

Complete chloroplast genome of *Cardamine hupingshanensis* K.M.Liu, L.B.Chen, H.F.Bai & L.H.Liu (Brassicaceae) in Enshi, Hubei

Xiuqing Liu^a and Siying Zhang^b

^aForestry College, Xinyang Agriculture and Forestry University, Xinyang, China; ^bTourism Management College, Xinyang Agriculture and Forestry University, Xinyang, China

ABSTRACT

Cardamine hupingshanensis K.M.Liu, L.B.Chen, H.F.Bai & L.H.Liu 2008, also called *Cardamine enshiensis*, belongs to the genus *Cardamine*, Brassicaceae. As a plant with selenium enrichment ability, it has high development value. Here, we analyzed the chloroplast genome of *C. hupingshanensis*. The complete chloroplast genome had a total size of 154,832 bp with a typical quadripartite structure, including a large single-copy region (LSC, 83,908 bp) and a small single-copy region (SSC, 17,938 bp), separated by a pair of inverted repeat regions (IRs, 26,493 bp). Genome annotation showed the chloroplast genome contained 113 unique genes, including 79 protein-coding genes, 30 tRNA genes, and four rRNA genes. A total of 143 SSRs were found in the chloroplast genome. Phylogenetic analysis showed that *C. hupingshanensis* was closer to the *C. circaeoides* and *C. lyrata*. This chloroplast genome resource will be useful for study of the phylogeny and evolution of *Cardamine* in the future.

ARTICLE HISTORY

Received 27 June 2022
Accepted 10 August 2022



KEYWORDS

Brassicaceae; *Cardamine hupingshanensis*; chloroplast genome; phylogeny

Cardamine hupingshanensis K.M.Liu, L.B.Chen, H.F.Bai & L.H.Liu 2008, also called *Cardamine enshiensis*, belongs to the Brassicaceae family, is a selenium (Se) hyperaccumulator plant in Enshi, Hubei Province, China (Shao et al. 2007; Bai et al. 2008; Cui et al. 2018). Selenium (Se) is an essential trace element for human and animals, which is necessary for the synthesis of selenoprotein and selenoenzymes (Butler et al. 1989; Behne and Kyriakopoulos 2001; Rafferty et al. 2003). For humans and animals, plants are the main source of selenium intake, so adequate selenium in plants is very important for health (Sargent 2008; Williams et al. 2009). Some previous studies showed that some Cabombaceae, Liliaceae, and Fabaceae plants had high ability to accumulate Se (Sharma et al. 2010; Freeman and Bañuelos 2011), such as the *Astragalus bisulcatus* (Pilon-Smits et al. 2009) and *Stanleya pinnata* (Freeman et al. 2006). In recent years, the researchers systematically analyzed the whole genome of *C. hupingshanensis*, and studied the mechanism underlying Se tolerance and hyperaccumulation (Huang et al. 2021). Nevertheless, the phylogeny and evolution of the chloroplast genome still need to be explored for *C. hupingshanensis*. Hence, we characterized the chloroplast genome of *C. hupingshanensis* that may contribute to understanding the phylogenetic relationship of *Cardamine*.

The material of *C. hupingshanensis* was collected from Tiechangba, Enshi, Hubei in China (110°08'43"E, 30°47'22"N, and altitude 1150 m). A specimen was deposited in the herbarium of the Forestry College, Xinyang

Agriculture and Forestry University (voucher number: ENS001, <https://www.xyafu.edu.cn/lxy/>, Zhang Siying, 2019300002@xyafu.edu.cn). The total genomic DNA of leaves was extracted using modified CTAB method (Doyle and Doyle 1987; Chen et al. 2014). Genomic DNA was sequenced using the Illumina NovaSeq platform of Biomarker Technologies Company (Beijing, China). A total of 3.3 Gb raw reads were generated, and low-quality sequences were filtered by using Trimmomatic v. 0.39 (Bolger et al. 2014) to obtain clean data. The clean data were quality-controlled by using FastQC v. 0.11.9 (Simon 2020). Then, the complete chloroplast genome was assembled by using GetOrganelle v. 1.7.5 (Jin et al. 2020), the parameters set as follows: -R=15, -F=embplant_pt, -k=21, 45, 65, 85, and 105. We used Bandage v. 0.8.1 (Wick et al. 2015) to check the integrity of the assembly result, and then used CPGAVAS2 (Shi et al. 2019) and PGA (Qu et al. 2019) for genome annotation. Finally, the complete chloroplast genome had a total size of 154,832 bp (ON322745) with a typical quadripartite structure, including a large single-copy region (LSC, 83,908 bp) and a small single-copy region (SSC, 17,938 bp), separated by a pair of inverted repeat regions (IRs, 26,493 bp). The total GC content of the chloroplast genome was 36.27%. The chloroplast genome encoded 113 unique genes, including 79 protein-coding genes, 30 tRNA genes, and four rRNA genes. MicroSatellite software (Beier et al. 2017) was used for simple repeat sequence (SSR) analysis in chloroplast genome of *C. hupingshanensis*, the number of repetitions was set to 10,

CONTACT Siying Zhang  2019300002@xyafu.edu.cn  Tourism Management College, Xinyang Agriculture and Forestry University, Xinyang, China

© 2022 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.

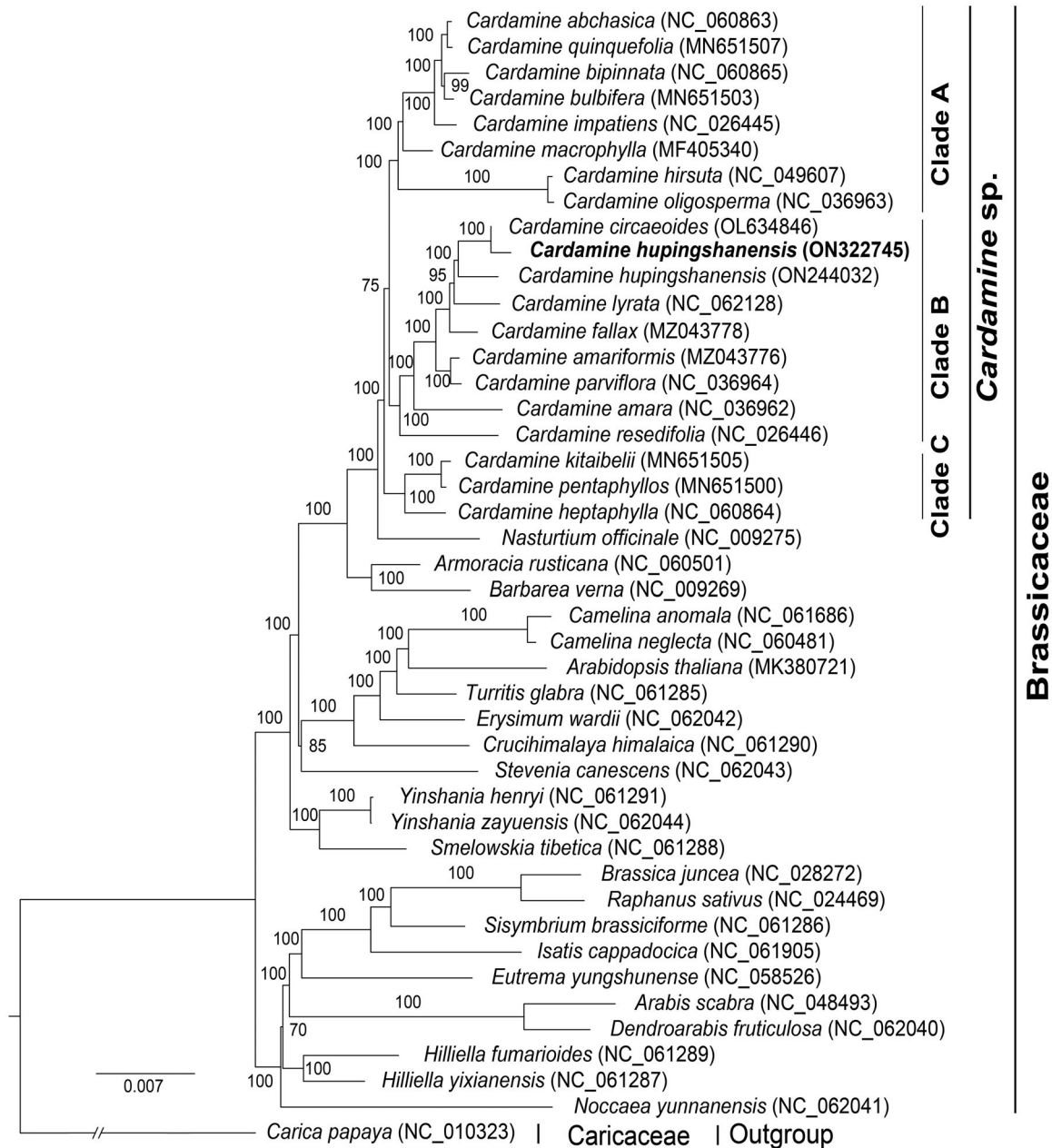


Figure 1. Phylogenetic tree reconstructed by maximum-likelihood (ML) based on the 79 protein-coding genes of chloroplast genome.

6, 5, 5, 5, and 5 from mononucleotide to hexanucleotide, respectively. The simple sequence repeat analysis showed that 143 SSRs were identified in the chloroplast genome of *C. hupingshanensis*. Moreover, the genome size, GC content, gene number, and gene order were similar to other *Cardamine* chloroplast genomes.

In order to explore the phylogenetic relationship of *C. hupingshanensis*, the complete chloroplast genomes of 42 species from Brassicaceae and one species from Caricaceae were obtained from the GenBank database. We used the PhyloSuite v. 1.2.2 (Zhang et al. 2020) to extract 79 protein-coding genes from the chloroplast genome annotation files. Each protein-coding gene sequence was aligned by using MAFFT v. 7.4 (Katoh and Standley 2013), and then 79 aligned sequences were concatenated by using PhyloSuite v. 1.2.2

(Zhang et al. 2020). With the *Carica papaya* as the outgroup, the phylogenetic tree (Figure 1) was constructed by maximum-likelihood (ML) method with IQ-TREE v. 2.1.2 (Nguyen et al. 2015) under the optimal model of GTR + F + R4. The bootstrap value was 1000. The analysis result showed that *Cardamine* and *Nasturtium* were closer, and were monophyletic groups to each other. There were three main branches in the genus *Cardamine*. Clade A included *C. abchasica*, *C. quinquefolia*, *C. bipinnata*, *C. bulbifera*, *C. impatiens*, *C. macrophylla*, *C. hirsuta*, and *C. oligosperma*. Clade B included *C. circaeoides*, *C. hupingshanensis*, *C. lyrata*, *C. fallax*, *C. amariformis*, *C. parviflora*, *C. amara*, and *C. resedifolia*. Clade C included *C. kitaibelii*, *C. pentaphyllos*, and *C. heptaphylla*. In the genus *Cardamine*, *C. hupingshanensis* was closer to *C. circaeoides* and *C. lyrata*. This finding was similar to previous research results (Hu et al. 2015;

Raman and Park 2021, 2022; Raman et al. 2021; Xu et al. 2022). Our results provide valuable data and shed light on the phylogenomic study of *Cardamine* and Brassicaceae.

Ethical approval

In this study, all experimental protocols relating to plant experiments were in accordance with the measures for the Wild Plant Protection Regulations of Henan Province (approved by the Henan Provincial Government in 2007) and Plant Protection Regulations of Hubei Province, China (approved by the Hubei Provincial Government in 2009). The Plant Herbarium of Xinyang Agriculture and Forestry University approved the collection and research of this material. All the research meets ethical guidelines and adheres to the legal requirements of the study country.

Author contributions

Zhang Siying conceptualized and designed research; Liu Xiuqing analyzed data and wrote the manuscript. All authors read and approved the final manuscript.

Disclosure statement

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

Funding

This work was supported by the Key Scientific Research Project Plan of Colleges and Universities in Henan Province [19A220005].

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. ON322745. The associated 'BioProject', 'Bio-Sample', and 'SRA' numbers are PRJNA830645, SAMN27735482, and SRR18884471, respectively.

References

Bai H, Chen L, Liu K, Liu L. 2008. A new species of *Cardamine* (Brassicaceae) from Hunan, China. *Novon*. 18(2):135–137.

Behne D, Kyriakopoulos A. 2001. Mammalian selenium-containing proteins. *Annu Rev Nutr*. 21:453–473.

Beier S, Thiel T, Münch T, Scholz U, Mascher M. 2017. MISA-web: a web server for microsatellite prediction. *Bioinformatics*. 33(16):2583–2585.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–2120.

Butler JA, Beilstein MA, Whanger PD. 1989. Influence of dietary methionine on the metabolism of selenomethionine in rats. *J Nutr*. 119(7):1001–1009.

Cui L, Zhao J, Chen J, Zhang W, Gao Y, Li B, Li Y-F. 2018. Translocation and transformation of selenium in hyperaccumulator plant *Cardamine enschiensis* from Enshi, Hubei, China. *Plant Soil*. 425(1–2):577–588.

Chen L, Song M, Zha H, Li Z. 2014. A modified protocol for plant genome DNA extraction. *Plant Divers Resour*. 36(3):375–380.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissues. *Phytochem Bull*. 19:11–15.

Freeman JL, Bañuelos GS. 2011. Selection of salt and boron tolerant selenium hyperaccumulator *Stanleya pinnata* genotypes and characterization

of Se phytoremediation from agricultural drainage sediments. *Environ Sci Technol*. 45(22):9703–9710.

Freeman JL, Zhang LH, Marcus MA, Fakra S, McGrath SP, Pilon-Smits EAH. 2006. Spatial imaging, speciation, and quantification of selenium in the hyperaccumulator plants *Astragalus bisulcatus* and *Stanleya pinnata*. *Plant Physiol*. 142(1):124–134.

Hu S, Sablok G, Wang B, Qu D, Barbaro E, Viola R, Li M, Varotto C. 2015. Plastome organization and evolution of chloroplast genes in *Cardamine* species adapted to contrasting habitats. *BMC Genomics*. 16(1):306.

Huang C, Ying H, Yang X, Gao Y, Li T, Wu B, Ren M, Zhang Z, Ding J, Gao J, et al. 2021. The *Cardamine enschiensis* genome reveals whole genome duplication and insight into selenium hyperaccumulation and tolerance. *Cell Discov*. 7(1):62.

Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol*. 21(1):241.

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

Pilon-Smits EAH, Quinn CF, Tapken W, Malagoli M, Schiavon M. 2009. Physiological functions of beneficial elements. *Curr Opin Plant Biol*. 12(3):267–274.

Qu X-J, Moore MJ, Li D-Z, Yi T-S. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods*. 15(1):50.

Rafferty TS, Green MHL, Lowe JE, Arlett C, Hunter JAA, Beckett GJ, McKenzie RC. 2003. Effects of selenium compounds on induction of DNA damage by broadband ultraviolet radiation in human keratinocytes. *Br J Dermatol*. 148(5):1001–1009.

Raman G, Park KT, Park S. 2021. The complete chloroplast genome of an endemic plant to Korea, *Cardamine amaraeformis* Nakai.: genome structure and phylogenetic analysis. *Mitochondrial DNA Part B*. 6(9):2725–2726.

Raman G, Park S. 2021. Complete chloroplast genome features and phylogenetic implications of *Cardamine fallax* (O. E. Schulz) Nakai. *Mitochondrial DNA Part B*. 6(9):2722–2724.

Raman G, Park S. 2022. Structural characterization and comparative analyses of the chloroplast genome of Eastern Asian species *Cardamine occulta* (Asian *C. flexuosa* With.) and other *Cardamine* species. *Front Biosci*. 27(4):124.

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.

Sargent M. 2008. Food-chain selenium and human health: spotlight on speciation. *Br J Nutr*. 100(2):254–268.

Sharma S, Bansal A, Dhillon SK, Dhillon KS. 2010. Comparative effects of selenate and selenite on growth and biochemical composition of rapeseed (*Brassica napus* L.). *Plant Soil*. 329(1–2):339–348.

Shao S, Zheng B, Su H, Luo C, Li XY. 2007. A new species of selenium hyperaccumulator identified in Yutangba se deposit area. *Acta Mineral Sin*. 27:567–570.

Simon A. 2020. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

Williams PN, Lombi E, Sun G-X, Scheckel K, Zhu Y-G, Feng X, Zhu J, Carey A-M, Adomako E, Lawgali Y, et al. 2009. Selenium characterization in the global rice supply chain. *Environ Sci Technol*. 43(15):6024–6030.

Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*. 31(20):3350–3352.

Xu X, Yao X, Zhang C, Xia P. 2022. Characterization of the complete chloroplast genome sequence of *Cardamine lyrata* Bunge (Brassicaceae). *Mitochondrial DNA Part B*. 7(6):936–937.

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour*. 20(1):348–355.