scientific reports

Check for updates

OPEN

Structural and phylogenetic comparisons of the complete mitochondrial genomes among taxa in genus *Toona*

Yan-Wen Lv[1,2], Zi-Yun Wang[1,2], Zi-Han He[1,2], Yu Xiao[1,2], Chao Wu[1,2] & Xin-Sheng Hu[1,2]✉

Toona **is of a high economic value and has a great potential for industrial and medicinal utilization. Currently, the mitochondrial genome information of this genus is incomplete. Here we sequenced the mitochondrial genomes of *T. rubriflora* and *T. microcarpa*, and re-sequenced four varieties of *T. ciliata*. The mitochondrial genome of *T. rubriflora* was 653,710 bp in length, with a typical circular structure and the GC content of 45.42%. The mitochondrial genome of *T. microcarpa* was successfully assembled in two circulars, designated as chromosome 1 and chromosome 2, respectively. Chromosome 1 had 474,320 bp and its GC content was 45.37%. Chromosome 2 had 166,958 bp and its GC content was 46.64%. Comparative analyses of *Toona* mitochondrial genomes revealed that species were highly conserved in GC content, coding gene sequences and codon usage frequency. Mononucleotide repeats were the dominant type of genome repeats. Small differences existed among species in RNA editing sites, intracellular genome homology between mitochondrial and chloroplast genomes and between mitochondrial and nuclear genomes. Most protein coding genes (PCGs) were under purifying selection. Species in *Toona* were relatively recently divergent, and the varieties of *T. ciliata* were well genetically mixed. Overall, this study provided comprehensive information on *Toona* mitochondrial genomes, which could be used for species identification and molecular studies on *Toona*.**

Keywords *Toona*, Mitochondrial genomes, Phylogeny, Purifying selection, *Toonaciliata*

The *Toona* genus belongs to the Meliaceae family and is widely distributed in Asia and Oceania[1]. In China, the species of *Toona* are distributed in southern and southwestern regions as well as across northern regions. The genus is composed of four species in China: *T. sinensis*, *T. ciliata*, *T. microcarpa*, and *T. rubriflora*. Five varieties of *T. ciliata* are recognized according to leaf and flower characteristics, including *T. ciliata* var. *ciliata*, *T. ciliata* var. *yunnanensis*, *T. ciliata* var. *henryi*, *T. ciliata* var. *sublaxiflora*, and *T. ciliata* var. *pubescens*. Historically, the genus was misclassified as the *Cedrela* genus and officially recognized as a separate genus by M.J. Roemer in 1846. The misclassifications can also be indicated from multiple synonyms of a species. From Royal Botanic Gardens at Kew (http://www.kew.org/, accessed on September 9, 2024), there were eighteen synonyms for *T. sinensis* and the classified genera were *Ailanthus*, *Cedrela*, *Mioptrila*, *Surenus*, and *Toona*. There were twenty-one synonyms for *T. ciliata* and the classified genera were *Cedrela*, *Surenus*, *Swietenia* and *Toona*. There were eight synonyms for *T. microcarpa* and the classified genera were *Cedrela*, *Surenus*, *Swietenia* and *Toona*. This taxonomic complexity probably arose from large morphological variation within and between species. For example, inhabitants in western Yunnan historically considered the purple buds of *T. sinensis* as"*T. microcarpa*"and occasionally as"*T. ciliata*"or"*T. ciliata* var. *pubescens*". In studying the nutrient composition of young leaves of *T. microcarpa* under two different cultivation conditions, Luo et al. (2014) mistakenly identified the local purple buds of *T. sinensis* as "*T. microcarpa*"[2]. Other processes, such as interspecific hybridization[3,4], may also lead to misclassification. A similar situation occurred in identifying varieties of *T. ciliata* based on the leaf and flower variations[1]. These varieties are sympatric or partially overlapping in geographical distribution and bring challenges to delimitate them. Thus, it is significant in taxonomy to elucidate the phylogenetic relationships among species in genus *Toona* and among varieties in *T. ciliata*.

Practically, the species of *Toona* are ideal materials for making furniture and interior decoration because of beautiful texture and excellent decay resistance[5–8]. They also have medicinal values in China[9–12]. *T. sinensis* has

[1]College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou 510642, China. [2]Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, Guangzhou 510642, China. ✉email: xinsheng@scau.edu.cn

been cultivated in China for more than 2000 years. *T. ciliata* is not as widely distributed as *T. sinensis* in China but has a great potential for plantations. The species has been classified as an endangered species at Level II in China due to over logging and low nature regeneration[13–15]. Both *T. microcarpa* and *T. rubriflora* are more locally distributed. Although *T. microcarpa* was not on the list of rare and endangered plants in China, Zou et al. pointed out that it would be potentially endangered because global warming could influence the survival and reproduction of this species[16]. Thus, it is practically important to clarify their genetic divergence for conservation of *Toona* genetic resources.

Previous studies have examined genetic divergence among partial species. Wang et al. examined the phylogenetic relationship between *T. sinensis* and *T. ciliata* using nuclear genome sequences and estimated their divergent time at about 15.06 (6–25) Mya[3,17,18]. Xiao et al. used nuclear ribosomal DNA (nrDNA) ITS and mitochondrial DNA (mtDNA) markers to assess population variation in terms of the *T. ciliata* complex[19]. Both marker analyses showed the presence of geographic variation[20,21].

Chloroplast genomes (cpDNA) were also applied to evaluating phylogenetic relationship among partial species in genus *Toona*. Three species (*T. sinensis*, *T. ciliata*, *T. microcarpa*) have comparable lengths of cpDNA sequences[22]. Xiao et al. examined phylogenetic relationships among *T. sinensis* and four varieties of *T. ciliata*[19]. Wang et al. used cpDNA markers to investigate geographic variation among populations in terms of the *T. ciliata* complex[23]. However, mitochondrial genome (mtDNA) has not been applied to assess phylogenetic relationships among species in genus *Toona* or among varieties of *T. ciliata*. Although mtDNA sequences of *T. ciliata* (NC_065060.1) and *T. sinensis* (NC_065061.1) were recorded in the literature, a comparison was only confined to these two species, without fully exploring the phylogeny of *Toona*[24]. It is well recognized that mtDNA is characterized by high variability in genomic structure, size, and rearrangement but conservative in coding sequences, which poses challenges for sequencing and assembly[25]. However, mtDNA is also used in comparative genomic analyses due to several key features, such as the high rate of evolution in animals or a lower mutation rate in plants compared with cpDNA, simple structure and maternal inheritance. Records showed that mtDNA was effective in elucidating phylogenetic connections between closely related species[26,27]. Here, we sequenced and assembled mitochondrial genomes of *T. rubriflora* and *T. microcarpa* and re-sequenced varieties of *T. ciliata* for assessing genomic variations among taxa in *Toona*.

The purposes of this study were (1) to compare the similarity and divergence in mitochondrial genome among species of genus *Toona*; (2) to analyze evolutionary mechanisms underlying species divergence; and (3) to reconstruct the phylogenetic relationships among species of genus *Toona* and among varieties of *T. ciliata*. The overall analyses provided complementary information to the existing taxonomic division of taxa and consolidated previous recommendations of genetic conservation on the *T. ciliata* complex from the perspective of mitochondrial genomes.

Results

Structural comparison and gene annotation

Substantial structural variation in mtDNA sequence occurred among species of *Toona*. For *T. ciliata*, we used the mtDNA sequence of individual NC_065060.1 of *T. ciliata* (GenBank Access NC_065060) as the reference sequence. The mtDNA sequences of twenty samples of *T. ciliata* could be categorized into two classes based on the coverage rate of sequencing reads against the reference sequence (Supplementary Table S1). Class A sequences had 100% coverage in alignment with the reference sequence. Small variation in the mtDNA sequences occurred among 16 class A samples. There were 119 single-nucleotide polymorphisms (SNPs) in total. Class B was represented by *T. ciliata* var. *yunnanensis* 1, with approximately 95% coverage in alignment with the reference sequence. Its mitochondrial genome was assembled in circular structure with 831,995 bp in length (Fig. 1A). The GC content was 45.40%, and the four base contents were 27.26% A, 27.33% T, 22.85% G, and 22.54% C (Table 1). There were 80 genes in total (Supplementary Table S2), including 35 PCGs, 42 transfer RNAs (tRNAs) and 3 ribosomal RNAs (rRNAs). Notably, *T. ciliata* var. *yunnanensis* 1 had a larger genome size compared to the other varieties of *T. ciliata* although comparable GC contents were observed. The mitochondrial genome of *T. rubriflora* was assembled in circular structure, with 653,710 bp in length (Fig. 1B). The GC content was 45.42%, and four base contents were comparable to those in mtDNA of *T. ciliata* var. *yunnanensis* 1 (Table 1). There were 70 genes in total (Supplementary Table S3), including 35 PCGs, 32 tRNAs, and 3 rRNAs (Table 2).

The mitochondrial genome of *T. microcarpa* was successfully assembled in two circulars, designated as chromosome 1 and chromosome 2, respectively (Figs. 1C and 1D). Chromosome 1 had 474,320 bp and its GC content was 45.37%. The nucleotide compositions were 27.33% A, 27.30% T, 22.88% G and 22.50% C (Table 1). There were 44 genes in total (Supplementary Table S4), including 25 PCGs, 18 tRNAs and 1 rRNA. Chromosome 2 had 166,958 bp, where A, T, G, and C contents were 27.35%, 45, 27.02%, 22.79%, and 22.85%, respectively. The GC content was 46.64%. There were 26 genes in total, including 10 PCGs, 14 tRNAs and 2 rRNAs.

There were 35 PCGs and 3 rRNAs genes shared among species, while the number of tRNAs slightly varied from 32 to 42 genes across species. Except for *T. sinensis*, the protein-coding sequences for each species were 31,788 bp in length (Table 1). Among 35 PCGs, there were 19 genes encoding the proteins that were related to electron transport and ATP synthesis, including 9 for complex I (*nad*1, *nad*2, *nad*3, *nad*4, *nad* 4 l, *nad*5, *nad*6, *nad*7, and *nad*9), 2 for complex II (*sdh*3 and *sdh*4), 1 for complex III (*cob*), 3 for complex IV (*cox*1, *cox*2, *cox*3), and 5 for complex V (*atp*1, *atp*4, *atp*6, *atp*8, and *atp*9). Furthermore, there were 9 genes associated with ribosomal proteins (*rpl*2, *rpl*5, *rpl*10, *rpl*16, *rps*1, *rps*3, *rps*4, *rps*10 and *rps*12), 4 with cytochrome c biogenesis (*ccm*B, *ccm*C, *ccm*FN, and *ccm*FC), and 2 with other functions (*mat*R and *mtt*B). In chromosome 1 of *T. microcarpa*, there were 4 genes for complex I (*nad*3, *nad* 4 l, *nad*7, and *nad*9), 1 gene for complex II (*sdh*4), 1 gene for complex III (*cob*), 3 genes for complex IV (*cox*1, *cox*2, *cox*3), and 4 genes for complex V (*atp*1, *atp*4, *atp*6, and *atp*8). Additionally, there were 8 ribosomal protein genes (*rpl*2, *rpl*5, *rpl*10, *rpl*16, *rps*1, *rps*3, *rps*10, *rps*12), 2 genes involved in cytochrome c biogenesis (*ccm*C and *ccm*FC), and 2 other genes (*mat*R and *mtt*B). In chromosome

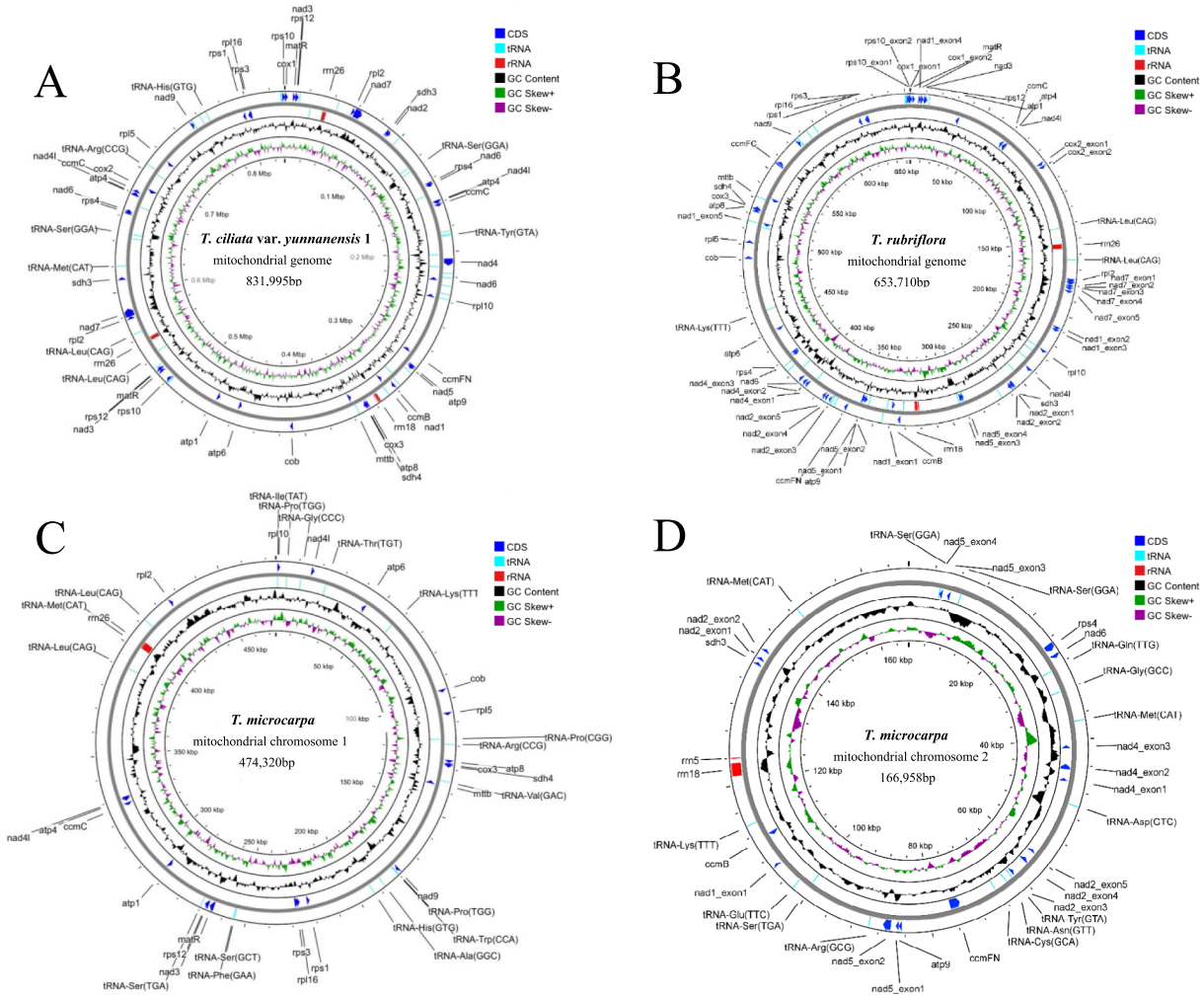

Fig. 1. The mitochondrial genome maps of three species of *Toona*. (**A**) *T. ciliata* var. *yunnanensis* 1; (**B**) *T. rubriflora*; (**C**) *T. microcarpa* (chromosome 1); (**D**) *T. microcarpa* (chromosome 2).

Species	Length/Percentage	A	T	G	C	Total	GC content
T. ciliata var. *yunnanensis* 1	Size (bp)	226,873	227,419	190,145	187,558	831,995	377,703
	Percentage (%)	27.26	27.33	22.85	22.54		45.40
T. rubriflora	Size (bp)	178,264	178,512	149,623	147,311	653,710	296,934
	Percentage (%)	27.27	27.31	22.89	22.53		45.42
T. microcarpa (chromosome 1)	Size (bp)	129,622	129,467	108,510	106,721	474,320	215,231
	Percentage (%)	27.33	27.30	22.88	22.50		45.37
T. microcarpa (chromosome 2)	Size (bp)	45,657	45,107	38,043	38,151	166,958	76,194
	Percentage (%)	27.35	27.02	22.79	22.85		45.64
T. ciliata (NC_065060.1)	Size (bp)	186,538	186,361	156,039	154,062	683,000	310,101
	Percentage (%)	27.31	27.92	22.85	22.56		45.40
T. sinensis (NC_065061.1)	Size (bp)	174,639	172,983	145,343	145,527	638,482	290,870
	Percentage (%)	27.35	27.09	22.76	22.79		45.56

Table 1. Comparison of mitochondrial genome structure among species of *Toona*.

Species		*T. ciliata* var. *yunnanensis* 1	*T. rubriflora*	*T. microcarpa*	*T. ciliata* (NC_065060.1)	*T. sinensis* (NC_065061.1)
PCGs	genes	35	35	35	35	35
	Total length(bp)	31,788	31,788	31,788	31,788	31,743
rRNA	genes	3	3	3	3	3
	Total length(bp)	8,205	5,189	5,189	5,189	5,189
tRNA	genes	42	32	32	33	34
	Total length(bp)	2,212	2,426	2,426	2,499	2,635
Total number of genes		80	70	70	71	72

Table 2. Comparison of coding genes in mitochondrial genomes among species of *Toona*.

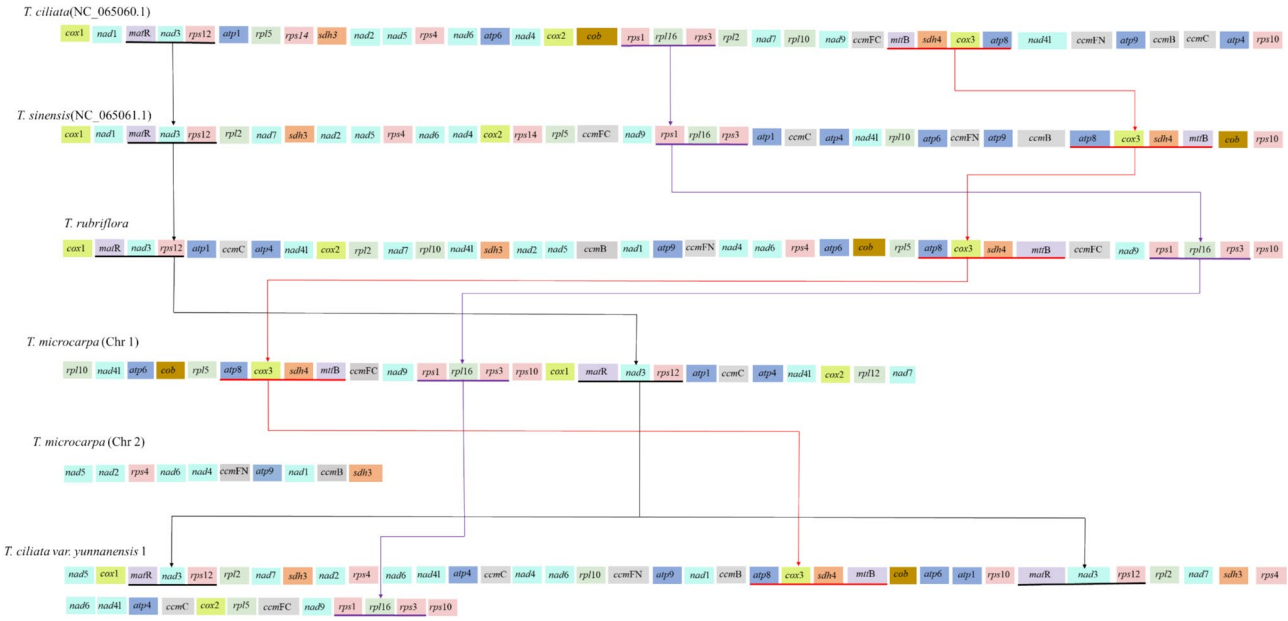

Fig. 2. Structural comparison of protein-coding genes in mitochondrial genomes across species of *Toona*.

2 of *T. microcarpa*, there were 5 genes for complex I (*nad*1, *nad*2, *nad*4, *nad*5, and *nad*6), 1 gene for complex II (*sdh*3), 1 gene for complex V (*atp*9), 1 gene for ribosomal protein (*rps*4), and 2 genes responsible for cytochrome c biogenesis (*ccm*B and *ccm*FN).

Protein coding genes and codon usage

Comparison of PCGs along genome revealed that only the *mat*R-*nad*3-*rps*12 and *rps*1-*rpl*16-*rps*3 gene fragments were fully conserved in all species (Fig. 2). The gene fragment *atp*8-*cox*3-*sdh*4-*mtt*B in *T. ciliata* was entirely reversed, compared with the other species. Notably, *T. ciliata* var. *yunnanensis* 1 exhibited a high frequency of gene rearrangements (*rps*12-*rpl*2-*nad*7-*sdh*3 and *rps*4-*nad*6-*nad*4l-*atp*4-*ccm*C).

Most PCGs had ATG as the start codon except for four genes (*rps*10, *rps*4, *nad*1, *nad* 4 l) where ACG was the start codon. Gene *rps*10 possessed a non-typical stop codon CGA. Figure 3A shows the frequencies of different amino acids in PCGs in three newly sequenced species. The Leu was the most frequent amino acid, followed by Ser and Ile. Codons TTT (Phe), TTC (Phe) and ATT (Ile) were most common among 64 codons. Each amino acid had its preferred codon except for Met (ATG) and Trp (TGG) that had only one codon. All species had 30 optimal codons (RSCU > 1) except for the start codons ATG and TGG that had the RSCU values of 1 (Fig. 3B). The GCT codon had the highest RSCU value. The RSCU values for all NNA and NNT codons exceeded 1.00 except for GCA, ATA, and CTA (Supplementary Table S5).

Prediction of RNA editing sites

We used PREP (predictive RNA editors for plants) to predict RNA editing sites in 35 PCGs, with a cutoff value of 0.2. *T. microcarpa* had the most RNA editing sites (277), followed by *T. ciliata* (275) and *T. sinensis* (275), and *T. rubriflora* (272) and *T. ciliata* var. *yunnanensis* 1(272). There were 270 editing sites shared among species. All editing sites were changed from base C to T, and the transition from TCA to TTA was the most frequent (Supplementary Table S6). Gene *ccm*B had the highest frequency of RNA editing sites, whereas genes *nad* 4 l, *rps*1 and *sdh*4 had only one editing site. No editing events were detected in 9 genes (*atp*1, *atp*6, *atp*8, *ccm*FN, *ccm*FC, *cox*3, *mtt*B, *nad*3 and *rps*12) (Fig. S1).

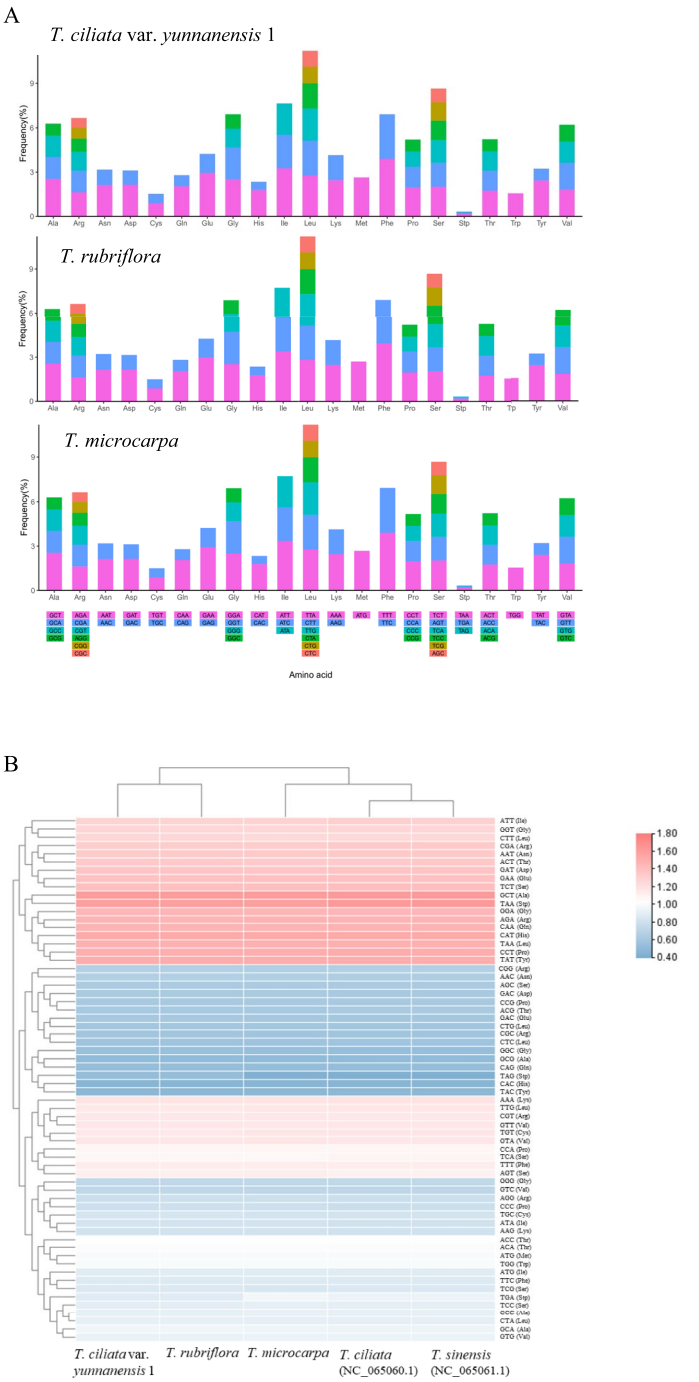

Fig. 3. Codon usage frequencies and RSCU values of amino acids in mitochondrial genomes of *Toona*. (**A**) Codon usage frequencies in protein-coding genes in *T. ciliata* var. *yunnanensis* 1, *T. rubriflora* and *T. microcarpa*; (**B**) A heatmap derived from RSCU value for each amino acid in each species of *Toona*.

Most editing sites altered amino acids, encompassing 15 types of amino acid changes in total. A major proportion of these changes were the conversion of hydrophilic amino acids to hydrophobic amino acids (48.5%). The most frequent conversions were from Ser to Leu, Pro to Leu, and Ser to Phe. Additionally, a few genes had codons transformed into the stop codons (TAA, TAG, TGA). One common editing was that one of codons for glutamine in *rpl*16 was converted to a stop codon (CAG to TAG) and another codon for arginine in *atp*9 was converted to a stop codon (CGA to TGA) in all species. There was one editing site converting arginine to a stop codon in gene *rps*10 only in *T. sinensis* and *T. ciliata* var. *yunnanensis* 1.

Repeat sequences

Figure 4A shows the frequencies of different types of SSRs in mitochondrial genomes across species, including mononucleotides, dinucleotides, trinucleotides, tetranucleotides and hexanucleotides repeats. Notably, the pentanucleotide repeats were absent, and the tetranucleotide repeats were identified exclusively in *T. ciliata*. Mononucleotide repeats were most abundant. The (A)n and (T)n repeats were more than the (G)n and (C)n repeats. The (AT)n and (TA)n repeats were more than the (GA)n and (AG)n repeats (Supplementary Table S7). All species exhibited minor variations in repeat abundance.

Most long repeats were between 30 and 40 bp in length, and the longest was 162,009 bp in *T. ciliata* var. *yunnanensis* 1(Fig. 4B). *T. ciliata* var. *yunnanensis* 1 had the most repeats (212), while *T. microcarpa* had the fewest repeats (61 in chromosome 1 and 9 in chromosome 2). Among these repeat sequences, most of them were forward and palindromic repeats and did not contain complementary repeat sequences. Reverse repeats were exclusively identified in *T. microcarpa* (genome 2). *T. ciliata* var. *yunnanensis* 1 had the longest forward repeat sequence (30,433 bp) and palindromic repeat sequence (735 bp) compared to other species.

Intracellular homologous comparisons

Intracellular horizontal transfer was analyzed between chloroplast and mitochondrial genomes in species of *Toona* (excluding *T. rubriflora*) using BLASTn. Analysis of genomic homology indicated that *T. ciliata* var. *yunnanensis* 1 had the longest homologous sequences (29,582 bp in total), accounting for about 3.6% of its mitochondrial genome (Fig. 5A). The total homologous sequences between *T. ciliata* and *T. sinensis* were 24,065 bp and 22,588 bp in length, representing 3.8% and 3.3% of their respective mitochondrial genomes. *T. microcarpa* had the shortest length of homologous sequences (21,677 bp in total), accounting for 3.3% of its mitochondrial genome (Fig. S2). Furthermore, we identified PCGs, tRNA and rRNA genes that were transferred between chloroplast and mitochondrial genomes in both *T. ciliata* var. *yunnanensis* 1 and *T. ciliata*. Eight genes with 100% homologous sequences were identified in both mitochondrial and chloroplast genomes in *T. ciliata* var. *yunnanensis* 1 and *T. ciliata*, including 7 tRNA genes (trnM-CAT, trnS-GGA, trnW-CCA, trnD-GTC, trnI-GAT, trnH-GTG, and trnN-GTT) and 1 PCG (*psb*F).

We also analyzed homologous sequences of the mitochondrial genomes of *T. ciliata* and *T. ciliata* var. *yunnanensis* 1 with the nuclear genome of *T. ciliata*, respectively. The results revealed that the mitochondrial genome of *T. ciliata* var. *yunnanensis* 1 was partially homologous with nuclear genomes across 28 chromosomes

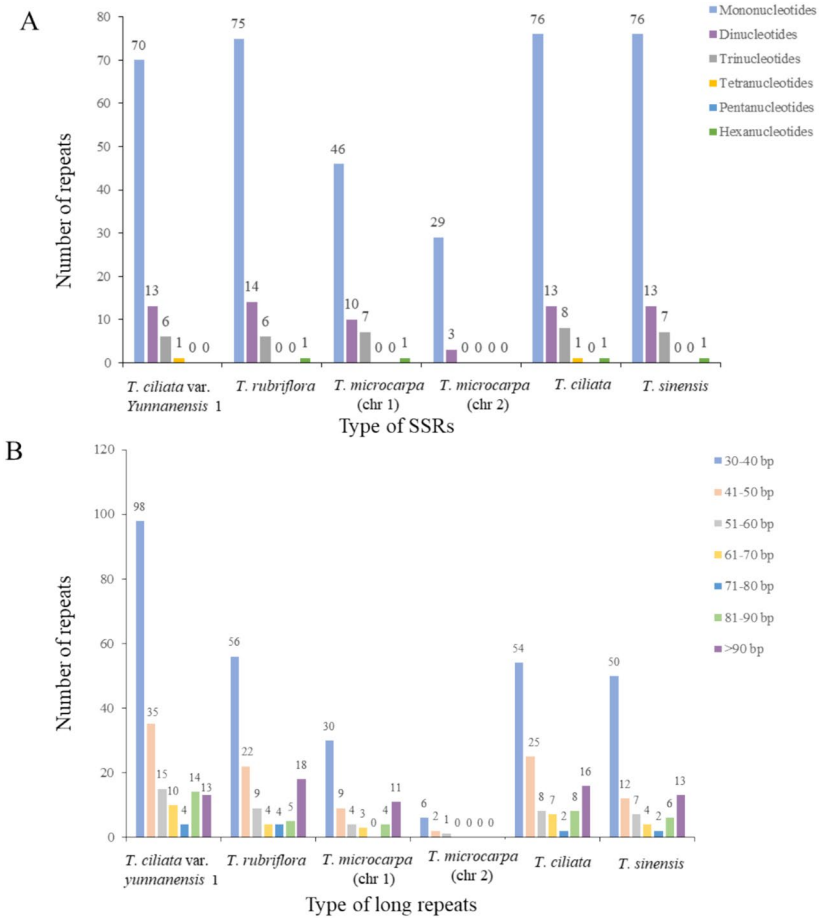

Fig. 4. Repeat sequences in mitochondrial genomes of *Toona*. **(A)** Simple sequence repeats; **(B)** Long sequence repeats.

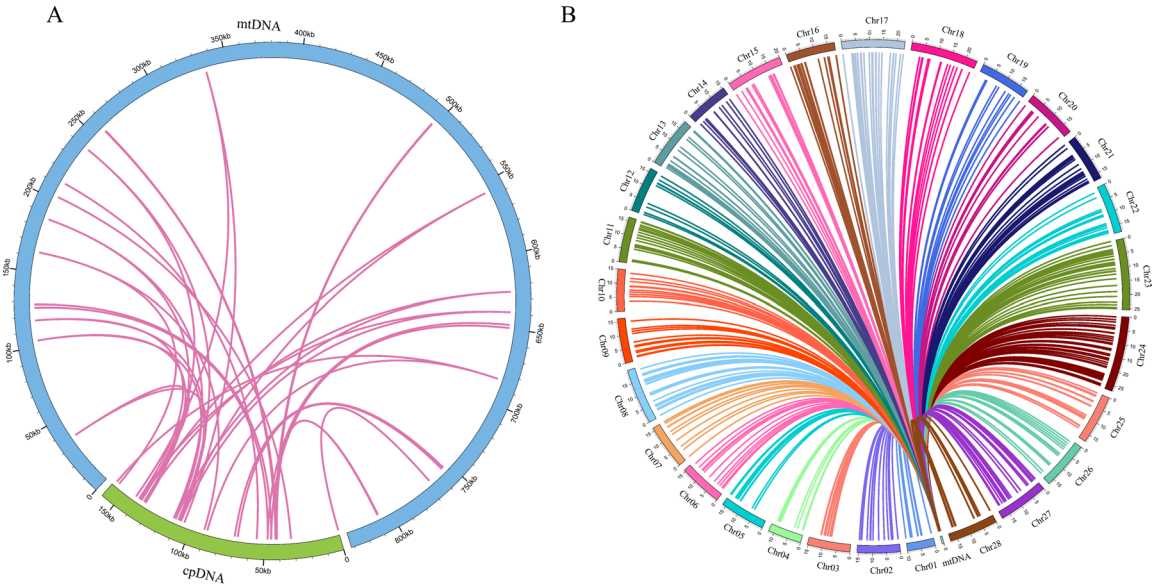

Fig. 5. Intracellular homologous comparisons. **(A)** Homology between chloroplast and mitochondrial genomes of *T. ciliata* var. *yunnanensis* 1. The mitochondrial genome is showed in blue and the chloroplast genome in green. The pink arcs represent homologous sequences; **(B)** Homology between the mitochondrial genome of *T. ciliata* var. *yunnanensis* 1 and the nuclear genome of *T. ciliata*. The homologous sequences were indicated in arcs in different colors.

of *T. ciliata* (Fig. 5B). A similar case was also identified in *T. ciliata* (Fig. S3).The most abundant homologous sequences occurred in the 24th nuclear chromosome, with 86 homologous fragments in *T. ciliata* var. *yunnanensis* 1 and 63 in *T. ciliata*, followed by the 16th nuclear chromosome (66 in *T. ciliata* var. *yunnanensis* 1 and 56 in *T. ciliata*), and the 23rd nuclear chromosome (56 in *T. ciliata* var. *yunnanensis* 1 and 43 in *T. ciliata*). There were additional 16 homologous fragments of more than 1,000 bp each in *T. ciliata* var. *yunnanensis* 1. The longest homologous fragment was 4,554 bp. The total length of homologous sequences was 195,805 bp in *T. ciliata* var. *yunnanensis* 1 (~ 23.5% of its mitochondrial genome) and 166,045 bp in *T. ciliata* (~ 24.3%).

Phylogeny and synteny analyses among taxa

Based on 35 PCG sequences, we constructed a phylogeny among species of *Toona* (Fig. 6A) where *Melia azedarach* was used as the outgroup. We estimated the species divergent times and 95% highest posterior density (HPD) based on two fossil calibration points using BEAST. The divergence time was 56.54 (39–74) Mya between *Toona* and *M. azedarach*. Within *Toona*, the minimum and maximum divergence times were approximately 4.42 Mya (0.04–14.41 Mya) and 21.99 Mya (5–39 Mya), respectively. *T. ciliata* (NC_065060.1) and *T. ciliata* var. *yunnanensis* 1 were clustered on one branch with a 100% supporting rate, consistent with the morphological classification. The Bayesian posterior probabilities for all clades were more than 99%, indicating a robust phylogenetic relationship constructed with mitochondrial PCG sequences. To elucidate the genetic similarity and differences among varieties of *T. ciliata*, we also conducted a phylogenetic relationship based on intergenic spacer (IGS) sequences. Figure 6B shows that *T. sinensis*, *T. rubriflora*, and *T. microcarpa* were distinctly separated into different branches, whereas the four varieties of *T. ciliata* were well mixed.

The collinearity analysis with mitochondrial genomes supported the phylogenetic relationships derived from PCG sequences among species of *Toona* (Fig. 7). The homologous collinearity was the highest between *T. ciliata* and *T. microcarpa* but generally low between *M. azedarach* and any species of *Toona*. These colinear fragments were short in size among species. One key feature was that the colinear fragments were arranged in quite different orders among species, indicating a genome-wide rearrangement occurred among them. The structure of mitochondrial genomes was not conserved among species.

Test of positive and negative selection

Selection underlying 35 PCGs was detected in terms of non-synonymous (dN) and synonymous (dS) substitution rates based on the phylogeny of *Toona* (Fig. 6A). Likelihood ratio tests (LRTs) with the branching model indicated that the dN/dS ratio (ω) was essentially the same across species (p-values > 5%) (Supplementary Table S8). When each PCG was separately analyzed, *sdh*4 had the highest dS value whereas *ccm*B had the lowest dS value (Fig. 8). For non-synonymous substitution rate (dN), *cox*2 had the highest dN values, while *atp*8, *atp*9, and *rps*12 had the lowest dN values. Most PCGs had the ratio (ω) of less than 1, indicating that purifying selection occurred throughout their evolutionary history. Four genes (*atp*4, *ccm*FN, *rps*1, and *rps*10) had ω values of greater than 1 and were under positive selection during their evolution. Additionally, the posterior probability for each amino acid across various site types was calculated using the BEB (Bayes Empirical Bayes) method, and eleven sites were identified to be under positive selection (Supplementary Table S8).

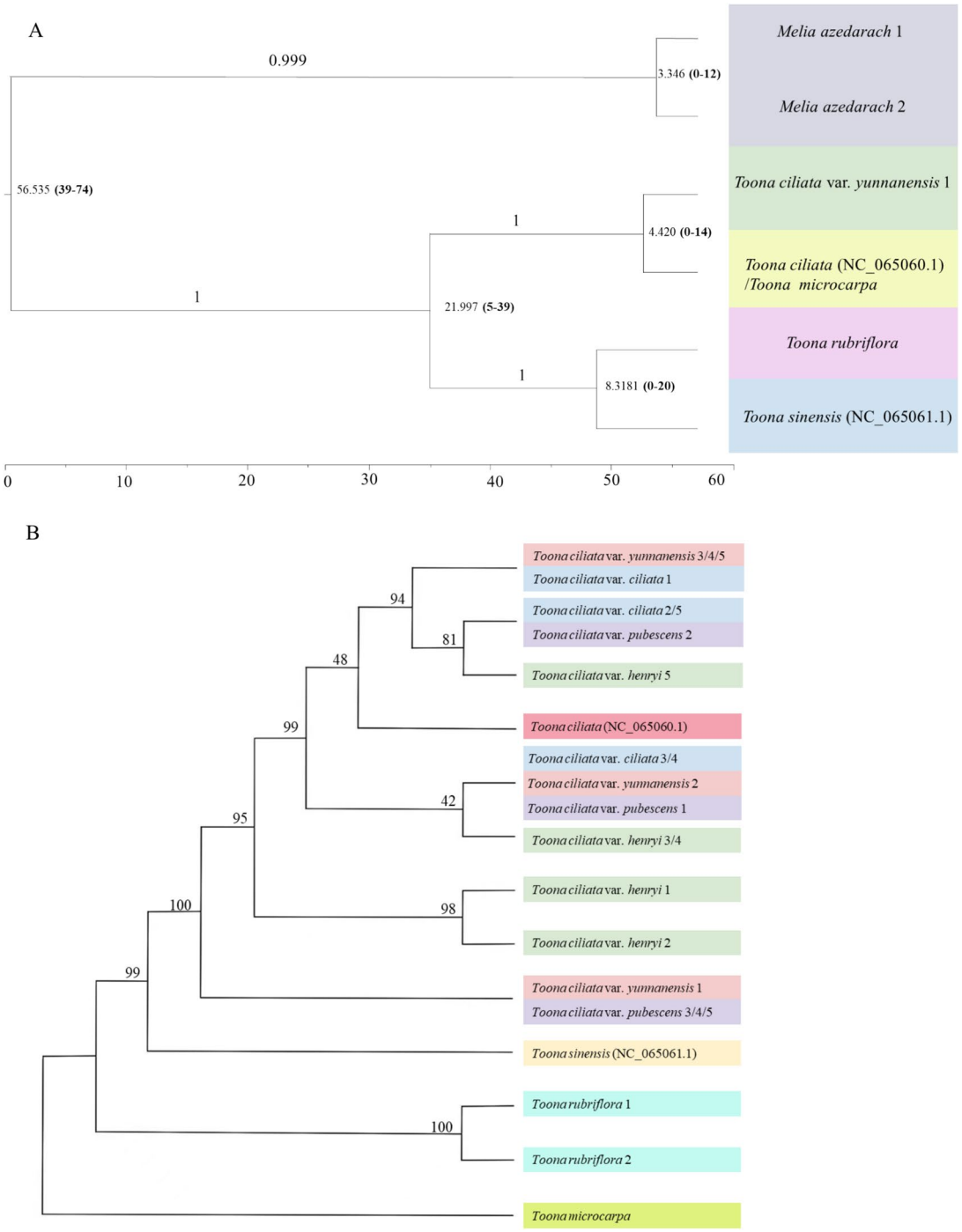

Fig. 6. The phylogenetic relationship among species of *Toona*. **(A)** The phylogenetic relationship derived from protein-coding genes. *M. azedarach* was used as the outgroup. The point estimates and 95% highest posterior densities (HPDs) were indicated for divergent times; **(B)** The phylogenetic relationship among taxa of *Toona* based on IGS sequences. Twenty samples of *T. ciliata* were included. Maximum likelihood tree was constructed with 1000 samples by bootstrapping.

Discussion

In this study, we provided evidence that species in genus *Toona* exhibited variation in mitochondrial genome structure (a high frequency of gene rearrangement) and genome size. The sequences of functional genes were conservative, like other plant mitochondrial genomes, such as comparable percentages of GC contents in *Aglaia odorata* (NC_084341.1) and *M. azedarach* (PP099859.1/PP099860.1). However, the proportions of coding sequences in *Toona* species were lower (3.81%–4.96%), compared with other angiosperms (9.54%)[28], especially

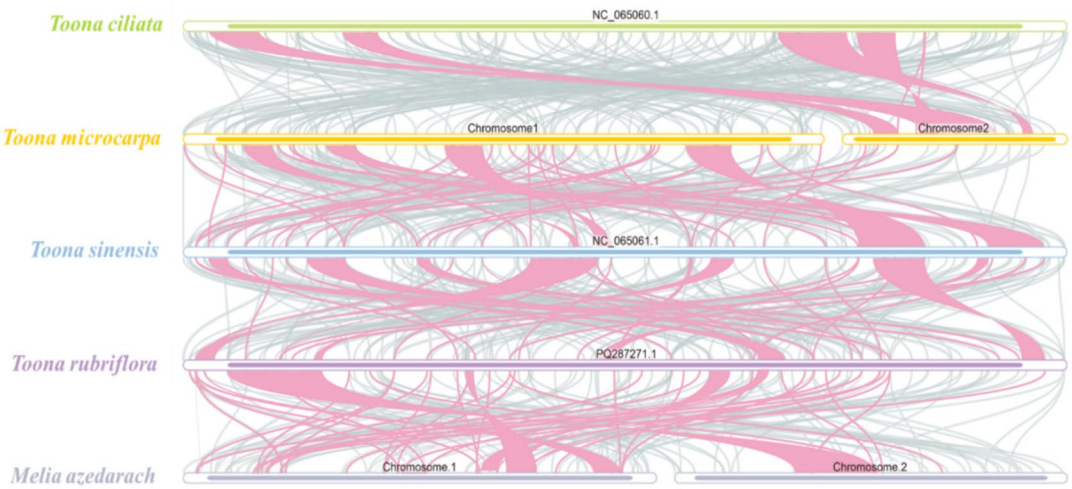

Fig. 7. Collinearity among species of *Toona* in mitochondrial genome. The bars in different colors are the mtDNA sequences of different species. The strips in grey stand for colinear homology and in pink for sequence inversions.

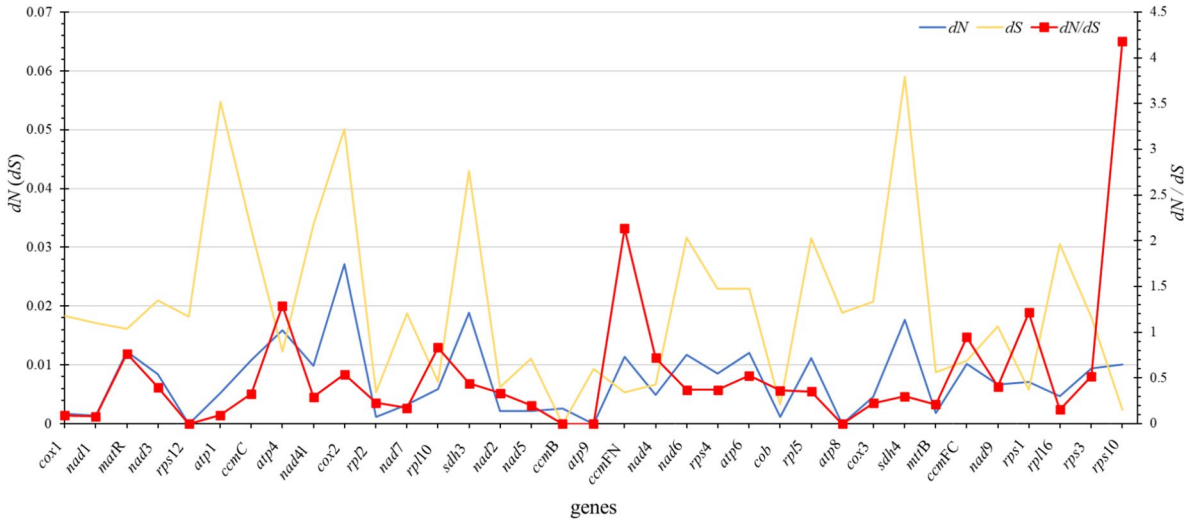

Fig. 8. The non-synonymous (dN) and synonymous (dS) substitution rates and their ratios (dN/dS) in 35 protein-coding genes in species of *Toona*.

for *T. ciliata* var. *yunnanensis* 1 with only 3.81%. This observation reflected distinct evolutionary features among the *Toona* species since their common ancestors[29].

T. microcarpa was assembled in two circulars, different from the other species of *Toona* that had a typical single circular. One circular of molecules ("master circle") was an oversimplified representation of plant mitochondrial genomes[30]. It is recognized that unlike animal mtDNA, plant mitochondrial genome structure is quite variable, such as two mtDNA circulars in *Neolamarckia cadamba* and *M. azedarach*[31,32], multi-chromosomal structures in *Dendrobium*[33], linear structure in *Lactuca sativa*[34] and branching structure in *Quercus acutissima*[35]. This phenomenon reflects the dynamic property of plant mitochondrial genomes that often undergo intracellular recombination and gene rearrangement. The presence of multiple repeat sequences provides the basis for forming ring-like structure and producing alternative conformations. For example, two pairs of repeat sequences in the mitochondrial genome of *M. azedarach* generate multiple conformations[32]. Chromosome 1 of *T. macrocarpa* (474,320 bp) had one long repeat sequence (> 1000 bp) and could likely have similar properties to a master circle. It was enriched with genes encoding complex III, complex IV, and ribosomal proteins, while chromosome 2 (166,958 bp) carried partial genes of complex I, complex V, and genes related to cytochrome c maturation. This gene distribution pattern could likely facilitate expression of some specific genes and affect the metabolic and physiological processes of *T. microcarpa*. The two-chromosome structure of *T. microcarpa* could enhance its adaptability and diversity in response to various environment changes and diverse habitats. This needs further verification in future study.

Besides the genome-wide rearrangement, a few other structural variations were observed in mitochondrial genomes of species of *Toona*. It is recognized that mitochondrial genes in angiosperms have the typical start codon ATG/GTG and the stop codon TAA/TAG/TGA[36,37]. PCGs of the mitochondrial genomes were quite variable among species of *Toona* and exhibited different start codons. This could be related to frequent RNA editing observed in plant mitochondrial genomes[38,39]. The relative usage of the synonymous codons was similar across *Toona* species. Amino acids Leu, Ser, and Ile were the most prevalent in *Toona*, which was consistent with patterns observed in many other plant species[31,40]. In *Toona*, the highest proportion of SSRs in the mitochondrial genome was the mononucleotides repeat. The high frequency of A/T in SSRs contributed to the overall abundance of AT observed in *Toona*, reflecting a common trend among land plants and suggesting a potentially parallel or convergent evolutionary trajectory in plant organelles[41].

RNA editing is an important process to alter gene expression and primary protein structure and could play a significant role in plant adaptation[42]. All species of *Toona* generally exhibited comparable editing sites, implying parallel evolution in their mitochondrial genomes. In *Toona*, a strong preference of RNA editing sites (the conversion from TCA to TTA) was identified in mitochondrial genome. Such editing patterns were also reported in other plant species, such as *Diospyros oleifera*[43]. Evidence was also recorded in the literature to show the editing of start and stop codons. For instance, genes *rps*10, *rps*4, *nad*1, and *nad* 4 l exhibited ACG as the start codons, while CGA in *rps*10 was edited as the stop codon[38,39]. Most editing sites within encoded genes were located at the second position of the triplet codons. The biological significance of such RNA editing remains to be explored in genus *Toona*.

In angiosperms, the abundance of repeat sequences influences mitochondrial genome size[44]. This was recorded in the literature, showing that genomic size differences between closely related species were attributed to large sequence duplications[45] or large insertions of exogenous DNA[46,47]. For example, *Cucumis melo* (2,740 kb) had larger mitochondrial genome size than *Cucumis sativus* (1,685 kb) due to abundant repeats sequences and a large proportion of DNA of nuclear origin[48,49]. Compared with *T. ciliata* (NC_065060.1), *T. ciliata* var. *yunnanensis* 1 had more repeat sequences and showed a tendency of genome expansion. The intracellular homologous analysis implicates that more horizontal transfers between mitochondrial and chloroplast genomes or between mitochondrial and nuclear genome are likely responsible for genome expansion in *T. ciliata* var. *yunnanensis* 1.

The divergent time among species of the genus *Toona* spanned from 4.42 to 21.99 Mya. This wide range could arise from the calibration point selection or the heterogeneity in mitochondrial genome mutation among sites. In this study, two fossil calibration points (the divergence times of *M. azedarach*-*Toona* and *T. sinensis*-*T. ciliata*) were used for molecular clock analysis. However, fossil record incompleteness or interpretive uncertainties likely contributed to the broad divergence time estimates. *T. ciliata*, a Class—II protected endangered species in China, has faced population bottlenecks or fragmentation due to over-exploitation and low natural regeneration. This would subsequently reduce its effective population size and the coalescent branch lengths and hence affect the estimation of divergent times. Moreover, the mitochondrial genomes of *Toona* had the highly rearranged structure but conservative coding sequences, which further compromised the accuracy of divergent time estimates.

Although mitochondrial genomes generally have lower mutation rates than chloroplast genomes in plant species[50], our results provided evidence that the PCG sequences were effective to distinguish species of *Toona*. This was also demonstrated in *N. cadamba*[31]. Compared with the results derived from cpDNA sequences, three species (*T. ciliata*, *T. sinensis* and *M. azedarach*) were separated with mitochondrial genomes but the varieties of *T. ciliata* were not. The estimated divergent time between *Toona* and *M. azedarach* (56.535Mya) was longer in terms of mtDNA sequences than in terms of cpDNA sequences (48.02Mya)[19]. This was due to the lower mutation rate of plant mitochondrial genome. Besides, the high frequency of rearrangement could be used to design molecular markers (e.g., RFLP) for analyzing genetic variation, such as in *Ficus carica* and *Symplocos laurina*[51,52]. Integration of mitochondrial, chloroplast, and nuclear genomes can improve resolution to phylogeny within genus *Toona*. The chloroplast and nuclear genomic data are expected to more effective for clarifying species-level relationships due to the higher mutation rate. The mitochondrial sequences are suitable for insights into deep evolutionary divergences and ancestral lineage dynamics[53]. In the Orchidaceae family, the phylogenetic positions of subgroups such as the subtribe Epipogiinae, tribe Gastrodieae, and subfamily Tropidieae were poorly supported or excluded from analyses. However, these issues were effectively resolved through mitochondrial data analysis[54]. The present study confirmed the appropriateness of mitochondrial genomes for elucidating species of genus *Toona*.

Like the previous studies on *T. ciliata* or other species using chloroplast genomes[20,55], a conflict phylogenetic relationship among varieties of *T. ciliata* was observed in terms of mitochondrial genome versus morphological traits. Both PCGs and IGS sequences could not distinguish varieties of *T. ciliata*, which also provided a biological basis to term the *T. ciliata* complex[20]. Since leaf and flower traits are predominantly controlled by nuclear genomes, such a conflict essentially reflects the cytonuclear conflict in taxonomic division, a common phenomenon in many plant phylogenies[56,57]. Besides the influences of the lower mutation rate, a few other processes could be involved in forming this conflict. One is the distinct genetic drift rates between nuclear and mitochondrial genomes[58]. In angiosperms, the genetic drift rate of the biparentally inherited nuclear genomes is a half of that of the maternally inherited mitochondrial genomes when male and female parents are equally abundant. The coalescent processes lead to reciprocal monophyly more swiftly in the mitochondrial genome compared with the nuclear genome[59]. Our phylogenetic relationship based IGS sequences probably reflected this situation (Fig. 6), when the divergent times are not long enough to reach the monophyletic phase. The difference between Fig. 6 of the present study and Fig. 4 of Xiao et al. for the varieties of *T. ciliata* mainly arose from lower mutation rates of mitochondrial genomes[19].

On the other hand, lineage sorting driven solely by genetic drift is slow, allowing for ancestral variations are shared across thousands to millions of generations in descendant lineages after speciation. Lineage sorting is still incomplete among varieties of *T. ciliata*. Like some *Quercus* species, the mitochondrial gene tree is discordant with the classification based on leaf shape and fruit morphology due to incomplete lineage sorting and ancient hybridization[60]. Similar cases of cytonuclear discordance in phylogeny were reported in the literature[55]. Given recent species divergence of *Toona*, lineage sorting among varieties of *T. ciliata* remains in early stage. The earlier few polymorphic sites in mtDNA sequences inherited from the common ancestors remained unaltered in IGS regions (more likely neutral sites) in varieties of *T. ciliata*. Consequently, only a few mutations were accumulated since speciation from common ancestral populations. This could be expected in closely related lineages that are not reciprocally monophyletic[61,62].

The second process is natural selection that impedes genomic divergence among taxa[63]. Almost all PCGs were under negative selection in mitochondrial genomes in *Toona*, like the case of chloroplast genomes[19]. There were no sequence variations in PCGs across varieties of *Toona*. However, adaptive selection could likely involve genetic divergence at genes associated with flower and leaf traits or other morphological traits, yielding morphological differentiation in *T. ciliata*. Alternative alleles were likely fixed due to selection in different varieties. This eventually leads to a conflict in phylogeny in terms of morphological traits versus the PCG sequences of mitochondrial genome, analogous to the findings in plant species of genus *Crataegus*[64,65] and animal species *Oophaga pumilio*[66].

One worthy point is that four PCGs (*atp*4, *ccm*FN, *rps*1, and *rps*10) were under positive selection ($dN/dS > 1$) at the species level. These genes participate in vital mitochondrial functions, including energy metabolism, the electron transport chain, and protein synthesis. Their adaptive evolution enhanced species divergence. Similar patterns of positive selection were observed in other plants. For instance, in the mitochondrial genome of the cave-endemic plant *Primulina hunanensis*[41], positive selection signals were also detected in the *atp*4 and *ccm*FN genes, which are associated with energy metabolism adaptation under low-light conditions. Species of *Toona* are distributed across different climatic zones in China. Positive selection could optimize ATP synthesis efficiency for adaptation to different temperatures and energy demands in different habitats. Nevertheless, these functional genes were under purifying selection in varieties of *T. ciliata*, impeding evolutionary divergence between varieties.

The third process is gene introgression among varieties of *T. ciliata* that could not seriously affect phylogeny based on mitochondrial genomes. Gene flow through pollination does not disperse maternally inherited genes. Only seed flow contributes to mitochondrial gene flow. However, gene flow through natural hybridization could impede nuclear genomic divergence among varieties of *T. ciliata*[67]. These varieties are sympatric or partially overlap in geographical distribution in China and natural hybridization could not be excluded among them. This process could effectively affect genomic divergence at neutral sites. Differential selection at adaptive sites, especially for the sites associated with speciation, would not be seriously affected in the presence of gene flow, except for strong gene flow that can swamp locally adaptive genes[68].

In general, this study provided comprehensive comparisons on mitochondrial genome structure of *Toona* and offered insights for identification and genetic research on the species of *Toona*. Owing to multiple properties of mitochondrial genomes, including uniparental inheritance, haploid, low mutation rates, a high frequency of recombination and gene rearrangement, these limit the power of elucidating species and recently diverging lineages. Natural hybridization could also reduce the power of identifying species and purifying selection impedes species divergence. This study confirmed the appropriateness for identifying species of *Toona* and the limitation for identifying varieties of *T. ciliata*. To address the unresolved phylogeny among varieties of *T. ciliata*, highly polymorphic markers or nuclear genome sequences are needed. Besides, future research could be expanded in following ways: (1) Improving sampling strategies by increasing the geographical coverage and number of populations, especially in the typical distribution areas of varieties, would help better understand genetic variation patterns; (2) Integrating quantitative analyses with transcriptomic data could provide new perspectives for clarifying taxonomic relationships in genus *Toona*, especially in the *T. ciliata* complex. These studies could further contribute to our understanding of the phylogeny and genetic diversity of genus *Toona*.

Methods

Plant material, DNA extraction and sequencing

Leaf samples of *T. rubriflora* and *T. microcarpa* were collected from healthy living plants in Fujian and Guangdong provinces, respectively (Table 3). Samples were then dried and stored in silica gel. From previous studies and literature of the varieties of *T. ciliata*[67], we collected five samples of each variety (*T. ciliata* var. *ciliata, T. ciliata* var. *yunnanensis, T. ciliata* var. *henryi, T. ciliata* var. *pubescens*) for resequencing. Note that *T. ciliata* var. *subaxaxflora* was not found in natural forests and hence was not included in this study. In addition, two samples of *M. azedarach* were collected from South China Agricultural University (SCAU) as the outgroup. The specimen was stored for records in Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, SCAU, Guangdong Province, China. The use of plant leaves in this study complies with institutional guidelines. Collection of the plant specimen was permitted by the University.

The total genomic DNA of each sample was extracted by CTAB[69], and the genomic DNA was used for library construction. For next-generation sequencing, the library was constructed with an insert length of 350 bp. It was sequenced on the Illumina Novaseq6000 platform (Illumina, USA). After library construction, the qPCR method and Agilent 2100 Bioanalyzer (Agilent Technologies, USA) were used for quality control. The qualified DNA libraries were sequenced using the Illumina Novaseq6000 (Illumina, USA) high-throughput sequencing platform with the PE150 (Pair-End 150) sequencing strategy. Raw sequencing data were filtered with fastp v.000.23.2[70] (-q 20 -l 150) to remove low-quality reads and adapters. The average Q20 and Q30 percentages were 98% and 94%, respectively, ensuring high-quality data for downstream analyses (Supplementary Table S9).

Species sample	Longitude (E°)	Latitude (N°)	Altitude (m)
Toona ciliata var. *ciliata* 1	100.65	26.41	1482.15
Toona ciliata var. *ciliata* 2	100.78	24.54	1181.68
Toona ciliata var. *ciliata* 3	100.71	24.66	1265.88
Toona ciliata var. *ciliata* 4	100.86	24.36	1141.4
Toona ciliata var. *ciliata* 5	100.90	24.35	1092.28
Toona ciliata var. *yunnanensis* 1	101.37	26.58	1047.75
Toona ciliata var. *yunnanensis* 2	102.65	25.62	1419.92
Toona ciliata var. *yunnanensis* 3	100.65	26.4	1490.97
Toona ciliata var. *yunnanensis* 4	100.65	26.41	1488.64
Toona ciliata var. *yunnanensis* 5	100.63	26.37	1438.94
Toona ciliata var. *pubescens* 1	102.65	25.65	1504.78
Toona ciliata var. *pubescens* 2	102.66	25.72	1448.33
Toona ciliata var. *pubescens* 3	101.22	26.33	1047.75
Toona ciliata var. *pubescens* 4	100.99	26.51	1361.8
Toona ciliata var. *pubescens* 5	101.21	26.49	1571.29
Toona ciliata var. *henryi* 1	101.04	24.13	1079.98
Toona ciliata var. *henryi* 2	100.98	24.17	1081.71
Toona ciliata var. *henryi* 3	100.80	24.54	1250.22
Toona ciliata var. *henryi* 4	100.80	24.53	1281.97
Toona ciliata var. *henryi* 5	100.98	24.19	1085.08
Toona rubriflora 1	117.51	26.48	296.24
Toona rubriflora 2	117.49	26.41	295.33
Toona microcarpa	113.36	23.19	46.32

Table 3. Localities of species samples of *Toona* and individual samples of varieties of *Toona ciliata*.

For the third-generation sequencing, the libraries of 10 kb fragments were constructed using the PacBio RS II platform (Pacific Biosciences, PN 101–853-100). and sequenced on the PacBio Revio system. High-quality HiFi data were obtained by extracting reads with the quality value greater than Q20 from the CCS data.

Genome assembly and annotation

The next-generation sequencing (NGS) data were aligned to the reference mitochondrial genome using Bowtie2 (v.2.4.2)[71]. The alignment results were sorted with Samtools (v.1.8)[72] to generate aligned BAM files. The assembled contig sequences were screened against related species' mitochondrial and chloroplast genomes using BLASTn (E-value ≤1e-10) to identify putative mitochondrial sequences. To enhance accuracy, only paired-end alignments were retained, and the final mitochondrial genome sequence was refined using Pilon (v.1.23)[73]. The third-generation sequencing data were assembled using Canu (v.2.2)[74] to obtain the genome sequence. The assembled sequences were compared to the NCBI nt database using Blastn (v.2.11.0 +)[75] to extract mitochondrial genome sequences. Non-cyclic tig sequences were manually spliced into a circular arrangement based on overlaps and paired-end relationships from second-generation sequencing data. Following assembly, NGS and third-generation sequencing were remapped to the mitochondrial genome assembly to validate coverage uniformity and confirm structural accuracy.

The mitochondrial genome was initially annotated using GeSeq-web[76]. Protein sequences obtained from the annotation were compared to the NCBI nr database using Blastp (v.2.11.0 +) to adjust the positions of coding genes. The rRNA annotations were validated and corrected by comparing to the NCBI core nt database using Blastn (v.2.11.0 +). The tRNA secondary structures were generated using ARWEN (v.1.2)[77], and any irregular tRNAs were further validated with tRNAscan-SE (v.2.0)[78]. The tRNAs with unreasonable lengths or incomplete structures were excluded from final annotation.

Codon usage and RNA editing analysis

The RSCU (relative synonymous codon usage) values of coding genes were calculated using MEGA 7.0.26[79]. No preference for an amino acid was indicated when RSCU was 1, while a preference for an amino acid was indicated when RSCU was greater than 1. Condon usage was represented through ggplot2 in R package. RNA editing sites were predicted for PCGs with online tool PREP suite (http://prep.unl.edu/).

Repeat and homologous sequence analysis

Simple sequence repeats (SSRs) were identified using MISA-web. Long repetitive sequences exceeding 30 bp were analyzed using REPuter[80]. The chloroplast genomes of *T. sinensis* (MF467522.1), *T. microcarpa* (MF467521.1), *T. ciliata* (OM135324.1), *T. ciliata* var. *yunnanensis* (OM135326.1) were downloaded from the NCBI database, respectively. The nuclear genome of *T. ciliata* was downloaded from the CNGB Nucleotide Sequence Archive (https://db.cngb.org/search/project/CNP0001985/) with accession number no. CNP0001985. The homologous

sequences were analyzed using BLASTn with parameters (- e value 1e—5 -outfmt 6 -out output_file) and visualized using the Circos software package (version 0.67)[81].

Positive and negative selection test

The non-synonymous (dN) and synonymous (dS) substitution rates and their ratios (dN/dS) were estimated for 35 PCGs in mitochondrial genome of *Toona* species using Codeml in the PAML package[82,83]. Branch models were used to estimate *dN/dS* for specific branches, and site models were applied to detect selection on PCGs. Likelihood ratio tests (LRTs) were conducted to compare alternative site models: M0(one ratio) vs. M1(neutral), M1(neutral) vs. M2(selection), and M7 (beta) vs. M8(beta and selection)[83]. Positively selective sites were identified using Bayesian posterior probabilities (BEB) with M2 and M8 models, with the probability greater than 0.95 indicating positive selection[84].

Phylogeny and synteny analysis

To construct phylogenetic tree of the *Toona* genus, we aligned common PCGs using MAFFT (v7.490)[85] and concatenated them into a "supergene" sequence set. The optimal substitution model (K3Pu + F + I) was determined using modelfinder in IQ-TREE (v1.5.6)[86], and a maximum likelihood tree was constructed with IQ-TREE v1.5.6 (1000 bootstrap replicates). Additionally, intergenic spacer (IGS) sequences were aligned using MAFFT, concatenated, and used to build another maximum likelihood tree with IQ-TREE. Bayesian inference was performed with BEAST (v2.7.7)[87] under the HKY model, using calibration times from Timetree (http://www.timetree.org/ accesss March 12, 2024) for divergence estimates. Four MCMC runs were executed for 20,000,000 iterations, with the ESS values greater than 200. For collinearity analysis, the mitochondrial genomes of *Toona* species were subjected to sequence alignment using BLASTn(- e value 1e—5 -outfmt 6 -out output_file). Individual colinear regions were identified using MCScanX[88] and visualized in Tbtools[89].

Conclusions

We sequenced the mitochondrial genomes of *T. rubriflora* and *T. microcarpa*, and re-sequenced four varieties of *T. ciliata*. A comprehensive analysis of *Toona* was conducted based on the mitochondrial genome, including codon preferences, repeat sequences, RNA editing events, genome homology and phylogenetic relationships. Comparative genomic analysis indicated the presence of substantial structural variation among species due to gene rearrangement. There were small variations in gene number, RNA editing sites, repeat sequences and genome homology. All species have comparable GC contents, conservative sequences of PCGs, codon occurrences, and RNA editing sites. There was no sequence divergence in PCGs among varieties of *T. ciliata* except for a few mutations in IGS regions. Phylogenetic analysis indicated that species of *Toona* were effectively identified in terms of PCG sequences. However, the varieties of *T. ciliata* were not well separated by both PCGs and IGS sequences, supporting the conservation strategy that considered the varieties of *T. ciliata* as a single conservation unit, termed as the *T. ciliata* complex. Most PCGs were predominantly under purifying selection across species. Species of *Toona* had about 4% mitochondrial genomes homologous with chloroplast genomes and about 24% homologous with nuclear genomes. Conflict phylogenies existed among varieties of *T. ciliata* in terms of mitochondrial sequences versus morphological (leaf and flower) traits. The overall results provided comprehensive information on *Toona* mitochondrial genomes, which could be used for species identification in genus *Toona* and conservation strategy for the *T. ciliata* complex.

Data availability

The datasets generated by this study are available in nucleotide database of GenBank (https://www.ncbi.nlm.nih.gov/nucleotide/) repository with accession numbers: Toona rubriflora (PQ287271.1) and Toona microcarpa (Chromosome 1: PQ318328.1, Chromosome 2: PQ318327.1). The re-sequencing data for the varieties of Toona ciliata were provided in the Supplementary Material.

Received: 28 October 2024; Accepted: 11 June 2025
Published online: 25 June 2025

References

1. Chen, S. K., Li, H. & Chen, P. Y. *Flora* of *China* Vol. 43 (Beijing Science Press, 1997) (**in Chinese**).
2. Luo, X. X., Xiao, Y. Y., He, P., Shi, Z. G. & Zhao, P. Comparisons of nutritional compositions in the young leaves of *Toona microcarpa* under two cultivation conditions. *J. Engr.* **28**, 66–68 (2014).
3. Wang, X. *et al.* A chromosome-level genome assembly of *Toona ciliata* (Meliaceae). *Genome Biol. Evol.* **14**, evac121 (2022).
4. Nie, P. et al. Trace identification and visualization of multiple benzimidazole pesticide residues on *Toona sinensis* leaves using terahertz imaging combined with deep learning. *Int. J. Mol. Sci.* **22**, 3425 (2021).
5. Pan, J. *et al.* Local patterns of arbuscular mycorrhizal fungal diversity and community structure in a natural *Toona ciliata* var. *pubescens* forest in South Central China. *Peer J.* **9**, e11331 (2021).
6. Hu, W., Liu, Y. R., Liao, M. Q., Gao, R. L. & Zhang, D. H. The research progress of the multipurpose *Toona sinensis*. *J. Fujian Sci. Technol.* **35**, 244–250 (2008).
7. Deng, Y., Zhang, C. H. & Qiu, J. Wood anatomical structure analysis of *Toona ciliata* and *Toona ciliata* var. *henryi*. *Guangxi Forestry Sci.* **46**, 89–92 (2017).
8. Li, P. et al. Development of an efficient regeneration system for the precious and fast-growing timber tree *Toona ciliata*. *Plant Biotechnol (Tokyo).* **35**, 51–58 (2018).
9. Xin, F. et al. Chemical constituents of *Toona microcarpa* (C. DC.) harms in Engl. (Meliaceae). *Biochem. Syst. Ecol.* **38**, 128–130 (2010).
10. Pramono, A. A., Palupi, E. R., Siregar, I. Z. & Kusmana, C. Characteristics of surian flower, fruit and seed productions (*Toona sinensis* (A. Juss.) M. Roem.) in Sumedang. *West Java. Trop. Life Sci. Res.* **27**, 77–91 (2016).

11. Liu, W. J. et al. Non-destructive measurements of *Toona sinensis* chlorophyll and nitrogen content under drought stress using near infrared spectroscopy. *Front. Plant Sci.* **12**, 809828 (2022).
12. Yang, Z. et al. Chemical composition and antibacterial activity of 12 medicinal plant ethyl acetate extracts using LC-MS feature-based molecular networking. *Phytochem. Anal.* **33**, 473–489 (2022).
13. Fu, L. G. & Chin, C. M. *China Plant Red Data Book-Rare and Endangered Plants* Vol. 1 (Beijing Science Press, 1992) (**in Chinese**).
14. Liang, R. L., Liao, R. Y. & Dai, J. Analysis of endangered causes and protection strategies of *Toona ciliata*. *Guangxi Sci.* **40**, 201–203 (2011).
15. Huang, H. L., Zhang, L. & Liao, C. K. Seed rain, soil seed bank, and natural regeneration of natural *Toona ciliata* var. *pubescens* forest. *Chinese J. Appl. Ecol.* **23**, 972–978 (2012).
16. Zou, C., Xu, W. & Bu, J. Influence of global warming on vegetation in Northeast China. *Chinese Geogr. Sci.* **7**, 68–78 (1997).
17. Ji, Y. T. et al. Long read sequencing of Toona sinensis (A. Juss) Roem: A chromosome-level reference genome for the family Meliaceae. *Mol. Ecol. Resour.* **21**, 1243–1255 (2021).
18. Wang, X. et al. Evolutionary divergence between *Toona ciliata* and *Toona sinensis* assayed with their whole genome sequences. *Genes* **13**, 1799 (2022).
19. Xiao, Y. et al. Assessing the phylogenetic relationship among varieties of *Toona ciliata* (Meliaceae) in sympatry with chloroplast genomes. *Ecol. Evol.* **13**, e10828 (2023).
20. Li, P. et al. Genetic diversity and population structure of *Toona ciliata* Roem. based on sequence-related amplified polymorphism (SRAP) markers. *Forests* **6**, 1094–1106 (2015).
21. Zhan, X. et al. Genetic diversity and population structure of *Toona ciliata* revealed by simple sequence repeat markers. *Biotechnol. Biotech. Eq.* **33**, 214–222 (2019).
22. Lin, N. et al. Complete plastome sequencing from *Toona* (Meliaceae) and phylogenomic analyses within Sapindales. *Appl. Plant Sci.* **6**, e1040 (2018).
23. Wang, Z. Y. et al. Population Structure and Genetic Diversity of the *Toona ciliata* (Meliaceae) Complex Assayed with Chloroplast DNA Markers. *Genes* **15**, 320 (2024).
24. Li, Y. L. et al. Sequencing and analysis of the complete mitochondrial genomes of *Toona sinensis* and *Toona ciliata* reveal evolutionary features of *Toona*. *BMC Genomics* **24**, 58 (2023).
25. Sloan, D. B. et al. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* **10**(1), e1001241 (2012).
26. Qiu, Y. L. et al. Angiosperm phylogeny inferred from sequences of four mitochondrial genes. *J. Syst. Evol.* **48**, 391–325 (2010).
27. Richardson, A. O., Rice, D. W., Young, G. J., Alverson, A. J. & Palmer, J. D. The, "fossilized" mitochondrial genome of *Liriodendron tulipifera*: ancestral gene content and order, ancestral editing sites, and extraordinarily low mutation rate. *BMC Biol.* **11**, 29 (2013).
28. Zhu, Q. & Ge, S. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytol.* **167**, 249–265 (2005).
29. Gao, H. & Kong, J. Distribution characteristics and biological function of tandem repeat sequences in the genomes of different organisms. *Zool Res.* **26**, 555–564 (2005).
30. Sloan, D. B. One ring to rule them all? Genome sequencing provides new insights into the 'master circle' model of plant mitochondrial DNA structure. *New Phytol.* **200**, 978–985 (2013).
31. Wang, X. et al. A complete sequence of mitochondrial genome of *Neolamarckia cadamba* and its use for systematic analysis. *Sci Rep.* **11**, 21452 (2021).
32. Hao, Z. G. et al. Complete mitochondrial genome of Melia azedarach L., reveals two conformations generated by the repeat sequence mediated recombination. *BMC Plant Biol.* **24**, 645 (2024).
33. Wang, M. T. et al. Rapid structural evolution of *Dendrobium* mitogenomes and mito-nuclear phylogeny discordances in *Dendrobium* (Orchidaceae). *J. Syst. Evol.* **61**, 790–805 (2023).
34. Kozik, A. et al. The alternative reality of plant mitochondrial DNA: one ring does not rule them all. *PLoS Genet.* **15**, e1008373 (2019).
35. Liu, D. et al. Complex physical structure of complete mitochondrial genome of *Quercus acutissima* (Fagaceae): a significant energy plant. *Genes* **13**, 1321 (2022).
36. Powell, W., Machray, G. C. & Provan, J. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1**, 215–222 (1996).
37. Paço, A., Freitas, R. & Vieira-da-Silva, A. Conversion of DNA sequences: from a transposable element to a tandem repeat or to a gene. *Genes (Basel).* **10**, 1014 (2019).
38. Knoop, V. The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.* **46**, 123–139 (2004).
39. Ichinose, M. & Sugita, M. RNA editing and its molecular mechanism in plant organelles. *Genes (Basel).* **8**, 5 (2016).
40. Li, Q. et al. Characterization and phylogenetic analysis of the complete mitochondrial genome of the medicinal fungus *Laetiporus sulphureus*. *Sci. Rep.* **8**, 9104 (2018).
41. Chen, L. et al. Assembly and comparative analysis of the initial complete mitochondrial genome of *Primulina hunanensis* (Gesneriaceae): a cave-dwelling endangered plant. *BMC Genomics* **25**, 322 (2024).
42. Edera, A. A., Gandini, C. L. & Sanchez-Puerta, M. V. Towards a comprehensive picture of C-to-U RNA editing sites in angiosperm mitochondria. *Plant Mol. Biol.* **97**, 215–231 (2018).
43. Xu, Y. et al. Characterization and phylogenetic analysis of the complete mitochondrial genome sequence of *Diospyros oleifera*, the first representative from the family Ebenaceae. *Heliyon* **8**, e09870 (2022).
44. Knoop, V., Volkmar, U., Hecht, J. & Grewe, F. *Plant Mitochondria* Vol. 1 Ch. 1 (New York Springer Bress, 2011).
45. Allen, J. O. et al. Comparisons among two fertile and three male-sterile mitochondrial genomes of maize. *Genetics* **177**, 1173–1192 (2007).
46. Goremykin, V. V., Lockhart, P. J., Viola, R. & Velasco, R. The mitochondrial genome of *Malus domestica* and the import-driven hypothesis of mitochondrial genome expansion in seed plants. *Plant J.* **71**, 615–626 (2012).
47. Rice, D. W. et al. Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm. *Amborella. Science* **342**, 1468–1473 (2012).
48. Alverson, A. J., Rice, D. W., Dickinson, S., Barry, K. & Palmer, J. D. Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* **23**, 2499–2513 (2012).
49. Rodríguez-Moreno, L. et al. 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* **12**, 1–14 (2012).
50. Wolfe, K. H., Li, W. H. & Sharp, P. M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci.* **84**, 9054–9058 (1987).
51. Khadari, B., Grout, C., Santoni, S. & Kjellberg, F. Contrasted genetic diversity and differentiation among Mediterranean populations of *Ficus carica* L.: a study using mtDNA RFLP. *Genet. Resour. Crop Evol.* **52**, 97–109 (2005).
52. Banu, S., Lagu, M. D. & Gupta, V. S. Phylogeographical studies in disjunct populations of *Symplocos laurina* Wall. using cytoplasmic PCR-RFLP approach. *Tree Genet. Genomes* **6**, 13–23 (2010).
53. Rydin, C., Wikström, N. & Bremer, B. Conflicting results from mitochondrial genomic data challenge current views of Rubiaceae phylogeny. *Am. J. Bot.* **104**, 1522–1532 (2017).
54. Li, Y. X. et al. Phylogenomics of Orchidaceae based on plastid and mitochondrial genomes. *Mol. Phylogenet. Evol.* **139**, 106540 (2019).

55. Hu, Y. et al. Advancing phylogeography with chloroplast DNA markers. *Biodiversity Sci.* **27**, 219–234 (2019).
56. Dong, S. S. et al. Plastid and nuclear phylogenomic incongruences and biogeographic implications of *Magnolia* sl (Magnoliaceae). *J. Syst. Evol.* **60**, 1–15 (2022).
57. Ke, X. R. et al. Nuclear and plastid phylogenomic analyses provide insights into the reticulate evolution, species delimitation, and biogeography of the Sino-Japanese disjunctive *Diabelia* (Caprifoliaceae). *J. Syst. Evol.* **60**, 1331–1343 (2022).
58. Freeland, J. R., Kirk, H. & Petersen, S. D. *Molecular ecology* 2nd edn (Oxgord Wiley-Blackwell Press, 2011).
59. Lv, Y. W. et al. Advances in lineage sorting theories and their detection methods. *Biodiversity Sci.* **32**, 23400 (2024).
60. Hipp, A. L. et al. Genomic landscape of the global oak phylogeny. *New Phytol.* **226**, 1198–1212 (2020).
61. Palumbi, S. R., Cipriano, F. & Hare, M. P. Predicting nuclear gene coalescence from mitochondrial data: The three-times rule. *Evolution* **55**, 859–868 (2001).
62. Hudson, R. R. & Coyne, J. A. Mathematical consequences of the genealogical species concept. *Evolution* **56**, 1557–1565 (2002).
63. Bazin, E. G. & Galtier, S. Population size does not influence mitochondrial genetic diversity in animals. *Science* **312**, 570–572 (2006).
64. Zhang, X. et al. Evaluation of molecular and phenotypic diversity of *Crataegus bretschneideri* CK Schneid. and related species in China. *Genet. Resour. Crop Evol.* **70**, 221–234 (2023).
65. Zhang, X. et al. Comparative genomic and phylogenetic analyses of mitochondrial genomes of hawthorn (Crataegus spp) in Northeast China. *Int. J. Biol. Macromol.* **272**, 132795 (2024).
66. Brwon, J. L., Maan, M. E., Cummings, M. E. & Summers, K. Evidence for selection on coloration in a Panamanian poison frog: a coalescent-based approach. *J. Biogeogr.* **37**, 891–901 (2010).
67. Zhang, C. H. Study on characteristics and suitable region distributions for *Toona ciliata* and Congenerics. PhD thesis. *China Academy of Forestry Sciences* Beijing (2018).
68. Hu, X. S. Mating system and the critical migration rate for swamping selection. *Genet. Res. Camb.* **93**, 233–254 (2011).
69. Li, J., Wang, S., Jing, Y., Ling, W. & Zhou, S. A modified CTAB protocol for plant DNA extraction. *Chinese Bull. Bot.* **48**, 72 (2013).
70. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
71. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie. *Nat. Methods* **9**, 357–359 (2012).
72. Danecek, P. *et al*. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
73. Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).
74. Koren, S., Walenz, B. P., Berlin, K., Miller, J. R. & Phillippy, A. M. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
75. Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7**, 203–214 (2000).
76. Tillich, M. et al. GeSeq–versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **45**, W6–W11 (2017).
77. Laslett, D. & Canbäck, B. ARWEN, a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics* **24**, 172–175 (2008).
78. Lowe, T. M. & Chan, P. P. tRNAscan-SE on-line: search and contextual analysis of transfer RNA genes. *Nucleic. Acids. Res.* **44**, W54–W57 (2016).
79. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2018).
80. Kurtz, S. et al. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic. Acids. Res.* **29**, 4633–4642 (2001).
81. Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
82. Yang, Z. H. & Bielawski, J. B. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**, 496–503 (2000).
83. Yang, Z. H. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
84. Yang, Z. H. & Swanson, W. J. Codon-substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol. Biol. Evol.* **19**, 49–57 (2002).
85. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **20**, 1160–1166 (2019).
86. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
87. Gavryushkina, A. et al. Bayesian total-evidence dating reveals the recent crown radiation of penguins. *Syst. Biol.* **66**, 57–73 (2012).
88. Wang, Y. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
89. Chen, C. et al. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol. Plant* **16**, 1733–1742 (2023).

Acknowledgements

We appreciate Zhi-Jian Feng at South China Agricultural University, You-Pai Zeng at South China Botanical Garden for providing information of species, Dai-Quan Ye and Yong Li at Yangkou State Forest Farm of Fujian for kind help during our field sample collection.

Author contributions

L.Y.W. conducted the experiment, analyzed the data, and wrote the manuscript; L.Y.W., W.Z.Y., and H.Z.H. participated in field sampling and experiment; X.Y. and W.C. provided experimental supports; H.X.S. revised and finalized the manuscript. All authors read and approved the final manuscript.

Funding

The study was supported jointly by Guangdong Basic and Applied Basic Research Foundation (2023A1515012137), National Natural Science Foundation (32171819), and the funding from the South China Agricultural University (4400-K16013).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-06816-x.

Correspondence and requests for materials should be addressed to X.-S.H.

Reprints and permissions information is available at www.nature.com/reprints.

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