

The complete chloroplast genome of *Camellia osmantha*, an edible oil *Camellia*

Yanju Liu, Yufen Xu and Xiaocheng Jia

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

ABSTRACT

Camellia osmantha is a new species of the *Camellia* genus discovered in Nanning, Guangxi, China, in 2012. It can be used as an excellent woody oil crop. There is little related research on this species in China and abroad, and its genome information is still lacking. In this study, the complete chloroplast genome sequence of *C. osmantha* was first reported (GenBank number: MZ128138). The whole chloroplast genome is 156,981 bp in length with a GC content of 37.28%, and it is composed of a large single copy (LSC) region of 86,647 bp, a small single copy (SSC) region of 18,284 bp, and a pair of inverted repeat (IR) regions of 26,025 bp each. The genome contains a total of 135 functional genes, including 37 transfer RNA genes, 90 protein-coding genes, and 8 ribosomal RNA genes. The maximum likelihood analysis based on 21 chloroplast genomes showed that *C. osmantha* and *C. oleifera* (MF541730.2) were the most closely related.

ARTICLE HISTORY

Received 24 June 2021
Accepted 26 September 2021

KEYWORDS

Camellia osmantha;
chloroplast genome;
phylogenetic analysis

Camellia osmantha is a new species discovered in Nanning, Guangxi, China, in 2012. It can be used as an excellent woody oil crop with the characteristics of large fruit quantity, early maturity, and high seed yield (Ouyang et al. 2020). Compared with ordinary *Camellia oleifera*, this new species shows stronger growth potential and stress resistance (Jiang et al. 2016). It is distributed in the tropics and the northern edge of the tropics and has physiological and ecological characteristics that make it well adapted to the tropical climate. The fruit of *C. osmantha* can be squeezed into oil and has high economic value, and the oil is rich in several fatty acids, including stearic, linoleic, and palmitic acids (Xia et al. 2014). Simultaneously, *C. osmantha* can be used as an ornamental tree species with a beautiful tree shape and fragrant flowers (Ma et al. 2012). There is little research on this species in China and abroad, and its genome information is still lacking. Therefore, we assembled and annotated the complete chloroplast (cp) genome (NCBI accession number: MZ128138) of *Camellia osmantha* by high-throughput sequencing, which can provide a theoretical basis for phylogenetics, comparative studies and chloroplast engineering of *C. osmantha*.

The fresh leaves of *C. osmantha* from the Coconut Research Institute of CATAS (Hainan, China; Coordinates: 19°32'4.92"N, 110°45'47.29"E) were used as research materials, and the certificate specimen (XH1) is stored in the *Camellia* Research Center, Coconut Research Institute of CATAS. Genomic DNA was extracted by Takara DNA extrac-

tion kit (Takara, Dalian, China); library construction and sequencing were performed on the Illumina HiSeq 2500 second-generation platform; and chloroplast assembly software GetOrganelle (Jin et al. 2020) was used for genome assembly. We used the online annotation software Geseq (Tillich et al. 2017) and CpGAVAS (Liu et al. 2012) to assemble the cp genome annotation, *Camellia pubicosta* (NC_024662.1) was used as reference genome. After proofreading and correction, the complete cp genome was submitted to GenBank (NCBI accession number: MZ128138).

The *C. osmantha* cp genome is a circular molecule with a total length of 156,981 bp, containing a large single copy (LSC) region of 86,647 bp, a small single copy (SSC) region of 18,284 bp, and a pair of inverted repeat (IR) regions of 26,025 bp each. The total GC content is 37.28%. The corresponding GC contents in the LSC, SSC, and IR regions are 35.3, 30.5, and 43.0%, respectively. The cp genome has 135 functional genes, including 37 transfer RNA genes, 90 protein-coding genes, and 8 ribosomal RNA genes.

The complete cp genomes of *C. osmantha* and 17 other *Camellia* species were used for phylogenetic analysis, and RAxML 8.2.12 (Stamatakis 2014) was used to construct a maximum likelihood (ML) tree in GTRGAMMA model, with three closely related *Camellia* species as the outgroups. As shown in Figure 1, in this study, all the *Camellia* plants gathered into an independent branch, and the relationship between

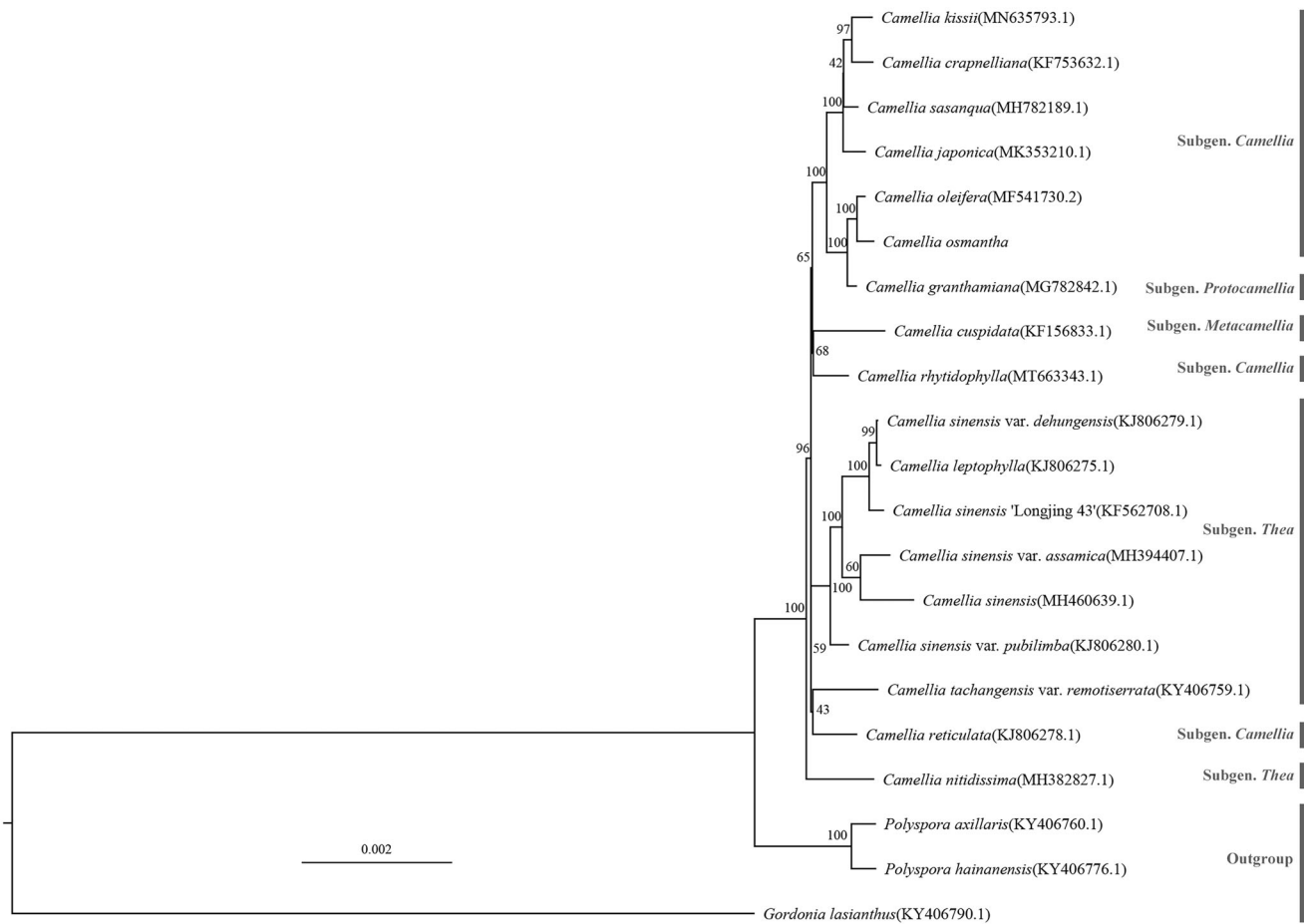


Figure 1. A phylogenetic tree of *Camellia* based on 21 complete chloroplast genome sequences. The bootstrap values from 1000 replicates are listed on each node.

C. osmantha and *C. oleifera* (MF541730.2) was found to be the closest.

Disclosure statement

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

Funding

This study was supported by Central Finance Forestry Science and Technology Popularization & Demonstration Project [QIONG [2020] TG06], Hainan Provincial Natural Science Foundation of China [319MS081] and National Key R&D Program of China [2019YFD1001602].

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. MZ128138. The associated BioProject, Bio-Sample, and SRA numbers are PRJNA725044, SAMN18928057, and SRR14469701 respectively.

References

- Jiang ZP, Wang DX, Hao BQ, Ye H. 2016. Screening of excellent rootstock line *Camellia osmantha*. *J Anhui Agri Sci.* 44(29):161–163.
- Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21(1):241.
- Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics.* 13:715 doi:10.1186/1471-2164-13-715. PMC: 23256920
- Ma JL, Ye H, Ye CX. 2012. A new species of *Camellia* sect. *Paracamellia*. *Guihaia.* 32 (6):753–755.
- Ouyang XL, Yang L, Huang L, Pan YM. 2020. Antitumor activity on human bladder cancer T-24 cells and composition analysis of the core of *Camellia osmantha* fruit. *Nat Prod Res.* 34(18):2689–2693.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9):1312–1313.
- 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. doi:10.1093/nar/gkx391. 28486635
- Xia EH, Jiang JJ, Huang H, Zhang LP, Zhang HB, Gao LZ. 2014. Transcriptome analysis of the oil-rich tea plant, *Camellia oleifera*, reveals candidate genes related to lipid metabolism. *PLoS One.* 9(8):e104150.