

Characterization of the complete chloroplast genome of *Malus baccata* var. *xiaojinensis*

Yuan Ren^{a,b}, Ming Yan^{a,b}, Xueqing Zhao^{a,b}, Xuesen Chen^c and Zhaohe Yuan^{a,b}

^aCo-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China; ^bCollege of Forestry, Nanjing Forestry University, Nanjing, China; ^cState Key Laboratory of Crop Biology, Shandong Agricultural University, Tai'an, Shandong, China

ABSTRACT

Malus baccata var. *xiaojinensis* belonging to Rosaceae is not only an ornamental tree but also an apple genotype that is highly tolerant to Fe deficiency. In this study, we reported the complete chloroplast (cp) genome of *M. baccata* var. *xiaojinensis* using ILLUMINA sequencing. The whole cp genome is 160,067 bp in length, containing a pair of inverted repeats (IRs) of 26,358 bp, a large single copy (LSC) region of 88,157 bp and a small single copy (SSC) region of 19,194 bp. And, the overall GC content of the cp genome was 36.56%. A total of 110 unique genes were found in the cp genome. 17 genes were duplicated in the IRs, including six protein-coding genes, seven tRNA genes, and four rRNA genes. Fourteen genes contained one intron, whereas three genes (*rps12*, *clpP*, and *ycf3*) contained two introns. The phylogenetic analysis demonstrated a close relationship between *M. baccata* var. *xiaojinensis* and *Malus hupehensis*.

ARTICLE HISTORY

Received 6 April 2019
Accepted 22 June 2019

KEYWORDS

Malus baccata var. *xiaojinensis*; complete chloroplast genome; phylogenetic analysis

Malus baccata var. *xiaojinensis* is a small wild apple with elegant shapes and bright flowers and red or yellow fruits (Zhi-Qin et al. 1999) that can be used as a rootstock with resistance to iron deficiency, low temperature, and drought stress (Han et al. 1994; Liu et al. 2018). Therefore, the anti-iron deficiency characteristics will make it the most valued rootstock resource in the future. Besides, the chloroplast genome has shown great potential for solving phylogenetic problems (Hao et al. 2016) and is applicable for elucidating the taxonomic characteristics of the species and origin of the hybrid progeny (Jansen et al. 2007; Givnish et al. 2010). The comparative analysis of the cp genomes toward related species helps to understand the relationship between the important traits under the control of plastid genome (Liu et al. 2013; Yan et al. 2019). In this study, we determined the complete chloroplast genome sequence of *M. baccata* var. *xiaojinensis* based on next-generation sequencing data and compared it with other genus chloroplast genome sequences. The complete cp genome has been deposited in GenBank with accession number MK434915.

Total genomic DNA was extracted from young leaf tissue collected from Shandong Agricultural University using the DNeasy Plant Mini Kit (Qiagen, Venlo, the Netherlands). The voucher specimen was preserved in Shandong Agricultural University (117.112, 36.194). Paired-end Illumina genomic library was prepared and sequenced using an HiSeq 2500 platform according to the manufacturer's instructions (Illumina, San Diego, CA). The raw paired-end reads were

filtered and using fastp program (Chen et al. 2018). The high-quality reads were applied to a de novo assembly performed using GetOrganelle (Jin et al. 2018). Annotation was completed by the online program GeSeq (Tillich et al. 2017) and the result was manually adjusted where necessary using Geneious (Kearse et al. 2012).

The complete cp genome was 160,067 bp in length, containing a pair of inverted repeats (IRs) of 26,358 bp, a large single copy (LSC) region of 88,157 bp and a small single copy (SSC) region of 19,194 bp. A total of 110 unique genes were found in the cp genome. Seventeen genes were duplicated in the IRs, including six protein-coding genes, seven tRNA genes, and four rRNA genes. Among these genes, 14 genes contained one intron, whereas three genes (*rps12*, *clpP*, and *ycf3*) contained two introns. The overall GC content of the cp genome was 36.56%, which the corresponding values in LSC, SSC, and IRs regions were 34.22, 30.39 and 42.69%, respectively.

The complete cp genome sequences of twelve species were downloaded from GenBank. Multiple sequences alignment was executed using MAFFT (Katoh and Toh 2010). The maximum-likelihood (ML) phylogenetic tree was constructed using the IQ-tree with the best-fit model determined using ModelFinder (Nguyen et al. 2015; Kalyaanamoorthy et al. 2017). *Pyrus spinosa* and *Pyrus pashia* were set as outgroups. The phylogenetic tree provided valuable information for understanding the relationship within genus *Malus*. Our results (Figure 1) indicated that *M. baccata* var. *xiaojinensis* is closely related to *M. hupehensis*.

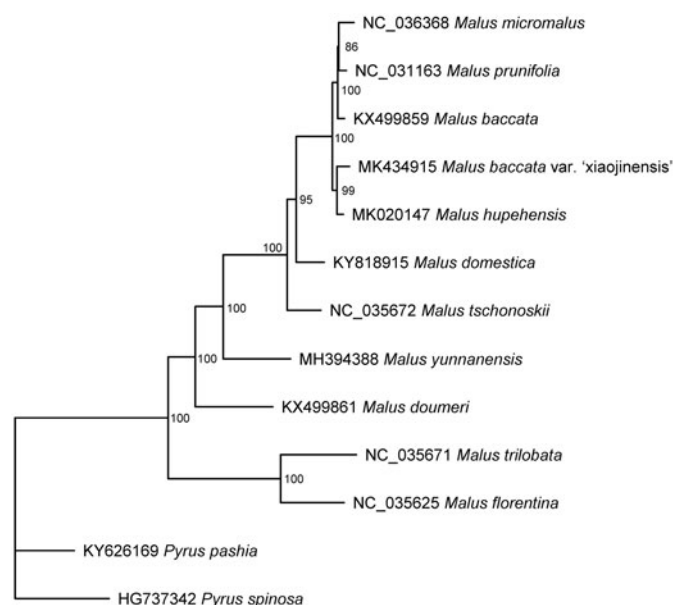


Figure 1. The phylogenetic tree based on the 13 complete chloroplast genomes sequences.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Initiative Project for Talents of Nanjing Forestry University [GXL2014070, GXL2018032], the Priority Academic Program Development of Jiangsu High Education Institutions (PAPD), and the Natural Science Foundation of Jiangsu Province [BK20180768].

References

- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34:i884–i890.
- Givnish TJ, Ames M, McNeal JR, McKain MR, Steele PR, dePamphilis CW, Graham SW, Pires JC, Stevenson DW, Zomlefer WB, et al. 2010. Assembling the tree of the monocotyledons: plastome sequence phylogeny and evolution of poales. *Ann Missouri Bot Garden*. 97:584–616.
- Han ZH, Wang Q, Shen T. 1994. Comparison of some physiological and biochemical characteristics between iron-efficient and iron-inefficient species in the genus *malus*. *J Plant Nutr*. 17:1257–1264.
- Hao Z, Cheng T, Zheng R, Xu H, Zhou Y, Li M, Lu F, Dong Y, Liu X, Chen J, Shi J. 2016. The complete chloroplast genome sequence of a relict conifer *Lyptostrobus pensilis*: comparative analysis and insights into dynamics of chloroplast genome rearrangement in Cupressophytes and Pinaceae. *PLoS One*. 11:e0161809.
- Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack J, Muller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, et al. 2007. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci USA*. 104:19369–19374.
- Jin J, Yu W, Yang J, Song Y, Yi T, Li D. 2018. GetOrganelle: a simple and fast pipeline for de novo assembly of a complete circular chloroplast genome using genome skimming data. *bioRxiv*. 256479.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 14:587–589.
- Katoh K, Toh H. 2010. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics*. 26:1899–1900.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz 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: 1647–1649.
- Liu W, Wu T, Li Q, Zhang X, Xu X, Li T, Han Z, Wang Y. 2018. An ethylene response factor (MxERF4) functions as a repressor of Fe acquisition in *Malus xiaojinensis*. *Sci Rep*. 8:1068.
- Liu Y, Huo N, Dong L, Wang Y, Zhang S, Young HA, Feng X, Gu YQ. 2013. Complete chloroplast genome sequences of Mongolia medicine *Artemisia frigida* and phylogenetic relationships with other plants. *PLoS One*. 8:e57533.
- Nguyen L, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*. 32:268–274.
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Res*. 45:W6–W11.
- Yan M, Zhao X, Zhou J, Huo Y, Ding Y, Yuan Z. 2019. The complete chloroplast genome of cultivated apple (*Malus domestica* Cv. 'Yantai Fuji 8'). *Mitochondrial DNA B*. 4:1213–1216.
- Zhi-Qin Z. 1999. The apple genetic resources in China: the wild species and their distributions, informative characteristics and utilisation. *Genetic Res Crop Evol*. 46:599–609.