

## The complete chloroplast genome and phylogenetic analysis of *Cuminum cyminum*

Jing Zhou<sup>a</sup> and Zhenwen Liu<sup>b</sup>

<sup>a</sup>School of Pharmaceutical Science & Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Kunming, China; <sup>b</sup>Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

### ABSTRACT

*Cuminum cyminum* (Apiaceae) is an economically important plant, whose fruits are usually used as flavoring, and also have pharmacological activities such as antioxidant, antibacterial, hypolipidemic, and so on. In this study, we assembled and annotated complete chloroplast (cp) genome sequence of *C. cyminum*. The results showed that the complete cp genome of *C. cyminum* was 157,839 bp in length, consisting of a large single-copy (LSC) region of 83,927 bp, a small single-copy (SSC) region of 17,598 bp, and two inverted repeat regions (IRa and IRb) of 28,157 bp. In total, 131 genes were annotated, comprising of 86 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. The phylogenetic analysis indicated that *C. cyminum* belongs to the tribe Scandiceae, and showed close relationship with *Daucus carota*.

### ARTICLE HISTORY

Received 8 January 2020  
Accepted 19 January 2020

### KEYWORDS

*Cuminum cyminum*;  
complete chloroplast  
genome; phylogeny



*Cuminum cyminum* L., an economically important plant belonging to the family Apiaceae, is indigenous to SW Asia and the Mediterranean region, but is today widely cultivated. Its aromatic fruits (cumin) are popular for its flavor, and has been used in Chinese and Ayurvedic medicine for the treatment of dyspepsia, diarrhea, and jaundice (Sheh et al. 2005; Srivastava et al. 2011). Studies on this species have focused on describing its chemical constituents (Yan et al. 2002; Hashemi et al. 2009; Zhou et al. 2011) and pharmacological activity (Lee 2005; Sultana et al. 2010; Chen et al. 2011), rarely involved in its genomes. Here, the complete chloroplast (cp) genome sequence of *C. cyminum* is characterized, and its phylogenetic relationships with related taxa in Apiaceae are revealed.

The total genomic DNA was extracted from the leaves of *C. cyminum* collected from Kuerle (41°36′5.34″N, 86°03′25.66″E), Xinjiang of China using a Universal Genomic DNA Extraction kit (Tiangen Biotech, Beijing, China) following the manufacturer's protocol. Voucher specimen was deposited in KUN (Kunming Institute of Botany, Chinese Academy of Sciences, ZJ1706). Then, the genome sequencing were performed with Illumina Hiseq 2500 (Majorbio, Shanghai, China) platform with pair-end (2 × 300) library. The raw data were filtered using Trimmomatic with default settings (Bolger et al. 2014). Then paired-end reads of clean data were assembled into circular contigs using GetOrganelle.py (Jin et al. 2018) with *Daucus carota* (No. NC\_008325) as reference. Finally, the plastome was annotated by the Dual Organellar Genome Annotator (DOGMA; <http://dogma.cccb.utexas.edu/>) (Wyman et al. 2004)

and tRNAscan-SE (Lowe and Chan 2016) with manual adjustment using Geneious (Kearse et al. 2012), and the physical map was drawn by OGDRAW (Lohse et al. 2007).

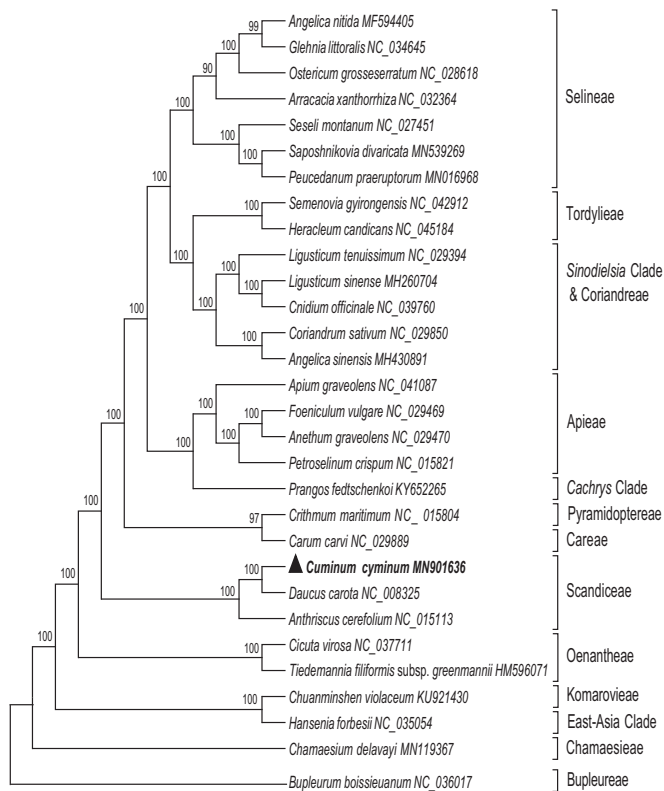
The plastomes of *C. cyminum* (GenBank accession number: MN901636) show the typical properties of Apiaceae and most other eudicot plastid DNAs in its structural organization, gene content, and gene arrangement. The total length of the chloroplast genome was 157,389 bp, with 37.80% overall GC content. With typical quadripartite structure, a pair of IRs (inverted repeats) of 28,157 bp was separated by a small single-copy (SSC) region of 17,598 bp and a large single-copy (LSC) region of 83,927 bp. The cp genome contained 131 genes, including 86 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Among these, 18 genes were duplicated in the inverted repeat regions.

To investigate its phylogenetic placement, a total of 30 cp genome sequences of Apiaceae were downloaded from the NCBI database. All sequences were aligned using the MAFFT (Katoh and Standley 2013) webserver (<http://mafft.cbrc.jp/alignment/server/>), a maximum likelihood (ML) analysis was constructed using RAxML (Stamatakis 2014) with 1000 bootstrap replicates, and *Bupleurum boissieuianum* used as outgroup (NC\_036017). The results showed that *C. cyminum* belongs to the tribe Scandiceae, and it was closely related to *Daucus carota* (Figure 1). Meanwhile, the phylogenetic relationship in Apiaceae was consistent with previous studies, and this will be beneficial to construct a more reasonable phylogeny of Apiaceae.

**CONTACT** Zhenwen Liu  [liuzw@mail.kib.ac.cn](mailto:liuzw@mail.kib.ac.cn)  Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, 650201, China

© 2020 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.** Maximum likelihood (ML) tree of 30 species in the family Apiaceae based on the complete chloroplast sequences using *Bupleurum boissieuianum* (NC\_036017) as an outgroup. Numbers on the nodes are bootstrap values from 1000 replicates.

## Disclosure statement

The authors declare no conflicts of interest and are responsible for the content of the article.

## Funding

This work was supported by the National Natural Science Foundation of China [No. 31460052 and 31960048] and the Hundred-Talent Program of Kunming Medical University [60118260127].

## References

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–2120.

- Chen QQ, Gan ZL, Dai YQ, Yang Y, Ni YY. 2011. Antioxidant activity and kinetics analysis of cumin oleoresin and its main components in vitro. *Sci Technol Food Industry*. 32:111–113.
- Hashemi P, Shamzadeh M, Badieli A, Ghiasvand A R, Azizi K. 2009. Study of the essential oil composition of cumin seeds by an amino ethyl-functionalized nanoporous SPME fiber. *Chromatographia*. 70(7–8): 1147–1157.
- Jin JJ, Yu WB, Yang JB, Song Y, Yi TS, Li DZ. 2018. GetOrganelle: a simple and fast pipeline for de novo assembly of a complete circular chloroplast genome using genome skimming data. *bioRxiv*. 1–11.
- Katoh K, Standley D. 2013. MAFFT multiple sequence alignment software-version improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- 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.
- Lohse M, Drechsel O, Bock R. 2007. Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet*. 52(5–6):267–274.
- Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res*. 44(W1): W54–W57.
- Lee HS. 2005. Cuminaldehyde: aldose reductase and  $\alpha$ -glucosidase inhibitor derived from *Cuminum cyminum* L. seeds. *J Agric Food Chem*. 53(7):2446–2450.
- Sheh ML, Pu FD, Pan ZH, Watson MF, Cannon JFM, Holmes-Smith I, Kljuykov EV, Phillippe LR, Pimenov MG. 2005. Apiaceae. In: *Flora of China* Editorial Committee, editor. *Flora of China*, vol. 14. St. Louis MI: Missouri Botanical Garden Press; p. 1–205.
- Srivastava R, Srivastava SP, Prakash S, Jaiswal N, Mishra A, Maurya R, Srivastava AK. 2011. Antidiabetic and antidiyslipidemic activities of *Cuminum cyminum* L. in validated animal models. *Med Chem Res*. 20(9):1656–1666.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.
- Sultana S, Ripa FA, Hamid K. 2010. Comparative antioxidant activity study of some commonly used spices in Bangladesh. *Pak J Biol Sci*. 13(7): 340–343.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20(17):3252–3255.
- Yan JH, Tang KW, Zhong M, Deng NH. 2002. Determination of chemical components of volatile oil from *Cuminum cyminum* L. by gas chromatography-mass spectrometry. *Se Pu*. 20(6):569–572.
- Zhou JW, Huang JX, Huang L, Liu L. 2011. Column chromatography separation and GC-MS determination of essential oil components of *Cuminum cyminum* L. *J Anhui Agri Sci*. 39:9840–9841.