

Complete chloroplast genome of a cultivated oil camellia species, *Camellia gigantocarpa*

Yufen Xu, Yanju Liu and Xiaocheng Jia

Hainan Key Laboratory of Tropical Oil Crops Biology/Coconut Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wenchang, China

ABSTRACT

Camellia gigantocarpa Hu et T. C. Huang, belonging to the Theaceae family, is an excellent landscape tree species with high ornamental value. It is also an important woody oil-bearing plant with high economic value. This study reports the first complete chloroplast genome sequence of *C. gigantocarpa* (GenBank accession number: MZ054232). Its whole chloroplast genome is 156,953 bp long with an overall GC content of 37.31%, which is composed of a large single copy region (86,631 bp), a small single copy region (18,402 bp), and a pair of inverted repeat regions (25,960 bp each). A total of 135 genes were predicted in this genome, namely eight ribosomal RNA genes, 37 transfer RNA genes, and 90 protein-coding genes. Based on maximum likelihood analysis results, we found that the *Camellia* species are clustered into a distinct branch, and the phylogenetic relationships among *C. gigantocarpa*, *C. crapnelliana*, and *C. kissii* were the closest.

ARTICLE HISTORY

Received 19 August 2021
Accepted 12 November 2021

KEYWORDS

Camellia; Theaceae;
complete chloroplast genome


In 1965, Hu Hsen-Hsu first published *Camellia gigantocarpa* Hu et T. C. Huang as a new species (Hu 1965). This species belongs to the Theaceae family, and is naturally distributed in the Bobai and Luchuan areas of Guangxi Province in China (Huang et al. 2021). *Camellia gigantocarpa* has a high ornamental value and is an excellent landscape tree species with a beautiful tree shape, large flowers, and large fruits. Moreover, seeds of this species have a high oil content and can be used for edible oil extraction with high economic value (Fan 2011). As an important woody oil-bearing plant, *C. gigantocarpa* is also one of the main cultivated species of oil-tea camellia in China (Dai et al. 2021). Research on *C. gigantocarpa* has mainly focused on genetic diversity (Fan et al. 2011; Peng et al. 2013), photosynthetic characteristics (Wu 2019), camellia seed oil composition, and soil heavy-metal remediation (Zhang et al. 2010); however, only a few studies on its taxonomy and phylogeny exist. In this report, we present the first complete chloroplast genome sequence of *C. gigantocarpa* and evaluate its phylogenetic relationships with related species.

Samples of *C. gigantocarpa* were collected from the Germplasm Repository of Oil Camellia at the Coconut Research Institute of the Chinese Academy of Tropical Agricultural Sciences (CATAS) (Hainan, China; coordinates: 19°32'4.80" N, 110°45'47.43" E). The sample (BB01) and voucher herbarium (BB1) were deposited in the Camellia Research Center of the Coconut Research Institute of CATAS.

Total genomic DNA was extracted from the leaf material using a modified CTAB method (Doyle 1987). The library was constructed with an insert length of 350 bp and paired-end sequenced with 150-bp reads on the Illumina HiSeq2500 second-generation sequencing platform. In total, approximately 10.79 GB of raw reads was generated. With *C. pubicosta* (NC_024662.1) as the reference, the toolkit GetOrganelle (Jin et al. 2020) was used to assemble the chloroplast genome de novo. Similar to gene annotation, the starting positions of the chloroplast genome and the inverted repeat (IR) region were determined using the online annotation software Geseq (Tillich et al. 2017) (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>) and CpGAVAS2 (Shi et al. 2019) (<http://www.herbalgenomics.org/cpgavas>). Finally, manual checking was performed to ensure that the annotations were correct, and the complete chloroplast genome of *C. gigantocarpa* was submitted to GenBank (accession number: MZ054232).

The entire chloroplast genome of *C. gigantocarpa* is a typical circular quadripartite structure 156,953 bp in length; the genome includes a large single copy region (LSC, 86,631 bp), a small single copy region (SSC, 18,402 bp), and a pair of IR regions (25,960 bp each). The overall GC content was 37.31%. According to the annotation results, the total number of functional genes was 135, consisting of eight ribosomal RNA genes, 37 transfer RNA genes, and 90 protein-coding genes. And the annotated gene names in chloroplast genome from

CONTACT Xiaocheng Jia  xcjia1@163.com  Hainan Key Laboratory of Tropical Oil Crops Biology/Coconut Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wenchang, China

 Supplemental data for this article is available online at [here](#).

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

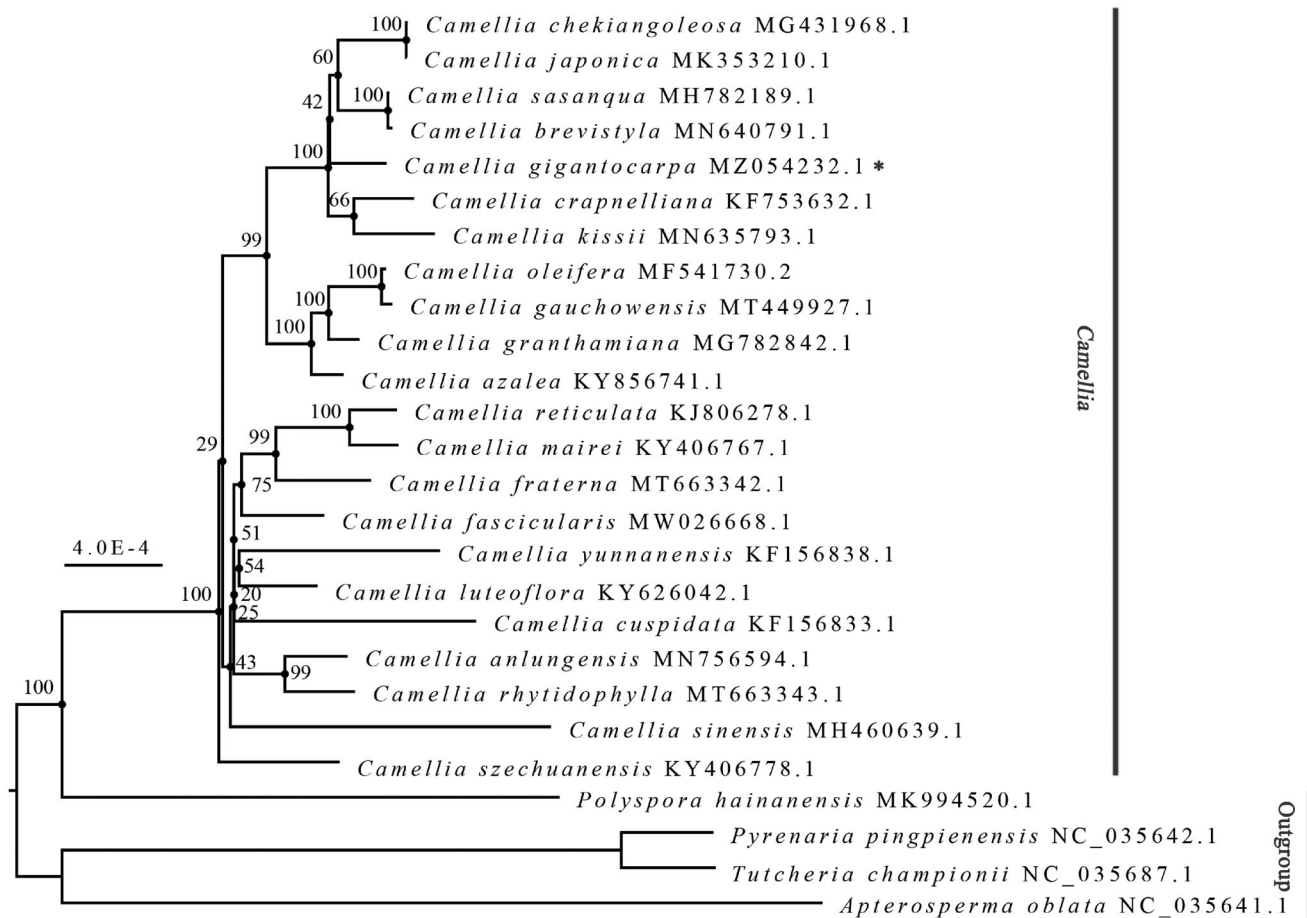


Figure 1. The maximum likelihood phylogenetic tree based on 26 chloroplast genome sequences of the Theaceae family. Bootstrap values based on 1,000 replicates are indicated at each branch node.

C. gigantocarpa were listed in Supplemental material 1. Similar to most angiosperms, the chloroplast genome of *C. gigantocarpa* lost characteristic chloroplast genes of algae, bryophytes, pteridophytes, and gymnosperms, that is *psaM*, *psb30*, *chlB*, *chlL*, *chlN*, and *rpl21*. However, unlike most angiosperms, the chloroplast genome of *C. gigantocarpa* lost *atpI* and *rpl36* genes as well. This may be attributable to the reorganization of the chloroplast genome during the evolution of different lineages (Mohanta et al. 2020).

To explore the phylogenetic relationships within different *Camellia* species, a maximum likelihood phylogeny analysis was performed using IQ-TREE 2.1.3 (Minh et al. 2020). The best model for the complete chloroplast genomes of *C. gigantocarpa* (MZ054232), 21 other *Camellia* species and 4 outgroup species from Theaceae was K3Pu+I; 1,000 bootstrap replicates were used for this purpose. As illustrated in Figure 1, all the 22 *Camellia* species distinctly clustered into one large branch with relatively short internal evolutionary branches, whereas the other four species formed a separate branch with the relatively long internal evolutionary branches. This indicates that chloroplast genomes of *Camellia* species are distinctly and genetically different from those of the four outgroup Theaceae species, while the chloroplast genomes of *Camellia* species have a relatively high degree of genetic similarity. Furthermore, our results indicate that the *Camellia* genus is a monophyletic group, which is consistent with the findings of Yu et al. (2017). We also found close

genetic relationships among *C. gigantocarpa*, *C. crapnelliana*, and *C. kissii*. We consider that the report provides a scientific basis for the molecular phylogeny and molecular breeding of *Camellia* species for research and industrial purposes.

Disclosure statement

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

Funding

This study was supported by Hainan Provincial Natural Science Foundation of China [319MS081] and Central Finance Forestry Science and Technology Popularization & Demonstration Project (QIONG [2020] TG06).

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/>) under the accession no. MZ054232. The associated BioProject, Bio-Sample, and SRA numbers are PRJNA725044, SAMN20667270, and SRR15371496, respectively.

References

Dai J, Yu J, Qi H, Zheng W, Wang J, Wu Y, Lai H, Hu X. 2021. DNA barcoding identification of different species in *Camellia* based on *trnH-psbA* and *matK* sequences. *Chinese J Trop Crops*. 42(3):611–619.

- Doyle J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19(1):11–15.
- Fan H. 2011. The study on the genetic diversity of *Camellia gigantocarpa* Hu et T. C. Huang by ISSR. *Central South University of Forestry and Technology.*
- Fan H, Cao F, Peng J, Long J, Deng M, Si S. 2011. Establishment and optimization of ISSR-PCR reaction system of *Camellia gigantocarpa*. *J Central South Univ Forest Technol.* 31(04):97–103.
- Hu X. 1965. New species and varietie of *Camellia* and Theopsis of China (1). *Acta Phytotaxonomica Sinica.* 10(2):131–142.
- Huang T, Wang J, Chen G, Ma J. 2021. Study on drying characteristics and drying schedule of *Camellia gigantocarpa* wood. *Guangxi Forest Sci.* 50(03):325–329.
- Jin J, Yu W, Yang J, Song Y, DePamphilis CW, Yi T, Li D. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21(1):241.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler 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.
- Mohanta T K, Mishra AK, Khan A, Hashem A, Abd Allah EF, Al-Harrasi A. 2020. Gene Loss and Evolution of the Plastome Genes. 11:1133.
- Peng J, Cao F, Fan H. 2013. Study on genetic diversity of *Camellia gigantocarpa* detected by ISSR markers. *J Central South Univ Forest Technol.* 33(07):62–66.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73.
- 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(W1):W6–W11.
- Wu X. 2019. The research on photosynthetic characteristics and assimilate distribution of six *Camellia* plants. *Central South University of Forestry and Technology.*
- Yu X, Gao L, Soltis DE, Soltis PS, Yang J, Fang L, Yang S, Li D. 2017. Insights into the historical assembly of East Asian subtropical evergreen broadleaved forests revealed by the temporal history of the tea family. *New Phytol.* 215(3):1235–1248.
- Zhang C, Chou S, Zhao J, Li X, Quan Y. 2010. Phytoremediation properties of three ornamental plants for cadmium absorption in soils. *Guangxi Agric Sci.* 41(10):1101–1103.