

## Complete plastome sequence of *Manilkara zapota* (L.) P.Royen (Sapotaceae)

Shu-Xiang Li, Xiao-Feng Zhang, Hong-Xin Wang, Zhi-Xin Zhu and Hua-Feng Wang

Hainan Key Laboratory for Sustainable Utilization of Tropical Bioresources, School of Life and Pharmaceutical Sciences, Hainan University, Haikou, China

### ABSTRACT

The plastome of *Manilkara zapota* is found to be 158,386 bp long with the typical quadripartite structure of angiosperms, contains two inverted repeats (IRs) of 26,099 bp each, a large single-copy (LSC) region of 87,745 bp, and a small single-copy (SSC) region of 18,443 bp. The plastome contains 114 genes, consisting of 80 unique protein-coding genes, 30 unique tRNA genes, and 4 unique rRNA genes. The overall A/T content in the plastome of *M. zapota* is 63.00%. The phylogenetic analysis indicated that *M. zapota* is close to *Sideroxylon wightianum* within Sapotaceae in this study.

### ARTICLE HISTORY

Received 13 August 2019  
Accepted 16 August 2019

### KEYWORDS

*Manilkara zapota*; plastome; phylogeny; genome structure; Sapotaceae

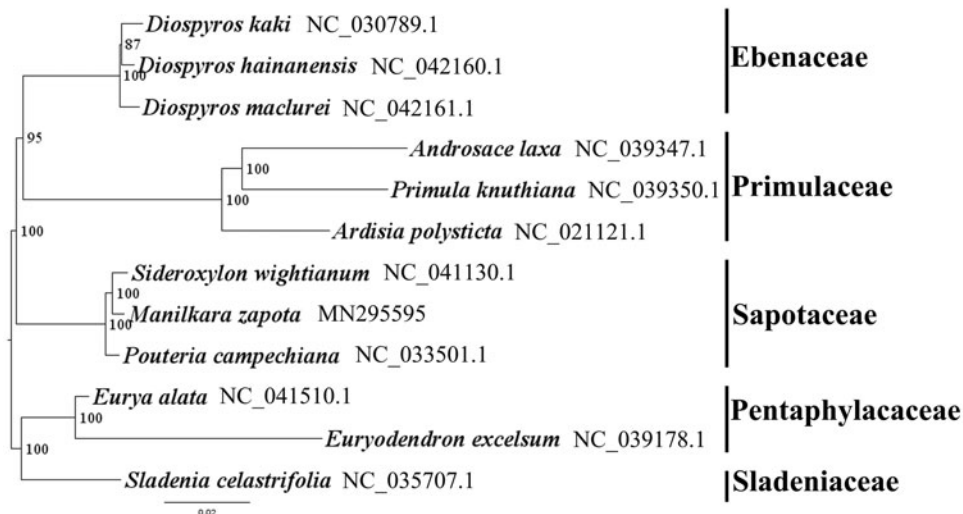
### Introduction

*Manilkara zapota* (L.) P.Royen (Sapotaceae, Ericales) is a kind of arbor. It is 15–20 m tall. It originates in tropical America, cultivated in Guangdong, Guangxi, and Yunnan (Xishuangbanna) of China. Its fruit is edible, sweet, and delicious, its trunk milk is raw material of chewing gum, its seed contains 20% oil and its bark contains alkaloids, which can cure fever (Editorial Committee of Flora of China, Chinese Academy of Sciences 1987). Consequently, the genetic and genomic information is urgently needed to promote its systematic research and the development of its conservation

value. Here, we report and characterize the complete plastome of *M. zapota* (GenBank accession number: MN295595). This is the first report of a complete plastome for *M. zapota*.

In this study, *M. zapota* was sampled from the greenhouse within Hainan University campus, Haikou, Hainan, China (110.327°E, 20.059°N). A voucher specimen (Wang et al., B254) was deposited in the Herbarium of the Institute of Tropical Agriculture and Forestry (HUTB), Hainan University, Haikou, China.

Around 6 Gb clean data were assembled against the plastome of *Camellia huana* (KY626040.1) (Wang et al. 2017) using MITObim v1.8 (Hahn et al. 2013). The plastome was



**Figure 1.** The best ML phylogeny recovered from 12 complete plastome sequences by RAxML. Accession numbers: *Manilkara zapota* MN295595, *Pouteria campechiana* NC\_033501.1, *Sideroxylon wightianum* NC\_041130.1, *Primula knuthiana* NC\_039350.1, *Ardisia polysticta* NC\_021121.1, *Androsace laxa* NC\_039347.1, *Diospyros hainanensis* NC\_042160.1, *Diospyros kaki* NC\_030789.1, *Diospyros maclurei* NC\_042161.1. Outgroups: *Eurya alata* NC\_041510.1, *Euryodendron excelsum* NC\_039178.1, *Sladenia celastrifolia* NC\_035707.1.

annotated using Geneious R8.0.2 (Biomatters Ltd., Auckland, New Zealand) against the plastome of *C. huana* (KY626040.1). The annotation was corrected with DOGMA (Wyman et al. 2004).

The plastome of *M. zapota* was found to be 158,386 bp long with the typical quadripartite structure of angiosperms, containing two inverted repeats (IRs) of 26,099 bp, a large single-copy (LSC) region of 87,745 bp, and a small single-copy (SSC) region of 18,443 bp. The plastome contains 114 genes, consisting of 80 unique protein-coding genes, 30 unique tRNA genes, and 4 unique rRNA genes. The overall A/T content in the plastome of *M. zapota* is 63.00%, with the corresponding values of the LSC, SSC, and IR regions equalling 65.10, 69.80, and 57.10%, respectively.

We used RAxML (Stamatakis 2006) with 1000 bootstraps under the GTRGAMMAI substitution model to reconstruct a maximum-likelihood (ML) phylogeny of eight published complete plastomes of three families (Ebenaceae, Primulaceae and Sapotaceae), using 3 species from Pentaphragmaceae and Sladeniaceae as outgroups. The phylogenetic analysis indicated that *M. zapota* is close to *Sideroxylon wightianum* within Sapotaceae in this study (Figure 1). Most nodes in the plastome ML tree were strongly supported. The complete plastome sequence of *M. zapota* will provide a useful resource for the conservation genetics of this species as well as for the phylogenetic studies of Sapotaceae.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This study was funded by National Natural Scientific Foundation of China [31660055 and 31660074], start-up fund from Hainan University [kyqd1633 and kydq(zr)1840] and by Postgraduate Innovation Project of Biological Science of Tropical Agriculture and Forest Institute, Hainan University.

## References

- Editorial Committee of Flora of China, Chinese Academy of Sciences. 1987. Flora of China. Vol. 60. Beijing (China): Science Press; p. 50.
- Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* 41:e129.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 22:2688–2690.
- Wang G, Luo Y, Hou N, Deng LX. 2017. The complete chloroplast genomes of three rare and endangered camellias (*Camellia huana*, *C. liberofilamenta* and *C. luteoflora*) endemic to southwest china. *Conservation Genet Resour.* 9:583–585.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.