

The complete chloroplast genome sequence of *Prunus salicina* 'Wushan plum'

Bo Fang^a, Yicen Xu^b and Jie Yu^b

^aChongqing Academy of Agricultural Sciences, Chongqing, China; ^bCollege of Horticulture and Landscape Architecture, Southwest University, Chongqing, China

ABSTRACT

Prunus salicina 'Wushan plum' is a local economic fruit crop. In this study, we reported the complete chloroplast genome sequence of *P. salicina* 'Wushan plum'. The genome has a circular structure of 157,921 bp containing a large single-copy region (LSC) of 86,184 bp, a small copy region (SSC) of 19,031 bp, and two inverted repeats (IR) of 26,353 bp by each. It harbors 110 unique genes, including 78 protein-coding genes, 4 ribosomal RNA genes, and 28 transfer RNA genes. The phylogenomic analysis shows that *P. salicina* 'Wushan plum' is clustered with *Prunus salicina*.

ARTICLE HISTORY

Received 22 February 2021
Accepted 13 March 2021

KEYWORDS

Prunus salicina 'Wushan plum'; chloroplast genome; phylogenetic analysis

Prunus salicina 'Wushan plum' is the local characteristic variety of Chongqing. It belongs to *Prunus* of Rosaceae. It is a kind of Chinese plum (*Prunus salicina* L.). 'Wushan plum' has a good yield and strong adaptability. The fruit has good characters in shape and color, fragrant, sour and sweet, and nutritious (Zhao et al. 2019).

P. salicina 'Wushan plum' was collected from Wushan, Chongqing. The DNA library was constructed using the Agilent 2100 and sequenced by using the Illumina NovaSeq 6000 sequencing platform.

The chloroplast genome was assembled from the clean data by GetOrganelle (v. 1.6.4) (Jin et al. 2020). The annotation of the chloroplast genome was conducted initially using CpGAVAS2 (Linchun et al. 2019). Furthermore, the annotations with problems were manually edited by using Apollo (Misra and Harris 2006). The genome sequence and annotations have been deposited in the GenBank with accession number MW406458.

The chloroplast genome of *P. salicina* 'Wushan plum' is 157,921 bp in size with a large single-copy region (LSC) of 86,184 bp, small copy region (SSC) of 19,031 bp and two inverted repeats (IRs) of 26,353 bp by each. The chloroplast genome of *P. salicina* 'Wushan plum' comprises 131 genes, among which, 110 are unique genes, including 78 protein-coding genes, 4 ribosomal RNA (rRNA) genes, and 28 transfer RNA (tRNA) genes. Among the 78 protein-coding genes annotated, nine unique genes contain only one intron, two genes (*ycf3*, *clpP*) contain two introns, and six tRNA genes contain one intron. The length of the protein-coding sequence (CDS) in the chloroplast genome of *P. salicina* 'Wushan plum' is 79,728 bp. In contrast, the length of the rRNA genes is 9,048 bp, and the length of the tRNA genes is

2,815 bp. The GC content analysis showed that the overall GC content is 36.72%. Note that the GC contents in IR regions (42.62%) are significantly higher than that in LSC (34.51%) and SSC regions (30.37%).

To examine the phylogenetic position of *P. salicina* 'Wushan plum', we constructed the maximum likelihood (ML) trees by using the *P. salicina* 'Wushan plum' and other 29 *Prunus* species and two outgroups complete chloroplast genome sequences. Two species, *Malus baccata* (Rosaceae) and *Prinsepia utilis* (Rosaceae) were selected as outgroups. The complete chloroplast genome sequences were aligned by using MAFFT (<https://mafft.cbrc.jp/alignment/server/>) online version 7.471 (John et al. 2019). These aligned sequences were used to construct the maximum likelihood tree by RaxML (v8.2.4) (Stamatakis 2014). The phylogenetic analysis showed *P. salicina* 'Wushan plum' was clustered with *Prunus salicina* (Figure 1).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The Performance Incentive & Guidance Project Special for the Scientific Research Institutions of Chongqing [cstc2019xjl80012].

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 number MW406458. The associated 'BioProject',

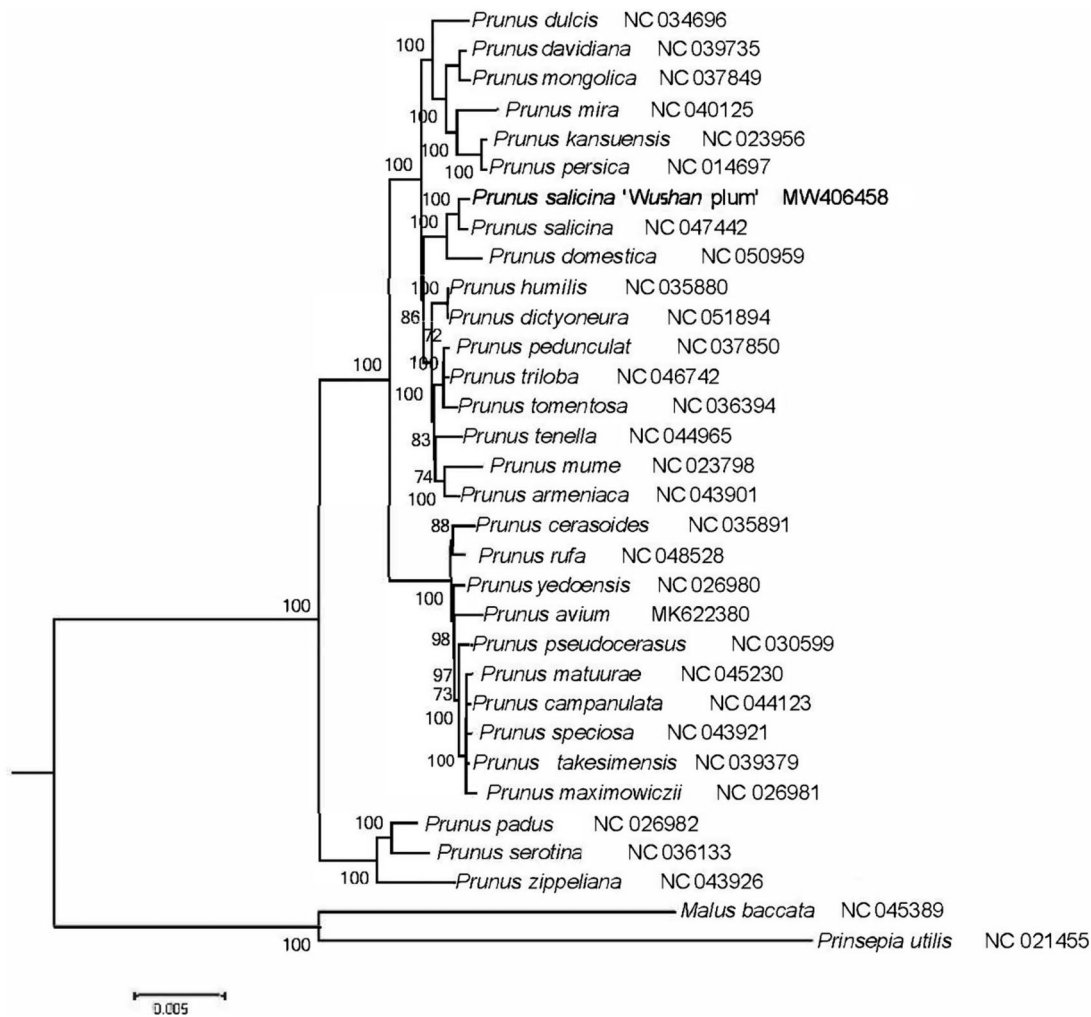


Figure 1. Maximum-likelihood phylogenetic tree for genus *Prunus* L. based on 32 complete chloroplast genomes. Values along branches correspond to ML bootstrap percentages.

'SRA', and 'Bio-Sample' numbers are PRJNA699156, SRR13627323, and SAMN17767268 respectively.

References

- 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.
- John R, Songling L, Mar AK, Standley DM, Kazutaka K. 2019. MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* 47 (W1):W5–W10.
- Linchun S, Haimei C, Mei J, Liqiang W, Xi W, Linfang H, Chang L. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73.
- Misra S, Harris N. 2006. Using Apollo to browse and edit genome annotations. *Curr Protoc Bioinformatics.* 12:9.5.1–9.5.28.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9): 1312–1313.
- Zhao Q, Fang B, Niu Y. 2019. Effects of 1-MCP treatment on fruit quality and cell wall components of 'Wushan' plum fruits during cold storage. *Southwest Chin J Agric Sci.* 32(12):2933–2938.