

Complete chloroplast genome of *Callicarpa formosana* Rolfe, a famous ornamental plant and traditional medicinal herb

Yongxi Du^a, Yanfeng Liu^b, Bo Liu^b and Tielin Wang^a

^aNational Resource Center for Chinese Meteria Medica, China Academy of Chinese Medical Sciences, Beijing, P. R. China; ^bState Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, P. R. China

ABSTRACT

Callicarpa formosana is a species of beauty-berry with large medicinal value belonging to the family Verbenaceae. In this study, the complete chloroplast genome of *C. formosana* was sequenced using Illumina HiSeq X Ten platform. The chloroplast genome was 1,54,210 bp in length, containing two short inverted repeat (IRa and IRb) regions of 25,701 bp, which was separated by a large single copy (LSC) region of 84,938 bp and a small single copy (SSC) region of 17,870 bp. The GC content of the whole chloroplast genome was 38.1%. The chloroplast DNA of *C. formosana* comprised 113 genes, including 79 protein-coding genes, 4 ribosomal RNA genes, and 30 transfer RNA genes. Phylogenetic analysis indicated that the genus *Callicarpa* L. was located in the basal position within the family Verbenaceae. The chloroplast genome (cpDNA) of *C. formosana* was closely related to *Callicarpa nudiflora*.

ARTICLE HISTORY

Received 15 August 2020
Accepted 2 September 2020

KEYWORDS

Callicarpa formosana;
chloroplast genome;
phylogenetic relationship

Callicarpa L. is a genus of shrubs and small trees in the family Verbenaceae. In recent years, research on the medicinal value of *Callicarpa* plants has increased. Four species of *Callicarpa* were included in the Pharmacopeia of the People's Republic of China. *Callicarpa formosana* Rolfe is a species of beauty-berry. It is native to China (type specimens were collected from Taiwan), Japan, and Philippines (Wu et al. 1993–2013; Nakashima et al. 2016). The species is cultivated as an ornamental plant that is adaptable to various conditions of cultivation and climates. The globous drupes of resembling tiny clusters of berries are of lavender color. All parts of the plant are utilized in the traditional Chinese medicine for various pathologies (National Pharmacopeia Committee 2020). Chloroplast genomes are important sources for phylogenetic analyses, genetic diversity evaluation, and plant molecular identification (Dong et al. 2018; Sun et al. 2020). In this study, we determined the complete chloroplast genome (cpDNA) sequence of *C. formosana* based on the next-generation sequencing method. The annotated cpDNA has been deposited into GenBank with the accession number MT830861.

Fresh samples of *C. formosana* were collected from Suining county, Hunan province, China (26°36'16"N, 110°8'18"E). Voucher specimen was deposited at the herbarium of Institute of Chinese Materia Medica (CMMI), China Academy of Chinese Medical Sciences with the specimen voucher number is 430527LY0206. Total genomic DNA from fresh leaves of a single individual was isolated using the method of Li et al. (2013). And the sequencing library was constructed and quantified following the methods introduced

by Dong et al. (2017). The whole genome sequencing was conducted with 150 bp paired-end reads on the Illumina HiSeq X Ten platform. Next-generation sequencing QC toolkit was used for quality control and to filter the low quality reads. Contigs were assembled from the high quality paired-end reads by using the SPAdes version 3.6.1 program (Kmer = 95) (Bankevich et al. 2012). The chloroplast genome contigs selected by the Blast program (Altschul et al. 1990), taken *Callicarpa nudiflora* (GenBank: MK783316) as the reference. The selected contigs were assembled using Sequencher version 4.10 (Gene Codes Corporation, Ann Arbor, MI USA, <http://www.genecodes.com>). Gene annotation of *C. formosana* was performed using DOGMA annotation (Wyman et al. 2004) and manually corrected for codons and gene boundaries using BLAST searches.

The circular cpDNA of *C. formosana* was 1,54,210 bp in length, containing two short inverted repeat (IRa and IRb) regions of 25,701 bp, which was separated by a large single copy (LSC) region of 84,938 bp and a small single copy (SSC) region of 17,870 bp. The GC content of the whole chloroplast genome was 38.1%. The cpDNA of *C. formosana* comprised 113 distinct genes, including 79 protein-coding genes, 4 ribosomal RNA genes, and 30 transfer RNA genes. In these genes, 19 were duplicated in the IR regions and 19 genes contained one or two introns. 17 harbored a single intron, and two (*ycf3* + *clpP*) contained double introns.

Eighteen chloroplast genome sequences were used for phylogenetic analysis to confirm the location of *C. formosana*, including four outgroup samples and 14 samples of

CONTACT Tielin Wang  wtl82@163.com  National Resource Center for Chinese Meteria Medica, China Academy of Chinese Medical Sciences, Beijing, P. R. China

© 2020 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.

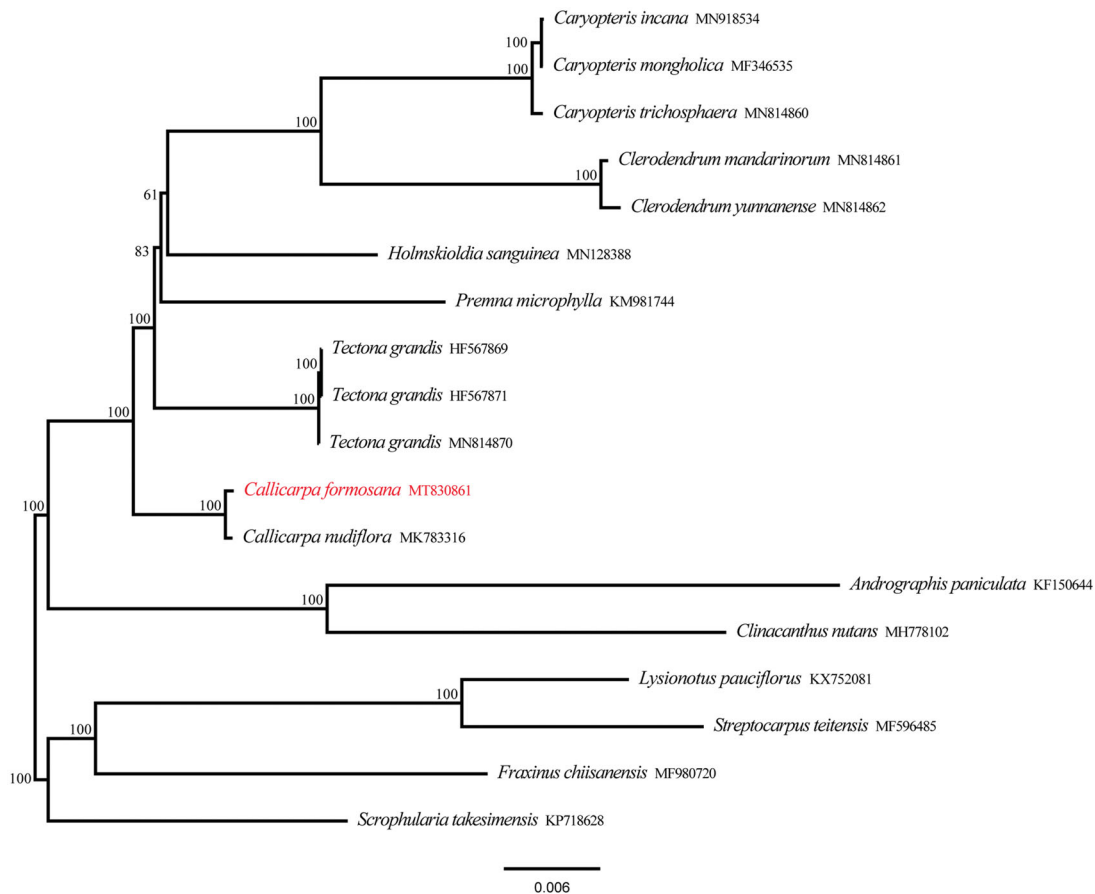


Figure 1. Phylogenetic tree reconstruction of 18 taxa using maximum likelihood (ML) methods based on protein-coding genes in the chloroplast genome sequences. ML bootstrap support value presented at each node.

Verbenaceae from the GenBank. We used 79 protein-coding genes to conduct a maximum likelihood (ML) analysis using IQ-tree under the GTR + G model with 1000 bootstrap replicates (Nguyen et al. 2015; Zhang et al. 2020). The phylogenetic analysis revealed that samples of Verbenaceae were strongly supported as monophyletic, the genus *Callicarpa* was located in the basal position additionally (Figure 1). The cpDNA of *C. formosana* is closely related to *C. nudiflora*. The complete chloroplast genome reported in this study will be a valuable resource for future studies on genetic diversity, taxonomy, and phylogeny of family Verbenaceae.

Disclosure statement

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

Funding

This work was supported by the Key Project at Central Government Level: the ability establishment of sustainable use for valuable Chinese medicine resources [2060302] and the Fundamental Research Funds for the Central Public Welfare Research Institutes [ZZXT202003].

Data availability statement

The data that support the findings of this study are openly available in GenBank of NCBI <https://www.ncbi.nlm.nih.gov/>, reference number MT830861, raw data submissionID: SUB7932473, BioProject ID: PRJNA657909.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215(3):403–410.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Dong W, Xu C, Li W, Xie X, Lu Y, Liu Y, Jin X, Suo Z. 2017. Phylogenetic resolution in Juglans based on complete chloroplast genomes and nuclear DNA sequences. *Front Plant Sci.* 8:1148.
- Dong W, Xu C, Wu P, Cheng T, Yu J, Zhou S, Hong DY. 2018. Resolving the systematic positions of enigmatic taxa: manipulating the chloroplast genome data of Saxifragales. *Mol Phylogenet Evol.* 126:321–330.
- Li JL, Wang S, Jing Y, Wang L, Zhou SL. 2013. A modified CTAB protocol for plant DNA extraction. *Chin Bull Bot.* 48:72–78.
- Nakashima S, Sarath E, Okada H, Ezaki K, Darnaedi D, Tsukaya H, Soejima A. 2016. Morphological and phylogenetic investigations for several cryptic ant-plants found in *Callicarpa* (Lamiaceae) from Borneo. *J Plant Res.* 129(4):591–601.
- Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.
- Sun J, Wang Y, Liu Y, Xu C, Yuan Q, Guo L, Huang L. 2020. Evolutionary and phylogenetic aspects of the chloroplast genome of *Chaenomeles* species. *Sci Rep.* 10(1):11466.
- Wu ZY, Raven PH, Hong DY. 1994 (17). *Flora of China.*
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20(17):3252–3255.
- Zhang D, Gao F, Jakovlic I, Zou H, Zhang J, Li WX, Wang GT. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour.* 20(1):348–355.