diversity

MDPI

*Article*

# Complete Chloroplast Genome of *Cnidium monnieri* (Apiaceae) and Comparisons with Other Tribe Selineae Species

**Ting Ren, Xueyimu Aou, Rongming Tian, Zhenbing Li, Chang Peng and Xingjin He ***

Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China; renting@stu.scu.edu.cn (T.R.); aouxueyimu@stu.scu.edu.cn (X.A.); tianrongming@stu.scu.edu.cn (R.T.); lizhenbing@stu.scu.edu.cn (Z.L.); 2017322040008@stu.scu.edu.cn (C.P.)
* Correspondence: xjhe@scu.edu.cn

**Abstract:** *Cnidium monnieri* is an economically important traditional Chinese medicinal plant. In this study, the complete chloroplast (cp) genome of *C. monnieri* was determined using the Illumina paired-end sequencing, the GetOrganelle de novo assembly strategy, as well as the GeSeq annotation method. Our results showed that the cp genome was 147,371 bp in length with 37.4% GC content and included a large single-copy region (94,361 bp) and a small single-copy region (17,552 bp) separated by a pair of inverted repeat regions (17,729 bp). A total of 129 genes were contained in the cp genome, including 85 protein-coding genes, 36 tRNA genes, and eight rRNA genes. We also investigated codon usage, RNA editing, repeat sequences, simple sequence repeats (SSRs), IR boundaries, and pairwise Ka/Ks ratios. Four hypervariable regions (*trnD-trnY-trnE-trnT*, *ycf2*, *ndhF-rpl32-trnL*, and *ycf1*) were identified as candidate molecular markers for species authentication. The phylogenetic analyses supported non-monophyly of *Cnidium* and *C. monnieri* located in tribe Selineae based on the cp genome sequences and internal transcribed spacer (ITS) sequences. The incongruence of the phylogenetic position of *C. monnieri* between ITS and cpDNA phylogenies suggested that *C. monnieri* might have experienced complex evolutions with hybrid and incomplete lineage sorting. All in all, the results presented herein will provide plentiful chloroplast genomic resources for studies of the taxonomy, phylogeny, and species authentication of *C. monnieri*. Our study is also conducive to elucidating the phylogenetic relationships and taxonomic position of *Cnidium*.

**Keywords:** *Cnidium monnieri*; chloroplast genome; comparative analysis; tribe Selineae; phylogenetic relationship



**Citation:** Ren, T.; Aou, X.; Tian, R.; Li, Z.; Peng, C.; He, X. Complete Chloroplast Genome of *Cnidium monnieri* (Apiaceae) and Comparisons with Other Tribe Selineae Species. *Diversity* **2022**, *14*, 323. https://doi.org/10.3390/d14050323

Academic Editor: Mario A. Pagnotta

Received: 15 March 2022
Accepted: 20 April 2022
Published: 21 April 2022



## 1. Introduction

*Cnidium monnieri* (L.) Cuss, an annual herb in Apiaceae with excellent medicinal and economic values, is mainly grown in the fields, roadsides, grasslands, and riverside wetlands of China [1]. The mature and dry fruit of *C. monnieri* is the traditional Chinese medicine "She Chuang Zi" (Figure 1) and is generally used for killing parasites, anti-itch, wind-expelling, and removing dampness [1,2]. Owing to the good medicinal value, previous studies about *C. monnieri* largely focused on their pharmacological effect, chemical extraction, and herb authentication [3–6]. In addition, *C. monnieri* is the type species of the genus *Cnidium*. Downie et al. [7] found *Cnidium* is not monophyletic and identified two clades (tribe Selineae and *Sinodielsia* Clade) within *Cnidium* based on nuclear ribosomal DNA internal transcribed spacer (ITS), which indicated that the taxonomy and phylogeny of this genus needs to be further studied. The determination of the phylogenetic position of type species can pave the way for resolving the phylogeny and taxonomy of *Cnidium*. However, there are few studies on the molecular phylogeny of *C. monnieri*, and these studies only involve this species instead of concentrating on it [7,8]. Therefore, it is particularly necessary to analyze the phylogeny of *C. monnieri*. Recently, with the increasing use of medicinal plants, species identification is especially important. However, Apiaceae plants

are famous for their difficulty to identify, because of the morphological similarity. Thus, it is urged to develop more discriminating molecular markers for species authentication of this species to assure medicinal quality.

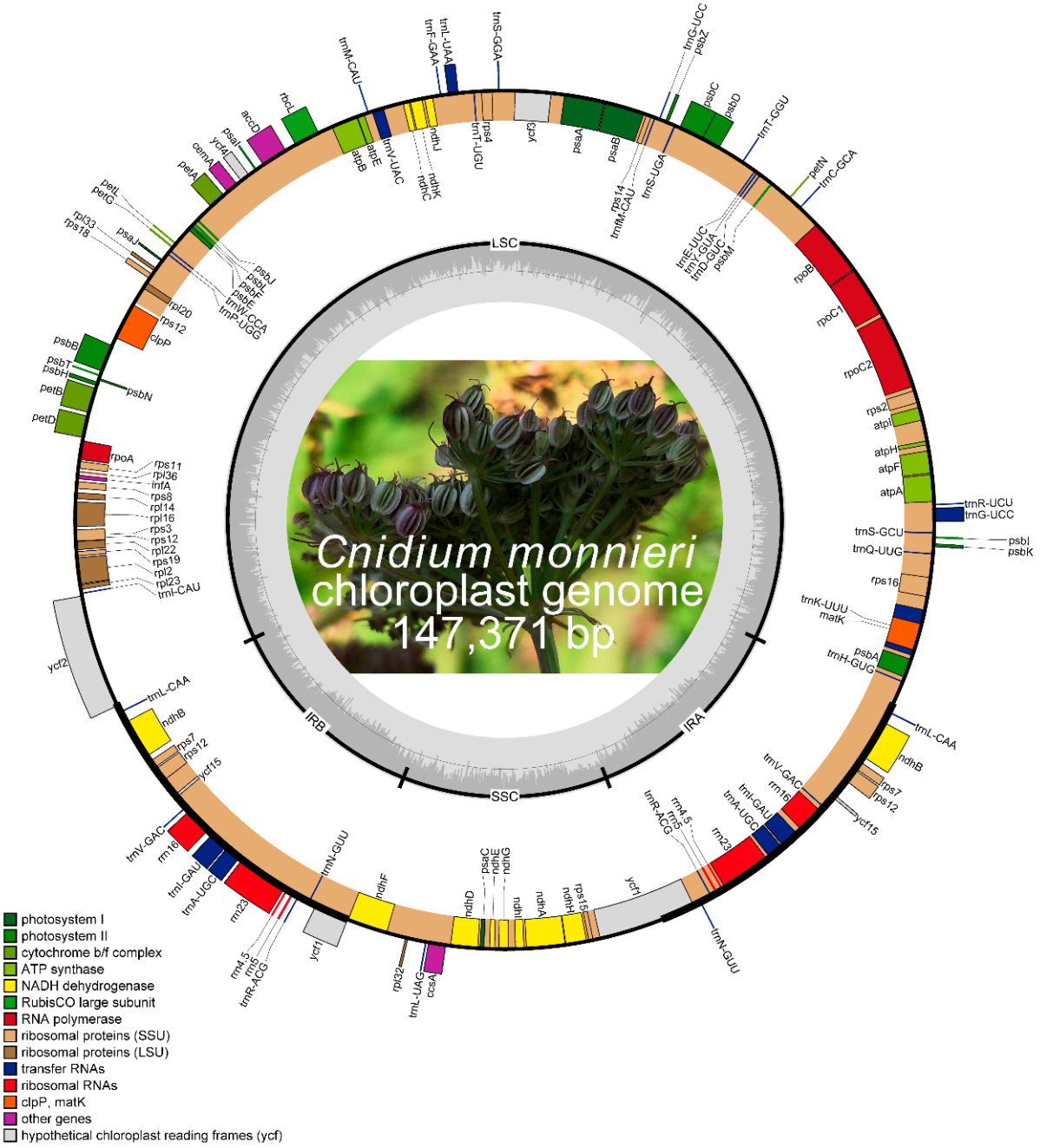


**Figure 1.** Gene map of the *C. monnieri* chloroplast genome. The genes drawn outside of the circle are transcribed clockwise, while those inside are counterclockwise. The genes belonging to different functional groups are color-coded.

The chloroplast is an important organelle for green plants, which participates in the photosynthesis, the biosynthesis of fatty acids, starch, pigments, and amino acids [9]. The chloroplast (cp) genome generally ranges from 115 to 165 kb and has a quadripartite structure comprising a pair of inverted repeat (IR) regions separated by the large single-copy (LSC) and small single-copy (SSC) regions in most angiosperms [10]. The uniqueness of the cp genome, including maternal inheritance, small size, conserved sequences, and simple structure [11], makes it suitable for divergence dating, DNA barcoding, taxonomy, and

phylogeny studies [12–15]. Currently, the high-throughput sequencing technology has been very mature and widely used. For this reason, more and more cp genomes of Apiaceae have been reported and performed in phylogenomic research, such as *Angelica* [16], *Bupleurum* [17,18], *Notopterygium* [19], *Ligusticum* [20], etc. There has been a lack of studies using the cp genome to resolve the phylogeny of *C. monnieri* with respect to Apiaceae.

In this study, we obtained the complete cp genome of *C. monnieri* by de novo assembly. We described the cp genome characteristics, including length, gene content, GC content, repeat sequences, SSRs, codon usage, and RNA editing. We simultaneously compared and analyzed the cp genome with 12 members of the tribe Selineae. These 12 species involved nine genera of published cp genome in the tribe Selineae. We additionally performed a phylogenetic analysis using cp genome sequences and internal transcribed spacer (ITS) sequences to reconstruct the phylogeny and infer the phylogenetic position of *C. monnieri*. In a word, our study is useful for the studies on phylogeny, taxonomy, and species authentication of *C. monnieri*, and is also conducive to elucidating the phylogenetic relationships and taxonomic position of *Cnidium*.

## 2. Materials and Methods

### *2.1. Taxon Sampling and DNA Sequencing*

The fresh plant leaves of *C. monnieri* were collected from Huanren, Liaoning, China (41°7′6.77″ N, 125°17′29.72″ E, October 2020). A voucher specimen (RT2020100801) was deposited at the herbarium of Sichuan University (Chengdu, China). Total DNA was extracted using the CTAB method [21]. 1% agarose gel electrophoresis and a Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, Carlsbad, CA, USA) were used to measure the DNA integrity and concentration. The total DNA sample was sent to Shanghai Personal Biotechnology Co., Ltd. (Shanghai, China) for library construction (400 bp) and Illumina sequencing. The TruSeq DNA Sample Preparation Kits (Illumina, San Diego, CA, USA) were used to construct an Illumina paired-end library. The paired-end (2 × 150 bp) sequencing was conducted on the Illumina NovaSeq platform (Illumina, San Diego, CA, USA), yielding approximately 5 Gb of raw data for *C. monnieri*.

### *2.2. Chloroplast Genome Assembly and Annotation*

Firstly, the obtained raw data were trimmed by removing adaptors and low-quality reads using AdapterRemoval v2 (trimwindows = 5; minlength = 50) [22]. Then, de novo assembly was performed using GetOrganelle v1.7.2 [23] with default settings (-R 15; -k 21, 45, 65, 85, 105). Finally, the annotation of the cp genome was performed using the GeSeq [24]. The Geneious v9.0.2 [25] was used to manually correct the positions of start and stop codons, as well as the boundaries of introns and exons by comparing the sequences with the related species to ensure the accuracy of the annotation results. The gene map of *C. monnieri* was plotted by OGDRAW [26]. The annotated complete cp genome of *C. monnieri* was deposited in GenBank with accession no. OL839918.

### *2.3. Codon Usage, RNA Editing, Repeat Sequence, and SSR of C. monnieri Cp Genome*

The codon usage of all protein-coding genes was analyzed using MEGA6.0 software [27]. RNA editing sites were detected by the PREP suite [28] with a cut-off value of 0.8. Simple sequence repeats (SSRs) were detected using MISA Perl script [29], with the following parameters: 10 repeat units for mononucleotide SSRs, 5 repeat units for dinucleotide repeat SSRs, 4 repeat units for trinucleotide repeat SSRs, and 3 repeat units for tetra-, penta-, and hexanucleotide repeat SSRs. The forward, reverse, palindromic and complementary repeats were detected by the online REPuter program [30] with a minimal size $\geq$ 30 bp, and a hamming distance = 3.

### *2.4. Comparative Chloroplast Genomic Analysis*

The IR/SC boundaries of *C. monnieri* and the other tribe Selineae species were compared to illustrate the IR expansion and contraction. These species involve nine genera

of published cp genome in the tribe Selineae. The 12 species are *A. dahurica* (NC_029392), *A. gigas* (KX118044), *A. laxifoliata* (NC_040122), *G. littoralis* (NC_034645), *L. buchtormensis* (NC_058871), *Lig. likiangense* (NC_049055) [20], *Lig. thomsonii* (NC_049058) [20], *M. pimpinelloideum* (NC_047428) [31], *P. japonicum* (NC_034644), *P. praeruptorum* (MN016968) [32], *S. divaricata* (MN857472) [33], and *Ses. montanum* (KM035851) [34]. The alignments of the 13 complete cp genomes were generated by mVISTA [35] using *C. monnieri* as a reference. The nucleotide diversity (Pi) of 13 cp genome sequences was calculated using DnaSP v5.1 [36]. The step size and window length were set to 200 bp and 600 bp in the sliding window method.

### *2.5. Evolutionary and Phylogenomic Analyses*

We used the KaKs_Calculator2.0 program [37] with the NG model to calculate the rates of non-synonymous substitutions (Ka), synonymous substitutions (Ks), and their ratio (Ka/Ks) using *C. monnieri* as a reference. The 79 protein-coding sequences were aligned using the ClustalW (Codons) method of the MEGA6.0 software [27], then axt format files were prepared using AXTConvertor.

The 48 complete cp genome sequences were used to construct phylogenetic trees to determine the phylogenetic position of *C. monnieri*, in which 45 completed cp genome sequences from Apioideae, and three Saniculoideae species were used as outgroup. We selected Maximum likelihood (ML) and Bayesian inference (BI) methods to perform phylogenomic analyses. 78 chloroplast protein-coding genes (PCGs) shared by 48 species were extracted using PhyloSuite v1.2.2 [38]. The shared 78 PCGs were aligned by MAFFT [39] with codon mode, trimmed by trimAl [40] with nogaps option, and were then concatenated by PhyloSuite v1.2.2 [38]. Moreover, corresponding 48 nuclear ITS sequences downloaded from the GenBank were also used to infer the phylogenetic relationships of *C. monnieri*. We selected the GTRGAMMA model for ML analysis implemented in RAxML v8.2.8 [41] with 1000 bootstrap replicates. MrBayes v3.1.2 [42] was used to perform the BI analysis with the nucleotide substitution model (GTR + I + G and SYM + I + G) selected by Modeltest v3.7 [43]. The Markov chain Monte Carlo (MCMC) algorithm was run for 5,000,000 generations (sampling every 100 generations) with two runs and four chains (three heated chains and one cold chain). The first 25% of trees were discarded as burn-in, and the remaining trees were used to construct the consensus tree.

## 3. Results

### *3.1. Characteristics of the C. monnieri Cp Genome*

The complete cp genome sequence of *C. monnieri* was 147,371 bp in length and had a typical quadripartite structure as found in most land plants (Figure 1). The cp genome included a LSC region of 94,361 bp and a SSC region of 17,552 bp, separated by a pair of IR regions with 17,729 bp (Table 1). The GC content in the whole genome, LSC, SSC, and IRs was 37.4, 35.9, 30.9, and 44.9%, respectively (Table 1). It was worth mentioning that GC content in IRs was higher than LSC and SSC regions. The GC content of the protein-coding regions (CDS) was 37.9%. Within CDS, the GC content was 45.8% for the first codon position, 38.2% for the second position, and 29.6% for the third position.

There are 129 genes in the *C. monnieri* cp genome, containing 85 protein-coding genes, 36 tRNA genes, and 8 rRNA genes (Figure 1, Table S1). Among these genes, 15 genes (*trnK-UUU*, *trnG-UCC*, *trnL-UAA*, *trnV-UAC*, *trnI-GAU*, *trnA-UGC*, *rps16*, *atpF*, *rpoC1*, *petB*, *petD*, *rpl16*, *rpl2*, *ndhA*, and *ndhB*) harbored a single intron and three genes (*ycf3*, *clpP*, and *rps12*) had two introns (Table S1). The *trnK-UUU* had the largest intron, up to 2527 bp (Figure 1). The *rps12* gene is trans-spliced (Figure 1). Its 5′ end and the duplicated 3′ end are located in the LSC and IR regions, respectively.

The cp genomes of 12 tribe Selineae species were selected to compare with the *C. monnieri* (Table 1). The GC content was similar in different species, while the total length and the gene number had differences. *M. pimpinelloideum* and *P. japonicum* had the

largest total length and the most gene number. The other species varied slightly in the total length and the gene number.

**Table 1.** Comparative analyses of chloroplast genomes among 13 tribe Selineae species.

| | Sequence Lengths (bp) | | | | Number of Genes | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Total** | **LSC** | **SSC** | **IR** | **Total** | **Protein-Coding** | **tRNA** | **rRNA** | **GC(%)** |
| *C. monnieri* | 147,371 | 94,361 | 17,552 | 17,729 | 129 | 85 | 36 | 8 | 37.4 |
| *A. dahurica* | 146,918 | 93,604 | 17,676 | 17,819 | 129 | 85 | 36 | 8 | 37.5 |
| *A. gigas* | 146,918 | 93,120 | 17,582 | 18,108 | 128 | 83 | 36 | 8 | 37.6 |
| *A. laxifoliata* | 147,026 | 93,191 | 17,493 | 18,171 | 129 | 85 | 36 | 8 | 37.5 |
| *G. littoralis* | 147,477 | 93,496 | 17,555 | 18,213 | 129 | 85 | 36 | 8 | 37.5 |
| *L. buchtormensis* | 147,036 | 91,969 | 17,469 | 18,799 | 127 | 83 | 36 | 8 | 37.6 |
| *Lig. likiangense* | 148,196 | 92,305 | 17,575 | 19,158 | 129 | 85 | 36 | 8 | 37.5 |
| *Lig. thomsonii* | 147,462 | 93,363 | 17,591 | 18,254 | 129 | 85 | 36 | 8 | 37.6 |
| *M. pimpinelloideum* | 164,431 | 76,445 | 17,565 | 35,211 | 144 | 99 | 37 | 8 | 37.5 |
| *P. japonicum* | 164,653 | 75,584 | 17,551 | 35,759 | 144 | 99 | 37 | 8 | 37.5 |
| *P. praeruptorum* | 147,197 | 92,161 | 17,610 | 18,713 | 128 | 84 | 35 | 8 | 37.6 |
| *S. divaricata* | 147,834 | 93,202 | 17,324 | 18,654 | 129 | 85 | 36 | 8 | 37.5 |
| *Ses. montanum* | 147,823 | 92,620 | 17,479 | 18,862 | 127 | 82 | 36 | 8 | 37.6 |

### 3.2. Repeat Sequence and SSRs

We detected a total of 47 repeats in the cp genome of *C. monnieri*, including 31 forward repeats (F), 14 palindromic repeats (P), and 2 reverse repeats (R), while no complement repeat (C) was found (Figure 2A, Table S2). Among them, 40 repeats were 30–45 bp, 4 repeats were 45–60 bp, 2 repeats were 60–75 bp, 1 repeat was more than 75 bp, and the longest repeat was 113 bp (Figure 2B). Most of these repeats (32) were located in the intron and intergenic spacer (IGS), and the minority were found in the extron. Among the extron, *ycf*2 possessed the highest number of repeats (14) and the longest repeat was located in this gene.

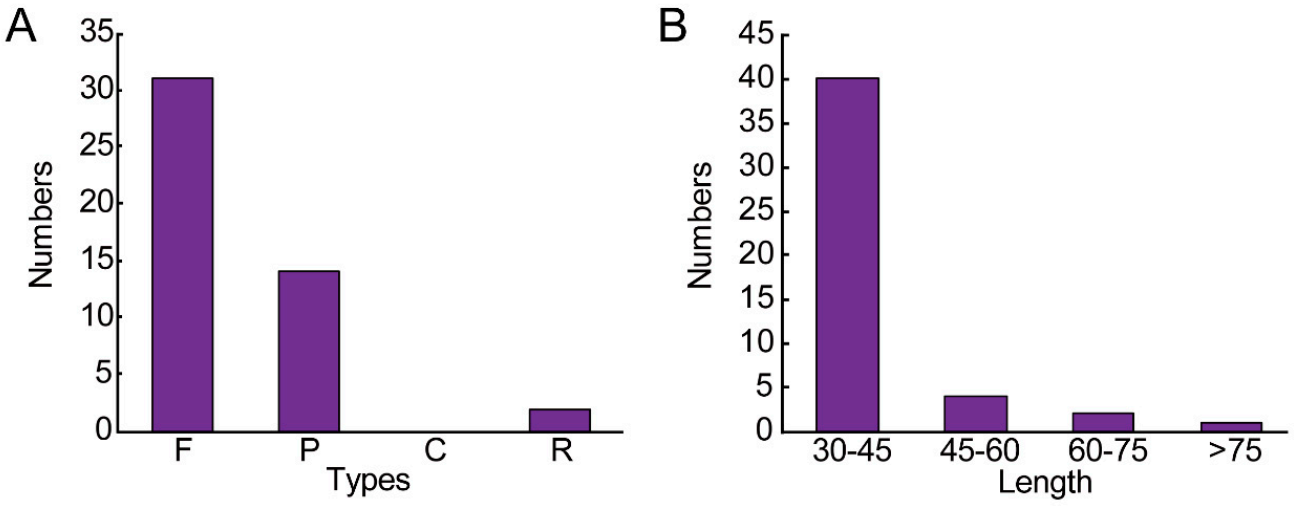


**Figure 2.** The repeat sequences in the *C. monnieri* chloroplast genome. (**A**) Total number of three repeat types; (**B**) Number of repeats divided by length. F: forward repeats; P: palindromic repeats; R: reverse repeats; C: complementary repeats.

In total, 79 SSRs with at least 10 bp long were observed in the *C. monnieri* cp genome, and they are AT-riched (Figure 3A, Table S2). More than half of the SSRs (45, 57.0%) were mononucleotide A/T repeats, only three C/G mononucleotide SSRs were present. Among dinucleotide SSRs, these are all AT/TA repeats (21). Trinucleotide SSRs (ATA) repeats, tetranucleotide SSRs (TTTA, AATA, TCCT, AGGT, TCTT, CAAT, CTAC) repeats

and pentanucleotide SSRs (AATCA, TTTTA) had one, respectively. Furthermore, the majority of SSRs located in the LSC (54) followed by SSC (15) and IRs (10), whereas 57 were located in the intergenic spacer, 6 in the intron, and 16 in the extron (Figure 3B).

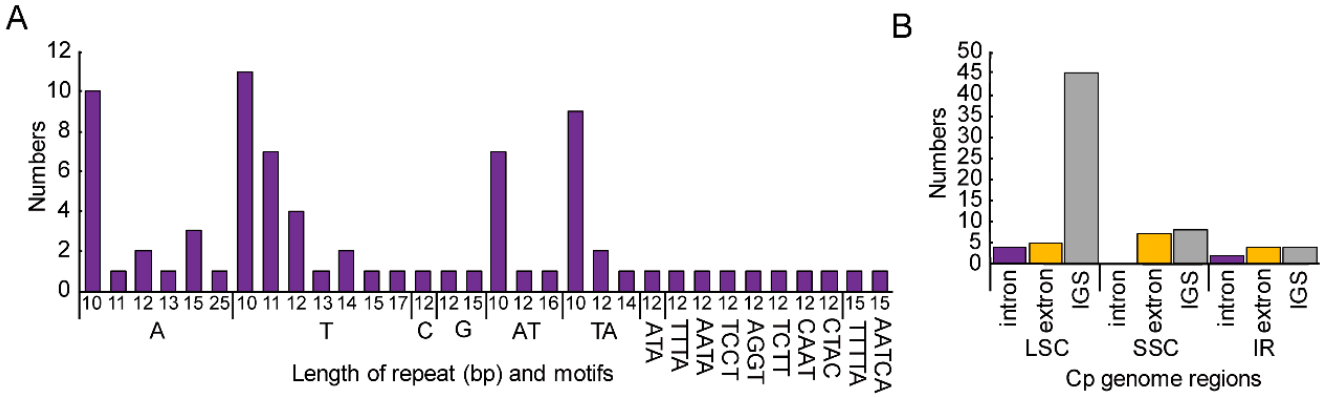


**Figure 3.** The simple sequence repeats (SSRs) in the *C. monnieri* chloroplast genome. (**A**) Types and numbers of SSRs; (**B**) SSRs distribution in different regions.

### *3.3. Codon Usage, RNA Editing, and Ka/Ks in Protein-Coding Genes*

Codon usage frequency of the *C. monnieri* cp genome was estimated and summarized (Table S3). All the protein-coding genes were encoded by 24,248 codons. In these codons, the most frequent amino acid was Leucine (2576, 10.62%) and the least was Cysteine (258, 1.06%). The start codon AUG was identified 572 times. All three stop codons were present with UAA used most frequently (46 times) followed by UAG 22 times and UGA 17 times. The GC content of the third codon positions is significantly lower than the first and second codon positions.

The PREP-cp program predicted 59 potential RNA editing sites for 22 protein-coding genes of the *C. monnieri* cp genome (Table S4). Of these 59 editing sites, 15 (25.4%) and 44 (74.6%) were located at the first and the second codon position, but no editing site was found at the third codon position. The genes *ndhB* (10), *rpoB* (6), *rpoC2* (6), *ndhD* (4), *ndhA* (4), *accD* (5), and *matK* (3) had a high number of RNA editing sites. Other genes, namely, *atpA*, *atpB*, *ccsA*, *ndhF*, *ndhG*, *petB*, *petG*, *psaI*, *psbE*, *psbF*, *rpl20*, *rpoC2*, *rpoA*, *rps2*, and *rps14*, contained two or one potential RNA editing site. The amino acid conversion S to L occurred most frequently, while A to V and R to C occurred least. 37 out of 59 RNA editing sites changed the encoded amino acid from polar to apolar.

We analyzed the synonymous and non-synonymous change rates of 79 protein-coding genes in tribe Selineae (Table S5). 11 genes (*ccsA*, *matK*, *ndhA*, *psbI*, *rbcL*, *rpl22*, *rpoA*, *rpoC2*, *rps3*, *ycf1*, and *ycf2*) were identified under positive selection (Ka/Ks > 1). This showed that although tribe Selineae face weak selection pressure, some are undergoing adaptations to their environment. Among the 11 genes, *ndhA*, *psbI*, *rpoC2*, *rps3*, and *rpl22* showed high rates for one species. The genes *ccsA*, *ycf2*, and *rpoA* presented high rates for the two species. The genes *matK*, *rbcL*, and *ycf1* presented high rates for 5, 6, and 9 species, respectively.

### *3.4. Comparative Analyses*

A comparison of the IR boundaries was performed among the 13 tribe Selineae species (Figure 4). JLB extended over the *petB* in *M. pimpinelloideum* and *P. japonicum*, while JLB extended over *ycf2* in the other 11 tribe Selineae species. The JSB was located in the *ycf1* or *ψycf1*. Moreover, the JSB expanded 1–70 bp into *ndhF* of *M. pimpinelloideum*, *G. littoralis*, and *A. dahurica*. The JSA was located in the *ycf1* and expanded 1655–2269 bp into the *ycf1* in the 13 tribe Selineae cp genomes. The *trnH* were all located in the LSC region and were 1632 bp away from the JLA. The *petD* were 1029–1472 bp away from JLA in *M. pimpinelloideum* and *P. japonicum*, while *trnL* were 510–1862 bp away from JLA in other 11 tribe Selineae species.

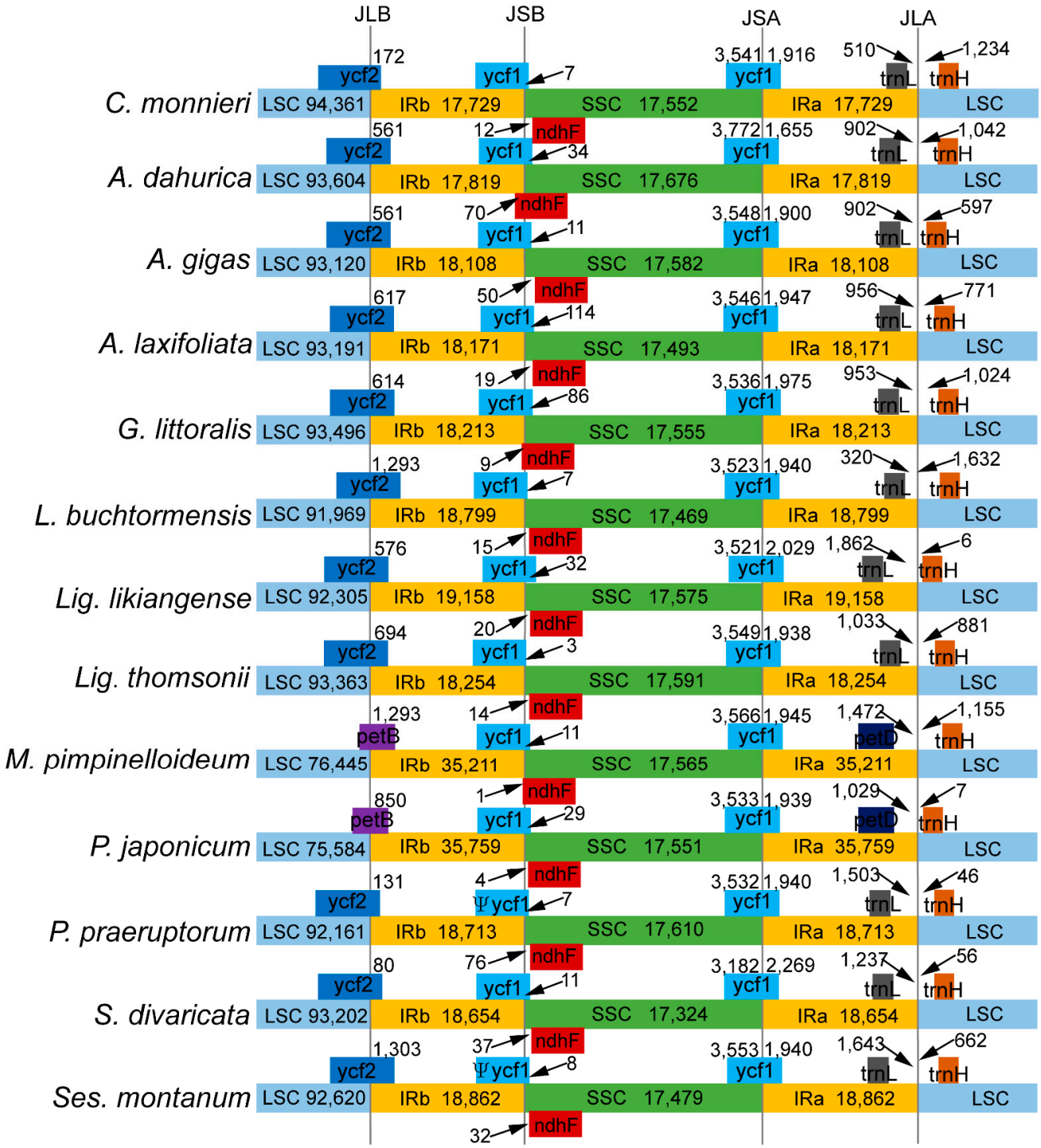


**Figure 4.** Comparison of the IR boundaries from 13 tribe Selineae chloroplast genomes. The ψ indicates a pseudogene. The figure is not drawn to scale.

The overall sequence identity of the 13 tribe Selineae cp genomes was plotted using the mVISTA program (Figure 5). The comparison demonstrated that the two IR regions were less divergent than the LSC and SSC regions. Non-coding regions were more divergent than coding regions, and the most divergent regions were localized in the intergenic spacers. The nucleotide diversity (Pi) was calculated to determine the sequence divergence levels (Figure 6). The analysis reached the same conclusion that the IR regions were more conserved than the SC regions. Moreover, we detected four highly variable regions (Pi > 0.02), namely *trnD-trnY-trnE-trnT*, *ycf2*, *ndhF-rpl32-trnL*, and *ycf1*.

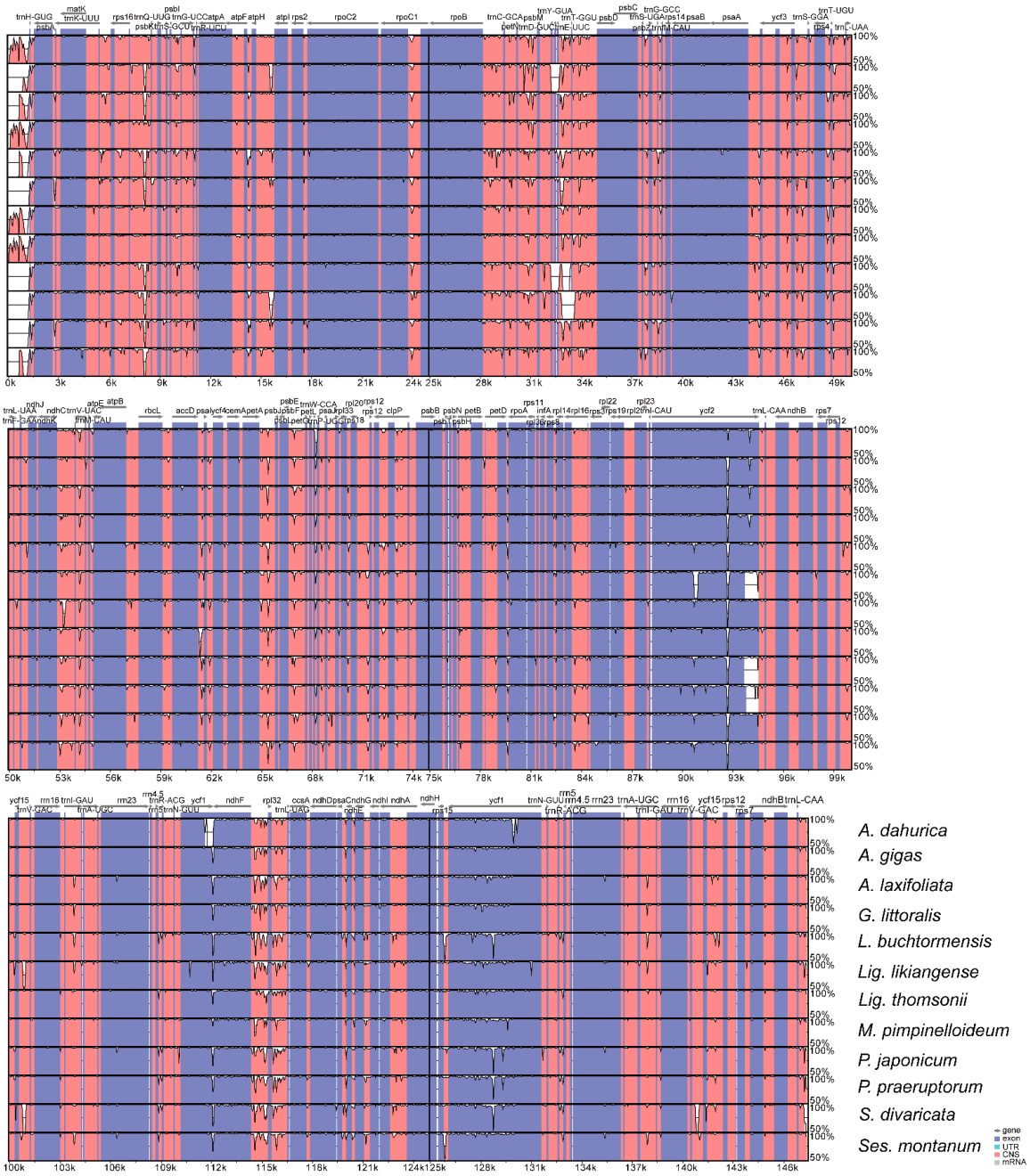


**Figure 5.** The sequence alignment of 13 tribe Selineae chloroplast genomes in mVISTA using *C. monnieri* as a reference. The vertical scale represents the percentage of identity ranging from 50 to 100%.

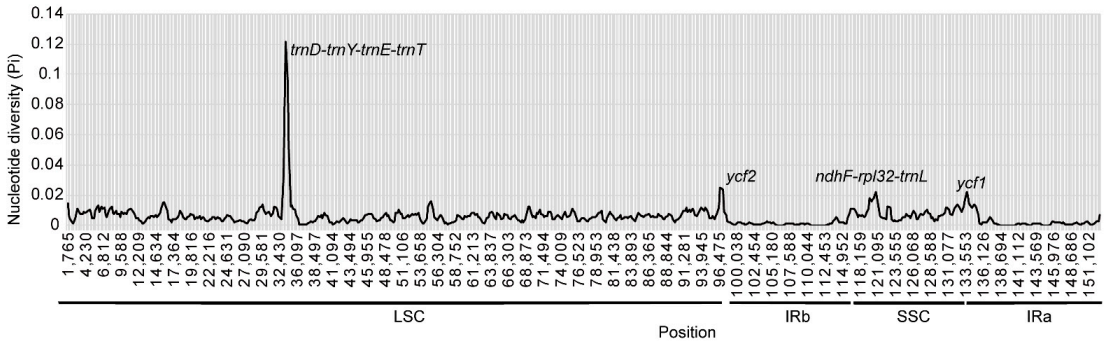


**Figure 6.** The nucleotide diversity (Pi) by sliding window analysis in the aligned 13 tribe Selineae chloroplast genomes.

### 3.5. Phylogeny Inference

The chloroplast genome sequences and nuclear ITS sequences were used to perform the phylogenetic analyses (Table S6). The cp tree and ITS tree produced incongruent tree topologies, while they all inferred the non-monophyly of *Cnidium* (Figures 7 and 8). *C. monnieri* and *C. officinale* fell into tribe Selineae and *Sinodielsia* Clade with 100% support values, respectively. The cp phylogenetic tree based on 78 shared PCGs with ML and BI methods had an identical topological structure with high support values (Figure 7). *C. monnieri* was related to *Angelica* group instead of other tribe Selineae species. The *Angelica* group also contained two species (*G. littoralis* and *M. pimpinelloideum*) in our study [44]. *C. officinale* was related to *L. sinense*, and was then clustered with *L. chuanxiong* and *L. jeholense* formed a clade. In the ITS tree, the tree topologies resulting from ML and BI analysis were sometimes different (Figure 8). *C. monnieri* was related to *Ses. Montanum* with weak support (BS < 50%), then they clustered with other tribe Selineae species except for *Angelica* group by ML analysis, whereas it was resolved as a sister to other tribe Selineae species except for *Angelica* group with moderate support by BI analysis (PP = 0.7). As for *C. officinale*, it was still clustered with *Ligusticum* species in *Sinodielsia* Clade with high support. Other clades of Apioideae were generally consistent with prior studies [7,45–47].

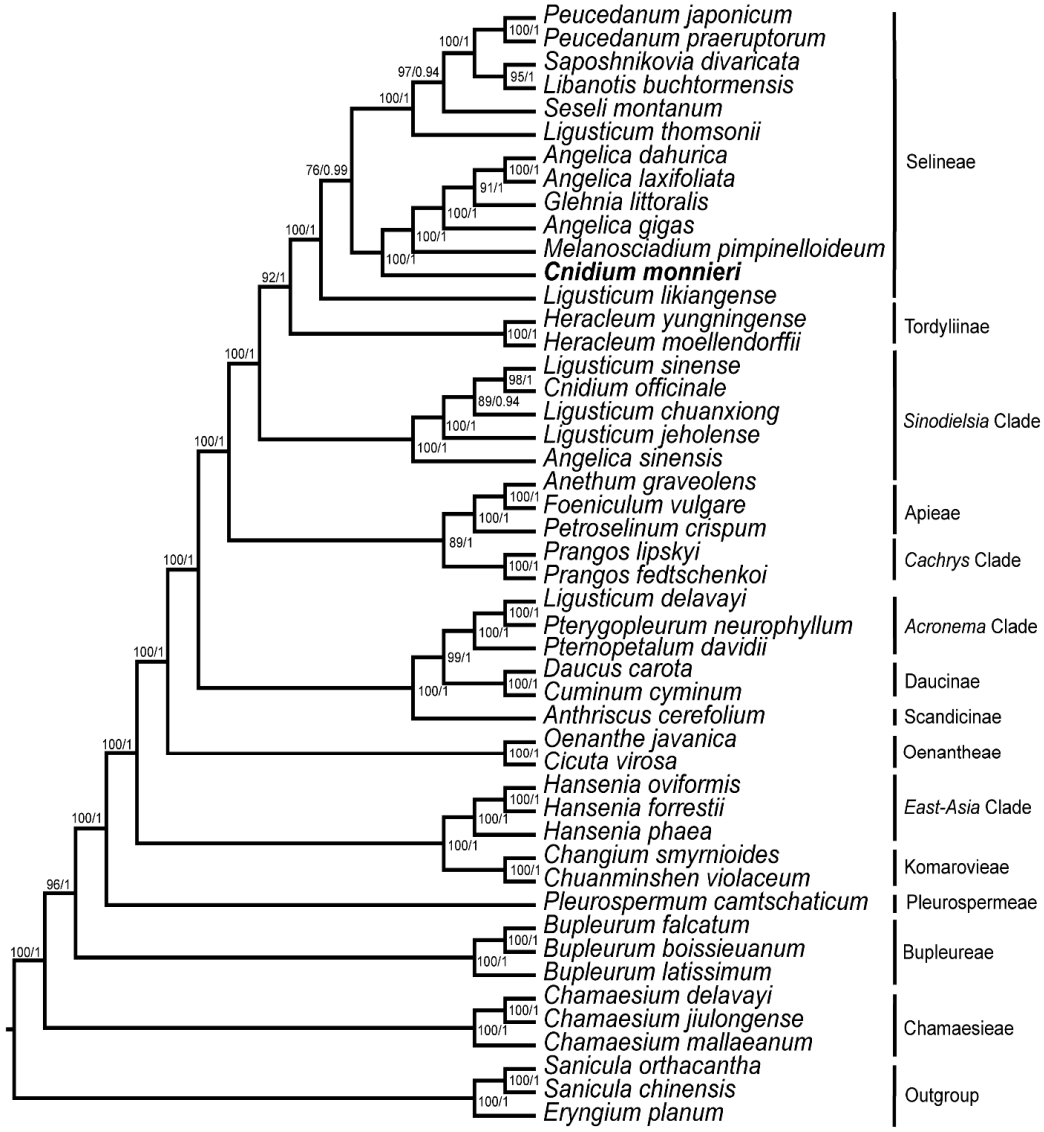


**Figure 7.** Phylogenetic relationships of 48 species inferred from 78 shared protein-coding genes. The numbered-above nodes are bootstrap support values and posterior probability values.

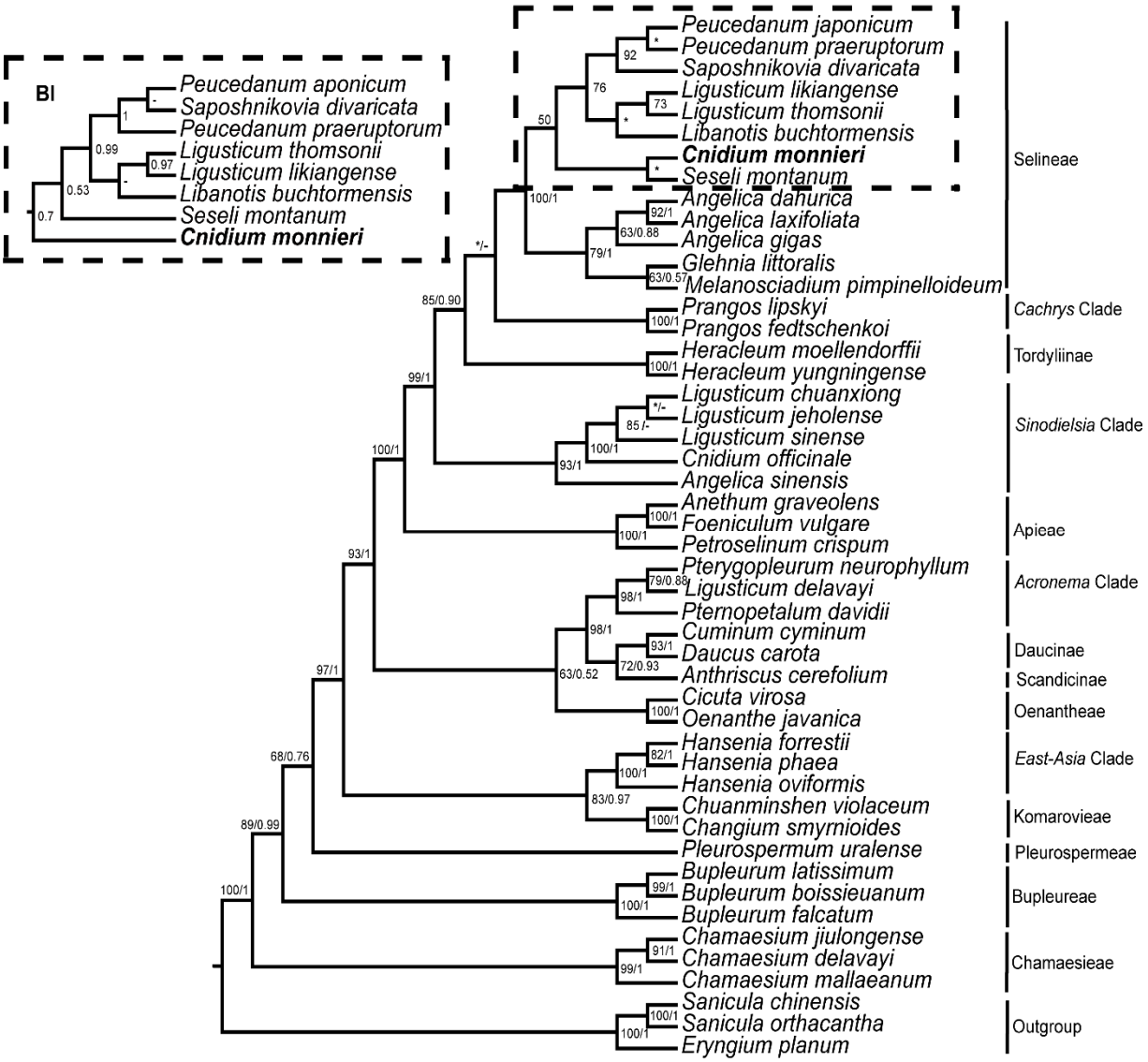


**Figure 8.** Phylogenetic relationships of 48 species inferred from nuclear ribosomal internal transcribed spacer (ITS) sequences. The numbered-above nodes are bootstrap support values and posterior probability values. "*" indicates bootstrap support values less than 50%, "-" indicates posterior probability values less than 0.5.

## 4. Discussion

In the present study, the cp genome of *C. monnieri* was sequenced, and comparative analyses were conducted with the other 12 tribe Selineae species to gain a complete view of the architecture of the *C. monnieri* cp genome. The *C. monnieri* cp genome herein revealed a typical quadripartite structure and an expected size range for angiosperm plants [11,48]. There was no gene loss in the cp genome of *C. monnieri*, yet gene loss actually often occurs in plant lineages. Many cp genes have been lost in different plants, such as *accD*, *infA*, *clpP*, *ccsA*, *rps16*, *rpl23*, *ndh* complex, and so on [49–52]. However, the *rpl20* gene was considered the most stable, for it was not lost in the study of 2511 cp genomes [49]. The high GC content in the IR region of *C. monnieri* might result from four rRNAs genes with a high GC content [53]. The length of the whole genome, IR, LSC, and SSC regions, as well as the gene number, showed obvious differences among the 13 tribe Selineae cp genomes, which was also observed in a recent study of Apiaceae [45].

Codons play an important role in spreading genetic information because it acts as a bridge to nucleic acids and proteins. The same amino acid has two or more codons, while

each has its preferred codon owing to codon usage bias (CUB). Several studies have shown that the CUB may be the result of natural selection, mutation, and genetic drift [54,55]. In *C. monnieri*, the most preferred codons end with A/U, which is consistent with various terrestrial plant cp genomes [50,56,57]. Perhaps high AT content in the cp genomes is the major reason for biased codons ending with A/U [58]. In higher plants, the protein-coding genes of chloroplast always occur in RNA editing and mostly convert cytidine (C) to uridine (U) [59]. RNA editing can correct DNA mutations on the RNA level, and is thereby essential for the growth and development of plants [60,61]. Here, 59 RNA editing sites were identified in 22 protein-coding genes of the cp genome of *C. monnieri*. Most RNA editing sites changed the encoded amino acid from polar to apolar, thus increasing the protein hydrophobicity. This tendency was also reported in other studies, which implies that the increased hydrophobicity can affect the protein structural features, protein–protein interactions, and transmembrane domains in the chloroplast protein complexes [59,61,62].

As for repeat sequences and SSRs, *C. monnieri* showed patterns comparable to other Apiaceae in numbers and location [17,20]. A high number of repeats were located in intergenic regions or *ycf2*. Repeats in the *ycf2* are commonly observed in *Primula* and *Cardiocrinum* [50,63]. Most mono- and dinucleotides are generally composed of A or T and infrequently contain C or G, which contribute to the high AT content in the cp genome [56,64]. As in other species [58,64], the distribution of SSRs is uneven in the *C. monnieri* cp genome. These abundant SSRs in cp genomes would be served as genetic markers applied in population genetics and phylogeographical studies of *C. monnieri*.

The change of the IR/SC boundary is a universal phenomenon in the cp genome evolution [65,66]. In the 13 tribe Selineae species, *C. monnieri*, *A. dahurica*, *A. gigas*, *A. laxifoliata*, *G. littoralis*, *L. buchtormensis*, *Lig. likiangense*, *Lig. thomsonii* and *S. divaricata* showed similar characteristics, and only the length flanking showed a little difference. The organization gene of JSB in *P. praeruptorum* and *Ses. montanum* is *ψycf1*, which is different from other species. *M. pimpinelloideum* and *P. japonicum* showed more differences than other species in organization genes of JLB and JLA, namely, *petB* and *petD*, as well as the length of IR and LSC (35,211–35,759 bp; 76,445–75,584 bp). The IR contraction and expansion is the main reason for the size change of cp genomes [66,67]. We therefore concluded that the IR expansion of *M. pimpinelloideum* and *P. japonicum* causes a longer length of the two cp genomes (164,431–164,653 bp). Except for contraction and expansion, the absence of IR region is not uncommon in plant lineages. A recent study showed that about 10.31% of the cp genomes have lost the IR region span across all the lineages [52]. In angiosperms, some species of *Erodium*, *Cassytha*, *Passiflora*, and legumes [68–71] have been described as IR-lacking lineages.

The non-synonymous (Ka) and synonymous (Ks) nucleotide substitution patterns are very vital markers in gene evolution studies. The ratio of Ka/Ks < 1 indicates purifying selection, while Ka/Ks > 1 indicates positive selection [72]. Non-synonymous nucleotide substitutions have occurred less frequently than synonymous substitutions in most protein-coding genes [73]. We calculated Ka/Ks ratios of the protein-coding genes in *C. monnieri* cp genome versus 12 other tribe Selineae species. Notably, most of the genes indicated purifying selection based on the Ka/Ks values. However, 11 genes (*ccsA*, *matK*, *ndhA*, *psbI*, *rbcL*, *rpl22*, *rpoA*, *rpoC2*, *rps3*, *ycf1*, and *ycf2*) were identified under positive selection. Among these, the Ka/Ks ratio of *ndhA*, *psbI*, *rpoC2*, *rps3*, *rpl22*, *ccsA*, *ycf2*, and *rpoA* genes in one or two of the 12 comparison groups, while the Ka/Ks ratio of *matK*, *rbcL*, and *ycf1* genes in 5, 6, and 9 of the 12 comparison groups, respectively. The *matK* gene with about 1500 bp in length encodes the only maturase (MatK) in the cp genome of land plants [74]. MatK regulates the expression of several essential genes related to protein biosynthesis [61], thus the *matK* gene is essential for cell viability. The *rbcL* gene encodes the large subunit of the ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) [11]. It is a modulator of photosynthetic electron transport and is essential for photosynthesis [75]. The *ycf1* gene is one of the largest genes in the cp genome, encoding a protein of about 1800 amino acids [11]. A study showed that *ycf1* gene encodes Tic214, a vital component of the *Arabidopsis* TIC

complex [76]. Positive selection on the three genes has also been observed in Dipsacales [77]. Furthermore, *matK*, *rbcL*, and *ycf1* genes are also known as fast-evolving genes compared with other chloroplast genes, and they therefore can be used as an excellent molecular marker to be widely applied in phylogenetic inferences [78–80].

The hypervariable regions can serve as potential molecular markers for species identification [81]. Four highly variable regions, *trnD-trnY-trnE-trnT*, *ycf2*, *ndhF-rpl32-trnL*, and *ycf1*, were detected. The highly variable tRNA cluster (*trnD-trnY-trnE-trnT*) was also identified in *Notopterygium* [19]. As a medicinal plant, the candidate DNA barcodes could be applied to species authentication for assuring medicinal quality.

We investigated the phylogenetic position of *C. monnieri* using cp genome sequences and ITS sequences, while the cp tree and ITS tree produced incongruent tree topologies. Nevertheless, they all inferred the non-monophyly of *Cnidium*, which has been demonstrated in other studies [7,8]. *C. monnieri* fell into tribe Selineae, consistent with previous studies [8,45–47,82]. In the cp tree, *C. monnieri* was related to *Angelica* group instead of other tribe Selineae species. In the ITS tree, *C. monnieri* was related to *Ses. montanum* then clustered with other tribe Selineae species except for *Angelica* group by ML analysis, whereas it was resolved as sister to other tribe Selineae species except for *Angelica* group by BI analysis. The incongruence of the phylogenetic position of *C. monnieri* between ITS and cp phylogenies is consistent with the previous study based on transcriptomes and chloroplast genomes [45,82], suggesting that *C. monnieri* may have experienced complex evolutions with hybrid and incomplete lineage sorting [82]. In addition, the incongruence may result from the different mutation rates and inherited background between cpDNA and ITS. The ITS is biparentally inherited with a higher mutation rate, while the cpDNA is maternally inherited with a lower mutation rate [83]. As for *C. officinale*, it fell into *Sinodielsia* Clade and formed a clade with *L. sinense*, *L. chuanxiong* and *L. jeholense*. This may indicate that *C. officinale* should be suggested to be transferred to *Ligusticum* [7,8]. This study is helpful to deepen the understanding of the phylogenetic position of *C. monnieri*, and provides new insight into the phylogeny and taxonomy of the genus *Cnidium*. In a word, our result will be beneficial to future phylogeny, taxonomy, and evolutionary studies of *C. monnieri*.

## 5. Conclusions

In this study, the chloroplast genome of *C. monnieri* was sequenced, assembled, and compared with other tribe Selineae species. The cp genome was a circular molecule of 147,371 bp with 37.4% GC content. *C. monnieri* cp genome harbored 129 genes, including 85 protein-coding genes, 36 tRNA genes, and eight rRNA genes. All the protein-coding genes were encoded by 24,248 codons. We detected 79 simple sequence repeats (SSRs), 47 repeat sequences, and 59 RNA editing sites. The IR boundary of *C. monnieri* did not show significant expansion and contraction relative to the other 12 species. Four hypervariable regions (*trnD-trnY-trnE-trnT*, *ycf2*, *ndhF-rpl32-trnL*, and *ycf1*) were identified as candidate molecular markers for species authentication. Eleven genes (*ccsA*, *matK*, *ndhA*, *psbI*, *rbcL*, *rpl22*, *rpoA*, *rpoC2*, *rps3*, *ycf1*, and *ycf2*) were identified under positive selection (Ka/Ks > 1). The phylogenetic analysis based on cp genome sequences and internal transcribed spacer (ITS) sequences from 48 species supported *C. monnieri* located in tribe Selineae with high support. The incongruence of the phylogenetic position of *C. monnieri* between ITS and cpDNA phylogenies suggested that *C. monnieri* might have experienced complex evolutions with hybrid and incomplete lineage sorting. This study provided precious genetic resources of *C. monnieri*, which can be used for the species authentication, phylogeny, and taxonomy studies of *C. monnieri*. This study would also be beneficial in elucidating the taxonomy and reconstructing the phylogeny of the genus *Cnidium*.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d14050323/s1, Table S1: Gene composition in the *C. monnieri* chloroplast genome. Table S2: The repeat sequences and simple sequence repeats (SSRs) in the *C. monnieri* chloroplast genome. Table S3: Codon usage and relative synonymous codon usage (RSCU) value

for protein-coding genes in the *C. monnieri* chloroplast genome. Table S4: RNA editing sites in the *C. monnieri* chloroplast genome. Table S5: The rates of Ka, Ks, and Ka/Ks of 79 protein-coding genes among 13 tribe Selineae species. Table S6: Species included in the phylogenetic analyses with their accession numbers.

**Author Contributions:** Conceptualization, X.H.; Data curation, T.R., X.A. and R.T.; Formal analysis, T.R., X.A., R.T. and Z.L.; Methodology, T.R.; Project administration, X.H.; Resources, T.R., R.T. and Z.L.; Software, X.A. and C.P.; Writing—original draft, T.R.; Writing—review and editing, C.P. and X.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant Nos. 32070221, 31872647), National Herbarium of China, National Herbarium resources teaching specimen database (Grant No. 2020BBFK01), the fourth national survey of traditional Chinese medicine resources (Grant No. 2019PC002).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The complete chloroplast genome sequence of *C. monnieri* was deposited at NCBI (GenBank accession number: OL839918).

**Acknowledgments:** We thank the reviewers who helped improve our manuscript. We also thank Shanghai Personal Biotechnology company for sequencing.

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

## References

1. Pu, F.D.; Watson, M.F. *Cnidium Cusson. Flora of China*; Science Press: Beijing, China, 2005; Volume 14, pp. 136–137.
2. China Pharmacopoeia Committee. *Pharmacopoeia of the People's Republic of China Part I*; Medical Science and Technology Press: Beijing, China, 2015.
3. Zhang, Q.; Qin, L.; He, W.; Van Puyvelde, L.; Maes, D.; Adams, A.; Zheng, H.; De Kimpe, N. Coumarins from *Cnidium monnieri* and their antiosteoporotic activity. *Planta Med.* **2007**, *73*, 13–19. [CrossRef] [PubMed]
4. Sun, Y.; Yang, A.W.H.; Lenon, G.B. Phytochemistry, ethnopharmacology, pharmacokinetics and toxicology of *Cnidium monnieri* (L.) Cusson. *Int. J. Mol. Sci.* **2020**, *21*, 1006. [CrossRef] [PubMed]
5. Li, Y.M.; Jia, M.; Li, H.Q.; Zhang, N.D.; Wen, X.; Rahman, K.; Zhang, Q.Y.; Qin, L.P. *Cnidium monnieri*: A review of traditional uses, phytochemical and ethnopharmacological properties. *Am. J. Chin. Med.* **2015**, *43*, 835–877. [CrossRef] [PubMed]
6. Shi, J.L.; LI, Q.; Zhang, C.Y. Physical and chemical identification of Fructus cnidii and its false *Apium graveolens* fruits. *Chin. J. Mod. Appl. Pharm.* **2001**, *18*, 196–197.
7. Downie, S.R.; Spalik, K.; Katz-Downie, D.S.; Reduron, J. Major clades within Apiaceae subfamily Apioideae as inferred by phylogenetic analysis of nrDNA ITS sequences. *Plant Div. Evol.* **2010**, *128*, 111–136. [CrossRef]
8. Zhou, J.; Gao, Y.Z.; Wei, J.; Liu, Z.W.; Downie, S.R. Molecular phylogenetics of *Ligusticum* (Apiaceae) based on nrDNA ITS sequences: Rampant polyphyly, placement of the Chinese endemic species, and a much-reduced circumscription of the genus. *Int. J. Plant. Sci.* **2020**, *181*, 306–323. [CrossRef]
9. Neuhaus, H.E.; Emes, M.J. Nonphotosynthetic metabolism in plastids. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2000**, *51*, 111–140. [CrossRef]
10. Ravi, V.; Khurana, J.P.; Tyagi, A.K.; Khurana, P. An update on chloroplast genomes. *Plant Syst. Evol.* **2008**, *271*, 101–122. [CrossRef]
11. 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]
12. Zhang, S.D.; Jin, J.J.; Chen, S.Y.; Chase, M.W.; Soltis, D.E.; Li, H.T.; Yang, J.B.; Li, D.Z.; Yi, T.S. Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytol.* **2017**, *214*, 1355–1367. [CrossRef]
13. Carbonell-Caballero, J.; Alonso, R.; Ibañez, V.; Terol, J.; Talon, M.; Dopazo, J. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Mol. Biol. Evol.* **2015**, *32*, 2015–2035. [CrossRef] [PubMed]
14. Krawczyk, K.; Nobis, M.; Myszczyński, K.; Klichowska, E.; Sawicki, J. Plastid super-barcodes as a tool for species discrimination in feather grasses (Poaceae: Stipa). *Sci. Rep.* **2018**, *8*, 1924. [CrossRef] [PubMed]
15. Zhao, F.; Chen, Y.P.; Salmaki, Y.; Drew, B.T.; Wilson, T.C.; Scheen, A.C.; Celep, F.; Bräuchler, C.; Bendiksby, M.; Wang, Q.; et al. An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Biol.* **2021**, *19*, 2. [CrossRef] [PubMed]
16. Wang, M.; Wang, X.; Sun, J.; Wang, Y.; Ge, Y.; Dong, W.; Yuan, Q.; Huang, L. Phylogenomic and evolutionary dynamics of inverted repeats across *Angelica* plastomes. *BMC Plant Biol.* **2021**, *21*, 26. [CrossRef] [PubMed]
17. Li, J.; Xie, D.F.; Guo, X.L.; Zheng, Z.Y.; He, X.J.; Zhou, S.D. Comparative analysis of the complete plastid genome of five *Bupleurum* species and new insights into DNA barcoding and phylogenetic relationship. *Plants* **2020**, *9*, 543. [CrossRef]

18. Huang, R.; Xie, X.; Chen, A.; Li, F.; Tian, E.; Chao, Z. The chloroplast genomes of four *Bupleurum* (Apiaceae) species endemic to Southwestern China, a diversity center of the genus, as well as their evolutionary implications and phylogenetic inferences. *BMC Genom.* **2021**, *22*, 714. [CrossRef]
19. Yang, J.; Yue, M.; Niu, C.; Ma, X.F.; Li, Z.H. Comparative analysis of the complete chloroplast genome of four endangered herbals of *Notopterygium*. *Genes* **2017**, *8*, 124. [CrossRef]
20. Ren, T.; Li, Z.X.; Xie, D.F.; Gui, L.J.; Peng, C.; Wen, J.; He, X.J. Plastomes of eight *Ligusticum* species: Characterization, genome evolution, and phylogenetic relationships. *BMC Plant Biol.* **2020**, *20*, 519. [CrossRef]
21. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem. Bull.* **1987**, *19*, 11–15.
22. Schubert, M.; Lindgreen, S.; Orlando, L. AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. *BMC Res. Notes* **2016**, *9*, 88. [CrossRef]
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*, 1–31. [CrossRef] [PubMed]
24. 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]
25. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef] [PubMed]
26. 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]
27. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef]
28. 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]
29. Thiel, T.; Michalek, W.; Varshney, R.K.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [CrossRef]
30. 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]
31. Tan, J.B. The complete chloroplast genome of *Melanosciadium pimpinelloideum* (Apiaceae), an endemic species of China. *Mitochondrial DNA B Resour.* **2020**, *5*, 2371–2372. [CrossRef]
32. Li, Y.; Geng, M.; Xu, Z.; Wang, Q.; Li, L.; Xu, M.; Li, M. The complete plastome of *Peucedanum praeruptorum* (Apiaceae). *Mitochondrial DNA B Resour.* **2019**, *4*, 3612–3613. [CrossRef]
33. Li, L.; Geng, M.; Li, Y.; Xu, Z.; Xu, M.; Li, M. Characterization of the complete plastome of *Saposhnikovia divaricata* (Turcz.) Schischk. *Mitochondrial DNA B Resour.* **2020**, *5*, 786–787. [CrossRef] [PubMed]
34. Samigullin, T.H.; Logacheva, M.D.; Terenteva, E.I.; Degtjareva, G.V.; Vallejo-Roman, C.M. Plastid genome of *Seseli montanum*: Complete sequence and comparison with plastomes of other members of the Apiaceae family. *Biochemistry* **2016**, *81*, 981–985. [CrossRef] [PubMed]
35. 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] [PubMed]
36. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
37. Wang, D.; Zhang, Y.; Zhang, Z.; Zhu, J.; Yu, J. KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinf.* **2010**, *8*, 77–80. [CrossRef]
38. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, 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.* **2020**, *20*, 348–355. [CrossRef]
39. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef]
40. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef]
41. Stamatakis, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analysis with thousands of taxa and mixed models. *Bioinformatics* **2006**, *22*, 2688–2690. [CrossRef]
42. Ronquist, F.; Teslenko, M.; Van Der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]
43. Posada, D.; Crandall, K.A. Modeltest: Testing the model of DNA substitution. *Bioinformatics* **1998**, *14*, 817–818. [CrossRef] [PubMed]
44. Liao, C.; Downie, S.R.; Yu, Y.; He, X. Historical biogeography of the *Angelica* group (Apiaceae tribe Selineae) inferred from analyses of nrDNA and cpDNA sequences. *J. Syst. Evol.* **2012**, *50*, 206–217. [CrossRef]

45. Wen, J.; Xie, D.F.; Price, M.; Ren, T.; Deng, Y.Q.; Gui, L.J.; Guo, X.L.; He, X.J. Backbone phylogeny and evolution of Apioideae (Apiaceae): New insights from phylogenomic analyses of plastome data. *Mol. Phylogenet. Evol.* **2021**, *161*, 107183. [CrossRef] [PubMed]
46. Zhou, J.; Peng, H.; Downie, S.R.; Liu, Z.W.; Gong, X. A molecular phylogeny of Chinese Apiaceae subfamily Apioideae inferred from nuclear ribosomal DNA internal transcribed spacer sequences. *Taxon* **2008**, *57*, 402–416.
47. Zhou, J.; Gong, X.; Downie, S.R.; Peng, H. Towards a more robust molecular phylogeny of Chinese Apiaceae subfamily Apioideae: Additional evidence from nrDNA ITS and cpDNA intron (*rpl16* and *rps16*) sequences. *Mol. Phylogenet. Evol.* **2009**, *53*, 56–68. [CrossRef]
48. Palmer, J.D. Comparative organization of chloroplast genomes. *Annu. Rev. Genet.* **1985**, *19*, 325–354. [CrossRef]
49. Schwarz, E.N.; Ruhlman, T.A.; Sabir, J.S.; Hajrah, N.H.; Alharbi, N.S.; Al-Malki, A.L.; Bailey, C.D.; Jansen, R.K. Plastid genome sequences of legumes reveal parallel inversions and multiple losses of *rps16* in papilionoids. *J. Syst. Evol.* **2015**, *53*, 458–468. [CrossRef]
50. Ren, T.; Yang, Y.C.; Zhou, T.; Liu, Z.L. Comparative plastid genomes of *Primula* species: Sequence divergence and phylogenetic relationships. *Int. J. Mol. Sci.* **2018**, *19*, 1050. [CrossRef]
51. Yao, G.; Jin, J.J.; Li, H.T.; Yang, J.B.; Mandala, V.S.; Croley, M.; Mostow, R.; Douglas, N.A.; Chase, M.W.; Christenhusz, M.J.M.; et al. Plastid phylogenomic insights into the evolution of Caryophyllales. *Mol. Phylogenet. Evol.* **2019**, *134*, 74–86. [CrossRef]
52. Mohanta, T.K.; Mishra, A.K.; Khan, A.; Hashem, A.; Abd_Allah, E.F.; Al- Harrasi, A. Gene loss and evolution of the plastome. *Genes* **2020**, *11*, 1133. [CrossRef]
53. Qian, J.; Song, J.Y.; Gao, H.H.; Zhu, Y.J.; Xu, J.; Pang, X.H.; Yao, H.; Sun, C.; Li, X.E.; Li, C.Y.; et al. The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. *PLoS ONE* **2013**, *8*, e57607. [CrossRef] [PubMed]
54. Akashi, H. Codon bias evolution in *Drosophila*. Population genetics of mutation-selection drift. *Gene* **1997**, *205*, 269–278. [CrossRef]
55. Chen, S.L.; Lee, W.; Hottes, A.K.; Shapiro, L.; McAdams, H.H. Codon usage between genomes is constrained by genome-wide mutational processes. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3480–3485. [CrossRef] [PubMed]
56. Dong, S.; Ying, Z.; Yu, S.; Wang, Q.; Liao, G.; Ge, Y.; Cheng, R. Complete chloroplast genome of *Stephania tetrandra* (Menispermaceae) from Zhejiang Province: Insights into molecular structures, comparative genome analysis, mutational hotspots and phylogenetic relationships. *BMC Genom.* **2021**, *22*, 880. [CrossRef]
57. Wang, X.; Zhou, T.; Bai, G.; Zhao, Y. Complete chloroplast genome sequence of *Fagopyrum dibotrys*: Genome features, comparative analysis and phylogenetic relationships. *Sci. Rep.* **2018**, *8*, 12379. [CrossRef]
58. Tian, S.; Lu, P.; Zhang, Z.; Wu, J.Q.; Zhang, H.; Shen, H. Chloroplast genome sequence of Chongming lima bean (*Phaseolus lunatus* L.) and comparative analyses with other legume chloroplast genomes. *BMC Genom.* **2021**, *22*, 194. [CrossRef]
59. He, P.; Huang, S.; Xiao, G.; Zhang, Y.; Yu, J. Abundant RNA editing sites of chloroplast protein-coding genes in *Ginkgo biloba* and an evolutionary pattern analysis. *BMC Plant Biol.* **2016**, *16*, 257. [CrossRef]
60. Takenaka, M.; Zehrmann, A.; Verbitskiy, D.; Härtel, B.; Brennicke, A. RNA editing in plants and its evolution. *Annu. Rev. Genet.* **2013**, *47*, 335–352. [CrossRef]
61. Pacheco, T.G.; da Silva, G.M.; de Santana, L.A.; de Oliveira, J.D.; Rogalski, J.M.; Balsanelli, E.; de Souza, E.M.; de Oliveira Pedrosa, F.; Rogalski, M. Phylogenetic and evolutionary features of the plastome of *Tropaeolum pentaphyllum* Lam. (Tropaeolaceae). *Planta* **2020**, *252*, 17. [CrossRef]
62. de Santana Lopes, A.; Pacheco, T.G.; Nimz, T.; do Nascimento Vieira, L.; Guerra, M.P.; Nodari, R.O.; de Souza, E.M.; de Oliveira Pedrosa, F.; Rogalski, M. The complete plastome of macaw palm [*Acrocomia aculeata* (Jacq.) Lodd. ex Mart.] and extensive molecular analyses of the evolution of plastid genes in Arecaceae. *Planta* **2018**, *247*, 1011–1030. [CrossRef]
63. 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]
64. Asaf, S.; Waqas, M.; Khan, A.L.; Khan, M.A.; Kang, S.M.; Imran, Q.M.; Shahzad, R.; Bilal, S.; Yun, B.W.; Lee, I.J. The complete chloroplast genome of wild rice (*Oryza minuta*) and its comparison to related species. *Front. Plant Sci.* **2017**, *8*, 304. [CrossRef] [PubMed]
65. 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] [PubMed]
66. Kaila, T.; Chaduvla, P.K.; Saxena, S.; Bahadur, K.; Gahukar, S.J.; Chaudhury, A.; Sharma, T.R.; Singh, N.K.; Gaikwad, K. Chloroplast genome sequence of Pigeonpea (*Cajanus cajan* (L.) Millspaugh) and *Cajanus scarabaeoides* (L.) Thouars: Genome organization and comparison with other Legumes. *Front. Plant Sci.* **2016**, *7*, 1847. [CrossRef]
67. 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]
68. Guisinger, M.M.; Kuehl, J.V.; Boore, J.L.; Jansen, R.K. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: Rearrangements, repeats, and codon usage. *Mol. Biol. Evol.* **2011**, *28*, 583–600. [CrossRef]
69. Sabir, J.; Schwarz, E.; Ellison, N.; Zhang, J.; Baeshen, N.A.; Mutwakil, M.; Jansen, R.; Ruhlman, T. Evolutionary and biotechnology implications of plastid genome variation in the inverted-repeat-lacking clade of legumes. *Plant Biotechnol. J.* **2014**, *12*, 743–754. [CrossRef]
70. Song, Y.; Yu, W.B.; Tan, Y.H.; Liu, B.; Yao, X.; Jin, J.; Michael, P.; Yang, J.B.; Corlett, R.T. Evolutionary comparisons of the chloroplast genome in Lauraceae and insights into loss events in the Magnoliids. *Genome Biol. Evol.* **2017**, *9*, 2354–2364. [CrossRef]

71. Cauz-Santos, L.A.; Da Costa, Z.P.; Callot, C.; Cauet, S.; Zucchi, M.I.; Bergès, H.; Van den Berg, C.; Vieira, M.L.C. A repertory of rearrangements and the loss of an inverted repeat region in *Passiflora* chloroplast genomes. *Genome Biol. Evol.* **2020**, *12*, 1841–1857. [CrossRef]
72. Yang, Z.; Nielsen, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* **2000**, *17*, 32–43. [CrossRef]
73. Makalowski, W.; Boguski, M.S. Evolutionary parameters of the transcribed mammalian genome: An analysis of 2820 orthologous rodent and human sequences. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 9407–9412. [CrossRef] [PubMed]
74. Hao, D.C.; Chen, S.L.; Xiao, P.G. Molecular evolution and positive Darwinian selection of the chloroplast maturase *matK*. *J. Plant Res.* **2010**, *123*, 241–247. [CrossRef] [PubMed]
75. Allahverdiyeva, Y.; Mamedov, F.; Mäenpää, P.; Vass, I.; Aro, E.M. Modulation of photosynthetic electron transport in the absence of terminal electron acceptors: Characterization of the *rbcL* deletion mutant of tobacco. *Biochim. Biophys. Acta* **2005**, *1709*, 69–83. [CrossRef]
76. Kikuchi, S.; Bédard, J.; Hirano, M.; Hirabayashi, Y.; Oishi, M.; Imai, M.; Takase, M.; Ide, T.; Nakai, M. Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* **2013**, *339*, 571–574. [CrossRef] [PubMed]
77. Fan, W.B.; Wu, Y.; Yang, J.; Shahzad, K.; Li, Z.H. Comparative chloroplast genomics of Dipsacales species: Insights into sequence variation, adaptive evolution, and phylogenetic relationships. *Front. Plant Sci.* **2018**, *9*, 689. [CrossRef]
78. Hilu, K.W.; Black, C.M.; Oza, D. Impact of gene molecular evolution on phylogenetic reconstruction: A case study in the rosids (Superorder Rosanae, Angiosperms). *PLoS ONE* **2014**, *9*, e99725. [CrossRef]
79. Dong, W.P.; Xu, C.; Li, C.H.; Sun, J.H.; Zuo, Y.J.; Shi, S.; Cheng, T.; Guo, J.J.; Zhou, S.L. *ycf1*, the most promising plastid DNA barcode of land plants. *Sci. Rep.* **2015**, *5*, 8348. [CrossRef]
80. Kress, W.J.; Wurdack, K.J.; Zimmer, E.A.; Weigt, L.A.; Janzen, D.H. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8369–8374. [CrossRef]
81. Nock, C.J.; Waters, D.L.; Edwards, M.A.; Bowen, S.G.; Rice, N.; Cordeiro, G.M.; Henry, R.J. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol. J.* **2011**, *9*, 328–333. [CrossRef]
82. Wen, J.; Yu, Y.; Xie, D.F.; Peng, C.; Liu, Q.; Zhou, S.D.; He, X.J. A transcriptome-based study on the phylogeny and evolution of the taxonomically controversial subfamily Apioideae (Apiaceae). *Ann. Bot.* **2020**, *125*, 937–953. [CrossRef]
83. Wolfe, K.H.; Li, W.H.; Sharp, P.M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 9054–9058. [CrossRef] [PubMed]