


Chloroplast genome features and phylogenomic placement of *Sorbus hupehensis* var. *paucijuga* (Rosaceae)

Dan Liu^{a,b}, Xiao-Man Xie^b, Yi-Zeng Lu^b , Feng Li^c, Jian Zheng^d, Hai-Ping Yang^b, Kun Liu^b, Xiao-bo Zhang^a and Wen-Qing Li^b

^aCollege of Biological Sciences and Biotechnology, Beijing Forestry University, Beijing, People's Republic of China; ^bShandong Forest Germplasm Resources Center, Ji'nan, Shandong Province, People's Republic of China; ^cYishan State owned Forest Farm of Yishui County, Linyi, Shandong Province, People's Republic of China; ^dCollege of Landscape Architecture, Beijing University of Agriculture, Beijing, People's Republic of China

ABSTRACT

The complete chloroplast genome of *Sorbus hupehensis* var. *paucijuga* was sequenced with Illumina HiSeq 2000 platform. It was a typical quadruple structure as other plants of *Sorbus* with 160,050 bp in length, including a large single-copy (LSC: 87,905 bp) region and a small single-copy (SSC: 19,325 bp) which were separated by a pair of inverted repeats (IRa, b: 26,410 bp) region. The overall GC content is 36.5%. A total of 130 genes was annotated which contained 85 protein-coding genes including the Trans splicing gene of *rps12*, 37 tRNA genes, and 8 rRNA genes. ML phylogenetic analysis compared with 7 expressed chloroplast genomes of Rosaceae revealed that *S. hupehensis* var. *paucijuga* was a sister to other *Sorbus* species. Six species of *Sorbus* were divided into two groups, the species of group one is distributed in Asia and the species of group two distributed in Europe. Among group one, *S. hupehensis* var. *paucijuga* had the closest genetic relationship with *S. ulleungensis* which is a New Endemic Species on Ulleung Island of Korea, and followed by *S. setschwanensis* which is only distributed in Sichuan and Guizhou of China. *Sorbus hupehensis* var. *paucijuga* has a relatively close relationship with the other three species of *Sorbus* in the group two. And, it has a relatively distant from other genera of *Prunus mongolica* and *Rosa rugosa*.

ARTICLE HISTORY

Received 2 September 2020
Accepted 20 February 2021

KEYWORDS

Chloroplast genome;
phylogenomic placement;
Sorbus hupehensis
var. *paucijuga*

Sorbus hupehensis var. *paucijuga* (D. K. Zang & P. C. Huang) L. T. Lu is an endangered endemic variety only distributed in Shandong Province of China (Li et al. 2016) and it is a multi-purpose tree species of *Sorbus* in Rosaceae. After the publication of this species (Zang and Huang 1992), scholars revised it as a variety of *S. hupehensis* according to its fruit color and other characteristics (Lu and Spongberg 2003; Lu 2000). However, due to the differences in habitat, the number of leaflets and the characteristics of leaf veins, the classification status of *Sorbus hupehensis* var. *paucijuga* is still controversial (Xiong 2019). Related studies showed that its karyotype was 2B (Peng 2016), and proper shading was beneficial to its seedling growth (Zhao et al. 2019), but there was no molecular biological evidence. The discovery of genomics resources of *S. hupehensis* var. *paucijuga* is helpful to further clarify its taxonomic status and its introduction and cultivation.

The fresh leaves of *S. hupehensis* var. *paucijuga* were collected from the living individual permanently conserved in the *Sorbus* gene bank (36.632°N, 117.178°E) and its Provenance was Laoshan Mountain in Shandong Province. The specimens were preserved in National Plant Specimen Resource Center (<http://www.cvh.ac.cn/>, barcode

SDF1005913). Total genomic DNA (saved in DNA library of Shandong Forest Germplasm Resources Center with the code of bjhj2020cp02) was extracted by the Plant DNA extraction Kit (TIANGEN, Beijing, China) according to the requirements of the reagent company. Paired-end reads were constructed according to the Illumina library preparation protocol and sequenced on an Illumina HiSeq 2000 platform. The whole chloroplast genome of *S. hupehensis* var. *paucijuga* was assembled by MITObim v1.8 (Hahn et al. 2013) and was annotated in DOGMA (<http://dogma.cccb.utexas.edu/>). The whole chloroplast genome of *S. hupehensis* var. *paucijuga* and other 8 published plastomes of Rosaceae were conducted by using MAFFT v7.429 (Kato and Standley 2013), which included five species from Maloideae, one species from the subfamily of Rosoideae and Prunoideae respectively, and with *Pteroceltis tatarinowii* as outgroup. Maximum-likelihood (ML) phylogenetic tree with 1000 bootstrap replicates was inferred using IQ-TREE v1.6.12 (Lam-Tung et al. 2015) and TVM + F + R2 model.

The chloroplast genome of *S. hupehensis* var. *paucijuga* (GenBank accession number MT916771) was also a typical quadruple structure with 160,050 bp in length that contains a

CONTACT Wen-Qing Li  wenqingli203@126.com  Shandong Forest Germplasm Resources Center, Ji'nan 250102, China; Jian Zheng  buazhengjian708@126.com  College of Landscape Architecture, Beijing University of Agriculture, Beijing 102206, People's Republic of 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-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

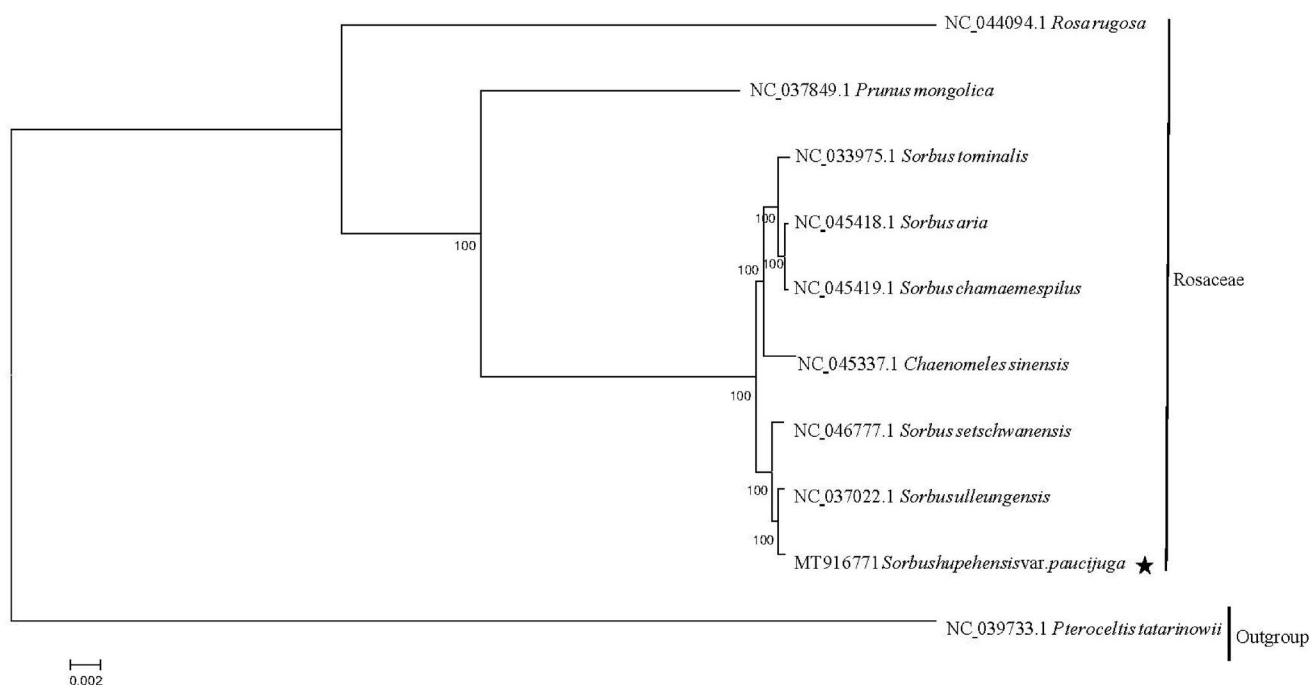


Figure 1. A maximum-likelihood (ML) tree of *S. hupehensis* var. *paucijuga* and other 7 related species based on the complete chloroplast genome sequences with *Pteroceltis tatarinowii* as outgroup. The accession numbers are showed in the figure, the numbers behind each node are bootstrap support values.

large single copy (LSC: 87,905 bp) region and a small single copy (SSC: 19,325 bp), which were separated by a pair of inverted repeats (IRa, b: 26,410 bp) region. The overall GC content was 36.5%. A total of 130 genes were annotated, including 85 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Among these, 15 genes had a single intron respectively, *clpP* and *ycf3* had two introns respectively. While *rps12* had Trans splicing function. ML phylogenetic analysis compared with 7 expressed chloroplast genomes of Rosaceae revealed that *S. hupehensis* var. *paucijuga* was a sister to other *Sorbus* species. Six species of *Sorbus* were divided into two groups, the species of group one is distributed in Asia and the species of group two distributed in Europe. Among group one, *S. hupehensis* var. *paucijuga* had the closest genetic relationship with *S. ulleungensis* (Chang and Gil 2014) which is a New Endemic Species on Ulleung Island of Korea, and followed by *S. setschwanensis* which only distributed in Sichuan and Guizhou of China. *S. hupehensis* var. *paucijuga* has a relatively close relationship with the other three species of *Sorbus* in the group two. And it has a relatively distant from other genera of *Prunus mongolica* and *Rosa rugosa* (Figure 1).

Disclosure statement

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

Funding

The study was financially supported by the National Natural Fund Project [No. 31770369]; Construction and Operation Project of Shared Service Platform for National Forest Tree Germplasm Resources [2005-DKA21003]; Wild Plant Protection and Management Project of 'Conservation of Germplasm Resources of Chinese Endemic Species of National Key Protected Wild Plants' [2019073031]; 'Collection,

Conservation, and Accurate Identification of Forest Tree Germplasm Resources' of Shandong Provincial Agricultural Elite Varieties Project [2019LZGC018].

ORCID

Yi-Zeng Lu  <http://orcid.org/0000-0002-5198-0155>

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. MT 916771. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA694828, SRR13528764, and SAMN17575002 respectively.

References

- Chang CS, Gil HY. 2014. *Sorbus ulleungensis*, a New Endemic Species on Ulleung Island, Korea. *Harvard Papers Bot.* 19(2):247–255.
- 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.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Lam-Tung N, 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.
- Li W, Zang D, Xie X. 2016. Rare and endangered protected tree species in Shandong. Beijing: Science Press.
- Lu L. 2000. Notes on species of some genera of Rosaceae, China. *Acta Phytotaxonom Sin.* 38(03):276–281.
- Lu L, Spongberg SA. 2003. *Sorbus* Linnaeus. In: Wu ZY, Raven PH, Hong DY, editors. *Flora of China: Rosaceae*. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press; p. 144–170.

- Peng H. 2016. Morphological and cytological research on species in *Sorbus* L. Nanjing of China's Jiangsu Province: Nanjing Forestry University.
- Xiong Z. 2019. Leaf characters of Section *Sorbus* from China and their taxonomic significance. Nanjing of China's Jiangsu Province: Nanjing Forestry University.
- Zang D, Huang P. 1992. New taxa of Rosaceae from Shandong. Bull Bot Res. 12(04):321–323.
- Zhao Y, Chen Y, Zhang X, Xiong Z, Chen X. 2019. Effects of shading on photosynthetic and chlorophyll fluorescence response of *Sorbus hupehensis* var. *paucijuga*. J Northeast Forest Univ. 47(01):30–34.