

The complete chloroplast genome sequence of *Hippophae rhamnoides* subsp. *sinensis*

Songfeng Diao^{a,b}, Guoyun Zhang^a, Caiyun He^a, Aiguo Duan^a and Jianguo Zhang^{a,c}

^aState Key Laboratory of Tree Genetics and Breeding & Key Laboratory of Tree Breeding and Cultivation, National Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China; ^bNon-timber Forestry Research and Development Center, Chinese Academy of Forestry, Key Laboratory of Non-timber Forest Germplasm Enhancement & Utilization of National Forestry and Grassland Administration, Zhengzhou, China; ^cCollaborative Innovation Center of Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China

ABSTRACT

The complete chloroplast genome sequence of *Hippophae rhamnoides* subsp. *sinensis* was characterized from Illumina pair-end sequencing. The chloroplast genome of *H. rhamnoides* subsp. *sinensis* was 156,355 bp in length, containing a large single-copy region (LSC) of 84,002 bp, a small single-copy region (SSC) of 19,055 bp, and two inverted repeat (IR) regions of 26,649 bp. The overall GC content is 36.6%, while the corresponding values of the LSC, SSC, and IR regions are 64.5%, 69.2%, and 60.1%, respectively. The genome contains 131 complete genes, including 88 protein-coding genes, 38 tRNA genes (29 tRNA species), and 8 rRNA genes (4 rRNA species). The neighbour-joining phylogenetic analysis showed that *H. rhamnoides* subsp. *sinensis* and *H. rhamnoides* clustered together as sisters to other *H. rhamnoides* species.

ARTICLE HISTORY

Received 8 January 2020
Accepted 19 January 2020

KEYWORDS

Hippophae rhamnoides subsp. *sinensis*; chloroplast genome; phylogenetic analysis; genetic information

Introduction

Hippophae rhamnoides subsp. *sinensis* is the most widely distributed subspecies of *Hippophae* Linn. in China (Jia et al. 2012). It not only has the characteristics of drought tolerance, cold resistance, salt and alkali resistance, and barren resistance, but also has the ecological effects of wind and sand fixation, water and soil conservation, and soil improvement (He et al. 2015). Its fruit is rich in vitamin, protein, amino acid, unsaturated fatty acid, flavonoid, a variety of trace elements and other 428 kinds of beneficial nutrients and bioactive substances, with lipid-lowering, anti-inflammatory, anti-cancer and treatment of coronary heart disease, gastrointestinal diseases, liver and kidney protection and other significant effects (Zhang et al. 2019) and enjoys the reputation of 'green gold' (Jimenez-Garcia et al. 2013).

The fresh leaves of *H. rhamnoides* subsp. *sinensis* were collected from Fengning County, Hebei province, P. R. China (41°14'37"N, 116°38'59"E). Fresh leaves were silica-dried and taken to the laboratory until DNA extraction. The voucher specimen (ZGSJ001) was laid in the Herbarium of Research Institute of Forestry and the extracted DNA was stored in the -80 °C refrigerator of the State Key Laboratory of Tree Genetics and Breeding, Key Laboratory of Silviculture of the State Forestry Administration. We extracted total genomic DNA from 25 mg silica-gel-dried leaf using a modified CTAB method (Doyle 1987). The whole-genome sequencing was

then conducted by Biodata Biotechnologies Inc. (Hefei, China) with Illumina HiSeq platform. The Illumina HiSeq 2000 platform (Illumina, San Diego, CA) was used to perform the genome sequence. We used the software MITObim 1.8 (Hahn et al. 2013) and metaSPAdes (Nurk et al. 2017) to assemble chloroplast genomes. We used *H. rhamnoides* (GenBank: NC_035548) as a reference genome. We annotated the chloroplast genome with the software DOGMA (Wyman et al. 2004), and then corrected the results using Geneious 8.0.2 (Campos et al. 2016) and Sequin 15.50 (<http://www.ncbi.nlm.nih.gov/Sequin/>).

The complete chloroplast genome of *H. rhamnoides* subsp. *sinensis* (GenBank accession number MN857418) was 156,355 bp in length, containing a large single-copy region (LSC) of 84,002 bp, a small single-copy region (SSC) of 19,055 bp, and two inverted repeat (IR) regions of 26,649 bp. The overall GC content is 36.6%, while the corresponding values of the LSC, SSC, and IR regions are 64.5%, 69.2%, and 60.1%, respectively. The genome contains 131 complete genes, including 88 protein-coding genes, 38 tRNA genes (29 tRNA species) and 8 rRNA genes (4 rRNA species).

We used the complete chloroplast genomes sequence of *H. rhamnoides* subsp. *sinensis* and other related species of Elaeagnaceae, *Cannabis sativa* and *Humulus lupulus* as out-group to construct phylogenetic tree. The 10 chloroplast genome sequences were aligned with MAFFT (Kato and Standley 2013), and then the Neighbour-joining tree was

CONTACT Jianguo Zhang  zhangjg@caf.ac.cn  State Key Laboratory of Tree Genetics and Breeding & Key Laboratory of Tree Breeding and Cultivation, National Forestry and Grassland Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China

© 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.

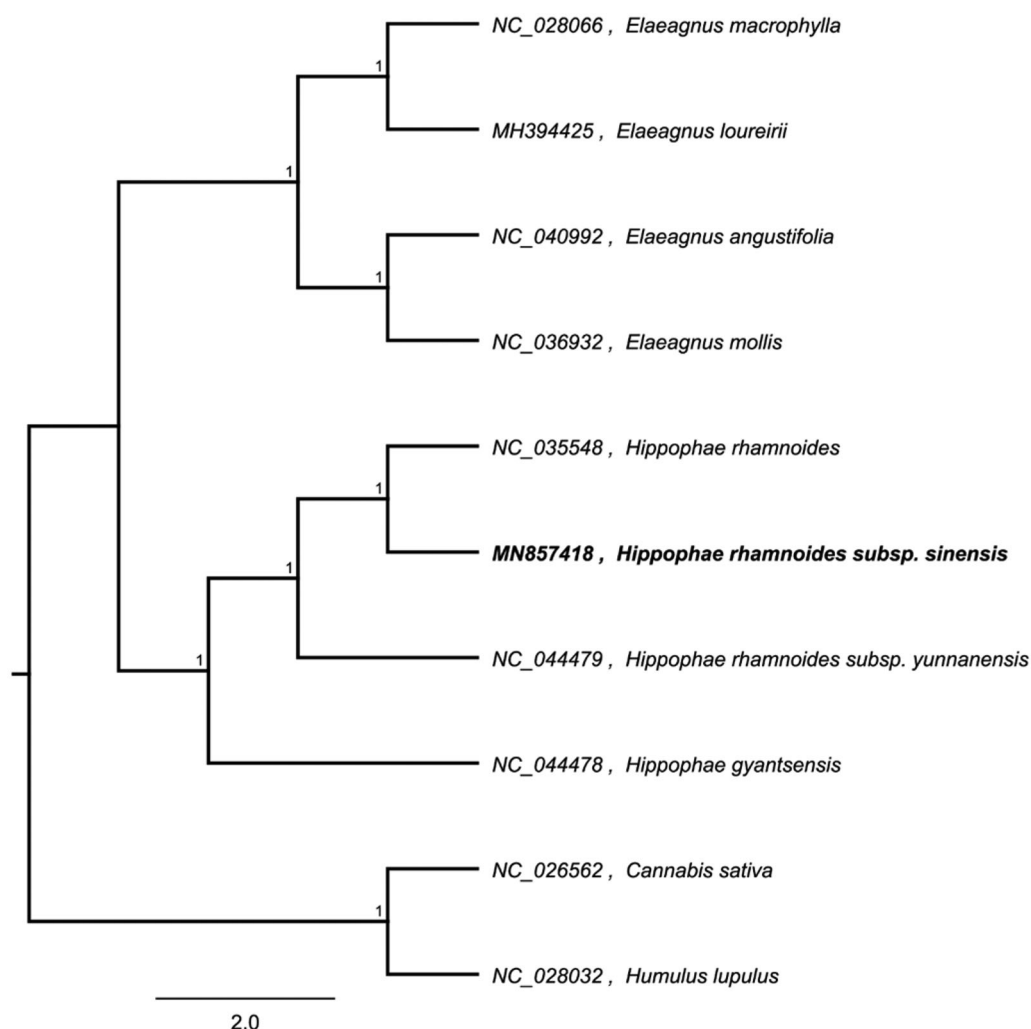


Figure 1. Neighbour-joining (NJ) analysis of *H. rhamnoides* subsp. *sinensis* and other related species based on the complete chloroplast genome sequence.

constructed by MEGA 7.0 (Kumar et al. 2016). The results confirmed that *H. rhamnoides* subsp. *sinensis* and *Hippophae rhamnoides* clustered together as sisters to other *H. rhamnoides* species (Figure 1).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

his research was provided by Special Fund from Key Laboratory of Tree Breeding and Cultivation, National Forestry and Grassland Administration [ZDRIF2019].

References

- Campos F, Kluge M, Franco A, Giongo A, Valdez F, Saggi T, Roehle P. 2016. Complete genome sequence of porcine parvovirus 2 recovered from swine sera. *Genome Announce*. 4(1):e01627–01615.
- Doyle J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19(1):11–15.
- Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads – a baiting and iterative mapping approach. *Nucleic Acids Res*. 41(13): e129–e129.
- He C, Li M, Luo H, Gao G, Zhang J. 2015. Comprehensive evaluation on drought resistance of *Hippophae rhamnoides*. *Forest Res*. 28(5): 634–639.
- Jia D, R, Abbott R, J, Liu T, L, Mao K, S, Bartish I, V, Liu J, Q. 2012. Out of the Qinghai–Tibet Plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of *Hippophae rhamnoides* (Elaeagnaceae). *New Phytologist*. 194(4):1123–1133.
- Jimenez-Garcia SN, Guevara-Gonzalez RG, Miranda-Lopez R, Feregrino-Perez A, A, Torres-Pacheco I, Vazquez-Cruz M. A. 2013. Functional properties and quality characteristics of bioactive compounds in berries: biochemistry, biotechnology, and genomics. *Food Res Int*. 54(1): 1195–1207.
- 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.
- Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. 27(5): 824–834.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20(17):3252–3255.
- Zhang G, Diao S, Zhang T, Chen D, He C, Zhang J. 2019. Identification and characterization of circular RNAs during the sea buckthorn fruit development. *RNA Biol*. 16(3):354–361.