agronomy

MDPI

Article

Complete Chloroplast Genome Sequence of a New Variety of *Brasenia schreberi*: Genome Characteristics, Comparative Analysis, and Phylogenetic Relationships

Yue Sun †, Mengyao Li †, Junying Ma, Maolin He and Yangxia Zheng *

College of Horticulture, Sichuan Agricultural University, Chengdu 611130, China
* Correspondence: zhengyx13520@sicau.edu.cn
† These authors contributed equally to this work.

Abstract: This study sequenced and assembled the chloroplast (cp) genome of *Brasenia schreberi* cv. 'Mahu Chuncai', a novel variety of *B. schreberi* rich in nutrients with distinctive characteristics, unlike other varieties in China. The cpDNA genome of 'Mahu Chuncai' has a typical quadripartite structure, with a full length of 158,973 bp, including 88 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The phylogenetic analysis revealed that all species can be divided into three main clades. Results from inverted repeats (IR) boundary analysis revealed substantial differences between *Brasenia* and *Cabomba* species. The cpDNA genome of *B. schreberi* identified was strongly related to *Brasenia* species but appeared to be a distant relative of *Cabomba aquatica* more than other species in *Cabombaceae*. In contrast with the species from *Cabombaceae*, 'Mahu Chuncai' was a close relative of *B. schreberi* MN315507.1, which was a distant relative of *C. aquatica* MG720559.1. Furthermore, we found four potential molecular markers, i.e., *ycf1* in the IR region, *psbT* in the LSC region, and *ndhF* and *rps15* in the SSC region. Collectively, our findings confirm the phylogenetic evolution and cultivation origin of *B. schreberi*. We identified genetic characteristics and nucleotide diversity hotspots, which provides a theoretical basis for additional research on variety identification, germplasm resources, and molecular breeding of the precious vegetable.

Keywords: *Brasenia schreberi*; nutrition quality; chloroplast genome; SSR; phylogenetic evolution; genetic distance analysis

Citation: Sun, Y.; Li, M.; Ma, J.; He, M.; Zheng, Y. Complete Chloroplast Genome Sequence of a New Variety of *Brasenia schreberi*: Genome Characteristics, Comparative Analysis, and Phylogenetic Relationships. *Agronomy* **2022**, *12*, 2972. https://doi.org/10.3390/agronomy12122972

Academic Editor: Matthew Hegarty

Received: 14 October 2022
Accepted: 23 November 2022
Published: 26 November 2022

1. Introduction

Brasenia schreberi is a member of the *Brasenia* genus distributed in temperate and tropical regions such as Eastern Asia, the West Indies, Australia, and America [1]. It is a perennial aquatic plant that mainly grows in unpolluted environments characterized by rich humus and pure water [2]. Its tender bud and leaves are tender and edible with a unique flavor [3]. The leaves are covered by a thick and transparent colloid, with a crystal-clear mucus containing acidic polysaccharides, hot water-soluble polysaccharides, proteins, polyphenols, and trace elements [4]. In addition, *B. schreberi* is rich in proteins, trace elements, mineral elements, dietary fiber, and polysaccharides, among other nutrients [5]. Since the Ming Dynasty in China, *B. schreberi* has been used as a medicinal plant and was first included in the Compendium of Materia Medica (Bencao Gangmu). Research reports indicate that the extracts of *B. schreberi* have excellent antibacterial and anticancer effects, enhance immune function, and reduce blood sugar [6,7]. In China and Japan, this plant is a popular precious aquatic vegetable, and its processed products are exported to Southeast Asian countries. However, its cultivated area is highly limited because of its unique environmental growth requirements, a phenomenon that limits yield production. China has only four major production areas, including Shizhu in Chongqing province, Lichuan in Hubei province, Leibo in Sichuan province, and Hangzhou in Zhejiang province. Moreover,

the cultivation of *B. schreberi* is facing various hurdles, including mixed germplasm, intense variety degradation, and poor stress resistance.

The chloroplast (cp) plays a crucial role in the photosynthesis of green plants. It produces chemical energy and oxygen using solar energy and carbon dioxide, which regulate plant life and maintain the global ecosystem [8]. Previous studies have demonstrated that the cpDNA genome structure of most flowering plants is highly conserved, and stable with a relatively small genome [9,10]. The cpDNA genome has a base substitution rate of about 1/2 to 1/7 or 1/10 that of nuclear DNA [11–13]. The maternal inheritance of chloroplast, which is typical for angiosperms, can avoid the interference of intramolecular recombination on genetic evolution [14]. Moreover, plastid DNA has a lower substitution rate than that of nuclear DNA, which makes it more conducive to analysis of genetic diversity, origin evolution, and plant phylogeny [15]. Therefore, the cpDNA genome is an effective tool for plant phylogeny research. On average, the cpDNA genome has a quadripartite structure of covalently closed circular double-stranded DNA, containing a large single-copy sequence (SSC), a small single-copy sequence (LSC), and a pair of inverted repeat regions (IR, IRa/IRb). In some members of *Fabaceae* such as *Trifolium subterraneum*, the cpDNA genomes exhibit a tripartite structure due to the loss of an inverted repeat sequence [16].

Successful sequencing of the whole cpDNA genomes of *Nicotiana tabacum* [17] and *Marchantia polymorpha* [18] in 1986 marked the beginning of cpDNA genome research. The application of cpDNA genome was restricted, to a certain extent, by traditional chloroplast sequencing methods, such as Sanger sequencing, due to its complexity and inefficiency [19]. Advancement of second-generation high-throughput sequencing technology, coupled with the development of tools for analyzing whole plant genome DNA have allowed rapid and efficient acquisition of complete sequences of the cp genome. The sequences of the cpDNA genome have gradually improved molecular identification and classification of new and existing plant species, as well as authentication of medicinal material. For instance, numerous studies have used the ITS sequence as a molecular marker for the analysis of phylogenetic relationships among plants, such as *Spartina anglica* [20]. To establish comprehensive reference databases, many plant DNA barcoding necessitates after marker-specific are authenticated. Most DNA barcodes have focused on the use of chloroplast DNA sequences, including *matK*, *rbcL* in the coding region and *trnH-psbA*, *trnL-trnF* in the intergenic region [21]. Among them, *matK*, *rbcL*, and *trnk-psbA* are the most commonly used to develop DNA barcodes. At present, next-generation sequencing technology is used to generate relevant information on cpDNA genomes of commercial crops, including *Spathiphyllum 'Parrish'* [22] and *Camellia chuongtsoensis* [23], thus providing abundant insights into the evolutionary characteristics of the cpDNA genome and molecular breeding.

B. schreberi is a plant that relies on asexual reproduction or outcrossing to sustain itself and reproduce, therefore, it is important to breed new varieties and maintain their genetic characteristics [2,24]. Consequently, researchers have used the resulting data to develop programs for the introduction, planting and cultivation, nutrition, and physiological mechanisms of *B. schreberi*. However, few studies have described germplasm resources, molecular genetics, and breeding. Several cpDNA genomes of *B. schreberi* have been sequenced but the systematic evolutionary relationship among these *B. schreberi* and species in the family *Cabombaceae* remains unknown. This necessitates further exploration into the origin of species, its evolution, and the relationship among different species [25,26]. A deeper understanding of the chloroplast gene pool is paramount to enhancing knowledge about species evolution, genetics, and phylogenetic relationships. In this study, we sequenced and assembled the cpDNA genome of a new variety of *B. schreberi* from Leibo country, Sichuan province. Subsequently, repetitive sequencing, nucleic acid diversity, and phylogenetic relationships among species were analyzed to clarify differences among the varieties of *B. schreberi*. Overall, our findings provide insights into the characteristics of the chloroplast genomes of *B. schreberi* and the phylogenetic relationships within *Brasenia* species.

2. Materials and Methods

2.1. Preparation of Plant Material

The plant material *B. schreberi* cv. 'Mahu Chuncai' was a national product of geographical indication of Leibo County in Sichuan province. The leaves of three varieties of *B. schreberi* ('Mahu Chuncai', 'Hangzhou Chuncai' from Zhejiang province and 'Shizhu Chuncai' from Hubei province) were harvested from the cultivation base (28°23′ N, 103°47′ E) in Mahu Township, Leibo County, Sichuan Province. The plants which had grown for 3 years, had been covered with transparent colloid, were disease-free, and had tender stems, were harvested. Leaf samples were collected from 12 plants and three leaves from each plant, and preserved at −80 °C after liquid nitrogen freezing for genome sequencing.

2.2. Physiological Indicators of the Leaves from Three Varieties

A total of 30 leaves from each variety were randomly selected for morphological index measurement. The leaf length and width were measured using a ruler. The electronic scale was used to weigh the number of clean leaves. The calipers were used to measure the colloid thickness. The chlorophyll content was measured as described by Li et al. [27]. Protein content was determined using the Coomassie brilliant blue G-250 dye-binding method. Anthrone colorimetric method was used to measure the soluble sugar. Vitamin C content was measured using the Iodine titration method.

2.3. DNA Extraction and Genome Sequencing

Plant genomic DNA samples were extracted using the improved CTAB method [28]. The quality and concentration of extracted DNA samples, sampling 1%, were evaluated using agarose gel electrophoresis and a NanoDrop ND1000 spectrophotometer. The qualified DNA was interrupted by the ultrasonic mechanical method. Thereafter, fragment purification, end repair, 3′ end addition of A, and connection of sequencing connectors were performed on the fragmented DNA. The fragment size was then selected using agarose gel electrophoresis, then amplified by PCR to form a sequencing library. The qualified libraries were sequenced by combining Illumina and Nanopore platforms with pairwise sequencing (PE) read length 150 bp at Genepioneer Biotechnologies Co, Ltd. (Nanjing, China).

2.4. Genome Sequence Assembly and Annotation

First, the quality of the original data was evaluated to filter out reads inclusive of unknown nucleotides and those that were low-quality. The fastp (version 0.20.0, https://github.com/OpenGene/fastp; accessed on 5 March 2022) software was used to filter the original data and acquire high-quality reads (Q20 > 97.68 and Q30 > 93.17). Bowtie2 (v2.24, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; accessed on 5 March 2022) in very-sensitive-local pattern was used to compare the chloroplast genome database, and the sequence was used as the cpDNA sequence of the sample. Complete cpDNA genome sequence of 'Mahu Chuncai' was obtained using SPAdes (v3.10.1, http://cab.spbu.ru/software/spades/; accessed on 5 March 2022) with iterative k-mer sizes of 55, 87, 121 [29]. For genome sequence assembly, the SPAdes software was used to assemble the cpDNA sequences to obtain the SEED sequence of the cpDNA genome. Gapfiller v2.1.1 (https://sourceforge.net/projects/gapfiller/; accessed on 5 March 2022) was used to complementd GAP into scaffolds, and genome sequences were aligned to pseudo genome and underwent genome correction. According to the chloroplast structure, the revisionary pseudo genome started a coordinate rearrangement to complete the chloroplast circular genome sequence. The chloroplast genome map was drawn using OGDRAW (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html; accessed on 5 March 2022).

To improve the accuracy of annotation, two methods were used to annotate the chloroplast genome i.e., first, Prodigal (v2 6.3, https://www.github.com/hyattpd/Prodigal; accessed on 7 March 2022) annotated the CDS of chloroplasts, then Hmmer software (v3.1b2, http://www.hmmer.org/; accessed on 7 March 2022) and Aragorn (v1.2.38, http://130.235.244.92/ARAGORN/; accessed on 7 March 2022) was used to predict rRNA and

tRNA, respectively. Secondly, based on the related species published on NCBI, we extracted gene sequences of related species, then compared the assembled sequence to obtain a second annotation result using blast (v2.6, https://blast.ncbi.nlm.nih.gov/Blast.cgi; accessed on 7 March 2022). The two annotation results were manually checked for different genes to remove wrong and redundant annotations. The multi-exon boundary was determined to acquire the final annotation. The complete cpDNA genome sequence of 'Mahu Chuncai' was uploaded to the NCBI database with GenBank (https://www.ncbi.nlm.nih.gov/; accessed on 7 March 2022) under accession No. MZ328718.

Codon frequency and relative synonymous codon usage (RSCU) were calculated based on gene-coding sequences. RSCU < 1, indicates a codon used less frequently than other synonymous codons; RSCU > 1 indicates a codon used more frequently than other synonymous codons called high-frequency codon; RSCU = 1, indicates a codon with no preference [30].

2.5. Chloroplast Genome Data

2.5.1. SSR and IRS Analysis

Simple Sequence Repeats (SSR) markers are types of tandem repeat sequences with tens of nucleotides, composed of several nucleotides (generally 1–6) as repeat units. SSR markers of the chloroplast genome were referred to as cpSSR markers. The MISA software (v1.0, http://pgrc.ipk-gatersleben.de/misa/misa.html; accessed on 10 March 2022) was used for cpSSR analysis. Parameters were with 8 times or more repeats for mono-nucleotides, 5 times for di-nucleotides, and 3 times for tri-nucleotides, tetra-nucleotides, penta-nucleotides, and hexa-nucelotides. Interspersed repetitive sequence (IRS) is another type of repeat different from SSR, with decentralized distribution in the genome. The IRS was performed in vmatch (v2.3.0, http://www.vmatch.de/; accessed on 10 March 2022) combined with Perl script with the parameters set as the minimum length to 30 bp and the hamming distance to 3 [31]. A total of four identification forms existed, including forward, palindromic, reverse, and complement.

2.5.2. KaKs Calculation and Nucleotide Diversity Analysis

Non-synonymous mutation occurs when a base change causes a change in the amino acid; synonymous mutation occurs in all other cases. Natural selection influences the non-synonymous mutation. The ratio of the non-synonymous mutation rate (Ka) to the synonymous mutation rate (Ks) indicates the affected selection. First, the mafft software (v7.310, https://mafft.cbrc.jp/alignment/software/; accessed on 11 March 2022) was used to align gene sequences. To estimate the KaKs ratio that reflects gene selection pressure, the KaKs calculator software (v2.0, https://sourceforge.net/projects/kakscalculator2/; accessed on 11 March 2022) was used to calculate the Ka and Ks for each pair of homologous genes. A ratio of Ka/Ks greater than 1 indicates a positive selection, and a ratio less than 1 indicates a purifying selection effect. Nucleotide diversity (Pi) can reveal variations in the size of nucleic acid sequences among different species, and highly variable regions provide potential molecular markers for population genetics. Based on the statistics formula, nucleotide diversity was compared between each marker at the species level [32]. Homologous gene sequences of different species were globally aligned using the software Mafft (v7.310, https://mafft.cbrc.jp/alignment/software/; accessed on 12 March 2022) in–auto method and Pi values were calculated for each gene using DnaSP5 (http://www.ub.edu/dnasp; accessed on 12 March 2022).

2.5.3. IR Boundary Analysis

In plant genome evolution, the IR boundary will expand and contract, and a few genes will enter the IR region or the single-copy region. Visualizing the boundary information, the SVG module in Perl was used to compare the IR boundary of the seven species of Cabombaceae. Afterward, the CGVIEW software (http://stothard.afns.ualberta.ca/cgview_server/, RELEASE-2017_09_19; accessed on 14 March 2022) was used to demon-

strate subsequent comparisons and analyses of cpDNA genome structure for the 7 species. The Mauve software (http://darlinglab.org/mauve, mauve_snapshot_2015-02-13; accessed on 14 March 2022) was used for the whole-genome architecture rearrangement and comparison under the default algorithm settings.

2.6. Comparative Analysis of cpDNA Genomes and Related Species

2.6.1. Comparative Analysis on the Structure of Chloroplast Related Species and Chloroplast Sequence Homology Analysis

A comparative analysis of cpDNA genome structure was performed for closely related species using the default parameters of the software CGVIEW (http://stothard.afns.ualberta.ca/cgview_server/; accessed on 15 March 2022). The cpDNA genome sequence was imported into Geneious to perform genome structure and component alignment. Thereafter, we used the Mauve alignment (http://darlinglab.org/mauve; accessed on 15 March 2022) with default parameters to compare visualization results to obtain the homology analysis data [33].

2.6.2. Genetic Distance and Phylogenetic Analysis

Genetic distance analysis was performed using MAFFT (v7.427,—auto method) to align 28 cpDNA genomes of *Nymphaeaceae*, *Cabombaceae*, and *Nelumbonaceae*. The genetic distance was then calculated using the R statistical programming language. Indels were removed to assess genetic distance. The common CDS sequences were used for phylogenetic tree analysis, and each sequence was aligned using MAFFT software. The aligned data were connected end-to-end, using trimAI (v1.4. rev15) pruning. The RAxML software (v8.2.10, https://cme.h-its.org/exelixis/software.html; accessed on 16 March 2022) was used to build the maximum likelihood evolutionary tree with 1000 bootstraps and a GTRGAMMA model.

3. Results

3.1. Comparison of Morphology and Nutrition in Different Varieties of B. schreberi

As shown in Figure 1A, leaves of three varieties had significant differences. The leaves color of 'Hangzhou Chuncai', 'Mahu Chuncai', and 'Shizhu Chuncai' were dark red, light red, and light green with red edges, respectively. The values of leaf length, width, weight, and chlorophyll content of 'Mahu Chuncai' were significantly higher than that of 'Shizhu Chuncai' and 'Hangzhou Chuncai', indicating that 'Mahu Chuncai' had greater growth and yield (Figure 1B–E). 'Mahu Chuncai' had a thicker colloid, more vitamin C, and more soluble sugar content than the other two varieties (Figure 1F–H). Overall, compared with plants in other producing areas, 'Mahu Chuncai' had apparent quality advantages of thick colloids, large leaves, excellent quality, and high yield.

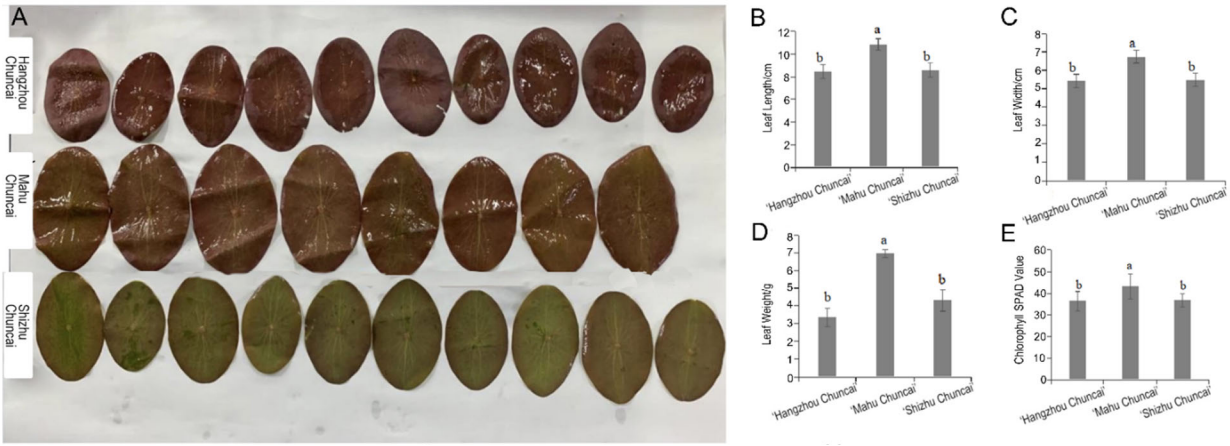

Figure 1. *Cont.*

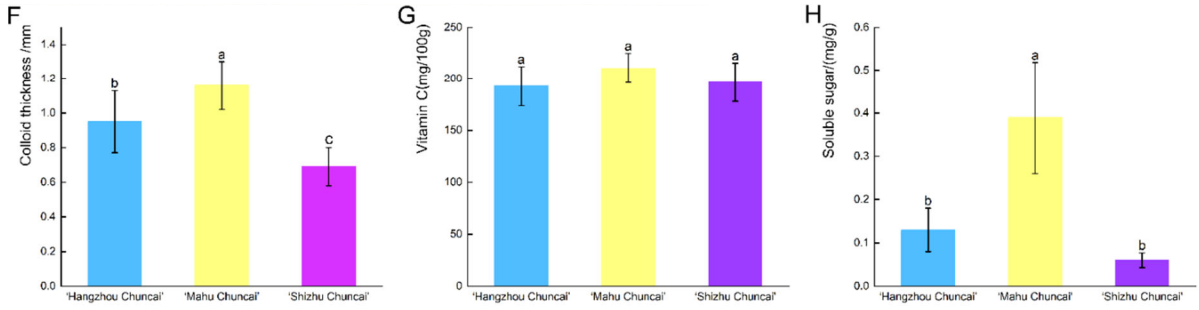

Figure 1. Comparison of morphology and nutrition in different varieties of *B. schreberi*. (**A**) Color and size of leaves of *B. schreberi*; (**B**–**E**) Determination of the growth index of varieties of *B. schreberi*; (**F**–**H**) Differences in contents of young bud of three varieties; different lowercase letters (a, b, c) indicate significant differences between different treatments at the 5% significant level. 30 leaves from each variety were randomly selected for each morphological index measurement.

3.2. Chloroplast Genome of 'Mahu Chuncai'

A total of 7,712,942,700 bases were obtained by paired-end sequencing, and 25,709,809 high-quality clean reads (the average sequencing depth was 3734X) were obtained after a strict filtering process (Table S1). Sequence assembly resulted in depths of coverage of 14.37X and covered 99.97% of the reference genome (KT705316.2). According to the Figure S1, the IGV screen capture showed that the chloroplast genome of 'Mahu Chuncai' didn't have heteroplasmy and the structure is reliable and unique. After assembly and editing, the cpDNA genome of 'Mahu Chuncai' was 158,973 bp in length with a highly conserved quadripartite structure. This genome included a large single copy sequence (LSC) of 88,779 bp, a small single copy sequence (SSC) of 19,512 bp, and a pair of inverted repeats (IRa and IRb) of 25,341 bp (Figure 2 and Table S2). Based on the annotation analysis, the genome had 133 functional genes, comprising 88 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Analysis of gene structure indicated that 18 genes contained introns, among which 15 contained only one intron, whereas 3 genes (*rps12*, *clpP*, and *ycf3*) contained two introns. Moreover, nine protein-coding and six tRNA genes had one intron, whereas no intron was found in the rRNA genes (Table 1). Notably, *Trnk-UUU* located in the LSC region contained the largest intron (2510 bp in length).

Table 1. List of genes in the *B. schreberi* cv. 'Mahu Chuncai' chloroplast genome.

Category	Gene Group	Gene Name
Photosynthesis	Subunits of photosystem I	*psaA*, *psaB*, *psaC*, *psaI*, *psaJ*
	Subunits of photosystem II	*psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*
	Subunits of NADH dehydrogenase	*NdhA* *, *ndhB* *(2), *ndhC*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhI*, *ndhJ*, *ndhK*
	Subunits of cytochrome b/f complex	*petA*, *petB* *, *petD* *, *petG*, *petL*, *petN*
	Subunits of ATP synthase	*atpA*, *atpB*, *atpE*, *atpF* *, *atpH*, *atpI*
	Large subunit of rubisco	*rbcL*
	Subunits photochlorophyllide reductase	-

Table 1. *Cont.*

Category	Gene Group	Gene Name
Self-replication	Proteins of large ribosomal subunit	*rpl14*, *rpl16* *, *rpl2* *(2), *rpl20*, *rpl22*, *rpl23*(2), *rpl32*, *rpl33*, *rpl36*
	Proteins of small ribosomal subunit	*rps11*, *rps12* **(2), *rps14*, *rps15*, *rps16* *, *rps18*, *rps19*, *rps2*, *rps3*, *rps4*, *rps7*(2), *rps8*
	Subunits of RNA polymerase	*rpoA*, *rpoB*, *rpoC1* *, *rpoC2*
	Ribosomal RNAs	*rrn16*(2), *rrn23*(2), *rrn4.5*(2), *rrn5*(2)
	Transfer RNAs	*trnA-UGC* *(2), *trnC-GCA*, *trnD-GUC*, *trnE-UUC*, *trnF-GAA*, *trnG-GCC* *, *trnG-UCC*, *trnH-GUG*, *trnI-CAU*(2), *trnI-GAU* *(2), *trnK-UUU* *, *trnL-CAA*(2), *trnL-UAA* *, *trnL-UAG*, *trnM-CAU*(2), *trnN-GUU*(2), *trnP-UGG*, *trnQ-UUG*, *trnR-ACG*(2), *trnR-UCU*, *trnS-GCU*, *trnS-GGA*, *trnS-UGA*, *trnT-GGU*, *trnT-UGU*, *trnV-GAC*(2), *trnV-UAC* *, *trnW-CCA*, *trnY-GUA*
Other genes	Maturase	*matK*
	Protease	*clpP* **
	Envelope membrane protein	*cemA*
	Acetyl-CoA carboxylase	*accD*
	c-type cytochrome synthesis gene	*ccsA*
	Translation initiation factor	*infA*
	other	-
Genes of unknown function	Conserved hypothetical chloroplast ORF	*orf42*(2), *ycf1*(2), *ycf2*(2), *ycf3* **, *ycf4*

Notes: Gene *: Gene with one introns, Gene **: Gene with two introns, Gene (2): Number of copies of multi-copy genes.

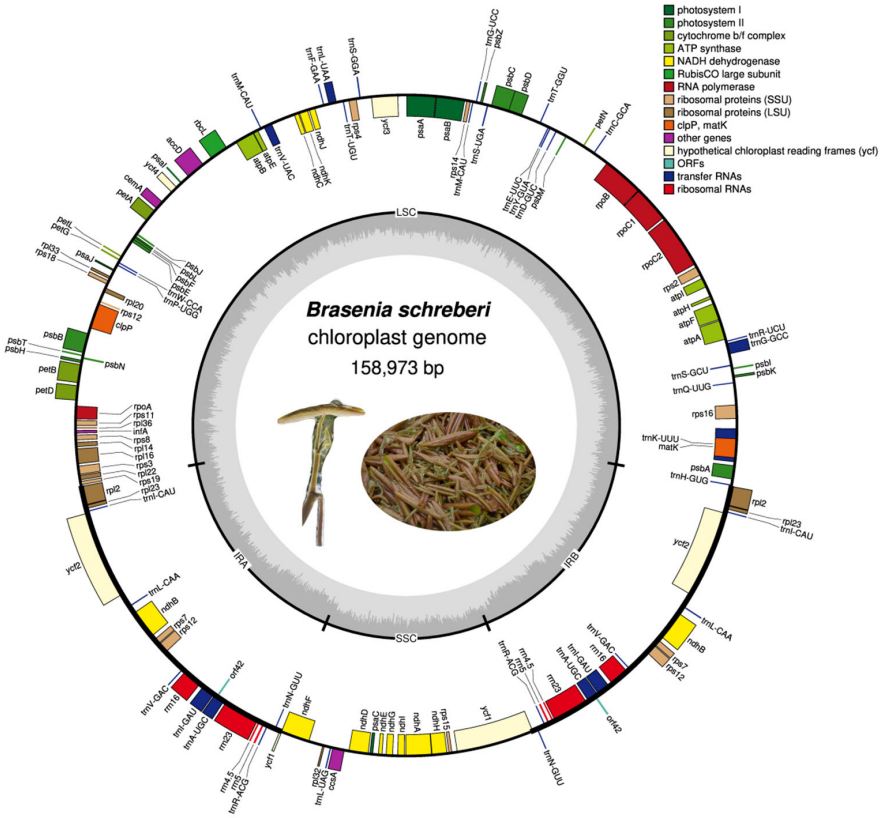

Figure 2. Physical map of the cpDNA genome of *B. schreberi* cv. 'Mahu Chuncai'. The rectangles of different width represent various genes. Genes outside the circle are forward encoded, while genes inside the circle are reverse encoded. The smallest internal grey circle in grey shows the GC content.

3.3. Codon Usage Pattern Analysis

The cpDNA genome of 'Mahu Chuncai' comprised 26,541 codons (Figure 3), among which 32 had an RSCU value greater than 1, whereas 29 codons ended with A or T. The codon encoding leucine was the most dominant, appearing 2739 times, whereas that encoding cysteine was the least common (311 times). At the same time, codons ATT and ATA appeared 1017 times and 1 time, respectively. Besides Trp which only has one codon, other amino acids have 2–6 synonymous codons. Stop codon TAA appeared 37 times, which was slightly higher than that of TAG and TGA. Moreover, ATA had the lowest relative synonymous use, with an RSCU value of 0.006, whereas ATG had the highest usage, as evidenced by an RSCU value of 3.9692. ATA and ATG are synonymous codons that encode Met. Overall, this indicates that the codon of the cpDNA gene of *B. schreberi* prefers codons ending with A or T.

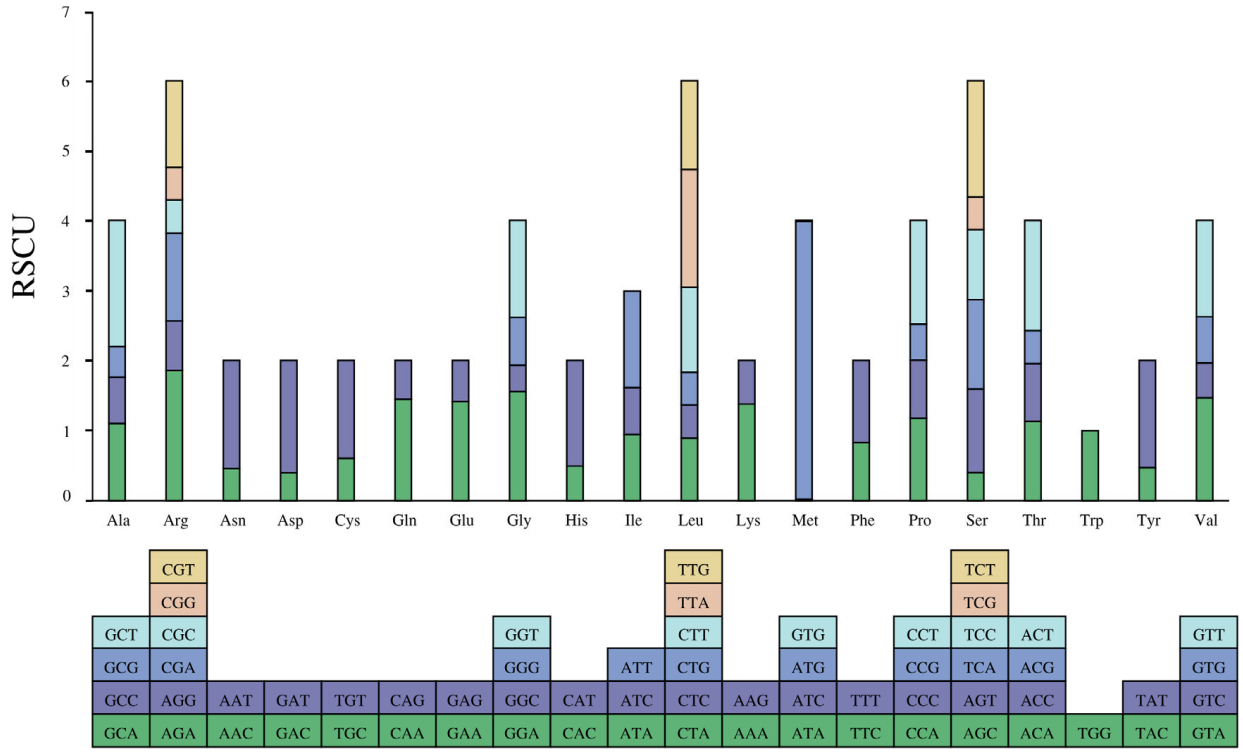

Figure 3. Relative synonymous codon usage of amino acids in cpDNA genome of *B. schreberi*. Rectangles with different colors refer to different amino acids codons, and RSCU value depends on codon repeating times. Arg, Leu, Ser all have six codons, while Trp has the least one codon.

3.4. Analysis of Repeat Sequences

Chloroplast repeats regulate genome rearrangement and recombination [34]. Simple sequence repeat (SSR) markers are types of tandem repeat sequences characterized by tens of nucleotides, which comprise several nucleotides (generally 1–6) as repeat units [35]. In the present study, 170 SSR sites were found in the chloroplast gene of *B. schreberi*. These included 71 mononucleotides, nine dinucleotides, 77 trinucleotides, and 13 tetranucleotides, with a length of 8–20 bp. Among the mononucleotide repeats, 66 sites consisted of A/T, 33 T repeats, and 34 A repeats, whereas G/C only occurred in five sites (Figure 4A). The dinucleotide sequence comprised AT/TA repeats, appearing nine times, followed by the trinucleotide sequence ATT/TTA repeats which appeared 11 times, indicating that the base composition of SSRs has a preference for AT, which is consistent with the fact that AT content (60.95%) was significantly high in the chloroplast genome. In addition, SSRs are mostly located in the intergenic region, with proportions of 19%, 17.3%, and 63.7% in SSC, IR, and LSC regions, respectively. Interspersed repetitive sequence (IRS), another type of

repeat different from SSR, exhibited a decentralized distribution in the genome. A total of 82 IRSs were detected, including 66 forward repeats (F), 12 palindromic repeats (P), two reverse repeats (R), and two complement repeats (C). The length of IRS ranged between 30–68 bp (Figure 4B). P was the longest sequence, with a length of 25,341bp. The F type (with a length of 31 bp) had the most repeats, appearing eight times, followed by 35 bp which appeared seven times.

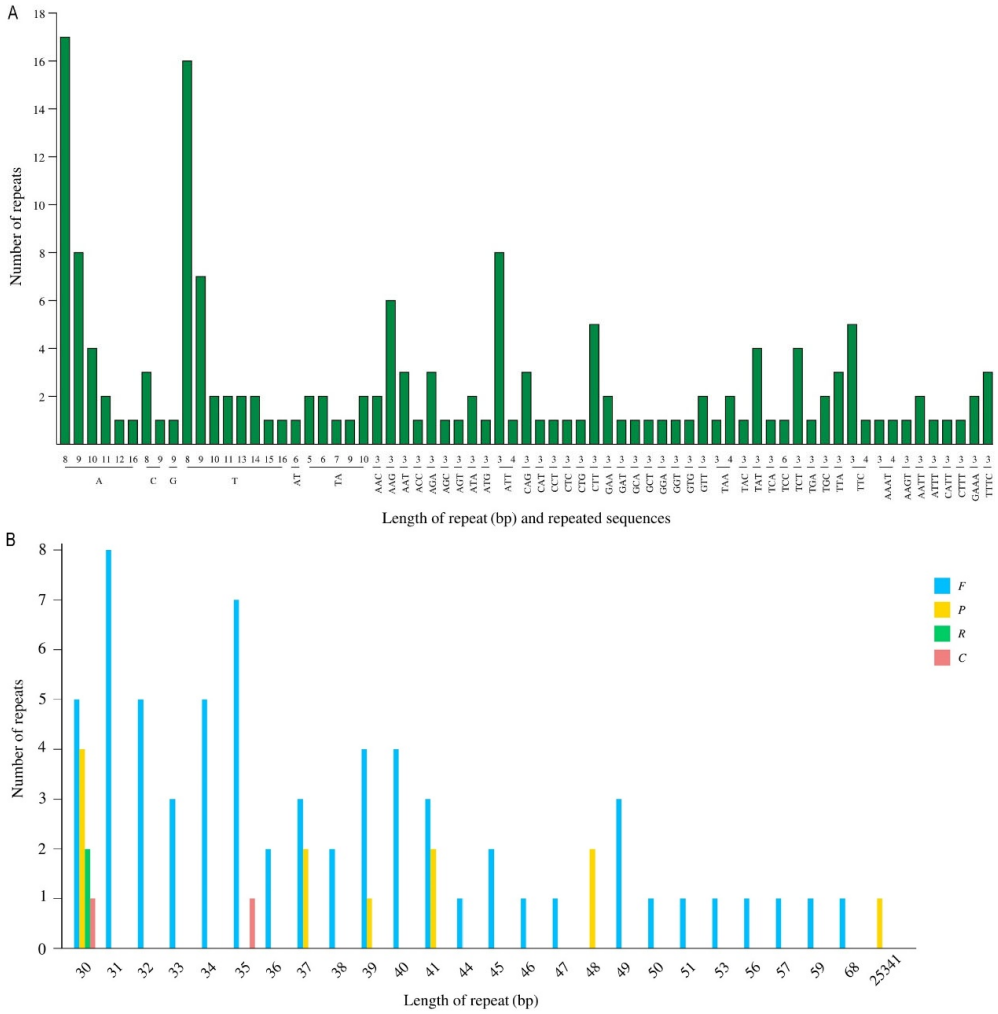

Figure 4. Analysis of repeat sequences in chloroplast genome of 'Mahu Chuncai'. (**A**) Number and distribution of SSRs. The X-axis represents length of repeat and shows the repeated sequences or repeat unit, while the Y-axis represents numbers of SSRs of each type. Four colors of rectangle boxes on behalf of four types of repeated sequences, brown for mononucleotide repeats, blue for dinucleotides repeats, red for trinucleotides repeats, purple for tetranucleotides repeats. (**B**) Number and distribution of IRSs. Length represents the length of the repeat sequence. Type represents the type of repeat sequence. The number of each type represents the number of each type. The X-axis shows the types of dispersed repeats, while the Y-axis shows the number of dispersed repeats. F: forward repeat, P: palindrome repeat, R: reverse repeat, C: complementary repeat.

3.5. KaKs Analysis

KaKs analysis was performed to determine the relationship between 'Mahu Chuncai' and other six species from *Brasenia* and *Cabomba*. Whole gene KaKs values of the six varieties ranged between 0–0.51, which are less than 1, indicating that they have been subject to purifying selection (Tables 2 and S3). The KaKs value between 'Mahu Chuncai' and *B. schreberi* KY392763.1 was 0, implying that 'Mahu Chuncai' have similar non-synonymous or synonymous substitutions with *B. schreberi* KY392763.1. Apart from 'Mahu Chuncai' vs

KY392763.1, the number of genes selected by purifying in other single gene comparison combinations was more than positive selection effect genes. This indicates that this variety is substantially purified during evolution. Among *B. schreberi* vs KT705317.3, MG720559.1, and MG967470.1, all the three comparative combinations had gene KaKs greater than 1, indicating that some genes of these varieties may have experienced positive selection in the evolutionary process.

Table 2. KaKs analysis among *B. schreberi* and other plants.

Groups	Each Gene			All Genes		
	KaKs > 1	KaKs = 1	KaKs < 1	Ka	Ks	KaKs
'Mahu Chuncai' vs KT705316.2	0	0	1	0.01417	0.02794	0.51
'Mahu Chuncai' vs KT705317.3	1	0	66	1.03607	4.31501	0.24
'Mahu Chuncai' vs KY392763.1	0	0	0	0	0	0
'Mahu Chuncai' vs MG720559.1	1	0	62	0.99297	4.34761	0.23
'Mahu Chuncai' vs MG967470.1	1	0	65	1.03967	4.50068	0.23
'Mahu Chuncai' vs MN315507.1	0	0	1	0.89994	2.34814	0.38

3.6. Analysis of Nucleotide Diversity (Pi)

Divergent hotspots on cp genomes are used to identify closely related species and provide information on phylogeny [36]. Pi reveals the variation of nucleic acid sequences of different plants, with regions of high variation used to select potential molecular markers for population genetics. We calculated the Pi values of 111 genes of 'Mahu Chuncai' and discovered that 86 genes had Pi values ranging from 0.00057 to 0.06282. Notably, *ycf1*, located in IR the region had the highest Pi value (0.06282), followed by *rps15* and *ndhF* located in the SSC region, and *psbT* located in the LSC region (Figure 5 and Table S4). Genes in the SSC region had the highest average Pi value, followed by those in the LSC and IR regions, indicating that the SSC region has the highest mutation level.

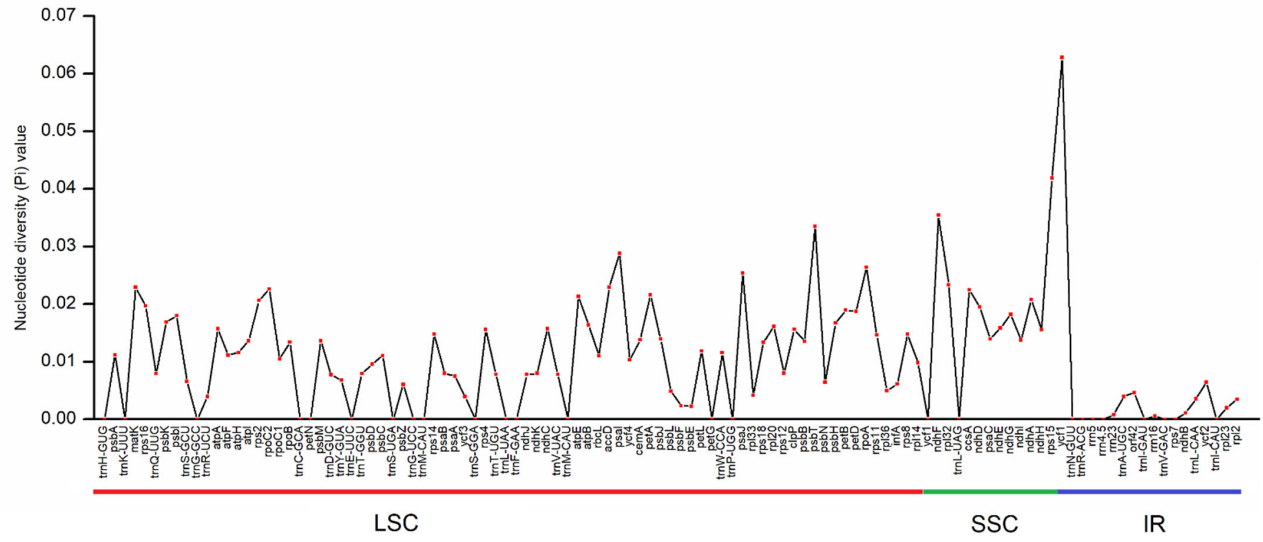

Figure 5. Nucleotide diversity of chloroplast genome of *B. schreberi.* The gene name and the area they existed are indicated in the X-axis, while the Pi value is shown in the Y-axis. The red point represent genes, and the location of points show the Pi value.

3.7. Contraction and Expansion of IR Regions

IRscope-based analysis revealed that the chloroplast genome has a circular structure, made of LSC, SSC, IRa, and IRb, with four boundaries, including LSC-IRb, IRb-SSC, SSC-IRa, and IRa-LSC [37]. As illustrated in Figure 6, analysis of the IR boundary across seven species from *Brasenia* and *Cabomba* in *Cabombaceae* revealed significantly different IR regions of species, including replacement across of *ycf1, trn23,* and *rrn4* genes, as well as expansion and contraction of *rps19* genes. We also found marked differences in genes or gene lengths at each IR boundary, and all seven chloroplast genomes exhibited visible divergences at the IRa-LSC and IRb-SSC borders in chloroplast genomes. The genes, *rps19*, *petD*, *trnH*, *rpl*, *rpoA*, *ycf1*, *ndhF*, *trnA*, and *trn23* were present at the juncture of the LSC-IRa, IRa-SSC, SSC-IRb, and IRb-LSC borders. Apart from *B. schreberi* MN315507.1, the other cp genomes of *Brasenia* exhibited little differences in IR boundaries. Both LSC-IR and SSC-IR regions in *B. schreberi* MN315507.1 contained *rpl2* and *tmA* genes in opposite directions. The difference between the two inverted repeat regions was that IRb-SSC spanned the gene *trn23*, whereas no evidence of the same gene was observed in SSC-IRa. Across the seven cp genomes, *B. schreberi* MN315507.1 exhibited different genes at the IRb-SSC and SSC-IR boundaries compared to that of the other species. Notably, the IR boundaries of the three species of the genus *Cabomba* were identical, except for *Cabomba caroliniana* KT705317.3.

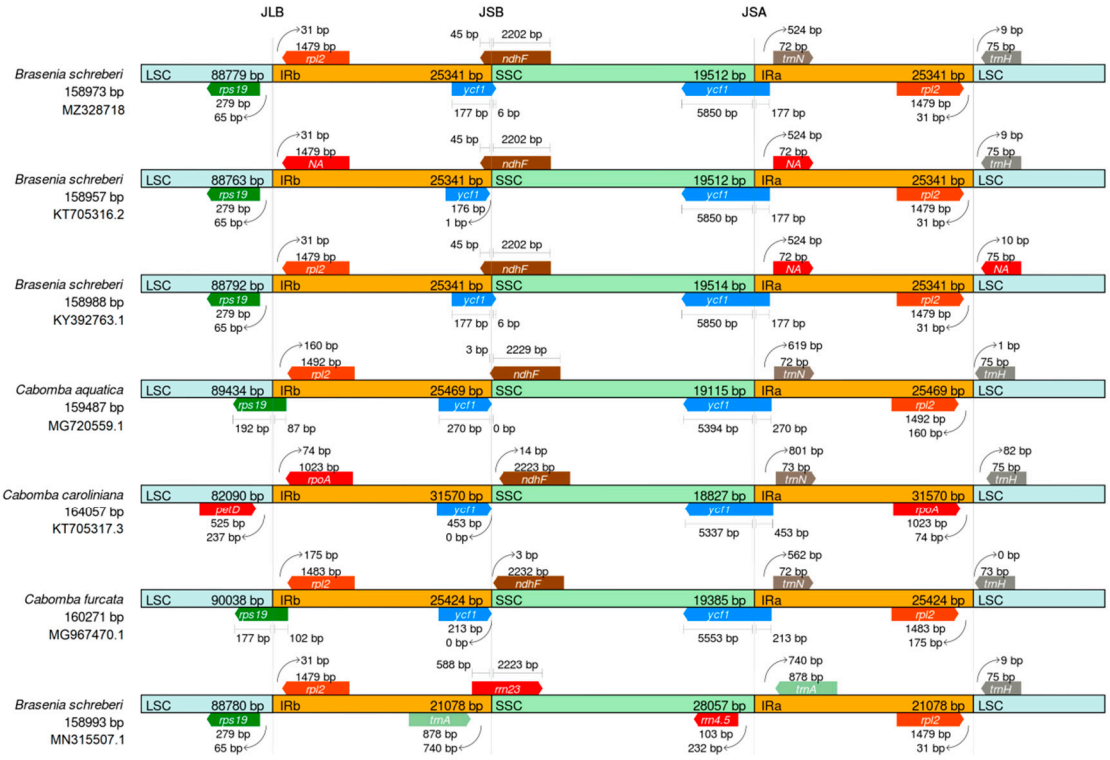

Figure 6. Boundary analysis of LSC, SSC, IRa-IRb regions of chloroplast genomes of seven species of Cabombaceae. Full lines between each colored boxes represent the point of junction. Four colors represent four areas in cpDNA genome. Numbers represent the length of each gene and the distance between the ends of genes and the border sites.

3.8. Genome Comparison and Collinearity Analysis

The CGVIEW software was used to compare and analyze the cpDNA genome structure of its close-related species, with the annotated cpDNA genome sequence of 'Mahu Chuncai' as a reference (Figure 7). Results revealed a high similarity in rRNA and tRNA coding regions across different species of *Brasenia* and *Cabomba*. Moreover, there were differences in the protein-coding regions, although the difference was not significant. Analysis of chloroplast sequence homology revealed that sequences of all other species, except *C. car-*

oliniana KT705317.3 and *B. schreberi* MN315507.1, had a high degree of collinearity (Figure 8), indicating gene rearrangements in the two species. Other species recorded a substantial degree of collinearity, indicating that the above two species have gene rearrangement and certain differences from the other five plants.

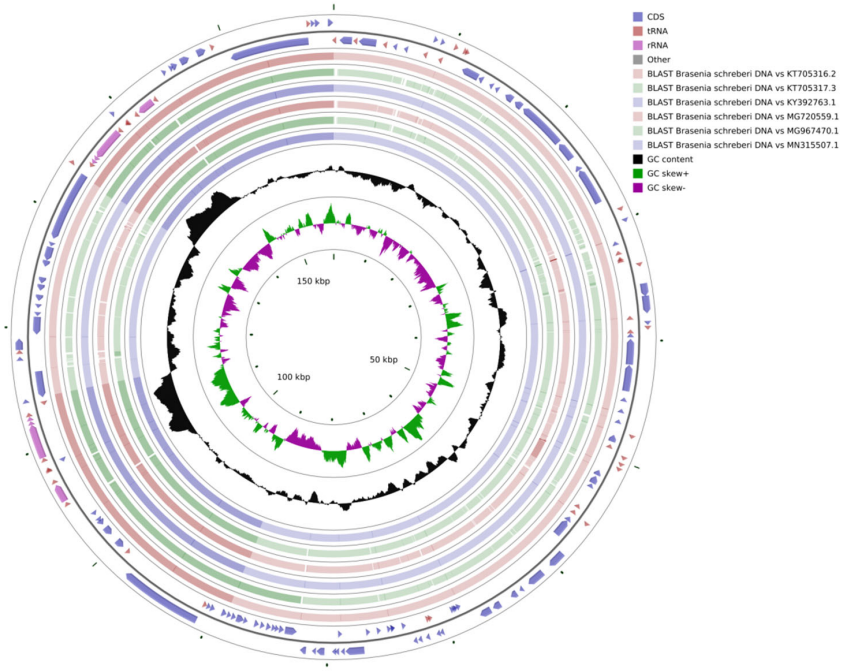

Figure 7. Comparative analysis of chloroplast structure of species from Brasenia and Cabomba of Cabombaceae family. The outer two circles in the figure describe the gene length and direction of the genome, while the inner circles represent the similarity results compared with other reference genomes, and the black circles represent the GC content.

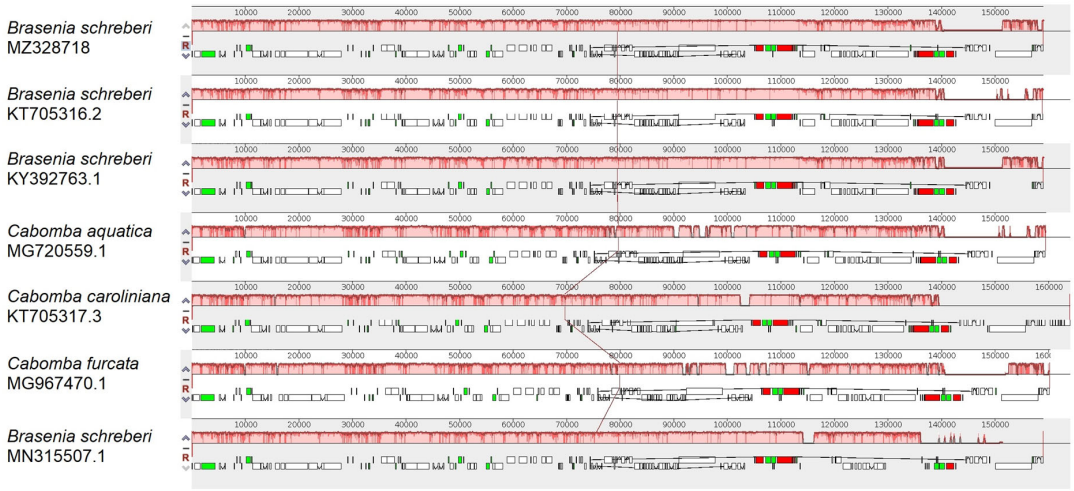

Figure 8. Chloroplast sequence homology analysis. Note: The long squares in the figure represent the similarity between genomes, and the connection between the long squares represents a collinear relationship. Short squares represent gene positions for each genome. Among them, white represents CDS, green represents tRNA, and red represents rRNA.

3.9. Phylogenetic Relationships

The maximum likelihood algorithm was used to construct a phylogenetic tree comprising 29 plant species of *Barclaya*, *Nymphaea*, *Euryale*, *Victoria*, and *Nuphar* from *Nymphaeaceae;*

Brasenia, and *Cabomba* from *Cabombaceae;* and *Nelumbo* from *Nelumbonaceae* to show phylogenetic relationships. The phylogenetic tree revealed that the main clades, i.e., *Barclaya, Nymphaea, Euryale, Nupharand, Victoria* were combined into one main clade; *Brasenia* and *Cabomba* to one main clade; and the *Nelumbo* to another main clade (Figure 9). All clades had high bootstrap (BS) supports, greater than 78%, with most of them showing 100% for robust supports. The four varieties of *B. schreberi* belonged to one monophyletic clade, whereas the three species of *Cabomba* belonged to another monophyletic clade. According to the results of contraction and expansion of IR regions and collinearity analysis, 'Mahu Chuncai' was similar to KT705316.2 and MN315507.1. However, 'Mahu Chuncai' showed a closest relative to MN315507.1 rather than KT705316.2. Notably, 'Mahu Chuncai' and *B. schreberi* MN315507.1 were supposed to have a closer relationship among four varieties of *B. schreberi*. The plants in *Nymphaeaceae* were paraphyletic clades, which indicated that plants of the family *Nymphaeaceae* had diverged. *Nuphar* of the family *Nymphaeaceae* showed tight relationships with plants of *Cabombaceae*. Moreover, we calculated genetic distances to determine differences across 28 species, and the results revealed that the genetic distances of the 28 species ranged from 0 to 0.973 (Table S5). The distance between *Nuphar* of the family *Nymphaeaceae* and 'Mahu Chuncai' was 0.0221~0.0226, and the distances between *Nuphar* and plants of *Cabombaceae* were shorter than 0.343.

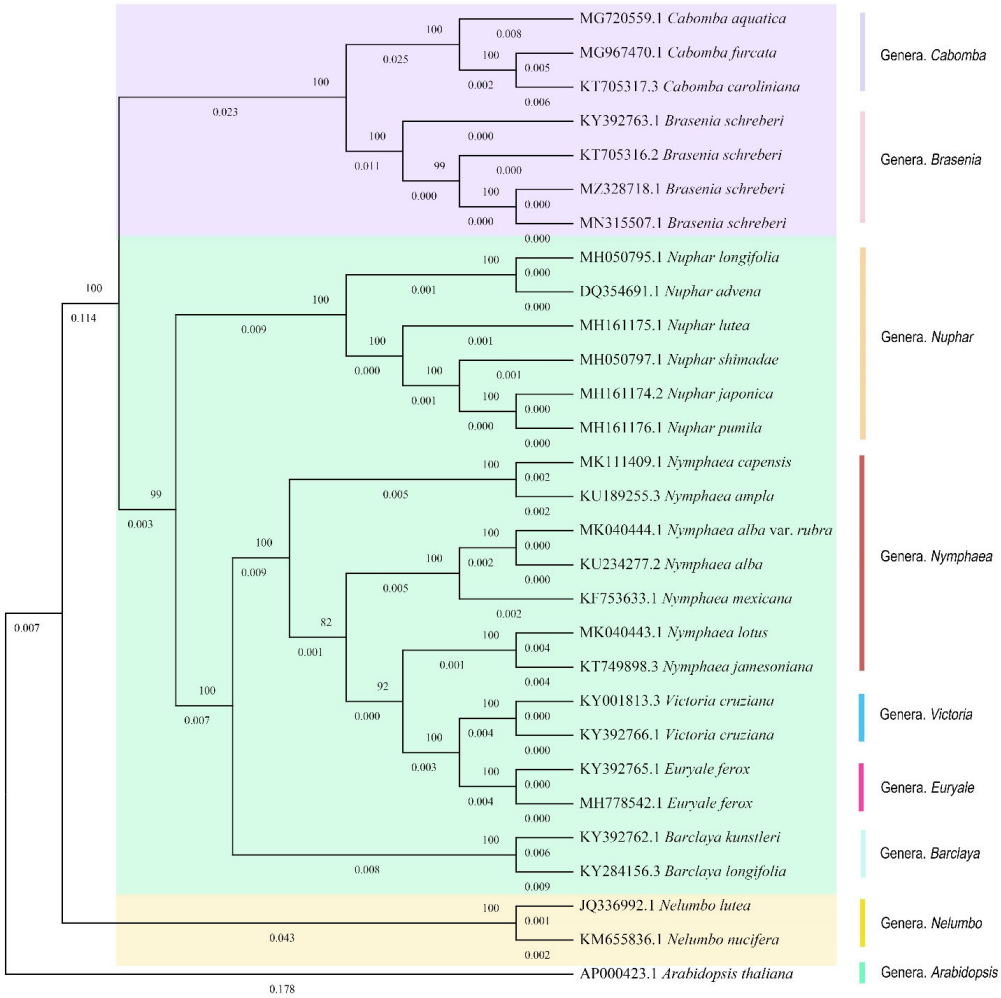

Figure 9. Maximum likelihood evolutionary tree of 29 species. The cp genome sequence of *Arabidopsis thaliana* (AP000423.1) was used as outgroup to root the trees. Number beyond each branch indicates the bootstrap support value. Three colored lines represent three families analyzed phylogenetic relationship, *Nymphaeaceae, Cabombaceae* and *Nelumbonaceae.* Nine colored lines indicate the nine genera.

4. Discussion

4.1. Analysis of the Chloroplast Genome of B. schreberi

The cpDNA genome structure of most flowering plants is characterized to be highly conserved, with a lower substitution rate than that of nuclear DNA. The perennial aquatic plant *B. schreberi*, which is a popular precious aquatic vegetable or medicinal plant, is only produced in four major production areas in China because of its unique environmental growth requirements. With a highly conserved quadripartite structure, 'Mahu Chuncai' was 158,973 bp in cpDNA genome length, which was consistent with the chloroplast genome structure of most higher plants [38,39]. A total of 18 genes contained introns was detected in gene structure, and *Trnk-UUU* located in the LSC region contained the largest intron (2510 bp in length), which was similar to those in other plants [40]. Recent chloroplast genome studies have provided data on the molecular phylogeny of plants, and this is because the chloroplast genome is characterized by a relatively slower rate of evolution compared with the nuclear genome, and the maternal inheritance of chloroplast prevents interference of intramolecular recombination on genetic evolution [41,42]. Another reason is that sequencing of chloroplast genes is relatively simple [43].

The codon of the cpDNA gene of *B. schreberi* prefers codons ending with A or T, which was consistent with previous studies showing that dicot plants prefer codons ending in A or T [44]. It has been shown that codon preferences are formed during the long-term evolution of organisms, and the preferences of codon usage vary among species and different genes [45]. However, codon preferences are shared by all organisms, and the mutation, genetic drift and selection are the main force affecting the codon preference [46]. Herein, 170 SSR sites were found in the chloroplast gene of *B. schreberi*. Among the repeats, the sites consisted of A or T were 86 times, which is consistent with the fact that AT content (60.95%) was significantly high in the chloroplast genome. The high AT content observed in SSR loci may be attributed to A or T polymers, consistent with results from previous studies [47]. The LSC (63.7%) and IR (17.3%) regions had the most and least SSRs loci, respectively, which is consistent with results reported in most angiosperms such as *Miscanthus* species [48]. A total of 82 IRSs repeats were detected in the *B. schreberi*; 49 repeat sequences were recorded in the *Houttuynia cordata* cpDNA genome [49]; and 60 repeats in the *Eupatorium catarium* [50]. This demonstrates that the repeats significantly varies across species. Four IRS types' lengths ranged between 30–68 bp with different times, indicating that repeat sequences are variable between lineages, suggesting their potential as genomic markers. The study fills a gap in the repeat sequence of the *B. schreberi* cpDNA genome.

4.2. Evolutionary and Phylogenetic Relatedness

The KaKs analysis of 'Mahu Chuncai' and other six species from *Brasenia* and *Cabomba*, showing whole gene KaKs values of the six varieties were all less than 1, indicating that they have been subject to purifying selection. *PsaI* formed a nose-shaped region of PSI located opposite the half-moon-shaped LHCI belt in the plant complex [51]. *PsaI* revealed a positive selection among KT705317.3, MG720559.1, and MG967470.1, potentially due to different habitats or distinctive needs and utilization of light intensity. This is because more unknown selective forces increase KaKs ratios, resulting in species divergence [52]. Genes in the SSC region had the highest average Pi value, indicating that the SSC region has the highest mutation level, followed by those in the LSC and IR regions, which is consistent with previous research [53]. Conversely, the IR region is the most conservative, which is consistent with GC content results that revealed that the rRNA sequence was comparatively conservative. The indel and SNP mutation events in the genome were non-random but clustered as "hotspots" [54]. Such mutational dynamics created highly variable regions in the genome. Four potential hotspots (*ycf1, ndhF, psbT, rps15*), the first four largest sequence divergence regions were found in the IR, LSC, and SSC regions, which are potential molecular markers. Previous studies identified regions with high evolutionary rates to resolve phylogenetic problems at the species level, or identify species using DNA sequences. Searching for additional regions with high evolutionary rates is critical for

plant phylogenetic analyses and DNA barcoding [55]. These molecular markers can be further used to distinguish varieties and provide a theoretical basis for germplasm resource identification and genetic breeding of *B. schreberi* varieties.

IR boundaries expand and contract during genome evolution, a phenomenon that allows certain genes to enter IR regions or single-copy regions [56]. The contraction and expansion of the IR regions modulate the variation of chloroplast genome length. Previous studies have shown that IR is the most conserved region in the chloroplast genome, with contraction and expansion of this boundary representing a common condition and important reason for chloroplast genome size variation during genome evolution [41]. Analysis of the IR boundary across seven species revealed significant differences, including replacement across of *ycf1*, *trn23*, and *rrn4* genes, as well as expansion and contraction of *rps19* genes. In summary, the conservation of the IR region of the selected variety was low, with some degree of IR boundary expansion or contraction. The selected variety had little conservation in the IR region, and all the species exhibited a certain degree of IR boundary expansion or contraction, or even gene replacement. It has been shown that the expansion and contraction events of cpDNA genomes' IR boundaries belonging to the same genus or the same family have different or large differences [57].

The compared and analyzed cpDNA genome structure of its close-related species, the 'Mahu Chuncai' results revealed a high similarity in rRNA and tRNA coding regions across different species of *Brasenia* and *Cabomba*. Apart from *C. caroliniana* KT705317.3 and *B. schreberi* MN315507.1, the other varieties had a high degree of collinearity, indicating gene rearrangements in the two species. Other species recorded a substantial degree of collinearity, indicating that the above two species have gene rearrangement and certain differences from the other five plants. In addition, the similar genomic features and gene co-linearity also demonstrated that the seven cp genomes have close taxonomic relationships. Previous studies have shown that insertion in the cpDNA genome of angiosperms occurs between closely related or within species, and the insertion event is considered a key event in plant evolution [58]. According to the results of contraction and expansion of IR regions and collinearity analysis, 'Mahu Chuncai' was similar to KT705316.2 and MN315507.1, which showed gene rearrangement and certain differences of different degrees in all these species. However, 'Mahu Chuncai' showed closest relaton to MN315507.1 rather than KT705316.2 via the results of phylogenetic relationships and genetic distance analysis, indicating that 'Mahu Chuncai' and *B. schreberi* MN315507.1 were supposed to have a closer relationship among four varieties of *B. schreberi*. The short distance between *Nuphar* and plants of *Cabombaceae* shows a significantly close phylogenetic relationship between the genus *Nuphar* and the species of family *Cabombaceae*. The present study strongly suggests that three potential scenarios of *Nuphar* for the phylogenetic position have been referred to recently for *Nuphar* as an early-diverging lineage of the *Nymphaeaceae*, as a sister to a clade formed by the *Cabombaceae* and the rest of *Nymphaeaceae*, thus forming a clade with the *Cabombaceae* [26]. In addition, *Nelumbo* (which belongs to the family *Nelumbonaceae*) had a distant relationship with other species in the *Nymphaeaceae* family. This research provides a certain experimental basis for the distinction of *Nelumbo* from the *Nymphaeaceae*, which is consistent with previous research results [59].

5. Conclusions

In conclusion, we assembled the chloroplast genome of *B. schreberi* cv. 'Mahu Chuncai', and analyzed repeat sequencing, codon preferences, and nucleic acid diversity, which provided data for cpDNA genome and evolutionary as well as phylogenetic relatedness of *B. schreberi*. The chloroplast genome of 'Mahu Chuncai' has a length of 158,973 bp, and contains some base differences compared with other varieties of *B. schreberi*. Phylogenetic analysis revealed that *B. schreberi* is closely related to *Cabomba* and *Brasenia*. Interestingly, we found four potential target regions, namely *ycf1*, *ndhF*, *psbT*, and *rps15* in the IR, LSC, and SSC regions, which can be used as potential molecular markers. These molecular markers, which have been used to scrutinize DNA sequence variations among animal and

plant species, can distinguish varieties and provide a theoretical basis for further research on germplasm resources and genetic breeding of *B. schreberi* varieties.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agronomy12122972/s1. Figure S1: IGV screen capture of the chloroplast genome mapping results; Table S1: The details of chloroplast sequencing data of *B. schreberi* cv. 'Mahu Chuncai'; Table S2: The detail characteristics of the complete cp genome of *B. schreberi*; Table S3: KaKs analysis among *B. schreberi* and other plants; Table S4: Nucleotide diversity of chloroplast genome of *B. schreberi*; Table S5: Genetic distances of 28 species.

Author Contributions: Conceptualization, M.L. and Y.Z.; data curation, Y.S.; formal analysis, Y.S. and M.H.; investigation, J.M.; software, M.H.; writing—original draft preparation, M.L. and Y.S.; writing—review and editing, M.L. and Y.Z.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Sichuan Science and Technology Program, grant number 2020ZHFP0122.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The cpDNA genome sequence of *Brasenia schreberi* cv. 'Mahu Chuncai' generated in this study were deposited at: https://www.ncbi.nlm.nih.gov/genbank/ (accessed on 7 March 2022) under the accession numbers MZ328718.

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

References

1. Kim, C.; Na, H.R.; Choi, H.K. Conservation genetics of endangered *Brasenia schreberi* based on RAPD and AFLP markers. *J. Plant Biol.* **2008**, *51*, 260–268. [CrossRef]
2. Li, Z.Z.; Gichira, A.W.; Wang, Q.F.; Chen, J.M. Genetic diversity and population structure of the endangered basal angiosperm *Brasenia schreberi (Cabombaceae)* in China. *Peer J.* **2018**, *6*, e5296. [CrossRef] [PubMed]
3. Elakovich, S.D.; Wooten, J.W. An examination of the phytotoxicity of the water shield, *Brasenia schreberi*. *J. Chem. Ecol.* **1987**, *13*, 1935–1940. [CrossRef] [PubMed]
4. Xie, C.; Li, J.; Pan, F.; Fu, J.; Zhou, W.; Lu, S.; Li, P.; Zhou, C. Environmental factors influencing mucilage accumulation of the endangered Brasenia schreberi in China. *Sci. Rep.* **2018**, *8*, 17955. [CrossRef] [PubMed]
5. Liu, P.; Liu, Y.; Yang, Y.; Chen, Z.; Li, J.; Luo, J. Mechanism of biological liquid superlubricity of *Brasenia schreberi* mucilage. *Langmuir* **2014**, *30*, 3811–3816. [CrossRef] [PubMed]
6. Li, J.; Yi, C.; Zhang, C.; Pan, F.; Xie, C.; Zhou, W.; Zhou, C. Effects of light quality on leaf growth and photosynthetic fluorescence of *Brasenia schreberi* seedlings. *Heliyon* **2021**, *7*, e06082. [CrossRef]
7. Feng, S.; Luan, D.; Ning, K.; Shao, P.; Sun, P. Ultrafiltration isolation, hypoglycemic activity analysis and structural characterization of polysaccharides from *Brasenia schreberi*. *Int. J. Biol. Macromol.* **2019**, *135*, 141–151. [CrossRef]
8. Yang, F.; Xiao, K.; Pan, H.; Liu, J. Chloroplast: The Emerging Battlefield in Plant-Microbe Interactions. *Front. Plant Sci.* **2021**, *12*, 637853. [CrossRef]
9. Daniell, H.; Lin, C.S.; Yu, M.; Chang, W.J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [CrossRef]
10. Fan, Y.; Jin, Y.; Ding, M.; Tang, Y.; Cheng, J.; Zhang, K.; Zhou, M. The complete chloroplast genome sequences of eight *Fagopyrum* species: Insights into genome evolution and phylogenetic relationships. *Front. Plant Sci.* **2021**, *12*, 799904. [CrossRef]
11. Tian, X.; Zheng, J.; Hu, S.; Yu, J. The rice mitochondrial genomes and their variations. *Plant Physiol.* **2006**, *140*, 401–410. [CrossRef]
12. 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]
13. Drouin, G.; Daoud, H.; Xia, J. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Mol. Phylogenet. Evol.* **2008**, *49*, 827–831. [CrossRef]
14. 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]
15. 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]
16. Cai, Z.; Guisinger, M.; Kim, H.G.; Ruck, E.; Blazier, J.C.; McMurtry, V.; Kuehl, J.V.; Boore, J.; Jansen, R.K. Extensive reorganization of the plastid genome of *Trifolium subterraneum (Fabaceae)* is associated with numerous repeated sequences and novel DNA insertions. *J. Mol. Evol.* **2008**, *67*, 696–704. [CrossRef]

17. Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashiuda, N.; Matsubayashi, T.; Zaita, N.; Chunwongse, J.; Obokata, J.; Yamaguchi-Shinosaki, K.; et al. The complete nucleotide sequence of the tobacco chloroplast genome: Its gene organization and expression. *EMBO J.* **1986**, *5*, 2043–2049. [CrossRef]
18. Ohyama, K.; Fukuzawa, H.; Kohchi, T.; Shirai, H.; Sano, T.; Sano, S.; Umesono, K.; Shiki, Y.; Takeuchi, M.; Chang, Z.; et al. Chloroplast gene organization deduced from complete sequence of liverwort, Marchantia polymorpha chloroplast DNA. *Nature* **1986**, *322*, 572–574. [CrossRef]
19. Bull, K.R.; Rimmer, A.J.; Siggs, O.M.; Miosge, L.A.; Roots, C.M.; Enders, A.; Bertram, E.M.; Crockford, T.L.; Whittle, B.; Potter, P.K.; et al. Unlocking the bottleneck in forward genetics using whole-genome sequencing and identity by descent to isolate causative mutations. *PLoS Genet.* **2013**, *9*, e1003219. [CrossRef]
20. Muhammad, B.L.; Ki, J.S. Hybrid origin of the invasive *Spartina anglica* inferred from chloroplast and nuclear ITS phylogenies. *Aquat. Botany* **2021**, *178*, 103484. [CrossRef]
21. Kolter, A.; Gemeinholzer, B. Plant DNA barcoding necessitates marker-specific efforts to establish more comprehensive reference databases. *Genome* **2021**, *64*, 265–298. [CrossRef] [PubMed]
22. Liu, X.F.; Zhu, G.F.; Li, D.M.; Wang, X.J. Complete chloroplast genome sequence and phylogenetic analysis of Spathiphyllum Parrish. *PLoS ONE* **2019**, *14*, e0224038. [CrossRef] [PubMed]
23. Yu, B.; Sun, Y.B.; Huang, L.L.; Xu, Y.C.; Zhao, C.Y.; Liu, X.F. The complete chloroplast genome sequence of Camellia chuongtsoensis. *Mitochondrial DNA B Resour.* **2021**, *6*, 247–249. [CrossRef] [PubMed]
24. Kim, C.; Jung, J.; Na, H.R.; Kim, S.W.; Li, W.; Kadono, Y.; Shin, H.; Choi, H.-K. Population Genetic Structure of the Endangered *Brasenia schreberi* in South Korea Based on Nuclear Ribosomal Spacer and Chloroplast DNA Sequences. *J. Plant Biol.* **2012**, *55*, 81–91. [CrossRef]
25. Yang, Y.; Li, J.W.; Wang, N.; Zou, X.X.; Zou, S.Q. The complete chloroplast genome sequence of Brasenia schreberi (Cabombaceae). Mitochondrial DNA. *Mitochondrial DNA Part B* **2019**, *4*, 3842–3843. [CrossRef] [PubMed]
26. Gruenstaeudl, M.; Nauheimer, L.; Borsch, T. Plastid genome structure and phylogenomics of Nymphaeales: Conserved gene order and new insights into relationships. *Plant Syst. Evol.* **2017**, *303*, 1251–1270. [CrossRef]
27. Li, M.Y.; Li, J.; Zhang, R.; Lin, Y.X.; Xiong, A.S.; Tan, G.F.; Luo, Y.; Zhang, Y.; Chen, Q.; Wang, Y.; et al. Combined Analysis of the Metabolome and Transcriptome to Explore Heat Stress Responses and Adaptation Mechanisms in Celery (*Apium graveolens L.*). *Int. J. Mol. Sci.* **2022**, *23*, 3367. [CrossRef]
28. Hanania, U.; Velcheva, M.; Sahar, N.; Perl, A. An improved method for isolating high-quality DNA from *Vitis vinifera* nuclei. *Plant Mol. Biol. Rep.* **2004**, *22*, 173–177. [CrossRef]
29. Menezes, A.P.A.; Resende-Moreira, L.C.; Buzatti, R.S.O.; Nazareno, A.G.; Carlsen, M.; Lobo, F.P.; Kalapothakis, E.; Lovato, M.B. Chloroplast genomes of *Byrsonima species (Malpighiaceae)*: Comparative analysis and screening of high divergence sequences. *Sci. Rep.* **2018**, *8*, 2210. [CrossRef]
30. Wang, H.; Ma, Z.H.; Mao, J.; Chen, B.H. Genome-wide identification and expression analysis of the EXO70 gene family in grape (*Vitis vinifera* L.). *Peer J.* **2021**, *9*, e11176. [CrossRef]
31. Li, M.; Song, Y.F.; Sylvester, S.P.; Sylvester, S.P.; Wang, X.R. Comparative analysis of the complete plastid genomes in *Prunus* subgenus *Cerasus* (Rosaceae): Molecular structures and phylogenetic relationships. *PLoS ONE* **2022**, *17*, e0266535.
32. Nei, M.; Li, W.H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad Sci. USA* **1979**, *76*, 5269–5273. [CrossRef]
33. Darling, A.C.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394–1403. [CrossRef]
34. Xie, D.F.; Yu, Y.; Deng, Y.Q.; Li, J.; Liu, H.Y.; Zhou, S.D.; He, X.J. Comparative analysis of the chloroplast genomes of the Chinese endemic genus Urophysa and their contribution to chloroplast phylogeny and adaptive evolution. *Int. J. Mol. Sci.* **2018**, *19*, 1847. [CrossRef]
35. Cavagnaro, P.F.; Senalik, D.A.; Yang, L.; Simon, P.W.; Harkins, T.T.; Kodira, C.D.; Huang, S.; Weng, Y. Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genomics* **2010**, *11*, 569. [CrossRef]
36. Gu, C.; Ma, L.; Wu, Z.; Chen, K.; Wang, Y. Comparative analyses of chloroplast genomes from 22 *Lythraceae* species: Inferences for phylogenetic relationships and genome evolution within Myrtales. *BMC Plant Biol.* **2019**, *19*, 281. [CrossRef]
37. Huang, Y.; Yang, Z.; Huang, S.; An, W.; Li, J.; Zheng, X. Comprehensive analysis of *Rhodomyrtus tomentosa* chloroplast genome. *Plants* **2019**, *8*, 89. [CrossRef]
38. Nguyen, H.Q.; Nguyen, T.; Doan, T.N.; Nguyen, T.; Phạm, M.H.; Le, T.L.; Sy, D.T.; Chu, H.H.; Chu, H.M. Complete chloroplast genome of novel Adrinandra megaphylla Hu species: Molecular structure, comparative and phylogenetic analysis. *Sci. Rep.* **2021**, *11*, 11731. [CrossRef]
39. Xiong, Y.; Xiong, Y.; Jia, S.; Ma, X. The complete chloroplast genome sequencing and comparative analysis of reed canary grass (*Phalaris arundinacea*) and hardinggrass (*P. aquatica*). *Plants* **2020**, *9*, 748. [CrossRef]
40. Liu, H.; Su, Z.; Yu, S.; Liu, J.; Yin, X.; Zhang, G.; Liu, W.; Li, B. Genome comparison reveals mutation hotspots in the chloroplast genome and phylogenetic relationships of *Ormosia* Species. *Biomed Res. Int.* **2019**, *2019*, 7265030. [CrossRef]
41. Dong, W.; Liu, J.; Yu, J.; Wang, L.; Zhou, S. Highly Variable Chloroplast Markers for Evaluating Plant Phylogeny at Low Taxonomic Levels and for DNA Barcoding. *PLoS ONE* **2012**, *7*, e35071. [CrossRef] [PubMed]

42. Xue, S.; Shi, T.; Luo, W.; Ni, X.; Iqbal, S.; Ni, Z.; Huang, X.; Yao, D.; Shen, Z.; Gao, Z. Comparative analysis of the complete chloroplast genome among *Prunus mume*, *P. armeniaca*, and *P. salicina*. *Hortic. Res.* **2019**, *6*, 89. [CrossRef] [PubMed]
43. Hurst, G.D.D.; Jiggins, F.M. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: The effects of inherited symbionts. *Proc. Biol. Sci.* **2005**, *272*, 1525–1534. [CrossRef] [PubMed]
44. Zhang, Y.; Shen, Z.; Meng, X.; Zhang, L.; Liu, Z.; Liu, M.; Zhang, F.; Zhao, J. Codon usage patterns across seven Rosales species. *BMC Plant Biol.* **2022**, *22*, 65. [CrossRef] [PubMed]
45. Zhao, Y.; Zheng, H.; Xu, A.; Yan, D.; Jiang, Z.; Qi, Q.; Sun, J. Analysis of codon usage bias of envelope glycoprotein genes in nuclear polyhedrosis virus (NPV) and its relation to evolution. *BMC Genom.* **2016**, *17*, 677. [CrossRef] [PubMed]
46. Pintó, R.M.; Bosch, A. The codon usage code for cotranslational folding of viral capsids. *Genome Biol. Evol.* **2021**, *13*, evab089. [CrossRef]
47. Li, M.Y.; Zhang, R.; Li, J.; Zheng, K.; Xiao, J.; Zheng, Y. Analyses of chloroplast genome of *Eutrema japonicum* provide new insights into the evolution of *Eutrema* species. *Agronomy* **2021**, *11*, 2546. [CrossRef]
48. Sheng, J.J.; Yan, M.; Wang, J.; Zhao, L.L.; Zhou, F.S.; Hu, Z.L.; Jin, S.R.; Diao, Y. The complete chloroplast genome sequences of five miscanthus species, and comparative analyses with other grass plastomes. *Ind. Crops Prod.* **2021**, *162*, 113248. [CrossRef]
49. Zhu, B.; Feng, Q.; Yu, J.; Yu, Y.; Zhu, X.; Wang, Y.; Guo, J.; Hu, X.; Cai, M. Chloroplast genome features of an important medicinal and edible plant: *Houttuynia cordata (Saururaceae)*. *PLoS ONE* **2020**, *15*, e0239823. [CrossRef]
50. Zhang, Y.; Li, L.; Yan, T.L.; Liu, Q. Complete chloroplast genome sequences of *Praxelis* (*Eupatorium catarium Veldkamp*), an important invasive species. *Gene* **2014**, *549*, 58–69. [CrossRef]
51. Plöchinger, M.; Torabi, S.; Rantala, M.; Tikkanen, M.; Suorsa, M.; Jensen, P.E.; Aro, E.M.; Meurer, J. The Low Molecular Weight Protein PsaI Stabilizes the Light-Harvesting Complex II Docking Site of Photosystem, I. *Plant Physiol.* **2016**, *172*, 450–463. [CrossRef]
52. Wen, F.; Wu, X.; Li, T.; Jia, M.; Liu, X.; Liao, L. The complete chloroplast genome of *Stauntonia chinensis* and compared analysis revealed adaptive evolution of subfamily Lardizabaloideae species in China. *BMC Genomics* **2021**, *22*, 161. [CrossRef]
53. Wang, L.; Wang, J.; He, C. Characterization and comparison of chloroplast genomes from two *sympatric Hippophae* species (Elaeagnaceae). *J. For. Res.* **2021**, *32*, 307–318. [CrossRef]
54. Du, Y.P.; Bi, Y.; Yang, F.P.; Zhang, M.F.; Chen, X.Q.; Xue, J.; Zhang, X.H. Complete chloroplast genome sequences of Lilium: Insights into evolutionary dynamics and phylogenetic analyses. *Sci. Rep.* **2017**, *7*, 5751. [CrossRef]
55. Dong, W.; Xu, C.; Cheng, T.; Lin, K.; Zhou, S. Sequencing angiosperm plastid genomes made easy: A complete set of universal primers and a case study on the phylogeny of Saxifragales. *Genome Biol. Evol.* **2013**, *5*, 989–997. [CrossRef]
56. Wang, X.; Dorjee, T.; Chen, Y.; Gao, F.; Zhou, Y. The complete chloroplast genome sequencing analysis revealed an unusual IRs reduction in three species of subfamily Zygophylloideae. *PLoS ONE* **2022**, *17*, e0263253. [CrossRef]
57. Wu, Z.; Gui, S.; Quan, Z.; Pan, L.; Wang, S.; Ke, W.; Liang, D.; Ding, Y. A precise chloroplast genome of *Nelumbo nucifera (Nelumbonaceae)* evaluated with Sanger, Illumina MiSeq, and PacBio RS II sequencing platforms: Insight into the plastid evolution of basal eudicots. *BMC Plant Biol.* **2014**, *14*, 289. [CrossRef]
58. Noutsos, C.; Richly, E.; Leister, D. Generation and evolutionary fate of insertions of organelle DNA in the nuclear genomes of flowering plants. *Genome Res.* **2005**, *15*, 616–628. [CrossRef]
59. Mukherjee, P.K.; Mukherjee, D.; Maji, A.K.; Rai, S.; Heinrich, M. The sacred lotus (*Nelumbo nucifera*)-phytochemical and therapeutic profile. *J. Pharm. Pharmacol.* **2009**, *61*, 407–422. [CrossRef]