


The complete chloroplast genome of the ephemeral plant *Isatis minima* (brassicaceae) of northwest China

Feng Song^a, Lu Jin^b, Xue-Jun Ge^a  and Ying Feng^c

^aKey Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China; ^bCollege of Life Sciences, South China Agricultural University, Guangzhou, China; ^cState Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, China

ABSTRACT

The complete chloroplast genome of *Isatis minima*, a typical ephemeral plant of Brassicaceae in the Central Asia desert, was sequenced and characterized in this study. The genome 153,642 bp in size, contains a typical quadripartite genome organization including LSC and SSC regions of 83,423 bp and 17,709 bp, and two copies of the IR regions of 26,255 bp. It has 113 unique genes, including 79 protein-coding, 30 tRNA, and four rRNA genes. Phylogenetic analysis fully resolved *I. minima* in a monophyletic clade with *I. tinctoria*. This bioinformatic data contributes to the phylogenetics systematics and evolutionary history of Brassicaceae.

ARTICLE HISTORY

Received 5 July 2021
Accepted 27 August 2021

KEYWORDS

Brassicaceae; chloroplast genome; ephemeral plants; *Isatis minima*


Isatis Tourn. ex L. (Brassicaceae) contains about 85 species native to the Mediterranean region east to central Asia (Mabberley 2017). There are four species in China. The roots and leaves of *Isatis* are used for medicinal purposes and are also sources of dye (Zhou et al. 2001). *Isatis minima* Bunge is a typical ephemeral plant that mainly grows in deserts, steppe, and roadsides areas of Xinjiang and Gansu (Zhou et al. 2001). In this study, we obtained the complete chloroplast (cp) genome of *I. minima* using genome skimming methods. We characterized the cp genome to determine its genome structure and provide resources for studying the phylogenetics and speciation of ephemeral plant classified in the Brassicaceae.

The specimen of *I. minima* was collected from Urumqi, Xinjiang Uygur Autonomous Region (43°48'57.45"N, 87°33'51.34"E), and dried with silica gel. A voucher was deposited at the Herbarium of South China Botanical Garden (IBSC, <http://herbarium.scbg.cas.cn/>, Xue-Jun Ge, xjge@scbg.ac.cn) under the voucher number IBSC 0812950. Total genomic DNA isolation protocols followed the cetyltrimethyl ammonium bromide (CTAB) method (Doyle and Doyle 1987). DNA extracts were fragmented for 300 bp short-insert library construction and sequenced -2×150 bp paired-end (PE) reads on an Illumina HiSeq X-Ten instrument at the Beijing Genomics Institute (BGI, Shenzhen, China). The complete cp genome was assembled using GetOrganelle (Jin et al. 2018), and the finished cp genome was annotated with GeSeq (Tillich et al. 2017) and adjusted manually using Geneious v 11.0.2 (Ripma et al. 2014), with *I. tinctoria* (MK637736) as reference. The annotated cp genome was submitted to

GenBank (Accession number: MZ488447). MISA-web v2.1 (Beier et al. 2017) was applied to identify simple sequence repeats sequences (SSR).

The complete cp genome of *I. minima* is 153,642 bp in length (36.5% GC contents), with a large single copy region (LSC) of 83,423 bp (34.3% GC contents), a small single copy region (SSC) of 17,709 bp (29.5% GC contents), and a pair of two inverted repeats regions (IR) of 26,255 bp (42.7% GC contents). The cp genome encoded 113 unique genes, including 79 protein-coding, 30 tRNA, four rRNA genes. The cp genome generated in this study exhibited typical plastome structure, gene order and GC content was similar to that of *I. tinctoria* (Yang and Wang 2017). A total of 68 SSRs were detected, 59 of which were mono-nucleotide (A/C/T, 86.76%), eight were di-nucleotides (AT/TA, 11.76%), one was tetra-nucleotide repeat (TTAA, 1.47%), respectively.

A phylogenetic analysis of the Brassicaceae was carried out on 31 complete cp genomes, and *Tarenaya hassleriana* (*Tarenaya*, Cleomaceae, Accession number: NC_034364.1) was designated as the outgroup. The DNA sequences for these 32 complete cp genomes (after removing one IR) were aligned using the default option implemented in MAFFT version 7 (Kato and Standley 2013). The GTR+G model was used for Maximum likelihood (ML) tree analyses using RAxML version 8.2.10 (Stamatakis 2014) and then was generated with 1000 replicates. Consistent with Li et al. (2014), the species of *Isatis* formed a monophyletic clade with 100% bootstrap support value (Figure 1) and shared a sister relationship to *Conringia planisiliqua* (Accession number: NC_049619).

CONTACT Ying Feng  luckfy@ms.xjb.ac.cn  State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, 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.

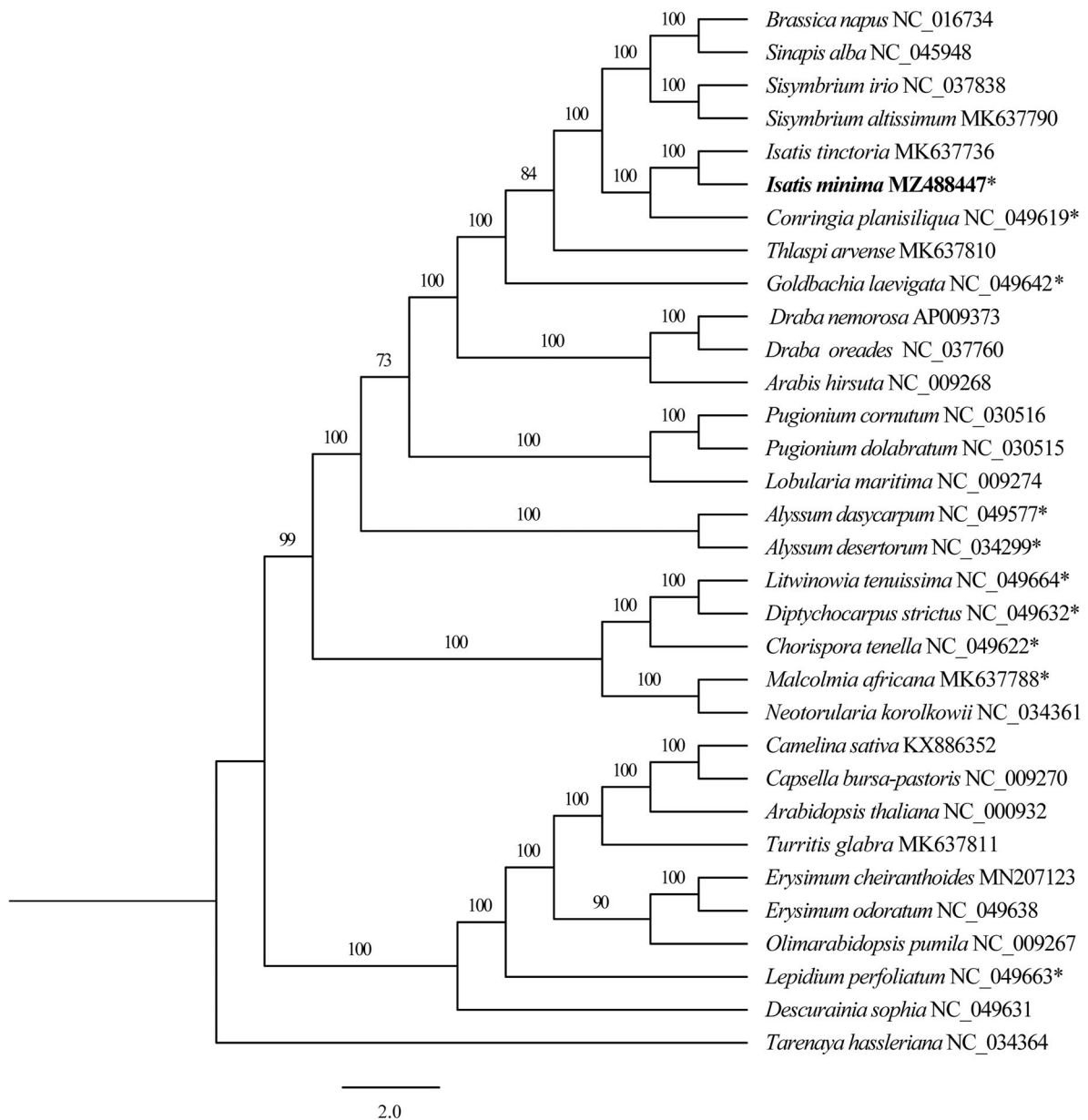


Figure 1. The maximum-likelihood (ML) phylogenetic tree based on 31 complete chloroplast genomes (only one IR region) of Brassicaceae (*Tarenaya hassleriana* was chosen as outgroup). “*” indicates ephemeral plants. The number above branches indicated the value of bootstrap values.

Disclosure statement

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

Funding

This study was financially supported by The National Basic Resource Survey of China [2017FY100201], The National Natural Science Foundation of China [31770227], The Science and Technology Basic Resources Survey special [2018FY100704].

ORCID

Xue-Jun Ge  <http://orcid.org/0000-0002-5008-9475>

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. MZ488447. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA742888, SRS9358470, and SAMN19986842 respectively.

References

- Beier S, Thiel T, Munch T, Scholz U, Mascher M. 2017. MISA-web: a web server for microsatellite prediction. *Bioinformatics*. 33(16): 2583–2585.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytoch Bull.* 19:11–15.
- Jin JJ, 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*. 4:256479.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Li Y, Feng Y, Liu B, Lv GH, Wang XY. 2014. Phylogenetic relationships and divergence time of Brassicaceae ephemeral plants in Xinjiang. *Arid Zone Research.* 31(6):1100–1108.

- Mabberley DJ. 2017. *Mabberley's plant-book*. 4th ed. Cambridge (UK): Cambridge University Press.
- Ripma LA, Simpson MG, Hasenstab-Lehman K. 2014. Geneious! Simplified genome skimming methods for phylogenetic systematic studies: a case study in *Oreocarya* (Boraginaceae). *Appl Plant Sci*. 2: 1400062.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9): 1312–1313.
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq - versatile and accurate annotation of organellar genomes. *Nucleic Acids Res*. 45(W1):W6–W11.
- Yang S, Wang ZZ. 2017. The complete chloroplast genome sequence of the medicinal and economic plant woad *Isatis indigotica* (Brassicaceae). *Mitochondrial DNA B Resour*. 2(2):514–515.
- Zhou TY, Lu LL, Yang G, Al-Shehbaz IA. 2001. Brassicaceae (Cruciferae). Wu ZY, Raven PH, Hong D, editors. Vol. 8, *Flora of China*. Beijing: Science Press, Missouri Botanical Garden Press.