

The complete chloroplast genome of *Castanopsis carlesii* (Hemsl.) Hay

Rong-Xi Sun, Xue-Min Ye, Zhen-Li Wang and Xiao-Fan Lin

Jiangxi Provincial Key Laboratory of Silviculture, College of Forestry, Jiangxi Agricultural University, Nanchang, China

ABSTRACT

Castanopsis carlesii (Hemsl.) Hay. is a widely distributed and dominant tree species with significant ecological and economical values. In this study, the complete chloroplast genome sequence of *C. carlesii* was reported using the Illumina Hiseq 2500 platform. The complete chloroplast genome was 160,205 bp forming a typical quadripartite structure, with a pair of inverted repeated (IRs) regions of 25,670 bp, a large single copy (LSC) region of 89,849 bp, and a small single copy (SSC) region of 19,016 bp. A total of 124 functional genes were annotated, including 79 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The ML phylogenetic analysis showed that the genus *Castanopsis* formed a clade except *Castanopsis fargesii*.

ARTICLE HISTORY

Received 17 June 2019
Accepted 22 June 2019

KEYWORDS

Castanopsis carlesii;
complete chloroplast
genome; phylogenetic
analysis

Castanopsis carlesii (Hemsl.) Hay. is a widely distributed and dominant tree species of the genus *Castanopsis* from the family Fagaceae in subtropical evergreen broadleaved forest of China, with significant ecological and economical values (Lin et al. 2017). However, few studies on the chloroplast genome of *C. carlesii* was reported, so, we assembled and characterized its complete chloroplast genome, in order to provide information for the study the origin of *C. carlesii*.

The plant material was collected in Ganzhou of Jiangxi Province, China (24°33'N, 114°25'E). The voucher specimen was deposited in Jiangxi agricultural university college of forestry (No. JXMIZHU1). The total genomic DNA of *C. carlesii* was extracted using the Plant Genomic DNA Kit and sequenced using Illumina Hiseq2500 platform. Approximately 6 GB of clean data were yielded. These trimmed reads were assembled by NOVOPlasty (Dierckx et al. 2017), and then the assembled genome was annotated using CpGAVAS (Liu et al. 2012) and DOGMA (Lohse et al. 2007).

The complete chloroplast genome of *C. carlesii* was 160,205 bp in length forming a typical quadripartite structure,

with a pair of inverted repeated (IRs) regions of 25,670 bp, a large single copy (LSC) region of 89,849 bp, and a small single copy (SSC) region of 19,016 bp. The GC content of the whole chloroplast genome was 36.82%. The proportions of GC contents in IRs, LSC, and SSC were 42.8%, 34.67%, and 30.82%, respectively. A total of 124 functional genes were annotated, including 79 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The majority of genes were single copy, whereas 19 functional genes were duplicated, including eight protein-coding genes, four rRNA genes, and seven tRNA genes. All genes were encoded by 25 997 codons. The annotated chloroplast genome of *C. carlesii* has been deposited in GenBank with accession number MK745999.

Eight of the genus of *Castanopsis* and *Castanea* chloroplast genome sequences were aligned by MAFFT (Kazutaka et al. 2005). *Corylus heterophylla* was used as the outgroups. Phylogenetic analysis was computed based on the alignment sequences chloroplast genomes by maximum likelihood (ML) analysis by RAXML based on Kimura 2-parameter model with 1000 bootstrap replicates (Alexandros et al. 2017). As shown

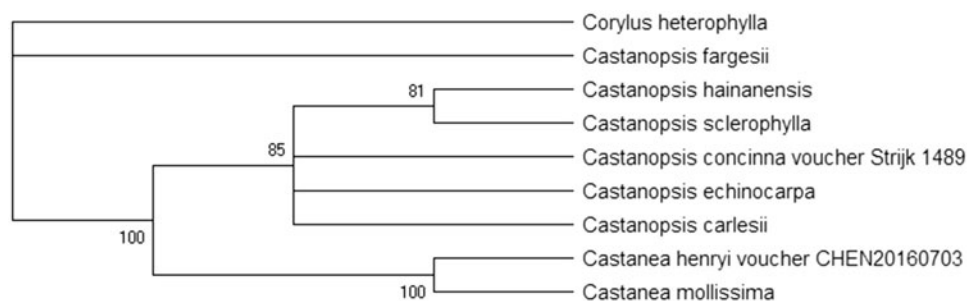


Figure 1. The ML phylogenetic tree based on nine complete chloroplast genomes. Accession number: *Corylus heterophylla* (KX822769), *Castanea henryi* voucher CHEN20160703 (KX954615), *Castanea mollissima* (HQ336406), *Castanopsis concinna* voucher Strijk 1489 (KT793041), *Castanopsis echinocarpa* (KJ001129), *Castanopsis hainanensis* (MG383644), *Castanopsis fargesii* (MK571045), *Castanopsis carlesii* (MK745999), and *Castanopsis sclerophylla* (MK387847). The number on each node indicates the bootstrap value.

in the ML phylogenetic tree (Figure 1), the genus *Castanea* formed a monophyletic clade with high bootstrap value, whereas, the genus *Castanopsis* formed a clade except *Castanopsis fargesii*. These results on complete chloroplast genome reported here will lay a basis for the study of phylogeny, phylogeography and population genetic diversity of *C. carlesii*.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Research Project of Jiangxi Provincial Department of Forestry under grant [No. 201801].

References

- Alexandros S, Paul H, Jacques R. 2017. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol.* 57:758.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45:e18.
- Kazutaka K, Kei-Ichi K, Hiroyuki T, Takashi M. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33:511–518.
- Lin K, Lyu M, Jiang M, Chen Y, Li Y, Chen G, Xie J, Yang Y. 2017. Improved allometric equations for estimating biomass of the three *Castanopsis carlesii* H. forest types in subtropical China. *New Forests.* 48:115–121.
- 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.* 13:715–715.
- Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet.* 52:267–274.