Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

First complete mitochondrial genome of star fruit (Averrhoa carambola L.): assembly, comparative analysis, and evolutionary implications

Junru Qu
Guangxi Subtropical Crops Research Institute
Liping Wang
Guangxi Subtropical Crops Research Institute
Hui Wen
Guangxi Subtropical Crops Research Institute
Liyan Lu
Guangxi Subtropical Crops Research Institute
Ying Liu
Guangxi Subtropical Crops Research Institute
Yangfan Zhu
Guangxi Subtropical Crops Research Institute
Kaikai Meng
mengkaikai2024@163.com
Guangxi Subtropical Crops Research Institute

Research Article

Keywords: Averrhoa carambola, Mitochondrial genome, Phylogenetic analysis

Posted Date: July 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10266342/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Averrhoa carambola L. belonging to the family Oxalidaceae of the order Oxalidales, is a commercially valuable fruit tree extensively grown in tropical and subtropical regions. However, its mitochondrial genome remains uncharacterized.

Results

In this study, we report the first complete assembly and comprehensive characterization of the *A. carambola* mitochondrial genome, providing an essential genomic resource for evolutionary and functional studies. The assembled mitochondrial genome is a circular, double-stranded DNA molecule of 421,040 bp with a GC content of 44.56%. It contains 39 protein-coding genes and exhibits a pronounced A/T bias in codon usage. We identified 163 simple sequence repeats (SSRs), 448 long repetitive sequences (≥ 30 bp), and 32 chloroplast-derived DNA fragments. C-to-U RNA editing was the predominant type, with 567 predicted editing sites, mainly concentrated in the *nad4* gene. Comparative collinearity analysis revealed extensive structural rearrangements relative to related species, and phylogenomic analysis robustly placed *A. carambola* within the order Oxalidales.

Conclusions

These findings significantly expand the available mitochondrial genomic resources for Oxalidales and provide a foundational framework for phylogenetic inference, germplasm authentication, and molecular breeding of *A. carambola*.

Background

Mitochondria are essential organelles in eukaryotic cells, serving as the primary sites of energy metabolism by supplying ATP through oxidative phosphorylation. They also participate in key biological processes, such as apoptosis and stress responses, playing irreplaceable roles in plant growth, development, and environmental adaptability [1]. Mitochondrial DNA (mtDNA) is characterized by maternal inheritance, a relatively high mutation rate, and limited recombination, making it a valuable tool for investigating species evolution, organelle interactions, and adaptive mechanisms [2].

Unlike their animal and fungal counterparts, plant mitochondrial genomes exhibit unique complexity, reflected in extreme size variation (typically 200 kb to 3 Mb in angiosperms), diverse structural conformations (circular, linear, multicyclic, or branched), a high proportion of non-coding sequences, and frequent inter-organellar gene transfer [3]. Despite their biological significance, the technical challenges associated with assembling plant mitochondrial genomes have resulted in far fewer fully resolved genomes compared to chloroplast genomes, leaving many species uncharacterized [4].

Current research on plant mitochondria focuses on four key areas: the evolutionary dynamics and structural plasticity of mitochondrial genomes; the molecular mechanisms governing RNA editing, intron splicing, and post-transcriptional regulation; the functional integration of mitochondrial metabolism with photosynthesis, photorespiration and stress tolerance; and mitochondrial retrograde signaling in development and environmental adaptation [5]. However, most studies have concentrated on model or well-studied plants, while tropical fruit trees such as *Averrhoa carambola* L. (Fig. 1) remain largely unexplored at the mitochondrial genome level [6].

A. carambola belonging to the family Oxalidaceae (order Oxalidales), is a widely cultivated economic fruit tree in tropical and subtropical regions [7]. Its fruits are rich in vitamins, minerals, and bioactive compounds, with high nutritional and medicinal value [8, 9]. The species also exhibits strong environmental adaptability, making it an ideal candidate for studying stress response mechanisms in tropical plants [7, 10, 11]. To date, research on *A. carambola* have primarily focused on cultivation practices, fruit-content composition, and pest and disease control [8, 12–14]. Genomic studies have been limited to nuclear genome and the chloroplast genome, with no complete mitochondrial genome assembly available [15–17].

In this study, we report the first complete assembly and annotation of the mitochondrial genome of *A. carambola*, which represents the first mitochondrial genome for any species within the order Oxalidales. This work fills a significant gap in the genomic resources of this order. Beyond providing a high-quality reference genome, our study offers new insights into genome structure, repeat dynamics, and inter-organellar gene transfer in a previously uncharacterized lineage, and establishes a robust molecular foundation for phylogenetic analyses and evolutionary studies both within Oxalidales and across related orders.

Materials and Methods

Plant Materials and Sequencing Data

Fresh leaves of *A. carambola* were collected from healthy mature plants cultivated at the Guangxi Subtropical Crops Research Institute, Nanning, China (22.90414709°N, 108.35362343°E). Immediately after collection, the leaf samples were snap-frozen in liquid nitrogen and stored at -80°C for subsequent nucleic acid extraction.

Total genomic DNA was extracted from the leaf tissues using a modified cetyltrimethylammonium bromide (CTAB) method [18]. The integrity, concentration, and purity of the extracted DNA were assessed by agarose gel electrophoresis combined with a Nanodrop 2000 ultraviolet spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Qualified DNA samples passing quality control were divided equally into two aliquots for the construction of Illumina short-read genomic libraries and Oxford Nanopore Technologies (ONT) long-read genomic libraries, respectively. Specifically, the short-read libraries were subjected to high-throughput sequencing on an Illumina NovaSeq 6000 sequencer (Illumina

Inc., San Diego, CA, USA), while the long-read libraries were sequenced on a PromethION sequencer (Oxford Nanopore Technologies, Oxford, UK).

For transcriptome sequencing, total RNA was extracted from the same batch of leaf samples using a commercial RNA extraction kit. RNA purity was assessed using a Nanodrop 2000 ultraviolet spectrophotometer, concentration was quantified using a Qubit 4 Fluorometer, and integrity was verified by agarose gel electrophoresis. mRNA was enriched from total RNA using oligo(dT) magnetic beads, and RNA-seq libraries were constructed following standard protocols. The libraries were sequenced on the Illumina NovaSeq 6000 platform, generating 150 bp paired-end reads.

The resulting raw sequencing data from all three platforms were subjected to quality control and used for the subsequent assembly and analysis of the *A. carambola* mitochondrial genome.

Mitochondrial Genome Assembly and Annotation

The mitochondrial genome was assembled based on long-read sequencing data using Flye v2.9 [19] with default parameters, generating graphical assembly results in the GFA format. All assembled contigs were constructed in a local database using makeblastdb. Subsequently, the BLASTn program [20] was employed to identify contig fragments containing the mitochondrial genome using mitochondrial genes from closely related species as query sequences. Since no complete mitochondrial genomes of species in the order Oxalidales have been published to date, four high-quality complete mitochondrial genomes from the phylogenetically closest orders Malpighiales and Celastrales were selected as reference sequences to identify mitochondrial-derived contig fragments via homologous alignment. The reference genomes included *S. dunnii* (NC_058734.1), *P. alba* (NC_041085.1), *E. alatus* (NC_053921.1), and *T. wilfordii* (NC_082972.1). The BLASTn analysis was performed with the following parameters: -evalue 1e-5 -outfmt 6 -max_hsps 10 -word_size 7 -task blastn-short.

Mitochondrial contigs were visualized and curated with Bandage v0.8.1 [21] to obtain a preliminary mitochondrial genome draft. Subsequently, long and short reads were mapped to the identified mitochondrial contigs using minimap2 v2.26-r1175 [22], and all reads aligned to mitochondrial sequences were extracted and retained. Finally, the high-quality cleaned long and short reads were subjected to hybrid co-assembly using Unicycler v0.5.0 [23], resulting in the complete circular mitochondrial genome of *A. carambola*.

The complete mitochondrial genome of *A. carambola* was initially annotated using the Plant Mitochondrial Genome Annotation (PMGA) platform [24] against a curated reference database comprising 319 published plant mitochondrial genomes. Transfer RNA (tRNA) genes were predicted using tRNAscan-SE v2.0.11 [25], while ribosomal RNA (rRNA) genes were identified via homologous alignment using BLASTn v2.13.0 [26]. All putative misannotations were manually inspected and corrected using Apollo v1.11.8 [27]. Finally, a schematic circular map of the annotated mitochondrial genome was generated using OGDRAW v1.3.1 [28].

Mitochondrial Genome Bioinformatic Analyses

Codon usage bias analysis: Protein-coding sequences (CDSs) were extracted from the annotated *A. carambola* mitochondrial genome using PhyloSuite v1.1.16 [29], and relative synonymous codon usage (RSCU) values were calculated using MEGA7 [30].

Repeat sequence analysis: Simple sequence repeats (SSRs) were identified using MISA v2.1 [31], with minimum repeat thresholds set to 10, 5, 4, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexanucleotide motifs, respectively. Tandem repeats were detected using Tandem Repeats Finder v4.09 [32] under default parameters. Interspersed repeats were analyzed using the REPuter online server [33] with a minimum repeat length of 30 bp and a Hamming distance of 3. All repeat analysis results were visualized using Excel 2021 and Circos v0.69.9 [34].

Inter-organellar sequence transfer analysis: The complete chloroplast genome of *A. carambola* was assembled using GetOrganelle v1.7.5 [35], annotated using CPGAVAS2 [36], and manually curated using CPGView [37]. Homologous fragments longer than 100 bp between the mitochondrial and chloroplast genomes were identified via BLASTn (E-value ≤ 1e-5) and visualized using Circos.

RNA editing prediction: Mitochondrial reads were extracted from the transcriptome dataset and uniquely mapped to the *A. carambola* mitochondrial genome. Sequence differences between genomic DNA and RNA transcripts were detected using REDItools v2.0 [38], and high-confidence RNA editing sites supported by at least 10 uniquely mapped reads, a base quality score ≥ 30 and an editing frequency ≥ 10% were retained for further analysis.

Collinearity analysis: Conserved collinear blocks between *A. carambola* and related species were identified via all-vs-all BLASTn with the following parameters: -evalue 1e-5 -word_size 9 -gapopen 5 -gapextend 2 -reward 2 -penalty – 3. Only homologous fragments longer than 500 bp were retained for subsequent analysis. Synteny plots were then constructed and visualized using MCscanX [39] based on the BLASTn alignment results.

Phylogenetic analysis: Complete mitochondrial genomes of closely related species were downloaded from the NCBI GenBank database. Single-copy orthologous genes shared by all species were extracted using PhyloSuite v1.1.16 [29] and aligned using MAFFT v7.505 [40]. A maximum likelihood (ML) phylogenetic tree was subsequently constructed using IQ-TREE v1.6.12 [41] with the parameters --alrt 1000 -B 1000. The resulting tree was visualized using iTOL v6 [42], with two mitochondrial genomes from the order Fabales (Psophocarpus tetragonolobus NC_086919.1 and Glycine max NC_020455.1) designated as outgroups for rooting.

Results

Structural Features and Functional Annotation

The mitochondrial genome of *A. carambola* is a single circular molecule of 421,040 bp. Its nucleotide composition is A: 27.42%, T: 28.02%, G: 22.18%, and C: 22.38%, resulting in a GC content of 44.56% (Fig. 2). The assembly achieved an average sequencing depth of 95×, confirming high coverage and assembly quality. A total of 62 mitochondrial genes were annotated, including 39 unique protein-coding genes (PCGs), 20 tRNA genes, and three rRNA genes (Table 1). Among the PCGs, 24 are core mitochondrial genes essential for oxidative phosphorylation and related biogenesis, while the remaining 15 are non-core genes involved in ribosomal function and auxiliary metabolism.

The core gene set includes five ATP synthase subunit genes (*atp1*, *atp4*, *atp6*, *atp8*, *atp9*); nine NADH ubiquinone oxidoreductase subunit genes (*nad1–nad9*, excluding *nad8*); four cytochrome c biogenesis factors (*ccmB*, *ccmC*, *ccmFC*, *ccmFN*); three cytochrome c oxidase subunits (*cox1–cox3*); one mitochondrial membrane transporter (*mttB*); one maturase (*matR*); and one apocytochrome b (*cob*). The non-core PCGs comprise four large ribosomal subunit proteins (*rpl2*, *rpl5*, *rpl10*, and *rpl16*), nine small ribosomal subunit proteins (*rps1*, *rps3*, *rps4*, *rps7*, *rps10*, *rps12*, *rps13*, *rps14*, and *rps19*), and two succinate dehydrogenase subunits (*sdh3*, *sdh4*). The tRNA genes encode 20 distinct anticodons, with five tRNAs present in multiple copies (*trnF-GAA×2*, *trnK-UUU×2*, *trnP-CGG×3*, *trnP-UGG×2*, and *trnV-GAC×2*). The rRNA genes include *rrn5* (×2), *rrn18* (×2), and *rrn26* (×1) (Table 1).

Codon Usage analysis of PCGs

Among the annotated protein-coding genes, both start and stop codons exhibit typical characteristics of angiosperm mitochondrial genomes, and functional modifications of codons mediated by post-transcriptional RNA editing are prevalent.

For initiation codons, most genes employ the canonical ATG as the start codon, while seven genes adopt non-canonical initiation patterns. The genomically encoded native start codons of *rps4*, *rps10*, *nad1*, *nad4L* and *sdh3* are all ACG. Among them, the start codons of *rps10* and *sdh3* are unambiguously annotated to be generated by C-to-U RNA editing, which converts the genomically encoded ACG codon into the canonical AUG initiation codon. For *rps4*, *nad1* and *nad4L*, their start codons are likewise genomically encoded as ACG, with corresponding C-to-U RNA editing sites located at the second position of the codon. However, due to conservative annotation criteria and varying strengths of supporting evidence, these codons are annotated as putative or undetermined. The remaining two genes, *mttB* and *rpl16* employ native non-canonical initiation codons that do not rely on RNA editing, with ATA for *mttB* and GTG for *rpl16*. Both codons are directly recognized by the mitochondrial translation machinery and initiate translation with methionine. With respect to termination codons, most genes harbor genomically encoded TAA or TAG as canonical termination signals.

The functional stop codons of four genes, namely *atp6*, *atp9*, *ccmFC* and *rps10*, are all generated by C-to-U RNA editing. Specifically, *atp6* and *atp9* feature the CAA-to-UAA editing pattern, while *ccmFC* and *rps10* feature the CGA-to-UGA editing pattern, with all editing sites located at the first position of the codon. Furthermore, *rps10* is the only gene in which both the start and stop codons are generated via C-to-U RNA editing (Table 1).

Codon usage analysis revealed a strong preference for A/T-ending codons among the PCGs (Fig. 3).. The relative synonymous codon usage (RSCU) values for the start codon AUG and tryptophan codon UGG were exactly 1.0. Among stop codons, UAA was the most frequently used (RSCU = 1.63), followed by UGA (RSCU = 0.86) and UAG (RSCU = 0.51). The most preferred codons were GCU (RSCU = 1.59), UAU and CAU (RSCU = 1.52), whereas GCG, CAC and UAC (RSCU = 0.48) were the least utilized (Fig. 3).

Table 1
Protein-coding genes in the mitochondrial genome of *A. carambola*

Group of genes	Name of genes	Length	Start codon	Stop codon	Amino acid
ATP synthase	*atp*1	1500	ATG	TAA	499
	*atp*4	597	ATG	TAG	198
	*atp*6	720	ATG	CAA	239
	*atp*8	480	ATG	TAA	159
	*atp*9	225	ATG	CAA	74
NADH dehydrogenase	*nad*1	978	ACG	TAA	325
	*nad*2	1467	ATG	TAA	488
	*nad*3	357	ATG	TAA	118
	*nad*4	1488	ATG	TAA	495
	*nad*4L	303	ACG	TAA	100
	*nad*5	2016	ATG	TAA	671
	*nad*6	681	ATG	TAA	226
	*nad*7	1185	ATG	TAA	394
	*nad*9	573	ATG	TAA	190
Cytochrome b	*cob*	1179	ATG	TAA	392
Cytochrome c biogenesis	*ccm*B	621	ATG	TAA	206
	*ccm*C	762	ATG	TAA	253
	*ccm*FC	1317	ATG	CGA	438
	*ccm*FN	1722	ATG	TAA	573
Cytochrome c oxidase	*cox*1	1584	ATG	TAA	527
	*cox*2	783	ATG	TAA	260
	*cox*3	798	ATG	TAA	265
Maturases	*mat*R	1968	ATG	TAA	655
Protein transport subunit	*mtt*B	939	ATA	TAA	312
Ribosomal protein large subunit	*rpl*2	897	ATG	TAA	298
	*rpl*5	558	ATG	TAA	185

Group of genes	Name of genes	Length	Start codon	Stop codon	Amino acid
	*rpl*10	489	ATG	TAA	162
	*rpl*16	435	GTG	TAA	144
Ribosomal protein small subunit	*rps*1	726	ATG	TAA	241
	*rps*3	1692	ATG	TAA	563
	*rps*4	1050	ACG	TAA	349
	*rps*7	447	ATG	TAA	148
	*rps*10	429	ACG	CGA	142
	*rps*12	378	ATG	TAA	125
	*rps*13	177	ATG	TAA	58
	*rps*14	303	ATG	TAA	100
	*rps*19	327	ATG	TAA	108
Succinate dehydrogenase	*sdh*3	525	ACG	TAA	174
	*sdh*4	426	ATG	TAA	141
Ribosome RNA	*rrn*5(×2)	118			
	*rrn*18(×2)	1938			
	*rrn*26	3026			
Transfer RNA	*trnC-GCA*	71			
	trnD-GUC	74			
	trnE-UUC	72			
	trnF-GAA(×2)	74/73			
	trnfM-CAU	74			
	trnG-GCC	72			
	trnH-GUG	74			
	trnI-CAU	77			
	trnK-UUU(×2)	73/75			
	trnM-CAU	73			
	trnN-GUU	72			
	trnP-CGG(×3)	74/93/96			

Group of genes	Name of genes	Length	Start codon	Stop codon	Amino acid
	trnP-UGG(×2)	75/68			
	trnQ-UUG	72			
	trnS-GCU	88			
	trnS-GGA	87			
	trnS-UGA	87			
	trnV-GAC(×2)	72			
	trnW-CCA	74			
	trnY-GUA	83			

(Note: The numbers in parentheses indicate the copy number of the corresponding gene.)

Characteristics of Repetitive Sequences

A total of 163 simple sequence repeats (SSRs) were identified (Fig. 4A, Table S1). Monomeric SSRs were the most abundant (65, 39.88%), followed by tetrameric (43, 26.38%), dimeric (29, 17.79%), trimeric (16, 9.82%), and pentameric (9, 5.52%) repeats. Only one hexameric repeat was detected. Among monomeric repeats, A-rich repeats dominated (37 out of 65, 56.92%), followed by T-rich repeats (26 out of 65, 40.00%). Only two C-rich mononucleotide repeats were identified, and no G-rich mononucleotide repeats were detected. Dinucleotide SSRs were primarily composed of AT and TA motifs, while only one CG dinucleotide repeat was found. The total length of each SSR locus ranged from 10 bp to 30 bp. The sole hexameric SSR reached the maximum length of 30 bp, whereas the majority of SSR loci fell within the length range of 10–16 bp.

A total of 448 interspersed repetitive sequences (≥ 30 bp) were identified (Fig. 4B, Table S2), including forward repeats (252, 56.25%), palindromic repeats (188, 41.96%), reverse repeats (7, 1.56%), and complementary repeats (1, 0.22%). The longest palindromic repeat spanned 11,551 bp, and the longest forward repeat was 3,626 bp. All E-values of these dispersed repeats were far below the standard significance threshold of 0.01, corroborating the authenticity of all detected repeat pairs. Specifically, six forward dispersed repeats with an E-value of 0 represented ultra-high-confidence homologous matches, with sequence lengths ranging from 143 bp to 3,626 bp. The interval distance between these high-confidence repeat pairs was zero, indicating that the two matching segments were adjacently positioned in the mitogenome with no intervening nucleotides. Additionally, 34 tandem repeat units were detected (Table S3), with core motif lengths ranging from 3 bp to 36 bp, sequence similarity of ≥ 79%, and copy numbers varying from 1.9 to 11.

A circular visualization of repeat distribution (Figure S1) revealed extensive intra-genomic homologous pairings, highlighting the structural plasticity of the *A. carambola* mitochondrial genome. In the diagram, orange lines represent forward repeats and purple lines represent palindromic repeats; the two prominent highlighted long lines colored red and blue correspond to complementary repeats and reverse repeats, respectively. Black segments on the second concentric circle indicate tandem repeat sequences, and those on the outermost ring denote microsatellite repeats.

DNA Transfer between Mitochondrial and Chloroplast Genomes

A total of 32 mitochondrial plastid DNA (MTPT) fragments were identified, spanning 29,617 bp in total and accounting for 7.03% of the mitochondrial genome length (Fig. 5, Table S4). Sequence identity of these homologous fragments relative to their chloroplast counterparts ranged from 96.39% to 100%, and the longest fragment spanned 3,959 bp. The majority of protein-coding genes within these transferred segments occurred as partial truncated fragments, suggesting pseudogenization subsequent to interorganellar transfer. In total, 16 complete genes were detected, including seven chloroplast-derived PCGs (*petG*, *psaJ*, *psbB*, *rpl20*, *rpl33*, *rps11*, *rps18*) and nine tRNA genes (*trnD-GUC*, *trnF-GAA*, *trnH-GUG*, *trnM-CAU*, *trnN-GUU*, *trnP-UGG*, *trnS-GGA*, *trnV-GAC*, *trnW-CCA*).

RNA Editing Sites

we systematically profiled RNA editing sites across all protein-coding genes (PCGs) using transcriptome sequencing data. A total of 567 high-confidence C-to-U RNA editing sites were identified in PCGs (Figure S2, Table S5), and no U-to-C reverse editing events were detected in this study. The amino acid conversions resulting from RNA editing exhibited a distinct preference, with Pro→Leu, Ser→Leu, and Ser→Phe being the predominant substitution types. Nearly all editing events were non-synonymous and altered the encoded amino acids; only 42 synonymous editing events were identified, involving 10 amino acid types: Ala, Gly, Ile, Leu, Phe, Pro, Ser, Thr, Tyr, and Val.

At the gene level, *nad4* harbored the largest number of editing sites (n = 40), followed by *ccmB* and *ccmFN* (n = 36 each), *nad7* (n = 33), and *nad2* and *nad5* (n = 31 each). In contrast, *atp8*, *rps14*, and *rps7* each contained only two editing sites. Notably, PCGs encoding subunits of respiratory complex I (the *nad* family) and cytochrome c biogenesis factors (the *ccm* family) collectively accounted for a large proportion of editing sites, whereas ribosomal protein genes (*rps* and *rpl*) possessed relatively fewer editing loci.

Phylogenetic Analysis

To resolve the phylogenetic placement of *A. carambola*, a maximum likelihood (ML) phylogeny was inferred from the concatenated nucleotide alignments of 14 conserved mitochondrial PCGs (*atp1*, *atp4*, *atp6*, *atp8*, *ccmB*, *ccmFC*, *cob*, *cox3*, *matR*, *nad2*, *nad3*, *nad5*, *nad6*, and *nad7*) from 24 angiosperm species representing four orders, with two Fabales species *Psophocarpus tetragonolobus* (L.) DC. and *Glycine max* (L.) Merr. as outgroups. The resulting tree (Fig. 6) fully supported the APG IV classification [43], with *A.*

carambola forming an independent monophyletic clade within Oxalidales. Consistent with the tree topology, Fabales occupied the basal position of the phylogeny as the earliest-diverging lineage. Oxalidales formed a sister clade to the combined Celastrales–Malpighiales group, while Celastrales and Malpighiales were resolved as reciprocally sister orders.

Collinearity Analysis

Comparative collinearity analysis of the *A. carambola* mitochondrial genome was performed against three representative angiosperm species (*Jatropha curcas*, *Euonymus alatus*, and *Psophocarpus tetragonolobus*). The results revealed extensive structural rearrangements, including large-scale inversions and genome-wide reshuffling of syntenic blocks (Fig. 7), indicating low structural conservation of the mitogenome. All homologous syntenic segments were supported by high-confidence BLAST alignments, with nearly all entries exhibiting an E-value of 0 and sequence identities ranging from 73.64% to 98.99% (Table S6). To improve visualization clarity, syntenic blocks shorter than 0.5 kb were excluded from the plot.

Discussion

Structural Characteristics of the A. carambola Mitochondrial Genome

The mitochondrial genome of *A. carambola* is a circular molecule of 421,040 bp with a GC content of 44.56%, features that are broadly consistent with the typical architecture of angiosperm mitochondrial genomes. Compared with the *A. carambola* chloroplast genome (155,965 bp) [17], the mitochondrial genome is substantially larger and exhibits a higher proportion of non-coding regions and repetitive sequences, reflecting two well-documented evolutionary trends in plant mtDNA: rampant size variation and structural complexity [44]. The GC content of 44.56% lies well within the narrow, evolutionarily conserved range of 43–45% reported for most flowering plants [45]. This compositional stability is not accidental; it results from strong purifying selection acting on evolutionarily conserved protein-coding genes, particularly those encoding core respiratory chain components. Mutations that alter GC content in these functionally constrained regions are likely to be deleterious and are therefore efficiently removed [2]. The conservation of GC content across diverse angiosperm lineages, including the newly sequenced *A. carambola*, underscores the long-term functional and structural constraints that govern plant mtDNA evolution [46]. As the first representative of Oxalidales, the *A. carambola* mitochondrial genome fills a critical phylogenetic gap and confirms that this lineage conforms to the same fundamental compositional rules observed across angiosperms.

Repeat Sequences in the A. carambola Mitochondrial Genome

Plant mitochondrial genomes are characterized by extraordinary structural flexibility, in which repetitive sequences serve as core substrates for homologous recombination and represent a primary driver of genome size expansion and architectural rearrangement [47]. In the *A. carambola* mitogenome, we identified 448 interspersed repetitive sequences (≥ 30 bp) and 34 tandem repeat units. Forward and

palindromic repeats accounted for the vast majority of interspersed repeats, with the longest palindromic repeat extending to 11,551 bp and the longest forward repeat reaching 3,626 bp. The abundance of long homologous repeats provides abundant templates for intra-genomic recombination, which directly underlies the high structural plasticity of the *A. carambola* mitogenome [48]. This repeat composition is consistent with the general evolutionary pattern observed in angiosperm mitochondrial genomes, where repeat-mediated recombination acts as a pervasive force shaping mitogenome architecture [49–50].

RNA Editing Events in the A. carambola Mitochondrial Genome

The distribution of C-to-U RNA editing sites in the *A. carambola* mitogenome aligns with the canonical pattern observed in most angiosperms [5]. The *nad4* gene harbored the largest number of sites (n = 40), and genes encoding NADH dehydrogenase subunits (complex I) and cytochrome c biogenesis factors generally contained more editing sites, whereas ribosomal protein genes possessed relatively fewer loci [51]. Editing frequencies varied widely across individual sites, indicating the coexistence of fully edited and partially edited transcripts, reflecting that core respiratory genes are under stronger functional constraint and thus retain more conserved editing events [52]. The presence of partially edited sites further implies potential spatiotemporal specificity in RNA editing regulation, which may fine-tune mitochondrial function across developmental stages or tissue types [53].

Plastid-to-Mitochondrion DNA Transfer and Organellar Co-Evolution

The transfer of chloroplast-derived sequences is a hallmark of plant mitochondrial genome evolution [54]. In this study, 32 chloroplast-derived homologous fragments were identified in the *A. carambola* mitochondrial genome, totaling 29,617 bp and representing 7.03% of the mitogenome. This substantial proportion confirms extensive chloroplast-to-mitochondrial DNA transfer in *A. carambola* [55]. These fragments include seven protein-coding genes and nine tRNA genes, which not only enrich the sequence diversity of the mitochondrial genome but also potentially provide functional complementation for mitochondria and improve the stability of the mitochondrial translation system [56]. Moreover, the ongoing acquisition and stabilization of chloroplast-derived sequences constitute a key mechanism driving structural innovation and adaptive functional refinement in the *A. carambola* mitogenome [47]. The predominance of pseudogenized protein-coding genes and the preferential retention of intact tRNA genes conform to the universal pattern of plastid-to-mitochondrion transfer in angiosperms [57]. The intact plastid-derived tRNA genes may be functionally recruited by the mitochondrial translation system, reflecting functional coordination and co-evolution between the two organellar genomes [58].

Phylogenetic Relationships of the A. carambola Mitochondrial Genome

Compared with nuclear and chloroplast genomes, plant mitochondrial genomes evolve more slowly at the nucleotide sequence level but exhibit pronounced structural dynamism, rendering them particularly powerful for resolving deep phylogenetic relationships among angiosperm lineages [5]. Phylogenetic reconstruction based on conserved mitochondrial protein-coding genes has therefore become a robust

complementary strategy for validating higher-level angiosperm clades, especially for recalcitrant nodes where plastid and nuclear datasets yield conflicting topologies [59]. The COM clade (Celastrales, Oxalidales, Malpighiales) is a well-delineated major lineage within the rosids, yet the inter-ordinal relationships among its three constituent orders have remained a long-standing point of systematic debate [60]. Our study provides independent genome-scale mitochondrial evidence that Oxalidales forms a sister clade to the combined Celastrales–Malpighiales group, with Celastrales and Malpighiales resolved as reciprocally sister lineages. These results corroborate the current taxonomic placement of *A. carambola* and provide new genome-scale mitochondrial evidence for inter-ordinal relationships within the COM clade. The complete mitochondrial genome of *A. carambola* obtained in this study, together with the phylogenetic analysis results, provides a solid foundation for further research on the origin, evolutionary diversification, and adaptive evolution of Oxalidales species.

Conclusions

We present the first complete, high-quality mitochondrial genome assembly of *A. carambola*. The mitogenome exhibits a typical circular molecular architecture with a total length of 421,040 bp and a GC content of 44.56%. Genome annotation revealed 39 protein-coding genes, 20 transfer RNA genes, and 3 ribosomal RNA genes. The genome harbors abundant simple sequence repeat (SSR) loci, as well as widely distributed forward and palindromic repeats; these repetitive sequences serve as core substrates for intramitochondrial homologous recombination and represent key drivers of genomic structural rearrangement. We detected extensive interorganellar DNA transfer from plastids to mitochondria, with 32 chloroplast-derived fragments (MTPTs) identified in the *A. carambola* mitogenome. Codon usage analysis revealed distinct usage bias. RNA editing events were overwhelmingly dominated by C-to-U conversions, with 567 high-confidence editing sites predicted across protein-coding genes. The *nad4* gene was identified as the primary editing hotspot, and the global editing distribution pattern is consistent with the conserved post-transcriptional regulatory paradigm observed across angiosperm mitochondria. Phylogenetic reconstruction based on 14 conserved mitochondrial protein-coding genes robustly places *A. carambola* within the Oxalidales clade, which is congruent with both classical morphological classification and the APG IV taxonomic framework. Comparative synteny analysis uncovered extensive genomic structural rearrangements, highlighting the highly dynamic and rapidly evolving nature of angiosperm mitochondrial genomes. The complete mitogenome of *A. carambola* reported herein enriches the available organellar genomic resources for Oxalidales, and provides an essential foundation for future studies on evolutionary biology, germplasm evaluation, and molecular breeding of *Averrhoa* species.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All raw sequencing data and the assembled mitogenome sequence of *A. carambola* are publicly available in NCBI repositories: the circular mitochondrial genome is deposited in GenBank (accession PZ577443), and the associated transcriptomic raw reads are hosted in the SRA database under BioProject PRJNA1465236.

Competing interests

The authors declare that they have no competing interests.

Funding

The Basic Scientific Research Project of Guangxi Academy of Agricultural Sciences (GNK2025YP135 and GNK2026YT017);The Science and Technology Development Project of Guangxi Academy of Agricultural Sciences (GNK2023ZX06).

Authors' contributions

K.M. and J.Q. conceived and designed the project. J.Q. and L.W. assembled the mitogenomes and performed multiple analyses. J.Q. drafted the manuscript. Y.L. and Y.Z. provided experimental materials. J.Q., L.W., K.M., H.W. and L.L. participated in the revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

Not applicable.

References

1. Millar AH, Small ID, Day DA, Whelan J. Mitochondrial biogenesis and function in Arabidopsis. Arabidopsis Book. 2008;6:e0111. https://doi.org/10.1199/tab.0111.
2. Xing BP, Lin LY, Wu Q. Application of mitochondrial genomes to species identification and evolution. ELECTRON J BIOTECHN. 2025;76:39–48. https://doi.org/10.1016/j.ejbt.2025.04.001.
3. Gualberto JM, Mileshina D, Wallet C, Niazi AK, Weber-Lotfi F, Dietrich A. The plant mitochondrial genome: Dynamics and maintenance. Biochimie. 2014;100:107–120. https://doi.org/10.1016/j.biochi.2013.09.016.
4. Wang J, Kan SL, Liao XZ, Zhou Jw, Tembrock LR, Daniell H, et al. Plant organellar genomes: much done, much more to do. Trends Plant Sci. 2024;29:754–769. https://doi.org/10.1016/j.tplants.2023.12.014.

5. Moller IM, Rasmusson AG, Van Aken O. Plant mitochondria–past, present and future. Plant J. 2021;108:912–959. https://doi.org/10.1111/tpj.15495.
6. Wang RZ, Li XL, Sun MY, Xue C, Korban SS, Wu J. Genomic insights into domestication and genetic improvement of fruit crops. Plant Physiol. 2023;192:2604–2627. https://doi.org/10.1093/plphys/kiad273.
7. Patil A, Koli S, Patil D, Phatak A. A comprehensive review of an important medicinal plant–*Averrhoa carambola* L.. Pharmacognosy. 2012;2:13–17. https://doi.org/10.5530/PC.2012.2.3.
8. Muthu N, Lee SY, Phua KK, Bhore SJ. Nutritional, medicinal and toxicological attributes of star-fruits (*Averrhoa carambola* L.): a review. Bioinformation. 2016;12:420. https://doi.org/10.6026/97320630012420.
9. Dhara J, Saha SK, Saha PS, Chakraborty R. Star Fruit (*Averrhoea Carambola*): Exploring its Nutritional Value, Health Benefits, and Consideration for Potential Risk and Future Application in Biotechnology and Medicine. Curr. Pharmacol. Rep. 2025;11:7. https://doi.org/10.1007/s40495-024-00387-7.
10. Wu PP, Wu CB, Zhou BY. Drought stress induces flowering and enhances carbohydrate accumulation in *Averrhoa Carambola*. Hortic. Plant J. 2017;3:60–66. https://doi.org/10.1016/j.hpj.2017.07.008.
11. Mustafa MA, Ali A, Seymour G, Tucker G. Enhancing the antioxidant content of carambola (*Averrhoa carambola*) during cold storage and methyl jasmonate treatments. Postharvest Biol. Technol. 2016;118:79–86. https://doi.org/10.1016/j.postharvbio.2016.03.021.
12. Núñez-Elisea R, Crane JH. Selective pruning and crop removal increase early-season fruit production of carambola (*Averrhoa carambola* L.). SCI HORTIC-AMSTERDAM. 2000;86:115–126. https://doi.org/10.1016/S0304-4238(00)00141-2.
13. George HL, Davies FS, Crane JH, Schaffer B. Root temperature effects on '*Arkin'carambola* (*Averrhoa carambola* L.) trees: I. Leaf gas exchange and water relations. SCI HORTIC-AMSTERDAM. 2002;96:53–65. https://doi.org/10.1016/S0304-4238(02)00090-0.
14. Yasawardene P, Jayarajah U, De Zoysa I, Seneviratne SL. Mechanisms of star fruit (*Averrhoa carambola*) toxicity: A mini-review. Toxicon. 2020;187:198–202. https://doi.org/10.1016/j.toxicon.2020.09.010.
15. Wu SS, Sun W, Xu Z, Zhai JW, Li XP, Li CR, et al. The genome sequence of star fruit (*Averrhoa carambola*). Hortic. Res. 2020;7:95. https://doi.org/10.1038/s41438-020-0307-3.
16. Fan YN, Sahu SK, Yang T, Mu WX, Wei JP, Cheng L, et al. Dissecting the genome of star fruit (*Averrhoa carambola* L.). Hortic. Res. 2020;7:94. https://doi.org/10.1038/s41438-020-0306-4.
17. Jo S, Kim HW, Kim YK, Cheon SH, Kim KJ. Complete plastome sequence of *Averrhoa carambola* L. (Oxalidaceae). MITOCHONDRIAL DNA B. 2016;1:609–611. https://doi.org/10.1080/23802359.2016.1209095.
18. Arseneau JR, Steeves R, Laflamme M. Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues. Mol Ecol Resour. 2017;17:686–693. https://doi.org/10.1111/1755-0998.12616.

19. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat. Biotechnol. 2019;37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
20. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J. Mol. Biol. 1990;215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
21. Wick RR, Schultz MB, Zobel J, Holt KE. Bandage: interactive visualization of de novo genome assemblies. BIOINFORMATICS. 2015;31:3350–3352. https://doi.org/10.1093/bioinformatics/btv383.
22. Li H. Minimap2: pairwise alignment for nucleotide sequences. BIOINFORMATICS. 2018;34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
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:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
24. Li JL, Ni Y, Lu QQ, Chen HM, Liu C. PMGA: A plant mitochondrial genome annotator. Plant Commun. 2025;6:101191. https://doi.org/10.1016/j.xplc.2024.101191.
25. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25:955–964. https://doi.org/10.1093/nar/25.5.955.
26. Chen Y, Ye WC, Zhang YD, Xu YS. High speed BLASTN: an accelerated Mega BLAST search tool. Nucleic Acids Res. 2015;43:7762–7768. https://doi.org/10.1093/nar/gkv784.
27. Lewis SE, Searle SMJ, Harris N, Gibson M, Iyer V, Richter J, et al. Apollo: a sequence annotation editor. Genome Biol. 2002;3:research0082.1. https://doi.org/10.1186/gb-2002-3-12-research0082.
28. Greiner S, Lehwark P, Bock R. Organellar Genome DRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 2019;47:W59–W64. https://doi.org/10.1093/nar/gkz238.
29. Zhang D, Gao FL, Jakovlić I, Zou H, Zhang J, Li WX, et al. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol. Ecol. Resour. 2020;20:348–355. https://doi.org/10.1111/1755-0998.13096
30. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 2016;33:1870–1874. https://doi.org/10.1093/molbev/msw054.
31. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. BIOINFORMATICS. 2017;33:2583–2585. https://doi.org/10.1093/bioinformatics/btx198.
32. Benson G. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 1999;27:573–580. https://doi.org/10.1093/nar/27.2.573.
33. 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;29:4633–4642. https://doi.org/10.1093/nar/29.22.4633.
34. Zhang HE, Meltzer P, Davis S. RCircos: an R package for Circos 2D track plots. BioMed Central Bioinformatics. 2013;14244. https://doi.org/10.1186/1471-2105-14-244.

35. 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;21:241. https://doi.org/10.1186/s13059-020-02154-5.
36. Shi LC, Chen HM, Jiang M, Wang LQ, Wu X, Huang LF, Liu C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 2019;47:W65–W73. https://doi.org/10.1093/nar/gkz345.
37. Liu SY, Ni Y, Li JL, Zhang XY, Yang HY, Chen HM, Liu C. CPGView: a package for visualizing detailed chloroplast genome structures**.** Mol. Ecol. Resour. 2023;23:694–704. https://doi.org/10.1111/1755-0998.13729.
38. Picardi E, Pesole G. REDItools: high-throughput RNA editing detection made easy. BIOINFORMATICS. 2013;29:1813–1814. https://doi.org/10.1093/bioinformatics/btt287.
39. Wang YP, Tang HB, DeBarry JD, Tan X, Li JP, Wang XY, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40:e49–e49. https://doi.org/10.1093/nar/gkr1293.
40. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 2013;30:772–780. https://doi.org/10.1093/molbev/mst010.
41. 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;32:268–274. https://doi.org/10.1093/molbev/msu300.
42. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47:W256–W259. https://doi.org/10.1093/nar/gkz239.
43. The Angiosperm Phylogeny Group. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 2016;181:1–20. https://doi.org/10.1111/boj.12385.
44. Morley SA, Nielsen BL. Plant mitochondrial DNA. FRONT BIOSCI-LANDMRK. 2017;22:1023–1032. https://doi.org/10.2741/4531.
45. Kubo T, Mikami T. Organization and variation of angiosperm mitochondrial genome. Physiol. Plant. 2007;129:6–13. https://doi.org/10.1111/j.1399-3054.2006.00768.x.
46. Chen H, Skylaris CK. Analysis of DNA Interactions and Gc Content with Energy Decomposition in Large-Scale Quantum Mechanical Calculations. Phys. Chem. Chem. Phys. 2021;23:8891–8899. https://doi.org/10.1039/D0CP06630C.
47. Gualberto JM, Newton KJ. Plant mitochondrial genomes: dynamics and mechanisms of mutation. Annu. Rev. Plant Biol. 2017;68:225–252. https://doi.org/10.1146/annurev-arplant-043015-112232.
48. Bock DG, Andrew RL, Rieseberg LH. On the adaptive value of cytoplasmic genomes in plants. Mol. Ecol. 2014;23:4899–4911. https://doi.org/10.1111/mec.12920.
49. Binder S, Brennicke A. Gene expression in plant mitochondria: Transcriptional and post–transcriptional control. Philos. Trans. R. Soc. Lond. B Biol. Sci. 2003;358:181–189.

https://doi.org/10.1098/rstb.2002.1179.

50. Zhou G, Qin M, Liu X, Qi Y, Ou X, Tang M. De Novo Assembly of the Mitochondrial Genome of *Glycyrrhiza glabra* and Identification of Two Types of Homologous Recombination Configurations Caused by Repeat Sequences. BMC Genom. 2025;26:13. https://doi.org/10.1186/s12864-024-11190-5.
51. Braun H, Binder S, Brennicke A, Eubel H, Fernie AR, Finkemeier I, et al. The life of plant mitochondrial complex I. MITOCHONDRION. 2014;19:295–313. https://doi.org/10.1016/j.mito.2014.02.006.
52. Maldonado M, Abe KM, Letts JA. A structural perspective on the RNA editing of plant respiratory complexes. Int J Mol Sci. 2022;23:684. https://doi.org/10.3390/ijms23020684.
53. Hammani K, Giege P. RNA metabolism in plant mitochondria. Trends Plant Sci. 2014;19:380–389. https://doi.org/10.1016/j.tplants.2013.12.008.
54. Korobeinikova AV, Garber MB, Gongadze GM. Ribosomal proteins: structure, function, and evolution. Biochem. (Mosc.). 2012;77:562–574. https://doi.org/10.1134/S0006297912060028.
55. Li J, Lu SS, Bi Y, Jiang YM, Feng LD, He J. Assembly and Analysis of the Mitochondrial Genome of *Hippophae rhamnoides* subsp. *sinensis*, an Important Ecological and Economic Forest Tree Species in China. Plants. 2025;14:2170. https://doi.org/10.3390/plants14142170.
56. Mower JP, Jain K, Hepburn NJ. The Role of Horizontal Transfer in Shaping the Plant Mitochondrial Genome. Adv. Bot. Res. 2012;63:41–69. https://doi.org/10.1016/B978-0-12-394279-1.00003-X.
57. Rice DW, Alverson AJ, Richardson AO, Young GJ, Sanchez-Puerta MV, Munzinger J, et al. Horizontal Transfer of Entire Genomes Via Mitochondrial Fusion in the *Angiosperm amborella*. Science. 2013;342:1468–1473. https://doi.org/10.1126/science.1246275.
58. Hou J, Liu M, Yang K, Liu B, Liu H, Liu J. Genetic variation for adaptive evolution in response to changed environments in plants. *J. Integr.* Plant Biol. 2025;67:2265–2293. https://doi.org/10.1111/jipb.13961.
59. Ou T, Wu Z, Liu Q, Tian C, Yang Y, Liu L, et al. Complete Mitochondrial Genome Of Medicago sativa ssp. Falcata (Papilionoideae, Fabaceae): Characterization and Phylogenetic Analysis. Planta. 2025;261:119. https://doi.org/10.1007/s00425-025-04698-x.
60. Sun M, Soltis DE, Soltis PS, Zhu XY, Burleigh JG, Zhiduan Chen ZD. Deep phylogenetic incongruence in the angiosperm clade Rosidae. Mol. Phylogenet. Evol. 2015;83:156–166. https://doi.org/10.1016/j.ympev.2014.11.003.

Figures



Figure 1

Morphology of the *A. carambola* plant and fruit. (A) Whole plant; (B) Inflorescence; (C) Whole fruit (left) and transverse section (right).

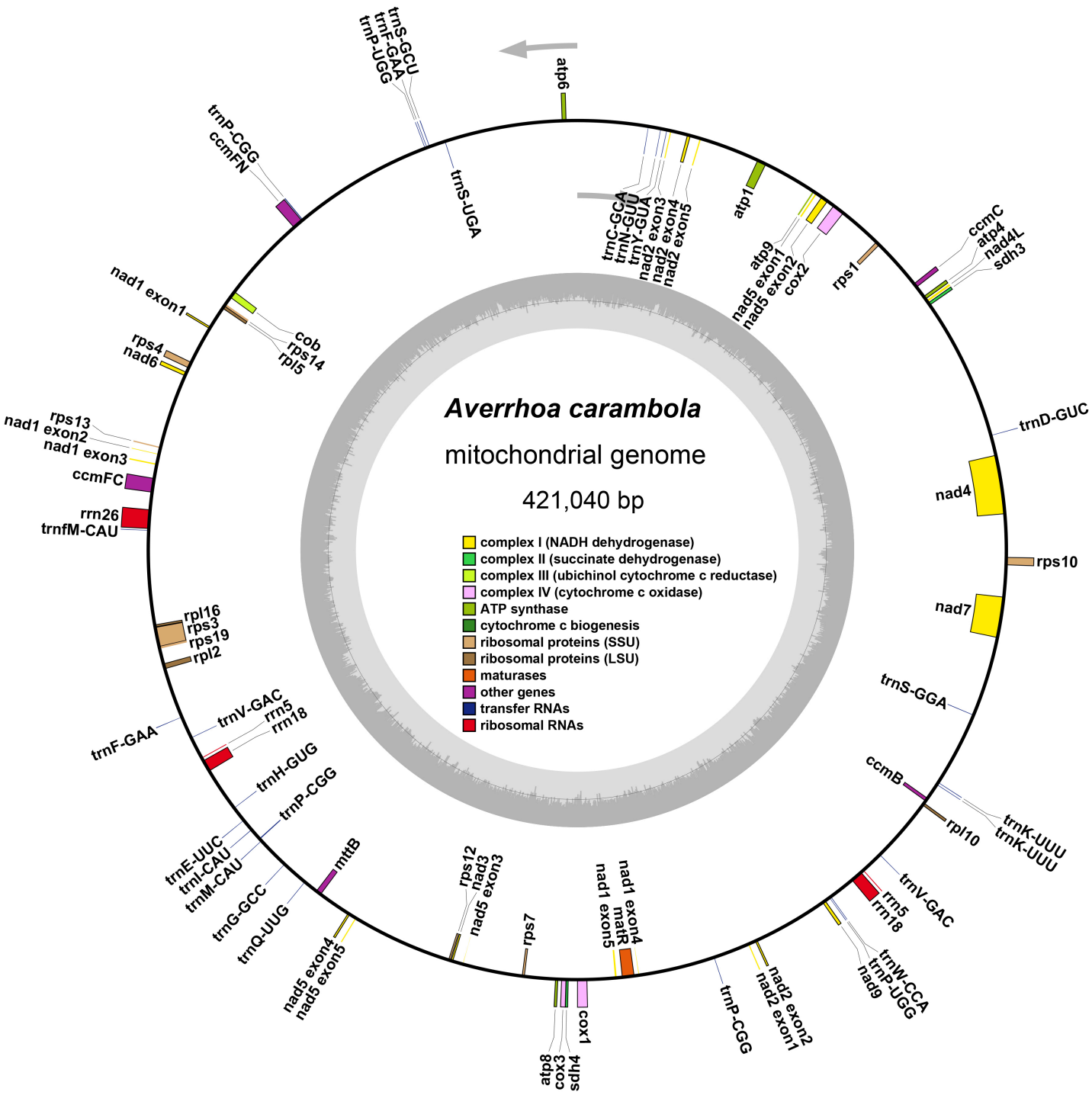


Figure 2

Circular map of the *A. carambola* mitochondrial genome.

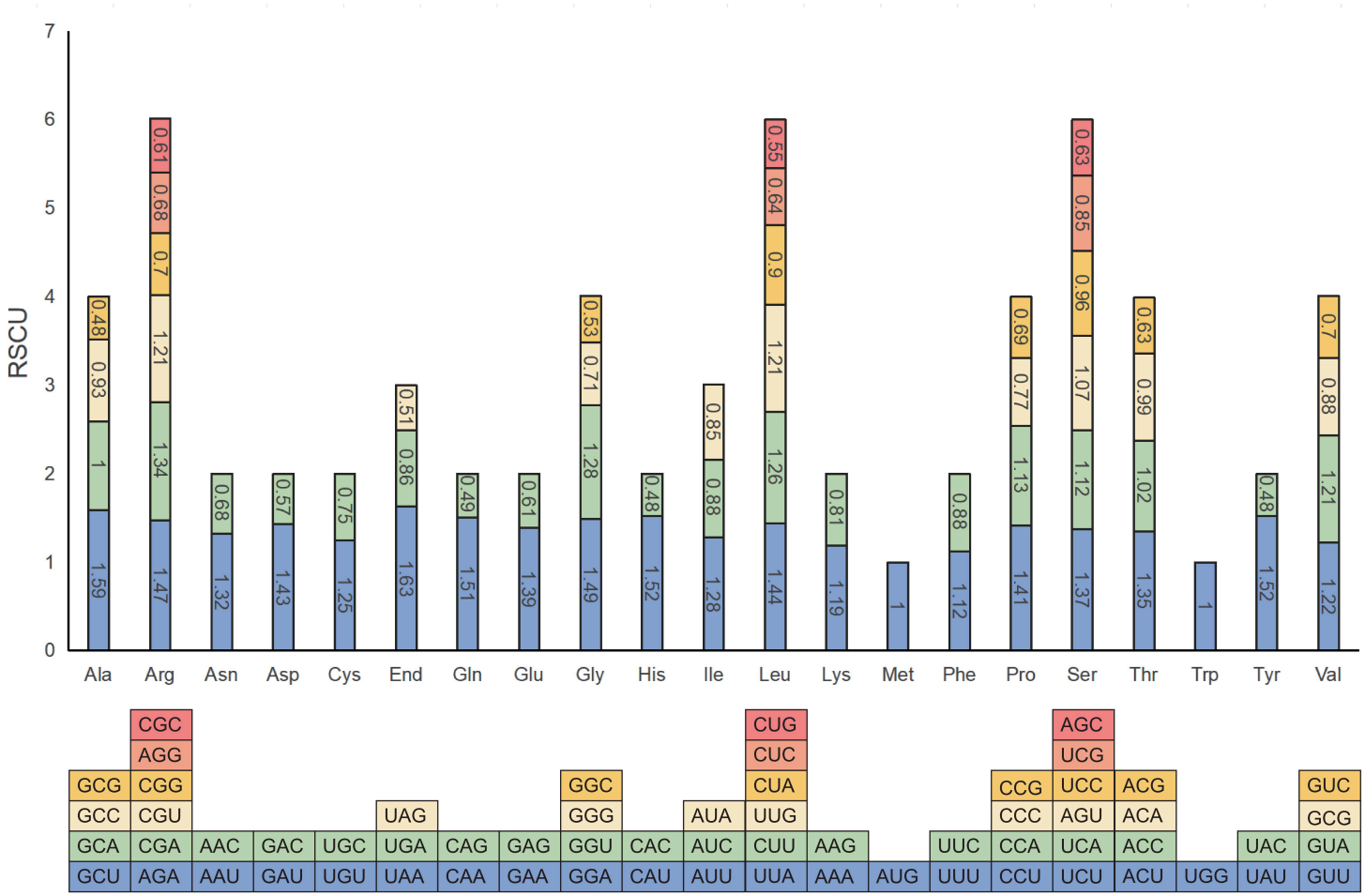


Figure 3

RSCU analysis of the *A. carambola* mitochondrial genome.

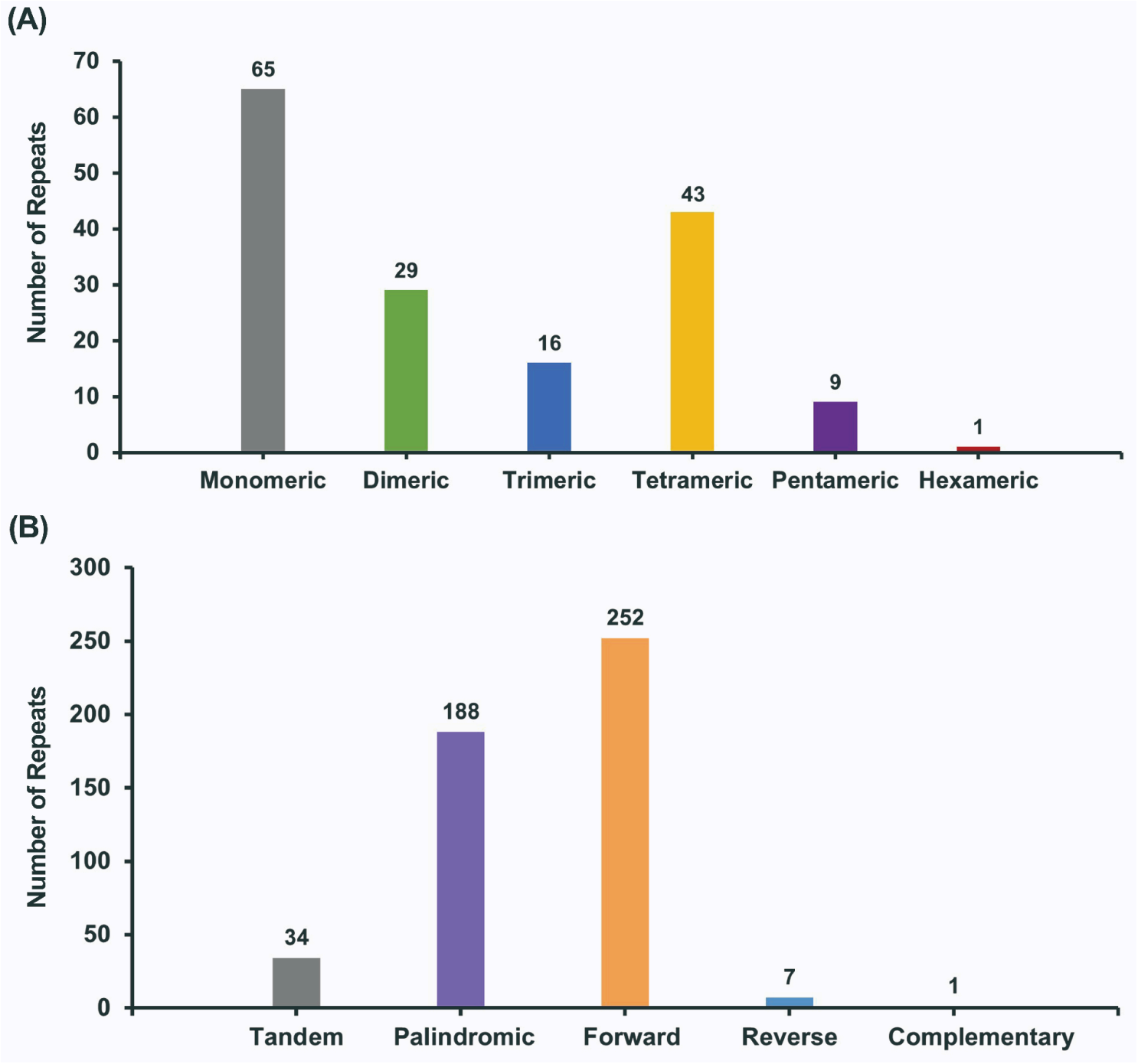


Figure 4

Analysis of repetitive sequences in *A. carambola*. (A) Distribution of different types of simple sequence repeats (SSRs). (B) Distribution of different types of interspersed repetitive sequences (length ≥30 bp).

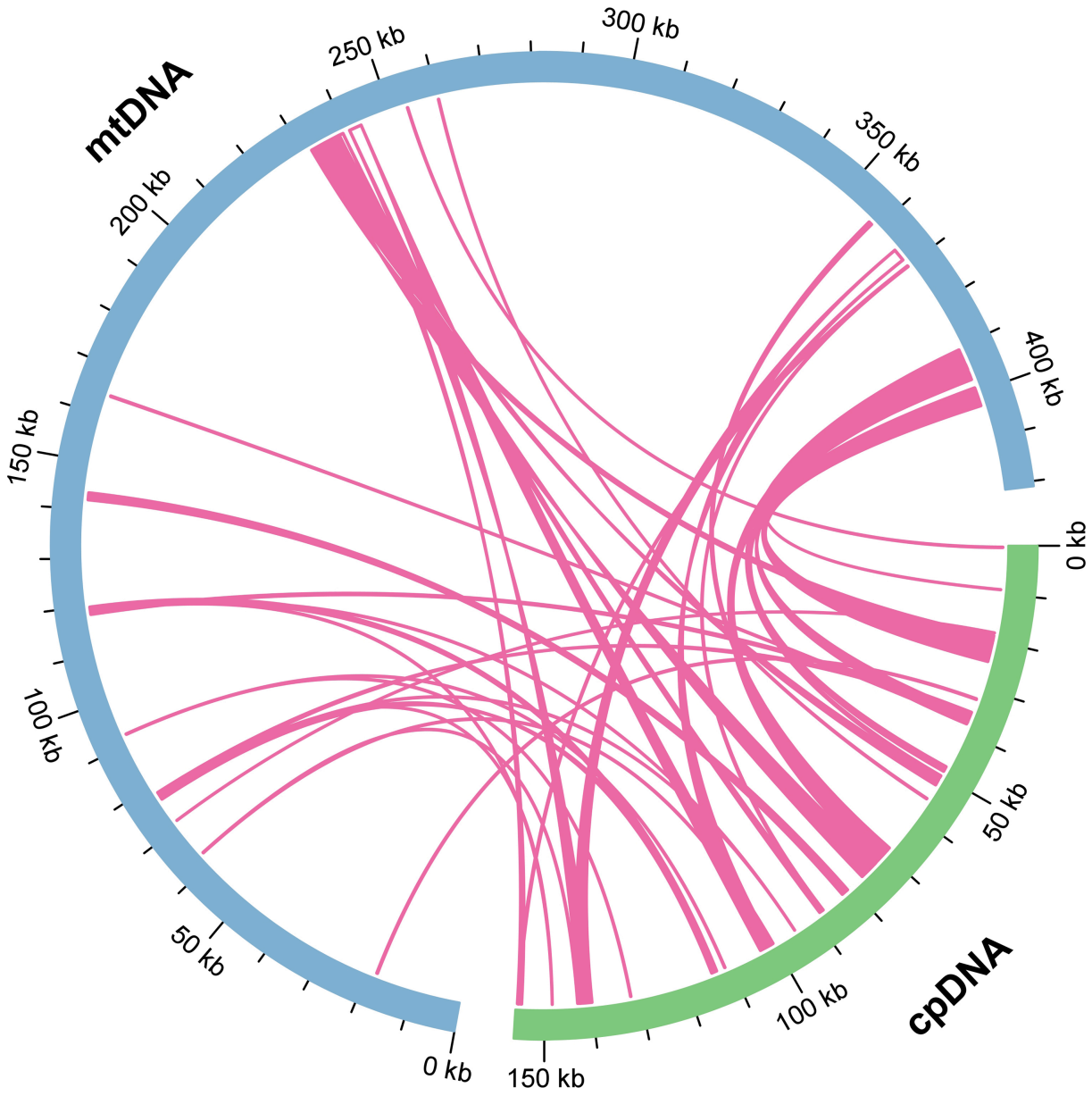


Figure 5

Distribution of mitochondrial-plastid homologous fragments in the mitochondrial genomes of *A. carambola*.

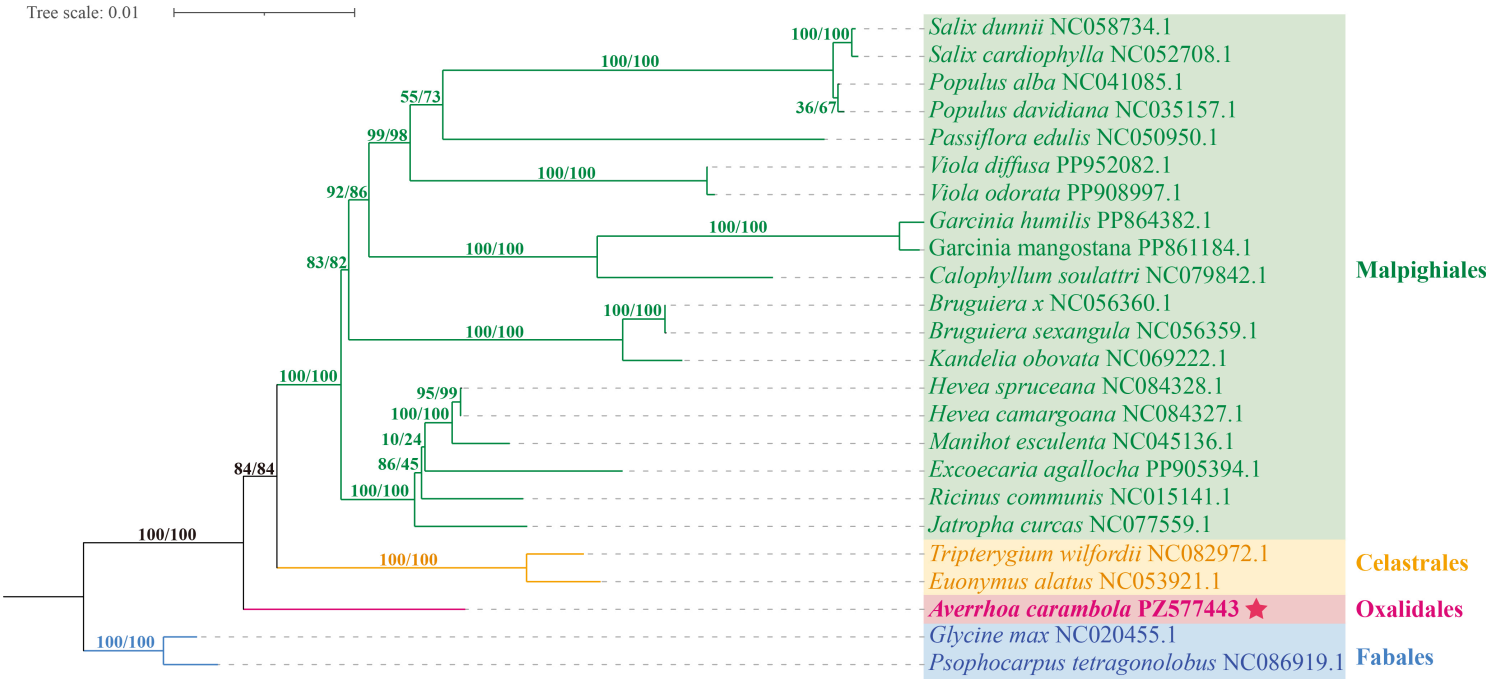


Figure 6

ML phylogenetic tree based on concatenated mitochondrial PCGs, with SH-aLRT test (1000 replicates, left) and ultrafast bootstrap (UFBoot, 1000 replicates, right) for branch support estimation.

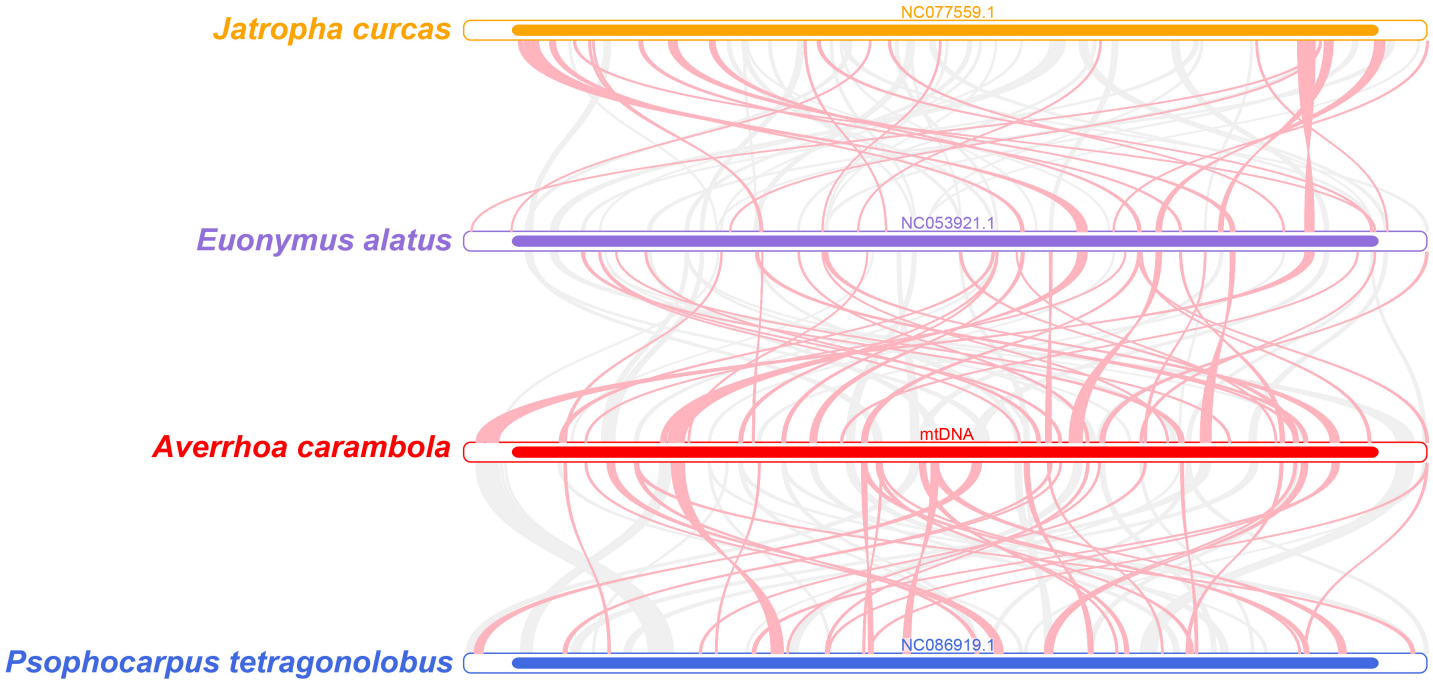


Figure 7

Collinearity comparison of mitochondrial genomes between *A. carambola* and related species. Red arcs indicate inverted regions, while gray arcs denote conserved homologous segments.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1SSRsinthemitochondrialgenomeofA.carambola.xlsx
- TableS2Dispersedrepeatsequencesinthemitochondrialgenomeof A.carambola.xlsx
- TableS3Tandemrepeatsequencesinthemitochondrialgenomeof A.carambola.xlsx
- TableS4ThehomologousDNAfragmentintheA.carambolamitochondrialgenome..xlsx
- TableS5RNAeditinganalysisinthemitochondrialgenomeofA.carambola..xlsx
- TableS6BLASTalignmentresultsofmitochondrialgenomecomparisonbetweenJ.curcasandE.alatus.xlsx
- FigureS1Circularmapofintragenomiccollinearityandrepeatpairings..jpg
- FigureS2DistributionofRNAeditingsitesacrossPCGs..jpg
- SupplementaryInformation.docx