Taylor & Francis Taylor & Francis Group

MITOGENOME ANNOUNCEMENT



The complete chloroplast genome sequence of Prunus simonii 'Weiwang'

Yicen Xua*, Bo Fangb* and Jie Yua

^aCollege of Horticulture and Landscape Architecture, Southwest University, Chongging, China; ^bChongging Academy of Agricultural Sciences, Chongqing, China

ABSTRACT

Prunus simonii 'Weiwang' is an important economic fruit crops. In this study, we reported the complete chloroplast genome sequence of P. simonii 'Weiwang'. The genome has a circular structure of 157,924 bp containing a large single-copy region (LSC) of 86,187 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 Prunus simonii 'Weiwang' is clustered with Prunus salicina.

ARTICLE HISTORY

Received 6 April 2021 Accepted 26 June 2021

KEYWORDS

Prunus simonii 'Weiwang'; chloroplast genome; phylogenetic analysis

Prunus simonii 'Weiwang' is a successful hybrid cultivated by American agronomists over half a century. It belongs to Prunus of Rosaceae. The fruit has good characters in shape and color, fragrant, sour and sweet in taste, and nutritious. It comes into the market in summer, with early ripening and high yield, which is favored by consumers and growers (Xianzhong et al. 2011).

Prunus simonii 'Weiwang' was collected from Sanjiao town, Qijiang District, 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 correctness of the assembly was confirmed by using Bowtie2 (v2. 0.1) (Langmead et al. 2009). 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 MW406463.

The chloroplast genome of P. simonii 'Weiwang' is 157,924 bp in size with a large single-copy region (LSC) of 86,187 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. simonii 'Weiwang' 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 (rps16, atpF, rpoC1, petB, petD, rpl16, rpl2, ndhB, ndhA) contain only one intron, two genes (ycf3, clpP) contain two introns, and six tRNA genes (trnK-UUU, trnG-GCC, trnL-UAA, trnV-UAC, trnl-GAU, trnA-UGC) contain one intron. 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. simonii 'Weiwang', we constructed the maximum likelihood (ML) trees using the complete chloroplast genomes of P. simonii 'Weiwang', other 29 Prunus species, and two 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. simonii 'Weiwang' 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 Chongging [cstc2019jxjl80012].

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 MW406463. The associated "BioProiect," "SRA," and "Bio-Sample" numbers are PRJNA719267, SRR14133449, and SAMN18593728 respectively."

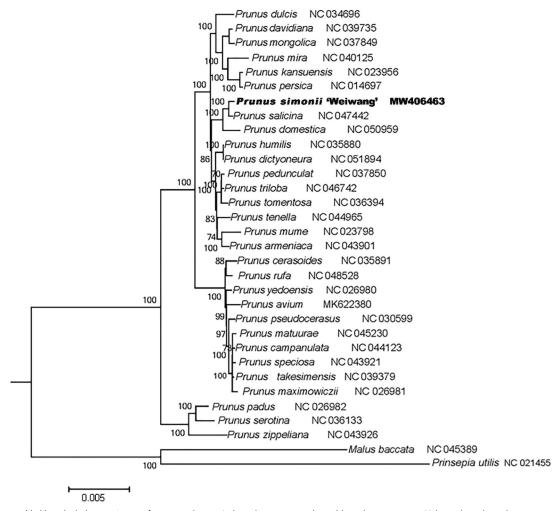


Figure 1. Maximum-likelihood phylogenetic tree for genus *Prunus* L. based on 32 complete chloroplast genomes. Values along branches correspond to ML bootstrap percentages.

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.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10(3):R25.

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.

Xianzhong L, Zhiming Y, Yi Y. 2011. Cultivation status and development countermeasure of *Prunus simonii* in China. Forest Invent Plann. 36(02):120–123.