



Article

# Complete Chloroplast Genomes and Comparative Analyses of *L. chinensis*, *L. anhuiensis*, and *L. aurea* (Amaryllidaceae)

Fengjiao Zhang <sup>1,2,\*</sup> , Tao Wang <sup>1,2</sup> , Xiaochun Shu <sup>1,2</sup>, Ning Wang <sup>1,2</sup>, Weibing Zhuang <sup>1,2</sup> and Zhong Wang <sup>1,2,\*</sup>

<sup>1</sup> Jiangsu Key Laboratory for the Research and Utilization of Plant Resources, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (Nanjing Botanical Garden Mem. Sun Yat-Sen), Nanjing 210014, China; johnwt1007@163.com (T.W.); islbe@163.com (X.S.); wangning813@njau.edu.cn (N.W.); weibingzhuangnj@sina.com (W.Z.)

<sup>2</sup> The Jiangsu Provincial Platform for Conservation and Utilization of Agricultural Germplasm, Institute of Botany, Jiangsu Province and Chinese Academy of Sciences (Nanjing Botanical Garden Mem. Sun Yat-Sen), Nanjing 210014, China

\* Correspondence: zfjiao@yeah.net (F.Z.); wangzhong19@163.com (Z.W.)

Received: 8 July 2020; Accepted: 3 August 2020; Published: 10 August 2020



**Abstract:** The genus *Lycoris* (about 20 species) includes important medicinal and ornamental plants. Due to the similar morphological features and insufficient genomic resources, germplasm identification and molecular phylogeny analysis are very limited. Here, we sequenced the complete chloroplast genomes of *L. chinensis*, *L. anhuiensis*, and *L. aurea*; they have very similar morphological traits that make it difficult to identify. The full length of their cp genomes was nearly 158k bp with the same guanine-cytosine content of 37.8%. A total of 137 genes were annotated, including 87 protein-coding genes, 42 tRNAs, and eight rRNAs. A comparative analysis revealed the conservation in sequence size, GC content, and gene content. Some variations were observed in repeat structures, gene expansion on the IR-SC (Inverted Repeat-Single-Copy) boundary regions. Together with the cpSSR (chloroplast simple sequence repeats), these genetic variations are useful to develop molecular markers for germplasm identification. Phylogenetic analysis showed that seven *Lycoris* species were clustered into a monophyletic group, and closed to *Narcissus* in Amaryllidaceae. *L. chinensis*, *L. anhuiensis*, and *L. longituba* were clustered together, suggesting that they were very likely to be derived from one species, and had the same ancestor with *L. squamigera*. Our results provided information on the study of genetic diversity, origins or relatedness of native species, and the identification of cultivars.

**Keywords:** *Lycoris*; chloroplast genome; comparative analysis; phylogeny

## 1. Introduction

The monocot genus *Lycoris* (Amaryllidaceae) is a herbaceous perennial species and commonly used as an important medicinal and ornamental plant. It is rich in Amaryllidaceae-type alkaloids that have been used in immunostimulatory, anti-malarial, tumor-suppressing, and viral activity studies [1,2]. It also has been used as a garden flower due to the colorful and special flowers; another feature is the absence of leaves while flowering, so it is also called the lily [3]. *Lycoris* is native to eastern Asia and confined to temperate China, Japan, and Korea [4].

*Lycoris* consists of about 20 species, of which 15 species and one variety are distributed in China according to the record of “Flora of China”, which represents the most comprehensive catalog. However, several new species continue to be reported, such as *L. hunanensis* [5], *L. hubeiensis* [6], and some hybrids, suggesting that more species are present in the genus than we now recognize. Highly frequent

natural hybridization caused the wide morphological variation in the genus, so there was a general assumption that *Lycoris* is of hybrid origin [7], and *L. radiata* was considered to be a possible ancestor of the genus *Lycori*, according to the phylogenetic analysis using chloroplast genome regions, but it is not consistent with the karyotype evolution study [8]. The complex genetic background makes it difficult to identify some species with high similarity on morphology and clarify the phylogenetic analysis. In the past years, to understand the genetic diversity, possible origin species, maternal donors, and an evolutionary relationship of *Lycoris*, morphological characters [9], plastid DNA sequences (such as *matk*, *rbcL*, *atpB-rbcL* IGS, *trnL-F*, *trnS-trnfM*, and *trnC-ycf6*) [8,10,11] and molecular markers, such as random amplification of polymorphic DNA (RAPD) markers [12], inter-simple sequence repeat (ISSR) markers [13], and Start codon targeted polymorphism (SCoT) markers [14] were reported. However, compared to the results of each method, some results were conflicting, suggesting that these selected loci and fragments were not sufficient for the phylogenetic information to elucidate the evolutionary relationships among *Lycoris* species.

In the past decade, the technology of Next Generation Sequencing (NGS) and the application of DNA barcodes have helped to improve our understanding of the species and relationships in many groups of plants [15,16]. In eukaryotes, the chloroplast is the core organelle for photosynthesis and carbon fixation [17]. Their genome sequences are becoming a very useful tool in identifying closely related plant species, and have been widely used for plant evolution and phylogeny [18–20]. The complete chloroplast genome information has enhanced our understanding of plant biology, diversity, and evolutionary relationships. Moreover, it could reveal considerable variation within species in terms of both sequence and structural variation. It has been widely used in many genera and family analysis, such as *Prunus* [21], *Aristolochia* [22], and *Araceae* [23]. Therefore, a more species' complete chloroplast genome will offer especially valuable information for species identification and conservation of valuable traits, then facilitate the breeding of closely related species and phylogenetic analysis [24]. To date, five *Lycoris* chloroplast genomes have been reported, with published whole sequences and simple clustering but there have been no reports on the structure and comparative analysis of the cp genomes, which need to be paid more attention to analyze their difference by interspecific comparison.

Here, we sequenced three complete chloroplast genomes of *Lycoris* species, *L. chinensis*, *L. anhuiensis*, and *L. aurea*, which have highly similar flowers and are difficult to identify by leaf and bulb morphology. To explore their genetic divergence and phylogenetic relationships, we compared the different characteristics of the cp genomes of these three species and other four reported *Lycoris* species, they were *L. squamigera* [25], *L. radiata* [26], *L. longitube* [27], and *L. sprengeri* [28]. An interspecific comparative analyses of the cp genomes provide valuable molecular markers for germplasm identification, and the phylogenetic analysis may significantly improve the low resolution in evolutionary relationships and contribute to the conservation, domestication, and utilization of *Lycoris* plants.

## 2. Results and Discussion

### 2.1. General Features of the *Lycoris* Chloroplast Genomes

Using next-generation sequencing technology, the total DNA of *L. chinensis*, *L. anhuiensis*, and *L. aurea* were sequenced. As a result, a total of 14 million, 8 million, and 7 million paired-end reads of 150 bp were used for chloroplast genome assembly, and the average organelle coverage reached 14,255×, 8624×, and 7497× sequencing depth, respectively. The organelle assembly NOVOPlasty Version 3.3 [29] was conducted to de novo assemble the cp genomes where the sizes of *L. chinensis*, *L. anhuiensis*, and *L. aurea* were 158,484 bp, 158,490 bp, and 158,367 bp, respectively (Table 1). The other four reported species are also presented in Table 1. The length of *L. radiata* and *L. aurea* was close, *L. sprengeri* was the longest, and the other four species were ranging from 158,480 to 158,490 bp. Their GC contents were 37.8%. Compared with *Allium* (Amaryllidaceae) [30], most of their genome size was between 152k bp and 155k bp, the GC content varied from 36.7% to 37.0%. Here, longer cp

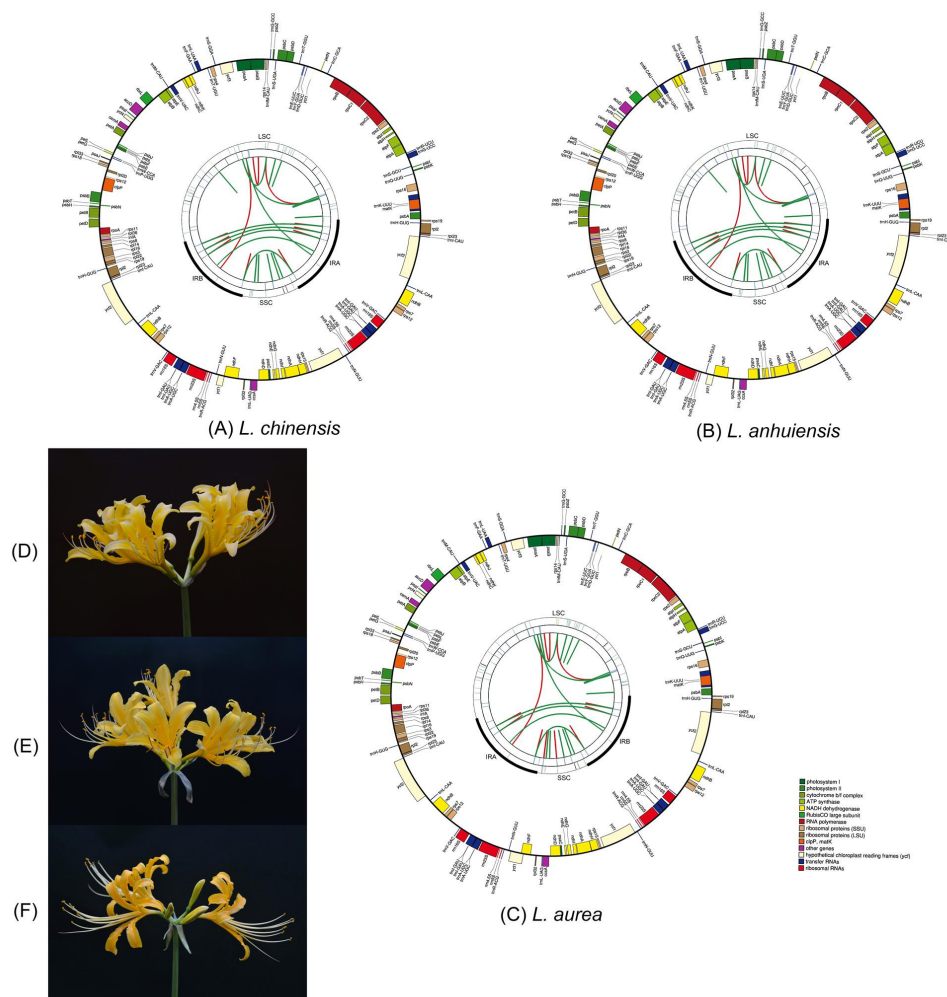
genome sequences and higher GC content were observed in *Lycoris* species. They exhibited the typical quadripartite structure, consisting of a large single-copy region (LSC) (86,458–86,612 bp), a small single-copy region (SSC) (16,620–18,540 bp), and a pair of inverted repeats (IRs) (26,731–27,568 bp) (Table 1 and Figure 1).

**Table 1.** Summary of seven chloroplast genomes of *Lycoris* species.

Genome Features	<i>L. chinensis</i>	<i>L. anhuiensis</i>	<i>L. aurea</i>	<i>L. radiata</i> [26]	<i>L. longituba</i> [27]	<i>L. sprengeri</i> [28]	<i>L. squamigera</i> [25]
Genome size (bp)	158,484	158,490	158,367	158,335	158,484	158,687	158,482
LSC size (bp)	86,458	86,464	86,611	86,612	86,458	86,489	86,454
SSC size (bp)	18,496	18,496	16,620	18,261	18,496	18,540	18,500
IR size (bp)	26,765	26,765	27,568	26,731	26,765	26,829	26,764
GC content (%)	37.8	37.8	37.8	37.8	37.8	37.8	37.8
No. of genes	137	137	137	137	137	137	159
No. of PCGs	87	87	87	87	87	87	105
No. of tRNAs	42	42	42	42	42	42	46
No. of rRNAs	8	8	8	8	8	8	8

PCGs: protein-coding genes.

In seven *Lycoris* species, a total of 137 genes were identified, including 87 protein-coding genes (PCGs), 42 tRNA genes, and eight rRNA genes, of these, 20 genes were duplicated in the different regions (Tables 1 and 2), suggesting the high conservation of cp genes in *Lycoris*. In the three species we provided in this study, the gene distribution in *L. aurea* was slightly different from *L. chinensis* and *L. anhuiensis*. In *L. aurea*, the LSC region contained 63 protein-coding genes (*rps12* duplicated) and 21 tRNA genes, of these genes, *rps19* distributed in the LSC and IR regions. In *L. chinensis* and *L. anhuiensis*, 62 PCGs (*rps12* duplicated) and 22 tRNA genes were distributed in the LSC region, which had no *rps19* but one more *trnN-GUU*. In three species, the gene distribution in the SSC region was consistent, which contained 13 PCGs (*ycf1* duplicated) and one tRNA, two *ycf1* and *ndhF* spanned the SSC and IR regions (Figure 1). The comparative analysis showed the various number of PCGs and tRNAs in different species, but the rRNA number was conserved, where there were 81–87 PCGs in *Allium* [30], 79 in *Amomum* [31], 80 in *Croomia* and *Stemona* [20], but they all had eight rRNA genes (duplicated four rRNAs in the IR regions).



**Figure 1.** The plasmome features and morphological characteristics of *L. chinensis*, *L. anhuiensis*, and *L. aurea*. (A,D) are the plasmome features and morphological characteristics of *L. chinensis*. (B,E) are *L. anhuiensis*. (C,F) are *L. aurea*. The schematic map contain four circles from the center going outward, the first circle shows the forward and reverse repeats connected with red and green arcs, respectively. The next circle shows the tandem repeats marked with short bars. The third circle shows the microsatellite sequences. The fourth circle shows the gene structure on the plasmome. The genes were colored based on their functional categories. (A–C) represented the *L. chinensis*, *L. anhuiensis*, and *L. aurea*, respectively.

**Table 2.** Gene composition of seven *Lycoris* chloroplast genomes.

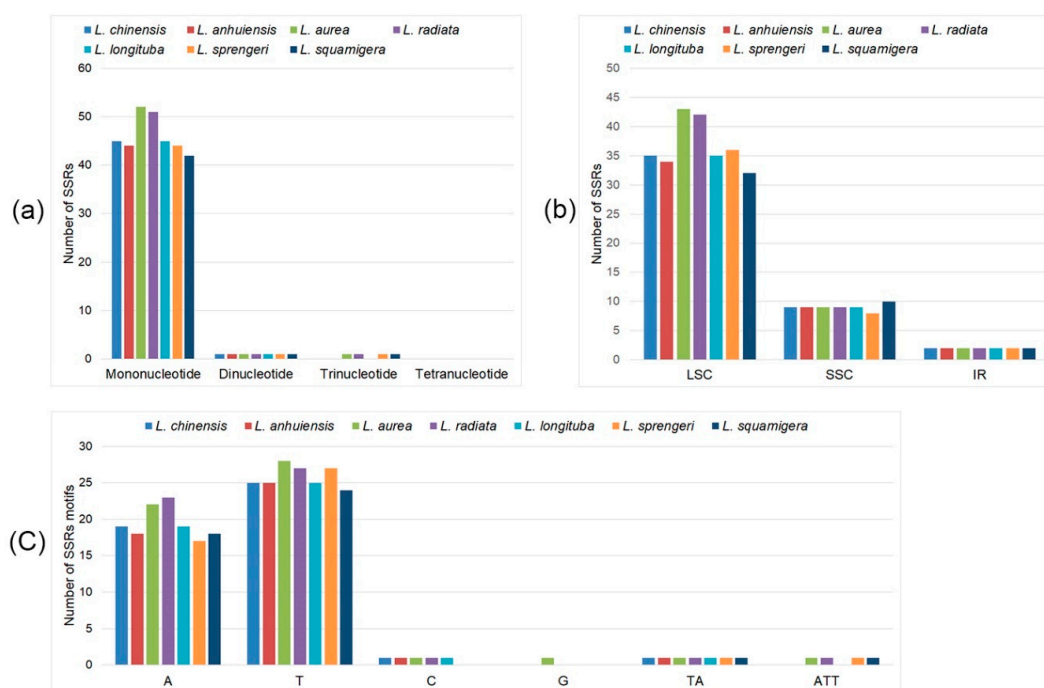
Category of Genes	Group of Genes	Name of Genes	
Genes for photosynthesis	Subunits of photosystem I	<i>psaA, psbA, psbC, psbD, psbE, psbF, psbH, psbI, psbJ</i>	
	Subunits of photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbN, psbT, psbZ, ycf3</i>	
	Subunits of NADH-dehydrogenase	<i>ndhA, ndhB (x2), ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>	
	Subunits of cytochrome b/f complex	<i>petA, petB, petD, petG, petL, petN</i>	
	Subunits of ATP synthase	<i>atpA, atpB, atpE, atpF, atpH, atpI</i>	
	Subunit of rubisco	<i>rbcL</i>	
Self-replication	Large subunit of ribosome	<i>rpl2 (x2), rpl14, rpl16, rpl20, rpl22, rpl23 (x2), rpl32, rpl33, rpl36</i>	
	DNA dependent RNA polymerase	<i>rpoA, rpoB, rpoC1, rpoC2</i>	
	Small subunit of ribosome	<i>rps2, rps3, rps4, rps7 (x2), rps8, rps11, rps12 (x2), rps14, rps15, rps16, rps18, rps19 (x2)</i>	
	Ribosomal RNAs	<i>rrn23S (x2), rrn16S (x2), rrn5S (x2), rrn4.5S (x2), trnA-UGC (x4), trnC-GCA, trnD-GUC, trnE-UUUC, trnF-GAA, trnM-CAU, trnG-GCC, trnG-UCC, trnH-GUG (x2), trnI-CAU (x2), trnI-GAU (x4), trnK-UUU, trnL-CAA (x2), trnL-UAA, trnL-UAG, trnM-CAU, trnN-GUU (x2), trnP-UGG, trnQ-UUG, trnR-ACG (x2), trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC (x2), trnV-UAC, trnW-CCA, trnY-GUA</i>	
	Transfer RNAs	Subunit of Acetyl-CoA-carboxylase	<i>accD</i>
		c-type cytochrome synthesis gene	<i>ccsA</i>
		Envelop membrane protein	<i>cemA</i>
		Protease	<i>clpP</i>
	Other genes	Translational initiation factor	<i>infA</i>
		Maturase	<i>matK</i>
Unknown	Conserved open reading frames	<i>ycf1 (x3), ycf2 (x2), ycf4</i>	

In most plant chloroplast genomes, Group II (G2) introns comprise the majority of noncoding DNA and could provide rich sequence characters for infrageneric and intrafamilial comparative analysis and phylogeny construction [32,33]. The sole group I intron, *trnL* intron, was not detected in *Lycoris* species, and the *trnL* gene itself was lost in their plastomes. In *Lycoris*, there were 17 same-splitting genes in *L. chinensis*, *L. anhuiensis*, *L. aurea*, and *L. longituba*, 18 in *L. radiata* and *L. sprengeri*, including one more *ndhE*, where the 5' end exon was located in the IR region and the 3' end exon and intron were located in the SSR regions. Compared with other *Lycoris* species, there were only 14 splitting genes in *L. squamigera*; however, it contained four specific genes (*trnS-CGA*, *rpoC2*, *trnC-ACA*, and *trnE-UUC*). In all seven cp genomes, there were nine identical splitting genes, which were *atpF*, *rpoC1*, *ycf3*, *trnL-UAA*, *clpP*, *rpl2*, *ndhB*, *trnA-UGC*, and *ndhA*; two of them (*ycf3* and *clpP*) contained two introns and three exons, and others consisted of one intron and two exons (Table S1). In seven *Lycoris* species, the *rps12* intron was lost, which was the same as the three members of *Anemone* [34]. In *L. squamigera*, the *rpl16* intron was absent, that also occurred in some members of Geraniaceae, Goodeniaceae, and Plumbaginaceae [35,36]. In *Lycoris*, the *rpoC1* intron was located in the LSC region, but it was missing in members of the Poaceae, Aizoaceae, Fabaceae, and in [37–39]. The differences of introns between species and conservation in the genus provided information on evolutionary relationships in infrageneric and intergeneric studies.

## 2.2. cpSSRs Analyses and Repeat Structures

Chloroplast simple sequence repeats (cpSSRs) are typically mononucleotide tandem repeat DNA sequences that are widely used as an important molecular marker for ecological and evolutionary processes in plants [40,41]. When there are no effective genomic resources, the simplicity of PCR amplification and the polymorphism of cpSSRs make it easy to use as a marker for characterizing genetic variation. They commonly show intraspecific variation in repeat numbers when located in the noncoding regions of the cp genome [42,43]. Here, we analyzed the number and type of cpSSRs in

seven *Lycoris* species. There were 46, 45, 54, 53, 46, 46, and 44 SSRs in the *L. chinensis*, *L. anhuiensis*, *L. aurea*, *L. radiata*, *L. longituba*, *L. sprengeri*, and *L. squamigera* cp genomes, respectively (Figure 2a and Table S2). The SSRs were largely distributed in LSC regions (72.7–79.6%), 9–10 SSRs were in the SSC region, and two SSRs in IR regions in seven species that were analyzed in the present study (Figure 2b). For the different unit size, the mononucleotide was the most frequent, accounting more than 95% of all types in the seven species, in which base T and A were the primary elements, one C motif in five species, and one G only in *L. aurea* (Figure 2c). There was one dinucleotide (TA) and no tetranucleotide repetition in seven species, showing the conservation among species. However, the trinucleotide (ATT) was detected in only four species, which were *L. aurea*, *L. radiata*, *L. sprengeri*, and *L. squamigera*, suggesting some variation in different species. Our results showed the intraspecific variation in repeat number, repeats distribution, and repeat motifs, the highly similar morphological characteristics of *L. chinensis* and *L. anhuiensis* presented minor SSRs changes. Compared with other published species, the cpSSR loci number ranged from a few to several hundred, such as nice in *Cryptomeria japonica* [44], 237 in *Magnolia kwangsiensis* [45]. The cpSSRs number of *Lycoris* was between 40 and 60, and the cpSSRs were primarily composed of short polyadenine or polythymine repeats but rarely tandem guanine or cytosine repeats, which was consistent with the previous hypothesis [45]. It is the first report about the SSRs of *Lycoris* complete cp genome, in the absence of genome resources of *Lycoris*, the cpSSRs were available markers for characterizing genetic variation and could offer unique insights into the evolutionary processes of *Lycoris* species and their relatives.



**Figure 2.** Analyses of simple sequence repeats of plastome in seven *Lycoris* species. (a) Numbers of different repeat types; (b) Frequency of repeat types in LSC, SSC, and IR regions. (c) Numbers of identified SSRs motifs.

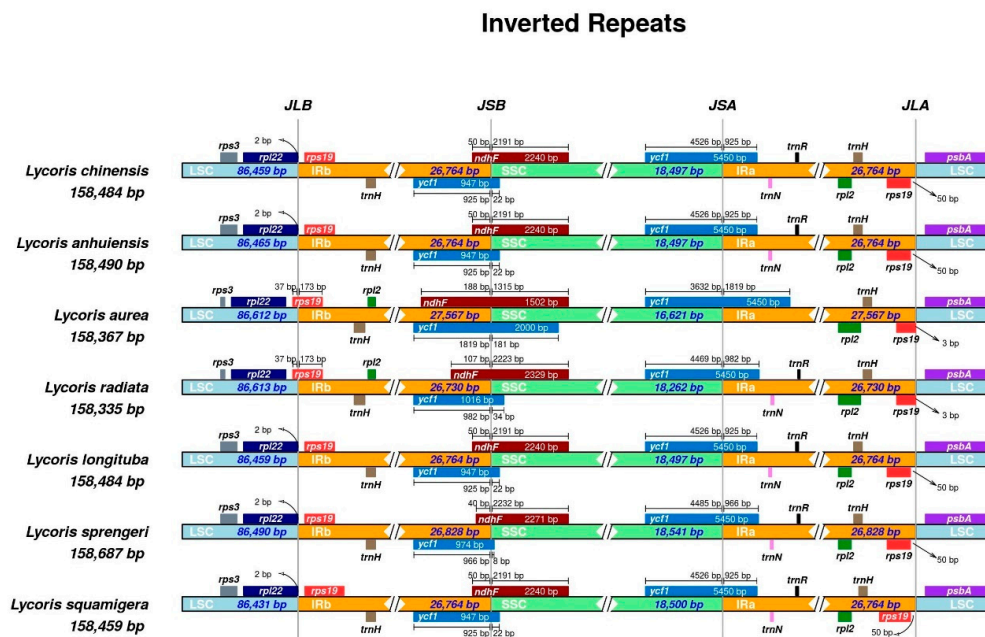
### 2.3. Codon Usage

Codon usage patterns and nucleotide composition of the cp genome could provide a theoretical basis for genetic modification of the chloroplast genome. We deduced the amino acid frequency, the number of codon usage, and the relative synonymous codon usage (RSCU) in the seven *Lycoris* species (Table S3). Seven species presented the 64 RSCU, and the number of codons ranged from 33,640 to 49,982 in seven species, the least in *L. squamigera*. *L. chinensis* and *L. anhuiensis* had the same patterns of codon usage of numbers and types. Leucine and cysteine were the most and least prevalent amino

acids, ATT (encoding isoleucine), and TGA (encoding translational stop) were the most and least used codons in *Lycoris*, respectively. Except for methionine and tryptophan, almost all amino acids had more than one synonymous codon; of these, leucine, arginine and serine had the most (six codon usages). When the RSCU value was  $>1.00$ , we defined that it was preferred, and vice versa. There were 32 preferred and 30 non-preferred (RSCU  $< 1.00$ ) codon usages were identified in five species, which were *L. chinensis*, *L. anhuiensis*, *L. aurea*, *L. radiata*, and *L. longituba*, 31 preferred and 30 non-preferred in *L. sprengeri*, 30 preferred and 31 non-preferred in *L. squamigera* (Table S3). All seven species had the three same stop codons (TAA, TAG, and TGA). Most of the preferred codon usages end with base A or T, and non-preferred codon usages end with base C or G.

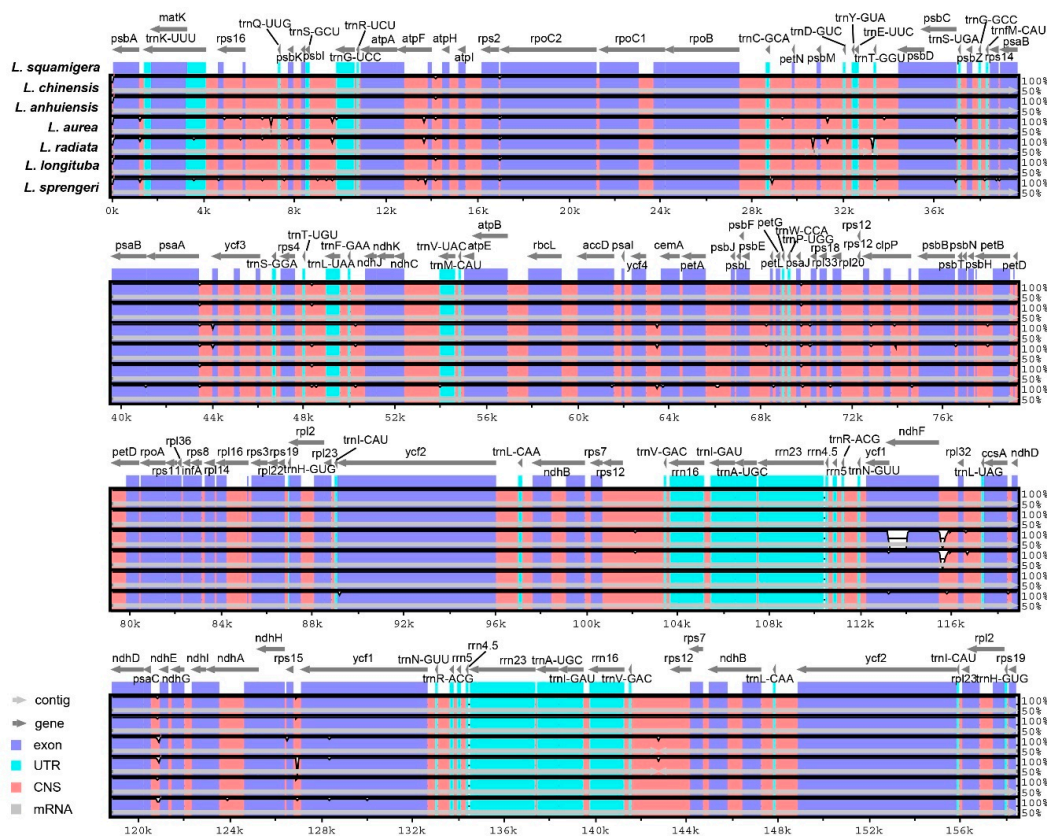
#### 2.4. Inverted Repeats Contraction, Expansion, and Interspecific Comparison

The circular structure of chloroplast genome makes four boundaries among IR, LSC and SSC, which are IRB-LSC, IRB-SSC, IRA-SSC, and IRA-LSC. With the evolution of the genome, the shrinkage and expansion of the IR boundary causes the different sizes of the cp genomes and certain genes to enter the IR region or single-copy region [46]. The IR regions of the seven *Lycoris* species are shown in Figure 3, where the gene number and order were conserved, but minor differences were still exhibited at the boundaries. The IR-LSC borders of *L. radiata* and *L. aurea* were within the *rps19*, where some of the sequences were in the LSC region, showing a partial duplication, but the other five species exhibited IR expansion, leading to the entire *rps19* duplication in the IR regions, which is consistent with some species, such as *Vigna radiata* [47] and *Phaseolus vulgaris* [48]. The IR-SSC boundaries of all species are on the *ycf1*, which is common in angiosperms whose complete copy of *ycf1* spanned the SSC-IR regions [49], 925 bp *ycf1* duplicated at the IR regions in four species, which were *L. chinensis*, *L. anhuiensis*, *L. longituba*, and *L. squamigera*. The *trnH* gene is duplicated at the IR regions in all seven species, but there is variation in *trnR* and *trnN* genes, which were not detected at the IRa region in *L. aurea*. By comparing the IRb/SSC and SSC/IRa regions of *ndhF*, and *ycf1*, high similarity was revealed in the *L. chinensis*, *L. anhuiensis*, *L. longituba* and *L. squamigera*, *L. aurea* showed the most obvious difference at IR-SSC boundary (Figure 3).



**Figure 3.** Comparison of boundary distance of large single-copy regions (LSC), small single-copy regions (SSC), and an inverted repeat (IR) among seven *Lycoris* chloroplast genomes.

To reveal the conservation and divergence in *Lycoris* species, interspecific comparisons were performed using online software mVISTA [50,51], which was commonly used to compare the cp genome sequences from two or more species and visualize the alignment results [31,52]. Here, we compared seven species using annotated *L. squamigera* plastome as a reference, including three species we offered in this study and four species previously reported (Figure 4). The alignment revealed high sequence conservation of the seven *Lycoris* species, especially in the IR regions, but non-coding regions showed more divergence than the coding regions. The exons of *ndhF* and *ycf1* were presented the most obvious difference with others, which was the same as the IR boundary analysis.



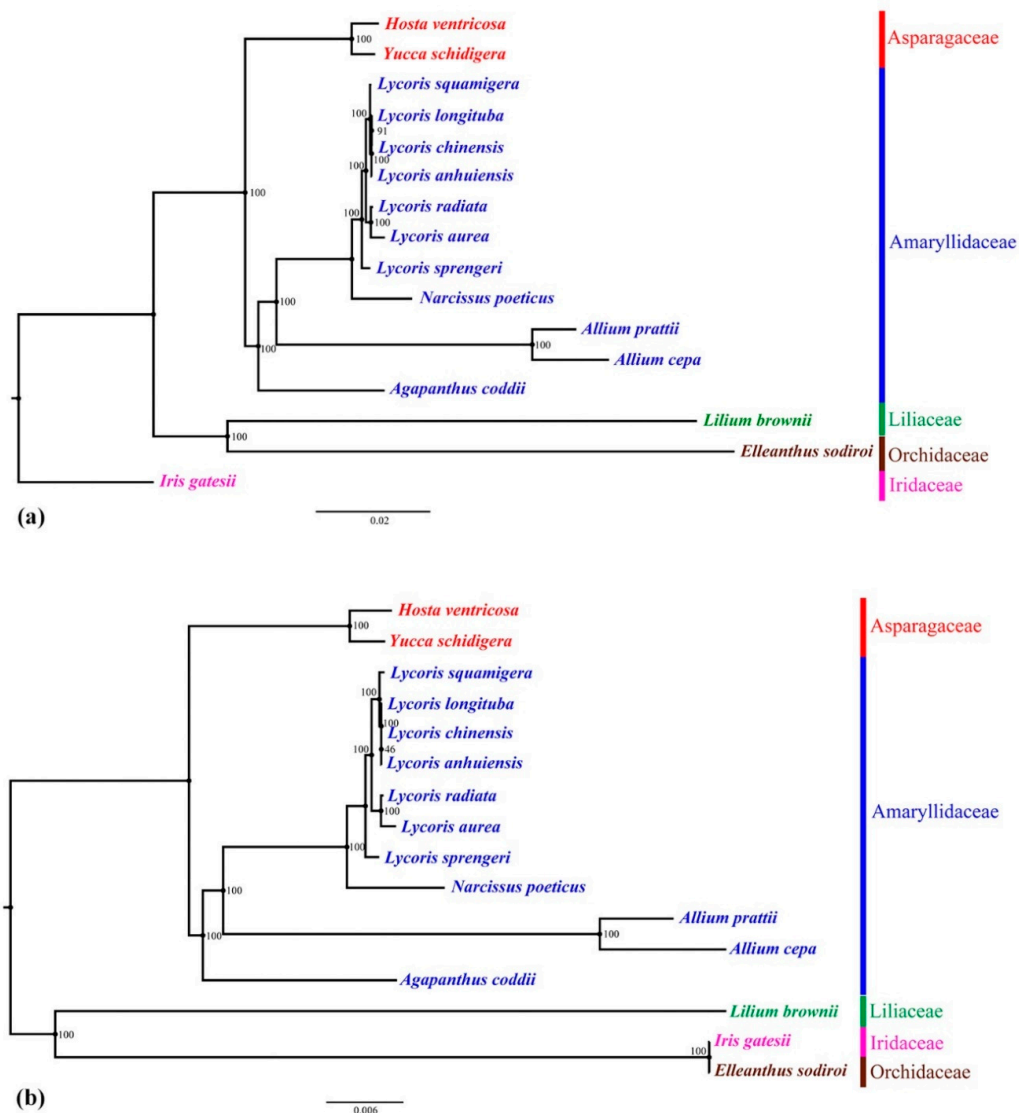
**Figure 4.** mVISTA identity plot based on the complete cp sequence alignment by Shuffle-LAGAN for seven *Lycoris* species. *L. squamigera* was selected as a reference. The pink regions are Conserved Non-Coding Sequences (CNS), the dark blue regions are exons, and the light-blue regions are UTRs. Arrows signifying genes are drawn above the graphs, pointing in the direction of the gene. 70% cut-off identity was used for the plots, and the Y-scale represents the percent identity from 50% to 100%.

### 2.5. Phylogenetic Analysis

To explore the phylogenetic positions and evolutionary relationships of *Lycoris* species, we selected 16 species from five families for the construction of maximum likelihood (ML) tree (Figure 5). The complete cp genome sequences and 72 common protein-coding genes were used for analysis, respectively. On the basis of the topologic structure, each of the five selected families (Amaryllidaceae, Asparagaceae, Orchidaceae, Liliaceae, Iridaceae) was clustered into a monophyletic branch respectively. Seven *Lycoris* species formed a monophyletic clade in both trees. They were close to *Narcissus* in the Amaryllidaceae. Our previous analysis using four complete cp sequences presented the close relationship between *L. longituba* and *L. squamigera* [27], here, because of the addition of three species, it showed a closer relationship among the *L. longituba*, *L. chinensis*, and *L. anhuiensis*, but it was still on a subclade with *L. squamigera*. The complete cp genome analysis supported a closer relationship between *L. chinensis* and *L. longituba* (Figure 5a), but the tree constructed by the PCGs showed a closer



relationship between *L. chinensis* and *L. anhuiensis* (Figure 5b), which was same with the morphological feature classification. The previous study considered that *L. anhuiensis* might be another variety of *L. longituba* [27]. In the present study, two phylogenetic trees showed a tiny difference in their relationship, suggesting that *L. anhuiensis*, *L. longituba* and *L. chinensis* were very likely to be derived from one species, and had the same ancestor with *L. squamigera*. The flower of *L. aurea* was similar to *L. chinensis*, but it was clustered with *L. radiata*, which shared a similar growth habit. *L. sprengeri* was separated into an independent branch, which was the same as our previous result [27]. Here, the three species with the most similar morphological characteristics were not clustered together, suggesting the inconsistent of phylogenetic relationships in morphology and genomics.



**Figure 5.** Phylogenetic relationships of the 16 species by maximum likelihood (ML) analyses. (a) The topology was constructed by the complete cp genome sequences. (b) the tree was constructed based on 72 protein-coding genes.

Before the complete cp sequences were published, interspecific relationship and phylogenetic analysis of *Lycoris* was commonly relied on rDNA internal transcribed spacer (ITS) sequences [53,54] and some specific plastid gene sequences, such as *matK*, *atpB*, *rbcL* [8], *trnS-trnfM*, and *trnC-ycf6* [11]. However, the result of ITS analysis was different from plastid DNA phylogeny, and different plastid fragments also derived differences in results, natural hybridization, and reticulation of *Lycoris* may

be the initial cause. Considering the *Lycoris* hybrid origin, maternal inheritance characteristics of chloroplast genome could provide more evidence to clarify their intraspecific evolution. In this study, we sequenced three cp genome of *Lycoris*, and performed the comparative and phylogenetic analyses with four reported sequences, the result revealed the more accurate relationship of these species. Although it was not included all species in the genus, our result provides more cp genome information of *Lycoris*, and makes sure the closest relationship between *L. chinensis*, *L. anhuiensis* and *L. longituba*.

### 3. Materials and Methods

#### 3.1. Sample Collection, DNA Extraction, and Sequencing

*L. chinensis*, *L. anhuiensis* and *L. aurea* were planted in Nanjing Botanical Garden, Mem. Sun Yat-sen (E118\_83, N32\_06), Nanjing, China. Their specimens were stored at the herbarium of Institute of Botany, Jiangsu Province and Chinese Academy of Sciences with the no. of NAS00591936, NAS00585494, and NAS00585496, respectively. Fresh leaves were collected, quick-frozen in liquid nitrogen, and stored at  $-80\text{ }^{\circ}\text{C}$  for DNA extraction. Genomic DNA was isolated using the Plant Genomic DNA Kit (Huayueyang, Beijing, China) according to the instructions. DNA integrity was examined by electrophoresis in 1% (*w/v*) agarose gel, the concentration was measured, and quantified accurately using NanoDrop spectrophotometer 2000 (Thermo Scientific; Waltham, MA, USA) and Qubit 2.0 respectively. Finally, high-quality DNA was used for libraries' construction and sequencing, which was sequenced at Novogene (Tianjin, China) using Nova-PE150 strategy with an insert size of 350 bp for high-throughput sequencing following the manufacturer's protocol.

#### 3.2. Cp Genome Assembly, Annotation and Structure Analysis

The complete cp genomes were assembled by the organelle assembly NOVOPlasty Version 3.3 [29]. The complete cp genome sequence of *L. radiata* (accession no. MN158120) was selected as a reference [26]. Assembled sequences were manually corrected and circularized by BLASTn comparison with the cp genome of *L. radiata*. GC content was analyzed using Geneious software (version R11, <http://www.geneious.com>). Then, the web server CPGAVAS2 (<http://www.herbalgenomics.org/cpgavas2>) [55] was used for cp genome annotation and visualization with the default parameters. The relative synonymous codon usage (RSCU) values were calculated using MEGAX [56]. The obtained cp genome of *L. chinensis*, *L. anhuiensis*, and *L. aurea* were deposited in the National Center for Biotechnology Information (NCBI), under the accession numbers of *L. chinensis* (MT700549), *L. anhuiensis* (MT700550) and *L. aurea* (MN158985).

#### 3.3. Interspecific Genome Comparison

Four reported complete cp genomes of *Lycoris* were downloaded from the NCBI database, which were *L. squamigera* (MH118290), *L. radiata* (MN158120), *L. sprengeri* (MN158986), and *L. longituba* (MN096601). Chloroplast simple sequence repeat (cpSSR) markers were identified using MISA [43]. The definition of microsatellites (unit size/minimum repeats) was (1/10) (2/6) (3/5) (4/5) (5/5) (6/5) and the maximum length of sequence between two SSRs was set as 100. IRscope (<https://irscope.shinyapps.io/irapp/>) was used for visualizing the genes' differences on the boundaries of the junction sites of the seven chloroplast genomes [57]. mVISTA web-interface [50] was performed to align and compare the seven *Lycoris* cp genome sequences. Shuffle-LAGAN was selected as the alignment program for detecting the rearrangements and inversions in sequences.

#### 3.4. Phylogenetic Analyses

In this study, we obtained three *Lycoris* cp genome sequences, and 13 complete cp sequences of related species were downloaded from the NCBI database, including four *Lycoris* species, four species in Amaryllidaceae, and five species in the other four families. Their Genbank accession number are shown in Table S4. Both the complete cp genome sequences and 72 common PCGs were used for

tree construction. The 72 common PCGs were *rps12*, *psbA*, *matK*, *psbK*, *psbI*, *atpA*, *atpF*, *atpH*, *atpI*, *rpoC2*, *rpoC1*, *rpoB*, *petN*, *psbM*, *psbD*, *psbC*, *psbZ*, *rps14*, *psaB*, *psaA*, *ycf3*, *rps4*, *ndhJ*, *ndhK*, *ndhC*, *atpE*, *atpB*, *rbcL*, *accD*, *psaI*, *ycf4*, *cemA*, *petA*, *psbJ*, *psbL*, *psbF*, *psbE*, *petL*, *petG*, *psaJ*, *rpl33*, *rps18*, *rpl20*, *clpP*, *psbB*, *psbT*, *psbN*, *psbH*, *petB*, *petD*, *rpoA*, *rpl36*, *rps8*, *rpl14*, *rpl16*, *rps3*, *rpl22*, *rpl2*, *rpl23*, *ycf2*, *ndhB*, *rps7*, *ndhF*, *ccsA*, *psaC*, *ndhE*, *ndhG*, *ndhI*, *ndhA*, *ndhH*, *rps15*, *ycf1*, which were extracted and aligned using MAFFT v7.458 [58,59], and manually verified. Maximum likelihood (ML) phylogenies based on the best-fit model of TVM+F+R2 was conducted using IQ-TREE v. 2.0.3 [60]. The Best-fit model by ModelFinder [61], according to Bayesian information criterion (BIC) and the robustness of the topology, was estimated using 1000 bootstrap replicates.

#### 4. Conclusions

In this study, the complete cp genome sequences of *L. chinensis*, *L. anhuiensis*, and *L. aurea* were reported, and interspecific comparison of complete cp genome sequences was analyzed for the first time in the genus *Lycoris*. Compared with the published four cp genome of *Lycoris* species, we found that the genome size, genomic structure, and gene content of the genus *Lycoris* were comparatively conserved, but the IR-SC boundary regions were distinct in seven cp genomes. In all seven species, *trnH* gene was duplicated at the IR regions, but the *trnR* and *trnN* genes were variations. Phylogenetic analyses clarified the interspecific relationship among the *Lycoris* species with whole cp genome sequences. Moreover, we identified the cpSSR sites in different species, together with the highly variable regions identified in interspecific comparison, which could be as useful barcoding markers for the species identification and improve the resolution of phylogenetic relationships within *Lycoris*.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1422-0067/21/16/5729/s1>, Table S1: Splitting genes with introns and exons in the chloroplast genomes of seven *Lycoris* species, Table S2: Chloroplast simple sequence repeats (cpSSRs) of seven *Lycoris* species, Table S3: Relative synonymous codon usage (RSCU) in the *Lycoris* chloroplast genomes, Table S4: The GenBank accession numbers of complete chloroplast genome sequences used for phylogenetic analysis.

**Author Contributions:** F.Z. designed the experiment and wrote the manuscript. T.W. contributed to the sampling. N.W. analyzed the data. X.S. and W.Z. performed the experiment. Z.W. proofed the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Science Foundation of Jiangsu Province (Grant No. BK20180310) and Jiangsu Agricultural Science and Technology Innovation Fund (Grand No. CX(19)3033).

**Acknowledgments:** We thank Feng Peng, Shulan He and Jianliang Wang helped to collect and plant the materials used for the experiments.

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

#### Abbreviations

CP	Chloroplast
NGS	Next-generation sequencing
LSC	Large single copy
SSC	Small single copy
IR	Inverted repeat
GC	Guanine-cytosine
PCGs	Protein-coding genes
cpSSRs	Chloroplast simple sequence repeats
ISSR	inter-simple sequence repeat
RSCU	Relative synonymous codon usage
CNS	Conserved non-coding sequences
RAPD	Random amplification of polymorphic DNA
SCoT	Start codon targeted polymorphism
ML	Maximum likelihood
BIC	Bayesian information criterion

## References

1. Jin, Z. Amaryllidaceae and Sceletium alkaloids. *Nat. Prod. Rep.* **2003**, *20*, 606–614. [[CrossRef](#)] [[PubMed](#)]
2. Wang, R.; Xu, S.; Wang, N.; Xia, B.; Jiang, Y.; Wang, R. Transcriptome analysis of secondary metabolism pathway, transcription factors, and transporters in response to methyl jasmonate in *Lycoris aurea*. *Front. Plant Sci.* **2017**, *7*, 1971. [[CrossRef](#)] [[PubMed](#)]
3. Benqi, Y.; Shoubiao, Z.; Qi, L.; Weihua, Q. Medicinal and ornamental values of genus *Lycoris*. *Chin. Wild Plant Resour.* **2006**, *2*, 29–32.
4. Tsi, Z.; Meerow, A. *Amaryllidaceae. Flora of China*; Beijing Science Press & Missouri Botanical Garden Press: Beijing, China, 2000; Volume 24, p. 264.
5. Quan, M.; Ou, L.; She, C. A new species of *Lycoris* (Amaryllidaceae) from Hunan, China. *Novon* **2013**, *22*, 307–310. [[CrossRef](#)]
6. Meng, W.; Zheng, L.; Shao, J.; Zhou, S.; Liu, K. A new natural allotriploid, *Lycoris* × *hubeiensis* hybr. nov. (Amaryllidaceae), identified by morphological, karyological and molecular data. *Nord. J. Bot.* **2018**, *36*. [[CrossRef](#)]
7. Shi, S.; Qiu, Y.; Li, E.; Wu, L.; Fu, C. Phylogenetic relationships and possible hybrid origin of *Lycoris* species (Amaryllidaceae) revealed by ITS sequences. *Biochem. Genet.* **2006**, *44*, 198–208. [[CrossRef](#)]
8. Hori, T.; Hayashi, A.; Sasanuma, T.; Kurita, S. Genetic variations in the chloroplast genome and phylogenetic clustering of *Lycoris* species. *Genes. Genet. Syst.* **2006**, *81*, 243–253. [[CrossRef](#)]
9. Tae, K.H.; Ko, S.C. A taxonomic study of the genus *Lycoris* (Amaryllidaceae) based on morphological characters. *Trop. Med. Int. Health* **1995**, *9*, 41–46. [[CrossRef](#)]
10. Yuan, J.; Sun, S.; Peng, F.; Feng, X.; Zheng, Y.; Xia, B. Genetic variations in trnL-F sequence and phylogenetic clustering of *Lycoris* species. *China J. Chin. Mater. Med.* **2008**, *33*, 1523–1527.
11. Shi, S.; Sun, Y.; Wei, L.; Lei, X.; Cameron, K.M.; Fu, C. Plastid DNA sequence data help to clarify phylogenetic relationships and reticulate evolution in *Lycoris* (Amaryllidaceae). *Bot. J. Linn. Soc.* **2014**, *176*, 115–126.
12. Yoo, Y.K.; Yuan, T.; Lee, J.S.; Lee, A.K.; Roh, M.S.; Kurita, S.; Suh, J.K. Species relationships of *Lycoris* endemic to Korea evaluated by RAPD and SNPs of nrDNA-ITS regions. *Hortic. Environ. Biote.* **2011**, *52*, 145–151. [[CrossRef](#)]
13. Shi, S.; Qiu, Y.; Wu, L.; Fu, C. Interspecific relationships of *Lycoris* (Amaryllidaceae) inferred from inter-simple sequence repeat data. *Sci. Hortic-amsterdam* **2006**, *110*, 285–291. [[CrossRef](#)]
14. Gao, Y.; Zhu, Y.; Tong, Z.; Xu, Z.; Jiang, X.; Huang, C. Analysis of genetic diversity and relationships among genus *Lycoris* based on start codon targeted (SCoT) marker. *Biochem. Syst. Ecol.* **2014**, *57*, 221–226. [[CrossRef](#)]
15. Kress, W.J.; Wurdack, K.J.; Zimmer, E.A.; Weigt, L.A.; Janzen, D.H. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8369–8374. [[CrossRef](#)]
16. Hollingsworth, P.M.; Li, D.; Der Bank, M.V.; Twyford, A.D. Telling plant species apart with DNA: From barcodes to genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2016**, *371*, 20150338. [[CrossRef](#)] [[PubMed](#)]
17. Palmer, J.D. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *Am. Nat.* **1987**, *130*. [[CrossRef](#)]
18. Moore, M.J.; Bell, C.D.; Soltis, P.S.; Soltis, D.E. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 19363–19368. [[CrossRef](#)]
19. Shaw, J.; Shafer, H.L.; Leonard, O.R.; Kovach, M.J.; Schorr, M.S.; Morris, A.B. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV. *Am. J. Bot.* **2014**, *101*, 1987–2004. [[CrossRef](#)]
20. Lu, Q.; Ye, W.; Lu, R.; Xu, W.; Qiu, Y. Phylogenomic and comparative analyses of complete plastomes of *Croomia* and *Stemona* (Stemonaceae). *Int. J. Mol. Sci.* **2018**, *19*, 2383. [[CrossRef](#)]
21. Xue, S.; Shi, T.; Luo, W.; Ni, X.; Iqbal, S.; Ni, Z.; Huang, X.; Yao, D.; Shen, Z.; Gao, Z. Comparative analysis of the complete chloroplast genome among *Prunus mume*, *P. armeniaca*, and *P. salicina*. *Hortic. Res. Engl.* **2019**, *6*, 89. [[CrossRef](#)]
22. Li, X.; Zuo, Y.; Zhu, X.; Liao, S.; Ma, J. Complete chloroplast genomes and comparative analysis of sequences evolution among seven *Aristolochia* (Aristolochiaceae) medicinal species. *Int. J. Mol. Sci.* **2019**, *20*, 1045. [[CrossRef](#)] [[PubMed](#)]

23. Kim, S.; Yang, J.; Park, J.; Yamada, T.; Maki, M.; Kim, S. Comparison of whole plastome sequences between thermogenic skunk cabbage *Symplocarpus renifolius* and Nonthermogenic *S. nipponicus* (Orontioideae; Araceae) in East Asia. *Int. J. Mol. Sci.* **2019**, *20*, 4678. [[CrossRef](#)] [[PubMed](#)]
24. Daniell, H.; Lin, C.; Yu, M.; Chang, W. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [[CrossRef](#)]
25. Jin, S.W.; Park, J.Y.; Kang, S.-J.; Park, H.-S.; Shim, H.; Lee, T.J.; Kang, J.H.; Sung, S.H.; Yang, T.-J. The complete chloroplast genome sequence of Magic Lily (*Lycoris squamigera*). *Mitochondrial DNA B* **2018**, *3*, 1210–1211. [[CrossRef](#)]
26. Zhang, F.; Shu, X.; Wang, T.; Zhuang, W.; Wang, Z.L. The complete chloroplast genome sequence of *Lycoris radiata*. *Mitochondrial DNA B* **2019**, *4*, 2886–2887. [[CrossRef](#)]
27. Zhang, F.; Tong, H.; Yang, H.; Wang, T.; Zhuang, W.; Shu, X.; Wang, Z.L. Characterisation of the complete chloroplast genome of *Lycoris longituba* (Amaryllidaceae). *Mitochondrial DNA B* **2019**, *4*, 3782–3783. [[CrossRef](#)]
28. Zhang, F.; Zhuang, W.; Shu, X.; Wang, T.; Wang, Z.L. Complete chloroplast genome of *Lycoris sprengeri* (Amaryllidaceae) and genetic comparison. *Mitochondrial DNA B* **2019**, *4*, 3577–3578. [[CrossRef](#)]
29. Dierckxsens, N.; Mardulyn, P.; Smits, G. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **2016**, *45*, e18.
30. Xie, D.-F.; Tan, J.-B.; Yu, Y.; Gui, L.-J.; Su, D.-M.; Zhou, S.-D.; He, X.-J. Insights into phylogeny, age and evolution of *Allium* (Amaryllidaceae) based on the whole plastome sequences. *Ann. Bot. Lond.* **2020**, *125*, 1039–1055. [[CrossRef](#)]
31. Cui, Y.; Chen, X.; Nie, L.; Sun, W.; Hu, H.; Lin, Y.; Li, H.; Zheng, X.; Song, J.; Yao, H. Comparison and phylogenetic analysis of chloroplast genomes of three medicinal and edible *Amomum* species. *Int. J. Mol. Sci.* **2019**, *20*, 4040. [[CrossRef](#)]
32. Jurica, M.S.; Stoddard, B.L. Homing endonucleases: Structure, function and evolution. *Cell. Mol. Life Sci.* **1999**, *55*, 1304–1326. [[CrossRef](#)] [[PubMed](#)]
33. Kelchner, S.A. Group II introns as phylogenetic tools: Structure, function, and evolutionary constraints. *Am. J. Bot.* **2002**, *89*, 1651–1669. [[CrossRef](#)] [[PubMed](#)]
34. Hoot, S.B.; Palmer, J.D. Structural rearrangements, including parallel inversions, within the chloroplast genome of *Anemone* and related genera. *J. Mol. Evol.* **1994**, *38*, 274–281. [[CrossRef](#)] [[PubMed](#)]
35. Downie, S.R.; Palmer, J.D. A chloroplast DNA phylogeny of the Caryophyllales based on structural and inverted repeat restriction site variation. *Syst. Bot.* **1994**, *19*, 236–252. [[CrossRef](#)]
36. Campagna, M.; Downie, S. The intron in chloroplast gene *rpl16* is missing from the flowering plant families Geraniaceae, Goodeniaceae, and Plumbaginaceae. *T. Illinois State Acad. Sci.* **1998**, *91*, 1–11.
37. Katayama, H.; Ogihara, Y. Phylogenetic affinities of the grasses to other monocots as revealed by molecular analysis of chloroplast DNA. *Curr. Genet.* **1996**, *29*, 572–581. [[CrossRef](#)]
38. Downie, S.R.; Llanas, E.; Katzdownie, D.S. Multiple independent losses of the *rpoC1* intron in angiosperm chloroplast DNA's. *Syst. Bot.* **1996**, *21*, 135–151. [[CrossRef](#)]
39. Thiede, J.; Schmidt, S.A.; Rudolph, B. Phylogenetic implication of the chloroplast *rpoC1* intron loss in the Aizoaceae (Caryophyllales). *Biochem. Syst. Ecol.* **2007**, *35*, 372–380. [[CrossRef](#)]
40. Parks, M.; Cronn, R.; Liston, A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.* **2009**, *7*, 84. [[CrossRef](#)]
41. Wheeler, G.L.; Dorman, H.E.; Buchanan, A.T.; Challagundla, L.; Wallace, L.E. A review of the prevalence, utility, and caveats of using chloroplast simple sequence repeats for studies of plant biology. *Appl. Plant Sci.* **2014**, *2*, 1400059. [[CrossRef](#)]
42. Ebert, D.; Peakall, R. Chloroplast simple sequence repeats (cpSSRs): Technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Mol. Ecol. Resour.* **2009**, *9*, 673–690. [[CrossRef](#)] [[PubMed](#)]
43. Yang, A.; Zhang, J.; Yao, X.; Huang, H. Chloroplast microsatellite markers in *Liriodendron tulipifera* (Magnoliaceae) and cross-species amplification in *L. chinense*. *Am. J. Bot.* **2011**, *98*. [[CrossRef](#)] [[PubMed](#)]
44. Hirao, T.; Watanabe, A.; Kurita, M.; Kondo, T.; Takata, K. A frameshift mutation of the chloroplast *matK* coding region is associated with chlorophyll deficiency in the *Cryptomeria japonica* virescent mutant Wogon-Sugi. *Curr. Genet.* **2009**, *55*, 311–321. [[CrossRef](#)] [[PubMed](#)]

45. Kuang, D.; Wu, H.; Wang, Y.; Gao, L.; Zhang, S.; Lu, L. Complete chloroplast genome sequence of *Magnolia kwangsiensis* (Magnoliaceae): Implication for DNA barcoding and population genetics. *Genome* **2011**, *54*, 663–673. [[CrossRef](#)] [[PubMed](#)]
46. Lin, C.; Wu, C.; Huang, Y.; Chaw, S. The complete chloroplast genome of *Ginkgo biloba* reveals the mechanism of inverted repeat contraction. *Genome Biol. Evol.* **2012**, *4*, 374–381. [[CrossRef](#)]
47. Tangphatsornruang, S.; Sangsrakru, D.; Chanprasert, J.; Uthaipaisanwong, P.; Yoocha, T.; Jomchai, N.; Tragoonrung, S. The chloroplast genome sequence of Mungbean (*Vigna radiata*) determined by high-throughput pyrosequencing: Structural organization and phylogenetic relationships. *DNA Res.* **2010**, *17*, 11–22. [[CrossRef](#)] [[PubMed](#)]
48. Ma, P.; Zhang, Y.; Zeng, C.; Guo, Z.; Li, D. Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (poaceae). *Syst. Biol.* **2014**, *63*, 933–950. [[CrossRef](#)]
49. Raubeson, L.A.; Peery, R.; Chumley, T.W.; Dziubek, C.; Fourcade, H.M.; Boore, J.L.; Jansen, R.K. Comparative chloroplast genomics: Analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genomics* **2007**, *8*, 174. [[CrossRef](#)]
50. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, 273–279. [[CrossRef](#)]
51. Mayor, C.; Brudno, M.; Schwartz, J.R.; Poliakov, A.; Rubin, E.M.; Frazer, K.A.; Pachter, L.; Dubchak, I. VISTA: Visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **2000**, *16*, 1046–1047. [[CrossRef](#)]
52. Silva, S.R.; Pinheiro, D.G.; Penha, H.A.; Plachno, B.J.; Michael, T.P.; Meer, E.J.; Miranda, V.F.O.; Varani, A.M. Intraspecific variation within the *Utricularia amethystina* Species morphotypes based on chloroplast genomes. *Int. J. Mol. Sci.* **2019**, *20*, 6130. [[CrossRef](#)]
53. Chen, Y.; Gao, Y.; Liao, W.; Tong, Z. Analysis on rDNA-ITS sequence and research of intra-specific phylogeny of *Lycoris albiflora*. *J. Plant Resour. Environ.* **2009**, *18*, 25–31.
54. Quan, M.; Ou, L.; She, C.; Wu, X.; Chen, D. rDNA internal transcribed spacer sequence analysis of *Lycoris* Hert. *Afr. J. Biotechnol.* **2012**, *11*, 7361–7365.
55. Shi, L.; Chen, H.; Jiang, M.; Wang, L.; Wu, X.; Huang, L.; Liu, C. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* **2019**. [[CrossRef](#)] [[PubMed](#)]
56. Kumar, S.; Stecher, G.; Li, M.; Nnyaz, C.; Tamura, K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)] [[PubMed](#)]
57. Amiryousefi, A.; Hyvonen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [[CrossRef](#)]
58. Kuraku, S.; Zmasek, C.M.; Nishimura, O.; Katoh, K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* **2013**, *41*, 22–28. [[CrossRef](#)]
59. Rozewicki, J.; Li, S.; Amada, K.M.; Standley, D.M.; Katoh, K. MAFFT-DASH: Integrated protein sequence and structural alignment. *Nucleic Acids Res.* **2019**, *47*, W5–W10. [[CrossRef](#)] [[PubMed](#)]
60. Nguyen, L.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)] [[PubMed](#)]
61. Kalyaanamoorthy, S.; Bui, M.Q.; Wong, T.K.F.; Von Haeseler, A.; Jermini, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [[CrossRef](#)] [[PubMed](#)]

