Complete chloroplast genome sequence and phylogenetic analysis of *Magnolia pilocarpa*, a highly ornamental species endemic in central China

Jiaojun Yu,a,b Hongjin Dong,a,b Jun Xiang,a,b Yunli Xiao,a,b and Yuanping Fang,a,b

aHubei Key Laboratory of Economic Forest Germplasm Improvement and Resources Comprehensive Utilization, Huanggang Normal University, Huanggang, China; bHubei Collaborative Innovation Center for the Characteristic Resources Exploitation of Dabie Mountains, Huanggang, China

**ABSTRACT**

*Magnolia pilocarpa* Z. Z. Zhao et Z. W. Xie is a species with high horticultural and medicinal value, which found only in Dabie Mountain of central China. In this study, we sequenced, assembled and annotated the complete chloroplast (*cp*) genome of *M. pilocarpa*. The length of *M. pilocarpa* complete *cp* genome was 160,104 bp, with a typical quadripartite structure comprising a pair of inverted repeat regions (*IRs*; 26,591 bp), a large single copy region (*LSC*; 88,110 bp) and a small single copy region (*SSC*; 18,812 bp). The whole *cp* genome contained 129 unique genes, including 79 protein-coding genes (*PCGs*), 30 transfer RNA (*tRNA*) genes and four ribosomal RNA (*rRNA*) genes. The maximum-likelihood (ML) phylogenetic analysis indicated that *M. pilocarpa* was relatively closed to the *M. denudata*. This chloroplast genome would provide a valuable genetic resource for *M. pilocarpa*.

*Magnolia pilocarpa* Z. Z. Zhao et Z. W. Xie is an endemic species of Magnoliaceae distributed in Dabie Mountains in central China, including Anhui, Henan and Hubei Province. It lives in the sunny slopes from 500 m to 1500 m in altitudes, with obvious sepal in March to April. It was grown as ornamental plant in the adjacent area and the buds were used as the substitutes of the traditional Chinese medicine ‘Xinyi’. Based on the flowering before leaf buds, deciduous vs evergreen (*Magnolia*) and other characteristics, it is separated from *Magnolia as Yulania* (Xia et al. 2008).

Genomic DNA was extracted from fresh leaves of *M. pilocarpa* from Wujiashan, Yingshan, Hubei, China (115°48’17.53”E, 31°12’00.92”N, 1450 m; Dong et al. -HGNU-0256, 2019-04-13; HTGC), the total genomic DNA was isolated according to a modified CTAB method (Doyle 1987). Total genome DNA of *M. pilocarpa* was sequenced by Illumina Hiseq 2500 Sequencing System (Illumina, Hayward, CA) to construct the shotgun library. About 10 Gb pair-end (150 bp) raw short sequence data were obtained and the NOVOPlasty software (Dierckxsens et al. 2017) was used to extract and assemble *cp* genome. The low-quality sequences were filtered out Using CLC Genomics Workbench v8.0 (CLC Bio, Aarhus, Denmark). The complete chloroplast genome of *M. pilocarpa* was annotated using Geneious ver. 10.1 (http://www.geneious.com (Matthew et al. 2012)) and online program Chloroplast Genome Annotation, Visualization, Analysis, and GenBank Submission (CPGAVAS) (Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China) (Liu et al. 2012) and then submitted to GenBank (accession no. MN614308). Finally, a physical map of the genome was drawn by using the online program Organelle Genome DRAW (OGDRAW) (Marc et al. 2013).

The size of chloroplast genome of *M. pilocarpa* is 160,104 bp, which exhibited a typical quadripartite structure including a large single-copy (*LSC*) region of 88,110 bp and a small single-copy (*SSC*) region of 18,812 bp separated by a pair identical inverted repeat regions (*IRs*) of 26,591 bp each. A total of 113 genes were successfully annotated containing 79 protein-coding genes, 30 *tRNA* genes and four *rRNA* genes.

To determine the phylogenetic location of *M. pilocarpa* with respect to the other 44 magnoliids and basal angiosperm species with fully sequenced chloroplast genomes, the complete plastome of *M. pilocarpa* was used to reconstruct the phylogenetic relationships. The sequences were initially aligned using MAFFT (Kazutaka and Standley 2013) and then visualized and manually adjusted using BioEdit (Hall 1999). IQ-TREE was used to evaluate and select the best-fitting models of nucleotide sequences (Lam-Tung et al. 2015). Take the plastome of *Amborella trichopoda* (Amborellaceae, Amborellales; AJ506156) as an out-group, a maximum-likelihood analysis was performed with RAxML version 8 program.
(Alexandros 2014) using 1000 bootstrap. The phylogenetic tree in Figure 1 revealed that *M. pilocarpa* was most closely related to *M. denudate*. All 35 taxa within the Magnoliaceae family were grouped into a clade and clustered into two subclades, one of which was formed by Illiciaceae, and the other subclade was formed by Magnoliaceae. The topological structure of which was consistent with the previously-published phylogeny (Gao et al. 2018; Cui et al. 2019; Li et al.

Figure 1. Maximum-likelihood phylogenetic tree for *Magnolia pilocarpa* based on 45 complete chloroplast genomes. The number on each node indicates bootstrap support value.
The results will be valuable for the genetic study on genus Magnolia.

Disclosure statement
No potential conflict of interest was reported by the authors.

Funding
This study was financially supported by the Startup Foundation for Doctors of Huanggang Normal University [2042019020 to JJY and 2042019013 to HJD], the High-level Training Program of Huanggang Normal University [204201913203 to JJY and 204201913303 to HJD], and Foundation for Assessment and Comprehensive Utilization of Characteristic Biological resources in Dabie Mountains [4022019006].

References
Alexandros S. 2014. Raxml version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.
Cui Y, Chong L, Yue Z, Yuan L, Wang Y, Dan L, Du FK. 2019. The complete chloroplast genome of siebold’s magnolia: Magnolia sieboldii (magnoliaceae), a highly ornamental species with attractive aromatic flowers. Conserv Genet Resour. 11(3):299–301.
Dierckxsens N, Mardulyn P, Smits G. 2017. Novoplasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.
Doyle JJ. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.
Gao J, Song Y, Zheng B. 2018. Complete chloroplast genome sequence of an endangered tree species, magnolia sieboldii (magnoliaceae). Mitochondrial DNA. 32(2):1261–1262.
Hall TA. 1999. Bioedit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/nt. Nucleic Acids Symposium Series (41):95–98.
Kazutaka K, Standley DM. 2013. Mafft multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
Lam-Tung N, Schmidt HA, Arndt VH, Minh BQ. 2015. Iq-tree: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.
Li Y, Sylvester S P, Li M, Zhang C, Li X, Duan Y, Wang X. 2019. The complete plastid genome of magnolia zenii and genetic comparison to magnoliaceae species. Molecules. 24(2):261.
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):715–715.
Marc L, Oliver D, Sabine K, Ralph B. 2013. Organellargenomedraw—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. 41:575–581.
Matthew K, Richard M, Amy W, Steven SH, Matthew C, Shane S, Simon B, Alex C, Sidney M, Chris D. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12):1647.
Xia NH, Liu YH, Nooteboom HP. 2008. Magnoliaceae. In: Wu Z, Raven PH, editors. Flora of China. Beijing: Science Press and St. Louis: Missouri Botanical Garden Press; p. 48–91.