Characterization of the complete chloroplast genome sequence of *Maesa hupehensis* Rehd (primulaceae)

Yan Li, Xuan Yao, Qinxiang Chang, Caijuan Zhang and Pengguo Xia

Department of Art Design, Taiyuan University, Taiyuan, China; Key Laboratory of Plant Secondary Metabolism and Regulation of Zhejiang Province, College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou, China

**ABSTRACT**

*Maesa hupehensis* Rehd 1916 is mainly planted in Hubei and Sichuan. In this study, we assembled and characterized the complete chloroplast genome of *M. hupehensis* as resources for the future study. The chloroplast genome was 157,005 bp in length, with 37.3% GC content, composing of one large single copy (87,628 bp) and one small single copy (18,111 bp), separated by two inverted repeats (25,633 bp). A total of 130 genes were predicted, including 8 rRNAs, 37 tRNAs, and 84 protein-coding genes. Phylogenetic analysis showed that *M. hupehensis* was closely related to *Maesa montana*.

The genus *Maesa* Forsk (Primulaceae) comprises ~200 species mainly distributed in the tropics of the eastern hemisphere. There are 29 species and one variety in China, which are distributed in the south of the Yangtze River Basin (Lai et al. 2002). In previous classification of this genus, it was placed under the family Myrsinaceae. However, it belongs to the family Primulaceae based on gross analysis of morphological and gene data (APG III 2009; Wang and Xia 2011). *Maesa hupehensis* belongs to the genus *Maesa*, is a native plant of the Sichuan, Guangdong provinces in China. At present, the complete chloroplast information and systematic position of *M. hupehensis* has not been studied and reported. In this study, we sequenced, assembled and annotated the complete chloroplast genome of *M. hupehensis*, which will contribute to the further research of *Maesa* Forsk.

The materials of *M. hupehensis* came from Zhongba Canyon, Shiquan County, Ankang City, Shaanxi Province (32°54′56.11″N, 108°7′18.57″E, and altitude 527 m). A specimen was deposited at the Herbarium of Xi’an Botanical Garden (voucher number: Xun Lulu et al. 01495, Lulu Xun, xunlulu20032006@126.com). The total genomic DNA of leaves was extracted by CTAB method (Doyle and Doyle 1987), sequenced using Illumina NovaSeq 6000 platform and filtered using Fastp software (Chen et al. 2018). The extracted DNA was deposited at Key Laboratory of Plant Secondary Metabolism and Regulation of Zhejiang Province, Zhejiang Sci-Tech University (http://sky.zstu.edu.cn) under the voucher number ZSTUX0101 (collected by Pengguo Xia and xpg_xpg@zstu.edu.cn). Resultant clean reads were assembled using GetOrganelle v1.7.0 (Jin et al. 2020). Geneious prime was used for gene annotation by having *Maesa montana* (KU569490) as the reference genome. Genomic data were deposited into NCBI GenBank with the accession number M2846203.

The complete chloroplast genome of *M. hupehensis* was 157,005 bp in length containing two inverted repeat regions of 25,633 bp, a large single copy region of 87,628 bp and the small single copy region of 18,111 bp. The GC content was 37.3%. The complete chloroplast genome encoded 130 genes, including 84 protein coding genes, 37 tRNA genes, and 8 rRNA genes.

In order to determine the phylogenetic position of *M. hupehensis* in Primulaceae, the complete chloroplast genome sequences of 13 species from Primulaceae and two species of Ericaceae as complex outgroup were downloaded from GenBank, and all sequences were aligned using MAFFT (Katoh and Standley 2013). The best nucleotide substitution model was the Transversion model with empirical base frequencies and FreeRate model (TVM + F + R2), revealed by ModelFinder (Kalyaanamoorthy et al. 2017). The IQ-TREE v2.1.2 (Minh et al. 2020) was used for maximum likelihood (ML) reconstruction based on the selected model. Branch supports were calculated using ultrafast bootstrap (UFBoot) and SH-like approximate likelihood ratio test (SH-aLRT) with 1000 replicates (Figure 1). *Maesa hupehensis* was closely related to *M. montana* (KU569490). Our results provide valuable data and shed light on the phylogenomic study of Primulaceae.
Ethical approval

Research and collection of plant material was conducted according to the guidelines provided by Herbarium of Xi’an Botanical Garden. Permission was granted by Herbarium of Xi’an Botanical Garden to carry out research on the species.

Author contributions

P.X. conceived and designed this study. Y.L. and C.Z. conducted analysis. P.X. and C.Z. contributed the analytical methods. X.Y., Q.C., and C.Z wrote the manuscript. P.X. edited the manuscript. All authors have read and agreed to the published version of the manuscript.

Disclosure statement

No conflict of interest exits in the submission of this manuscript, and manuscript is approved by all authors for publication. And we declared comply 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.

Funding

This work was supported by Key Research and Development Project of the Ningxia Hui Autonomous Region [Project No.2021BBF02008].

ORCID

Pengguo Xia http://orcid.org/0000-0003-3572-7616

Data availability statement

The data that support the findings of this study are openly available in NCBI (https://www.ncbi.nlm.nih.gov) GenBank with the accession number (MZ846203). The associated BioProject, SRA, and BioSample numbers are PRJNA756405, SRR15533156, and SAMN20865416 respectively.

References

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Bot J Linnean Soc. 161:948–121.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 34(17):i884–i890.

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

Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21(1):241–241.

Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14(6):587–589.

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

Lai GF, Chen JJ, Wang YF, Cao GX, Wang HY, Luo SD. 2002. Advances in studies on Maesa Forsk. Chinese Trad Herbal Drugs. 33(6):562–563.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Haeseler MD, Von A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 37(5):1530–1534.

Wang J, Xia NH. 2011. A new variety of Maesa Forsk. (Maesacerae) from Kachin State of Myanmar. J Trop Subtrop Bot. 19(03):225–227.