Comparative Genomic Analysis of Mitochondrial Genomes from Two Lychee Cultivars

Jingyi You
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Ning Xu
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Ronghui Wen
State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bio-Resources, College of Life Science and Technology, Guangxi University, Nanning 530004, China
Yanjie Hou
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Dongbo Li
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Hongye Qiu
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Chen Fang
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Shuwei Zhang
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Xianquan Qin
Institute of Horticulture, Guangxi Academy of Agricultural Sciences
Hongli Li
lee2003ty@gxaas.net

Institute of Horticulture, Guangxi Academy of Agricultural Sciences

Research Article

Keywords: Litchi chinensis Sonn., Mitochondrial genome, Bioinformatics analysis, Phylogenetics

Posted Date: December 22nd, 2025

DOI: https://doi.org/10.21203/rs.3.rs-8259668/v1

Additional Declarations: No competing interests reported.

Abstract

Background

Lychee fruits are sweet and juicy, yet mitochondrial genomic data for this species remains scarce, limiting in-depth studies of its genetic and evolutionary characteristics. To address this gap, in this study, the small-seeded cultivar ‘Xianjinfeng’(XJF) and the large-seeded cultivar ‘Xinqiumili’ (XQML) were selected for analysis. Using third-generation sequencing technology, we sequenced, assembled, and annotated their mitochondrial genomes, and compared their structural characteristics and evolutionary relationships.

Result

Assembly revealed mitochondrial genome sizes of 579,270 bp for XJF and 579,261 bp for XQML, both with 45.41% GC content. Each genome encoded 62 genes, comprising 22 tRNAs, 3 rRNAs, and 35 protein-coding genes. Further analysis revealed 15 homologous sequences originating from chloroplasts in both mitochondrial genomes, totaling 12,194 bp (2.11% of the mitochondrial genome). These included 9 tRNA genes, 4 rRNA genes, and partial protein-coding sequences. Additionally, 184 simple sequence repeats (SSR) were identified in both cultivars, whereas 564 and 563 RNA editing sites were detected in XJF and XQML, respectively, indicating subtle genetic differences between the cultivars. This study also analyzed codon usage preferences, nucleotide diversity, and chloroplast-to-mitochondria gene transfer events. Collinearity and comparative genomics results indicate that lychee is closely related to *Nephelium lappaceum* L. and *Xanthoceras sorbifolium* Bunge within the Sapindaceae family. Phylogenetic analysis based on the mitochondrial genome further confirmed the evolutionary position of lychee within this family.

Conclusion

In this study, two high-quality lychee mitochondrial genomes were successfully assembled and annotated, enriching the mitochondrial genome resources of Sapindaceae plants and laying a foundation for future lychee phylogenetic and evolutionary studies of closely related species.

Introduction

Lychee (*Litchi chinensis* Sonn.) is a perennial fruit tree in the Sapindaceae family[1]. It is a significant tropical and subtropical fruit tree and is widely cherished for its bright red, sweet, and juicy fruit. The palatable flavor and high sugar content of its fruit[2, 3] confer substantial economic and horticultural value to this three. Lychee pulp is rich in various bioactive compounds, including polysaccharides, polyphenols, and gamma-aminobutyric acid (GABA). These components have been reported to exhibit antioxidant, antiobesity, hepatoprotective, and immunomodulatory effects[4].

The genome serves as the foundation for any omics analysis. The lychee genome of the Guangdong cultivated variety Feizixiao was decoded in 2020, significantly advancing the application of other lychee omics studies[5]. [6]Since the discovery of the genetic material within mitochondria[6], researchers subsequently identified all the substances required for replication, transcription, and protein translation within mitochondria, including DNA polymerase, RNA polymerase, transfer RNA, and ribosomal RNA. These findings collectively demonstrated that mitochondria possess a relatively independent genetic transcription system[7].

Mitochondria are ubiquitous, semiautonomous, powerhouse organelles of eukaryotic cells. Their genomes play pivotal roles in plant evolution, cellular metabolism, and species adaptation. Mitochondrial genomes synthesize adenosine triphosphate (ATP) through the tricarboxylic acid cycle and oxidative phosphorylation, providing energy for cellular life activities[8]. Mitochondria play critical roles in cellular differentiation, programmed cell death, growth, development, and male sterility[9]. Compared with animal mitochondrial genomes, plant mitochondrial genomes have larger sizes, higher proportions of repetitive sequences, and more complex structural and dynamic characteristics[10]. Their genome lengths vary widely, ranging from 66 kb to 11.3 Mb[11, 12], and typically encode 19–41 protein-coding genes (excluding repeat genes and open reading frames)[11] with extremely low point mutation rates [13, 14]. Plant mitochondrial genomes exhibit high structural diversity, presenting configurations such as single-stranded, multistranded, linear, or multichromosomal forms, with large variations in gene order and content across species[15]. This structural complexity primarily stems from intramolecular recombination mediated by repetitive sequences (such as forward, reverse, and palindromic repeats)[16, 17], the insertion of exogenous sequences, and horizontal gene transfer between chloroplasts and mitochondria[18]. Furthermore, RNA editing—particularly C→U editing—is a common feature of gene expression in plant mitochondria. This process can alter start codons, stop codons, and amino acid sequences, thereby influencing the structure and function of the final proteome[19]. Given the high sequence and structural variability of plant mitochondrial genomes and their rich phylogenetic information, they demonstrate significant potential for applications in species identification, phylogenetic reconstruction, and evolutionary studies[20]. Mitochondrial and chloroplast genomes are frequently used as models for phylogenetic reconstruction, studying DNA transfer from chloroplasts to mitochondria, understanding population differentiation, and assessing hybridization in angiosperms because of their nonrecombining and maternal inheritance[21, 22]. *Arabidopsis thaliana* was the first higher plant whose mtDNA was sequenced[21, 22]. However, despite the prominence of lychee in horticultural and breeding research, studies on its mitochondrial genome remain relatively underdeveloped. The severe scarcity of lychee mitochondrial genome-related information in public databases limits an in-depth understanding of the phylogenetic evolution and genetic mechanisms within this genus.

In recent years, the advancement of high-throughput sequencing technologies, particularly the widespread application of third-generation sequencing platforms, has enabled the de novo assembly and annotation of plant mitochondrial genomes. To date, mitochondrial genomes from several woody plants, including *Punica granatum*[24], *Abies alba* Mill[25], *Crataegus* spp.[26], and *Diospyros kaki*[27], have been

successfully sequenced, providing valuable resources for comparative genomics research. However, systematic investigations into the structural features, gene composition, RNA editing, codon usage preferences, and organelle-to-organelle gene transfer of the lychee mitochondrial genome remain largely unexplored.

In this study, PacBio HiFi third-generation sequencing technology was used to sequence, assemble, and annotate the mitochondrial genomes of the primary lychee cultivars ‘Xianjinfeng’ (XJF) and ‘Xinquimili’ (XQML). By systematically analyzing their genomic structures, repetitive sequences, codon usage patterns, RNA editing sites, and chloroplast-derived sequence migration, this study further explored the evolutionary relationships and phylogenetic positions of lychee relative to those of other species through comparative genomics and phylogenetic analyses. These findings fill a gap in mitochondrial genome data for lychee plants, providing crucial genetic resources and theoretical foundations for studies on the systematics, speciation, and adaptation of Sapindaceae plants.

Materials and Methods

Plant Samples and Mitochondrial Genome Sequencing

Two 10-year-old lychee cultivars, ‘Xianjinfeng’ (XJF) and ‘Xinqiumili’ (XQML) were cultivated at the experimental orchard of Guangxi Academy of Agricultural Sciences, Nanning, Guangxi, China (22.84°N, 108.32°E), under standardized horticultural management. The 'XJF' cultivar was jointly developed by the Institute of Fruit Tree Research, Guangdong Academy of Agricultural Sciences and Zengcheng Agricultural Technology Extension Center, and subsequently introduced to Nanning, Guangxi by the Guangxi Academy of Agricultural Sciences. The 'XQML' cultivar was developed by the Chinese Academy of Tropical Agricultural Sciences and introduced to Nanning, Guangxi by the Guangxi Academy of Agricultural Sciences. Young, healthy, fresh leaves were harvested from both cultivars, rapidly frozen in liquid nitrogen, and stored at − 80°C until mitochondrial genome sequencing. Sequencing was performed by a commercial biotechnology company using the PacBio Sequel II platform in HiFi sequencing mode.

Mitochondrial Genome Assembly and Annotation

The HiFi data were aligned to the assembled sequences using minimap2 software. NextPolish was then applied to refine the results, yielding the final assembled sequences and adjusting the corresponding gfa. Using BLAST, coding proteins and rRNA were annotated by alignment with published plant mitochondrial sequences serving as references, followed by manual refinement based on closely related species. tRNAs were annotated using tRNAscan-SE[28] (http://lowelab.ucsc.edu/tRNAscan-SE/). ORFs were annotated using Open Reading Frame Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). RNA editing sites were initially predicted using PmtREP (http://112.86.217.82:9919/#/tool/alltool/detail/336). Genome maps were generated with OGDRAW. The final annotation results were obtained after reviewing and manually correcting the initial outputs.

Codon Usage and Repeat Sequence Analysis

Simple sequence repeats (SSR) were identified using misa software (v1.0, parameters: 1–10 2–5 3–4 4 – 3 5 – 3 6 – 3). Tandem repeats were identified using trf software (trf409.linux64; parameters: 2 7 7 80 10 50 2000 -f -d -m). Dispersed repeats were identified using blastn (v2.10.1, parameters: -word_size 7, evalue 1e-5, with redundancy removal and tandem repeat exclusion). Visualization was performed using Circos v0.69-5.

Phylogenetic Tree Construction

Maximum likelihood phylogenetic trees were constructed using CDSs. Multiple sequence alignment was performed with MAFFT software (v7.427, --auto mode). The aligned sequences were concatenated end-to-end. The sequences were trimmed using trimAl (v1.4.rev15) (parameter: -gt 0.7). Model prediction was performed using jmodeltest-2.1.10 software, confirming the GTR model. RAxML v8.2.10 (https://cme.h-its.org/exelixis/software.html) software was subsequently used with the GTRGAMMA model and bootstrap = 1000 to construct the maximum likelihood phylogenetic tree.

Sequence Divergence, Ka/Ks Ratio, and Phylogenetic Analysis

First, homologous gene pairs were extracted and then aligned using MAFFT v7.427 (https://mafft.cbrc.jp/alignment/software/). After alignment, the Ka and Ks values for each gene pair were calculated using KaKs_Calculator v2.0 (https://sourceforge.net/projects/kakscalculator2/) and the MLWL method. Finally, the Ka/Ks values for each gene pair were summarized, and box plots were generated.

Chloroplast and Mitochondrial Homologous Sequence Analysis

BLAST software was used to identify homologous sequences between the chloroplast and mitochondrial genomes, with E-value set to 1e-5. The results were visualized using Circos v0.69-5. BLAST with an E value of 1e-5 was used to align the amino acid sequences of the chloroplast genome against those of the plant mitochondrial genome, clearly demonstrating which chloroplast-encoded sequences were transferred to the mitochondrial genome.

Results

Structural Organization and Characteristics of the Mitochondrial Genome

In this study, the mitochondrial genomes of the major lychee cultivars ‘XJF’ and ‘XQML’ were sequenced, assembled, and annotated. Through third-generation sequencing data assembly (Table S1), the lychee mitochondrial genomes were successfully assembled into linear DNA molecules (Fig. 1). The total lengths of the two mitochondrial genomes were 579,270 bp and 579,261 bp, respectively, with a GC content of 45.41% (Table 1). The relevant sequences will be submitted to GenBank (GenBank accession are not yet available).

Table 1
Table of Assembly Results Summary

ID	Type	Length	GC%
Litchi_chinensis_XJF-1	linear	579270	45.41
Litchi_chinensis_XQML-1	linear	579261	45.41

A total of 62 genes were annotated, including 35 protein-coding genes, 22 tRNA genes, and 3 rRNA genes (rrn5, rrn18, and rrn26). The protein-coding genes were classified into the following 10 functional categories: ATP synthase genes (atp1, atp4, atp6, atp8, and atp9), NADH dehydrogenase genes (nad1–nad9), cytochrome c synthesis-related genes (ccmB, ccmC, ccmFC, and ccmFN), cytochrome c oxidase genes (cox1, cox2, and cox3), a mature enzyme-encoding gene (matR), a membrane transporter gene (mttB), and a pantothenol-cytochrome c reductase gene (cob). The variable genes included four large-subunit ribosomal protein genes (rpl10, rpl16, rpl2, and rpl5), five small-subunit ribosomal protein genes (rps1, rps10, rps12, rps4, and rps13), and two succinate dehydrogenase genes (sdh3 and sdh4) (Table 2).

Table 2
Genes identified in the two lychee mitochondrial genomes

Group of genes	Gene name
ATP synthase	atp1 atp4 atp6 atp8 atp9
Cytohrome c biogenesis	ccmB ccmC ccmFc* ccmFn
Ubiquinol cytochrome c reductase	cob
Cytochrome c oxidase	cox1 cox2* cox3
Maturases	matR
Transport membrane protein	mttB
NADH dehydrogenase	nad1**** nad2**** nad3 nad4*** nad4L nad5**** nad6 nad7**** nad9
Ribosomal proteins (LSU)	rpl10 rpl16 rpl2* rpl5
Ribosomal proteins (SSU)	#rps14 #rps3 rps1 rps10* rps12 rps13 rps4
Succinate dehydrogenase	sdh3 sdh4
Ribosomal RNAs	rrn18 rrn26 rrn5
Transfer RNAs	trnC-GCA(2) trnD-GTC trnE-TTC trnF-GAA trnG-GCC trnH-GTG trnI-TAT* trnK-TTT trnK-TTT* trnM-CAT(3) trnN-GTT trnP-TGG(2) trnQ-TTG trnS-GCT(2) trnS-TGA trnW-CCA trnY-GTA

Identification of RNA Editing Sites in Protein-Coding Regions of Mitochondrial Genes

Using online prediction tools, a total of 564 and 563 potential RNA editing sites were identified in the 32 protein-coding genes of XJF and XQML, respectively, all of which were C→U edits. The nad4 gene contained the greatest number of editing sites (44), followed by ccmFn (41 sites), while nad3 had the lowest number (1 site) (Fig. 2).

Codon Preference and Analysis of Repeat Structures in the Mitochondrial Genomes of Two Lychee Species

Among the two lychee varieties, leucine (Leu), serine (Ser), and arginine (Arg) were the most frequent amino acids, whereas tryptophan (Trp) and methionine (Met) appeared least frequently.

Relative Synonymous Codon Usage (RSCU)analysis revealed that 4,989–5,072 codons (58.12%–59.87% of total codons) in the XJF and XQML mitochondrial genomes exhibited RSCU values greater than 1, indicating stronger usage preferences for these codons relative to their synonymous counterparts. Most codons ending in A or U had RSCU values greater than 1, suggesting that their usage frequency exceeded random expectations. Conversely, codons ending in C or G generally had RSCU values less than 1, indicating a significant preference for A/U-terminated codons in the lychee mitochondrial genome (Fig. 3).

Multiple types of repetitive sequences, including forward, palindromic, reverse, and complementary repeats, were identified in each lychee mitochondrial genome. Overall, 396 repetitive sequences were identified, comprising 47 tandem repeats and 165 dispersed repeats (Table 3). SSR lengths ranged from 10 to 20 bp, with the vast majority concentrated between 10 and 14 bp (78.3–81.6%). The 12-bp repeat unit was the most common in both cultivars, and 46 pairs of homologous SSRs with ≥ 90% similarity were identified, spanning lengths from 10 to 16 bp (Figure S1).

Table 3
Repetitive Sequences in Two Lychee Genomes

Sample	SSR nu	Tandem nu	Dispersed nu	Total
Litchi_chinensis_XJF	184	47	165	396
Litchi_chinensis_XQML	184	47	165	396

Phylogenetic and Collinearity Analysis

A phylogenetic tree constructed on the basis of 15 conserved mitochondrial protein-coding genes (PCGs) indicated that lychee is closely related to rambutan(Fig. 4). Collinearity analysis further supported the presence of numerous homologous blocks between the two species (Figure S2). Litchi is relatively close to rambutan and Chinese cornel, revealing that mitochondrial protein-coding genes serve as ideal materials for elucidating phylogenetic relationships among different plant species.

Ka/Ks Analysis

Amino acid changes caused by base mutations are termed nonsynonymous mutations, whereas those causing no change are synonymous mutations. Nonsynonymous mutations are typically subject to natural selection. The ratio of the nonsynonymous mutation rate (Ka) to the synonymous mutation rate (Ks) indicates the type of selection pressure. A Ka/Ks ratio greater than 1 indicates positive selection, whereas a ratio less than 1 indicates purifying selection. We aligned protein-coding genes from two lychee mitochondrial genomes with those from other species and calculated Ka/Ks values (Table S2). As shown in Fig. 5, 35 genes underwent negative selection during evolution, indicating that most protein-coding genes in the mt genome are relatively conserved. The protein-coding genes ccmB, rpl10, mttB, and rps13 all exhibited Ka/Ks ratios greater than 1, suggesting that these genes may have undergone

positive selection. Genes with Ka/Ks > 1 likely play crucial roles in species adaptation to environments, functional differentiation, or coevolution.

Migration of chloroplast DNA in the mitochondrial genome

Among the two lychee varieties, 15 homologous segments of chloroplast origin were identified, totaling 12,194 bp in length and accounting for 2.11% of the mitochondrial genome. These segments included 9 tRNA genes, 4 rRNA genes, and partial protein-coding sequences. This finding indicates frequent gene transfer between chloroplasts and mitochondria (Table S3).

Discussion

Structure and Characteristics of the Lychee Mitochondrial Genome

The diversity of plant mitochondrial genome structures is among their most distinctive features. In this study, the mitochondrial genomes of XJF and XQML were successfully assembled as linear DNA molecules. This differs from the circular structure of hawthorn[26], the multilinear structure of the pomegranate “Taishan Hong[26],” and the typical circular structure of the persimmon “Taishu[27],” potentially reflecting a unique genomic organization characteristic of Sapindaceae plants. Notably, the mitochondrial genome of *Aquilegia amurensis* has a hybrid configuration of one circular chromosome + two linear chromosomes (total length 538.7 kb)[29], whereas *Fagopyrum dibotrys* has a more complex structure comprising seven circular chromosomes (total length 370 kb)[30]. These findings indicate that medium-sized (370–580 kb) mitochondrial genomes exhibit significant structural plasticity in angiosperms and that the presence of linear segments is not unique to lychee. Compared with the genomes of conifers, the lychee genome is smaller but has a more compact repetitive sequence structure: the *Abies alba* mitochondrial genome spans 1.43 Mbp across 11 scaffolds, with repetitive elements constituting 0.168% of its sequence[25]. It contains nine giant repeats exceeding 10 kb, directly driving genome expansion and multiconformation transitions. The396 repeat units in lychee (accounting for approximately 68.4% of the genome) exhibit higher density but lack such megamatches. This explains its relatively simple structure while maintaining high genetic variability potential.

Codon Preferences and Repetitive Sequence Analysis

Codon usage preferences are influenced by natural selection and other factors. In both lychee varieties, leucine (Leu), serine (Ser), and arginine (Arg) are high-frequency amino acids, whereas tryptophan (Trp) and methionine (Met) are less frequent. This pattern resembles the amino acid usage characteristics of hawthorn, reflecting the shared protein synthesis requirements among closely related genera and species. RSCU analysis revealed that 58.12%-59.87% of codons in both varieties presented RSCU values exceeding 1, predominantly ending in A/U. Conversely, C/G-terminated codons frequently fell below 1, aligning with preferences observed in pomegranate and persimmon. This may stem from greater tRNA availability for A/U-terminated codons, facilitating efficient protein synthesis. With respect to repetitive

sequences, the two varieties share 396 repeats (47 tandem and 165 scattered). SSRs predominantly cluster within 10–14 bp (78.3–81.6%), with 12-bp repeats being the most common, comprising 46 pairs of highly similar homologous SSRs. This differs markedly from that in hawthorn (222–447 repeats, 65–112 SSRs) and pomegranate (188 pairs of scattered repeats, 141 SSRs), likely because of variations in repeat sequence generation and amplification rates. Homologous SSRs may serve as markers for genetic analysis.

Characteristics of RNA Editing Sites

RNA editing is a crucial post-transcriptional modification in plant mitochondria. Among the 32 protein-coding genes in the two lychee cultivars, 564 and 563 C→U editing sites were identified, which is consistent with the editing patterns observed in hawthorn, pomegranate, and persimmon[25]. This finding indicates that C→U editing represents a conserved key mechanism. There were significant differences in gene editing site counts: nad4 (44 sites) had the greatest number, followed by ccmFn (41 sites), while nad3 (1 site) had the lowest. This pattern resembles the high number of nad4 editing sites in pomegranate, likely because the involvement of nad4 in energy metabolism requires more editing to ensure function, whereas nad3 sequences are more conserved. Among editing types, hydrophilic→hydrophobic substitutions were most prevalent (44.15%-44.23%), followed by hydrophobic→hydrophobic substitutions (35.52%-35.64%), mirroring the distribution pattern observed in pomegranate. This distribution may alter protein conformation and interactions, facilitating environmental adaptation.

Gene Transfer between Chloroplasts and Mitochondria

The transfer of genes from chloroplasts to mitochondria is a common phenomenon. The two lychee varieties share 15 homologous chloroplast fragments (12,194 bp, accounting for 2.11% of the mitochondrial genome), including 9 tRNAs, 4 rRNAs, and partial protein-coding sequences. Compared with hawthorn (12–19 fragments), pomegranate (22 fragments, 5.20%), and persimmon (5.71%), differences in fragment numbers and proportions reflect varying frequencies of organelle gene exchange among species. Transferred genes predominantly consist of tRNAs and rRNAs, as seen in hawthorn and pomegranate. These genes may increase mitochondrial protein synthesis efficiency; some protein-coding sequences may have become pseudogenes because of mutations, providing raw material for genomic evolution. Gene transfer may also alter mitochondrial genome structure and influence evolutionary rates.

Phylogenetic and Collinearity Analysis

A phylogenetic tree constructed on the basis of 15 PCGs revealed that lychee is most closely related to rambutan, which is consistent with the phylogenetic clustering logic observed in hawthorn and pomegranate. These findings indicate that conserved mitochondrial PCGs reliably reflect evolutionary relationships among species. The limited phylogenetic resolution between the two cultivated lychee varieties stems from differences of only 9 bases and 1 RNA editing site. Collinearity analysis revealed extensive homologous blocks shared between lychee, rambutan, and Chinese cornel, indicating that

conserved genomic regions potentially harbor core functional genes. Similar collinearity patterns observed among closely related species such as hawthorn and pomegranate, along with minor gene rearrangements, reflect independent evolutionary trajectories following speciation. These findings provide a foundation for lychee systematics and origin/divergence studies while advancing the understanding of mitochondrial evolutionary patterns in fruit trees.

Conclusion

In this study, the mitochondrial genomes of the lychee cultivars XJF and XQML were successfully assembled and annotated using third-generation sequencing technology, providing critical foundational data for genetic and evolutionary research in the genus *Litchi*. The mitochondrial genomes of XJF and XQML are linear DNA molecules measuring 579,270 bp and 579,261 bp, respectively, with identical GC content (45.41%). A total of 62 genes were annotated, comprising 35 protein-coding genes, 22 tRNA genes, and 3 rRNA genes (rrn5, rrn18, and rrn26). The protein-coding genes included core functional categories such as ATP synthase and NADH dehydrogenase, similar to the core gene composition of fruit tree mitochondrial genomes, including hawthorn (*Crataegus* spp.), pomegranate (*Punica granatum*), and persimmon (*Diospyros kaki*). It also exhibits species-specific gene distribution patterns, such as variations in the number of specific tRNA genes. These findings suggest that while the lychee mitochondrial genome maintains fundamental physiological functions, it has developed unique genetic characteristics adapted to its own growth and development. A total of 396 repetitive sequences (47 tandem repeats and 165 scattered repeats) and numerous SSR loci were identified. Among these, 46 pairs of highly similar homologous SSRs provide abundant candidate loci for genetic diversity analysis and molecular marker development in lychee. RNA editing sites were exclusively C→U substitutions, totaling 563–564 instances primarily distributed in genes such as nad4 and ccmFn. Hydrophilic→hydrophobic edits accounted for the greatest proportion, indicating that RNA editing plays a crucial role in optimizing mitochondrial gene function in lychee. Both lychee varieties presented 15 chloroplast-derived homologous fragments in their mitochondrial genomes (total length of 12,194 bp, accounting for 2.11% of the mitochondrial genome), including 9 tRNA genes, 4 rRNA genes, and partial protein-coding sequences. These findings confirm frequent gene transfer between lychee chloroplasts and mitochondria, providing direct evidence for understanding the evolution of lychee organelle genomes. Phylogenetic and colinearity analyses based on 15 conserved mitochondrial protein-coding genes revealed a close relationship between lychee and *Nephelium lappaceum*. Colinearity analysis further confirmed extensive homologous blocks between the two species, clarifying the evolutionary position of lychee within the plant taxonomic system. In summary, this study elucidates the structural features, genetic element patterns, and evolutionary relationships of the lychee mitochondrial genome. It not only enriches the plant mitochondrial genome database but also provides crucial theoretical support and data references for germplasm resource conservation, genetic breeding, and systematic classification studies of *Litchi* species.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflict of interest.

Funding

This research was funded by Science and Technology Development Fund of Guangxi Academy of Agricultural Sciences(Grant No. GuiNongKe 2025YP001),Guangxi Science and Technology Major Program-Breeding and Demonstration Application of New Litchi Varieties in Guangxi (Grant No. Gui Ke AA23023007), China Agriculture Research System of MOF and MARA (Grant No. CARS-32-02).

Author Contribution

Conceptualization, analysis of literary data, writing, J.Y.; conceptualization and resources, N.X.,R.W.,Y.H.and D.L.; supervision, and review, H.Q., C.F., S.Z., and editing, J.Y.,X.Q. and H.L. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

Not applicable.

Data Availability

The mt genomes of XJF (GenBank accession: PX659691) and XQML(GenBank accession: PX659692) were assembled with the assembly software and have been submitted to NCBI (https://www.ncbi.nlm.nih.gov/).

References

1. Menzel C. The lychee crop in Asia and the Pacific. FAO Regional Office for Asia and the Pacific; 2002.

2. Chen X, et al. Metabolic variations in the pulp of four litchi cultivars during pulp breakdown. Food Res Int. 2020;140:110080.
3. Yao P et al. Litchi (*Litchi chinensis* Sonn.): a comprehensive review of phytochemistry, medicinal properties, and product development. Food Funct. 12(20): pp. 9527–48.
4. Reichel M, et al. Changes in Flavonoids and Nonphenolic Pigments during On-Tree Maturation and Postharvest Pericarp Browning of Litchi (*Litchi chinensis* Sonn.) As Shown by HPLC-MS ~ n. J Agric Food Chem. 2011;59(8):3924–39.
5. Hu G, et al. Two divergent haplotypes from a highly heterozygous lychee genome suggest independent domestication events for early and late-maturing cultivars. Nat Genet. 2022;54(1):73–83.
6. NASS MM, NASS S. Intramitochondrial fibers with dna characteristics. I. Fixation and electron staining reactions. J Cell Biol. 1963;19(3):593–611.
7. Chen XJ, Butow RA. The organization and inheritance of the mitochondrial genome. Nat Rev Genet. 2005;6(11):815–25.
8. Gualberto JM, et al. The plant mitochondrial genome: dynamics and maintenance. Biochimie. 2014;100:107–20.
9. Aken OV, Breusegem FV. Licensed to Kill: Mitochondria, Chloroplasts, and Cell Death. Trends Plant Sci. 2015;20(11):754–66.
10. Kozik A, et al. The alternative reality of plant mitochondrial DNA: One ring does not rule them all. PLoS Genet. 2019;15(8):e1008373.
11. Skippington E et al. Miniaturized mitogenome of the parasitic plant Viscum scurruloideum is extremely divergent and dynamic and has lost all nad genes. Proceedings of the National Academy of Sciences - PNAS, 2015. 112(27): pp. E3515-E3515.
12. Sloan DB, et al. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 2012;10(1):e1001241.
13. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci U S A. 1987;84(24):9054–8.
14. PALMER, et al, et al. Plant mitochondrial-dna evolves rapidly in structure, but slowly in sequence. J MOL EVOL; 1989.
15. Mower JP, Sloan DB, Alverson AJ. Plant Mitochondrial Genome Diversity: The Genomics Revolution. 2012.
16. Cole LW et al. High and Variable Rates of Repeat-Mediated Mitochondrial Genome Rearrangement in a Genus of Plants. Mol Biol Evol, 2018(11): p. 11.
17. Davila JI, et al. Double-strand break repair processes drive evolution of the mitochondrial genome in Arabidopsis. BMC Biol. 2011;9:64.
18. Cusimano N, Wicke S. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. New Phytol. 2016;210(2):680–93.

19. Mower JP, Palmer JD. Patterns of partial RNA editing in mitochondrial genes of Beta vulgaris. Mol Genet Genomics. 2006;276(3):285–93.
20. Zubaer A, Wai A, Hausner G. The mitochondrial genome of Endoconidiophora resinifera is intron rich. Sci Rep. 2018;8(1):17591.
21. Jin LY, Ram UY, Seong KYMC. Characterization and comparative analysis of the complete organelle genomes of three red macroalgae species (Neoporphyra dentata, Neoporphyra seriata, and Neopyropia yezoensis) and development of molecular makers for their identification. Genes genomics. 2024;46(3):355–65.
22. Liu H, et al. A large-scale population based organelle pan-genomes construction and phylogeny analysis reveal the genetic diversity and the evolutionary origins of chloroplast and mitochondrion in *Brassica napus* L. BMC Genomics. 2022;23(1):1–14.
23. Unseld M, et al. The mitochondrial genome of Arabidopsis thaliana contains 57 genes in 366,924 nucleotides. Nat Genet. 1997;15(1):57–61.
24. Feng L, et al. Multichromosomal mitochondrial genome of *Punica granatum*: comparative evolutionary analysis and gene transformation from chloroplast genomes. BMC Plant Biol. 2023;23(1):512–512.
25. Kersten B, et al. The mitochondrial genome sequence of *Abies alba* Mill. reveals a high structural and combinatorial variation. BMC Genomics. 2022;23(1):776–15.
26. Zhang X, et al. Comparative genomic and phylogenetic analyses of mitochondrial genomes of hawthorn (*Crataegus spp.*) in Northeast China. Int J Biol Macromol. 2024;272(Part1):11.
27. Yang Y, Duan C. Mitochondrial genome features and systematic evolution of diospyros kaki thunb 'Taishuu'. BMC Genomics. 2024;25(1):285.
28. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic acids research, 2019. 1997.25.5(5): pp. 955 – 64.
29. Xu L, et al. Characterizing complete mitochondrial genome of *Aquilegia amurensis* and its evolutionary implications. BMC Plant Biol. 2024;24(1):142.
30. Feng G, et al. Complete mitochondrial genome assembly and comparative analysis of Fagopyrum dibotrys (Golden Buckwheat). BMC Plant Biol. 2025;25(1):985.

Figures

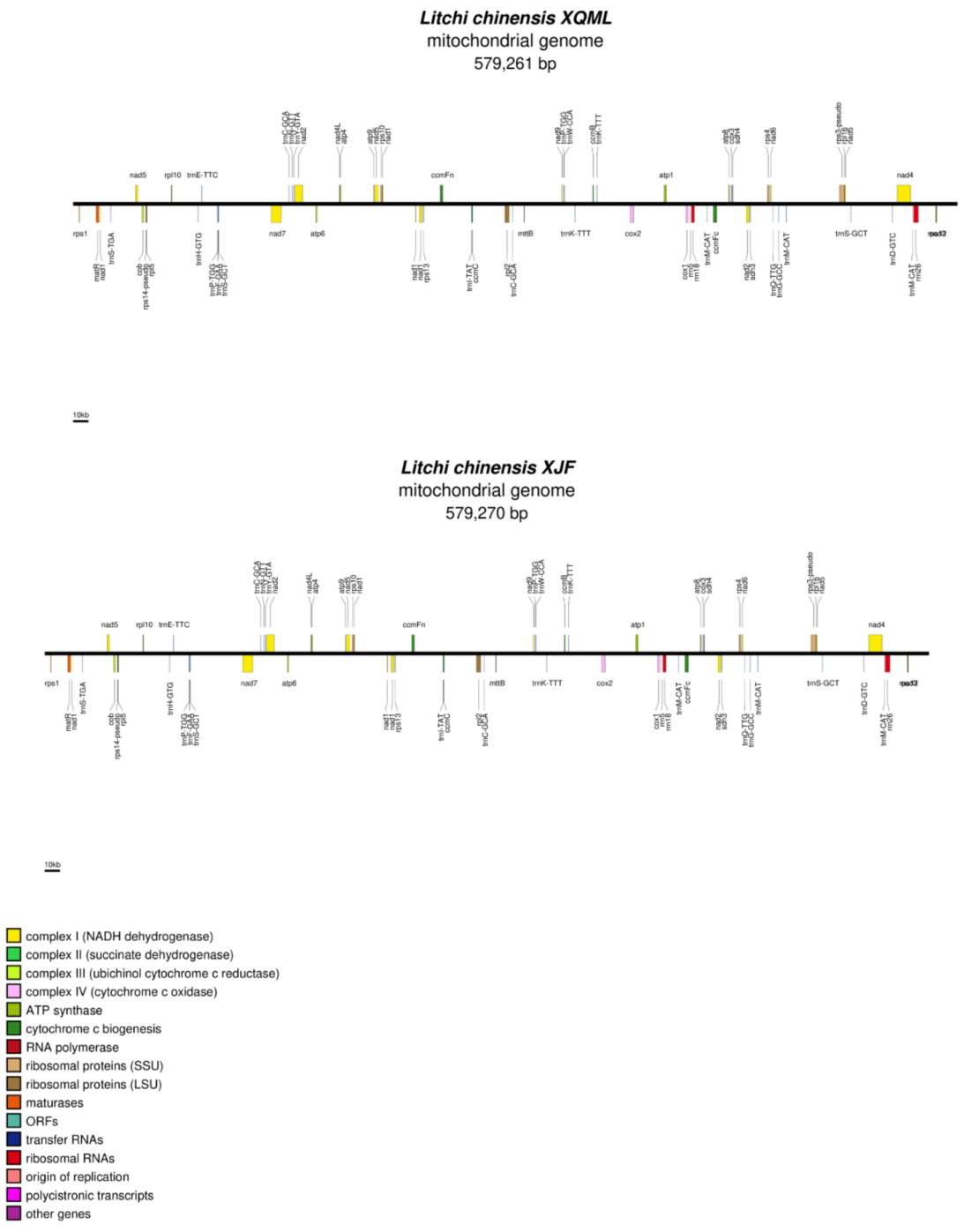

Figure 1

Two lychee linear mitochondrial genome maps; forward-encoding genes are positioned above the circle, while reverse-encoding genes are located below the circle.

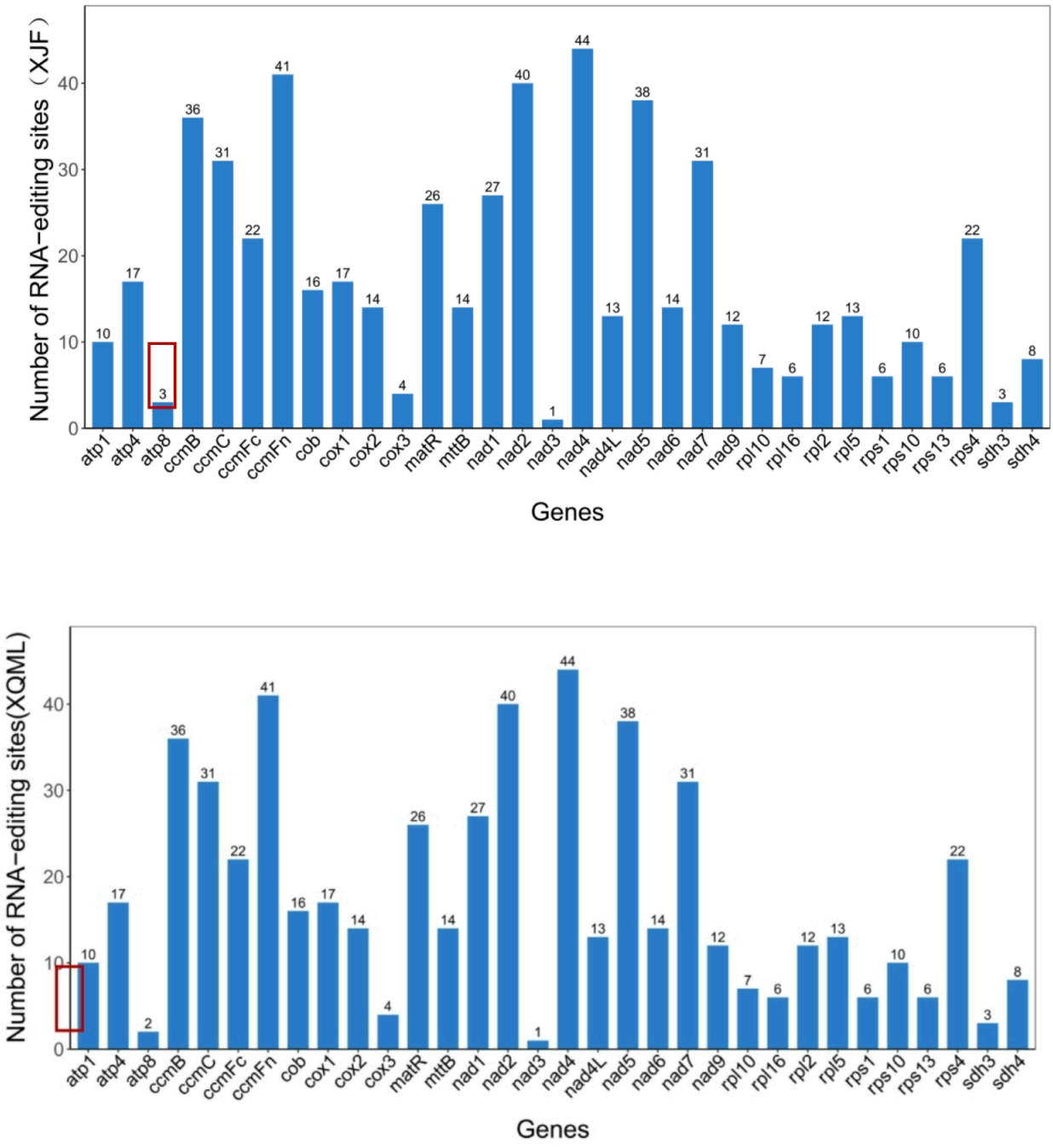

Figure 2

RNA editing sites in the two lychee mitochondrial genomes

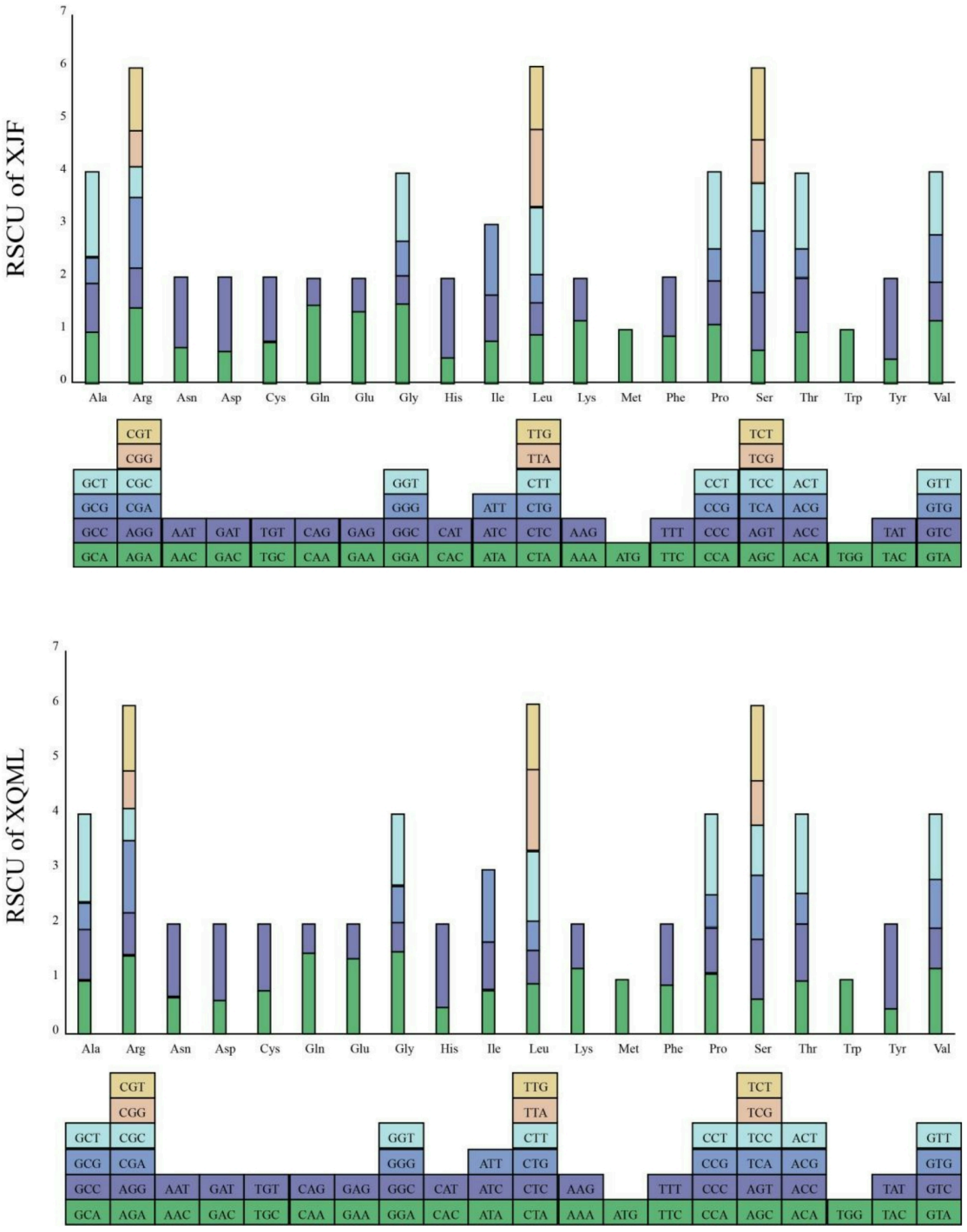

Figure 3

RSCU bar chart of the two lychee mitochondrial genomes. The lower blocks represent all codons encoding each amino acid, while the height of the upper bars indicates the total RSCU value for all codons.

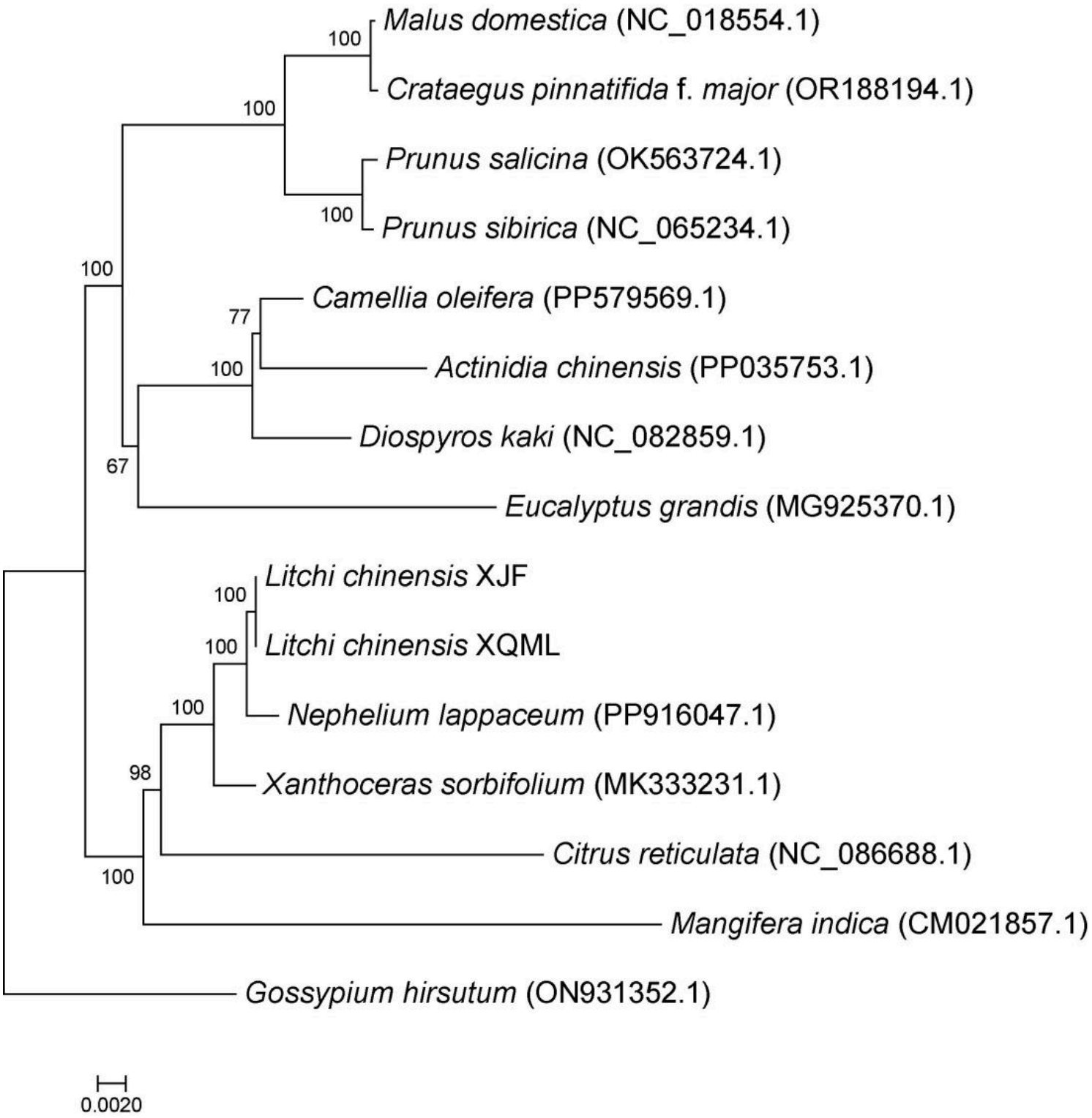

Figure 4

Mitochondrial Evolutionary Analysis. Evolutionary branch length, also known as genetic variation or evolutionary distance, represents the degree of branch variation; the shorter branch length represents a smaller difference and closer evolutionary distance. Distance scale: unit length of numerical differences between organisms or sequences, equivalent to the scale of the evolutionary tree. Bootstrap value is marked at the node position and used to evaluate the credibility of the branch

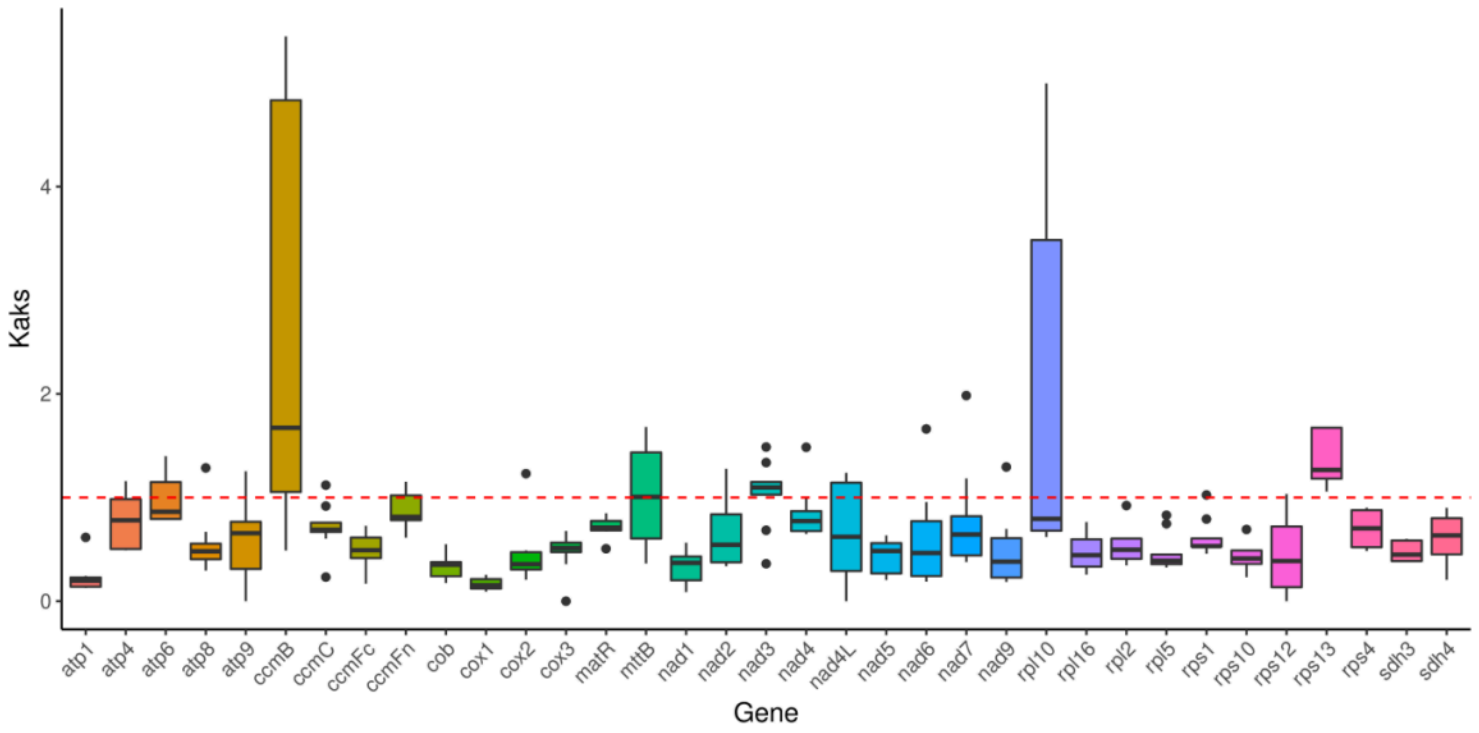

Figure 5

KaKs Box Plot Analysis Across Species. The horizontal axis represents gene names, and the vertical axis represents Ka/Ks values. In the box plot, the upper and lower endpoints of the vertical line within the rectangle represent the upper and lower boundaries of the data, respectively. The thick line inside the rectangle represents the median, the upper and lower edges represent the upper and lower quartiles, respectively, and the dots beyond the upper and lower boundaries of the data represent outliers