The complete chloroplast genome and phylogenetic analysis of Dahlia pinnata Cavanilles 1791(Asteraceae: Dahlia)

Shan-De Duan
Agricultural University of Hebei
Kua-Yue Zheng
Agricultural University of Hebei
Yang Liu
Agricultural University of Hebei
Xing-Ru Li
Agricultural University of Hebei
Di-Ying Xiang
Agricultural University of Hebei
Li-Hong Hao
Agricultural University of Hebei
Duan-Fen Chen
Agricultural University of Hebei
Shan-Ce Niu
niushance@163.com
Agricultural University of Hebei

Article

Keywords: Dahlia pinnata, chloroplast, phylogeny

Posted Date: July 5th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3075823/v1

Additional Declarations: No competing interests reported.

Abstract

Dahlia pinnata Cavanilles 1791 is an important ornamental plant worldwide. The chloroplast genome has obvious advantages in studies of systematic evolution at the plant classification and species level, making it an important resource for phylogenetic research.Here, we sequenced the full chloroplast genome from *D. pinnata* 'Chocolate' and found that it exhibited a typical tetrad structure. The full-length *D. pinnata* chloroplast genome was 152,107 bp, with a GC content of 38.45%. The genome included an 83,704 bp large single-copy (LSC) region, an 18,347 bp small single-copy (SSC) region, and a pair of 25,028 bp inverted repeats (*IRa* and *IRb*). A total of 134 genes were annotated, including 86 protein-coding genes, 38 transfer RNA genes, 8 ribosomal RNA genes, and 2 pseudogenes. Analysis of password preferences shows that passwords chosen by *D. pinnata* tend to end with A/U. A total of 161 SSR markers were detected in the simple sequence repeat (SSR) analysis. Phylogenetic analysis that the *Dahlia* species formed a monophyly. *Dahlia* was clustered with *Cosmos-Bidens*, which differed from studies using nuclear genomic DNA. We suggest that nuclear-cytoplasmic incongruences may be widespread in Asteraceae, and should be thoroughly evaluated in order to understand the true evolutionary history of this economically-important group of plants.

Introduction

Dahlia(*Dahlia pinnata* Cav.) is a perennial bulbous flower of the Asteraceae family, also known as Mexican marigold, *chrysanthemum*, sweet potato flower, and oriental *chrysanthemum*. It is native to Mexico and is the national flower of Mexico[1]. Due to its long blooming period, large flower diameter, abundant flowers, and diverse colors and shapes, *Dahlia* is known as a world-famous flower. Its tubers are rich in inulin, making it valuable for both medicinal and culinary purposes. In recent years, with the continuous exploration of new cut flower resources, *Dahlia* has become increasingly popular as a new cut flower internationally, in addition to traditional potted plants and floral arrangements. As the primary material for autumn cut flowers, *D. pinnata* is highly valued in countries such as the Netherlands and Australia [2].

The chloroplast genome consists of multiple copies of a circular DNA fragment (110 ~ 210 kb) [3], The chloroplast DNA of plants mostly has a typical quadripartite structure [4], including a large single copy(LSC), a small single copy(SSC) and a inverted repeat(IR).The chloroplast genome of most land plants contains 110–130 genes [5]. Chloroplast is a semi-autonomous organelle in plant cells, inherited from the maternal line. It has a small size, easy assembly, low recombination rate, and is more easily analyzed. Therefore, chloroplast genes belong to maternal lineage genes, and using them to construct a phylogenetic tree can effectively avoid interference from paternal lineage genes. Additionally, compared with nuclear genomes, chloroplast genomes evolve slowly and have higher sequence conservation [6], which is a major advantage for selecting chloroplast genomes for tree construction. Therefore, the chloroplast genome has obvious advantages in studies such as plant classification and species-level phylogenetics, and is an important resource for phylogenetic research [7, 8, 9, 10].

Given the horticultural importance of Asteraceae and the status of chloroplast genome in phylogenetic research, studying the chloroplast genome of *chrysanthemum* can improve its photosynthesis and quality, thus better meeting market demand.Using the chloroplast genome of *D. pinnata* for phylogenetic analysis can help us better understand the Heliantheae alliance, especially the evolutionary relationships between *D. pinnata* and other species, and further improve the phylogenetic relationships of Asteraceae plants.

Result

Chloroplast genome structure of Dahlia. The chloroplast genome was found to be a double-stranded circular DNA molecule exhibiting a typical tetrad structure (Fig. 1) and an average GC content of 38.45%. The full sequence was 152,107 bp, containing an 83,704 bp large single-copy (LSC) region, an 18,347 bp small single-copy (SSC) region, and an 25,028 bp pair of inverted repeats (IR). A total of 134 genes were annotated, including 86 protein-coding genes, 38 tRNA genes, 8 rRNA genes, and 2 pseudogenes. The annotation results were compared with four other *D. pinnata* chloroplast genomes (published, but not released) downloaded from NCBI. Of note, we found that the sequence MW589541.1 lacked *IRb*, suggesting an incomplete assembly. Overall, our results resulted in the annotation of two protein-coding genes not found in MW900254.1, and one tRNA not found in MW900253.1 or MW900255.1 (Table. 1).

Genes drawn inside the circle are transcribed clockwise, and those outsides are counter clockwise [20].

Table 1
Comparison of annotation results of 5 *D. pinnata* chloroplast genomes

Plastome	LSC	SSC	*IRa*	*IRb*	GC Content	Number of Genes	Protein-Coding Genes	tRNA	rRNA	Pseudogenes
OP006582.1	**83704**	**18347**	**25028**	**25028**	**38.45%**	**134**	**86**	**38**	**8**	**2**
MW900254.1	83742	18383	25022	25022	38.44%	131	84	37	8	2
MW900253.1	83710	18345	25029	25029	38.43%	133	86	37	8	2
MW900255.1	83708	18347	25028	25028	38.43%	133	86	37	8	2
MW589541.1	83309	18383	25025	0	38.47%	114	79	29	4	2

We further analyzed 134 genes annotated in the chloroplast genome, including 88 protein-coding genes, of which 46 were related to photosynthesis, 29 to self-replication, and 13 to other or unknown functions (*rps19* and one *ycf1* were pseudogenes) (Table. 2). In addition, out of all genes, 23 contained introns, with 19 having a single intron, two having double introns (Table. 3), and two having reverse splicing events *rps12*.

Table 2
Functional list of chloroplast genes in *D. pinnata*

Category of genes	Group of genes	Name of genes	Number of genes
Genes for photosynthesis	Subunits of ATP synthase	*atpA, atpB, atpE, atpF, atpH, atpI*	46
	Subunits of photosystem I	*psaA, psaB, psaC, psaI, psaJ*	
	Subunits of photosystem II	*psb*A, *psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ, ycf*3	
	Subunits of NADH-dehydrogenase	*ndhA, ndhB, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK*	
	Subunits of cytochrome b/f complex	*petA, petB, petD, petG, petL, petN*	
	Subunit of rubisco	*rbcL*	
Self replication	RNA polymerase	*rpoA, rpoB, rpoC1, rpoC2*	29
	Small subunit of ribosome	*rps11, rps12, rps12, rps14, rps15, rps16, rps18, rps19, rps2, rps3, rps4, rps7, rps7, rps8*	
	Large subunit of ribosome	*rpl14, rpl16, rpl2, rpl2, rpl20, rpl22, rpl23, rpl23, rpl32, rpl33, rpl36*	
Other genes	Subunit of Acetyl-CoA-carboxylase	*accD*	13
	c-type cytochrom synthesis gene	*ccsA*	
	Envelop membrane protein	*cemA*	
	Protease	*clpP*	
	translational initiation factor 1	*infA*	
	Maturase	*matK*	
Genes of unknown function	Conserved open reading frames	*ycf1, ycf1, ycf15, ycf15, ycf2, ycf2, ycf4*	

Table 3
Lengths of introns and exons of chloroplast genes in *D. pinnata* 'Chocolate'

Gene	Strand	chocolate		ExonI	IntronI	ExonII	IntronII	ExonIII
		Start	End	/bp	/bp	/bp	/bp	/bp
trnK-UUU	-	1790	4385	36	2522	37		
rps16	-	5199	6323	40	864	215		
rpoC1	+	16296	19119	432	730	1638		
atpF	-	27089	28350	145	707	410		
trnG-UCC	-	30338	31134	23	724	49		
ycf3	-	42281	44243	124	705	230	751	153
trnL-UAA	+	47197	47718	37	435	50		
trnV-UAC	-	51067	51715	38	574	37		
clpP	-	69566	71568	71	783	294	617	226
petB	+	74515	75924	6	762	642		
petD	+	76108	77297	8	707	475		
rpl16	-	80746	82169	9	1016	399		
rpl2	-	83860	85349	391	665	434		
ndhB	-	93857	96059	777	670	756		
trnI-GAU	+	101627	102480	43	775	35		
trnA-UGC	+	102545	103437	38	820	35		
ndhA	-	118517	120672	553	1060	539		
trnA-UGC	-	132375	133267	38	820	35		
trnI-GAU	-	133332	134185	43	775	35		
ndhB	+	139753	141955	777	670	756		
rpl2	+	150463	151952	391	665	434		

Codon preference analysis. The use of codons shows a preference in the process of plant evolution [21]. Relative synonymous codon usage (RSCU) refers to the ratio of the observed frequency of a specific codon to its expected frequency, and is an effective indicator for detecting codon preference [22]. There are 64 codons encoding 20 known amino acids in the chloroplast genome of 'Chocolate' (Fig. 1). When RSCU < 1.0, it means no preference, RSCU in the range of 1.0 ~ 1.2 shows low preference, RSCU in the range of 1.2 ~ 1.3 shows moderate preference, and when RSCU > 1.3, it represents high preference [23]. Figure (Fig. 1) shows that there are 30, 29, and 29 codons with RSCU > 1.0, 11, 10, and 8 codons with RSCU in the range of 1.0 ~ 1.2, 7, 8, and 9 codons with RSCU in the range of 1.2 ~ 1.3, and 19, 19, and 21 codons show strong preference. Among the codons that show preference, except for UUG which ends with G, the rest of the codons end with A/U, which is a common phenomenon in plant chloroplast genomes [24, 25].

Simple sequence repeat (SSR) analysis. Detection of simple repetitive sequence SSRs [26] in the newly sequenced chloroplast genome sequence of *D. pinnata* 'Chocolate' in this study using MISA software (http://pgrc.ipk-gatersleben.de/misa/misa.html). SSRs of mononucleotides, dinucleotides, trinucleotides, tetranucleotides, pentanucleotides, and hexanucleotides were detected with a repeat threshold of 8, 5, 4, 3, 3, and 3, respectively.

In the chloroplast genome of the *D. pinnata* 'Chocolate', SSR sites with 1–6 nucleotides were detected. The number of SSR sites with 1, 2, 3, 4, 5, and 6 nucleotides are 113, 9, 4, 12, 2, and 1, respectively. Please refer to Table 4 for the summary of SSR sites.

Table 4
Summary table of simple repetitive sequences of *D. pinnata* chloroplast genome

variety	mononucleotide	Dinucleotide	trinucleotide	tetranucleotide	pentanucleotide	Hexanucleotide	Total number of identified SSRs
Chocolate	133	9	4	12	2	1	161

Phylogenetic analysis. We obtained a phylogenetic tree with a high support rate (Fig. 3). Phylogenetic analysis that the *Dahlia* species formed a monophyly. Otherwise, our analysis revealed an interesting incongruency, resulting in Bidens being positioned as the sister to Cosmos (Fig. 2), while nuclear genomic analysis indicates that *Dahlia* is the sister to Cosmos ;The chloroplast genome tree result shows that Eupatorieae is more closely related to Heliantheae, while the nuclear genome tree places Heliantheae as sister to Corepsideae and Eupatorieae as more closely related to Millerieae [27]. Our analysis also revealed the systematic relationship among Millerieae, Heliantheae, and Coreopsideae. Our results are consistent with the results based on chloroplast analysis of Liu et al. [28], indicating that Heliantheae and Millerieae are the most closely related tribes in the Heliantheae alliance. However, our results are inconsistent with the results based on nuclear genome analysis of Zhang et al, which indicated that Heliantheae and Coreopsideae were the most closely related tribes. We suggest that inconsistent phylogenetic relationshipsmay be widespread in Asteraceae, and should be thoroughly evaluated in order to understand the true evolutionary history of this economically-important group of plants.

Discussion

Using chloroplast gene information to construct phylogenetic studies has been one of the key focuses of research in recent decades. At the same time, because early studies had lower levels of resolution and obtaining complete genomes was difficult, small and easily obtained chloroplast genes became more prominent. For early evolutionary studies in Asteraceae, markers from the chloroplast genome were used [29, 30, 31]. Compared with constructing phylogenetic trees using a single gene or a few markers, some families have used the complete chloroplast genome CDS gene set to reconstruct their phylogenetic relationships. Examples include Polygonaceae [32], Dennstaedtiaceae [33], and Brassicaceae [14].The use of the complete chloroplast genome CDS genes to reconstruct phylogenetic relationships has opened the way for the combination of long sequences and multiple loci to construct family-level phylogenetic relationships. It has also laid a foundation for constructing phylogenetic relationships using long sequences. Phylogenetic trees constructed based on the complete chloroplast genome contain more sequence information and are more conducive to comprehensively demonstrating chloroplast genome mutations. Asteraceae is the largest plant family and requires comprehensive phylogenetic research.

Our results show that there are three inconsistent phylogenetic sites in the 15 chloroplast genome phylogenetic studies of the Heliantheae supertribe in the Asteraceae subfamily Asteroideae. Therefore, we speculate that if more sampling sites are added and the scope of phylogenetic research is expanded, more inconsistent phylogenetic sites may be discovered. Given its position in the plant kingdom, comprehensive research on Asteraceae is conducive to improving our understanding of the family's evolutionary relationships and revealing the secrets of its genetic evolution.

Materials and methods

Experimental materials. Material collection in accordance with relevant regulations. For this experiment, fresh samples of *D. pinnata* 'Chocolate' (Fig. 4) were collected from the greenhouse test park (38°82 ' N, 115°44 ' E) located within the east campus of Hebei Agricultural University in Baoding City, Hebei Province. For future research, a specimen was deposited at the College of Horticulture, Hebei Agricultural University (https://yuanyi.hebau.edu.cn/, Shan-Ce Niu, niushance@163.com) under the voucher number DLH20190822. See Fig. 4.

Extraction of chloroplast genomic DNA and establishment of sequence library. The processed leaves were physically disrupted (by ultrasonic oscillation) to break the genomic DNA into fragments of 350bp in size. Then, a small fragment sequencing library was constructed through steps such as end repairing, adding 'A', adding adapters, selecting target fragments and PCR amplification. The 2100 Bioanalyzer and Q-PCR were used to detect the size and quantity of library fragments, to determine whether the library met the

sequencing standards. The library was fixed onto the sequencing chip via bridge PCR. The library was subjected to paired-end 150bp (PE150) sequencing using the Illumina sequencer, generating approximately 10Gb of raw data.

Assembly and annotation of the chloroplast genome. The chloroplast genome was assembled using GetOrganelle [11]. After the initial annotation using the online annotation tool [12], CPGAVAS2 (http://47.96.249.172:16019/analyzer/annotate), inappropriate annotation results were manually edited. The annotated genome was uploaded to NCBI GenBank with accession number OP006582.1.

Reconstructing the phylogenetic tree. In order to examine the phylogenetic relationships among the Heliantheae alliance, the chloroplast genomes from *D. pinnata*-1 and 51 other species were used to construct a phylogenetic tree. The genomes downloaded from NCBI and used for comparison were *Helianthus hirsutus* MT302566.1, *H. tuberosus* MG696658.1, *H. grosseserratus* MT302568.1, *H. schweinitzii* MW381301.1, *Pappobolus lanatus* MT700543.1, T*ithonia diversifolia* MT576958.1, *Aldama anchusifolia* MN337902.1, *Xanthium sibiricum* NC 042232.1, *X. strumarium* OP606043.1, *X. spinosum* MT668935.1, *Ambrosia artemisiifolia* MF362689.1, *A. trifida* NC 036810.2, *Echinacea purpurea* KX548224.1, *E. pallida* KX548218.1, *Acmella paniculata* MZ292978.1, *Rudbeckia laciniata* MN518844.1, *Silphium perfoliatum* MN445187.1, *Eclipta alba* NC 039774.1, *E. prostrata* NC 030773.1, *Chromolaena odorata* NC 050055.1, *Praxelis clematidea* NC 023833.1, *Adenostemma lavenia* MW583043.1, *Ageratum conyzoides* MK905238.1, *Eupatorium fortunei* OK545755.1, *Stevia rebaudiana* NC 066036.1, *Stevia sp. OlivelRa* MT793846.1, *Guizotia abyssinica* NC 010601.1[13], *Sigesbeckia orientalis* NC 053700.1[14], *Smallanthus sonchifolius* MT872397.1, *Galinsoga parviflora* MK737938.1, *G. quadrlRadiata* KX752097.1, *Lasthenia burkei* KM360047.1, *Tagetes erecta* MN203535.1, *T. minuta* NC 065038.1, *T. lemmonii* NC 061912.1, *Flaveria bidentis* MK836182.1, *Bidens menziesii* NC 047273.1, *B. micrantha* NC 047277.1, *B. sandvicensis* NC 047276.1, *B. campylotheca* NC 047270.1, *B. alba* MW551955.1, *B. bipinnata* MW551957.1, *B. biternata* MW551958.1, *B. pilosa* MW551953.1[15], *Cosmos bipinnatus* MN518845.1[16], *D. pinnata-1* OP006582, *D. pinnata-4* MW900255.1, *D. pinnata-2* MW900253.1, *D. pinnata-3* MW900254.1, *Centipeda minima* NC 065155.1, *Blumea balsamifera* BK013127.1, *Ligularia fischeri* NC 039352.1[17], with *C. minima* NC 065155.1, *B. balsamifera* BK013127.1 and *L. fischeri* NC 039352.1 used as an outgroup. All sequences were compared using MAFFT version 7[18], and Mega_11.0.13 was used to determine the optimal model (GTR + R). Finally, the maximum likelihood method was used for phylogenetic analysis[19].

Declarations

Acknowledgements

This work was supported by Hebei Provincial Key R&D Program Project (20326814D), and Hebei Provincial Colleges and Universities Basic Research Task Fee Research Project (KY2021055).

Author contributions statement

S.C.N., and D.F.C. were involved in the conception and design. S.D.D., K.Y.Z., Y.L, and X.R.L. conducted experimentsand analyzed data, wrote the manuscript. Li-Hong Hao and Di-Ying Xiang cultivated, managed plant samples and provided photos. All authors agree to be accountable for all aspects of the work and the final approval of the version to be published.

Data availability

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov] under the accession no. OP006582. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA861061, SAMN29881931, and SRR20392866 respectively.

References

1. Wen Yi He Jin-rong Jiang Hao. 2004. Dahlia. Beijing China Forestry Publishing House.
2. Yan G X. 2020. Overview and enlightenment of Dahlia industry development. Anhui Agricultural Science Bulletin. 26(10): 38-39. in Chinese
3. Sugiura M. 1992. The chloroplast genome. Plant Molecular Biology 19 (1): 149 – 168.
4. Palmer J D Thompson W F. 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell 29 (2): 537 – 550.

5. Lin C P Wu C S Huang Y Y Chaw S M. 2012. The complete chloroplast genome of Ginkgo biloba reveals the mechanism of inverted repeat contraction. Genome Biology and Evolution 4 (3): 374 – 381.
6. Kang H I Lee H O Lee I H Kim I S Shim D. 2019. Complete chloroplast genome of Pinus densiflora Siebold & Zucc. and comparative analysis with five pine trees. Forests 10 (7): 600.
7. Jansen R K Cai Z Raubeson L A Daniell H Depamphilis C W Leebens-Mack J Müller K F Guisinger-Bellian M Haberle R C Hansen A K Chumley T W Lee S B Peery R McNeal J R Kuehl J V Boore J L. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proceedings of the National Academy of Sciences 104 (49): 19369 – 19374.
8. Moore M J Soltis P S Bell C D Burleigh J G Soltis D E. 2010. Phylogenetic analysis of 83 plastid genes further resolves the early diversification of eudicots. Proceedings of the National Academy of Sciences 107 (10): 4623 – 4628.
9. Olsson S Grivet D Vian J C. 2018. Species-diagnostic markers in the genus Pinus: evaluation of the chloroplast regions matK and ycf1. Forest systems 27 (3): 2.
10. Wu F Li M Liao B Shi X Xu Y. 2019. DNA Barcoding Analysis and Phylogenetic Relation of Mangroves in Guangdong Province China. Forests 10 (1): 56.
11. Jin J, Yu W B, Yang J B, et al. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology 21(1): 1-31.
12. Shi L, Chen H, Jiang M, et al. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Research. 47(W1): W65-W73.
13. Dempewolf H, Kane N C, Ostevik K L, et al. 2010. Establishing genomic tools and resources for Guizotia abyssinica (Lf) Cass. -the development of a library of expressed sequence tags, microsatellite loci, and the sequencing of its chloroplast genome. Molecular Ecology Resources. 10(6): 1048-1058.
14. Liu J, Jiang M, Wang L, et al. 2020. The complete chloroplast genome sequence of the medicinal plant Siegesbeckia orientalis L., the first in the genus Siegesbeckia. Mitochondrial DNA Part B. 5(1): 145-146.
15. Lin Y, Duan R, Tan Z, et al. 2020. The complete chloroplast genome of the invasive and Cd-hyperaccumulator herb Bidens pilosa L. (Asteraceae). Mitochondrial DNA Part B. 5(1): 436-438.
16. Jiang M, Chen H, Wang L, et al. 2019. The complete chloroplast genome sequence of Cosmos bipinnatus, the first of the genu s Cosmos[J]. Mitochondrial DNA Part B. 4(2): 4204-4206.
17. Lee J, Lee H, Lee S C, et al. 2016. The complete chloroplast genome sequence of Ligularia fischeri (Ledeb.) Turcz. (Asteraceae). Mitochondrial DNA Part B. 1(1): 4-5.
18. Katoh K, Rozewicki J, Yamada K D. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Briefings in Bioinformatics. 20(4): 1160-1166.
19. Tamura K, Stecher G, Kumar S. 2021. MEGA11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution. 38(7): 3022-3027.
20. Lohse M Drechsel O Kahlau S Bock R. 2013. OrganellarGenomeDRAW-a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Research 41: W575 – W581.
21. Liu Q Xue Q. 2005. Comparative studies on codon usage pattern of chloroplasts and their host nuclear genes in four plant species. Journal of Genetics 84 (1): 55 – 62.
22. Sharp P M Li W H. 1987. The codon adaptation index-a measure of directional synonymous codon usage bias and its potential applications. Nucleic Acids Research 15 (3): 1281 – 1295.
23. Zuo L H Shang A Q Zhang S Yu X Y Ren Y C Yang M S Wang J M. 2017. The first complete chloroplast genome sequences of Ulmus species by de novo sequencing: Genome comparative and taxonomic position analysis. PLoS ONE 12 (2): e0171264
24. Redwan R M Saidin A Kumar S V. 2015. Complete chloroplast genome sequence of MD-2 pineapple and its comparative analysis among nine other plants from the Subclass Commelinidae. BMC Plant Biology 15 (196): 1 – 20.
25. Yan C Du J Gao L Li Y Hou X. 2019. The complete chloroplast genome sequence of watercress (Nasturtium officinale R. Br.): Genome organization adaptive evolution and phylogenetic relationships in Cardamineae. Gene 699: 24 – 36.
26. Beier S Thiel T Munch T Scholz U Mascher M. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics 33: 2583 – 2585.

27. Zhang C, Huang C H, Liu M, et al. 2021. Phylotranscriptomic insights into Asteraceae diversity, polyploidy, and morphological innovation. Journal of Integrative Plant Biology. 63(7): 1273-1293.
28. Liu M, Zhang C, Huang J X, et al. 2015. Reconstruction of phylogenetic relationships between Aster subfamily Asteraceae using low-copy nuclear genes. Chin Bull Bot. 50(5): 549-564. (in Chinese
29. Panero, J.L., and Funk, V.A. (2008). The value of sampling anomalous taxa in phylogenetic studies: Major clades of the Asteraceae revealed. Mol. Phylogen. Evol. 47: 757–782.
30. Panero, J.L., Freire, S.E., Ariza Espinar, L., Crozier, B.S., Barboza, G.E., and Cantero, J.J. (2014). Resolution of deep nodes yields an improved backbone phylogeny and a new basal lineage to study early evolution of Asteraceae. Mol. Phylogen. Evol. 80: 43–53.
31. Shahzadi I, Mehmood F, Ali Z, et al. Chloroplast genome sequences of Artemisia maritima and Artemisia absinthium: Comparative analyses, mutational hotspots in genus Artemisia and phylogeny in family Asteraceae[J]. Genomics, 2020, 112(2): 1454-1463.
32. Zhang H J, Zhang X, Landis J B, et al. Phylogenomic and comparative analyses of Rheum (Polygonaceae, Polygonoideae)[J]. Journal of Systematics and Evolution, 2022.
33. Wang T, Li T Z, Chen S S, et al. Untying the Gordian knot of plastid phylogenomic conflict: A case from ferns[J]. Frontiers in Plant Science, 2022, 13.

Figures

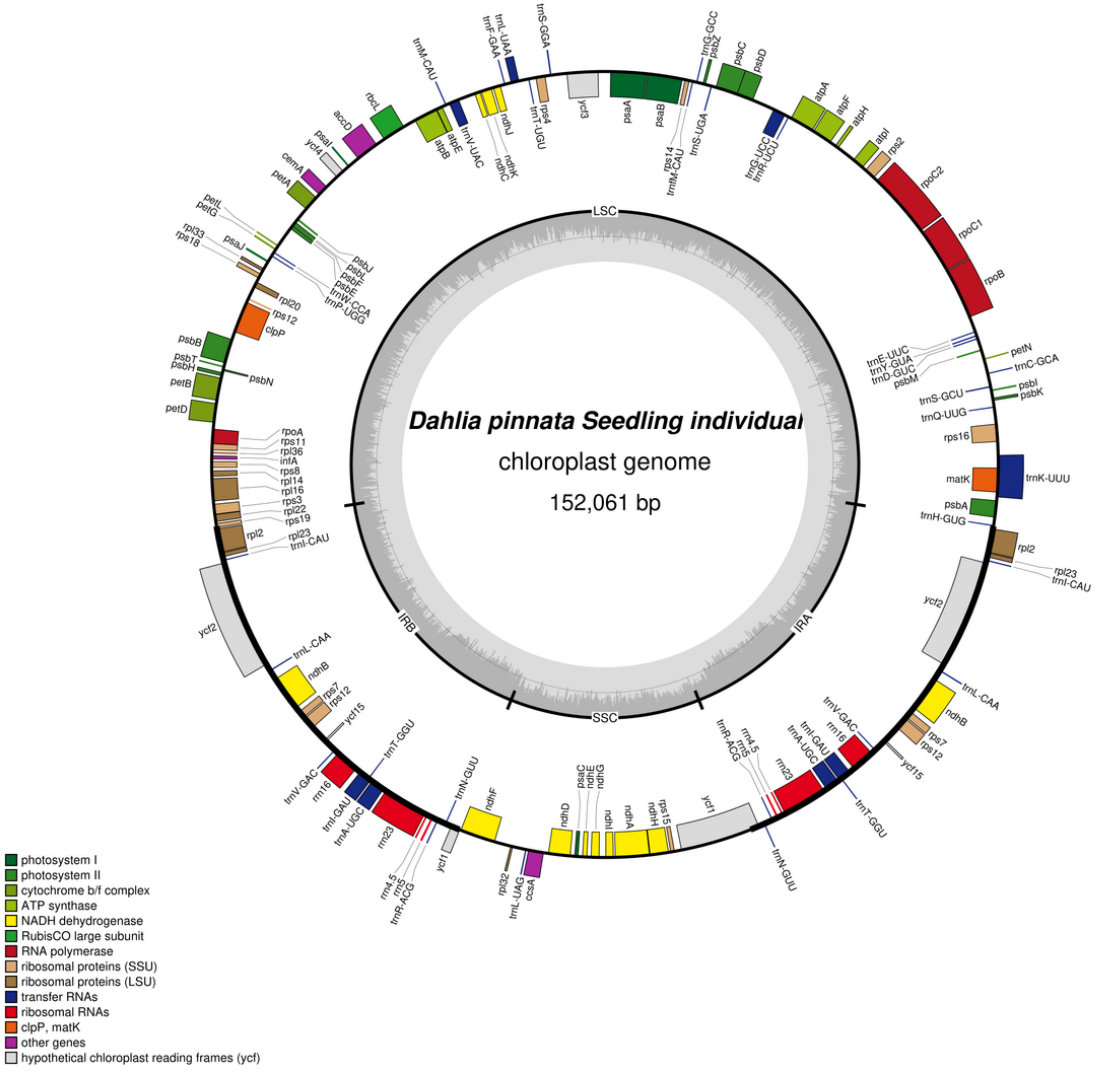

Figure 1

Chloroplast genome map of *D. pinnata* Seedling individual

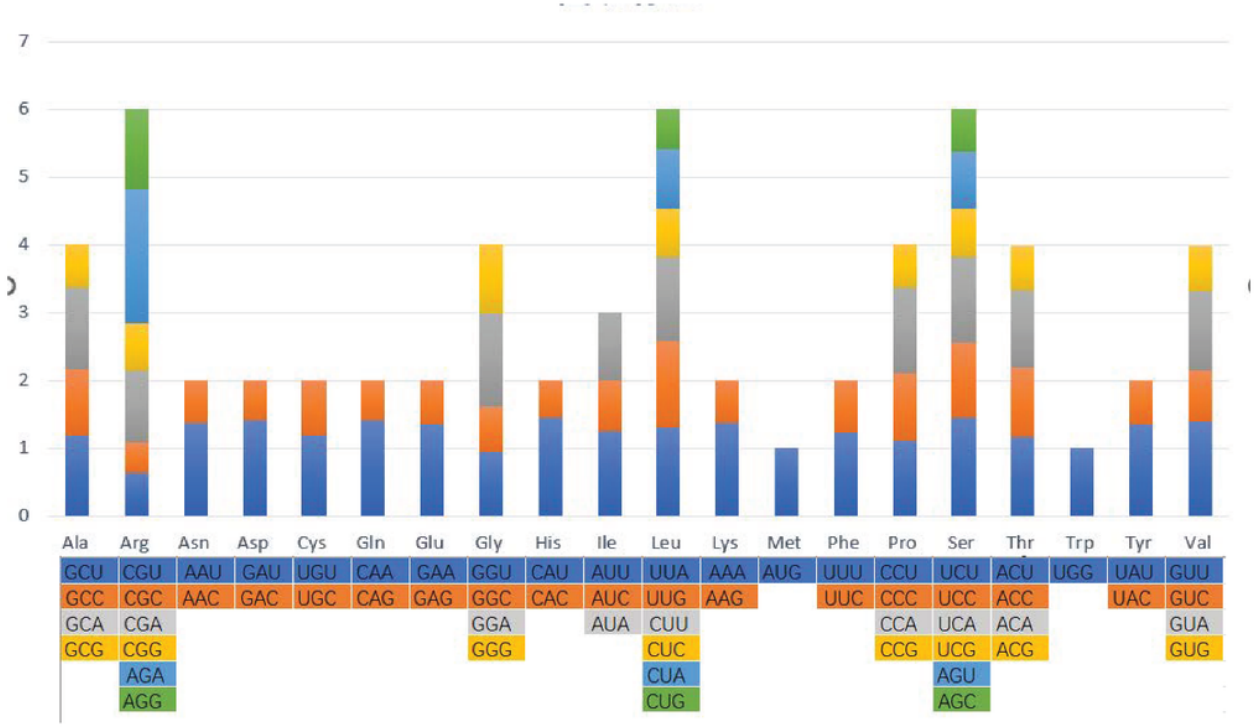

Figure 2

Heat map of codon preference in chloroplast genome of *D. pinnata* 'Chocolate'

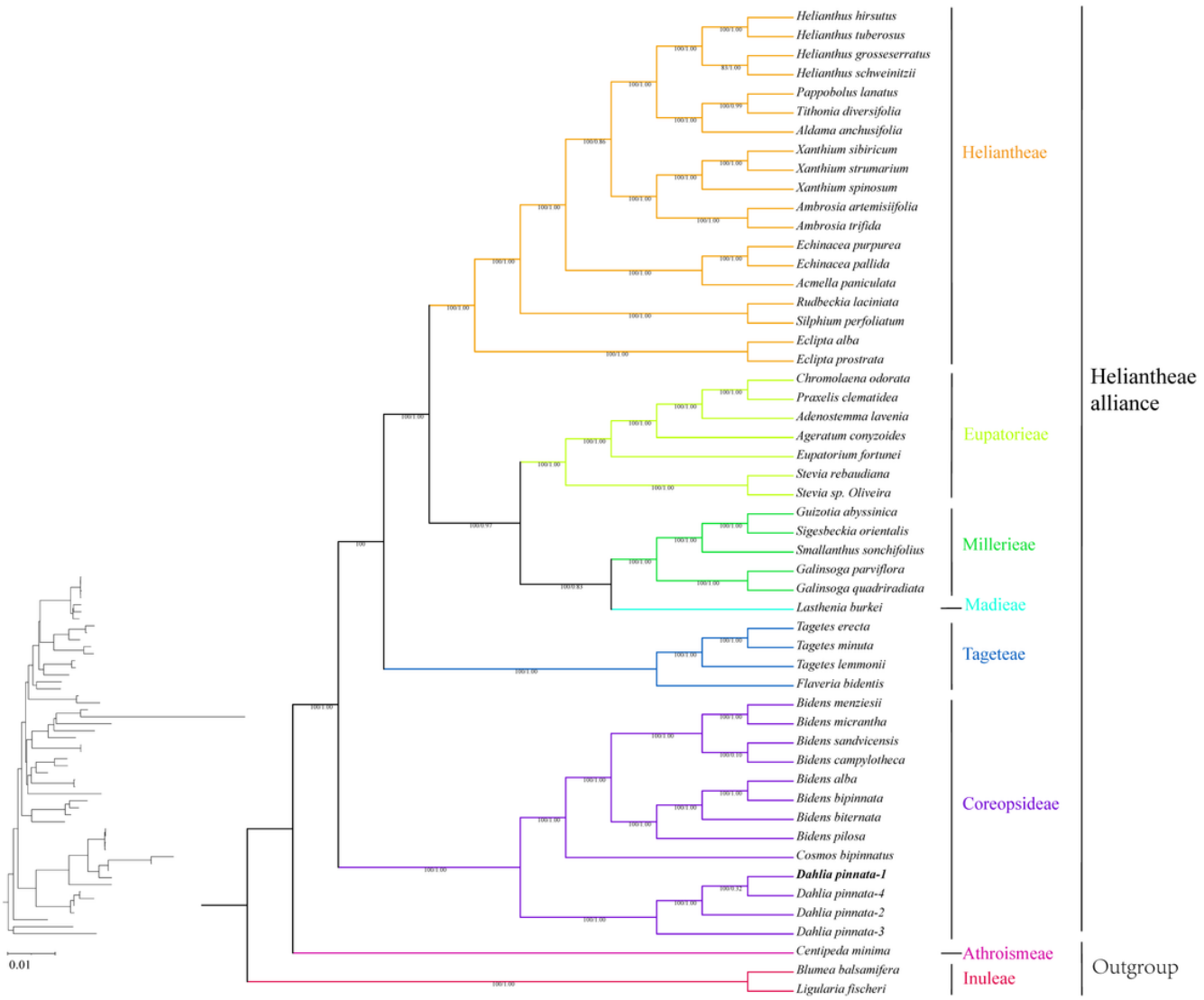

Figure 3

Phylogenetic tree taxa using maximum likelihood (ML) and Bayesian inference (BI) based on 52 shared genes. ML bootstrap values (BS) and posterior probabilities (PP) are shown at nodes. Branches with no values listed have 100% BS and PP of 1.0.

Figure 4

D. pinnata 'Chocolate'. This photo was taken by Di-Ying Xiang and Li-Hong Hao on August 21, 2021.