

## Characterization and analysis of complete chloroplast genome of *Clausena anisum-olens* (Blanco) Merrill

Jin-jin Li<sup>a</sup>, Chang Liu<sup>b</sup>, Dan Wang<sup>a</sup>, Dong-qin Guo<sup>a</sup> and Hui-hui Du<sup>a</sup>

<sup>a</sup>College of Biology and Food Engineering, Chongqing Three Gorges University, Chongqing, China; <sup>b</sup>Nanjing Institute for Comprehensive Utilization of Wild Plants, Nanjing, P. R. China

### ABSTRACT

In this study, the chloroplast genome (Cp) of *Clausena anisum-olens* (Blanco) Merrill was sequenced by high-throughput sequencing technology. We found that the length of the chloroplast genome of *Clausena anisum-olens* was 159,753 bp, and the total GC content was 38.7%, including a large single-copy (LSC) region of 87,623 bp, a small single-copy (SSC) region of 17,986 bp, and 27,072 bp pairs of reverse repeats (IRS). The Cp genome encoded 129 genes, containing 86 protein-coding, 37 tRNA, and 6 rRNA genes. Phylogenetic analysis of the genome sequence showed that *Clausena anisum-olens* was closely related to *Micromelum minutum* and *Glycosmis mauritiana*.

### ARTICLE HISTORY

Received 30 August 2021  
Accepted 15 October 2021

### KEYWORDS

*Clausena anisum-olens* (Blanco) Merrill; chloroplast genome; sequence; gene; phylogenetic analysis

*Clausena anisum-olens* (Blanco) Merrill is an important *Clausena* genus in Rutaceae. The wild species are distributed in Taiwan of China and are cultivated in Guangdong, Guangxi, and Yunnan of China (Su et al. 2011), *Clausena anisum-olens* belongs to the same genus *Clausena* but has more wild resources than *Clausena lansium* (Su et al. 2011). *Clausena anisum-olens* fruit exerts phlegm and cough relief. The leaves and branches are used as a herbal medicine to relieve rheumatic pains (Wu 1995). Most of the studies on this species have been focused on their pharmacological activities and chemical composition analysis, but few involved genome analysis (Su and Liang 2010).

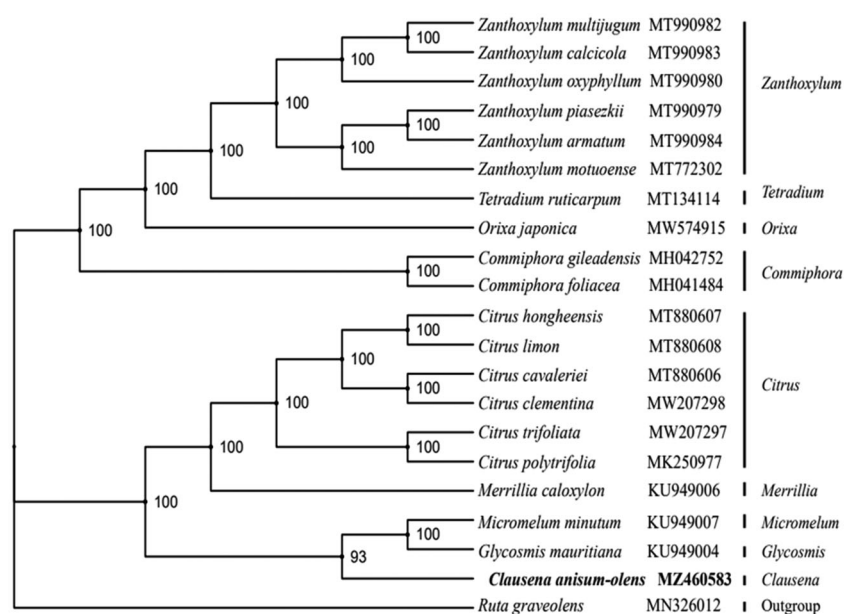
The chlorophyll in the leaves of *Clausena anisum-olens* has attracted considerable attention because of its beneficial pharmacological activities, such as abundant content of fat-soluble natural pigments, which have antiseptic, deodorization, antipyretic, and hemostatic effects. Herein, we report the identification and characterization of the complete chloroplast genome of *Clausena anisum-olens*, which provides valuable information on a large number of sequences. Our present findings will facilitate further investigations of the genome and the phylogenetic relationships of the Rutaceae family.

Fresh and healthy wampee leaves (104°65'68.07"E, 23°12'99.61"N) were collected from Wenshan, of Yunnan Province. The certified herbarium specimen (No. ZDQ17012) was collected and stored by Kunming Zhifeng Biotechnology Co., Ltd. Total DNA was isolated from dried leaf materials according to a previously reported modified CTAB method

(Yang et al. 2014) and sequenced by second-generation sequencing using Illumina HiSeq 2500 platform (Novogene, Tianjin, China). To decrease the redundant data, the original reads were filtered by Trimmomatic v.0.32 software with default parameters (Bolger et al. 2014). Then, the obtained clean reads were assembled into circular contigs using GetOrganelle (Jin et al. 2020) with *Clausena excavata* (No. NC\_032685) as the reference. Finally, the cpDNA was annotated by the Dual Organellar GenoMe Annotator GeSeq (Tillich et al. 2017) and CpGAVAS2 (Shi et al. 2019). Finally, the annotated chloroplast genome was submitted to the GenBank (accession number: MZ460583).

The whole genome of *C. anisum-olens* has a typical tetragonal structure and a length of 159,753 bp. It consists of 87,623 bp large single-copy (LSC) region, 17,986 bp small single-copy (SSC) region and 27,072 bp reverse repeat regions (IRA and IRB). The total GC content was 37.8%. The CP genome was found to be composed of 129 genes, including 86 protein-coding, 37 tRNA, and 6 rRNA genes.

To determine the phylogenetic position of *C. anisum-olens*, 20 complete chloroplast genome sequences were downloaded from the NCBI database. The genome sequences were aligned with MAFFT version 7.427 (Katoh and Standley 2013), and then the maximum-likelihood (ML) tree was constructed by the RAxML (Stamatakis 2014) program with 1000 bootstrap replicates and the GTRGAMMAI model. Our results showed that *C. anisum-olens* was closely related to *Micromelum* and *Glycosmis* in the family of Rutaceae (Figure 1).



**Figure 1.** Phylogenetic tree reconstructed by maximum-likelihood (ML) analysis based on the complete chloroplast genome sequences of 21 species with *Ruta graveolens* as an outgroup. Bootstrap support values (1000 replicates) are displayed next to the nodes.

## Acknowledgments

We express our gratitude to Kunming Zhifeng Biotechnology Co., Ltd for the provision of the plant material for molecular experiments.

## Disclosure statement

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

## Funding

This work was supported by the Chongqing Science and Technology Commission [cstc2018jcyjAX0639] and the Scientific and Technological Research Program of Chongqing Municipal Education Commission [KJCX2020046; KJQN201901222].

## Data availability statement

The genome sequence data that support the findings of this study are openly available in the GenBank of NCBI at (<https://www.ncbi.nlm.nih.gov/>) under the accession no. MZ460583. The associated BioProject, SRA, and BioSample numbers are PRJNA758156, SRR15647457, and SAMN21013949, respectively.

## References

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

- Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CM, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 21(1):241.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- 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.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9):1312–1313.
- Su XF, Liang ZY. 2010. Chemical components of volatile oil from *Clausena Anisum-olens* by GC-MS. *Food Res Develop*. 31(12):176–178.
- Su XF, Lu ZH, Chen CQ. 2011. Optimization of chlorophyll extraction process in the leaves of *Clausena anisum-olens* by orthogonal design. *Med Plant*. 2(01):52–53 + 56.
- Su XF, Wei LX, Huang LJ, Feng PZ. 2011. Antimicrobial activity of ethanol extracts and volatile oil from the bark of *Clausena anisum-olens*. *Hubei Agric Sci*. 50(13):2668–2670.
- 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.
- Wu, Z (Ed.), 1995. *Institutum Botanicum Kunmingense Academiae Sinicae Edita. Flora Yunnanica*. Vol. 6. Beijing: Science Press; p. 767.
- Yang JB, Li DZ, Li HT. 2014. Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. *Mol Ecol Resour*. 14(5):1024–1031.