terrestrial plant genomes. To date, the mitogenomes of Zingiberales have not been publicly released, limiting direct comparisons among them.

Codons are essential for translating genetic information into proteins [42]. Some studies have found that codon usage rates vary widely across species, likely resulting from a long-term equilibrium shaped by evolutionary selection [43]. In *C. amarissima*, the majority of PCGs initiated with the standard ATG codon, and the amino acid composition aligned with that of other angiosperms

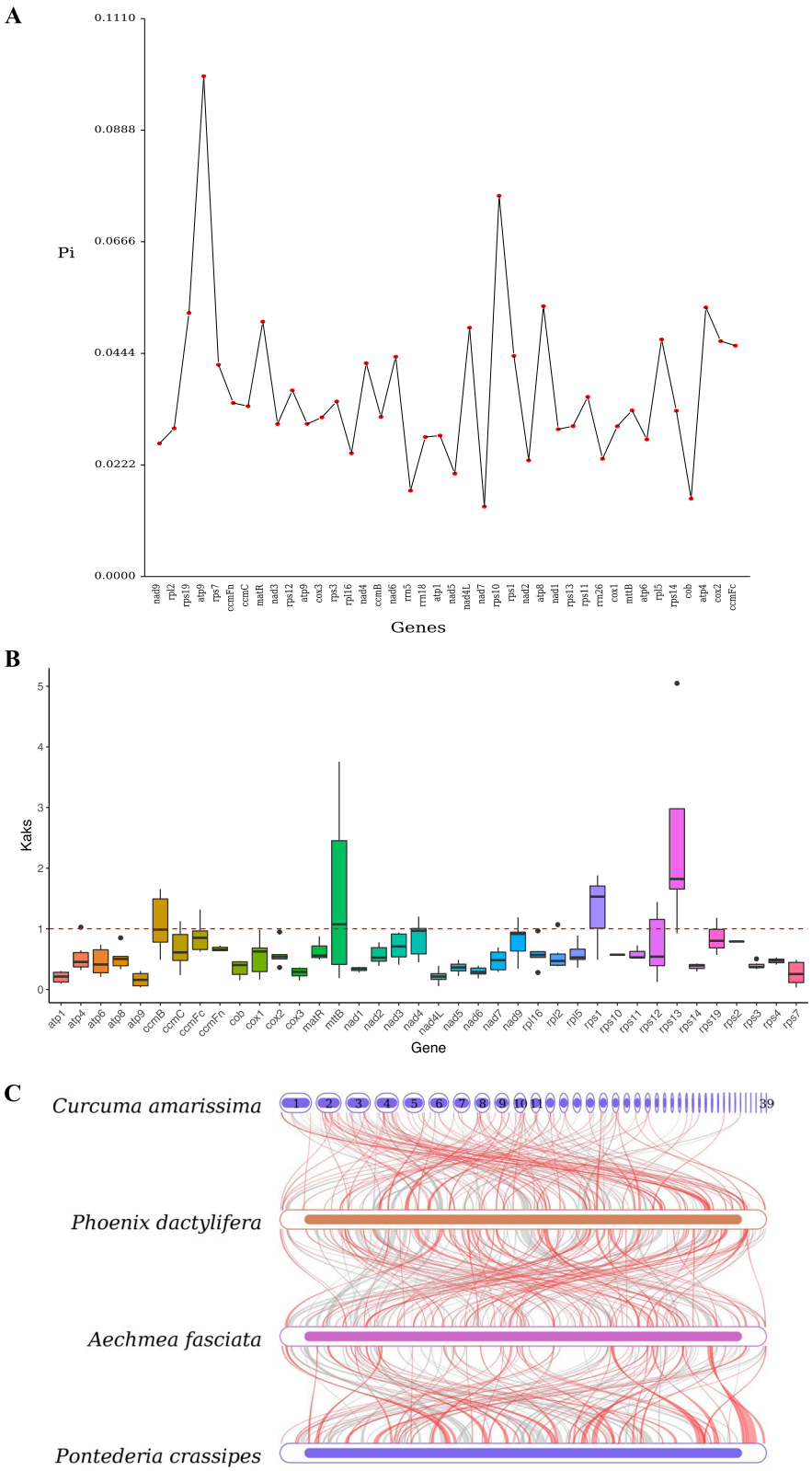

Fig. 4 Variation in mitochondrial genes and the evolutionary characteristics of *C. amarissima*. **A** Ka/Ks ratio calculated for the PCGs. **B** nucleotide diversity of the PCGs. **C** Collinear analysis of *C. amarissima* and other Commelinanae species. The red arcs indicate inverted regions, while the gray arcs indicate better homologous regions

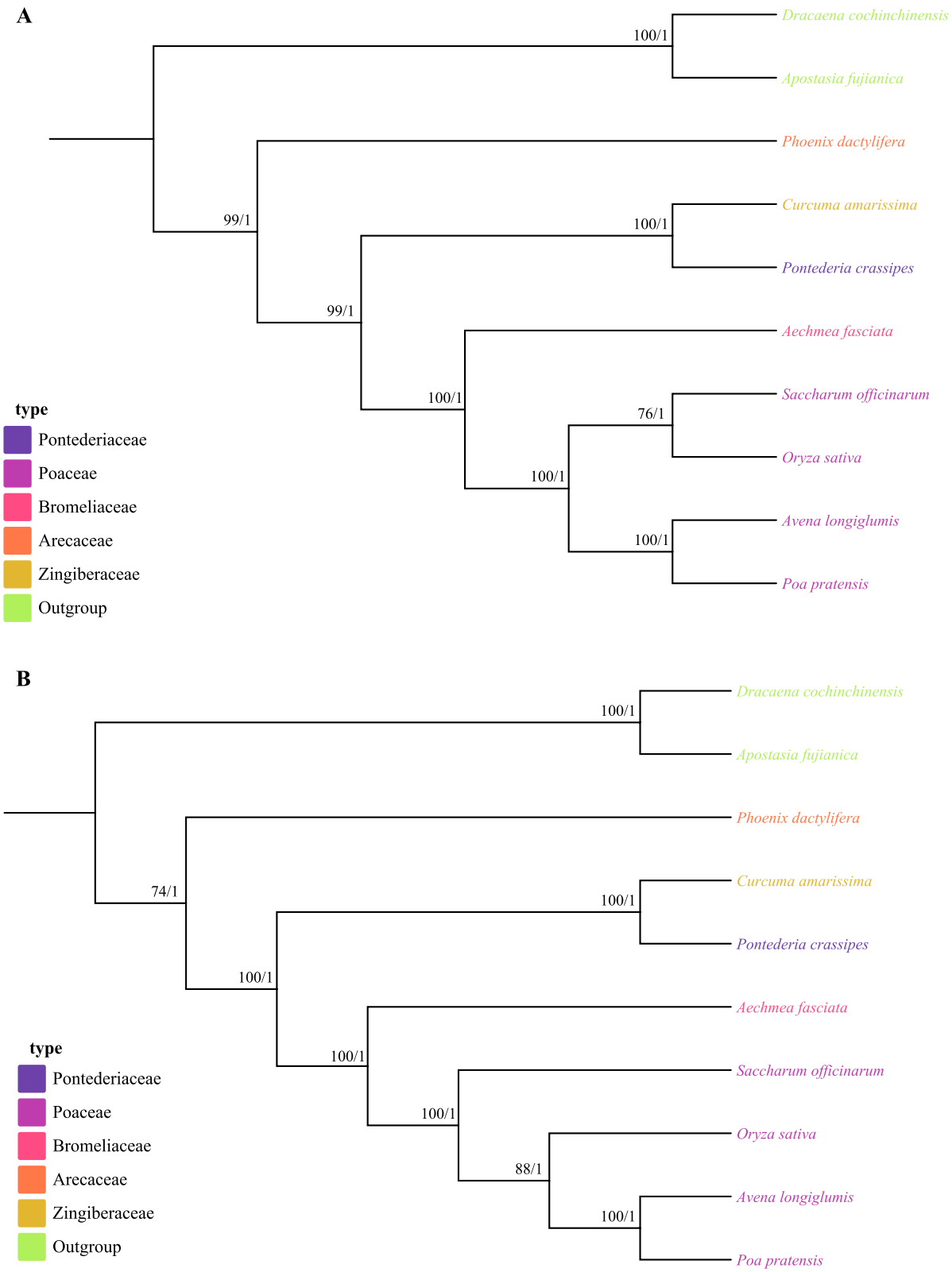

Fig. 5 Molecular phylogenetic analysis was conducted using sequences from both mitochondrial and chloroplast genomes. **A** A phylogenetic tree based on conserved mitochondrial proteins was analyzed using ML and BI. ML support values and BI posterior probabilities are shown at the nodes. **B** A similar tree was constructed from conserved chloroplast proteins using the same methods

[44]. Analysis of codon usage indicated a weak preference in the *C. amarissima* mitogenome, with 30 codons exhibiting an RSCU > 1. Most of these codons ended in A or T, reflecting a notable A/T bias at the third position, a trend often observed in plant mitogenomes [45].

RNA editing is a posttranscriptional modification occurring in the mitochondrial and chloroplast genomes of higher plants, altering mRNA sequences [46, 47]. Identifying RNA editing sites enhances our understanding of gene expression in plant mitochondria [48]. In this study, 492 RNA editing sites were predicted in the *C. amarissima* mitogenome, the numerous of RNA editing sites is similar to most angiosperm species eg. *Solanum muricatum* (585) [49], and lower than gymnosperms such as *Taxus cuspidata* (974) [50]. Besides, RNA editing is also common in mitochondrial transcripts and plays a vital role in generating functional proteins [51]. This process often introduces non-synonymous changes, which frequently result in codons that code for hydrophobic amino acids [52]. These edits help enhance protein folding and functionality, often leading to the production of more stable or functional proteins. In the NJCMS2B *soybean* maintenance line, two editing sites within a conserved region were identified that convert hydrophilic serine to hydrophobic leucine, impacting the orientation of the protein's transmembrane segments. In contrast, no RNA editing was observed in the *atp9* transcript of the sterility line NJCMS2 A [53]. In our study, all identified editing sites were C-to-T transitions located in 35 PCGs, with most resulting in amino acid changes favoring hydrophobic residues. Moreover, RNA editing in the *C. amarissima* mitogenome may lead to premature termination of translation, potentially affecting gene function. We identified 307 RNA editing sites that occurred at the second position of the triplet codon, including 207 that resulted in a hydrophilic-to-hydrophobic shift and 110 indicated a hydrophobic-to-hydrophobic change. These positions were the most susceptible to RNA editing, with a preference for leucine in the resulting amino acids. However, further investigation is required to fully understand their functional significance.

Repeated sequences are abundant in mitogenomes and serve as valuable resources for developing markers used in population and evolutionary studies. These repeats also play a critical role in genome recombination, which may lead to dynamic structural changes in the mitogenome [54]. In this study, we examined the three categories of repetitive sequences present in the *C. amarissima* mitogenome as previously described. These repeats are crucial for mitogenome recombination, as they contribute to variations in its size and structure [55]. The presence of such repeated sequences increases the likelihood of recombination, potentially resulting in structural changes that may influence gene expression and function [56]. In comparison to other mitogenomes, the *C. amarissima* mitogenome exhibits a high number of repetitive sequences, indicating that frequent intermolecular recombination may occur, potentially resulting in dynamic structural changes during evolution [57]. The identified monomer SSRs primarily consisted of A and T bases, which are linked by two hydrogen bonds. Previous studies have shown that this configuration requires less energy to break than the bonds in GC pairs, aiding in the structural stability of the mitogenome [58].

Previous studies have shown that some nuclear, mitochondrial, and chloroplast genes, as well as heterologous sequences, have been identified in many plant mitogenomes [8, 59]. A portion of these sequences are either noncoding or pseudogenes, though some are within coding regions. Understanding the fate of these transferred genes is crucial for tracking evolutionary changes in the mitogenome [8, 56]. A considerable portion of these sequences consists of noncoding regions or pseudogenes, although some are found within coding regions [60]. Some mitochondrial genes in plants contain introns, possibly due to intergenomic transfer (IGT); however, these features have minimal influence on the size and variability of plant mitogenome [56]. In the mitogenome of *C. amarissima*, the transfer of sequences from chloroplast to mitogenomes is conserved, showing a lower proportion of chloroplast sequence fragments (1.06%; 23,053 bp) than observed in other angiosperms. Additionally, 18 complete tRNA genes have migrated from the chloroplast DNA of *C. amarissima*. Previous studies have shown that the transfer of tRNA genes from the chloroplast to the mitogenome is common in angiosperms [61]. In *C. amarissima*, those tRNA genes were transferred from the chloroplast genome as complete sequences, suggesting they may contribute to regular mitochondrial functions [62]. Since comprehensive genome data for *C. amarissima* are currently lacking, the potential influence of nuclear genome transfer on the expansion of the mitogenome has not been fully explored. Moreover, earlier research has suggested two main scenarios regarding the fate of foreign genes within the mitogenome: (1) Transferred genes often become nonfunctional, while functional native copies remain [63]; and (2) Endogenous genes may be lost, allowing exogenous copies to take on roles in supporting normal cellular functions [64]. Understanding these patterns of sequence transfer is crucial for tracking historical recombination events and structural changes in plant mitogenomes, emphasizing the need for further investigation.

In genetics, the Ka/Ks ratio is an important measure used to evaluate the selection pressure acting on PCGs throughout evolution [65]. In our study, four PCGs

under positive selection were identified, which is in line with previous findings [12, 66–68]. The Ka/Ks analysis revealed that most conserved PCGs in the mitochondrial genome have been retained, indicating a strong selective pressure to maintain their function across evolutionary time. The genes under positive selection have shaped the evolutionary trajectory of local adaptation. For instance, *rps1* (with a Ka/Ks ratio greater than 1) plays a crucial role in activating the heat stress response in higher plants, potentially facilitating communication that initiates nuclear gene expression necessary for heat tolerance. Our study also identified two RNA editing sites in the *rps1* gene within the mitogenome of *C. amarissima*. This suggests that *rps1* may have been transformed from nonfunctional to translatable mRNA through RNA editing, enhancing the plant's heat resistance. Identifying these RNA editing sites in the genes under positive selection also could provide vital insights into the gene functions with new codons, improving our understanding of gene expression and local adaptation in the plant mitogenome.

The mitochondrial genome assembled in this study provides a valuable new resource for resolving the complex evolutionary relationships within the genus *Curcuma* and the broader Zingiberaceae family. Previous phylogenetic studies relying primarily on chloroplast or nuclear genomes have encountered challenges due to inter- and intraspecific morphological similarities within *Curcuma* [69, 70]. Despite recent advancements from chloroplast [71, 72] and nuclear genomes [73, 74], the lack of mitochondrial genome data in Zingiberaceae limits its comprehensive evolutionary analyses. Our mitochondrial genome data from *C. amarissima* offers preliminary insights into the evolutionary relationships and highlights possible structural rearrangements that could influence phylogenetic inference. However, given the current limited mitochondrial genomic resources for Zingiberaceae and the potential existence of hybridization events, our conclusions remain preliminary. Additional mitochondrial genome assemblies from various species within the family are necessary to resolve existing conflicts and fully understand the evolutionary history of Zingiberaceae and Zingiberales [75–79].

Conclusions

In this study, we characterized the first complete mitochondrial genome of *C. amarissima*, expanding mitochondrial genomic resources available for the Zingiberaceae family. The mitogenome, spanning 6,505,655 bp, exhibited a unique multi-branched structure. Codon usage analysis revealed weak bias, suggesting natural selection influences mitochondrial gene evolution. Comparative genomics analyses highlighted distinct structural rearrangements and frequent sequence migration events between mitochondrial and chloroplast genomes. Phylogenetic analysis provided valuable preliminary insights into evolutionary relationships within the Zingiberaceae family, suggesting that mitochondrial genome data offer distinct advantages compared to chloroplast genomes. However, given the limited availability of mitochondrial genomic data within Zingiberaceae, the phylogenetic conclusions drawn here should be viewed as preliminary. Future studies involving additional mitochondrial genome sequencing within Zingiberaceae and related taxa are essential to provide comprehensive insights into the evolutionary history and genetic relationships within this important plant family.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11540-x.

Supplementary Material 1: Fig. S1 Genomic map of the C. amarissima mitogenome. The genome consisted of 39 segments

Supplementary Material 2: Table S1 Genes encoding and organization of the mitogenome of C. amarissima

Supplementary Material 3: Table S2 Relative synonymous codon usage for each amino acid pair in the mitochondrial genome of C. amarissima

Supplementary Material 4: Table S3 The statistics in ENC of C. amarissima

Supplementary Material 5: Table S4 Prediction of RNA editing sites of C. amarissima

Supplementary Material 6: Table S5 The statistics in Dispersed repeats of C. amarrisa

Supplementary Material 7: Table S6 The statistics in SSRs of C. amarissima

Supplementary Material 8: Table S7 The statistics in Tandem repeats of C. amarissima

Supplementary Material 9: Table S8 The statistics of homologous fragments between mitochondria and chloroplasts in C. amarissima

Acknowledgements

We thank Lei Xu (Genepioneer Biotechnologies Co. Ltd, Nanjing) for assistance with the data analysis.

Authors' contributions

G.G. designed and supervised the project. H.L. wrote the manuscript. H.L. annotated and analyzed the genomes. H.Q., J.D., and Y.W. prepared the samples and performed the experiments. H.Q., J.D., and Y.W. analyzed the data. G.G., L.H., R.W. and J.D. revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the Natural Science Foundation of China (32001390), Supported by State Key Laboratory of Resource Insects (SKLRI—ORP202420) and the Basic scientific research business expenses of HAAS (ITH2024ZD02).

Data availability

The accession number of Curcuma amarissima mitogenome in GeneBank are PQ442956-PQ442994.

Declarations

Ethics approval and consent to participate
We collected fresh leaf materials of *Curcuma amarissima* for this study. The plant samples and experimental research comply with relevant institutional, national, and international guidelines and legislation. No specific permissions or licenses were required.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Institute of Tropical Horticulture Research, Hainan Academy of Agricultural Sciences, Haikou 571100, China. [2]Faculty of Agriculture, Forestry and Food Engineering, Yibin University, Yibin, Sichuan, China. [3]College of Chuanjiu, Sichuan Vocational College of Chemical Technology, Luzhou 646300, China. [4]College of Life Science, Sichuan Agricultural University, Ya'an, Sichuan Province 625014, China. [5]School of Life Sciences, Technical University of Munich, Freising 85354, Germany.

Received: 5 November 2024 Accepted: 27 March 2025
Published online: 05 April 2025

References

1. Nass MM. Mitochondrial DNA: Advances, Problems, and Goals: Studies of size and structure of mitochondrial DNA relate to biogenesis and function of this organelle. Science. 1969;165(3888):25–35.
2. Liberatore KL, Dukowic-Schulze S, Miller ME, Chen C, Kianian SF. The role of mitochondria in plant development and stress tolerance. Free Radical Biol Med. 2016;100:238–56.
3. Atteia A, van Lis R, Tielens AG, Martin WF. Anaerobic energy metabolism in unicellular photosynthetic eukaryotes. Biochim Biophys Acta. 2013;1827(2):210–23.
4. Antico Arciuch VG, Elguero ME, Poderoso JJ, Carreras MC. Mitochondrial regulation of cell cycle and proliferation. Antioxid Redox Signal. 2012;16(10):1150–80.
5. Seo BJ, Yoon SH, Do JT. Mitochondrial dynamics in stem cells and differentiation. Int J Mol Sci. 2018;19(12):3893.
6. Hampl V, Roger AJ. The evolutionary origin of mitochondria and mitochondrion-related organelles. In: Endosymbiotic Organelle Acquisition: Solutions to the Problem of Protein Localization and Membrane Passage. Springer;2024:89–121.
7. Lu G, Wang W, Zhang S, Yang G, Zhang K, Que Y, Deng L. The first complete mitochondrial genome of Grossulariaceae: Molecular features, structure recombination, and genetic evolution. BMC Genomics. 2024;25(1):744.
8. Wang J, Kan S, Liao X, Zhou J, Tembrock LR, Daniell H, Jin S, Wu Z. Plant organellar genomes: much done, much more to do. Trends Plant Sci. 2024;29(7):754–69.
9. Zhou Q, Ni Y, Li J, Huang L, Li H, Chen H, Liu C. Multiple configurations of the plastid and mitochondrial genomes of Caragana spinosa. Planta. 2023;258(5):98.
10. Yang H, Ni Y, Zhang X, Li J, Chen H, Liu C. The mitochondrial genomes of Panax notoginseng reveal recombination mediated by repeats associated with DNA replication. Int J Biol Macromol. 2023;252:126359.
11. Jackman SD, Coombe L, Warren RL, Kirk H, Trinh E, MacLeod T, Pleasance S, Pandoh P, Zhao Y, Coope RJ. Complete mitochondrial genome of a gymnosperm, Sitka spruce (Picea sitchensis), indicates a complex physical structure. Genome Biol Evol. 2020;12(7):1174–9.
12. Feng G, Jiao Y, Ma H, Bian H, Nie G, Huang L, Xie Z, Ran Q, Fan W, He W. The first two whole mitochondrial genomes for the genus Dactylis species: assembly and comparative genomics analysis. BMC Genomics. 2024;25(1):235.
13. Jain K, Panigrahi M, Nayak SS, Rajawat D, Sharma A, Sahoo SP, Bhushan B, Dutt T. The evolution of contemporary livestock species: Insights from mitochondrial genome. Gene. 2024;927:148728.
14. Shan Y, Li J, Zhang X, Yu J. The complete mitochondrial genome of Amorphophallus albus and development of molecular markers for five Amorphophallus species based on mitochondrial DNA. Front Plant Sci. 2023;14:1180417.
15. Hao Z, Jiang X, Pan L, Guo J, Chen Y, Li J, Liu B, Guo A, Luo L, Jia R. The complete mitochondrial genome of Pontederia crassipes: using HiFi reads to investigate genome recombination and gene transfer from chloroplast genome. Front Plant Sci. 2024;15:1407309.
16. Mau J-L, Lai EY, Wang N-P, Chen C-C, Chang C-H, Chyau C-C. Composition and antioxidant activity of the essential oil from Curcuma zedoaria. Food Chem. 2003;82(4):583–91.
17. Selvam RM, Athinarayanan G, Nanthini AUR, Singh AR, Kalirajan K, Selvakumar PM. Extraction of natural dyes from Curcuma longa, Trigonella foenum graecum and Nerium oleander, plants and their application in antimicrobial fabric. Ind Crops Prod. 2015;70:84–90.
18. Di Lorenzo R, Grumetto L, Sacchi A, Laneri S, Dini I. Dermocosmetic evaluation of a nutricosmetic formulation based on Curcuma. Phytother Res. 2023;37(5):1900–10.
19. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100.
20. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37(5):540–6.
21. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 2017;27(5):737–46.
22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.
23. 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.
24. Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. Insights into the evolution of mitochondrial genome size from complete sequences of Citrullus lanatus and Cucurbita pepo (Cucurbitaceae). Mol Biol Evol. 2010;27(6):1436–48.
25. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3. 1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47(W1):W59-W64.
26. Hershberg R, Petrov DA. Selection on codon bias. Annu Rev Genet. 2008;42(1):287–99.
27. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 70 for bigger datasets. Mol Biol Evol. 2016;33(7):1870–4.
28. Sharp PM, Li W-H. Codon usage in regulatory genes in Escherichia coli does not reflect selection for 'rare'codons. Nucleic Acids Res. 1986;14(19):7737–49.
29. Covello PS, Gray MW. RNA editing in plant mitochondria. Nature. 1989;341(6243):662–6.
30. Lenz H, Knoop V. PREPACT 2.0: predicting C-to-U and U-to-C RNA editing in organelle genome sequences with multiple references and curated RNA editing annotation. Bioinformatics and biology insights 2013;7:BBI. S11059.
31. Chen Y, Ye W, Zhang Y, Xu Y. High speed BLASTN: an accelerated MegaBLAST search tool. Nucleic Acids Res. 2015;43(16):7762–8.
32. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004;14(7):1394–403.
33. Katoh K, Misawa K. Kuma Ki, Miyata T: MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66.
34. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34(12):3299–302.
35. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35(21):4453–5.
36. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 32: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–42.

37. Martin W, Herrmann RG. Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol. 1998;118(1):9–17.
38. Rodríguez-Moreno L, González VM, Benjak A, Martí MC, Puigdomènech P, Aranda MA, Garcia-Mas J. Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. BMC Genomics. 2011;12:1–14.
39. Fan W, Liu F, Jia Q, Du H, Chen W, Ruan J, Lei J, Li DZ, Mower JP, Zhu A. Fragaria mitogenomes evolve rapidly in structure but slowly in sequence and incur frequent multinucleotide mutations mediated by microinversions. New Phytol. 2022;236(2):745–59.
40. 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.
41. Liang H, Qi H, Chen J, Wang Y, Liu M, Sun X, Wang C, Xia T, Feng X, Feng S, et al. Assembly and analysis of the first complete mitochondrial genome sequencing of main Tea-oil Camellia cultivars Camellia drupifera (Theaceae): revealed a multi-branch mitochondrial conformation for Camellia. BMC Plant Biol. 2025;25(1):13.
42. Zhou Z, Dang Y, Zhou M, Li L. Yu C-h, Fu J, Chen S, Liu Y: Codon usage is an important determinant of gene expression levels largely through its effects on transcription. Proc Natl Acad Sci. 2016;113(41):E6117–25.
43. Parvathy ST, Udayasuriyan V, Bhadana V. Codon usage bias. Mol Biol Rep. 2022;49(1):539–65.
44. Zhou P, Zhang Q, Li F, Huang J, Zhang M. Assembly and comparative analysis of the complete mitochondrial genome of Ilex metabaptista (Aquifoliaceae), a Chinese endemic species with a narrow distribution. BMC Plant Biol. 2023;23(1):393.
45. Guan D-L, Ma L-B, Khan MS, Zhang X-X, Xu S-Q, Xie J-Y. Analysis of codon usage patterns in Hirudinaria manillensis reveals a preference for GC-ending codons caused by dominant selection constraints. BMC Genomics. 2018;19:1–14.
46. Bonnard G, Gualberto JM, Lamattina L, Grienenberger JM, Brennlcke A. RNA editing in plant mitochondria. Crit Rev Plant Sci. 1992;10(6):503–24.
47. Small ID, Schallenberg-Rüdinger M, Takenaka M, Mireau H, Ostersetzer-Biran O. Plant organellar RNA editing: what 30 years of research has revealed. Plant J. 2020;101(5):1040–56.
48. Hammani K, Giegé P. RNA metabolism in plant mitochondria. Trends Plant Sci. 2014;19(6):380–9.
49. Li Z, Liu J, Liang M, Guo Y, Chen X, Wu H, Jin S. De novo assembly of the complete mitochondrial genome of pepino (Solanum muricatum) using PacBio HiFi sequencing: insights into structure, phylogenetic implications, and RNA editing. BMC Plant Biol. 2024;24(1):1–13.
50. Kan S-L, Shen T-T, Gong P, Ran J-H, Wang X-Q. The complete mitochondrial genome of Taxus cuspidata (Taxaceae): eight protein-coding genes have transferred to the nuclear genome. BMC Evol Biol. 2020;20:1–17.
51. Gott JM, Emeson RB. Functions and mechanisms of RNA editing. Annu Rev Genet. 2000;34(1):499–531.
52. Xia X, Li W-H. What amino acid properties affect protein evolution? J Mol Evol. 1998;47:557–64.
53. Jiang W, Yang S, Yu D, Gai J. A comparative study of ATPase subunit 9 (Atp9) gene between cytoplasmic male sterile line and its maintainer line in soybeans. Afr J Biotech. 2011;10(51):10387–92.
54. Sharbrough J, Bankers L, Cook E, Fields PD, Jalinsky J, McElroy KE, Neiman M, Logsdon JM Jr, Boore JL. Single-molecule sequencing of an animal mitochondrial genome reveals chloroplast-like architecture and repeat-mediated recombination. Mol Biol Evol. 2023;40(1):msad007.
55. 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.
56. Gualberto JM, Newton KJ. Plant mitochondrial genomes: dynamics and mechanisms of mutation. Annu Rev Plant Biol. 2017;68(1):225–52.
57. Cheng Y, He X, Priyadarshani S, Wang Y, Ye L, Shi C, Ye K, Zhou Q, Luo Z, Deng F. Assembly and comparative analysis of the complete mitochondrial genome of Suaeda glauca. BMC Genomics. 2021;22:1–15.
58. Giudice E, Várnai P, Lavery R. Base pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations. Nucleic Acids Res. 2003;31(5):1434–43.
59. Liu M, Yang M, Liang H, Luo B, Deng J, Cao L, Zheng D, Chen C. Polyploidy drives autophagy to participate in plant-specific functions. iMeta. 2024;3:e252.
60. Wang D, Rousseau-Gueutin M, Timmis JN. Plastid sequences contribute to some plant mitochondrial genes. Mol Biol Evol. 2012;29(7):1707–11.
61. Knie N, Polsakiewicz M, Knoop V. Horizontal gene transfer of chlamydial-like tRNA genes into early vascular plant mitochondria. Mol Biol Evol. 2015;32(3):629–34.
62. Bi C, Paterson AH, Wang X, Xu Y, Wu D, Qu Y, Jiang A, Ye Q, Ye N. Analysis of the complete mitochondrial genome sequence of the diploid cotton Gossypium raimondii by comparative genomics approaches. Biomed Res Int. 2016;2016(1):5040598.
63. Kubo T, Newton KJ. Angiosperm mitochondrial genomes and mutations. Mitochondrion. 2008;8(1):5–14.
64. Wang Y, Liu X-Y, Yang Y-Z, Huang J, Sun F, Lin J, Gu Z-Q, Sayyed A, Xu C, Tan B-C. Empty Pericarp21 encodes a novel PPR-DYW protein that is required for mitochondrial RNA editing at multiple sites, complexes I and V biogenesis, and seed development in maize. PLoS Genet. 2019;15(8):e1008305.
65. Zhang Z, Li J, Yu J. Computing Ka and Ks with a consideration of unequal transitional substitutions. BMC Evol Biol. 2006;6:1–10.
66. Qin N, Yang S, Wang Y, Cheng H, Gao Y, Cheng X, Li S. The de novo assembly and characterization of the complete mitochondrial genome of bottle gourd (Lagenaria siceraria) reveals the presence of homologous conformations produced by repeat-mediated recombination. Front Plant Sci. 2024;15:1416913.
67. Yang Y, Duan C. Mitochondrial genome features and systematic evolution of diospyros kaki thunb'Taishuu'. BMC Genomics. 2024;25(1):285.
68. Miao X, Yang W, Li D, Wang A, Li J, Deng X, He L, Niu J. Assembly and comparative analysis of the complete mitochondrial and chloroplast genome of Cyperus stoloniferus (Cyperaceae), a coastal plant possessing saline-alkali tolerance. BMC Plant Biol. 2024;24(1):628.
69. Liang H, Deng J, Gao G, Ding C, Zhang L, Yang R. Incompatibility Phylogenetic Signals between Double-Digest Restriction Site-Associated DNA Sequencing and Plastid Genomes in Chinese Curcuma (Zingiberaceae)—A Recent Qinghai-Tibetan Plateau Diversification Genera. Forests. 2022;13(2):280.
70. Liang H, Deng J, Gao G, Ding C, Zhang L, Xu K, Wang H, Yang R. Inferring the Phylogeny and Divergence of Chinese Curcuma (Zingiberaceae) in the Hengduan Mountains of the Qinghai-Tibet Plateau by Reduced Representation Sequencing. Forests. 2021;12(5):520.
71. Liang H, Chen J. Comparison and phylogenetic analyses of nine complete chloroplast genomes of Zingibereae. Forests. 2021;12(6):710.
72. Liang H, Zhang Y, Deng J, Gao G, Ding C, Zhang L, Yang R. The complete chloroplast genome sequences of 14 Curcuma species: insights into genome evolution and phylogenetic relationships within Zingiberales. Front Genet. 2020;11:802.
73. Yin Y, Xie X, Zhou L, Yin X, Guo S, Zhou X, Li Q, Shi X, Peng C, Gao J. A chromosome-scale genome assembly of turmeric provides insights into curcumin biosynthesis and tuber formation mechanism. Front Plant Sci. 2022;13:1003835.
74. Liao X, Ye Y, Zhang X, Peng D, Hou M, Fu G, Tan J, Zhao J, Jiang R, Xu Y. The genomic and bulked segregant analysis of Curcuma alismatifolia revealed its diverse bract pigmentation. Abiotech. 2022;3(3):178–96.
75. Kress WJ, Prince LM, Williams KJ. The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. Am J Bot. 2002;89(10):1682–96.
76. Kress WJ, Liu AZ, Newman M, Li QJ. The molecular phylogeny of Alpinia (Zingiberaceae): a complex and polyphyletic genus of gingers. Am J Bot. 2005;92(1):167–78.
77. Jiang D, Cai X, Gong M, Xia M, Xing H, Dong S, Tian S, Li J, Lin J, Liu Y. Complete chloroplast genomes provide insights into evolution and phylogeny of Zingiber (Zingiberaceae). BMC Genomics. 2023;24(1):30.
78. Kress WJ. The phylogeny and classification of the Zingiberales. Ann Missouri Botanical Garden. 1990;77:698–721.
79. Carlsen MM, Fér T, Schmickl R, Leong-Škorničková J, Newman M, Kress WJ. Resolving the rapid plant radiation of early diverging lineages in the tropical Zingiberales: pushing the limits of genomic data. Mol Phylogenet Evol. 2018;128:55–68.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.