

The complete chloroplast genome sequences and phylogenetic analysis of *Crotalaria pallida* (Leguminosae)

Nong Zhou^{a,b*}, Hai-Ling Li^a, You Zhou^{a,b} and Dong-Qin Guo^{a,b*}

^aCollege of Biology and Food Engineering, Chongqing Three Gorges University, Chongqing, PR China; ^bEngineering Laboratory of Chongqing for Green Planting and Deep Processing of Genuine Medicinal Materials in Three Gorges Reservoir Area, Chongqing Three Gorges University, Chongqing, PR China

ABSTRACT

The complete chloroplast genome of *Crotalaria pallida* was obtained using the high-throughput sequencing technology in this article. The complete chloroplast genomes of this species were 152,658 bp in length, consisting of a large single-copy region (LSC) of 83,652 bp and a small single-copy region (SSC) of 18,028 bp, which were separated by a pair of inverted repeat (IRs) regions of 25,489 bp. The chloroplast genome contained 111 unique genes, including 77 protein-coding genes, 30 *tRNA* genes, and four *rRNA* genes. The phylogenomic relationship analysis suggested that *C. pallida* was closely related to *Lupinus* in the family of Leguminosae.

ARTICLE HISTORY

Received 20 October 2020
Accepted 24 November 2020

KEYWORDS

Crotalaria pallida; complete chloroplast genome; Illumina sequencing; phylogeny

Crotalaria pallida, also named 'Zhu-Shi-Dou', is a species of genus *Crotalaria*, family Leguminosae, which is widely distributed in the tropical and subtropical regions of China, India as well as Africa and America countries (Editorial Committee of Flora of China 1998). *Crotalaria pallida* is used as folk medicine to treat scrofula, mastitis, and dysentery with stems and leaves, and to treat neurasthenia, dizziness, leucorrhea, and cancers with seeds in China (Lian 1986; Hu et al. 2017). Researches about bioactive constituents in *C. pallida* have been carried out, such as flavonoids, alkaloids, unexplored proteins, and anti-tyrosinase compounds (Weng et al. 2003; Ko et al. 2004; Ukil et al. 2017; Cheng et al. 2020). In the previous study, the analysis based on its complete chloroplast genome is still lack. Here, to provide a solid foundation for further phylogenetic studies, we assembled and characterized the complete plastome of *C. pallida*.

In this study, the sample of *C. pallida* was collected from Dali, Yunnan province, China (100°15'42.75"E, 25°31'33.38"N). The voucher specimen (NO. ZSY109) was preserved at the Herbarium of Medicinal Plants and Crude Drugs of the College of Pharmacy, Dali University. We used the modified CTAB method (Doyle 1987; Yang et al. 2014) to extract the total genomic DNA from the dry and healthy leaves. A pair-end (2 × 300 bp) library was constructed and sequenced via the Illumina Hiseq 2500 (Novogene, Tianjing, China) platform. After the sequencing, the raw data were filtered via Trimmomatic version 0.32 with default settings (Bolger et al. 2014). Subsequently, the filtered sequences were assembled into circular contigs utilizing GetOrganelle (Jin et al. 2020) with the cp genome of closely related species *Lupinus*

atlanticus (KU726827) as the reference. Finally, the assembled genome was annotated and corrected manually using Geneious R version 11.0.2 (Kearse et al. 2012). The annotated complete chloroplast genome of *C. pallida* was submitted to GenBank under the accession number of MT920364.

The complete chloroplast genomes of *C. pallida* were 152,658 bp in length. The typical quadripartite structure consisted of a pair of inverted repeat regions (IRa and IRb, 25,489 bp), a large single-copy region (LSC, 83,652 bp), and a small single-copy region (SSC, 18,028 bp). The total GC content was 36.7%, while those of IR regions (42.9%) were higher than LSC (34.3%) and SSC (30.6%) regions. A total of 111 genes was successfully annotated, containing 77 coding protein genes, 30 *tRNA* genes, and four *rRNA* genes.

To further ascertain the phylogenetic position of *C. pallida*, a phylogenetic analysis was carried out among 23 complete chloroplast genomes in Leguminosae with *Eriobotrya malipoensis* (Rosaceae) as an outgroup. The genome sequences were aligned with MAFFT version 7.427 (Katoh and Standley 2013) and then the maximum-likelihood (ML) tree was conducted by RAxML (Stamatakis 2014) program with 1000 bootstrap replicates and the GTRGAMMAI model. The result showed that *C. pallida* was closely related to *Lupinus* in the family of Leguminosae (Figure 1).

Acknowledgments

We thank Dequan Zhang in Dali University for affording plant material of molecular experiments.

CONTACT Dong-Qin Guo  guodongqin1997@163.com  College of Biology and Food Engineering, Engineering Laboratory of Chongqing for Green Planting and Deep Processing of Genuine Medicinal Materials in Three Gorges Reservoir Area, Chongqing Three Gorges University, Chongqing, PR China

*These two authors had equal contribution to the manuscript.

© 2021 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.

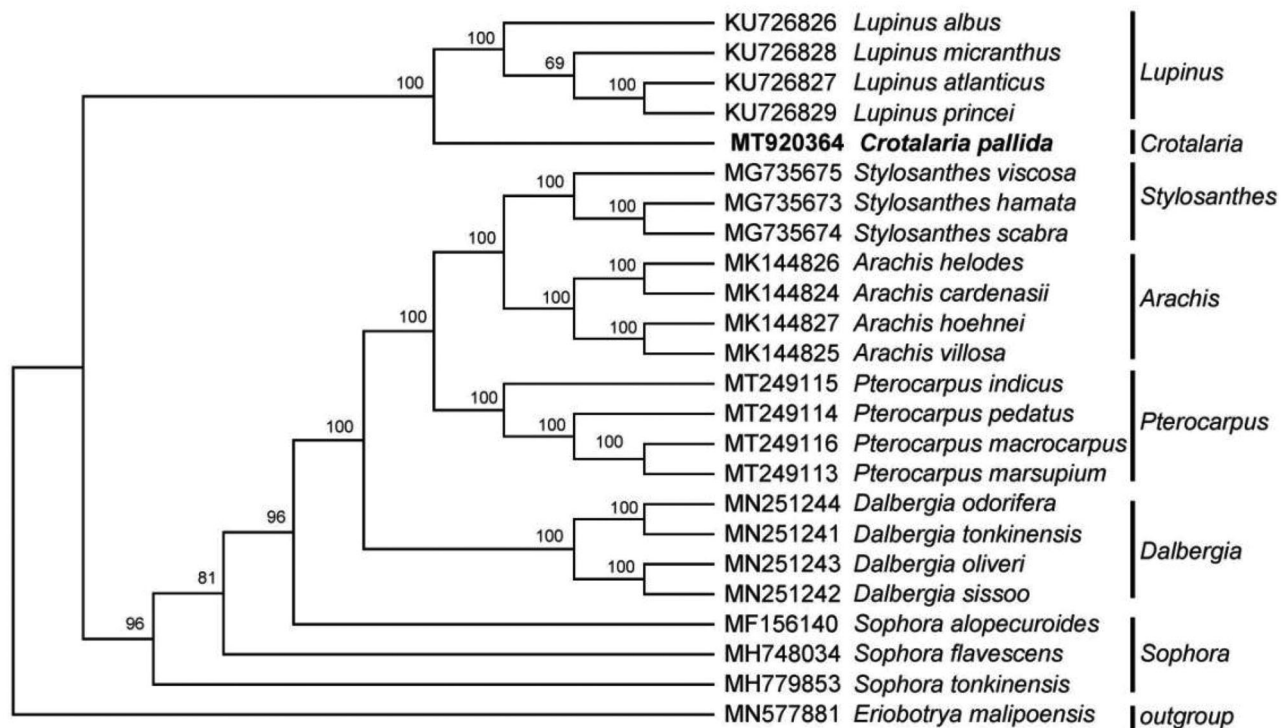


Figure 1. ML phylogenetic tree based on the complete chloroplast genome sequences of 24 species with *Eriobotrya malipoensis* as an outgroup. Bootstrap support values (1000 replicates) are shown next to the nodes.

Disclosure statement

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

Funding

This study was supported by the science and technology research program of Chongqing Education Commission of China [grant number KJQN202001242 and KJQN201801210].

Data availability statement

The data supporting the finding of this study is available in GenBank. The accession number is MT920364: <https://www.ncbi.nlm.nih.gov/nucleotide/MT920364>.

References

- ## Disclosure statement
- No potential conflict of interest was reported by the author(s).
- ## Funding
- This study was supported by the science and technology research program of Chongqing Education Commission of China [grant number KJQN202001242 and KJQN201801210].
- ## Data availability statement
- The data supporting the finding of this study is available in GenBank. The accession number is MT920364: <https://www.ncbi.nlm.nih.gov/nucleotide/MT920364>.
- ## References
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–2120.
- Cheng ZY, Sun Q, Yang PY, Huang XX, Song SJ. 2020. Isolation and structure elucidation of anti-tyrosinase compounds from the seeds of *Crotalaria pallida*. *J Asian Nat Prod Res*. 6:1–7.
- Doyle J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19:11–15.
- Editorial Committee of Flora of China. 1998. *Flora of China*. Vol. 42, Beijing, China: Science Press; p. 349.
- Hu XR, Chou GX, Zhang CG. 2017. Flavonoids and alkaloids from the seeds of *Crotalaria pallida* and their cytotoxicity and anti-inflammatory activities. *Phytochemistry*. 143:64–71.
- 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 D. 2013. MAFFT multiple sequence alignment software version 7: 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, Mark-Owitz 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.
- Ko H, Weng JR, Tsao LT, Yen MH, Wang JP, Lin CA. 2004. Anti-inflammatory flavonoids and pterocarpanoid from *Crotalaria pallida* and *C. assamica*. *Bioorg Med Chem Lett*. 14(4):1011–1014.
- Lian WY. 1986. A survey on the medicinal plants of the genus *Crotalaria*. *J Chin Med Mat*. 9:19–22.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9):1312–1313.
- Ukil S, Laskar S, Bandyopadhyay D. 2017. Isolation, purification and partial characterization of *Crotalaria pallida* Aiton seed proteins. *Int J Pept Res Ther*. 23(4):461–467.
- Weng JR, Tsao LT, Yen MH, Wang JP, Lin CA. 2003. Anti-inflammatory constituents and new pterocarpanoid of *Crotalaria pallida*. *J Nat Prod*. 66(3):404–407.
- 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.