

The complete nucleotide molecular sequence of plastid genome of *Zanthoxylum armatum* (Rutaceae)

Wei Sa^a, Jian Liang^a, Qian-Han Shang^a, Mi-Li Liu^b, Ma-Li Wang^b and Zhong-Hu Li^b

^aState Key Laboratory of Plateau Ecology and Agriculture, College of Ecological and Environmental Engineering, Qinghai University, Xining, China; ^bKey Laboratory of Resources Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an, China

ABSTRACT

Zanthoxylum armatum DC. (Rutaceae) is a shrub and/or tree species with the important medicinal and economic values. In this study, the plastid genome of *Z. armatum* was characterized by Illumina Hiseq 2500 sequencing platform. In total, the plastid genome is 158,557 bp in length, and comprises a large single copy region of 85,752 bp, a small single copy region of 17,605 bp, and two inverted repeat regions of 27,600 bp. The complete plastid genome contains 87 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Phylogenetic analysis suggested that *Z. armatum* and the congeneric *Z. simulans* clustered into an evolutionary clade with the high support.

ARTICLE HISTORY

Received 24 July 2020
Accepted 9 September 2020

KEYWORDS

Zanthoxylum armatum;
plastid genome;
phylogenetic analysis



Zanthoxylum armatum DC. (Rutaceae) is widely distributed in the temperate and tropics regions of east Asia. The extract of bark, fruits and seeds of *Z. armatum* is commonly used in the folklore medicine (Kalia et al. 1999). Due to multiple utilities and high cultural value, unsustainable harvest of above ground parts, and low regeneration in nature (Kala 2010), the natural population resources of *Z. armatum* has been rapidly decreased (Samant et al. 2007). It is urgent to protect and manage natural resources of *Z. armatum*. In plant, plastid genome provided valuable phylogenetic information, owing to its conserved genome structures and comparatively high evolutionary rates (Wu and Ge 2012). In this study, we characterized the complete nucleotide sequence of plastid genome of *Z. armatum* (GenBank accession number: MN080708) based on the Illumina next-generation sequencing technology.

The individual of *Z. armatum* was collected from the Qinling Mountains (N: 33.9862; E:108.7231), Shaanxi, China. The dried plant samples and voucher specimens (No. ZALZH2018015) were deposited in the Key Laboratory of Resource Biology and Biotechnology in Western China (Shaanxi, China). The total genomic DNA was isolated from leaf materials using modified CTAB procedure (Doyle and Doyle 1987). Then, the fragmented DNAs were subjected to Illumina samples preparation, and pair-end sequenced using an Illumina Hiseq 2500 platform by Novogene Bioinformatics Technology Co., Ltd (Beijing, China). 1,065,639 reads were obtained and assembled with the program MITObim v1.8 (Hahn et al. 2013) using the congeneric *Zanthoxylum simulans* (NC_037482) as reference. The plastid genome annotation was conducted using the newly developed biosoftware

CPGAVAS2 (<http://47.96.249.172:16019/analyzer/home>, Shi et al. 2019).

The complete plastid genome of *Z. armatum* was 158,557 bp in length and comprises a large single copy (LSC) region with 85,752 bp, a small single copy (SSC) region with 17,605 bp, and two inverted repeat regions (IRs) of 27,600 bp, which formed a typical quadripartite DNA molecular structure. The complete plastid genome contains 132 genes, including 87 protein-coding genes (PCGs), 37 transfer RNA genes (tRNA) and 8 ribosomal RNA genes (rRNA). A total of 16 genes (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl2*, *rpl16*, *rps12*, *rps16*, *rpoC1*, *trnA-UGC*, *trnI-GAU*, *trnG-UCC*, *trnK-UUU*, *trnL-UAA*, *trnV-UAC*) contained one intron, and 2 genes (*ycf3*, *clpP*) contained 2 introns. The overall GC content of *Z. armatum* plastid genome is 38.5%, while the corresponding values of LSC, SSC, and IR regions are 36.8, 33.6, and 42.5%, respectively.

In order to determine phylogenetic position of *Z. armatum*, the whole plastid genome sequences of 16 species were available downloaded from NCBI. All of the 17 plastid sequences were aligned using the software MAFFT (Katoh and Standley 2013) with the default parameter settings. We used the Modeltest v3.7 (Posada and Crandall 1998) to determine the best-fitting model. The phylogenetic analysis was conducted using the program RAxML (Stamatakis 2014) with 1000 bootstrap replicates. The results revealed that *Z. armatum*, *Z. simulans*, *Z. bungeanum*, *Z. schinifolium* and *Z. piperitum* presented a group with high bootstrap support value (100%) (Figure 1).

CONTACT Zhong-Hu Li  lizhonghu@nwu.edu.cn  Key Laboratory of Resources Biology and Biotechnology in Western China, Ministry of Education, College of Life Sciences, Northwest University, Xi'an, 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.

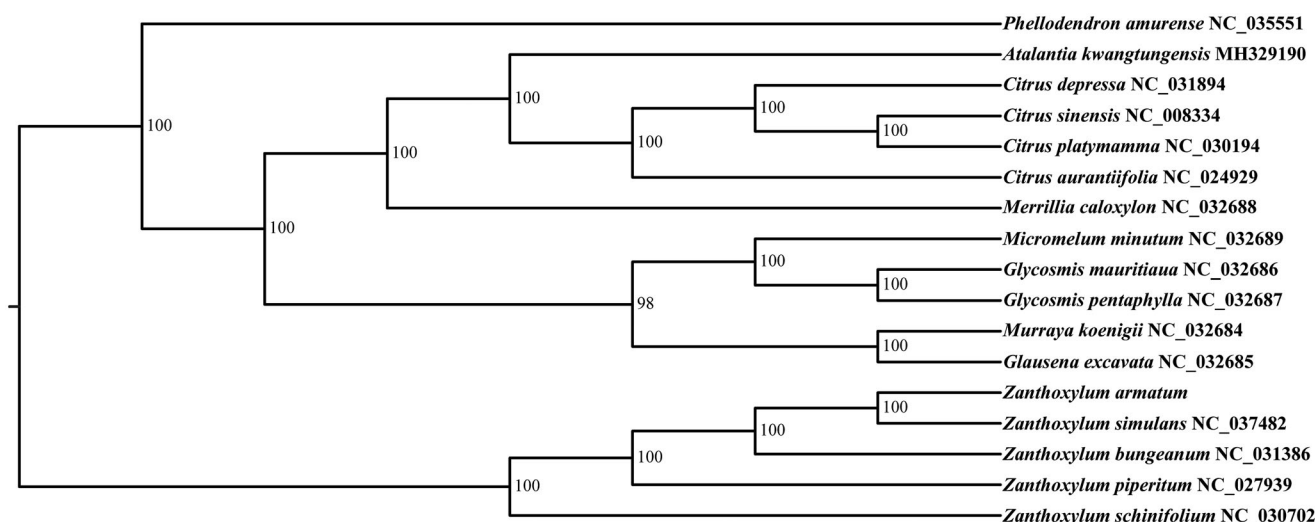


Figure 1. Phylogenetic tree based on seventeen nucleotide sequences of whole plastid genomes.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was co-supported by the Youth program of Natural Science Foundation of Science and Technology Department of Qinghai Province [2018-ZJ-963Q], the Key R & D and transformation projects of science and Technology Department of Qinghai Province [2017-NK-151], the Fourth National Survey of Traditional Chinese Medicine Resources (Grants no. 2017-66, 2018-43 and 2019-39) and the Open Foundation of Key Laboratory of Resource Biology and Biotechnology in Western China (Ministry of Education) [Grants no. ZSK2017007 and ZSK2019008].

Data availability statement

The data that support the results of this research are openly available in GenBank at <https://www.ncbi.nlm.nih.gov>, reference number MN080708.

References

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19:11–15.

Hahn C, Bachmann L, Chevreur B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucl Acids Res.* 41(13): e129–e129.

Kala CP. 2010. Assessment of availability and patterns in collection of Timroo (*Zanthoxylum armatum* DC.): a case study of Uttarakhand Himalaya. *Med Plnts Int Jnl Phyt Rela Ind.* 2(2):91–96.

Kalia NK, Singh B, Sood RP. 1999. A new amide from *Zanthoxylum armatum*. *J Nat Prod.* 62(2):311–312.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.

Posada D, Crandall KA. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics.* 14(9):817–818.

Samant SS, Butola JS, Sharma A. 2007. Assessment of conservation diversity, a status and preparation distribution of management plan for medicinal plants in the catchment area of Parbati hydroelectric project stage-III in northwestern Himalaya. *J Mt Sci.* 4:34–56.

Shi LC, Chen HM, Jiang M, Wang LQ, Wu X, Huang LF, 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.

Wu ZQ, Ge S. 2012. The phylogeny of the BEP clade in grasses revisited: evidence from the whole-genome sequences of chloroplasts. *Mol Phylogenet Evol.* 62(1):573–578.