

The complete chloroplast genome sequence of monotypic *Cyphotheca* (Melastomataceae), an endemic genus in China

Zhenying Wen^a, Sijin Zeng^{b,c}, Tingzhang Li^b, Guoqiang Zhang^b and Donghui Peng^a

^aCollege of Landscape Architecture, Fujian Agriculture and Forestry University, Fuzhou, China; ^bShenzhen Key Laboratory for Orchid Conservation and Utilization, The National Orchid Conservation Center of China and The Orchid Conservation and Research Center of Shenzhen, Shenzhen, China; ^cCollege of Forestry and Landscape Architecture, South China Limestone Plants Research Center, South China Agricultural University, Guangzhou, China

ABSTRACT

Cyphotheca, a shrub narrowly endemic to Yunnan, southwest China, is a monotypic genus of Sonerileae in Melastomataceae. The species, *Cyphotheca montana*, is dwindling and is listed as 'Near Threatened (NT)' by IUCN Red List of China. Herein, we report the complete chloroplast genome (cpDNA) sequence assembled from Illumina pair-end sequencing data. The complete chloroplast genome was 156,424 bp in length, which includes two inverted repeat regions (IRs) of 26,776 bp each, which were separated by a large single copy region (LSC) 85,900 bp and a small single copy region (SSC) 16,972 bp. The chloroplast genome contained 130 genes, including 85 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. The overall GC content in the chloroplast genome of *C. montana* was 37.0%. The phylogenomic analysis showed that *C. montana* is sister to *Barthea barthei*.

ARTICLE HISTORY

Received 26 April 2019
Accepted 17 May 2019

KEYWORDS

Cyphotheca montana; Melastomataceae; chloroplast genome; Illumina sequencing; phylogenetic analysis

Cyphotheca (Melastomataceae: Sonerileae), a shrub originally described by Diels (1932) based on *Cyphotheca montana* Diels, is a monotypic genus. It is narrowly endemic to south and southwestern Yunnan, China (Wu 1979; Chen 1984; Chen and Renner 2007), growing only in dense broad-leaved forests, bamboo forests, mountain slopes, valleys, hillsides, stream banks. It is characterized by having dense stellate trichomes in inflorescences, pedicel and bracts, funnellform-campanulate hypanthium, obtusely 4-sided and 8-ribbed, orbicular calyx lobes, winged on ridges, white to pink petals, 8 stamens, curved and shortly beaked, unequal whorls in length, inflated at middle connective. Because of its rarity, *C. montana* was categorized as 'Near Threatened (NT)' in the latest version (2013) of IUCN Red List of China (<http://rep.iplant.cn/protolist/4>). The previous study on the chemical constituents indicated that the ethyl acetate and *n*-butanol extracts obtained from the alcohol syrup of the aerial parts of *C. montana* had antibacteria activity (Luo et al. 2003). Its systematic position was first evaluated by Zhou et al. (2018) and showed a close relationship to *Phyllagathis fengii*. To better understand the evolution history and genetic information of *C. montana*, the complete chloroplast genome sequence was characterized.

The leaf sample was acquired from Jinping, Yunnan, China, and a voucher specimen (specimen code SJ361) is deposited at the South China Agricultural University Herbarium (CANT). Total genomic DNA was extracted from

young and healthy leaves using a modified CTAB method (Doyle and Doyle 1987) and sequenced based on the Illumina pair-end technology. The clean reads were firstly aligned to *Opisthocentra clidemioides* (GenBank accession KX826828) and then assembled into contigs in the software CLC Genomics Workbench v8.0 (CLC Bio, Aarhus, Denmark). The assembled chloroplast genome was annotated using DOGMA, and the annotation was corrected using Geneious (Kearse et al. 2012). The physical map of the *C. montana* chloroplast genome was drawn using OGDRAW (Lohse et al. 2013). The accurate new annotated complete chloroplast genome was submitted to GenBank with accession number MK814185. The complete chloroplast genome of *C. montana* was a circular molecule of 156,424 bp long, containing a large single-copy (LSC) region of 85,900 bp and a small single-copy (SSC) region of 16,972 bp, separated by a pair of inverted repeat regions of 26,776 bp each. It was predicted to contain 130 genes, including 85 protein-coding genes, 37 transfer RNA (tRNA) genes, and 8 ribosomal RNA (rRNA) genes. The overall GC content of the whole plastome was 37.0%.

For phylogenetic reconstruction, the chloroplast genome sequences of *C. montana* was aligned with 19 other complete chloroplast genomes of Melastomataceae (Reginato et al. 2016; He et al. 2017; Ng et al. 2017; Zhou, Ng, et al. 2018), and *Eucalyptus globulus* (GenBank accession AY780259) as an outgroup. The sequences were aligned using HomBlocks pipeline (Bi et al. 2018). A maximum likelihood phylogenetic

