

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

Mei Jiang^{a*}, Haimei Chen^{a*}, Liqiang Wang^a, Jinwen You^b and Chang Liu^a 

^aKey Laboratory of Bioactive Substances and Resource Utilization of Chinese Herbal Medicine from Ministry of Education, Engineering Research Center of Chinese Medicine Resources from Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P. R. China; ^bInstitute of Chinese Herbal Medicine, Hubei Academy of Agricultural Sciences, Enshi, Hubei, P. R. China

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 region (SSC) 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*.

ARTICLE HISTORY

Received 30 October 2019
Accepted 9 November 2019

KEYWORDS

Cosmos bipinnatus;
chloroplast genome;
phylogeny


Cosmos bipinnatus, commonly called the garden cosmos or Mexican aster, belongs to the *Cosmos* genus, Asteraceae family (NCBI:txid51277 2019). Chemical components with diverse structures, including monoterpenes (Olajuyigbe and Ashafa 2014), sesquiterpene (Sohn et al. 2013), triterpene alcohols (Akihisa et al. 1996), and flavonoids (Saito 1974) had been isolated from *C. bipinnatus*. Pharmacological studies indicated the extracts of *C. bipinnatus* has anti-inflammatory activity by inhibiting the expression of inducible nitric oxide synthase, cyclooxygenase-2, and pro-inflammatory cytokines (Sohn et al. 2013). In contrast, there is little research on the molecular genetics of *C. bipinnatus*, which limited the resource conservation, species identification, and drug development of *C. bipinnatus*. In this study, we reported the first complete chloroplast genome sequence of *C. bipinnatus*.

Fresh leaf samples were collected from the Central China Medicinal Botanical Garden, Enshi, Hubei China (E30°17'84", N109°74'39") and identified as from *C. bipinnatus* by Professor Jinwen You. The genomic DNA was extracted with plant genomic DNA kit (Tiangen Biotech, Beijing, China), the genome sequence was completed using the Hiseq 2500 platform (Illumina, San Diego, CA, USA). A voucher specimen and its DNA (accession number: 201808281) were deposited at Institute of Medicinal Plant Development. The raw sequence data were assembled into a chloroplast genome by NOVOPlasty (version 2.7.2) (Dierckxsens et al. 2017).

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 Asteroideae 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

CONTACT Chang Liu  cliu6688@yahoo.com  Key Laboratory of Bioactive Substances and Resource Utilization of Chinese Herbal Medicine from Ministry of Education, Engineering Research Center of Chinese Medicine Resources from Ministry of Education, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, 100193, P. R. China

*These authors have contributed equally to this work.

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

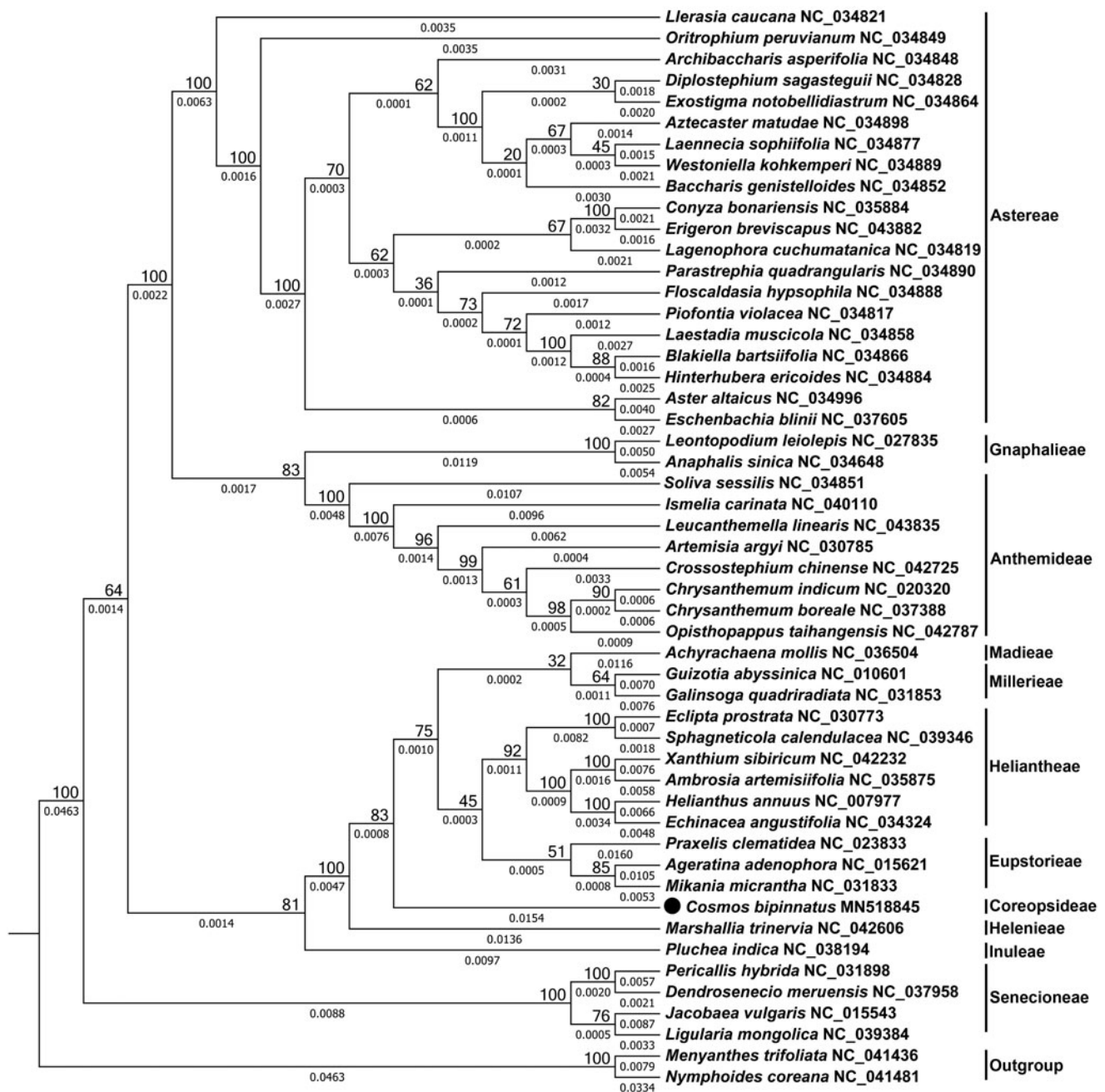


Figure 1. The phylogenetic tree of species from Asteroideae subfamily constructed using the maximum likelihood (ML) method using the 57 shared proteins, RPS12, PSBA, MATK, PSBI, PSBM, RPOB, RPOC2, RPS2, ATPH, ATPF, ATPA, PSBD, PSAB, PSAA, NDHJ, NDHK, NDHC, ATPE, ATPB, RBCL, PSAl, YCF4, CEMA, PETA, PSBF, PSBE, PETL, PETG, RPL33, RPS18, RPL20, PSBT, PSBN, PSBH, PETB, PETD, RPS11, RPL36, INFA, RPS8, RPL14, RPL16, RPS3, RPL23, NDHB, RPS7, RPS15, NDHH, NDHA, NDHI, NDHG, NDHE, PSAC, NDHD, CCSA, RPL32, and NDHF. The number below each line represents the branch length, the number above each node represents the bootstrap values. The GenBank accession number for the corresponding chloroplast sequences were shown after the Latin name of the species.

branches. All species from the same tribe were grouped together, consistent with the current taxonomic classification. The phylogenetic 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-I2M-1-013, 2016-I2M-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.