

Comparative analysis of the complete chloroplast genomes of *Cirsium japonicum* from China and Korea

Lixia Tian*, Mei Jiang*, Haimei Chen, Jinglin Li, Linfang Huang and Chang Liu 

Key Research Laboratory of Traditional Chinese Medicine Resources Protection, Administration of Traditional Chinese Medicine, National Administration of Traditional Chinese Medicine, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

ABSTRACT

Cirsium japonicum (*C. japonicum*) is a traditional Chinese medicine belonging to the family Asteraceae. The previous studies have indicated that the chemical compound content of *C. japonicum* from different places was different. To distinguish *C. japonicum* from different geographies, the chloroplast genome of *C. japonicum* from China was sequenced and compared with that from Korea. The total length of this genome is 152,602 bp, similar to that of Korea (152,606 bp). It has a conservative quartile structure which is composed of a large single-copy (LSC) region, a small single-copy (SSC) region and a pair of inverted repeats (IRs) regions, with lengths of 83,487 bp, 18,721 bp, and 25,197 bp, respectively. It encodes 79 protein-coding, 27 transfer RNAs, and 4 ribosomal RNA genes. The overall GC content of the genome is 37.70%. A total of 20 single nucleotide polymorphisms and 6 insertions and deletions were identified between the chloroplast genome of *C. japonicum* from China and Korea. These results can be applied to develop molecular markers to distinguish *C. japonicum* from different geographical origins.

ARTICLE HISTORY

Received 5 November 2020
Accepted 29 March 2021

KEYWORDS

Cirsium japonicum;
chloroplast genome;
identification; phylogenetic
analysis

C. japonicum, belonging to *Cirsium* genus, is a perennial herb distributed in China, as well as Japan and Korea, which has a long history of being used as functional food and herb (Zeng et al. 2016). Previous reports have shown that it has multiple pharmacologic effects, including antidiabetic, antitumor, and anti-inflammatory (Ma et al. 2016; Yang et al. 2018; Zhao et al. 2018; Jang et al. 2020). The chemical compound content of *C. japonicum* from different places is different, which may affect clinical efficacy (Chen and Gong 2013). The chloroplast genome is a perfect option for developing molecular markers to discriminate the different geographical areas of the same species (Jiang et al. 2018). Hence, the chloroplast genome of *C. japonicum* from China was sequenced and compared with that from Korea to find the potential marker.

The fresh and healthy leaves of the *C. japonicum* were harvested from the Central China medicinal botanical garden, Hubei, China (30°10'42"N, 109°44'55"E). The specimen was stored under voucher number Implad201808173 at the Herbarium of Institute of Medicinal Plant Development in Beijing, China. DNA was extracted followed the modified CTAB method. Subsequently, the integrity and concentration of extracted DNA were measured by electrophoresis in 1% (w/v) agarose gel and spectrophotometer (Nanodrop 2000, Thermo Fisher Scientific, Waltham, MA). The DNA library with

insert sizes of 500 bases was constructed with 1 µg DNA using the library preparation kit (New England Biolabs, Ipswich, MA), then was subjected to high-throughput sequencing using an Illumina HiSeq2000 sequencer (Illumina Inc., San Diego, CA). Clean reads were assembled using NOVOPlasty v.2.7.2 (Dierckxsens et al. 2017). The assembled chloroplast genome was annotated with CPGAVAS2 (Shi et al. 2019) and then edited using Apollo (Misra and Harris 2006). The genome sequence and annotations have been deposited in GenBank with accession numbers MW035606.

The chloroplast genome of *C. japonicum* from China was 152,602 bp, similar to that from Korea (152,606 bp), comprising of a large single-copy (LSC) region of 83,487 bp, a small-copy (SSC) region of 18,721 bp regions, and a pair of inverted repeats (IRs) regions of 25,197 bp. The GC content of the total genome, LSC, SSC, and IR regions is 37.70%, 35.85%, 31.38%, and 43.10%, respectively. Besides, the genome encodes a total of 110 genes, including 79 protein-coding genes, 27 tRNA genes and 4 rRNA genes, respectively. Overall, the genome is highly consistent with the one previously reported.

To detect the variation between the chloroplast genome of *C. japonicum* from China and Korea, the DnaSP v6.0 (Rozas et al. 2017) software was employed. A total of 20 single nucleotide polymorphisms (SNPs) and 6 insertions and

CONTACT Chang Liu  cliu@implad.ac.cn; Linfang Huang:  lfhuang@implad.ac.cn  Key Research Laboratory of Traditional Chinese Medicine Resources Protection, Administration of Traditional Chinese Medicine, National Administration of Traditional Chinese Medicine, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

*These authors contributed equally to this work.

© 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.

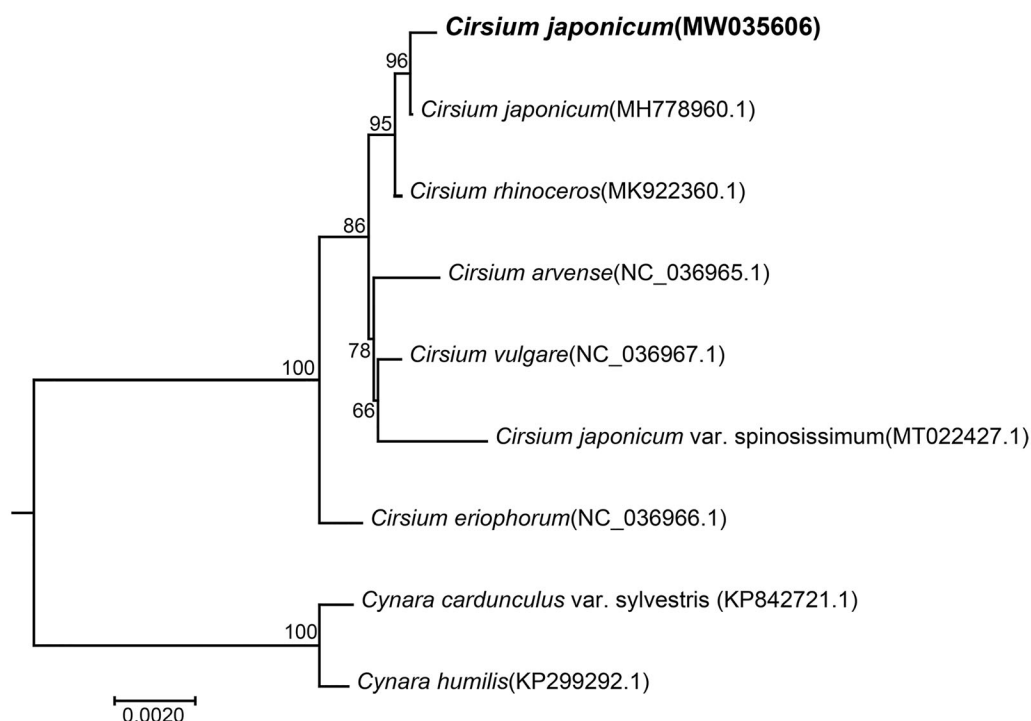


Figure 1. Phylogenetic relationships of *Cirsium* species inferred using maximum-likelihood (ML) method. The phylogenetic tree constructed using the shared protein sequences among the 9 samples. The number on the branch indicates the bootstrap value. *Cynara humilis* (KP299292.1) and *Cynara cardunculus* var. *sylvestris* (KP842721.1) were used as outgroups.

deletions (InDels) were detected. Most SNPs were detected in the LSC and SSC regions, with 10 and 6, respectively, and 4 were found in the IR regions. Among the protein-coding genes, the SNPs were found in *ycf1*, *rps16*, and *trnE-UUC* genes, resulting in the changes of three amino acids (836nd, Ser-Arg; 1239nd, His-Asn; 1342nd, Leu-termination codon), one (259nd, Ser-Leu) and one (89nd, Arg-Lys), respectively. The InDels were found in the *trnK-UUU* and *matK* genes. Our results showed that 5 protein-coding genes (*ycf1*, *rps16*, *trnE-UUC*, *trnK-UUU*, and *matK*) might be effective in distinguishing *C. japonicum* from China and Korea.

Finally, we inferred phylogenetic relationships among 6 *Cirsium* species based on the complete chloroplast genome sequences. Two *Cynara* species were used as outgroups. The chloroplast genome sequences of 6 species were downloaded from GenBank (the accession numbers are shown in Figure 1). Shared protein sequences were aligned using the CLUSTALW2 (v2.0.12) program. The phylogenetic tree was constructed using the maximum-likelihood (ML) method implemented in RaxML (v8.2.4) (Stamatakis 2014). The bootstrap analysis was performed with 1000 replicates. In the phylogenetic tree, most nodes have high bootstrap support values that are >80%. Figure 1 shows that the two *C. japonicum* species are clustered together, as expected. And they are closely related to *Cirsium rhinoceros*. The present study indicated that the chloroplast genome can be used to identify *C. japonicum* from different origins.

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 (No. 81473315), National Science & Technology Fundamental Resources Investigation Program of China (2018FY100701), Sichuan Province Science and Technology Plan Project (2018JZ0028), the Open Research Fund of Chengdu University of Traditional Chinese Medicine Key Laboratory of Systematic Research of Distinctive Chinese Medicine Resources in Southwest China (003109034001), and CAMS Innovation Fund for Medical Sciences (CIFMS) No. 2016-I2M-3-015.

ORCID

Chang Liu  <http://orcid.org/0000-0003-3879-7302>

Data availability statement

The data that support the findings of this study are openly available in GenBank at <https://www.ncbi.nlm.nih.gov/nuccore/MW035606>. Raw sequencing data used in this study have been deposited in SRA with accession number SRR12975765.

References

- Chen Q, Gong QF. 2013. Determination of pectolarigenin in carbonized *Cirsium japonicum* DC. from different places. *Northwest Pharm J.* 28(5): 449–451.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45(4):e18.
- Jang M, Kim KH, Kim GH. 2020. Antioxidant capacity of Thistle (*Cirsium japonicum*) in various drying methods and their protection effect on neuronal PC12 cells and *Caenorhabditis elegans*. *Antioxidants.* 9(3):200.
- Jiang M, Chen H, He S, Wang L, Chen AJ, Liu C. 2018. Sequencing, characterization, and comparative analyses of the plastome of *Caragana rosea* var. *rosea*. *Int J Mol Sci.* 19(5):1419.

- Ma Q, Wang LH, Jiang JG. 2016. Hepatoprotective effect of flavonoids from *Cirsium japonicum* DC on hepatotoxicity in comparison with silymarin. *Food Funct.* 7(5):2179–2184.
- Misra S, Harris N. 2006. Using Apollo to browse and edit genome annotations. *Curr Protoc Bioinformatics.* 12(5):1–28.
- Rozas J, Ferrer-Mata A, Sanchez-Delbarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol.* 34(12):3299–3302.
- 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.
- Yang X, Shao H, Chen Y, Ding N, Yang A, Tian J, Jiang Y, Li G, Jiang Y. 2018. In renal hypertension, *Cirsium japonicum* strengthens cardiac function via the intermedin/nitric oxide pathway. *Biomed Pharmacother.* 101:787–791.
- Zeng QH, Zhao JB, Wang JJ, Zhang XW, Jiang JG. 2016. Comparative extraction processes, volatile compounds analysis and antioxidant activities of essential oils from *Cirsium japonicum* Fisch. ex DC and *Cirsium setosum* (Willd.) M.Bieb. *LWT – Food Sci Technol.* 68:595–605.
- Zhao ZW, Chang JC, Lin LW, Tsai FH, Chang HC, Wu CR. 2018. Comparison of the hepatoprotective effects of four endemic *Cirsium* species extracts from Taiwan on CCl₄-induced acute liver damage in C57BL/6 mice. *IJMS.* 19(5):1329.