The complete chloroplast genome of *Mimusops elengi* (Sapotaceae: Sapoteae)

Qiyi Huang\(^a\) and Lijuan He\(^b\)

\(^a\)School of Civil Engineering, Xiamen University Tan Kah Kee College, Fujian, China; \(^b\)Xiamen Overseas Chinese Subtropical Plant Introduction Garden, Plant Introduction & Quarantine and Plant Product Key Laboratory of Xiamen City, Fujian, China

**ABSTRACT**

The first complete chloroplast genome sequences of *Mimusops elengi* Linaeus, 1753 (Sapotaceae: Sapoteae) were reported in this study. The cpDNA of *M. elengi* is 159,719 bp in length, contains a large single-copy region (LSC) of 88,935 bp and a small single-copy region (SSC) of 18,606 bp, which were separated by a pair of inverted repeat (IR) regions of 26,089 bp. The genome contains 132 genes, including 87 protein-coding genes, 8 ribosomal RNA genes, and 37 transfer RNA genes. The overall GC content of the whole genome is 36.8%. Phylogenetic analysis of 8 chloroplast genomes within the tribe Sapoteae suggests that the sister relationship of *Autranella* and *Tieghemella* are strongly supported. *Minusops* genus is close to *Autranella* and *Tieghemella*, although the support value is still low.

**ARTICLE HISTORY**

Received 16 March 2022
Accepted 6 June 2022

**KEYWORDS**

*Mimusops elengi*; chloroplast genome; phylogenetic analysis

---

*Mimusops elengi* Linaeus, 1753 belongs to the tribe Sapoteae of Sapotaceae family, a. It is native to South and Southeast Asia and northern Australia, and has been known for its myriad values. Its timber is hard and durable. Its flower is aromatic. Its fruit is edible. It is also widely used in traditional medicine (Gupta 2013). It was introduced in China in 1962 from India and now is sparsely cultivated in Fujian, Guangdong, Guangxi, Hainan and Yunnan (Bai et al. 2011). This tribe includes 12 genera. But its taxonomy and phylogenetic relationship remained controversial due to low resolution of limited molecular evidence (Swenson and Anderberg 2005; Kümper et al. 2016). Till now, 8 plastomes representative of 7 genera from this tribe have been reported. Therefore, we sequenced the complete chloroplast genome of *M. elengi*. It cp genome sequence would contribute to the further study of this species and also provide genetic information for the construction of the phylogenetic relationship among Sapoteae.

The silica-dried material of *M. elengi* was obtained and from Xiamen Overseas Subtropical Plant Introduction Garden, Fujian province in China (latitude 24.4485°N and longitude 118.0717°E). The specimens were deposited in its herbarium, OSBG with specimen code YZY20170264 (Mrs. He, helj2013@126.com). Total genomic DNA was extracted using Tiangen Plant Genomic DNA Kit (Tiangen Biotech Co., Beijing, China) and sheared into ca. 350 bp fragments. Libraries for paired-end (PE) Illumina sequencing were conducted following the standard protocol of manufacture (TruSeq Library Construction Kit) and sequenced from both ends of 150 bp fragments on the Illumina HiSeq 2000 platform at Novogene Co., Ltd (Beijing, China). The complete chloroplast genome was assembled, circumscribed and annotated using GENIOUS PRIME 2019.1.1 (Kearse et al. 2012) with cp genome of *M. coriacea* (MW846242) as reference. The chloroplast genome was submitted to GenBank (http://www.ncbi.nlm.nih.gov/) with accession numberOK458682.

Chloroplast genome size of *M. elengi* is 159,719 bp with the GC content of 36.8%. It consists of the traditional quadripartite structure: a large single-copy region of 88,935 bp, a small single copy region of 18,606 bp, and two inverted repeat regions of 26,089 bp. The new sequence possesses a total of 130 genes, containing 85 protein-coding genes (PCGs), 8 ribosomal RNA (rRNA) genes, and 37 transfer RNA (tRNA) genes.

In previous studies, the intrafamily relationship within Sapotaceae is difficult to resolve with limited markers (Swenson and Anderberg 2005). To achieve a reliable phylogenetic relationship, maximum likelihood tree was constructed based on 78 single copy protein-coding genes with 12 other Sapotaceae species using RAxML HPC v.8.2.4 (Stamatakis 2014). GTR + I + G model was selected as the best-fitting model determined by jModelTest v.2.1.1. *M. elengi* is revealed to be close to *Autranella* and *Tieghemella*, although the support value is still low (Figure 1). The sister relationship of the latter two genera are strongly supported. Our study shows the high strength of complete chloroplast genomes in uncovering the deep relationship within Sapotaceae.

**Ethical approval**

The data collection of plants was carried out with permission of related institution, and complied with national or international guidelines and legislation.
Author contributions
Q-Y H. and L-J H. were involved in the conception and design. Q.Y. H. contributed the drafting of the paper. L.J. H. performed the data analysis and revised it critically for intellectual content. All authors were involved in the final approval of the version to be published. All authors agree to be accountable for all aspects of the work.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Funding
This study was supported by by grants from the National Science Foundation of China [grant number 31900193].

Data availability statement
The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov/nuccore/OK458682 under the accession no. OK458682. The associated BioProject, SRA, and BioSample numbers are PRJEB49293, ERS8701945, and SAMEA11052235 respectively.

References
Gupta PC. 2013. *Mimusops elengi* Linn. (Bakul) – a potential medicinal plant: a review. Int J Pharm Phytopharmacol Res. 2:332–339.
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(12): 1647–1649.
Kumpers BM, Richardson JE, Anderberg AA, Wilkie P, Ronse De Craene LP. 2016. The significance of meristic changes in the flowers of Sapotaceae. Bot J Linn Soc. 180(2):161–192.
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9): 1312–1313.
Swenson U, Anderberg AA. 2005. Phylogeny, character evolution, and classification of Sapotaceae (Ericales). Cladistics. 21(2):101–130.

Figure 1. Maximum likelihood tree of *M. elengi* based on concatenated protein sequences of 78 protein-coding sequences shared by 13 Sapotaceae species. Numbers at the nodes indicates bootstrap values (1000 replicates).