



# Phylogenetic analysis of *Fritillaria cirrhosa* D. Don and its closely related species based on complete chloroplast genomes

Qi Chen<sup>1</sup>, Xiaobo Wu<sup>1</sup> and Dequan Zhang<sup>1,2</sup>

<sup>1</sup> College of Pharmacy and Chemistry, Dali University, Dali, Yunnan, China

<sup>2</sup> Institute of Materia Medica, Dali University, Dali, Yunnan, China

## ABSTRACT

*Fritillaria cirrhosa* D. Don, whose bulb is used in a well-known traditional Chinese medicine to relieve cough and eliminate phlegm, is one of the most important medicinal plants of *Fritillaria* L. The species is widely distributed among the alpine regions in southwestern China and possesses complex morphological variations in different distributions. A series of newly related species were reported, based on obscure morphological differences. As a result, *F. cirrhosa* and its closely related species constitute a taxonomically complex group. However, it is difficult to accurately identify these species and reveal their phylogenetic relationships using traditional taxonomy. Molecular markers and gene fragments have been adopted but they are not able to afford sufficient phylogenetic resolution in the genus. Here, we report the complete chloroplast genome sequences of *F. cirrhosa* and its closely related species using next generation sequencing (NGS) technology. Eight plastid genomes ranged from 151,058 bp to 152,064 bp in length and consisted of 115 genes. Gene content, gene order, GC content, and IR/SC boundary structures were highly similar among these genomes. SSRs and five large repeat sequences were identified and the total number of them ranged from 73 to 79 and 63 to 75, respectively. Six highly divergent regions were successfully identified that could be used as potential genetic markers of *Fritillaria*. Phylogenetic analyses revealed that eight *Fritillaria* species were clustered into three clades with strong supports and *F. cirrhosa* was closely related to *F. przewalskii* and *F. sinica*. Overall, this study indicated that the complete chloroplast genome sequence was an efficient tool for identifying species in taxonomically complex groups and exploring their phylogenetic relationships.

Submitted 9 April 2019  
Accepted 15 July 2019  
Published 21 August 2019

Corresponding author  
Dequan Zhang,  
zhangdeq2008@126.com

Academic editor  
Robert Toonen

Additional Information and  
Declarations can be found on  
page 20

DOI 10.7717/peerj.7480

© Copyright  
2019 Chen et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Genetics, Plant Science, Taxonomy

**Keywords** *Fritillaria cirrhosa* D. Don, Complete chloroplast genome, Closely related species, Taxonomically complex groups, Phylogenetic relationship

## INTRODUCTION

*Fritillaria* L. is one of the most important genera in Liliaceae, which includes approximately 140 species of perennial herbaceous plants (Day et al., 2014; Tekşen, Aytaç & Pınar, 2010). Almost all of the species are distributed in the temperate regions of the northern hemisphere (Rix et al., 2001). There are 24 species in China and most of them possess

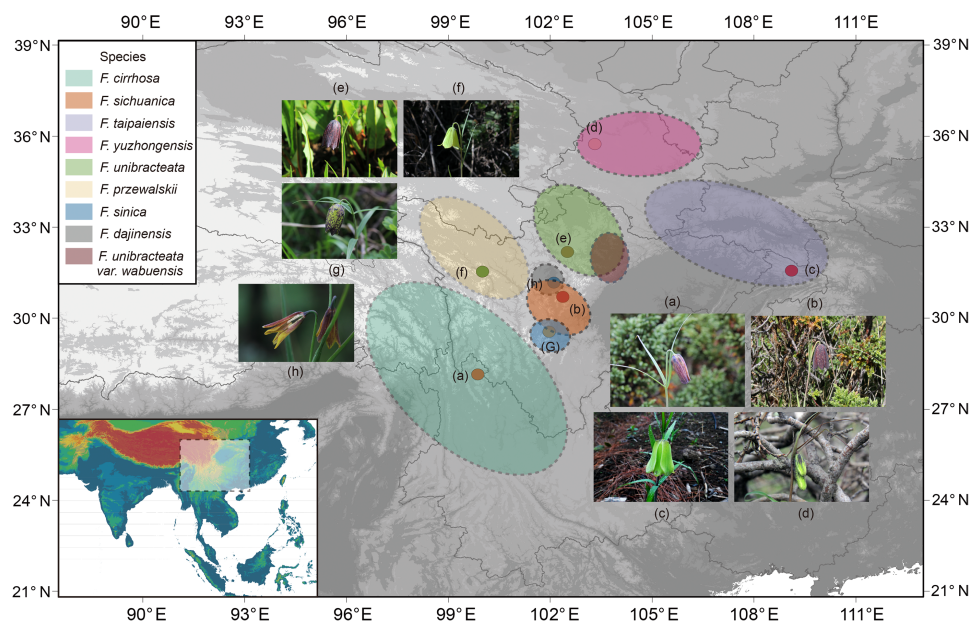
important medicinal properties; these species include *F. cirrhosa* D. Don., *F. ussuriensis* Maxim., *F. walujewii* Regel., *F. thunbergii* Miq., and others (Chen & Helen, 2000; National Pharmacopoeia Committee, 2015). Of these species, *F. cirrhosa* is one of the major original plants of Fritillariae Cirrhosae bulbus, a famous traditional Chinese medicine, which is used to relieve cough and eliminate phlegm (National Pharmacopoeia Committee, 2015). It is mainly found in areas of high altitude in the southwest of China and grows in moist environments near bushes, meadows, and other similar habitats (Chen & Helen, 2000). However, this species exhibits complicated variations in morphology in different regions, especially in flower color and apex shape of bracts. Based on uncertain morphological differences, a series of newly related species were reported. As a result, *F. cirrhosa* and its closely related species constitute a taxonomically complex group that is difficult to be clearly distinguished based on morphological traits alone (Luo & Chen, 1996; Zhang, Zhang & Sun, 2001; Zhang & Cheng, 1998). Luo & Chen (1996) had proposed the concept of a “complex group of *F. cirrhosa*” which was composed of four species, namely *F. cirrhosa*, *F. sichuanica*, *F. taipaiensis* and *F. yuzhongensis*, based on uncertain morphological characteristics and geographical distributions. They also pointed out that *F. sichuanica* was possibly a hybrid among *F. cirrhosa*, *F. przewalskii* and *F. unibracteata*. This posits the question of whether the theory is reasonable and what would then be the phylogenetic relationships of the species in the so-called “complex group of *F. cirrhosa*” and their close relatives.

Over the past few decades, molecular methods have been widely used in plant evolution and phylogeny due to rapid development of molecular technologies. As a result, the well-known APG (Angiosperm Phylogeny Group) classification system was constructed based on the latest progress in plant molecular phylogenetics (Byng et al., 2016; The Angiosperm Phylogeny Group, 1998; The Angiosperm Phylogeny Group, 2003; The Angiosperm Phylogeny Group, 2009). Although molecular markers were also used for phylogenetic inference in complicated groups (Simmons et al., 2007), only a few species of *Fritillaria* adopted them to explore phylogeny (Çelebi et al., 2008; Wietsma et al., 2015). Nevertheless, gene fragments, especially nrITS and cpDNA genes, gained more attention due to the rapid development of DNA sequencing (Gao et al., 2010; Kress et al., 2005; Mishra et al., 2015). Day et al. (2014) elaborated on the evolutionary relationships of 92 *Fritillaria* species based on three plastid regions but most of the species were not well distinguished. Huang et al. (2018) used three plastid markers (*matK*, *rbcL* and *rpl 16*) and nuclear ITS to explore the phylogeny with 191 taxa in the tribe Lilieae (including 57 *Fritillaria* species) but the boundaries between a few species of *Fritillaria* remained ambiguous and needed further research. Meanwhile, the combination of nrITS and cpDNA genes was also adopted to reveal inter-specific relationships and to discriminate between the species of *Fritillaria*. Although these gene fragments had a preliminary resolution on certain species in *Fritillaria*, they could not be effective in discriminating between the closely related species (Khourang et al., 2014; Rønsted et al., 2005; Zhang et al., 2016). Overall, it is probably sufficient to use individual or combined regions based on Sanger sequencing in order to explore the phylogenetic relationships of major genera, but they are generally insufficient for complex groups or closely related species (Liu et al., 2017). Fortunately, with the emergence and development of next-generation sequencing (NGS), the complete chloroplast genome

might be a better tool for discriminating between species and revealing the phylogenetic relationships of complex groups (Mardis, 2008; Parks, Cronn & Liston, 2009; Shendure & Ji, 2008; Tangphatsornruang et al., 2009).

In plants, chloroplasts (cp) are photosynthetic organelles providing the necessary energy for growth and are fundamental in the biosynthesis of starch, fatty acids, pigments, and amino acids (Gao et al., 2010; Neuhaus & Emes, 2000). Typically, angiosperm chloroplasts have a circular genome ranging from 72 to 217 kb and quadripartite structure composed of a large single copy region (LSC), a small single copy region (SSC), and a pair of inverted repeats (IRs) (Moore et al., 2010; Sugiura, 1992; Wang et al., 2015; Yurina & Odintsova, 1998). In contrast with nuclear and mitochondrial genomes, cp genomes are largely conserved in the gene content, organization, and structure. Moreover, they are typically inherited maternally in the angiosperm, which is beneficial in genetic engineering due to the lack of cross-recombination (Maliga, 2002; Tangphatsornruang et al., 2009). The initial cp genomes were sequenced from tobacco and liverwort (Ohyama et al., 1986; Shinozaki et al., 1986). The plastid genome exhibits an overwhelming superiority for use in species discrimination of complex taxa and has been widely used to reveal their unresolved phylogenetics (Bayly et al., 2013; Du et al., 2017; Henriquez et al., 2014), such as in Araceae, Arundinarieae, Lemnoideae, Myrtaceae, Nelumbonaceae, *Amborella*, *Nymphaea*, *Citrus*, *Gossypium* and *Oncidium* (Bayly et al., 2013; Carbonell-Caballero et al., 2015; Ding et al., 2017; Goremykin et al., 2003; Goremykin et al., 2004; Henriquez et al., 2014; Li et al., 2014a; Ma et al., 2014; Wu et al., 2010; Xue et al., 2012). In a recent study, Park et al. (2017) evaluated the relationships between *F. ussuriensis* and *F. cirrhosa* based on the chloroplast genome. Bi et al. (2018) explored the phylogenetic relationship of eight species representing each subgenus of *Fritillaria* using the complete chloroplast genome. Li et al. (2018) also adopted the plastid genomes to reveal inter-specific relationships among seven *Fritillaria* species that were mainly distributed in the Xinjiang province in China. Although these studies had partly revealed the classification and phylogenetics of *Fritillaria* and showed the power of a complete chloroplast genome, *F. cirrhosa* and its closely related species were not included and their relationships were still unresolved. Hence, we attempted to report the complete chloroplast genomes of *F. cirrhosa* and its related species, as well as to explore their phylogenetic relationships.

Here, we presented complete chloroplast genome sequences of several *Fritillaria* species using Illumina sequencing technology and performed comparative analyses of genomic information. Our aims were as follows: (1) to investigate the global structure patterns of eight plastid genomes in *Fritillaria* in this study; (2) to examine the variations of simple sequence repeats (SSRs) and other repeats (tandem, palindrome, forward, reverse, and complement repeats) among the eight *Fritillaria* plastid genomes; (3) to discover highly divergent regions that could be used as specific DNA barcodes for *Fritillaria*; and (4) to reveal phylogenetic relationships between *F. cirrhosa* and its closely related species. The study might provide better understanding of phylogenetic relationships of the complex group and afford sufficient genomic information to use in further research and the application of these medicinal species.



**Figure 1** Distribution of *Fritillaria cirrhosa* and its closely related species. The distribution area of each species is drawn according to the records of Luo & Chen (1996), Liu, Wang & Chen (2009) and some existing voucher specimens (<http://www.cvh.ac.cn/>). Photos of representative living plants of eight *Fritillaria* species: (A) *F. cirrhosa*, (B) *F. sichuanica*, (C) *F. taipaiensis*, (D) *F. yuzhongensis*, (E) *F. unibracteata*, (F) *F. przewalskii*, (G) *F. sinica*, (H) *F. dajinensis*. Topographic data digital elevation modeling (DEM) data were required from the USGS website (<https://glovis.usgs.gov/app?tour>) with a 90-m spatial resolution grid.

Full-size DOI: 10.7717/peerj.7480/fig-1

## MATERIALS & METHODS

### Material sampling

Eight species, including *Fritillaria cirrhosa* and its closed relatives were collected and used in this study (Fig. 1, Table S1). *F. thunbergii* was used as a supplemental outgroup for phylogenetic analysis. The related species were collected in the wilds of the Hengduan Mountains but *F. thunbergii* was cultivated in the Zhejiang province in China. Fresh, unblemished leaves were sampled from healthy, mature individuals and then dried with allochroic silicagel during the field work. Meanwhile, 3–5 individuals with flowers were collected and preserved as voucher specimens that were then used for morphological analysis and taxonomic identification. During the field work, geographic information, such as latitude, longitude, and altitude etc. was determined by Global Position System (GPS, Garmin) and morphologic traits (especially of the flower) were described immediately. All voucher specimens of *Fritillaria* were identified carefully by Dr. Dequan Zhang and deposited at the Herbarium of Medicinal Plants and Crude Drugs of the College of Pharmacy and Chemistry, Dali University.

### DNA extraction, sequencing, and assembly

Total genomic DNA was extracted from about 100 mg of dried leaf material according to a modified CTAB method (Doyle, 1987; Yang, Li & Li, 2014). DNA quality was checked



by electrophoresis on 1.2% agarose gel and then its concentration was determined using SmartSpec™ Plus Spectrophotometer (Bio-Rad, Hercules, CA, USA). DNA extracts were fragmented for 300 bp short-insert library construction and sequenced  $2 \times 150$  bp paired-end (PE) reads on an Illumina HiSeq X-Ten instrument at Beijing Genomics Institute (BGI, Shenzhen, China).

The raw data was filtered using Trimmomatic v.0.32 (Bolger, Lohse & Usadel, 2014) with default settings. Then paired-end reads of the clean data were filtered and assembled into contigs using GetOrganelle.py (Jin et al., 2018) with *Fritillaria cirrhosa* (accession number: KF769143) as a reference (Li et al., 2014b), calling the bowtie2 v., blastN v. and SPAdes v.3.10 (Bankevich et al., 2012). The *de novo* assembly graphs were visualized and edited using Bandage Window dynamic v.8.0 (Wick et al., 2015) and then a whole or nearly whole circular chloroplast genome was generated.

### Genome annotation and sequence submission

The plastid genomes were annotated by aligning to the complete chloroplast genome sequence published in GenBank (*Fritillaria cirrhosa*, accession number: KF769143) using MAFFT (Katoh & Standley, 2013) with default parameters, coupled with manual adjustment using Geneious v.10.1.3 (Kearse et al., 2012). The circular genome map was generated with OGDRAW v.1.2 (Lohse et al., 2013). Finally, the annotated chloroplast genomes of the nine *Fritillaria* species were submitted to GenBank (Table 1, Table S1).

### Genome comparative analysis

In this study, the multiple sequence alignment of chloroplast genome sequences was performed using MAFFT v.7.129 with default settings and adjusted manually in BioEdit v.7.0.9 (Hall, 1999; Katoh & Standley, 2013). The mVISTA software was used to compare the complete chloroplast genome of *F. cirrhosa* with eight other *Fritillaria* species, taking annotation of the chloroplast genome of *F. cirrhosa* (accession number: KF769143) as a reference. Default parameters were utilized to align the chloroplast genomes in Shuffle-LAGAN mode (Frazer et al., 2004). P-distance, GC content, and codon use were analyzed by the software MEGA v.7.0.26 (Kumar, Stecher & Tamura, 2016). DnaSP v.6.11 (Rozas et al., 2017) was adopted to calculate the variable and parsimony information sites and nucleotide diversity of five regions (whole chloroplast genome, large single copy, small single copy, inverted repeat regions, and protein coding genes). Additionally, the IR contraction/expansion regions were compared among the eight *Fritillaria* species.

### Characterization of repeat sequences and SSRs

REPuter software was used to identify repeat sequences, including palindromic, complement, reverse, and forward repeats within the chloroplast genome. The following conditions for repeat identification were used in the analysis: (i) Hamming distance of 3, (ii) 90% or greater sequence identity, and (iii) a minimum repeat size of 30 bp (Kurtz et al., 2001). MISA was adopted to evaluate SSRs. The minimum thresholds were set to ten repeat units for mononucleotide SSRs, five repeat units for dinucleotide SSRs, four repeat units for trinucleotide, and three repeat units for tetranucleotide, pentanucleotide, hexanucleotide SSRs (Murat et al., 2011). In addition, tandem repeats in eight *Fritillaria*

**Table 1** Summary of complete chloroplast genomes for eight *Fritillaria* species.

	Total (bp)	Large single copy (LSC, bp)	Small single copy (SSC, bp)	Inverted repeat (IR, bp)	GC%	Total genes	Protein coding genes	tRNA	rRNA	Accession number in GenBank
<i>F. cirrhosa</i>	151,998	81,755	17,545	26,349	36.9%	115	78	30	4	<a href="#">MH244906</a>
<i>F. sichuanica</i>	151,958	81,726	17,542	26,345	37.0%	115	78	30	4	<a href="#">MH244907</a>
<i>F. przewalskii</i>	151,983	81,744	17,539	26,350	36.9%	115	78	30	4	<a href="#">MH244908</a>
<i>F. unibracteata</i>	151,058	81,339	17,539	26,090	37.0%	115	78	30	4	<a href="#">MH244909</a>
<i>F. taipaiensis</i>	151,707	81,451	17,552	26,352	37.0%	115	78	30	4	<a href="#">MH244910</a>
<i>F. yuzhongensis</i>	151,645	81,417	17,526	26,351	37.0%	115	78	30	4	<a href="#">MH244911</a>
<i>F. sinica</i>	152,064	81,827	17,537	26,350	36.9%	115	78	30	4	<a href="#">MH244912</a>
<i>F. dajinensis</i>	151,991	81,723	17,540	26,364	36.9%	115	78	30	4	<a href="#">MH244913</a>

species chloroplast genomes were identified using Tandem Repeats Finder v.4.09 with the following settings: 80, 10, 50, and 500 for match probability, indel probability, minimum alignment score, and maximum period size, respectively (Benson, 1999).

### Chloroplast genome analysis by sliding window

After using MAFFT v.7.129 to align the chloroplast genome sequences, BioEdit software was used to adjust the sequences manually (Hall, 1999; Katoh & Standley, 2013). A sliding window analysis was conducted for nucleotide variability (Pi) in the whole chloroplast genome using the DanSP. The step size was set to 200 bp, with a 600 bp window length (Rozas et al., 2017).

### Phylogenetic analyses

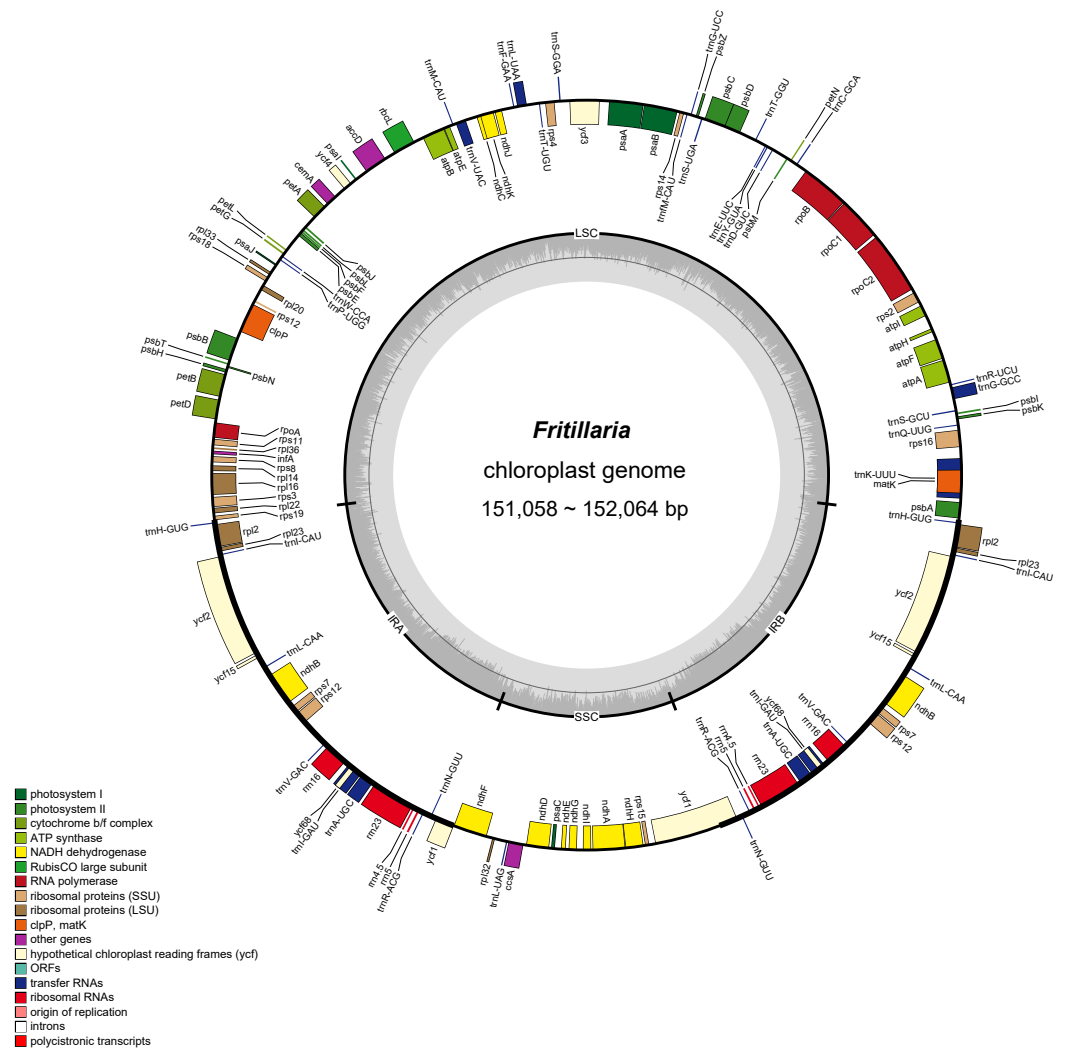
The eight species of *F. cirrhosa* and its closely related species were used for phylogenetic analysis, to be supplemented with *F. thunbergii* (accession number: MH244914) and *Lilium brownii* F. E. Brown ex Mieliez (accession number: NC\_035588) as outgroups (Du et al., 2017). Furthermore, the available chloroplast genome sequence of *F. unibracteata* var. *wabuensis* (KF769142), which was a variety of *F. unibracteata*, was downloaded from GenBank for our phylogenetic analysis (Li et al., 2016). Phylogenetic inference was performed based on the following five data sets: (1) chloroplast genome sequence (only containing one IR), (2) large single copy region, (3) small single copy region, (4) inverted repeat region, and (5) protein-coding genes. The sequences were aligned using MAFFT and then edited by BioEdit manually (Hall, 1999; Katoh & Standley, 2013). Lengths of aligned sequences were shown in Table S8. In order to explore the phylogenetic relationship of *F. cirrhosa* and its closely related species, Bayesian inference (BI), Maximum parsimony (MP) and Maximum likelihood (ML) methods were adopted for phylogenetic inference, respectively.

MEGA v.7.0.26 was used for MP analysis with 1,000 bootstrap replicates (Kumar, Stecher & Tamura, 2016). For BI and ML analysis, the best substitution models were tested based on Akaike information criterion (AIC) by jModelTest v.2.1.7 (Darriba et al., 2012). The best-fitting models in the analysis were GTR+I+G for LSC and SSC region, and GTR+I for others (Table S8). ML analysis was performed with RAxML v.8.2.4 (Stamatakis, 2014). And 1,000 replications were adopted to calculate the local bootstrap probability of each branch. BI analysis was conducted in MrBayes v.3.2.6 (Ronquist et al., 2012). The Markov Chain Monte Carlo (MCMC) algorithm was calculated for 1,000,000 generations with a sampling of trees every 1,000 generations. The first 25% of generations were discarded as burn-in. Stasis was considered to be reached when the average standard deviation of split frequencies was <0.01 and a consensus tree was constructed using the remaining trees.

## RESULTS

### Chloroplast genome organization of *Fritillaria* chloroplast genomes

Nucleotide sequences of the eight *Fritillaria* chloroplast genomes ranged from 151,083 bp in *F. unibracteata* to 152,064 bp in *F. sinica* and shared the typical quadripartite structure, composed of a pair of IRs (26,090-26,364 bp) separated by the LSC (81,339-81,827 bp) and



**Figure 2** Gene map of *Fritillaria* chloroplast genomes. Genes outside the circle are transcribed clockwise, and genes shown on the inside of the circle are counter-clockwise. Genes belonging to functional group are color-coded. The darker gray in the inner corresponds to GC content, and the lighter gray corresponds to AT content.

Full-size DOI: 10.7717/peerj.7480/fig-2

SSC (17,526-17,545 bp) regions (Table 1, Fig. 2). GC content of the complete chloroplast genomes was 36.9%-37.0% (Table 1). The content of the IR regions (42.5%) was higher than that of whole genome (36.9%), LSC (34.9%), and SSC (30.5%) in *F. cirrhosa* due to the presence of eight rRAN (55%) sequences in these regions (Table 2).

In the eight whole chloroplast genomes, a total of 115 genes were found, including 78 protein coding genes, 30 tRNA genes, four rRNA genes, and three pseudogenes (*infA*, *ycf15* and *ycf68*) (Table 1, Fig. 2). The protein coding genes present in the chloroplast genome of eight *Fritillaria* genomes included nine genes for large ribosomal proteins (*rpl2*, *rpl14*, *rpl16*, *rpl20*, *rpl22*, *rpl23*, *rpl32*, *rpl33*, *rpl36*), 12 genes for small ribosomal proteins (*rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps11*, *rps12*, *rps14*, *rps15*, *rps16*, *rps18*, *rps19*), five genes for photosystem I



**Table 2** Base composition in *Fritillaria cirrhosa* chloroplast genome.

	T/U%	C%	A%	G%	AT%	Length (bp)
Genome	31.9	18.8	31.1	18.1	63.1	151,998
LSC	33.3	17.9	31.9	17.0	65.1	81,755
SSC	35.0	16.1	34.5	14.4	69.5	17,545
IR	28.5	20.5	29	22.0	57.5	26,349
tRNA	25.0	23.7	21.9	29.4	46.9	2,877
rRNA	18.9	23.5	26.0	31.5	45.0	9,052
Protein Coding genes	31.7	17.3	31.0	20.0	62.7	68,234
1st position codon	24.6	18.1	30.9	26.4	55.5	22,745
2nd position codon	32.2	19.9	29.9	18.1	62.0	22,745
3rd position codon	38.3	14.0	32.1	15.6	70.4	22,744

(*psaA*, *psaB*, *psaC*, *psaI*, *psaJ*), 15 genes for photosystem II (*psbA*, *psbB*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbN*, *psbT*, *psbZ*), and six genes for ATP synthase (*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *atpI*) (Table 3, Fig. 2). Furthermore, 20 duplicated genes were found in the IR regions, as well as five protein coding genes, 11 tRNA genes and four rRNA genes. 26 protein coding genes possessed introns (Fig. 2).

Protein coding genes, rRNA and tRNA were encoded by 44.89%, 5.96%, and 1.89% in the *F. cirrhosa* whole chloroplast genome, respectively, and the remaining 47.26% was non-coding regions. The 20 amino acids crucial for protein biosynthesis were encoded by 30 tRNA. Moreover, protein coding genes included 78 protein genes and the length was 68,234 bp, which comprised 22,396 codons (Table 2). Interestingly, among all of the encoded amino acids, leucine (10.32%) and cysteine (1.57%) were the maximum and minimum commonly detected amino acids, respectively (Table S2). Within the protein coding regions, the AT percentages for the first, second, and third codons were 55.5%, 62.0% and 70.4% in *F. cirrhosa*, respectively (Table 2).

### SSR analysis of *Fritillaria* chloroplast genomes

Numerous SSR loci were found through the MISA analysis of nine *Fritillaria* chloroplast genome sequences. In total, five types of SSR (mononucleotide, dinucleotide, trinucleotide, tetranucleotide, and pentanucleotide repeats) were detected based on the comparison of eight *Fritillaria* cp genomes. A total of 78 perfect SSRs were found in *F. cirrhosa* (Fig. 3A). Similarly, 73, 74, 78, 79, 77, 76, and 75 SSRs were detected in *F. sichuanica*, *F. przewalskii*, *F. unibracteata*, *F. taipaiensis*, *F. yuzhongensis*, *F. sinica*, and *F. dajinensis*. Lengths of those SSRs ranged from 10 to 22 bp (Table S3). The most abundant type of SSR were mononucleotide repeats ranging from 51 bp in *F. sichuanica* to 56 bp in *F. unibracteata*, followed by dinucleotide repeats, tetranucleotide repeats, trinucleotide repeats, and pentanucleotide repeats (Fig. 3A). In the cp genome of *F. cirrhosa*, all mononucleotide repeats are composed of A (47.27%) and T (52.72%) motifs in the majority of dinucleotide repeats are AT (64.29%) (Fig. 3B).

Further analysis revealed that most of the microsatellites were located in the LSC region, with a small portion distributed through the SSC and IR regions (Fig. 3C). Moreover, the

**Table 3** Gene contents in eight *Fritillaria* chloroplast genome.

Category for gene	Group of genes	Name of genes
Self-replication	Large subunit of ribosome	<i>rpl2</i> <sup>1*</sup> , <i>rpl14</i> , <i>rpl16</i> <sup>*</sup> , <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> <sup>1</sup> , <i>rpl32</i> , <i>rpl33</i> , <i>rpl36</i>
	Small subunit of ribosome	<i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> <sup>1</sup> , <i>rps8</i> , <i>rps11</i> , <i>rps12</i> <sup>1*</sup> , <i>rps14</i> , <i>rps15</i> , <i>rps16</i> <sup>*</sup> , <i>rps18</i> , <i>rps19</i>
	DNA dependent RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> <sup>*</sup> , <i>rpoC2</i>
	rRNA gene	<i>rrn4.5</i> <sup>1</sup> , <i>rrn5</i> <sup>1</sup> , <i>rrn16</i> <sup>1</sup> , <i>rrn23</i> <sup>1</sup>
	tRNA gene	<i>trnK-UUU</i> <sup>*</sup> , <i>trnI-GAU</i> <sup>1*</sup> , <i>trnA-UGC</i> <sup>1*</sup> , <i>trnG-GCC</i> <sup>*</sup> , <i>trnV-UAC</i> <sup>*</sup> , <i>trnL-UAA</i> <sup>*</sup> , <i>trnS-UGA</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnY-GUA</i> , <i>trnC-GCA</i> , <i>trnL-CAA</i> <sup>1</sup> , <i>trnL-UAG</i> , <i>trnH-GUG</i> <sup>1</sup> , <i>trnD-GUC</i> , <i>trnM-CAU</i> , <i>trnW-CCA</i> , <i>trnP-UGG</i> , <i>trnI-CAU</i> <sup>1</sup> , <i>trnR-ACG</i> <sup>1</sup> , <i>trnI-CAU</i> <sup>1</sup> , <i>trnE-UUC</i> , <i>trnT-UGU</i> , <i>trnF-GAA</i> , <i>trnQ-UUG</i> , <i>trnR-UCU</i> , <i>trnT-GGU</i> , <i>trnM-CAU</i> , <i>trnV-GAC</i> <sup>1</sup> , <i>trnN-GUU</i> <sup>1</sup> , <i>trnN-GUU</i> <sup>1</sup> , <i>trnV-GAC</i> <sup>1</sup> , <i>trnG-UCC</i>
Gene for photosynthesis	Subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>
	Subunits of 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>
	Subunits of NADH-dehydrogenase	<i>ndhA</i> <sup>*</sup> , <i>ndhB</i> <sup>1*</sup> , <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i>
	Subunits of cytochrome b/f complex	<i>petA</i> , <i>petB</i> <sup>*</sup> , <i>petD</i> <sup>*</sup> , <i>petG</i> , <i>petL</i> , <i>petN</i>
	Subunit for ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> <sup>*</sup> , <i>atpH</i> , <i>atpI</i>
	Large subunit of rubisco	<i>rbcL</i>
Other genes	Translational initiation factor	<i>infA</i>
	Maturase	<i>matK</i>
	Protease	<i>clpP</i> <sup>*</sup>
	Envelope membrane protein	<i>cemA</i>
	Subunit of Acetyl-carboxylase	<i>accD</i>
	C-type cytochrome synthesis gene	<i>ccsA</i>
	Open reading frames(ORF,ycf)	<i>ycf1</i> , <i>ycf2</i> <sup>1</sup> , <i>ycf3</i> <sup>*</sup> , <i>ycf4</i> , <i>ycf15</i> <sup>1</sup> , <i>ycf68</i> <sup>1</sup>

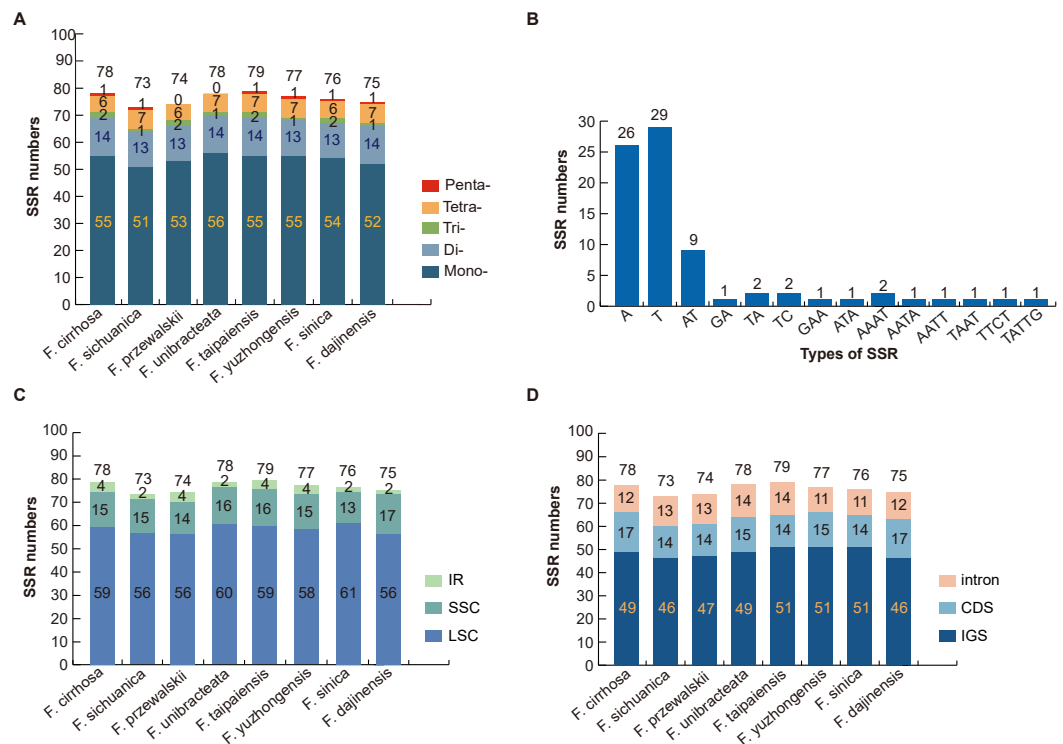
**Notes.**

The I label after gene names reflect genes located in IR regions. Intron containing gene is indicated by one asterisk.

SSRs in the genomes were distributed mainly in the intergenic spacer (IGS), with others dispersed at similar levels in introns and protein coding genes (CDS) (Fig. 3D). Seven protein coding genes in the SSR loci were *rpoC2*, *cemA*, *ndhD*, *ndhG*, *ndhH*, *ycf1*, and *ycf2* in the CDS regions of the *Fritillaria* cp genome (Table S3).

**Other repeats analysis of *Fritillaria* chloroplast genomes**

A total of 63 repeats including tandem, palindrome, forward, reverse, and complement repeats were found in the *F. cirrhosa* chloroplast genome. Similarly, 65, 66, 70, 66, 73, 75, and 73 repeats were detected in *F. sichuanica*, *F. przewalskii*, *F. unibracteata*, *F. taipaiensis*, *F. yuzhongensis*, *F. sinica*, and *F. dajinensis*, respectively (Fig. 4A). Among these, tandem repeats, which had larger numbers than others, were mainly distributed in the intergenic spacer (IGS), with others dispersed in protein coding genes (CDS) and introns (Fig. 4B). The tandem repeats in the CDS regions were located in five protein coding genes (*trnK-UUU*, *rps11*, *rps16*, *ycf1*, and *ycf2*) of the plastid genomes (Table S4) and mainly ranged from 10 to 29 bp in length, whereas only one tandem repeat longer than 40 bp was found in the



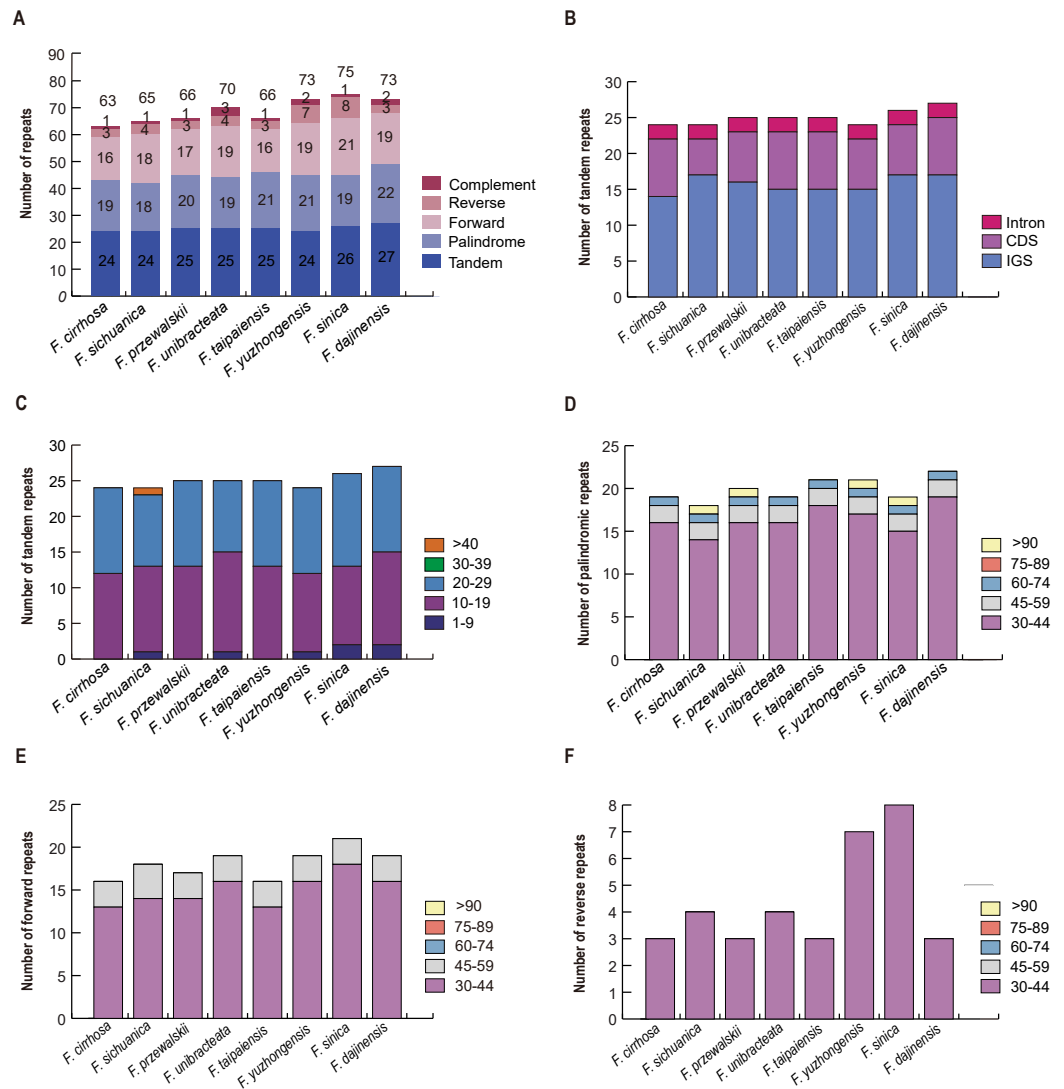
**Figure 3** Analysis of simple sequence repeat (SSR) in eight *Fritillaria* cp genomes. (A) Number different SSRs type detected in nine genomes; (B) frequency of SSR motifs in different repeat types of *F. cirrhosa* cp genome; (C) frequency of identified SSR in LSC, SSC, and IR regions; (D) frequency of identified SSR in IGS, CDS, and intron.

Full-size DOI: 10.7717/peerj.7480/fig-3

*F. sichuanica* genome (Fig. 4C). In the remaining four repeats, most occurred in the regions of the intergenic spacer, whereas some were found in the protein coding genes and intron (Table S5). Copy lengths with 30–44 bp were the most common. Moreover, the length of palindrome repeats more than 90 bp were found in four plastid genomes (*F. sichuanica*, *F. przewalskii*, *F. yuzhongensis*, and *F. sinica*). However, almost all of the lengths of the forward and reverse repeats were less than 59 bp in eight *Fritillaria* chloroplast genomes (Fig. 4D–4F, Tables S4–S6).

### Comparison of chloroplast genome among *F. cirrhosa* and related species

The annotation of *F. cirrhosa* (accession number: KF769143) was used as a reference for visualization analysis of the pairwise chloroplast genomic alignment between *F. cirrhosa* and its closely related species using mVISTA (Fig. 5). The alignment revealed a high sequence similarity across eight *Fritillaria* plastid genomes, which showed that the genomes were highly conserved. Furthermore, a vast majority of sequence variations were concentrated in the single copy regions, compared with the least number in the IR regions. This indicated that there were higher divergence levels in the single copy regions than that in the IR regions. Moreover, coding regions were less divergent than non-coding regions. Similarly,



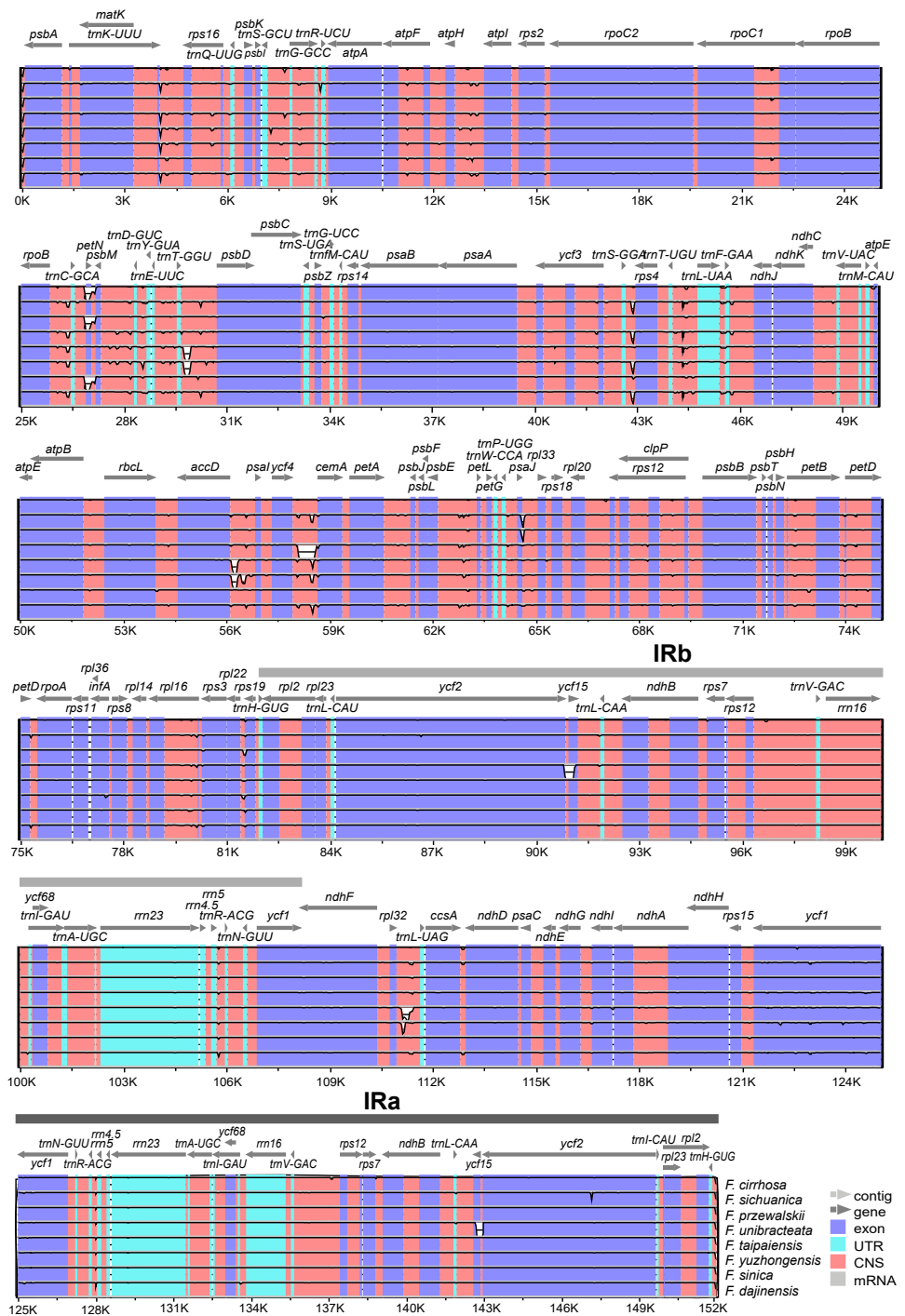
**Figure 4** Analysis of large repeat sequences in eight *Fritillaria* cp genomes. (A) Total of five repeat types; (B) frequency of tandem repeats in IGS, CDS, and intron; (C) frequency of tandem repeats by length; (D) frequency of palindromic repeats by length; (E) frequency of forward repeats by length; (F) frequency of reverse repeats by length.

Full-size DOI: 10.7717/peerj.7480/fig-4

sequence divergence in the intron was higher than that in the exon. Highly divergent regions among eight *Fritillaria* chloroplast genomes were mainly located in the intergenic spacers, including *atpH-atpI*, *rpoB-trnC-GCA*, *petN-psbM*, *psbM-trnD-GUC*, *trnT-GGU-psbD*, *trnS-GGA-rps4*, *trnT-UGU-trnL-UAA*, *accD-psaI*, *ycf4-cemA*, and *psbE-petL*, but others (*matK* and *ycf1*) were distributed in protein coding regions.

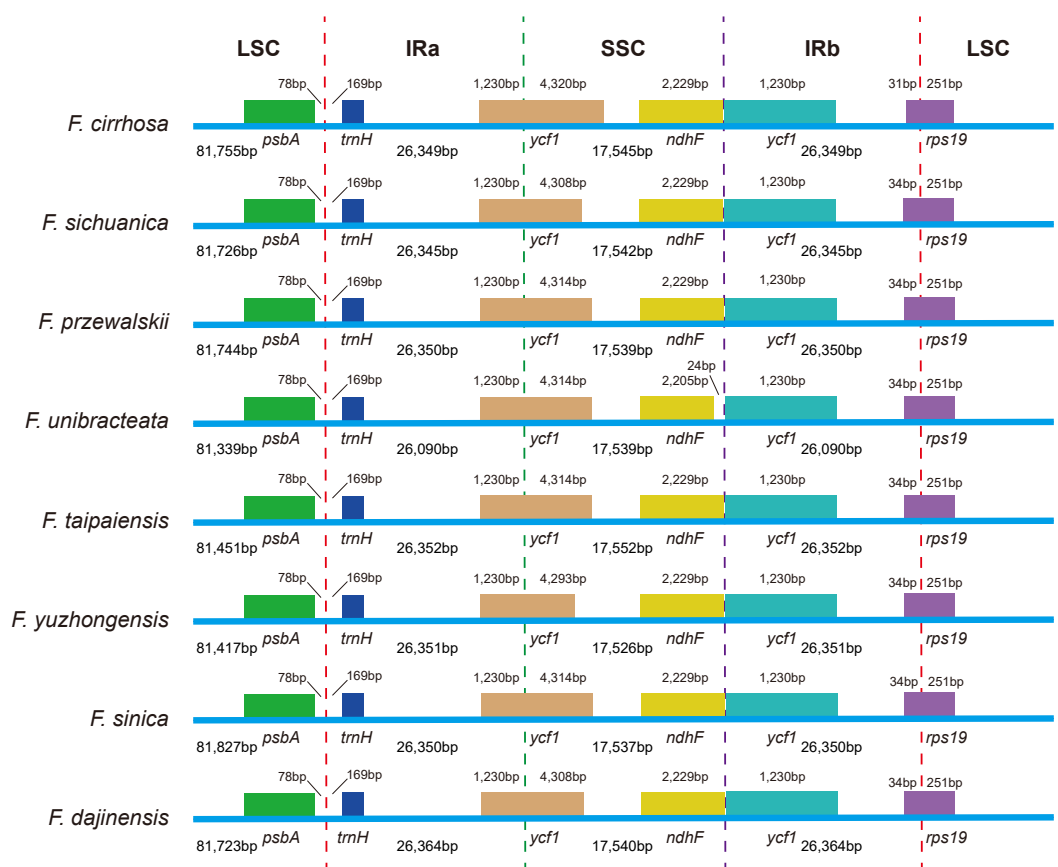
Expansion and contraction at the boundaries of IR regions of eight *Fritillaria* chloroplast genomes were revealed and a detailed comparison of four junctions of two IRs between *F. cirrhosa* and its closely related species was performed (Fig. 6). There were some differences in length compared with each region among the *Fritillaria* chloroplast genomes, but they





**Figure 5** Visualization alignment of nine *Fritillaria* cp genomes. VISTA-based identify plot showing sequence identify among eight *Fritillaria* species using *Fritillaria cirrhosa* D. Don as a reference. The thick black line shows the inverted repeats (IRs) in the chloroplast genomes.

Full-size [DOI: 10.7717/peerj.7480/fig-5](https://doi.org/10.7717/peerj.7480/fig-5)



**Figure 6** Comparison of LSC, SSC, and IR border regions among eight *Fritillaria* cp genomes. Colored boxes for genes represent the gene position.

Full-size DOI: 10.7717/peerj.7480/fig-6

exhibited striking similarities on the IR borders. Although IR regions were highly conserved, subtle structure variation was still observed in the chloroplast genomes. In contrast, *ycf1* was mainly located in the SSC region ranging from 4,293 bp to 4,320 bp and others 1,230 bp in IRa region. The border between IRb/LSC extended into the *rps19*, but there were only 31 bp in the IRb region of *F. cirrhosa*. Moreover, variation was found in *F. unibracteata*, and *ndhF* was 24 bp away from the SSC/IRb border.

Evolutionary divergences and differences among the eight *Fritillaria* chloroplast genomes were compared using sequence distance and nucleotide substitutions. Across all the species, p-distance was 0.0003–0.0023, and the value of the nucleotide differences was 52–340 (Table 4). The p-distance in three *Fritillaria* (*F. cirrhosa*, *F. przewalskii*, and *F. sinica*) was between 0.0005–0.0008.

### Divergence region in chloroplast genome of the *F. cirrhosa* and related species

Nucleotide diversity of highly variable regions was calculated with a sliding window (step size was set to 200 bp, with a 600 bp window length) to estimate the divergence level of different regions in the eight *Fritillaria* plastid genomes. Of these, the SSC region exhibited

**Table 4** Number of nucleotide substitutions and sequence distance in eight complete chloroplast genomes.

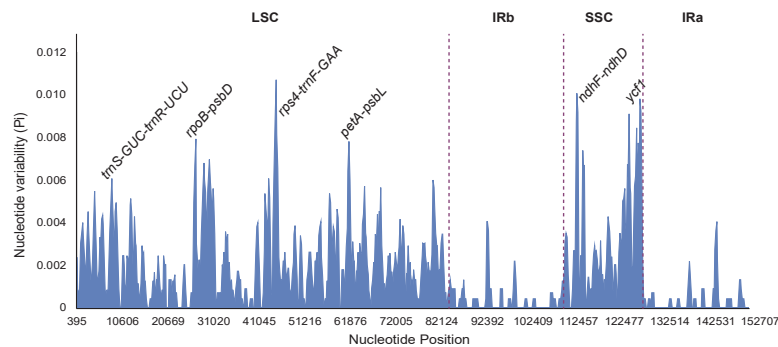
	<i>F. cirrhosa</i>	<i>F. sichuanica</i>	<i>F. przewalskii</i>	<i>F. unibracteata</i>	<i>F. taipaiensis</i>	<i>F. yuzhongensis</i>	<i>F. sinica</i>	<i>F. dajinensis</i>
<i>F. cirrhosa</i>		311	112	314	335	310	117	311
<i>F. sichuanica</i>	0.0021		328	95	290	261	331	52
<i>F. przewalskii</i>	0.0007	0.0022		317	340	314	81	328
<i>F. unibracteata</i>	0.0021	0.0006	0.0021		277	252	320	105
<i>F. taipaiensis</i>	0.0022	0.0019	0.0023	0.0018		169	337	294
<i>F. yuzhongensis</i>	0.0021	0.0017	0.0021	0.0017	0.0011		313	261
<i>F. sinica</i>	0.0008	0.0022	0.0005	0.0021	0.0022	0.0021		333
<i>F. dajinensis</i>	0.0021	0.0003	0.0022	0.0007	0.0020	0.0017	0.0022	

**Notes.**

The upper triangle shows number of nucleotide substitutions and the lower triangle indicates genetic distance in complete cp genomes among species.

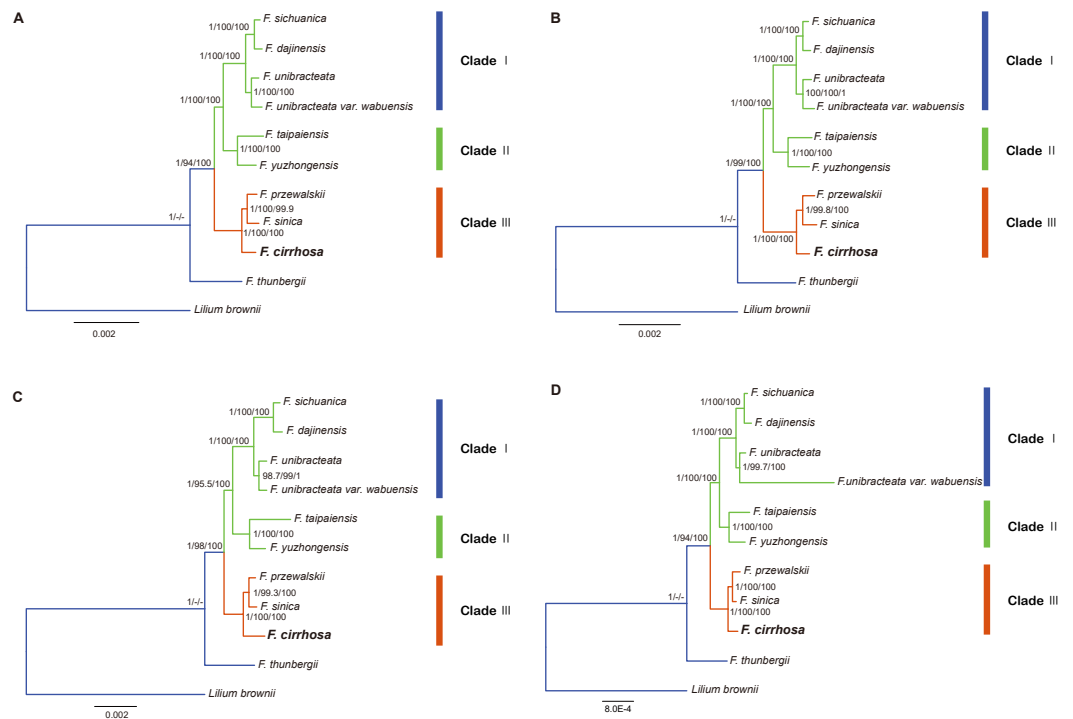
**Table 5** Variable site analysis in *Fritillaria* chloroplast genomes.

	Number of sites	Number of variable sites	Number of parsimony information sites	Nucleotide diversity
Complete cp genome	152,707	728	342	0.00172
LSC	82,378	514	243	0.00223
SSC	17,582	162	74	0.00332
IR	26,372	27	13	0.00038
Protein coding genes	68,709	237	112	0.00129

**Figure 7** Sliding window analysis of eight *Fritillaria* cp genomes (window length: 600 bp, step size: 200 bp). X-axis: position of the midpoint of a window; Y-axis: nucleotide diversity of each window.

Full-size [DOI: 10.7717/peerj.7480/fig-7](https://doi.org/10.7717/peerj.7480/fig-7)

the highest divergence levels (0.00332) and IR regions had the least (0.00038) (Table 5). Furthermore, six regions with a relatively high variability, including 5 intergenic regions (*trnS-GCU-trnR-UCU*, *rpoB-psbD*, *rps4-trnF-GAA*, *petA-psbL*, and *ndhF-ndhD*) and one gene region (*ycf1*) from the genomes, were selected as potentially suitable gene fragments for the study of species identification and phylogenetics in *Fritillaria* (Fig. 7). All highly divergent sequences were found in the SC regions whereas no higher variable loci were found in the IR regions. The six highly variable regions included 257 variable sites which



**Figure 8** Phylogenetic relationship of nine *Fritillaria* species inferred from Bayesian analyses (BI), maximum parsimony (MP), and maximum likelihood (ML) of different datasets. (A) Chloroplast genome (Only contains one IR); (B) LSC region; (C) SSC region; (D) protein coding region. Number above nodes are support values with Bayesian posterior probabilities (PP) values on the left, MP bootstrap values in the middle, ML bootstrap values on the right.

Full-size DOI: 10.7717/peerj.7480/fig-8

possessed 116 parsimony informative sites and their nucleotide diversity values ranged from 0.00455 to 0.00935 (Table S7). The *petA-psbL* showed the highest variability, the next more variable regions were *rps4-trnF-GAA*, *ndhF-ndhD*, *ycf1* and *rpoB-psbD*, but that of *trnS-GCU-trnR-UCU* was the lowest.

### Phylogenetic relationship of *F. cirrhosa* and related species

In this study, five datasets extracted from the eleven plastid genomes were used for phylogenetic analysis (*Fritillaria thunbergii* and *Lilium brownie* were used as outgroups). BI, MP, ML analyses were performed to construct phylogenetic trees using the datasets (Table S8) and topology structures of the previous three trees were nearly identical. Finally, the BI tree was adopted to present phylogenetic results, with the addition of support values from MP and ML analyses. The phylogenetic tree based on different datasets achieved higher support values, except the IR dataset (Fig. 8, Fig. S1). According to the trees, the eight species of *Fritillaria* were obviously divided into three clades (clade I, II and III). Clade I contained four species with strong support, namely *F. sichuanica*, *F. dajinensis*, *F. unibracteata*, and *F. unibracteata* var. *wubuensis*. It was revealed that *F. sichuanica* had a close relationship with *F. dajinensis*. *Fritillaria taipaiensis*, and *F. yuzhongensis*, both of which were distributed in the northern edge of the complex group and were clustered into



one clade (II). The last clade (III) was composed of *F. cirrhosa*, *F. przewalskii* and *F. sinica* which revealed that *F. cirrhosa* was a sister species to the latter two species.

## DISCUSSION

### Comparative analysis of *Fritillaria* chloroplast genomes

Eight plastid genomes in this study ranged from 151,009 bp to 152,064 bp, consisting of 115 genes with a GC content of 36.9%-37.0% (Table 1, Table 2 and Fig. 2). In the chloroplast genome of *F. cirrhosa*, the GC content of the IR regions (42.5%) was highest, which could be attributed to the presence of eight rRNA (55%) sequences in these regions (Table 2). The present results were similar to previous reports with a higher GC content in the IR regions (Bi et al., 2018; Li et al., 2018; Park et al., 2017). These studies might be beneficial for systematically recognizing the gene number, gene order, and chloroplast genome structure of *Fritillaria*. Furthermore, protein coding genes of the *F. cirrhosa* genome were encoded by 44.89% and the AT percentage of the third codon in them was 70.4%. Preference for a higher AT content at the third codon position has been also observed in other terrestrial plant chloroplast genomes (Asaf et al., 2016; Liu & Xue, 2005; Qian et al., 2013; Tangphatsornruang et al., 2009).

SSRs in the chloroplast genome (cpSSRs), which are 1-6 bp repeating sequences and distributed throughout the genome, have been used for the study of population genetics because of their high variability (Asaf et al., 2016; Pauwels et al., 2012; Powell et al., 1995). In this study, certain parameters were set as microsatellites of more than 10 bp are prone to slipped-strand mispairing (Raubeson et al., 2007; Rose & Falush, 1998). Five types of SSR (mononucleotide, dinucleotide, trinucleotide, tetranucleotide, and pentanucleotide repeats) were detected and the number of them ranged from 73 to 79 (Fig. 3). The detected SSRs were located in seven protein coding genes (*rpoC2*, *cemA*, *ndhD*, *ndhG*, *ndhH*, *ycf1* and *ycf2*) of the *Fritillaria* plastid genomes. In the previous study, Bi et al. (2018) observed that five types of SSRs were located in nine protein coding genes (*matK*, *rpoC1*, *rpoC2*, *cemA*, *ndhD*, *ndhG*, *ndhH*, *ycf1* and *ycf2*). Lu, Li & Qiu (2016) found that 15 SSRs were located in eight protein coding genes (*rpoC2*, *cemA*, *rpl22*, *ndhD*, *ndhE*, *ndhH*, *ycf1* and *ycf2*) of three *cardiocrinum* plastid genomes. Therefore, all the studies strongly indicated that the chloroplast genome could be used for developing lineage-specific cpSSR markers that could help for studies on population genetics of the *Fritillaria* species.

Repeat sequences play an important role in genome rearrangement and variation due to the illegitimate recombination and slipped-strand mispairing in the chloroplast genome (Cavalier-Smith, 2002; Lu, Li & Qiu, 2016; Yuan et al., 2017). In the present repeat analysis, five types of repeats including tandem, palindrome, forward, reverse, and complement repeats were identified (Fig. 4). Among them, tandem repeats had the largest numbers and were mainly distributed in the intergenic spacer (IGS). Although substantial repeats have been distinguished in the chloroplast genome of higher plants, the mechanism for the origin of these tandem repeats was unclear (Vieira et al., 2014; Yi et al., 2013). Significant correlations have been observed among DNA rearrangement, mutation, and gene duplication (Cosner et al., 1997; Do, Kim & Kim, 2014; Vieira et al., 2014;

Yi et al., 2013). It was reported that repeat sequences made sense for population genetics because of their significance in rearrangement (Cavalier-Smith, 2002). Most of the remaining four repeats occurred in intergenic spacer regions and the lengths ranged from 9 to 95 (4D–4F, Tables S5–S6). The results for the locations and sequence lengths of the four major repeats were similar to the latest studies (Bi et al., 2018; Park et al., 2017). The research also revealed that repeat sequences were caused by illegitimate recombination and slipped-strand mispairing in the genome (Cavalier-Smith, 2002; Lu, Li & Qiu, 2016; Yuan et al., 2017). Furthermore, the region where the repeats existed was a potential hotspot for genomic reconfiguration (Gao et al., 2009). Additionally, these repeat motifs might provide some informative sources to develop genetic markers for analysis on population genetics (Nie et al., 2012).

Expansion and contraction at the boundaries on the IR regions of the chloroplast genome are important factors that cause size variations and this plays a major role in structural stability and evolution (Asaf et al., 2018; Dang et al., 2014; Wang et al., 2008). In this study, a detailed comparison of four junctions of two IRs between *F. cirrhosa* and its closely related species was performed. The IR regions are highly conserved and structure variation was not significant in the eight *Fritillaria* chloroplast genomes (Fig. 6).

### Identification of highly variable regions

Highly variable regions of the chloroplast genomes could not only be used for resolving phylogeny and identifying species at the species level, but also provide crucial information to explore species divergence and population structure at the population level (Dang et al., 2014; Du et al., 2017). Nucleotide diversity was calculated with a sliding window to estimate the divergence of different regions in eight *Fritillaria* cp genomes. Of these regions, the SSC region exhibited the highest value (0.00332) and the IR regions had the least (0.00038) (Table 5). Once again it indicated that IR regions were conserved in eight *Fritillaria* cp genomes. Similar results related to these regions have been reported in the latest studies of *Fritillaria* (Bi et al., 2018; Park et al., 2017) and have also been found in *Lilium* (Du et al., 2017). Furthermore, six relatively highly variable regions, including 5 intergenic regions (*trnS-GCU-trnR-UCU*, *rpoB-psbD*, *rps4-trnF-GAA*, *petA-psbL*, and *ndhF-ndhD*) and one gene region (*ycf1*) from the chloroplast genomes, were selected as potentially suitable gene fragments to study species identification and phylogenetics in *Fritillaria* (Fig. 8). The region of *petA-psbL* possessed the highest variability, followed by *rps4-trnF-GAA*, *ndhF-ndhD*, *ycf1* and *rpoB-psbD*, whereas *trnS-GCU-trnR-UCU* was the lowest. Therefore, the regions with rich variation and suitable length, such as *petA-psbL*, *rps4-trnF-GAA*, *ndhF-ndhD*, *ycf1* and *rpoB-psbD* could be used as a prior choice of species identification for *Fritillaria*. Meanwhile, all of the highly variable regions are judged to be suitable for revealing phylogenetic relationships and genetic structure at the species and population level in *Fritillaria*.

### Phylogenetic analysis

In the present study, the four datasets from the plastid genomes, unanimously clustered *Fritillaria cirrhosa* and its closely related species into three clades (clade I, II and III)

based on BI, MP, and ML analysis (Fig. 7). First, clade I was composed of two parts: *F. sichuanica* and *F. dajinensis*, as well as *F. unibracteata* and its variety. It was surprising to find that *F. sichuanica* was so closely related to *F. dajinensis* because they were obviously different in flower traits (Fig. 1) (Chen & Helen, 2000). One possible reason might be the conflict between molecules and morphology that was also observed in other taxa (Anand et al., 2016). The two species that were located at the northeastern edge of geographical distribution of the whole group, namely *F. taipaiensis* and *F. yuzhongensis* were clustered into clade II. Finally, *F. cirrhosa* was the most closely related to *F. unibracteata* and *F. sinica*. Although *F. sichunica* is thought to be a hybrid among *F. cirrhosa*, *F. unibracteata* and *F. przewalskii* (Luo & Chen, 1996), they seemly did not show close relationships. Our results were highly supportive of those from other studies and agreed with Huang et al. (2018) in the phylogeny of *Fritillaria* at the species level. However, among *F. sichunica* and its relatives, the phylogenetic inference of these results remained ambiguous. Molecular data from the nuclear genome and genetic analysis on population level might be necessary to further explore phylogenetic relationships among these related species.

Moreover, this study preliminarily indicated that the concept of the “complex group of *F. cirrhosa*” suggested by Luo & Chen (1996) might not contain four species but include other species as well (Fig. 8). *Fritillaria cirrhosa* is widely distributed in the alpine and subalpine regions of SW China, and exhibits complicated variations in morphology among the different distributions. Luo & Chen (1996) proposed the concept of a “complex group of *Fritillaria cirrhosa*”, including four species, namely *F. cirrhosa*, *F. sichuanica*, *F. taipaiensis*, and *F. yuzhongensis* based on obscure morphological traits and rough geographical distributions. However, there are no obvious borderlines among species within the complex group, as well as between the group and their closely related species. For example, *F. sichunica* is extremely similar to *F. unibracteata* except for subtle differences in length of the stigma lobes (Chen & Helen, 2000). So, the concept might be unreasonable and should be revised based on more detailed research.

### Super and specific DNA barcodes

Potential DNA barcodes are generally used in species identification and phylogenetic studies of plants, but they could not provide enough informative sites to resolve the relationships among *F. cirrhosa* and its closely related species (Burke et al., 2016; Percy et al., 2014; Zhang et al., 2016). In recent research, the complete chloroplast genome as a super-barcode has been proven to be an effective tool for species discrimination in some complicated groups, and specific DNA barcodes are a trade-off for species identification of those groups based on highly variable regions of the plastid genome (Chen et al., 2018; Ma et al., 2018). In the genus *Fritillaria*, the complete chloroplast genomes were much better at uncovering the phylogeny of *Fritillaria* species (Bi et al., 2018; Li et al., 2018; Park et al., 2017). Similarly, clear phylogenetic relationships among *F. cirrhosa* and its close relatives were indicated based on these tools with extremely high bootstrap values in this study (Fig. 8). Thus, using the whole chloroplast genome as a super-barcode might be suitable for the species identification of *Fritillaria*. Meanwhile, highly variable regions observed in this study could be also used as specific barcodes for identifying species in *Fritillaria*.

## CONCLUSION

The chloroplast genomes of *F. cirrhosa* and its closely related species were sequenced using NGS technology and their genetic information was primarily revealed. The eight genomes exhibited a typical circular quadripartite structure and shared a high similarity in gene order and genomic structure, but still provided rich genetic information for research on the *Fritillaria* species. The position change of the IR/SC junction was not obvious among the eight cp genomes. SSRs, large repeat sequences, and pairwise sequence divergences were determined. Highly variable loci and divergent regions were identified as possible ways to develop genetic markers which could be used for further study on population genetics. Moreover, phylogenetic analyses revealed that the eight *Fritillaria* species were divided into three clades with high support values based on the genome-scale datasets. The results indicated that *F. cirrhosa* was the close relative to *F. unibracteata* and *F. sinica*; thus, it indicated that the concept of a “complex group of *F. cirrhosa*” might be inappropriate and need further revision. Furthermore, the complete chloroplast genomes and highly variable regions were very promising for identifying the species and resolving phylogeny in *F. cirrhosa* which meant that they could be used as super-barcode and specific barcodes of the genus. Overall, the study would be beneficial to facilitate our understanding on phylogeny and evolution in *Fritillaria*.

## ACKNOWLEDGEMENTS

We thank Junbo Yang, Tingshuang Yi, Rong Zhang and Zhirong Zhang in Kunming Institute of Botany (Chinese Academy of Sciences, CAS) for their help in molecular experiment and data analysis of complete chloroplast genome in this study.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This study was co-supported by the National Natural Science Foundation of China (31660081) and Yunnan Provincial Science and Technology Department (Grant No. 2016FB144).

### Grant Disclosures

The following grant information was disclosed by the authors:

National Natural Science Foundation of China: 31660081.

Yunnan Provincial Science and Technology Department: 2016FB144.

### Competing Interests

The authors declare there are no competing interests.

### Author Contributions

- Qi Chen performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables.
- Xiaobo Wu performed the experiments.



- Dequan Zhang conceived and designed the experiments, contributed reagents/materials/analysis tools, authored or reviewed drafts of the paper, approved the final draft.

### DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

All the sequences are accessible at GenBank via accession numbers [MH244906](#), [MH244907](#), [MH244908](#), [MH244909](#), [MH244910](#), [MH244911](#), [MH244912](#) and [MH244913](#). The sequencing data are also available as a [Supplemental File](#).

### Data Availability

The following information was supplied regarding data availability:

Specimens are available at the Herbarium of Medicinal Plants and Crude Drugs of the College of Pharmacy and Chemistry, Dali University under accession numbers ZDQ130053, ZDQ15022, ZDQ130018, ZDQ13030, HCB1, ZDQ14003, ZDQ15023, ZDQ15021, and ZDQ15009.

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.7480#supplemental-information>.

## REFERENCES

- Anand KK, Jena SN, Chaudhary LB, Singh M. 2016.** Conflict between morphological and molecular data: a case study of *Ficus krishnae* (Moraceae). *Phytotaxa* 247:143–147 DOI [10.11646/phytotaxa.247.2.7](#).
- Asaf S, Khan AL, Khan AR, Muhammad W, Kang SM, Khan MA, Seok-Min L, In-Jung L. 2016.** Complete chloroplast genome of *Nicotiana otophora* and its comparison with related species. *Frontiers in Plant Science* 7:843 DOI [10.3389/fpls.2016.00843](#).
- Asaf S, Khan AL, Khan MA, Shahzad R, Lubna , Kang SM, Alharrasi A, Alrawahi A, Lee IJ. 2018.** Complete chloroplast genome sequence and comparative analysis of loblolly pine (*Pinus taeda* L.) with related species. *PLOS ONE* 13(3):e0192966 DOI [10.1371/journal.pone.0192966](#).
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012.** SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19:455–477 DOI [10.1089/cmb.2012.0021](#).
- Bayly MJ, Rigault P, Spokevicius A, Ladiges PY, Ades PK, Anderson C, Bossinger G, Merchant A, Udovicic F, Woodrow IE. 2013.** Chloroplast genome analysis of Australian eucalypts—*Eucalyptus*, *Corymbia*, *Angophora*, *Allosyncarpia* and *Stockwellia* (Myrtaceae). *Molecular Phylogenetics and Evolution* 69:704–716 DOI [10.1016/j.ympev.2013.07.006](#).
- Benson G. 1999.** Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research* 27:573–580 DOI [10.1093/nar/27.2.573](#).

- Bi Y, Zhang MF, Xue J, Dong R, Du YP, Zhang XH. 2018. Chloroplast genomic resources for phylogeny and DNA barcoding: a case study on *Fritillaria*. *Scientific Reports* 8:1184 DOI 10.1038/s41598-018-19591-9.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120 DOI 10.1093/bioinformatics/btu170.
- Burke SV, Wysocki WP, Zuloaga FO, Craine JM, Pires JC, Edger PP, Mayfieldjones D, Clark LG, Kelchner SA, Duvall MR. 2016. Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae). *BMC Plant Biology* 16:140 DOI 10.1186/s12870-016-0823-3.
- Byng JW, Chase MW, Christenhusz MJ, Fay MF, Judd WS, Mabberley DJ, Sennikov AN, Soltis DE, Soltis PS, Stevens PF. 2016. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181:1–20 DOI 10.1111/boj.12385.
- Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J. 2015. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Molecular Biology and Evolution* 32:2015–2035 DOI 10.1093/molbev/msv082.
- Cavalier-Smith T. 2002. Chloroplast evolution: secondary symbiogenesis and multiple losses. *Current Biology* 12:R62–R64 DOI 10.1016/S0960-9822(01)00675-3.
- Çelebi A, Tekşen M, Açık L, Aytaç Z. 2008. Taxonomic relationships in genus *Fritillaria* (Liliaceae): Evidence from RAPD-PCR and SDS-PAGE of seed proteins. *Acta Botanica Hungarica* 50:325–343 DOI 10.1556/ABot.50.2008.3-4.11.
- Chen SC, Helen VM. 2000. *Fritillaria* L. In: *Flora of China*. 24th. Beijing: Science Press.
- Chen XL, Zhou JG, Cui YX, Wang Y, Duan BZ, Yao H. 2018. Identification of *Ligularia* herbs using the complete chloroplast genome as a super-barcode. *Frontiers in Pharmacology* 9:695 DOI 10.3389/fphar.2018.00695.
- Cosner ME, Jansen RK, Palmer JD, Downie SR. 1997. The highly rearranged chloroplast genome of *Trachelium caeruleum* (Campanulaceae): multiple inversions, inverted repeat expansion and contraction, transposition, insertions/deletions, and several repeat families. *Current Genetics* 31:419–429 DOI 10.1007/s002940050225.
- Dang YY, Yang Y, Li Q, Lu JJ, Li XW, Wang YT. 2014. Complete chloroplast genome sequences of poisonous and medicinal plant *Datura stramonium*: organizations and implications for genetic engineering. *PLOS ONE* 9(11):e110656 DOI 10.1371/journal.pone.0110656.
- Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and high-performance computing. *Nature Methods* 9:772 DOI 10.1038/nmeth.2109.
- Day PD, Madeleine B, Laurence H, Fay MF, Leitch AR, Leitch IJ, Kelly LJ. 2014. Evolutionary relationships in the medicinally important genus *Fritillaria* L. (Liliaceae). *Molecular Phylogenetics and Evolution* 80:11–19 DOI 10.1016/j.ympev.2014.07.024.
- Ding YQ, Fang Y, Guo L, Li ZD, He KZ, Zhao Y, Zhao H. 2017. Phylogenetic study of Lemnoideae (duckweeds) through complete chloroplast genomes for eight accessions. *Peer J* 5:e4186 DOI 10.7717/peerj.4186.

- Do HDK, Kim JS, Kim JH. 2014.** A *trnL*\_CAU triplication event in the complete chloroplast genome of *Paris verticillata* M.Bieb. (Melanthiaceae, Liliales). *Genome Biology and Evolution* **6**:1699–1706 DOI [10.1093/gbe/evu138](https://doi.org/10.1093/gbe/evu138).
- Doyle J. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19**:11–15.
- Du YP, Bi Y, Yang FP, Zhang MF, Chen XQ, Xue J, Zhang XH. 2017.** Complete chloroplast genome sequences of *Lilium*: insights into evolutionary dynamics and phylogenetic analyses. *Scientific Reports* **7**:5751 DOI [10.1038/s41598-017-06210-2](https://doi.org/10.1038/s41598-017-06210-2).
- Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004.** VISTA: computational tools for comparative genomics. *Nucleic Acids Research* **32**:W273–W279 DOI [10.1093/nar/gkh458](https://doi.org/10.1093/nar/gkh458).
- Gao L, Yi X, Yang YX, Su YJ, Wang T. 2009.** Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. *BMC Evolutionary Biology* **9**:130 DOI [10.1186/1471-2148-9-130](https://doi.org/10.1186/1471-2148-9-130).
- Gao T, Hui Y, Song JY, Zhu YJ, Chang L, Chen SL. 2010.** Evaluating the feasibility of using candidate DNA barcodes in discriminating species of the large Asteraceae family. *BMC Evolutionary Biology* **10**:324 DOI [10.1186/1471-2148-10-324](https://doi.org/10.1186/1471-2148-10-324).
- Goremykin VV, Hirschernst KI, Wolf S, Hellwig FH. 2003.** Analysis of the *Amborella trichopoda* chloroplast genome sequence suggests that *Amborella* is not a basal angiosperm. *Molecular Biology and Evolution* **20**:1499–1505 DOI [10.1093/molbev/msg159](https://doi.org/10.1093/molbev/msg159).
- Goremykin VV, Hirschernst KI, Wöfl S, Hellwig FH. 2004.** The chloroplast genome of *Nymphaea alba*: whole-genome analyses and the problem of identifying the most basal angiosperm. *Molecular Biology and Evolution* **21**:1445–1454 DOI [10.1093/molbev/msh147](https://doi.org/10.1093/molbev/msh147).
- Hall TA. 1999.** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**:95–98.
- Henriquez CL, Arias T, Pires JC, Croat TB, Schaal BA. 2014.** Phylogenomics of the plant family Araceae. *Molecular Phylogenetics & Evolution* **75**:91 DOI [10.1016/j.ympev.2014.02.017](https://doi.org/10.1016/j.ympev.2014.02.017).
- Huang J, Yang LQ, Yu Y, Liu YM, Xie DF, Li J, He XJ, Zhou SD. 2018.** Molecular phylogenetics and historical biogeography of the tribe Lilieae (Liliaceae): bi-directional dispersal between biodiversity hotspots in Eurasia. *Annals of Botany* **122**(7):1245–1262 DOI [10.1093/aob/mcy138](https://doi.org/10.1093/aob/mcy138).
- Jin JJ, Yu WB, Yang JB, Song Y, Yi TS, Li DZ. 2018.** GetOrganelle: a simple and fast pipeline for de novo assembly of a complete circular chloroplast genome using genome skimming data. *bioRxiv* DOI [10.1101/256479](https://doi.org/10.1101/256479).
- Katoh K, Standley D. 2013.** MAFFT multiple sequence alignment software version improvements in performance and usability. *Molecular Biology & Evolution* **30**:772–780 DOI [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010).
- Kearse M, Moir R, Wilson A, Stoneshavas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C. 2012.** Geneious basic: an integrated and extendable

- desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649 DOI [10.1093/bioinformatics/bts199](https://doi.org/10.1093/bioinformatics/bts199).
- Khourang M, Babaei A, Sefidkon F, Naghavi MR, Asgari D, Potter D. 2014.** Phylogenetic relationship in *Fritillaria* spp. of Iran inferred from ribosomal ITS and chloroplast *trnL-trnF* sequence data. *Biochemical Systematics & Ecology* **57**:451–457 DOI [10.1016/j.bse.2014.10.001](https://doi.org/10.1016/j.bse.2014.10.001).
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. 2005.** Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences of the United States of America* **102**:8369–8374 DOI [10.1073/pnas.0503123102](https://doi.org/10.1073/pnas.0503123102).
- Kumar S, Stecher G, Tamura K. 2016.** MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology & Evolution* **33**:1870–1874 DOI [10.1093/molbev/msw054](https://doi.org/10.1093/molbev/msw054).
- Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001.** REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Research* **29**:4633–4642 DOI [10.1093/nar/29.22.4633](https://doi.org/10.1093/nar/29.22.4633).
- Li PB, Li ZH, Liu HM, Hua JP. 2014a.** Cytoplasmic diversity of the cotton genus as revealed by chloroplast microsatellite markers. *Genetic Resources & Crop Evolution* **61**:107–119 DOI [10.1007/s10722-013-0018-9](https://doi.org/10.1007/s10722-013-0018-9).
- Li Y, Li QS, Li XW, Song JY, Sun C. 2016.** Complete chloroplast genome sequence of *Fritillaria unibracteata* var. *wabuensis* based on SMRT sequencing technology. *Mitochondrial DNA* **27**:3757–3758 DOI [10.3109/19401736.2015.1079892](https://doi.org/10.3109/19401736.2015.1079892).
- Li QS, Li Y, Song JY, Xu HB, Xu J, Zhu YJ, Li XW, Gao HH, Dong LL, Qian J, Sun C, Chen SL. 2014b.** High-accuracy de novo assembly and SNP detection of chloroplast genomes using a SMRT circular consensus sequencing strategy. *New Phytologist* **204**:1041–1049 DOI [10.1111/nph.12966](https://doi.org/10.1111/nph.12966).
- Li Y, Zhang ZR, Yang JB, Lv GH. 2018.** Complete chloroplast genome of seven *Fritillaria* species, variable DNA markers identification and phylogenetic relationships within the genus. *PLOS ONE* **13**(3):e0194613 DOI [10.1371/journal.pone.0194613](https://doi.org/10.1371/journal.pone.0194613).
- Liu ZF, Ci XQ, Li L, Li HW, Conra JG, Li J. 2017.** DNA barcoding evaluation and implications for phylogenetic relationships in Lauraceae from China. *PLOS ONE* **12**(4):e0175788 DOI [10.1371/journal.pone.0175788](https://doi.org/10.1371/journal.pone.0175788).
- Liu ZD, Wang S, Chen SC. 2009.** A taxonomic note of *Fritillaria wabuensis* (Liliaceae). *Acta Botanica Yunnanica* **31**:145 DOI [10.3724/SP.J.1143.2009.09056](https://doi.org/10.3724/SP.J.1143.2009.09056).
- Liu QP, Xue QZ. 2005.** Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. *Journal of Genetics* **84**:55–62 DOI [10.1007/bf02715890](https://doi.org/10.1007/bf02715890).
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013.** OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research* **41**:W575–W581 DOI [10.1093/nar/gkt289](https://doi.org/10.1093/nar/gkt289).
- Lu RS, Li P, Qiu YX. 2016.** The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: comparative genomic and phylogenetic analyses. *Frontiers in Plant Science* **7**:2054 DOI [10.3389/fpls.2016.02054\(2016\)](https://doi.org/10.3389/fpls.2016.02054(2016)).

- Luo YB, Chen SC. 1996.** A revision of *Fritillaria* L. (Liliaceae) in the hengdian mountains and adjacent regions, China(1)—a study of *Fritillaria cirrhosa* D. Don and its related species. *Acta Phytotaxonomica Sinica* **34**:304–312.
- Ma PF, Zhang YX, Zeng CX, Guo ZH, Li DZ. 2014.** Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (poaceae). *Systematic Biology* **63**:933–950 DOI [10.1093/sysbio/syu054](https://doi.org/10.1093/sysbio/syu054).
- Ma SJ, Zhou JG, Li Y, Chen XL, Wu ML, Sun W, Li YH, Song JY, Yao H. 2018.** Complete chloroplast genomes of *Dioscorea opposita* and *D. collettii* and screening specific DNA barcodes (in Chinese). *Scientia Sinica Vitae* **48**:571–582 DOI [10.1360/N052017-00160](https://doi.org/10.1360/N052017-00160).
- Maliga P. 2002.** Engineering the plastid genome of higher plants. *Current Opinion in Plant Biology* **5**:164–172 DOI [10.1016/S1369-5266\(02\)00248-0](https://doi.org/10.1016/S1369-5266(02)00248-0).
- Mardis ER. 2008.** The impact of next-generation sequencing technology on genetics. *Trends in Genetics* **24**:133–141 DOI [10.1016/j.tig.2007.12.007](https://doi.org/10.1016/j.tig.2007.12.007).
- Mishra P, Kumar A, Nagireddy A, Mani DN, Shukla AK, Tiwari R, Sundaresan V. 2015.** DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. *Plant Biotechnology Journal* **14**:8–21 DOI [10.1111/pbi.12419](https://doi.org/10.1111/pbi.12419).
- Moore MJ, Soltis PS, Bell CD, Burleigh JG, Soltis DE. 2010.** Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. *Proceedings of the National Academy of Sciences of the United States of America* **107**:4623–4628 DOI [10.1073/pnas.0907801107](https://doi.org/10.1073/pnas.0907801107).
- Murat C, Riccioni C, Belfiori B, Cichocki N, Labbe J, Morin E, Tisserant E, Paolucci F, Rubini A, Martin F. 2011.** Distribution and localization of microsatellites in the Perigord black truffle genome and identification of new molecular markers. *Fungal Genetic Biology* **48**:592–601 DOI [10.1016/j.fgb.2010.10.007](https://doi.org/10.1016/j.fgb.2010.10.007).
- National Pharmacopoeia Committee. 2015.** *Pharmacopoeia of the People's Republic of China*. 2015 edition. Beijing: Chemical Industry Press.
- Neuhaus HE, Emes MJ. 2000.** Nonphotosynthetic metabolism in plastids. *Annual Review Plant Physiology and Plant Molecular Biology* **51**:111–140 DOI [10.1146/annurev.arplant.51.1.111](https://doi.org/10.1146/annurev.arplant.51.1.111).
- Nie XJ, Lv SZ, Zhang YX, Du XH, Wang L, Biradar SS, Tan XF, Wan FH, Song WN. 2012.** Complete chloroplast genome sequence of a major invasive species, crofton weed (*Ageratina adenophora*). *PLOS ONE* **7**:e36869 DOI [10.1371/journal.pone.0036869](https://doi.org/10.1371/journal.pone.0036869).
- Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z. 1986.** Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* **322**:572–574 DOI [10.1038/322572a0](https://doi.org/10.1038/322572a0).
- Park I, Kim WJ, Yeo SM, Choi G, Kang YM, Piao RZ, Moon BC. 2017.** The complete chloroplast genome sequences of *Fritillaria ussuriensis* Maxim. and *Fritillaria cirrhosa* D. Don, and comparative analysis with other *Fritillaria* species. *Molecules* **22**:982 DOI [10.3390/molecules22060982](https://doi.org/10.3390/molecules22060982).



- Parks M, Cronn R, Liston A. 2009.** Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology* 7:84 DOI [10.1186/1741-7007-7-84](https://doi.org/10.1186/1741-7007-7-84).
- Pauwels M, Vekemans X, Godé C, Frérot H, Castric V, Saumitoullaprade P. 2012.** Nuclear and chloroplast DNA phylogeography reveals vicariance among European populations of the model species for the study of metal tolerance, *Arabidopsis halleri* (Brassicaceae). *New Phytologist* 193:916–928 DOI [10.1111/j.1469-8137.2011.04003.x](https://doi.org/10.1111/j.1469-8137.2011.04003.x).
- Percy DM, Argus GW, Cronk QC, Fazekas AJ, Kesanakurti PR, Burgess KS, Husband BC, Newmaster SG, Barrett SC, Graham SW. 2014.** Understanding the spectacular failure of DNA barcoding in willows (*Salix*): does this result from a trans-specific selective sweep? *Molecular Ecology* 23:4737–4756 DOI [10.1111/mec.12837](https://doi.org/10.1111/mec.12837).
- Powell W, Morgante M, Mcdevitt R, Vendramin GG, Rafalski JA. 1995.** Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proceedings of the National Academy of Sciences of the United States of America* 92:7559–7763 DOI [10.1073/pnas.92.17.7759](https://doi.org/10.1073/pnas.92.17.7759).
- Qian J, Song JY, Gao HH, Zhu YJ, Xu J, Pang XH, Yao H, Sun C, Li XE, Li CY. 2013.** The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*. *PLOS ONE* 8:e57607 DOI [10.1371/journal.pone.0057607](https://doi.org/10.1371/journal.pone.0057607).
- Raubeson AL, Peery R, Chumley TW, Chris D, Matthew FH, Boore JL, Jansen RK. 2007.** Comparative chloroplast genomics: analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* 8:174 DOI [10.1186/1471-2164-8-174](https://doi.org/10.1186/1471-2164-8-174).
- Rix M, Frank E, Webster G, Group AGSF. 2001.** *Fritillaria: a revised classification: together with an updated list of species*. UK: Fritillaria Group of the Alpine Garden Society.
- Ronquist F, Teslenko M, Van DMP, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012.** MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61:539–542 DOI [10.1093/sysbio/sys029](https://doi.org/10.1093/sysbio/sys029).
- Rønsted N, Law S, Thornton H, Fay MF, Chase MW. 2005.** Molecular phylogenetic evidence for the monophyly of *Fritillaria* and *Lilium* (Liliaceae; Liliales) and the infrageneric classification of *Fritillaria*. *Molecular Phylogenetics and Evolution* 35:509–527 DOI [10.1016/j.ympev.2004.12.023](https://doi.org/10.1016/j.ympev.2004.12.023).
- Rose O, Falush D. 1998.** A threshold size for microsatellite expansion. *Molecular Biology and Evolution* 15:613–615 DOI [10.1093/oxfordjournals.molbev.a025964](https://doi.org/10.1093/oxfordjournals.molbev.a025964).
- Rozas J, Ferrer-mata A, Sánchez-delbarrio JC, Guirao-rico S, Librado P, Ramoson-sins SE, Sánchez-gracia A. 2017.** DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution* 34:3299–3302 DOI [10.1093/molbev/msx248](https://doi.org/10.1093/molbev/msx248).
- Shendure J, Ji H. 2008.** Next-generation DNA sequencing. *Nature Biotechnology* 26:1135–1145 DOI [10.1038/nbt1486](https://doi.org/10.1038/nbt1486).
- Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchishinozaki K. 1986.** The complete nucleotide



- sequence of the tobacco chloroplast genome: its gene organization and expression. *Plant Molecular Biology Reporter* 5:2043–2049.
- Simmons MP, Zhang LB, Webb CT, Müller K. 2007.** A penalty of using anonymous dominant markers (AFLPs, ISSRs, and RAPDs) for phylogenetic inference. *Molecular Phylogenetics and Evolution* 42:528–542 DOI 10.1016/j.ympev.2006.08.008.
- Stamatakis A. 2014.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313 DOI 10.1093/bioinformatics/btu033.
- Sugiura M. 1992.** The chloroplast genome. *Plant Molecular Biology* 19:149–157 DOI 10.1007/BF00015612.
- Tangphatsornruang S, Sangsrakru D, Chanprasert J, Uthaipaisanwong P, Yoocha T, Jomchai N, Tragoonrung S. 2009.** The chloroplast genome sequence of Mungbean (*Vigna radiata*) determined by High-throughput pyrosequencing: structural organization and phylogenetic relationships. *DNA Research* 17:11–22 DOI 10.1093/dnares/dsp025(2010).
- Tekşen M, Aytaç Z, Pınar NM. 2010.** Pollen morphology of the genus *Fritillaria* L. (Liliaceae) in Turkey. *Turkish Journal of Botany* 34:397–416 DOI 10.3906/bot-0907-93.
- The Angiosperm Phylogeny Group. 1998.** An ordinal classification for the families of flowering plants. *Annals of the Missouri botanical Garden* 85(4):531–553 DOI 10.2307/2992015.
- The Angiosperm Phylogeny Group. 2003.** An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* 141:399–436 DOI 10.1046/j.1095-8339.2003.t01-1-00158.x.
- The Angiosperm Phylogeny Group. 2009.** An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society* 161:105–121 DOI 10.1111/j.1095-8339.2009.00996.x.
- Vieira LDN, Faoro H, Rogalski M, Fraga HPDF, Cardoso RLA, Souza EMD, Pedrosa FDO, Nodari RO, Guerra MP. 2014.** The complete chloroplast genome sequence of *podocarpus lambertii*: genome structure, evolutionary aspects, gene content and SSR detection. *PLOS ONE* 9(3):e90618 DOI 10.1371/journal.pone.0090618.
- Wang MX, Cui LC, Feng KW, Deng PC, Du XH, Wan FH, Song WN, Nie XJ. 2015.** Comparative analysis of Asteraceae chloroplast genomes: structural organization, RNA editing and evolution. *Plant Molecular Biology Reporter* 33:1526–1538 DOI 10.1007/s11105-015-0853-2.
- Wang RJ, Cheng CL, Chang CC, Wu CL, Su TM, Chaw SM. 2008.** Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. (Research article). *BMC Evolutionary Biology* 8:36 DOI 10.1186/1471-2148-8-36.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015.** Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 31:3350–3352 DOI 10.1093/bioinformatics/btv383.

- Wietsma WA, Deinum D, Teunissen HAS, Berg RGVD. 2015. Phylogenetic relationships within *Fritillaria* section *Petilium* based on AFLP fingerprints. *Plant Systematics and Evolution* 301:1043–1054 DOI 10.1007/s00606-014-1135-4.
- Wu FH, Chan MT, Liao DC, Hsu CT, Lee YW, Daniell H, Duvall MR, Lin CS. 2010. Complete chloroplast genome of *Oncidium* Gower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. *BMC Plant Biology* 10:68 DOI 10.1186/1471-2229-10-68.
- Xue JH, Dong WP, Cheng T, Zhou SL. 2012. Nelumbonaceae: systematic position and species diversification revealed by the complete chloroplast genome. *Journal of Systematics and Evolution* 50:477–487 DOI 10.1111/j.1759-6831.2012.00224.x.
- Yang JB, Li DZ, Li HT. 2014. Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. *Molecular Ecology Resources* 14:1024–1031 DOI 10.1111/1755-0998.12251.
- Yi X, Gao L, Wang B, Su YJ, Wang T. 2013. 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 Biology and Evolution* 5:688–698 DOI 10.1093/gbe/evt042.
- Yuan C, Zhong W, Mou F, Gong Y, Pu D, Ji P, Huang H, Yang Z, Zhang C. 2017. The complete chloroplast genome sequence and phylogenetic analysis of *Chuanminshen* (*Chuanminshenviolaceum* Sheh et Shan). *Physiology & Molecular Biology of Plants* 23:1–7 DOI 10.1007/s12298-016-0395-6.
- Yurina NP, Odintsova MS. 1998. Comparative structural organization of plant chloroplast and mitochondrial genomes. *Genetika* 34:5–22.
- Zhang YJ, Cheng L. 1998. Taxonomic Studies of *Fritillaria* (Liliaceae) from Gansu Province. *Journal of Lanzhou University (Natural Sciences)* 32:84–91 DOI 10.13885/j.issn.0455-2059.1998.02.017.
- Zhang DQ, Mo XC, Xiang JY, Zhou N. 2016. Molecular identification of original plants of *Fritillariae cirrhosae* bulbus, a traditional chinese medicine (TCM) using plant dna barcoding. *African Journal of Traditional Complementary & Alternative Medicines Ajtcam* 13:74–82 DOI 10.21010/ajtcam.v13i6.12.
- Zhang LN, Zhang YJ, Sun JZ. 2001. A taxometric analysis of the variation patterns of characters of *Fritillaria przewalskii* (Liliaceae). *Acta Botanica Boreali-occidentalia Sinica* 21:844–856.