Article

Complete Chloroplast Genome Sequences of Four Species in the *Caladium* Genus: Comparative and Phylogenetic Analyses

Yuanjun Ye [1,2], Jinmei Liu [1,2], Yiwei Zhou [1,2], Genfa Zhu [1,2], Jianjun Tan [1,2] and Yechun Xu [1,2,*]

1   Guangdong Provincial Key Lab of Ornamental Plant Germplasm Innovation and Utilization, Environmental Horticulture Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, China
2   Key Laboratory of Urban Agriculture in South China, Ministry of Agriculture, Guangzhou 510640, China
*   Correspondence: tahua@126.com

Abstract: Caladiums are promising colorful foliage plants due to their dazzling colors of the leaves, veins, stripes, and patches, which are often cultivated in pots or gardens as decorations. Four wild species, including *C. bicolor*, C. humboldtii, *C. praetermissum,* and *C. lindenii,* were employed in this study, where their chloroplast (cp) genomes were sequenced, assembled, and annotated via high-throughput sequencing. The whole cp genome size ranged from 162,776 bp to 168,888 bp, and the GC contents ranged from 35.09% to 35.91%. Compared with the single large copy (LSC) and single small copy (SSC) regions, more conserved sequences were identified in the inverted repeat regions (IR). We further analyzed the different region borders of nine species of Araceae and found the expansion or contraction of IR/SSC regions might account for the cp genome size variation. Totally, 131 genes were annotated in the cp genomes, including 86 protein-coding genes (PCGs), 37 tRNAs, and eight rRNAs. The effective number of codons (ENC) values and neutrality plot analyses provided the foundation that the natural selection pressure could greatly affect the codon preference. The GC_3 content was significantly lower than that of GC_1 and GC_2, and codons ending with A/U had higher usage preferences. Finally, we conducted phylogenetic relationship analysis based on the chloroplast genomes of twelve species of Araceae, in which *C. bicolor* and *C. humboldtii* were grouped together, and *C. lindenii* was furthest from the other three *Caladium* species occupying a separate branch. These results will provide a basis for the identification, development, and utilization of *Caladium* germplasm.

Keywords: *Caladium*; ornamental foliage plant; chloroplast genome; natural selection pressure; phylogenetic analysis

1. Introduction

The genus *Caladium* Vent. (family Araceae) includes perennial herbs native to the tropical regions of South and Central America [1–3]. Most *Caladium* species are distributed in the Amazon rainforest either in open areas or beside streams [4,5]. *Caladium* spp. is regarded as the most promising colorful foliage plant with great variation in leaf color [6,7]. Because of the dazzling colors of the leaves, veins, stripes, and patches, as well as the long ornamental period, they are often cultivated in pots or gardens as decorations [8]. The ornamental effect for urban areas is excellent, and hence this genus is collectively known as the "Queen of Foliage Plants" [9,10]. In many countries of Europe and America, *Caladium* is grown as a replacement for traditional grasses and flowers for the purposes of arranging flower beds and flower borders or creating a unique landscape effect of "no flowers is better than flowers" [11–13]. In recent years, *Caladium* has become a new favorite foliage plant due to its colorful leaves, short production cycle, and high selling profit which conferred its high popularity in domestic and foreign markets [14,15].

Over nearly 150 years of selection and breeding, more than 2000 varieties have been cultivated [16–18]. At present, there are more than 90 varieties on the market, among which

more than 50 varieties are grown for large-scale flower production [19]. Since the beginning of this century, China has successively introduced some varieties of colorful *Caladium* as ornamental plants, which are greatly appreciated by retailers and young people [20].

The number of native species of the genus *Caladium* has always been controversial, ranging from 7 to 17. It has been reported earlier that *C. marmoratum*, *C. picturatum,* and *C. steudnirifolium* should be more appropriately merged into *C. bicolor* [21]. However, Croat recommended keeping the three species separate [22]. In 2013, *C. clavatum* and *C. praetermissum* were added to the list of species of wild *Caladium*, jointly released by Kew Gardens in the UK and the Missouri Botanical Garden in the US [23]. This increased the number of native species of *Caladium* to 14. Cao and Deng investigated the genome size and chromosome numbers of ten native species and 53 cultivars of *Caladium* [24]. They found that there are two genome types, thus providing some evidence for the evolutionary origin of *Caladium* germplasm.

Genetic and kinship analyses of *Caladium* germplasm have also received some attention. Loh et al. used 17 pairs of amplified fragment length polymorphism (AFLP) markers to accurately distinguish two wild species (*C. bicolor* and *C. schomburgkii*) and six cultivars of *Caladium* [25]. Deng et al. used simple sequence repeat (SSR) markers to analyze the genetic relationships among 45 major cultivars and seven wild species of *Caladium* [23]. The authors reported that the genetic variation rate among cultivars was low (44.4%). However, the two wild species *C. bicolor* and *C. schomburgkii,* exhibited high genetic consistencies with the cultivars, whereas the genetic consistencies between other wild species and cultivars were low. Moreover, *C. steudnirifolium* and *C. lindenii* were less genetically related to other wild species.

The chloroplast is the site of photosynthetic activity and is also one of the most important organelles of green plants, playing an important role in the inheritance and expression of plant genetic material [26–29]. The chloroplast has its own genome with an independent genetic system consisting of multiple copies of circular DNA molecules (110–210 kb), although few plant species have multiple linear copies of chromosomes [30–32]. The chloroplast genome has a typical tetrad structure featuring a large single-copy region, a small single-copy region, and a pair of inverted repeats (IR) regions [33,34]. The differences among plant species are attributed to the expansion and contraction of the IR regions [35]. Compared with the nuclear genome, the chloroplast genome shows maternal uni-parental inheritance, highly conserved structure, and less gene recombination [36–38]. Therefore, the chloroplast genome sequence is widely used in plant systematics, phylogeny, and population genomics [39,40].

Next-generation sequencing (NGS) technology has been widely used to identify functional genes and DNA markers massively due to its high throughput, high accuracy, and low cost. With the recent improvement of sequencing technology, the ongoing upgrading of sequencing platforms, and the development of a series of dynamic assembly and annotation software such as plasmaSPAdes, NOVOPlasty, and GetOrganelle, the chloroplast genomes of more than 3000 plant species have been sequenced and analyzed [41–44]. Remarkably, the chloroplast genome of *Caladium* has not yet been sequenced, and few phylogenetic studies based on chloroplast genomics have been reported.

In this study, four wild species of *Caladium* were used, and their chloroplast genomes were sequenced, assembled, and annotated via high-throughput sequencing. The characteristics of IR boundary and spacer loci in the chloroplast genomes were analyzed, and their genetic diversity and phylogenetic relationship were investigated. These results will provide a basis for the identification, development, and utilization of *Caladium* germplasm.

2. Results

2.1. The Structure of the Chloroplast Genomes of the Four Caladium Species

The chloroplast genome sizes of *C. bicolor* and *C. humboldtii* showed high similarity (162,933 bp and 162,776 bp, respectively) with a 157 bp difference, whereas those of *C. praetermissum* and *C. lindenii* were slightly larger (165,286 bp and 168,888 bp, respectively)

(Figure 1). The GC contents of the four cp genomes were 35.87%, 35.91%, 35.64%, and 35.09%, respectively (Table 1). The cpDNA of the four samples showed a typical tetrad ring structure, consisting of a pair of inverted repeat regions (IRa and IRb; 26,277, 26,277, 26,484 and 26,472 bp in length, respectively), a large single copy region (LSC; 89,209, 88,986, 91,168 and 93,162 bp in length, respectively) and a small single copy region (SSC; 21,170, 21,236, 21,150 and 22,782 bp in length, respectively). Table 1 also shows that the LSC region of *C. humboldtii* was also significantly shorter than that of the other species, whereas that of *C. lindenii* was the longest. In terms of SSC length, *Zamioculcas zamiifolia* showed the smallest, but *C. lindenii* had the greatest SSC. For IR size, *Z. amazonica* exhibited the shortest while *Z. zamiifolia* had the longest IR. Moreover, we compared the differences in chloroplast genome sequences between the ON7070731 and NC_060474 (Table S1). Results showed there were only some variations in the noncoding region and one SNP in the coding region, which did not affect the genetic structure.
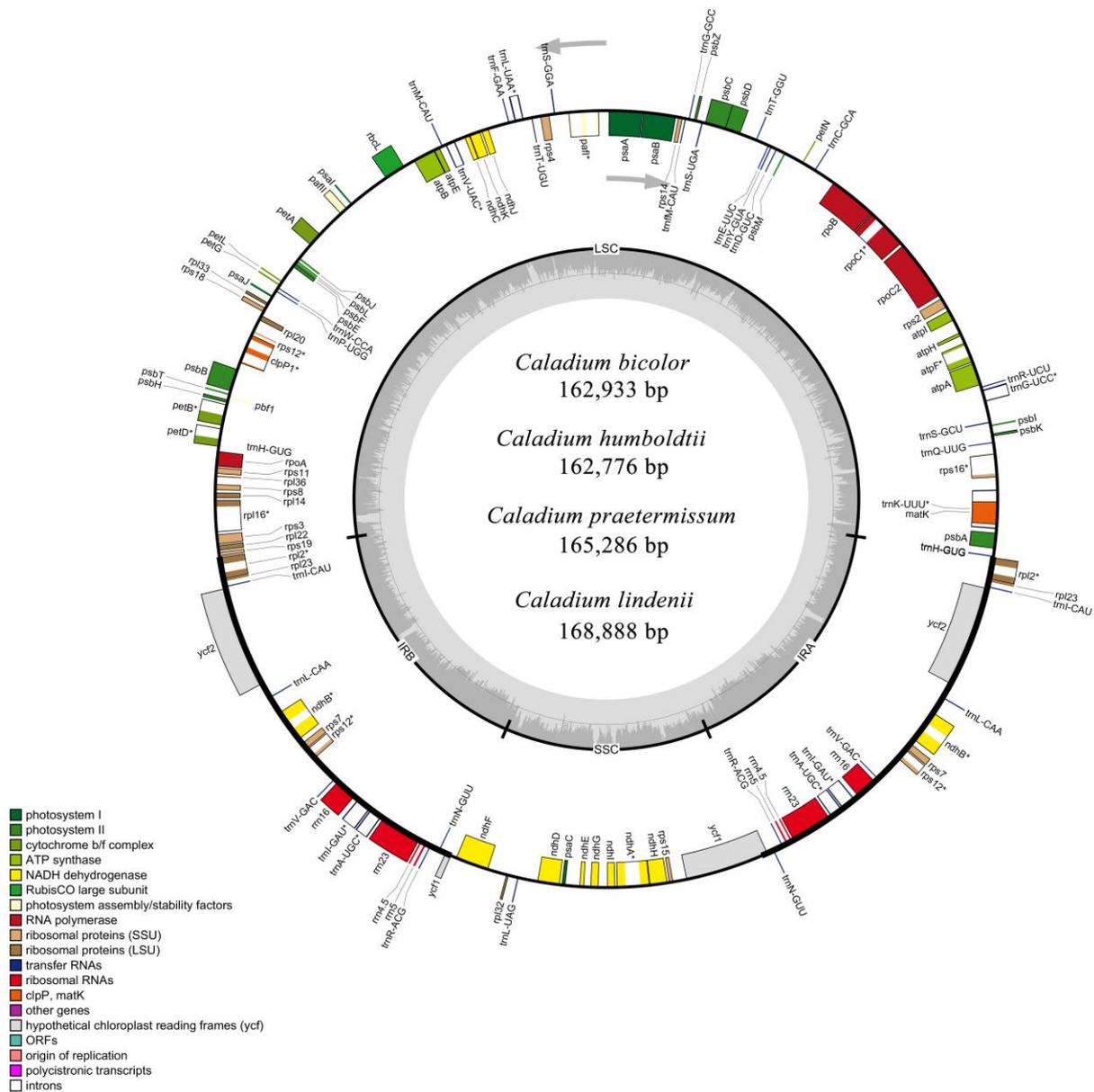
Figure 1. Circular gene map of the complete chloroplast genomes of four *Caladium* species.

A total of 131 genes were identified in the chloroplast genome of *Caladium*, including 86 protein-coding genes (PCGs), 37 tRNAs, and eight rRNAs (Table S2). Most genes appeared in the LSC or SSC region in single copy form. Among them, 12 genes were assigned to the SSC region, including 11 PCGs (*ndhF, rpl32, ccsA, ndhD, psaC, ndhE, ndhG, ndhI, ndhA, ndhH,* and *rps15*) and one tRNA (*trnL-UAG*). There were 83 genes in the LSC region, including 61 PCGs and 22 tRNAs. Only 16 genes were detected in the IR region, including five PCGs (*rpl2, rpl23, ycf2, ndhB,* and *rps7*), seven tRNAs (*trnI-CAU, trnL-CAA, trnV-GAC, trnI-GAU, trnA-UGC, trnI-ACG,* and *trnA-GUU*) and four rRNAs (*rrn16, rrn23, rrn4.5* and *rrn5*). The *ycf1* gene spanned the SSC region and the IR region. The *rps12* had two copies, each having three exons, and the two copies shared the first exon, which was in the LSC region, while the other two exons were in the IR region.

Table 1. Comparison of the general information of nine species of Araceae.

Species Names	Size (bp)/GC Content (%)				Number of Gene			Total Genes
	Genome	LSC	SSC	IR	PCGs	RNA/rRNA	RNA/tRNA	
C. bicolor	162,933/35.87	89,209/34.17	21,170/29.07	26,277/41.5	86	37	8	131
C. humboldtii	162,776/35.91	88,986/34.26	21,236/28.99	26,277/41.5	86	37	8	131
C. praetermissum	165,286/35.64	91,168/33.76	21,150/29.27	26,484/41.42	86	37	8	131
C. lindenii	168,888/35.09	93,162/33.26	22,782/27.87	26,472/41.43	86	37	8	131
Syngonium angustatum	164,929/35.72	90,714/33.94	21,559/28.98	26,328/41.53	86	37	8	131
Zomicarpella amazonica	162,729/35.82	90,811/34.33	21,656/28.76	25,131/41.55	85	37	8	130
Zamioculcas zamiifolia	167,405/35.7	91,357/34.02	19,326/29.47	28,361/40.52	86	37	8	131
Xanthosoma helleborifolium	164,418/35.84	90,833/33.88	20,705/41.53	26,440/41.53	86	37	8	131
Xanthosoma sagittifolium	165,169/35.67	91,122/33.79	21,079/29.36	26,484/41.42	86	37	8	131

2.2. Analysis of Contraction and Expansion of the IR Region

By comparing the gene distribution of IR/LSC and IR/SSC border regions in the chloroplast genomes of four *Caladium* species and those of related species, the expansion or contraction of IR/SSC boundary regions was assessed. As shown in Figure 2, the nine sequences had similar gene structures and sequences and the genes distributed near the boundaries of IR/LSC and IR/SSC were *rps19, rpl22, rpl2, ycf1, trnH,* and *psbA*. Among them, the IR boundaries of *C. bicolor* and *C. humboldtii* were identical, whereas *C. praetermissum* differed only in the SSC/IRa boundary. In *C. bicolor* and *C. humboldtii*, the sizes of the *ycf1* gene were 5241 and 422 bp in the SSC and IRa regions, respectively, while these sizes in *C. praetermissum* were 5292 bp and 422 bp, respectively. As for *C. lindenii*, it was quite different from the other three *Caladium* species, and its boundaries were different, indicating that this species may have undergone a unique evolutionary process.

2.3. GView Analysis

To gain a deeper understanding of the phylogenetic relationships among the different species of *Caladium* and the differences from other closely related species, we used the GView tool to create a circle map of the chloroplast genomes with the assembled *C. bicolor* genome as a reference. The characteristics and structural variation of all chloroplast genomes were evaluated. Figure 3 shows that the nine genomes had similar structures. Compared with the IR region, the LSC and SSC regions varied greatly among different species. Even within the same genus, the chloroplast genomes of the studied four species of *Caladium* showed inconsistencies. Particularly, there existed tiny variations among the intergenic regions. In terms of the genome structures, *C. bicolor* and *C. humboldtii* were identical. However, the cp genome of *C. praetermissum* was similar to that of *Xanthosoma sagittifolium*. Moreover, the cp genome of *C. lindenii* was similar to that of *Syngonium angustatum*.
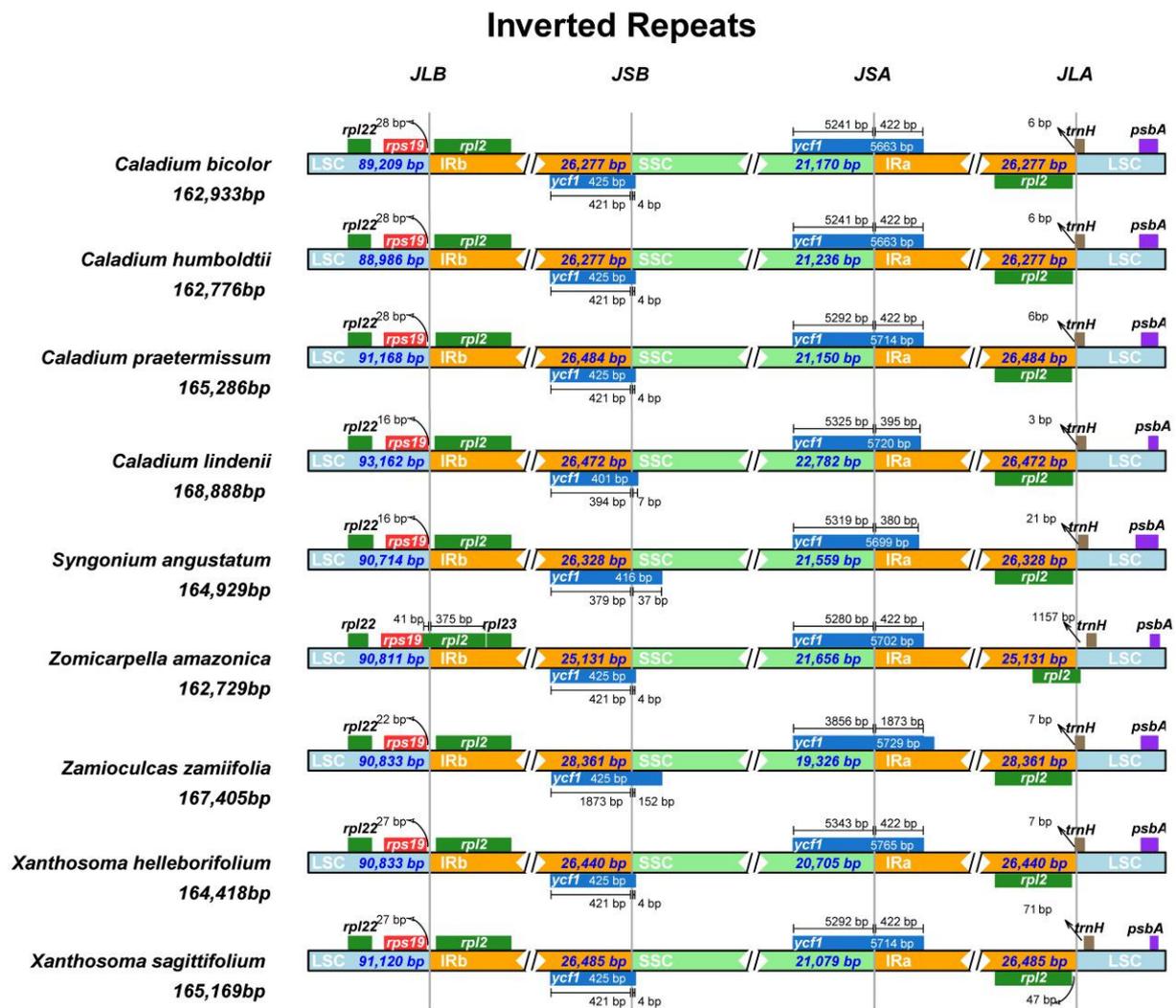
Figure 2. Comparison of the borders of the LSC, SSC, and IR regions of nine species of Araceae. The numbers above the gene features denote the distance between the gene borders, either the start or end of genes, and the junction sites.

2.4. Analysis of Chloroplast Microsatellites and Repeat Sequences

In this study, we analyzed the distribution of SSRs in the cp genomes of four *Caladium* species. As shown in Figure 4, there were only three types of SSRs, including single-base, two-base, and three-base repeats. Three-base repeats were only present in *C. lindenii* but not in the other three species. Most SSRs were single-base repeats, accounting for 63.0–81.7% of the total SSRs, which was more than all other repeat types combined. The total numbers of SSRs in *C. bicolor* and *C. humboldtii* were similar (91 and 89, respectively). *C. praetermissum* had the least number of SSRs, whereas *C. lindenii* had the highest. The SSRs of the four *Caladium* genomes were mostly distributed in the LSC region and the non-coding region, compared to other regions. In terms of repeating units, A/T repeats were significantly more than C/G repeats. All two-base and three-base repeats were AT/AT repeats and AAT/ATT, respectively, but there were no other types of repeats.
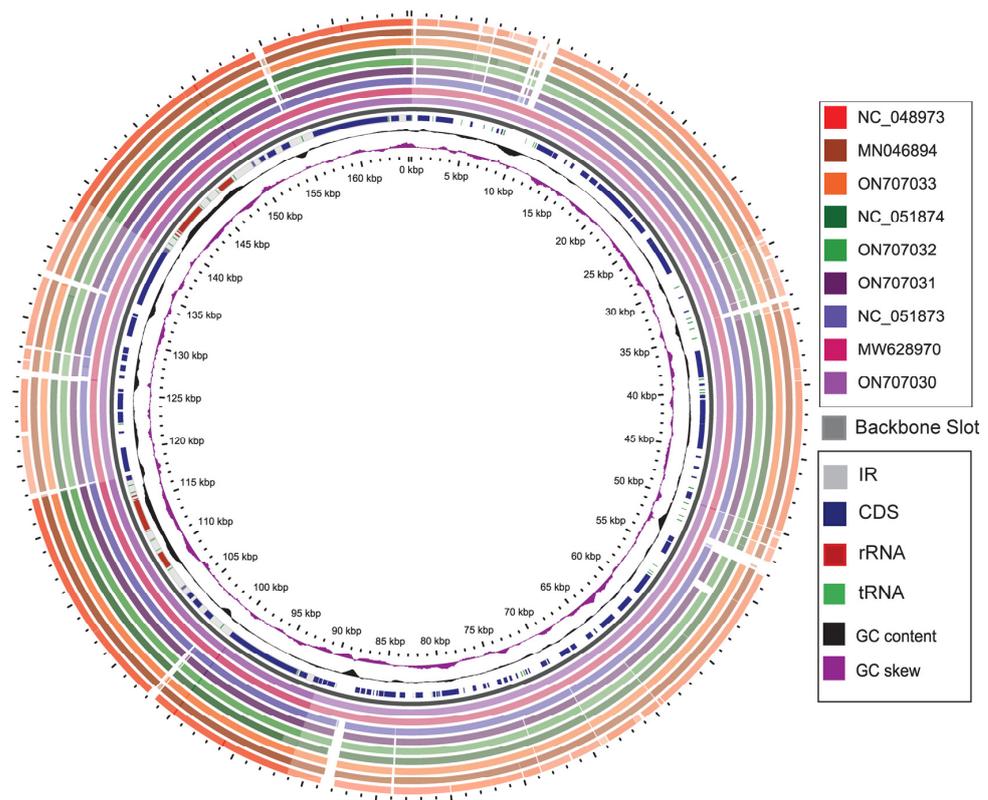
Figure 3. The graphical map shows nine circular plastome assemblies using *C. bicolor* plastome alignment as a reference. The innermost ring shows the genome size in kbp, followed by GC skew (purple) and GC content (black). From the inner to outer: *C. praetermissum* (ON707030), *X. sagittifolium* (MW628970), *X. helleborifolium* (NC_051873), *C. bicolor* (ON707031), *C. humboldtii* (ON707032), *Z. amazonica* (NC_051874), *C. lindenii* (ON707033), *S. angustatum* (MN046894), *Z. zamiifolia* (NC_048973). The similar and divergent locations are represented by continuous and interrupted track lines, respectively.
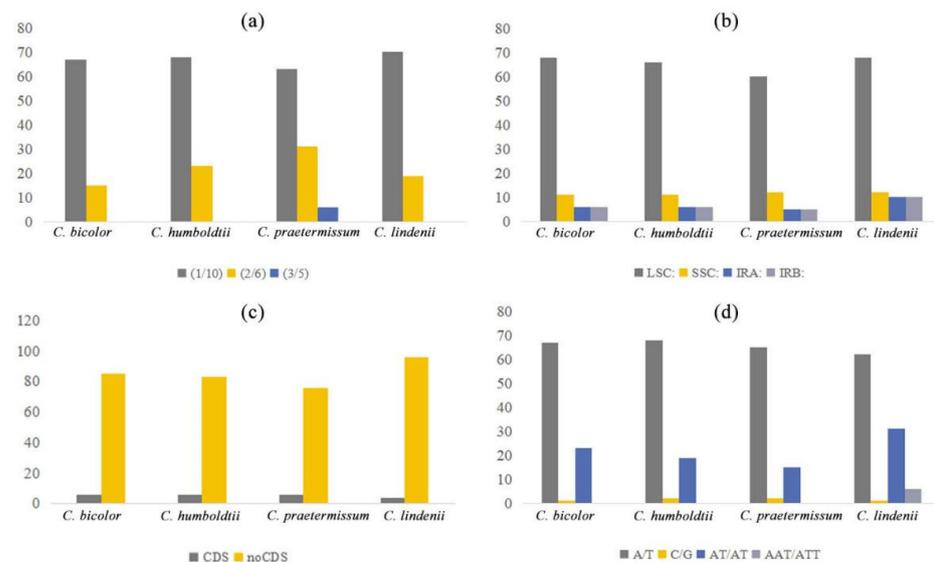
Figure 4. Comparison of repeat sequences in chloroplast genome distribution of four *Caladium* species. (**a**) The frequencies of different SSR repeat classes; (**b**) The SSR distribution in different chloroplast genome regions; (**c**) The SSR distribution in CDS and noCDS; (**d**) The SSR distributions of mono-, di-, and trinucleotide motifs.

2.5. Analysis of Selection Pressure and Codon Bias

The relative synonymous codon usage (RSCU) tool was used to evaluate the use of synonymous codons in coding regions, where a larger RSCU indicated a stronger bias. Our data showed that the content of leucine was highest in the chloroplast genomes, followed by serine and arginine, whereas the number of codons of tryptophan was the least. As shown in Figure 5, all amino acids except tryptophan used two or more synonymous codons. For example, isoleucine was encoded by three synonymous codons (alanine, glycine, and proline). Threonine and valine were encoded by four synonymous codons, and leucine, serine, and arginine were encoded by six synonymous codons. There were 32 RSCU values greater than one, of which 13 of them ended with A and 16 end with U. These findings were consistent with previous studies, which showed that codons ending with A/U in plants had higher usage preference. As shown in Table S3, the codon preferences of different *Caladium* species showed high conservation, where two codons exhibited the consistent RSCU value (AUG, 1.997; GUG, 0.003) and represented the extreme value in the four *Caladium* species. However, there were still some discrepancies among different materials, which mainly focused on the number of codons. For most codons, the codon preferences of *C. bicolor* and *C. humboldtii* were almost identical, significantly differing from those of the other two species, especially in *C. lindenii*. t
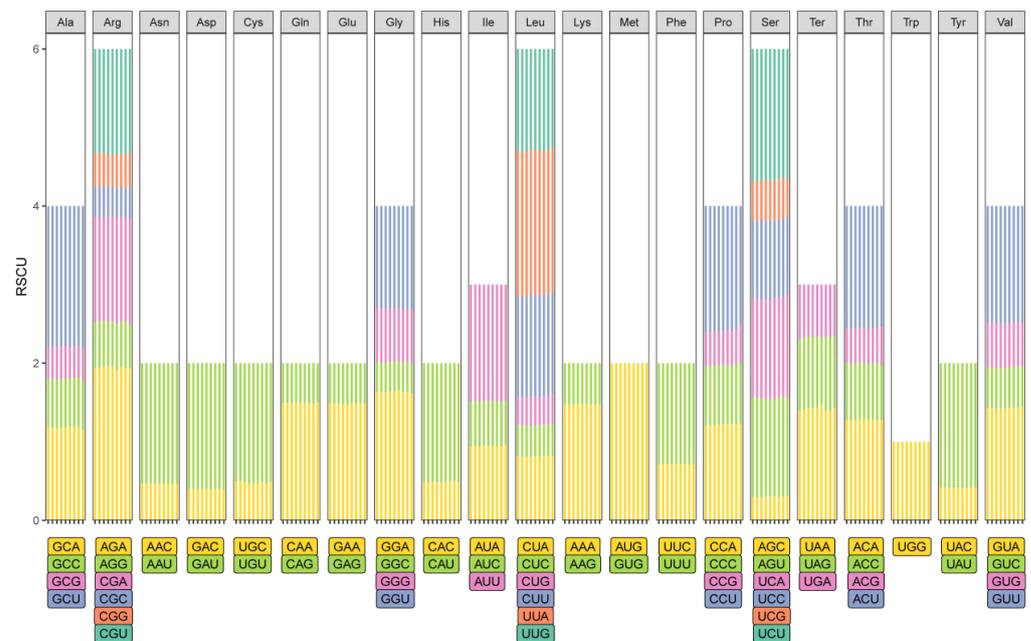
Figure 5. Relative synonymous codon usage (RSCU) analysis in chloroplast genes of *Caladium* species.

In order to analyze the trend of codon usage bias in the cp genomes of nine species, the values of the effective number of codons (ENC) were investigated. As in Table S4, the ENC values varied from 17.158 to 61.000, showing different extents of codon preferences among the species and indicating that the codon preference was weak. To further explore the details, the distribution of the ENC values of the coding genes in the genomes was exhibited in Figure S1.

GC_{1-3} means the GC content of three different positions of each codon. The overall GC content of the cp genomes varied among the nine species and ranged from 29.06% to 46.21% (Table S5). As expected, the GC_1, GC_2, and GC_3 contents varied significantly across species and also among genes in the genomes. The average value of GC_3 in the cp genomes was 28.63%, and the GC_1 and GC_2 were 46.07%, and 40%, respectively. The GC_3 content was significantly lower than that of GC_1 and GC_2. We also found that the greatest difference in GC content existed in GC_3, which was widely applied to better illustrate the codon usage variation. The neutrality plot was shown in Figure S2, which revealed little

correlation between GC_3 and GC_{12}. These results provided the foundation that the natural selection pressure could greatly affect codon preference.

2.6. Phylogenetic Relationship Analysis

We selected the chloroplast genomes of twelve species of Araceae to explore the genetic relationship between *Caladium* and its relatives. The phylogenetic tree was constructed based on the complete chloroplast genome sequences using maximum likelihood (ML) and Bayesian inference (BI) methods. The results showed that the topological structures of the ML and BI analyses were identical and that most clades had high posterior probabilities and bootstrap values. As shown in Figure 6, four *Caladium* species were divided into different branches, in which two *C. bicolor* (ON7070731 and NC_060474) and *C. humboldtii* were grouped together, with *Z. amazonica* being relatively close. Furthermore, *C. praetermissum* and *X. sagittifolium* were categorized into a branch, while *C. lindenii* was furthest from the other three *Caladium* species occupying a separate branch and being relatively closely related to *Syngonium angustatum*. These findings are consistent with previous reports, in which *C. lindenii* was later classified as *Caladium* genus, and its appearance and resistance were more similar to those of *S. angustatum*.
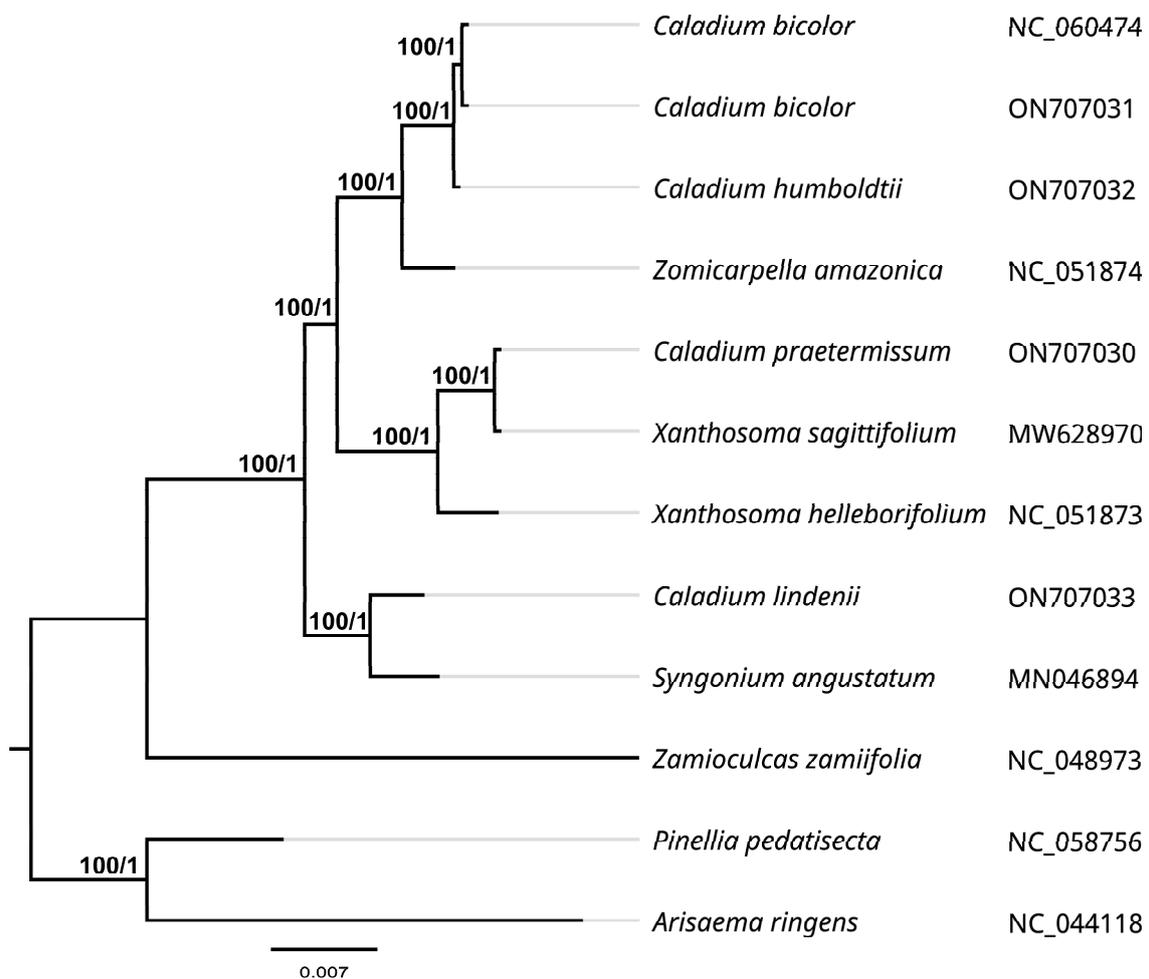
Figure 6. Maximum likelihood (ML) phylogenetic tree construction, including twelve species based on concatenated sequences from all chloroplast genomes. The obtained bootstrap values (BS) and Bayesian inference posterior probabilities (PP) are marked at the tree node (BS/PP).

3. Discussion

The cp genomes of the four *Caladium* species all had the typical circular tetrad structure of angiosperm cp genomes and were quite different in length. In terms of genome sequence

length and number of annotated genes, except for *Z. amazonica*, which has 130 genes, the other *Caladium* species, as well as the related species, had 131 genes, indicating that the cp genome of *Caladium* had a certain degree of conservation. The GC content was reported as an important indicator for judging the genetic relationship among species [45,46]. In our study, the gene types, numbers, and order of genes encoded by the genomes of the four *Caladium* species were identical, with highly similar G + C content. Based on the G + C content in the region sequence, the rank from high to low was IR, LSC, and SSC, which is a ubiquitous phenomenon in many plant species [47–49]. The IR boundary was different among different species, and the fluctuation of the IR boundary was the main reason for the difference [50]. The cp genomes of *C. bicolor* and *C. humboldtii* showed the smallest difference in the IR regions among those of the four *Caladium* species. There was no significant difference in the contraction and expansion of the IR regions between the two. Therefore, the GC content and the IR region boundary conditions, to a large extent, indicate that *C. bicolor* and *C. humboldtii* are very closely related but are relatively distinct from *C. humboldtii* and *C. lindenii*.

Codon bias affects translation initiation, elongation, and accuracy, as well as mRNA splicing and protein folding [51–53]. Therefore, codon preference can also reflect kinship to a certain extent. Among the chloroplast genes of the four *Caladium* species, the number of codons of leucine was the largest, the number of tryptophan occurrences was the least, and codons ending in A/T were preferred. This feature is consistent with that of most plant species [54,55]. The codon preferences of *C. bicolor* and *C. humboldtii* showed almost no difference, suggesting that the two species are closely related, while the codon preferences of *C. humboldtii* and *C. lindenii* were markedly different, indicating that they may have unique evolutionary positions.

Microsatellite sequences (microsatellite DNA), also known as SSR, are 1–6 bp repeats that are widely distributed in the cp genome. SSR is highly polymorphic and specific and is a valuable marker for studying gene flow, population genetics, and genetic mapping [56]. Except for *C. lindenii*, the repeat types and distribution numbers of the other three *Caladium* species were roughly the same. Identification of these SSR loci can provide candidate molecular markers for research on the genetic diversity and conservation genetics of *Caladium*. The nodes in the LSC, SSC, and IR regions of the four *Caladium* species were highly conserved, indicating that the cp genome structure of *Caladium* is highly conserved. Phylogenetic analysis classified *C. lindenii* as a single clade that was far from the other three *Caladium* species, which is consistent with previous reports [23–25]. As for the limited research on the phylogenetic analysis, AFLP and SSR markers were often used to identify their kinship among *Caladium* cultivars. In this paper, we provided more accurate findings on the phylogenetic relationship analysis among different species of Caladieae. Based on the above results, we conclude that *C. lindenii* has greater specificities in cp genome structure, IR region contraction and expansion, SSR distribution, and codon preference, being relatively close to *S. angustatum*. Therefore, this species is suggested to be classified into the genus *Synaptocarpus*.

4. Materials and Methods

4.1. Plant Materials

Plants of four *Caladium* species were collected from the Environmental Horticulture Research Institute of Guangdong Academy of Agricultural Sciences (23° 23′ N, 113° 26′ E), namely, *C. bicolor*, *C. humboldtii* 'Mini White', *C. praetermissum* 'Hilo Beauty,' and *C. lindenii*. The phenotypic characteristics are shown in Figure 7. Young leaves were collected and rinsed thoroughly with tap water. Subsequently, they were washed several times with sterile water and dried quickly in a sampling bag containing silica gel. The samples were then stored at −80 °C until used.
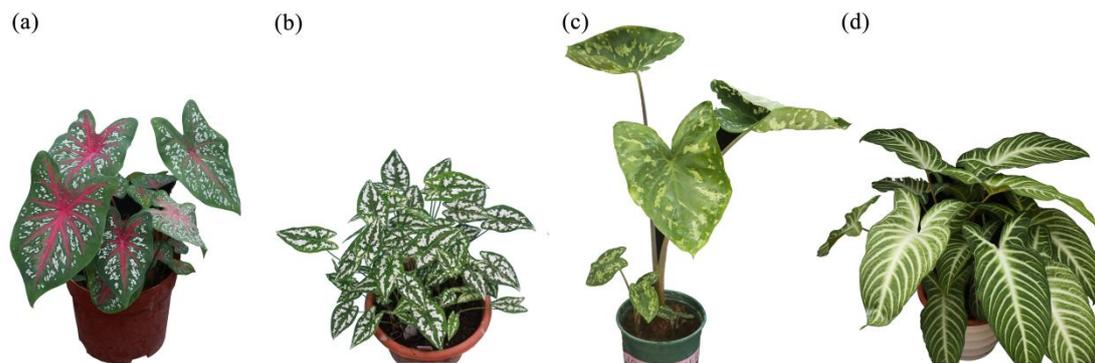
Figure 7. The four *Caladium* species used in the present study. (**a**) *C. bicolor*; (**b**) *C. humboldtii*; (**c**) *C. praetermissum*; (**d**) *C. lindenii*.

4.2. DNA Extraction and High-Throughput Sequencing

Total DNA was extracted from frozen leaf samples by the modified cetyltrimethylammonium bromide (CTAB) method, and the quality of DNA was assessed by 1.5% agarose gel electrophoresis. The DNA was fragmented by mechanical interruption (ultrasound) and then purified, and end repaired. PolyA tails were added to the 3′ ends, and the fragments were ligated with sequencing adapters. The required fragment size was selected by agarose gel electrophoresis. PCR amplification was performed to form a sequencing library, and the qualified library was sequenced using the BGISEQ-500 platform with PE150 read lengths according to the manufacturer's instructions. DNA extraction and sequencing were all performed by Guangzhou Bio&Data Biotechnologies Co., Ltd. (Guangzhou, China).

4.3. Chloroplast Genome Assembly and Annotation

At least 5 G of raw data were obtained for each species. After data filtering, adapter sequences and low-quality reads were removed to obtain high-quality clean data. First, NOVOPlasty software (k-mer = 39) was used for assembly and splicing, where the size of the insert was set to 250 bp [57]. Subsequently, the online program GeSeq was employed to annotate the chloroplast genome sequence, and Geneious v9.0.2 was used for visualizing the annotated sequence with manual corrections [58,59]. The annotated sequencing data for the four species were uploaded to the NCBI database with serial numbers ON707030, ON707031, ON707032, and ON707033, respectively. Finally, with the help of the online program Organellar Genome DRAW, the genome maps of the *Caladium* chloroplast genomes were constructed [60].

4.4. Comparative Analysis of Chloroplast Genomes

Relative synonymous codon usage (RSCU) analysis for every codon in each genome was conducted to determine codon bias [61]. The expansion and contraction of the IR borders of the four chloroplast genomes of *Caladium* were mapped with the aid of an IRscope [62]. Chloroplast genome similarity was assessed using BLAST Atlas on the GView server (http://server.gview.ca/, accessed on 26 October 2022) with 50 kbp connection windows with C. bicolor genome as a reference [63]. The Perl program provided by MIcroSAtellite Identification Tool (MISA) was used to analyze simple repeat sequence (SSR) sites. For mononucleotides, dinucleotides, trinucleotides, tetranucleotides, pentanucleotides, and hexanucleotides, the repetition thresholds were set to 10, 5, 4, 3, 3, and 3, respectively.

4.5. Phylogenetic Analysis

Published and fully annotated chloroplast genome-wide data of another eight species of Araceae were downloaded from the NCBI database, including *C. Bicolor* (NC_060474), *Syngonium angustatum* (MN046894), *Zomicarpella amazonica* (NC_051874), *Zamioculcas zamiifolia* (NC_048973), *Xanthosoma helleborifolium* (NC_051873), *Xanthosoma sagittifolium* (MW628970),

Pinellia pedatisecta (NC_058756) and *Arisaema ringens* (NC_044118). A phylogenetic tree was constructed based on the complete chloroplast genome sequences of 12 species with maximum likelihood (ML) and Bayesian inference (BI) methods [64]. The ML tree was conducted using IQ-TREE v.2.1.4 [65]. The best-fitting nucleotide substitution model TVM+F+R3 was determined using the Bayesian Information Criterion (BIC) by ModelFinder in the IQ-TREE package, and 1000 bootstrap replicates [66]. The Bayesian inference was performed with MrBayes v.3.2.7, employing the GTR+G model of nucleotide substitution [67]. After the phylogenetic tree was exported, it was viewed using FigTree version 1.4.2.z.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/genes13122180/s1, Figure S1: Analysis of ENC-plot in the cp genomes of nine species of Araceae; Figure S2: Neutrality analysis performed by plotting GC_{12} values against GC_3 values for the cp genomes of nine species of Araceae. The diagonal line on the neutrality plot shows that the value of GC_{12} is equal to GC_3; Table S1: The comparison between the ON707031 and NC_060474; Table S2: Functional classification of Caladium chlroplast genome; Table S3: Determination of optimal codons in the chloroplast genomes of Caladium; Table S4: The ENC values of the coding genes in the cp genomes of nie species of Araceae; Table S5: The GC contents of three positions of codons (GC1, GC2 and GC3) in the cp genomes of nine species of Araceae.

Author Contributions: Conceptualization, Y.Y. and Y.X.; methodology, Y.Z. and J.L.; software, Y.Z. and Y.Y.; validation, G.Z., J.T., Y.Z. and Y.Y.; formal analysis, Y.Y.; investigation, J.T. and J.L.; resources, Y.X.; data curation, Y.Z.; writing-original draft preparation, Y.Y.; writing-review and editing, Y.Y.; visualization, Y.Y.; supervision, G.Z. and Y.X.; project administration, Y.X.; funding acquisition, G.Z. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The four-chloroplast genome sequence data generated in this study are available in GenBank of the National Center for Biotechnology Information (NCBI) under the access numbers: ON707030-ON707033.

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

References

1. Zhang, Y.S.; Chen, J.J.; Cao, Y.M.; Duan, J.X.; Cai, X.D. Induction of tetraploids in 'Red Flash' *Caladium* using colchicine and oryzalin: Morphological, cytological, photosynthetic and chilling tolerance analysis. *Sci. Hortic.* **2020**, *275*, 109524. [CrossRef]
2. Cao, Z.; Deng, Z.A. Morphological, cytological and molecular marker analyses of 'Tapestry' *Caladium* variants reveal diverse genetic changes and enable association of leaf coloration pattern loci with molecular markers. *Plant Cell Tiss. Org.* **2020**, *143*, 363–375. [CrossRef]
3. Chen, J.J.; Zhang, Y.S.; Duan, J.X.; Cao, Y.M.; Cai, X.D. Morphological, cytological, and pigment analysis of leaf color variants regenerated from long-term subcultured *Caladium* callus. *In Vitro Cell Dev. Biol.-Plant* **2021**, *57*, 60–71. [CrossRef]
4. Croat, T.B.; Delannay, X.; Ortiz, O.O.; Jiménez, P.D. A review of the Aroid Tribe Caladieae with the description of three new species of *Caladium* and seven new species of *Syngonium* (Araceae). *Novon J. Bot. Nomencl.* **2019**, *27*, 38–64. [CrossRef]
5. Akhigbemen, A.M.; Ozolua, R.I.; Bafor, E.E.; Okwuofu, E.O. Evaluation of some neuropharmacological effects of *Caladium bicolor* aiton (Araceae) leaf extracts in mice. *Metab. Brain Dis.* **2019**, *34*, 537–544. [CrossRef]
6. Isah, T. Changes in the biochemical parameters of albino, hyperhydric and normal green leaves of *Caladium bicolor* cv. "Bleeding hearts" in vitro long-term cultures. *J. Photoch. Photobiol. B* **2019**, *191*, 88–98. [CrossRef]
7. Deng, Z.A.; Harbaugh, B.K. Independent inheritance of leaf shape and main vein color in *Caladium*. *J. Am. Soc. Hortic. Sci.* **2006**, *131*, 53–58. [CrossRef]
8. Ahmed, E.U.; Hayashi, T.; Yazawa, S. Leaf color stability during plant development as an index of leaf color variation among micropropagated *Caladium*. *HortScience* **2004**, *39*, 328–332. [CrossRef]
9. Kokalis-Burelle, N.; Brito, J.A.; Hartman, R.D. Susceptibility of seven *Caladium* (*Caladium × hortulanum*) cultivars to *Meloidogyne arenaria*, *M. enterolobii*, *M. floridensis*, *M. incognita*, and *M. javanica*. *J. Nematol.* **2017**, *49*, 457–461.

10. Cao, Z.; Deng, Z.A. De novo assembly, annotation, and characterization of root transcriptomes of three *Caladium* cultivars with a focus on necrotrophic pathogen resistance/defense-related genes. *Int. J. Mol. Sci.* **2017**, *18*, 712. [CrossRef]

11. Hussain, R.; Younis, A.; Riaz, A.; Tariq, U.; Ali, S.; Ali, A.; Raza, S. Evaluating sustainable and environment friendly substrates for quality production of potted *Caladium*. *Int. J. Recycl. Org.* **2017**, *6*, 13–21. [CrossRef]

12. Cao, Z.; Sui, S.Z.; Cai, X.D.; Yang, Q.; Deng, Z.A. Somaclonal variation in 'Red Flash' *Caladium*: Morphological, cytological and molecular characterization. *Plant Cell Tiss. Org.* **2016**, *126*, 269–279. [CrossRef]

13. Cao, Z.; Sui, S.Z.; Yang, Q.; Deng, Z.A. Inheritance of rugose leaf in *Caladium* and genetic relationships with leaf shape, main vein color, and leaf spotting. *J. Am. Soc. Hortic. Sci.* **2016**, *141*, 527–534. [CrossRef]

14. Ekeke, C.; Agbagwa, I.O. Anatomical characteristics of Nigerian variants of *Caladium bicolor* (Aiton) Vent. (Araceae). *Afr. J. Plant Sci.* **2016**, *10*, 121–129.

15. Ahmed, E.U.; Hayashi, T.; Yazawa, S. Auxins increase the occurrence of leaf-colour variants in *Caladium* regenerated from leaf explants. *Sci. Hortic.* **2004**, *100*, 153–159. [CrossRef]

16. Deng, Z.A.; Peres, N.A. 'Sea Foam Pink' *Caladium*. *HortScience* **2019**, *54*, 1637–1640. [CrossRef]

17. Yu, J.L.; Boyd, N.S. Tolerance of *Caladium* cultivars florida cardinal and florida fantasy to sulfonylurea herbicides. *HortScience* **2018**, *53*, 850–858. [CrossRef]

18. Cai, X.D.; Cao, Z.; Xu, S.X.; Deng, Z.A. Induction, regeneration and characterization of tetraploids and variants in 'Tapestry' *Caladium*. *Plant Cell Tiss. Org.* **2015**, *120*, 689–700. [CrossRef]

19. Maia, A.C.D.; Schlindwein, C. *Caladium bicolor* (Araceae) and *Cyclocephala celata* (Coleoptera, Dynastinae): A well-established pollination system in the northern atlantic rainforest of pernambuco, Brazil. *Plant Biol.* **2006**, *8*, 529–534. [CrossRef]

20. Deng, Z.A.; Goktepe, F.; Harbaugh, B.K. Inheritance of leaf Spots and their genetic relationships with leaf shape and vein color in *Caladium*. *J. Am. Soc. Hortic. Sci.* **2008**, *133*, 78–83. [CrossRef]

21. Madison, M. Notes on *Caladium* (Araceae) and its allies. *Selbyana* **1981**, *5*, 342–377.

22. Croat, T. Taxonomic status of neotropical Araceae. *Aroideana* **1994**, *17*, 33–60.

23. Deng, Z.A.; Goktepe, F.; Harbaugh, B.K.; Hu, J.G. Assessment of genetic diversity and relationships among *Caladium* cultivars and species using molecular markers. *J. Am. Soc. Hortic. Sci.* **2007**, *132*, 219–229. [CrossRef]

24. Cao, Z.; Deng, Z.A. Interspecific genome size and chromosome number variation shed new light on species classification and evolution in *Caladium*. *J. Am. Soc. Hortic. Sci.* **2014**, *139*, 449–459. [CrossRef]

25. Loh, J.P.; Kiew, R.; Kee, A.; Gan, L.H.; Gan, Y.Y. Amplified fragment length polymorphism (AFLP) provides molecular markers for the identification of *Caladium bicolor* cultivars. *Ann. Bot.* **1999**, *84*, 155–161. [CrossRef]

26. Chumley, T.W.; Palmer, J.D.; Mower, J.P.; Fourcade, H.M.; Calie, P.J.; Boore, J.L.; Jansen, R.K. The complete chloroplast genome sequence of *Pelargonium* × *hortorum*: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol. Biol. Evol.* **2006**, *23*, 2175–2190. [CrossRef]

27. Firetti, F.; Zuntini, A.R.; Gaiarsa, J.W.; Oliveira, R.S.; Lohmann, L.G.; Van Sluys, M.A. Complete chloroplast genome sequences contribute toplant species delimitation: A case study of the *Anemopaegma* species complex. *Am. J. Bot.* **2017**, *104*, 1493–1509. [CrossRef]

28. Lu, R.S.; Li, P.; Qiu, Y.X. The complete chloroplast genomes of three *Cardiocrinum* (Liliaceae) species: Comparative genomic and phylogenetic analyses. *Front. Plant Sci.* **2017**, *7*, 2054. [CrossRef]

29. Lee, S.R.; Kim, K.; Lee, B.Y.; Lim, C.E. Complete chloroplast genomes of all six *Hosta* species occurring in Korea: Molecular structures, comparative, and phylogenetic analyses. *BMC Genom.* **2019**, *20*, 833. [CrossRef]

30. Daniell, H.; Lin, C.S.; Yu, M.; Chang, W.J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [CrossRef]

31. Wu, M.L.; Li, Q.; Hu, Z.G.; Li, X.W.; Chen, S.L. The complete *Amomum kravanh* chloroplast genome sequence and phylogenetic analysis of the commelinids. *Molecules* **2017**, *22*, 1875. [CrossRef] [PubMed]

32. Zhou, Y.X.; Nie, J.; Xiao, L.; Hu, Z.G.; Wang, B. Comparative chloroplast genome analysis of rhubarb botanical origins and development of specific identification markers. *Molecules* **2018**, *23*, 2811. [CrossRef] [PubMed]

33. Sun, J.H.; Shi, S.; Li, J.L.; Jing, Y.; Wang, L.; Yang, X.Y.; Guo, L.; Zhou, S.L.; Sun, F.J. Phylogeny of Maleae (Rosaceae) based on multiple chloroplast regions: Implications to genera Circumscription. *Biomed. Res. Int.* **2018**, *2018*, 7627191. [CrossRef] [PubMed]

34. Lin, M.M.; Qi, X.J.; Chen, J.Y.; Sun, L.M.; Zhong, Y.P.; Fang, J.B.; Hu, C.G. The complete chloroplast genome sequence of *Actinidia arguta* using the PacBio RS II platform. *PLoS ONE* **2018**, *13*, e0197393. [CrossRef]

35. Li, D.M.; Zhao, C.Y.; Liu, X.F. Complete chloroplast genome sequences of *Kaempferia galanga* and *Kaempferia elegans*: Molecular structures and comparative analysis. *Molecules* **2019**, *24*, 474. [CrossRef]

36. Fonseca, L.H.M.; Lohmann, L.G. Plastome rearrangements in the "*Adenocalymma-Neojobertia*" clade (Bignonieae, Bignoniaceae) and its phylogenetic implications. *Front. Plant Sci.* **2017**, *8*, 1875. [CrossRef]

37. Thode, V.A.; Lohmann, L.G. Comparative chloroplast genomics at low taxonomic levels: A case study using *Amphilophium* (Bignonieae, Bignoniaceae). *Front. Plant Sci.* **2019**, *10*, 796. [CrossRef]

38. Chiapella, J.O.; Barfuss, M.H.J.; Xue, Z.Q.; Greimler, J. The plastid genome of *Deschampsia cespitosa* (Poaceae). *Molecules* **2019**, *24*, 216. [CrossRef]

39. Cui, Y.X.; Chen, X.L.; Nie, L.P.; Sun, W.; Hu, H.Y.; Lin, Y.L.; Li, H.T.; Zheng, X.L.; Song, J.Y.; Yao, H. Comparison and phylogenetic analysis of chloroplast genomes of three medicinal and edible *Amomum* species. *Int. J. Mol. Sci.* **2019**, *20*, 4040. [CrossRef]

40. Cui, Y.X.; Nie, L.P.; Sun, W.; Xu, Z.C.; Wang, Y.; Yu, J.; Song, J.Y.; Yao, H. Comparative and phylogenetic analyses of ginger (*Zingiber offcinale*) in the family Zingiberaceae based on the complete chloroplast genome. *Plants* **2019**, *8*, 283. [CrossRef]

41. Barrett, C.F.; Baker, W.J.; Comer, J.R.; Conran, J.G.; Lahmeyer, S.C.; Leebens-Mack, J.H.; Li, J.; Lim, G.S.; MayfieldJones, D.R.; Perez, L.; et al. Plastid genomes reveal support for deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. *N. Phytol.* **2016**, *209*, 855–870. [CrossRef]

42. Ferrarini, M.; Moretto, M.; Ward, J.A.; Surbanovski, N.; Stevanovic, V.; Giongo, L.; Viola, R.; Cavalieri, D.; Velasco, R.; Cestaro, A.; et al. An evaluation of the PacBio RS platform for sequencing and de novo assembly of a chloroplast genome. *BMC Genom.* **2013**, *14*, 670. [CrossRef] [PubMed]

43. Wu, Z.H.; Gui, S.T.; Guan, Z.W.; Pan, L.; Wang, S.Z.; Ke, W.D.; Liang, D.Q.; Ding, Y. A precise chloroplast genome of *Nelumbo nucifera* (Nelumbonaceae) evaluated with Sanger, Illumina MiSeq, and PacBio RS II sequencing platforms: Insight into the plastid evolution of basal eudicots. *BMC Plant Biol.* **2014**, *14*, 289. [CrossRef] [PubMed]

44. Li, X.W.; Hu, Z.G.; Lin, X.H.; Li, Q.; Gao, H.H.; Luo, G.A.; Chen, S.L. High-throughput pyrosequencing of the complete chloroplast genome of *Magnolia officinalis* and its application in species identification. *Acta Pharm. Sin.* **2012**, *47*, 124–130.

45. Li, Y.F.; Sylvester, S.P.; Li, M.; Zhang, C.; Li, X.; Duan, Y.F.; Wang, X.R. The complete plastid genome of *Magnolia zenii* and genetic comparison to Magnoliaceae species. *Molecules* **2019**, *24*, 261. [CrossRef] [PubMed]

46. Brunkard, J.O.; Runkel, A.M.; Zambryski, P.C. Chloroplast extend stromules independently and in response to internal redox signals. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10044–10049. [CrossRef]

47. Gao, C.M.; Deng, Y.F.; Wang, J. The complete chloroplast genomes of *Echinacanthus* species (Acanthaceae): Phylogenetic relationships, adaptive evolution, and screening of molecular markers. *Front. Plant Sci.* **2019**, *9*, 1989. [CrossRef]

48. Pauwels, M.; Vekemans, X.; Godé, C.; Frérot, H.; Castric, V.; Saumitou-Laprade, P. Nuclear and chloroplast DNA phylogeography reveals vicariance among european populations of the model species for the study of metal tolerance, *Arabidopsis halleri* (Brassicaceae). *N. Phytol.* **2012**, *193*, 916–928. [CrossRef]

49. Wu, L.W.; Nie, L.P.; Wang, Q.; Xu, Z.C.; Wang, Y.; He, C.N.; Song, J.Y.; Yao, H. Comparative and phylogenetic analyses of the chloroplast genomes of species of Paeoniaceae. *Sci. Rep.* **2021**, *11*, 14643. [CrossRef]

50. Li, L.; Wu, Q.P.; Fang, L.; Wu, K.L.; Li, M.Z.; Zeng, S.J. Comparative chloroplast genomics and phylogenetic analysis of *Thuniopsis* and closely related genera within Coelogyninae (Orchidaceae). *Front. Genet.* **2022**, *13*, 850201. [CrossRef]

51. Hanson, G.; Coller, J. Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* **2017**, *19*, 2030. [CrossRef] [PubMed]

52. Tang, D.; Wei, F.; Cai, Z.; Wei, Y.; Khan, A.; Miao, J.; Wei, K. Analysis of codon usage bias and evolution in the chloroplast genome of *Mesona chinensis* Benth. *Dev. Genes Evol.* **2021**, *231*, 1–9. [CrossRef] [PubMed]

53. Wei, L.; He, J.; Jia, X.; Qi, Q.; Liang, Z.; Zheng, H.; Ping, Y.; Liu, S.; Sun, J. Analysis of codon usage bias of mitochondrial genome in *Bombyx moriand* its relation to evolution. *BMC Evol. Biol.* **2014**, *14*, 262. [CrossRef] [PubMed]

54. Zheng, G.; Wei, L.L.; Ma, L.; Wu, Z.Q.; Gu, C.H.; Chen, K. Comparative analyses of chloroplast genomes from 13 *Lagerstroemia* (Lythraceae) species: Identification of highly divergent regions and inference of phylogenetic relationships. *Plant Mol. Biol.* **2020**, *102*, 659–676. [CrossRef] [PubMed]

55. Li, N.; Sun, M.H.; Jiang, Z.S.; Shu, H.R.; Zhang, S.Z. Genome-wide analysis of the synonymous codon usage patterns in apple. *J. Integr. Agr.* **2016**, *15*, 983–991. [CrossRef]

56. Ye, Y.J.; Xu, Y.C.; Li, D.M.; Tan, J.J.; Liu, J.M. Characterization of EST-SSR markers in *Curcuma kwangsiensis* S. K. Lee & C. F. Liang based on RNA sequencing and its application for phylogenetic relationship analysis and core collection construction. *Genet. Resour. Evol.* **2021**, *68*, 1503–1516.

57. Dierckysens, N.; Mardulyn, P.; Smits, G. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **2017**, *45*, e18.

58. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq-versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [CrossRef]

59. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [CrossRef]

60. Greiner, S.; Lehwark, P.; Bock, R. Organellar Genome DRAW (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **2019**, *47*, W59–W64. [CrossRef]

61. Meade, J.C.; Shah, P.H.; Lushbaugh, W.B. Trichomonas vaginalis: Analysis of codon usage. *Exp. Parasitol.* **1997**, *87*, 73–74. [CrossRef] [PubMed]

62. Amiryousefi, A.; Hyvonen, J.; Poczai, P. Irscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2008**, *34*, 3030–3031. [CrossRef] [PubMed]

63. Petkau, A.; Stuart-Edwards, M.; Stothard, P.; Domselaar, G.V. Interactive microbial genome visualization with GView. *Bioinformatics* **2010**, *26*, 3125–3126. [CrossRef] [PubMed]

64. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]

65. Nguyen, L.T.; Schmidt, H.A.; Haeseler, A.V.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef]

66. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **2017**, *14*, 587–589. [CrossRef] [PubMed]
67. Ronquist, F.; Teslenko, M.; van der Mark, P.; Ayres, D.L.; Darling, A.; Höhna, S.; Larget, B.; Liu, L.; Suchard, M.A.; Huelsenbeck, J.P. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **2012**, *61*, 539–542. [CrossRef]