horticulturae

MDPI

Article

Plastomes of Seven *Coelogyne* s.l. (Arethuseae, Orchidaceae) Species: Comparative Analysis and Phylogenetic Relationships

Songkun Lin [1,†], Ruyi Li [2,†], Shuling Tang [3], Yuming Chen [1], Yin Yan [1], Xuyong Gao [2,*] and Xiaokang Zhuo [4,*]

[1] College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China; 17859563113@163.com (S.L.); 12304022001@fafu.edu.cn (Y.C.); 52304022100@fafu.edu.cn (Y.Y.)
[2] Key Laboratory of National Forestry and Grassland Administration for Orchid Conservation and Utilization at College of Landscape Architecture, Fujian Agriculture and Forestry University, Fuzhou 350002, China; liruyi0720@163.com
[3] Zhangzhou Institute of Technology, Zhangzhou 363000, China; fjzzztsl@126.com
[4] College of Horticulture, Fujian Agriculture and Forestry University, Fuzhou 350002, China
* Correspondence: gaoxuyong11@126.com (X.G.); zhuoxiaokang@fafu.edu.cn (X.Z.)
† These authors contributed equally to this work.

Abstract: The *Coelogyne* s.l. is one of the emblematic genera of the Asian orchids, with high horticultural and medicinal values. However, the phylogenetic relationships of the genus inferred from previous studies based on a limited number of DNA markers remain ambiguous. In this study, we newly sequenced and assembled the complete plastomes of seven *Coelogyne* species: *C. bulleyia*, *C. fimbriata*, *C. flaccida*, *C. prolifera*, *C. tricallosa*, *C. uncata*, and an unknown taxa, *Coelogyne* sp. The plastomes of *Coelogyne* exhibited a typical quadripartite structure, varying in length between 157,476 bp and 160,096 bp, accompanied by a GC content spanning from 37.3% to 37.5%. A total of 132 genes were annotated for each plastome, including 86 protein-coding genes, eight rRNA genes, and 38 tRNA genes. Among these, 19 genes underwent duplication within the inverted repeat (IR) regions, and 18 genes exhibited the presence of introns. Additionally, we detected 54 to 69 simple sequence repeats (SSRs) and 30 to 49 long repeats. In terms of codon usage frequency, leucine (Leu) emerged with the highest frequency, while cysteine (Cys) exhibited the lowest occurrence. Furthermore, eight hypervariable regions (*atpB-rbcL*, *psbK-psbI*, *rps8-rpl14*, *rps16-trnQ^{UUG}*, *psaC-ndhE*, *ndhF-rpl32*, *psbB-psbT*, and *ycf1*) were identified. Phylogenetic analyses using complete plastomes and protein-coding genes indicated that *Coelogyne* s.l. was monophyletic. Moreover, the results robustly supported the division of *Coelogyne* s.l. into five clades. This study provides a comprehensive analysis of the structural variation and phylogenetic analysis of the *Coelogyne* s.l. based on plastome data. The findings offer significant insights into the plastid genomic characteristics and the phylogenetic relationships of *Coelogyne* s.l., contributing to a deeper understanding of its evolutionary history.

Keywords: *Coelogyne*; plastid genome; hypervariable region; phylogenetic analysis

check for updates

Academic Editor: Adriana F. Sestras

Received: 24 December 2024
Revised: 26 January 2025
Accepted: 27 January 2025
Published: 30 January 2025

Citation: Lin, S.; Li, R.; Tang, S.; Chen, Y.; Yan, Y.; Gao, X.; Zhuo, X. Plastomes of Seven *Coelogyne* s.l. (Arethuseae, Orchidaceae) Species: Comparative Analysis and Phylogenetic Relationships. *Horticulturae* **2025**, *11*, 144. https://doi.org/10.3390/horticulturae11020144

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Orchidaceae is one of the largest flowering plant families, comprising more than 28,000 species in five subfamilies [1,2]. Epidendroideae, the largest and most diverse subfamily, encompasses around 21,000 species, constituting roughly 75% of the total number of Orchidaceae [1]. Arethuseae is a key tribe located at the base of the "Higher Epidendroideae", including two subtribes (Arethusinae and Coelogyninae) with approximately 763 species from 14 genera [1,3–5]. *Coelogyne* s.l. *sensu* Chase et al. (2021) [3], belonging to

Coelogyninae, is an expanded genus including 14 genera, such as *Bulleyia*, *Dendrochilum*, *Panisea*, and *Pholidota*, which contains approximately 560 species [1]. It is mainly distributed throughout tropical and subtropical Asia to Oceania [6]. There are about 70 species distributed in China, mainly in Sichuan, Yunnan and Tibet [7]. Most species are epiphytes, rare lithophytes, or terrestrials, occurring in tropical lowlands and montane rainforests [8]. Certain *Coelogyne* species, such as *C. cristata* and *C. ovalis*, possess medicinal values [9,10].

The phylogenetic analysis of the genus *Coelogyne* has remained unclear since it was first described by Lindley [11], and it is generally considered to be one of the most complex taxa of Coelogyninae. Gravendeel et al. defined the taxa as two clades based on the combined analysis of three DNA regions (ITS, RFLPs, and *matK*) [12]. Subsequently, phylogenetic analysis of the subtribe Coelogyninae constructed by van den Berg et al., Górniak et al., and Li et al. based on different numbers of molecular markers, all of which have a small number of species and collapsed topology [13–15]. Based on three combined DNA regions (ITS, *matK*, and *trnL*), Li et al. defined the group into two clades [4]. Huang et al. conducted the phylogenetic analysis of Coelogyninae employing six sequences (*matK*, *trnL-F*, *rbcL*, *rpoC1*, *rpl32-trnL*, and *ycf1*), which revealed that the group can be divided into four clades [5]. The phylogenetic tree of *Coelogyne* and related taxa has been constructed using a limited number of markers [4,5,12–15], leading to low bootstrap values and unstable topologies across all studies.

Recently, advancements in high-throughput sequencing technology have significantly facilitated the acquisition of plastome (plastid genome) data, making it more accessible and cost-effective. The plastome of most angiosperms is maternally inherited [16], with rare instances of paternal or biparental inheritance, which can greatly reduce the impact of complex phenomena such as hybridization and genetic recombination on phylogenetic relationships. In contradistinction to nuclear and mitochondrial genomes, plastomes present distinct advantages, including conserved structure, easy sequencing and assembly, and moderate evolutionary rate [17]. Consequently, they have been extensively utilized for phylogenetic analysis [18–22]. Givnish et al. constructed the phylogenetic tree of Orchidaceae utilizing 75 plastid genes for 39 representative species, and the results showed high support, supporting the division of Orchidaceae into five subfamilies [18]. Tu et al. selected 46 representative species from the *Goodyera–Cheirostylis* clade of the Goodyerinae and reconstructed the phylogenetic relationships of these taxa using both plastid protein-coding genes and the whole plastome. The results successfully elucidated the intricate phylogenetic relationships within this group, which had previously been challenging to clarify using limited molecular markers [21]. Li et al. conducted the first comprehensive study of genetic variation in the genus *Pholidota* based on plastid genomic data and systematically analyzed its phylogenetic analysis and evolution. The findings provided a solid foundation for future investigations into the evolutionary mechanisms and classification of this economically and medicinally important genus [22].

In this study, we successfully sequenced and assembled the plastomes of seven *Coelogyne* species (*C. bulleyia*, *C. fimbriata*, *C. flaccida*, *C. prolifera*, *C. tricallosa*, *C. uncata*, and an unknown taxa, *Coelogyne* sp.). Our research was driven by three primary objectives: (1) detailed characterization and comparative analysis of seven *Coelogyne* plastomes; (2) exploration of genetic diversity and identification of molecular markers; (3) construction of a robust phylogenetic framework *Coelogyne* s.l. *sensu* Chase et al. (2021) [3]. The results of this study not only enrich the genetic resources available for *Coelogyne*, but also lay a solid foundation for future molecular identification and evolutionary investigations. The characterized plastomes and the identified molecular markers will facilitate a deeper understanding of the genetic diversity, population structure, and evolutionary history of *Coelogyne* species, as well as promote the development of its horticultural and medicinal values.

2. Materials and Methods

2.1. Taxon Sampling, DNA Extraction and Sequencing, Plastome Assembly and Annotation

A total of 45 samples from six genera of Arethuseae were employed, including seven newly sequenced plastomes and 38 plastomes obtained from the GenBank (Supplementary Table S1). To establish a robust phylogenetic framework, ten species from five genera (*Arundina*, *Bletilla*, *Pleione*, *Thunia*, and *Thuniopsis*) were selected as the outgroup based on previous studies [5,22]. The seven species of *Coelogyne* were collected from the Forest Orchid Garden greenhouse at Fujian Agriculture and Forestry University, Fuzhou, Fujian Province, China. We employed the methods described in previous studies for DNA extraction, sequencing, plastome assembly, and annotation [22].

2.2. Plastome Comparative and Codon Usage Analysis

The identification of simple sequence repeats (SSRs) in seven sequenced plastomes was conducted using the MIcroSAtellite identification tool (MISA) (https://webblast.ipk-gatersleben.de/misa/, accessed on 13 October 2024) [23]. The parameters for minimum repeat numbers were set as follows: mononucleotide (mono), dinucleotide (di), trinucleotide (tri), tetranucleotide (tetra), pentanucleotide (penta), and hexanucleotide (hexa), with values of 10, 5, 4, 3, 3, and 3, respectively. Four long repeat types (forward, reverse, complement, and palindrome) were identified in the plastomes of seven *Coelogyne* species using REPuter (hamming distance = 3, maximum computed repeats = 50 bp, and minimum repeat size = 30 bp) (https://bibiserv.cebitec.uni-bielefeld.de/reputer, accessed on 13 October 2024) [24]. The ProgressiveMauve algorithm in Geneious v11.1.5 [25] was used to visualize rearrangements among seven *Coelogyne* plastome sequences. The online tool, IRscope (https://irscope.shinyapps.io/irapp/, accessed on 15 October 2024) [26], was employed to visualize alterations in the boundaries of large single-copy (LSC), inverted repeat (IR), and small single-copy (SSC).

The protein-coding sequences were extracted from the plastomes, resulting in 79 sequences. These sequences were aligned using MAFFT v7.487 [27], and the alignment files were converted from fasta to aln format using the ALTER (http://www.sing-group.org/ALTER/, accessed on 15 October 2024) [28] online tool. We estimated the non-synonymous (Ka) and synonymous (Ks) substitution rates for each gene, as well as the Ka/Ks ratio, using the NG method and 11-Bacterial and Plant Plastid Code genetic code in KaKs_Calculator 2.0 [29]. A Ka/Ks ratio of <1, =1, and >1 indicates purifying selection, neutral evolution, and positive selection, respectively. The analysis of relative synonymous codon usage (RSCU) for seven *Coelogyne* species was conducted using DAMBE 7 [30].

2.3. Sequence Divergence Analysis

To visualize the variation hotspot map of seven *Coelogyne* plastomes, we employed the Shuffle-LAGAN alignment program through the online tool mVISTA (https://genome.lbl.gov/vista/mvista/submit.shtml, accessed on 17 October 2024) [31]. The plastome of *C. bulleyia* was utilized as the reference for this comparative analysis. The nucleotide diversity (Pi) within both the 79 protein-coding genes and intergenic regions of *Coelogyne* was computed through the utilization of DnaSP 6 [32].

2.4. Phylogenetic Analyses

We reconstructed the phylogenetic relationships within *Coelogyne* by employing diverse analytical methods, including maximum likelihood (ML), maximum parsimony (MP), and Bayesian inference (BI). These analyses were conducted based on two distinct data matrixes: whole plastome sequence and 79 protein-coding genes. The whole plastomes sequence was aligned by MAFFT v7.487 [27], and subsequent trimming was conducted using

TrimAL v1.4 [33]. The extraction of the 79 protein-coding genes was carried out using PhyloSuite v1.2.2 [34] and subsequently aligned using the MAFFT v7.487 [27]. Concatenation of the 79 protein-coding genes was performed utilizing PhyloSuite v1.2.2 [34].

Phylogenetic reconstructions were performed using the CIPRES Science Gateway website (RAxML-HPC2 on XSEDE 8.2.12, PAUP on XSEDE 4.a165, and MrBayes on XSEDE 3.2.7a) [35]. For the ML analysis, the GTRGAMMA model was employed for all datasets [36], and bootstrap values were calculated from 1000 bootstrap replicates through heuristic searches [37]. For the MP analysis, a heuristic search with 1000 random addition sequence replicates and tree-bisection-reconnection (TBR) branch swapping was performed. For the BI analysis, the GTR + I + Γ model inference was employed for all datasets, with the following parameter configurations: ngen = 2,000,000; samplefreq = 100; burnin = 500,000.

3. Results

3.1. Plastome Features

The plastomes of seven *Coelogyne* species demonstrated variability in size, varying from 157,476 bp in *C. uncata* to 160,096 bp in *Coelogyne* sp. (Table 1). The plastomes of seven *Coelogyne* species exhibited a typical quadripartite structure (Figure 1). The LSC region spanned from 86,886 bp in *C. uncata* to 87,869 bp in *C. prolifera*, the SSC region ranged from 17,987 bp in *C. fimbriata* to 18,826 bp in *Coelogyne* sp., and the IR region ranged from 26,117 bp in *C. uncata* to 26,746 bp in *C. flaccida*. Noteworthy variations in GC content were observed among the complete plastomes (37.3–37.5%), LSC (35.2–35.3%), SSC (30.3–31.1%), and IR regions (43.2–43.3%) regions, respectively (Table 1).

Table 1. Plastome features of *Coelogyne*.

Species	Plastome		LSC		SSC		IR		Number of Genes	Protein-Coding Genes	tRNA Genes	rRNA Genes
	Length (bp)	GC (%)	Length (bp)	GC (%)	Length (bp)	GC (%)	Length (bp)	GC (%)				
C. bulleyia	159,575	37.4	87,557	35.2	18,714	30.4	26,652	43.3	132	86	38	8
C. fimbriata	158,056	37.5	87,373	35.3	17,987	31.1	26,348	43.3	132	86	38	8
C. flaccida	159,901	37.4	87,606	35.3	18,803	30.5	26,746	43.3	132	86	38	8
C. prolifera	159,967	37.3	87,869	35.2	18,782	30.4	26,658	43.3	132	86	38	8
C. tricallosa	159,244	37.3	87,109	35.2	18,783	30.3	26,676	43.3	132	86	38	8
C. uncata	157,476	37.3	86,886	35.2	18,356	30.4	26,117	43.2	132	86	38	8
Coelogyne sp.	160,096	37.3	87,848	35.2	18,826	30.4	26,711	43.3	132	86	38	8

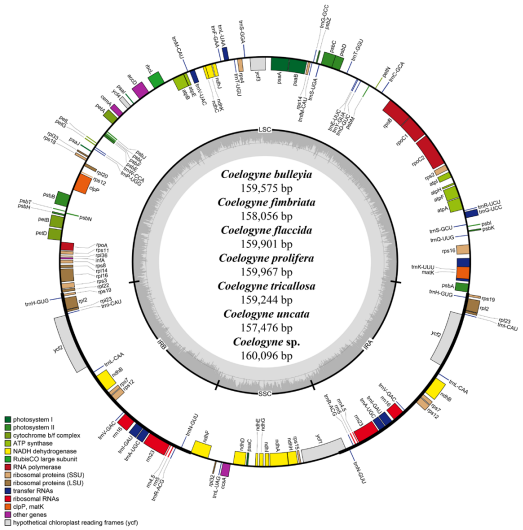

Figure 1. The plastome annotation map of seven *Coelogyne* species (*C. bulleyia*, *C. fimbriata*, *C. flaccida*, *C. prolifera*, *C. tricallosa*, *C. uncata*, and *Coelogyne* sp.). The transcription direction varies between genes inside and outside the circle, with clockwise orientation for inner-circle genes and counterclockwise orientation for outer-circle genes. Distinct gene colors are employed to signify their respective functions, and the darker gray in the inner circle represents the GC content.

The plastomes of seven *Coelogyne* species were found to have a consistent gene content, comprising 132 genes in total. These genes were categorized into three groups: 86 protein-coding genes, eight ribosomal RNA (rRNA) genes, and 38 transfer RNA (tRNA) genes. Notably, 19 genes underwent duplication within the IR regions. Furthermore, among the annotated genes, 18 genes exhibited the presence of introns (Supplementary Table S2). Mauve analysis showed that there were no obvious rearrangements among the examined plastomes of the seven *Coelogyne* species (Supplementary Figure S1).

3.2. Codon Usage Analysis

The codon usage frequency was computed using the 79 protein-coding genes extracted from the plastomes of seven *Coelogyne* species (Figure 2, Supplementary Table S3). These 79 genes of seven *Coelogyne* species were encoded by 22,840 (*C. fimbriata*) to 22,949 (*C. flaccida*) codons. Leucine (Leu) was the most abundant among these codons, ranging from 2383–2399 codons, accounting for a frequency of 10.42–10.48%. Isoleucine (Ile) followed, with 1941–1958 codons and a frequency of 8.49–8.57%. Conversely, cysteine (Cys) had the lowest frequency (excluding stop codons), with 271–278 codons and a frequency of 1.19–1.21%. RSCU values varied between 0.346 and 1.857, encompassing 30 codons exceeding 1, 32 codons below 1, and 2 codons (AUG and UGG) possessing RSCU values of 1. The GCU and GCG codons exhibited the highest and lowest RSCU values, ranging from 1.849 to 1.857 and 0.346 to 0.375, respectively. In terms of stop codons (UAA, UAG, and UGA), UAA demonstrated a higher preference than UAG and UGA, ranging from 1.291 to 1.367. Additionally, most codons terminated with A/U, ranging from 70.01% to 70.19%.

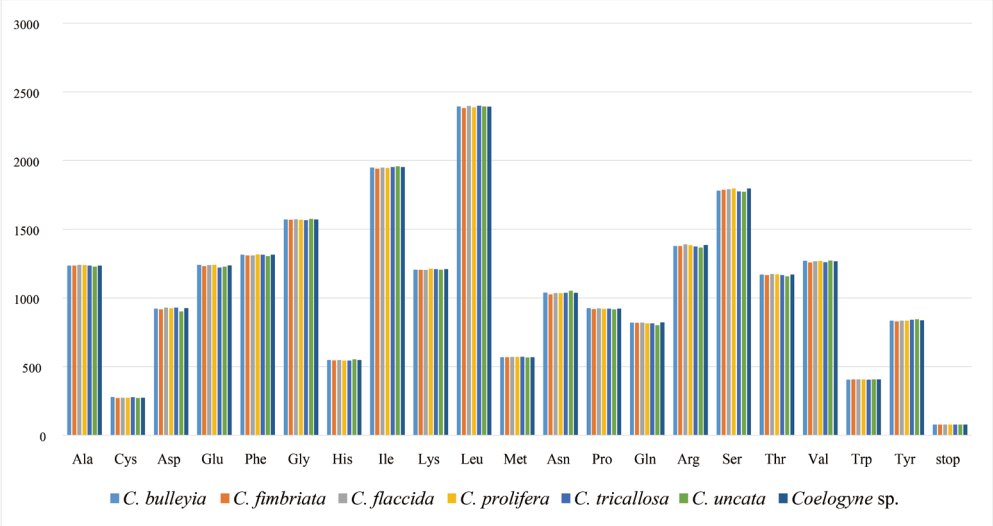

Figure 2. The number of amino acids and stop codons in seven *Coelogyne* species.

3.3. Repeat Sequence Analysis

In the seven *Coelogyne* plastomes, a comprehensive analysis revealed a total of 432 simple sequence repeats (SSRs) and 314 long repeats. Regarding SSRs, the counts were as follows: 59 in *C. bulleyia*, 60 in *C. fimbriata*, 64 in *C. flaccida*, 61 in *C. prolifera*, 69 in *C. tricallosa*, 54 in *C. uncata*, and 65 in *Coelogyne* sp. (Figure 3, Supplementary Table S4). *C. tricallosa* (69) exhibited the highest number of SSRs, while the lowest was in *C. uncata* (54). Mononucleotide repeats, ranging from 30 to 45, were the most prevalent among the identified SSRs, followed by dinucleotide (excluding *C. fimbriata*), ranging from 10 to 13. For hexanucleotide repeats, five species appeared at least once, while the remaining two species did not (Figure 3A). Most mononucleotide repeats primarily consisted of A/T, ranging from 27 to 45. AT/AT repeats were the most prevalent among dinucleotide repeats, varying

from 6 to 11 (Supplementary Table S4). The LSC region contained the majority of SSRs (Figure 3B), with 42 (59), 45 (60), 46 (64), 44 (61), 48 (69), 45 (54), and 46 (65), in *C. bulleyia*, *C. fimbriata*, *C. flaccida*, *C. prolifera*, *C. tricallosa*, *C. uncata*, and *Coelogyne* sp., respectively.

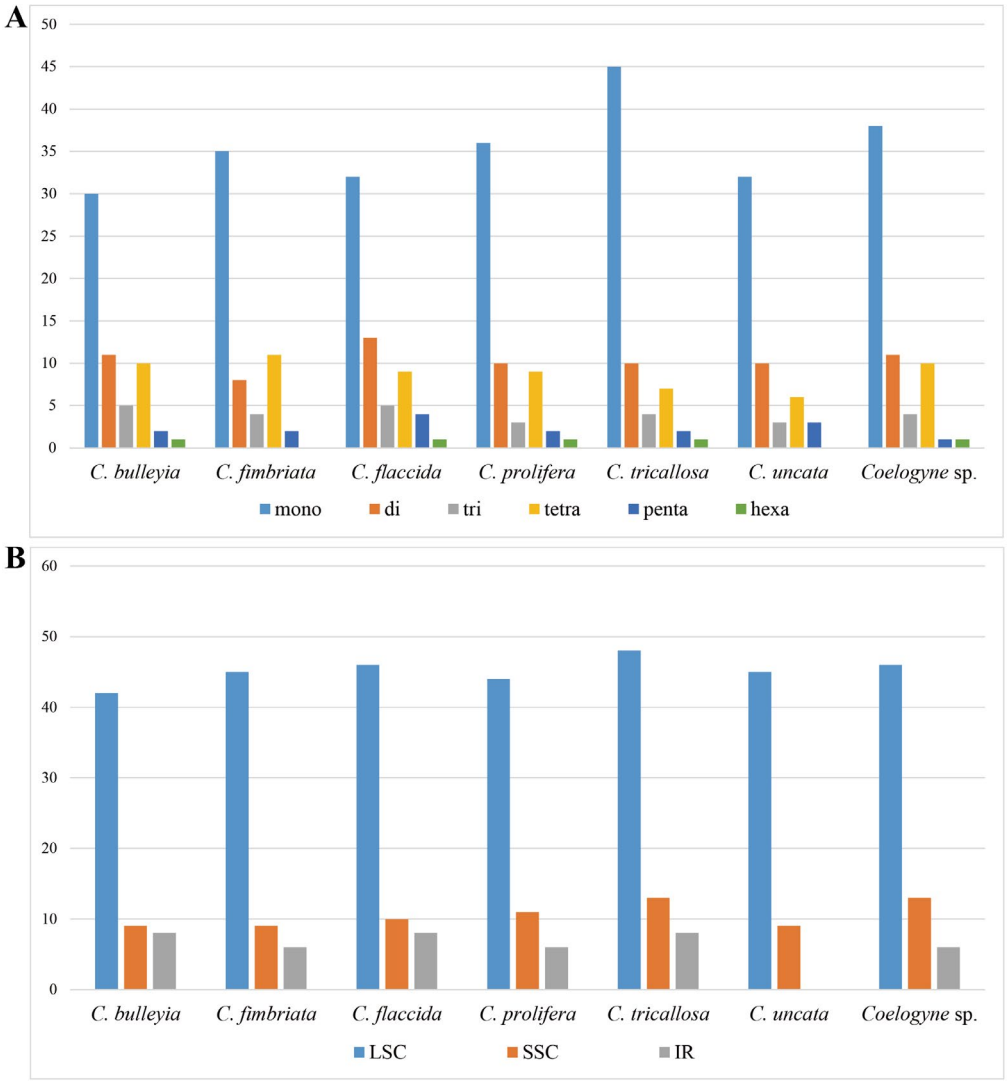

Figure 3. SSRs analyses were conducted on the seven *Coelogyne* plastomes. (**A**) Number and type of SSRs. (**B**) Frequency of SSRs in LSC, IR, and SSC regions.

Concerning long repeats, a total of 314 long repeats were identified and categorized into palindrome, forward, reverse, and complementary types. The long repeats of the *C. bulleyia*, *C. fimbriata*, *C. flaccida*, *C. prolifera*, *C. tricallosa*, *C. uncata*, and *Coelogyne* sp., were 30, 49, 49, 49, 49, 39, and 49, respectively (Figure 4, Supplementary Table S5). Among these long repeats, palindrome repeats exhibited the highest abundance and ranged from 19 (*C. bulleyia*) to 32 (*Coelogyne* sp.), followed by forward repeats, ranging from 10 (*C. bulleyia*) to 20 (*C. tricallosa*). Reverse repeats were exclusively detected in *C. bulleyia*, *C. fimbriata*, and *C. prolifera*, while the complementary repeats were not identified in any species (Figure 4A). Based on the length of the repeat sequences, all repeat sequences were divided into three categories. Most of the long repeats in all species were in the range of 30–39 bp, succeeded by the 40–49 bp and 50+ bp categories, except for *C. prolifera*, where there were more long repeats in the 50+ bp range than in the 40–49 bp range (Figure 4B).

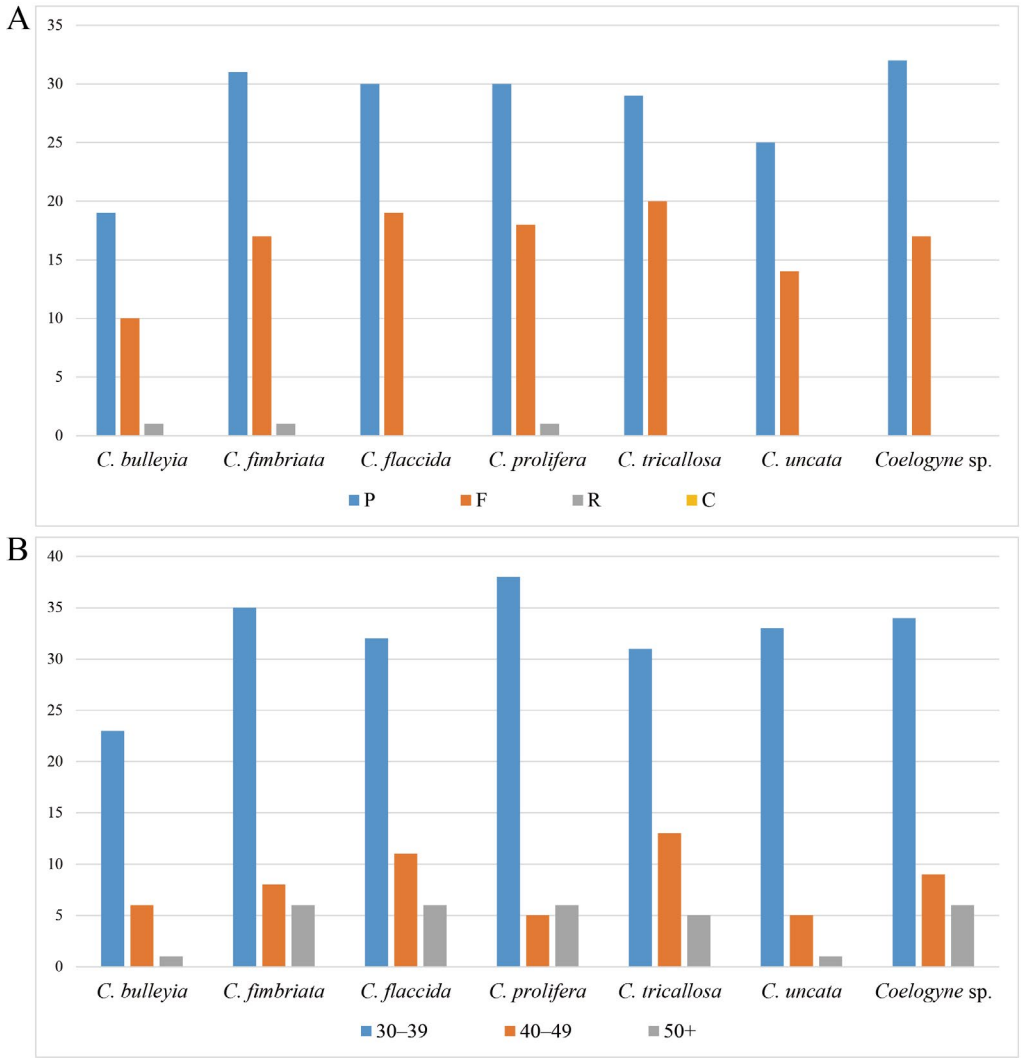

Figure 4. Long repeats analysis of seven *Coelogyne* plastomes. (**A**) Number and type of long repeats. (**B**) Frequency of long repeats by length.

3.4. IR Expansion and Contraction

Comparison of the LSC/IRb/SSC/IRa (LSC/IRb (JLB), IRb/SSC (JSB), SSC/IRa (JSA), and IRa/LSC (JLA)) boundaries for seven *Coelogyne* species (Figure 5). In general, these boundaries were relatively conserved in the genus *Coelogyne*, but some differences were observed. The JLB boundary was located between the *rpl22* and *rps19* genes, except for *C. uncata*, where the *rpl22* gene spanned across the JLB boundary by 51 bp. At the JLA boundary, the *psbA* gene was located to its right with variable distances spanning from 88 bp to 136 bp. The *ndhF* gene extended across the JSB boundary and was primarily positioned within the SSC region, exhibiting lengths spanning from 2188 bp to 2204 bp. The *ycf1* gene spanned the JSA boundary, with its length differing between the IRa and SSC regions. In the SSC region, the range spanned from 4427 bp (*C. uncata*) to 4566 bp (*C. bulleyia*), while in the IRa region, it varied from 990 bp (*C. tricallosa*) to 1036 bp (*C. uncata*) (Figure 5).

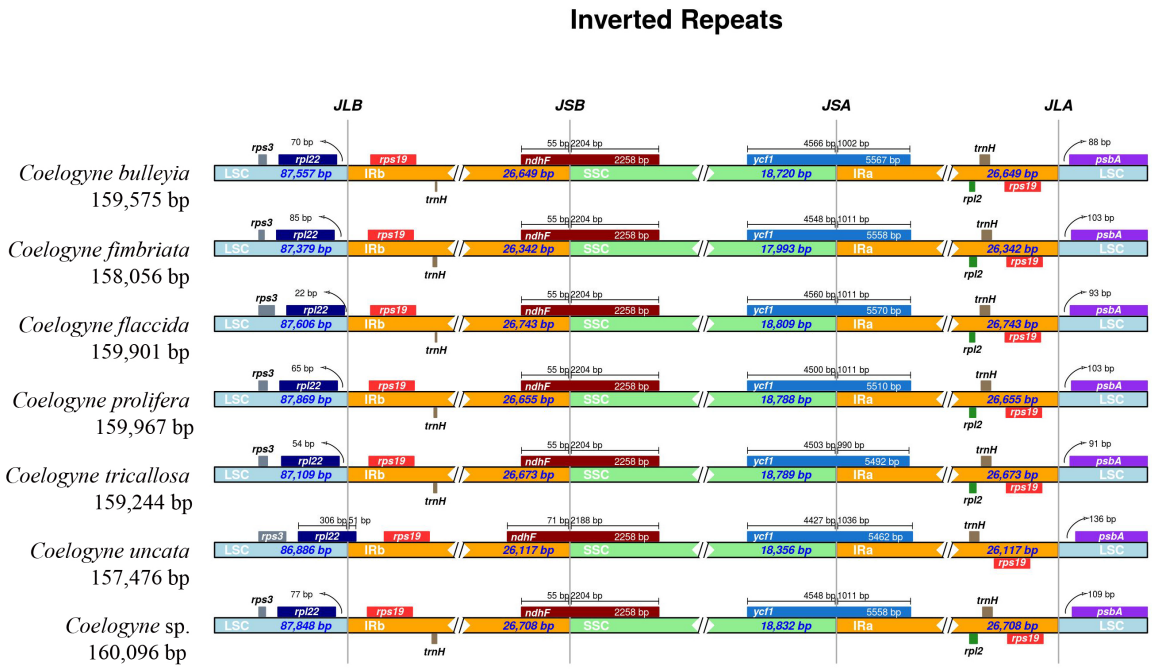

Figure 5. Comparison of the LSC/IR/SSC boundaries for seven *Coelogyne* plastomes.

3.5. Sequence Divergence Analysis

To analyze sequence divergence, the plastomes of seven *Coelogyne* species were compared using mVISTA, with *C. bulleyia* serving as the reference (Supplementary Figure S2). The findings indicated higher conservation in the IR regions, with protein-coding genes exhibiting lower divergence than intergenic sequences.

To identify highly variable hotspots of the *Coelogyne*, both protein-coding and intergenic regions were extracted for the computation of nucleotide diversity (Pi) (Figure 6, Supplementary Table S6). The findings displayed Pi values spanning from 0 to 0.08467 (*ycf1*) for protein-coding genes and from 0 to 0.13290 (*psbB-psbT*) for intergenic regions. The intergenic sequences and protein-coding genes within the SSC regions exhibited elevated average Pi values, with the former registering an average Pi of 0.09036, and the latter recording an average Pi of 0.02714. The LSC region followed, with 0.01521 for protein-coding genes and 0.06366 for intergenic sequences. The IR region showed the lowest nucleotide diversity, with 0.00494 and 0.01689 for protein-coding genes and intergenic sequences, respectively. For the protein-coding genes, three highly variable (Pi > 0.04) genes were detected: *rpl20* (0.04145), *rpl32* (0.04242), and *ycf1* (0.08467) (Figure 6A). For the intergenic regions, seven highly divergent regions (Pi > 0.09) were identified: *atpB-rbcL* (0.09647), *psbK-psbI* (0.09746), *rps8-rpl14* (0.09914), *rps16-trnQ*^{UUG} (0.10348), *psaC-ndhE* (0.11394), *ndhF-rpl32* (0.12537), and *psbB-psbT* (0.13290) (Figure 6B).

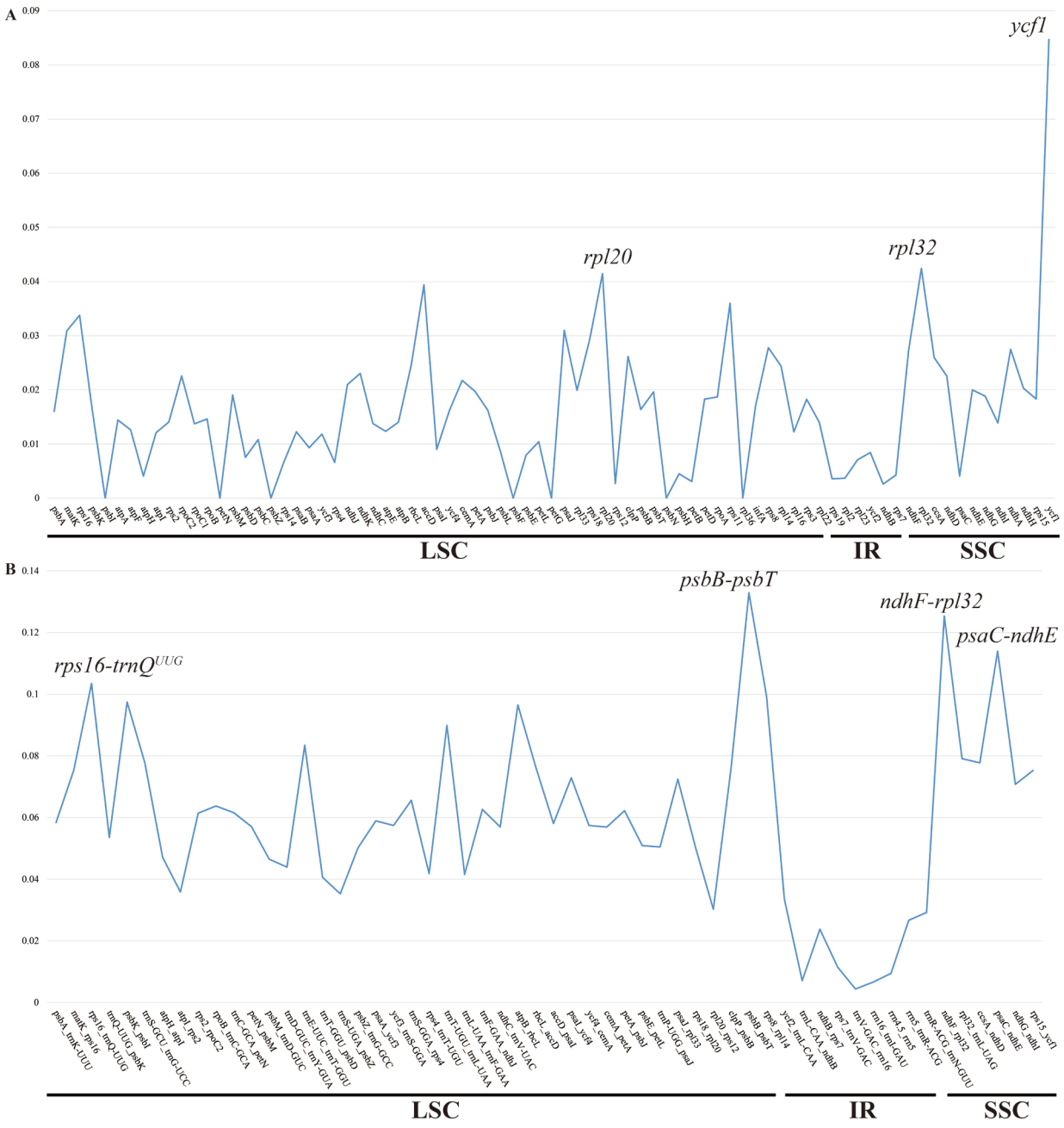

Figure 6. The Pi values for both protein-coding and intergenic sequences. (**A**) protein-coding genes; (**B**) intergenic sequences.

3.6. Selection Pressure Analysis

The non-synonymous substitution rate (Ka), synonymous substitution rate (Ks), and the Ka/Ks ratio of 79 protein-coding genes across the plastomes of seven species in the *Coelogyne* were calculated to evaluate the selective pressure acting on these genes (Supplementary Table S7). The results showed that the ranges for Ka, Ks, and the Ka/Ks ratio were found to be 0 to 0.0498, 0 to 0.1359, and 0 to 4.0934, respectively. In a comparative analysis of seven *Coelogyne* species, *C. uncata*, *C. tricallosa*, and *C. fimbriata* were found to have the highest average Ka values. Similarly, the species that exhibited the highest average Ks values were *C. uncata*, *C. tricallosa*, and *Coelogyne* sp. When considering the ratio of Ka/Ks, *C. fimbriata*, *Coelogyne* sp., and *C. bulleyia* emerged as the top three species with the highest values.

3.7. Phylogenetic Analysis

Phylogenetic reconstructions were conducted employing complete plastomes and 79 protein-coding genes from 45 samples; the results showed that *Coelogyne* s.l. *sensu* Chase et al. (2021) was monophyletic (Figure 7, Supplementary Figure S3). Furthermore, the results robustly supported the division of *Coelogyne* s.l. into five clades: *C. rochussenii*, *C. sulphurea*, *C. ventricose*, *C. uncata*, *C. apoensis*, and *C. cootesii* formed clade I, which was strongly supported as sister to the remaining taxa of *Coelogyne* (BS = 100, PP = 1.00). The members of Clade II were *C. missionariorum*, *C. cantonensis*, *C. bulleyia*, *C. leveilleana*, *C. wenshanica*, and *C. kouytcheensis*. Clade III contained six species, including *C. protracta*, *C. chen-tsii*, *C. niana*, *C. mandarinorum*, *C. tricallosa*, and *C. uniflora*. Clade IV contained *C. prolifera*, *C. barbata*, *C. fimbriata*, *Coelogyne* sp., and *C. ovalis*. Lastly, clade V consisted of the remaining species, *C. flaccida*, *C. porrecta*, *C. fusca*, *C. chinensis*, *C. articulata*, *C. cristata*, *C. viscosa*, *C. corymbosa*, *C. punctulata*, *C. gardneriana*, *C. pallida*, and *C. imbricata*.

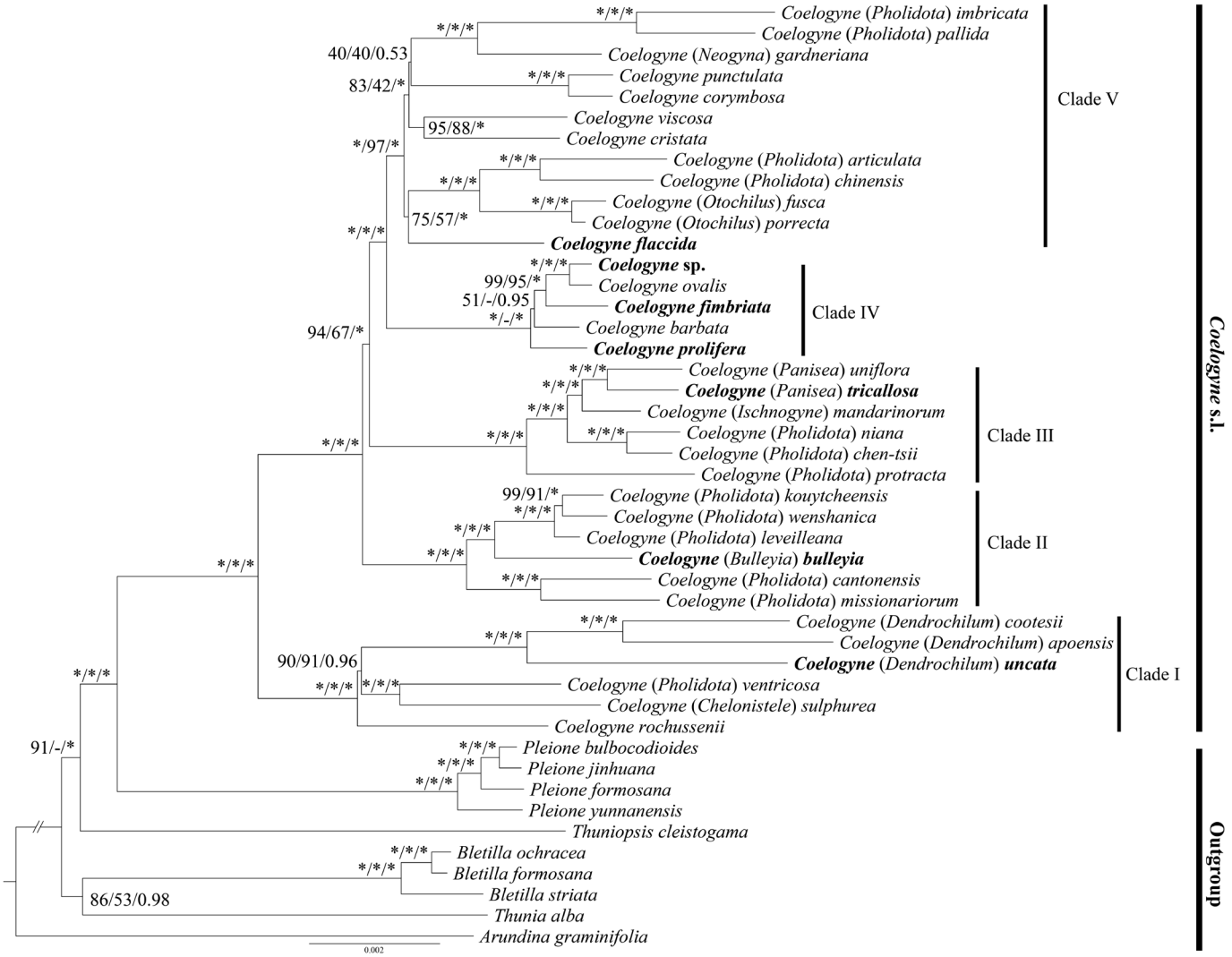

Figure 7. The ML tree of *Coelogyne* s.l. was constructed using the complete plastome. Bootstrap percentages and Bayesian posterior probabilities (BS_{ML}, BS_{MP}, and PP) are denoted near the nodes, with asterisks (*) representing nodes supported by 100% bootstrap or 1.00 posterior probability. The symbol "-" signifies discrepancies observed among ML, MP, and BI tree topologies. The bracket indicates the accepted generic name before Chase et al. (2021) [3]. The data in this study are highlighted in black.

4. Discussion

4.1. Plastome Characteristics

In the present investigation, seven plastomes of *Coelogyne* (*C. bulleyia*, *C. fimbriata*, *C. flaccida*, *C. prolifera*, *C. tricallosa*, *C. uncata*, and an unknown taxa, *Coelogyne* sp.) were newly sequenced and assembled. All plastomes of *Coelogyne* exhibited a typical quadripartite structure (Figure 1), consistent with the most previously reported plastome structure of Orchidaceae [19,21,22,38–40]. The plastomes of the seven species of *Coelogyne* encoded the same number of genes (132 genes), including 86 protein-coding genes, eight rRNA genes, and 38 tRNA genes (Table 1), the observed gene content was similar to the other studies [22,39]. In previous studies, the reported plastome lengths of Arethuseae ranged from 156,939 bp (*Bletilla sinensis*) (ON243844) to 160,433 bp (*C. flaccida*) (OK624421). The size of *Coelogyne* plastomes varied between 157,476 bp (*C. uncata*) and 160,096 bp (*Coelogyne* sp.), consistent with the previously reported plastomes range in Arethuseae. The LSC region ranged from 86,886 to 87,869 bp, the SSC ranged from 17,987 to 18,834 bp, and the IR region ranged from 26,117 to 26,746 bp, which were similar to previously reported closely related species [22,39].

The GC content of the complete plastomes exhibited a range of 37.3% to 37.5%. This range aligns with the documented GC content range of Arethuseae, spanning from 37.1% in *Mengzia foliosa* (OR159908) to 37.5% in *C. cantonensis* (OQ985352). The descending order of GC content was as follows: IR (43.2–43.3%), LSC (35.2–35.3%), and SSC (30.3–31.1%) (Table 1). These results were primarily attributed to the localization of all rRNA genes within the IR regions, which were consistent with the previously reported plastomes of Orchidaceae [22,39,41]. In this study, 18 genes contained one or two introns among the seven *Coelogyne* species (Supplementary Table S2), which was consistent with the other Orchidaceae [41,42]. Previous studies had shown that gene loss and pseudogenization were prevalent in the plastomes of Orchidaceae, with the most common being the NADH dehydrogenase (*ndh*) genes [43], such as Aeridinae [19], *Bulbophyllum* [38], *Dendrobium* [44], and *Renanthera* [45]. There were changes in the plastome structure of some Orchidaceae species, such as the inversion of $trnS^{GCU}$-$trnS^{GCA}$ in *Apostasia wallichii* [46] and the inversion of *clpP-accD* in *Uncifera acuminata* [19]. These phenomena were not detected in *Coelogyne*. In general, the plastome of *Coelogyne* exhibited relatively conserved gene order, GC content, and structure.

The contraction and expansion of the IR regions were widespread phenomena in angiosperms and were regarded as the primary reason for plastome size changes [47]. A comparison of the LSC/IRb/SSC/IRa boundaries of seven *Coelogyne* plastomes showed that IR/SC boundaries were conserved (Figure 5). However, some differences in the JLB boundary were observed; the *rpl22* gene in *C. uncata* spanned the JLB boundary, while in the remaining species, the *rpl22* gene was located to its left. The JSA boundary was positioned within the *ycf1* gene, and the JLA boundary was positioned to the left of the *psbA* gene. These phenomena were consistent with previous studies for other related taxa [22,39].

The ratio of Ka/Ks was widely used to evaluate the selective pressures on genes and their rates of evolution [48]. This metric played a crucial role in advancing the understanding of the mechanisms underlying adaptive evolution. Our results showed that the majority of genes demonstrated Ka/Ks ratios less than one (Ka/Ks < 1), indicating that purifying selection had been the primary evolutionary force influencing most protein-coding genes (Supplementary Table S7).

4.2. Codon Usage and Repeat Sequences Analysis

Codon usage bias (CUB) analysis was commonly employed to unravel the genetic structure and evolutionary patterns across various species [49]. The findings revealed a distinct conservation of codon usage patterns within the *Coelogyne* plastomes. Leucine

(Leu) was the most abundant among these codons, while cysteine (Cys) exhibited the lowest occurrence (Figure 2). The GCU and GCG codons exhibited the highest and lowest RSCU values, respectively. For stop codons (UAA, UAG, and UGA), UAA demonstrated a higher preference than UAG and UGA. Most codons terminated with A/U (Supplementary Table S3). These results aligned consistently with prior investigations of CUB within the Orchidaceae [39,45,50].

Repeat sequences are essential for various applications, such as species identification, genetic diversity, and phylogenetic studies [51–53]. In the seven *Coelogyne* plastomes, a total of 432 SSRs were calculated. Mononucleotide repeats emerged as the predominant type among the detected SSRs, while hexanucleotide repeats were the least frequent. Most mononucleotide repeats primarily consisted of A/T. The LSC region harbored the majority of SSRs, while fewer were observed in the SSC and IR regions (Figure 3, Supplementary Table S4). These results were similar to previous reports in other related taxa [22,39]. Additionally, our study identified a total of 314 long repeats. The most prevalent among these long repeats was the palindrome, with forward repeats following closely. The majority of the long repeats in all species fell within the 30–39 bp range (Figure 4, Supplementary Table S5). These results were in concordance with the reported results in other Orchidaceae [54,55].

4.3. Molecular Markers Investigation

The mVISTA and nucleotide diversity (Pi) analysis of *Coelogyne* plastomes showed higher conservation in the IR regions, with protein-coding genes exhibiting lower divergence than intergenic sequences (Supplementary Figure S2). These results aligned with earlier studies [22,39,45,50].

Molecular markers are invaluable for species identification, enabling the differentiation of distinct taxa and clarifying interspecific relationships. They also facilitate the assessment of genetic diversity in populations or individual species, providing essential data for evolutionary research and conservation strategies. Li et al. constructed the phylogenetic trees of Arethuseae based on three molecular markers (ITS, *matK*, and *trnL*), and the results showed that *Thuniopsis* is a distinct genus [4].

With advancements in sequencing technology, many plastomes of Orchidaceae have been sequenced in recent years, leading to the proposal of various plastid markers for Orchidaceae [22,39,56,57]. For instance, Li et al. identified eight molecular markers (*atpB-rbcL*, *psaC-ndhE*, *trnW*^{CCA}*-trnP*^{UGG}, *ndhF-rpl32*, *rps16-trnQ*^{UUG}, *petN-petM*, *ndhA*, and *ndhJ*) in *Thuniopsis* and closely related genera [39]. Additionally, they selected six highly variable regions (*rps16-trnQ*^{UUG}, *ndhF-rpl32*, *rpoB-trnC*^{GCA}, *rps11-rpl36*, *ycf1*, and *ndhA*) in *Pholidota* (synonym of *Coelogyne*) [22], and four mutation hotspot regions (*ycf1*, *ndhF-rp132*, *psaC-ndhE*, and *rp132-trnL*^{UAG}) as effective molecular markers in Coelogyninae [57]. In this study, we identified eight hypervariable regions (*atpB-rbcL*, *psbK-psbI*, *rps8-rpl14*, *rps16-trnQ*^{UUG}, *psaC-ndhE*, *ndhF-rpl32*, *psbB-psbT*, and *ycf1*) by selecting those with the highest Pi values (Figure 6, Supplementary Table S6). Notably, six regions (*atpB-rbcL*, *rps16-trnQ*^{UUG}, *psaC-ndhE*, *ndhF-rpl32*, *psbB-psbT*, and *ycf1*) were previously reported in *Coelogyne* and related taxa [22,39,56,57]. However, two hypervariable regions (*psbK-psbI* and *rps8-rpl14*) were reported here for the first time in *Coelogyne*. These highly variable regions are pivotal in species identification and phylogenetic analyses of *Coelogyne*.

4.4. Phylogenetic Analysis

The phylogenetic analysis of the genus *Coelogyne* was unclear, and it was commonly regarded as one of the most complex taxa of the subtribe Coelogyninae. Previous phylogenetic studies of *Coelogyne* and related taxa based on a few nuclear or plastid sequences showed low support and unstable topology [3–5,12–15]. Gravendeel et al. defined the taxa as two clades based on the combined analysis of three DNA regions (ITS, RFLPs, and

matK) [12]. Based on three combined DNA regions (ITS, *matK*, and *trnL*), Li et al. defined the group as two clades [4]. In order to solve the problem of monophyly in the genus *Coelogyne*, Chase et al. expanded the genus *Coelogyne* to include 14 genera, such as *Bulleyia*, *Dendrochilum*, *Panisea*, and *Pholidota* [3]. The above studies were based on a few traditional molecular markers, which were inconsistent in clades and species composition within clades, and there were problems of low support and unstable topology.

Li et al. reconstructed the phylogenetic tree of Coelogyninae based on complete plastomes and protein-coding genes; the results showed that the clade of *Coelogyne* s.l. could be divided into five subclades [57]. Han et al. constructed the phylogenetic tree of Arethuseae using the protein-coding genes. The results indicated the subdivision of *Coelogyne* into four clades with high support and stable topology [58]. In this study, phylogenetic reconstructions were conducted employing complete plastomes and 79 protein-coding genes (Figure 7, Supplementary Figure S3). These results all showed that *Coelogyne* s.l. was monophyletic, consistent with previous studies [3,4,57,58]. *Coelogyne rochussenii*, *C. sulphurea*, *C. ventricose*, *C. uncata*, *C. apoensis*, and *C. cootesii* formed clade I, which was strongly supported as sister to the remaining members of *Coelogyne* (BS = 100, PP = 1.00). The remaining members of *Coelogyne* were divided into four clades, which were consistent with Li et al. [22,57] but in conflict with the results of some other studies [4,5,12]. This study supported the idea that plastomes were effective molecular markers for elucidating the phylogenetic relationships of these intricate taxa.

5. Conclusions

In this study, seven representative plastomes of *Coelogyne* (*C. bulleyia*, *C. fimbriata*, *C. flaccida*, *C. prolifera*, *C. tricallosa*, *C. uncata*, and an unknown taxa, *Coelogyne* sp.) were newly assembled and annotated. Our results showed that gene order, GC content, and structure of *Coelogyne* plastomes were conserved. The majority of genes demonstrated Ka/Ks ratios less than one (Ka/Ks < 1), indicating purifying selection. Additionally, eight hypervariable regions (*atpB-rbcL*, *psbK-psbI*, *rps8-rpl14*, *rps16-trnQ*^{UUG}, *psaC-ndhE*, *ndhF-rpl32*, *psbB-psbT*, and *ycf1*) were identified, and these identified hypervariable regions hold significant potential for conservation and horticulture. Phylogenetic analysis indicated that *Coelogyne* s.l. *sensu* Chase et al. (2021) was monophyletic. Furthermore, the results robustly support the division of *Coelogyne* s.l. into five clades. The results are expected to offer a theoretical foundation for further research on germplasm identification, phylogenetic analysis, evolutionary patterns, and resource application of *Coelogyne* and Arethuseae. Based on these findings, future research could employ molecular markers within breeding initiatives to develop superior ornamental cultivars.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/horticulturae11020144/s1, Figure S1: Mauve analysis of seven *Coelogyne* plastomes; Supplementary Figure S2: The mVISTA analysis of seven *Coelogyne* plastomes; Supplementary Figure S3: Phylogenetic analysis of 79 protein-coding genes, with asterisks (*) representing nodes supported by 1.00 posterior probability. The data in this study are highlighted in black; Supplementary Table S1: GenBank accession numbers of the species used in this study; Supplementary Table S2: The plastome annotation gene information of *Coelogyne*; Supplementary Table S3: The relative synonymous codon usage values of all 64 codons for the seven *Coelogyne* plastomes; Supplementary Table S4: SSRs identified on the plastome of seven *Coelogyne* species; Supplementary Table S5: Long repeat sequences identified on the plastome of seven *Coelogyne* species; Supplementary Table S6: The Pi values of protein-coding and intergenic sequences; Supplementary Table S7: The Ka, Ks, and Ka/Ks values of 79 protein-coding genes.

Author Contributions: Conceptualization, X.G. and X.Z.; methodology, S.L., R.L. and S.T.; software, S.L. and R.L.; data curation, S.L., R.L., Y.C. and Y.Y.; writing—original draft preparation, S.L., R.L., S.T., Y.C. and Y.Y.; writing—reviewing and editing, S.L., R.L., S.T., Y.C. and Y.Y.; validation, X.G., X.Z. and S.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2023YFD1600505, and the Nature Science Foundation of Fujian Province, China, grant number 2024J01394.

Data Availability Statement: The seven plastomes generated in this study are available in NCBI (https://www.ncbi.nlm.nih.gov) with accession numbers PQ723622-PQ723628. For the raw data, interested researchers can contact the corresponding author directly to obtain it. The voucher specimens were deposited in the Herbarium of Fujian Agriculture and Forestry University, with voucher numbers MHLi or172, MHLi or 167, MHLi or150, MHLi or155, MHLi or164, MHLi or161, and MHLi or149.

Acknowledgments: We acknowledge the technical support of laboratory staff during the conducting of laboratory experiments.

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

References

1. Chase, M.W.; Cameron, K.M.; Freudestein, J.V.; Pridgeon, A.M.; Salazar, G.; Van den Berg, C.; Schuiteman, A. An update classification of Orchidaceae. *Bot. J. Linn. Soc.* **2015**, *177*, 151–174. [CrossRef]
2. Christenhusz, M.J.M.; Byng, J.W. The number of known plants species in the world and its annual increase. *Phytotaxa* **2016**, *261*, 201–217. [CrossRef]
3. Chase, M.W.; Gravendeel, B.; Sulistyo, B.P.; Wti, R.K.; Schuiteman, A. Expansion of the orchid genus *Coelogyne* (Arethuseae; Epidendroideae) to include *Bracisepalum*, *Bulleyia*, *Chelonistele*, *Dendrochilum*, *Dickasonia*, *Entomophobia*, *Geesinkorchis*, *Gynoglottis*, *Ischnogyne*, *Nabaluia*, *Neogyna*, *Otochilus*, *Panisea* and *Pholidota*. *Phytotaxa* **2021**, *510*, 94–134. [CrossRef]
4. Li, L.; Ye, D.P.; Niu, M.; Yan, H.F.; Wen, T.L.; Li, S.J. *Thuniopsis*: A new orchid genus and phylogeny of the tribe Arethuseae (Orchidaceae). *PLoS ONE* **2015**, *10*, e0132777. [CrossRef]
5. Huang, W.C.; Liu, Z.J.; Jiang, K.; Luo, Y.B.; Jin, X.H.; Zhang, Z.; Xu, R.H.; Muchuku, J.K.; Musungwa, S.S.; Yukawa, T.; et al. Phylogenetic analysis and character evolution of tribe Arethuseae (Orchidaceae) reveal a new genus *Mengzia*. *Mol. Phylogenet. Evol.* **2022**, *167*, 107362. [CrossRef]
6. Chen, S.C.; Liu, Z.J.; Zhu, G.H.; Lang, K.Y.; Ji, Z.H.; Luo, Y.B.; Jin, X.H.; Cribb, P.J.; Wood, J.J.; Gale, S.W.; et al. Orchidaceae. In *Flora of China*; Wu, Z., Raven, P.H., Hong, D., Eds.; Science Press: Beijing, China, 2009; pp. 315–325.
7. Zhou, X.; Cheng, Z.; Liu, Q.; Zhang, J.; Hu, A.; Huang, M.; Hu, C.; Tian, H. An updated checklist of Orchidaceae for China, with two new national records. *Phytotaxa* **2016**, *276*, 1–148. [CrossRef]
8. Pridgeon, A.M.; Cribb, P.J.; Chase, M.W.; Rasmussen, F.N. *Genera Ochidacearum Volume 4: Epidendroideae (Part One)*; Oxford University Press: Oxford, UK, 2005.
9. Pramanick, D.D. Pharmacognostic studies on the pseudobulb of *Coelogyne cristata* Lindl. (Orchidaceae)-An epiphytic orchid of ethno-medicinal importance. *J. Pharmacogn. Phytochem.* **2016**, *5*, 120–123.
10. Singh, N.; Kumaria, S. A combinational phytomolecular-mediated assessment in micropropagated plantlets of *Coelogyne ovalis* Lindl.: A horticultural and medicinal orchid. *Proc. Natl. Acad. Sci. India Sec. B Biol. Sci.* **2020**, *90*, 455–466. [CrossRef]
11. Lindley, J. *Collectanea Botanica*; Richard and Arthur Taylor: London, UK, 1821; Volume 4.
12. Gravendeel, B.; Chase, M.W.; de Vogel, E.F.; Roos, M.C.; Mes, T.H.; Bachmann, K. Molecular phylogeny of *Coelogyne* (Epidendroideae; Orchidaceae) based on plastid RFLPs, *matK*, and nuclear ribosomal ITS sequences: Evidence for polyphyly. *Am. J. Bot.* **2001**, *88*, 1915–1927. [CrossRef]
13. Van den Berg, C.; Goldman, D.H.; Freudenstein, J.V.; Pridgeon, A.M.; Cameron, K.M.; Chase, M.W. An overview of the phylogenetic relationships within Epidendroideae inferred from multiple DNA regions and recircumscription of Epidendreae and Arethuseae (Orchidaceae). *Am. J. Bot.* **2005**, *92*, 613–624. [CrossRef]
14. Górniak, M.; Paun, O.; Chase, M.W. Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, *Xdh*: Congruence with organellar and nuclear ribosomal DNA results. *Mol. Phylogenet. Evol.* **2010**, *56*, 784–795. [CrossRef] [PubMed]
15. Li, M.H.; Zhang, G.Q.; Lan, S.R.; Liu, Z.J.; China Phylogeny Consortium. A molecular phylogeny of Chinese orchids. *J. Syst. Evol.* **2016**, *54*, 349–362. [CrossRef]

16. Jansen, R.K.; Ruhlman, T.A. Plastid Genomes of Seed Plants. In *Genomics of Chloroplasts and Mitochondria*; Springer: Dordrecht, The Netherlands, 2012.
17. Wicke, S.; Schneeweiss, G.M.; Depamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [CrossRef] [PubMed]
18. Givnish, T.J.; Spalink, D.; Ames, M.; Lyon, S.P.; Hunter, S.J.; Zuluaga, A.; Iles, W.J.; Clements, M.A.; Arroyo, M.T.; Leebens-Mack, J.; et al. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. Biol. Sci. B* **2015**, *282*, 2108–2111. [CrossRef]
19. Liu, D.K.; Tu, X.D.; Zhao, Z.; Zeng, M.Y.; Zhang, S.; Ma, L.; Zhang, G.Q.; Wang, M.M.; Liu, Z.J.; Lan, S.R.; et al. Plastid phylogenomic data yield new and robust insights into the phylogeny of *Cleisostoma–Gastrochilus* clades (Orchidaceae, Aeridinae). *Mol. Phylogenet. Evol.* **2020**, *145*, 106729. [CrossRef]
20. Serna-Sánchez, M.A.; Pérez-Escobar, O.A.; Bogarín, D.; Torres-Jimenez, M.F.; Alvarez-Yela, A.C.; Arcila-Galvis, J.E.; Hall, C.F.; de Barros, F.; Pinheiro, F.; Dodsworth, S.; et al. Plastid phylogenomics resolves ambiguous relationships within the orchid family and provides a solid timeframe for biogeography and macroevolution. *Sci. Rep.* **2021**, *11*, 6858. [CrossRef]
21. Tu, X.D.; Liu, D.K.; Xu, S.W.; Zhou, C.Y.; Gao, X.Y.; Zeng, M.Y.; Zhang, S.; Chen, J.L.; Ma, L.; Zhou, Z.; et al. Plastid phylogenomics improves resolution of phylogenetic relationship in the *Cheirostylis* and *Goodyera* clades of Goodyerinae (Orchidoideae, Orchidaceae). *Mol. Phylogenet. Evol.* **2021**, *164*, 107269. [CrossRef]
22. Li, L.; Wang, W.; Zhang, G.; Wu, K.; Fang, L.; Li, M.; Liu, Z.; Zeng, S. Comparative analyses and phylogenetic relationships of thirteen *Pholidota* species (Orchidaceae) inferred from complete chloroplast genomes. *BMC Plant Biol.* **2023**, *23*, 269. [CrossRef]
23. Beier, S.; Thiel, T.; Münch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [CrossRef]
24. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. REPuter: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [CrossRef]
25. Darling, A.E.; Mau, B.; Perna, N.T. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **2010**, *5*, e11147. [CrossRef] [PubMed]
26. Amiryousefi, A.; Hyvonen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [CrossRef] [PubMed]
27. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]
28. Glez-Peña, D.; Gómez-Blanco, D.; Reboiro-Jato, M.; Fdez-Riverola, F.; Posada, D. ALTER: Program-oriented format conversion of DNA and protein alignments. *Nucleic Acids Res.* **2010**, *38*, W14–W18. [CrossRef]
29. Wang, D.; Zhang, Y.; Zhang, Z.; Zhu, J.; Yu, J. KaKs_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* **2010**, *8*, 77–80. [CrossRef]
30. Xia, X. Dambe7: New and improved tools for data analysis in molecular biology and evolution. *Mol. Biol. Evol.* **2018**, *35*, 1550–1552. [CrossRef]
31. Brudno, M.; Malde, S.; Poliakov, A.; Do, C.B.; Couronne, O.; Dubchak, I.; Batzoglou, S. Glocal alignment: Finding rearrangements during alignment. *Bioinformatics* **2003**, *19*, i54–i62. [CrossRef]
32. Rozas, J.; Sánchez-DelBarrio, J.C.; Messeguer, X.; Rozas, R. Dnasp, dna polymorphism analyses by the coalescent and other methods. *Bioinformatics* **2003**, *19*, 2496–2497. [CrossRef]
33. Capella-Gutierrez, S.; Silla-Martinez, J.M.; Gabaldon, T. Trimal: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef]
34. Zhang, D.; Gao, F.; Jakovlić, I.; Zou, H.; Zhang, J.; Li, W.X.; Wang, G.T. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol. Ecol. Resour.* **2020**, *20*, 348–355. [CrossRef]
35. Miller, M.A.; Pfeiffer, W.; Schwartz, T. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. In Proceedings of the 2010 Gateway Computing Environments Workshop, GCE, New Orleans, LA, USA, 14 November 2010.
36. Stamatakis, A.; Hoover, P.; Rougemont, J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst. Biol.* **2008**, *57*, 758–771. [CrossRef] [PubMed]
37. Felsenstein, J. Confidence limits on phylogenies: An approach using the bootstrap. *Evol. Int. J. Org. Evol.* **1985**, *39*, 783–791. [CrossRef]
38. Zavala-Páez, M.; Vieira, L.D.N.; Baura, V.A.D.; Balsanelli, E.; Souza, E.M.D.; Cevallos, M.C.; Smidt, E.D.C. Comparative plastid genomics of neotropical *Bulbophyllum* (Orchidaceae; Epidendroideae). *Front. Plant Sci.* **2020**, *11*, 799. [CrossRef] [PubMed]
39. Li, L.; Wu, Q.; Fang, L.; Wu, K.; Li, M.; Zeng, S. Comparative chloroplast genomics and phylogenetic analysis of *Thuniopsis* and closely related genera within Coelogyninae (Orchidaceae). *Front. Genet.* **2022**, *13*, 850201. [CrossRef]
40. Zhang, J.Y.; Liao, M.; Cheng, Y.H.; Feng, Y.; Ju, W.B.; Deng, H.N.; Li, X.; Plenkovic-Moraj, A.; Xu, B. Comparative chloroplast genomics of seven endangered *Cypripedium* species and phylogenetic relationships of Orchidaceae. *Front. Plant Sci.* **2022**, *13*, 911702. [CrossRef]

41. Tao, L.; Duan, H.; Tao, K.; Luo, Y.; Li, Q.; Li, L. Complete chloroplast genome structural characterization of two *Phalaenopsis* (Orchidaceae) species and comparative analysis with their alliance. *BMC Genom.* **2023**, *24*, 359. [CrossRef]
42. Guo, Y.Y.; Yang, J.X.; Bai, M.Z.; Zhang, G.Q.; Liu, Z.J. The chloroplast genome evolution of venus slipper (*Paphiopedilum*): IR expansion, SSC contraction, and highly rearranged SSC regions. *BMC Plant Biol.* **2021**, *21*, 248. [CrossRef]
43. Lin, C.S.; Chen, J.J.W.; Chiu, C.C.; Hsiao, H.C.W.; Yang, C.J.; Jin, X.H.; Leebens-Mack, J.; de Pamphilis, C.W.; Huang, Y.T.; Yang, L.H.; et al. Concomitant loss of NDH complex-related genes within chloroplast and nuclear genomes in some orchids. *Plant J.* **2017**, *90*, 994–1006. [CrossRef]
44. Zhitao, N.; Shuying, Z.; Jiajia, P.; Ludan, L.; Jing, S.; Xiaoyu, D. Comparative analysis of *Dendrobium* plastomes and utility of plastomic mutational hotspots. *Sci. Rep.* **2017**, *7*, 2073. [CrossRef]
45. Xiao, T.; He, L.; Yue, L.; Zhang, Y.; Lee, S.Y. Comparative phylogenetic analysis of complete plastid genomes of *Renanthera* (Orchidaceae). *Front. Genet.* **2022**, *13*, 998575. [CrossRef]
46. Niu, Z.; Pan, J.; Zhu, S.; Li, L.; Xue, Q.; Liu, W.; Ding, X. Comparative analysis of the complete plastomes of *Apostasia wallichii* and *Neuwiedia singapureana* (Apostasioideae) reveals different evolutionary dynamics of IR/SSC boundary among photosynthetic orchids. *Front. Plant Sci.* **2017**, *8*, 1713. [CrossRef] [PubMed]
47. Weng, M.L.; Ruhlman, T.A.; Jansen, R.K. Expansion of inverted repeat does not decrease substitution rates in *Pelargonium* plastid genomes. *New Phytol.* **2017**, *214*, 842–851. [CrossRef] [PubMed]
48. Kimura, M. The neutral theory of molecular evolution and the world view of the neutralists. *Genome* **1989**, *31*, 24–31. [CrossRef] [PubMed]
49. Yu, X.; Liu, J.; Li, H.; Liu, B.Y.; Zhao, B.Q.; Ning, Z.Y. Comprehensive analysis of synonymous codon usage bias for complete genomes and E2 gene of atypical porcine pestivirus. *Biochem. Genet.* **2021**, *59*, 799–812. [CrossRef]
50. Jiang, H.; Tian, J.; Yang, J.; Dong, X.; Zhong, Z.; Mwachala, G.; Zhang, C.; Hu, G.; Wang, Q. Comparative and phylogenetic analyses of six Kenya *Polystachya* (Orchidaceae) species based on the complete chloroplast genome sequences. *BMC Plant Biol.* **2022**, *22*, 177. [CrossRef]
51. Doorduin, L.; Gravendeel, B.; Lammers, Y.; Ariyurek, Y.; Chin-A-Woeng, T.; Vrieling, K. The complete chloroplast genome of 17 individuals of pest species *Jacobaea vulgaris*: SNPs, microsatellites and barcoding markers for population and phylogenetic studies. *DNA Res.* **2011**, *18*, 93–105. [CrossRef]
52. Yang, A.H.; Zhang, J.J.; Yao, X.H.; Huang, H.W. Chloroplast microsatellite markers in *Liriodendron tulipifera* (Magnoliaceae) and cross-species amplification in *L. chinense*. *Am. J. Bot.* **2011**, *98*, e123–e126. [CrossRef]
53. Kartzinel, T.R.; Shefferson, R.P.; Trapnell, D.W. Relative importance of pollen and seed dispersal across a Neotropical mountain landscape for an epiphytic orchid. *Mol. Ecol.* **2013**, *22*, 6048–6059. [CrossRef]
54. Zhao, Z.; Zeng, M.Y.; Wu, Y.W.; Li, J.W.; Zhou, Z.; Liu, Z.J.; Li, M.H. Characterization and comparative analysis of the complete plastomes of five *Epidendrum* (Epidendreae, Orchidaceae) species. *Int. J. Mol. Sci.* **2023**, *24*, 14437. [CrossRef]
55. Zhou, C.Y.; Zeng, M.Y.; Gao, X.Y.; Zhao, Z.; Li, R.Y.; Wu, Y.H.; Liu, Z.J.; Zhang, D.Y.; Li, M.H. Characteristics and comparative analysis of seven complete plastomes of *Trichoglottis* s.l. (Aeridinae, Orchidaceae). *Int. J. Mol. Sci.* **2023**, *24*, 14544. [CrossRef]
56. Jiang, K.; Miao, L.Y.; Wang, Z.W.; Ni, Z.Y.; Hu, C.; Zeng, X.H.; Huang, W.C. Chloroplast genome analysis of two medicinal *coelogyne* spp. (Orchidaceae) shed light on the genetic information, comparative genomics, and species identification. *Plants* **2020**, *9*, 1332. [CrossRef] [PubMed]
57. Li, L.; Wu, Q.P.; Zhai, J.W.; Wu, K.L.; Fang, L.; Li, M.Z.; Zeng, S.J.; Li, S.J. Comparative chloroplast genomics of 24 species shed light on the genome evolution and phylogeny of subtribe Coelogyninae (Orchidaceae). *BMC Plant Biol.* **2024**, *24*, 31. [CrossRef] [PubMed]
58. Han, S.; Wang, R.; Hong, X.; Wu, C.; Zhang, S.; Kan, X. Plastomes of *Bletilla* (Orchidaceae) and phylogenetic implications. *Int. J. Mol. Sci.* **2022**, *23*, 10151. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.