

The complete chloroplast genome sequence of *Zanthoxylum undulatifolium* Hemsl. (Rutaceae)

Chong Sun^{a,b}, Xia Liu^b, Houlin Zhou^c, Jing Liu^a, Xiaoying Li^a, Hailang Liu^a and Can He^a

^aCollege of Horticulture and Gardening, Yangtze University, Hubei, China; ^bCollege of Landscape Architecture and Life Science, Chongqing University of Arts and Sciences, Chongqing, China; ^cChongqing Wulipo National Nature Reserve Management Office, Chongqing, China

ABSTRACT

Zanthoxylum undulatifolium is an excellent economic tree species with important medical value. This study reports the first complete chloroplast genome sequence of *Z. undulatifolium*. Its whole chloroplast genome is 158,400 bp in length, including a large single-copy (LSC) region of 85,898 bp, a small single-copy (SSC) region of 17,610 bp, and two inverted repeat (IR) regions of 27,446 bp. The chloroplast genome contains a total of 132 genes, comprising 87 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The overall GC content of the chloroplast genome is 38.46%, with the corresponding values in the LSC, SSC, and IR regions are 36.87%, 33.51%, and 42.55%, respectively. Phylogenetic analysis revealed the sister relationship between *Z. undulatifolium* and *Z. bungeanum*.

ARTICLE HISTORY

Received 13 December 2021
Accepted 2 February 2022

KEYWORDS

Zanthoxylum undulatifolium;
complete chloroplast
genome; phylogenetic
analysis

In 1895, W. Botting Hemsley, F. R. S. first published a description of *Zanthoxylum undulatifolium* Hemsl. as a new species (Hemsley 1895). *Z. undulatifolium* is a member of the genus *Zanthoxylum* L. in the family Rutaceae. It is a rare medicinal plant that narrowly distributed in the 1,600–2,300 m mountain forests or vegetation thicket areas of Southwest China (Editing Committee of Chinese Flora 1993). *Z. undulatifolium* is an excellent ecological tree species for soil and water conservation (Editing Committee of Chinese Flora 1993). No genomic information of *Z. undulatifolium* has been reported thus far. In this report, we present the first complete chloroplast genome sequence of *Z. undulatifolium* and construct its phylogenetic relationships with related species.

Samples of *Z. undulatifolium* was collected from Wushan County Goddess Peak, Chongqing, China (31.0461°N, 110.0281°E), and a voucher specimen was deposited at the Chongqing University of Arts and Sciences Herbarium (LYWS) under accession number CUAS-LY20180518 (Xia Liu, liuxiavip8@163.com). The genomic DNA was extracted from silica-dried leaf tissue using a modified CTAB method (Doyle and Doyle 1987). The DNA library was sequenced by Hefei Bio&Data Biotechnologies Inc. (Hefei, China) on the BGISEQ-500 platform with PE150 read lengths. The clean reads were used for the de novo assembly of the chloroplast genome using the SPAdes Assembler v3.9.0 (Bankevich et al. 2012). The annotation of the complete genome was performed using CpGAVAS (Liu et al. 2012) and GeSeq software (Michael et al. 2017). After a manual check and adjustment, the annotated chloroplast genome sequence of *Z. undulatifolium* was submitted to GenBank (MZ676708).

The chloroplast genome of *Z. undulatifolium* exhibited a typical angiosperm circular structure with a length of

158,400 bp and consisted of a large single-copy region (LSC: 85,898 bp), a small single-copy region (SSC: 17,610 bp), and two inverted repeat regions (IRs: 27,446 bp). The overall GC content of *Z. undulatifolium* is 38.46% and the values in the LSC, SSC and IR regions are 36.87%, 33.51%, and 42.55%, respectively. The chloroplast genome encodes a total of 132 genes (87 protein-coding, 37 tRNA, and 8 rRNA genes), with 18 duplicated genes (7 protein-coding, 7 tRNA, and 4 rRNA genes). Nineteen genes contain two exons and four protein-coding genes (*ycf3*, *clpP*, and two *rps12*) contain three exons.

We performed a phylogenetic analysis based on the complete chloroplast genomes of 12 species and then constructed a phylogenetic tree to explore the phylogenetic relationships of *Z. undulatifolium* (Figure 1). The 12 complete chloroplast genome sequences were subjected to multiple sequence alignment using MAFFT software (Katoh and Standley 2013). A maximum likelihood (ML) phylogenetic tree was built using the RAxML version 8 program (Alexandros 2014) with 1,000 bootstrap replicates. Phylogenetic analysis showed that *Z. paniculatum* and *Z. madagascariense* at the base of the phylogenetic tree are the oldest species among the selected *Zanthoxylum* species. *Z. undulatifolium* is most closely related to *Z. bungeanum*, and a sister to *Z. sp. NH018* and *Z. simulans*, with 100% bootstrap support values.

A range of previous studies attempted to study the evolutionary relationships of *Zanthoxylum* species with molecular markers such as random amplified polymorphic DNA (Medhi et al. 2014), inter simple sequence repeat (Feng et al. 2015), isozyme (Li et al. 2004), simple sequence repeat (Kim et al. 2017), amplified fragment length polymorphism (Gupta and Mandi 2013), and chloroplast DNA markers (Appelhans et al. 2018). The limited number of polymorphic loci produced by these

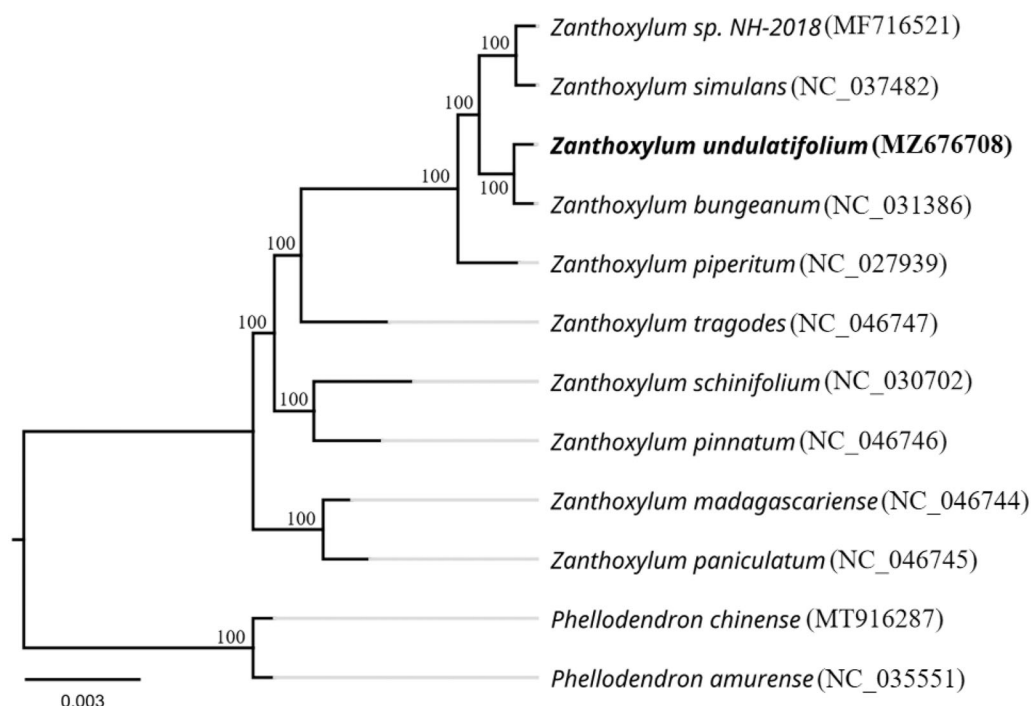


Figure 1. Maximum-likelihood phylogenetic tree of *Z. undulatifolium* and other related species based on the complete chloroplast genome sequences. The number on each node indicates the bootstrap support value.

low-resolution markers hinders phylogenetic research on *Zanthoxylum* species. Therefore, (i) we should obtain more samples of *Zanthoxylum* from China, North America and Japan, and (ii) we should study the evolutionary relationship based on the complete chloroplast genome of these *Zanthoxylum* species to examine the evolution relationships of *Zanthoxylum* species. In this paper, the complete chloroplast genome sequence of a representative *Zanthoxylum* species provides important insights into the evolution of *Zanthoxylum* in eastern Asia.

Acknowledgements

We thank Xia Gong, Yinming Wu and Zheng Chen (Sichuan Academy of Botanical Engineering, China) for their help in sample collections.

Ethical statement

The present study was approved by the authors' institution (the Chongqing University of Arts and Sciences) and national. The research does not involve a threatened/endangered species. All the research meets ethical guidelines and adheres to the legal requirements of the study country. The collection of plant material has been carried out in accordance with the International Union for Conservation of Nature (IUCN) policies research involving species at risk of extinction, the Convention on Biological Diversity and the Convention on the Trade in Endangered Species of Wild Fauna and Flora.

Author contributions statement

Chong Sun, Xia Liu and Jing Liu designed the study, writing and revised the manuscript; Hailang Liu and Can He involved in the process of sequences editing and phylogenetic analyses; Houlin Zhou, Xia Liu and Xiaoying Li participated in the collection and identification of plant material. All authors read and approved the final manuscript, and agreed to be accountable for all aspects of the work.

Disclosure statement

The authors declare that they have no conflict interests.

Funding

This work was supported by the [National Natural Science Foundation of China #1] under Grant [number 31901324]; [Natural Science Foundation of Chongqing of China #2] under Grant [number cstc2019jcyj-msxmX0693]; and [Science and Technology Research Program of Chongqing Municipal Education Commission #3] under Grant [number KJQN202101320].

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at (<https://www.ncbi.nlm.nih.gov/>) under the accession no. MZ676708. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA680256, SRR17163936, and SAMN23766596, respectively.

References

- Alexandros S. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.
- Appelhans MS, Reichelt N, Groppo M, Paetzold C, Wen J. 2018. Phylogeny and biogeography of the pantropical genus *Zanthoxylum* and its closest relatives in the proto-Rutaceae group (Rutaceae). *Mol Phylogenet Evol*. 126:31–44.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 19(5):455–477.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19(1):11–15.
- Editing Committee of Chinese Flora. 1993. 1993.. *Flora of China*. Beijing: Science Press, 43(2):p. 46.

- Feng S, Yang T, Li X, Chen L, Liu Z, Wei A. 2015. Genetic relationships of Chinese prickly ash as revealed by ISSR markers. *Biologia*. 70(1):45–51.
- Gupta DD, Mandi SS. 2013. Species specific AFLP markers for authentication of *Zanthoxylum acanthopodium* & *Zanthoxylum oxyphyllum*. *J Med Plants Stud*. 1(6):1–9.
- Hemsley WB. 1895. Descriptions of some New Plants from Eastern Asia, chiefly from the Island of Formosa, presented by Dr. Augustine Henry, F.L.S., to the Herbarium, Royal Gardens, Kew. *Ann Bot*. 9(1):143–160.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- Kim YM, Jo A, Jeong JH, Kwon YR, Kim HB. 2017. Development and characterization of microsatellite primers for *Zanthoxylum schinifolium* (Rutaceae). *Appl Plant Sci*. 5(7):1600145.
- Li XL, Liang GL, Guo QG, Shang HL. 2004. Isoenzyme analysis of main species of cultivated *Zanthoxylum* L. in China (in Chinese). *J Southwest Agric Univ (Nat Sci)*. 26(4):445–447.
- Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics*. 13(1):715.
- Medhi K, Sarmah DK, Deka M, Bhau BS. 2014. High gene flow and genetic diversity in three economically important *Zanthoxylum* Spp. of upper Brahmaputra valley zone of NE India using molecular markers. *Meta Gene*. 2:706–721.
- Michael T, Pascal L, Tommaso P, Elena SUJ, Axel F, Ralph B, Stephan G. 2017. GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Res*. 45(W1):W6–W11.