The complete chloroplast genome sequence of *Cosmos bipinnatus*, the first of the genus *Cosmos*

Mei Jiang, Haimei Chen, Liqiang Wang, Jinwen You and Chang Liu

**ABSTRACT**

*Cosmos bipinnatus* has been used widely in traditional medicines. Here, we sequenced and assembled the complete chloroplast genome of *C. bipinnatus*. This genome is 150,356 bp in size with a pair of inverted repeats (IRs) of 25,082 bp, a large single-copy (LSC) region of 83,003 bp, and a small single-copy (SSC) region of 18,397 bp. It contains 112 unique genes, including 80 protein-coding, 4 rRNA, and 28 tRNA genes. The phylogenomic analysis showed the *C. bipinnatus* and species belonging to the Madieae, Millerieae, Heliantheae, and Eupatorieae tribes were clustered together. The availability of chloroplast genome provided valuable information for future conservation, taxonomy, evolution, and differentiation studies of *C. bipinnatus*.

**CpGAVAS2** was used to annotate the chloroplast genome (Shi et al. 2019).

The chloroplast genome of *C. bipinnatus* (GenBank accession number: MN518845) is 150,356 bp in size with a pair of inverted repeats (IRs) of 25,082 bp separated by a large single-copy (LSC) region of 83,003 bp and a small single-copy (SSC) region of 18,397 bp. The chloroplast genome encoded 129 genes, of which 112 are unique genes including 80 protein-coding, 4 ribosome RNA (rRNA), and 28 transfer RNA (tRNA) genes. Among them, seven protein-coding genes had one intron and two protein-coding genes had two introns. Six tRNA genes were found to contain one intron. The GC content of the whole genome was 37.57%, of which the protein-coding, the rRNA, and the tRNA genes were 37.83, 54.67, and 53.15%, respectively. Within the protein-coding regions, the GC contents for the first, second and third positions of the codons were 45.53, 38.04, and 29.92%, respectively.

We obtained 51 complete chloroplast genome sequences to explore the phylogenomic relationships among the species from the Asteroidae subfamily. A total of 57 shared proteins present among these chloroplast genomes were subjected to multiple sequence alignment using CLUSTALW2 (version 2.0.12) (Thompson et al. 2002). Then, the phylogenomic tree was constructed using the maximum likelihood method implemented in RaxML (version 8.2.4) (Stamatakis 2014). As shown (Figure 1), the 11 tribes were divided into four
branches. All species from the same tribe were grouped together, consistent with the current taxonomic classification. The phylogenomic tree showed the *C. bipinnatus* and species belonging to the Madieae, Millerieae, Heliantheae, and Eupatorieae tribes were clustered together with a support value of 83. The availability of chloroplast genome provided valuable information for future conservation, taxonomy, evolution, and differentiation studies of *C. bipinnatus*.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by Chinese Academy of Medical Sciences, Innovation Funds for Medical Sciences (CIFMS) [2017-12M-1-013, 2016-12M-3-016], National Science Foundation Funds [81872966], and National Science &Technology Fundamental Resources Investigation Program of China [2018FY100705]. The funders were not involved in the study design, data collection, and analysis, decision to publish, or manuscript preparation.

**ORCID**

Chang Liu http://orcid.org/0000-0003-3879-7302
References

Akihisa T, Yasukawa K, Oinuma H, Kasahara Y, Yamanouchi S, Takido M, Kumaki K, Tamura T. 1996. Triterpene alcohols from the flowers of compositae and their anti-inflammatory effects. Phytochemistry. 43(6):1255–1260.

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

NCBI Taxonomy. (2019, September 28). Cosmos bipinnatus. Retrieved from https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=51277.

Olajuyigbe O, Ashafa A. 2014. Chemical composition and antibacterial activity of essential oil of Cosmos bipinnatus Cav. Leaves from South Africa. Iran J Pharm Res. 13(4):1417–1423.

Saito K. 1974. Distribution of flavonoids and related compounds in various parts of Cosmos bipinnatus. Z Pflanzenphysiol. 71(1):80–82.

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.

Sohn SH, Yun BS, Kim SY, Choi WS, Jeon HS, Yoo JS, Kim SK. 2013. Anti-inflammatory activity of the active components from the roots of Cosmos bipinnatus in lipopolysaccharide-stimulated RAW 264.7 macrophages. Nat Prod Res. 27(11):1037–1040.

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

Thompson JD, Gibson TJ, Higgins DG. 2002. Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinformatics. 00:2.3.1–2.3.22.