

Comparative analysis of chloroplast genomes of endangered heterostylous species *Primula wilsonii* and its closely related species

Yan-Ping Xie¹, Ganggang Yang², Chan Zhang², Xingwang Zhang¹, and Xianfeng Jiang³

¹HuaiBei Normal University

²Henan Normal University

³Dali University

July 14, 2022

Abstract

Primula, well known for its heterostyly, is the largest genus in the family Primulaceae with more than 500 species. The considerable species number has introduced a huge challenge for taxonomy. Although several phylogenetic constructions have been carried out thoroughly, the relationships between *Primula* species were remained incompletely understood, especially for the relationship among sections within Chinese species. *P. wilsonii* Dunn is a PSESP (plant species with extremely small populations) with very limited genetic information to explore its endangered mechanism and conservation. In this study, we sequenced and assembled the complete chloroplast genomes of *P. wilsonii* using Illumina sequencing and compared its genomic sequences with those of four related *Primula* species. The chloroplast genomes of *Primula* species were similar in the basic structure, gene order and GC content. The detected 38 SSRs loci and 17 hyper-variable regions had many similarities in *P. wilsonii*, *P. anisodora*, *P. miyabeana* and *P. poissonii*, but showed a significant difference compared with those in *P. secundiflora*. Slight variations were observed among *Primula* chloroplast genomes, in consideration of the relatively stable patterns of IR contraction and expansion. Phylogenetic analysis based on chloroplast genomes confirmed three major clades in Chinese *Primula*, but the infrageneric sections were not in accordance with morphological traits. The *P. poissonii* complex was confirmed here and *P. anisodora* was the species that was most closely related to *P. wilsonii*. Overall, the chloroplast genome sequences provided useful genetic and evolutionary information for phylogeny, population genetics and conservation studies on Chinese *Primula* species.

Comparative analysis of chloroplast genomes of endangered heterostylous species *Primula wilsonii* and its closely related species

Yan-Ping Xie ¹, Gang-Gang Yang ², Chan Zhang ², Xing-Wang Zhang ^{1,*} and Xian-Feng Jiang ^{3,*}

¹ School of Life Sciences, Huaibei Normal University, Huaibei 253000, China; xieyanping0001@163.com;

² School of Life Sciences, Henan Normal University, Xinxiang 453007, China; yangganggang@htu.deu.cn; zhangchan0705@163.com

³ College of Agriculture and Bioscience, Dali University, Dali 671003, China

* Correspondence: zhangxingwang79@126.com and jiangxianfeng@dali.edu.cn

Abstract: *Primula*, well known for its heterostyly, is the largest genus in the family Primulaceae with more than 500 species. The considerable species number has introduced a huge challenge for taxonomy. Although several phylogenetic constructions have been carried out thoroughly, the relationships between *Primula* species were remained incompletely understood, especially for the relationship among sections

within Chinese species. *P. wilsonii* Dunn is a PSESP (plant species with extremely small populations) with very limited genetic information to explore its endangered mechanism and conservation. In this study, we sequenced and assembled the complete chloroplast genomes of *P. wilsonii* using Illumina sequencing and compared its genomic sequences with those of four related *Primula* species. The chloroplast genomes of *Primula* species were similar in the basic structure, gene order and GC content. The detected 38 SSRs loci and 17 hyper-variable regions had many similarities in *P. wilsonii*, *P. anisodora*, *P. miyabeana* and *P. poissonii*, but showed a significant difference compared with those in *P. secundiflora*. Slight variations were observed among *Primula* chloroplast genomes, in consideration of the relatively stable patterns of IR contraction and expansion. Phylogenetic analysis based on chloroplast genomes confirmed three major clades in Chinese *Primula*, but the infrageneric sections were not in accordance with morphological traits. The *P. poissonii* complex was confirmed here and *P. anisodora* was the species that was most closely related to *P. wilsonii*. Overall, the chloroplast genome sequences provided useful genetic and evolutionary information for phylogeny, population genetics and conservation studies on Chinese *Primula* species.

Keywords: plastid genome; Chinese *Primula*; phylogeny; sequence divergences; plant species with extremely small populations

1. Introduction

As a genetically controlled floral polymorphism, heterostyly has evolved numerous times in at least 30 families of angiosperm plants [1-3], and this number is increasing with the emergence of new reports [4-7]. Heterostylous plants have a reciprocal arrangement of stigma and anther heights, which are referred to as reciprocal herkogamy and are believed to promote outcrossing [8, 9]. Reciprocal herkogamy is often linked with a self-incompatibility system that prevents or strongly limits self- and intramorph fertilization [10]. Heterostylous plants may comprise distylous (long style and short style) or tristylous (long style, middle style, and short style) morphs [11]. Since the heterostyle was described by Darwin in his classic book *The different forms of flowers on plants of the same species*, it has intrigued numerous researchers for centuries [9, 12-14].

Primula is famous for its mating with the heterostyly system and high ornamental value. This group comprises about 500 species, with more than 300 species distributed in China, particularly clustering in the Himalaya-Hengduan Mountains [15]. The exceptionally high species diversity of *Primula* was presumed to be triggered by the extensive uplifts of the Qinghai-Tibet Plateau and climate oscillations during the Quaternary [16]. The taxonomic study of *Primula* has been revised many times according to morphological characters. The infrageneric system with a total of 31 sections of *Primula* was initially proposed by Smith and Fletcher [17]. Later, a revised system with seven subgenera was proposed that considered some putative reticulate evolutionary relationships [18]. Then, a six-subgenera system was amended in *Primula* by Richards [15], whereas, the subgenus concept was not adopted in the delimitation of the Chinese *Primula*; a total of 24 sections were delimited [19]. Moreover, many key diagnostic traits, such as the shape of calyx, are slightly different and nonquantitative. The frequently happened natural hybridization and gene introgression also confuse species boundaries in this genus [20-24]. Although an increasing number of phylogenetic constructions have been carried out previously and have greatly advanced our understanding of the evolution of *Primula*, the phylogenetic relationships within the genus *Primula* species have remained incompletely resolved.

Primula wilsonii Dunn is a perennial herb in Sect. *Proliferae* Pax of *Primula* (Primulaceae). The most closely related species of *P. wilsonii* are *P. miyabeana*, *P. poissonii* and *P. anisodora*, these species clustered in a well-supported complex in Sect. *Proliferae* based on phylogeny construction using *rbcL* + *matK* + *ITS* markers, which is considered to be the first choice to barcode *Primula* plants at present [25]. *P. wilsonii* only sporadically scatters in limited areas of Sichuan and Yunnan provinces, unlike the other widespread *Primula*s that are common in the fields and gardens. Considering the limited distribution areas and small size of populations in the fields, *P. wilsonii* was identified as a PSESP (plant species with extremely small populations) and in need of protection eagerly [26]. However, the mechanisms leading to an endangered situation and measures to promote the conservation of *P. wilsonii* are very scarce.

Chloroplasts are the photosynthetic organelles in plant cells, with highly conserved genomes that are inherited maternally in major angiosperms [27-29]. The quadripartite structure of the angiosperm chloroplast genome consists of a large single copy (LSC) region and a small single copy (SSC) region, separated by two inverted repeats (IRs) [30]. The lack of recombination and slow evolutionary rate as compared with mitochondrial and nuclear genomes make it suitable for phylogeny at genus and family level, species barcoding, population genetics and the conservation of endangered species [28, 30-35]. Additionally, up till now, taking advantage of the development of high-throughput sequencing technology, more than 7,000 chloroplast genomes of land plants have been publicly known [36] since the publication of the first plastid genome sequences for *Nicotiana tabacum* and *Marchantia polymorpha* [37,38].

Unfortunately, few chloroplast genomes have been reported in *Primula* species. In this study, we determine the complete chloroplast genomes of a PSESP species in this genus [39]. In addition, we explore simple sequence repeats (SSRs) loci and identify highly variable regions by comparing the genome contents and structures between the PSESP specie *P. wilsonii* and its widely spread closely related species. Here, we aim to: (1) investigate the molecular structures of the chloroplast genomes of *P. wilsonii*, (2) examine the variations of SSRs and highly divergent regions that could be used as specific DNA markers for *P. wilsonii* and its close relatives, (3) evaluate the evolution of several *Primula* chloroplast genomes, and (4) facilitate the conservation and systematics of the Chinese *Primula* species.

2. Materials and Methods

2.1. Plant materials, DNA extraction and sequencing

The plant materials used in this study were collected from Wuxuhai, Sichuan Province, China. The voucher specimen (voucher accession number XYP202007016) was deposited in the Key Laboratory of Plant Resource and Biology at Huaibei Normal University. Total genomic DNA was extracted from silica-dried leaves with a modified cetyltrimethylammonium bromide (CTAB) method [36]. The quality and quantity of the DNA samples were determined using agarose gel electrophoresis and the Qubit dsDNA BR assay (Life Technologies, Carlsbad, CA, USA). The qualified PCR-amplified library was sequenced with the Illumina NovaSeq Tenplatform (Nanjing Genepioneer Biotechnologies Inc., Nanjing, China).

2.2. Genome assembly, annotation and analysis

The low quality reads were assessed and filtered using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and Trimmomatic v 0.36 to obtain high-quality clean data [41]. The chloroplast genome sequence of *P. wilsonii* was assembled by a de nova method using SPAdes Genome Assembler v3.10.1 [42]. The k-mers were set to 55, 87 and 121 to achieve optimal assembly. The whole chloroplast genomes were assembled using high-coverage and long-sequenced contigs. Then, the SSPACE v2.0 was used to construct the scaffold of the chloroplast genomes [43] and Gapfiller v2.1.1 was used to fill the gaps [44]. Finally, Bowtie2 v2.2.4 was used to validate the genome assembly by mapping the initial reads to the assembled genome sequence [45]. The chloroplast genomes were annotated in two ways. Prodigal v2.6.3 [46], Hmmer v3.1b2 [47] and ARAGORN v1.2.38 [48] were used to detect the CDS, rRNA and tRNA, respectively. Furthermore, BLAST (NCBI) was used to check the annotation, followed by manual correction through comparison with the other closely related chloroplast genomes. The circular chloroplast genome map of *P. wilsonii* was then generated using OGDRAW [49]. The repeating sequences including forward, reverse, complement and palindrome repeats were identified using the online REPuter program [50], with three for Hamming distance and 30 for minimal repeat size. Simple sequence repeats (SSRs) were detected using MISA software (<https://webblast.ipk-gatersleben.de/misa/>) [51]. Thresholds for a minimum number of repeat units were set as follows: 10 for mono-nucleotide repeat units; 5 for di-; 4 for tri-; and 3 for tetra-, penta-, or hexa-nucleotide repeat units.

2.3. Comparative plastome analysis

Gene distribution and the percentage of sequence identity were compared and visualized using mVISTA software [52] with the LAGAN mode [53] in chloroplast genomes of five *Primula* species, with *P. secundiflora*

, *P. poissonii*, *P. anisodora*, and *P. miyabeana* selected as close relatives of *P. wilsonii*. All the chloroplast genomes were obtained from NCBI, except for *P. wilsonii*. The annotation of *P. wilsonii* served as a reference. Nucleotide variability values (P_i) were calculated using the same alignment. Nucleotide diversity was detected by sliding window analysis conducted in DnaSP v.6.11.01 with a step size of 200 bp and window length of 600 bp [54]. The expansion or contraction of the inverted repeat (IR) regions in the chloroplast genomes of the five related *Primula* species was investigated and visualized using IRscope program [55].

2.4. Phylogenetic analyses

To investigate the phylogenetic relationship in the sections of Chinese *Primula* and the resolution ability of chloroplast genomes, phylogenetic analysis was performed for the 60 species representing 20/24 sections of Chinese *Primula* based on the whole chloroplast genomes downloaded from NCBI. Two outgroup species (*Androsace bulleyana* and *Ardisia solanacea*) were sampled from a closely related genus *Androsace* and a more distantly related genus *Ardisia* of the Primulaceae in the sense of phylogenetic relationships [56]. The complete chloroplast sequences were firstly aligned by MAFFT v7.307 [57]. jModelTest 2.0 was used to find the best nucleotide-substitution model before phylogenetic construction [58]. Trees were then constructed using the maximum likelihood (ML) method by online RaxML BlackBox software [59]. ML was implemented starting from random trees, using 1,000 rapid bootstrap inferences with a General Time Reversible (GTR) nucleotide-substitution model as suggested. The final phylogenetic results were viewed by using FigTree v.1.6.1.

3. Results

3.1. Chloroplast genomes features

The assembled chloroplast genome of *P. wilsonii* was 151,677 bp in length, exhibiting a typical circular chloroplast structure like most angiosperms. As shown in Figure 1, it contains a large single-copy (LSC) region of 83,510 bp, a small single-copy (SSC) region of 17,765 bp, and a pair of inverted repeats (IRa and IRb) of 25,201 bp. The average coverage depth of the assembled chloroplast genome was 1249 ×. Overall, the GC content was 36.99%. We found an uneven distribution of GC content across the regions of the genomes, which was 34.89%, 30.18% and 42.87% for the LSC, SSC and IR regions, respectively. The chloroplast genome structures, the size of each region, and the GC contents of the five chloroplast genomes of *Primula* species were compared and are shown in Table 1.

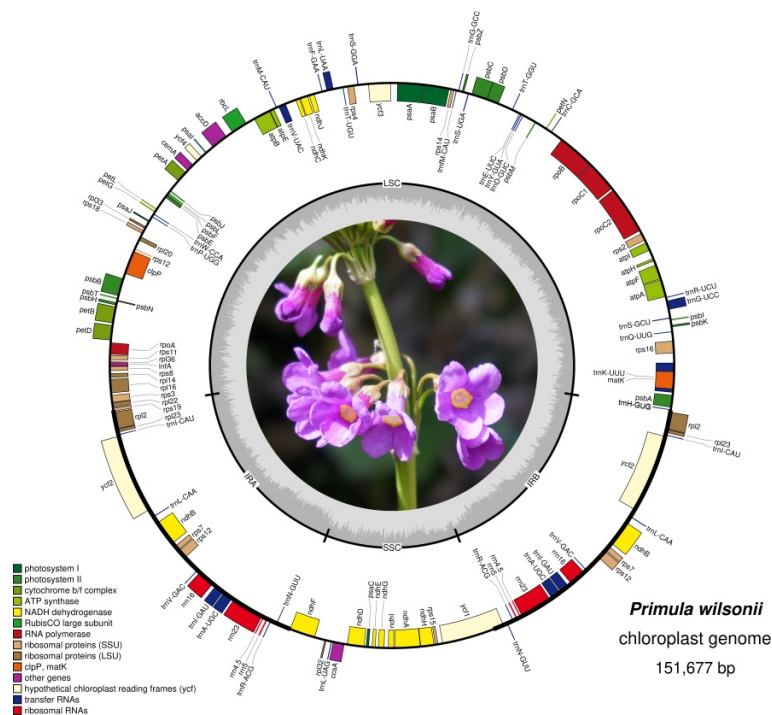


Figure 1. Circular gene map of the chloroplast genome of *Primula wilsonii*. The thick lines indicate the inverted repeat regions (IRa and IRb) that separate the genome into small (SSC) and large (LSC) single copy regions. Genes on the inside of the map are transcribed clockwise, while those on the outside are transcribed counterclockwise. Genes grouped by different functions are shown in different colors.

Table 1. Comparison of general features of *Primula* chloroplast genomes.

Genome features	P. wilsonii	P. anisodora	P. miyabeana	P. poissonii	P. secundiflora
GenBank number	MW442886	NC053578	MT317303	NC024543	MT317261
Genome size (bp)	151677	151674	151706	151664	151583
LSC length (bp)	83510	83498	83520	83444	83362
SSC length (bp)	17765	17772	17747	17822	17839
IR length (bp)	25210	25202	25219	25199	25191
Total GC content (%)	36.99	37.02	37.00	37.02	36.98
Total genes	113	113	112	115	112
Protein-coding genes	79	80	79	81	79
tRNA genes	30	29	29	30	29
rRNA genes	4	4	4	4	4

Regardless of the duplicate genes, the chloroplast genome of *P. wilsonii* contained 113 genes, including 79 protein-coding genes, 30 tRNA genes, and four rRNA genes. Among them, six protein-coding genes, seven tRNAs, and all the four rRNAs were completely duplicated within IRs (Figure 1). The *rps12* gene was

found to be trans-spliced in chloroplast genomes. It had two duplicates in IRs and one of its exons located in the LSC region (Figure 1). A total of 18 genes, including 12 protein-coding genes and six tRNA genes, had introns. It was noted that 16 genes among them had only one intron while the other two genes had two introns. The comparisons between the genome of *P. wilsonii* and the other *Primula* chloroplasts are shown in Table 1. The results indicated that the gene contents were similar among the *Primula* species, excluding the loss of the *ycf15* gene in the *P. wilsonii* chloroplast genome. The chloroplast genomes of *P. poissonii* and *P. wilsonii* possessed the *infA* gene, whereas the others did not have it. Moreover, the *trnG* -GCC sequence only appeared in the chloroplast genome of *P. wilsonii*.

3.2. Repeat sequences and SSRs analysis

Except for the largest repeat in each genome (IR regions), a total of 111 repeat pairs no more than 60 bp in length were identified in the five *Primula* chloroplast genomes. Only two types of repeat sequences, forward and palindromic repeats, were detected (Figure 2). There were 22 (nine forward, 13 palindrome), 22 (nine forward, 13 palindrome), 27 (12 forward, 15 palindrome), 13 (13 forward), and 27 (12 forward, 15 palindrome) repeats in *P. wilsonii*, *P. anisodora*, *P. miyabeana*, *P. poissonii*, and *P. secundiflora*, respectively. Among them, the chloroplast genomes of *P. miyabeana* and *P. secundiflora* had the largest number of repeats, whereas *P. poissonii* had the fewest. The results indicated that forward repeats were more common, accounting for 54.47% (67 of 123 repeats). In these five genomes, the length of the repeats was mainly in the range of 30-39 bp, with a percentage of 61.26% (68 of 111 repeats), followed by 40-49 bp, contributing 29.73% (33 of 111 repeats).

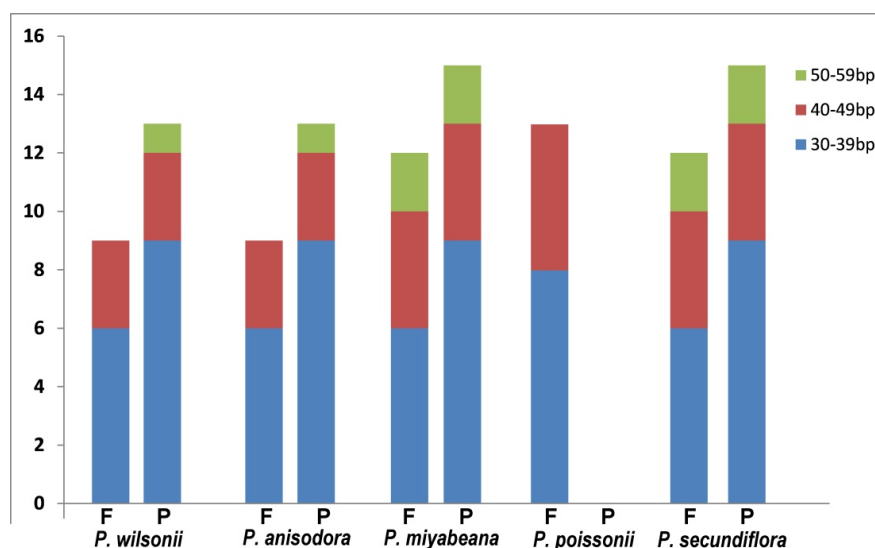


Figure 2. Types and numbers of large repeat sequences in the chloroplast genomes of five *Primula* species.

A total of 38 Microsatellites in the *P. wilsonii* chloroplast genome were detected, including 28 mono-, five di-, two tri- and three tetra-nucleotide repeats. In *P. poissonii*, 38 SSRs loci were detected, consisting of 27 mono-, six di-, two tri- and three tetra-nucleotide repeats, respectively (Table 2). Moreover, 41 SSRs loci were spotted in *P. secundiflora*, and the numbers of different types are listed in Table 2. Here, we found that the one-base SSRs loci of *P. wilsonii* and *P. poissonii* were only A/T repeats. The dinucleotide repeats were AT/TA in all the five *Primula* species. It is shown that the trinucleotide repeats were AAT/ATT in *P. wilsonii*, *P. anisodora*, *P. miyabeana* and *P. poissonii*, but they were not present in *P. secundiflora*. The tetra-nucleotide repeats AAAT/AATA were present in all the *Primula* species. However, the AAAG/CTTT and AATT repeats only appeared in *P. secundiflora*. Among these SSRs loci, 30 (78.95%) were in the LSC

region, six (15.79%) were in SSC region, and two (5.26%) were in the IR region of the *P. wilsonii* chloroplast genome.

Table 2. Types and numbers of SSRs in the chloroplast genomes of five *Primula* species.

SSRs type	Repeat unit	P. wilsonii	P. anisodora	P. miyabeana	P. poissonii	P. secundiflora
Mono-	A/T	28	24	30	27	30
	C/G	-	1		-	1
Di-	AT/TA	5	5	5	6	5
Tri-	AAT/ATT	2	2	2	2	-
Tetra-	AAAT/ATTT	3	3	3	3	2
	AAAG/CTTT	-			-	1
	AATT/AATT	-			-	1
Penta-	AAATC/ATTTG	-			-	1
Total		38	35	40	38	41

3.3. Sequence divergence

The *Primula* chloroplast genomes exhibited moderate sequence divergence. Furthermore, the results showed that the sequence of coding regions and the two IR regions were significantly more conserved than that of LSC and SSC regions. In the coding regions, most genes were relatively conserved, except for *matK*, *rpl22*, *ndhF* and *ycf1*. In contrast, the intergenic regions were shown to be highly divergent (Figure 3). Then, we found that the variation level of DNA polymorphism was from 0.00067 to 0.02633. The greatest differences among the five *Primula* species were located in the two SC regions, while IR regions were the least different. About 17 hyper-variable regions were discovered with a nucleotide diversity (P_i) value that was greater than 0.01 (Figure 4). Some relatively high P_i value were detected in CDS, such as *psbA* (0.02633), *ycf1* (0.01533), *psaJ* (0.01333), *rpl32* (0.01233), *petL* (0.01167) and *ndhA* (0.01033). Consistently, the gene *ycf1* exhibited higher diversity and showed abundant variation. In addition, we revealed five highly divergent regions among noncoding regions, including *ccsA-ndhD* (0.01733), *ndhF-rpl32* (0.01583), *petA-psbJ* (0.01567), *trnE* (UUC)-*trnT* (GGU) (0.01267) and *rps15-ycf1* (0.01167). These mutational hotspots can serve as potential loci when developing novel DNA barcodes for plant classification.

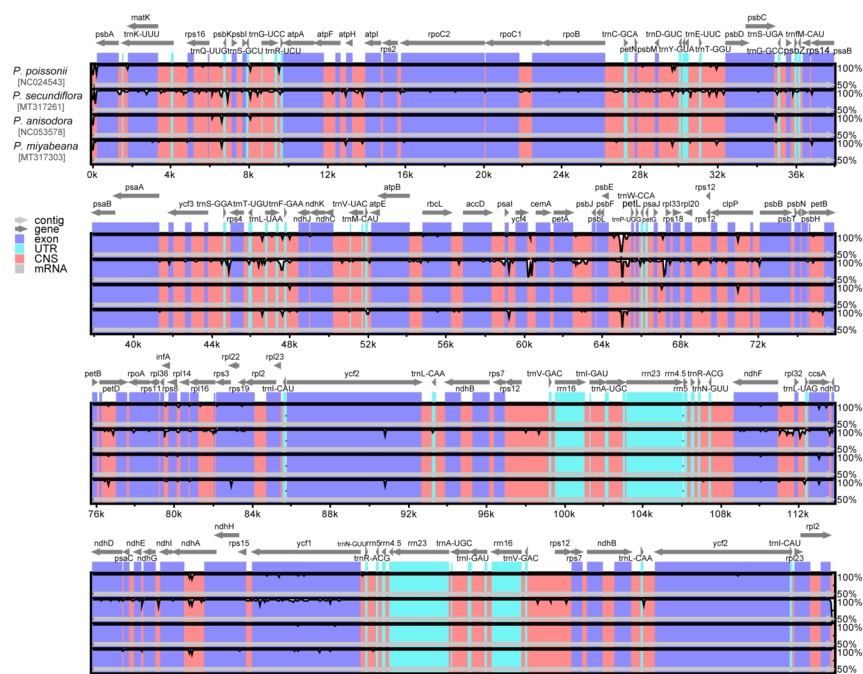


Figure 3. Comparison of the complete chloroplast genome of five *Primula* species, with *P. wilsonii* as a reference. Gray arrows and thick black lines above the alignment indicate gene orientation. The y-axis represents the percentage identity ranging from 50% to 100%.

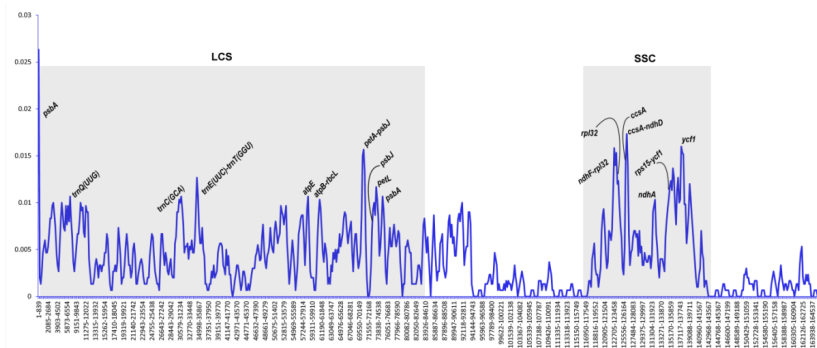


Figure 4. Sliding window analysis of the whole chloroplast genomes of five *Primula* species. The nucleotide variability of the five *Primula* species was assessed by DnaSP with window length of 600 bp and step size of 200 bp. The X-axis represents the regions of chloroplast genome and Y-axis represents nucleotide diversity for each window.

3.4. IR expansion and contraction analyses

The IR/SC junction regions showed slight differences in gene content and gene order. As shown in Figure 5, genes *rps19* /*rpl2* ,*ndhF* , *ycf1* and *rpl2* /*trnH* were present at the junction of the LSC/IRb, IRb/SSC, SSC/IRa and IRa/LSC borders, respectively. The LSC/IRb boundary was located in the *rps19* region, which crossed into the IRb region in all the six chloroplast genomes, resulting in a variable expansion (14–102 bp) of IRb region toward the *rps19* gene. The *ndhF* gene was entirely located in the SSC region in the chloroplast genomes of *P. secundiflora* and *P. poissonii* , whereas the IRb region extended 42 bp into the *ndhF* gene in

all the other *Primula* chloroplast genomes. The SSC/IRa junction was situated in the *ycf1* coding region, which crossed into the IRa region in all five chloroplast genomes. However, the length of *ycf1* in the IRa region varied from 5,468 to 5,483 bp, indicating the dynamic variation of the SSC/IRa junctions. The gene *trnH* was located in LSC, 0–14 bp away from the IRa/LSC border (Figure 5). Taken together, these data indicate that the contractions and expansions of the IR regions exhibited relatively stable patterns within these *Primula* species, with slight variations.

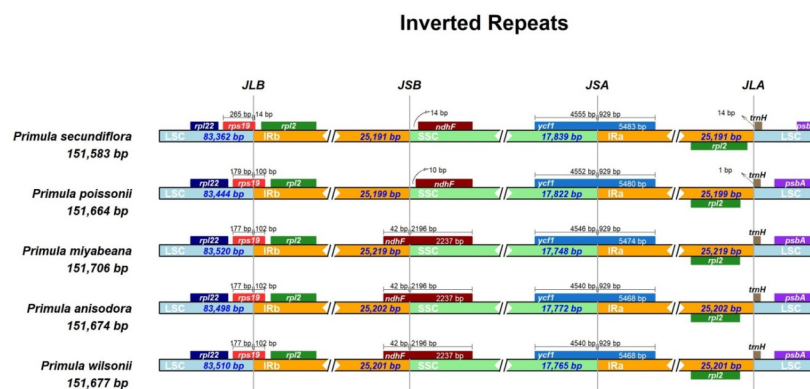


Figure 5. Comparisons of the LSC, IR, and SSC boundary regions among the five *Primula* chloroplast genomes. Genes at the IR/SC borders are denoted by the colored boxes. The numbers within the boxes represent the distances from the boundaries.

3.5. Phylogenetic analysis

Our sampling represented 20 of 24 recognized sections of the genus *Primula* in China [19]. The phylogenetic tree identified that all the samples in the genus *Primula* could be divided into three major clades with high support (Figure 6). Clade A included Sects. *Crystallophlomis*, *Ranunculooides*, *Amethyatina*, *Petiolares*, and *Proliferae*. Clade B contained Sects. *Primula*, *Souliei*, *Sikkimensis*, *Aleuritia*, *Denticulata*, *Capitatae*, *Soldanelloides*, and *Minutissimae*. Clade C combined Sects. *Auganthus*, *Obconicolisteri*, *Dryadifolia*, *Carolinella*, *Bullatae*, *Monocarpicae* and *Cortusoides* species. Our results found that several sections were not monophyletic groups, such as Sects. *Monocarpicae*, *Crystallophlomis*, *Obconicolisteri*, *Denticulata*, and *Proliferae*. It is worth noting that *P. wilsonii* was closest to *P. anisodora* with very high support value in Sect. *Proliferae*. The *P. poissonii* complex was further confirmed, which included *P. wilsonii*, *P. anisodora*, *P. poissonii* and *P. miyabeana* (Figure 6). However, the monophyly of Sect. *Proliferae* suggested in previous studies was not supported here [25].

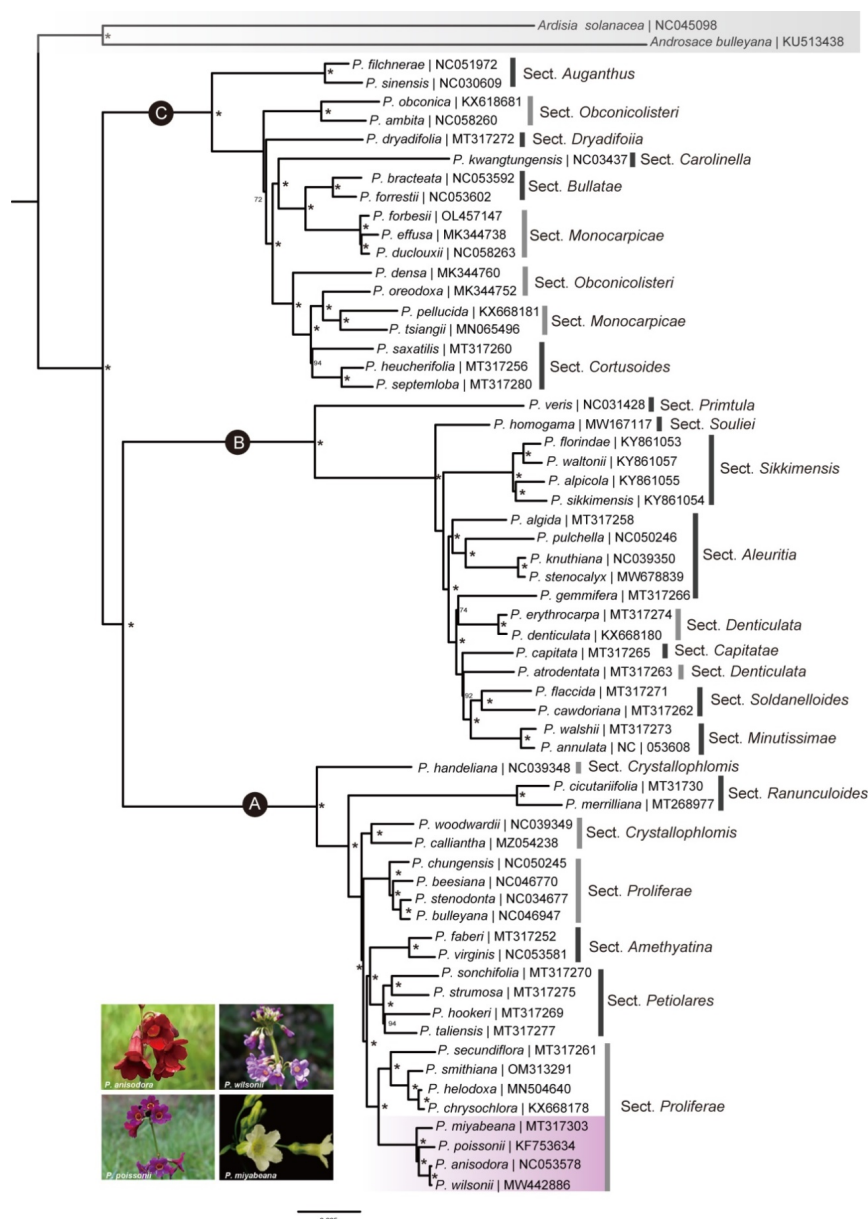


Figure 6. Maximum likelihood tree of *Primula* species based on chloroplast genomes. Bootstrap support values over 95% are labeled with asterisks. Outgroups and *P. poissonii* complex are highlighted with gray and purple shadings, respectively.

4. Discussion

4.1. The general characteristics of *Primula* chloroplast genomes

As with most angiosperms, the chloroplast genomes were conserved in *Primula* species, with similar GC content and typical quadripartite structures, including small and large single copy (SSC and LSC) regions separated by two inverted repeats (IRs) regions [60]. However, gene loss was found here. The *infA* gene, which encodes translation initiation factor 1 [61], was present in the chloroplast genome of *P. poissonii* and *P. wilsonii*, but was not present in the related *Primula* chloroplast genomes in our study. Additionally, these

findings are consistent with the results of some *Primula* species and other groups in angiosperm chloroplast genomes in previous studies [62, 63]. Remarkably, the *ycf15* gene was only missing in the chloroplast genomes of *P. wilsonii*. *ycf15* was located in the IR region and was highly conserved. The absence of *ycf15* was also reported in many other plants, such as *Colchicum* genus [64]. However, the function of the *ycf15* gene remains unclear and needs to be further investigated.

The patterns of gene loss we revealed here could be used for phylogeny reconstruction and species identification. The loss of *ycf15* gene in colchicine plants successfully determined the infrageneric relationship in the expanded *Colchicum* genus [64]. Thus, the non-presence of *ycf15* we found here might be a valuable molecular marker to separate *P. wilsonii* from *P. poissonii*, which is morphologically similar to *P. wilsonii*. Both of the two species are perennial herbs with candelabra inflorescence and purple flowers, so some scholars argue *P. wilsonii* should be merged into *P. poissonii* or treated as a subspecies of *P. poissonii*. Here we suggested that the missing *ycf15* gene in the *P. wilsonii* chloroplast genome could be extremely useful for distinguishing the two confusing species at the molecular level.

4.2. The evolution of the chloroplast genomes in *Primula*

IR regions are highly conserved in most angiosperm chloroplast genomes. However, the contraction and expansion of IR regions are not rare [65]. In this study, gene orders at the boundaries of SC/IR regions were the same among the five chloroplast genomes of *Primula*. However, the accurate positions of the genes at the SC/IR border were slightly varied, such as the genes *rps19*, *ndhF*, *ycf1*, *rpl2* and *trnH* [63]. In addition, some genes normally located in the SC region, such as *ndhF*, had moved to IR region due to the expansion of the IR region. It was reported that the chloroplast genomes' size, the LSC/SSC length, the gene numbers and pseudogene origination could vary among different species due to the expansion and contraction of IR regions [66, 67]. Moreover, the loss of IR regions has been occasionally detected in some taxa [68]. This could be the reason that the chloroplast genome size of *P. miyabeana* was the largest among the five *Primula* species with the longest IR region, and the chloroplast genome size of *P. secundiflora* was the smallest with the shortest IR region. Furthermore, a large number of studies also confirmed that the length of IR region greatly affected the chloroplast genome size [69, 70].

Species of *Primula* are famous for their ornamental value and heterostyly phenomenon in Southwest China. More genomic resources are needed to deeply investigate the phylogeny, biogeography, genetics and heterostyly evolution of *Primula* species. In addition, considering that *P. wilsonii* is a plant species with extremely small populations (PSESP), we need more genetic information for the conservation of germplasm resources. The numbers and distributions of repeat sequences, especially large repeats that are longer than 20 bp and 60 bp, may play important roles in the arrangement and recombination of the plastid genome [71, 72]. A total of 123 repeats were detected in the six *Primula* chloroplast genomes. All the repeat sequences appeared to be shorter than 60 bp in length. These findings are consistent with the results in other *Primula* species [63, 73], but not in agreement with the results of some other angiosperm plants, such as herbaceous *Alpinia* species [74] or woody *Aquilaria* species [70]. Our study detected very high levels of polymorphism in the large repeat sequences among the six *Primula* species in terms of both the types or numbers. Therefore, these large repeats might be an informative source for developing genetic markers for population genetics and phylogenetic constructions of *Primula* [75]. SSRs markers are a valuable genetic resource for phylogenetic investigations, population genetics assessment and species discrimination due to their abundant polymorphism and codominant inheritance [70, 76]. The SSRs markers detected here were mostly A/T mono-nucleotide repeats (28/38), similar to the results of other *Primula* species [63] and some other angiosperm species [77, 78]. The vast majority of SSRs loci were in SC regions (78.95% in LSC regions and 15.79% in SSC regions), yet few of them were present in IR regions. Moderate sequence divergence with greater variability in the SC region of *Primula* chloroplast genomes was displayed, which corresponded with previous studies [79]. Since the hyper-variable regions of the chloroplast genome are useful for phylogenetic construction, population genetics and DNA barcoding, the 17 highly polymorphic loci and the SSRs markers found in our study could serve as potential genetic markers for further phylogenetic and biogeographic analyses, population genetics and conservation analysis of *Primula* species.

4.3. Phylogenetic relationships of Chinese *Primula*

A total of 60 species representing 20 of 24 sections in Chinese *Primula* were sampled in our phylogenetic construction using chloroplast genome sequences based on ML method. Three major clades of *Primula* were detected with high internal support in this study, which was in accordance with previous studies [25, 73, 80]. Several sections did not exhibit monophyletic taxa, such as Sects. *Monocarpicae*, *Crystallophlomis*, *Obconicolisteri*, *Denticulata* and *Proliferae*, which were partly or entirely confirmed by the previous viewpoints [25, 73, 80]. A decision on the monophyly of Sect. *Proliferae* requires additional consideration. It has been treated as a monophyletic group based on the concatenation of ITS, *matK* and *rbcL* sequences [25, 73]. However, the chloroplast transcripts and protein coding sequences from chloroplast genomes analyses strengthen the assumption that Sects. *Amethyatina* and *Petiolares* species are nested within Sect. *Proliferae* [80]. This assumption is additionally supported by the results based on the whole chloroplast genome analysis in our investigation. This is corroborated by morphological traits such as an umbel with multiple flowers, campanulate calyx, and globose capsule. On the one hand, the conflicting phylogenetic diagnoses of nuclear and chloroplast sequences are common in plants [81]. On the other hand, the adaptive radiation caused by heterostyly, polyploidization and natural hybridization, or gene introgression might complicate the phylogenetic relationships under *Primula* [20–24]. This would explain why quite a few sections in *Primula* didn't belong to monophyletic group according to morphological characters.

P. wilsonii, together with *P. poissonii*, *P. anisodora*, and *P. miyabeana* (endemic to Taiwan) form to *P. poissonii* complex, which was one of the taxonomically challenging groups in Sect. *Proliferae*. The close relationship of these species has been revealed in studies, and *P. wilsonii* was closest to *P. miyabeana* based on *rbcL* + *matK* + ITS sequences, with low support [25]. However, the closest relative species was *P. anisodora* with very high support based on chloroplast genomic sequences in this study. Therefore, we suggest that the phylogenetic relationships between *Primula* species need to be further studied based on more genetic information, especially at the genome level, and we may come to the conclusion that chloroplast genomes sequences could provide a valuable resource for phylogenetic constructing of *Primula*.

5. Conclusions

This study compared the basic characteristics of the chloroplast genomes from several Chinese *Primula* species. We assessed the variation and IR boundaries evolution among these species. Furthermore, we constructed the phylogenetic relationships of the genus *Primula* covering a wide range of samples based on their chloroplast genomic sequences. In addition, we determined the conserved and variable regions in the chloroplast genomes. The large repeat sequences, SSRs loci, and 17 hypervariable regions were detected here, which could be used for population genetics and phylogenetic analysis in the future. Three major clades in *Primula* were confirmed, yet the sections were not in accordance with morphological traits, reflecting in the non-monophyletic nature of several sections. Therefore, we suggest that chloroplast genomes provide useful genetic and evolutionary information for studies on the phylogeny, population genetics, and conservation of *Primula* species.

Funding: This research was funded by National Natural Science Foundation of China (32100169) and Natural Science Foundation of Anhui Province (2108085QC104), Natural Science Foundation from Educational Commission of Anhui Province (KJ2020B25) to Y.P. Xie, the Henan Province Youth Talent Lift Project of China (212102310242) and the Key Scientific and Technological Project of Henan Province (2021HYTP042) to G.-G. Yang.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in the study are depositing in the NCBI repository, and the accession numbers are shown in the article.

Acknowledgments: The authors wish to thank Jian-Li Zhao and Li Li for help in collecting samples and operation of software; Heng-Yi Shao and MPDI for their advices on the language organization and English editing.

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

References

Barrett, S.C.H.; Shore, J.S. New insights on heterostyly: comparative biology, ecology and genetics. In *Self-incompatibility in flowering plants-evolution, diversity and mechanisms* ; Franklin-Tong, V.E., Ed.; Springer-Verlag: Berlin, Germany, 2008; pp 3–32.

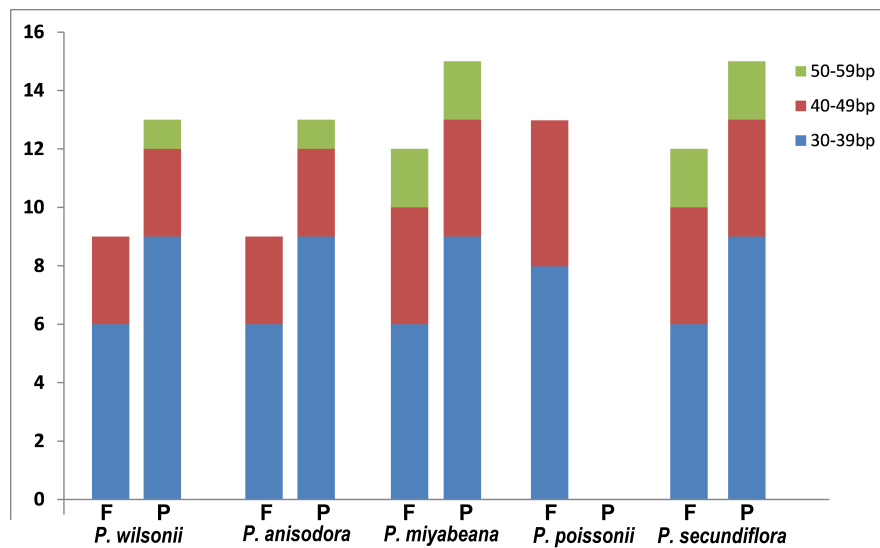
1. Chen, M.L.; You, Y.L.; Zhang, X.P. Advances in the research of heterostyly. *Acta. Pratacult. Sin.* **2010** , *19* , 226–239.
2. Ganders, F.R. The biology of heterostyly. *Nea. Zeal. J. Bot.* **1979** , *17* , 607-635.
3. Watanabe, K.; Yang, T.Y.A.; Nishihara, C.; Huang, T.L.; Nakamura, K.; Peng, C.I.; Sugawara, T. Distyly and floral morphology of *Psychotria cephalophora* (Rubiaceae) on the oceanic Lanyu (Orchid) Island, Taiwan. *Bot. Stud.* **2015** , *56* , 10.
4. Xu, X.Y., Zhou, L.L., Wang, Z.K., Zhuang, L. Flower distyly and breeding system of *Limonium chrysocomum*. *Bull. Bot. Res.* **2015** , *35* , 883-890.
5. Yang, C.; He, X.; Gou, G. *Ophiorrhiza guizhouensis* (Rubiaceae), a new species from Guizhou Province, southwestern China. *PhytoKeys* **2018** , *95* , 121.
6. Zhang, C.; Wang, L.L.; Duan, Y.W.; Lan, D.; Yang, Y.P. Pollination ecology of *Arnebia szechenyi* (Boraginaceae), a Chinese endemic perennial characterized by distyly and heteromorphic self-incompatibility. *Ann. Bot. Fenn.* **2014** , *51* , 297-304.
7. Webb, C.J.; Lloyd, D.G. The avoidance of interference between the presentation of pollen and stigmas in angiosperms II. Herkogamy. *Nea. Zeal. J. Bot.* **1986** , *24* , 163-178.
8. Darwin, C. The different forms of flowers on plants of the same species. John Murray: London, UK, 1877
9. Barrett, S.C.H.; Cruzan, M.B. Incompatibility in heterostylous plants. In *Genetic control of self-incompatibility and reproductive development in flowering plants* ; Williams, E.G., Clarke, A.E., Knox, R.B. Eds.; Springer Netherlands: Dordrecht, Holland, 1994; pp 189-219.
10. Barrett, S.C.H. The evolution of plant sexual diversity. *Nat. Rev. Genet.* **2002** , *3* , 274-284.
11. Barrett, S.C.H. Heterostylous genetic polymorphisms: model systems for evolutionary analysis. In *Evolution and function of heterostyly* ; Barrett, S.C.H., Ed.; Springer-Verlag: Berlin, Germany, 1992; pp 1-29.
12. Weller, S.G. The different forms of flowers – what have we learned since Darwin? *Bot. J. Linn. Soc.* **2009** , *160* , 249–261.
13. Cohen, J.I. A case to which no parallel exists’: the influence of Darwin’s *Different Forms of Flowers*. *Am. J. Bot.* **2010** , *97* , 701–716.
14. J. R. *Primula (new edition)* . Timber Press: Oregon, USA, 2003
15. Wen, J.; Zhang, J.Q.; Nie, Z.L.; Zhong, Y.; Sun, H. Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. *Front. Genet.* **2014** , *5* , 4
16. Smith, W.W.; Fletcher, H.R. XVII. —The genus *Primula* : Sections *Obconica* , *Sinenses* , *Reinii* , *Pinnatae* , *Malacoides* , *Bullatae* , *Carolinella* , *Grandis* and *Denticulata*. *Trans. R. Soc. Edinb.* **1947** , *61* , 415–478.
17. Wendelbo, P. Studies in Primulaceae. II. An account of *Primula* subgenus *Sphondylia* (Syn. Sect. *Floribundae*) with a review of the subdivisions of the genus. *Matematisk-Naturvitenskapelig Ser.* 1961, *11* , 1-46.
18. Hu, C.M., *Primulaceae* . Science Press: Beijing, China, 1990; Vol. 59, p 2.
19. De Vos, J.M.; Hughes, C.E.; Schneeweiss, G.M.; Moore, B.R.; Conti, E. Heterostyly accelerates diversification via reduced extinction in primroses. *P. Roy. Soc. B-Biol. Sci.* **2014** , *281* , 20140075.
20. Guggisberg, A.; Mansion, G.; Conti, E. Disentangling reticulate evolution in an arctic-alpine polyploid complex. *Syst. Biol.* **2009** , *58* , 55-73.
21. Ma, Y.P.; Tian, X.; Zhang, J.; Wu, Z.K.; Sun, W.B. Evidence for natural hybridization between *Primula beesiana* and *P. bulleyana*, two heterostylous primroses in NW Yunnan, China. *J. Syst. Evol.* **2014** , *52* , 500–507.

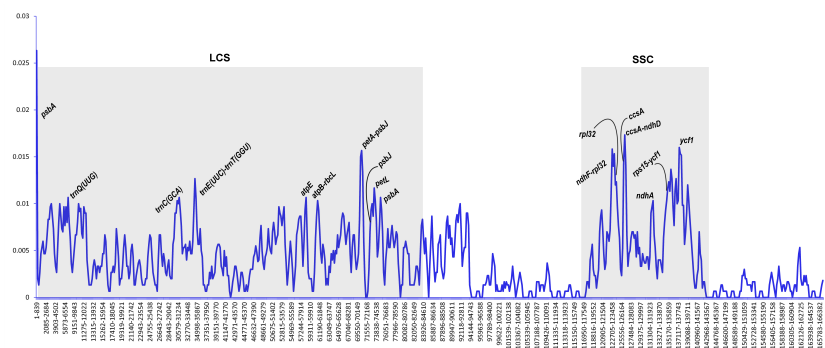
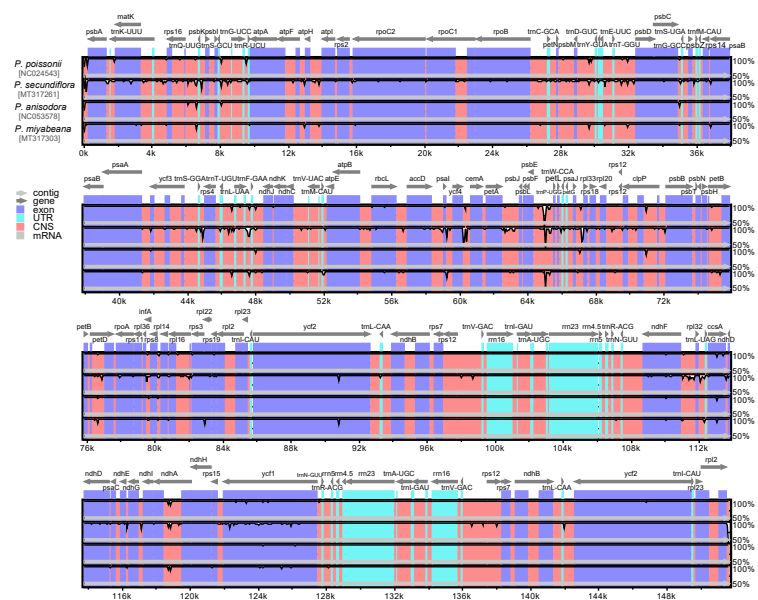
22. Xie, Y.P.; Zhu, X.F.; Ma, Y.P.; Zhao, J.L.; Li, L.; Li, Q.J. Natural hybridization and reproductive isolation between two *Primula* species. *J. Integr. Plant Biol.* **2017**, *59*, 526-530.
23. Xie, Y.P.; Zhao, J.L.; Zhu, X.F.; Li, L.; Li, Q.J. Asymmetric hybridization of *Primula secundiflora* and *P. poissonii* in three sympatric populations. *Biodiv. Sci.* **2017**, *25*, 647-653.
24. Yan, H.F.; Liu, Y.J.; Xie, X.F.; Zhang, C.Y.; Hu, C.M.; Hao, G.; Ge, X.J. DNA barcoding evaluation and its taxonomic implications in the species-rich genus *Primula* L. in China. *PLoS One* **2015**, *10*, e0122903.
25. Sun, W.B.; Yang, J.; Dao, Z.L., Study and conservation of plant species with extremely small populations (PSESP) in Yunnan province, China. Science Press: Beijing, China, 2019; p 176
26. Moore, M.J.; Soltis, P.S.; Bell, C.D.; Burleigh, J.G.; Soltis, D.E. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4623–4628.
27. Daniell, H.; Lin, C.S.; Yu, M.; Chang, W.J. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 1-29.
28. Daniell, H. Transgene containment by maternal inheritance: Effective or elusive? *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6879–6880.
29. Wang, L.; Wuyun, T.-N.; Du, H.Y.; Wang, D.P.; Cao, D.M. Complete chloroplast genome sequences of *Eucommia ulmoides*: genome structure and evolution. *Tree Genet. Genomes* **2016**, *12*, 12.
30. Palmer, J.D.; Jansen, R.K.; Michaels, H.J.; Chase, M.W.; Manhart, J.R. Chloroplast DNA variation and plant phylogeny. *Ann. Mo. Bot. Gard.* **1988**, *75*, 1180-1206.
31. Wambugu, P.W.; Brozynska, M.; Furtado, A.; Waters, D.L.; Henry, R.J. Relationships of wild and domesticated rices (*Oryza* AA genome species) based upon whole chloroplast genome sequences. *Sci. Rep.* **2015**, *5*, 13957.
32. Nguyen, V.B.; Park, H.S.; Lee, S.C.; Lee, J.; Park, J.Y.; Yang, T.J. Authentication markers for five major *Panax* species developed via comparative analysis of complete chloroplast genome sequences. *J. Agric. Food Chem.* **2017**, *65*, 6298-6306.
33. Henriquez, C.L.; Arias, T.; Pires, J.C.; Croat, T.B.; Schaal, B.A. Phylogenomics of the plant family Araceae. *Mol. Phylogenet. Evol.* **2014**, *75*, 91-102.
34. Zhai, W.; Duan, X.S.; Zhang, R.; Guo, C.C.; Li, L.; Xu, G.X.; Shan, H.Y.; Kong, H.Z.; Ren, Y. Chloroplast genomic data provide new and robust insights into the phylogeny and evolution of the Ranunculaceae. *Mol. Phylogenet. Evol.* **2019**, *135*, 12-21.
35. Kirill, A.; Alexander, U.; Maksim, M.; Vladimir, K.; Vera, G. Comparative analysis of chloroplast genomes of seven perennial *Helianthus* species. *Gene* **2021**, *774*, 145418.
36. Shinozaki, K.; Ohme, M.; Tanaka, M.; Wakasugi, T.; Hayashida, N.; Matsubayashi, T.; Zaita, N.; Chunwongse, J.; Obokata, J.; Yamaguchi-Shinozaki, K.; et al. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *EMBO J.* **1986**, *5*, 2043–2049.
37. 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. *Nat. Rev. Genet.* **1986**, *322*, 572-574.
38. Xie, Y.P.; Jiang, X.F.; Yang G.G. The complete plastome of *Primula wilsonii*, a heterostylous ornamental species. *MITOCHONDRIAL DNA PART B* **2021**, *6*, 1324-1325.
39. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* **1987**, *19*, 11-15.
40. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120.
41. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455-477.
42. Boetzer, M.; Henkel, C.V.; Jansen, H.J.; Butler, D.; Pirovano, W. Scaffolding preassembled contigs using SSPACE. *Bioinformatics* **2011**, *27*, 578-579.
43. Boetzer, M.; Pirovano, W. Toward almost closed genomes with GapFiller. *Genome Biol.* **2012**, *13*,

R56

44. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012** , *9* , 357–359.
45. Hyatt, D.; Chen, G.L.; Locascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **2010** , *11* , 119.
46. Wu, S.T.; Zhu, Z.W.; Fu, L.M.; Niu, B.F.; Li, W.Z. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* , **2011** , *12* , 444
47. Laslett, D.; Canback, B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **2004** , *32* , 11-16.
48. Lohse, M.; Drechsel, O.; Kahlau, S.; Bock, R. OrganellarGenomeDRAW - a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **2013** , *41* , W575–W581.
49. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001** , *29* , 4633-4642.
50. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: a web server for microsatellite prediction. *Bioinformatics* **2017** , *33* , 2583-2585.
51. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004** , *32* , 273-279.
52. Brudno, M.; Do, C.B.; Cooper, G.M.; Kim, M.F.; Davydov, E.; NISC Comparative Sequencing Program; Green, E.D.; Sidow, A.; Batzoglou, S. LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **2003** , *13* , 721-731.
53. Rozas, J.; Ferrer-Mata, A.; Sánchez-Delbarrio, J.C.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sánchez-Gracia, A. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* **2017** , *34* , 3299-3302.
54. Amiryousefi, A.; Hyvönen, J.; Poczai, P. IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018** , *34* , 3030-3031.
55. Chen, Z.D.; Yang, T.; Lin, L.; Lu, L.M.; Li, H.L.; Sun, M.; Liu, B.; Chen, M.; Niu, Y.T.; Ye, J.F.; et al. Tree of life for the genera of Chinese vascular plants. *J. Syst. Evol.* **2016** , *54* , 277-306.
56. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013** , *30* , 772–780.
57. Darriba, D.; Taboada G.L.; Doallo, R.; Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* , **2012** , *9* , 772.
58. Stamatakis, A.; Hoover, P.; Rougemont, J. A rapid bootstrap algorithm for the RAxML web servers. *Syst. Biol.* **2008** , *57* , 758-771.
59. Xin, T.Y.; Zhang, Y.; Pu, X.D.; Gao, R.R.; Xu, Z.C.; Song, J.Y. Trends in herbgonomics. *Sci. China Life Sci.* **2019** , *62* , 288- 308.
60. 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.
61. Gichira, A.W.; Li, Z.Z.; Saina, J.K.; Long, Z.C.; Hu, G.W.; Gituru, R.W.; Wang, Q.F.; Chen, J.M. The complete chloroplast genome sequence of an endemic monotypic genus *Hagenia* (Rosaceae): Structural comparative analysis, gene content and microsatellite detection. *Peer J* **2017** , *5* , e2846.
62. Ren, T.; Yang, Y.C.; Zhou, T.; Liu, Z.L. Comparative plastid genomes of *Primula* species: sequence divergence and phylogenetic relationships. *Int. J. Mol. Sci.* **2018** , *19* , 1050.
63. Nguyen, P.A.T.; Kim, J.S.; Kim, J.H. The complete chloroplast genome of colchicine plants (*Colchicum autumnale* L. and *Gloriosa superba* L.) and its application for identifying the genus. *Planta* **2015** , *242* , 223-237.
64. Kim, K.J.; Lee, H.L. Complete chloroplast genome sequences from Korean Ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Research* **2004** , *11* , 247-261.
65. Menezes, A.P.A.; Resende-Moreira, L.C.; Buzatti, R.S.O.; Nazareno, A.G.; Carlsen, M.; Lobo, F.P.; Ka-

- lapothakis, E.; Lovato, M.B. Chloroplast genomes of *Byrsonima* species (Malpighiaceae): comparative analysis and screening of high divergence sequences. *Sci. Rep.* **2018** , 8 , 1-12.
66. Saina, J.K.; Li, Z.Z.; Gichira, A.W.; Liao, Y.Y. The complete chloroplast genome sequence of tree of heaven (*Ailanthus altissima* (mill.)) (sapindales: Simaroubaceae), an important pantropical tree. *Int. J. Mol. Sci.* **2018** ,19 ,
67. Yi, X.; Gao, L.; Wang, B.; Su, Y.J.; Wang, T. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): Evolutionary comparison of *Cephalotaxus* chloroplast DNAs and insights into the loss of inverted repeat copies in Gymnosperms. *Genome Biol. Evol.* **2013** , 5 , 688–698.
68. Sun, Y.X.; Moore, M.J.; Lin, N.; Adelalu, K.F.; Meng, A.; Jian, S.G.; Yang, L.S.; Li, J.Q.; Wang, H.C. Complete plastome sequencing of both living species of Circaeasteraceae (Ranunculales) reveals unusual rearrangements and the loss of the *ndh* gene family. *BMC Genomics* **2017** , 18 , 592.
69. Ren, F.M.; Wang, L.Q.; Li, Y.; Zhuo, W.; Xu, Z.C.; Guo, H.J.; Liu, Y.; Gao, R.R.; Song, J.Y. Highly variable chloroplast genome from two endangered Papaveraceae lithophytes *Corydalis tomentella* and *Corydalis saxicola*. *Ecol. Evol.* **2021** ,11 , 4158-4171.
70. Pombert, J.F.; Lemieux, C.; Turmel, M. The complete chloroplast DNA sequence of the green alga *Oltmannsiellopsis viridis* reveals a distinctive quadripartite architecture in the chloroplast genome of early diverging ulvophytes. *BMC Biol.* **2006** ,4 , 3.
71. Guisinger, M.M.; Kuehl, J.V.; Boore, J.L.; Jansen, R.K. Extreme reconfiguration of plastid genomes in the angiosperm family Geraniaceae: rearrangements, repeats, and codon usage. *Mol. Biol. Evol.* **2011** , 28 , 583–600.
72. Xu, W.B.; Xia, B.S.; Li, X.W. The complete chloroplast genome sequences of five pinnate-leaved *Primula* species and phylogenetic analyses. *Sci. Rep.* **2020** ,10 , 20782.
73. Li, D.M.; Zhu, G.F.; Xu, Y.C.; Ye, Y.J.; Liu, J.M. Complete chloroplast genomes of three medicinal *Alpinia* species: genome organization, comparative analyses and phylogenetic relationships in family Zingiberaceae. *Plants* **2020** , 9 , 286.
74. Hishamuddin, M.S.; Lee, S.Y.; Ng, W.L.; Ramlee, S.I.; Lamasudin, D.U.; Mohamed, R. Comparison of eight complete chloroplast genomes of the endangered *Aquilaria* tree species (Thymelaeaceae) and their phylogenetic relationships. *Sci. Rep.* **2020** , 10 , 13034
75. Gulzar, K.; Zhang, F.Q.; Gao, Q.B.; Fu, P.C.; Zhang, Y.; Chen, S.L. Spiroides shrubs on Qinghai-Tibetan Plateau: multilocus phylogeography and palaeodistributional reconstruction of *Spiraea alpina* and *S. Mongolica* (Rosaceae). *Mol. Phylogenet. Evol.* **2018** , 123 , 137-148.
76. Shen, J.; Zhang, X.; Jacob, B.L.; Zhang, H.J.; Deng, T.; Sun, H.; Wang, H.C. Plastome Evolution in *Dolomiaea* (Asteraceae, Cardueae) using Phylogenomic and Comparative analyses. *Front. Plant Sci.* **2020** , 11 , 376.
77. Wang, L.Y.; Wang, J.; He, C.Y.; Zhang, J.G.; Zeng, Y.F. Characterization and comparison of chloroplast genomes from two sympatric *Hippophae* species (Elaeagnaceae). *J. Forestry Res.* **2021** , 32 , 307-318.
78. Zhu, A.D.; Guo, W.H.; Gupta, S.; Fan, W.S.; Mower, J.P. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. *New Phytol.* **2016** , 209 , 1747-1756.
79. Liu, T.J. A transcriptomic phylogenomic study of *Primula* L. (Primulaceae). Doctor, South China Agreicultural University, Guangzhou, China, **2018**.
80. Degna, J.H.; Rosenberg, N.A. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* **2009** , 24 , 332-340.





Inverted Repeats

