![forests logo] *forests*

MDPI

Article

Complete Chloroplast Genome of *Michelia shiluensis* and a Comparative Analysis with Four Magnoliaceae Species

Yanwen Deng [1,2] ![ORCID], **Yiyang Luo** [1,2], **Yu He** [1,2] ![ORCID], **Xinsheng Qin** [2], **Chonggao Li** [3] and **Xiaomei Deng** [1,2,*] ![ORCID]

1   Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, Guangzhou 510642, China
2   College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou 510642, China
3   Guangzhou Hengsheng Construction Engineering Co., Ltd., Guangzhou 510000, China
*   Correspondence: xmdeng_scau@163.com; Tel.: +86-137-1116-8219

![check for updates]

Abstract: *Michelia shiluensis* is a rare and endangered magnolia species found in South China. This species produces beautiful flowers and is thus widely used in landscape gardening. Additionally, its timber is also used for furniture production. As a result of low rates of natural reproduction and increasing levels of human impact, wild *M. shiluensis* populations have become fragmented. This species is now classified as endangered by the IUCN. In the present study, we characterized the complete chloroplast genome of *M. shiluensis* and found it to be 160,075 bp in length with two inverted repeat regions (26,587 bp each), a large single-copy region (88,105 bp), and a small copy region (18,796 bp). The genome contained 131 genes, including 86 protein-coding genes, 37 tRNAs, and 8 rRNAs. The guanine-cytosine content represented 39.26% of the overall genome. Comparative analysis revealed high similarity between the *M. shiluensis* chloroplast genome and those of four closely related species: *Michelia odora*, *Magnolia laevifolia*, *Magnolia insignis*, and *Magnolia cathcartii*. Phylogenetic analysis shows that *M. shiluensis* is most closely related to *M. odora*. The genomic information presented in this study is valuable for further classification, phylogenetic studies, and to support ongoing conservation efforts.

Keywords: Hainan Province; endemic species; conservation; codon usage; sequence divergence; phylogeny

1. Introduction

Michelia shiluensis Chun and Y. F. Wu (Magnoliaceae) is an endangered flowering plant that is sparsely distributed throughout Hainan Province, China [1]. It is characterized by leafy branches and beautiful flowers, and is, therefore, widely used in landscape gardening [2]. This species is also a source of excellent quality wood which is in demand for furniture production [3]. In recent decades, there has been a serious decline in wild populations of this species as a result of the illegal harvesting to supply both the timber and horticultural markets [4]. Moreover, this species naturally has a low seeding rate and its wild populations are declining [5]. Consequently, *M. shiluensis* is categorized as a Class II National Key Protected Species in China [6] and is considered endangered (EN) by the International Union for Conservation of Nature [7]. Currently, most studies on *M. shiluensis* have focused on its use in landscape gardening and its protection in China [5]; however, there remains a lack of evolutionary and phylogenetic research.

The chloroplast is an important organelle in plants with its own genome (hereafter, cp genome) and participates in photosynthesis and other functions [8]. The cp genome of most land plants has a circular structure, including four segments: A large single-copy (LSC), a small single-copy (SSC), and two invert repeats (IRs) [9]. Although the cp genome is generally conserved, it has undergone intra- and inter-species rearrangement during evolution [10,11], including IR expansion and contraction. The information obtained from sequence rearrangements can be applied in phylogenetic analyses to solve taxonomic problems, such as low-level classifications, using genome comparison [12–17]. In the section *Michelia*, complete cp genomes have been reported for only *Magnolia alba* (NC037005), *Magnolia laevifolia* (NC035956), and *Michelia odora* (NC023239). Therefore, analysis of the cp genomes of other *Michelia* plants is necessary because of the similarity of morphology among Magnoliaceae species [18].

In the present study, we characterized the cp genome of *M. shiluensis* and compared its sequence features with four closely related species (*M. odora*, *M. laevifolia*, *Magnolia insignis*, and *Magnolia cathcartii*). The phylogenetic relationships among 28 Magnoliaceae species were constructed based on 79 protein-coding gene (PCG) sequences and show that *M. shiluensis* is most closely related to *M. odora*.

2. Materials and Methods

2.1. Plant Material and DNA Extraction

Fresh leaves of *M. shiluensis* were collected in the South China Botanical Garden (113°21′ E, 23°10′ N), China and transported to the laboratory at the South China Agricultural University. Total genomic DNA was isolated from the leaves using the CTAB method [19].

2.2. Genome Sequencing and Annotation

An Illumina shotgun library was established according to the manufacturer's protocol, and high-throughput sequencing was conducted using the HiSeq X TEN platform (Illumina, San Diego, CA, USA). After filtration using SOAPnuke [20], 4.93 GB of clean data were generated. Filtered reads were assembled de novo using SPAdes (version 3.10.1) [21] by referencing them against the cp genome sequence of *M. odora* (NC037005.1) using BLAST v2.2.30 (National Center for Biotechnology Information, Bethesda, MD, USA). Gene annotation was performed using GeSeq [22]. The cp genome map was generated using Organellar Genome DRAW (version 1.2) [23]. The annotated sequence was submitted to GenBank (accession number MN418056).

2.3. Sequence and Repeat Analysis

We used the Editseq v7.1.0 [24] software to calculate the guanine-cytosine (GC) content. MEGA v7.0.26 [25] was used to generate the relative synonymous codon usage (RSCU) values based on 79 PCGs. RNA editing sites in PCGs were predicted using the PREP suite [26] with the default settings.

The REPuter [27] online service was used to identify repeats (forward, reverse, complement, and palindromic) in the cp genome with default parameters. Chloroplast simple sequence repeats (cpSSRs) were identified using MISA-web [28] with minimal repeat numbers of 8, 5, 4, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats, respectively.

2.4. Genome Comparison and Sequence Divergence

Comparisons between five Magnoliaceae cp genomes were visualized using online mVISTA software [29] with the annotation of *M. shiluensis* as the reference in Shuffle-LAGAN mode. The borders of four different regions among the five cp genomes of Magnoliaceae were visualized using IRscope [30]. The nucleotide diversity (Pi), the rate of nonsynonymous (dN) substitutions, the rate of synonymous (dS) substitutions were determined using DNAsp v6.12.03 [31] to investigate the nucleotide diversity of sequences and genes that are considered to be under selection pressure.

2.5. Phylogenetic Analysis

To research the phylogenetic relationships and allow for comparisons among Magnoliaceae species, a maximum likelihood tree was constructed using RAxML [32], with 1000 bootstrap replicates, based on the PCG sequences found in 28 Magnoliaceae cp genomes. All 28 Magnoliaceae cp genome sequences were downloaded from the NCBI nucleotide database.

3. Result

3.1. Structures and Features of M. shiluensis Chloroplast Genome

The complete cp genome of *M. shiluensis* was 160,075 bp in length and comprised two IR regions of 26,587 bp each, separated by an LSC region of 88,105 bp, and an SSC region of 18,796 bp. The cp genome had the following base proportions: Adenine (A), 29.99%; thymine (T), 30.75%; cytosine (C), 19.98%; and guanine (G), 19.28%. Therefore, the total GC content was 39.26%. The GC content of the LSC, SSC, and IR regions were 37.95%, 34.28%, and 43.20%, respectively (Table 1).

Table 1. Summary of the chloroplast genomes of *Michelia shiluensis* and four closely related species (*Michelia odora*, *Magnolia laevifolia*, *Magnolia insignis*, and *Magnolia cathcartii*).

	M. shiluensis	*M. odora*	*M. laevifolia*	*M. insignis*	*M. cathcartii*
Accession	MN418056	NC023239	NC035956	NC035657	NC023234
			Genome		
Length (bp)	160,075	160,070	160,120	160,117	159,950
GC (%)	39.26	39.26	39.24	39.24	39.22
			LSC		
length (bp)	88,105	88,098	88,145	88,195	88,142
GC (%)	37.95	37.95	37.9	37.92	37.91
Length (%)	55.04	55.04	55.05	55.08	55.11
			SSC		
length (bp)	18,796	18,800	18,799	18,782	18,790
GC (%)	34.28	34.28	34.32	34.25	34.13
Length (%)	11.74	11.74	11.74	11.73	11.75
			IR		
length (bp)	26,587	26,586	26,588	26,570	26,509
GC (%)	43.2	43.2	43.2	43.19	43.2
Length (%)	16.61	16.61	16.61	16.59	16.57
			No. of Genes (duplicated in IR)		
Genes	131(18)	131(18)	131(18)	131(18)	131(18)
PCGs	86(7)	86(7)	86(7)	86(7)	86(7)
tRNA	37(7)	37(7)	37(7)	37(7)	37(7)
rRNA	8(4)	8(4)	8(4)	8(4)	8(4)
With introns	16(5)	16(5)	16(5)	16(5)	16(5)

Notes: GC: Guanine-cytosine; LSC: Large single-copy; IR: Invert repeat; SSC: Small single-copy; PCG: Protein-coding gene.

The cp genome of *M. shiluensis* contained 113 unique genes, including 79 PCGs, 30 tRNAs, and four rRNAs (Table 1, Figure 1). A total of 58 genes were found to be involved in self-replication, 12 genes encoded small ribosomal subunit proteins, eight genes encoded large ribosomal subunit proteins, 30 genes encoded tRNA, and four genes encoded RNA polymerase subunits. A total of 44 genes were found to be involved in photosynthesis, including six genes for ATP synthase, 11 genes for NADH dehydrogenase, six genes for the cytochrome b/f complex, five genes for photosystem I, 15 genes for photosystem II, and one gene for the large chain of Rubisco. In total, 18 genes were duplicated in the cp

genome of *M. shiluensis*, including seven PCGs, seven tRNA genes, and four rRNA genes, all of which were located in the IR region (Table 2). None of the genes contained stop codons in coding sequences, therefore, no pseudogenes were detected.
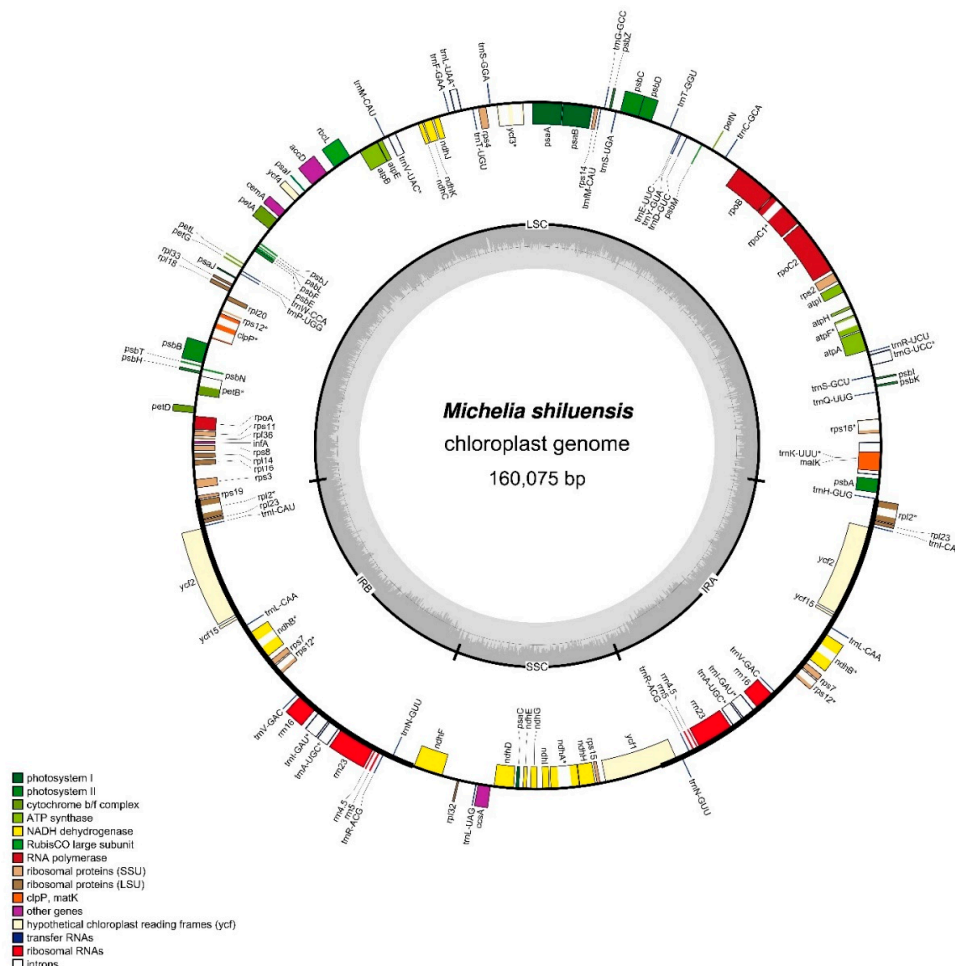

Figure 1. Gene map of the *Michelia shiluensis* chloroplast genome. Genes on the outside of the circle are transcribed counter-clockwise, while genes on the inside are transcribed clockwise. Different colors represent different kinds of functional genes. The guanine-cytosine content is indicated by darker gray and the adenine-thymine content is indicated by light gray.

A total of 16 genes were found to have introns, including 10 PCGs and six tRNA genes. Of these genes, *clpP*, *trnA-UGC*, *trnI-GAU*, and *ycf3* had two introns, whereas *atpF*, *ndhA*, *ndhB*, *petB*, *rpl2*, *rpoC1*, *rps12*, *rps16*, *trnG-UCC*, *trnK-UUU*, *trnL-UAA*, and *trnV-UAC* had one intron. The *rps12* gene which encodes the 40S ribosomal protein S12, was trans-spliced, with one exon located in the LSC region, and the other two exons located in the IR region. The largest intron was located in the *trnK* gene (2490 bp) with the *matK* gene inside; *trnL-UAA* had the smallest intron (491 bp) (Table 3).

We compared the basic cp genome features of *M. shiluensis* with four Magnoliaceae species. The cp genome lengths of *M. laevifolia* and *M. insignis* were 45 and 42 bp longer than that of *M. shiluensis*, respectively, while the cp genome lengths of *M. odora* and *M. cathcartii* were five and 125 bp shorter, respectively. Compared with *M. shiluensis*, the variation in the lengths of the LSC, SSC, and IR regions ranged from 7 to 90, 3 to 14, and 1 to 78 bp, respectively. In addition, the GC content of the whole genome and of each region of *M. shiluensis* were highly similar to those of the other four species. Moreover, there was no variation with respect to the total number of genes, PCGs, tRNA genes, rRNA genes, and genes with introns (Table 1).

Table 2. List of the annotated genes in the *Michelia shiluensis* chloroplast genome.

Category of Genes	Subcategory of Genes	Gene Names			
Self-replication	**rRNA genes**	*rrn 4.5* *	*rrn5* *	*rrn16* *	*rrn23* *
	tRNA genes	*trnA-UGC* *	*trnC-ACA*	*trnD-GUC*	*trnE-UUC*
		trnF-GAA	*trnfM-CAU*	*trnG-GCC*	*trnG-UCC*
		trnH-GUG	*trnI-CAU*	*trnI-GAU* *	*trnK-UUU*
		trnL-CAA *	*trnL-UAA*	*trnL-UAG*	*trnM-CAU* *
		trnN-GUU *	*trnP-UGG*	*trnQ-UUG*	*trnR-ACG* *
		trnR-UCU	*trnS-GCU*	*trnS-GGA*	*trnS-UGA*
		trnT-GGU	*trnT-UGU*	*trnV-GAC* *	*trnV-UAC*
		trnW-CCA	*trnY-GUA*		
	Small subunit of ribosome	*rps2*	*rps3*	*rps4*	*rps7* *
		rps8	*rps11*	*rps12* *	*rps14*
		rps15	*rps16*	*rps18*	*rps19*
	Large subunit of ribosome	*rpl2* *	*rpl14*	*rpl16*	*rpl20*
		rpl23 *	*rpl32*	*rpl33*	*rpl36*
	RNA polymerase subunits	*rpoA*	*rpoB*	*rpoC1*	*rpoC2*
Photosynthesis	ATP synthase gene	*atpA*	*atpB*	*atpE*	*atpF*
		atpH	*atpI*		
	NADH dehydrogenase	*ndhA*	*ndhB* *	*ndhC*	*ndhD*
		ndhE	*ndhF*	*ndhG*	*ndhH*
		ndhI	*ndhJ*	*ndhK*	
	Cytochrome b/f complex	*petA*	*petB*	*petD*	*petG*
		petL	*petN*		
	Photosystem I	*psaA*	*psaB*	*psaC*	*psaI*
		psaJ			
	Photosystem II	*psbA*	*psbB*	*psbC*	*psbD*
		psbE	*psbF*	*psbH*	*psbI*
		psbJ	*psbK*	*psbL*	*psbM*
		psbN	*psbT*	*psbZ*	
	Large chain of Rubisco	*rbcL*			
Other genes	ATP-dependent protease	*clpP*			
	Cytochrome c biogenesis	*ccsA*			
	Acetyl-CoA carboxylase	*accD*			
	Translation initiation factor IF-1	*infA*			
	Membrane protein	*cemA*			
	Maturase	*matK*			
Unknown function	Hypothetical chloroplast reading frame	*ycf1*	*ycf2* *	*ycf3*	*ycf4*
		ycf15 *			

Note: "*" indicates duplicated genes.

Table 3. Characteristics of the genes that contain introns in the cp genome of *Michelia shiluensis*.

Gene	Location	Exon I (bp)	Intron I (bp)	Exon II (bp)	Intron II (bp)	Extron III (bp)
trnK-UUU	LSC	35	2490	37		
rps16	LSC	217	825	44		
trnG-UCC	LSC	24	768	48		
atpF	LSC	411	706	144		
rpoC1	LSC	1624	722	434		
ycf3	LSC	154	729	227	739	126
trnL-UAA	LSC	35	491	50		
trnV-UAC	LSC	37	584	56	565	39
clpP	LSC	246	630	291	781	69
petB	LSC	5	786	641		
rpl2	IR	432	658	387		
rps12	IR/LSC	114	-	25	536	232
ndhB	IR	756	700	777		
trnI-GAU	IR	42	936	35		
trnA-UGC	IR	38	799	35		
ndhA	SSC	541	1078	551		

3.2. Codon Usage and RNA Analysis

Based on the PCGs, 22,791 codons were detected (excluding the stop codons) (Table 4). The three most abundant amino acids were leucine (2423 codons), isoleucine (2085 codons), and serine (1719 codons),

and the three least abundant amino acids were cysteine (314 codons), tryptophan (427 codons), and methionine (602 codons) (Figure S1). Of the 30 most frequent codons (RSCU > 1), most of them end with A or U, and only the UUG codon ends with G. In contrast, most of the 32 least frequent codons (RSCU < 1) end with C or G. In addition, two codons, AUG and UGG, have no codon bias (RSCU = 1).

PREP suite was used to edit predictions in the genome of *M. shiluensis* by manipulating the first codon position of the first nucleotide (Table S1). A total of 106 RNA editing sites were detected from the PCGs in *M. shiluensis*; with the majority of the amino acid conversions involving the conversion of serine to leucine. Most of the RNA editing sites were located on the *ndhB* gene (14 sites), followed by *ndhD* (11 sites), and *matK* (nine sites). Most of the conversions changed from a polar group to a nonpolar group, while only two sites changed from a nonpolar group to a polar group (proline to serine); one of these was located on the *psbE* gene while the other was located on the *ccsA* gene.

Table 4. Relative synonymous codon usage (RSCU) in the chloroplast genome of *Michelia shiluensis*.

Codon	Amino Acid	Count	RSCU	Codon	Amino Acid	Count	RSCU
UUU	Phe	714	1.1	UAU	Tyr	600	1.47
UUC	Phe	586	0.9	UAC	Tyr	214	0.53
UUA	Leu	671	1.66	UAA	*	147	0.98
UUG	Leu	510	1.26	UAG	*	137	0.91
CUU	Leu	439	1.09	CAU	His	429	1.41
CUC	Leu	206	0.51	CAC	His	178	0.59
CUA	Leu	400	0.99	CAA	Gln	560	1.41
CUG	Leu	197	0.49	CAG	Gln	237	0.59
AUU	Ile	893	1.28	AAU	Asn	712	1.44
AUC	Ile	545	0.78	AAC	Asn	278	0.56
AUA	Ile	647	0.93	AAA	Lys	777	1.37
AUG	Met	602	1	AAG	Lys	359	0.63
GUU	Val	474	1.37	GAU	Asp	631	1.54
GUC	Val	202	0.58	GAC	Asp	189	0.46
GUA	Val	487	1.41	GAA	Glu	780	1.42
GUG	Val	223	0.64	GAG	Glu	318	0.58
UCU	Ser	457	1.6	UGU	Cys	214	1.36
UCC	Ser	266	0.93	UGC	Cys	100	0.64
UCA	Ser	373	1.3	UGA	*	166	1.11
UCG	Ser	192	0.67	UGG	Trp	427	1
CCU	Pro	345	1.41	CGU	Arg	256	1.15
CCC	Pro	213	0.87	CGC	Arg	72	0.32
CCA	Pro	307	1.25	CGA	Arg	289	1.3
CCG	Pro	117	0.48	CGG	Arg	129	0.58
ACU	Thr	426	1.49	AGU	Ser	324	1.13
ACC	Thr	237	0.83	AGC	Ser	107	0.37
ACA	Thr	358	1.26	AGA	Arg	410	1.84
ACG	Thr	120	0.42	AGG	Arg	179	0.8
GCU	Ala	537	1.82	GGU	Gly	542	1.33
GCC	Ala	185	0.63	GGC	Gly	171	0.42
GCA	Ala	335	1.13	GGA	Gly	628	1.54
GCG	Ala	126	0.43	GGG	Gly	291	0.71

Note: "*" indicates the stop codon.

3.3. Repeat Sequence Analysis

The REPuter results show that the *M. shiluensis* cp genome contains a total of 49 repeats: 23 palindromic, 18 forward, and eight reverse repeats (Figure 2). The repeat size ranged from 18 to 33 bp. The most abundant repeats were 18 bp (12 sites) followed by 20 bp (10 sites) (Figure S2).
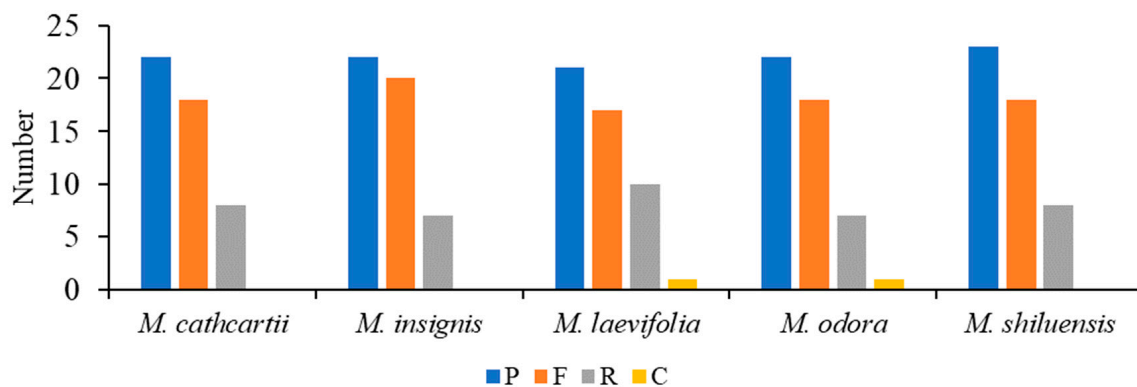
Figure 2. Comparison of repeats among five Magnoliaceae species: *Michelia shiluensis*, *Michelia odora*, *Magnolia laevifolia*, *Magnolia insignis*, and *Magnolia cathcartii*. (F: Forward; R: Reverse; C: Complement; and P: Palindromic).

In the first location, 46.9% of repeats were detected in the intergenic spacers (IGSs), while 34.7% were in the PCGs, and 18.4% were in the tRNA genes. Of all the PCGs, the *ycf2* gene had five forward repeats and four palindromic repeats and was the gene with the most repeats (Table S2). Comparison of the repeat types with the other four species revealed no substantial variation among the five species (Figure 2). *Michelia shiluensis* had the highest frequency of palindromic repeats (23), while *M. laevifolia* had the lowest (21). *Magnolia cathcartii*, *M. odora*, and *M. shiluensis* had the same number of forward repeats (18), while *M. shiluensis* and *M. cathcartii* had eight reverse repeats. In addition, only one complement repeat was found in the genomes of *M. laevifolia* and *M. odora*, whereas no complements were identified in the cp genomes of *M. shiluensis* and *M. insignis*. A total of 141 cpSSRs were found in the cp genome of *M. shiluensis* (Table S3). The majority of them were mononucleotide repeats (118), followed by tetranucleotide repeats (9), and dinucleotide repeats (8) (Figure 3). No pentanucleotide repeats were detected in the cp genome of *M. shiluensis*. The longest repeat was 18 bp while the shortest was 8 bp. Noncoding regions, including IGSs (97) and introns (19), contained most of the SSRs while only 25 repeats were located in coding regions, including *cemA*, *ndhD*, *ndhF*, *psbC*, *rpoB*, *rpoC1*, *rpoC2*, *rps19*, *rps3*, *ycf1*, *ycf2*, and *ycf4* (Table 5). The cpSSRs were mainly distributed in the LSC region (72.34%), followed by the SSC region (17.73%), with just 4.96% in the IR. The cpSSRs in *M. shiluensis* had base bias towards A-T bases. In total, 113 SSRs had A or T bases, accounting for 80.14% of the total SSRs (Figure 3). Comparison among the five species of Magnoliaceae show high similarity in SSR type and distribution. The variation in the total amount, mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats among the five species was 5, 4, 0, 2, 0, 1, and 1, respectively (Figure 3). The number of SSRs in the IR were the same among the five species while the counts in different locations and regions were highly conserved.

Table 5. Distribution of single sequence repeats in different locations and regions among five Magnoliaceae species: *Michelia shiluensis*, *Michelia odora*, *Magnolia laevifolia*, *Magnolia insignis*, and *Magnolia cathcartii*.

Species	Number	Location			Regions		
		LSC	IR	SSC	CDS	Intron	IGS
M. cathcartii	143	103	7	26	25	18	100
M. insignis	141	102	7	25	26	15	100
M. laevifolia	141	101	7	26	27	18	96
M. odora	138	99	7	25	24	19	95
M. shiluensis	141	102	7	25	25	19	97

Notes: LSC: Large single copy; IR: Invert region; SSC: Small single copy; CDS: Coding sequence; IGS: Intergenic spacer.
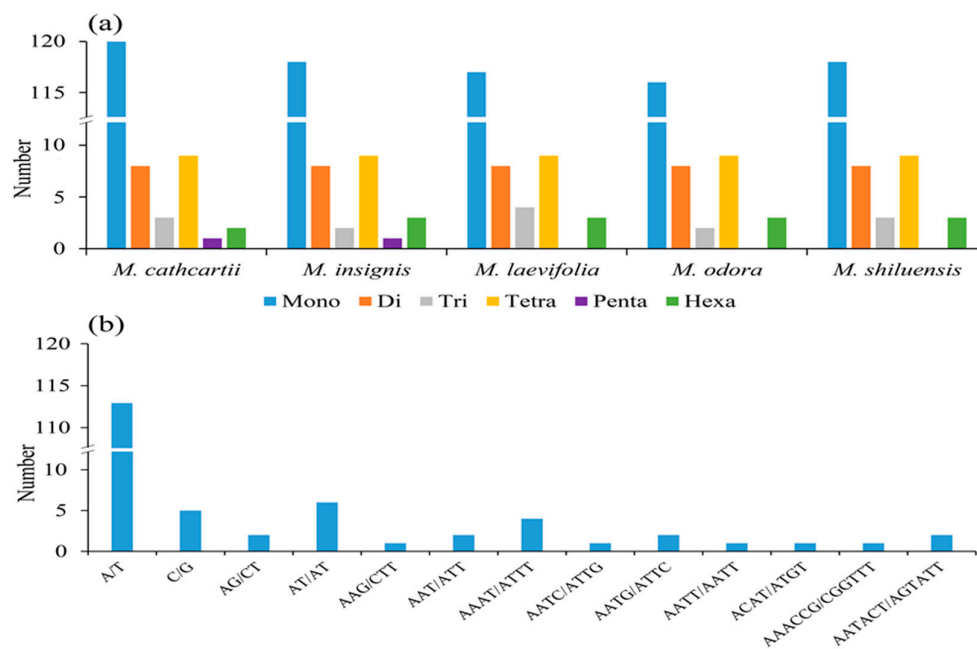
Figure 3. Single sequence repeats (SSRs) in the chloroplast genome of *Michelia shiluensis*. (**a**) Comparison of the SSRs among five Magnoliaceae species (*M. shiluensis*, *Michelia odora*, *Magnolia laevifolia*, *Magnolia insignis*, and *Magnolia cathcartii*); (**b**) base composition of SSRs in the cp genome of *M. shiluensis*.

3.4. Genome Comparison and Sequence Divergence

The mVISTA online software was used to compare the variation in the whole cp genome among the five species (Figure 4). The alignments indicated that the whole cp genome of the five species was highly conserved, especially in the IR region. The noncoding sequences had relatively more divergence than the coding sequences. The noncoding sequences that contained high levels of divergence were *rps16-trnQ*, *atpH-atpI*, *trnT-psbD*, *petA-psbJ*, and *ndhF-trnL*. In the coding sequences, only *ycf1* show relatively more variation than the other genes. No obvious insertions were found among the five species.

The four junctions in the regions of the cp genomes of the five species were shown using IRscope (Figure 5). There was a conserved structure on each border, and slight distance differences among the five species. Gene *rps19* was fully located in the LSC at a distance of 1–6 bp from the LSC/IRb border, while gene *rpl2* was fully located in the IRb. The *ndhF* gene was found in the SSC region and was 61 bp away from the IRb/SSC border in *M. odora*, *M. laevifolia*, and *M. insignis*, while it was 21 bp longer in *M. shiluensis*, and 7 bp shorter in *M. cathcartii*. The SSC/IRa border was inside the *ycf1* gene in all five species. Compared to *M. shiluensis* and *M. odora*, the part of the *ycf1* gene in the SSC region of *M. laevifolia* and *M. insignis* was almost 20 bp longer, and this resulted in the differences in gene length. However, the *ycf1* gene of *M. cathcartii* was almost 30 bp shorter on both sides of the SSC/IRa border; thus, the *ycf1* gene of *M. cathcartii* was almost 60 bp shorter than those of the other four species. The distance from the *trnH* gene to the IRa/LSC border was 11 bp in *M. shiluensis*, *M. odora*, and *M. laevifolia*, while it was 16 bp in *M. insignis* and 9 bp in *M. cathcartii*. Due to the short length in the IR region, the whole length of the *M. cathcartii* cp genome was significantly shorter than those of the other four species.

To detect the selective pressures on the PCGs in the *M. shiluensis* cp genome, the rate of nonsynonymous (dN) substitutions, the rate of synonymous (dS) substitutions, and their ratio (dN/dS) were calculated based on the 79 PCG sequences of the five Magnoliaceae species (Figure 6). Only four genes had a dN/dS ratio greater than 1 (*accD* in *M. insignis* vs. *M. cathcartii*, score 1.14; *ndhD* in *M. shiluensis* vs. *M. cathcartii*, score 1.29; *ndhF* in *M. odora* vs. *M. cathcartii*, score 1.89; and *rpoC2* in *M. laevifolia* vs. *M. cathcartii*, score 2.50), which indicates that most genes are under the influence of negative selection, while only a few genes are under the influence of positive selection.
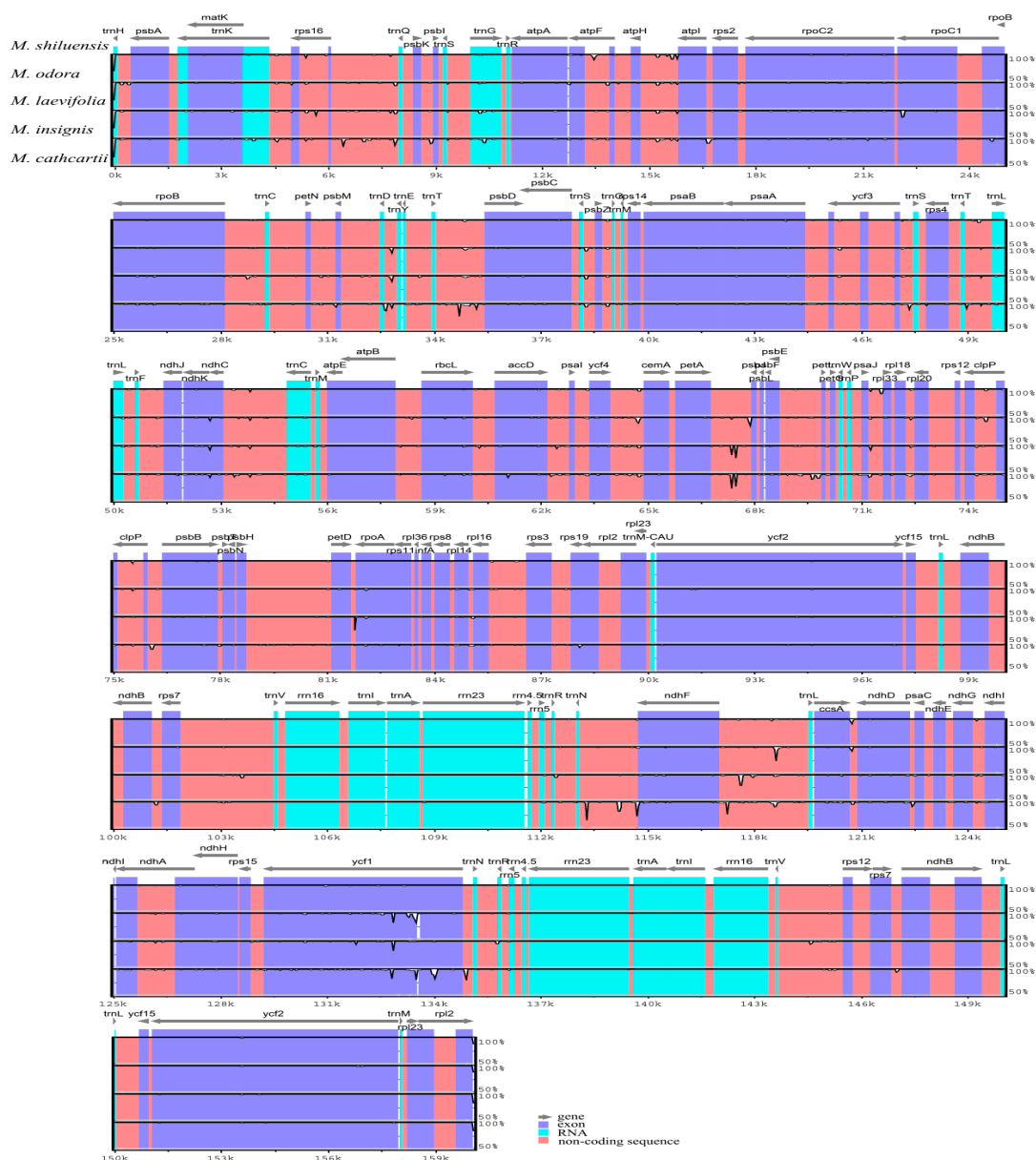
Figure 4. Sequence alignment of five whole chloroplast genomes in Magnoliaceae (*Michelia shiluensis*, *Michelia odora*, *Magnolia laevifolia*, *Magnolia insignis*, and *Magnolia cathcartii*) using *M. shiluensis* as a reference in *mVISTA*.
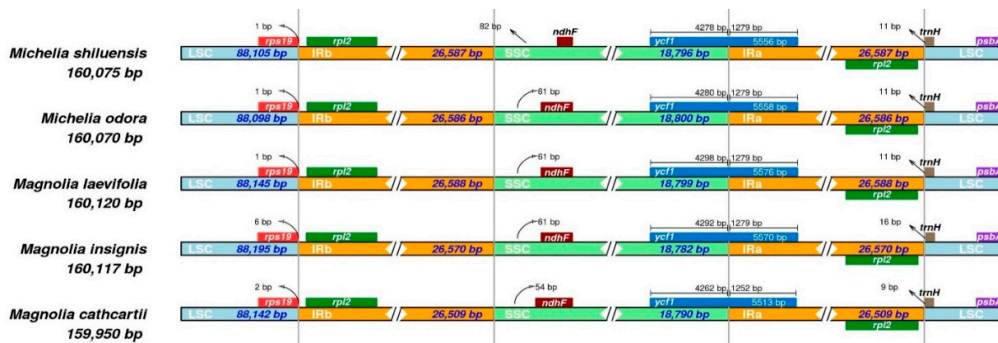

Figure 5. The four junctions of the regions in the chloroplast genomes of the five *Magnoliaceae* species, determined using *IRscope*.
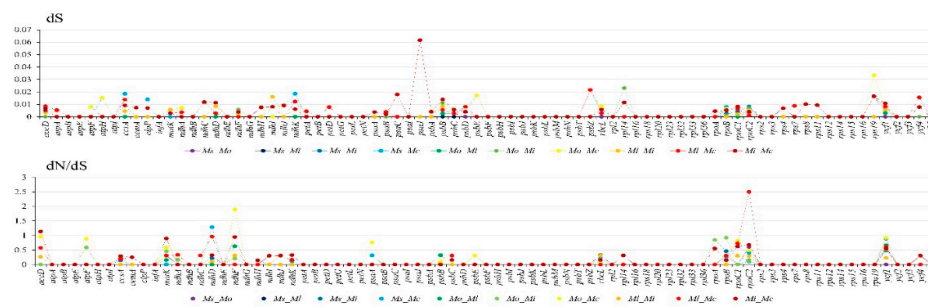
Figure 6. The synonymous (dS) and nonsynonymous substitutions (dN)/dS ratio values of 79 protein-coding genes from five Magnoliaceae chloroplast genomes (Ms: *Michelia shiluensis*; Mo: *Michelia odora*; Ml: *Magnolia laevifolia*; Mi: *Magnolia insignis*; Mc: *Magnolia cathcartii*).

In the coding region, the mean Pi in the PCGs was 0.00117 (ranging from 0 to 0.00606); and the mean values of Pi in the LSC, IR, and SSC regions were 0.001192, 0.000186, and 0.001634, respectively (Figure 7). Meanwhile, in the IGSs, the mean Pi value was 0.00295 (ranging from 0 to 0.02416); and the mean Pi value in the LSC, IR, and SSC regions were 0.0297, 0.00045, and 0.00731, respectively. This result indicates that the Pi value in the coding region is lower than that in the IGSs. The results also demonstrate that the IR region is the most conserved region among the five species, followed by the LSC and SSC regions. In total, 20 mutation sites (Pi > 0.005) were identified, including 19 sites in IGSs and one site in a PCG. The mutation sites in IGSs were as follows: *trnH-psbA*, *psbK-psbI*, *atpA-atpF*, *rps2-rpoC2*, *trnT-psbD*, *ycf3-trnS*, *ndhJ-ndhK*, *ndhK-ndhC*, *accD-psaI*, *psbL-psbF*, *petL-petG*, *trnW-trnP*, *trnP-psaJ*, *rpl18-rpl20*, *ndhF-trnL*, *ccsA-ndhD*, *ndhD-psaC*, *ndhG-ndhI*, and *ndhI-ndhA*. One gene, *psaJ*, was unique and had a Pi value greater than 0.005.
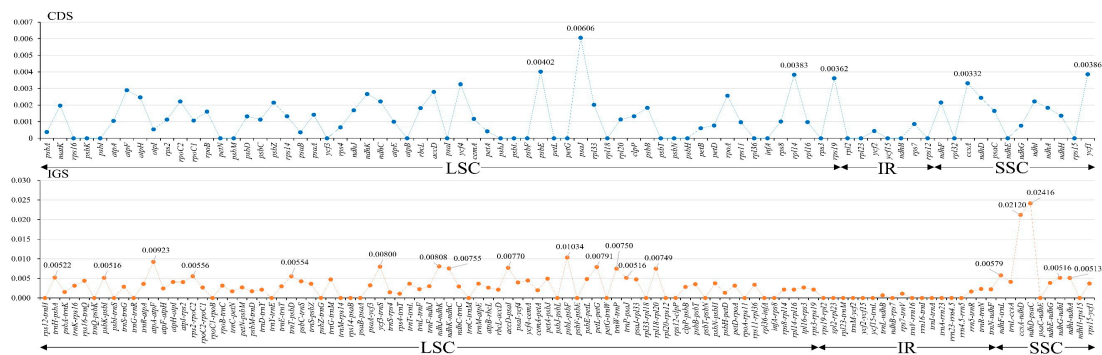

Figure 7. Nucleotide diversity (Pi) in the chloroplast genome of five Magnoliaceae species (*Michelia shiluensis*; *Michelia odora*; *Magnolia laevifolia*; *Magnolia insignis*; and *Magnolia cathcartii*).

3.5. Phylogenetic Analysis

To reveal the evolutionary relationships between the investigated species and to enable comparison with traditional phylogenies, a maximum likelihood phylogenetic tree was constructed using RAxML (with 1000 bootstrap replicates) based on the PCG sequences found in 28 Magnoliaceae cp genomes (Figure 8). The phylogenetic tree generated 25 nodes; most of which had 100% bootstrap support. This result strongly supports the notion that *M. shiluensis* is most closely related to *M. odora*.
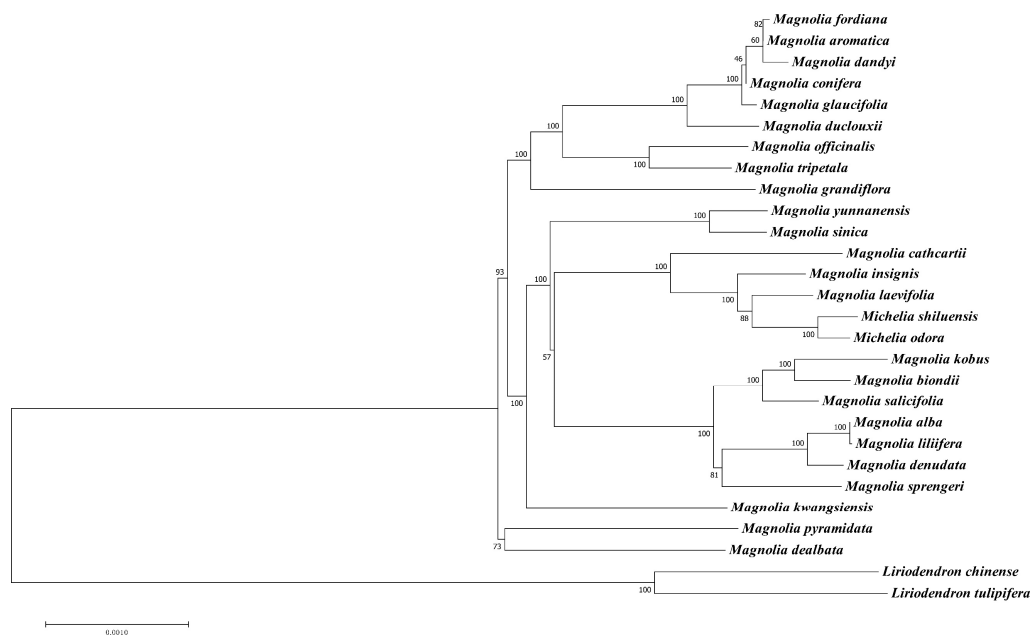
Figure 8. Maximum likelihood tree with 1000 bootstrap replicates constructed using RAxML based on chloroplast genomes of 28 Magnoliaceae species (26 species of *Magnolia* and *Michelia*, and two species of *Liriodendron* as outgroups). Bootstrap values (%) are shown above branches. Accession numbers: *Magnolia alba* NC037005, *Magnolia liliflora* NC023238, *Magnolia denudata* NC018357, *Magnolia sprengeri* NC023242, *Magnolia salicifolia* NC023240, *Magnolia biondii* KY085894, *Magnolia kobus* NC023237, *Michelia odora* NC023239, *Michelia shiluensis* MN418056, *Magnolia laevifolia* NC035956, *Magnolia insignis* NC035657, *Magnolia cathcartii* NC023234, *Magnolia yunnanensis* NC024545, *Magnolia sinica* NC023241, *Magnolia kwangsiensis* NC015892, *Magnolia conifera* NC037001, *Magnolia dandyi* NC037004, *Magnolia aromatica* NC037000, *Magnolia fordiana* MF990562, *Magnolia glaucifolia* NC037003, *Magnolia duclouxii* NC037002, *Magnolia tripetala* NC024027, *Magnolia officinalis* NC020316, *Magnolia grandiflora* NC020318, *Magnolia pyramidata* NC023236, *Magnolia dealbata* NC023235, *Liriodendron chinense* NC030504, and *Liriodendron tulipifera* NC008326.

4. Discussion

In this study, we characterized the complete cp genome of *M. shiluensis*, an endangered and valuable species (Figure 1). By comparing five closely related species, we found that gene content, order, structure, and other features were highly conserved among them (Figures 3–5). *Michelia shiluensis* was shown to be most closely related to *M. odora* (Figure 8). This finding can help to further our understanding of the characterization of the *M. shiluensis* cp genome and reveal information concerning the evolution, population genetics, and phylogeny of this species.

Normally, the length of the cp genome in higher plants is in the range of about 120–160 kb, with a stable structure and conserved sequence [8,33]. The cp genome of *M. shiluensis* displayed a typical quadripartite structure, with an LSC and an SSC which were separated by two IR regions (Figure 1). The whole length of this genome was 160,075 bp, with 39.26% GC content, and containing 113 unique genes and 16 genes with one or two introns (Tables 1–3). Among the 26 Magnoliaceae species, the length of the cp genome ranged from 158,177 to 160,183 bp, the GC content ranged from 39.15% to 39.30%, and they collectively contained 112 common genes, including 79 PCGs, 29 tRNA, and four rRNA genes; also, one or two introns were found among these 16 genes. The results for the *M. shiluensis* cp genome were consistent with a previous analysis of 26 Magnoliaceae species [34], except for the number of genes, one additional tRNA gene (*trnV-GAC*) was detected in *M. shiluensis*. Similar to other angiosperms, a high GC content was detected in the IR region of *M. shiluensis*, which may be a result of the existence of high-GC rRNA sequences [9,35–37]. Introns play a vital role in selective gene splicing [38]. However, introns have been lost among some species during their evolution [39,40].

In this study, no introns were lost in the cp genome of *M. shiluensis* during evolution, which reflects the fact that the cp genome is highly conserved in Magnoliaceae [34].

In total, 22,791 codons were found in the cp genome of *M. shiluensis* (Table 4), among which, the codons for leucine were the most abundant (10.63%). This result has also been observed in *Ailanthus altissima* [41] and *Justicia flava* [42]. Among the preferred codons (RSCU > 1), we found that most of them ended with A or U, except UUG. This is not unique to the *M. shiluensis* cp genome as similar findings have been observed in *Papaver rhoeas* and *P. orientale* [36], *Ageratina adenophora* [43], and *Oryza sativa* [44]. Of the PCGs in the *M. shiluensis* cp genome, 106 possible sites for RNA editing were detected (Table S1). The majority of the amino acid conversion was from serine to leucine, and the *ndhB* gene accounted for a high number of editable sites (14 of the total 106 sites). Similar results have been obtained for *Forsythia suspensa* [45] and *Sanionia uncinata* [46].

Repeat sequences play an important role in genomic structural variation, expansion, and rearrangement [8,40]. Previous research has indicated that most of the repeat sequences are located in the IGS regions followed by the coding regions [47,48]. A similar result was found in this study, with 46.9% of repeats detected in the IGS regions, followed by 34.7% in the coding regions, and the remainder in the tRNAs (Table S2). The cpSSR is an effective marker [49,50] that is widely used in population genetics, biogeographic studies, and phylogenetic evaluation [51,52]. In the cp genome of *M. shiluensis*, over 80% of the SSRs consisted of A or T bases, and over 80% were mononucleotide repeats. Similar results have been observed in other studies [48,50,53]. The majority of SSRs are found in the SSC and LSC regions [50] and, in this respect, *M. shiluensis* is no exception (Table 5). These two regions accounted for 90.07% of the SSRs, and only seven SSRs were found in the IR region.

Although the cp genome of angiosperms is relatively conserved in structure and size [54,55], the expansion and contraction of the IR region, as caused by evolutionary events, has resulted in minor changes in the IR boundary and size of the genome [39,56], thus increasing the chloroplast genetic diversity of angiosperms [57,58]. In this study, comparative analysis of five Magnoliaceae species revealed that the IR lengths were similar in *M. shiluensis, M. odora*, and *M. laevifolia* (Figure 5). However, the IR region of *M. insignis* had completely lost 11 bp in the *rps12-trnV*, *rrn23-rrn4.5* IGSs, while *M. cathcartii* had lost 5 bp in the *rpl2* intron, 6 bp in the *rps12* intron, 26 bp in *ycf1*, and 41 bp in the *trnN-ndhF* IGS. The losses of these bases resulted in differences in the lengths of the IR regions among the five species.

DNA barcoding is a technique that is widely applied in plant identification studies [59,60]. However, only a few regions have been used for the DNA barcoding of Magnoliaceae [61–63]. We used mVISTA to compare the genomes of five Magnoliaceae species and revealed that the IR region is more highly conserved than the LSC and SSC regions, and that the coding region was more highly conserved than the noncoding region (Figure 4), consistent with other angiosperms [9,38]. Five regions in the *M. shiluensis* cp genome had high levels of variation (four on IGSs and one on a PCG). The Pi value was also investigated among the 79 PCGs and 125 IGSs (Figure 7), and only 20 regions were found to have a Pi value greater than 0.005; which confirmed the low base substitution rate in Magnoliaceae [64]. Regions with a high degree of variation can be used to develop high resolution DNA barcoding for identification.

Due to the high morphological similarity among Magnoliaceae species [18,65], there have been some difficulties with respect to the classification of the family. Thus, the classification of Magnoliaceae has always been controversial [66–72]. The cp genome contains sufficient information and has been shown to be more effective than cpDNA fragments for clarifying low level phylogenetic relationships in plants [53,73]. In this study, the phylogenetic results of 28 Magnoliaceae plants based on PCG sequences revealed that *M. shiluensis* is most closely related to *M. odora* (Figure 8), which is consistent with phylogenetic results based on the *ndhF* sequence [70]. According to traditional morphological classification, *M. insignis* has been placed in the subgenera *Manglietia*, and *M. alba* has been placed in the section *Michelia* [69,74]. However, the phylogenetic relationship results based on the cp genome show that *M. insignis* is located in the section *Michelia* clade and *M. alba* is located in the subgenera

Yulania clade. This result differs from that of traditional morphological classification [69,74] and the results of three nuclear gene sequences [75]. These findings confirm that not even a complete cp genome can distinguish species in young evolutionary lineages, and that phylogenetic conclusions may require consideration of certain features in the nuclear genome [76].

5. Conclusions

The complete cp genome provided by this study can be used for in-depth genetic research on *M. shiluensis* and Magnoliaceae species in general, and may also play an important role in the development of new conservation and management strategies to ultimately aid species conservation efforts.

References

1. Francisco-Ortega, J.; Wang, F.-G.; Wang, Z.-S.; Xing, F.-W.; Liu, H.; Xu, H.; Xu, W.-X.; Luo, Y.-B.; Song, X.-Q.; Gale, S. Endemic seed plant species from Hainan Island: A checklist. *Bot. Rev.* **2010**, *76*, 295–345. [CrossRef]

2. Bao, S.; Zhou, S.; Yu, Y. Comparative anatomy of the leaves in *Michelia*. *Guangxi Zhi Wu* **2002**, *22*, 140–144.

3. Xiong, J.; Wang, L.-J.; Qian, J.; Wang, P.-P.; Wang, X.-J.; Ma, G.-L.; Zeng, H.; Li, J.; Hu, J.-F. Structurally diverse sesquiterpenoids from the endangered ornamental plant *Michelia shiluensis*. *J. Nat. Prod.* **2018**, *81*, 2195–2204. [CrossRef]

4. Chen, B. Forest wild biospecies diversity and its preservation in China. *Chin. J. Ecol.* **1993**, *12*, 39–43. [CrossRef]

5. Wei, Y.; Hong, F.; Yuan, L.; Kong, Y.; Shi, Y. Population distribution and age structure characteristics of *Michelia shiluensis*, an endangered and endemic species in Hainan Island. *Chin. J. Trop. Crops* **2017**, *38*, 2280–2284. [CrossRef]

6. State Forestry Administration of China. List of national key protected wild plants (first batch). *State Counc. Bull.* **1999**, *13*, 6.

7. Bhandari, M. International Union for Conservation of Nature. In *The Wiley-Blackwell Encyclopedia of Globalization*; Ritzer, G., Ed.; John Wiley & Sons: Hoboken, NJ, USA, 2012. [CrossRef]

8. Wicke, S.; Schneeweiss, G.M.; dePamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [CrossRef] [PubMed]

9. Chen, Y.; Hu, N.; Wu, H. Analyzing and characterizing the chloroplast genome of *Salix wilsonii*. *BioMed Res. Int.* **2019**, *2019*, 5190425. [CrossRef] [PubMed]

10. Tangphatsornruang, S.; Uthaipaisanwong, P.; Sangsrakru, D.; Chanprasert, J.; Yoocha, T.; Jomchai, N.; Tragoonrung, S. Characterization of the complete chloroplast genome of *Hevea brasiliensis* reveals genome rearrangement, RNA editing sites and phylogenetic relationships. *Gene* **2011**, *475*, 104–112. [CrossRef] [PubMed]

11. Walker, J.F.; Jansen, R.K.; Zanis, M.J.; Emery, N.C. Sources of inversion variation in the small single copy (SSC) region of chloroplast genomes. *Am. J. Bot.* **2015**, *102*, 1751–1752. [CrossRef]

12. Johansson, J.T. There large inversions in the chloroplast genomes and one loss of the chloroplast generps16 suggest an early evolutionary split in the genusAdonis (Ranunculaceae). *Plant Syst. Evol.* **1999**, *218*, 133–143. [CrossRef]

13. Lee, H.-L.; Jansen, R.K.; Chumley, T.W.; Kim, K.-J. Gene relocations within chloroplast genomes of Jasminum and Menodora (Oleaceae) are due to multiple, overlapping inversions. *Mol. Biol. Evol.* **2007**, *24*, 1161–1180. [CrossRef] [PubMed]

14. Jansen, R.K.; Wojciechowski, M.F.; Sanniyasi, E.; Lee, S.-B.; Daniell, H. Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of rps12 and clpP intron losses among legumes (Leguminosae). *Mol. Phylogenet. Evol.* **2008**, *48*, 1204–1217. [CrossRef] [PubMed]

15. Shaw, J.; Lickey, E.B.; Schilling, E.E.; Small, R.L. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *Am. J. Bot.* **2007**, *94*, 275–288. [CrossRef]

16. Mardanov, A.V.; Ravin, N.V.; Kuznetsov, B.B.; Samigullin, T.H.; Antonov, A.S.; Kolganova, T.V.; Skyabin, K.G. Complete sequence of the duckweed (*Lemna minor*) chloroplast genome: Structural organization and phylogenetic relationships to other angiosperms. *J. Mol. Evol.* **2008**, *66*, 555–564. [CrossRef]

17. Park, I.; Kim, W.-J.; Yang, S.; Yeo, S.-M.; Li, H.; Moon, B.C. The complete chloroplast genome sequence of *Aconitum coreanum* and *Aconitum carmichaelii* and comparative analysis with other *Aconitum* species. *PLoS ONE* **2017**, *12*, e0184257. [CrossRef]

18. Li, J.; Conran, J. Phylogenetic relationships in Magnoliaceae subfam. Magnolioideae: A morphological cladistic analysis. *Plant Syst. Evol.* **2003**, *242*, 33–47. [CrossRef]

19. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.

20. Chen, Y.; Shi, C.; Huang, Z.; Zhang, Y.; Li, S.; Li, Y.; Ye, J.; Yu, C.; Li, Z. SOAPnuke: A MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* **2017**, *7*, gix120. [CrossRef]

21. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef]

22. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq-versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [CrossRef] [PubMed]

23. Lohse, M.; Drechsel, O.; Kahlau, S.; Bock, R. OrganellarGenomeDRAW—A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **2013**, *41*, W575–W581. [CrossRef] [PubMed]

24. Burland, T.G. DNASTAR's Lasergene sequence analysis software. In *Bioinformatics Methods and Protocols*; Misener, S., Krawetz, A.K., Eds.; Springer: Basel, Switzerland, 2000; pp. 71–91.

25. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]

26. Mower, J.P. The PREP suite: Predictive RNA editors for plant mitochondrial genes, chloroplast genes and user-defined alignments. *Nucleic Acids Res.* **2009**, *37*, W253–W259. [CrossRef]

27. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [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. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, W273–W279. [CrossRef]

30. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [CrossRef]

31. 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]

32. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]

33. Palmer, J.D. Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* **1985**, *19*, 325–354. [CrossRef] [PubMed]

34. Shen, Y.; Chen, K.; Gu, C.; Zheng, S.; Ma, L. Comparative and phylogenetic analyses of 26 Magnoliaceae species based on complete chloroplast genome sequences. *Can. J. For. Res.* **2018**, *48*, 1456–1469. [CrossRef]

35. Asaf, S.; Khan, A.L.; Khan, A.R.; Waqas, M.; Kang, S.-M.; Khan, M.A.; Lee, S.-M.; Lee, I.-J. Complete chloroplast genome of *Nicotiana otophora* and its comparison with related species. *Front. Plant Sci.* **2016**, *7*, 843. [CrossRef]

36. Zhou, J.; Cui, Y.; Chen, X.; Li, Y.; Xu, Z.; Duan, B.; Li, Y.; Song, J.; Yao, H. Complete chloroplast genomes of *Papaver rhoeas* and *Papaver orientale*: Molecular structures, comparative analysis, and phylogenetic analysis. *Molecules* **2018**, *23*, 437. [CrossRef] [PubMed]

37. Meng, J.; Li, X.; Li, H.; Yang, J.; Wang, H.; He, J. Comparative analysis of the complete chloroplast genomes of four Aconitum Medicinal Species. *Molecules* **2018**, *23*, 1015. [CrossRef] [PubMed]

38. Zhong, Q.; Yang, S.; Sun, X.; Wang, L.; Li, Y. The complete chloroplast genome of the Jerusalem artichoke (*Helianthus tuberosus* L.) and an adaptive evolutionary analysis of the ycf2 gene. *PeerJ* **2019**, *7*, e7596. [CrossRef]

39. He, L.; Qian, J.; Li, X.; Sun, Z.; Xu, X.; Chen, S. Complete chloroplast genome of medicinal plant *Lonicera japonica*: Genome rearrangement, intron gain and loss, and implications for phylogenetic studies. *Molecules* **2017**, *22*, 249. [CrossRef]

40. Maréchal, A.; Brisson, N. Recombination and the maintenance of plant organelle genome stability. *New Phytol.* **2010**, *186*, 299–317. [CrossRef]

41. Saina, J.; Li, Z.-Z.; Gichira, A.; Liao, Y.-Y. The complete chloroplast genome sequence of tree of heaven (*Ailanthus altissima* (Mill.) (Sapindales: Simaroubaceae), an important pantropical tree. *Int. J. Mol. Sci.* **2018**, *19*, 929. [CrossRef]

42. Yaradua, S.S.; Alzahrani, D.A.; Albokhary, E.J.; Abba, A.; Bello, A. Complete chloroplast genome sequence of *Justicia flava*: Genome comparative analysis and phylogenetic relationships among Acanthaceae. *BioMed Res. Int.* **2019**, *2019*, 4370258. [CrossRef]

43. Nie, X.; Lv, S.; Zhang, Y.; Du, X.; Wang, L.; Biradar, S.S.; Tan, X.; Wan, F.; Weining, S. Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*). *PLoS ONE* **2012**, *7*, e36869. [CrossRef]

44. Liu, Q.; Xue, Q. Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. *J. Genet.* **2005**, *84*, 55–62. [CrossRef] [PubMed]

45. Wang, W.; Yu, H.; Wang, J.; Lei, W.; Gao, J.; Qiu, X.; Wang, J. The complete chloroplast genome sequences of the medicinal plant *Forsythia suspensa* (Oleaceae). *Int. J. Mol. Sci.* **2017**, *18*, 2288. [CrossRef] [PubMed]

46. Park, M.; Park, H.; Lee, H.; Lee, B.-h.; Lee, J. The complete plastome sequence of an Antarctic bryophyte *Sanionia uncinata* (Hedw.) Loeske. *Int. J. Mol. Sci.* **2018**, *19*, 709. [CrossRef] [PubMed]

47. 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]

48. Zhou, T.; Chen, C.; Wei, Y.; Chang, Y.; Bai, G.; Li, Z.; Kanwal, N.; Zhao, G. Comparative transcriptome and chloroplast genome analyses of two related *Dipteronia* species. *Front. Plant Sci.* **2016**, *7*, 1512. [CrossRef]

49. Provan, J.; Powell, W.; Hollingsworth, P.M. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. *Trends Ecol. Evol.* **2001**, *16*, 142–147. [CrossRef]

50. Ebert, D.; Peakall, R. Chloroplast simple sequence repeats (cpSSRs): Technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Mol. Ecol. Res.* **2009**, *9*, 673–690. [CrossRef]

51. Mohammad-Panah, N.; Shabanian, N.; Khadivi, A.; Rahmani, M.-S.; Emami, A. Genetic structure of gall oak (*Quercus infectoria*) characterized by nuclear and chloroplast SSR markers. *Tree Genet. Genomes* **2017**, *13*, 70. [CrossRef]

52. Zeng, J.; Chen, X.; Wu, X.F.; Jiao, F.C.; Xiao, B.G.; Li, Y.P.; Tong, Z.J. Genetic diversity analysis of genus *Nicotiana* based on SSR markers in chloroplast genome and mitochondria genome. *Acta Tabacaria Sin.* **2016**, *22*, 89–97. [CrossRef]

53. Dong, W.; Xu, C.; Li, W.; Xie, X.; Lu, Y.; Liu, Y.; Jin, X.; Suo, Z. Phylogenetic resolution in *Juglans* based on complete chloroplast genomes and nuclear DNA sequences. *Front. Plant Sci.* **2017**, *8*, 1148. [CrossRef] [PubMed]

54. Liu, L.-X.; Li, R.; Worth, J.R.; Li, X.; Li, P.; Cameron, K.M.; Fu, C.-X. The complete chloroplast genome of Chinese bayberry (*Morella rubra*, Myricaceae): Implications for understanding the evolution of Fagales. *Front. Plant Sci.* **2017**, *8*, 968. [CrossRef] [PubMed]

55. Lu, R.-S.; Li, P.; Qiu, Y.-X. The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: Comparative genomic and phylogenetic analyses. *Front. Plant Sci.* **2017**, *7*, 2054. [CrossRef] [PubMed]

56. Raubeson, L.A.; Peery, R.; Chumley, T.W.; Dziubek, C.; Fourcade, H.M.; Boore, J.L.; Jansen, R.K. Comparative chloroplast genomics: Analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genom.* **2007**, *8*, 174. [CrossRef]

57. 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]

58. Huang, H.; Shi, C.; Liu, Y.; Mao, S.-Y.; Gao, L.-Z. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: Genome structure and phylogenetic relationships. *BMC Evol. Biol.* **2014**, *14*, 151. [CrossRef]

59. Zhang, W.; Fan, X.; Zhu, S.; Zhao, H.; Fu, L. Species-specific identification from incomplete sampling: Applying DNA barcodes to monitoring invasive *Solanum* plants. *PLoS ONE* **2013**, *8*, e55927. [CrossRef]

60. Dick, C.W.; Webb, C.O. Plant DNA barcodes, taxonomic management, and species discovery in tropical forests. In *DNA Barcodes. Methods in Molecular Biology (Methods and Protocols)*; Kress, W., Webb, C.O., Eds.; Humana Press: Totowa, NJ, USA, 2012; Volume 858, pp. 379–393.

61. Yu, H.; Wu, K.; Song, J.; Zhu, Y.; Yao, H.; Luo, K.; Dai, Y.; Xu, S.; Lin, Y. Expedient identification of Magnoliaceae species by DNA barcoding. *Plant Omics* **2014**, *7*, 47.

62. Huan, H.V.; Trang, H.M.; Toan, N.V. Identification of DNA barcode sequence and genetic relationship among some species of *Magnolia* Family. *Asian J. Plant Sci.* **2018**, *17*, 56–64. [CrossRef]

63. Huan, H.V.; Nguyen, L.T.T.; Quang, N.M. To create DNA barcode data of *Magnolia chevalieri* (Dandy) V.S. Kumar for identification species and researching genetic diversity. *J. For. Sci. Technol.* **2019**, *18*, 3–9.

64. Azuma, H.; Thien, L.B.; Kawano, S. Molecular phylogeny of *Magnolia* (Magnoliaceae) inferred from cpDNA sequences and evolutionary divergence of the floral scents. *J. Plant Res.* **1999**, *112*, 291–306. [CrossRef]

65. Bentley, R.A. *Manual of Botany: Including the Structure, Classification, Properties, Uses, and Functions of Plants*; J. & A. Churchill: London, UK, 1882.

66. Azuma, H.; García-Franco, J.G.; Rico-Gray, V.; Thien, L.B. Molecular phylogeny of the Magnoliaceae: The biogeography of tropical and temperate disjunctions. *Am. J. Bot.* **2001**, *88*, 2275–2285. [CrossRef] [PubMed]

67. Dandy, J.E. A revised survey of the genus Magnolia together with *Manglietia* and *Michelia*. In *Magnolias*; Treseder, N.G., Ed.; Faber & Faber: London, UK, 1978; pp. 27–37.

68. Liang, C.B.; Nooteboom, H.P. Notes on Magnoliaceae III: The Magnoliaceae of China. *Ann. Mo. Bot. Gard.* **1993**, *80*, 999–1104. [CrossRef]

69. Figlar, R.B.; Nooteboom, H.P. Notes on Magnoliaceae IV. *Blumea* **2004**, *49*, 87–100. [CrossRef]

70. Kim, S.; Park, C.W.; Kim, Y.D.; Suh, Y. Phylogenetic relationships in family Magnoliaceae inferred from ndhF sequences. *Am. J. Bot.* **2001**, *88*, 717–728. [CrossRef] [PubMed]

71. Kim, S.; Nooteboom, H.P.; Park, C.-W.; Suh, Y. Taxonomic revision of *Magnolia* section *Maingola* (Magnoliaceae). *Blumea* **2002**, *47*, 319–339.

72. Nooteboom, H. Different looks at the classification of the Magnoliaceae. In Proceedings of the International Symposium on the Family Magnoliaceae, Guangzhou, China, 18–22 May 1998; Science Press: Beijing, China, 2000; pp. 26–37.

73. Zhao, M.-L.; Song, Y.; Ni, J.; Yao, X.; Tan, Y.-H.; Xu, Z.-F. Comparative chloroplast genomics and phylogenetics of nine *Lindera* species (Lauraceae). *Sci. Rep.* **2018**, *8*, 8844. [CrossRef]

74. Noteboom, H. Notes on Magnoliaceae with a revision of *Pachylarnax* and *Elmerrillia* and the Malasian species of *Manglietia* and *Michelia*. *Blumea* **1985**, *31*, 65–121.

75. Nie, Z.-L.; Wen, J.; Azuma, H.; Qiu, Y.-L.; Sun, H.; Meng, Y.; Sun, W.-B.; Zimmer, E.A. Phylogenetic and biogeographic complexity of Magnoliaceae in the Northern Hemisphere inferred from three nuclear data sets. *Mol. Phylogenet. Evol.* **2008**, *48*, 1027–1040. [CrossRef]

76. Ruhsam, M.; Rai, H.S.; Mathews, S.; Ross, T.G.; Graham, S.W.; Raubeson, L.A.; Mei, W.; Thomas, P.I.; Gardner, M.F.; Ennos, R.A. Does complete plastid genome sequencing improve species discrimination and phylogenetic resolution in *Araucaria*? *Mol. Ecol. Resour.* **2015**, *15*, 1067–1078. [CrossRef]