Phyton-
International Journal of
Experimental Botany

Tech Science Press

http://doi.org/10.32604/phyton.2026.079853

ARTICLE

Comparative Analysis of Chloroplast Genomes and Phylogenetic Relationships of Eight *Quercus* Species

Xin Li[#], Xiaoming Tian[*,#], Guangfeng Xiang, Lan Zhou, Gaofei Li, Lu Zhu, Chao Liu, Xiangpeng Li, Li Liu, Tianyi Jiang, Cun Mou and Hao Lv[*]

Institute of Plant Conservation, Hunan Botanical Garden, Changsha, China
[*]Corresponding Authors: Xiaoming Tian. Email: tianxiaoming1986@126.com; Hao Lv. Email: lvhao0203@163.com
[#]These authors contributed equally to this work

Received: 29 January 2026; Accepted: 06 May 2026; Published: 29 June 2026

ABSTRACT: *Quercus* is widely distributed globally and holds significant ecological and economic value. However, the morphological classification of this genus has long been controversial, with the core issue being whether the *Cyclobalanopsis* should be treated as an independent genus or as a subgenus within *Quercus*. In this study, the chloroplast genomes of eight *Quercus* species were determined and analyzed. Among these, the complete chloroplast genome of *Quercus pachyloma* is reported for the first time, alongside newly published chloroplast genome data for seven other *Quercus* species collected from Hunan and Jiangxi provinces in China. All eight species exhibited a typical quadripartite chloroplast genome structure, with genome sizes ranging from 160,716 to 160,842 bp. The gene composition and arrangement were highly conserved, with each species containing 131 genes. Three highly variable regions (*psaI*, *infA*, *rpl22*) were identified as candidate DNA barcodes. The total number of simple sequence repeats (SSRs) was similar across species, with all types distributed among them; however, dispersed repetitive sequences exhibited distinct regional characteristics and interspecific differences, providing potential marker resources for distinguishing closely related species. Phylogenetic analysis revealed that all eight species nested within *Quercus*, with *Quercus shennongii* and *Cyclobalanopsis edithae* clustered together as a distinct branch with 100% bootstrap support. Samples from different geographic sources within the same species did not cluster directly, suggesting the potential for intraspecific genetic variation and geographic population structure differentiation within *Quercus*. This study provides new chloroplast genomic evidence for the taxonomic status of *Cyclobalanopsis*. The phylogenetic results are compatible with treating *Cyclobalanopsis* as a subgenus within *Quercus* but may not be the sole explanation. Final taxonomic treatment requires careful evaluation integrating nuclear genomic, phytogeographic, and population data.

KEYWORDS: Chloroplast genome; *quercus*; *cyclobalanopsis*; phylogeny; molecular markers

1 Introduction

Chloroplasts (cp) are essential matrilineal genetic organelles for green plant cells with independent circular genomes and play an important role in plant photosynthesis [1]. The chloroplast genome of terrestrial plants exhibits a typical quadripartite structure, consisting of a large single-copy region (LSC) and a small single-copy region (SSC) separated by two inverted repeat regions (IR), independent of the nuclear genome [2–4]. Chloroplast genome is highly conservative with low recombination and mutation rates, making it an effective tool for studying plant evolution and taxonomy [5,6]. In recent years, chloroplast DNA has played an increasingly important role in systematic geography and phylogenetic studies [7,8]. Under the premise of ensuring the full availability of cp DNA data, it is very efficient to compare sister

species or other possible species through cp genome information to create phylogenetic relationships, and can provide data for the study of population dynamics and phylogeographic evolution patterns of late evolved species [9]. Chloroplast genes are also regarded as "DNA barcodes" and are expected to become an efficient tool for species identification, contributing not only to molecular identification and differentiation of multiple species, but also to biodiversity analysis [10].

Fagaceae is a key foundational family in the forest ecosystems of the Northern Hemisphere [11]. This family comprises core groups including *Quercus*, *Fagus*, *Castanopsis*, and *Lithocarpus*. Among these, *Fagus* is considered an early-diverging basal lineage within Fagaceae, while *Quercus* represents a core branch resulting from later radiation and evolution within the family [12–14]. *Quercus* is an important member of middle subtropical evergreen broad-leaved forest, possessing important ecological and economic value, and is widely distributed throughout tropical and subtropical Asia [15–17]. Currently, *Quercus* is generally considered to comprise two main branches consisting of 5–6 subgenera [18,19]. However, its infrageneric classification, particularly the taxonomic status of the *Cyclobalanopsis* group, remains a subject of long-standing debate. Traditional morphological classification primarily relies on cupule scale morphology: the cupule scales of *Cyclobalanopsis* are usually connate into concentric rings, whereas those of typical *Quercus* are imbricately arranged [20,21]. Based on this significant difference, many taxonomic systems have treated the *Cyclobalanopsis* group as a distinct genus. However, recent molecular phylogenetic studies have consistently shown that this group is stably nested within *Quercus*, forming a monophyletic clade, which supports its treatment as a subgenus of *Quercus* [22–24]. This conflict between morphological and molecular phylogenetic evidence renders the generic delimitation and infrageneric classification of *Quercus* an urgent issue requiring clarification.

Currently, phylogenetic studies of *Quercus* based on chloroplast genomes have largely focused on European and North American taxa, with insufficient attention given to Asian lineages, particularly the rich endemic species in China, which limits our comprehensive understanding of the evolutionary patterns within this genus [25,26]. To address this gap and provide new genomic insights into the aforementioned taxonomic controversy, this study selected eight *Cyclobalanopsis* species distributed in central-southern China for comparative chloroplast genome analysis. The selection of these species was based on the following rationales: (1) Representativeness: they encompass the morphological diversity of cupule forms within *Cyclobalanopsis* [27]; (2) Geographical focus: they target the understudied subtropical regions of China; (3) Novelty: in this study, the complete chloroplast genome of *Quercus pachyloma* is reported for the first time. Although chloroplast genome data for *Quercus jenseniana* and the other six species have been previously published, this study provides independent genomic resources from new geographical populations (Hunan and Jiangxi Province, China) for each of them. In summary, this study presents, for the first time, a systematic characterization of the complete chloroplast genomes and their annotations for eight *Cyclobalanopsis* species. Through high-throughput sequencing, comparative genomics, and phylogenetic analysis of these key lineages, this study aims to elucidate their genetic relationships, evaluate the efficacy of chloroplast genomes in resolving taxonomic issues within *Cyclobalanopsis*, and provide new data to support the systematic classification of *Quercus*.

2 Methods

2.1 Sample Materials Collection, DNA Extraction, and Sequencing

Fresh leaves of 8 species of *Quercus* were collected from the native environment (Table 1). The collected leaf samples were immediately dried and preserved in silica gel. All samples were identified by the plant taxonomist Lihong Yan (Hunan Botanical Garden) in the field and at the herbarium based on authoritative

taxonomic works, including the the Flora of China [27,28]. Voucher specimens for all samples were prepared and deposited in the Hunan Botanical Garden Herbarium (HNBG); their corresponding collection numbers are listed in Table 1.

Genomic DNA was isolated by the modified method CTAB. Agarose gel electrophoresis and a microspectrophotometer (OD-1000, Shanghai Cytoeasy Biotech Co., Ltd., Shanghai, China) were used to detect DNA integrity and quality. The sequencing of cp genome was conducted by Gene pioneer Biotechnologies Co., Ltd. (Nanjing, China) using the Illumina NovaSeq6000 Sequencing System.

Table 1: Information on the collection of samples from eight species of *Quercus*.

Plant Name	Collection Place	Latitude/Longitude	Altitude (m)	Voucher Specimen
Quercus pachyloma	Tongdao Dong Nationality Autonomous County, Hunan Province	26°4′27″ N, 109°30′45″ E	506 m	HNBG20190419
Quercus jenseniana	Yizhang County, Hunan Province	25°23′57″ N, 112°57′5″ E	575 m	HNBG20190512
Quercus fleuryi	Dayu County, Jiangxi Province	25°33′29″ N, 114°26′05″ E	700 m	HNBG20200502
Quercus gilva	Jingzhou County, Hunan Province	26°34′48″ N, 109°40′48″ E	650 m	HNBG20200505
Quercus glauca	Yongshun County, Hunan Province	29°0′48″ N, 109°57′53″ E	780 m	HNBG20200519
Quercus shennongii	Longshan County, Hunan Province	29°38′4″ N, 109°46′8″ E	550 m	HNBG20200530
Quercus multinervis	Guzhang County, Hunan Province	28°37′10″ N, 109°56′53″ E	750 m	HNBG20200628
Quercus myrsinifolia	Guzhang County, Hunan Province	28°37′10″ N, 109°56′53″ E	605 m	HNBG20200630

2.2 Chloroplast Genome Assembly and Annotation

The data quality control is carried out by using FASTQ software, and the clean data obtained by quality control is assembled by GetOrganelle software [29]. The original data is map to the assembled sequence to check the coverage, and combined with Gapcloser software v1.12 [30] (https://sourceforge.net/projects/soapdenovo2/files/GapCloser/) to fill the gaps in assembly. The cp genome of *Q. stewardiana* (MN199023) was uploaded to NCBI (https://www.ncbi.nlm.nih.gov/) as the reference sequences, and annotated by CpGAVAS2 (http://www.herbalgenomics.org/cpgavas/) [31]. The annotated data were manually adjusted and modified by Geneious Prime 2021 (Biomatters Ltd., Auckland, New Zealand). The physical structure maps of chloroplast genomes were drawn by OGDRAW v1.3.132 (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html). Finally, the annotated data were submitted to NCBI. The sequence accession numbers of *Q. jenseniana*, *Q. pachyloma*, *Q. fleuryi*, *Q. gilva*, *Q. glauca*, *Q. shennongii*, *Q. multinervis* and *Q. myrsinifolia* are respectively: OP442516.1, OP442517.1, OP450821.1, OP450822.1, OP450818.1, OP450820.1, OP450823.1, OP450819.1.

2.3 Codon Usage Analysis

The relative synonymous codon usage of chloroplast genomes of 8 species of *Quercus* was determined and analyzed by CodonW1.4.2 (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=codonw), and the heat map was drawn by RSCU. The RSCU value of >1 indicates that the codon is used more frequently; equal to 1 indicates that the codon has no usage preference; <1 indicates that the codon is used less frequently.

2.4 Comparative Analysis of Quercus Chloroplast Genomes

The comparative cp genome includes the sequence of 8 species of *Quercus*. The using DNA Baser Sequence Assembler v5.152 to quantificat the base content, the 8 chloroplast genomes were aligned using the MAFFT v7.475 [32] with default parameters to identify highly variable regions. The number of polymorphic sites and nucleotide variability (Pi) were evaluated using a sliding window with 200-bp step size and a

600-bp window length implemented in DnaSP v5.10.1 [33]. IRscope (https://irscope.shinyapps.io/irapp/) was used to map the chloroplast genomic IR boundaries [34].

2.5 Repeat Sequences Analysis of Quercus Chloroplast Genomes

Four types of repeat sequences: forward, reverse, complementary, and palindromic, were identified using REPuter with a Hamming distance of 3 and a minimum repeat size of 30 bp [35]. In addition, tandem repeats were identified with Tandem Repeats Finder v4.09 with default parameters [36]. The overlapped repeats of the results were removed manually.

2.6 Phylogenetic Analysis of Quercus Chloroplast Genomes

To construct the phylogenetic tree, we assembled a dataset comprising the complete chloroplast genome sequences of the eight species from this study and 39 closely related species downloaded from the NCBI database, totaling 47 species. To enhance taxonomic representativeness, the analysis included two accessions each for *Quercus fleuryi*, *Quercus myrsinifolia*, and *Quercus gilva*: one newly sequenced in this study (accession numbers OP450821.1, OP450819.1, and OP450822.1, respectively, collected from Jiangxi and Hunan) and one previously published from different geographical regions—*Quercus fleuryi* (MZ502291.1, Fujian), *Quercus myrsinifolia* (MN199025.1, Anhui), and *Quercus gilva* (MK986651.1, Fujian) [37–39]. Species from the genera *Corylus* and *Alnus* of Betulaceae were selected as the familial outgroup [40,41], while species of *Fagus* (Fagaceae) were chosen as the intra-familial basal outgroup to jointly root the phylogenetic tree [1,14,23].

The tree reconstruction procedure was as follows: all circular chloroplast genome sequences were linearized by aligning them to a common starting point. Multiple sequence alignments were performed using MAFFT software (v7.427, —auto mode). The resulting alignments were trimmed using trimAl (v1.4.rev15) to remove poorly aligned regions. Finally, a maximum likelihood (ML) phylogenetic tree was constructed using RAxML v8.2.10 under the GTRGAMMA model with 1000 rapid bootstrap replicates. Node support was assessed using bootstrap values (BS).

3 Results

3.1 Chloroplast Genomes of Eight Species of Quercus

In this study, we obtained the complete chloroplast genomes of eight *Quercus* species, including *Quercus pachyloma* and *Quercus jenseniana*. Among these, the complete chloroplast genome of *Q. pachyloma* is reported here for the first time. Although chloroplast genome information for *Q. jenseniana* and the other six *Quercus* species has been previously published, this study provides independent genomic resources from new geographical populations (Hunan and Jiangxi) for each of them. All genomes exhibited a typical quadripartite structure (Fig. 1; Supplementary Fig. S1). The complete chloroplast genome of *Q. pachyloma* had a total length of 160,778 bp, comprising a LSC of 90,205 bp, an SSC of 18,889 bp, separated by two 25,842 bp IRs. The total length of chloroplast genomes of *Q. jenseniana* was 160,811 bp and consisted of one LSC (90,237 bp), one SSC (18,906 bp), and two IRs (25,834 bp). The cp genomes of *Q. pachyloma*, *Q. jenseniana*, and other 6 species all contained 131 genes, including 85 protein-coding genes, 37 tRNA genes, 8 rRNA genes, and one pseudogene (Table 2).

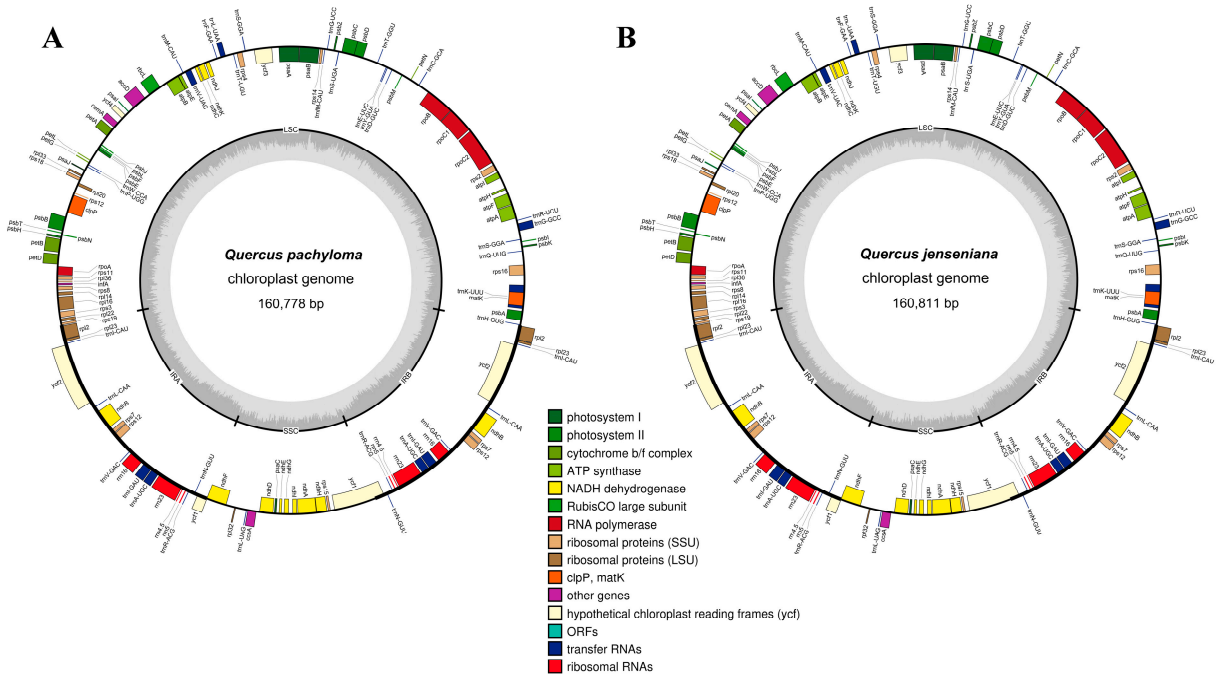

Figure 1: (**A**) Chloroplast genome maps of *Q. pachyloma*; (**B**) Chloroplast genome maps of *Q. jenseniana*. Genes are colored based on different function. The inner circle (dashed gray area) indicates the proportional GC content of the corresponding genes. Regions of the LSC, SSR, IRa, and IRb are indicated.

In addition, the GC content of the complete cp genomes, as well as those of the LSC, SSC, and IR regions, were similar among the eight species, ranging from 36.89% to 36.90%, 34.74% to 34.75%, 31.06% to 31.11%, and 42.77% to 42.80%, respectively (Table 2). A total of 23 genes contained introns in their complete cp genomes, among which 21 genes (13 protein-coding genes and 8 tRNA genes) harbored a single intron, while two genes (*ycf3* and *clpP*) contained two introns. Furthermore, 15 protein-coding genes, 11 tRNA genes, and 8 rRNA genes were duplicated in the IR regions. The LSC region harbored 62 protein-coding genes and 22 tRNA genes, whereas the SSC region contained 12 protein-coding genes and 1 tRNA gene.

To comprehensively evaluate the genomic characteristics obtained in this study, we conducted comparisons at two levels: First, we performed a direct comparison with previously published cp genomes of the same species. We collected published cp genome annotation data for seven species—*Q. jenseniana*, *Q. fleuryi*, *Q. gilva*, *Q. glauca*, *Q. shennongii*, *Q. multinervis*, and *Q. myrsinifolia* (accession numbers are provided in Supplementary Table S1) [37–39,42–45], and compared them in detail with the corresponding newly assembled genomes from this study (Supplementary Table S2). The comparison revealed that, for the same species, the newly assembled genomes differed only minimally from the published versions in total length and the length of LSC, SSC, and IR regions (average difference < 0.1%), and the GC content of each region were nearly identical (Table 2; Supplementary Table S2). This suggests the stability of the overall genome structure within species. However, while all species in this study consistently contained 131 genes, the gene counts in previously published genomes of the same species fluctuated between 128 and 134. For instance, *Q. glauca* (KX852399.1) contained 134 genes, whereas *Q. fleuryi* (MZ502291.1) contained 129 genes. Second, a comparison of gene numbers with closely related species. Compared to the 131 genes identified in the eight species of this study, previously reported *Quercus* species generally possessed slightly higher gene numbers. For example, Yang et al. reported 134 genes for five species including *Quercus variabilis* [10]; Huang et al. reported 133 genes for four species including *Q. ningangensis* [46]; and Li et al.

reported 132 genes for four species including *Q. disciformis* [47]. Given that this study employed a uniform and rigorous annotation workflow, these discrepancies more likely reflect inconsistencies in annotation standards or assembly completeness across different studies.

In summary, these comparisons indicate that the cp genome data obtained in this study are of high quality and internal consistency. The observed differences in gene counts suggest the importance of adopting uniform standards in comparative genomic research, while also potentially revealing genuine microevolutionary variations among different lineages within *Quercus*.

Table 2: Comparison of chloroplast genomes among eight species of *Quercus*.

Species	*Quercus jenseniana*	*Quercus pachyloma*	*Quercus fleuryi*	*Quercus gilva*	*Quercus glauca*	*Quercus shennongii*	*Quercus multinervis*	*Quercus myrsinifolia*
Genome size (bp)	160,811	160,778	160,778	160,779	160,814	160,716	160,776	160,842
Length of LSC (bp)	90,237	90,205	90,205	90,219	90,256	90,213	90,232	90,235
Length of SSC (bp)	18,906	18,889	18,889	18,880	18,900	18,881	18,882	18,939
Length of IRs (bp)	25,834	25,842	25,842	25,840	25,829	25,811	25,831	25,834
Number of protein-coding genes	85	85	85	85	85	85	85	85
Number of tRNA genes	37	37	37	37	37	37	37	37
Number of rRNA genes	8	8	8	8	8	8	8	8
Total number of genes	131	131	131	131	131	131	131	131
GC content of overall (%)	36.89	36.90	36.90	36.90	36.89	36.90	36.90	36.89
GC content of LSC (%)	34.74	34.74	34.75	34.75	34.74	34.75	34.74	34.75
GC content of SSC (%)	31.10	31.10	31.11	31.10	31.08	31.07	31.11	31.06
GC content of IRs (%)	42.77	42.77	42.77	42.77	42.77	42.80	42.78	42.77

Note: for comparative data on the characteristics of the seven published genomes of the same species, please refer to Supplementary Table S2.

3.2 Relative Synonymous Codons Usage

For the two species newly reported in this study, 26,400 and 26,402 codons were obtained from *Q. jenseniana* and *Q. pachyloma*, respectively. A total of 66 codons were detected across the cp genomes of the eight *Quercus* species. Excluding the three stop codons, the remaining codons encode 20 amino acids. Serine was the most frequently used amino acid, while cysteine was the least frequent. Among the encoded amino acids, tryptophan is encoded exclusively by the single codon UGG, with a relative synonymous codon usage (RSCU) value of 1.00, indicating no obvious preference. The codon AUG, encoding methionine, exhibited the highest RSCU value of 6.99. Analysis of the cp genomes of the eight species revealed that multiple codons encode a single protein (Fig. 2). Among these, three amino acids (leucine, arginine, and serine) are encoded by six codons, four amino acids (alanine, glycine, threonine, and valine) are encoded by four codons, one amino acid (isoleucine) is encoded by three codons, and the remaining 11 amino acids are each encoded by two codons.

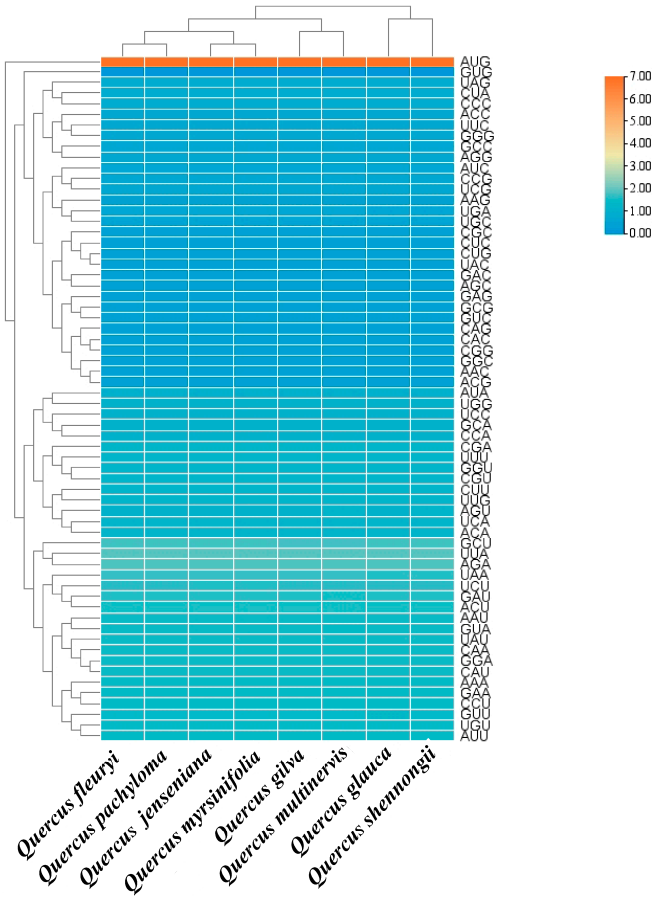

Figure 2: Heat map of the Relative synonymous codons usage of the cp genomes in the eight *Quercus* species.

3.3 Repeat Sequences Analysis

Approximately 290 SSR loci were identified in each *Quercus* species, yielding a total of 2328 SSRs across all species. Each species contained over 160 single-nucleotide repeats, far exceeding the numbers of dinucleotides, trinucleotides, tetranucleotides, and pentanucleotides. The predominant motif types among the eight species were A/T, C/G, AT/AT, ATA/TAT, and TAA/TTA, while tetranucleotides and pentanucleotides were less abundant than other types. Furthermore, all detected SSR types were present in all eight species, with no unique SSR types observed (Fig. 3).

Among the eight species, P repeats (palindromic) were the most common, followed by F repeats (forward). R repeats (reverse) were present in only four species, and C repeats (complement) were found in only three species. In the SSC region, only F and P repeats were present, with *Q. myrsinifolia* and *Q. jenseniana* exhibiting the highest numbers of F repeats, while the number of P repeats was generally similar across species. In the LSC region, four types of repeats (P, F, R, C) were present in *Q. pachyloma*, *Q. glauca*, and *Q. gilva*; three types were present in *Q. jenseniana* (P, F, R); and only F and P repeats were found in the remaining species. Among these, *Q. gilva* had the highest number of F repeats, while the number of P repeats remained largely consistent across species. In the IR region, only F and P repeats were detected, and with the exception of *Q. shennongii*, the numbers of F and P repeats were identical across the other species (Fig. 3).

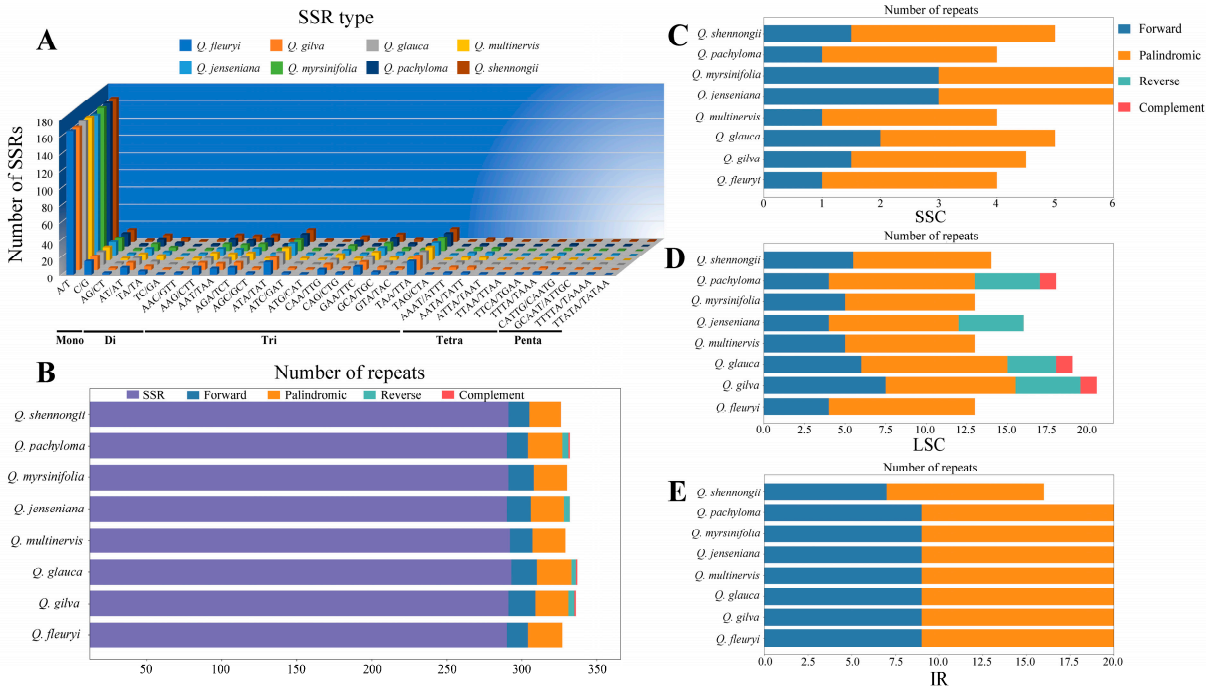

Figure 3: (**A**) Number of different SSR types in eight *Quercus* species; (**B**) Number of SSRs and dispersed repeats in eight *Quercus* species; (**C**) Number of different dispersed repeat types in the SSC region; (**D**) Number of different dispersed repeat types in the LSC region; (**E**) Number of different dispersed repeat types in the IR region.

3.4 Comparative Analysis of the Cp Genome

In the cp genomes of the eight *Quercus* species, the boundaries of LSC/IRb, SSC/IRb, SSC/IRa, and LSC/IRa were stably positioned, although the distances to adjacent genes exhibited some variation. For example, the JSB boundary was located 7 bp from the *ndhF* gene in *Q. pachyloma*, but only 1 bp away in *Q. glauca* (Fig. 4). Furthermore, we compared nucleotide variation across all genes and intergenic regions among the eight cp genomes. Sliding window analysis revealed that nucleotide diversity (Pi) values ranged from 0 to 0.004. Three mutational hotspot regions were identified, encompassing the *psaI*, *infA*, and *rpl22* genes (Fig. 5). These regions represent potential markers for population genetic studies and species identification within *Quercus*.

3.5 Phylogenetic Analysis

In this study, a phylogenetic tree was constructed using the cp genomes of 47 species, including the eight newly sequenced *Quercus* species from this study (Fig. 6). The overall topological structure shows that *Quercus* species (including the traditional *Cyclobalanopsis* group) formed a monophyletic clade together with their closely related genera *Lithocarpus*, *Castanopsis*, and *Castanea*. However, the internal topology of *Quercus* reveals a complex basal structure, detailed as follows: *Quercus rubra* and *Quercus griffithii* formed a clade located at the most basal position of the entire group. Above this clade, *Quercus engleriana* and *Quercus senescens* formed a second basal branch. Further above, *Quercus tarokoensis*, *Quercus phillyreoides*, and *Quercus bawanglingensis* formed a third basal branch. Above these basal *Quercus* lineages, *Castanea*, *Castanopsis*, and *Lithocarpus* formed an independent monophyletic branch, clearly separated from *Quercus*. The crown *Quercus* clade, which includes the eight species from this study and other

typical *Quercus* species, was located at the top of the phylogenetic tree and formed a monophyletic group (bootstrap support = 100%).

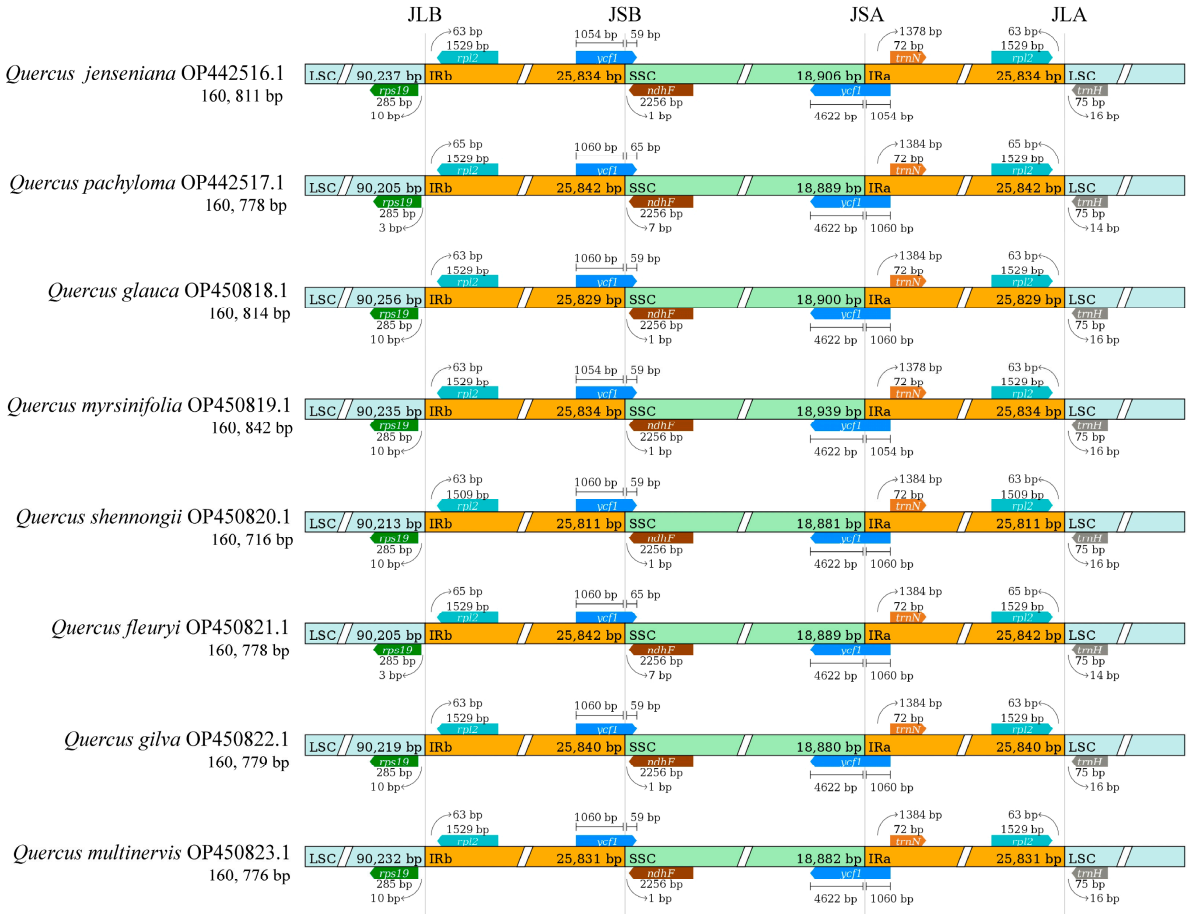

Figure 4: Analyses of expansion and contraction of inverted repeats in the eight *Quercus* chloroplast genomes.

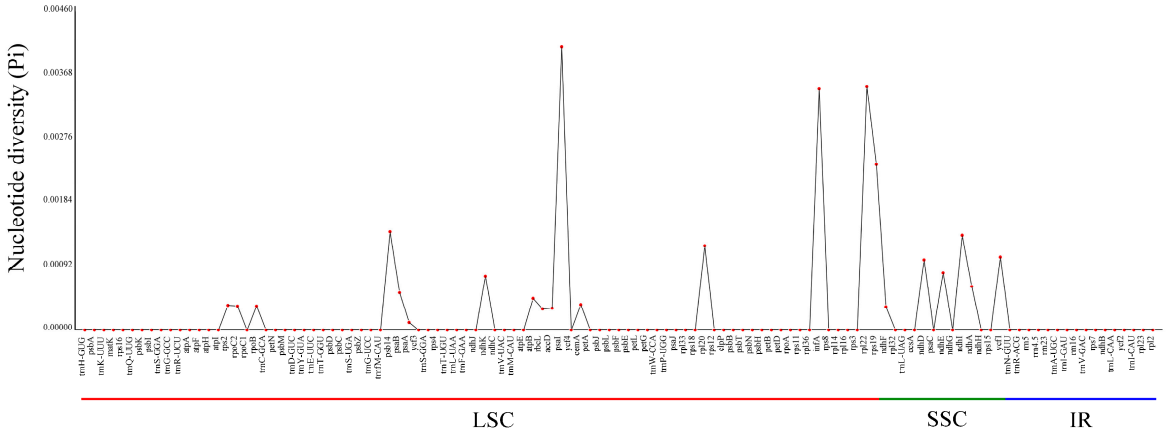

Figure 5: Nucleotide diversity (Pi) values among the eight *Quercus* species.

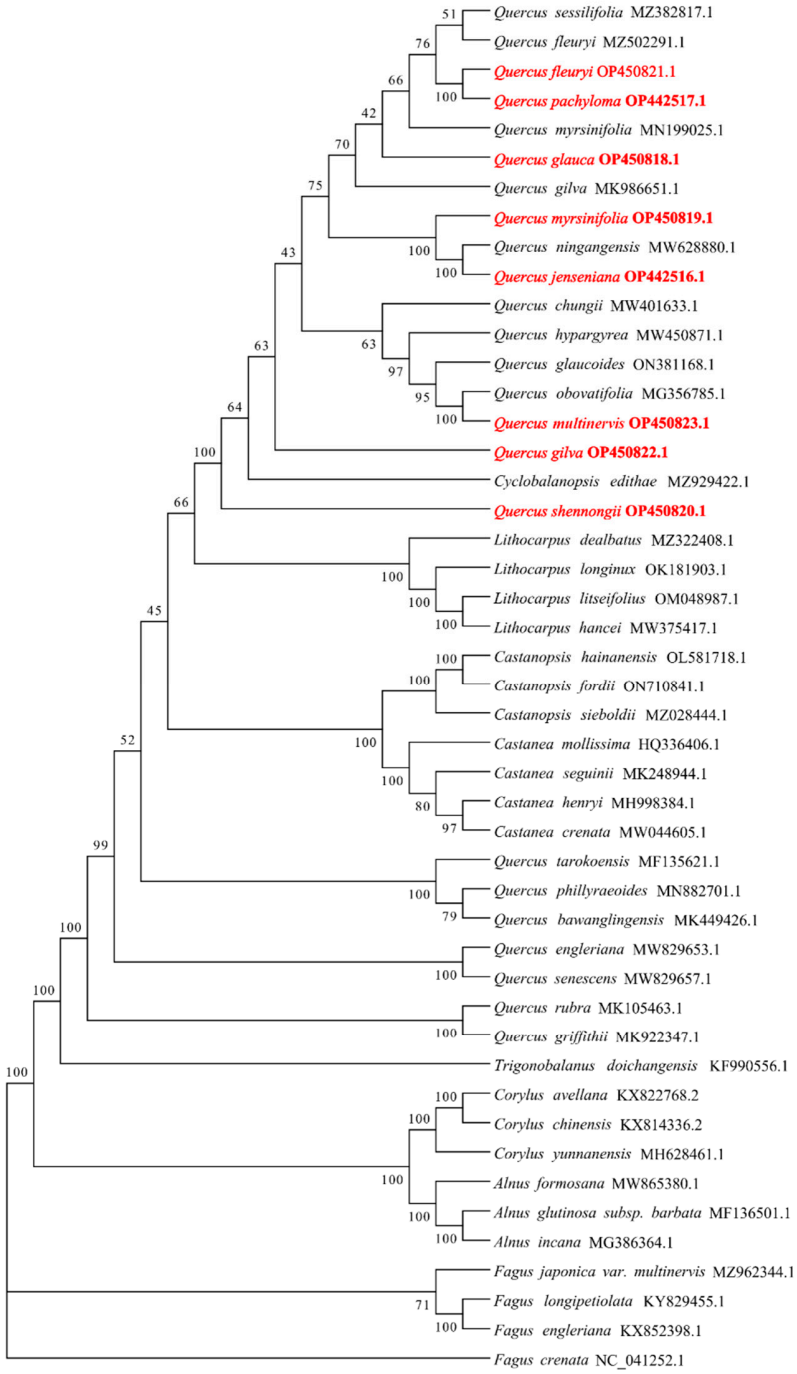

Figure 6: Phylogenetic relationships of 47 species inferred from maximum likelihood method.

Within this crown clade, the following topological structure was observed: First, *Q. fleuryi* (OP450821.1) and *Q. pachyloma* (OP442517.1) formed a sister group (BS = 100%), while another previously published sample of *Q. fleuryi* from Fujian (MZ502291.1) was positioned in a neighboring branch within the same major clade [39]. Second, the *Q. myrsinifolia* sample from this study (OP450819.1) and another previously published sample from Anhui (MN199025.1) did not cluster directly together but were placed in different positions within the same branch [37]; the two samples of *Q. gilva* exhibited a similar pattern [38]. Third, *Q. shennongii* (OP450820.1) and *Cyclobalanopsis edithae* (MZ929422.1) clustered together as a single branch

(BS = 100%), and this branch did not group with the other seven species from this study. The remaining species, including *Q. jenseniana*, *Q. gilva*, *Q. glauca*, and *Q. multinervis*, each formed independent terminal branches. These results provide a preliminary plastome-based phylogenetic framework for investigating the phylogenetic relationships within *Quercus*.

4 Discussion

In this study, we compared the chloroplast genomes of eight *Quercus* species. Among these, the complete chloroplast genome of *Quercus pachyloma* is reported here for the first time. Although chloroplast genome information for *Q. jenseniana* and the other six species has been previously published, this study provides independent genomic resources from new geographical populations (Hunan and Jiangxi) for each of them. Comparative analysis revealed that the chloroplast genomes of the eight species were highly similar in size and composition, ranging from 160,716 to 160,842 bp in total length. The LSC regions ranged from 90,205 to 90,237 bp, the SSC regions from 18,880 to 18,939 bp, and the IR regions from 25,811 to 25,842 bp. The GC content of the IR regions was higher than that of the LSC and SSC regions, and all GC content were below 50%. All eight species contained 131 genes. Notably, as shown in Section 3.1, compared to the consistent gene number of 131 obtained in this study using a uniform workflow, previously published data for the same species exhibited some fluctuation in gene numbers (ranging from 128 to 134) (Supplementary Table S2) [10,37–39,42–47]. This suggests that annotation consistency may be an important factor when conducting cross-study comparisons of chloroplast genomes. The chloroplast genome structure of *Quercus* species is highly conserved [48]. We detected several genes located at boundary regions including *rps19*, *rpl2*, *ycf1*, *ndhF*, *trnN*, and *trnH* in all eight species, a finding consistent with previous studies on other *Quercus* species [1,47]. Although the copy numbers and positions of these genes varied slightly among species, their localization at boundary regions showed no significant differences. Clear codon usage bias was observed in all eight species, and the codon usage patterns were highly consistent across species, with the methionine-encoding codon AUG exhibiting the highest RSCU value.

In *Quercus*, the frequent occurrence of transitional morphologies resulting from hybridization and environmentally driven convergent evolution has substantially compromised the reliability of morphological characteristics for systematic classification [49]. Therefore, the development of efficient and stable molecular markers is crucial for establishing reliable species identification systems [50]. Repetitive sequences play an important role in genetic variation and can serve as molecular markers for phylogenetic analysis and population structure studies [50]. In this study, the eight *Quercus* species exhibited similar SSR abundances, and all detected SSR types were distributed across all species, with no species-specific SSRs identified. This indicates that overall SSR abundance is highly conserved within *Quercus*, and species identification based on SSRs should rely on allele frequency differences rather than simple presence/absence variations. Additionally, all species showed the highest proportion of A/T combinations, a bias that may be associated with lower energy consumption during information transfer [51]. In contrast to the conservatism of SSRs, the distribution of dispersed repeats exhibited pronounced regional characteristics and interspecific variation. In the SSC and IR regions, only F and P repeats were present, while the repeat types in the LSC region were more complex. Certain species, such as *Q. gilva* and *Q. pachyloma*, possessed all four types of repeats in the LSC region, whereas *Q. shennongii* showed a markedly different number of repeats in the IR region compared to other species. These differences provide potential marker resources for molecular identification of closely related *Quercus* species. This study compared the abundance, types, and distribution patterns of chloroplast SSRs and dispersed repeats across eight *Quercus* species, providing foundational data for molecular marker development, species identification, and conservation genetics research [30].

Mutational hotspots are hypervariable regions identified through comparative analysis of chloroplast genomes across different species, making them valuable candidates for DNA barcoding in species identification [52]. The comparison of cp genes differences among species in this study showed that the cp genes of 8 *Quercus* species had high similarity and there might be no obvious differentiation among them. However, there are still three hypervariable regions, namely: *psaI*, *infA*, *rpl22*. Among these, *rpl22* has been consistently detected as a hypervariable region in both this study and previous investigations, suggesting its potential as a candidate DNA barcode for *Quercus* [1,10]. These findings could provide reference information for taxonomic classification and genetic resource development in *Quercus*. However, it should be noted that comparisons based on only eight species may not fully capture the overall characteristics of this diverse genus, and further studies incorporating more taxa and broader population-level sampling are needed to validate these observations.

Phylogenetic analysis plays an essential role in elucidating genetic relationships among species, and understanding these interspecific relationship is of great significance for the rational development, utilization, and conservation of plant resources [53]. Chloroplast genomes provide a molecular perspective for resolving difficult issues in morphological classification and have been widely used in phylogenetic studies [48,54]. However, due to the highly conserved nature of chloroplast genomes (with the highest nucleotide diversity Pi value in this study being only approximately 0.004), their resolving power at the species level is limited. Therefore, although the use of chloroplast genomes to elucidate the genetic relationships of *Quercus* species can provide clues for taxonomic research, it should be integrated with other lines of evidence for comprehensive assessment.

The chloroplast phylogenetic tree constructed in this study (Fig. 6) reveals a complex basal topology within *Quercus*. Notably, three basal *Quercus* lineages appear sequentially at the most basal positions of the phylogenetic tree: *Q. rubra* with *Q. griffithii*, *Q. engleriana* with *Q. senescens*, and *Q. tarokoensis* with *Q. phillyreoides* and *Q. bawanglingensis*. Above these basal Quercus lineages, the clades of *Lithocarpus*, *Castanopsis*, and *Castanea* appear, while the core *Quercus* clade (including the eight species from this study) is located at the most crown position of the phylogenetic tree. This topology warrants further investigation. Liu et al. demonstrated that there is strong cytonuclear discordance at deep nodes within *Quercoideae* and *Quercus*, wherein *Quercus* is non-monophyletic in the chloroplast phylogenetic tree, mainly due to frequent historical gene flow. Therefore, the "basal *Quercus* lineages-closely related genera-core *Quercus* clade" topology observed in this study may not fully reflect the true species divergence order, but rather more likely reflects extensive chloroplast capture or ancient hybridization events between *Quercus* and its closely related genera during early diversification [55]. Indeed, Liu et al. confirmed the existence of extensive gene flow between *Quercus* and major lineages of *Quercoideae* during early evolution, with these ancient hybridization events dating back to the early-to-middle Eocene [55]. This interpretation is also echoed by another observation in this study: some *Quercus* species (e.g., *Q. tarokoensis* and its relatives) are placed below the closely related genera in the chloroplast tree, which may reflect that the chloroplast genomes of these species experienced capture events from the ancestors of closely related genera during their evolutionary history.

Phylogenetic analysis revealed that samples of *Q. fleuryi*, *Q. myrsinifolia*, and *Q. gilva* from different geographical origins did not form monophyletic groups. This phenomenon is relatively common in recently radiated and rapidly diversifying woody genera [56,57]. Considering the characteristics of the chloroplast genomes and the evolutionary background of *Quercus*, the causes of this pattern may be related to the following two factors. On the one hand, the highly conserved nature of the chloroplast genome (with the highest nucleotide diversity Pi value in this study being only approximately 0.004) results in

a slow rate of molecular differentiation. Meanwhile, the core *Quercus* clade belongs to a recently and rapidly radiated group [56]. The ancestral population's chloroplast haplotype polymorphism may not have been completely and randomly fixed during species divergence, leading to incomplete lineage sorting. In this scenario, different geographical populations of the same species may carry different ancestral haplotypes, resulting in their dispersal across different branches in the phylogenetic tree and forming a "non-monophyletic" topological pattern. On the other hand, this result may reflect intraspecific genetic variation and geographical population structure differentiation within *Quercus* [26,58]. Due to the maternal uniparental inheritance of the chloroplast genome, its phylogeographic structure may be more pronounced, with different geographical populations potentially retaining distinct plastid haplotypes [48,59]. This geographically driven haplotype differentiation may lead to substantial differences in the chloroplast genomes of different geographical populations of the same species, resulting in a "same species, different branches" clustering pattern. This may imply the existence of gene flow or complex phylogeographic histories within these widely distributed species [26]. Furthermore, the phenomenon of conspecific samples failing to form monophyletic groups is not rare in chloroplast phylogenetic trees. Ogishima et al. found that different populations of the same species in the genus *Isodon* often appear paraphyletically distributed in chloroplast trees, suggesting that chloroplast capture may be a cause of this phenomenon [60]. Therefore, the non-monophyly of conspecific samples observed in this study may reflect the influence of evolutionary events commonly found in *Quercus*, such as chloroplast capture and incomplete lineage sorting. These findings suggest that when using chloroplast genomes for species-level phylogenetic inference, it is crucial to fully consider the geographical representativeness of samples and potential intraspecific variation. Future studies integrating more extensive population-level sampling with nuclear genomic data will be necessary to more comprehensively elucidate the evolutionary dynamics and species boundaries of these taxa.

Within the crown clade, the highly supported clustering of *Q. shennongii* with *Cyclobalanopsis edithae* is another topological feature worthy of attention. This clustering pattern suggests that the chloroplast genome of *Q. shennongii* is highly similar to that of *C. edithae*, while being differentiated from other species. However, interpretation of this result requires caution. In *Quercus*, a genus characterized by frequent hybridization and introgression, chloroplast capture is a widely documented phenomenon [48,59,61,62] The strongly supported clustering of *Q. shennongii* with *C. edithae* could reflect either close common ancestry or, alternatively, the possibility of historical interspecific hybridization resulting in chloroplast genome introgression from one lineage into another. Chloroplast data alone cannot distinguish between these two hypotheses. This result is compatible with, but does not uniquely support, the treatment of *Cyclobalanopsis* as a subgenus of *Quercus*; it represents only one of several possible explanations. Notably, we observed no significant structural differentiation in the chloroplast genomes, further suggesting that phylogenetic inferences based on this uniparentally inherited marker require comprehensive evaluation integrating multiple lines of evidence, including nuclear genomic data.

The phylogenetic results based on chloroplast genomes in this study provide a new molecular perspective on the traditional taxonomic issues within *Quercus*, while also highlighting conflicts with morphological evidence. First, regarding the discordance between molecular and morphological evidence, the chloroplast phylogeny supports the inclusion of *Cyclobalanopsis* (characterized by concentric cupule scales) within *Quercus*, contradicting traditional morphological classifications that recognize it as a distinct genus based on its imbricate cupule scales [20,21]. Second, concerning evidence supporting the nesting of *Cyclobalanopsis* within *Quercus*, the two groups exhibit extensive overlap and transitional forms in their distribution ranges and morphological features other than cupule morphology, suggesting shared ancestry and co-evolutionary history [63]. Finally, it is essential to recognize methodological limitations: the high

bootstrap support values obtained in this study (e.g, BS = 100%) do not necessarily equate to genuine phylogenetic signal. The complete chloroplast genome is approximately 160 kb in length, and the large alignment size, combined with non-independence among sites, may lead to systematic overestimation of bootstrap values. Therefore, high support alone cannot substitute for independent validation of phylogenetic topology stability. In summary, the molecular evidence presented in this study adds chloroplast genomic data to clarify the phylogenetic position of *Cyclobalanopsis*. However, due to the highly conserved nature of the chloroplast genome (with the highest Pi value of approximately 0.004), its resolving power at the species level is limited, and the uniparentally inherited plastid tree may be influenced by evolutionary events such as chloroplast capture. Consequently, ultimately resolving this controversy will require integration of nuclear genomic data, broader population sampling, and morphological and biogeographical evidence to better elucidate the evolutionary relationships between *Cyclobalanopsis* and other *Quercus* lineages.

5 Conclusion

In this study, we compared the chloroplast genomes of eight *Quercus* species. Among these, the complete chloroplast genome of *Quercus pachyloma* is reported for the first time, and for the remaining seven species, we provide independent genomic resources from new geographical populations (Hunan and Jiangxi). *Quercus* species are widely distributed globally and play important ecological and economic roles; however, their taxonomic system has long remained unresolved. Our results revealed that the chloroplast genomes of these eight *Quercus* species are highly conserved in structure, with highly consistent composition and gene content. The eight species in this study consistently contained 131 chloroplast genes, which may reflect annotation consistency rather than biological variation across studies. Three highly variable regions: *psaI*, *infA*, and *rpl22* were identified through comparative analysis. The phylogenetic tree indicates that this study reconstructed a phylogenetic relationship with high node support among the eight species based on complete chloroplast genome data. However, due to the highly conserved nature of the chloroplast genome (with the highest Pi value of approximately 0.004), its resolving power at the species level is limited. Therefore, this result should be regarded as preliminary molecular evidence. The distinctive clustering of *Quercus shennongii* and *Cyclobalanopsis edithae* provides referable molecular evidence for further exploring the taxonomic status of *Cyclobalanopsis*. This result is compatible with, but does not uniquely support, treating *Cyclobalanopsis* as a subgenus of *Quercus*. Given the genetic characteristics of chloroplast genomes and their complexity in *Quercus* evolution, final taxonomic decisions require careful evaluation integrating multiple lines of evidence, including nuclear genomic, morphological, and biogeographical data.

Acknowledgement: None.

Funding Statement: This research was funded by the Changsha Natural Science Foundation, grant number kq2202356 and the Hunan Forestry Science and Technology Innovation Plan Project, grant number XLK202106-2 and XLK202616.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Xin Li, Xiaoming Tian and Hao Lv; methodology, Xiaoming Tian and Gaofei Li; software, Lu Zhu and Chao Liu; validation, Xin Li, Guangfeng Xiang and Lan Zhou; investigation, Xin Li, Li Liu and Tianyi Jiang; resources: Hao Lv and Cun Mou; data curation, Xiangpeng Li and Li Liu; writing—original draft preparation, Xin Li and Xiaoming Tian; writing—review and editing, Xin Li, Hao Lv and Lan Zhou. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Supplementary Materials: The supplementary material is available online at https://www.techscience.com/doi/10.32604/phyton.2026.079853/s1.

References

1. Li Y, Wang TR, Kozlowski G, Liu MH, Yi LT, Song YG. Complete chloroplast genome of an endangered species *Quercus litseoides*, and its comparative, evolutionary, and phylogenetic study with other *Quercus section Cyclobalanopsis* Species. Genes. 2022;13(7):1184. [CrossRef].
2. Bobik K, Burch-Smith TM. Chloroplast signaling within, between and beyond cells. Front Plant Sci. 2015;6:781. [CrossRef].
3. Daniell H, Lin CS, Yu M, Chang WJ. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. Genome Biol. 2016;17(1):134. [CrossRef].
4. Liu X, Chang EM, Liu JF, Huang YN, Wang Y, Yao N, et al. Complete chloroplast genome sequence and phylogenetic analysis of *Quercus bawanglingensis* Huang, Li et Xing, a vulnerable oak tree in China. Forests. 2019;10(7):587. [CrossRef].
5. Ravi V, Khurana JP, Tyagi AK, Khurana P. An update on chloroplast genomes. Plant Syst Evol. 2008;271(1):101–22. [CrossRef].
6. Kassem MA. Comparative analysis of chloroplast genomes across 20 plant species reveals evolutionary patterns in gene content, codon usage, and genome structure. Int J Plant Biol. 2025;16(3):105. [CrossRef].
7. Ji Y, Yang L, Chase MW, Liu C, Yang Z, Yang J, et al. Plastome phylogenomics, biogeography, and clade diversification of *Paris* (Melanthiaceae). BMC Plant Biol. 2019;19(1):543. [CrossRef].
8. Fan R, Ma W, Liu S, Huang Q. Integrated analysis of three newly sequenced fern chloroplast genomes: Genome structure and comparative analysis. Ecol Evol. 2021;11(9):4550–63. [CrossRef].
9. Shi SL, Liu YQ, Xia RX, Qin L. Comprehensive analysis of codon usage in *Quercus* Chloroplast genome and focus on psbA gene. Genes. 2022;13(11):2156. [CrossRef].
10. Yang Y, Zhou T, Duan D, Yang J, Feng L, Zhao G. Comparative analysis of the complete chloroplast genomes of five *Quercus* Species. Front Plant Sci. 2016;7:959. [CrossRef].
11. Zhou BF, Yuan S, Crowl AA, Liang YY, Shi Y, Chen XY, et al. Phylogenomic analyses highlight innovation and introgression in the continental radiations of Fagaceae across the Northern Hemisphere. Nat Commun. 2022;13:1320. [CrossRef].
12. Manos PS, Steele KP. Phylogenetic analyses of "higher" Hamamelididae based on plastid sequence data. Am J Botany. 1997;84(10):1407–19. [CrossRef].
13. Manos PS, Cannon CH, Oh SH. Phylogenetic relationships and taxonomic status of the paleoendemic Fagaceae of western North America: Recognition of a new genus, *Notholithocarpus*. Madroño. 2008;55(3):181–90. [CrossRef].
14. Oh SH, Manos P. Molecular phylogenetics and cupule evolution in Fagaceae as inferred from nuclear CRABS CLAW sequences. TAXON. 2008;57(2):434–51. [CrossRef].
15. Curtu AL, Gailing O, Finkeldey R. Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. BMC Evol Biol. 2007;7(1):218. [CrossRef].
16. Ortego J, Espelta JM, Armenteras D, Díez MC, Muñoz A, Bonal R. Demographic and spatially explicit landscape genomic analyses in a tropical oak reveal the impacts of late Quaternary climate change on Andean montane forests. Mol Ecol. 2023;32(12):3182–99. [CrossRef].
17. Xiong S, Zhao Y, Chen Y, Gao M, Wang M, Wang Y, et al. Transcriptome and proteome analyses reveal a critical role of the sucrose anabolic pathway in regulating *Quercus fabri* Hance axillary bud outgrowth. Ind Crops Prod. 2023;194:116284. [CrossRef].
18. Denk T, Tekleva MV. Pollen morphology and ultrastructure of *Quercus* with focus on Group *Ilex* (=*Quercus* subgenus *Heterobalanus* (Oerst.) Menitsky): implications for oak systematics and evolution. Grana. 2014;53(4):255–82. [CrossRef].
19. Hubert F, Grimm GW, Jousselin E, Berry V, Franc A, Kremer A. Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus*. Syst Biodivers. 2014;12(4):405–23. [CrossRef].

20. Nixon KC. Infrageneric classification of *Quercus* (Fagaceae) and typification of sectional names. Ann For Sci. 1993;50:25s–34s. [CrossRef].
21. Menitsky YL. Oaks of Asia. New York, NY, USA: Science Publishers; 2005.
22. Luo Y, Zhou ZK. Leaf architecture in *Quercus* subgenus *Cyclobalanopsis* (Fagaceae) from China. Bot J Linn Soc. 2002;140(3):283–95. [CrossRef].
23. Deng M, Zhou ZK, Li QS. Taxonomy and systematics of *Quercus* subgenus *Cyclobalanopsis*. Int Oaks. 2013;24:48–60.
24. Deng M, Jiang XL, Hipp AL, Manos PS, Hahn M. Phylogeny and biogeography of East Asian evergreen oaks (*Quercus section Cyclobalanopsis*; Fagaceae): Insights into the Cenozoic history of evergreen broad-leaved forests in subtropical Asia. Mol Phylogenet Evol. 2018;119:170–81. [CrossRef].
25. Cavender-Bares J, Gonzalez-Rodriguez A, Pahlich A, Koehler K, Deacon N. Phylogeography and climatic niche evolution in live oaks (*Quercus series* Virentes) from the tropics to the temperate zone. J Biogeogr. 2011;38(5):962–81. [CrossRef].
26. Simeone MC, Grimm GW, Papini A, Vessella F, Cardoni S, Tordoni E, et al. Plastome data reveal multiple geographic origins of *Quercus* Group *Ilex*. PeerJ. 2016;4:e1897. [CrossRef].
27. Wu ZY, Raven Peter H. Flora of China. Vol. 4. Beijing, China: Science Press; 1999.
28. Huang CJ. Flora of China. Beijing, China: Science Press; 1998. (In Chinese).
29. Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, Yi TS, et al. Get Organelle: A fast and versatile toolkit for accurate *de novo* assembly of organelle genomes. Genome Biol. 2020;21(1):241. [CrossRef].
30. Lan Z, Shi Y, Yin Q, Gao R, Liu C, Wang W, et al. Comparative and phylogenetic analysis of complete chloroplast genomes from five *Artemisia* species. Front Plant Sci. 2022;13:1049209. [CrossRef].
31. Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, et al. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC Genom. 2012;13(1):715. [CrossRef].
32. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80. [CrossRef].
33. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009;25(11):1451–2. [CrossRef].
34. Amiryousefi A, Hyvönen J, Poczai P. IRscope: An online program to visualize the junction sites of chloroplast genomes. Bioinformatics. 2018;34(17):3030–1. [CrossRef].
35. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. REPuter: The manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 2001;29(22):4633–42. [CrossRef].
36. Benson G. Tandem repeats finder: A program to analyze DNA sequences. Nucleic Acids Res. 1999;27(2):573–80. [CrossRef].
37. Li Y, Wang L, Fang Y. The complete chloroplast genome sequence of *Quercus myrsinifolia* (Fagaceae). Mitochondrial DNA Part B. 2019;4(2):3139–40. [CrossRef].
38. Zeng QM, Liu B, Lin RQ, Jiang YT, Liu ZJ, Chen SP. The complete chloroplast genome sequence of *Quercus Gilva* (Fagaceae). Mitochondrial DNA Part B. 2019;4(2):2493–4. [CrossRef].
39. Lu Y, Ye DQ, Weng YH, Zheng XY, Li Y, Liu P. Chloroplast genome and characteristics analysis of Cyclobalanopsis fleuryi. J Jiangsu For Sci Technol. 2024;51(4):1–7. (In Chinese). [CrossRef].
40. Li JQ. On the phylogeny of the fagaceae. Acta Phytotaxon Sin. 1996;34(6):597–609. (In Chinese).
41. Yang YY, Stull GW, Qu XJ, Deng M, Zhao L, Hu Y, et al. Genome duplications, genomic conflict, and rapid phenotypic evolution characterize the Cretaceous radiation of Fagales. J Integr Plant Biol. 2025;67(11):2929–44. [CrossRef].
42. Yang Y, Zhu J, Feng L, Zhou T, Bai G, Yang J, et al. Plastid genome comparative and phylogenetic analyses of the key Genera in Fagaceae: Highlighting the effect of codon composition bias in phylogenetic inference. Front Plant Sci. 2018;9:82. [CrossRef].
43. Li Y, Wang L, Liu Q, Fang Y. The complete plastid genome sequence of *Quercus ciliaris* (Fagaceae). Mitochondrial DNA Part B. 2020;5(2):1954–5. [CrossRef].

44. Yang Y, Zhou T, Qian Z, Zhao G. Phylogenetic relationships in Chinese oaks (Fagaceae, *Quercus*): Evidence from plastid genome using low-coverage whole genome sequencing. Genomics. 2021;113(3):1438–47. [CrossRef].
45. Huang T. Chloroplast genome analysis and phylogenetic of Section Glauca of Subgen. Cyclobalanopsis from China [dissertation]. Nanchong, China: China West Normal University; 2023. (In Chinese).
46. Huang T, Tang M, Chen XL, Li BY, Zhang XM. Comparison of chloroplast genomes and phylogenic analysis of four species of *Quercus* subg. *Cyclobalanopsis*. Guihaia. 2023;43(4):741–54. (In Chinese). [CrossRef].
47. Li B, Huang K, Chen X, Qin C, Zhang X. Comparative and phylogenetic analysis of chloroplast genomes from four species in *Quercus* section *Cyclobalanopsis*. BMC Genom Data. 2024;25(1):57. [CrossRef].
48. Zhou YR, Li Y, Yang LH, Kozlowski G, Yi LT, Liu MH, et al. The adaptive evolution of *Quercus section Ilex* using the chloroplast genomes of two threatened species. Sci Rep. 2024;14:20577. [CrossRef].
49. Yang J, Guo YF, Chen XD, Zhang X, Ju MM, Bai GQ, et al. Framework phylogeny, evolution and complex diversification of Chinese oaks. Plants. 2020;9(8):1024. [CrossRef].
50. Zhou B, Yao W, Guo C, Bian L, Ding Y, Lin S. Chloroplast genome variation and phylogenetic analyses of seven dwarf ornamental bamboo species. Forests. 2022;13(10):1671. [CrossRef].
51. Kim GB, Lim CE, Kim JS, Kim K, Lee JH, Yu HJ, et al. Comparative chloroplast genome analysis of *Artemisia* (Asteraceae) in East Asia: Insights into evolutionary divergence and phylogenomic implications. BMC Genom. 2020;21(1):415. [CrossRef].
52. Niu Y, Su T, Wu C, Deng J, Yang F. Complete chloroplast genome sequences of the medicinal plant *Aconitum transsectum* (Ranunculaceae): Comparative analysis and phylogenetic relationships. BMC Genom. 2023;24(1):90. [CrossRef].
53. Zhang JY, Liao M, Cheng YH, Feng Y, Ju WB, Deng HN, et al. Comparative chloroplast genomics of seven endangered *Cypripedium* Species and phylogenetic relationships of Orchidaceae. Front Plant Sci. 2022;13:911702. [CrossRef].
54. Wang LL, Li Y, Zheng SS, Kozlowski G, Xu J, Song YG. Complete chloroplast genomes of four oaks from the section *Cyclobalanopsis* Improve the phylogenetic analysis and understanding of evolutionary processes in the genus *Quercus*. Genes. 2024;15(2):230. [CrossRef].
55. Liu SY, Yang YY, Tian Q, Yang ZY, Li SF, Valdes PJ, et al. An integrative framework reveals widespread gene flow during the early radiation of oaks and relatives in *Quercoideae* (Fagaceae). J Integr Plant Biol. 2025;67(4):1119–41. [CrossRef].
56. Zhou BF. A study on phylogeny and genomic divergence of Fagaceae [dissertation]. Guangzhou, China: South China Botanical Garden, Chinese Academy of Sciences; 2021. (In Chinese).
57. Guo C, Ma PF, Yang GQ, Ye XY, Guo Y, Liu JX, et al. Parallel ddRAD and genome skimming analyses reveal a radiative and reticulate evolutionary history of the temperate bamboos. Syst Biol. 2021;70(4):756–73. [CrossRef].
58. Zhang RS, Yang J, Hu HL, Xia RX, Li YP, Su JF, et al. A high level of chloroplast genome sequence variability in the Sawtooth Oak *Quercus acutissima*. Int J Biol Macromol. 2020;152:340–8. [CrossRef].
59. Sousa F, Viegas MB, Costa J, Marques I, Pina-Martins F, Simões F, et al. Haplotype diversity patterns in *Quercus suber* (Fagaceae) inferred from cpDNA sequence data. Plant Syst Evol. 2023;309(6):42. [CrossRef].
60. Ogishima M, Horie S, Kimura T, Yamashiro T, Dohzono I, Kawaguchi L, et al. Frequent chloroplast capture among *Isodon* (Lamiaceae) species in Japan revealed by phylogenies based on variation in chloroplast and nuclear DNA. Plant Species Biol. 2019;34(3):127–37. [CrossRef].
61. Dumolin-Lapègue S, Kremer A, Petit RJ. Are chloroplast and mitochondrial DNA variation species independent in oaks? Evolution. 1999;53(5):1406–13. [CrossRef].
62. Yang YY, Qu XJ, Zhang R, Stull GW, Yi TS. Plastid phylogenomic analyses of Fagales reveal signatures of conflict and ancient chloroplast capture. Mol Phylogenet Evol. 2021;163:107232. [CrossRef].
63. Denk T, Grimm GW, Manos PS, Deng M, Hipp AL. An updated infrageneric classification of the oaks: Review of previous taxonomic schemes and synthesis of evolutionary patterns. In: Oaks physiological ecology. Exploring the functional diversity of genus Quercus L. Cham, Switzerland: Springer International Publishing; 2017. p. 13–38. [CrossRef].