

Complete chloroplast genomes of five classical Wuyi tea varieties (*Camellia sinensis*, Synonym: *Thea bohea* L.), the most famous Oolong tea in China

Li Fan*, Li Li*, Yunfei Hu*, Yibiao Huang, Yongcong Hong and Bo Zhang

College of Tea and Food Science, Wuyi University, Wuyishan, People's Republic of China

ABSTRACT

Wuyi tea (*Camellia sinensis*, Synonym: *Thea bohea* L.) is recognized as the most prestigious oolong tea in China. For germplasm identification and protection, the complete chloroplast genomes of five classical Wuyi tea varieties were determined by next-generation sequencing. These chloroplast genomes showed highly conserved structures and are 157,024–157,126 bp in length, consisting of a pair of reverse repeats (IR) regions of 25,944–26,095 bp, one large single-copy (LSC) region of 86,594–86,859 bp, and one small single copy (SSC) region of 18,276–18,291 bp. A total of 137 genes were observed and overall GC contents were all about 37.3%. Phylogenetic analysis revealed Wuyi tea varieties did not cluster together, suggesting that these Wuyi tea varieties might have diverged early in their evolutionary history and the complete chloroplast genome could be used as a super-barcode to identify these varieties. This study will be valuable for future studies of evolution and intraspecific identification in Wuyi tea.

ARTICLE HISTORY

Received 26 September 2021
Accepted 29 March 2022

KEYWORDS

Wuyi tea; chloroplast genome; phylogenetic analysis; super-barcode

Wuyi tea belongs to *Camellia sinensis* species. It usually named after the English transliteration (Bohea) of Wuyi, and was named *Bohea* or *Var Bohea* (*Thea bohea* Linnaeus, 1762, GBIF Secretariat 2019). Wuyi tea was cultivated in rocky soil of Wuyi Mountain, where is a UNESCO World Heritage site in northern China's Fujian province and considered as the birthplace of Oolong tea. Recognized as the most prestigious oolong tea in China, Wuyi tea is renowned for its rich flavor and long-lasting fragrance, so-called 'rock charm and floral fragrance' (Lou et al. 2017; Chen et al. 2018). Wuyi tea has a long history of more than 1500 years, and there were thousands of varieties in history (Xiao 2017). Nowadays, due to the lack of effective protection, a large number of historical Wuyi tea varieties have been lost or difficult to find, causing irreversible losses. Therefore, it is urgent to protect and improve the germplasm resources of Wuyi tea.

The chloroplast (cp) genome can provide valuable information for phylogenetic studies and species identification (Bi et al. 2018). So far, only one cp genome of Wuyi tea has been reported (Li et al. 2021), which is not enough to provide effective molecular information to distinguish more rare varieties of Wuyi tea. Here, we present a complete assembly of cp genomes of five representative Wuyi tea cultivars, including *C. sinensis* cv. *Baijiguan* (BJG), *C. sinensis* cv. *Rougui* (RG), *C. sinensis* cv. *Shuijingui* (SJG), *C. sinensis* cv. *Tieluohan* (TLH) and *C. sinensis* cv. *Bantianyao* (BTY), which were known as their long history and good quality.

In this study, the young leaves of five cultivars were collected for genomic DNA isolation using the CTAB method. All five cultivars were originally from Wuyi Mountain, Fujian Province (27°43'42.46" N, 118°0'14.40" E) and then asexually propagated and planted in the Tea Germplasm Resources Museum of Wuyi University (The voucher number: BJG, No. 20200318; RG, No. 20200320; SJG, No. 20200321; TLH, No. 20200322 and BTY, No. 20200323). The genomic DNA were preserved at –80 °C at the Key Laboratory of Tea germplasm Genetic Resources of Wuyi University.

The complete cp genome was sequenced using a combination of PacBio RS and Illumina sequencing platforms by Biozeron Co., Ltd. (Shanghai, China). For PacBio sequencing, the average coverage reached 263×, 667×, 2370×, 1939×, 1665× and 2060× sequencing depth for five cp genomes assembly (BJG, RG, SJG, TLH and BTY), respectively. For Illumina sequencing, the average coverage reached 39313×, 30796×, 32018×, 36436× and 34766× sequencing depth, respectively. Briefly, PacBio and Illumina reads were mapped against the published cp genomes of *C. sinensis* var. *sinensis* (Accession number: KJ806281; Huang et al. 2014) to filter out the cp reads and be de-novo assembled using CLC Genomics Workbench 11.0.1. Cp genes were annotated by cpGAVAS (Liu et al. 2012) and sequenced coordinates of annotated cp genes available on NCBI using BLAST search (Acland et al. 2014). The final annotated cp genome sequences were deposited to NCBI GenBank (Accession number: BJG,

CONTACT Yongcong Hong  10817788@qq.com; Bo Zhang  3157814624@qq.com  College of Tea and Food Science, Wuyi University, Wuyishan, People's Republic of China

*These authors have contributed equally to this work.

© 2022 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.

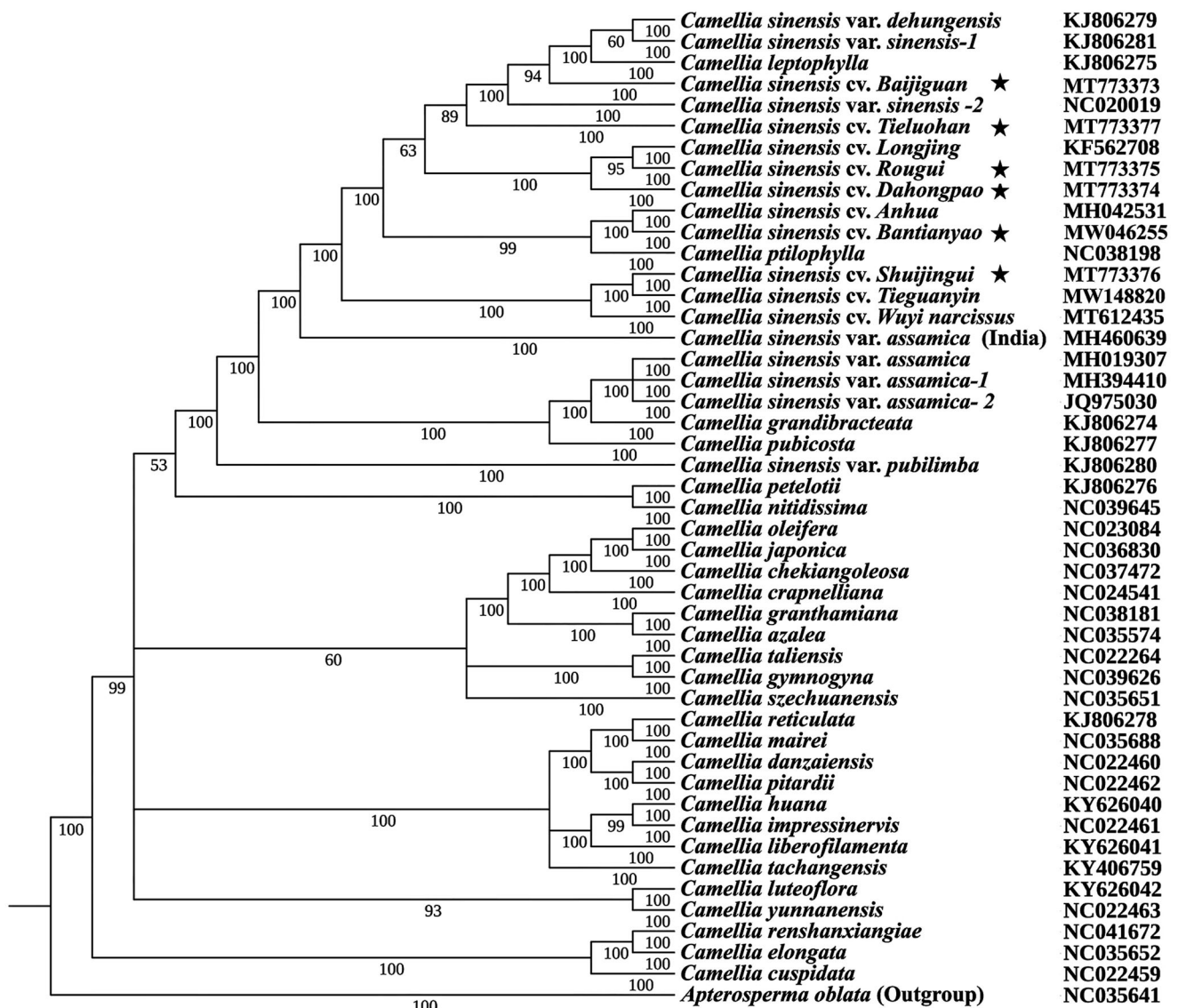


Figure 1. Phylogeny of *Camellia sinensis* based on the complete chloroplast genome. Numbers associated with nodes indicated Bayesian inference (BI) posterior probabilities values. The position of Wuyi tea varieties were marked with black asterisk and GenBank number was listed behind each variety name.

MT773373; RG, MT773375; SJG, MT773376; TLH, MT773377 and BTY, MW046255).

The five cp genome lengths ranged from 157,024 bp (BTY) to 157,126 bp (SJG). The cp genomes consisted of circular double-stranded DNA and showed a typical quadripartite structure, including an LSC region of 86,594–86,859 bp, an SSC region of 18,276–18,291 bp, and a pair of the IR regions of 25,944–26,095 bp. A total of 137 coding genes were identically annotated in the same order, consisting of 92 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Except for two introns in the *ycf3* and *clpP* genes, all other genes contained only one intron. Overall GC contents were all about 37.3%. To explore the phylogenetic positions and evolutionary relationships of Wuyi tea varieties, 46 *Camellia* species and an outgroup (*Apterospasma oblata*) were selected for the construction of Bayesian tree (Figure 1). The best-fitting models were determined by the GTR as implemented in Modeltest 3.7. Bayesian analyses were performed with four

independent chains running for two million generations, sampling a tree every 100 generations, the first 2500 trees were removed as burn-in and the remaining trees were used calculate Bayesian posterior probabilities. The result showed Wuyi tea cultivars did not cluster together, but intermingled with other *C. sinensis* var. *sinensis* varieties, reflecting the rich genetic diversity of Wuyi tea varieties. All these Wuyi tea cultivars could be effectively distinguished with high support, suggesting that Wuyi tea cultivars might have diverged early in their evolutionary history and the complete cp genome can be used as a super-barcode to identify them.

Ethical approval

The study is based on DNA sequence analysis and the samples used in this study were originally from Mount Wuyi and then asexually propagated and planted in the Tea

Germplasm Resources Museum of our university, and therefore we did not need any permission to use these samples.

Acknowledgement

We are very grateful to Mr. Luo Shengcai (Fujian Tea Tree Excellent Germplasm Resources Conservation Zone, China) for his guidance.

Disclosure statement

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

Funding

This study was financially supported by Scientific Research Launch Fund of Wuyi University [YJ201902], Collaborative education mechanism mode and practice based on new ecology of industrial College [XGK202001], Construction Fund of Wuyi Tea Industrial Technology Research Institute [2018N2004], Collection and innovative utilization of oolong tea and black tea germplasm resources across the Taiwan Strait [2019I0020], Construction project of First-class undergraduate major (tea Science) in Fujian [SJZY2019004] and Special Fund for Science and Technology Innovation of Fujian Zhang Tianfu Tea Development Foundation [FJZTF02].

CRedit authorship statement

Conception and design: **Fan Li**; Analysis and interpretation: **Li Li**, **Yunfei Hu** and **Yibiao Huang**; Drafting and revising: **Fan Li** and **Li Li**; Final approval: **Yongcong Hong** and **Bo Zhang**. All authors agree to be accountable for all aspects of the work.

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 NCBI accession number: MT773373 (BJG), MT773375 (RG), MT773376 (SJG), MT773377 (TLH) and MW046255 (BTY). The associated BioProject numbers are PRJNA644760 (BJG, RG, SJG, TLH) and PRJNA666051 (BTY). The associated SRA numbers are SRS7200201 (BJG), SRS7200203 (RG), SRS7200204 (SJG), SRS7200205 (TLH), and SRS9680346

(BTY). The associated Bio-Sample numbers are SAMN15804988 (BJG), SAMN15804990 (RG), SAMN15804991 (SJG), SAMN15804992 (TLH), and SAMN20564595 (BTY).

References

- Acland A, Agarwala R, Barrett T, Beck J, Benson D, Bollin C, Bolton E, Bryant S, Canese K, Church D, Clark K, DiCuccio M, Dondoshansky I, Federhen S, Feolo M, Geer L, Gorenkov V, Hoepfner M, Johnson M, Zbicz K. 2014. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 42:D7–D17. DOI: [10.1093/nar/gkt1146](https://doi.org/10.1093/nar/gkt1146).
- Bi Y, Zhang M-F, Xue J, Dong R, Du Y-P, Zhang X-H. 2018. Chloroplast genomic resources for phylogeny and DNA barcoding: a case study on *Fritillaria*. *Sci Rep.* 8(1):1184.
- Chen S, Li MH, Zheng GY, Wang TT, Lin J, Wang SS, Wang XX, Chao QL, Cao SX, Yang ZB, et al. 2018. Metabolite profiling of 14 Wuyi rock tea cultivars using UPLC-QTOF MS and UPLC-QqQ MS combined with chemometrics. *Molecules.* 23(2):104.
- GBIF Secretariat. 2019. *Thea bohea* L., 1762 in The International Plant Names Index Collaborators. International plant names index. Checklist dataset. <https://doi.org/10.15468/uhllmw> (accessed via GBIF.org on 2021-06-09).
- Huang H, Shi C, Liu Y, Mao S-Y, Gao L-Z. 2014. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships. *BMC Evol Biol.* 14: 151.
- Li L, Hu Y, Wu L, Chen R, Luo S. 2021. The complete chloroplast genome sequence of *Camellia sinensis* cv. Dahongpao: a most famous variety of Wuyi tea (Synonym: *Thea bohea* L.). *Mitochond DNA B Resour.* 6(1): 3–5.
- 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.
- Lou YX, Fu XS, Yu XP, Ye ZH, Cui HF, Zhang YF. 2017. Stable isotope ratio and elemental profile combined with support vector machine for provenance discrimination of Oolong Tea (Wuyi-Rock Tea). *J Anal Methods Chem.* 2017:1–8.
- Xiao KB. 2017. The taste of tea: Material, embodied knowledge and environmental history in northern Fujian, China. *J Mater Cult.* 22(1): 3–18.