


The complete chloroplast genome sequence of *begonia* (Begoniaceae)

Shikai Guan^a, Qian Song^a, Jinye Zhou^a, Yun Liu^b, Haixia Yan^a, Dayan Tao^a  and Shuming Luo^{a,c}

^aFlower Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China; ^bHorticultural Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China; ^cPlant Breeding Institute, Faculty of Agriculture and Environment, University of Sydney, Cobbitty, Australia

ABSTRACT

Begonia L. (Begoniaceae) is the sixth largest genus in the world which consists of more than 2,039 species. Many species of *Begonia* have highly ornamental leaves and flowers, so they are mainly used for ornamental purposes, and some species can also be used as medicines or vegetables. *Begonia gulongshanensis* is a newly discovered species in 2018 which occurs exclusively in Jingxi county in Southern China, however, there are few studies on the molecular biology and phylogeny of this species currently. Therefore, we report its complete chloroplast genome sequence for the first time, hoping to provide a foundation for its future phylogenetic analysis. The chloroplast genome of *B. gulongshanensis* was 169,153 bp in size, which contained a large single-copy region of 75,998 bp, a small single-copy region of 18,063 bp, and two inverted repeat regions of the same 37,546 bp. The total GC content was 35.51%. The genome encodes 42 transfer RNA genes, 8 ribosomal RNA genes and 90 protein-coding genes. The phylogenetic analysis indicated that the genetic relationship between *B. gulongshanensis* and the other three *Begonia* species was very close, but there was still certain distance.

ARTICLE HISTORY

Received 30 September 2021
Accepted 12 December 2021

KEYWORDS

Begonia gulongshanensis;
chloroplast genome;
genome assembly;
phylogenetic tree

Begonia L. (Begoniaceae) was mainly distributed in tropical and subtropical areas (Tian et al. 2018), and 2,039 species in *Begonia* L. have been published (Hughes et al. 2015–Present). As the sixth largest genus in the world, it contains tuberous, semi-tuberous, moniliform, rhizomatous, trailing (scandent), and shrub-like species, and there is a huge diversity in habitat, habit, plant size, leaf type, flower and fruit morphology between different species (Tian et al. 2018). Many species of *Begonia* have highly ornamental leaves and flowers, so they are mainly used for ornamental purposes, and some species can also be used as medicine or vegetable (Guan et al. 2007). In the last 20 years, many new species have been discovered because of increasing attention from people (Tian et al. 2018; Dong and Liu 2019; Van et al. 2021) predicted that the number of species of *Begonia* may actually in total between 2000 and 2500. However, the classification of species in *Begonia* is still controversial, due to the traditional classification method is mainly based on morphological observations. In recent years, taxonomists have solved some taxonomic problems through a combination of morphology and molecular technology, but there are still a large number of species that lack molecular data, and there is still a lot of work remains to be done.

Begonia gulongshanensis is a newly reported species in 2018 which occurs exclusively in Jingxi county of Guangxi province in Southern China (Chen et al. 2018). It is only

grows on the moist surface of shady cliffs at the entrance to shallow caves in a deep valley, with a white area on the leaves (Chen et al. 2018). Due to the rare distribution in the wild, *B. gulongshanensis* was listed as class II protected plant in the List of National Key Protected Wild Plants in September 2021 (National Forestry and Grassland Administration & Ministry of Agriculture and Rural Affairs of the People's Republic of China 2021). So far, there have been few studies on the molecular and phylogeny of this species. The purpose of this study was to complete the sequencing of the chloroplast genome of *B. gulongshanensis* and study its phylogeny, so as to provide reference for the future research on the genetic relationship between it and other *Begonia* species, and provide a foundation for its conservation, exploitation and utilization, and to provide reference for subsequent intraspecific classification of *Begonia*.

We collected the leaf samples of *B. gulongshanensis* from Jingxi county of Guangxi province (106°39'30"E, 23°0'35"N, 550 m above sea level). The collected reference samples were deposited in the Flower Research Institute, Guangxi Academy of Agricultural Sciences (<http://www.gxaas.net/>), and the voucher number was GXAAS-B00158 (collected by Jinye Zhou, ahzhoujy@163.com). The Illumina NovaSeq 6000 platform was used to paired-end sequencing of extracted genomic DNA, and the sequencing read length was 150 bp. Then, we assembled the chloroplast genome by SPAdes v3.10.1

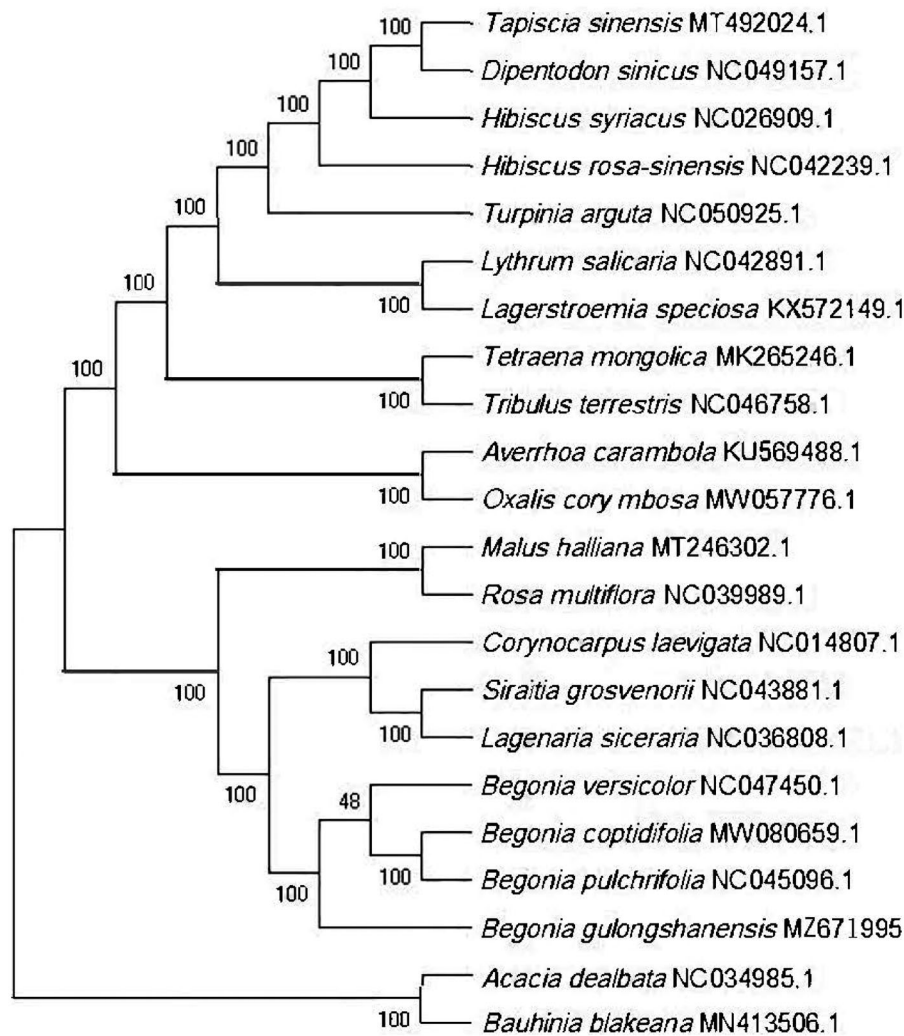


Figure 1. The phylogenetic tree based on the complete chloroplast genome of *B. gulongshanensis* and 21 other published species. The bootstrap support values are displayed next to the nodes.

(Bankevich et al. 2012), finally, annotated the chloroplast genome *B. gulongshanensis* by CPGAVAS 2 (Shi et al. 2019). We submitted the assembled chloroplast genome and the annotations to GenBank, and obtained the accession number MZ671995. In order to construct the phylogenetic tree of *B. gulongshanensis*, another 21 species (3 species from *Begonia* L., and 18 species from other families of Rosanae) were compared and analysed by via MAFFT v.7.271 (Kato et al. 2005) and RAxML v8.0 (Stamatakis 2014).

The assembled chloroplast genome of *B. gulongshanensis* was 169,153 bp in size, with a typical quarter structure which contained a large single-copy region (LSC) of 75,998 bp, a small single-copy region (SSC) of 18,063 bp, and two inverted repeat regions (IRA and IRB) of the same 37,546 bp. The total GC content is 35.51%. The annotated results showed that there were 140 genes in *B. gulongshanensis* chloroplast genome, including 42 transfer RNA genes (tRNAs), 8 ribosomal RNA genes (rRNAs) and 90 protein-coding genes (PCGs). The phylogenetic analysis showed that Begoniaceae was closest to Cucurbitaceae and Corynocarpaceae, which also belong to Cucurbitales, followed by Rosales, and had the most distant with Fabales. Although *B. gulongshanensis* was clustered with

the other three begonias, there was still a certain distance between them. We speculated that this was because the *B. gulongshanensis* belonged to the section *Coelocentrum* while the other three begonias belonged to the section *Platycentrum* (Figure 1).

Acknowledgment

The authors thank APEX BIO Co., Ltd (Shanghai, China) for assistance in sequencing.

Disclosure statement

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

Funding

This work was supported by the Guangxi Natural Science Foundation under Grant [No. 2020GXNSFAA259064], [No. 2020GXNSFAA297068], [No. 2020GXNSFBA297074]; and the Research and Development Fund of Guangxi Academy of Agricultural Sciences under Grant [GUINONGKE-2021YT135], [GUINONGKE-2021YM157], [NKYCZ202107].

ORCID

Dayan Tao  <http://orcid.org/0000-0002-5840-2307>

Data availability statement

The complete sequence of the chloroplast genome of *B. gulongshanensis* have been deposited in GenBank at <https://www.ncbi.nlm.nih.gov/>, and the accession no. was MZ671995. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA770846, SRR16311536, and SAMN22241481, respectively.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Chen WH, Radbouchoom S, Nguyen HQ, Nguyen HT, Nguyen KS, Shui YM. 2018. Seven new species of *Begonia* (Begoniaceae) in Northern Vietnam and Southern China. *PK.* 94:65–85.
- Dong LN, Liu Y. 2019. Supplement to *Begonia* L. in Flora of Guangxi. *Guihaia.* 39(1):16–39.
- Guan KY, Yamaguchi H, Li JX, Li HZ, Ma H. 2007. Traditional uses of begonias (Begoniaceae) in China. *Acta Bot Yunnanica.* 29 (1), 58e66.
- Hughes M, Moonlight PW, Jara-Muñoz A, Tebbitt MC, Wilson HP, Pullan M. 2015–Present. *Begonia* resource centre. [Accessed 10 September 2021]. Available from: <http://padme.rbge.org.uk/begonia/>
- Katoh K, Kuma K-I, Toh H, Miyata T. 2005. MAFFT version 5: Improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33(2): 511–518.
- National Forestry and Grassland Administration & Ministry of Agriculture and Rural Affairs of the People's Republic of China 2021. List of national key protected wild plants. <http://www.forestry.gov.cn/main/3954/20210908/163949170374051.html>
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9): 1312–1313.
- Tian DK, Xiao Y, Tong Y, Fu NF, Liu QQ, Li C. 2018. Diversity and conservation of Chinese wild begonias. *Plant Divers.* 40(3):75–90.
- Van TP, Quang DD, Van CN, Ngoc BT, Duc DN, The HN, Che WL. 2021. *Begonia* of Vietnam: an updated checklist, including a new species and a new record. *Phytotaxa.* 507(2):144–154.