


## The complete chloroplast genome of *Thalictrum baicalense* Turcz. ex Ledeb

Yanfeng He<sup>a,b</sup>, Ruinan Wang<sup>a,b</sup>, Xiangyun Gai<sup>a,b</sup>, Pengcheng Lin<sup>a,b</sup>, Yongxi Li<sup>c</sup> and Jiuli Wang<sup>c</sup> 

<sup>a</sup>Key Laboratory for Tibet Plateau Phytochemistry of Qinghai Province, Qinghai Nationalities University, Xining, China; <sup>b</sup>Key Laboratory for Tibet Plateau Phytochemistry of Qinghai Province, Xining, China; <sup>c</sup>The College of Ecological Environment and Resources, Qinghai Nationalities University, Xining, China

### ABSTRACT

*Thalictrum baicalense* Turcz. ex Ledeb. is a well-known herbaceous perennial that has been used as a traditional medicine to treat influenza, hepatitis, and detoxification. In this study, we release and detail the complete chloroplast genome sequences of *T. baicalense*. The whole chloroplast genome was 155,859 bp in length and comprised 130 genes, including 84 protein-coding genes, 37 tRNA genes, eight rRNA genes. The *T. baicalense* chloroplast genome had a GC content of 38.39%. The phylogenetic relationships inferred that *T. baicalense*, *T. tenue*, *T. minus* and *T. petaloideum* are closely related to each other within the genus *Thalictrum*.

### ARTICLE HISTORY

Received 25 October 2020  
Accepted 28 December 2020

### KEYWORDS

*Thalictrum baicalense*;  
complete chloroplast  
genome; phylogenetic  
analysis

*Thalictrum baicalense* Turcz. ex Ledeb. (Ranunculaceae) is a herbaceous perennial distributed in China, Korea, Russia and other north warm regions (Wang et al. 2001). The dried root of *T. baicalense* have the effects of clearing heat, anti-inflammatory and antiviral. It has been used in the treatment of hepatitis, dysentery and detoxification in traditional Chinese medicine. In addition, it is reported in the literature that the genus *Thalictrum* comprises about 150 species worldwide. There are 76 species throughout China, of which about 30 species have been used as traditional folk medicines to treat influenza, cancer, and conjunctivitis (Chen et al. 2003; Xue et al. 2020).





The complete chloroplast (cp) genome can provide valuable genomic information for the phylogenetics study and the conservation of rare species (Weng et al. 2014). However, the cp genome of *T. baicalense* has not been fully sequenced. In this study, we sequenced and analyzed the cp genome of *T. baicalense* based on Illumina pair-end sequencing and compared it with other genus cp genome sequences. It would not only promote the phylogenetics study in family Ranunculaceae but also provide useful genetic information for the protection of *T. baicalense* and other related species.

The samples of *T. baicalense* were collected in the Qingshishan mountain (102°01'6.42"E, 36°17'14.25"N) and the voucher specimens (He202008) are deposited in the Herbarium of College of Pharmacy, Qinghai Nationalities University, Xining, China. The total DNA of *T. baicalense* was extracted using Plant Genomic DNA Kit (DP350; TIANGEN Biotech(Beijing) Co.,Ltd.). High-throughput sequencing was performed using an Illumina NovaSeq 6000 series sequencer

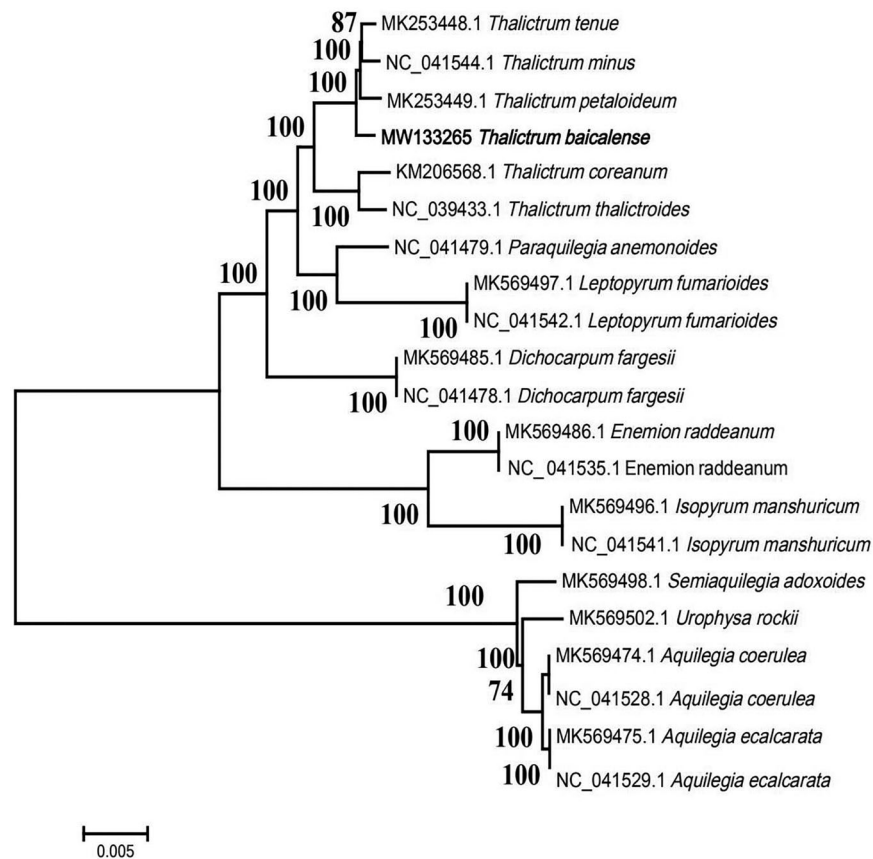
(PE150) by Nanjing Genepioneer Biotechnologies Inc. (Nanjing, China), and 8.04 GB of raw data was generated. The raw paired-end reads were filtered using the fastp program (Chen et al. 2018). The high-quality reads were applied to a *de novo* assembly performed using the program SPAdes assembler v3.10.1 (Bankevich et al. 2012). The assembled genome was annotated using CPGAVAS2 (Shi et al. 2019).

The complete cp genome (MW133265) of *T. baicalense* was 155,859 bp in length having 38.39% of total GC content. This cp genome has a typical quadripartite structure, containing a large single-copy region (LSC) of 85,258 bp, a small single-copy region (SSC) of 17,637 bp, and 2 inverted repeat (IR) regions of 26,482 bp. A total of 130 genes are successfully annotated, including 84 protein-coding genes, 37 tRNA genes and 8 rRNA genes. The tRNA genes are distributed throughout the whole genome with 22 in the LSC, 1 in the SSC, and 14 in the IR regions, while rRNAs are only situated in the IR regions. In IR, six protein-coding genes (*ndhB*, *rpl2*, *rpl23*, *rps7*, *rps12*, and *ycf2*), seven tRNA genes (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*), and four rRNA genes (*rrn4.5*, *rrn5*, *rrn16*, and *rrn23*) have been coded. Among the protein-coding genes, two genes (*clpP* and *ycf3*) contained two introns, and other nine genes (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, *rpoC1*, *rps12* and *rps16*) had one intron each.

In order to reveal the phylogenetic position of *T. baicalense* with its close relatives of Ranunculaceae, a phylogenetic analysis was performed based on 21 complete chloroplast genomes of Ranunculaceae. The complete chloroplast genomes were aligned by MAFFT v7.307 (Katoh and Standley

**CONTACT** Yanfeng He  [yfhe\\_629@163.com](mailto:yfhe_629@163.com)  Key Laboratory for Tibet Plateau Phytochemistry of Qinghai Province, Qinghai Nationalities University, Xining, Qinghai 810007, China; Jiuli wang  [wang\\_jiul@163.com](mailto:wang_jiul@163.com)  The College of Ecological Environment and Resources, Qinghai Nationalities University, Xining, Qinghai 810007, 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.** The maximum-likelihood (ML) phylogenetic tree of 21 complete chloroplast genome sequences.

2013) and the maximum-likelihood tree (Figure 1) was built using MEGA7 (Kumar et al. 2016). Using the Tamura-Nei model the ML phylogenetic analysis were conducted with MEGA v7.0.26 generating 1000 bootstrap replicates to determine measures of nodal support with each run initiating from a random starting tree. According to the result (Figure 1), *T. baicalense* has close relationship with the species *T. tenue*, *T. minus* and *T. petaloideum*.

### Disclosure statement

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

### Funding

This study was supported by the National Natural Science Foundation of China [81960785], the Applied Basic Research Project of Qinghai Province [2020-ZJ-717], the Development Project of Qinghai Provincial Key Laboratory [2020-ZJ-Y19], the Science and Technology Projects of Qinghai Nationalities University [2019XJG06] and Qinghai '1000 Talents' program.

### ORCID

Jiuli Wang  <http://orcid.org/0000-0002-7426-8101>

### Data availability statement

The data that support the results of this study are openly available in GeneBank (<https://www.ncbi.nlm.nih.gov/genbank/>) under the accession

numbers MW133265 after the paper has been published. The associated 'BioProject,' 'SRA,' and 'Bio-Sample' numbers are PRJNA680221, SRR13178662, and SAMN16872741 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.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 34(17):i884–890.
- Chen SB, Chen SL, Xiao PG. 2003. Ethnopharmacological investigations on *Thalictrum* plants in China. *J Asian Nat Prod Res.* 5(4):263–271.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33(7): 1870–1874.
- Shi L, Chen H, Jiang M, Wang I, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73.
- Wang WC, Fu DZ, Li LQ, Bruce B, Anthony RB, Bryan ED, Michael GG, Yuichi K, Orbélia RR, Michio T, et al. 2001. Ranunculaceae. In: Wu Z, Raven PH, editors. *Flora of China*. Beijing (China): Science Press and Missouri Botanical Garden; p. 133–438.
- Weng ML, Blazier JC, Govindu M, Jansen RK. 2014. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. *Mol Biol Evol.* 31(3):645–659.
- Xue JJ, Jiang CY, Zou DL, Li BJ, Lu JC, Li DH, Lin B, Li ZL, Hua H-M. 2020. Baicalensines A and B, Two Isoquinoline Alkaloids from the roots of *Thalictrum baicalense*. *Org Lett.* 22(19):7439–7442.