

Extensive gene loss in the plastome of holoparasitic plant *Cistanche tubulosa* (Orobanchaceae)

Wanqi Xu^{a*}, Haimei Chen^{a*}, Lixia Tian^a, Mei Jiang^a, Qiaoqiao Yang^a, Liqiang Wang^a, Bashir Ahmad^b and LinFang Huang^a

^aKey Research Laboratory of Traditional Chinese Medicine Resources Protection from Ministry of Education, Engineering Research Center of Chinese Medicine Resource, National Administration of Traditional Chinese Medicine, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, China; ^bCenter for Biotechnology and Microbiology, University of Peshawar, Peshawar, Pakistan

ABSTRACT

Extensive photosynthetic gene loss and rapid evolutionary rate occur in the plastomes of parasitic plants. The holoparasitic plant *Cistanche tubulosa* of Orobanchaceae is an important medicinal resource that are distributed in arid areas. In this study, the complete plastome of *C. tubulosa* has been sequenced, assembled and analyzed. The total plastome of *C. tubulosa* was 75,375 bp in length, consisting of a pair of inverted repeats (IRs, 6,593 bp), a large single-copy region (LSC, 32,470 bp) and a small single-copy region (SSC, 29,719 bp). It contained 24 intact protein coding genes, nine pseudogenes, and 44 missing genes. In addition, all the protein-coding genes, which were related to photosynthesis and energy production, were pseudogenised or lost. Four rRNA genes and 24 tRNA genes were intact meanwhile five tRNA genes were missing. Phylogenetic tree indicated that *C. tubulosa* was closely related to *C. phelypaea*. Our results may improve understanding of the plastome organization, classification, and evolution of parasitic plants.

ARTICLE HISTORY



Received 7 June 2020
Accepted 20 June 2020

KEYWORDS

Cistanche tubulosa;
plastome; gene loss;
holoparasitic plant;
phylogenetic analysis

Orobanchaceae is a special family that comprises all level growth behavior plants consisting of nonparasitic, hemiparasitic and holoparasitic plants (Xi et al. 2013). *Cistanche tubulosa*, a holoparasitic plant of the Orobanchaceae, absorbs water and organic and inorganic nutrition from the roots of its hosts. *Cistanche tubulosa* is traditionally used as nourishing herbs. A lot of compounds have been isolated from *C. tubulosa*, including phenylethanoid glycosides, carbohydrates, lignans, iridoids, echinacoside, verbascoside, chlorogenic acid, acteoside, and luteoloside (Yong and Peng-Fei 2009; Fu et al. 2018; Gao et al. 2019). These isolated compounds have exhibited abundant pharmacologic effects, such as neuroprotective, immunomodulatory, anti-senescence, anti-inflammatory, anti-osteoporosis, hepatoprotection, anti-oxidative, anti-bacterial, anti-tumour and glucose-tolerance-improving effects (Morikawa et al. 2019). Compared with the more than 3000 plastomes of autotrophic organisms that can be obtained from the National Center for Biotechnology Information database, data from the plastomes of parasitic plants are limited. To depict the characteristic of gene losses and provide new insights into the overall evolutionary process, we analyzed the plastome of *C. tubulosa*.

Fresh scale leaves of *C. tubulosa* were collected from the Hotan Prefecture (Xinjiang Uygur Autonomous Region), China (78°12'36"E, 36°19'12"N). The voucher specimens were deposited in the herbarium of Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College (IMD), with accession numbers of XJ20170502. Approximately 500 ng DNA was used to construct a library with an insert size of 400 bp and sequenced according to the manufacturer's instructions for HiSeq 4000 platform. The clean paired-end reads were filtered against all the plastomes of plants recorded in the GenBank database by using BLASTn with an e-value cutoff of 1e−5. The extracted reads were assembled using SPAdes (v. 3.10.1) (Bankevich et al. 2012). The boundary region was validated using primer pairs spanning the boundary region followed by PCR amplification of the regions and Sanger sequencing of the PCR production. Genes were annotated by using CpGAVAS2 web service (Shi et al. 2019) and edited manually by using Apollo genome editor (Lewis et al. 2002). The circle map of plastome was generated using OrganellarGenomeDRAW (Lohse et al. 2013), and GC content was analyzed using CGView Server (Grant and Stothard 2008). In comparison with the nonparasitic plant of

CONTACT LinFang Huang  lfhuang@implad.ac.cn  Key Research Laboratory of Traditional Chinese Medicine Resources Protection from Ministry of Education, Engineering Research Center of Chinese Medicine Resource, National Administration of Traditional Chinese Medicine, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100193, China

*These authors contributed equally to this work.

© 2020 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.

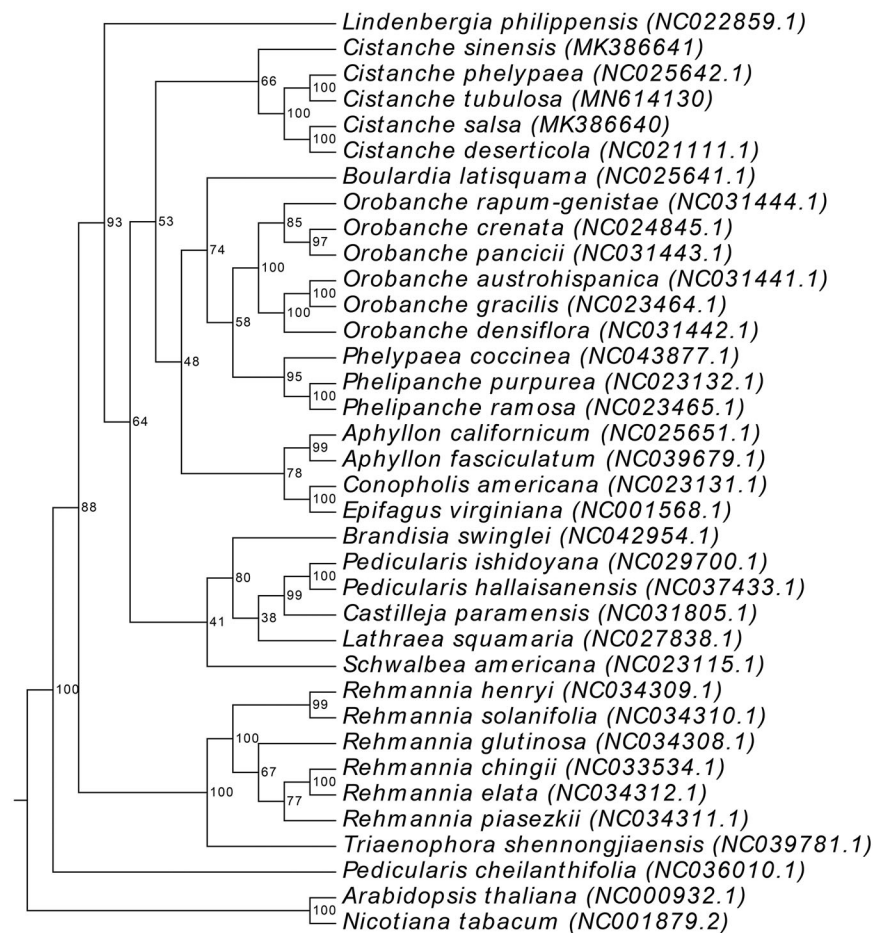


Figure 1. Phylogenetic tree based on 36 complete plastomes of Orobanchaceae.

Orobanchaceae *Rehmannia glutinosa* (NC_034308), genes that were similar to known protein-coding genes but truncated or contained one or more frameshift mutations, were classified as pseudogenes (Cusimano and Wicke 2016; Xi et al. 2013). The genome assembly and annotation results were deposited in GenBank, with accession numbers of MN614130.

The plastome of *C. tubulosa* was 75,375 bp in length and showed a typical quadripartite structure, including a pair of IRs (6593 bp) separated by LSC (32,470 bp) and SSC (29,719 bp) region. The total GC content was 34.95% and the SSC region had a higher GC content (38.00%) than the IR (33.22%) and LSC (32.86%) region. The plastome of *C. tubulosa* retained 27 intact protein-coding genes, nine genes became pseudogenes, and 44 genes were missing completely. All genes related to encoding photosynthetic proteins were pseudogenised or lost, which supports the holoparasitic lifestyle of *C. tubulosa*. All four rRNAs were universally localized in the SSC region. A total of 24 tRNAs remained, and up to five tRNAs were lost throughout the evolution.

To analyze the phylogenetic position of *C. tubulosa* within the Orobanchaceae family, a molecular evolutionary tree was constructed using 36 species. Twelve shared protein sequences were concatenated and aligned using the ClustalW program (Thompson et al. 2002). The phylogenetic tree was constructed using the Randomized Axelerated Maximum Likelihood (RAxML) software (Stamatakis 2014) and the ML

method, with *Arabidopsis thaliana* and *Nicotiana tabacum* as the outgroups. The phylogenetic tree showed that *C. tubulosa* and *C. phelypaea* were grouped together (Figure 1).

Disclosure statement

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

Funding

This work was supported by National Natural Science Foundation of China [81473315, U1812403-1-1], National Science & Technology Fundamental Resources Investigation Program of China [2018FY100701] and CAMS Innovation Fund for Medical Sciences [2016-I2M-3-015], which are gratefully acknowledged.

Data availability statement

The data that support the findings of this study are deposited in GenBank, with accession numbers of MN614130. <https://www.ncbi.nlm.nih.gov/nucleotide/MN614130.1/>.

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.

- Cusimano N, Wicke S. 2016. Massive intracellular gene transfer during plastid genome reduction in nongreen Orobanchaceae. *New Phytol.* 210(2):680–693.
- Fu Z, Fan X, Wang X, Gao X. 2018. Cistanches Herba: an overview of its chemistry, pharmacology, and pharmacokinetics property. *J Ethnopharmacol.* 219:233–247.
- Gao Y, Guo LN, Ma SC, Liu J, Zheng J, Zan K. 2019. Comparative studies of three *Cistanche* species based on UPLC specific chromatogram and determination of main components. *Zhongguo Zhong Yao Za Zhi.* 44(17):3749–3757.
- Grant JR, Stothard P. 2008. The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Res.* 36(Web Server):W181–W184.
- Lewis SE, Searle SMJ, Harris N, Gibson M, Lyer V, Richter J, Wiel C, Bayraktaroglu L, Birney E, Crosby MA, et al. 2002. Apollo: a sequence annotation editor. *Genome Biol.* 3(12):RESEARCH0082.
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41(Web Server issue):W575–W581.
- Morikawa T, Xie H, Pan Y, Ninomiya K, Yuan D, Jia X, Yoshikawa M, Nakamura S, Matsuda H, Muraoka O. 2019. A review of biologically active natural products from a desert plant *Cistanche tubulosa*. *Chem Pharm Bull.* 67(7):675–689.
- 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.
- Thompson JD, Gibson TJ, Higgins DG. 2002. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics.* Chapter 2: 2.3.1–2.3.22.
- Xi L, Ti-Cao Z, Qin Q, Zhumei R, Jiayuan Z, Takahiro Y, Masami H, Crabbe MJC, Jianqiang L, Yang Z. 2013. Complete chloroplast genome sequence of holoparasite *Cistanche deserticola* (Orobanchaceae) reveals gene loss and horizontal gene transfer from its host *Haloxylon ammodendron* (Chenopodiaceae). *PLoS One.* 8(3):e58747
- Yong J, Peng-Fei T. 2009. Analysis of chemical constituents in *Cistanche* species. *J Chromatogr A.* 1216(11):1970–1979.