

Article

The Complete Plastid Genome of *Rhododendron pulchrum* and Comparative Genetic Analysis of *Ericaceae* Species

Jianshuang Shen ¹, Xueqin Li ¹, Xiangtao Zhu ¹, Xiaoling Huang ¹ and Songheng Jin ^{1,2,*}

¹ Jiyang College, Zhejiang A&F University, Zhuji 311800, China; shenjianshuang18@163.com (J.S.); lxqin@zafu.edu.cn (X.L.); zxt8202@163.com (X.Z.); JYXYhxl@aliyun.com (X.H.)

² State Key Laboratory of Subtropical Silviculture, School of Forestry and Biotechnology, Zhejiang A&F University, Lin'an 311300, China

* Correspondence: shjin@zafu.edu.cn; Tel.: +86-575-87760007

Received: 19 January 2020; Accepted: 29 January 2020; Published: 31 January 2020



Abstract: Background and Objectives: *Rhododendron pulchrum* Sweet (*R. pulchrum*) belongs to the genus *Rhododendron* (*Ericaceae*), a valuable horticultural and medicinal plant species widely used in Western Europe and the US. Despite its importance, this is the first member to have its cpGenome sequenced. **Materials and Methods:** In this study, the complete cp genome of *R. pulchrum* was sequenced with NGS Illumina HiSeq2500, analyzed, and compared to eight species in the *Ericaceae* family. Results: Our study reveals that the cp genome of *R. pulchrum* is 136,249 bp in length, with an overall GC content of 35.98% and no inverted repeat regions. The *R. pulchrum* chloroplast genome encodes 73 genes, including 42 protein-coding genes, 29 tRNA genes, and two rRNA genes. The synonymous (*Ks*) and nonsynonymous (*Ka*) substitution rates were estimated and the *Ka/Ks* ratio of *R. pulchrum* plastid genes were categorized; the results indicated that most of the genes have undergone purifying selection. A total of 382 forward and 259 inverted long repeats, as well as 221 simple-sequence repeat loci (SSR) were detected in the *R. pulchrum* cp genome. Comparison between different *Ericaceae* cp genomes revealed significant differences in genome size, structure, and GC content. **Conclusions:** The phylogenetic relationships among eight *Ericaceae* species suggested that *R. pulchrum* is closely related to *Vaccinium oldhamii* Miq. and *Vaccinium macrocarpon* Aiton. This study provides a theoretical basis for species identification and future biological research of *Rhododendron* resources.

Keywords: *Rhododendron pulchrum*; chloroplast genome; phylogenetic analysis

1. Introduction

Rhododendron species belong to family *Ericaceae* and have been widely used as valuable horticultural and medicinal plants in China, Western Europe, the US, and Japan due to their beautiful vegetative forms and bright-colored flowers [1,2]. The family of *Ericaceae* consists of nine subfamilies, including exclusively autotrophic species, fully mycoheterotrophic (MH) species, and partially MH species [3]. The association between autotrophy and heterotrophy is related to drastic changes in plant morphology (such as the loss and/or reduction of vegetative organs) [4,5], physiology (for example, loss of chlorophyll and high stomatal conductivity) [5,6], genome (including rampant sequence divergence and gene loss) [7,8]. Chloroplast genome has been widely employed for studying the transition from autotrophy to heterotrophy due to its conserved size, gene content, linear gene order, and structure.

Rhododendron contains between 600–1000 species, and is among the largest genera of *Ericaceae* [9]. Previous *Rhododendron* classification was mainly based on phenotypic characteristics, and the classification systems proposed by David Chamberlain (Edinburgh Botanic Garden, UK) and

Sleumer (Germany) were the most popular. In recent years, the genetic diversity and structure of the nuclear and plastid genomes serves as another useful strategy to classify *Ericaceae* species [10,11]. The DNA sequences of nuclear genes have been adopted in studies of phylogeny, population genetics, molecular evolution, and genetic map construction in *Rhododendrons* [12].

According to the published data, the molecular marker technology and morphological analyses yield consistent results for *Rhododendron* classification [13]. Recently, complete plastid genome sequences have found their applications in plant studies, especially in the field of plant taxonomy; these include dissecting the diversity and phylogenetic relationships among different species [10,14], utilizing specific gene functions in agriculture and horticulture via genetic engineering [15], and estimating population genetics via DNA barcoding [16].

The plastid genome of higher plants is a circular molecule of double-stranded DNA, range from 72 to 217 kb in size and contain approximately 130 genes [17,18]. The plastid genomes (plastomes) of most photosynthetic land plants are between 140 to 160 kb in size and contain about 113 genes; in addition, these plastomes are conserved in size, gene content, gene order, gene structure, and genome structure [19,20]. Most plastomes contain four typical regions, including a large single-copy region (LSC) a small single-copy (SSC) region, and two identical inverted repeat regions (IR) [14]. Plastomes can provide valuable genetic information for studies on species identification, phylogeny, biology, and photosynthetic gene degradation [10,21–23] owing to their conserved structure and relatively high substitution rate [24].

The complete plastome sequences of eight *Ericaceae* species are currently available at the National Center for Biotechnology Information (NCBI); these include six fully MH species (*Monotropeoideae*) and two exclusively autotrophic species (*Vaccinioideae*). By contrast, most previous studies on *Rhododendron* have focused on their biology and physiology, whereas their plastid genomes have not been reported [25]. As an important horticultural species in the genus *Rhododendron*, *Rhododendron pulchrum* (*R. pulchrum*) is widely distributed in the temperate regions of Europe, Asia and North America [26]. Herein, we report the exhaustively analyzed cp genome of *R. pulchrum* previously published as a brief note with minimal analysis [27]. The size and structure of the *R. pulchrum* cp genome were analyzed, and the *Ka/Ks* ratio was determined to study the relationships between *Ericaceae* species and their closest autotrophic relatives, as well as to infer their evolutionary patterns. The cp genome information reported here is valuable for future phylogenetic and evolution studies, molecular marker development, and genetic improvement through genetic engineering in *Rhododendron*.

2. Materials and Methods

2.1. Sampling, DNA Extraction and Sequencing

Fresh leaves were collected from a single *R. pulchrum* tree grown in the Jiyang College of Zhejiang A&F University, Zhejiang, China. The harvested tissues were stored at the *Rhododendron* Germplasm Resource nursery in Hangzhou Botanical Garden (N30°15'8.65", E120°7'11.09") under accession number HZ041286. Total genomic DNA was extracted using the modified CTAB method [28]. The integrity and quality of the DNA were evaluated using agarose gel electrophoresis and a one drop spectrophotometer (OD-1000, Shanghai Cytoeasy Biotech Co., Ltd., Shanghai, China). The whole cp genome of *R. pulchrum* was sequenced via 250 bp pair-ended sequencing on an Illumina HiSeq 2500 Platform (Manufacturer, Nanjing, China), and at least 3.68 GB of clean sequencing data were obtained.

2.2. Chloroplast Genome Assembly, Gene Annotation and Plastomes Analysis

The reads were de novo assembled into complete cp genome by the NOVOPlasty 3.7.2 using genomes of the *Ericaceae* species as references [29]. CpGAVAS 2 (<http://www.herbalgenomics.org/cpgavas>) was used to annotate the sequences; DOGMA (<http://dogma.cccb.utexas.edu/>) and BLAST were used to manually check the annotation results [30,31]. The complete and annotated cp genome sequences were deposited to GenBank (accession number MN182619). A physical map of the genome

was generated by OGDRAWv1.2 (<http://ogdraw.mpimp-golm.mpg.de/>) and the online software REPuter [32,33] was used to identify repeat sequences.

Relative synonymous codon usage (RSCU) and codon usage were examined by CodonW 1.4.4 [34]. GC content of the complete cp genomes and the coding sequences (CDS) was analyzed. MISA [35] and REPuter [33] were used to visualize the simple sequence repeats (SSRs) and long repeats, respectively. (MISA. Available from: <http://pgrc.ipk-gatersleben.de/misa/misa.html>; REPuter. Available from: <https://bibiserv.Cebitec.uni-bielefeld.de/reputer>). The sequences were initially aligned using MAFFT version 5 [36], the pi value of each gene were calculated through alignment of each gene CDS sequences of different species using VCFtools [37], and the ratios of non-synonymous (Ka) to synonymous (Ks) substitutions (Ka/Ks) in protein-coding genes were determined by the KaKs_Calculator.

2.3. Whole Plastid Genome Comparison

To investigate the phylogenetic position of *R. pulchrum*, MUMmer [38] was used for pairing sequence alignment of the cp genomes, and the mVISTA (<http://genome.lbl.gov/vista/mvista>) [39] program was employed for comparing the complete cp genome of *R. pulchrum* to eight other related species whose cp genomes are sequenced. These eight species were divided into two groups at the genus level, including *Vaccinieae* (*Vaccinium macrocarpon*, NC_019616.1 and *Vaccinium oldhamii*, NC_042713.1) and *Monotropeae* (*Monotropa hypopitys* Linn., NC_029704.1; *Allotropa virgata* Torr. & A. Gray, NC_035580.1; *Hemitomes congestum* A. Gray, NC_035581.1; *Monotropa uniflora* L., NC_035582.1; *Monotropsis odorata* Schwein. ex Elliott, NC_035583.1; and *Pityopus californicus* (Eastw.) Copeland, NC_035584.1).

2.4. Phylogenetic Analysis

The *Ericaceae* cp genomes were obtained from the Organelle Genome and Nucleotide Resources database on NCBI. The sequences were initially aligned using MAFFT [36] (version 7, <https://mat.cbrc.jp/alignment/software/>) and the resulting multiple sequence alignment was visualized and manually adjusted in BioEdit [40]. *Actinidia deliciosa* C.F.Liang & A.R.Ferguson (NC_026691.1) and *Actinidia chinensis* Planch (NC_026690.1) were used as outgroups. The phylogenetic tree was constructed by the GTRGAMMA model implemented in RAxML.

3. Results

3.1. Features of the *R. pulchrum* cp Genome

A total of 3.68 Gb of clean data consisting 12.28 million pair-end reads were produced. All reads were deposited to NCBI Sequence Read Archive (SRA) under accession number MN182619. The complete *R. pulchrum* cp genome is 136,249 bp in length (Figure 1), and it does not take the form of a typical quadripartite structure due to the lack of inverted repeats (IR). The overall GC content of the *R. pulchrum* cp genome is 35.98% (Table 1) and we identified 73 functional genes, including two rRNA genes, 29 tRNA genes, and 42 protein-coding genes (Table 2). We then estimated codon usage frequency based on the protein-coding and tRNA genes. As shown in Figure 2, the cpGenome was composed of 8,693 codons (65 different types) encoding 20 amino acids, among which leucine (Leu) was the most frequently used amino acid (948 in number, 10.90%) and cysteine (Cys) was the least abundant (76 in number, 0.87%) (Table S1). The results suggest the *R. pulchrum* cp genome prefers synonymous codons ended with A or U with a relative synonymous codon usage value (RSCU) > 1.

Table 1. Base composition of the *R. pulchrum* cp genome.

Region	A (%)	T (U) (%)	C (%)	G (%)	A + T (%)	G + C (%)
Cp genome	31.16	32.85	18.66	17.32	64.01	35.98

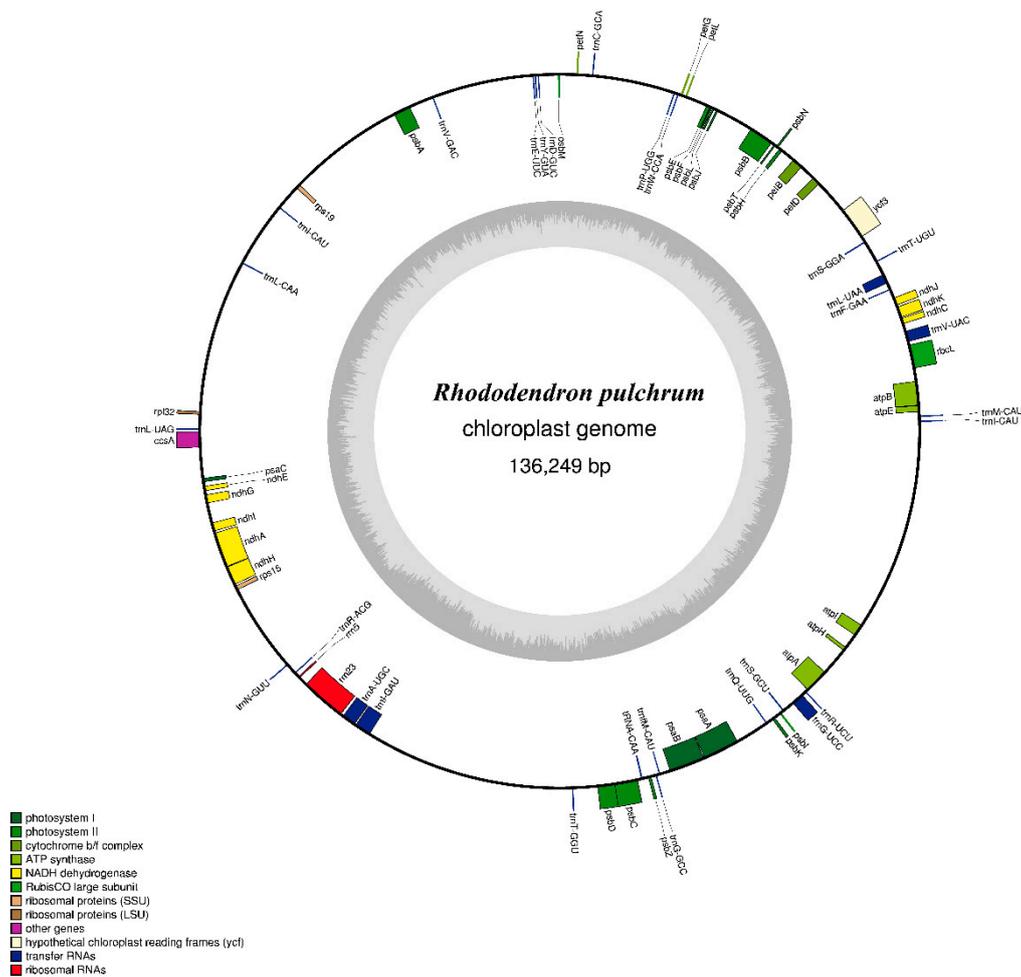


Figure 1. Gene map of the *R. pulchrum* cp genome. (Color figure online). Genes reside in the inside and outside of the outer circle are in the forward and reverse directions, respectively. The dark and light gray bars in the inner circle denote G+C and A+T contents, respectively.

Table 2. Annotated genes of the *R. pulchrum* cp genome.

Function	Genes
RNAs, transfer	<i>trnA-UIG*</i> , <i>trnA-CAA</i> , <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnE-UUC</i> , <i>trnF-GAA</i> , <i>trnG-UCC*</i> , <i>trnI-GAU*</i> , <i>trnI-CAU</i> , <i>trnL-CAA</i> , <i>trnL-UAG</i> , <i>trnL-UAA*</i> , <i>trnM-CAU</i> , <i>trnN-GUU</i> , <i>trnP-UGG</i> , <i>trnQ-UUG</i> , <i>trnR-ACG</i> , <i>trnR-UCU</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnT-UGU</i> , <i>trnT-GGU</i> , <i>trnV-UAC*</i> , <i>trnV-GAC</i> , <i>trnW-CCA</i> , <i>trnY-GUA</i>
RNAs, ribosomal	<i>rrn5</i> , <i>rrn23</i>
Small subunit of ribosome	<i>rps19</i> , <i>rps15</i>
Large subunit of ribosome	<i>rpl32</i>
ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpH</i> , <i>atpI</i>
Photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>ycf3**</i>
Photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbN</i> , <i>psbT</i> , <i>psbZ</i>
Calvin cycle	<i>rbcl</i>
Cytochrome complex	<i>petB</i> , <i>petD</i> , <i>petG</i> , <i>petL</i> , <i>petN</i>
NADH dehydrogenase	<i>ndhA*</i> , <i>ndhC</i> , <i>ndhE</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
Others (C-type cytochrome synthesis gene)	<i>ccsA</i>

Note: * Genes with one intron; ** genes with two introns.

3.5. Comparative Analysis of Gene Content and Genome Structure

Differences in cp sequences can help to infer the gene flow between species [41]. Thus, the complete *R. pulchrum* cp genome was compared with those of exclusively autotrophic and mycoheterotrophic species of *Ericaceae*, respectively. As shown in Table S5, gene content of *R. pulchrum* is quite distinct from the other eight *Ericaceae* species. We determined that while 9–12 (*rps*) genes were found among the other eight species, only two (*rps15* and *rps19*) were found in *R. pulchrum*. *rps15* was only annotated in the three exclusively autotrophic species (*R. pulchrum*, *V. macrocarpon* and *V. oldhamii*); *rps19* was annotated in all nine *Ericaceae* species. There have 5–8 (*rpl*) genes were found among the other eight species, only one gene (*rpl32*) were found in *R. pulchrum*. There have 4 (*rrn*) genes were found among the other eight species, only two gene (*rrn23* and *rrn5*) were found in *R. pulchrum*. However, gene content of *R. pulchrum* has common characteristics with the other eight species. For example, no *rpo* genes coding RNA polymerase subunits were found in *R. pulchrum* and six species of *Monotropeae*, while 4 genes were found in two species of *Vaccinieae*; 28–30 (*trn*) genes coding Transfer RNA were found in three exclusively autotrophic species (*R. pulchrum*, *V. macrocarpon* and *V. oldhamii*), while only 15–18 genes were found in six species of *Monotropeae*. Photosynthesis-related genes (*psb*, *psa*, *pet*, *atp*, *rbc*, *ndh*) were lost in the cp genomes of the six nonphotosynthetic species of *Monotropeae*. Moreover, 37, 41 and 44 photosynthesis-related genes were found in genome of *R. pulchrum*, *V. macrocarpon* and *V. oldhamii*, respectively, with 34 of these genes in common. Four genes (*atpF*, *ndhG*, *ndhK* and *PsbZ*) were lost in *V. macrocarpon*, and seven genes (*atpF*, *ndhB*, *ndhD*, *ndhF*, *PetA*, *psal* and *psaJ*) were lost in *R. pulchrum*. Besides self-replication and photosynthesis related genes, there have eight other genes (*accD*, *ccsA*, *cemA*, *clpP*, *infA*, *lhbA*, *matK* and *rp3*) were found in the other eight *Ericaceae* species. Among which, only one gene (*ccsA*) were found in genome of *R. pulchrum*, and three genes (*accD*, *clpP* and *infA*) were only found in six nonphotosynthetic species of *Monotropeae*.

As shown in (Figure 4, Figure 5 and Figure S1, the *R. pulchrum* cp genome is quite distinct from the eight *Ericaceae* species. Specifically, the *R. pulchrum* cp genome is more divergent than mycoheterotrophic *Ericaceae* species; sequence diversity is higher in the noncoding than coding regions (Figures 4 and 5). Only one gene (*trnL-CAU*) in *R. pulchrum* exhibited higher similarity to that in the six mycoheterotrophic *Ericaceae* species (Figure 4A). By contrast, twenty *R. pulchrum* cp genes showed relatively higher similarity to those of exclusively autotrophic species of *Ericaceae* (Figure 4B). Together, these data revealed a high level of genetic variation among species of different genera within *Ericaceae*, especially between mycoheterotrophic (*Monotropeae*) and exclusively autotrophic species (*Vaccinieae* and *Rhododendron*).

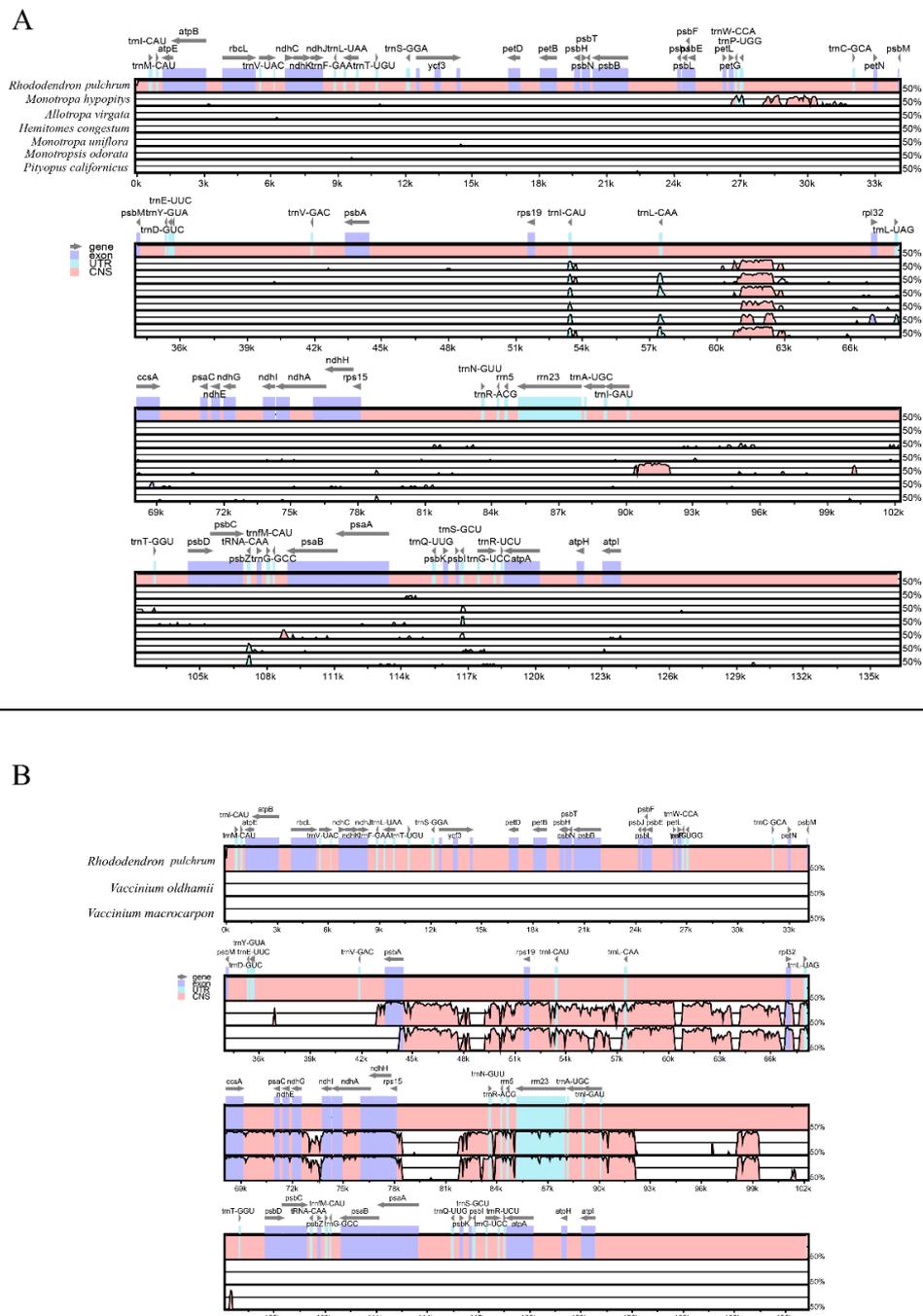


Figure 4. Comparison of the eight complete cp genomes of *Ericaceae* species. The grey arrows and thick black lines above the alignment denote gene orientations. The Y-axis represents sequence similarity. **A.** Comparison between the complete cp genomes of *R. pulchrum* and *Monotropeae* (*Monotropa hypopitys*, NC_029704.1; *Allotropa virgata*, NC_035580.1; *Hemitomes congestum*, NC_035581.1; *Monotropa uniflora*, NC_035582.1; *Monotropopsis odorata*, NC_035583.1; and *Pityopus californicus*, NC_035584.1). **B.** Comparison between the complete *R. pulchrum* and *Vaccinieae* (*Vaccinium macrocarpon*, NC_019616.1 and *Vaccinium oldhamii*, NC_042713.1) cp genomes. The cp genome of *Rhododendron pulchrum* was used as the reference.

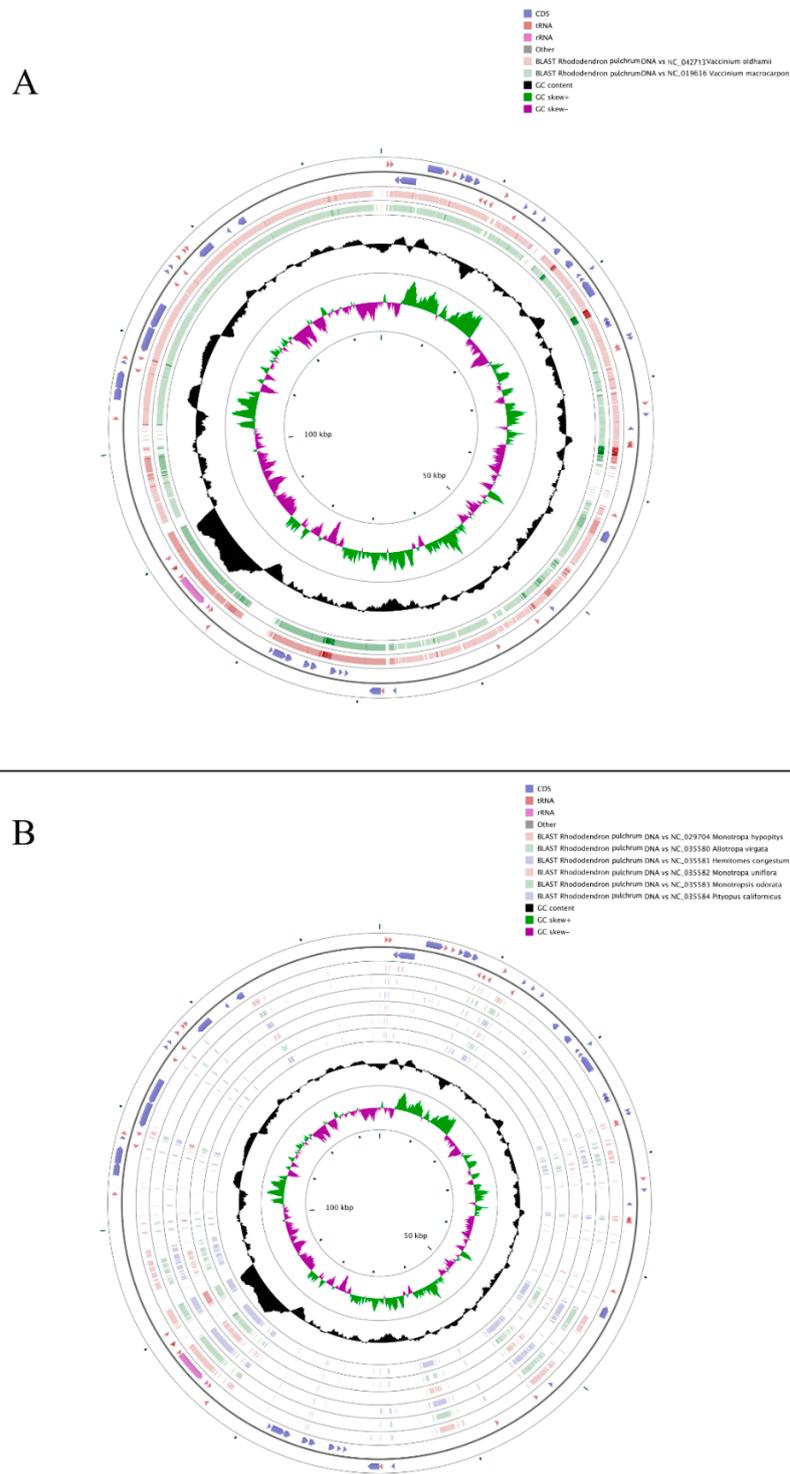


Figure 5. Comparison between the cp genomes of *R. pulchrum* and related species. The two outermost circles describe gene length and the direction of genome; the two inner circles represent the similarity between the *R. pulchrum* cp genome and the reference genomes of *Ericaceae* species; the black circle shows GC content. **A.** The alignment of *Vaccinium macrocarpon*, *Vaccinium oldhamii*, and *Rhododendron pulchrum* cp genomes. **B.** The alignment of *Rhododendron pulchrum* cp genome to that of *Monotropa hypopitys* (NC_029704.1), *Allotropa virgata* (NC_035580.1), *Hemitomes congestum* (NC_035581.1), *Monotropa uniflora* (NC_035582.1), *Monotropis odorata* (NC_035583.1), and *Pityopus californicus* (NC_035584.1).

3.6. The Phylogenetic Tree of Ericaceae

Phylogenetic analysis was performed based on an alignment of concatenated nucleotide sequences of all ten angiosperm cp genomes (Figure 6). A phylogenetic tree was built by using the Gtrgamma model and the Bayesian inference (BI) method based on RAxML, with *Actinidia deliciosa* and *Actinidia chinensis* as outgroups. All relationships inferred from these cp genomes received high supports with the support values ranging between 83 and 100. It is worth noticing that the nine species from family Ericaceae did not form a clade. Six heterotrophic species in genera *Monotropeae* clustered into one clade, and the three exclusively autotrophic *Ericaceae* species (*R. pulchrum*, *V. oldhamii* and *V. macrocarpon*) formed another (Figure 6).

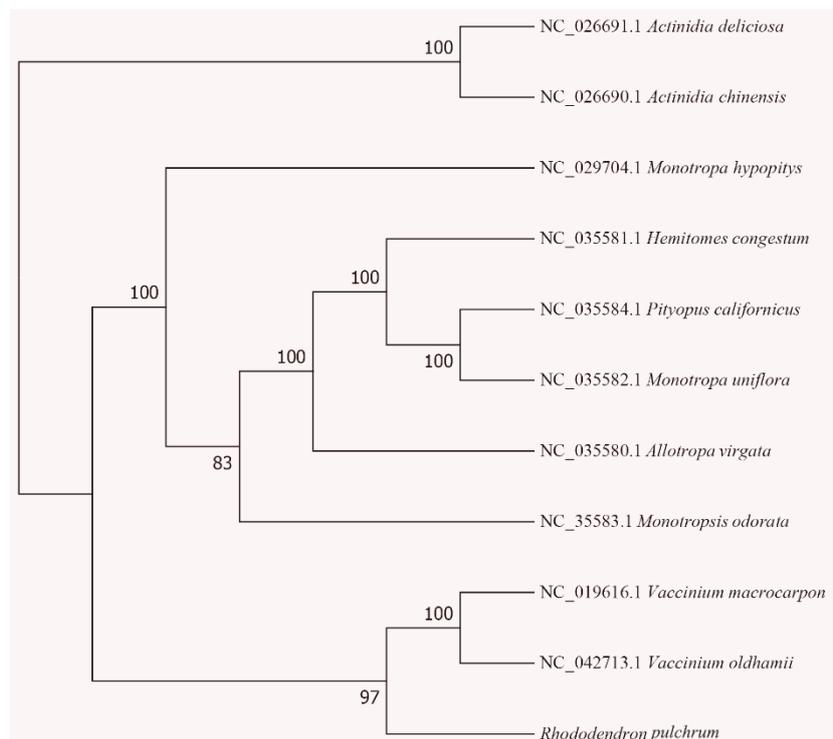


Figure 6. Phylogenetic tree of the nine *Ericaceae* species based on complete cp genomes. *Actinidia deliciosa* (NC_026691.1) and the *Actinidia chinensis* (NC_026690.1) were used as outgroups.

4. Discussion

The complete cp genome of *R. pulchrum* differs significantly from those of the other eight *Ericaceae* species with regard to genome size, structure, GC content, genes structure, but was similar to those of the two *Vaccinieae* species. Compared with exclusively autotrophic *Ericaceae* species (photosynthetic, *Vaccinieae* and *Rhododendron*), the cp genomes of mycoheterotrophic species (nonphotosynthetic, *Monotropeae*) are substantially smaller in size (ca. 33–41 kb) and gene content [10,11,42]. The *R. pulchrum* cp genome was found to have 73 functional genes, 110 and 133 genes were annotated in those of *Vaccinium macrocarpon* and *Vaccinium oldhamii*, whereas 40–45 genes were annotated in species of *Monotropeae* respectively—suggesting that *Ericaceae* cp genomes are highly variable. Gene content of *R. pulchrum* is quite distinct from the eight *Ericaceae* species, such as most of the self-replication related genes (*rps*, *rpl*, *rrn*) being missing compared to the other eight species, but similar as it has maintained most of the *trn* self-replicating genes. Interestingly, both *R. pulchrum* and six *Monotropeae* species are missing the *rpo* genes that are present in the two *Vaccinieae* species. Photosynthesis-related genes (*psb*, *psa*, *pet*, *atp*, *rbc*, *ndh*) were lost in the cp genomes of nonphotosynthetic *Ericaceae* species [10], while 37, 41 and 44 photosynthesis-related genes were found in genome of *R. pulchrum*, *V. macrocarpon*

and *V. oldhamii*, respectively. Thus, supporting the position of *Rhododendron* between *Vaccinieae* and *Monotropeae* during evolution. According to the published data, the plastid genomes of most photosynthetic land plants are conserved in size (140–160 kb) and display the typical quadripartite structure by showing the LSC, LSC, and two IRs [14,19,20]. However, our results indicate that the *R. pulchrum* cp genome lacks the IRs, which has also been reported in *Monotropeae* [10], *Medicago* [43,44] and *Erodium* plants (*Erodium carvifolium* HQ713469.1).

The *R. pulchrum* cp genome has an overall GC content of 35.98%, which is lower than exclusively autotrophic *Ericaceae* species including *Vaccinium oldhamii* (36.75%) and *Vaccinium macrocarpon* (36.80%) but higher than heterotrophic *Ericaceae* species (*Monotropeae*), including *Pityopus californicus* (34.37%), *Monotropa hypopitys* (34.31%), *Hemitomes congestum* (33.71%), *Allotropa virgata* (33.09%), *Monotropsis odorata* (31.20%), and *Monotropa uniflora* (28.47%). In plants, GC content is often associated with the degree of primitiveness of a taxon [45], whereas in MH species, it is often related to plastome degradation level [7]. Further, consistent with that observed with *Q. acutissima* [46], the *ycf3* and *trnL-UAA* genes in *R. pulchrum* have the largest intron and the smallest intron, respectively. The *ycf3* gene has been reported to be necessary for stabilizing the accumulation of photosystem I complexes [47], and the gain of intron is usually considered have closely relationship with the evolution of photosynthesis [46]. This result indicates that *ycf3* is likely a key player in photosynthesis in *R. pulchrum*. Thus, the complete cp genomes of genera *Vaccinieae*, *Monotropeae*, and *Rhododendron* are useful for studying *Ericaceae* species evolution, photosynthetic gene degradation, and the phylogenetic relationships among *Ericaceae* species.

In this study, a total of 221 cp SSRs were identified in *R. pulchrum*, more than that has been reported for other *Ericaceae* species (62) [11]. This may be a result of the complex genetic background of *Rhododendron*. Meanwhile, the cpSSRs identified here rarely contain tandem G or C repeats, which is in line with the published data [48]. These cpSSR markers offer an alternative to morphological features and nuclear DNA markers in defining *Rhododendron* subsections. The cpSSRs yielded from this study can be employed for genetic structure determination, as well as the diversity, differentiation, and maternity studies of *R. pulchrum* and its related species.

Compared with the eight *Ericaceae* species, only two genes *rps15* and *psbZ*—which encode the small subunit of ribosome and the core complex of photosystem II protein, respectively—showed signs of positive selection. Co-annotated genes were found in all the eight *Ericaceae* species during the transformation of *Ericaceae* from exclusively autotrophic (*Rhododendron*, *R. pulchrum*) to heterotrophic (*Monotropeae*, six species). In addition, we observed a high level of genetic diversity among *Ericaceae* genera, especially between mycoheterotrophic (*Monotropeae*) and exclusively autotrophic species (*Vaccinieae* and *Rhododendron*), which is consistent with the genomic rearrangement events in *Ericaceae* plants reported by previous studies [11]. In general, the cp genomes of species from the same family are conserved [23], and cp sequences have been successfully applied to the phylogenetic studies of angiosperms [49,50]. Our phylogenetic analysis clustered the nine *Ericaceae* species (from three genera) and two outgroup species into three distinct phylogenetic groups, including one outgroup (*Actinidia deliciosa* and *Actinidia chinensis*), one exclusively autotrophic group (*R. pulchrum*, *Vaccinium macrocarpon*, and *V. oldhamii*), and one heterotrophic group (*Monotropa hypopitys*, *Allotropa virgata*, *Hemitomes congestum*, *Monotropa uniflora*, *Monotropsis odorata*, and *Pityopus californicus*). In addition, the cp genome of *R. pulchrum* is closely related to those of *Vaccinium oldhamii* and *Vaccinium macrocarpon* species. Taken together, our results revealed the potential of *Ericaceae* family as suitable materials for studying plastid genome development and evolution.

5. Conclusions

In this study, we report the characterization of the complete cp genome of *R. pulchrum*, a horticultural and medicinal species cultivated worldwide. The *R. pulchrum* cp genome displays unique characteristics compared with species from other genera in *Ericaceae*, especially genus *Monotropeae*. The phylogenetic analysis revealed that *R. pulchrum* was closely related to species of *Vaccinium oldhamii*

and *Vaccinium macrocarpon*. The complete cp genome assembly of *R. pulchrum* provided in this research is valuable for future population genomic and phylogenomic studies, and will benefit *Rhododendron* conservation and utilization.

Supplementary Materials: The following are available online at <http://www.mdpi.com/1999-4907/11/2/158/s1>, Table S1. Codon-anticodon recognition patterns and codon usage of the *R. pulchrum* cp genome. Table S2. *Ka/Ks* ratios of the cp genes from *R. pulchrum* and related species (NA = not available). Table S3. Long repeats identified from the *R. pulchrum* cp genome (F = forward, I = inverted repeat; IGS = intergenic space). Table S4: Simple sequence repeats (SSRs) identified in the *R. pulchrum* cp genome. Table S5: Annotated genes of the nine species of Ericaceae cp genomes.

Author Contributions: The experiments were conceived and designed by S.J.; J.S., X.L., X.Z., and X.H. were involved in the collection of the study materials and participated in the DNA extraction and data analyses; J.S. wrote and S.J. revised the manuscript. All authors read and approved the final manuscript.

Funding: The work was funded by the Zhejiang Provincial Natural Science Foundation of China (grant number: LY16C160011), the National Natural Science Foundation of China (grant number: 31971641), Project of State Key Laboratory of Subtropical Silviculture, Zhejiang A & F University (grant number: ZY20180208, ZY20180308), and Jiyang College of Zhejiang A&F University under Grant (grant number: RQ1911B07).

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

References

1. Popescu, R.; Kopp, B. The genus *Rhododendron*: An ethnopharmacological and toxicological review. *J. Ethnopharmacol.* **2013**, *147*, 42–62. [[CrossRef](#)] [[PubMed](#)]
2. Chamberlain, D.; Hyam, R.; Argent, G.; Fairweather, G.; Walter, K.S. *The Genus Rhododendron: Its Classification and Synonymy*; Royal Botanic garden Edinburgh: Edinburgh, UK, 1996; p. 181.
3. Kron, K.A.; Judd, W.S.; Stevens, P.F.; Crayn, D.M.; Anderberg, A.A.; Gadek, P.A.; Quinn, C.J.; Luteyn, J.L. Phylogenetic Classification of Ericaceae: Molecular and Morphological Evidence. *Bot. Rev.* **2002**, *68*, 335–423. [[CrossRef](#)]
4. Leake, J.R. The biology of myco-heterotrophic ('saprophytic') plants. *New Phytol.* **1994**, *127*, 171–216. [[CrossRef](#)]
5. Manen, J.-F.; Habashi, C.; Jeanmonod, D.; Park, J.-M.; Schneeweiss, G.M. Phylogeny and intraspecific variability of holoparasitic Orobanchaceae (Orobanchaceae) inferred from plastid *rbcL* sequences. *Mol. Phylogenetics Evol.* **2004**, *33*, 482–500. [[CrossRef](#)]
6. Stewart, G.R.; Press, M.C. The physiology and biochemistry of parasiticangiosperms. *Annu. Rev. Plant Phys.* **1990**, *41*, 127–151.
7. Wicke, S.; Müller, K.F.; Depamphilis, C.W.; Quandt, D.; Bellot, S.; Schneeweiss, G.M. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9045–9050. [[CrossRef](#)]
8. Graham, S.W.; Lam, V.K.Y.; Merckx, V.S.F.T. Plastomes on the edge: The evolutionary breakdown of mycoheterotroph plastid genomes. *New Phytol.* **2017**, *214*, 48–55. [[CrossRef](#)]
9. Craven, L.; Dăneț, F.; Veldkamp, J.; Goetsch, L.; Hall, B. *Vireya Rhododendrons: Their monophyly and classification (Ericaceae, Rhododendron section Schistanthe)*. *Blumea - Biodiversity, Evol. Biogeogr. Plants* **2011**, *56*, 153–158. [[CrossRef](#)]
10. Thomas, W.A.B.; Michael, B.B.; Sasa, S.; John, V.F. On the brink: The highly reduced plastomes of nonphotosynthetic Ericaceae. *New Phytol.* **2017**, *216*, 254.
11. Fajardo, D.; Senalik, D.; Ames, M.; Zhu, H.Y.; Steffan, S.A.; Rebecca, H.; James, P.; Nicholi, V.; Emily, G.; Kathy, K.; et al. Complete plastid genome sequence of *Vaccinium macrocarpon*: Structure, gene content, and rearrangements revealed by next generation sequencing. *Tree Genet. Genomes* **2013**, *9*, 489–498. [[CrossRef](#)]
12. Zheng, S.; Tian, X.; Huang, C.; Wang, L.; Feng, Y.; Zhang, J. Molecular and morphological evidence for natural hybridization between *Rhododendron decorum* and *R. delavayi* (Ericaceae). *Biodivers. Sci.* **2017**, *25*, 627–637. [[CrossRef](#)]
13. Tsai, C.; Huang, S.; Chen, C.; Tseng, Y.; Huang, P.; Tsai, S.; Hou, C. Genetic relationships of *Rhododendron* (Ericaceae) in Taiwan based on the sequence of the internal transcribed spacer of ribosomal DNA. *J. Hortic. Sci. Biotechnol.* **2003**, *78*, 234–240. [[CrossRef](#)]

14. Jansen, R.K.; Sasaki, C.; Lee, S.; Hansen, A.K.; Daniell, H. Complete plastid genome sequences of three rosids (Castanea, Prunus, Theobroma): Evidence for at least two independent transfers of rpl22 to the nucleus. *Mol. Biol. Evol.* **2011**, *28*, 835–847. [[CrossRef](#)] [[PubMed](#)]
15. Maliga, P.; Svab, Z. Engineering the plastid genome of *Nicotiana glauca*, a diploid model species for plastid genetics. *Plant Chromosom. Eng. Met. Protoc.* **2011**, *701*, 37–50.
16. Fazekas, A.J.; Burgess, K.S.; Kesanakurti, P.R.; Graham, S.W.; Newmaster, S.G.; Husband, B.C.; Percy, D.M.; Hajibabaei, M.; Barrett, S.C.H. Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. *PLoS One* **2008**, *3*, e2802. [[CrossRef](#)]
17. Sugiura, M. The chloroplast genome. *Plant Mol. Boil.* **1992**, *19*, 149–168. [[CrossRef](#)]
18. Sugiura, M. The chloroplast genome. *Essays Biochem.* **1995**, *30*, 49–57.
19. Ravi, V.; Khurana, J.P.; Tyagi, A.K.; Khurana, P. An update on chloroplast genomes. *Plant Syst. Evol.* **2008**, *271*, 101–122. [[CrossRef](#)]
20. Wicke, S.; Schneeweiss, G.M.; DePamphilis, C.W.; M€uller, 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](#)]
21. Zhu, Q.L. Comparative genomic research on the organellar genome of watermelon and melon. Ph.D. Thesis, Northeast Agricultural University: Harbin, China, 2018.
22. Kang, H.I.; Lee, H.O.; Lee, I.H.; Kim, I.S.; Lee, S.W.; Yang, T.J.; Shim, D. Complete chloroplast genome of *Pinus densiflora* Siebold & Zucc. and comparative analysis with five pine trees. *Forests* **2019**, *10*, 600.
23. Liu, X.; Chang, E.-M.; Liu, J.-F.; Huang, Y.-N.; Wang, Y.; Yao, N.; Jiang, Z.-P. Complete Chloroplast Genome Sequence and Phylogenetic Analysis of *Quercus bawanglingensis* Huang, Li et Xing, a Vulnerable Oak Tree in China. *Forests* **2019**, *10*, 587. [[CrossRef](#)]
24. Wu, Z.-Q.; Ge, S. The phylogeny of the BEP clade in grasses revisited: Evidence from the whole-genome sequences of chloroplasts. *Mol. Phylogenetics Evol.* **2012**, *62*, 573–578. [[CrossRef](#)] [[PubMed](#)]
25. Zhang, C.Y. Study on cross compatibility of *Rhododendron*, L. Ph.D. Thesis, Nanjing Forestry University, Nanjing, China, 2012.
26. Galle, F.C. *Azaleas—Plants, Habits, Flowers and Leaves*; Timber Press: Portland, OR, USA, 1985.
27. Shen, J.S.; Li, X.Q.; Zhu, X.T.; Huang, X.L.; Jin, S.H. Complete chloroplast genome of *Rhododendron pulchrum*, an ornamental medicinal and food tree. *Mitochondrial DNA Part B* **2019**, *4*, 3527–3528. [[CrossRef](#)]
28. Doyle, J.J. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
29. Dierckxsens, N.; Mardulyn, P.; Smits, G. Novoplasty: De novo assembly of organelle genomes from whole genome DNA. *Nucleic Acids Res.* **2016**, *45*, e18.
30. Liu, C.; Shi, L.; Zhu, Y.; Chen, H.; Zhang, J.; Lin, X.; Guan, X. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genom.* **2012**, *13*, 715. [[CrossRef](#)]
31. Wyman, S.K.; Jansen, R.K.; Boore, J.L. Automatic annotation of organellar genomes with DOGMA. *Bioinform.* **2004**, *20*, 3252–3255. [[CrossRef](#)]
32. Kurtz, S.; Schleiermacher, C. REPuter: Fast computation of maximal repeats in complete genomes. *Bioinform.* **1999**, *15*, 426–427. [[CrossRef](#)]
33. Kurtz, S. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)]
34. Peden, J.F. Analysis of codon usage. *Biosystems.* **1999**, *5*, 45–50.
35. Beier, S.; Thiel, T.; M€unch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinform.* **2017**, *33*, 2583–2585. [[CrossRef](#)] [[PubMed](#)]
36. Katoh, K.; Kuma, K.-I.; Toh, H.; Miyata, T. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **2005**, *33*, 511–518. [[CrossRef](#)] [[PubMed](#)]
37. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinform.* **2011**, *27*, 2156–2158. [[CrossRef](#)]
38. Kurtz, S.; Phillippy, A.; Delcher, A.L.; Smoot, M.; Shumway, M.; Antonescu, C.; Salzberg, S.L. Versatile and open software for comparing large genomes. *Genome Boil.* **2004**, *5*, R12. [[CrossRef](#)] [[PubMed](#)]

39. Mayor, C.; Brudno, M.; Schwartz, J.R.; Poliakov, A.; Rubin, E.M.; Frazer, K.A.; Pachter, L.S.; Dubchak, I. VISTA: Visualizing global DNA sequence alignments of arbitrary length. *Bioinform.* **2000**, *16*, 1046–1047. [[CrossRef](#)] [[PubMed](#)]
40. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.
41. Cavender Bares, J.; González Rodríguez, A.; Eaton, D.A.R.; Hipp, A.A.L.; Beulke, A.; Manos, P.S. Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Mol Ecol.* **2015**, *24*, 3668–3687. [[CrossRef](#)]
42. Kim, S.-C.; Baek, S.-H.; Lee, J.-W.; Hyun, H.J. Complete chloroplast genome of *Vaccinium oldhamii* and phylogenetic analysis. *Mitochondrial DNA Part B* **2019**, *4*, 902–903. [[CrossRef](#)]
43. Tao, X.; Ma, L.; Zhang, Z.; Liu, W.; Liu, Z. Characterization of the complete chloroplast genome of alfalfa (*Medicago sativa*) (Leguminosae). *Gene Rep.* **2017**, *6*, 67–73. [[CrossRef](#)]
44. Yang, C.; Wu, X.; Guo, X.; Bao, P.; Chu, M.; Zhou, X.; Liang, Z.; Ding, X.; Yan, P. The complete chloroplast genome of *Medicago sativa* cv. Hangmu No.1, a plant of space mutation breeding. *Mitochondrial DNA Part B* **2019**, *4*, 603–604. [[CrossRef](#)]
45. Schmid, P.A.; Flegel, W. Codon usage in vertebrates is associated with a low risk of acquiring nonsense mutations. *J. Transl. Med.* **2011**, *9*, 87. [[CrossRef](#)] [[PubMed](#)]
46. Li, X.; Li, Y.; Zang, M.; Li, M.; Fang, Y. Complete Chloroplast Genome Sequence and Phylogenetic Analysis of *Quercus acutissima*. *Int. J. Mol. Sci.* **2018**, *19*, 2443. [[CrossRef](#)]
47. Boudreau, E.; Takahashi, Y.; Lemieux, C.; Turmel, M.; Rochaix, J. The chloroplast *ycf3* and *ycf4* open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. *EMBO J.* **1997**, *16*, 6095–6104. [[CrossRef](#)] [[PubMed](#)]
48. Wang, L.; Wuyun, T.-N.; Du, H.; Wang, D.; Cao, D. Complete chloroplast genome sequences of *Eucommia ulmoides*: Genome structure and evolution. *Tree Genet. Genomes* **2016**, *12*, 12. [[CrossRef](#)]
49. Jansen, R.K.; Cai, Z.; Raubeson, L.A.; Daniell, H.; Depamphilis, C.W.; Leebens-Mack, J.; Müller, K.F.; Guisinger-Bellian, M.; Haberle, R.C.; Hansen, A.K.; et al. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19369–19374. [[CrossRef](#)]
50. Goremykin, V.V.; Holland, B.; Hirsch-Ernst, K.I.; Hellwig, F.H. Analysis of *Acorus calamus* Chloroplast Genome and Its Phylogenetic Implications. *Mol. Biol. Evol.* **2005**, *22*, 1813–1822. [[CrossRef](#)] [[PubMed](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).