The complete chloroplast genome and phylogenetic analysis of *Saussurea wettsteiniana* (Compositae)

Deng-Li Luo\(^{a,b}\), Ying-Min Zhang\(^{a}\), Ti-Cao Zhang\(^{b}\) and Guo-Dong Li\(^{a,b}\)

\(^{a}\)Yunnan Key Laboratory for Dai and Yi Medicines, Yunnan University of Chinese Medicine, Kunming, People's Republic of China; \(^{b}\)Faculty of Traditional Chinese Pharmacy, Yunnan University of Chinese Medicine, Kunming, People's Republic of China

**ABSTRACT**

*Saussurea wettsteiniana* is a medicinally important herb endemic to Hengduan Mountains. Here, we report and characterize the complete chloroplast genome sequence of *S. wettsteiniana* to provide genomic resources useful for future study. The complete chloroplast genome is 152,631 bp in length, consisting of a large single copy and a small single copy of 83,552 bp and 18,637 bp, which were separated by a pair of inverted repeats of 25,221 bp. Totally 133 genes were annotated, including 87 protein-coding genes, 36 tRNA genes, and eight rRNA genes. We also detected two pseudo-genes (*ycf1* and *rps19*). The overall GC content of the whole genome is 37.7%. The phylogenetic tree based on 23 complete plastomes indicated that *S. wettsteiniana* was closely related to *S. involucrata* of Compositae.

**CONTACT**

Guo-Dong Li \(\text{gammar116@163.com}\); Yunnan Key Laboratory for Dai and Yi Medicines, Yunnan University of Chinese Medicine, Kunming, People's Republic of China; Faculty of Traditional Chinese Pharmacy, Yunnan University of Chinese Medicine, Kunming, People's Republic of China

**ARTICLE HISTORY**

Received 13 January 2021
Accepted 30 May 2021

**KEYWORDS**

Saussurea wettsteiniana; chloroplast genome; medicinal plant; phylogenomics analysis

*Saussurea wettsteiniana* Hand. - Maz. is a perennial herb belonging to the Compositae family, endemic to Hengduan Mountains, and it is distributed in the elevation of 3400–4450 m of the hillside forests, forest edge, wet grassland and meadows (Flora of China Editorial Committee of Chinese Academy of Sciences 1999). It has been used in Tibetan folk medicine for the tonifying kidney, promoting blood circulation and strengthening Yang with a long history (Chen 2018). However, because of its low natural reproduction rate, overexploitation, and it is hard to cultivate artificially, natural populations of this species have suffered rapid declines. Besides, there is an extreme lack of research on the phylogenetic and conservation genetics of *S. wettsteiniana*. Here, we report and characterize the complete chloroplast genome sequence of *S. wettsteiniana*. It will provide reference data for the research on the phylogenetics and conservation genetics of *S. wettsteiniana*.

The fresh leaves of *S. wettsteiniana* were collected from Shangri-La County (27°52'N, 99°36'E), Yunnan province of China. A specimen was deposited at the Herbarium of Yunnan University of Chinese Medicine (YUNC) (Guo-Dong Li, gammar116@163.com; Cong-Wei Yang, congweiyang@foxmail.com) under the voucher number MW110638. The cpDNA sequence with complete annotation information was deposited at GenBank database under the accession number MW110638.

The complete chloroplast genome of *S. wettsteiniana* was a quadripartite circular and 152,631 bp in length, comprising a large single copy (LSC) region of 83,552 bp and a small single copy (SSC) region of 18,637 bp, separated by two inverted repeat (IR) regions of 25,221 bp. The chloroplast genome contained 133 genes, including 87 protein-coding genes, 36 tRNA genes, and eight rRNA genes. We also detected two pseudo-genes (*ycf1* and *rps19*). The overall GC-content of the whole cp genome is 37.7%, and those in the LSC, SSC, and IR regions are 35.8%, 38.6%, and 43.1%, respectively.

To identify the phylogenetic positions of *S. wettsteiniana*, as well as Compositae, we inferred the phylogenetic relationships based on the complete chloroplast genomes of 21 species from the whole range of the family Asteraceae and two Codonopsis species as outgroups. The alignment was carried out by software MAFFT v.7 (Katoh and Standley 2013). The maximum-likelihood (ML) analyses were constructed using IQ-TREE v1.6.10 (Nguyen et al. 2015) with...
the best-fit model selected by Model Finder and 1000 bootstrap replicates for each branch node (Figure 1). The phylogenetic tree topology from the ML analysis showed that *S. wettsteiniana* and *S. involucrata* from Compositae formed a monophyletic clade with 100% bootstrap value. The complete chloroplast genome of *S. wettsteiniana* will provide a useful resource for the conservation genetics of this species as well as for the phylogenetic studies of Compositae.

**Disclosure statement**

The authors declare that there is no conflict of interest regarding the publication of this article. The authors alone are responsible for the content and writing of the paper.

**Funding**

This research was funded by the High-end talent of Yunnan province ‘Youth Top-notch Talent Support Program’ (2020) to Guo-dong Li; the
Reserve Talent for Academic and Young Technical Leaders of Yunnan Province (2020) to Guo-dong Li; Special Subsidies for Public Health services of “Fourth Chinese Materia Medica Resources Survey” [DSS, MOF. 43/2018].

**Data availability statement**

The *Saussurea wettsteiniana* data that support the findings of this study are openly available in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov/](https://www.ncbi.nlm.nih.gov/) under the accession no. MW110638. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA721598, SUB9469822, and SAMN18721751, respectively.

**References**

Chen J. 2018. A systematic study of Saussurea subgen. Amphilaena (Asteraceae). Lanzhou University.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.

Flora of China Editorial Committee of Chinese Academy of Sciences. 1999. Flora of China. Vol. 78. Beijing: Science Press; Missouri: Missouri Botanical Garden Press.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Nguyen L, Schmidt HA, Von HA, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.

Patel RK, Jain M. 2012. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. PLoS One. 7(2):e30619.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq - versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11.