


Characterization of the complete plastid genome of *Gaultheria griffithiana* (Ericaceae)

Yi-Rong Li^a, Yan-Ling Xu^a, Xin-Yu Du^b , Shu-Da Yang^a and Lu Lu^a 

^aSchool of Pharmaceutical Sciences and Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Kunming, China; ^bPlant Germplasm and Genomics Center, Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China

ABSTRACT

Gaultheria griffithiana is an evergreen shrub in the family Ericaceae. It is used as a source of the Chinese traditional medicine, Tougucao, with distribution of the junction of eastern Himalaya and Hengduan Mountain. The chloroplast genome of *G. griffithiana* is 175,649 bp in length with 135 genes, including eight rRNA genes, 39 tRNA genes, and 85 protein-coding genes. Phylogenetic analysis has converged on the placement of *G. griffithiana* as sister to *G. praticola*, *G. nummularioides*, and *G. hookeri* within the Leucothoides clade of *Gaultheria* in this study.

ARTICLE HISTORY

Received 28 January 2021
Accepted 3 April 2021

KEYWORDS

Chloroplast genome;
Ericaceae; *Gaultheria griffithiana*; Leucothoides



Gaultheria griffithiana Wight belongs to the Leucothoides clade of *Gaultheria* of the tribe Gaultherieae within Ericaceae (Lu et al. 2019), which has been indigenously of use to treat rheumatism and anti-inflammatories due to containing benzoic acid derivatives, anthraquinones and alkaloids (Liu et al. 2013). As an evergreen shrub, *G. griffithiana* has a unique geographical distribution in the junction of eastern Himalaya and Hengduan Mountain including northeast India, Indochina (Myanmar, Vietnam), and southwest China, with elevations ranging from 2000 m to 3600 m (Fang and Stevens 2005; Fritsch et al. 2008). Based on the combined data of multiple genes from ITS and plastid loci, *G. griffithiana* was not monophyletic in the Leucothoides clade and its phylogenetic position was unresolved in the work of Lu et al. (2010). However, this species was subsequently resolved as a sister to most species such as *G. nummularioides* D. Don, *G. praticola* C. Y. Wu, and *G. hookeri* C. B. Clarke from the Leucothoides clade in the work of Fritsch et al. (2011).

The plastid genome data has been proved to apparently well improve the phylogenetic resolution of species within *Gaultheria* series *Trichophyllae* from distinct subclade within the Leucothoides clade (Zhang et al. 2017). Herein, we characterized the complete plastid genome sequence of *G. griffithiana* using Illumina sequencing data, for the purposes to understanding the phylogenetic position of the species, further evolutionary and pharmacological studies.

The leaf sample of *G. griffithiana* was collected from Cangshan Mountain of Dali City (25°52'12"N, 100°1'12"E) in the southwest of Yunnan Province, China. The voucher specimen (number: KUN1248996) was deposited in the herbarium at Kunming Institute of Botany (KUN). Total genomic DNA

was extracted with the CTAB (Cetyltrimethyl Ammonium Bromide) protocol (Doyle and Doyle 1987), followed by insert size of 150 bp paired-end sequencing using Illumina HiSeq X-Ten Sequencing System (the Molecular Biology Experiment Center, the Germplasm Bank of Wild Species in Southwest China) with standard Illumina sequencing protocols (Shendure and Ji 2008). The complete chloroplast genome was assembled using SPAdes v3.10.1 (Bankevich et al. 2012) and improved by the GetOrganelle pipeline (Jin et al. 2018) with *Vaccinium macrocarpon* Aiton (GenBank accession: NC019616.1) as reference, and then annotated using the Geneious R8 (Kearse et al. 2012). The phylogenetic tree of 19 species selected within Ericaceae was reconstructed based on complete chloroplast genome sequences which were downloaded from NCBI GenBank, with full-length sequences aligned by MAFFT v7 software (Katoh and Standley 2013). We used maximum likelihood analysis (option '-f a') with 1000 rapid bootstrap replicates under GTRGAMMA model on Cipres Science Gateway (Miller et al. 2015, available at www.phylo.org).

The complete chloroplast genome sequence of *G. griffithiana* is 175,649 bp in length with a large single-copy region (LSC) of 107,114 bp, a small single-copy region (SSC) of 3,693 bp, and two inverted repeat regions (IRa and IRb) of 32,421 bp each. Furthermore, the genome encodes 135 genes including 85 protein-coding genes, 39 tRNAs, and eight rRNAs. The overall GC content of the chloroplast genome is 36.6%, and the corresponding GC values in LSC, SSC, and IR regions are 35.6%, 27.8% and 38.7%, respectively. Annotated chloroplast genome sequence was submitted to GenBank with an accession number MW528025. The phylogenetic

CONTACT Lu Lu  lulukmu@163.com  School of Pharmaceutical Sciences and Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Kunming 650500, China

© 2021 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.

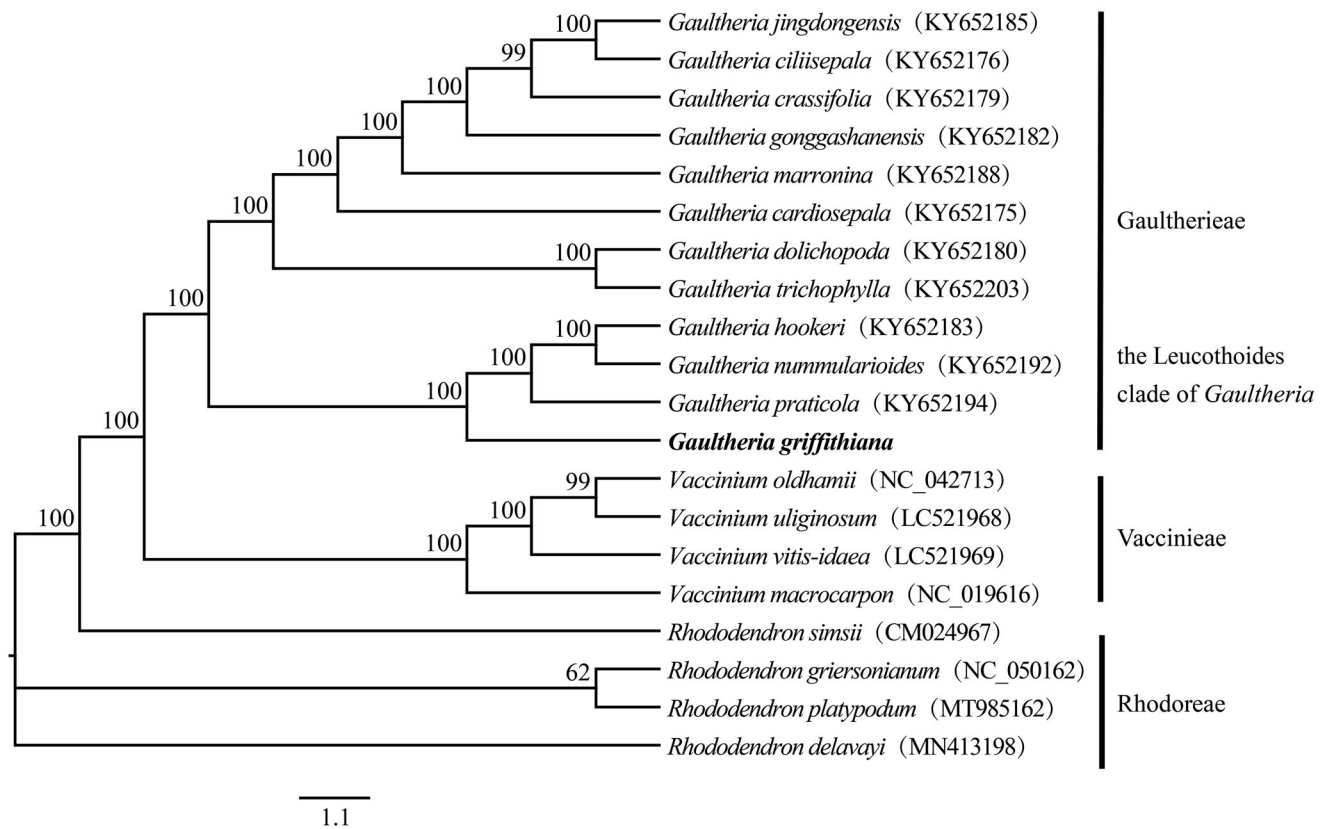


Figure 1. Phylogenetic tree based on 20 complete chloroplast genome sequences of Ericaceae. The bootstrap values are presented above each branch (only those > 50% are shown).

analysis has converged on the placement of *G. griffithiana* as sister to *G. praticola*, *G. nummularioides*, and *G. hookeri* within the Leucothoides clade of *Gaultheria* in this study (Figure 1), which was consistent with the topology based on ML analysis of combined chloroplast, and chloroplast plus nuclear sequence data in Fritsch et al. (2011). This provides additional rationale for using chloroplast data when reconstructing species relationships in *Gaultheria*.

Disclosure statement

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

Funding

This research was financially supported by the National Natural Science Foundation of China [31960080], the Yunnan Applied Basic Research Projects-Joint Special Project [202001AY070001-018], and the Reserve talents of young and middle-aged academic and technical leaders in Yunnan province [202005AC160020].

ORCID

Xin-Yu Du  <http://orcid.org/0000-0002-5721-7812>

Lu Lu  <http://orcid.org/0000-0002-7276-5358>

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. MW528025. The associated BioProject, SRA, and

Bio-Sample numbers are PRJNA703716, SRX10145200, and SAMN18022285, respectively.

References

- 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.
- Doyle JJ, Doyle JD. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19:11–15.
- Fang RZ, Stevens PF. 2005. *Gaultheria*. In: Wu ZY, Raven PH, Hong DY, editors. *Flora of China, Apiaceae through Ericaceae*, vol. 14. Beijing/St. Louis: Science Press/Missouri Botanical Garden Press; p. 464–475.
- Fritsch PW, Lu L, Bush CM, Cruz BC, Kron KA, Li DZ. 2011. Phylogenetic analysis of the wintergreen group (Ericaceae) based on six genic regions. *Syst Bot.* 36 (4):990–1003.
- Fritsch PW, Zhou L, Bartholomew B. 2008. The flowering plant genus *Gaultheria* (Ericaceae) in the Gaoligong Shan, along the border region of China and Myanmar: California Academy of Sciences. *Proc Calif Acad Sci Ser.* 4 (59):147–214.
- Jin J-J, Yu WB, Yang JB, Song Y, Yi TS, Li DZ. 2018. GetOrganelle: a simple and fast pipeline for de novo assembly of a complete circular chloroplast genome using genome skimming data. *BioRxiv.* 256479.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):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(12):1647–1649.
- Liu WR, Qiao WL, Liu ZZ, Wang XH, Jiang R, Li SY, Shi RB, She GM. 2013. *Gaultheria*: phytochemical and pharmacological characteristics. *Molecules.* 18(10):12071–12108.

- Lu L, Fritsch PW, Cruz BC, Wang H, Li DZ. 2010. Reticulate evolution, cryptic species, and character convergence in the core East Asian clade of *Gaultheria* (Ericaceae). *Mol Phylogenet Evol.* 57(1):364–379.
- Lu L, Fritsch PW, Matzke NJ, Wang H, Kron KA, Li DZ, Wiens JJ. 2019. Why is fruit colour so variable? Phylogenetic analyses reveal relationships between fruit-colour evolution, biogeography and diversification. *Global Ecol Biogeogr.* 28(7):891–903.
- Miller MA, Schwartz T, Hoover P, Yoshimoto K, Sivagnanam S, Majumdar A. 2015. The CIPRES workbench: a flexible framework for creating science gateways. In *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure*.
- Shendure J, Ji H. 2008. Next-generation DNA sequencing. *Nat Biotechnol.* 26(10):1135–1145.
- Zhang MY, Fritsch PW, Ma PF, Wang H, Lu L, Li DZ. 2017. Plastid phylogenomics and adaptive evolution of *Gaultheria* series *Trichophyllae* (Ericaceae), a clade from sky islands of the Himalaya-Hengduan Mountains. *Mol Phylogenet Evol.* 110:7–18.