

The complete chloroplast genome of *Prunus domestica* L. (Rosaceae) and its phylogenetic implication

Wenjuan Geng^{a,b}, Liting Ouyang^{a,b}, Min Xu^{a,b} and Jianyong Jing^{a,b}

^aCollege of Horticulture and Forestry, Xinjiang Agricultural University, Urumqi, China; ^bLaboratory of Germplasm Resources and Efficient Production of Horticultural Crops in Xinjiang, Urumqi, China

ABSTRACT

Prunus domestica commonly known as European plum is one of the most important wild fruit tree resources, and up to date, its wild community has only been found to be distributed in wild fruit forest area of the Tianshan Mountains in Ili region of Xinjiang, China. Despite its agricultural importance and long history of cultivation, many questions remain about the origin of this species, due to absence of genome data. In this study, the complete chloroplast (cp) genome sequence of *P. domestica* was determined using next-generation sequencing. The entire cp genome was determined to be 157,395 bp in length. It contained large single-copy (LSC) and small single-copy (SSC) regions of 85,744 and 18,949 bp, respectively, which were separated by a pair of 26,351 bp inverted repeat (IR) regions. The genome contained 130 genes, including 85 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The overall GC content of the genome is 36.8%. A phylogenetic tree reconstructed by 39 chloroplast genomes reveals that *P. domestica* is most closely related to *Prunus salicina*.

ARTICLE HISTORY

Received 2 May 2020
Accepted 9 May 2020

KEYWORDS

Prunus domestica; complete chloroplast genome; phylogenetic analysis; Rosaceae

Prunus L. s.l. containing six subgenera, is mainly distributed in the northern hemisphere, including *P. domestica*, one of the most important wild fruit tree resources (Chase et al. 2016; Zhang et al. 2017). Up to date, its wild community has only been found to be distributed in wild fruit forest area of the Tianshan Mountains in Ili region of Xinjiang, China (Lin and Shi 1989; Zhang et al. 1998). Despite its agricultural importance and long history of cultivation, many questions remain about the origin of this species, due to absence of genome data. So, it is necessary to develop genomic resources for *P. domestica* to provide basic intragenic information for the further study on phylogeny and breeding for genus *Prunus*.

Healthy leaf samples were collected from a wild *P. domestica* plant (43.38889 N, 83.59198 E, altitude 1235 m). The total genomic DNA was extracted from the fresh leaves of *P. domestica* using the DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). The voucher specimen was deposited at Xinjiang agricultural university characteristic fruit tree research center (19sd-c-022). The whole genome sequencing was conducted on the Illumina HiSeq 4000 Sequencing System (Illumina, Hayward, CA). The filtered sequences were assembled using the program SPAdes assembler 3.10.0 (Bankevich et al. 2012). Annotation was performed using the DOGMA (Wyman et al. 2004), and tRNAscan (Schattner et al. 2005).



The plastome of *P. domestica* was determined to comprise double stranded, circular DNA of 157,395 bp containing two inverted repeat (IR) regions of 26,351 bp each, separated by

large single-copy (LSC) and small single-copy (SSC) regions of 85,744 and 18,949 bp, respectively (Genome Warehouse acc. no. GWHALOD01000000; GenBank acc. no. MT302569). The genome contained 130 genes, including 85 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The six protein-coding genes, six tRNA genes and four rRNA genes were duplicated in IR region. nineteen genes contained two exons and four genes (*clpP* and *ycf3* and two *rps12*) contained three exons. The overall GC content of *P. domestica* cp genome is 36.8% and the corresponding values in LSC, SSC and IR regions are 34.6, 30.5 and 42.6%, respectively.

To investigate its taxonomic status, a maximum likelihood (ML) was reconstructed based on whole chloroplast genomes from 37 *Prunus* plants and two outgroup plant (*Neillia gracilis* and *Neillia serratisepala*) by Mafft version 1.4 and FastTree version 2.1.10 (Price et al. 2010). The ML phylogenetic tree shows that *P. domestica* is most closely related to *Prunus salicina*, with bootstrap support values of 100%. Chloroplast genomes of *P. domestica* adds valuable information for understanding the phylogenetic position of *P. domestica* in the subgenera *Prunus*. And more genome data of *Prunus* are needed to reveal the origin of *P. domestica* (Figure 1).

Disclosure statement

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

CONTACT Wenjuan Geng  gwj0526@163.com  College of Horticulture and Forestry, Xinjiang Agricultural University, Urumqi, China; Laboratory of Germplasm Resources and Efficient Production of Horticultural Crops in Xinjiang, Urumqi, 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.

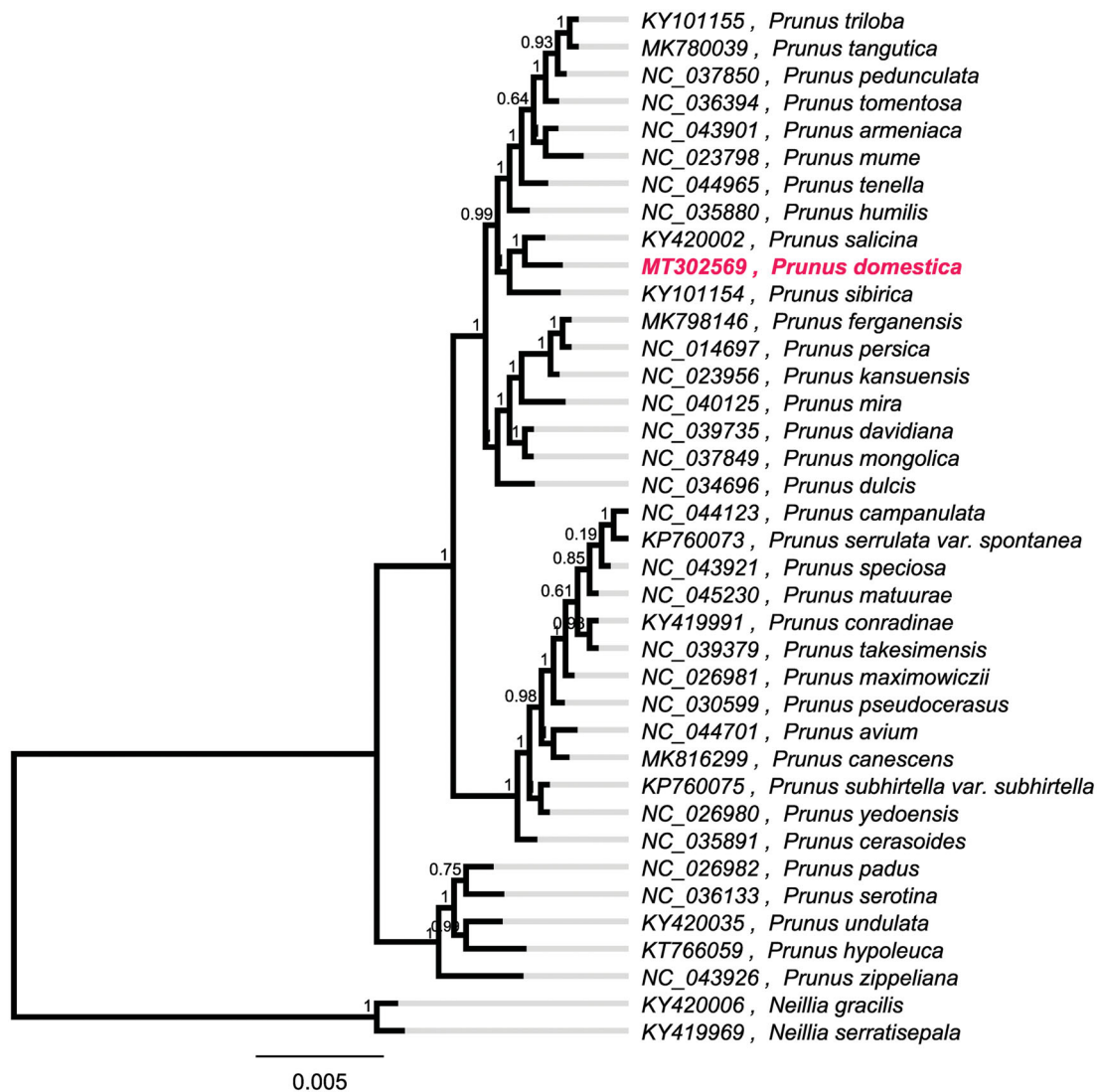


Figure 1. Maximum-likelihood phylogenetic tree based on whole chloroplast genomes from 37 *Prunus* plants and two outgroup plant (*Neillia gracilis* and *Neillia serratisepala*) and the support values are shown at the branches.

Data availability

The complete chloroplast genome sequence of *Prunus domestica* is deposited in the Genome Warehouse (<https://bigd.big.ac.cn/search?dbld=gwh&q=GWHALOD01000000&page=1>).

Funding

This research was supported by The National Natural Science Foundation of China [31560545], Xinjiang Uygur Autonomous Region Horticulture Key Discipline Fund [2016-10758-3], Xinjiang Agricultural University characteristic fruit development and regulation innovation team.

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.

Chase MW, Christenhusz MJM, Fay MF, et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc.* 181(1):1–20.

Lin P, Shi L. 1989. The discovery and distribution of Ili wild *Prunus domestica* (*P. communis* Fritsch) in Xinjiang. *Proceedings of International Symposium on Horticultural Germplasm, Cultivated and Wild*; Beijing (China). p. 282–286.

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2-approximately maximum-likelihood trees for large alignments. *PLoS One.* 5(3):e9490.

Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* 33(Web Server issue):W686–W689.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20(17):3252–3255.

Zhang JY, Liu WS, Lin PJ, Blanchet P. 1998. Preliminary report about wild *Prunus domestica* L. in Xinjiang (Western China). *Acta Hortic.* 478(478):347–350.

Zhang S-D, Jin J-J, Chen S-Y, Chase MW, Soltis DE, Li H-T, Yang J-B, Li D-Z, Yi T-S. 2017. Diversification of Rosaceae since the Late Cretaceous based on plastid phylogenomics. *New Phytol.* 214(3):1355–1367.