

## Characterization of the complete chloroplast genome of *Camellia granthamiana* (Theaceae), a Vulnerable species endemic to China

Weixi Li<sup>a</sup>, Xianggang Shi<sup>b</sup>, Wuxia Guo<sup>a</sup>, Achyut Kumar Banerjee<sup>a</sup>, Qun Zhang<sup>c</sup> and Yelin Huang<sup>a\*</sup>

<sup>a</sup>State Key Laboratory of Biocontrol and Guangdong Provincial Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China; <sup>b</sup>Experiment Teaching Centre of Marine Sciences, School of Marine Sciences, Sun Yat-sen University, Zhuhai, Guangdong, China; <sup>c</sup>Institute of Hydrobiology, Jinan University, Guangzhou, Guangdong, China

### ABSTRACT

*Camellia granthamiana* is a wild camellia resource endemic to China and is listed as a Vulnerable species globally. Here, we reported and characterized its complete chloroplast (cp) genome by using Illumina pair-end sequencing data. The total chloroplast genome size was 157,001 bp, including inverted repeats (IRs) of 26,042 bp, separated by a large single copy (LSC) and a small single copy (SSC) of 86,622 and 18,295 bp, respectively. A total of 131 genes, including 36 tRNA, 8 Rrna, and 87 protein-coding genes were identified. Phylogenetic analysis showed that *C. granthamiana* is sister to *C. sinensis* with 100% value support.

### ARTICLE HISTORY

Received 3 August 2018  
Accepted 31 August 2018

### KEYWORDS

*Camellia granthamiana*;  
Theaceae; complete  
chloroplast genome;  
Vulnerable species;  
phylogenomic analysis

*Camellia*, a genus of about 250 species belonging to the family Theaceae, is notable for having species with economic (e.g. *C. sinensis*) and ornamental (e.g. *C. japonica*) purposes (Gao et al. 2005; Vijayan et al. 2009). *Camellia granthamiana* is a rare wild camellia resources endemic to China with a restricted distribution in the narrow regions of Guangdong province of mainland China and Hong Kong (Ming and Bartholomew 2007). Due to its multiple large and persistent bracteoles and sepals, the species is considered as a primitive taxa (Chang 1981). Furthermore, the plant is polyploid and can be used for cross-breeding, which makes them highly appreciated by gardeners. In recent years, habitat degradation and loss have resulted in the decrease of the population size of the species leading to its consideration as a Vulnerable species globally (IUCN: <http://www.iucnr.edlist.org>) and is under protection in Country Parks in Hong Kong now. In this context, a better insight into its genomics may contribute to our understanding of conservation of this species, and to achieve this objective, we assembled a complete chloroplast genome of *C. granthamiana*.

Total DNA was isolated from fresh leaves of an individual of *C. granthamiana*, collected from the field in the Sun Yat-sen University Botanic Garden. The source population of this plant comes from Hong Kong. The voucher specimen (Shixg 171204) was deposited in the Sun Yat-sen University Herbarium (SYS). Genome sequencing was performed on an Illumina HiSeq X Ten platform with paired-end reads of

150 bp. Total 6.79 Gb short sequence data were obtained, which was used to assemble the chloroplast genome in NOVOPlasty (Dierckxsens et al. 2017). The chloroplast *atpB* sequence of *Entandrophragma excelsum* (GenBank accession number HQ158555) was used as the seed sequence. The genes in the chloroplast genome were annotated using the DOGMA program (Wyman et al. 2004). The circular chloroplast genome map was drawn using OGDRAW (Lohse et al. 2007).

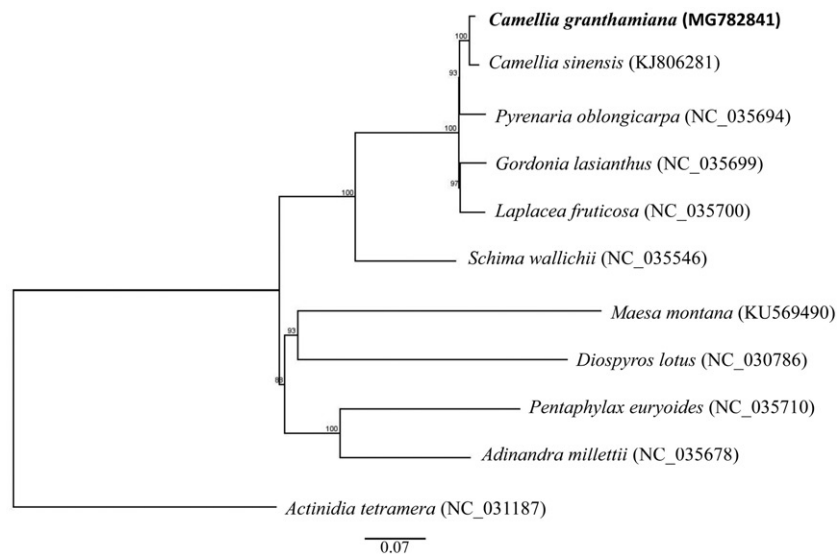
The complete cp genome of *C. granthamiana* (GeneBank accession number MG782842) was 157,001 bp in size, containing a pair of inverted repeats (IRs) of 26,042 bp, which separated a large single copy region (LSC) of 86,622 bp and a small single copy region (SSC) of 18,295 bp. The cp genome contained 131 genes, including 87 protein-coding genes, 36 transfer RNA genes, and 8 ribosomal RNA genes. Most of the genes occurred as a single-copy in the LSC or SSC, while 19 genes had two copies in the IRs. The overall GC content of the cp genome was 37.3%.

To perform a phylogenomic analysis, 11 complete chloroplast genomes within the order of Ericales were considered along with *C. granthamiana*. One species from Ebenaceae (*Diospyros lotus*) was set as the outgroup species. The chloroplast genome sequences were aligned using MAFFT (Katoh and Standley 2013). Phylogenetic analysis using the maximum likelihood algorithm was conducted with RAxML (Stamatakis 2014) implemented in Geneious ver. 10.1 ([\*\*CONTACT\*\* Qun Zhang !\[\]\(56549452e01ca28bdf2500ced9653143\_img.jpg\) \[tqzhang@jnu.edu.cn\]\(mailto:tqzhang@jnu.edu.cn\) !\[\]\(235f8f87c36d896db1ddff2848125c86\_img.jpg\) Institute of Hydrobiology, Jinan University, Guangzhou, Guangdong, China; Yelin Huang !\[\]\(4c4afa2dcab95991c60f31fce713d641\_img.jpg\) \[lsshyl@mail.sysu.edu.cn\]\(mailto:lsshyl@mail.sysu.edu.cn\) !\[\]\(1205f6f8a2eee5164f325ab7908828d5\_img.jpg\) State Key Laboratory of Biocontrol and Guangdong Provincial Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-sen University, Guangzhou, Guangdong, China](http://</a></p></div><div data-bbox=)

\*Present address: School of Life Sciences, Sun Yat-sen University, 135 Xingang West Road, 510275 Guangzhou, China. TEL: +86-20-84111164; FAX: +86-20-34022356.

© 2018 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.** Maximum-likelihood tree based on the sequences of 11 complete chloroplast genomes. Numbers in the nodes are bootstrap support values from 1000 replicates. The position of *Camellia granthamiana* is shown in bold and GenBank accession numbers are listed behind each species name.

[www.geneious.com](http://www.geneious.com), Kearse et al. 2012). The result showed that all Theaceae species are clustered into a monophyletic group and *C. granthamiana* is sister to *C. sinensis* with 100% value support (Figure 1).

### Disclosure statement

The authors declare that there is no conflict of interest regarding the publication of this article. The authors alone are responsible for the content and writing of the paper.

### Funding

This project was supported by grants from the National Natural Science Foundation of China [Grant Nos. 41776166 and 31700178], the Special Fund for Science and Technology Development of Guangdong Province [2017A030303014], [Grant No 2017A030313159], the Science and Technology Program of Guangzhou [Grant No. 201707020035] and the Chang Hungta Science Foundation of Sun Yat-sen University.

### References

Chang HT. 1981. A taxonomy of the genus *Camellia*. Acta Sci Nat Univ Sunyatseni, Monographic Series. 1:1–180.

- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45:e18.
- Gao J, Parks CR, Du YQ. 2005. Collected species of the genus *Camellia* and illustrated outline. Zhejiang: Zhejiang Science and Technology Press; p. 1–302.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30:772–780.
- 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: 1647–1649.
- 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.
- Ming TL, Bartholomew B. 2007. Theaceae. In: Wu ZY and Raven PH, editors. *Flora of China*. Vol.12. St. Louis: Science Press, Beijing and Missouri Botanical Garden Press.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30:1312–1313.
- Vijayan K, Zhang WJ, Tsou CH. 2009. Molecular taxonomy of *Camellia* (Theaceae) inferred from nrITS sequences. *Am J Bot.* 96:1348–1360.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organelle genomes with DOGMA. *Bioinformatics.* 20:3252–3255.