



## Research Article

# Comparison Analysis Based on Complete Chloroplast Genomes and Insights into Plastid Phylogenomic of Four *Iris* Species

Jing-lu Feng,<sup>1</sup> Li-wei Wu,<sup>1,2</sup> Qing Wang,<sup>1,2</sup> Yun-jia Pan,<sup>1</sup> Bao-li Li,<sup>1</sup> Yu-lin Lin <sup>1</sup>,  
and Hui Yao <sup>1,2</sup>

<sup>1</sup>Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China

<sup>2</sup>Engineering Research Center of Chinese Medicine Resource, Ministry of Education, Beijing 100193, China

Correspondence should be addressed to Yu-lin Lin; [linyulin2022@163.com](mailto:linyulin2022@163.com) and Hui Yao; [scauyaoh@sina.com](mailto:scauyaoh@sina.com)

Received 20 March 2022; Revised 6 June 2022; Accepted 5 July 2022; Published 25 July 2022

Academic Editor: Xiaoxuan Tian

Copyright © 2022 Jing-lu Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Iris* species, commonly known as rainbow flowers because of their attractive flowers, are extensively grown in landscape gardens. A few species, including *Belamcanda chinensis*, the synonym of *I. domestica* and *I. tectorum*, are known for their medicinal properties. However, research on the genomes and evolutionary relationships of *Iris* species is scarce. In the current study, the complete chloroplast (CP) genomes of *I. tectorum*, *I. dichotoma*, *I. japonica*, and *I. domestica* were sequenced and compared for their identification and relationship. The CP genomes of the four *Iris* species were circular quadripartite with similar lengths, GC contents, and codon usages. A total of 113 specific genes were annotated, including the *ycf1* pseudogene in all species and *rps19* in *I. japonica* alone. All the species had mononucleotide (A/T) simple sequence repeats (SSRs) and long forward and palindromic repeats in their genomes. A comparison of the CP genomes based on mVISTA and nucleotide diversity (Pi) identified three highly variable regions (*ndhF-rpl32*, *rps15-ycf1*, and *rpl16*). Phylogenetic analysis based on the complete CP genomes concluded that *I. tectorum* is a sister of *I. japonica*, and the subgenus *Pardanthopsis* with several *I. domestica* clustered into one branch is a sister of *I. dichotoma*. These findings confirm the feasibility of superbarcodes (complete CP genomes) for *Iris* species authentication and could serve as a resource for further research on *Iris* phylogeny.

## 1. Introduction

*Iris* (L.) is a genus of flowering plants, including 300 species of the Iridaceae family classified into six subgenera (subg.) [1, 2]. These species, commonly called rainbow flowers, are found in the northern hemisphere's temperate regions and are widely used in landscape gardens because of their beautiful and colorful flowers [3]. Most *Iris* species can adapt to dry environments, such as deserts, semideserts, or rocky habitats, and a few live in mesic and wetland areas [4]. *Iris* species are also used as medicinal plants. Several pharmacological studies have shown that the rhizome extracts of *Iris* species have anticancer, anti-inflammatory, and  $\alpha$ -glucosidase inhibitory effects and can reduce human infarct volume [5–7]. Few species are used to treat throat-swelling diseases [8]. The dried rhizomes of *I. tectorum* and *I. domestica*, referred to as “Chuan She Gan” and “She Gan,” respectively,

are used in traditional Chinese medicine, but “She Gan” is often adulterated with the dried rhizomes of *I. dichotoma* and *I. japonica*. Therefore, identifying these four species is needed for clinical safety.

*Iris* species are characterized by fan-shaped leaves, three colorful outer perianth segments, three inner perianth segments, three petaloid stigmas with a bifid crest, and underground tuberous organs [9]. However, these species have similar leaf shapes, flower shapes, and rhizome morphological characteristics. Therefore, identification based on morphological features alone is complicated, especially during the nonflowering period. The development of *I. domestica* and *I. dichotoma* hybrids has also made species identification challenging owing to the similarities between the hybrids and female parents [10]. Molecular phylogeny combined with palynology suggested that *I. tectorum* is far away from *I. japonica* [11], which is inconsistent with classical

TABLE 1: Length and composition of the CP genomes of *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*.

Types/species	<i>I. tectorum</i>	<i>I. japonica</i>	<i>I. dichotoma</i>	<i>I. domestica</i>
Accession number	MW201731	OK448493	OK448492	OK448491
Total length (bp)	153,253	152,443	153,658	153,736
SSC (bp)	18,562	18,490	18,150	18,168
LSC (bp)	82,833	83,237	83,116	83,140
IRs (bp)	51,858	50,716	52,392	52,428
CDS (bp)	78,957	78,507	79,050	79,059
Total GC (%)	37.89	37.85	37.87	37.85
GC of SSC (%)	31.42	31.40	31.49	31.46
GC of LSC (%)	36.16	36.13	36.00	35.97
GC of IRa (%)	42.97	43.03	43.04	43.05
GC of IRb (%)	42.97	43.03	43.04	43.05
GC of CDS (%)	38.15	38.08	38.03	38.02
AT at the 1 <sup>st</sup> position (%)	54.42	54.48	54.43	54.44
AT at the 2 <sup>nd</sup> position (%)	61.77	61.81	61.75	61.77
AT at the 3 <sup>rd</sup> position (%)	69.36	69.46	69.73	69.72

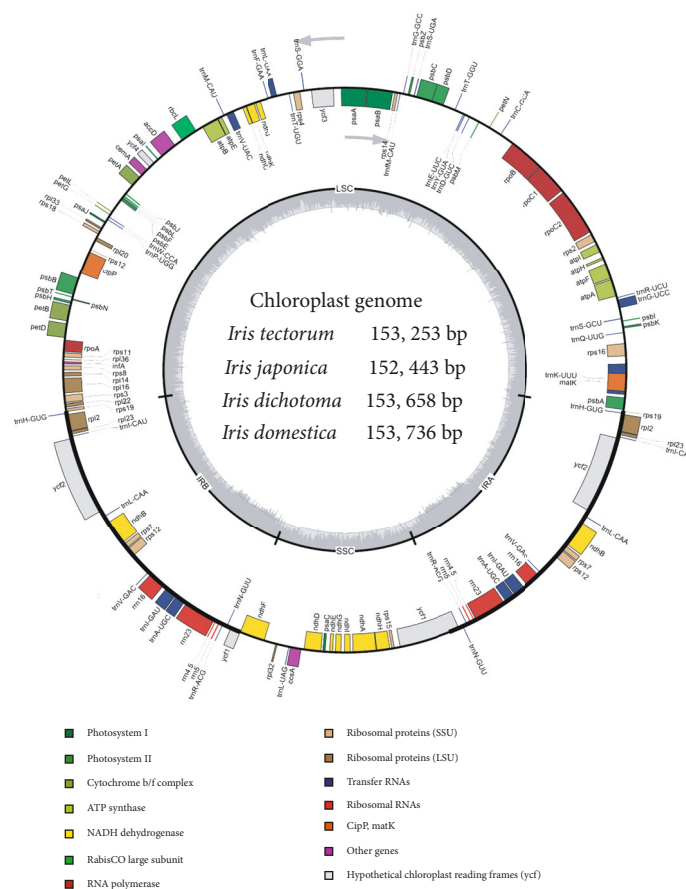


FIGURE 1: Chloroplast genome map of *Iris tectorum*. Arrows represent the transcription direction of genes. The dark (GC) and light (AT) gray areas are nucleotide contents.

taxonomy that shows the two species with a close relationship. *I. tectorum* is a species of section (sect.) *Lophiris* of subg. *Limniris* sect. *Lophiris* contains 13 species distributed in Eastern Asia; Dykes included this rank in sect. *Evansia*

[12], but this rank was later amended by Lawrence to subsection *Evansia* [13], by Rodionenko to subg. *Crossiris* [14], and finally by Mathew to sect. *Lophiris* of subg. *Limniris* [2]. Molecular phylogeny placed *I. domestica* in subg.

TABLE 2: Genes in the CP genomes of *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*.

Functional group	Genes	Number of genes
Photosystem I	<i>psaA, psaB, psaC, psaI, psaJ</i>	5
Photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>	15
Cytochrome b/f complex	<i>petA, petB *</i> , <i>petD *</i> , <i>petG, petL, petN</i>	6
ATP synthase	<i>atpA, atpB, atpE, atpF *</i> , <i>atpH, atpI</i>	6
NADH dehydrogenase	<i>ndhA *</i> , <i>ndhB *</i> (×2), <i>ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>	12
RubisCO large subunit	<i>rbcL</i>	1
RNA polymerase	<i>rpoA, rpoB, rpoC1 *</i> , <i>rpoC2</i>	4
Ribosomal proteins (SSU)	<i>rps2, rps3, rps4, rps7</i> (×2), <i>rps8, rps11, rps12 **</i> (×2), <i>rps14, rps15, rps16 *</i> , <i>rps18, rps19<sup>ψ</sup></i> (×2)	15
Ribosomal proteins (LSU)	<i>rpl2 *</i> (×2), <i>rpl14, rpl16 *</i> , <i>rpl20, rpl22, rpl23</i> (×2), <i>rpl32, rpl33, rpl36</i>	11
Other genes	<i>accD, clpP **</i> , <i>matK, ccsA, cemA, infA</i>	6
Proteins of unknown function	<i>ycf1<sup>ψ</sup>, ycf2</i> (×2), <i>ycf3 **</i> , <i>ycf4</i>	6
Transfer RNAs	38 tRNAs (8 in the IRs (×2), 6 contain one intron)	38
Ribosomal RNAs	<i>rrn4.5</i> (×2), <i>rrn5</i> (×2), <i>rrn16</i> (×2), <i>rrn23</i> (×2)	8

×2 indicates two gene copies. \* and \*\* indicate genes that contain 1 and 2 introns, respectively. <sup>ψ</sup> indicates a *pseudogene*.

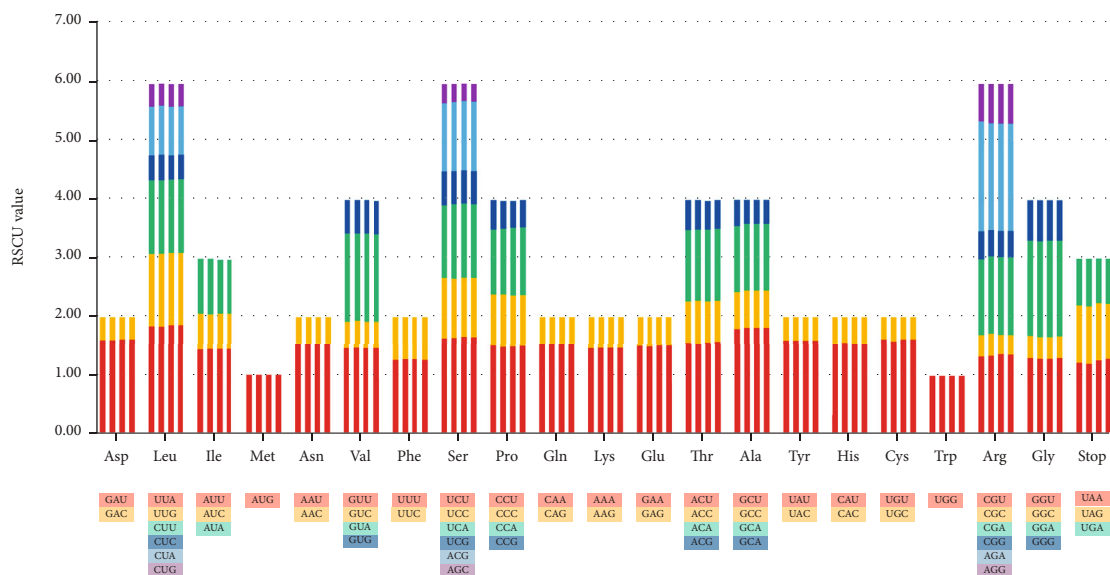


FIGURE 2: Codon usage of 20 amino acids and stop codons of the CDS in the CP genomes of *Iris* species. The four histograms from left to right in each amino acid represent *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*.

*Pardanthopsis* [15–17] with high support rates. Goldblatt and Mabberley confirmed that *Belamcanda chinensis* is a synonym of *Iris domestica* based on molecular, karyotype, and type specimen analyses [18]. Furthermore, karyotype analysis of *Iris* species revealed that their chromosomal genetics are abundant because of their complex origin [10, 19–23]. A few taxa of *Iris* species were identified using DNA barcodes [24–27]. Wilson [28–30] made considerable progress on molecular identification and phylogeny in *Iris* species. However, taxonomy of the *Iris* species still remains complicated [10, 11, 31, 32].

Angiosperms have a circular tetrameric chloroplast (CP) genome, consisting of a pair of inverted repeats (IRs), a small single copy (SSC) region, and a large single copy (LSC) region [33, 34]. The CP genomes serve as promising

tools in identifying species and analyzing phylogeny owing to their small and simple structure, conserved sequences, and moderate nucleotide substitution rate [35–37]. Few researchers analyzed the molecular phylogenies of *Iris* based on CP or nuclear DNA fragments; however, studies based on complete CP genomes are limited. Approximately 20 complete CP genomes about *Iris* species were documented in NCBI. However, the data need to be enriched to provide detailed information on the phylogeny [26, 38–45].

The current study sequenced the complete CP genomes of *I. tectorum*, *I. dichotoma*, *I. japonica*, and *I. domestica*. The study’s major objectives were to (1) characterize the complete CP genome structure and functional genes, (2) analyze the codon usage, (3) identify the SSRs and long repeats, and (4) compare the whole CP genomes of *Iris*

TABLE 3: SSRs in the CP genomes of four *Iris* species.

SSR types	Repeat units	Number				Proportion (%)			
		①	②	③	④	①	②	③	④
Mono	A/T	38	22	34	32	100.0	100.0	97.1	97.0
	C/G	—	—	1	1	—	—	2.9	3.0
Di	AT/AT	9	8	11	10	81.8	80.0	84.6	83.3
	AG/CT	2	2	2	2	18.2	20.0	15.4	16.7
Tri	AAG/CTT	2	2	2	2	50.0	50.0	66.7	66.7
	AAT/ATT	2	2	1	1	50.0	50.0	33.3	33.3
Tetra	AAAT/ATTT	2	3	3	3	66.7	75.0	75.0	75.0
	AATG/ATTC	1	1	1	1	33.3	25.0	25.0	25.0
Penta	AACTT/AAGTT	1	1	1	1	33.3	100.0	50.0	33.3
	AAAAT/ATTTT	1	—	—	1	33.3	—	—	33.3
	AAAAC/GTTTT	1	—	—	—	33.3	—	—	—
	AATAT/ATATT	—	—	—	1	—	—	—	33.3
Hexa	ACTAT/AGTAT	—	—	1	—	—	—	50.0	—
	AACAAG/CTTGTT	—	1	1	1	—	100.0	100.0	100.0

①: *I. tectorum*; ②: *I. japonica*; ③: *I. dichotoma*; ④: *I. domestica*; —: the absence of a particular type.

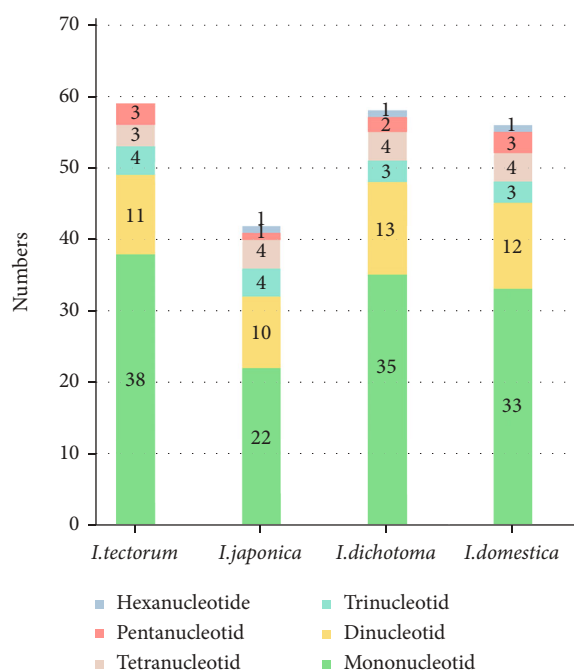


FIGURE 3: Distribution of the six types of SSRs in the CP genomes of four *Iris* species.

species to screen highly variable regions. The genomes were further used to uncover the phylogeny relationship among *Iris* species. The findings will lay a foundation for classifying the species and elucidating the phylogeny in Iridaceae.

## 2. Materials and Methods

**2.1. Sample Collection.** Leaves (fresh) from *I. tectorum*, *I. dichotoma*, and *I. domestica* were collected from the Institute of Medicinal Plant Development (IMPLAD), Beijing (40°2'5''N, 116°16'14''E), and those of *I. japonica* were from the

Chengdu University of Traditional Chinese Medicine, Chengdu (30°24'36''N, 103°28'48''E). The leaves were stored in a -80°C freezer, and Professor Yulin Lin identified the species. Voucher specimens were deposited in the herbarium of IMPLAD, the Chinese Academy of Medical Sciences, and the Peking Union Medical College.

**2.2. DNA Extraction and Sequencing.** Total DNA was extracted from the leaf samples by using the DNeasy Plant Mini Kit (Qiagen Co., Hilden, Germany). DNA quality was detected by agarose gel (1%) electrophoresis. The libraries (insert size average, 350 bp) were generated from total DNA and sequenced on an Illumina NovaSeq 6000 system.

**2.3. CP Genome Assembly and Annotation.** Filtered reads (low quality) from raw data were generated by Fastp version 0.23.2 [46], and clean data were assembled to generate the CP genome in GetOrganelle version 1.7.5.1 [47]. The genes were annotated using GeSeq version 2.03 [48], followed by manual correction. The genome circular map was drawn by OrganellarGenomeDRAW version 1.3.1 [49]. The whole CP genome sequences of *I. japonica* (OK448493), *I. tectorum* (MW201731), *I. dichotoma* (OK448492), and *I. domestica* (*B. chinensis*; OK448491) were submitted to NCBI.

**2.4. Genome Structure and Codon Usage Analyses.** Furthermore, MEGA X [50] was used to examine the GC content of the genome. CodonW version 1.4.2 was used to calculate the codon usage using the relative synonymous codon usage (RSCU) value as follows: there is no preference in codon usage (RSCU = 1), the codon usage frequency is less than expected (RSCU > 1), and the codon usage frequency is more than expected (RSCU < 1) [51, 52].

**2.5. SSR and Long Repeat Sequence Analyses.** The SSRs were examined by using the Microsatellite Identification tool version 2.1 [53, 54], with the parameters mentioned by Cui





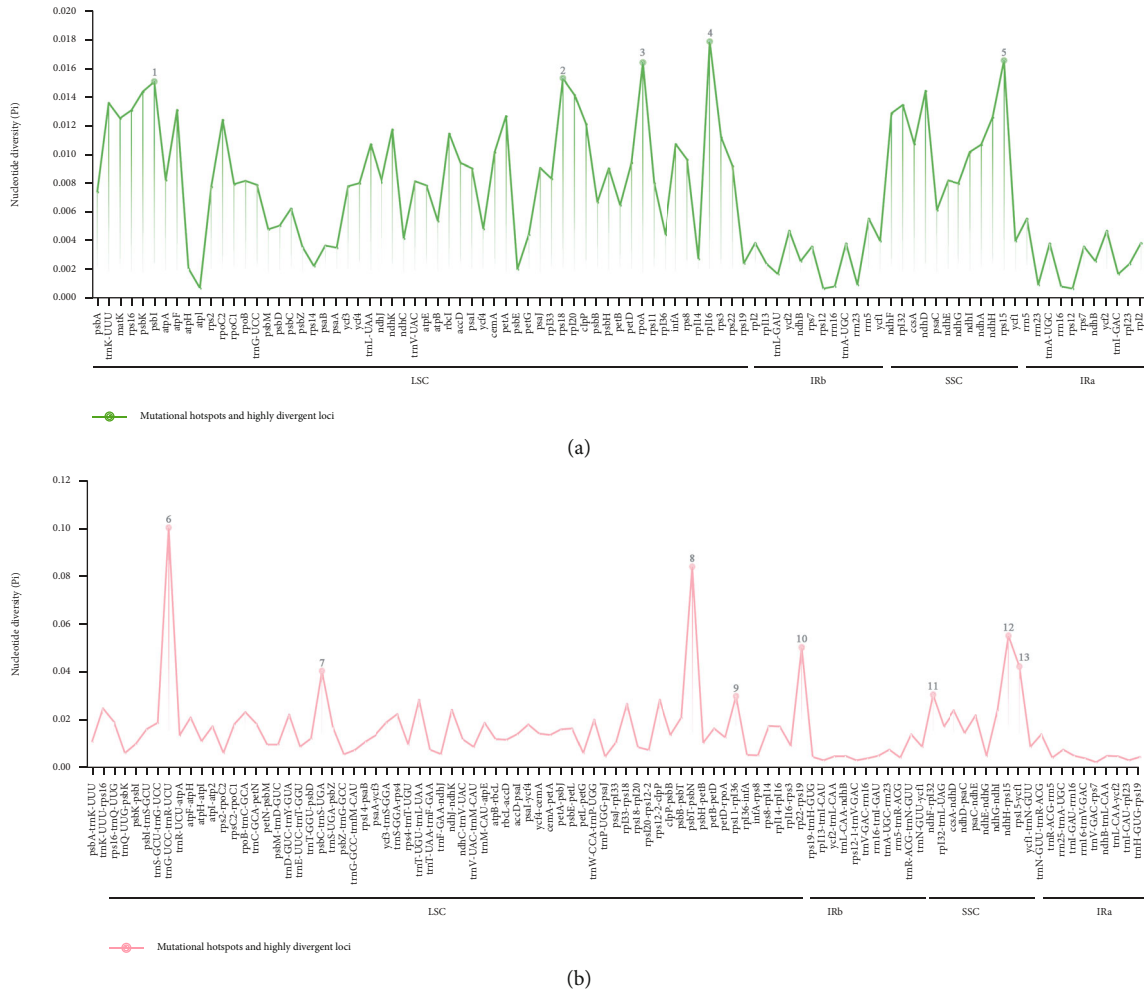


FIGURE 7: Nucleotide diversity of the gene (a) and intergenic spacer regions (b) on the whole CP genomes of four *Iris* species (Pi > 0, length > 100 bp).

26,169, 26,350, and 26,353 amino acid codons, respectively. The analysis of 64 codons encoding 20 amino acids (Figure 2 and Table S3) revealed that six codon types encoded leucine (Leu), serine (Ser), and arginine (Arg); these amino acids had maximum codons. However, one codon type encoded methionine (Met) and tryptophan (Try), and these amino acids had the least number of codons. Leucine was the most frequently coded amino acid (*I. tectorum*, 2696, 10.24%; *I. japonica*, 2661, 10.17%; *I. dichotoma*, 2692, 10.22%; and *I. domestica*, 2692, 10.22%), whereas cysteine (Cys) was the least coded (*I. tectorum*, 305, 1.16%; *I. japonica*, 303, 1.16%; *I. dichotoma*, 304, 1.15%; and *I. domestica*, 305, 1.16%).

Furthermore, the RSCU value was measured to determine nonuniform synonymous codon usage [51]. Most codons demonstrated preferences except for AUG (Met) and UGG (Try), which had RSCU values of 1. RSCU analysis revealed the presence of A or U at the third position of the preferred synonymous codons in the four *Iris* species. Other than the UGA stop codon, the CUA of leucine, and the AUA of isoleucine (Ile), the codons with A or U at the third position had RSCU values greater than 1, indicating

the preferential usage of A or U. The RSCU values of the UUA of leucine were 1.84, 1.83, 1.86, and 1.86, in the CP genomes of *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*, respectively. Similarly, the RSCU values of the AGA of arginine (Arg) were 1.88, 1.83, 1.83, and 1.83, and those of the GCU of alanine (Ala) were 1.79, 1.81, 1.81, and 1.81 in *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*, respectively (Table S3). Thus, the preferential codon usage patterns were similar among these four species, which was probably due to the codon usage bias toward A/T. These similarities in codon choice also reveal the related relationship in the four species. The observed codon pattern is consistent with the CP genomes of *Amomum* [68], *Panax* [69], *Dipterygium* and *Cleome* [70], and various other species [71–73].

**3.3. SSR and Long Repeat Sequences.** CP SSRs have been used as molecular markers in species authentication, population genetics, and phylogeny analysis owing to their high substitution rates [74–76]. A total of 59, 42, 58, and 56 SSRs were detected in the CP genomes of *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*, respectively (Table 3 and

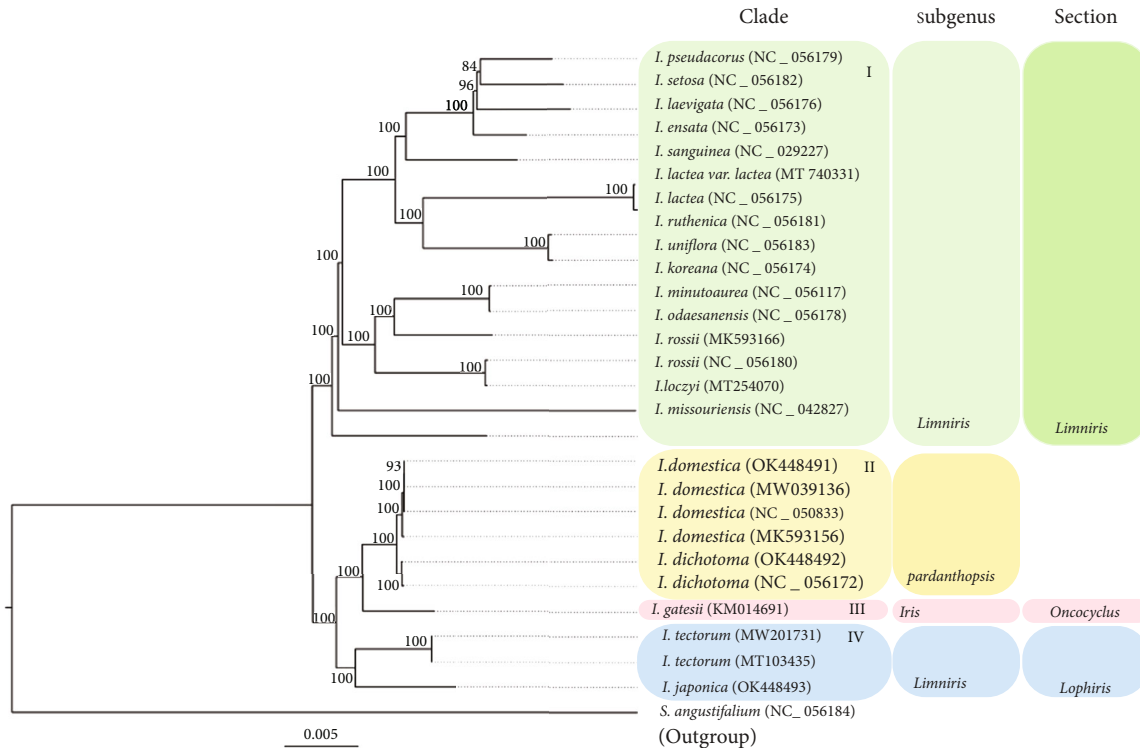


FIGURE 8: ML tree constructed based on the complete CP genomes of 26 *Iris* species and *S. angustifolium* (outgroup). Bootstrap support value is shown at each node.

Table S4), including 38, 22, 35, and 33 mononucleotide SSRs; 11, 10, 13, and 12 dinucleotide SSRs; 4, 4, 3, and 3 trinucleotide SSRs; 3, 4, 4, and 4 tetranucleotide SSRs; 3, 1, 2, and 3 pentanucleotide SSRs; and 0, 1, 1, and 1 hexanucleotide SSRs, respectively (Table S4 and Figure 3). The mononucleotide repeats of *I. tectorum* and *I. japonica* had no C/G type. All four species had one AACTT/AAGTT pentanucleotide repeat. Additionally, an AAAAT/ATTTT pentanucleotide repeat was present in *I. tectorum* and *I. domestica*, whereas none was seen in *I. japonica* and *I. dichotoma*. Moreover, *I. tectorum*, *I. dichotoma*, and *I. domestica* had one specific pentanucleotide (AAAAC/GTTTT, ACTAT/AGTAT, and AATAT/ATATT, respectively). The hexanucleotide repeat (ACAAG/CTTGTT) was found in all species except *I. tectorum* (Table 3). The analysis uncovered that A/T mononucleotide repeats were mostly SSRs and account for 100.0% in *I. tectorum* and *I. japonica*, 97.1% in *I. dichotoma*, and 97.0% in *I. domestica*. Moreover, A or T base was the most frequent in the SSRs, which is similar to the base preference observed in the CP genomes of *Symplocos* [77], *Achnatherum* [78], and other species [79, 80]. These previous studies were all researched between close taxa. Therefore, the SSRs identified in this study might address the relationship among closely related *Iris* species.

Long repeat sequences (F, P, R, and C types) are  $\geq 30$  bp long sequences and are generally located in the IGS and intron; these repeat sequences are responsible for CP genome rearrangement and genetic diversity in populations and used as sources to uncover phylogeny relationships

[81, 82]. The current study analyzed the number of long repeats within *Iris* species (Figure 4). A total of 38, 34, 43, and 67 long repeats were identified in *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*, respectively. Most of the long repeats were F and P types, accounting for 97.37% in *I. tectorum*, 100.00% in *I. japonica*, 88.37% in *I. dichotoma*, and 77.61% in *I. domestica*. The 30–39 bp long F and P types were the majority in the *Iris* species: >50% for *I. tectorum*, *I. japonica*, and *I. domestica* and 44% for *I. dichotoma*. Moreover, the repeats with  $\geq 70$  bp were all F and P types. None of the species had a C repeat, and *I. japonica* had no R repeat. In addition, *I. tectorum*, *I. dichotoma*, and *I. domestica* had 1, 5, and 15 R types, respectively. The distribution of repeats in the *Iris* species was similar to that of *Camellia* [83], *Saraca* [84], and various other species [85–87]. These repeats, one of the CP genome's various origins, are used in elucidating the phylogeny relationships of *Iris* species.

**3.4. Inverted Repeat Expansion and Contraction.** The comparison of boundaries in the CP genomes from *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica* revealed highly conserved LSC/IR/SSC conjunctural regions in the four species; however, variations were detected in the *rps19*, *ndhF*, and *ycf1* genes (Figure 5). The *rps19* gene was located 45, 34, and 45 bp away from the LSC/IRb boundary in *I. tectorum*, *I. dichotoma*, and *I. domestica*, respectively. In *I. japonica*, the *rps19* gene extended into the IRb region (72 bp), creating the *rps19* pseudogene in the IRa region. The *ndhF* gene crossed the SSC/IRb boundary in all species.



Moreover, the *ycf1* gene was located in the SSC/IRa boundary, resulting in a *pseudogene* 895 bp long in *I. tectorum*, 892 bp in *I. japonica*, and 893 bp in *I. domestica* and *I. dichotoma* in the IRb region. These observations suggest that the incomplete duplications at the boundaries probably knocked down the coding potential of the *rps19* gene in *I. japonica* and the *ycf1* gene in all four *Iris* species; these expansions in IR boundaries are consistent with those in *Passiflora* [88], *Lagerstroemia* [89], and various other species [90, 91]. Divergence variations due to IR expansion among interspecies will help distinguish closely related *Iris* species.

**3.5. Identification of Highly Variable Regions.** The complete CP genomes of the four *Iris* species were compared by using the mVISTA [57] program with those available sequences of *I. tectorum* (MT103435), *I. dichotoma* (NC\_056172), *I. domestica* (MW039136), *I. domestica* (NC\_050833), and *I. domestica* (MK593156) downloaded from GenBank. The annotated genome sequence of *I. tectorum* (MW201731) was used as the reference (Figure 6). *I. domestica* had the biggest genome (153,736 bp), and *I. japonica* had the smallest genome (152,443 bp). The reference *I. tectorum* genome (153,253 bp) was the third in size. The coding regions had less divergence than the noncoding sequence regions owing to the variable regions [92–94]. The IR regions were more conserved, whereas the LSC and SSC regions were more divergent.

Furthermore, the average Pi values [95, 96] were calculated separately for the shared genes and IGS to compare the DNA polymorphisms and identify the highly variable regions (Figure 7). The average Pi value of the gene regions was 0.00733 (Figure 7(a)), and that of the IGSs was 0.01629 (Figure 7(b)). LSC and SSC were higher than the IR regions in Pi values, similar to other plants, such as *Handroanthus* [97], *Speirantha* [98], and Combretaceae [99]. Consistent with earlier reports on other species, 13 mutational hotspots and highly divergent loci were examined in the SSC and LSC regions (Pi > 0.03 for IGS and Pi > 0.015 for gene regions), which is helpful for species authentication. The most remarkable divergent loci were *trnG-UCC-trnR-UCU* (Pi = 0.10078) and *rpl16* (Pi = 0.0178) in the IGS and gene regions, respectively. Finally, the combination of the mVISTA plots (divergent regions indicated in white) and the Pi values screened two IGSs, *ndhF-rpl32* (Figure 7(b), 11) and *rps15-ycf1* (Figure 7(b), 13), and the *rpl16* gene (Figure 7(a), 4). These regions with large white plots and high Pi values will serve as potential DNA barcodes for *Iris* species authentication.

**3.6. Phylogenetic Analysis.** CP genomes have been used to determine evolutionary relationships [100–104]. In the present study, a ML tree was constructed using 27 whole CP genome sequences to determine the evolutionary relationships of *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica* with *S. angustifolium* as the outgroup (Figure 8). The phylogenetic analysis revealed the relationships between *I. tectorum* and *I. japonica* and between *I. domestica* and *I. dichotoma*. Subg. *Limniris* was divided into two clades: I (sect. *Limniris*) and IV (sect. *Lophiris*). Here, sect. *Limniris*

showed a sister relationship with three clades, comprising subg. *Pardanthopsis* (clade II), subg. *Iris* (clade III), and sect. *Lophiris* (clade IV), including *I. tectorum* and *I. japonica*. These three monophyletic clades (clades I, II, and IV) were highly supported (bootstrap 100%). Moreover, subg. *Pardanthopsis* was a sister to subg. *Iris*, including *I. gatesii* of sect. *Oncocyclus* (bootstrap value of 100%); *I. domestica* and *I. dichotoma* in clade II were closely related sister species. Additionally, *I. domestica* (OK448491, *B. chinensis*) was clustered with the other three *I. domestica* sequences. This finding was consistent with the findings of Goldblatt and Mabberley [18], Mavrodiev et al. [105], and Wilson [28] who indicated that *B. chinensis* is a synonym of *I. domestica*. In addition, two *I. dichotoma* sequences (previous and present) were clustered into a branch, similar to the two sequences of *I. tectorum*. These results mutually corroborated the accuracy of the sequences. Notably, the four species were separated into distinct groups. Thus, for the first time, the present study deduced the relationship among the four *Iris* species based on complete CP genomes following the ML method. These results are consistent with the molecular phylogeny by Wilson [28], Guo and Wilson [11], Kang et al. [26], and Xiao et al. [106] based on different plastid fragments. Thus, the phylogenetic analysis uncovers that the CP genomes could be used to verify the subdivisions of *Iris* species, especially at the subgenus and section ranks.

The ML tree based on common protein-coding sequences (Figure S4) was similar to that based on the complete CP genomes (Figure 8), except for two branches, i.e., branch of *I. pseudacorus*, *I. setosa*, *I. laevigata*, and *I. ensata* species and branch of *I. domestica* and *I. dichotoma* species. In detail, *I. ensata*, in both trees, was the most primitive taxon among four species, but the *I. pseudacorus*, *I. setosa*, and *I. laevigata* demonstrated different relationships in these two trees. Meanwhile, *I. domestica* could be distinguished from *I. dichotoma* in the tree based on the complete chloroplast genomes, but the tree based on common protein-coding sequences could not differentiate *I. domestica* from *I. dichotoma*. The complete chloroplast genome has been commonly used as superbarcoding for species identification in researches, such as *Dipterygium* and *Cleome* [70] and *Zantedeschia* [91]. In the present study, the result of species authentication based on complete CP genomes among four medicinal *Iris* species also proved the efficacy of superbarcoding. The usage of complete CP genomes was more efficient than the usage of common protein-coding sequences for *Iris* species identification, probably derived from more variant regions contained in intergenic regions of the complete chloroplast genome [98, 104].

## 4. Conclusions

The present research sequenced and analyzed the complete CP genomes of four *Iris* species, namely, *I. tectorum*, *I. dichotoma*, *I. japonica*, and *I. domestica*. CP genome sizes, GC contents, codon usages, SSRs, and long repeats were examined, and the genome conservation and differences among the four *Iris* species were compared. Furthermore, comparing these species' genomes with other Iridaceae

species revealed a few variable regions; however, the use of these markers in DNA barcoding needs to be tested. The study also generated an ML phylogenetic tree that depicted the evolutionary relationship of *Iris* species and confirmed that *B. chinensis* is a synonym of *I. domestica*; however, the whole CP genomes of the 13 taxa of sect. *Lophiris* need to be included in one robust phylogenetic analysis. The study's findings confirm that CP genomes are a worthy genetic resource for identifying Iridaceae species and analyzing their phylogeny.

## Abbreviations

CP:	Chloroplast
CDS:	Protein-coding genes
SSR:	Simple sequence repeat
Pi:	Nucleotide diversity
subg:	Subgenera
sect:	Section
SSC:	Small single copy
LSC:	Large single copy
IR:	Inverted repeat
NCBI:	National Center for Biotechnology Information
RSCU:	Relative synonymous codon usage
ML:	Maximum likelihood
IGS:	Intergenic spacers.

## Data Availability

The data supporting the study's findings are publicly available in NCBI under the accession numbers MW201731, OK448491, OK448492, and OK448493. The associated data are available in Sequence Read Archive (SRA) under the BioSample, BioProject, and SRA numbers of *Iris tectorum* (SAMN17169715, PRJNA688136, and SRR13311445), *Iris domestica* (SAMN25087045, PRJNA798580, and SRR17692213), *Iris dichotoma* (SAMN25087046, PRJNA798580, and SRR17692212), and *Iris japonica* (SAMN25087047, PRJNA798580, and SRR17692211). The sequence data are available from <https://dataview.ncbi.nlm.nih.gov/object/SRR13311445>, <https://dataview.ncbi.nlm.nih.gov/object/SRR17692213>, <https://dataview.ncbi.nlm.nih.gov/object/SRR17692212> and <https://dataview.ncbi.nlm.nih.gov/object/SRR17692211>. The accession numbers of others used in the present study are shown in Table S5, and these were released from NCBI.

## Conflicts of Interest

The authors report no conflict of interest.

## Authors' Contributions

Project conception was realized by Yu-lin Lin and Hui Yao. Experiment design and data analysis were conducted by Jing-lu Feng. Plant material collection and identification were done by Bao-li Li and Yu-lin Lin, respectively. Experiment was performed by Jing-lu Feng and Yun-jia Pan. Bioinformatic analysis was carried out by Li-wei Wu and Qing

Wang. Manuscript draft was prepared by Jing-lu Feng. All authors approved the manuscript. Jing-lu Feng and Li-wei Wu contributed equally to this work and share the first authorship.

## Acknowledgments

This work was supported by the National Science & Technology Fundamental Resources Investigation Program of China (grant number 2018FY100700).

## Supplementary Materials

Figure S1: CP genome map of *Iris japonica*. Figure S2: CP genome map of *Iris dichotoma*. Figure S3: CP genome map of *Iris domestica*. Figure S4: ML tree constructed based on common protein-coding genes of 26 *Iris* species and *S. angustifolium* (outgroup). Bootstrap support value is shown at each node. Table S1: gene content and gene order in the chloroplast genomes of four *Iris* species. Table S2: gene catalog of *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*. Table S3: codon usage of four *Iris* species. Table S4: simple sequence repeats in the complete chloroplast genomes of *I. tectorum*, *I. japonica*, *I. dichotoma*, and *I. domestica*. Table S5: GenBank accession numbers of the complete chloroplast genome sequences used in the phylogenetic analysis. (*Supplementary Materials*)

## References

- [1] C. Linnaeus, *Species plantarum*, Ray Society, London, 1957.
- [2] B. Mathew, *The Iris*, Batsford, London, 1989.
- [3] B I S S Group, *A Guide to Species Irises: Their Identification and Cultivation*, Cambridge University Press, 1997.
- [4] P. Goldblatt and J. C. Manning, *The Iris Family: Natural History & Classification*, Timber Press, Portland, 2008.
- [5] A. Amin, S. H. Wani, T. A. Mokhdomi et al., "Investigating the pharmacological potential of *Iris kashmiriana* in limiting growth of epithelial tumors," *Pharmacognosy Journal*, vol. 5, no. 4, pp. 170–175, 2013.
- [6] A. Mocan, G. Zengin, A. Mollica et al., "Biological effects and chemical characterization of *Iris schachtii* Markgr. extracts: a new source of bioactive constituents," *Food and Chemical Toxicology*, vol. 112, pp. 448–457, 2018.
- [7] A. Jalsrai, A. Reinhold, and A. Becker, "Ethanol *Iris tenuifolia* extract reduces brain damage in a mouse model of cerebral ischaemia," *Phytotherapy Research*, vol. 32, no. 2, pp. 333–339, 2018.
- [8] N. P. Committee, *Pharmacopoeia of People's Republic of China*, China Medical Science Press, Beijing, 2020.
- [9] C. A. Wilson, J. Padiernos, and Y. Sapir, "The royal irises (*Iris* subg. *Iris* sect. *Oncocyclus*): plastid and low-copy nuclear data contribute to an understanding of their phylogenetic relationships," *Taxon*, vol. 65, no. 1, pp. 35–46, 2016.
- [10] X. Lian, G. Luo, H. Li, W. Xu, Y. Xiao, and X. Bi, "Reciprocal difference of interspecific hybridization between three different colours of *Iris dichotoma* and *I. domestica*," *The Journal of Horticultural Science and Biotechnology*, vol. 91, no. 5, pp. 483–490, 2016.

- [11] J. Guo and C. A. Wilson, "Molecular phylogeny of crested Iris based on five plastid markers (Iridaceae)," *Systematic Botany*, vol. 38, no. 4, pp. 987–995, 2013.
- [12] W. R. Dykes, *The Genus Iris*, Dover Publications, New York, 1974.
- [13] G. H. M. Lawrence, "A reclassification of the genus Iris," *Genes Herbarum*, vol. 8, pp. 346–371, 1953.
- [14] G. I. Rodionenko, *The Genus Iris L.: (Questions of Morphology, Biology Evolution and Systematics)*, Mitchell Beazley, 1987.
- [15] N. Tillie, M. W. Chase, and T. Hall, "Molecular studies in the genus Iris L.: a preliminary study," *Annali di Botanica*, vol. 58, 2000.
- [16] C. A. Wilson, "Phylogeny of Iris based on chloroplast matK gene and trnK intron sequence data," *Molecular Phylogenetics and Evolution*, vol. 33, no. 2, pp. 402–412, 2004.
- [17] C. A. Wilson, "Patterns in evolution in characters that define Iris subgenera and sections," *Aliso: A Journal of Systematic Evolutionary Botany*, vol. 22, no. 1, pp. 425–433, 2006.
- [18] P. Goldblatt and D. J. Mabberley, "Belamcanda included in Iris, and the new combination *I. domestica* (Iridaceae: Iridaceae)," *Novon*, vol. 15, no. 1, pp. 128–132, 2005.
- [19] E. Anderson, "The species problem in Iris," *Annals of the Missouri Botanical Garden*, vol. 23, no. 3, pp. 457–509, 1936.
- [20] Y. W. Park, D. M. Kim, Y. J. Hwang, K. B. Lim, and H. H. Kim, "Karyotype analysis of three Korean native Iris species," *Horticulture, Environment and Biotechnology*, vol. 47, no. 1, pp. 51–54, 2006.
- [21] E. Maugini and L. B. Maleci, "Further investigations on the karyotype of several 40 chromosome dwarf bearded irises (*I. chamaeiridis* Bertol. S.L.)," *Caryologia*, vol. 27, no. 1, pp. 117–127, 1974.
- [22] X. F. Yu, H. Q. Zhang, M. Yuan, and Y. H. Zhou, "Karyotype studies on ten Iris species (Iridaceae) from Sichuan, China," *Caryologia*, vol. 62, no. 3, pp. 253–260, 2009.
- [23] S. Jozghasemi, V. Rabiei, A. Soleymani, and A. Khalighi, "Karyotype analysis of seven Iris species native to Iran," *Caryologia*, vol. 69, no. 4, pp. 351–361, 2016.
- [24] D. O. Castro, D. E. Guacchio, D. E. Iorio et al., "Barcoding helps threatened species: the case of *Iris marsica* (Iridaceae) from the protected areas of the Abruzzo (Central Italy)," *Plant Biosystems*, vol. 154, no. 6, pp. 961–972, 2020.
- [25] Y. Lu, S. Liu, B. Ju, and J. Wang, "Genetic analysis of the *Blechnum* *Rhizoma* and its common adulterants using internal transcribed spacer 2 molecular barcoding," *European J Med Plants*, vol. 20, no. 2, pp. 1–8, 2017.
- [26] Y. J. Kang, S. Kim, J. Lee, H. Won, G. H. Nam, and M. Kwak, "Identification of plastid genomic regions inferring species identity from de novo plastid genome assembly of 14 Korean-native Iris species (Iridaceae)," *PLoS One*, vol. 15, no. 10, article e0241178, 2020.
- [27] M. Li, H. Cao, P. P. H. But, and P. C. Shaw, "Identification of herbal medicinal materials using DNA barcodes," *Journal of Systematics and Evolution*, vol. 49, no. 3, pp. 271–283, 2011.
- [28] C. A. Wilson, "Subgeneric classification in Iris re-examined using chloroplast sequence data," *Taxon*, vol. 60, no. 1, pp. 27–35, 2011.
- [29] C. A. Wilson, "Sectional relationships in the Eurasian bearded Iris (subgen. Iris) based on phylogenetic analyses of sequence data," *Systematic Botany*, vol. 42, no. 3, pp. 392–401, 2017.
- [30] C. A. Wilson, "Two new species in Iris series Chinenses (Iridaceae) from south-central China," *PhytoKeys*, vol. 161, pp. 41–60, 2020.
- [31] R. Liu, Y. Gao, L. Ruan, Z. Fan, and C. Li, "Variation of flower opening and closing times in hybrids of evening flowering species *Iris dichotoma* and daytime flowering species *Iris domestica*," *Plant Breeding*, vol. 137, no. 6, pp. 920–927, 2018.
- [32] J. Guo, "Comparative micromorphology and anatomy of crested sepals in Iris (Iridaceae)," *International Journal of Plant Sciences*, vol. 176, no. 7, pp. 627–642, 2015.
- [33] M. A. C. Bautista, W. Tao, Y. Zheng, Y. Deng, T. Chen, and S. Miao, "Chloroplast genome organization and phylogeny of *Gynochthodes cochinchinensis* (DC.) Razafim. & B. Bremer (Rubiaceae)," *Mitochondrial DNA Part B*, vol. 6, no. 1, pp. 261–262, 2021.
- [34] S. Khayi, F. Gaboun, S. Pirro et al., "Complete chloroplast genome of *Argania spinosa*: structural organization and phylogenetic relationships in Sapotaceae," *Plants (Basel)*, vol. 9, no. 10, p. 1354, 2020.
- [35] H. R. Bai, O. Oyebanji, R. Zhang, and T. S. Yi, "Plastid phylogenomic insights into the evolution of subfamily Dialioideae (Leguminosae)," *Plant Divers*, vol. 43, no. 1, pp. 27–34, 2021.
- [36] D. A. Alzahrani, S. S. Yaradua, E. J. Albokhari, and A. Abba, "Complete chloroplast genome sequence of *Barleria prionitis*, comparative chloroplast genomics and phylogenetic relationships among Acanthoideae," *BMC Genomics*, vol. 21, no. 1, p. 393, 2020.
- [37] A. L. Khan, S. Asaf, A. A.-R. Lubna, and A. Al-Harrasi, "Decoding first complete chloroplast genome of toothbrush tree (*Salvadora persica* L.): insight into genome evolution, sequence divergence and phylogenetic relationship within Brassicales," *BMC Genomics*, vol. 22, no. 1, p. 312, 2021.
- [38] C. A. Wilson, "The complete plastid genome sequence of *Iris gatesii* (section *Oncocycclus*), a bearded species from south-eastern Turkey," *Aliso: A Journal of Systematic Evolutionary Botany*, vol. 32, no. 1, pp. 47–54, 2014.
- [39] H. J. Lee, G. H. Nam, K. Kim, C. E. Lim, J. H. Yeo, and S. Kim, "The complete chloroplast genome sequences of *Iris sanguinea* *donn ex Hornem*," *Mitochondrial DNA A DNA Mapp Seq Anal*, vol. 28, no. 1, pp. 15–16, 2017.
- [40] E. M. Joyce, D. M. Crayn, V. K. Y. Lam, W. K. Gerelle, S. W. Graham, and L. Nauheimer, "Evolution of *Geosiris* (Iridaceae): historical biogeography and plastid-genome evolution in a genus of non-photosynthetic tropical rainforest herbs disjunct across the Indian Ocean," *Australian Systematic Botany*, vol. 31, no. 6, pp. 504–522, 2018.
- [41] T. Y. Choi, S. H. Oh, C. G. Jang, H. W. Kim, A. Kim, and S. R. Lee, "The complete chloroplast genome sequences of the *Iris loczyi* *kanitz* (Iridaceae)," *Mitochondrial DNA B Resour*, vol. 5, no. 3, pp. 2876–2877, 2020.
- [42] Z. Liu, X. Yu, P. Cui, and X. Tian, "The complete chloroplast genome of *Iris tectorum* (Iridaceae)," *Mitochondrial DNA Part B*, vol. 5, no. 2, pp. 1561–1562, 2020.
- [43] X. Cai, B. Zhang, S. Wang, Y. Cheng, and H. Wang, "Characterization and phylogenetic analysis of the chloroplast genome of *Iris lactea* *var. chinensis*," *Mitochondrial DNA B Resour*, vol. 6, no. 4, pp. 1490–1491, 2021.
- [44] J. Feng, Y. Pan, Y. Lin, and H. Yao, "The complete chloroplast genome sequence of *Iris tectorum*," *Mitochondrial DNA B Resour*, vol. 6, no. 12, pp. 3331–3332, 2021.

- [45] C. Li, S. Hu, Y. Ding, G. Bi, C. Su, and Z. Xia, "The complete chloroplast genome of Chinese medicinal herb *Belamcanda chinensis* (L.) Redouté (Iridaceae)," *Mitochondrial DNA B Resour*, vol. 6, no. 2, pp. 331–332, 2021.
- [46] S. Chen, Y. Zhou, Y. Chen, and J. Gu, "Fastp: an ultra-fast all-in-one FASTQ preprocessor," *Bioinformatics*, vol. 34, no. 17, pp. i884–i890, 2018.
- [47] J. J. Jin, W. B. Yu, J. B. Yang et al., "GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes," *Genome Biology*, vol. 21, no. 1, p. 241, 2020.
- [48] M. Tillich, P. Lehwark, T. Pellizzer et al., "GeSeq - versatile and accurate annotation of organelle genomes," *Nucleic Acids Research*, vol. 45, no. 1, pp. W6–W11, 2017.
- [49] S. Greiner, P. Lehwark, and R. Bock, "OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes," *Nucleic Acids Research*, vol. 47, no. 1, pp. W59–W64, 2019.
- [50] S. Kumar, G. Stecher, M. Li, C. Nkya, and K. Tamura, "MEGA X: molecular evolutionary genetics analysis across computing platforms," *Molecular Biology and Evolution*, vol. 35, no. 6, pp. 1547–1549, 2018.
- [51] P. M. Sharp and W. H. Li, "The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications," *Nucleic Acids Research*, vol. 15, no. 3, pp. 1281–1295, 1987.
- [52] A. M. Butt, I. Nasrullah, R. Qamar, and Y. Tong, "Evolution of codon usage in Zika virus genomes is host and vector specific," *Emerging Microbes & Infections*, vol. 5, no. 10, article e107, 2016.
- [53] S. Beier, T. Thiel, T. Münch, U. Scholz, and M. Mascher, "MISA-web: a web server for microsatellite prediction," *Bioinformatics*, vol. 33, no. 16, pp. 2583–2585, 2017.
- [54] T. Thiel, W. Michalek, R. K. Varshney, and A. Graner, "Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.)," *Theoretical and Applied Genetics*, vol. 106, no. 3, pp. 411–422, 2003.
- [55] Y. Cui, L. Nie, W. Sun et al., "Comparative and phylogenetic analyses of ginger (*Zingiber officinale*) in the family Zingiberaceae based on the complete chloroplast genome," *Plants (Basel)*, vol. 8, no. 8, p. 283, 2019.
- [56] S. Kurtz, J. V. Choudhuri, E. Ohlebusch, C. Schleiermacher, J. Stoye, and R. Giegerich, "REPuter: the manifold applications of repeat analysis on a genomic scale," *Nucleic Acids Research*, vol. 29, no. 22, pp. 4633–4642, 2001.
- [57] K. A. Frazer, L. Pachter, A. Poliakov, E. M. Rubin, and I. Dubchak, "VISTA: computational tools for comparative genomics," *Nucleic Acids Research*, vol. 32, pp. W273–W279, 2004.
- [58] K. Katoh, J. Rozewicki, and K. D. Yamada, "MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization," *Briefings in Bioinformatics*, vol. 20, no. 4, pp. 1160–1166, 2019.
- [59] J. Rozas, A. Ferrer-Mata, J. C. Sánchez-DelBarrio et al., "DnaSP 6: DNA sequence polymorphism analysis of large data sets," *Molecular Biology and Evolution*, vol. 34, no. 12, pp. 3299–3302, 2017.
- [60] S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermini, "ModelFinder: fast model selection for accurate phylogenetic estimates," *Nature Methods*, vol. 14, no. 6, pp. 587–589, 2017.
- [61] B. Q. Minh, H. A. Schmidt, O. Chernomor et al., "IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era," *Molecular Biology and Evolution*, vol. 37, no. 5, pp. 1530–1534, 2020.
- [62] P. Besse, *Molecular Plant Taxonomy*, Springer, 2021.
- [63] B. S. Gaut, S. V. Muse, and M. T. Clegg, "Relative rates of nucleotide substitution in the chloroplast genome," *Molecular Phylogenetics and Evolution*, vol. 2, no. 2, pp. 89–96, 1993.
- [64] E. M. Golenberg, M. T. Clegg, M. L. Durbin, J. Doebley, and D. P. Ma, "Evolution of a noncoding region of the chloroplast genome," *Molecular Phylogenetics and Evolution*, vol. 2, no. 1, pp. 52–64, 1993.
- [65] W. Tan, H. Gao, H. Zhang et al., "The complete chloroplast genome of Chinese medicine (*Psoralea corylifolia*): molecular structures, barcoding and phylogenetic analysis," *Plant Gene*, vol. 21, article 100216, 2020.
- [66] T. Wang, R. P. Kuang, X. H. Wang et al., "Complete chloroplast genome sequence of *Fortunella venosa* (Champ. ex Benth.) C.C.Huang (Rutaceae): comparative analysis, phylogenetic relationships, and robust support for its status as an independent species," *Forests*, vol. 12, no. 8, p. 996, 2021.
- [67] L. Wu, L. Nie, Q. Wang et al., "Comparative and phylogenetic analyses of the chloroplast genomes of species of Paeoniaceae," *Scientific Reports*, vol. 11, no. 1, p. 14643, 2021.
- [68] L. Yang, C. Feng, M. M. C. Cai, J. H. Chen, and P. Ding, "Complete chloroplast genome sequence of *Amomum villosum* and comparative analysis with other Zingiberaceae plants," *Chinese Herbal Medicines*, vol. 12, no. 4, pp. 375–383, 2020.
- [69] K. J. Kim and H. L. Lee, "Complete chloroplast genome sequences from Korean ginseng (*Panax schinseng* Nees) and comparative analysis of sequence evolution among 17 vascular plants," *DNA Research*, vol. 11, no. 4, pp. 247–261, 2004.
- [70] D. Alzahrani, E. Albokhari, S. Yaradua, and A. Abba, "Complete chloroplast genome sequences of *Dipterygium glaucum* and *Cleome chrysantha* and other Cleomaceae species, comparative analysis and phylogenetic relationships," *Saudi Journal of Biological Sciences*, vol. 28, no. 4, pp. 2476–2490, 2021.
- [71] D. A. Alzahrani, E. J. Albokhari, S. S. Yaradua, and A. Abba, "Comparative analysis of chloroplast genomes of four medicinal Capparaeae species: genome structures, phylogenetic relationships and adaptive evolution," *Plants (Basel)*, vol. 10, no. 6, p. 1229, 2021.
- [72] V. O. Wanga, X. Dong, M. A. Oulo et al., "Complete chloroplast genomes of *Acanthochlamys bracteata* (China) and *Xerophyta* (Africa) (Velloziaceae): comparative genomics and phylogenomic placement," *Frontiers in Plant Science*, vol. 12, article 691833, 2021.
- [73] H. Q. Nguyen, T. N. L. Nguyen, T. N. Doan et al., "Complete chloroplast genome of novel *Adrinandra megaphylla* Hu species: molecular structure, comparative and phylogenetic analysis," *Scientific Reports*, vol. 11, no. 1, p. 11731, 2021.
- [74] K. K. Tiwari, N. J. Thakkar, D. T. Dharajiya et al., "Genome-wide microsatellites in amaranth: development, characterization, and cross-species transferability," *Biotech*, vol. 11, no. 9, p. 395, 2021.
- [75] L. Liu, S. L. Low, S. Sakaguchi et al., "Development of nuclear and chloroplast polymorphic microsatellites for *Crossostephium chinense* (Asteraceae)," *Molecular Biology Reports*, vol. 48, no. 9, pp. 6259–6267, 2021.

- [76] P. Leontaritou, F. N. Lamari, V. Papatziropoulos, and G. Iatrou, "Exploration of genetic, morphological and essential oil variation reveals tools for the authentication and breeding of *Salvia pomifera* subsp. *calycina* (Sm.) Hayek," *Phytochemistry*, vol. 191, article 112900, 2021.
- [77] S. C. Kim, J. W. Lee, and B. K. Choi, "Seven complete chloroplast genomes from *Symplocos*: genome organization and comparative analysis," *Forests*, vol. 12, no. 5, p. 608, 2021.
- [78] X. Wei, X. Li, T. Chen et al., "Complete chloroplast genomes of *Achnatherum inebrians* and comparative analyses with related species from Poaceae," *FEBS Open Bio*, vol. 11, no. 6, pp. 1704–1718, 2021.
- [79] W. Zhang, H. Wang, J. Dong, T. Zhang, and H. Xiao, "Comparative chloroplast genomes and phylogenetic analysis of *Aquilegia*," *Applications in Plant Sciences*, vol. 9, no. 3, article e11412, 2021.
- [80] S. Qian, Y. Zhang, and S. Y. Lee, "Comparative analysis of complete chloroplast genome sequences in *Edgeworthia* (Thymelaeaceae) and new insights into phylogenetic relationships," *Frontiers in Genetics*, vol. 12, article 643552, 2021.
- [81] D. M. Li, G. F. Zhu, Y. C. Xu, Y. J. Ye, and J. M. Liu, "Complete chloroplast genomes of three medicinal *Alpinia* species: genome organization, comparative analyses and phylogenetic relationships in family Zingiberaceae," *Plants (Basel)*, vol. 9, no. 2, p. 286, 2020.
- [82] G. Hu, Y. Wang, Y. Wang, S. Zheng, W. Dong, and N. Dong, "New insight into the phylogeny and taxonomy of cultivated and related species of *Crataegus* in China, based on complete chloroplast genome sequencing," *Horticulturae*, vol. 7, no. 9, p. 301, 2021.
- [83] L. Li, Y. Hu, M. He et al., "Comparative chloroplast genomes: insights into the evolution of the chloroplast genome of *Camellia sinensis* and the phylogeny of *Camellia*," *BMC Genomics*, vol. 22, no. 1, p. 138, 2021.
- [84] M. A. Ali, T. K. Pan, A. B. Gurung et al., "Plastome of *Saraca asoca* (Detarioideae, Fabaceae): annotation, comparison among subfamily and molecular typing," *Saudi Journal of Biological Sciences*, vol. 28, no. 2, pp. 1487–1493, 2021.
- [85] F. Dong, Z. Lin, J. Lin, R. Ming, and W. Zhang, "Chloroplast genome of rambutan and comparative analyses in Sapindaceae," *Plants (Basel)*, vol. 10, no. 2, p. 283, 2021.
- [86] S. Guo, L. Guo, W. Zhao et al., "Complete chloroplast genome sequence and phylogenetic analysis of *Paeonia ostii*," *Molecules*, vol. 23, no. 2, p. 246, 2018.
- [87] Y. Huo, L. Gao, B. Liu et al., "Complete chloroplast genome sequences of four *Allium* species: comparative and phylogenetic analyses," *Scientific Reports*, vol. 9, no. 1, p. 12250, 2019.
- [88] S. O. Rabah, B. Shrestha, N. H. Hajrah et al., "Passiflora plastome sequencing reveals widespread genomic rearrangements," *Journal of Systematics and Evolution*, vol. 57, no. 1, pp. 1–14, 2019.
- [89] G. Zheng, L. Wei, L. Ma, Z. Wu, C. Gu, and K. Chen, "Comparative analyses of chloroplast genomes from 13 *Lagerstroemia* (Lythraceae) species: identification of highly divergent regions and inference of phylogenetic relationships," *Plant Molecular Biology*, vol. 102, no. 6, pp. 659–676, 2020.
- [90] A. M. Darshetkar, S. Maurya, C. Lee et al., "Plastome analysis unveils inverted repeat (IR) expansion and positive selection in sea lavenders (*Limonium*, Plumbaginaceae, Limonioidae, Limoniaceae)," *PhytoKeys*, vol. 175, pp. 89–107, 2021.
- [91] S. He, Y. Yang, Z. Li, X. Wang, Y. Guo, and H. Wu, "Comparative analysis of four *Zantedeschia* chloroplast genomes: expansion and contraction of the IR region, phylogenetic analyses and SSR genetic diversity assessment," *PeerJ*, vol. 8, article e9132, 2020.
- [92] L. Wang, J. Liang, W. Sa, and L. Wang, "Sequencing and comparative analysis of the chloroplast genome of *Ribes odoratum* provide insights for marker development and phylogenetics in *Ribes*," *Physiology and Molecular Biology of Plants*, vol. 27, no. 1, pp. 81–92, 2021.
- [93] C. Luo, Y. Li, R. Budhathoki et al., "Complete chloroplast genomes of *Impatiens cyanantha* and *Impatiens monticola*: insights into genome structures, mutational hotspots, comparative and phylogenetic analysis with its congeneric species," *PLoS One*, vol. 16, no. 4, article e0248182, 2021.
- [94] L. Gu, T. Su, M. T. An, and G. X. Hu, "The complete chloroplast genome of the vulnerable *Oreocharis esquirolii* (Gesneriaceae): structural features, comparative and phylogenetic analysis," *Plants (Basel)*, vol. 9, no. 12, p. 1692, 2020.
- [95] B. Loeuille, V. Thode, C. Siniscalchi, S. Andrade, M. Rossi, and J. R. Pirani, "Extremely low nucleotide diversity among thirty-six new chloroplast genome sequences from *Aldama* (Heliantheae, Asteraceae) and comparative chloroplast genomics analyses with closely related genera," *PeerJ*, vol. 9, article e10886, 2021.
- [96] S. J. Choi, Y. Kim, and C. Choi, "Chloroplast genome-based hypervariable markers for rapid authentication of six Korean *Pyropia* species," *Diversity*, vol. 11, no. 12, p. 220, 2019.
- [97] M. B. Sobreiro, L. D. Vieira, R. Nunes et al., "Chloroplast genome assembly of *Handroanthus impetiginosus*: comparative analysis and molecular evolution in Bignoniaceae," *Planta*, vol. 252, no. 5, p. 91, 2020.
- [98] G. Raman and S. Park, "The complete chloroplast genome sequence of the *Speirantha gardenii*: comparative and adaptive evolutionary analysis," *Agronomy*, vol. 10, no. 9, p. 1405, 2020.
- [99] Y. Zhang, H. L. Li, J. D. Zhong, Y. Wang, and C. C. Yuan, "Chloroplast genome sequences and comparative analyses of Combretaceae mangroves with related species," *BioMed Research International*, vol. 2020, 2020.
- [100] Y. Cui, X. Chen, L. Nie et al., "Comparison and phylogenetic analysis of chloroplast genomes of three medicinal and edible *Amomum* species," *International Journal of Molecular Sciences*, vol. 20, no. 16, p. 4040, 2019.
- [101] X. Chen, Y. Cui, L. Nie et al., "Identification and phylogenetic analysis of the complete chloroplast genomes of three ephedra herbs containing ephedrine," *BioMed Research International*, vol. 2019, 2019.
- [102] P. W. Schafran, E. A. Zimmer, W. C. Taylor, and L. J. Musselman, "A whole chloroplast genome phylogeny of diploid species of *Isoetes* (Isoëtaceae, Lycopodiophyta) in the southeastern United States," *Castanea*, vol. 83, no. 2, pp. 224–235, 2018.
- [103] K. M. Alanazi, M. A. Ali, S. Y. Kim et al., "The cp genome characterization of *Adenium obesum*: gene content, repeat organization and phylogeny," *Saudi Journal of Biological Sciences*, vol. 28, no. 7, pp. 3768–3775, 2021.
- [104] Y. Zhang, Z. Wang, Y. Guo, S. Chen, X. Xu, and R. Wang, "Complete chloroplast genomes of *Leptodermis scabrida* complex: comparative genomic analyses and phylogenetic relationships," *Gene*, vol. 791, article 145715, 2021.

- [105] E. V. Mavrodiev, M. Martínez-Azorín, P. Dranishnikov, and M. B. Crespo, "At least 23 genera instead of one: the case of *Iris* L. s.l. (Iridaceae)," *PLoS One*, vol. 9, no. 8, article e106459, 2014.
- [106] Y. E. Xiao, F. Y. Yu, and X. F. Zhou, "A new natural hybrid of *Iris* (Iridaceae) from Chongqing, China," *PhytoKeys*, vol. 174, pp. 1–12, 2021.