Scientific and Social Research, 2024, Volume 6, Issue 5
https://ojs.bbwpublisher.com/index.php/SSR

Comparative Analysis of Chloroplast Genome Variation and Evolutionary Analysis of the Endangered Species *Dendrobium flexicaule* and its Relative Species

Le Wang[1,2], Xue Liu[2], Jie Zhang[2,3], Xingjia Ming[2]*, Shu Shu[2], Xianyou Qu[2], Jiayi Zhang[2], Junsheng Qi[1]*

[1]College of Life Science and Food Engineering, Chongqing Three Gorges University, Wanzhou, Chongqing, 404000, China

[2]Chongqing Institute of Traditional Chinese Medicine, Nanan, Chongqing 400065, China

[3]School of Pharmacy, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

Corresponding authors: Xingjia Ming, xjming123@sina.com; Junsheng Qi, grut0518@163.com

Abstract: Globally, the genus *Dendrobium* (Family: *Orchidaceae*) exhibits widespread distribution, predominantly occupying the tropical and subtropical areas of Asia. Although the chloroplast genomes (cpDNAs) of certain *Dendrobium* species have been documented, the evolutionary relationships among these species remain largely unexplored. In this pioneering investigation, a comprehensive analysis of the complete cpDNAs from *Dendrobium* species is presented, including *Dendrobium flexicaule* along with three closely related taxa (*D. nobile*, *D. officinale*, and *D. huoshanense*). The size of the novel complete cpDNAs ranged from 150,529 to 152,588 bp, and their cpDNA structure exhibits a unique quadrifoliate structure reminiscent of female quadrivalents, including large single copy, small single copy (SSC), and inverted repeat regions. A total of 175 simple sequence repeats were identified in the *D. flexicaule* cpDNA. A few pseudogenes of the NA(D)H dehydrogenase-like (NDH) genes were found in the annotated cpDNAs of the four *Dendrobium* species. Within the shared pseudogenes *ndhF* and *ndhG*, *ndhF* frequently undergoes mutations. Additionally, the phylogenetic analysis showed that *D. flexicaule* and *D. scoriarum* were close sister species. The research findings provide crucial genetic information for the molecular phylogenetic system of *D. flexicaule*.

Keywords: *Dendrobium flexicaule*; Chloroplast genome; Phylogeny

Online publication: June 13, 2024

1. Introduction

Chloroplasts are essential organelles in plants that play a vital role in producing energy for the whole organism as a part of the photosynthetic process [1]. Moreover, chloroplast genomes (cpDNAs) possess abundant genetic information of a self-replication mechanism, are highly conserved, and slightly more variable than nuclear

genomes, which are usually used to investigate the evolutionary history of plants and analyze plant taxonomy and phylogeography [2–4]. Previous studies revealed that most of the cpDNA was highly conserved, with a length ranging from 120 to 160 kb [5–6]. The structure of the cpDNA comprises four parts: a large single copy (LSC), a small single copy (SSC), and a pair of inverted repeats (IRa and IRb) [5,7]. The majority of the protein-coding genes (PCGs) in the cpDNA are involved in photosynthesis and other biochemical pathways [8]. There is partial PCG loss, during migration from the chloroplast to the mitochondrial organelles, or alteration to pseudogenes in the evolution of cpDNA. Noncoding DNA sequences from numerous regions of the cpDNA have provided a significant source of characteristics for phylogenetic studies in seed plants [9–10]. Besides, research suggested several genetic mutations in cpDNA, such as gene or intron sequence loss, divergence owing to the expansion or contraction of partial sequences, as well as size variation and gene rearrangements in IR regions, which makes the cpDNA an effective tool for phylogenetic studies among the plant species [11–14].

Dendrobium flexicaule, an intricate member of the *Dendrobium* genus, is widely distributed in China on wet cliffs at an elevation of 1200–2000 m [15]. However, overcollection of *Dendrobium* species due to its vital economic, medicinal, and nutritional benefits for human beings has destroyed its wild resources. In addition, climate change is impacting *Dendrobium* species survival, and thus, it has been listed as a nationally protected endangered wild plant [16]. Meanwhile, the morphological traits of *D. flexicaule* bear high similarity with *Dendrobium huoshanense*, *Dendrobium officinale*, and *Dendrobium moniliforme*; thus, it is a challenge to differentiate them and ensure their safety and efficacy for medicinal use [17]. Currently, the majority of reports on *D. flexicaule* focus on the basic biological characteristics, tissue culture, DNA molecular evidence, and sterile germination of seeds. For example, Zhang et al. reported that factors, such as temperature, humidity, terrain, and companion plants, influence the cultivation and survival of *D. flexicaule* [18]. Some studies have improved the rapid propagation system of *D. flexicaule* [19]. Studies have reported divergence in polysaccharide composition and D-mannose content to allow the identification of *D. flexicaule* and its relative species [20]. Based on recombinant DNA internal transcribed spacer sequences and morphological features, *D. flexicaule*, and other *Dendrobium* species were accurately identified via sequence alignment [17]. Although several research reports exist on *D. flexicaule*, the genetic information on the cpDNA of *D. flexicaule* remains inadequate. Moreover, compared with other *Dendrobium* species, *D. flexicaule* is distributed at higher altitudes, making it an ideal material for studying the evolution and adaptability of cpDNA in the genus *Dendrobium*.

This study completed the sequencing, assembly, and annotation of the complete cpDNA of the four *Dendrobium* species (*D. flexicaule*, *D. officinale*, *D. nobile*, and *D. huoshanense*) and analyzed basic characteristics, relative synonymous codon usage, IR boundaries, simple sequence repeats (SSRs), nucleotide polymorphism, and nonsynonymous/synonymous substitution rate (Ka/Ks). The primary objectives of this study were as follows: to further differentiate *D. flexicaule* from other *Dendrobium* species and to reconstruct the phylogeny of *D. flexicaule*. This study provides a molecular basis for *Dendrobium* species identification and a valuable resource of genetic history, thereby developing an understanding of the breeding, protection, and basic biological characteristics of *D. flexicaule*.

2. Materials and methods

2.1. Collection of the plant materials and DNA extraction

The plant material of the four *Dendrobium* species (*D. flexicaule*, *D. huoshanense*, *D. nobile*, and *D. officinale*) was acquired from the Chengkou *Dendrobium* species base (31° 95 N, 108° 67 E), Chongqing Province, China. The voucher specimens with accession numbers of the four *Dendrobium* species (*D. flexicaule*, 500229210731001LY; *D. huoshanense*, 500229210731002LY; *D. nobile*, 500229210731003LY;

and *D. officinale*, 500229210731004LY) were deposited at the Chongqing Institute of Traditional Chinese Medicine, Chongqing, China. Fresh leaves were stored in liquid nitrogen at -80°C until further use. The cetyltrimethylammonium bromide method was used to extract the cpDNA of *Dendrobium* species [21]. The quality and concentration of the DNA samples were evaluated using 0.75% agarose gel electrophoresis, a NanoDrop One spectrophotometer (Thermo Fisher Scientific), and a Qubit 3.0 Fluorometer (Life Technologies, Carlsbad, CA, USA).

2.2. Sequencing, assembly, and annotation of the chloroplast genome

A total genome sequencing with paired-end (PE) 150 bp was performed using the Illumina Hiseq-2500 platform to acquire a complete cpDNA sequence. The SOAPnuke v. 1.3.0 was used to filter the raw data and the SPAdes v. 3.13.0 was performed to assemble clean reads. The final sequence was obtained using the Gapcloser v. 1.12 to fill the gap. Then, the annotation of the complete cpDNAs was performed using the PGA program [22]. The annotated results were manually checked and the annotated cpDNA was visualized online using the OrganellarGenomeDRAW (OGDRAW) [23]. The final assembly and annotation of the cpDNA were deposited in NCBI (Accession No: *D. flexicaule*, OQ360111; *D. huoshanense*, OR387325; *D. nobile*, OR387323; and *D. officinale*, OR387324).

2.3. Codon usage and IR boundary analysis

The CodonW software v.1.4.2 with default parameters was used to calculate the codon usage number and obtain the relative synonymous codon usage (RSCU) value for the four *Dendrobium* cpDNAs [24]. An RSCU = 1 indicates that the codon has no preference; an RSCU > 1 indicates that the codon is more frequently used; and an RSCU < 1 indicates the codon is less frequently used [25]. The effective number of codons (ENC) was calculated using the CodonW1.4.2 software. In the ENC plot, the two-dimensional scatter plot was mapped using ENC values as the longitudinal coordinate and guanine and cytosine content at the third codon position (GC3) values as the horizontal coordinate, and these points (or genes) on the standard curve calculated with the formula ENC = 2 + GC3 + 29/[GC32 + (1 − GC3)2] were used to determine the codons influenced by natural selection based on whether the gene falls above or slightly below the standard curve [26]. The CPJSdraw was used to draw the IR boundaries map of the four *Dendrobium* species to visualize the cpDNA [27].

2.4. Pseudogenes of the NDH gene

This study downloaded eight published *Cymbidium* species (*Cymbidium aloifolium*_ON943060, *Cymbidium ensifolium*_KT722983, *Cymbidium faberi*_KR919606, *Cymbidium goeringii*_KT722982, *Cymbidium kanran*_MK848038, *Cymbidium sinense*_MW582689, *Cymbidium tracyanum*_OP142287, and *Cymbidium wenshanense*_MK848057) and seven published *Dendrobium* species (*Dendrobium densiflorum*_MW007719, *Dendrobium longicornu*_MN227146, *Dendrobium moschatum*_OM161978, *Dendrobium sinense*_OM792979, *Dendrobium thyrsiflorum*_MN306203, *Dendrobium williamsonii*_OK173601, and *Dendrobium heterocarpum*_OM049526) from NCBI, with cpDNAs of *Dendrobium heterocarpum*_OM049526 as a reference sequence. The Geneious software was used to select pseudogenes within the annotated cpDNAs of the four *Dendrobium* species. Then, the same approach was applied to select for pseudogenes in the other 13 species. Functional protein-coding genes (PCGs) were aligned using Geneious with the translation align option [28–29]. Then, the variations in these functional PCGs were determined using the Annotate and Predict option in Geneious.

2.5. Simple sequence repeats analysis

MISA software was used to detect SSRs, with the following parameter settings: (1) mononucleotides ≥10; (2)

dinucleotides ≥5; (3) trinucleotides, tetranucleotides, pentanucleotides, and hexanucleotides ≥3. The minimum distance between the two SSR sites was set to 100 bp [30]. MISA (https://webblast.ipk-gatersleben.de/misa/), TRF (https://tandem.bu.edu/trf/trf.unix.help.html), and REPuter network server were used for the analysis of short tandem repeats, tandem repeats, and dispersed repeats, respectively [31–33]. The Origin 2018 software was used to visualize the map [34].

2.6. Comparative and phylogenetic analyses of the chloroplast genome

The cpDNAs of the four *Dendrobium* species were compared using the online comparison tool mVISTA in Shuffle LAGAN mode with default parameters, and the complete cpDNAs of *D. officinale*, *D. flexicaule*, and *D. nobile* were compared with the complete cpDNA of *D. huoshanense* as a reference. DnaSP v.6.0 was used to calculate the nucleotide diversity values (Pi) among the cpDNAs [35]. The Ks and Ka substitution rates for the shared PCGs of the four *Dendrobium* species were calculated using the KaKs_Calculator v.2.0 [36]. The map was visualized using Origin 2018.

The study downloaded the complete cpDNAs of 31 *Dendrobium* species and 11 orchid complete cpDNAs from the NCBI database. Then, the MATTF software with the default parameters was used to conduct the whole-genome sequence alignment with the newly sequenced four *Dendrobium* species cpDNA, and the MEGA-X tool was used to construct the Neighbor-Joining (NJ) phylogenetic tree [37–38]. The resulting NJ tree was visualized using the ITOL software v.4.0 [39].

3. Results

3.1. Basic structural features of the chloroplast genomes

Based on the Illumina sequencing technology, the cpDNAs of the four *Dendrobium* species (*D. flexicaule*, *D. nobile*, *D. huoshanense*, and *D. officinale*) were sequenced and assembled. The cpDNA structure exhibited a typical quadrifoliate structure and the size ranged from 150,602 to 152,588 bp, including large single copy (LSC, 84,781–85,270 bp), small single copy (SSC, 13,799–14,634 bp), and a pair of inverted repeat regions (IR, 51,566–52,684 bp) for the four *Dendrobium* species (**Figure 1**). The GC content ranged from 37% to 38%, which indicated extremely conserved regions among the four *Dendrobium* species (**Table 1**). A total of 493 predicted functional genes were obtained from the cpDNAs of the four *Dendrobium* species, ranging from 75 to 81 in PCGs, with the same number of transfer RNA (tRNA: 38) and ribosomal RNA (rRNA: 8, 16S, 23S, 4.5S, and 5S). The *ndhG* and *ndhF* pseudogenes were present in the cpDNAs of *D. nobile*.

A total of 122 PCGs were identified in the cpDNA of *D. flexicaule*, including 58 PCGs and 19 tRNA genes in the LSC region, six PCGs (*ndhB*, *rpl2*, *rpl23*, *rps7*, *rps19*, and *ycf2*), eight tRNA genes (*trnA-UGC*, *trnE-UCC*, *trnH-GUG*, *trnL-CAA*, *trnM-CAU*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*), four rRNA genes (*rrn4.5*, *rrn5*, *rrn16*, and *rrn23*), and ten PCGs and one tRNA gene were identified in the SSC region. The GC content in the different regions of the chloroplast genome was <50%, which indicated that *D. flexicaule* cpDNA has a tendency for A/T bases and A/T end codons usage.

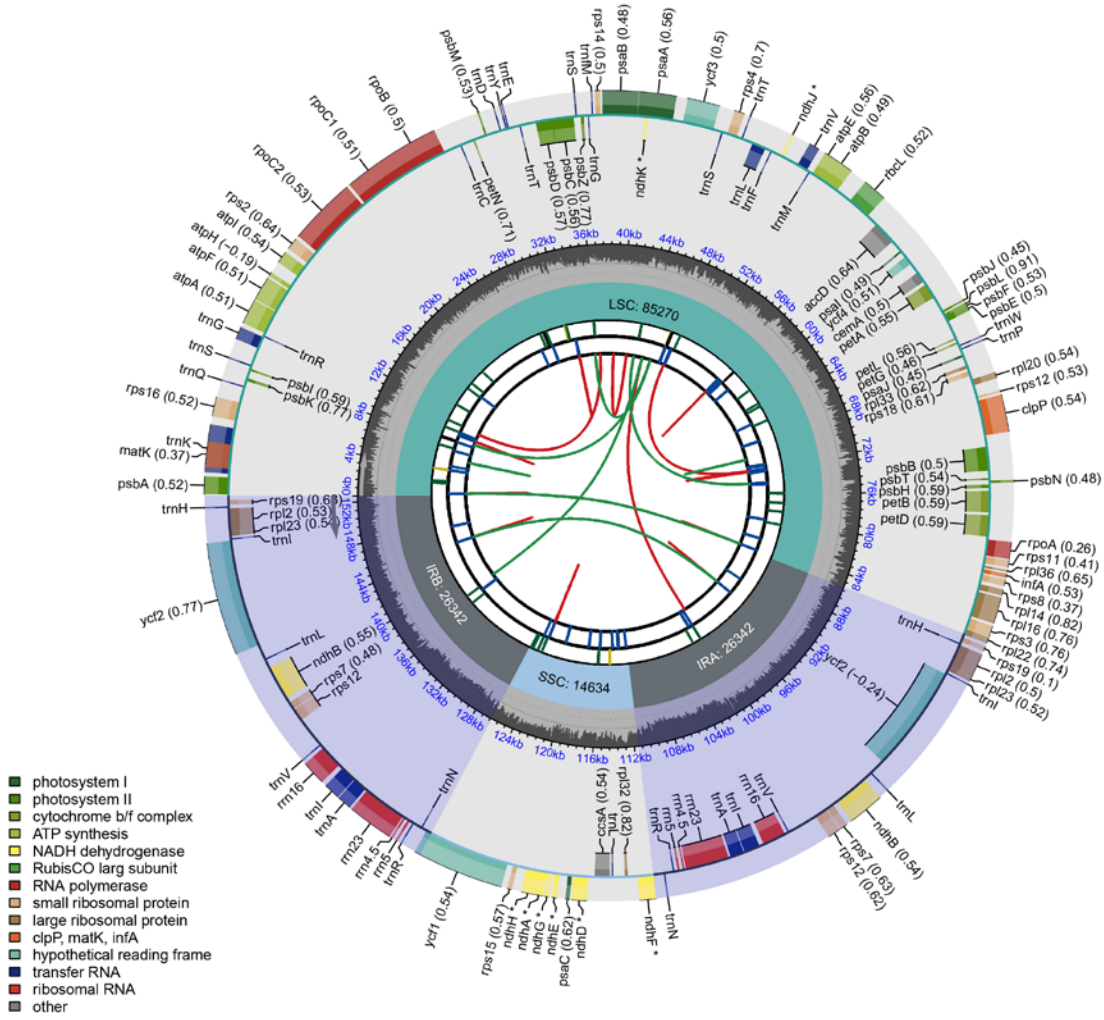

Figure 1. Chloroplast genome map of *D. flexicaule*

Table 1. Comparison of the characteristics of the chloroplast genome structure of the four *Dendrobium* species

Genomic features	*D. flexicaule*	*D. officinale*	*D. huoshanense*	*D. nobile*
Accession number	OQ360111	OR387324	OR387325	OR387323
Total genome length/bp	152,588	152,213	150,529	150,602
LSC region length/bp	85,270	85,146	84,792	84,781
Length of SSC region/bp	14,634	14,449	14,171	13,799
IR region length/bp	52,684	52,618	51,566	52,022
Total number of genes	122	121	123	127
Number of PCGs	76	75	77	81
Number of tRNA genes	38	38	38	38
Number of rRNA genes	8	8	8	8
GC content	37%	37%	38%	38%
GC content in the LSC region	35%	35%	35%	35%
GC content in the SSC region	30%	30%	30%	30%
GC content in the IR region	43%	43%	43%	43%

3.2. Analysis of codon usage bias based on the PCGs

The 76 PCGs of the *D. flexicaule* cpDNA included 50,862 codons. Leucine (Leu, 5,005), arginine (Arg, 3,229), and serine (Ser, 4,784) were the most abundant amino acids, accounting for 9.841.% (Leu), 6.35% (Arg), and 9.4% (Ser) of the codons, whereas tryptophan (Trp) was the rare amino acid, with 703 (1.38%) codons. Moreover, the relative synonymous codon usage (RSCU) value proved codon usage bias (**Figure 2**), with 33 codons having an RSCU > 1, except tryptophan. Most of the amino acids of *D. flexicaule* cpDNA have at least one codon, and Leu has six codons. To further investigate whether codon usage bias occurred among *Dendrobium* species, PCGs from the other three *Dendrobium* species, including *D. nobile* (50,200), *D. officinale* (50,737), and *D. huoshanense* (50,176) were identified. The RSCU values indicated a similar tendency to *D. flexicaule*, which suggested the conservation of codon usage preference in the *Dendrobium* species.

The ENC plots showed the main factors that influence codon usage preference. To determine the role of natural selection and mutation in the evolution of the cpDNA and codon usage pattern, PCGs of the cpDNAs of the four *Dendrobium* species were analyzed (**Figure 3**). The ENC plots indicated that the PCGs of the four *Dendrobium* species clearly showed a similar distribution in the plots. Most of the genes were distributed below the curve, whereas a small portion of the genes were distributed above the standard curve. This result indicated that the pressure of natural selection was the main factor that influenced codon usage preference in the cpDNA evolution of the *Dendrobium* species. Meanwhile, the distribution of a few genes on both sides of the standard curve suggested the involvement of other factors influencing codon usage bias.

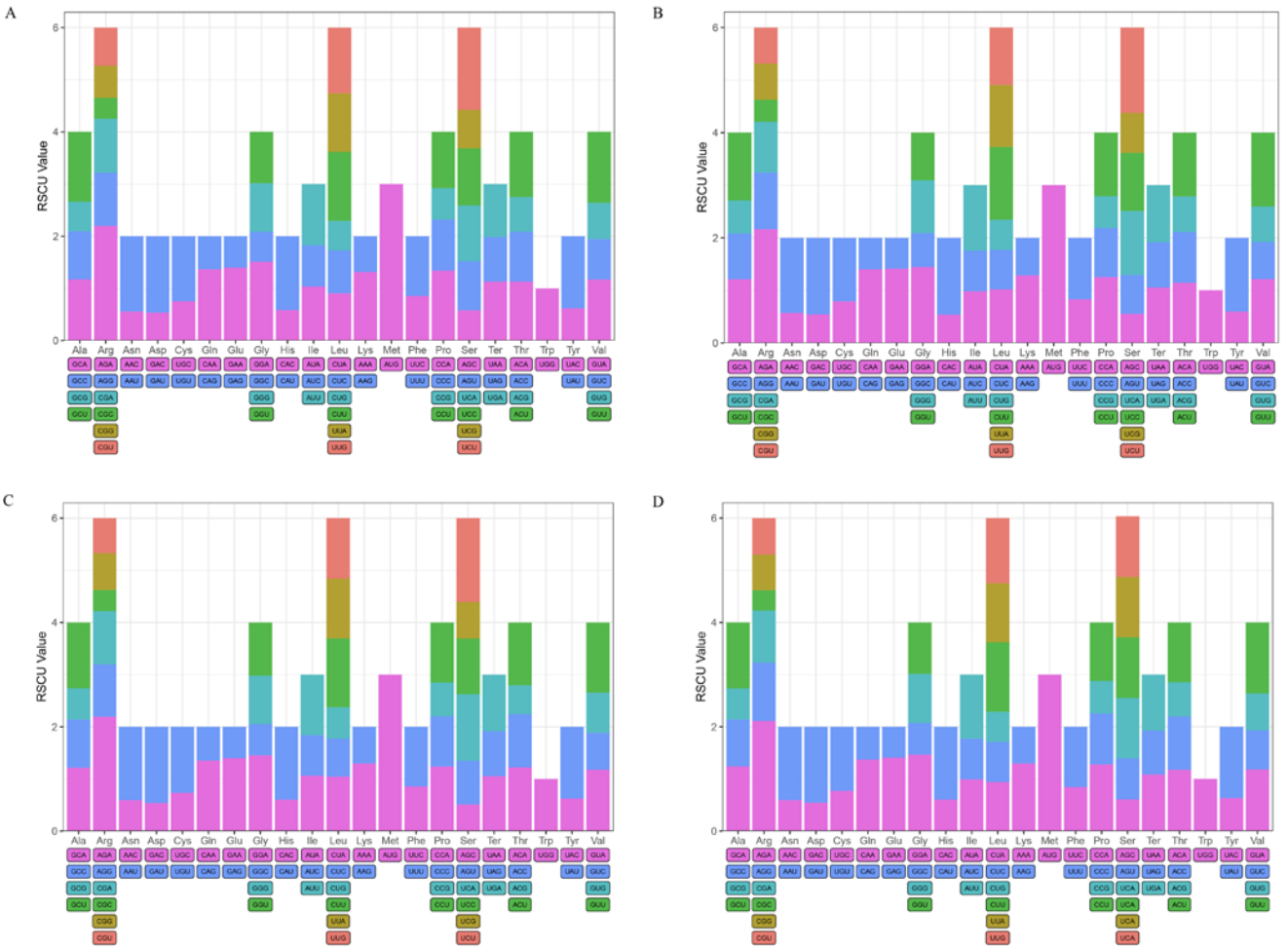

Figure 2. RSCU value map for the *Dendrobium flexicaule* (A), *Dendrobium huoshanense* (B), *Dendrobium nobile* (C) and *Dendrobium officinale* (D)

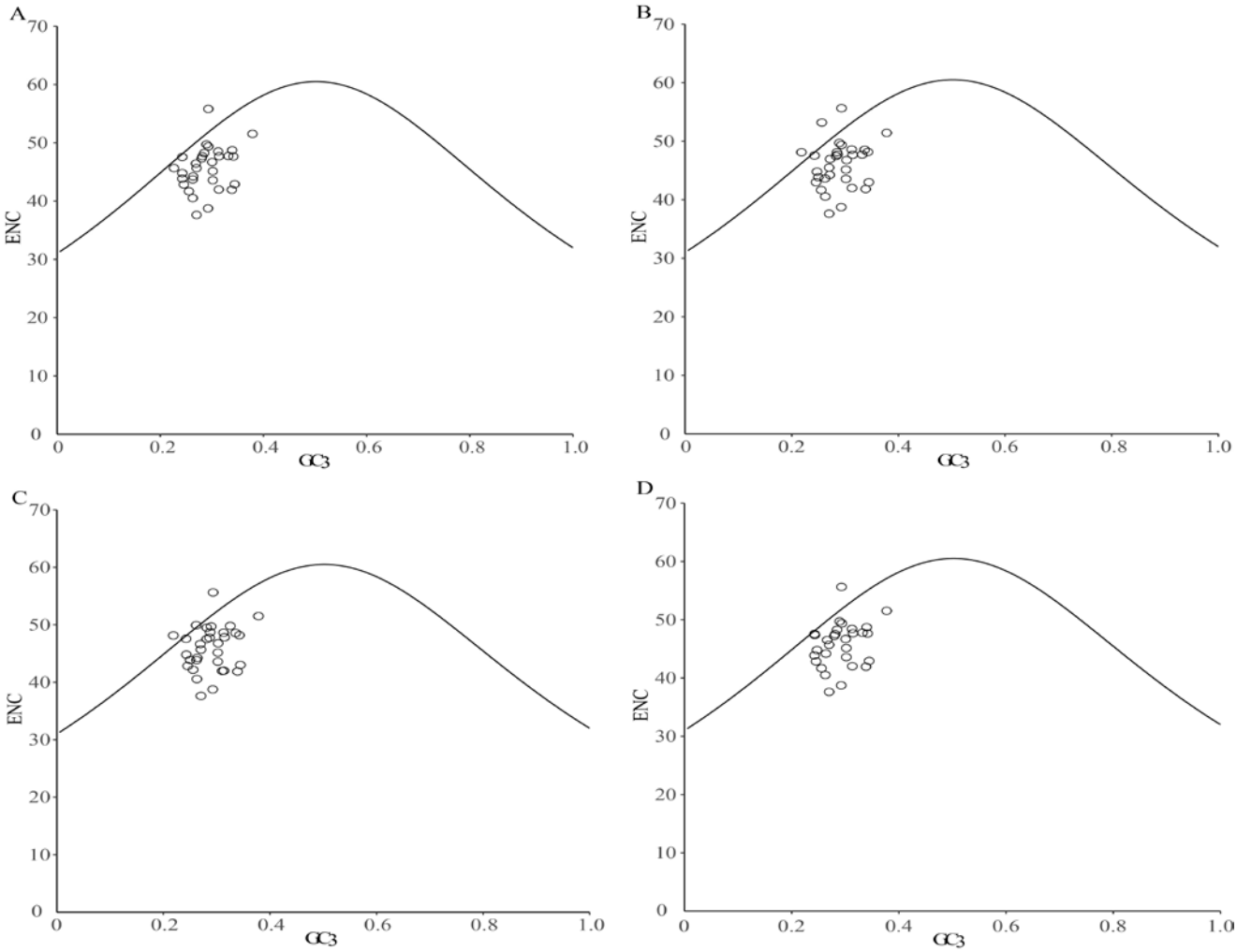

Figure 3. ENC-plot analysis of the *Dendrobium flexicaule* (A), *Dendrobium huoshanense* (B), *Dendrobium nobile* (C), and *Dendrobium officinale* (D); the small circles represent protein-coding genes

3.3. IR boundary expansion and contraction in the *Dendrobium* species

cpDNA was constructed from the LSC/IRb (JLB), IRb/SSC (JSB), SSC/IRa (JSA), and IRa/LSC (JLA) boundaries/junctions of the four *Dendrobium* species, and it was extremely conserved with minor variation among the four *Dendrobium* species (**Figure 4**). The lengths of LSC, SSC, and IRs from the chloroplast genomes of the four *Dendrobium* species were similar. However, the IRb/SSC boundaries of the *D. nobile* cpDNA species were distributed with *rpl32* genes in the SSC region that were different than those in the other three *Dendrobium* species. Further analysis revealed that the *ndhF* gene had become a pseudogene in the evolution of the chloroplast genome of *D. nobile*. *trnN* and *ndhF* were mainly distributed on the two sides of the IRb/SSC boundary lines or on the IRb/SSC boundary lines of the other three *Dendrobium* species. In contrast, the sizes of *ndhF* and *ycf1* genes of the other three *Dendrobium* species underwent varying degrees of expansion and contraction in the IRb/SSC, SSC/Ira, and IRa/*psbA* regions.

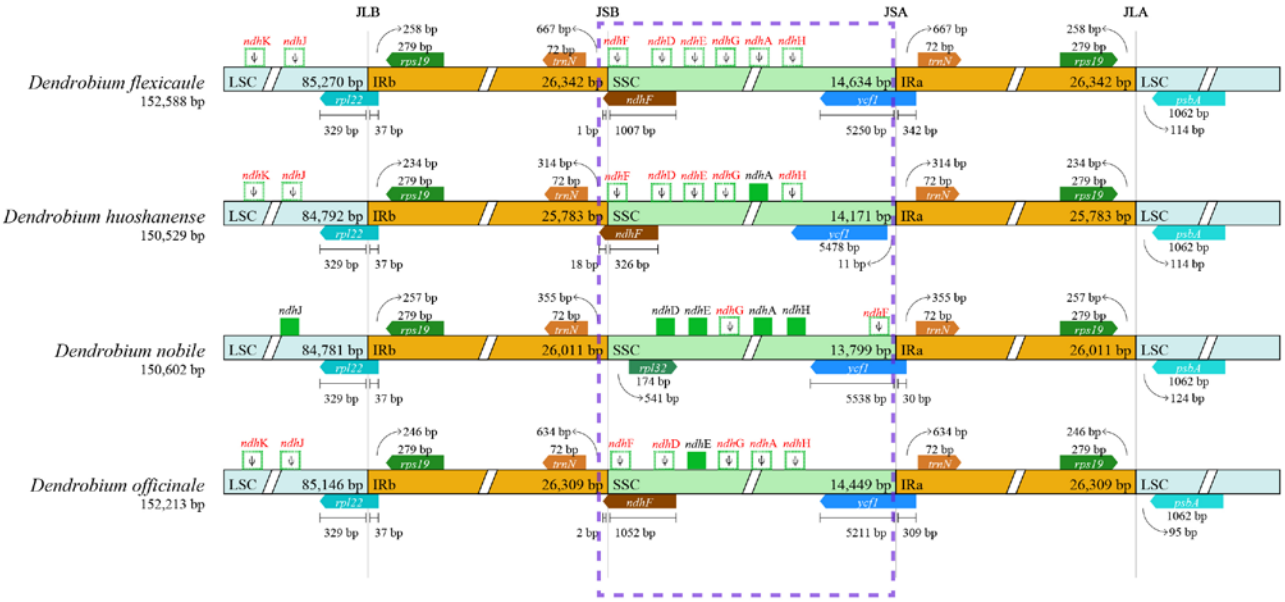

Figure 4. IR boundary expansion and contraction of the four *Dendrobium* chloroplast genomes

3.4. Pseudogenization and evolutionary history of *Dendrobium*

Partial pseudogenes were identified through the annotation of the cpDNAs for the four *Dendrobium* species, including *D. flexicaule* (*ndhA*, *ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhJ*, and *ndhK*), *D. huoshanense* (*ndhD*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhJ*, and *ndhK*), *D. officinale* (*ndhA*, *ndhD*, *ndhF*, *ndhG*, *ndhH*, *ndhJ*, and *ndhK*), and *D. nobile* (*ndhF* and *ndhG*). Among these, the majority of the pseudogenes located in the SSC region of the cpDNA showed great divergence in size according to the IR boundaries **(Figure 4**). In contrast, *ndhA* was functional only in *D. huoshanense* and *D. nobile*. *ndhE* gene was functional only in *D. nobile*, and *D. officinale*. *ndhD*, *ndhF*, *ndhG*, *ndhH*, and *ndhJ* genes were functional only in *D. nobile*. Moreover, *ndhK* was lost in *D. nobile*. A comparison of the cpDNAs of the two different genera within the *Orchidaceae* family (18 species) showed substantial variations based on the number and length of the NDH pseudogenes. An in-depth analysis of the pseudogenes *ndhF* and *ndhG*, shared among the four *Dendrobium* species, has unveiled intriguing data. In the context of *D. nobile*, *ndhF* exhibits five notable mutations in the coding DNA sequence (CDS): aAa (phenylalanine)-aTa (tyrosine), CgA (alanine)-AgT (serine), aTA (tyrosine)-aAG (phenylalanine), TTa (aspartic acid)- GAa (leucine), and AGG (serine)-GAA (leucine). Remarkably, *ndhG* in *D. nobile* not only demonstrates a substantial number of gene mutations but also manifests certain amino acid changes, resulting in premature termination codons. Notably, *ndhG* undergoes both insertions and deletions in the other three *Dendrobium* species, further emphasizing its dynamic nature and evolutionary significance.

3.5. Sequence repeat identification

A total of 175 SSRs from the cpDNA of *D. flexicaule* were obtained. Among these, most of the SSRs were mainly A/T (107, 61.14%), C/G monomeric repeats (4, 2.29%), dimeric repeats (54, 30.86%), trimeric repeats (4, 2.29%), tetrameric repeats (5, 2.86%), and hexameric repeats (1, 0.57%) (**Figure 5**). The longest monomeric repeats were 14 bp (*ycf1*) in the SSC region. The pentameric repeats were not identified in the cpDNA. Moreover, 39 dispersed repeats were observed in the cpDNA, with the majority of them located in the LSC and IR regions. A total of 16 forward repeats were identified with the longest being 39 bp (*ycf3*-intron1 and IGS: *rps12*, *trnV-GAC*). Twenty-one palindromic repeats were identified, with the longest being 59 bp and

located in the IGS (*rps19-2*, *psbA*) from the LSC region. Two reverse repeats were identified from the LSC and SSC regions of the cpDNA, including 31 bp (IGS: *rpl32* and *trnL-UAG*) in the SSC region and 30 bp (IGS: *rbcL* and *accD*) in the LSC region. The complementary repeats were not identified in the *D. flexicaule* cpDNA. To identify variation in the sequence repeats of the *Dendrobium* species, repeat fragments of the other three *Dendrobium* species were identified by comparing them with those of *D. flexicaule*. The number of SSRs ranged from 163 to 175. The nucleotide repeat types were similar to *D. flexicaule*, suggesting that sequence repeats were conserved in the evolution of the cpDNA. Ninety-nine dispersed repeats from the three *Dendrobium* species were identified, including palindromic repeats (18-23), forward repeats (7-10), and reverse repeats (2-2). The four complementary repeats were only observed for the *D. officinale* (1) and *D. huoshanense* (3) cpDNA in the LSC region.

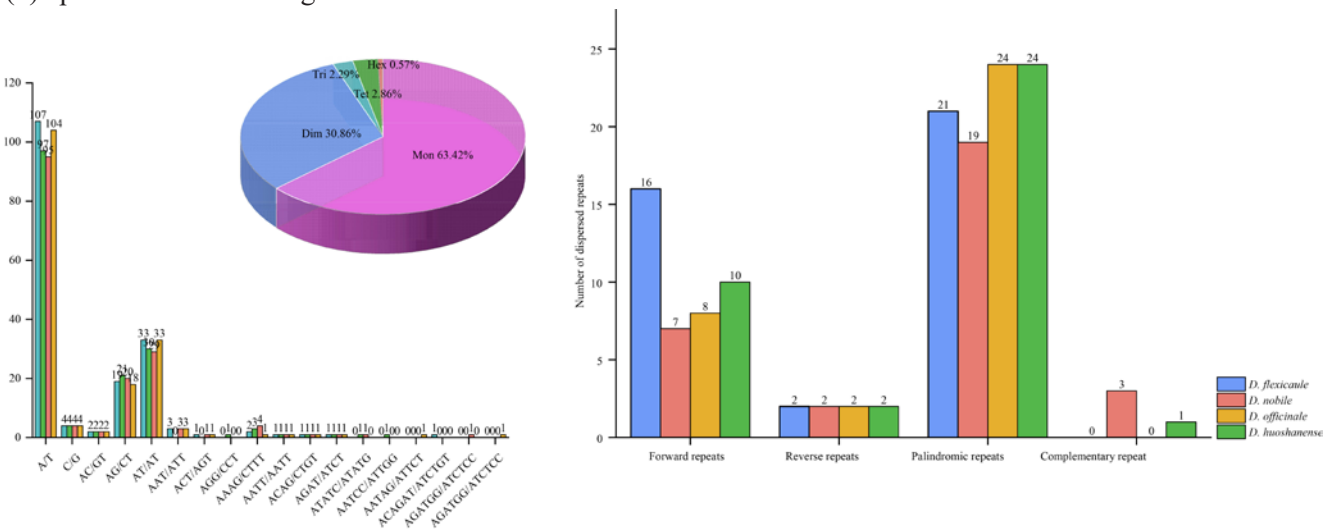

Figure 5. Comparative analysis of the types of sequence repeats among the four *Dendrobium* species

3.6. Comparative analysis of the chloroplast genome of *D. flexicaule* with the other three *Dendrobium* species

The cpDNAs from the four *Dendrobium* species were used to gain insight into the distinction of the four regions in the genomes using the mVISTA program according to the alignment of the four sequences, with *D. huoshanense* cpDNA as a reference sequence (**Figure 6**). The final alignment results showed that variation in IR regions was lower than in SSC and LSC regions, and the coding regions of the four *Dendrobium* cpDNAs were more highly conserved than the conserved noncoding regions. The different regions in the exons rarely exhibited any difference in the sequences of the four *Dendrobium* species.

In addition, the Pi of the cpDNA for the four *Dendrobium* species was analyzed to determine the divergence in the cpDNAs (**Figure 7**). The Pi plot indicated that the LSC and SSC regions had more variation than the IRs regions, and the LSC region was hypervariable compared with the SSC region. Moreover, the four hypervariable sites were distributed in different regions, including *psbI-trnS-GGA* (LSC region), *trnL-UAA* (LSC region), *rbcL* (LSC region), and *ycf1* (SSC region). The highest Pi was 0.25667 in *rbcL*. These hypervariable sites might undergo rapid nucleotide substitution, which is also used for the identification of molecular markers and the construction of the phylogenetic tree.

To further investigate the influence of the environment in the evolution of the cpDNA of *D. flexicaule*, 17 shared functional genes (*accD*, *atpB*, *ccsA*, *matK*, *psaB*, *rbcL*, *rpl20*, *rpl22*, *rpoA*, *rpoB*, *rpoC1*, *rpoC2*, *rps16*, *rps2*, *rps8*, *ycf1*, and *ycf2*) among the four *Dendrobium* species with *D. officinale* as a reference were identified using the Ka/Ks substitution rate (**Figure 8**). Among them, the Ka/Ks ratio of *matK*, *rpoC2*, *ycf1*, and *ycf2*

genes was > 1, which suggested that these genes were under positive selection in the evolution of cpDNA. The Ka/Ks ratio of the remaining 13 PCGs (*accD*, *atpB*, *ccsA*, *psaB*, *rbcL*, *rpl20*, *rpl22*, *rpoA*, *rpoB*, *rpoC1*, *rps16*, *rps2*, and *rps8*) was < 1, suggesting that these PCGs were under purifying selection during evolution. The above results show that PCGs under selection pressure in *Dendrobium* cpDNAs are mainly attributed to maturase, DNA-dependent RNA polymerase, and conserved open reading frames, which are closely related to RNA transcription and translation. Moreover, the Ka/Ks ratio of *D. huoshanense* (*matK*, *ycf1*, and *ycf2*), *D. flexicaule* (*ycf1*), and *D. nobile* (*rpoC2*, *ycf1*, and *ycf2*) genes was>1, suggesting that the cpDNA of the four *Dendrobium* species was under positive selection.

Figure 6. Visualization of the chloroplast genome sequence differences between *Dendrobium flexicaule* and its relatives

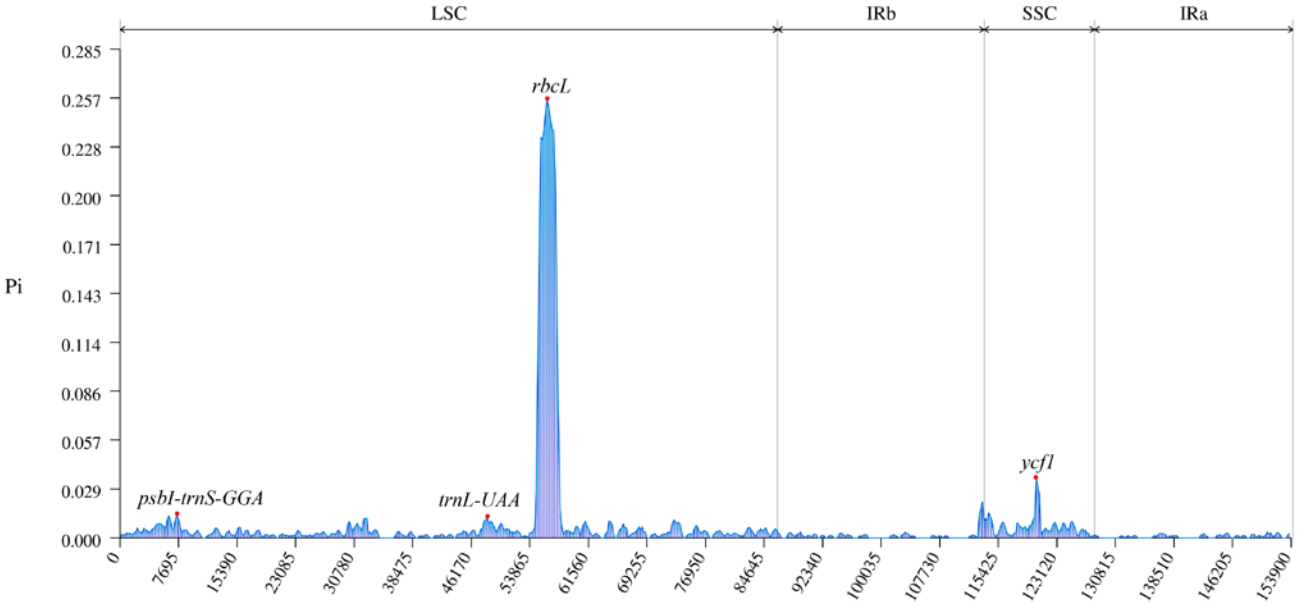

Figure 7. Nucleotide variation values in each region of the four *Dendrobium* species

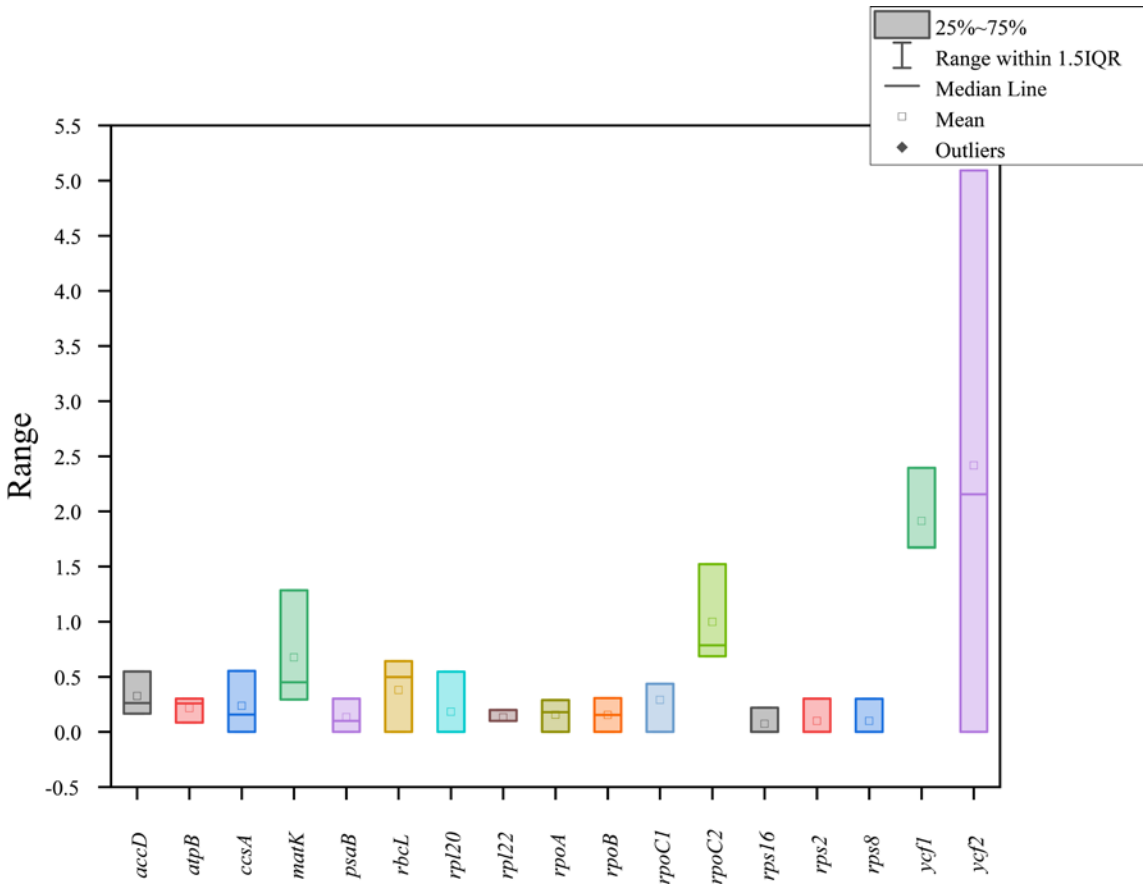

Figure 8. Ka/Ks ratio based on the shared PCGs of the four *Dendrobium* species

3.7. Phylogenetic analysis of *Dendrobium* species cpDNA

An NJ phylogenetic tree was constructed using 42 *Orchidaceae* species, 2 Amaryllidaceae species (*Allium ovalifolium* and *Allium tuberosum*), and 2 outgroup species (*Phoenix canariensis* and *Phoenix rupicola*) to gain further insight into the successive sister relationship of *D. flexicaule* with the *Dendrobium* genus (**Figure 9**). With the exception of *D. flexicaule*, *D. nobile*, *D. huoshanense*, and *D. officinale*, complete cpDNA sequences of the other 42 species were downloaded from the NCBI database.

The *Orchidaceae* species were categorized into three separate clades: the *Cypripedium clade*, the *Calanthe clade*, and the *Dendrobium clade*. *Calanthe* and *Cypripedium* were found to be successive sister clades, whereas Calanthe was distantly related to *Dendrobium*. Furthermore, the *Dendrobium* species were divided into multiple clades. Among these clades, *D. flexicaule* and *D. scoriarum* were closely related with strong support (Bootstrap = 1), whereas *D. huoshanense* was successively related to *D. wilsonii*. However, *D. officinale* and *D. nobile* formed two separate clades, indicating that different geographical locations and ecological environments may be the primary factors contributing to this divergence.

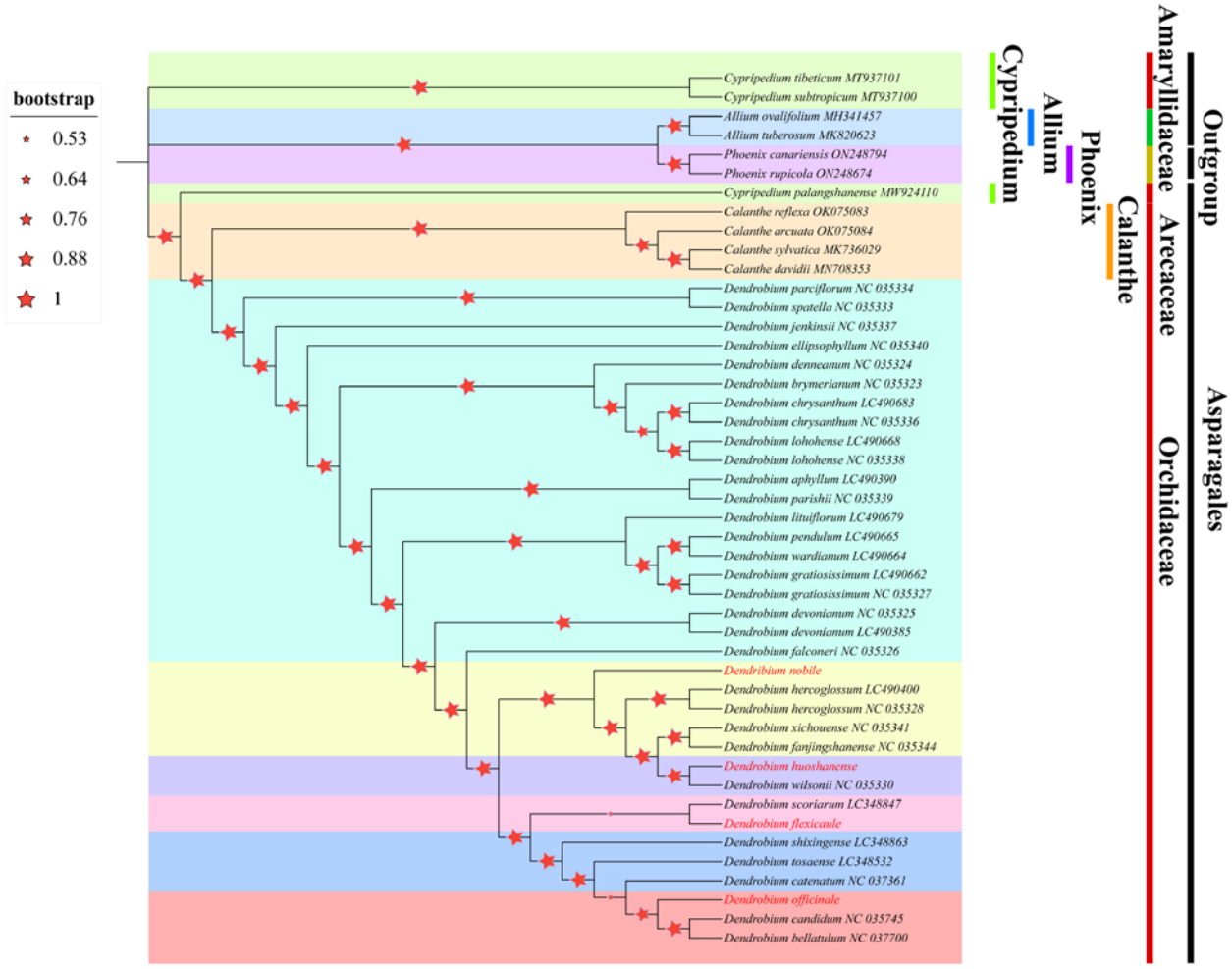

Figure 9. Phylogenetic tree construction based on the complete chloroplast genomes of the 46 *Orchidaceae* species

4. Discussion

As a vital organelle in seed plant cells, chloroplasts play a crucial role in photosynthesis, carbon dioxide fixation, and energy transformation [40]. The majority of the research reports indicated that the complete cpDNA has a typical quadripartite structure and conserved coding sequences in terrestrial plants [41]. In this paper, cpDNAs of *D. flexicaule* and its related species (*D. nobile*, *D. officinale*, and *D. huoshanense*) were sequenced, assembled, and annotated, and they exhibited similar structure as shown in previous studies, with size ranging from 150,529 to 152,588 bp. The length of *D. flexicaule* cpDNA was larger than that of its related species. Moreover, all of the 493 PCGs were found in the four *Dendrobium* species cpDNA. Among them, the total number of genes, GC content, and GC content in different regions (LSC, SSC, and IRs) had a similar tendency, which suggested that the genetic evolution of cpDNA was highly conserved in flowering plants (**Table 1**) [42]. Further investigation on the gene distribution in *D. flexicaule* cpDNA revealed that of 122 PCGs, 77 functional genes were distributed in the LSC region, accounting for 63% of the total PCGs. Then, comparative analysis of the GC content in the four *Dendrobium* species cpDNA revealed a tendency in the usage of A/T bases and A/T end codons that accounted for 61%–62%, which was similar to previous reports of cpDNA of *Panicum miliaceum* and *Amaranthus hypochondriacu* [43–44].

GC content and codon usage preference are closely related to the adaptability and frequency of tRNA usage. Codon usage bias was regarded as a response to adapt to the environment in evolutionary history [45].

As a common characteristic in terrestrial plants, reports have revealed codon usage bias of crucial value to investigate the mechanism of molecular evolution [46–47]. A total of 50,862 codons of PCGs were obtained from *D. flexicaule* cpDNA. Among them, 33 had codon usage preference (RSCU < 1), and UGG (tryptophan) indicated no codon usage bias (RSCU = 1). Notably, the comparative analysis of codon usage bias in cpDNAs of *D. flexicaule* with the other three *Dendrobium* species suggested they exhibited a similar tendency, demonstrating that codon usage preference was highly conserved among the *Dendrobium* species cpDNAs. The ENC-plot analysis of the four *Dendrobium* species cpDNA indicated consistency in the distribution of the functional genes in the standard curve, and the majority of the PCGs were distributed below the standard curve, which suggested that natural selection was responsible for the codon usage pattern. Additionally, some other factors influenced codon usage preference as few genes of the cpDNAs of the four *Dendrobium* species were identified on both sides of the standard curve, consistent with previous research reports [48]. It is necessary to understand the molecular mechanism of biological adaption and the evolutionary status of flowering plants to explore and gain insight into the different factors that impact codon usage patterns [49].

Repeat sequence in partial seed plants accounted for 90% of the genome length, which generated major divergence in genomes and showed evolutionary distances among seed plants [50]. Moreover, reports indicate that frequent recombination in the genome was mediated by the repeats [51]. Studies have revealed that SSRs in the cpDNA of seed plants are greatly hypervariable and have been used for population genetic and phylogenetic analyses at the intraspecific and interspecific levels [52–53]. The complete cpDNA of *D. flexicaule* revealed 175 SSRs and 39 dispersed repeats with 16 forward repeats, 21 palindromic repeats, and 2 reverse repeats. The majority of the SSRs were identified in the LSC regions. The repeat sequences of the other three *Dendrobium* species exhibited similar tendencies with mainly A/T SSRs types, indicating that the cpDNA was highly conserved in evolution. These findings were consistent with a previous study [54]. The four complementary repeats were only observed in the cpDNAs of *D. officinale* and *D. huoshanense* in the LSC region. An insight into the evolutionary origin, distribution, and chromosomal position in repeat sequences is vital to understanding the evolution, behavior, and organization of repetitive fragments in flowering plants [50].

The contraction and expansion of IRs were closely related to divergence in the four regions (LSC, SSC, IRa, and IRb) of cpDNA and its size [54–55]. Based on the complete cpDNA of the four *Dendrobium* species, an IR boundary was constructed. Although the four regions of the cpDNA were similar, *ndhF* was not identified in the IRb/SSC region and *rpl32* was distributed in the SSC region of *D. nobile* cpDNA. Then, further annotation of *ndhF* revealed that it had become a pseudogene during the evolution of cpDNA. Ycf1 underwent varying degrees of expansion and contraction in length in *D. flexicaule*, *D. huoshanense*, and *D. officinale*. However, *ycf1* was identified in the relative analysis of the IR boundary. To investigate the variation in the four regions of *Dendrobium* species cpDNA, the sequence alignment suggested that the divergence of SSC and LSC regions was higher than IR regions, and the conservation of coding regions was higher than noncoding regions. Extensive research has unveiled the pivotal physiological roles of the chloroplast NDH complex. This intricate molecular machinery not only participates in the complex process of cyclic electron transfer but also plays a vital role in alleviating excessive reduction and oxidative stress within the chloroplast stroma. In this paper, different numbers of pseudogenes have been identified in the cpDNAs of the four *Dendrobium* species. For example, eight pseudogenes were found in *D. flexicaule*, seven in *D. officinale* and *D. huoshanense*, and only two were found in *D. nobile*. *ndhK* was lost in *D. nobile*. Most of these genes were located in the SSC region of the cpDNA, with only a few located in the LSC region. This suggests that the structural instability of the SSC region might contribute to the accumulation of pseudogenes and subsequent NDH gene loss. Further analysis of the shared pseudogenes, *ndhF* and *ndhG*, in our *Dendrobium* species revealed multiple mutations

in *ndhF* of *D. nobile*, such as aAa (phenylalanine)-aTa (tyrosine), CgA (alanine)-AgT (serine), aTA (tyrosine)-aAG (phenylalanine), TTa (aspartic acid)-GAa (leucine), and AGG (serine)-GAA (leucine). *ndhG* in *D. nobile* not only exhibited a substantial number of gene mutations but also showed specific amino acid changes that resulted in premature termination codons. Furthermore, *ndhG* insertions and deletions were also observed in the other three *Dendrobium* species, indicating the dynamic nature and evolutionary significance of NDH genes. The random nature of NDH gene decay, its association with genomic rearrangements, and its potential adaptive value in specific ecological niches highlight the complex interplay between genomic evolution and plant adaptation in the context of photosynthesis. The Pi of the cpDNA of the four *Dendrobium* species was calculated, and it revealed that the LSC region had more hypervariable sites than the SSC region. The four hypervariable sites (*psbI-trnS-GGA*, *trnL-UAA*, *rbcL*, and *ycf1*) were acquired through comparison and screening. Among these hypervariable sites, the Pi value (0.25667) of *rbcL* was the highest. These genes could be beneficial for the development of molecular markers in *Dendrobium* species. Moreover, *rbcL* was frequently used for the identification of the genetic material of traditional Chinese medicinal substances.

In the current study, 17 shared PCGs (*accD*, *atpB*, *ccsA*, *matK*, *psaB*, *rbcL*, *rpl20*, *rpl22*, *rpoA*, *rpoB*, *rpoC1*, *rpoC2*, *rps16*, *rps2*, *rps8*, *ycf1*, and *ycf2*) from the four *Dendrobium* species were identified to calculate the Ka/Ks ratio. The Ka/Ks value of the four genes (*matK*, *rpoC2*, *ycf1*, and *ycf2*) was > 1, indicating that these genes were under positive selection, whereas the Ka/Ks ratio of the other 13 genes was < 1, which indicated that these PCGs were under purifying selection. More than one of four genes was closely associated with the transcription and translation of RNA. All the 42 plant species were used to construct the phylogenetic tree. The results indicated that two species (*D. flexicaule* and *D. scoriarum*) were clustered close to the *Dendrobium* species.

5. Conclusion

The differences in the cpDNAs between *D. flexicaule* and its three related species were investigated using the comparative analysis of the complete cpDNA sequence. The results showed that there was a preference for A/T bases at the SSR sites of single nucleotide repeat and dinucleotide repeat types in the four species of *Dendrobium*. Simultaneously, some differences in the IR/SC boundaries of the four species of *Dendrobium* were also observed. The IRb/SSC nodes of *D. officinale*, *D. flexicaule*, and *D. huoshanense* expanded into *ndhF*, whereas the IRb/SSC nodes of *D. flexicaule*, *D. officinale*, and *D. nobile* were between *ycf1* and *ndhF*. Interestingly, *ndhF* was converted to a pseudogene during the evolution of the *D. nobile* cpDNA. In addition, this study used the Pi value in combination with the mVISTA map to determine four positions with higher variation intervals: *psbI-trnS-GGA* (LSC region), *trnL-UAA* (LSC region), *rbcL* (LSC region), and *ycf1* (SSC region), which can be used as potential sites for the molecular identification of *D. flexicaule*. The results of the phylogenetic analysis showed that *D. flexicaule* and *D. scoriarum* clustered together. In summary, the findings of the present study provide a new reference for species identification, molecular marker development and utilization, and phylogenetic research on *D. flexicaule*.

Funding

This research was supported by Reproductive Ecology and Risk Factors of *Dendrobium flexicaule* Project, China (JBKY20210006), Ecological and Genetic Basis of the Special Growth Characteristics of *Dendrobium flexicaule* Project, China (CSTC2021 JSCX-DXWT BX0015), Research on the medicinal plants of Yintiaoling.

Author contribution

Study methodology conceptualization and software: Le Wang, Xue Liu
Data curation: Le Wang
Draft writing and preparation: Le Wang
Study visualization and investigation: Jie Zhang, Xingjia Ming
Project supervision: Shu Shu, Xianyou Qu
Software and data validation: Jie Zhang, Junsheng Qi
Article writing, reviewing, and editing: Xue Liu

Data availability

Data generated or analyzed during this study are available in the NCBI repository, accession numbers OQ360111, OR387324, OR387325, and OR387323, respectively.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Yu X, Zuo L, Lu D, et al., 2019, Comparative Analysis of Chloroplast Genomes of Five *Robinia* Species: Genome Comparative and Evolution Analysis. Gene, 2019(689): 141–151. https://doi.org/10.1016/j.gene.2018.12.023

[2] Konhar R, Debnath M, Vishwakarma S, et al., 2019, The Complete Chloroplast Genome of *Dendrobium nobile*, an Endangered Medicinal Orchid from North-east India and its Comparison with Related *Dendrobium* Species. PeerJ, 2019(7): e7756. https://doi.org/10.7717/peerj.7756

[3] Liang H, Zhang Y, Deng J, et al., 2020, The Complete Chloroplast Genome Sequences of 14 *Curcuma* Species: Insights into Genome Evolution and Phylogenetic Relationships Within Zingiberales. Frontiers in Genetics, 2020(11): 802. https://doi.org/10.3389/fgene.2020.00802

[4] Tian X, Li D, 2002, Application of DNA Sequences in Plant Phylogenetic Study. Acta Botanica Yunnanica, 24(2): 170–184.

[5] Palmer JD, 1985, Comparative Organization of Chloroplast Genomes. Annual Review of Genetics, 19(1): 325–354. https://doi.org/10.1146/annurev.ge.19.120185.001545

[6] Raubeson LA, Jansen RK, 2005, Chloroplast Genomes of Plants, in Plant Diversity and Evolution: Genotypic and Phenotypic Variation in Higher Plants. CABI Publishing, Wallingford, 45–68. http://doi/10.1079/9780851999043.0045

[7] Bendich AJ, 2004, Circular Chloroplast Chromosomes: the Grand Illusion. Plant Cell, 16(7): 1661–1666. https://doi.org/10.1105/tpc.160771

[8] Frailey DC, Chaluvadi SR, Vaughn JN, et al., 2018, Gene Loss and Genome Rearrangement in the Plastids of Five Hemiparasites in the Family Orobanchaceae. BMC Plant Biology, 18(1): 30. https://doi.org/10.1186/s12870-018-1249-x

[9] Böhle UR, Hilger H, Cerff R, et al., 1994, Non-coding Chloroplast DNA for Plant Molecular Systematics at the Infrageneric Level, in Molecular Ecology and Evolution: Approaches and Applications. Birkhäuser, Basel, 391–403.

[10] Small RL, Lickey EB, Shaw J, et al., 2005, Amplification of Noncoding Chloroplast DNA for Phylogenetic Studies in Lycophytes and Monilophytes with a Comparative Example of Relative Phylogenetic Utility from Ophioglossaceae.

Molecular Phylogenetics and Evolution, 36(3): 509–522. https://doi.org/10.1016/j.ympev.2005.04.018

[11] Blazier JC, Ruhlman TA, Weng ML, et al., 2016, Divergence of RNA Polymerase α Subunits in Angiosperm Plastid Genomes is Mediated by Genomic Rearrangement. Scientific Reports, 6(1): 24595. https://doi.org/10.1038/srep24595

[12] Chumley TW, Palmer JD, Mower JP, et al., 2006, The Complete Chloroplast Genome Sequence of *Pelargonium hortorum*: Organization and Evolution of the Largest and Most Highly Rearranged Chloroplast Genome of Land Plants. Molecular Biology and Evolution, 23(11): 2175–2190. https://doi.org/10.1093/molbev/msl089

[13] Gao L, Su YJ, Wang T, 2010, Plastid Genome Sequencing, Comparative Genomics, and Phylogenomics: Current Status and Prospects. Journal of Systematics and Evolution, 48(2): 77–93. https://doi.org/10.1111/j.1759-6831.2010.00071.x

[14] Guisinger MM, Kuehl JV, Boore JL, et al., 2011, Extreme Reconfiguration of Plastid Genomes in the Angiosperm Family Geraniaceae: Rearrangements, Repeats, and Codon Usage. Molecular Biology and Evolution, 28(1): 583–600. https://doi.org/10.1093/molbev/msq229

[15] Liu J, Hu Z, Cheng Z, et al., 2021, Study on Seed Germination of Endangered Plant *Dendrobium flexicaule*. Seed, 2021(40): 48–51 + 56.

[16] Lu W, 2021, The New List of Wild Plants under Key State Protection was Officially Released. Land Green, 48–49.

[17] Ding X, Xu L, Xu H, et al., 2001, Morphological and DNA Molecular Evidence for Authentication of *Dendrobium flexicaule* from its Allied Species of *Dendrobium*. Acta Pharmaceutica Sinica, 2001(36): 868–873.

[18] Zhang S, 2016, Effect of Environmental Conditions on Growth and Propagation of *Dendrobium flexicaule* in Funiu Mountain Area. Xiandai Horticulture, 2016(21): 13–14. https://doi.org/10.14051/j.cnki.xdyy.2016.21.110

[19] Hu ZF, 2021, Optimization of Rapid Propagation System of *Dendrobium flexicaule*, thesis, China Three Gorges University.

[20] Sun T, Yang Y, Hu S, et al., 2022, Identification and Quality Evaluation of *Dendrobium flexicaule* and its Related Species. Chinese Journal of Experimental Traditional Medical Formulae, 2022(28): 128–134.

[21] Wu T, Deng G, Yin Q, et al., 2023, Characterization and Molecular Evolution Analysis of *Periploca forrestii* Inferred from its Complete Chloroplast Genome Sequence. Genome, 66(2): 34–50. https://doi.org/10.1139/gen-2022-0050

[22] Qu XJ, Moore MJ, Li DZ, et al., 2019, PGA: a Software Package for Rapid, Accurate, and Flexible Batch Annotation of Plastomes. Plant Methods, 15(1): 50. https://doi.org/10.1186/s13007-019-0435-7

[23] Greiner S, Lehwark P, Bock R, 2019, OrganellarGenomeDRAW (OGDRAW) version 1.3.1: Expanded Toolkit for the Graphical Visualization of Organellar Genomes. Nucleic Acids Research, 47(W1): 59–64. https://doi.org/10.1093/nar/gkz238

[24] Sharp PM, Li WH, 1986, Codon Usage in Regulatory Genes in *Escherichia coli* does not Reflect Selection for "Rare" Codons. Nucleic Acids Research, 14(19): 7737–7749. https://doi.org/10.1093/nar/14.19.7737

[25] Sharp PM, Li WH, 1986, An Evolutionary Perspective on Synonymous Codon Usage in Unicellular Organisms. Journal of Molecular Evolution, 24(1–2): 28–38. https://doi.org/10.1007/BF02099948

[26] Wright F, 1990, The Effective Number of Codons Used in a Gene. Gene, 87(1): 23–29. https://doi.org/10.1016/0378-1119(90)90491-9

[27] Li H, Guo Q, Xu L, et al., 2023, CPJSdraw: Analysis and Visualization of Junction Sites of Chloroplast Genomes. PeerJ, 2023(11): e15326. https://doi.org/10.7717/peerj.15326

[28] Kearse M, Moir R, Wilson A, et al., 2012, Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data. Bioinformatics, 28(12): 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

[29] Li X, Yang JB, Wang H, et al., 2021, Plastid NDH Pseudogenization and Gene Loss in a Recently Derived Lineage from the Largest Hemiparasitic Plant Genus *Pedicularis* (Orobanchaceae). Plant and Cell Physiology, 62(6): 971–

984. https://doi.org/10.1093/pcp/pcab074

[30] Du Z, Yang L, Zhang Z, et al., 2021, Characteristics of the Complete Chloroplast Genome of *Dendrobium ochreatum* and its Comparative Analysis. Chinese Journal of Tropical Crops, 2021(42): 3111–3119. https://10.0.15.129/j.issn.1000-2561.2021.11.009

[31] Beier S, Thiel T, Münch T, et al., 2017, MISA-web: a Web Server for Microsatellite Prediction. Bioinformatics, 33(16): 2583–2585. https://doi.org/10.1093/bioinformatics/btx198

[32] Benson G, 1999, Tandem Repeats Finder: a Program to Analyze DNA Sequences. Nucleic Acids Research, 27(2): 573–580. https://doi.org/10.1093/nar/27.2.573

[33] Kurtz S, 2001, REPuter: The Manifold Applications of Repeat Analysis on a Genomic Scale. Nucleic Acids Research, 29(22): 4633–4642. https://doi.org/10.1093/nar/29.22.4633

[34] Moberly JG, Bernards MT, Waynant KV, 2018, Key Features and Updates for Origin 2018. Journal of Cheminformatics, 10(1): 5. https://doi.org/10.1186/s13321-018-0259-x

[35] Rozas J, Ferrer-mata A, Sánchez-delbarrio JC, et al., 2017, DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and Evolution, 34(12): 3299–3302. https://doi.org/10.1093/molbev/msx248

[36] Zhang Z, Li J, Zhao XQ, et al., 2006, KaKs_Calculator: Calculating Ka and Ks Through Model Selection and Model Averaging. Genomics, Proteomics & Bioinformatics, 4(4): 259–263. https://doi.org/10.1016/S1672-0229(07)60007-2

[37] Kumar S, Stecher G, Li M, et al., 2018, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution, 35(6): 1547–1549. https://doi.org/10.1093/molbev/msy096

[38] Nakamura T, Yamada KD, Tomii K, et al., 2018, Parallelization of MAFFT for Large-scale Multiple Sequence Alignments. Bioinformatics, 34(14): 2490–2492. https://doi.org/10.1093/bioinformatics/bty121

[39] Letunic I, Bork P, 2019, Interactive Tree of Life (iTOL) v4: Recent Updates and New Developments. Nucleic Acids Research, 47(W1): 256–259. https://doi.org/10.1093/nar/gkz239

[40] Verma D, Daniell H, 2007, Chloroplast Vector Systems for Biotechnology Applications. Plant Physiology, 145(4): 1129–1143. https://doi.org/10.1104/pp.107.106690

[41] Shi N, Yang Z, Miao K, et al., 2023, Comparative Analysis of the Medicinal Plant *Polygonatum kingianum* (Asparagaceae) with Related Verticillate Leaf Types of the Polygonatum Species based on Chloroplast Genomes. Frontiers in Plant Science, 2023(14): 1202634. https://doi.org/10.3389/fpls.2023.1202634

[42] Shahzadi I, Abdulla H, Mehmood F, et al., 2020, Chloroplast Genome Sequences of *Artemisia maritima* and *Artemisia absinthium*: Comparative Analyses, Mutational Hotspots in Genus Artemisia and Phylogeny in Family Asteraceae. Genomics, 112(2): 1454–1463. https://doi.org/10.1016/j.ygeno.2019.08.016

[43] Feng R, Mei C, Wang H, et al., 2019, Analysis of Codon Usage in the Chloroplast Genome of Grain Amaranth (*Amaranthus hypochondriacus L.*). Chinese Journal of Grassland, 2019(41): 8–15. https//doi.org/10.16742/j.zgcdxb.20180227

[44] Liu H, Wang M, Yue W, et al., 2017, Analysis of Codon Usage in the Chloroplast Genome of Broomcorn millet (*Panicum miliaceum L.*). Plant Science Journal, 35(3): 362–371. https://doi.org/10.11913/PSJ.2095-0837.2017.30362

[45] Botzman M, Margalit H, 2011, Variation in Global Codon Usage Bias among Prokaryotic Organisms is Associated with their Lifestyles. Genome Biology, 12(10): R109. https://doi.org/10.1186/gb-2011-12-10-r109

[46] Nie X, Deng P, Feng K, et al., 2014, Comparative Analysis of Codon Usage Patterns in Chloroplast Genomes of the Asteraceae Family. Plant Molecular Biology Reporter, 32(4): 828–840. https://doi.org/10.1007/s11105-013-0691-z

[47] Sharp PM, Matassi G, 1994, Codon Usage and Genome Evolution. Current Opinion in Genetics & Development, 4(6): 851–860. https://doi.org/10.1016/0959-437X(94)90070-1

[48] Ma M, Zhang W, Meng H, et al., 2021, Codon Bias Analysis of Chloroplast Genome in Medicinal Plants of *Amomum Roxb*. Chinese Traditional and Herbal Drugs, 52(12): 3661–3670. https://doi.org/10.7501/

j.issn.0253-2670.2021.12.024

[49] Angellotti MC, Bhuiyan SB, Chen G, et al., 2007, CodonO: Codon Usage Bias Analysis within and Across Genomes. Nucleic Acids Research, 35(2): 132–136. https://doi.org/10.1093/nar/gkm392

[50] Mehrotra S, Goyal V, 2014, Repetitive Sequences in Plant Nuclear DNA: Types, Distribution, Evolution and Function. Genomics, Proteomics & Bioinformatics, 12(4): 164–171. https://doi.org/10.1016/j.gpb.2014.07.003

[51] Varré JS, D'Agostino N, Touzet P, et al., 2019, Complete Sequence, Multichromosomal Architecture and Transcriptome Analysis of the *Solanum tuberosum* Mitochondrial Genome. International Journal of Molecular Sciences, 20(19): 4788. https://doi.org/10.3390/ijms20194788

[52] Gupta PK, Varshney RK, 2000, The Development and Use of Microsatellite Markers for Genetic Analysis and Plant Breeding with Emphasis on Bread Wheat. Euphytica, 113(3): 163–185. https://doi.org/10.1023/A:1003910819967

[53] Kang J, Lu J, Qiu S, et al., 2015, *Dendrobium* SSR Markers Play a Good Role in Genetic Diversity and Phylogenetic Analysis of *Orchidaceae* Species. Scientia Horticulturae, 183(12): 160–166. https://doi.org/10.1016/j.scienta.2014.12.018

[54] Song Y, Wang S, Ding Y, et al., 2017, Chloroplast Genomic Resource of Paris for Species Discrimination. Scientific Reports, 7(1): 3427. https://doi.org/10.1038/s41598-017-02083-7

[55] Niu ZT, Zhu SY, Pan JJ, 2017, Comparative Analysis of *Dendrobium plastomes* and Utility of Plastomic Mutational Hotspots. Scientific Reports, 7(1): 2073. https://doi.org/10.1038/s41598-017-02252-8