

## The complete chloroplast genome sequence of *Crassocephalum crepidioides* (Benth.) S. Moore. (Asteraceae)

Song Wei-Cai<sup>a</sup>, Liu Xiao-Fan<sup>a</sup>, Shuo Wang<sup>a</sup>, Zhang Rui-Qing<sup>a</sup>, Shi Chao<sup>a</sup> and Yun-Jiao Zhang<sup>b</sup>

<sup>a</sup>College of Marine Science and Biological Engineering, Qingdao University of Science and Technology, Qingdao, China; <sup>b</sup>Kunming Medical University Haiyuan College, Kunming, China

### ABSTRACT

The complete chloroplast (cp) genome of *Crassocephalum crepidioides* was sequenced and assembled for the first time. In this study, the total genome size is 150,596 bp in length and demonstrates a typical quadripartite structure containing a large single copy (LSC, 82,575 bp) and a small single copy (SSC, 18,293 bp), separated by a pair of inverted repeats (IRa, IRb) of 24,864 bp. The G + C content of this cp genome was 37.21%. Gene annotation analysis identified 130 genes including 85 protein-coding genes, 37 transfer RNA, and 8 ribosomal RNA genes. The maximum-likelihood phylogenetic analysis result showed that *C. crepidioides* was closely related to *Nannoglottis ravida* in the phylogenetic relationship.

### ARTICLE HISTORY

Received 18 March 2021  
Accepted 19 May 2021

### KEYWORDS

Asteraceae; *Crassocephalum crepidioides*; chloroplast genome; phylogenetic relationship

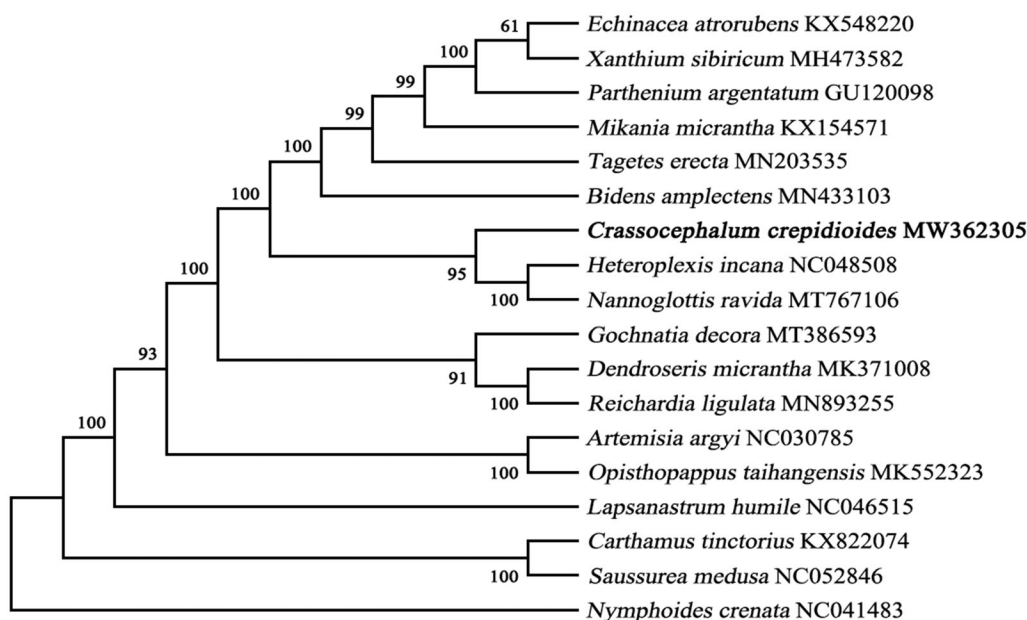
Chloroplasts is a double membrane-bounded organelle in of plants. It plays important metabolic roles, including photosynthesis, amino acid and lipid synthesis (Daniell et al. 2016; Mehmood et al. 2020a). Chloroplast genomes are inherited from the mother, which helps to study phylogenetic relationships (Mehmood et al. 2020b). *Crassocephalum crepidioides* (Benth.) S. Moore. (Henderson 1973) is an annual edible plant that is widely distributed in tropical and subtropical areas (Rajesh 2011). It is an erect, less branched herb, about 40–100 cm tall. The stem is stout, soft, angular, apex with short, thick hair (Kostermans et al. 1987), the leaves are elliptic to ovate, and the seeds consist of floating balls of many silky white hairs, which can be blown away by the wind. This plant grows abundantly in tree crop plantations (Dairo and Adanlawo 2007). Edible leaves and stems are often used to treat indigestion or as a laxative (Asif 2016). Also, extracts from this plant have been shown to have chemoprophylactic and anti-inflammatory properties for cancer (Hou et al. 2007). Studying the chloroplasts of *C. crepidioides* is of great significance to further taxonomic and population genetics studies of the species (Ahmed et al. 2012; Guo et al. 2017).

Fresh leaves of *C. crepidioides* were collected from Panlong District, Kunming City, Yunnan Province, China (24°23'N, 102°10'E), and the voucher specimen and DNA were deposited at Qingdao University of Science and Technology (Chao Shi, [chch1111@aliyun.com](mailto:chch1111@aliyun.com)) under a voucher number: HY0516. Total genomic DNA was extracted from fresh leaves using modified CTAB (Allen et al. 2006), the high-quality DNA was sent to construct a genomic library and sequenced using the Illumina HiSeq platform in Novogene (Nanjing, China). About 4.5 Gb high quality,

2 × 150 bp pair-end reads were obtained and were used to assemble the complete chloroplast genome of *C. crepidioides* (Wang et al. 2018). The *rbcl* gene of *C. crepidioides* (Genbank accession no. MN268502) was used as a seed to assemble the complete chloroplast genome of *C. crepidioides* (Genbank accession no. MW362305) by NOVOPlasty4.2 (Dierckxsens et al. 2017). We also deposited the raw sequencing reads in SRA with Accession no. SRR13823287. Gene annotation was performed with the GeSeq (Michael et al. 2017) and manually corrected for codons and gene boundaries using the Sequin.

The complete chloroplast genome reported here is 150,596 bp in length and exhibits a typical quadripartite structure, consisting of a pair of inverted repeat regions (IRa and IRb) with same length (24,864 bp) separated by the large single copy (LSC, 82,575 bp) and small single copy (SSC, 18,293 bp) regions. The overall GC content is 37.21%, and the corresponding values of the LSC, SSC and IR regions are 35.35%, 30.31%, and 42.84%, respectively. The chloroplast genome of *C. crepidioides* comprised 130 genes, including 85 protein-coding genes, 37 transfer RNA, and 8 ribosome RNA. Noticeably, nine protein-coding genes (*rps16*, *rpoC1*, *atpF*, *petB*, *petD*, *rpl16*, *rpl2*, *ndhB*, and *ndhA*) were disrupted by one intron, and three genes (*clpP*, *rps12*, and *ycf3*) by two.

An alignment comprising the complete chloroplast genome sequences of *C. crepidioides* and other 17 related taxa of Asteraceae was performed in MAFFT version 7.407 (Nakamura et al. 2018; Yupeng et al. 2020). *Nymphoides crenata* was selected as the outgroup of the phylogenetic tree (Journal 2011). Model selected process in Mega version X (Kumar et al. 2018), and GTR + G + I was selected as the optimal model by the Akaike Information Criterion. Phylogenetic tree



**Figure 1.** A maximum-likelihood (ML) tree illustrates the phylogenetic position of *C. crepidioides* among part of Asteraceae species. The number on each node indicates bootstrap support value. After species is the chloroplast genome sequence login number used by GenBank.

was constructed using the maximum-likelihood (ML) method and the bootstrap was set to 1000 times iteration in the Mega version X (Figure 1). The phylogenetic analysis (ML) results clearly showed that *C. crepidioides* was belonged to Asteraceae and closer to *Nannoglottis ravida* and *Heteroplexis incana*, these findings further enriched the phylogenetic relationship of the family Asteraceae and will provide useful genetic information for promoting the evolutionary studies of Asteraceae species.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by the National Natural Science Foundation of China [No. 31801022] and Shandong Province Natural Science Foundation of China [No. ZR2019BC094].

## Data availability statement

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. MW362305. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA705820, SRR13823287, and SAMN18103693, respectively.

## References

Ahmed I, Biggs PJ, Matthews PJ, Collins LJ, Hendy MD, Lockhart PJ. 2012. Mutational dynamics of aroid chloroplast genomes. *Genome Biol E*. 4(12):1316–1323.

- Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethyl ammonium bromide. *Nat Protoc*. 1(5):2320–2325.
- Asif M. 2016. Mini review on important biological properties of benzofuran derivatives. *J Anal Pharm Res*. 3(2):00050.
- Hou CC, Chen Y-P, Wu J-H, Huang C-C, Wang S-Y, Yang N-S, Shyur L-F. 2007. A galactolipid possesses novel cancer chemo preventive effects by suppressing inflammatory mediators and mouse B16 melanoma. *Cancer Res*. 67(14):6907–6915.
- Dairo FAS, Adanlawo IG. 2007. Nutritional quality of *Crassocephalum crepidioides* and *Senecio bialafrae*. *Pak J Nutr*. 6(1):35–39.
- Daniell H, Lin CS, Yu M, Chang WJ. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol*. 17(1):1–29.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 45(4):e18.
- Guo X, Liu J, Hao G, Zhang L, Mao K, Wang X, Zhang D, Ma T, Hu Q, Al-Shehbaz IA, et al. 2017. Plastome phylogeny and early diversification of Brassicaceae. *BMC Genomics*. 18(1):1–9.
- Henderson RJF. 1973. *Crassocephalum crepidioides* (Benth.) S. Moore in Australia: in invasive species compendium. *Proc R Soc Queensland*. 84(4):55–60.
- Journal I. 2011. A phylogenetic study of the order Asterales using one morphological and three molecular data sets. *Int J Plant Sci*. 164(4): 553–578.
- Kostermans AJGH, Wirjahardja S, Dekker RJ. 1987. The weeds: description, ecology and control. Weeds of rice in Indonesia. Jakarta (Indonesia): Balai Pustaka. p. 24–565.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol*. 35(6):1547–1549.
- Mehmood F, Abdullah , Shahzadi I, Ahmed I, Waheed MT, Mirza B. 2020a. Characterization of *Withania somnifera* chloroplast genome and its comparison with other selected species of Solanaceae. *Genomics*. 112(2):1522–1530.
- Mehmood F, Abdullah , Ubaid Z, Bao Y, Pocza P, Mirza B. 2020b. Comparative plastomics of ashwagandha (*Withania, solanaceae*) and identification of mutational hotspots for barcoding medicinal plants. *Plants*. 9(6):1–20.

- Michael T, Pascal L, Tommaso P, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. Geseq-versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 45(1):6–11.
- Nakamura T, Yamada KD, Tomii K, Katoh K. 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics.* 34(14):2490–2492.
- Rajesh KJ. 2011. Terpene composition of *Crassocephalum crepidioides* from Western Ghats region of India. *Int J Nat Prod Res.* 1(2):19–22.
- Wang X, Cheng F, Rohlsen D, Bi C, Wang C, Xu Y, Wei S, Ye Q, Yin T, Ye N, et al. 2018. Organellar genome assembly methods and comparative analysis of horticultural plants. *Hortic Res.* 5(1):3.
- Yupeng W, Jiyuan L, Zhengqi F, Dongyang W, Hengfu Y, Xinlei L. 2020. Characterization of the complete chloroplast genome of *Camellia brevistyla*, an oil-rich and evergreen shrub. *Mitochondrial DNA Part B.* 5(1):386–387.