

Article

The Complete Chloroplast Genome Sequence of *Machilus chuanchienensis* (Lauraceae): Genome Structure and Phylogenetic Analysis

Xue Bai, Juan Peng, Yongyi Yang and Biao Xiong *

College of Tea Science, Guizhou University, Guiyang 550000, China

* Correspondence: bxiong@gzu.edu.cn

Abstract: *Machilus chuanchienensis* is an ecological tree distributed in southwestern China. It has a significant valuation with making Hawk tea using its leaves, an ethnic traditional tea-like beverage with a long history in Chinese tea culture. The whole chloroplast (cp) genome is an ideal model for the phylogenetic study of Lauraceae because of its simple structure and highly conserved features. There have been numerous reports of complete cp genome sequences in Lauraceae, but little is known about *M. chuanchienensis*. Here, the next-generation sequencing (NGS) was used to sequence the *M. chuanchienensis* cp genome. Then, a comprehensive comparative genome analysis was performed. The results revealed that the *M. chuanchienensis*'s cp genome measured 152,748 base pairs (bp) with a GC content of 39.15% and coded 126 genes annotated, including comprising eight ribosomal RNA (rRNA), 36 transporter RNA (tRNA), and 82 protein-coding genes. In addition, the cp genome presented a typical quadripartite structure comprising a large single-copy (LSC; 93,811) region, a small single-copy (SSC; 18,803) region, and the inverted repeats (IRs; 20,067) region and contained 92 simple sequence repeat (SSR) locus in total. Phylogenetic relationships of 37 species indicated that *M. chuanchienensis* was a sister to *M. balansae*, *M. melanophylla*, and *M. minutiflora*. Further research on this crucial species may benefit significantly from these findings.



Citation: Bai, X.; Peng, J.; Yang, Y.; Xiong, B. The Complete Chloroplast Genome Sequence of *Machilus chuanchienensis* (Lauraceae): Genome Structure and Phylogenetic Analysis. *Genes* **2022**, *13*, 2402. <https://doi.org/10.3390/genes13122402>

Academic Editor: Zhiqiang Wu

Received: 20 November 2022

Accepted: 15 December 2022

Published: 18 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: *Machilus chuanchienensis* S. Lee; comparative plastid genomics; Illumina sequencing; genome skimming

1. Introduction

Machilus chuanchienensis S. Lee, an indecuate tree belonging to the Lauraceae family, is distributed in low-altitude montane forests in southeastern China. It contains substances such as terpenoids and flavonoids, quercitrin, kaempferol, hyperin, astragaloside, isoquercitrin, and quercetin [1]. As a tea-like plant, *M. chuanchienensis* can be utilized to make a traditional folk beverage Hawk tea [2], one of the non-Camellia teas, which has yellowish-red tea soup with camphor-aromatic smell [3]. According to *the Compendium of Materia Medica*, a classic book of traditional Chinese medicine, *M. chuanchienensis* can be used as Chinese medicine because of rich in polyphenols, flavonoids, vitamins, minerals, and other compounds and is free from caffeine. Moreover, due to the pharmacological effects of Hawk tea, such as antidiabetic [4], acting antioxidant, hypolipidemic, and anti-inflammatory properties [5], consumers sincerely like this beverage [6], and in some places, its consumption is even higher than that of green tea [7].

Lauraceae species are extensively spread in the world's tropical and subtropical regions, which contain about 2500–3000 species from around 50 genera [8]. Lauraceae is an evolutionarily complex and taxonomically controversial group in which the phylogenetic location in the disputed genera (such as *Actinodaphne* Nees and *Sassafras* J. Presl) has been controversial [9,10]. Although *Litsea coreana* var. *lanuginosa* is one of the most researched primary raw materials for making Hawk tea [11], there is a lack of relevant research on other plants belonging to Lauraceae species [12], especially for *Machilus chuanchienensis*,

making it challenging to explore Hawk tea germplasm resources and study the phylogenetic relationships of its material-species.

As important plastids in plant cells, chloroplasts are involved in the main functions, including photosynthesis [13], carbon sequestration, starch storage, nitrogen metabolism, fatty acids, and nucleic acid synthesis [14]. In addition, they self-replicate organelles with their DNA and RNA and have relatively independent genetic replication mechanisms like mitochondria [15]. Although the structure of the cp genome is remarkably conserved [16], its size and gene content have polymorphism related to genes reorganization and IRs region growth and contraction [17], which has been used for molecular identification, population genetics, endangered species conservation, and phylogenetic analysis [18,19]. High-throughput technologies' successful development has made the sequencing of chloroplast genomes economical and efficient, significantly developing chloroplast-based related studies. One of the most valuable methods at the moment is building a phylogenetic tree of the Lauraceae utilizing the cp genomes, provided that the technology for sequencing and assembling nuclear and mitochondrial genomes of Lauraceae is not to be universal [20]. However, using whole genome sequencing (WGS) technologies to research *M. chuanchienensis*' genome has yet to be reported, and its taxonomic position within Lauraceae has not been precisely placed.

In this study, we first assembled the whole cp genome sequence of *M. chuanchienensis* using NGS technology. In addition, a comprehensive analysis of the whole cp genomic features was performed. Then, a phylogenetic tree was constructed using the cp genome sequences to explore the phylogenetic relationship between *M. chuanchienensis* and other species, which provided some theoretical basis for understanding the genomic features of *M. chuanchienensis* and its phylogenetic relationships.

2. Materials and Methods

2.1. Experimental Materials and DNA Extraction

Ya'an city, Sichuan Province, China's habitats were the locations where the samples were taken (103°00' E, 29°98' N), and the fresh leaves were picked in the form of one bud and two leaves. The voucher specimen was stored at the Herbarium of Forestry College, Guizhou University. The accession number is YA202108MC02. The genomic DNA of *M. chuanchienensis* was extracted from healthy fresh young leaves by modified CTAB method [21] and was detected by ultraviolet spectrophotometer and 1% agarose gel electrophoresis [22].

2.2. DNA Sequencing, Genome Assembly and Annotations

Genomic DNA qualified for library construction that constructs a pair-end library with an insertion size of 150 bp was sequenced using the Illumina NovaSeq 6000 platform in NGS. Sequencing data was acquired for quality control, and then the GetOrganelle script was run on a Linux system to assemble the genome [23]. Bandage software was used to visualize if it is a circle [24]. The *Neolitsea homilantha* cp genome sequence was used as a reference to annotate genome using the cp genome annotation website CPGAVAS2 and the annotation results provided collinear analysis, gene function classification information, and intron information tables [25]. The tRNA was then identified by the tRNAscan-SE 1.21 program [26], and a circle diagram of the cp genome was drawn using the OrganelleGenomeDRAW v.1.3.1 [27].

2.3. Simple Sequence Repeats (SSRs) and Repeat Sequences Analysis

SSRs in the cp genome of *M. chuanchienensis* were identified by the microsatellite identification online tool (MISA) [28]. The search parameters for a minimum number of repeats were set as follows: 10 for mononucleotide repeats, five for dinucleotide repeats, four for trinucleotide repeats, and three for tetra-, penta-, and hexanucleotide repeats. In addition, '100 bp' was selected as the minimum distance between the two SSRs. The two

SSRs were considered to be a compound microsatellites if the distance was less than 100 bp. REP-UPer [29,30] was used to identify repeat sequences with the parameters reported [31–33].

2.4. Putative RNA Editing Site and Codon Usage

The coding sequences (CDS) were obtained by the CPGAVAS2 online software to anticipate the potential RNA editing sites in the *M. chuanchienensis* chloroplast genome. Then, they were submitted to the predictive RNA editors for the plant chloroplast (PREP-cp) database [34]. The *M. chuanchienensis*'s relative synonymous codon usage (RCSU) and codon usage count were examined using MEGA X [35].

2.5. Genomic Comparison with Other Species in *Machilus*

The *M. chuanchienensis* cp genome and four related species were homogeneously compared using the mVISTA program [36] made a homogeneity comparison. Of the *M. chuanchienensis* cp genome and the four related species, including *M. balansae*, *M. grijsii*, *M. robusta*, and *M. yunnanensis*, all of which were in the genus *Machilus*. Among them, the *M. chuanchienensis* cp genome was selected as the reference with the Shuffle-LAGAN.

The PhyloSuite software [37] extracted 76 common protein-coding genes from *M. balansae*, *M. grijsii*, *M. chuanchienensis*, *M. robusta*, and *M. yunnanensis*. The extracted gene sequences were initially aligned using the MAFFT software [38,39]. Then, the ratios of nonsynonymous (Ka) to synonymous (Ks) substitutions (Ka/Ks) were calculated by the DnaSP software [40]. The expansion and contraction of IR boundaries were detected using the web program IRscope [41]. Analyzing the nucleotide diversity (Pi) value also utilized the DnaSP program [40]. 500 bp was chosen as the step size and window length.

2.6. Phylogenetic Analysis

To create a maximum likelihood (ML) phylogenetic tree, we retrieved the cp genome sequences of 34 species from the Lauraceae family and two outgroup species from the Calycanthaceae, including *Chimonanthus nitens* (NC_042745) and *Idiospermum australiense* (NC_042743). Eighty-one CDS_NCU genes were extracted, aligned, and merged using PhyloSuite software [37]. The ML phylogenetic tree was conducted by IQ-TREE version 2 [42], with a TVM + F + I + G4 model chosen based on the Bayesian Information Criterion [43] from the result of ModelFinder.

3. Results

3.1. Structure and Characteristics of the *M. chuanchienensis* Chloroplast Genome

The *M. chuanchienensis* cp genome, like that of most angiosperms, is a covalently closed double-stranded cyclic molecule with a total length of 152,748 bp, including a small single-copy (SSC) region (18,803 bp), a large single-copy (LSC) region (93,811 bp) and a pair of inverted repeats (IRs) regions (20,067 bp) (Figure 1). The total GC content of the chloroplast genome was 39.15%, and the AT content was 60.85%, which had evident AT bias. In addition, there were some differences in the GC content for the IR, LSC, and SSC regions. The IR region had the highest GC content (44.43%), followed by the LSC region (37.94%) and the SSC region (33.92%) because the IR region's rRNA genes have a high GC content. (Figure 2).

The chloroplast genome annotation results showed that *M. chuanchienensis* contained 126 functional genes, of which 82, 8, and 36, respectively, were protein-coding, rRNA, and tRNA genes. Of these genes, six tRNA genes (*trnA-UGC*, *trnL-CAA*, *trnI-GAU*, *trnR-ACG*, *trnV-GAC*, *trnN-GUU*), three protein-coding genes (*rps7*, *rps12*, *ndhB*), and four rRNA genes (*rrn5*, *rrn23*, *rrn4.5*, *rrn16*) were located in the IR region, all of which were duplicated once in the IRs regions (Table 1). In addition, the *rps12* gene had a trans-spliced structure, with its 5' end in the LSC region and its 3' end in the IR region.

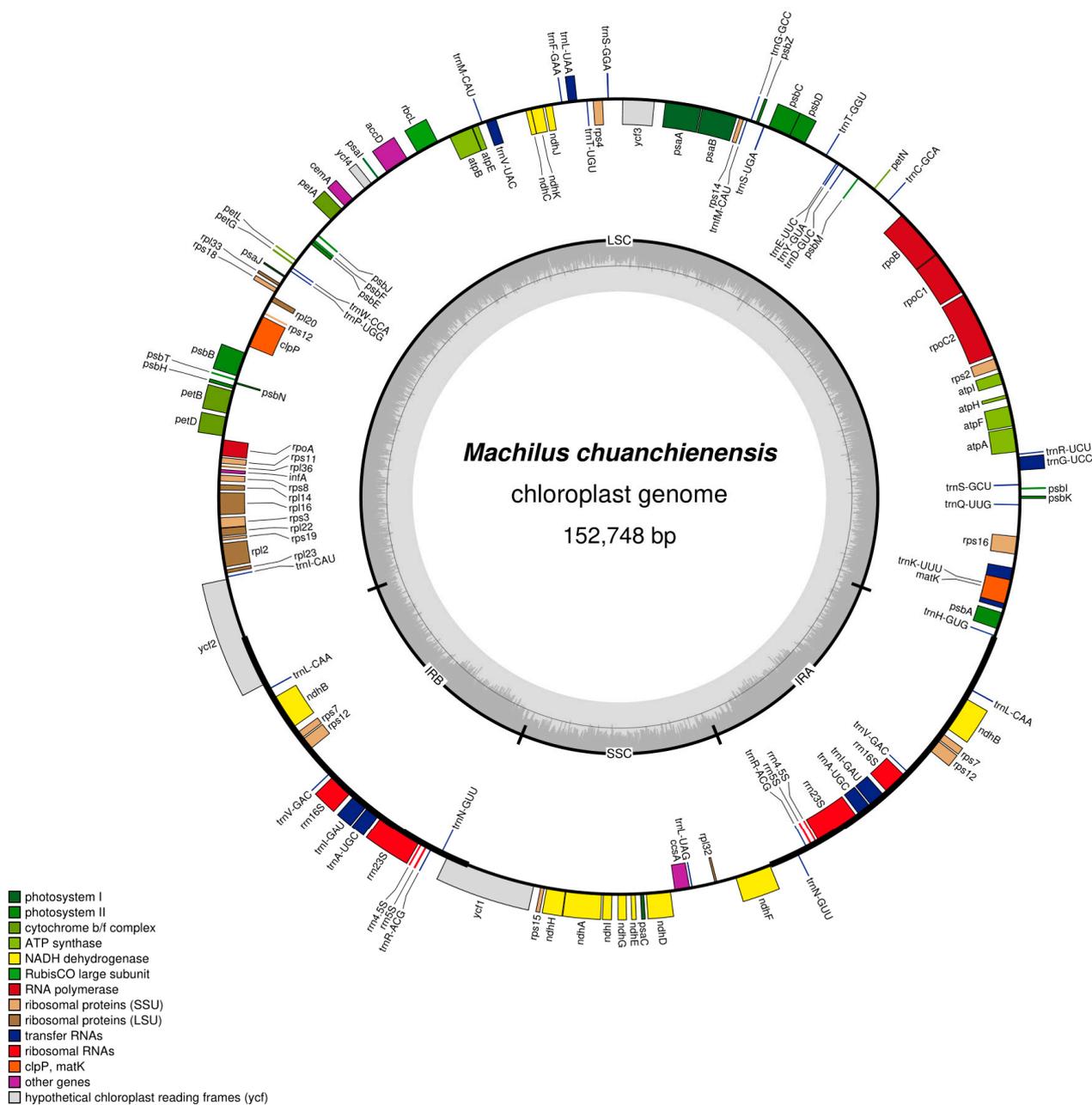


Figure 1. Chloroplast genome map of *M. chuanchienensis*. Genes on the outside of the main circle are transcribed clockwise, while genes on the inside are transcribed counterclockwise. The Organellar Genome Draw (OGDraw) online software was used to draw this map. Genes with different functions are represented by different colors. The gray portion of the inner circle indicates the GC content of the chloroplast genome.

Introns contribute significantly to the regulation of gene expression. There are 18 intron-containing genes in the *M. chuanchienensis* cp genome, including six tRNA (*trnG-UCC*, *trnI-GAU*, *trnA-UGC*, *trnK-UUU*, *trnL-UAA*, *trnV-UAC*) and 12 protein-coding genes (*rps12*, *ycf3*, *rps16*, *rpl16*, *petB*, *rpl2*, *petD*, *atpF*, *clpP*, *ndhA*, *rpoC1*, *ndhB*). The *clpP*, *rps12*, and *ycf3* contain two introns, and the others contain one intron (Table 2). The *trnK-UUU* gene contains the protein-coded gene *matK* and the maximum intron with a length of 2507 bp, which has similar properties to other green plants [44].

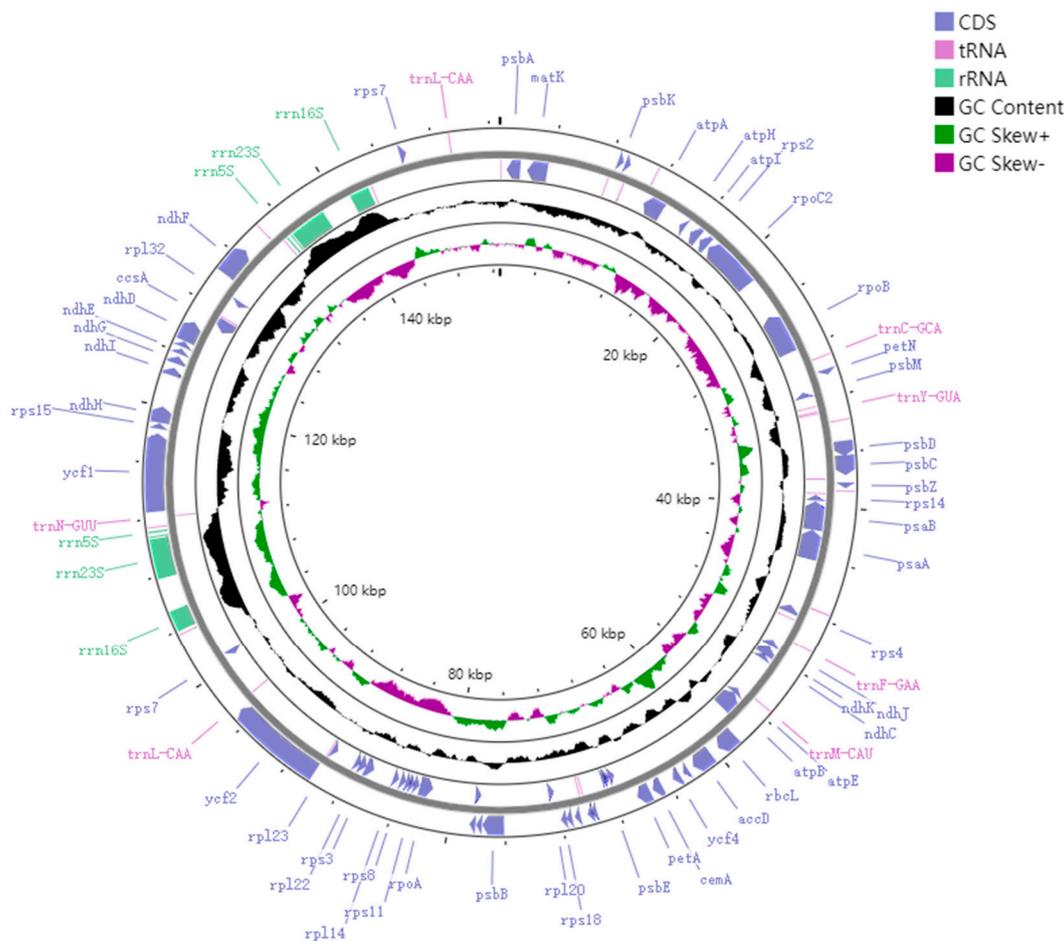


Figure 2. GC content of the *M. chuanchienensis*. This map was created using the web program GCView Server. The black portion represents changes in GC content in different regions of the genome. The deviation of G and C content in each single strand is called GC skew. The specific calculation method is $(nG - nC)/(nG + nC)$, so GC skew + (green portion) means that the content of G is greater than that of C, and GC skew - (magenta portion) means that the content of G is less than that of C.

Table 1. Genes found in the assembled *M. chuanchienensis* chloroplast genome.

Category of Genes	Group of Genes	Name of Genes
RNA genes	Transfer RNA	<i>trnH-GUG, trnK-UUU^b, trnQ-UUG, trnS-GCU, trnG-UCC^b, trnR-UCU, trnC-GCA, trnD-GUC, trnY-GUA, trnE-UUC, trnT-GGU, trnS-UGA, trnG-GCC, trnM-CAU, trnS-GGA, trnT-UGU, trnL-UAA^b, trnF-GAA, trnV-UAC^b, trnM-CAU, trnW-CCA, trnP-UGG, trnI-CAU, trnL-CAA (×2), trnV-GAC (×2), trnI-GAU^b (×2), trnA-UGC^b (×2), trnR-ACG (×2), trnN-GUU (×2), trnL-UAG</i>
	Ribosomal RNA	<i>rrn23 (×2), rrn16 (×2), rrn5 (×2), rrn4.5(×2)</i>
Transcription and translation related genes	DNA dependent RNA polymerase	<i>rpoA, rpoB, rpoC1^b, rpoC2</i>
	Large subunit of ribosome	<i>rpl2^b, rpl14, rpl16^b, rpl20, rpl22, rpl23, rpl32, rpl33, rpl36</i>
	Small subunit of ribosome	<i>rps2, rps3, rps4, rps7 (×2), rps8, rps11, rps12^{ac} (×2), rps14, rps15, rps16^b, rps18, rps19</i>
Photosynthesis-related genes	ATP synthase	<i>atpA, atpB, atpE, atpF^b, atpH, atpI</i>
	Photosystem I	<i>psaA, psaB, psaC, psal, psaj</i>
	Photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbI, psbJ, psbK, psbM, psbN, psbT, psbZ</i>
	Cytochrome b/f complex	<i>petA, petB^b, petD^b, petG, petL, petN</i>
	Large subunit of rubisco	<i>rbcL</i>

Table 1. Cont.

Category of Genes	Group of Genes	Name of Genes
Other genes	NADH dehydrogenase	<i>ndhA^b, ndhB^b (×2), ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	Translational initiation factor	<i>infA</i>
	Acetyl-CoA carboxylase	<i>accD</i>
	Maturase	<i>matK</i>
	Protease	<i>clpP^a</i>
	Envelop membrane protein	<i>cemA</i>
Unknown function	c-type cytochrom synthesis gene	<i>ccsA</i>
	Conserved open reading frames	<i>ycf1, ycf2, ycf3^a, ycf4</i>

Note: ^a Gene containing two introns; ^b gene containing a single intron; ^c gene divided into two independent transcription units; (×2) gene with two copies.

Table 2. Location and length of genes containing introns in *M. chuanchienensis* chloroplast.

Gene	Location	ExonI(bp)	IntronI(bp)	ExonII (bp)	IntronII (bp)	ExonIII (bp)
<i>ndhA</i> ⁺	SSC	553	1119	539		
<i>atpF</i> [−]	LSC	145	725	410		
<i>clpP</i> [−]	LSC	71	776	294	652	244
<i>petB</i> ⁺	LSC	6	790	642		
<i>petD</i> ⁺	LSC	8	716	475		
<i>rpoC1</i> [−]	LSC	453	718	1620		
<i>rps16</i> [−]	LSC	40	852	230		
<i>trnG-UCC</i> ⁺	LSC	23	765	48		
<i>rpl16</i> [−]	LSC	9	976	396		
<i>rpl2</i> [−]	LSC	392	672	430		
<i>trnK-UUU</i> [−]	LSC	37	2507	35		
<i>trnL-UAA</i> ⁺	LSC	35	479	50		
<i>rps12</i> [#]	LSC	114	—	232	536	26
<i>trnV-UAC</i> [−]	LSC	39	589	35		
<i>ycf3</i> [−]	LSC	124	734	230	730	153
<i>ndhB</i> ⁺	IR	721	702	758		
<i>ndhB</i> [−]	IR	721	702	758		
<i>trnA-UGC</i> ⁺	IR	38	798	35		
<i>trnA-UGC</i> [−]	IR	38	798	35		
<i>trnI-GAU</i> ⁺	IR	37	945	35		
<i>trnI-GAU</i> [−]	IR	37	945	35		

Note: ⁺ Exon is transcribed counterclockwise in Figure 1; [−] Exon is transcribed clockwise in Figure 1; — spliceosomal intron; [#] *rps12* is a trans-spliced gene with the 5' end located in the large single copy (LSC) region; it is duplicated in the 3' end in the IR regions.

3.2. Analysis of SSRs and Long Repeats

SSRs analysis showed that there were 92 SSRs loci, including 67 mononucleotides, ten dinucleotides, three trinucleotides, ten tetranucleotides, one pentanucleotide, and one hexanucleotide repeats. Mononucleotide SSRs were the most abundant, accounting for 72.83%. Moreover, A/T, AT/AT, and AAAT/ATTT motifs were 80.43%, indicating that SSRs of *M. chuanchienensis* preferred to use A and T bases (Table S1). Except for SSRs, a repeat (≥ 30 bp) is considered a long repeat sequence. The *M. chuanchienensis* chloroplast genome had 31 long repeats in total, including 13 forward, 5 reverses, 2 complement, and 11 palindrome repeats (Table S2). The size of the repeats ranged from 30 to 72 bp, of which the longest repeat resided in the LSC region (72 bp).

3.3. Codon Usage and Putative RNA Editing Site within *M. chuanchienensis*

The *M. chuanchienensis* cp genome had 23,598 codons in its all protein-coding genes. (Table S3). Among these codons, the three most numerous amino acids were leucine (2396, 10.15%), isoleucine (2003, 8.49%), and serine (1847, 7.83%), while the three least numerous

amino acids were cysteine (272, 1.15%), tryptophan (405, 1.72%), and methionine (557, 2.36%). Based on the calculation of relative synonymous codon use (RSCU), 31 codons with an A or T ending had RSCU >1 except TTG (Leucine, 1.27) and TCC (Serine, 1.03), and 31 codons had RSCU < 1, the vast majority of which ended in C or G, with only CTA (Leucine, 0.89) and ATA (Isoleucine, 0.92) ending in A. In addition, ATG and TGG had no codon bias (RSCU = 1) (Figure 3).

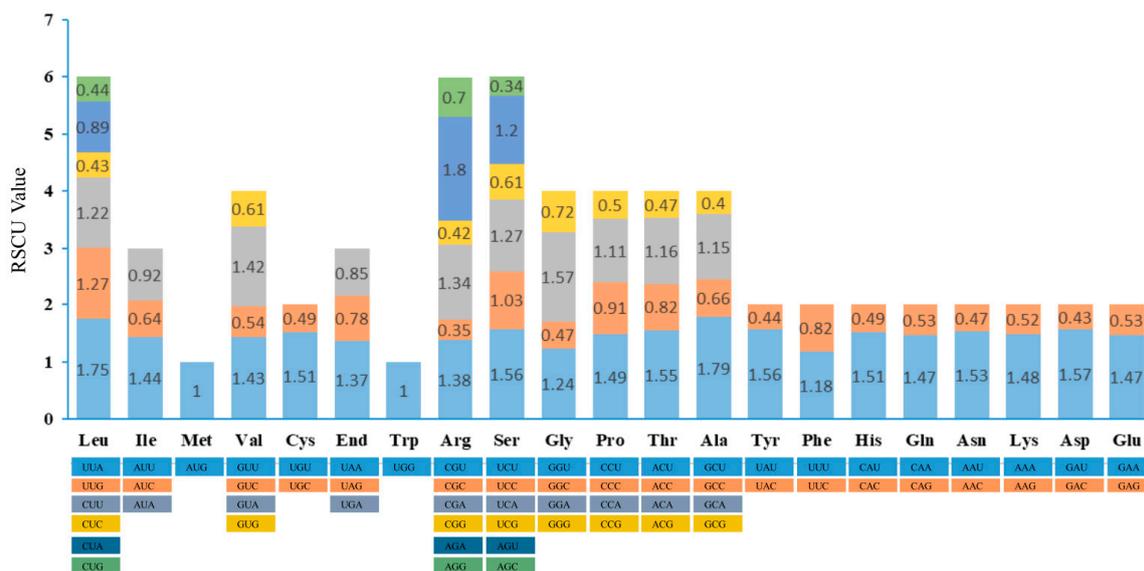


Figure 3. RSCU value of 20 amino acids and stop codons in all protein-coding genes of *M. chuanchienensis* chloroplast genome.

Prediction of putative RNA editing sites for *M. chuanchienensis* cp genes revealed a total of 123 editing sites, all with expected C-U transitions (Table S4). Of the 82 protein-coding genes, 32 genes had RNA edits. All editing caused amino acid changes, of which the S→L transformed form occurred most frequently. Among all genes in which editing occurred, the *ndhB* had the most abundant editing sites (up to 15). In addition, analysis of the codon positions where editing occurred revealed that editing occurred at the first and second codon positions but not at the third.

3.4. Genomic Comparison with Other Species in *Machilus*

The *M. chuanchienensis* chloroplast genome (NC_062133) was used as a reference for the global comparison via the online genome comparison tool mVISTA [36], and a comparison of the chloroplast genome sequences of its four related species revealed that these sequences were little changed overall, except for individual sequences in certain regions. First, compared to the conservative protein-coding regions, the intergenic spacer regions in the genomic sequences of the five chloroplasts were significantly more variable. There was little alteration in the rRNA genes, which were largely conserved. The rRNA genes were highly conserved with little variation. The genes in coding regions, such as *ndhF*, *ycf1*, *ccsA*, *rps15*, *rpl23*, and *ycf2*, were highly variable and visualized with large differences in peak maps (Figure 4). In the intergenic regions, *psbA-trnH-GUG*, *trnQ-UUG-rps16*, *trnD-GUC-trnY-GUA*, *ndhK-atpB*, *ycf4-cemA*, *rbcL-accD*, *psbE-petL*, *petA-psbJ*, *ndhH-ndhA*, *rpl32-trnL-UAG*, and *rpl32-ndhF* had a higher divergence. In addition, compared to the IR region, the LSC and SSC regions had much higher genetic variability.

The chloroplast genes of *M. chuanchienensis* with the other four species in *Machilus* were compared using the nonsynonymous (Ka) and synonymous (Ks) replacement rates to determine whether selection had taken place. (Table S5). We calculated the Ka/Ks values of 76 common protein-coding genes, and those with a Ka or Ks value of 0 were not included in the statistics. The results showed that the Ka/Ks ratio of *M. chuanchienensis* to *M. balansae*

3.5. Phylogenetic Analyses

In order to clarify the phylogenetic status and evolutionary relationship of *M. chuanchienensis* in Lauraceae, the whole chloroplast genome sequences of 36 reported species were selected to construct the ML phylogenetic tree with two Calycanthaceae species as outgroups. According to the findings, two distinct groupings can be made up of all Lauraceae species: the genera *Machilus*, *Neocinnamomum*, and *Cassytha* clustered into one group, while *Cryptocarya*, *Endiandra*, and *Beilschmiedia* clustered into another. *M. chuanchienensis* was the first to be separated from the sister clade of the genus *Machilus* with a 100% bootstrap value (Figure 7).

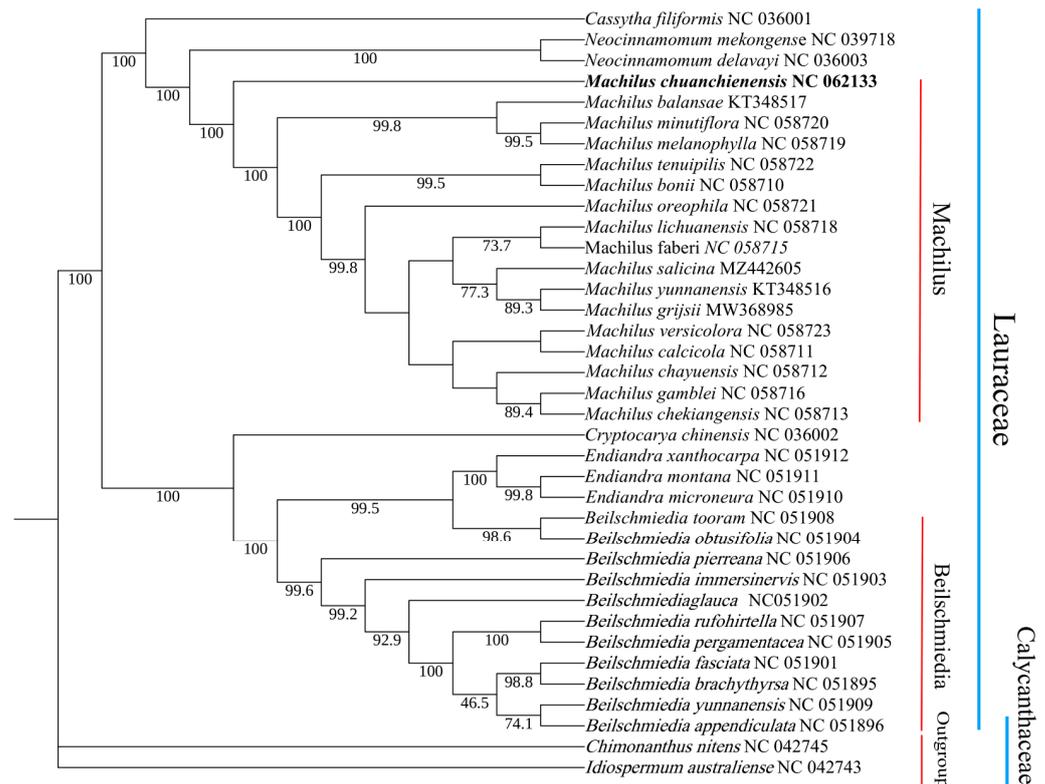


Figure 7. Maximum-likelihood phylogenetic tree of *M. chuanchienensis* based on the whole chloroplast genomes of 36 previously reported species (Numbers on the nodes are bootstrap values from 1000 replicates).

4. Discussion

The published lengths of the cp genomes of the genus *Machilus* ranged from 152,621 to 153,943 bp [19,45,46]. In comparison, the total length of the whole cp sequence of *M. chuanchienensis* assembled in this study was 152,748 bp, indicating that its cp genome size was in line with the traits of the *Machilus* species. Furthermore, the cp genome of the *Machilus* was found to be highly conserved when the *M. chuanchienensis* and the reported *Machilus* species in the Lauraceae family were compared. Additionally, it has been reported that the genus *Machilus* has 113–128 total cp genes. In this study, *M. chuanchienensis* cp genome was annotated to 126 genes (82 protein-coding genes, 36 tRNA genes, and eight rRNA genes). The cp genome GC content of *Machilus* species was similar at 39.15% to 39.16%. However, the GC content in the LSC region (37.93% to 33.95%) and SSC region (33.90% to 34.04%) was significantly lower than that in the IR region (41.43%), probably because the eight rRNA genes with higher GC content were distributed in the IR region in this study, which is similar to previous studies on other angiosperms cp genomes [47,48]. Furthermore, similar to other studies [49,50], three genes (*clpP*, *rps12*, and *ycf3*) in the cp genome of *Machilus* contained two intron regions. *ClpP* encodes the Clp proteolytic

enzyme subunit, whose function is primarily responsible for the degradation of abnormal proteins and is associated with maintaining the normal metabolism of chloroplasts [51]. Moreover, Boudreau et al. [52] showed that the gene *ycf3* interacts with the PSI subunits at the post-translational level [53] and is required for the accumulation of the photosystem I (PSI) complex. Therefore, more research on these genes is required.

The variation in SSR copy number in chloroplasts is an essential molecular marker with a more significant taxonomic distance than nuclear and mitochondrial microsatellites. It has various applications in plant population genetics, polymorphism, and evolutionary studies [54,55]. A total of 92 simple sequence repeats and 31 long repeats were obtained through in-line software analysis, which can provide candidate molecular markers for related studies such as genetic diversity and conservation genetics of *M. chuanchienensis*. According to studies, *M. chuanchienensis*' cp genome's SSRs were relatively abundant in polyadenine (poly-A) or polythymine (poly-T) repeats and seldom contained tandem guanine (G) or cytosine (C), which was similar to other plant cp genomes that have been reported [32,56,57]. According to the results of this study, mononucleotide repeats (72.83%) were the most repeated, which was in line with previous studies [58]. Additionally, a high percentage of forward repetitions (41.94%) were discovered among the four types of repeats, consistent with other studies that demonstrated forward repeats to be the most prevalent [59].

Different species' genomes exhibit varying relative synonymous codon usage (RSCU). There are biases in codon usage, which can provide critical information for studying species evolution [60]. In addition, codons play a role in vector design for chloroplast genetic engineering and are generally optimized first for vector design [61]. The main reason for codon preference selection is that some preferred codons are more efficient in translation [62]. This study, 23,598 codons were found in all protein-coding genes in the *M. chuanchienensis* cp genome. The most used codons were AAA, GAA, AUU, and AAU, similar to the previous studies in other angiosperms [31,63,64].

The most conserved portion of the cp genome is the IR region, as is widely known. The IR, LSC, and SSC regions' growth and contraction are frequent evolutionary occurrences and the main factor influencing variations in the length of the cp genome [65–67]. This study showed that the length of the IR regions of the cp genome among the related species was less different (20,067–20,092 bp). In addition, studies have shown that repetitive sequences are the leading cause of fragment duplication, deletion, and rearrangement of the cp genome [68]. The *ycf1* and *ndhF* genes are significantly rearranged in the *M. chuanchienensis* compared with the *M. robusta*, *M. salicina*, *M. bonil*, and *M. calcicole*. It was also found in *Dendrobium thyrsiflorum* species [69], and it might be a variant in the cp genome's boundary area that changed the structure of the cp gene, which may be a variation in the boundary region of the cp genome that led to changes in chloroplast gene structure [70]. The analysis results based on the mVISTA software showed that the noncoding regions occurred at a relatively higher level of divergence than the coding regions [63]. We identified that in the intergenic regions, *psbA-trnH-GUG*, *trnQ-UUG-rps16*, *trnD-GUC-trnY-GUA*, *ndhK-atpB*, *rbcL-accD*, *ycf4-cemA*, *petA-psbJ*, *psbE-petL*, *ndhH-ndhA*, *rpl32-trnL-UAG*, and *rpl32-ndhF* had a higher divergence. At the species level, these areas could undergo faster replacement. Understanding and mastering these mutation hotspots will make it easier to comprehend the evolutionary characteristics of the genus *Machilus* cp genome and allow the design of molecular markers based on these sequence fragments to identify molecular DNA barcode screening in the genus [71].

The K_a/K_s ratio mainly reflects the selection pressure of protein-coding genes, which is a meaningful way to detect whether protein-coding sequences have evolved. In this study, the vast majority of K_a/K_s ratios were less than 1% (97%), indicating that most of the genes in the Lauraceae family undergo purification selection, which was consistent with the results of previous studies in the Lauraceae family [72]. Moreover, the results showed that two K_a/K_s ratios were more significant than 1 (*matK*, *ndhA*), indicating that they were significantly positively selected. The *ndhA* belongs to the NADH dehydrogenase

subunit maturase gene. The chloroplast NDH monomer sensitive to bright light may have undergone dramatic changes, resulting in the development of new anti-stress functions, and positive selection has existed in the study of *Quercus* [73]. The *matK* gene is located in an intron between two exons with a highly conserved chloroplast lysine tRNA gene (*trnK*), with a sequence length of about 1500 bp. It is a single-copy coding gene encoding a mature enzyme (maturase) involved in the cleavage of type II introns in RNA transcripts [74]. *matK* is often used as a phylogenetic signal to address evolutionary relationships due to its high amino acid replacement rates and nucleotide [74]. However, a positive selection site in *matK* of *Machilus* suggests that this positive selection corrects beneficial variation of *Machilus*, and positive selection has existed in the study of *Chrysosplenium* [75]. Unfortunately, there are few K_a/K_s analyses of Lauraceae. Previous studies found that only two genes, *rpl16* and *ycf2*, had K_a/K_s values greater than 1 through the analysis of nine Lauraceae species [72]. Sequence mutational hotspots, also known as hyper-variable regions, provide a reference for designing accurate and efficient molecular markers and species barcodes [76]. The analysis of nucleotide diversity value calculated by the DnaSP software revealed that the SSC regions have high variability, which has also been found in other Lauraceae [77]. It also proves SSCs generally have a higher nucleotide replacement rate than IRS in land plants [78]. The *ndhA* requires our special attention and in-depth study, which has both a high K_a/K_s value and a P_i value. These may indicate that the *ndhA* has undergone a considerable mutation, which is crucial for the evolutionary process of the *Machilus* species.

The plant cp genome is second only to the nuclear genome and has much genetic information [79]. Therefore, whole-genome sequencing technology provides a new platform and idea to study the evolution system of medicinal plants [80]. We employed next-generation sequencing technologies for our sequencing. It is less expensive than first-generation sequencing technologies and does not need cloning, DNA sequence amplification, or strand termination, thus increasing sequencing speed and throughput [81]. Currently, there are two main methods to obtain cp genomes: one is to rely on traditional methods, first isolate chloroplasts, then extract chloroplast DNA and send it for sequencing, and finally assemble and splice to obtain cp genomes. Since the content of chloroplasts in plants is already small, it is not easy to completely separate chloroplast DNA and nuclear genomic DNA. This method is challenging to operate and takes a long time, so it is limited in terms of application. The second is the more commonly used method we adopt, the total DNA of the extract species is sequenced with high throughput, the chloroplast sequence of the species' close relatives is found as the reference sequence, and the results of the sequencing are compared to find the reads belonging to the chloroplast, and finally assembled. This method first breaks the shackles of traditional methods, eliminates the step of isolating chloroplasts, reduces the time and expense of the experiment, and improves the accuracy of the experiment. In this experiment, we used the second method to obtain the complete cp genome of *M. chuanchienensis* and perform a phylogenetic analysis. As in previous studies [82], the *Machilus* genera came together in this study. The results here showed that all Lauraceae species could be divided into two broad groups, genera *Machilus*, *Neocinnamomum*, and *Cassytha* clustered into one group; *Cryptocarya*, *Endiandra*, and *Beilschmiedia* clustered into another. The phylogenetic relationship of the *Machilus* genera obtained in this study was consistent with the results obtained by Wu et al. [45]. The results of this study will contribute to the subsequent phylogenetic studies and species identification of *Machilus* genera.

5. Conclusions

This study yielded the first complete sequence of the *M. chuanchienensis* cp genome. A comparative analysis of the cp genomes in five *Machilus* species was performed. The findings showed that all genomes of the species mentioned in this study exhibited a degree of relative conservation in terms of their content, gene order, and structure. However, the *ycf1* and *ndhF* in the *M. chuanchienensis* were significantly rearranged. The position of *M. chuanchienensis* inside a phylogenetic tree created using the whole cp genome was

evident. In addition, 92 SSRs that can be employed in breeding, population genetics, and evolutionary research were found. These findings may offer a clear foundation for the phylogenetic relationships of the *M. chuanchienensis* and provide essential data for exploring and utilizing tea-like species resources.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13122402/s1>, Table S1: Number of different SSRs types in *M. chuanchienensis*; Table S2: Long repeat sequences in the *M. chuanchienensis* chloroplast genome; Table S3: Codon usage of *M. chuanchienensis* chloroplast genome (used Universal Genetic code); Table S4: Putative RNA editing sites in the *M. chuanchienensis* chloroplast genes; Table S5: The nonsynonymous (Ka) and synonymous (Ks) replacement rates of the *M. chuanchienensis* chloroplast genome.

Author Contributions: Conceptualization, B.X.; methodology, X.B.; software, X.B.; validation, Y.Y.; resources, B.X.; data curation, J.P.; writing—original draft preparation, X.B.; writing—review and editing, B.X.; visualization, J.P.; supervision, B.X.; project administration, X.B.; funding acquisition, B.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (31900272; 32260086), Guizhou Science and Technology Support Project (Qiankehe Support [2020]1Y146), Guizhou Science and Technology Plan Project (Qiankehe Basics-ZK [2021] General 151), Cultivation Project of Guizhou University (Gzu. 2020 No. 65) and Guizhou Provincial Postgraduate Research Fund (YJSKYJJ [2021] 005).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are openly available in GeneBank at <https://www.ncbi.nlm.nih.gov/> (accessed on 24 January 2022). The complete chloroplast genome has been deposited in GeneBank with accession number OL685165. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA797553, SRS11758196, and SAMN25010768, respectively.

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

References

1. Liu, J.-C.; Chan, P.; Hsu, F.-L.; Chen, Y.-J.; Hsieh, M.-H.; Lo, M.-Y.; Lin, J.-Y. The In vitro inhibitory effects of crude extracts of traditional Chinese herbs on 3-Hydroxy-3-Methylglutaryl-Coenzyme a reductase on Vero cells. *Am. J. Chin. Med.* **2002**, *30*, 629–636. [[CrossRef](#)] [[PubMed](#)]
2. Tan, L.-H.; Zhang, D.; Wang, G.; Yu, B.; Zhao, S.-P.; Wang, J.-W.; Yao, L.; Cao, W.-G. Comparative analyses of flavonoids compositions and antioxidant activities of Hawk tea from six botanical origins. *Ind. Crops Prod.* **2015**, *80*, 123–130. [[CrossRef](#)]
3. Jia, X.; Dong, L.; Yang, Y.; Yuan, S.; Zhang, Z.; Yuan, M. Preliminary structural characterization and antioxidant activities of polysaccharides extracted from Hawk tea (*Litsea coreana* var. *lanuginosa*). *Carbohydr. Polym.* **2013**, *95*, 195–199. [[CrossRef](#)] [[PubMed](#)]
4. Lu, Y.X.; Zhang, Q.; Li, J.; Sun, Y.X.; Wang, L.Y.; Cheng, W.M.; Hu, X.Y. Antidiabetic effects of total flavonoids from *Litsea Coreana* leve on fat-fed, streptozotocin-induced type 2 diabetic rats. *Am. J. Chin. Med.* **2010**, *38*, 713–725. [[CrossRef](#)]
5. Liu, Y.; Wang, X.; Chen, Q.; Luo, L.; Ma, M.; Xiao, B.; Zeng, L. *Camellia sinensis* and *Litsea coreana* Ameliorate Intestinal Inflammation and Modulate Gut Microbiota in Dextran Sulfate Sodium-Induced Colitis Mice. *Mol. Nutr. Food Res.* **2020**, *64*, e1900943. [[CrossRef](#)]
6. Wei, C.; He, P.; He, L.; Ye, X.; Cheng, J.; Wang, Y.; Li, W.; Liu, Y. Structure characterization and biological activities of a pectic polysaccharide from cupule of *Castanea henryi*. *Int. J. Biol. Macromol.* **2018**, *109*, 65–75. [[CrossRef](#)]
7. Chen, Z.; Zhang, D.; Guo, J.-J.; Tao, W.; Gong, R.-X.; Yao, L.; Zhang, X.-L.; Cao, W.-G. Active Components, Antioxidant, Inhibition on Metabolic Syndrome Related Enzymes, and Monthly Variations in Mature Leaf Hawk Tea. *Molecules* **2019**, *24*, 657. [[CrossRef](#)]
8. Chandrabali, A.S.; Werff, H.; Renner, S. Phylogeny and historical biogeography of Lauraceae: Evidence from the chloroplast and nuclear genomes. *Ann. Mo. Bot. Gard.* **2001**, *88*, 104–134. [[CrossRef](#)]
9. Li, L.; Madriñán, S.; Li, J. Phylogeny and biogeography of *Caryodaphnopsis* (Lauraceae) inferred from low-copy nuclear gene and ITS sequences. *Taxon* **2016**, *65*, 433–443. [[CrossRef](#)]
10. Song, Y.; Yao, X.; Tan, Y.; Gan, Y.; Corlett, R. Complete chloroplast genome sequence of the avocado: Gene organization, comparative analysis, and phylogenetic relationships with other Lauraceae. *Can. J. For. Res.* **2016**, *46*, 1293–1301. [[CrossRef](#)]
11. Jia, X.J.; Ding, C.B.; Yuan, S.; Zhang, Z.W.; Chen, Y.E.; Du, L.; Yuan, M. Extraction, purification and characterization of polysaccharides from Hawk tea. *Carbohydr. Polym.* **2014**, *99*, 319–324. [[CrossRef](#)] [[PubMed](#)]
12. Liu, H.; Na, W.; Liu, Z.; Chen, X.; Su, X. A novel turn-on fluorescent strategy for sensing ascorbic acid using graphene quantum dots as fluorescent probe. *Biosens. Bioelectron.* **2017**, *92*, 229–233. [[CrossRef](#)] [[PubMed](#)]

13. Shinozaki, K.; Hayashida, N.; Sugiura, M. Nicotiana chloroplast genes for components of the photosynthetic apparatus. *Photosynth. Res.* **1988**, *18*, 7–31. [[CrossRef](#)] [[PubMed](#)]
14. Martin, G.; Baurens, F.-C.; Cardi, C.; Aury, J.-M.; D'Hont, A. The Complete Chloroplast Genome of Banana (*Musa acuminata*, Zingiberales): Insight into Plastid Monocotyledon Evolution. *PLoS ONE* **2013**, *8*, e67350. [[CrossRef](#)]
15. Fu, P.C.; Zhang, Y.Z.; Geng, H.M.; Chen, S.L. The complete chloroplast genome sequence of *Gentiana lawrencei* var. *farreri* (Gentianaceae) and comparative analysis with its congeneric species. *PeerJ* **2016**, *4*, e2540.
16. Gray, M.; Sankoff, D.; Cedergren, R.J. On the evolutionary descent of organisms and organelles: A global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Res.* **1984**, *12*, 5837–5852. [[CrossRef](#)]
17. Sloan, D.B.; Triant, D.A.; Forrester, N.J.; Bergner, L.M.; Wu, M.; Taylor, D.R. A recurring syndrome of accelerated plastid genome evolution in the angiosperm tribe *Sileneae* (*Caryophyllaceae*). *Mol. Phylogenetics Evol.* **2013**, *72*, 82–89. [[CrossRef](#)]
18. Raman, G.; Park, S. The Complete Chloroplast Genome Sequence of the *Speirantha gardenii*: Comparative and Adaptive Evolutionary Analysis. *Agronomy* **2020**, *10*, 1405. [[CrossRef](#)]
19. Song, Y.; Dong, W.; Liu, B.; Xu, C.; Yao, X.; Gao, J.; Corlett, R.T. Comparative analysis of complete chloroplast genome sequences of two tropical trees *Machilus yunnanensis* and *Machilus balansae* in the family Lauraceae. *Front. Plant Sci.* **2015**, *6*, 662. [[CrossRef](#)]
20. Liu, Z.-F.; Ma, H.; Ci, X.-Q.; Li, L.; Song, Y.; Liu, B.; Li, H.-W.; Wang, S.-L.; Qu, X.-J.; Hu, J.-L.; et al. Can plastid genome sequencing be used for species identification in Lauraceae? *Bot. J. Linn. Soc.* **2021**, *197*, 1–14. [[CrossRef](#)]
21. Doyle, J.J.; Doyle, J.L. Isolation of plant DNA from fresh tissue. *Focus.* **1990**, *12*, 13–15.
22. Wang, Y.; Wang, S.; Liu, Y.; Yuan, Q.; Sun, J.; Guo, L. Chloroplast genome variation and phylogenetic relationships of *Atractylodes* species. *BMC Genom.* **2021**, *22*, 1–12. [[CrossRef](#)] [[PubMed](#)]
23. Jin, J.-J.; Yu, W.-B.; Yang, J.-B.; Song, Y.; Depamphilis, C.W.; Yi, T.-S.; Li, D.-Z. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **2020**, *21*, 241. [[CrossRef](#)] [[PubMed](#)]
24. Wick, R.R.; Schultz, M.B.; Zobel, J.; Holt, K.E. Bandage: Interactive visualization of de novo genome assemblies. *Bioinformatics* **2015**, *31*, 3350–3352. [[CrossRef](#)] [[PubMed](#)]
25. Shi, L.; Chen, H.; Jiang, M.; Wang, L.; Wu, X.; Huang, L.; Liu, C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* **2019**, *47*, W65–W73. [[CrossRef](#)]
26. Schattner, P.; Brooks, A.N.; Lowe, T.M. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* **2005**, *33*, W686–W689. [[CrossRef](#)]
27. Greiner, S.; Lehwark, P.; Bock, R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **2019**, *47*, W59–W64. [[CrossRef](#)]
28. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
29. Kurt, S. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2011**, *29*, 4633–4642. [[CrossRef](#)]
30. Kurtz, S.; Schleiermacher, C. REPuter: Fast computation of maximal repeats in complete genomes. *Bioinformatics* **1999**, *15*, 426–427. [[CrossRef](#)]
31. Li, Y.; Sylvester, S.P.; Li, M.; Zhang, C.; Li, X.; Duan, Y.; Wang, X. The Complete Plastid Genome of *Magnolia zenii* and Genetic Comparison to *Magnoliaceae* species. *Molecules* **2019**, *24*, 261. [[CrossRef](#)] [[PubMed](#)]
32. Liu, H.-Y.; Yu, Y.; Deng, Y.-Q.; Li, J.; Huang, Z.-X.; Zhou, S.-D. The Chloroplast Genome of *Lilium henrici*: Genome Structure and Comparative Analysis. *Molecules* **2018**, *23*, 1276. [[CrossRef](#)] [[PubMed](#)]
33. Liu, L.; Wang, Y.; He, P.; Li, P.; Lee, J.; Soltis, D.E.; Fu, C. Chloroplast genome analyses and genomic resource development for epilithic sister genera *Oresitrophe* and *Mukdenia* (Saxifragaceae), using genome skimming data. *BMC Genom.* **2018**, *19*, 235. [[CrossRef](#)]
34. Mower, J.P. The PREP suite: Predictive RNA editors for plant mitochondrial genes, chloroplast genes and user-defined alignments. *Nucleic Acids Res.* **2009**, *37* (Suppl. S2), W253–W259. [[CrossRef](#)]
35. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)]
36. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32* (Suppl. S2), W273–W279. [[CrossRef](#)] [[PubMed](#)]
37. Zhang, D.; Gao, F.; Jakovlić, I.; Zhou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2019**, *20*, 348–355. [[CrossRef](#)]
38. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2017**, *20*, 1160–1166. [[CrossRef](#)]
39. Kuraku, S.; Zmasek, C.M.; Nishimura, O.; Katoh, K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* **2013**, *41*, W22–W28. [[CrossRef](#)]
40. Rozas, J.; Ferrer-Mata, A.; Sánchez-DelBarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)]
41. Amirouf, A.; Hyvönen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [[CrossRef](#)] [[PubMed](#)]

42. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)] [[PubMed](#)]
43. Zhu, J.; Wen, D.; Yu, Y.; Meudt, H.M.; Nakhleh, L. Bayesian inference of phylogenetic networks from bi-allelic genetic markers. *PLoS Comput. Biol.* **2018**, *14*, e1005932. [[CrossRef](#)] [[PubMed](#)]
44. Zhang, H.; Li, C.; Miao, H.; Xiong, S. Insights from the Complete Chloroplast Genome into the Evolution of *Sesamum indicum* L. *PLoS ONE* **2013**, *8*, e80508. [[CrossRef](#)] [[PubMed](#)]
45. Wu, Y.F.; Zheng, Y.G.; Liu, X.L.; Dai, X.Y.; Li, S.X.; Xu, H. The complete chloroplast genome sequence of *Machilus robusta* W. W. Smith (Lauraceae) from Jiangxi Province, China. *Mitochondrial DNA Part B* **2021**, *6*, 1890–1892. [[CrossRef](#)] [[PubMed](#)]
46. Xiaoxuan, L.; Shi, W.B.; Song, W.C.; Han, W.Q.; Yang, G.W.; Wang, S. Complete chloroplast genome sequence and annotation of *Machilus salicina* Hance, 1885 (Lauraceae). *Mitochondrial DNA Part B* **2022**, *7*, 780–782. [[CrossRef](#)] [[PubMed](#)]
47. Yang, J.-B.; Yang, S.-X.; Li, H.T.; Yang, J.; Li, D.-Z. Comparative Chloroplast Genomes of *Camellia* Species. *PLoS ONE* **2013**, *8*, e73053. [[CrossRef](#)]
48. Raman, G.; Park, V.; Kwak, M.; Lee, B.; Park, S. Characterization of the complete chloroplast genome of *Arabis stellari* and comparisons with related species. *PLoS ONE* **2017**, *12*, e0183197. [[CrossRef](#)]
49. Raman, G.; Choi, K.S.; Park, S. Phylogenetic Relationships of the Fern *Cyrtomium falcatum* (Dryopteridaceae) from Dokdo Island Based on Chloroplast Genome Sequencing. *Genes* **2016**, *7*, 115. [[CrossRef](#)]
50. Shen, J.; Li, X.; Chen, X.; Huang, X.; Jin, S. The Complete Chloroplast Genome of *Carya cathayensis* and Phylogenetic Analysis. *Genes* **2022**, *13*, 369. [[CrossRef](#)]
51. Yu, A.Y.H.; Houry, W.A. ClpP: A distinctive family of cylindrical energy-dependent serine proteases. *FEBS Lett.* **2007**, *581*, 3749–3757. [[CrossRef](#)] [[PubMed](#)]
52. Boudreau, E.; Takahashi, Y.; Lemieux, C.; Turmel, M.; Rochaix, J. The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. *EMBO J.* **1997**, *16*, 6095–6104. [[CrossRef](#)] [[PubMed](#)]
53. Naver, H.; Boudreau, E.; Rochaix, J.-D. Functional Studies of Ycf3: Its Role in Assembly of Photosystem I and Interactions with Some of Its Subunits. *Plant Cell* **2001**, *13*, 2731. [[CrossRef](#)] [[PubMed](#)]
54. Ahmed, I.; Matthews, P.J.; Biggs, P.J.; Naeem, M.; McLenachan, P.A.; Lockhart, P.J. Data from: Identification of chloroplast genome loci suitable for high-resolution phylogeographic studies of *Colocasia esculenta* (L.) Schott (Araceae) and closely related taxa. *Mol. Ecol. Resour.* **2013**, *13*, 929–937. [[CrossRef](#)] [[PubMed](#)]
55. Cavender-Bares, J.; González-Rodríguez, A.; Eaton, D.A.R.; Hipp, A.A.L.; Beulke, A.; Manos, P.S. Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Mol. Ecol.* **2015**, *24*, 3668–3687. [[CrossRef](#)]
56. Biju, V.C.; Shidhi, P.R.; Vijayan, S.; Rajan, V.S.; Sasi, A.; Janardhanan, A.; Nair, A.S. The Complete Chloroplast Genome of *Trichopus zeylanicus*, and Phylogenetic Analysis with Dioscoreales. *Plant Genome* **2019**, *12*, 190032. [[CrossRef](#)]
57. Kuang, D.-Y.; Wu, H.; Wang, Y.-L.; Gao, L.-M.; Zhang, S.-Z.; Lu, L. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): Implication for DNA barcoding and population genetics. *Genome* **2011**, *54*, 663–673. [[CrossRef](#)]
58. Jeon, J.-H.; Kim, S.-C. Comparative Analysis of the Complete Chloroplast Genome Sequences of Three Closely Related East-Asian Wild Roses (*Rosa* sect. *Synstylae*; Rosaceae). *Genes* **2019**, *10*, 23. [[CrossRef](#)]
59. Abdullah, S.I.; Mehmood, F.; Ali, Z.; Malik, M.S.; Waseem, S.; Mirza, B.; Ahmed, I.; Waheed, M.T. Comparative analyses of chloroplast genomes among three Firmiana species: Identification of mutational hotspots and phylogenetic relationship with other species of *Malvaceae*. *Plant Gene* **2019**, *19*, 100199. [[CrossRef](#)]
60. Shen, X.; Wu, M.; Liao, B.; Liu, Z.; Bai, R.; Xiao, S.; Li, X.; Zhang, B.; Xu, J.; Chen, S. Complete Chloroplast Genome Sequence and Phylogenetic Analysis of the Medicinal Plant *Artemisia annua*. *Molecules* **2017**, *22*, 1330. [[CrossRef](#)]
61. Daniell, H.; Lin, C.-S.; Yu, M.; Chang, W.-J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [[CrossRef](#)] [[PubMed](#)]
62. Hershberg, R.; Petrov, D.A. Selection on Codon Bias. *Annu. Rev. Genet.* **2008**, *42*, 287–299. [[CrossRef](#)] [[PubMed](#)]
63. Li, X.; Li, Y.; Zang, M.; Li, M.; Fang, Y. Complete Chloroplast Genome Sequence and Phylogenetic Analysis of *Quercus acutissima*. *Int. J. Mol. Sci.* **2018**, *19*, 2443. [[CrossRef](#)] [[PubMed](#)]
64. Jian, H.-Y.; Zhang, Y.-H.; Yan, H.-J.; Qiu, X.-Q.; Wang, Q.-G.; Li, S.-B.; Zhang, S.-D. The Complete Chloroplast Genome of a Key Ancestor of Modern Roses, *Rosa chinensis* var. *spontanea*, and a Comparison with Congeneric Species. *Molecules* **2018**, *23*, 389. [[CrossRef](#)]
65. Dugas, D.V.; Hernandez, D.; Koenen, E.J.; Schwarz, E.; Straub, S.; Hughes, C.E.; Jansen, R.K.; Nageswara-Rao, M.; Staats, M.; Trujillo, J.T.; et al. Mimosoid legume plastome evolution: IR expansion, tandem repeat expansions and accelerated rate of evolution in clpP. *Sci. Rep.* **2015**, *5*, 16958. [[CrossRef](#)]
66. Drescher, A.; Ruf, S.; Calsa, T., Jr.; Carrer, H.; Bock, R. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *Plant J.* **2000**, *22*, 97–104. [[CrossRef](#)]
67. Zhu, B.; Qian, F.; Hou, Y.; Yang, W.; Cai, M.; Wu, X. Complete chloroplast genome features and phylogenetic analysis of *Eruca sativa* (Brassicaceae). *PLoS ONE* **2021**, *16*, e0248556. [[CrossRef](#)]
68. Li, B.; Zheng, Y. Dynamic evolution and phylogenomic analysis of the chloroplast genome in Schisandraceae. *Sci. Rep.* **2018**, *8*, 9285. [[CrossRef](#)]

69. Zhu, B.; Gan, C.C.; Wang, H.C. Characteristics of the *Dendrobium thyrsiflorum* and its phylogenetic relationship analysis. *Biotechnol. Bull.* **2021**, *37*, 38–47.
70. Kim, K.J.; Lee, H.L. Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Res.* **2004**, *11*, 247–261. [[CrossRef](#)]
71. Zhang, J.-B.; Li, R.-Q.; Xiang, X.-G.; Manchester, S.R.; Lin, L.; Wang, W.; Wen, J.; Chen, Z.-D. Integrated Fossil and Molecular Data Reveal the Biogeographic Diversification of the Eastern Asian-Eastern North American Disjunct Hickory Genus (*Carya* Nutt.). *PLoS ONE* **2013**, *8*, e70449. [[CrossRef](#)] [[PubMed](#)]
72. Song, W.; Chen, Z.; Shi, W.; Han, W.; Feng, Q.; Shi, C.; Engel, M.S.; Wang, S. Comparative Analysis of Complete Chloroplast Genomes of Nine Species of *Litsea* (Lauraceae): Hypervariable Regions, Positive Selection, and Phylogenetic Relationships. *Genes* **2022**, *13*, 1550. [[CrossRef](#)] [[PubMed](#)]
73. 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](#)] [[PubMed](#)]
74. Hilu, K.W.; Liang, G. The matK gene: Sequence variation and application in plant systematics. *Am. J. Bot.* **1997**, *84*, 830–839. [[CrossRef](#)]
75. Wu, Z.; Liao, R.; Yang, T.; Dong, X.; Lan, D.; Qin, R.; Liu, H. Analysis of six chloroplast genomes provides insight into the evolution of *Chrysosplenium* (Saxifragaceae). *nBMC Genom.* **2020**, *21*, 621. [[CrossRef](#)]
76. Niu, Y.; Gao, C.; Liu, J. Comparative analysis of the complete plastid genomes of *Mangifera* species and gene transfer between plastid and mitochondrial genomes. *PeerJ* **2021**, *9*, e10774. [[CrossRef](#)]
77. Chen, C.H.; Zheng, Y.J.; Liu, S.A.; Zhong, Y.D.; Wu, Y.; Li, J.; Xu, L.-A.; Xu, M. The complete chloroplast genome of *Cinnamomum camphora* and its comparison with related Lauraceae species. *PeerJ* **2017**, *5*, e3820. [[CrossRef](#)]
78. Zhang, Y.-J.; Ma, P.-F.; Li, D.-Z. High-Throughput Sequencing of Six Bamboo Chloroplast Genomes: Phylogenetic Implications for Temperate Woody Bamboos (Poaceae: Bambusoideae). *PLoS ONE* **2011**, *6*, e20596. [[CrossRef](#)]
79. Spalik, K.; Downie, S.R. Intercontinental disjunctions in *Cryptotaenia* (Apiaceae, Oenantheae): An appraisal using molecular data. *J. Biogeogr.* **2007**, *34*, 2039–2054. [[CrossRef](#)]
80. Du, Y.-P.; Bi, Y.; Yang, F.-P.; Zhang, M.-F.; Chen, X.-Q.; Xue, J.; Zhang, X.-H. Complete chloroplast genome sequences of *Lilium*: Insights into evolutionary dynamics and phylogenetic analyses. *Sci. Rep.* **2017**, *7*, 5751. [[CrossRef](#)]
81. Buermans, H.P.J.; Den Dunnen, J.T. Next generation sequencing technology: Advances and applications. *Biochim. Biophys. Acta (BBA)-Mol. Basis Dis.* **2014**, *1842*, 1932–1941. [[CrossRef](#)] [[PubMed](#)]
82. Kong, D.; Ma, C.; Zhang, Q.; Li, L.; Chen, X.; Zeng, H.; Guo, D. Leading dimensions in absorptive root trait variation across 96 subtropical forest species. *New Phytol.* **2014**, *203*, 863–872. [[CrossRef](#)] [[PubMed](#)]