


Article

Comparative Chloroplast Genomics of *Fritillaria* (Liliaceae), Inferences for Phylogenetic Relationships between *Fritillaria* and *Lilium* and Plastome Evolution

Jiao Huang ^{1,2}, Yan Yu ¹, Yan-Mei Liu ¹, Deng-Feng Xie ¹ , Xing-Jin He ^{1,*} and Song-Dong Zhou ^{1,*}

¹ Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life Science, Sichuan University, Chengdu 610065, Sichuan, China; Fritillaria123@sina.com (J.H.); yuyu@scu.edu.cn (Y.Y.); 15208198824@163.com (Y.-M.L.); df_xie2017@163.com (D.-F.X.)

² College of Life Science, Leshan Normal University, Leshan 614000, Sichuan, China

* Correspondence: xjhe@scu.edu.cn (X.-J.H.); zsd@scu.edu.cn (S.-D.Z.); Tel.: +86-28-85415006 (X.-J.H.)

Received: 19 December 2019; Accepted: 20 January 2020; Published: 21 January 2020



Abstract: *Fritillaria* is a genus that has important medicinal and horticultural values. The study involved the most comprehensive chloroplast genome samples referring to Old and New World clades of *Fritillaria* for marker selection and phylogenetic studies. We reported and compared eleven newly sequenced whole-plastome sequences of *Fritillaria* which proved highly similar in overall size (151,652–152,434 bp), genome structure, gene content, and order. Comparing them with other species of Liliales (6 out of 10 families) indicated the same similarity but showed some structural variations due to the contraction or expansion of the inverted repeat (IR) regions. A/T mononucleotides, palindromic, and forward repeats were the most common types. Six hypervariable regions (*rps16-trnQ*, *rbcL-accD*, *accD-psaI*, *psaJ-rpl33*, *petD-rpoA*, and *rpl32-trnL*) were discovered based on 26 *Fritillaria* whole-plastomes to be potential molecular markers. Based on the plastome data that were collected from 26 *Fritillaria* and 21 *Lilium* species, a phylogenomic study was carried out with three *Cardiocrinum* species as outgroups. *Fritillaria* was sister to *Lilium* with a high support value, and the interspecies relationships within subgenus *Fritillaria* were resolved very well. The six hypervariable regions can be used as candidate DNA barcodes of *Fritillaria* and the phylogenomic framework can guide extensive genomic sampling for further phylogenetic analyses.

Keywords: *Fritillaria*; Liliaceae; plastome genome; genome structure; phylogenomics

1. Introduction

Plastid genome of angiosperm have a typical quadripartite structure with a pair of inverted repeat (IR) regions separated by a small single-copy (SSC) region and a large single-copy (LSC) region [1,2]. In contrast to mitochondrial and nuclear genomes, plastid genome is highly conserved and generally contains 110–130 distinct genes, ranging in size from 120–160 kb [3]. Although plastid genomes are reported as a highly conserved sequence and structure in most angiosperms, they have been showed considerable variation in many taxa [2]. These structural variations are always involved in the contraction or extension of the IR region [1], appearance of large inversions or deletions [4,5], and genes gain or loss [2,6]. Some hotspot regions with single nucleotide polymorphisms could be identified, which may be used for species identification in terms of enough information [7]. Due to low rates of nucleotide substitutions, lack of recombination, and uniparental inheritance, many plastid

DNA sequences have been used for inferring plant phylogenies and population genetic analyses [8,9], such as *matK*, *rbcL*, and *trnH-psbA* [10].

Currently, whole chloroplast genomes have been increasingly used for phylogenetic analyses and inferring phylogeographic histories, as the advent and fast development of next-generation sequencing (NGS) technologies. They can provide plenty of variable sites among their entire size for phylogenetic analyses [3]. Thus, whole chloroplast genomes show the potential for resolving evolutionary relationships and have been employed to generate highly resolved phylogenies and genetic diversity, especially in the unresolved relationship of some complex taxa or at low taxonomic levels [2,11–13]. Because different regions of the whole chloroplast genomes differed with their evolutionary rates, partitioning the genome by regions or genes might be preferable for phylogenomic analysis [3]. The concatenated coding genes were most widely used for plastome phylogenomics until now [14,15].

The genus *Fritillaria* L. (Liliaceae) contains 130 to 140 species which are mostly distributed in the temperate regions of the Northern Hemisphere. The centers of speciation and diversity of the genus may be in the Qinghai-Tiber Plateau (QTP), Irano-Turanian region, and the Mediterranean Basin [16,17]. Bulbs of *Fritillaria* have long been used as herbs in Asian countries, especially in traditional Chinese medicine, which are called “Bei-mu” in Chinese. Bulbs of *Fritillaria* are important medicine materials, which have been used in relieving cough, clearing heat, eliminating phlegm, analgesic, and detoxifying, and have even been used in treating cancers [18,19]. As the chemical compounds and pharmacological effects greatly differ among *Fritillaria* species, it is crucial to accurately identify them based on morphological and molecular data [18]. Different barcoding such as plastid DNA (*matK*, *rpl16*, *trnL-trnF*) and nuclear DNA internal transcribed spacer (ITS), single nucleotide polymorphisms (SNPs), and genomic inter-simple sequence repeats (ISSRs) have been performed for identifying *Fritillaria* species [18,20]. According to previous classification analysis, *Fritillaria* comprises eight subgenera, which are further confirmed monophyletic excepting the subgenus *Fritillaria* based on molecular phylogenetic analyses [17,21]. However, our previous study (using three cpDNA markers) [17] detected that the genus *Fritillaria* was not a monophyly and divided into the New World and Old World clades, and the species of *Lilium* were nested within *Fritillaria* with moderate support, although the ITS data revealed a poor monophyly of *Fritillaria*. Day et al. attempted to adopt low-copy nuclear gene regions as barcoding, but none of the target regions experimented could be amplified across all species tested [16]. These uncertainties about phylogenetic relationships between the genus *Fritillaria* and its neighbors are most likely due to few genetic barcodings used in previous studies, even though confounding influences such as hybridization in the evolutionary processes cannot be excluded. So far, a total of fifteen plastid genomes of genus *Fritillaria* have been sequenced and are available on GenBank. These comparative analyses with different *Fritillaria* plastid genomes have provided a basic understanding of the plastid genome characteristics of this genus [22–25].

In this study, plastomes of 11 *Fritillaria* species were sequenced, and a genomic comparative analysis was performed combined with other 15 *Fritillaria* and 21 *Lilium* plastomes available from Genbank. The specific purposes of this study were (1) to compare and analyze the plastome structural of the eleven *Fritillaria* species; (2) to identify highly divergent regions of all 26 *Fritillaria* plastomes; and (3) to resolve the evolutionary relationships between genus *Fritillaria* and *Lilium*. Overall, this study would be helpful to further understand the plastome evolution and phylogeny of *Fritillaria* species.

2. Results

2.1. Plastome Features of *Fritillaria* and Comparison with Other Liliales Taxa

A total of 1.02–2.24 G clean base of the each *Fritillaria* species was collected and the plastomes size ranged from 151,652 bp (*Fritillaria yuzhongensis*) to 152,434 bp (*Fritillaria maximowiczii*) (Table 1). All these plastomes showed the typical quadripartite structure akin to other seed plants, consisting of a pair of IRs (26,232–26,574 bp) separated by the SSC regions (17,310–17,684 bp) and the LSC

(81,424–81,976 bp). The whole GC content of these 11 plastomes was very similar (36.9–37.1%). The eleven *Fritillaria* plastid genomes contained about 131–136 predicted functional genes, including 85–90 protein-coding genes, 38 transfer RNA (tRNA) genes, and 8 ribosomal RNA (rRNA) genes (Figure 1, Table 1 and Table S1). Six to eight protein-coding genes, eight tRNA genes, and all four rRNA genes were duplicated in the IR regions. Among these genes, two CDS (*clpP* and *ycf3*) possessed two introns, while ten CDS genes and six tRNA genes contained a single intron (Table S1). The gene *rps12* was a trans-spliced gene with the 5' end exon located in the LSC region and the 3' exon and intron located in the IR regions. Four pseudogenes (*ψinfA*, *ψycf1*, *ψycf15*, and *ψycf68*) were identified from five *Fritillaria* plastomes (*F. crassicaulis*, *F. dajinensis*, *F. delavayi*, *F. sichuanica*, *F. unibracteata*), and the *ψycf15* and *ψycf68* were duplicated and located in IRA and IRB regions, respectively. The *rps19* gene at the LSC-IR_A border in *F. maximowiczii* was also identified as a pseudogene. In addition, the *F. przewalskii* and *F. yuzhongensis* plastid genome contained two pseudogenes (*ψinfA* and *ψycf1*). The genes *ycf68*, *ycf15*, and *infA* contained several internal stop codons and indicated that they may be identified as pseudogenes [9,22]. Pseudogene for *ycf1* was found at the SSC-IR_B junctions and has lost its protein-coding ability because of incomplete gene duplication. The similar phenomenon was also observed in the *rps19* gene at the LSC-IR_A border.

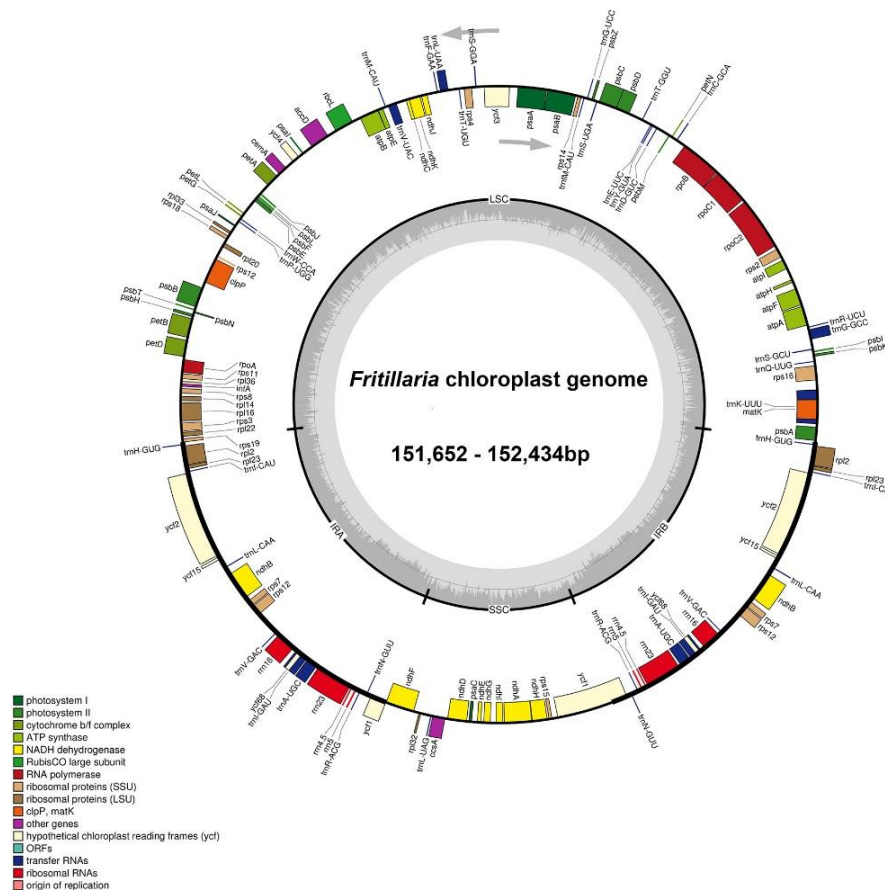


Figure 1. Circular gene map of *Fritillaria* species (*Fritillaria anhuiensis*, *F. crassicaulis*, *F. dajinensis*, *F. davidii*, *F. delavayi*, *F. maximowiczii*, *F. monantha*, *F. przewalskii*, *F. sichuanica*, *F. unibracteata*, and *F. yuzhongensis*) chloroplast genomes. The genes inside and outside of the circle are transcribed in clockwise and counterclockwise directions, respectively. Genes belonging to different functional groups are shown in different colors. The dashed darker gray area in the inner circle indicates genome GC content, while the lighter gray area shows AT content. IR = inverted repeat; SSC = small single copy; LSC = large single copy.

Table 1. Summary of major characteristics of eleven *Fritillaria* plastomes, including aspects of genome size, G-C content, and gene number (per type and location).

Species	Genome Size (bp)	LSC Length (bp)	IR Length (bp)	SSC Length (bp)	G-C (%)	Number of Genes				Number of CDS			
						Total	CDS	rRNAs	tRNAs	LSC	IR _A	SSC	IR _B
<i>F. anhuiensis</i>	152,196	81,950	26,340	17,566	36.9	131[18]	85[6]	8[4]	38[8]	60	7	11	7
<i>F. crassicaulis</i>	151,852	81,615	26,345	17,547	37.0	136[20]	90[8]	8[4]	38[8]	60(1)	9(3)	11	9(2)
<i>F. dajinensis</i>	151,971	81,728	26,351	17,541	37.0	136[20]	90[8]	8[4]	38[8]	60(1)	9(3)	11	9(2)
<i>F. davidii</i>	152,044	81,896	26,232	17,684	37.0	131[18]	85[6]	8[4]	38[8]	60	7	11	7
<i>F. delavayi</i>	151,948	81,683	26,357	17,551	36.9	136[20]	90[8]	8[4]	38[8]	60(1)	9(3)	11	9(2)
<i>F. maximowiczii</i>	152,434	81,976	26,574	17,310	37.1	132[18]	86[6]	8[4]	38[8]	60(1)	7	11	7
<i>F. monantha</i>	152,158	81,895	26,350	17,563	37.0	131[18]	85[6]	8[4]	38[8]	60	7	11	7
<i>F. przewalskii</i>	151,688	81,448	26,352	17,536	37.0	132[18]	86[6]	8[4]	38[8]	60(1)	7(1)	11	7
<i>F. sichuanica</i>	151,967	81,727	26,350	17,540	37.0	136[20]	90[8]	8[4]	38[8]	60(1)	9(3)	11	9(2)
<i>F. unibracteata</i>	151,971	81,730	26,350	17,541	36.9	136[20]	90[8]	8[4]	38[8]	60(1)	9(3)	11	9(2)
<i>F. yuzhongensis</i>	151,652	81,424	26,351	17,526	37.0	132[18]	86[6]	8[4]	38[8]	60(1)	7(1)	11	7

Abbreviations: CDS, protein-coding sequences/genes; LSC, large single-copy region; SSC, small single-copy region; IR, inverted repeat (A or B) regions. Numbers in brackets mean the number of duplicated genes, e.g., 131[18] means there were 131 total genes in the plastome of which 18 were duplicated in the IRs. Numbers in parentheses mean the number of pseudogenes, e.g., 61(1) means there were 61 unique genes observed in LSC of which one was pseudogene.

The IR/SC junction regions with full annotations were compared and had the almost same relative positions among the eleven newly sequencing *Fritillaria* chloroplast genomes (Figure 2). Except for *F. anhuiensis*, all the LSC-IR_B junctions were located within the *rps19* gene, resulting in the IR_B region expanded by a part (28–140 bp) toward the *rps19* gene. In *F. anhuiensis*, the junction was located at the intergenic spacer region (IGS) between *rps19* and *trnH-GUG*. The SSC-IR_A borders in the eleven plastomes were located in the *ycf1* gene and part of this gene was duplicated from 1147–1230 bp in the IR_B. In addition, the *ycf1* gene in four species (*F. anhuiensis*, *F. monantha*, *F. davidii*, and *F. maximowiczii*) extended 24–122 bp into the SSC region, which also had a 27–33 bp overlap with *ndhF*. In five species (*F. crassicaulis*, *F. dajinensis*, *F. delavayi*, *F. sichuanica*, *F. unibracteata*), the SSC-IR_B boundary positions were located at the junction between the pseudogene *ψycf1* and *ndhF*. However, for the other two species (*F. przewalskii* and *F. yuzhongensis*), the distance between *ψycf1* and *ndhF* varied from 58 bp (Figure 2). The LSC-IR_A border was located at the IGS between *trnH-GUG* and *psbA* in all but one species (*F. maximowiczii*). In *F. maximowiczii*, the junction was located in the *ψrps19-psbA* spacer.

Comparison of the 11 newly sequencing *Fritillaria* plastomes with 10 other species of Liliales showed obvious expansion and contraction of the IRs in several families (Figure S1). The LSC-IR_B boundary was located in the *rps19* gene in most Liliales species. However, this boundary shifted into the IGS between the *rpl22* and *rps19* genes in Campynemataceae and it expanded a part toward the *rpl22* gene in Smilacaceae. It further shifted into a part of *rps3* in Melanthiaceae (*Paris*). Long *ψycf1* fragment located at the IR_B varied from 2555 to 992 bp, and the IR has expanded ~1.0–2.0 kb at the SSC-IR_S border in Colchicaceae and Campynemataceae compared to the most Liliales species (see Table S2). Meanwhile, the IR region had a relative contraction in Melanthiaceae (*Veratrum*) (Figure S1). It is also notable that the SSC region of Campynemataceae had a different orientation (inversion) relative to the other Liliales species.

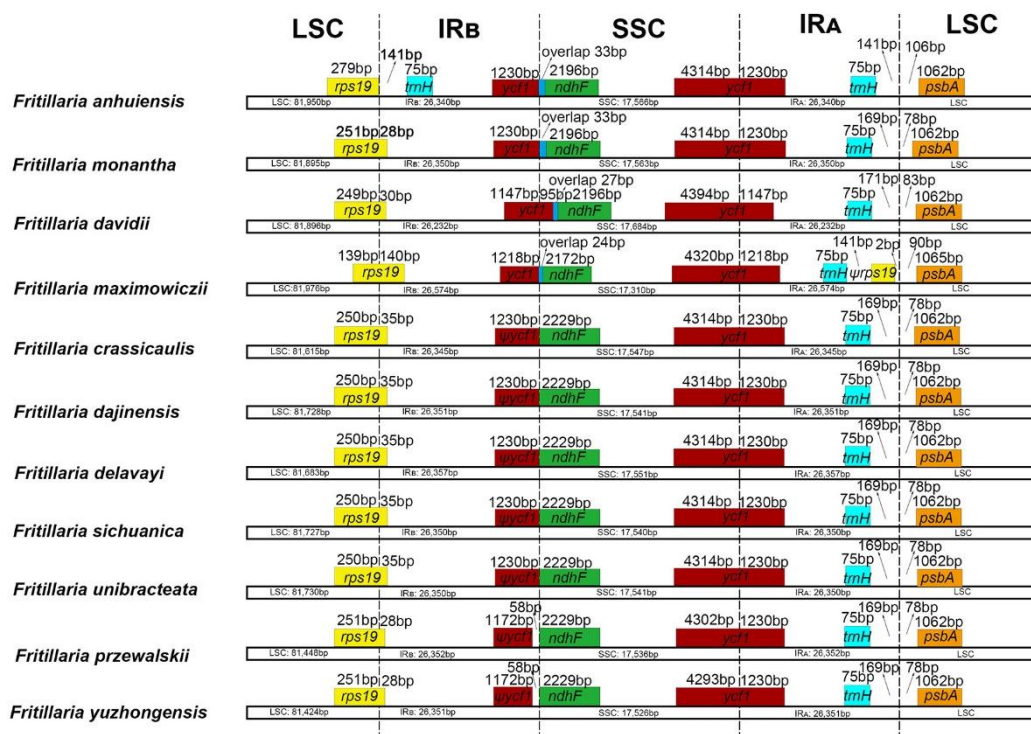


Figure 2. Comparison of LSC, IRs, and SSC junction positions among eleven *Fritillaria* chloroplast genomes.

2.2. Comparative Genomic Analysis and Divergence Hotspot Regions

We investigated the comprehensive sequence divergence of the 11 newly sequencing *Fritillaria* plastid genomes using mVISTA with *F. cirrhosa* as the reference. The aligned sequences revealed high sequence similarities, and several regions of high sequence length polymorphism were revealed (Figure S2). As expected, IR regions exhibited comparatively fewer sequence divergence than LSC and SSC regions. We identified 130 regions (18 introns, 53 intergenic spacers, and 59 coding regions) with more than 200 bp in length from the 26 *Fritillaria* plastid genomes. Of the 59 protein-coding regions (CDS), nucleotide variability (P_i) for each locus ranged from 0.00036 (*ndhB*) to 0.0112 (*rps19*) and the average of 0.00453. Thereby, seven regions (*matK*, *rps16*, *accD*, *ycf4*, *rps19*, *psbH*, *ndhD*) had remarkably high values ($P_i > 0.007$; Figure 3A and Table S3). For the 71 noncoding (intergenic spacer and intron) regions, P_i values ranged from 0.0000 (*ycf4-cemA* and *trnV-rrn16*) to 0.05996 (*rps16-trnQ*) and the average of 0.01163. Six of those regions also exhibited considerable high values ($P_i > 0.02$; i.e., *rps16-trnQ*, *rbcL-accD*, *accD-psaI*, *psaJ-rpl33*, *petD-rpoA*, and *rpl32-trnL*; see Figure 3B and Table S4). The results proved that the IR regions had more sequence conservation than the SC regions, and the average value of P_i in the non-coding regions was more than twice as much as in the coding regions. Specifically, *rbcL-accD*, *accD-psaI*, and *rpl32-trnL* had the highest resolution and similar topology according to the NJ trees. These six divergence hotspot regions in the noncoding regions should provide a useful resource for future phylogenetic and phylogeographic analyses as well as species identification of genus *Fritillaria*.

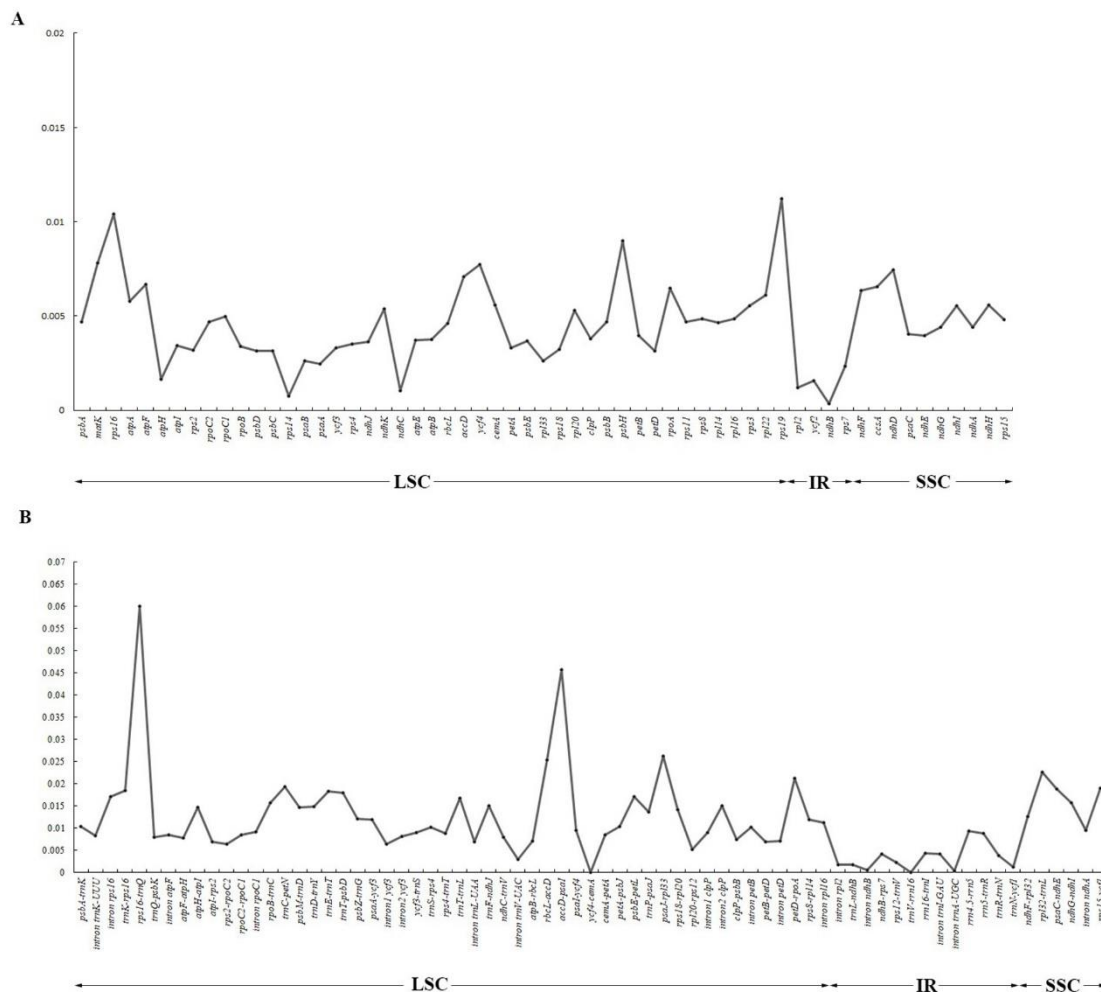


Figure 3. Comparison of nucleotide variability (Pi) values in 26 *Fritillaria* plastomes. (A) Pi values among protein-coding genes (CDS); (B) Pi values among intergenic spacer (IGS) regions.

2.3. SSRs Analysis and Repeat Sequences

With MISA analysis, a total of 848 SSRs were identified across the 11 newly sequencing *Fritillaria* plastomes. Each plastome/species was found to contain the number of SSRs ranging from 72 (*F. delavayi*) to 86 (*F. monantha*), with 20 SSRs shared between all 11 plastomes (Table S5, Figure 4A). Six types of SSRs were detected and the overall length ranged from 10 to 23 bp (Figure 4A, Table S5). Mono- (all A/T), di- (mostly AT/AT and AG/CT), and tetra-nucleotide (AAAG/CTTT, AAAT/ATTT, and AATT/AATT) SSRs were present in each plastome. Tri-nucleotide (AAG/CTT) SSRs were present in nine plastomes except for *F. przewalskii* and *F. yuzhongensis*, and (AAT/ATT) SSRs were present in eight plastomes except for *F. dajinensis*, *F. unibracteata*, and *F. sichuanica*. Penta-nucleotide (AATAT/ATATT) SSRs were in seven plastomes except for *F. monantha*, *F. unibracteata*, *F. delavayi*, and *F. maximowiczii*. By contrast, hexa-nucleotides (AATTAT/AATTAT) were only found in *F. anhuiensis* and *F. monantha* and penta- (AAAAT/ATTTT) and hexa-nucleotides (AGGGAT/ATCCCT) were only in *F. unibracteata*. Furthermore, penta- (AGAGG/CCTCT) nucleotides were only found in *F. davidii*, and di- (AC/GT) and penta-nucleotides (AAAGT/ACTTT) were only in *F. maximowiczii* (Figure 4B). Overall, mono-nucleotide SSRs were the most abundant type (68.28%) appearing among all 848 SSRs (Figure 4A). In the total 848 SSRs, most loci were situated in the intergenic spacer (IGS) regions, but some were found in *rpoC2*, *rpoB*, *accD*, *cemA*, *rpl22*, *ycf2*, *ndhD*, *ndhG*, *ndhH*, and *ycf1* coding genes (Table S5). Moreover, most of the SSR loci were situated in the LSC region (75.8%), followed by the SSC region (19.2%) and a minimum in the IR regions (5.0%) (Table S5). In addition, we found 41 polymorphic SSRs between the eleven

species of *Fritillaria* (excluding mononucleotide SSRs) (Table S6), which could be useful for further population genetic studies.

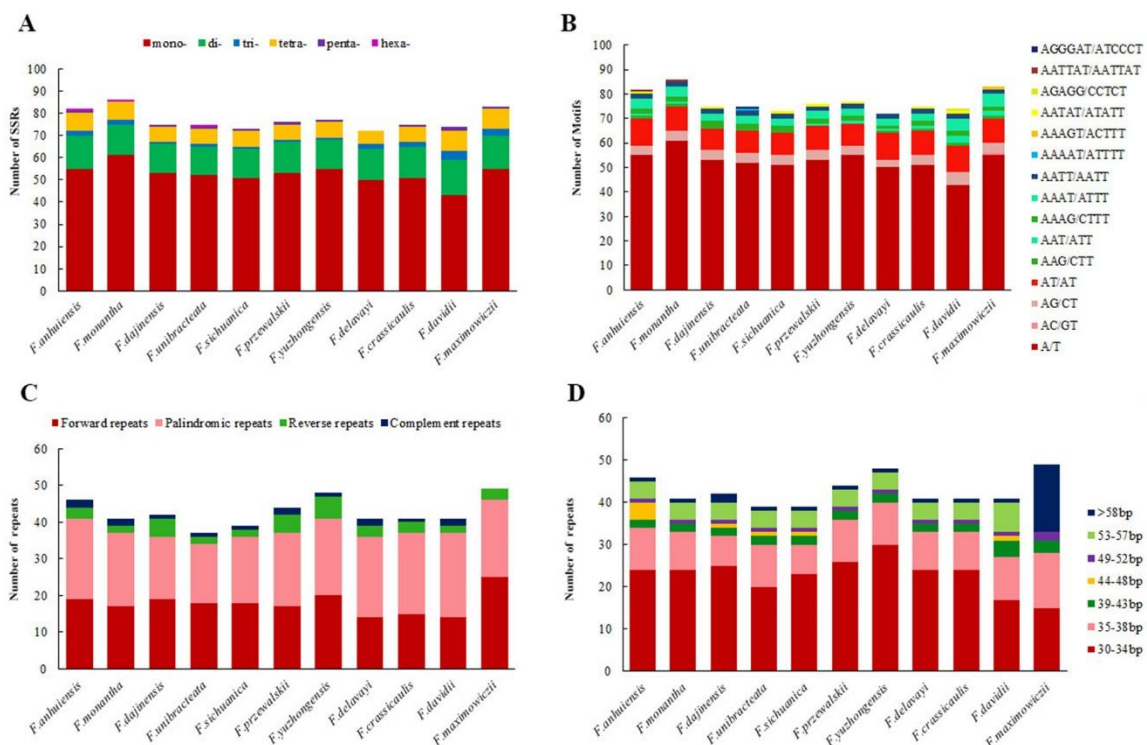


Figure 4. Analysis of different repeats in eleven chloroplast genomes of *Fritillaria*. (A) Number of different simple sequence repeat (SSR) types detected; (B) total numbers of different SSR motifs; (C) the number of four repeats types; (D) frequency of repeats by length.

We identified a total of 471 repeats including 196 forward, 224 palindromic, 36 reverse, and 15 complement in the eleven newly sequencing *Fritillaria* plastomes using REPUTER (Table S7). *F. maximowiczii* and *F. yuzhongensis* possessed the greatest total number of repeats (49 and 48), while *F. unibracteata* and *F. sichuanica* contained the fewest (39) (Figure 4C). Each *Fritillaria* plastid genome contained the number of large repeat sequences ranging from 30 to 57 and repeats with highest proportion ranging in size between 30 and 39 bp (Figure 4D). Repeats situated in homologous regions with identical lengths were recognized as shared repeats [13,26]. Under this criterion, there were 18 repeats shared by all eleven *Fritillaria* species, 24 repeats shared by ten of the *Fritillaria* species (except for *F. maximowiczii*), as well as 27 repeats shared by nine of the *Fritillaria* species (except for *F. maximowiczii* and *F. davidii*). Additionally, *F. maximowiczii* owned the most unique repeats (23), whereas four species had no unique repeats, namely *F. monantha*, *F. sichuanica*, *F. przewalskii*, and *F. crassicaulis* (Table S7).

2.4. Phylogenetic Analyses

Both Bayesian inference (BI) and maximum likelihood (ML) methods based on the 64 common CDS shared among the 52 plastomes (28 *Fritillaria*, 21 *Lilium*, and 3 *Cardiocrinum* species) generated almost identical topologies (Figure 5). *Lilium* was strongly supported as a monophyletic group (ML bootstrap support, BS = 100%, posterior probability, PP = 1). However, the monophyly of *Fritillaria* was moderately supported (BS = 81%, PP = 0.71). Within the Old World clade of *Fritillaria*, relationships between five subgenera (except subgenus *Korolkowia* and *Japonica*) were fully supported with generally high support values. The phylogenetic trees in this study also indicated subgenus *Fritillaria* was not monophyletic which divided into two clades, and *Fritillaria* clade B formed three monophyletic

subclades (*F. pallidiflora*, *F. thunbergii*, and *F. cirrhosa* subclades). This was in accordance with our previous phylogenetic analysis based on three plastid markers [17]. Within *Lilium*, two strongly supported lineages formed the backbone of the phylogeny. However, the monophyly of four sections (section *Sinomartagon*, *Martagon*, *Leucolirion*, and *Pseudolirium*) were not supported based on our limited samples. In addition, the topologies from the single-copy genes and the whole complete chloroplast genome sequences were similar to that from the CDS sequences, and most lineages possess high bootstrap values. Rather unexpectedly, however, the monophyly of *Fritillaria* was weakly supported (BS = 67%, PP = 0.52) based on 64 single-copy genes (Figure S3) and not recovered based on whole genome sequences (Figure S4).

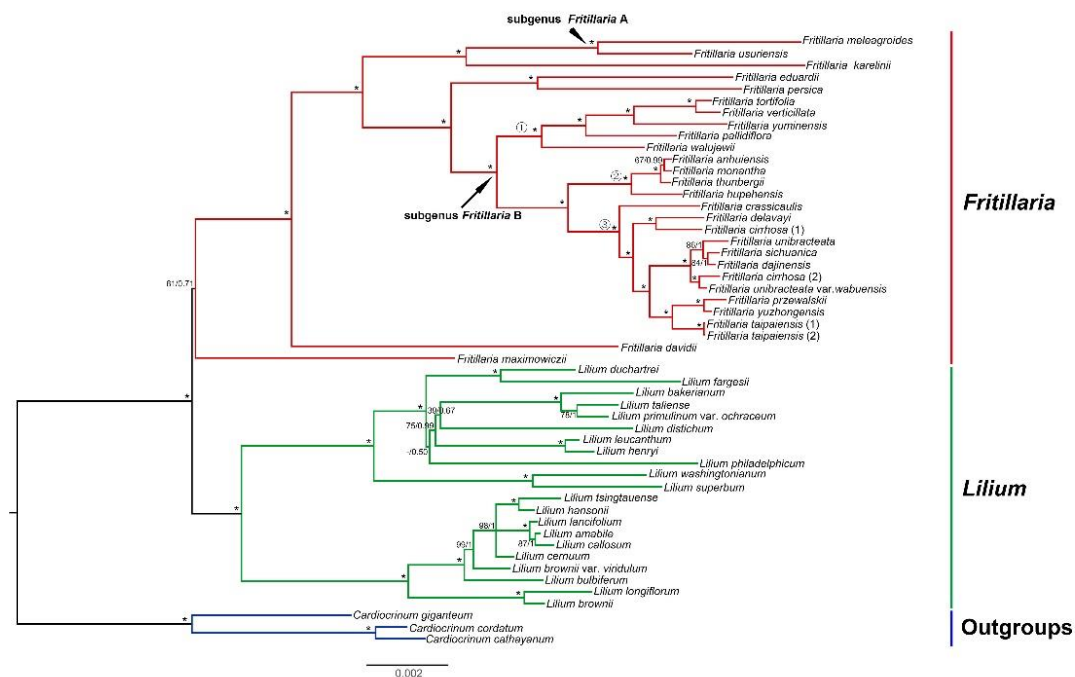


Figure 5. Phylogenetic relationship of the 52 species inferred from maximum likelihood (ML) and Bayesian inference (BI) analyses based on 64 shared protein-coding genes. Support values marker above the branches follow the order BS (bootstrap support)/pp (posterior probability), and “*” indicates 100% support values in both ML and BI trees. Accessions from different genera are written using different colors. ①, ②, and ③ indicate *F. pallidiflora*, *F. thunbergii*, and *F. cirrhosa* subclades, respectively.

3. Discussion

3.1. Comparison of *Fritillaria* Plastomes and Phylogenetically Informative Markers

Our results revealed that the whole plastome sequences of 11 newly sequenced *Fritillaria* species were conserved not only in genome size, but also in gene content, structure, and order (Figure 1), which was consistent with previously published *Fritillaria* plastomes studies [23–25]. Nevertheless, some CDS genes exhibited high variations, mainly in some pseudogenes, for example $\psi ycf15$, $\psi ycf68$, and $\psi infA$, which were also reported in other angiosperm plastid genomes [7,9,26,27]. It is worth noting that there are only *F. crassicaulis*, *F. dajinensis*, *F. delavayi*, *F. sichuanica*, and *F. unibracteata* having $\psi ycf15$ and $\psi ycf68$ in our results, indicating a few common aspects in their evolutionary processes and functions. The absence or presence of $\psi ycf15$ and $\psi ycf68$ in other *Fritillaria* species has also been reported by Li et al. [25]. The same phenomenon has also been observed in *Ipomoea* [28]. Additionally, the functions of pseudogene $ycf68$ and $ycf15$ are equivocal in all kinds of land plants [28], for example, in Podophylloideae, the $ycf15$ gene was detected as non-functional because of the presence of an interrupted sequence where large amounts of stop codons were located [2]. The $infA$ gene, which involved in codes of translation initiation factor 1 and an assembly of the translation initiation

complex [29,30], was detected as a pseudogene in seven *Fritillaria* plastomes while completely lost in other four *Fritillaria* plastomes (Table S1). Similar phenomena have been also observed in other angiosperm plastid genomes, for instance, those of *Primula* species, only the plastid genome of *Primula poissonii* contains the $\psi infA$ [31]. Our study indicates that the pseudogenization of *infA* may be a synapomorphy for the *F. cirrhosa* subclade.

Plastomes sequenced in our study are also largely similar in overall gene content and structure when compared with some previously published plastomes in Liliales [9,26,32–37]. However, the LSC-IR borders of Smilacaceae, Melanthiaceae (*Paris*), and Campynemataceae vary from those of *Fritillaria* (and other Liliales) by showing expansions (Figure S1), which may assist to stabilize the structure of the entire plastome as well as the prevention of gene gain and gene loss phenomenon [2,38–40]. By contrast, the SSC-IR boundaries of Melanthiaceae (*Veratrum*) feature contractions when compared to *Fritillaria* and other Liliales (Figure S1). Our results included 6 out of 10 families in the order Liliales (Table S2) and are in accordance with those of Do et al. [33], who used four families in Liliales and Do et al. [41], and five families in Liliales. Wang et al. [42] suggested that *trnH-rps19* clusters were located in the IR/LSC junctions within Liliales taxa, but different patterns of junction were found in Liliales from the present study (Figure S1). Furthermore, more studies covering all of 10 families should be performed to investigate the whole evolutionary trends of plastomes in Liliales. In general, the contraction and expansion of IR regions are relatively common evolutionary events in plants and have been used as evolutionary loci for phylogenetic relationships [1,31,43–45]. Li et al. reported that IR/LSC junctions expanded into *rps19* and suggested the events seem to be an ancestral symplesiomorphy of Liliaceae [26]. Here, we also found the feature but there was no obvious phylogenetic implication and further evidence was needed using sufficient genera of Liliaceae.

The cpDNA sequences from a variety of intergenic spacers (*trnL-trnF*) and genes (*matK*, *rbcL*, *rpl16*, and *atpB*) have often been used to infer the phylogeny of genus *Fritillaria* [16,17,46–48]. From our results, these frequently used plastid barcodes, such as *atpB*, *rpl16*, and *rbcL*, are among the relatively least informative genes (Figure 3). Therefore, the previous phylogenetic analysis based on these barcodes generated low resolution, especially for deep phylogenetic relationships with short internodes and fast rates, such as subgenus *Fritillaria*. By contrast, our sliding window analyses find six intergenic spacer regions with the highest P_i (>0.02) values from 26 *Fritillaria* species (Figure 3). We suggest these hotspot regions are valuable loci to understand the phylogenetic relationships for *Fritillaria* lineages (subgeneric levels) which have experienced rapid radiation.

3.2. Phylogenetic Relationships and Implications

Phylogenetic analyses based on 64 protein-coding genes generated a well resolved tree (Figure 5). Genus *Lilium* was sister to *Fritillaria*, and the monophyly of *Fritillaria* was recovered although moderate support values were detected. These results were different from a previous study, which showed that genus *Fritillaria* was paraphyletic and *Lilium* was nested within *Fritillaria* with moderate support values [17]. This difference may be attributed to our smaller sampling size, especially only one sample (*Fritillaria maximowiczii*) in the New World Clade (subgenus *Liliorhiza*). However, the phylogenetic results from single-copy genes and the whole chloroplast genomes showed that the monophyly of *Fritillaria* was weakly supported (BS = 67%, PP = 0.52) (Figure S3) and not recovered (Figure S4). Therefore, the relationships between the New World and the Old World clades of *Fritillaria* and *Lilium* still remain unsolved. In comparison with previous studies [24,25], our results detected further phylogenetic relationships between *Fritillaria* and *Lilium* with more plastomes data, and the subgenus *Fritillaria* was confirmed polyphyletic, which has a high resolution than previous studies (Figure 5) [16,17]. Our results also proved that the plastome data are of great advantages in phylogenetic analyses, although more extensive plastid genomic sampling was needed for further resolution of *Fritillaria* and *Lilium*.

4. Materials and Methods

4.1. Taxon Sampling, DNA Extraction, and Sequencing

We generated new plastome data for 11 species of *Fritillaria*. Fresh leaves of each species were collected in the field and dried with silica gel. All samples were obtained in China and field sampling was permitted by Natural Reserves in the Heilongjiang, Sichuan, Yunnan, Gansu, Anhui, and Zhejiang provinces. The vouchers were identified by Dr. Jiao Huang and were deposited in the Herbarium of Sichuan University (SZ); the voucher information is presented in Table S8. Total genomic DNAs were extracted from leaf material following the protocols using a plant genomic DNA kit (Tiangen Biotech, Beijing, China). The isolated genomic DNA was used to generate average 350 bp paired-end (PE) library according to the Illumina Hiseq platform (Illumina, San Diego, CA, USA), and sequenced by an Illumina genome analyser (Hiseq PE150).

4.2. Chloroplast Genome Assembly, Annotation, and Structural Analyses

FastQC v0.11.7 was used to assess the quality of sequenced raw reads [49]. Then, we sieved the chloroplast genome related reads by mapping all raw reads to the chloroplast genome sequences downloaded from NCBI of the genera *Fritillaria* and *Lilium* (38 species). Contigs, assembled from all related reads using SOAPdenovo2 [50], were sorted and joined into a single-draft sequence with *Fritillaria cirrhosa* (KF769143), *Fritillaria thunbergii* (NC034368), *Fritillaria taipaiense* (KC543997), and *Lilium leucanhum* (NC035590) as reference species in the software Geneious v11.0.4 [51]. Some gaps in the assembled plastomes were corrected using Sanger sequencing. The primers were designed by Primer 5.0 (Premier, Palo Alto, CA, USA). Primer synthesis and the sequencing of the polymerase chain reaction products were performed by Sangon Biotech (Shanghai, China). The amplifications and primers are shown in Table S9. Annotations of the complete plastomes were conducted using Geneious v11.0.4. The draft annotation was checked and edited manually following the reference genome to accurately confirm the start/stop codons and the exon/intron borders of genes. Furthermore, the schematic diagram of the circular plastome map was generated utilizing OGDRAW [52]. Eleven newly sequenced plastomes have been deposited in GenBank (accession numbers: MK258138-MK258148). The gene order and structure of 11 *Fritillaria* plastomes were compared using Geneious v11.0.4. We used the plastome of *Fritillaria crassicaulis* as a representation to further compare with plastomes of other 10 Liliales species (see Table S2).

4.3. Genome Comparative Analysis and Identification of Hypervariable Regions

Chloroplast genome comparisons across the eleven *Fritillaria* species was performed by the mVISTA program [53], using *F. cirrhosa* (KF769143) as the reference. For identifying hypervariable regions in *Fritillaria* and facilitating its utilization for future genetic population and species identification studies, multiple sequence alignments of 26 *Fritillaria* species (15 species were downloaded from NCBI; Table S10) were performed in MAFFT v.7 [54], and the software MEGA v.6 was used to adjust manually where necessary [55]. A sliding window analysis in DNASP v5.10 [56] was conducted for the sequence alignment to evaluate nucleotide diversity (π) including all protein-coding and noncoding (intron and intergenic spacer) regions. The extraction was made under the following two criteria: (a) mutation site >0 ; and (b) an aligned length >200 bp. Neighbor-Joining (NJ) trees were constructed by hypervariable markers in the noncoding regions using MEGA v.6 based on a k2P distance model.

4.4. Characterization of SSRs and Repeat Sequences

MISA perl script [57] was used to detect simple sequence repeat (SSR) loci of the 11 *Fritillaria* plastid genomes. The minimum numbers (thresholds) of repeats were 10, 5, 4, 3, 3, and 3 for mononucleotide, dinucleotides, trinucleotides, tetranucleotides, pentanucleotides, and hexanucleotides, respectively. The size and position of repeat sequences are assessed by REPuter [58], including inverted (palindromic), direct (forward), reverse, and complement repeats. The following constraint sets for repeat identification

were used: (1) 90% greater sequence identity; (2) hamming distance equal to 3; and (3) a minimum repeat size of 30 bp.

4.5. Phylogenetic Analyses

Phylogenetic analyses were performed for the 28 *Fritillaria* species (11 species sequenced here) and 21 *Lilium* species, using three *Cardiocrinum* species as outgroups based on previous studies [17] (Table S10). Two accessions were included for *F. cirrhosa* and *F. taipaiense*. The analysis was performed based on an alignment of 64 protein-coding genes from the plastid genomes of the 52 species. The sequences of the 64 common CDS were extracted and aligned using MAFFT v.7 [54]. MEGA v.6 was used to adjust manually where necessary [55]. Topologies were constructed using both Bayesian inference (BI) and maximum likelihood (ML) methods. We used the Akaike Information Criterion (AIC) in JModeltest v.2.1.7 to determine the best-fitting models of nucleotide substitutions [59]. The GTR+I+G models were most suitable for both datasets. ML analyses were performed using RAxML-HPC2 [60] on XSEDE of the CIPRES Science Gateway [61], and statistical node supports were estimated via a bootstrap analysis. Bayesian inference (BI) analyses were conducted in MrBayes v3.2 [62]. Two independent Markov Chain Monte Carlo chains were run simultaneously for five million generations. The trees were sampled every 100 generations with the first 25% of calculated trees discarded as burn-in. The consensus tree was constructed from the remaining trees to estimate posterior probabilities (PPs). Additionally, the complete chloroplast genome sequences and 52 extracted single-copy genes were also collected to perform the phylogenetic analyses, respectively.

5. Conclusions

Our study is the first to report the eleven whole plastomes of *Fritillaria* of which the organization is described. The comparison of these plastid genomes among each other as well as with other species of Liliales showed high similarities in entire structure and content. The synteny of gene order for eleven *Fritillaria* plastomes is also rather conserved. However, these plastid genomes show some structural variations at the junctions of their four regions due to the expansion or contraction of the IRs, especially obvious in Liliales. The six noncoding-cpDNA regions are identified as the fastest evolving loci from the comparison of DNA sequence divergence among 26 *Fritillaria* plastomes. Therefore, these highly variable loci, and 471 repeat sequences from eleven *Fritillaria* plastoms can be used for future phylogenetic and phylogeographic analysis. Furthermore, the plastid genomes SSRs with rich diversity identified herein could be useful for further population genetic studies of *Fritillaria*.

This study performed a new level of phylogenomic sampling, and found that *Fritillaria* was monophyletic and sister to *Lilium*. Future phylogenomic studies require more extensive taxonomic sampling, especially subgenus *Liliorhiza* of genus *Fritillaria* to discern the relationships among the Old and New World clades of *Fritillaria* and *Lilium*. Moreover, some efforts should focus on the evolutionary and adaptive researches of *Fritillaria* species in the future as the recent studies suggested [63,64]. In conclusion, our results will be valuable to understand the evolutionary relationship between *Fritillaria* and *Lilium*, especially plastid gene evolution of *Fritillaria*.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2223-7747/9/2/133/s1>, Table S1: Gene composition in the eleven *Fritillaria* plastid genomes. Table S2: Summary of accession numbers and genome sizes in Liliales. Table S3: Comparison of nucleotide diversity (Pi) values among protein-coding genes (CDS) in 26 *Fritillaria* plastomes. Table S4: Comparison of nucleotide diversity (Pi) values among intergenic spacer (IGS) and intron of 26 *Fritillaria* plastomes. Table S5: Statistics of simple sequence repeats in each species. Table S6: Simple sequence repeat (SSR) polymorphism in eleven *Fritillaria* chloroplast genomes. Table S7: Analyses of repeat sequences in eleven *Fritillaria* chloroplast genomes. Table S8: Voucher information for species newly sequenced for this study, including collector's name and collection number, geographical locality and herbarium acronym where the specimen is deposited. Table S9: Primers used for gap closure in this study. Table S10: Accession numbers of chloroplast genome sequences included in phylogenetic analyses. Figure S1: Comparison of LSC, IRs, SSC junction positions and gene loss among Liliales plastomes. Figure S2: Sequence identity plots between eleven *Fritillaria* chloroplast genomes, with *F. cirrhosa* as a reference. Annotated genes are displayed along the top. The vertical scale represents the percent identity between 50 and 100%. Genome regions are color coded as protein coding, rRNA coding, tRNA coding, or conserved non-coding sequences (CNS). Figure S3: Phylogenetic

relationship of the 52 species inferred from ML and BI analyses based on 64 shared single-copy genes. Support values marker above the branches follow the order BS (bootstrap support)/pp (posterior probability), and “*” indicates 100% support values in both ML and BI trees. Accessions from different genera are written using different colors. ①, ②, and ③ indicate *F. pallidiflora*, *F. thunbergii*, and *F. cirrhosa* subclades, respectively. Figure S4: Phylogenetic relationship of the 52 species inferred from ML and BI analyses based on whole genome sequences. Support values marker above the branches follow the order BS (bootstrap support)/pp (posterior probability), and “*” indicates 100% support values in both ML and BI trees. Accessions from different genera are written using different colors. ①, ②, and ③ indicate *F. pallidiflora*, *F. thunbergii*, and *F. cirrhosa* subclades, respectively.

Author Contributions: Conceptualization, J.H., X.-J.H., and S.-D.Z.; data curation, Y.Y.; formal analysis, J.H.; investigation, J.H. and Y.-M.L.; methodology, Y.Y. and D.-F.X.; supervision, X.-J.H.; writing—original draft, J.H.; writing—review & editing, J.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Grant numbers 31570198, 31872647), the National Infrastructure of Natural Resources for Science and Technology (Grant No. 2005DKA21403-JK), and the Science and Technology Basic Work (Grant No. 2013FY112100).

Acknowledgments: We are grateful to Jin-Bo Tan, Qun-Ying Xiao, Jun-Pei Chen, Chuan Xie, Yi-Qi Deng, Juan Li, Fang-Yu Jin, Hai-Ying Liu, Hao Li, and Xin Yang for the help in the collection of the field materials and data treating.

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

References

- Zhu, A.; Guo, W.; Gupta, S.; Fan, W.; 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. [[CrossRef](#)] [[PubMed](#)]
- Ye, W.Q.; Yap, Z.Y.; Li, P.; Peter, H.C.; Qiu, Y.X. Plastome organization, genome-based phylogeny and evolution of plastid genes in Podophylloideae (Berberidaceae). *Mol. Phylogenet. Evol.* **2018**, *127*, 978–987. [[CrossRef](#)] [[PubMed](#)]
- Fu, C.N.; Li, H.T.; Milne, R.; Zhang, T.; Ma, P.F.; Yang, J.; Li, D.Z.; Gao, L.M. Comparative analyses of plastid genomes from fourteen Cornales species: Inferences for phylogenetic relationships and genome evolution. *BMC Genom.* **2017**, *18*, 956. [[CrossRef](#)] [[PubMed](#)]
- Downie, S.R.; Jansen, R.K. A Comparative Analysis of Whole Plastid Genomes from the Apiales: Expansion and Contraction of the Inverted Repeat, Mitochondrial to Plastid Transfer of DNA, and Identification of Highly Divergent Noncoding Regions. *Syst. Bot.* **2015**, *40*, 336–351. [[CrossRef](#)]
- Hsu, C.Y.; Wu, C.S.; Chaw, S.M. Birth of Four Chimeric Plastid Gene Clusters in Japanese Umbrella Pine. *Genome Biol. Evol.* **2016**, *8*, 1776–1784. [[CrossRef](#)]
- Sun, Y.; Moore, M.J.; Lin, N.; Adelalu, K.F.; Meng, A.; Jian, S.; Yang, L.; Li, J.; Wang, H. Complete plastome sequencing of both living species of Circaeasteraceae (Ranunculales) reveals unusual rearrangements and the loss of the *ndh* gene family. *BMC Genom.* **2017**, *18*, 592. [[CrossRef](#)]
- Zhou, T.; Wang, J.; Jia, Y.; Li, W.; Xu, F.; Wang, X. Comparative Chloroplast Genome Analyses of Species in *Gentiana* section *Cruciata* (Gentianaceae) and the Development of Authentication Markers. *Int. J. Mol. Sci.* **2018**, *19*, 1962. [[CrossRef](#)]
- Shaw, J.; Lickey, E.B.; Beck, J.T.; Farmer, S.B.; Wusheng, L. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* **2005**, *92*, 142–166. [[CrossRef](#)]
- Lu, R.S.; Li, P.; Qiu, Y.X. The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: Comparative genomic and phylogenetic analyses. *Front. Plant Sci.* **2017**, *7*, 2054–2066. [[CrossRef](#)]
- Group, C.P.W. A DNA barcode for land plants. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12794–12797. [[CrossRef](#)]
- Hu, H.; Hu, Q.J.; Al-Shehbaz, I.A.; Luo, X.; Zeng, T.T.; Guo, X.Y.; Liu, J.Q. Species Delimitation and interspecific relationships of the genus *Orychophragmus* (Brassicaceae) inferred from whole chloroplast genomes. *Front. Plant Sci.* **2016**, *7*, 1826–1836. [[CrossRef](#)] [[PubMed](#)]
- Zhang, S.D.; Jin, J.J.; Chen, S.Y.; Chase, M.W.; Soltis, D.E.; Li, H.T.; Yang, J.B.; Li, D.Z.; Yi, T.S. Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytol.* **2017**, *214*, 1355–1368. [[CrossRef](#)] [[PubMed](#)]
- Xu, W.Q.; Losh, J.; Chen, C.; Pan, L.; Wang, R.H.; Zhao, Y.P.; Qiu, Y.X.; Fu, C.X. Comparative genomics of figworts (*Scrophularia*, Scrophulariaceae), with implications for the evolution of *Scrophularia* and Lamiales. *J. Syst. Evol.* **2018**, *57*, 55–65. [[CrossRef](#)]

14. Xi, Z.X.; Ruhfel, B.R.; Schaefer, H.; Amorim, A.M.; Sugumara, M. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17519–17524. [[CrossRef](#)] [[PubMed](#)]
15. Zhang, N.; Wen, J.; Zimmer, E.A. Another look at the phylogenetic position of the grape order Vitales: Chloroplast phylogenomics with an expanded sampling of key lineages. *Mol. Phylogenet. Evol.* **2016**, *101*, 216–223. [[CrossRef](#)]
16. Day, P.D.; Berger, M.; Hill, L.; Fay, M.F.; Leitch, A.R.; Leitch, I.J.; Kelly, L.J. Evolutionary relationships in the medicinally important genus *Fritillaria* L. (Liliaceae). *Mol. Phylogenet. Evol.* **2014**, *80*, 11–19. [[CrossRef](#)]
17. Huang, J.; Yang, L.Q.; Yu, Y.; Liu, Y.M.; Xie, D.F. Molecular phylogenetics and historical biogeography of the tribe Liliaeae (Liliaceae): Bi-directional dispersal between biodiversity hotspots in Eurasia. *Ann. Bot.* **2018**, *122*, 1245–1262. [[CrossRef](#)]
18. Hao, D.C.; Xiao-Jie, G.U.; Xiao, P.G. Phytochemical and biological research of *Fritillaria* medicinal resources. *Chin. J. Nat. Med.* **2013**, *11*, 330–344. [[CrossRef](#)]
19. Wang, D.; Li, Z.; Zhang, L.; Atanasov, A.G.; Wang, S. Characterization of the Isosteroidal Alkaloid Chuanbeinone from Bulbus of *Fritillaria pallidiflora* as Novel Antitumor Agent In Vitro and In Vivo. *Planta Med.* **2016**, *82*, 195–204. [[CrossRef](#)]
20. Altinordu, F.; Peruzzi, L.; Yu, Y.; He, X.J. A tool for the analysis of chromosomes: KaryoType. *Taxon* **2016**, *65*, 586–592. [[CrossRef](#)]
21. Rix, E.M. *Fritillaria: A Revised Classification Together with an Updated List of Species*; The *Fritillaria* Group of the Alpine Garden Society Press: Pershore, UK, 2001.
22. Li, Q.S.; Li, Y.; Song, J.Y.; Xu, H.B.; Xu, J. High-accuracy de novo assembly and SNP detection of chloroplast genomes using a SMRT circular consensus sequencing strategy. *New Phytol.* **2014**, *204*, 1041–1049. [[CrossRef](#)] [[PubMed](#)]
23. Park, I.; Kim, W.J.; Yeo, S.M.; Choi, G.; Kang, Y.M. The complete chloroplast genome sequences of *Fritillaria ussuriensis* Maxim. and *Fritillaria cirrhosa* D. Don, and comparative analysis with other *Fritillaria* species. *Molecules* **2017**, *22*, 982. [[CrossRef](#)] [[PubMed](#)]
24. Bi, Y.; Zhang, M.F.; Xue, J.; Dong, R.; Du, Y.P.; Zhang, X.H. Chloroplast genomic resources for phylogeny and DNA barcoding: A case study on *Fritillaria*. *Sci. Rep.* **2018**, *8*, 1184–1196. [[CrossRef](#)] [[PubMed](#)]
25. Li, Y.; Zhang, Z.; Yang, J.; Lv, G.H. Complete chloroplast genome of seven *Fritillaria* species, variable DNA markers identification and phylogenetic relationships within the genus. *PLoS ONE* **2018**, *13*, e0194613. [[CrossRef](#)]
26. Li, P.; Lu, R.S.; Xu, W.Q.; Ohi-Toma, T.; Cai, M.Q.; Qiu, Y.X.; Cameron, K.M.; Fu, C.X. Comparative Genomics and Phylogenomics of East Asian Tulips (*Amansia*, Liliaceae). *Front. Plant Sci.* **2017**, *8*, 451–463. [[CrossRef](#)]
27. Raman, G.; Park, S. Analysis of the complete chloroplast genome of a medicinal plant, *Dianthus superbus* var *longicalyncinus*, from a comparative genomics perspective. *PLoS ONE* **2015**, *10*, e0141329. [[CrossRef](#)]
28. Yan, L.; Lai, X.; Li, X.; Wei, C.; Tan, X.; Zhang, Y. Analyses of the complete genome and gene expression of chloroplast of sweet potato [*Ipomoea batata*]. *PLoS ONE* **2015**, *10*, e0124083. [[CrossRef](#)]
29. Millen, R.S.; Olmstead, R.G.; Adams, K.L.; Palmer, J.D.; Lao, N.T.; Heggie, L.; Kavanagh, T.A.; Hibberd, J.M.; Gray, J.C.; Morden, C.W. Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell* **2001**, *13*, 645–658. [[CrossRef](#)]
30. Wicke, S.; Schneeweiss, G.M.; Depamphilis, C.W.; Kai, F.M.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [[CrossRef](#)]
31. Ren, T.; Yang, Y.; Zhou, T.; Liu, Z.L. Comparative plastid genomes of *Primula* species: Sequence divergence and phylogenetic relationships. *Int. J. Mol. Sci.* **2018**, *19*, 1050. [[CrossRef](#)]
32. Liu, J.; Qi, Z.C.; Zhao, Y.P.; Fu, C.X.; Xiang, Q.Y. Complete cpDNA genome sequence of *Smilax china* and phylogenetic placement of Liliales—influences of gene partitions and taxon sampling. *Mol. Phylogenet. Evol.* **2012**, *64*, 545–562. [[CrossRef](#)] [[PubMed](#)]
33. Do, H.D.; Kim, J.S.; Kim, J.H. Comparative genomics of four Liliales families inferred from the complete chloroplast genome sequence of *Veratrum patulum* O. Loes. (Melanthiaceae). *Gene* **2013**, *530*, 229–235. [[CrossRef](#)] [[PubMed](#)]

34. Kim, J.S.; Kim, J.H. Comparative genome analysis and phylogenetic relationship of order Liliales insight from the complete plastid genome sequences of two Lilies (*Lilium longiflorum* and *Alstroemeria aurea*). *PLoS ONE* **2013**, *8*, e68180. [CrossRef] [PubMed]
35. Mennes, C.B.; Lam, V.K.Y.; Rudall, P.J.; Lyon, S.P.; Graham, S.W.; Smets, E.F.; Merckx, V.S.F.T. Ancient Gondwana break-up explains the distribution of the mycoheterotrophic family Corsiaceae (Liliales). *J. Biogeogr.* **2015**, *42*, 1123–1136. [CrossRef]
36. Huang, Y.; Li, X.; Yang, Z.; Yang, C.; Yang, J.; Ji, Y. Analysis of complete chloroplast genome sequences improves phylogenetic resolution in Paris (Melanthiaceae). *Front. Plant Sci.* **2016**, *7*, 1797. [CrossRef] [PubMed]
37. 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. [CrossRef]
38. Palmer, J.D.; Thompson, W.F. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. *Cell* **1982**, *29*, 537–550. [CrossRef]
39. Xiao, H.; Jiang, N.; Schaffner, E.; Stockinger, E.J.; Knaap, E.V.D. A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. *Science* **2008**, *319*, 1527–1530. [CrossRef]
40. Yang, M.; Zhang, X.W.; Liu, G.M.; Yin, Y.X.; Chen, K.F. The complete chloroplast genome sequence of date palm (*Phoenix dactylifera* L.). *PLoS ONE* **2010**, *5*, e12762. [CrossRef]
41. Do, H.D.K.; Kim, J.H. A dynamic tandem repeat in monocotyledons inferred from a comparative analysis of chloroplast genomes in Melanthiaceae. *Front. Plant Sci.* **2017**, *8*, 693. [CrossRef]
42. Wang, R.J.; Cheng, C.L.; Chang, C.C.; Wu, C.L.; Su, T.M.; Chaw, S.M. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evol. Biol.* **2008**, *8*, 36. [CrossRef] [PubMed]
43. Plunkett, G.M.; Downie, S.R. Expansion and contraction of the chloroplast inverted repeat in Apiaceae subfamily Apioideae. *Syst. Bot.* **2000**, *25*, 648–667. [CrossRef]
44. Kim, K.J.; Lee, H.L. Complete chloroplast genome sequences from Korean Ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants. *DNA Res.* **2005**, *11*, 247–261. [CrossRef] [PubMed]
45. Dong, W.L.; Wang, R.N.; Zhang, N.Y.; Fan, W.B.; Fang, M.F.; Li, Z.H. Molecular evolution of chloroplast genomes of Orchid species: Insights into phylogenetic relationship and adaptive evolution. *Int. J. Mol. Sci.* **2018**, *19*, 716. [CrossRef]
46. Rønsted, N.; Law, S.; Thornton, H.; Fay, M.F.; Chase, M.W. Molecular phylogenetic evidence for the monophyly of *Fritillaria* and *Lilium* (Liliaceae; Liliales) and the infrageneric classification of *Fritillaria*. *Mol. Phylogenet. Evol.* **2005**, *35*, 509–527. [CrossRef]
47. Turktas, M.; Aslay, M.; Kaya, E.; Ertugrul, F. Molecular characterization of phylogenetic relationships in *Fritillaria* species inferred from chloroplast trnL-trnF sequences. *Turk. J. Biol.* **2012**, *36*, 552–560.
48. Kim, J.S.; Hong, J.K.; Chase, M.W.; Fay, M.F.; Kim, J.H. Familial relationships of the monocot order Liliales based on a molecular phylogenetic analysis using four plastid loci: matK, rbcL, atpB and atpF-H. *Bot. J. Linn. Soc.* **2013**, *172*, 5–21. [CrossRef]
49. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data 2015. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 6 October 2019).
50. Luo, R.; Liu, B.; Xie, Y.; Li, Z.; Huang, W.; Yuan, J.; He, G.; Chen, Y.; Pan, Q.; Liu, Y. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience* **2012**, *1*, 18. [CrossRef]
51. Kears, M.; Moir, R.; Wilson, A.; Havas, S.S.; Cheung, M. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]
52. Lohse, M.; Drechsel, O.; Kahlau, S.; Bock, R. Organellar Genome DRAW—A suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **2013**, *41*, W575. [CrossRef]
53. Frazer, K.A.; Lior, P.; Alexander, P.; Rubin, E.M.; Inna, D. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32* (Suppl. S2), W273–W279. [CrossRef]
54. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]

55. Tamura, K.; Stecher, G.; Peterson, D.; Filipinski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [[CrossRef](#)] [[PubMed](#)]
56. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [[CrossRef](#)]
57. Thiel, T.; Michalek, W.; Varshney, R.K.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [[CrossRef](#)] [[PubMed](#)]
58. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)] [[PubMed](#)]
59. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [[CrossRef](#)] [[PubMed](#)]
60. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [[CrossRef](#)] [[PubMed](#)]
61. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the 2010 Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA, 14 November 2010; pp. 1–8.
62. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572–1574. [[CrossRef](#)]
63. Xie, D.F.; Yu, H.X.; Price, M.; Xie, C.; Deng, Y.Q.; Chen, J.P.; Yu, Y.; Zhou, S.D.; He, X.J. Phylogeny of Chinese *Allium* species in section Daghestanica and adaptive evolution of *Allium* (Amaryllidaceae, Alliioideae) species revealed by the chloroplast complete genome. *Front. Plant Sci.* **2019**, *10*, 460. [[CrossRef](#)]
64. Xie, D.F.; Yu, Y.; Deng, Y.Q.; Li, J.; Liu, H.Y.; Zhou, S.D.; He, X.J. Comparative analysis of the chloroplast genomes of the chinese endemic genus *Urophysa* and their contribution to chloroplast phylogeny and adaptive evolution. *Int. J. Mol. Sci.* **2018**, *19*, 1847. [[CrossRef](#)] [[PubMed](#)]



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