The complete chloroplast genome of *Castanopsis mekongensis* A. Camus (Fagaceae)

Jingyu Peng\(^a\), Lin Yang\(^a\), Anan Duan\(^a\), Dawei Wang\(^a\) and Siguang Li\(^b\)

\(^a\)Key Laboratory for Forest Resources Conservation and Utilization in the Southwest Mountains of China Ministry of Education, Southwest Forestry University, Kunming, China; \(^b\)Yunnan Academy of Forestry and Grassland, Kunming, China

**ABSTRACT**

*Castanopsis mekongensis* is highly valued in the furniture industry for its good wood texture, an endemic species in Yunnan province, southwest China. In our study, the chloroplast genome of *C. mekongensis* was assembled and characterized based on Illumina pair-end sequencing data. The total genome size of *C. mekongensis* was 160,699 bp, displayed a typical quadripartite structure including a pair of inverted repeat (IRs, 25,714 bp) regions separated by a large single-copy (LSC, 90,278 bp) region and a small single-copy (SSC, 18,963 bp) region. The overall guanine–cytosine (GC) content was 36.8%. We annotated 130 genes in the genome, containing 85 protein-coding genes, 37 tRNA genes, and eight rRNA genes, 12 genes contain a single intron, and two genes have two introns. The result of phylogenetic analysis based on maximum-likelihood (ML) phylogenetic tree indicated that *C. mekongensis* was most closely related to *C. hainanensis*.

---

*Castanopsis mekongensis* is an evergreen plant that belongs to the family Fagaceae, which is mainly distributed in Yunnan province, southwest China (Xia and Tong 1998). *Castanopsis mekongensis* is highly valued in the furniture industry for its good wood texture (Duanmu 1996). However, the yield of wild *C. mekongensis* gradually decreases due to the slow growth rate and excessive excavation (Duanmu 1996). To restore its genetic diversity and to solve the market demand, it is necessary to protect the remaining wild resources. As such, to better develop effective protecting strategies, we should understand the phylogenetic relationships between *C. mekongensis* and other *Castanopsis* species in the first step.

The chloroplast owning one circular DNA molecule is a semi-autonomous cell organelle in higher plants, which play a significant role in photosynthesis (Isao et al. 2010). Due to its stable structure and uniparental inheritance, the chloroplast genome is extremely useful for studying the origin and evolution of plants at different taxonomic levels (Osuna-Mascaró et al. 2018). In recent years, the chloroplast genome of many species was used to demonstrate considerable phylogenetic information (Xue et al. 2019; Zhang et al. 2019), while the chloroplast genomes of *C. mekongensis* is still unknown. In this paper, we reported the complete chloroplast genome sequence of *C. mekongensis* employing high-throughput sequencing. Phylogenetic analysis was conducted to explore its phylogenetic relationships with other relatives.

The leaf samples of *C. mekongensis* were collected from Kunming, Yunnan, China (102°44′38″ E, 25°8′6″ N, 1928 m). Total DNA was extracted using a Wolact Magnetic beads Plant DNA Purification Kit (KLink Biotechnologies, Shenzhen, China), the voucher specimens of *C. mekongensis* were deposited at the herbarium of Southwest Forestry University, Kunming, China (accession number: SWFU-FAG-FCM-5825). Genomic DNA was sequenced using the Illumina Novaseq 6000 platform at Annoroad Gene Technology (Beijing, China). About 5.0 GB high-quality clean reads were generated with adaptors trimmed. Then, the complete chloroplast genome was assembled by Get Organelle v1.6.2 (Jin et al. 2020). Genome annotation was performed by Geneious R8 (Kearse et al. 2012) and manually adjusted by comparing it to the reference chloroplast genome of *C. hainanensis* (GenBank accession number MG383644.1). The annotated genomic sequence has been submitted to GenBank (accession number: MW043480).

The complete chloroplast genomes of *C. mekongensis* displayed the typical quadripartite structure with the length of 160,699 bp, including a large single copy (LSC) region of 90,278 bp, a small single copy (SSC) region of 18,963 bp, and a pair of inverted repeats (IRs) region of 25,714 bp. The guanine–cytosine (GC) content of chloroplast genomes is 36.8%; the GC content of the IR regions is 42.8% and those of LSC and SSC regions 34.6% and 30.9%, respectively. The chloroplast genomes of *C. mekongensis* contain 130 functional genes, including 85 protein-coding genes, eight tRNA genes, and 37...
tRNA genes. There are 12 genes (atpF, ndhB, petB, rpl2, rpcC1, rps16, trnA-UGG, trnG-GCC, trnI-UGU, trnL-UAA, and trnV-UAC) contain a single intron, and two genes (ycf3 and clpP) have two introns.

The chloroplast genomes of *C. mekongensis* in this study and 23 previously reported species of Fagales were aligned by MAFFT v7 using default settings (Katoh and Standley 2013). The maximum-likelihood (ML) phylogenetic tree based on complete chloroplast genome sequences was performed by IQ-tree 1.5.5 (Nguyen et al. 2015). The optimal nucleotide substitution model (TVM + F + R7) was selected by ModelFinder (Kalyaanamoorthy et al. 2017). The ML analyses were conducted with 1000 rapid bootstrap replicates along with a search for the best-scoring ML tree (Gao et al. 2018). The phylogenetic analysis showed that *C. mekongensis* appeared to be most closely related to *C. hainanensis* (Figure 1). The chloroplast genome of *C. mekongensis* will provide useful genetic information for understanding the evolutionary relationships and conservation management.

**Disclosure statement**

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

**Funding**

The study was supported by the Forestry Science and Technology Development Project of National Forestry and Grassland Administration of China [KJZXSA2019036] and Special Funds for Local Scientific and Technological Development Guided by the Central Government [219001].

**ORCID**

Jingyu Peng [http://orcid.org/0000-0002-6870-5686](http://orcid.org/0000-0002-6870-5686)

**Data availability statement**

The voucher specimens of *Castanopsis mekongensis* were deposited at the herbarium of Southwest Forestry University, Kunming, Yunnan, China (accession number: SWFU-FAG-FCM-5825). The complete chloroplast genome sequence of *C. mekongensis* in this study has been submitted to National Center for Biotechnology Information (NCBI) ([https://www.ncbi.nlm.nih.gov/](https://www.ncbi.nlm.nih.gov/)) and obtained the GenBank accession number (MW043480).

**References**

Duanmu X. 1996. Comprehensive utilization of *Castanopsis* resources in China. Forest Chem Newsl. 5:38–40.

Gao F, Shen J, Liao F, Cai W, Lin S, Yang H, Chen S. 2018. The first complete genome sequence of narcissus latent virus from *Narcissus*. Arch Virol. 163(5):1383–1386.

Isao M, Motomichi M, Kiyoshi K. 2010. Extensive frameshift at all AGG and CCC codons in the mitochondrial cytochrome c oxidase subunit 1 gene of *Perkinsus marinus* (Alveolata; Dinoflagellata). Nucleic Acids Res. 38(18):86–94.
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.

Kalyaanamoorthy S, Minh BQ, Wong TKF, 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.

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.

Nguyen LT, 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(1):268–274.

Osuna-Mascaró C, Rubio de Casas R, Perfectti F. 2018. Comparative assessment shows the reliability of chloroplast genome assembly using RNA-seq. Sci Rep. 8(1):1–12.

Xia NH, Tong YH. 1998. Flora Reipublicae Popularis Sinicae. Sci Press. 22:31.

Xue S, Shi T, Luo W, Ni X, Iqbal S, Ni Z, Huang X, Yao D, Shen Z, Gao Z. 2019. Comparative analysis of the complete chloroplast genome among Prunus mume, P. armeniaca, and P. salicina. Hortic Res. 6(1):89.

Zhang Y, Zong D, Zhou A, He C. 2019. The complete chloroplast genome of Populus tomentosa narrow crown, native to Northern China. Mitochondrial DNA B. 4(1):1208–1210.