MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 11, 1487–1491
https://doi.org/10.1080/23802359.2024.2420846

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

The complete genome sequence of the chloroplast of *Bidens aurea*

Hongyu Mao[a], Ziyi Wang[a], Yuanyu Shan[a], Xin Cheng[a] and Jie Yu[a,b]

[a]College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China; [b]Key Laboratory of Horticulture Science for Southern Mountainous Regions, Ministry of Education, Chongqing, China

ABSTRACT

Bidens aurea (Asteraceae), a native of tropical America is now widespread in Asia and the Americas. We explored the *B. aurea* chloroplast genome and conducted a phylogenetic analysis. The chloroplast genome was circular, consisting of a large single copy (LSC) of 83,909 base pairs (bp), a small single copy (SSC) of 18,407 bp, and two inverted repeat regions (IR) of 24,729 bp each. Phylogenetic analysis showed that the 19 *Bidens* taxa were divided into five major clades, and *B. aurea* was most closely related to two species. Our findings offer a high-quality B. aurea chloroplast genome, aiding DNA barcode development and evolutionary history reconstruction.

ARTICLE HISTORY
Received 22 August 2024
Accepted 18 October 2024

KEYWORDS
Bidens aurea; assembly and annotation; chloroplast genome; phylogenetic analysis

1. Introduction

Bidens aurea (Aiton) Sherff 1915, a member of the Coreopsidinae tribe in the Asteraceae, is primarily found in tropical and subtropical regions, and is particularly widespread in subtropical, tropical, and warm-temperate North and South America. This genus comprises c. 240 species, usually annuals and perennials (Walliser et al. 2022). The term 'Bidens' and its Chinese counterpart describe the barbellate or ciliate achenes of the genus. Current *Bidens* research is predominantly centered on its chemical composition and medicinal attributes, while molecular phylogenetics research is relatively limited. Besides their medicinal value, numerous *Bidens* species are also cultivated for their ornamental appeal. Notably, *B. aurea* has gained prominence in the horticultural market for its diverse and vibrant varieties. *B. aurea* is an erect, shrubby plant indigenous to Mexico and the Southern United States. Despite their perennial nature, they will die at the end of the growing season. Pollinators are attracted to the abundance of flowers (Walliser et al. 2022). Our study aim was to explore the *B. aurea* chloroplast genome and conduct a phylogenetic analysis to clarify its evolutionary relationships with other *Bidens* species, which could significantly contribute to future research in medicinal and horticultural applications.

The chloroplast genome plays a decisive role in plant photosynthesis and is the cornerstone for phylogenetic studies and species identification (Szabò and Spetea 2017). With its relatively stable structure and content, it is becoming an excellent model for evolutionary and comparative genomics studies (Shin et al. 2016). Compared to nuclear genomes, the chloroplast genome is relatively conserved, so it also harbors highly variable regions extensively employed as molecular markers. (Liu et al. 2018). Some sequences are recognized as global DNA barcodes for species identification, such as *rbcL, matK, and trnH-psbA* (de Vere et al. 2015).

We sequenced the *B. aurea* chloroplast genome and conducted an extensive comparative genomic analysis with 18 other recognized *Bidens* species. The complete chloroplast genome can then be used as a super-barcode to ascertain *Bidens* taxa phylogenetic position.

2. Materials and methods

2.1. Sample collection

At Beibei Flower Market located in the Beibei District of Chongqing, China (coordinates: N29°48′59.5″, E106°24′51.8″), we carefully collected fresh young *B. aurea* leaves, which were dried using silica gel beads and then dissected (Figure 1). Voucher specimens were stored at the Key Laboratory of Horticulture Science for Southern Mountainous Regions, Ministry of Education, Chongqing, China. For further inquiries, contact Jie Yu (yujie1982@swu.edu.cn) and refer to voucher number GZC0321.

2.2. DNA extraction and sequencing

Total genomic DNA was extracted from dried leaf samples utilizing a DNeasy Plant Mini Kit (Qiagen, USA). DNA sample integrity and purity were evaluated by agarose gel electrophoresis and DNA spectrophotometry. Sample DNA showed clear bands on agarose gels at concentrations above 100 ng/μl with ratios from

CONTACT Jie Yu yujie1982@swu.edu.cn College of Horticulture and Landscape Architecture, Southwest University, Chongqing, China
Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2420846.

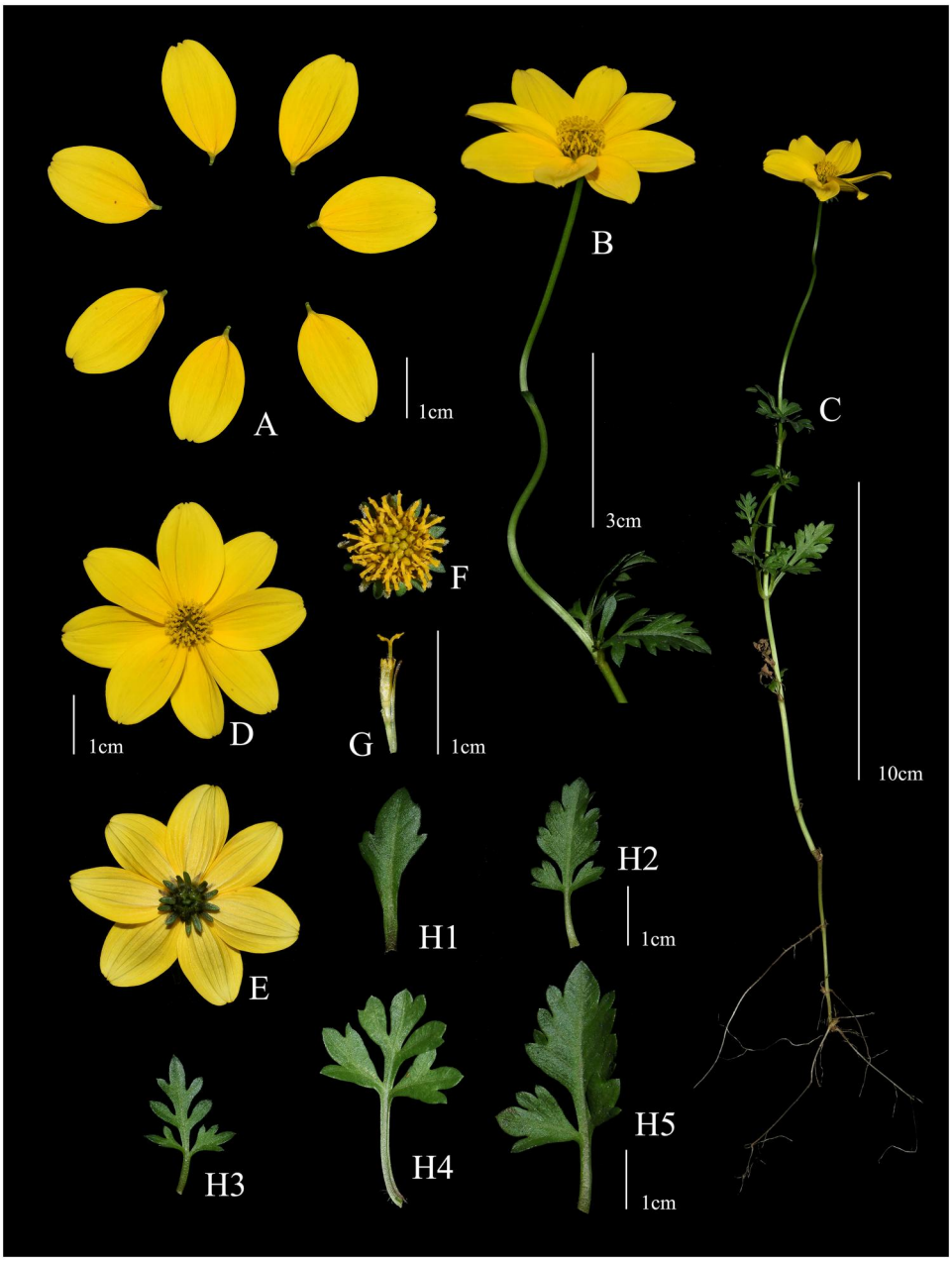

Figure 1. A-ligulate flower, B-flower stem, C-plant, D-flower, E-flower back view, F- capitulum without ligulate flower, G-tubular flower, H1-H5 - different types of leaves.
(Photographed by Ziyi Wang on 2024-03-23).

1.8 to 2.0 and 2.0 to 2.2 for A260/280 and A260/230, respectively. High-quality DNA sequencing libraries were constructed utilizing a TruSeq Nano DNA Sample Preparation Kit (Illumina, USA) and sequenced on the DNBSEQ platform to produce paired-end reads, each 150 bp long. The entire *B. aurea* chloroplast genome was assembled with NOVOPlasty (Dierckxsens et al. 2017). To obtain clean data, we filtered the raw sequences, removing those with a quality score (Q) $<$ 19 and exceeding 50% of the total bases, as well as those containing over 5% ambiguous 'N' bases (Shan et al. 2021).

2.3. Genome assembly and annotation

The clean data were subjected to genome assembly using NOVOPlasty version 2.7.2 (Dierckxsens et al. 2017), with the selection of a 39 bp k-mer length and a coverage depth of 7,870 (Figure S1). To verify spliced alleles accuracy, we utilized Genius version 8.1 (Kearse et al. 2012), aligning the assembled genome sequence to all clean reads. The *B. aurea* chloroplast genome gene composition was annotated with the GeSeq program (Tillich et al. 2017), while the gene map was constructed using OGDRAW v1.3.1 (Greiner et al. 2019). Additionally, the entire *B. aurea* chloroplast genome sequence was submitted to GenBank.

2.4. Phylogenetic analysis

We extracted seventy-five protein-coding genes (PCGs) from the chloroplast genomes of nineteen *Bidens* species. *Chrysanthemum* $\times$ *morifolium* was used as an outgroup. All 19

complete chloroplast genome sequences were sorted by default using the MAFFT version 7.450 online tool (Rozewicki et al. 2019). We then created phylogenetic trees from the sequence using the maximum-likelihood (ML) in RAxML version 8.2.4. (Stamatakis 2014) with 'raxmlHPC-PTHREADS-SSE3 -f a -N 1000 -m GTRGAMMA -x 551314260 -p 551314260' as the corresponding parameters (Shan et al. 2021). Accession numbers of the species in the phylogenetic tree are listed in Table S2.

3. Results

3.1. General features of chloroplast genomes

The complete *B. aurea* chloroplast genome was 151,774 bp long. It included an LSC, (83,909 bp), a small-single copy (SSC, 18,407 bp), and two inverted repeat regions (IR, 24,729 bp). The chloroplast genome comprised a total of 156 genes, among these 114 were unique, containing 80 PCGs, 30 tRNAs, and 4 rRNAs. This chloroplast genome contained 95 protein coding, 49 tRNA, and 12 rRNA genes (Figure 2). 17 regions were fully duplicated, including *trnR, rrn5, rrn4.5, rrn23, trnA, trnI, trnI, rrn16, trnV, ycf15, rps12, rps7, ndhB, trnL, trnI, rpl23,* and *rpl2*. There were no genes with nonstandard stop codons. The genome contained 36 genes with introns, with 30 having 1 intron, and 6 having 2 introns (Table S1). 11 cis-splicing genes had 1 intron, and 2 cis-splicing genes had 2 introns (Supplementary Figure S2). The *rps12* gene was a trans-splicing gene (Supplementary Figure S3). Figure 2 was the result of CPGAVAS2 (Shi et al. 2019). We used cpgview to demonstrate genome annotation correctness (Liu et al. 2023). The *B. aurea* genomic sequence data is publicly accessible on GenBank through the NCBI database with accession number PP596860.

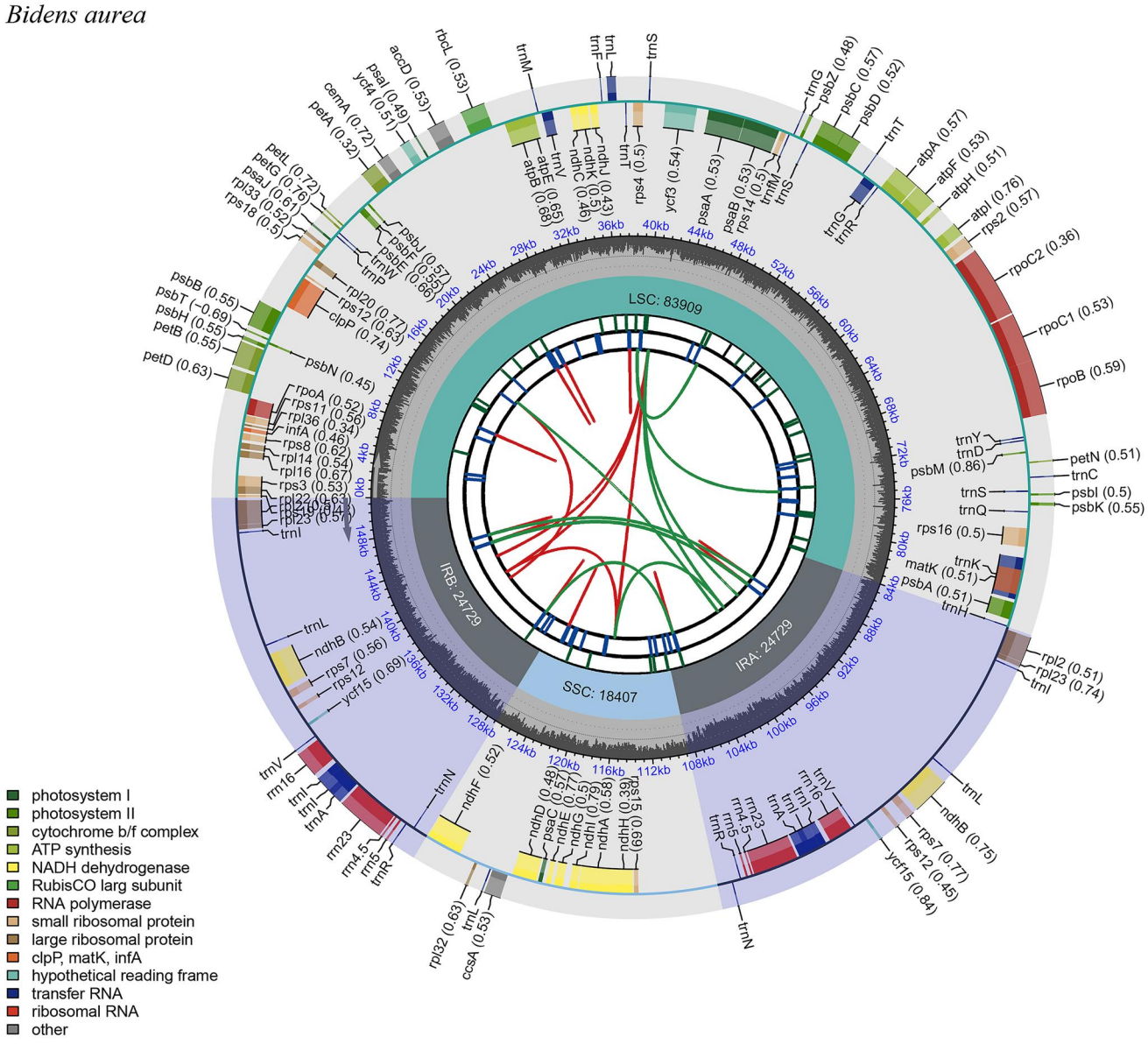

Figure 2. The map consists of four concentric rings: the first ring displays forward and reverse repeats, indicated by red and green arcs; the second ring shows concatenated repeats marked by short lines; the third ring represents microsatellite arrays identified by MISA; and the fourth ring, created with DrawGenemap, illustrates the gene structure on the chloroplast genome, with genes color-coded according to their functional class, as depicted in the bottom left corner. Use symbols to mark genes with introns (*).The correctness of the genome annotation has been demonstrated by using cpgview.

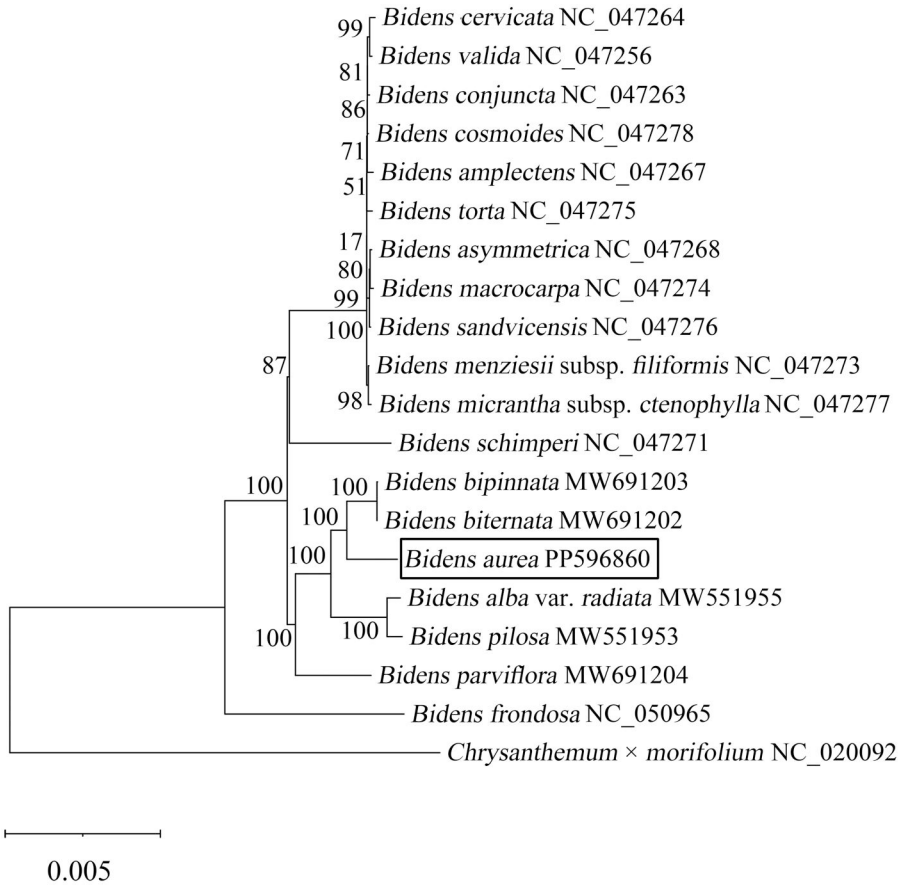

Figure 3. Phylogenetic relationships of Bidens species were inferred using maximum-likelihood (ML) method. The phylogenetic tree was constructed using complete chloroplast genome sequences. The number 0.005 at the bottom of the scale indicates the length of the branches, representing the base substitution frequency of 0.005 at each locus of the genome. The bootstrap value is calculated from 1000 repetitions.
The accession number of the species in the tree: *Bidens bipinnata* (No reference); *Bidens biternata* (No reference); *Bidens aurea* (This study); *Bidens alba var radiata* (No reference); *Bidens pilosa* (No reference); *Bidens parviflora* (No reference); *Bidens valida* (Knope et al.,2020); *Bidens cervicata* (Knope et al.,2020); *Bidens conjuncta* (Knope et al.,2020); *Bidens cosmoides* (Knope et al.,2020); *Bidens amplectens* (Knope et al.,2020), *Bidens macrocarpa* (Knope et al.,2020); *Bidens asymmetrica* (Knope et al.,2020); *Bidens sandvicensis* (Knope et al.,2020); *Bidens torta* (Knope et al.,2020); *Bidens micrantha* subsp ctenophylla (Knope et al.,2020); *Bidens menziesii* subsp filiformis (Knope et al.,2020); *Bidens schimperi* (Knope et al.,2020); *Bidens frondosa* (Li et at.,2020); *Chrysanthemum x morifolium* (Li et al.,2012).

3.2. Phylogenetic analysis

Both Maximum Likelihood and Bayesian reference methods produced congruent tree topologies. The phylogenetic analysis delineated five distinct clades, each further subdivided into multiple subclades. Among these clades, six species (*B. bipinnata*, *B. biternata*, *B. aurea*, *B. alba* var. *radiata*, *B. pilosa*, *B. parviflora*) formed a distinct cluster (depicted in blue in Figure 3), and the other 13 *Bidens* taxa clustered together. Among the species we sequenced, *B. aurea* exhibited the closest relationship with *B. bipinnata* and *B. biternata*. These results illustrate the usefulness of a whole chloroplast genome sequence as a super barcode for species identification, providing high resolution at the species level.

4. Discussion

We provide the first description of the *B. aurea* chloroplast genome genomic sequence. Upon consulting the NCBI nucleotide database, we observed minimal variation in size among *Bidens* species chloroplast genomes, while maintaining conservation of genetic structure, gene composition, and genetic process. These observations corroborated previous studies indicating maternal inheritance and high plastome structure conservation in most angiosperms (Birky 1995). The chloroplast gene rarely disappears without a trace, rather they are either placed in nuclear genomes or, in principle, replaced by genomes (Wang et al. 2018). Earlier studies reported an absence of *clpP* in *B. bipinnata* and *psbD* in *B. alba* and *B. bipinnata* (Zhang et al. 2023). In our investigation, *clpP* was similarly lost from *B. aurea*, while *psbD* remained intact, and *psbN*, *ycf3*, and *ycf4* were absent. Notably, the removal of a *clpP* intron has been observed in *Vicia sepium* (Li et al. 2020). Therefore, further studies examining more Bidens species are needed to explore genetic variation and develop molecular markers for each species.

5. Conclusion

We conducted *B. aurea* chloroplast whole-genome sequencing. The chloroplast genome structure included a large single copy (LSC) region of 83,909 base pairs (bp), a small single copy (SSC) region of 18,407 bp, and two inverted repeat (IR) regions, each 24,729 bp. The chloroplast genome comprised 156 genes, of which 114 were unique, consisting of 80 protein-coding genes (PCGs), 30 transfer RNA genes (tRNAs), and

4 ribosomal RNA genes (rRNAs). Notably, this chloroplast genome contained 95 protein-coding genes, 49 transfer RNA genes, and 12 ribosomal RNA genes.

Overall, our results contribute to the *Bidens* chloroplast genomes dataset and provide a foundation for genus phylogenetic reconstruction. We provide preliminary data for exploring the chloroplast genomes of *B. aurea* cultivars, which is significant for tracing genetic variations within the genus. Moreover, we conducted in-depth studies on plastid genes to enhance our understanding of plastid gene diversity within *Bidens* taxa.

Acknowledgments

The authors are thankful for the financial support provided by Southwest University.

Authors' contributions

Hongyu Mao conducted the experiments, analyzed the data, prepared figures and tables, wrote or revised paper drafts, and approved the final draft. Ziyi Wang prepared figures and tables, collected and determined the samples, drafted and reviewed drafts of the article, and approved the final draft. Yuanyu Shan conceived the conception and design, authored or reviewed drafts of the article, and approved the final draft. Xin Cheng analyzed the data, drafted or reviewed the paper, and approved the final draft. Jie Yu performed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The National Training Program of Innovation and Entrepreneurship for Undergraduates [202310635011].

Data availability statement

The following information was supplied regarding data availability: Raw data is available in the Supplementary Files. 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/. The Gene bank accession number is PP596860. The associated 'BioProject' 'SRA', and 'Bio-Sample' numbers are PRJNA1117199, SRR29203336, and SAMN41562697 respectively. The sample has been deposited in the herbarium of Southwest University, Chongqing, China with the accession number: HWA120002.

References

Birky CW. Jr., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. Proc Natl Acad Sci U S A. 92(25):11331–11338. doi:10.1073/pnas.92.25.11331.

de Vere N, Rich TC, Trinder SA, Long C. 2015. DNA barcoding for plants. Methods Mol Biol. 1245:101–118. doi:10.1007/978-1-4939-1966-6_8.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18. doi:10.1093/nar/gkw955.

Greiner S, Lehwark P, Bock R. 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47(W1):W59–w64. doi:10.1093/nar/gkz238.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12): 1647–1649. doi:10.1093/bioinformatics/bts199.

Li C, Zhao Y, Xu Z, Yang G, Peng J, Peng X. 2020. Initial characterization of the chloroplast genome of vicia sepium, an important wild resource plant, and related inferences about its evolution. Front Genet. 11:73. doi:10.3389/fgene.2020.00073.

Liu ML, Fan WB, Wang N, Dong PB, Zhang TT, Yue M, Li ZH. 2018. Evolutionary analysis of plastid genomes of seven Lonicera L. species: implications for sequence divergence and phylogenetic relationships. Int J Mol Sci. 19(12):4039. doi:10.3390/ijms19124039.

Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, Liu C. 2023. CPGView: a package for visualizing detailed chloroplast genome structures. Mol Ecol Resour. 23(3):694–704. doi:10.1111/1755-0998.13729.

Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. 2019. MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. 47(W1):W5–w10. doi:10.1093/nar/gkz342.

Shan Y, Pei X, Yong S, Li J, Qin Q, Zeng S, Yu J. 2021. Analysis of the complete chloroplast genomes of Scutellaria tsinyunensis and Scutellaria tuberifera (Lamiaceae). Mitochondrial DNA B Resour. 6(9): 2672–2680. doi:10.1080/23802359.2021.1920491.

Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 47(W1):W65–w73. doi:10.1093/nar/gkz345.

Shin DH, Lee JH, Kang SH, Ahn BO, Kim CK. 2016. The complete chloroplast genome of the hare's ear root, bupleurum falcatum: its molecular features. Genes (Basel). 7(5):20. doi:10.3390/genes7050020.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313. doi:10.1093/bioinformatics/btu033.

Szabò I, Spetea C. 2017. Impact of the ion transportome of chloroplasts on the optimization of photosynthesis. J Exp Bot. 68(12):3115–3128. doi:10.1093/jxb/erx063.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq - versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–w11. doi:10.1093/nar/gkx391.

Walliser B, Marinovic S, Kornpointner C, Schlosser C, Abouelnasr M, Hutabarat OS, Haselmair-Gosch C, Molitor C, Stich K, Halbwirth H. 2022. The (Bio)chemical base of flower colour in *Bidens ferulifolia*. Plants (Basel). 11(10):1289. doi:10.3390/plants11101289.

Wang YH, Wicke S, Wang H, Jin JJ, Chen SY, Zhang SD, Li DZ, Yi TS. 2018. Plastid genome evolution in the early-diverging legume subfamily cercidoideae (fabaceae). Front Plant Sci. 9:138. doi:10.3389/fpls.2018.00138.

Zhang D, Tu J, Ding X, Guan W, Gong L, Qiu X, Huang Z, Su H. 2023. Analysis of the chloroplast genome and phylogenetic evolution of *Bidens pilosa*. BMC Genomics. 24(1):113. doi:10.1186/s12864-023-09195-7.