

## Complete chloroplast genome of *Prunus canescens*: an endemic shrub in China

Xin Chen<sup>a</sup>, Xiaojuan Zong<sup>a</sup>, Qingzhong Liu<sup>a</sup>, Boqiang Tong<sup>b</sup>, Li Xu<sup>a</sup>, Lisi Zhang<sup>a</sup> and Po Hong<sup>a</sup>

<sup>a</sup>Shandong Provincial Key Laboratory of Fruit Tree Biotechnology Breeding, Shandong Institute of Pomology, Shandong Academy of Agricultural Sciences, Taian, Shandong, P. R. China; <sup>b</sup>Shandong Forest Germplasm Resources Center, Jinan, Shandong, P. R. China

### ABSTRACT

*Prunus canescens* is an endemic cherry species in China, which is distributed in Shaanxi, Gansu, Hubei, and Sichuan provinces of China. The chloroplast (cp) genome of *P. canescens* is 157,890 bp in size containing 125 unique genes, including 8 rRNA genes, 37 tRNA genes, and 80 protein-coding genes (PCGs). Phylogenetic analysis exhibited that *P. canescens* is most related to *P. pseudocerasus*.

### ARTICLE HISTORY

Received 11 May 2019  
Accepted 17 May 2019

### KEYWORDS

*Prunus canescens*;  
chloroplast genome;  
Illumina sequencing

*Prunus canescens*, an endemic species in China, is a species of cherry of the family Rosaceae. It is a deciduous shrubby tree and reaches a height of 2 m. It is distributed in Shaanxi, Gansu, Hubei, and Sichuan province of China (Istituto Botanico Boreali-Occidentali Academiae Sinicae 1974). The altitude of its natural mountain habitat is between 1300 and 1600 m. *Prunus canescens* is a parent of 'GiSeLa', which is one of the most popular rootstocks all over the world (Whiting et al. 2005). Hence, the genomic sequence information is urgently needed to promote molecular evolution, systematics research, conservation, and utilization of *P. canescens*. The objectives of the present study were to reconstruct the cp genome of *P. canescens* and assess phylogenetic relationships among different species within the family Rosaceae.

Leaves were sampled from a mature *P. canescens* tree at Maiji Mountain, Maiji, Tianshui, Gansu, China (34°21'7.91"N, 106°0'14.89"E) and chilled with liquid nitrogen immediately. The voucher specimen (accession no. TS\_2019\_Maiji\_Taian) was stored at -80°C in Shandong Institute of Pomology (SDIP). Genomic DNA (gDNA) was obtained from homogenized leaf tissues using a modified CTAB protocol (Doyle and Doyle 1987). The quantity and quality of the purified gDNA were detected by Nanodrop 8000 and via the Agilent 2100 Bioanalyzer. A library with 350 bp fragments inserted was constructed with 1 µg purified DNA and high-throughput sequenced with paired end (PE) reads of 2 × 150 bp on Illumina HiSeq 2500 platform. Raw reads were filtered and trimmed to remove low quality and contaminated reads by trim\_galore v0.4.4. Totally 7.9 Gb of clean data were aligned to the *Prunus cerasoides* complete cp genome (GenBank no. NC\_035891) (Xu et al. 2018) as a reference using bowtie2 v2.2.4 (Langmead and Salzberg 2012) and assembled with SPAdes v3.10.1 (Bankevich et al. 2012). The final cp genome

was annotated using DOGMA (Boore et al. 2004), HMMER 3.1b2 (Finn et al. 2011), and ARAGORN v1.2.38 (Laslett and Canback 2004).

The cp genome of *P. canescens* (GenBank no. MK816299) is 157,890 bp in size with total AT content of 63.3%. It contains a 19,147 bp small and 85,909 bp large single-copy regions with AT contents 65.4 and 69.9%, respectively, and two 26,417 bp inverted repeat regions with AT content 57.5%. In the cp genome of *P. canescens*, there are 125 unique genes, including 8 rRNA genes, 37 tRNA genes, and 80 PCGs. Thirteen genes, including seven PCGs (*rps19*, *rpl23*, *rpl2*, *psbA*, *ndhC*, *ndhI* and *atpA*), harbour one intron each, while two protein-coding genes (*psaA* and *rpl20*) harbour two introns each.


To perform the molecular phylogenetic analysis, 15 published complete cp genomes were aligned by MAFFT v7.307 (Kato and Standley 2013). Finally, a maximum likelihood (ML) tree was constructed using RAxML v7.2.6 with the GTRGAMMA model (Stamatakis 2006). The ML phylogenetic tree shows that *P. canescens* is most related to *P. pseudocerasus* (Figure 1).

### Disclosure statement

No potential conflict of interest was reported by the authors.

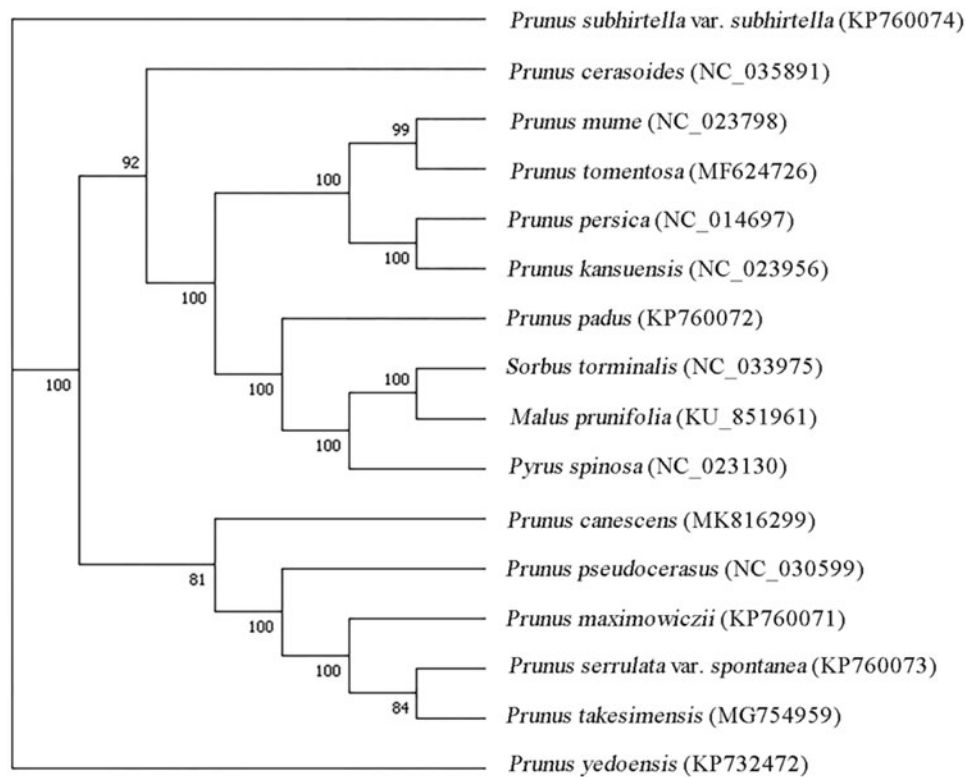
### Funding

This work was supported by the National Natural Science Foundation of China [No. 31601732] and Forestry Science & Technology Innovation Project of Shandong Province [LYCX01-2018-01 and LYCX04-2018-19]. Special Found for Fruit Innovation Team of Shandong Modern Agricultural Technology System, Shandong, China [Grant/Award Number: SDAIT-06-04].

**CONTACT** Xiaojuan Zong  285609317@qq.com; Qingzhong Liu  qzliu001@126.com  Shandong Provincial Key Laboratory of Fruit Tree Biotechnology Breeding, Shandong Institute of Pomology, Shandong Academy of Agricultural Sciences, Taian, Shandong 271000, P. R. China

© 2019 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.** Phylogenetic tree based on 16 complete cp genome sequences. The bootstrap support values are shown next to the branches.

## 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.* 19:455–477.
- Boore JL, Jansen RK, Wyman SK. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.
- Doyle JJ, Doyle JD. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19:11–15.
- Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39:W29–W37.
- Instituto Botanico Boreali-Occidentali Academiae Sinicae. 1974. *Flora Tsinlingensis*, (Tomus I): Spermatophyta (Pars 2). Beijing: Science Press.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30:772–780.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods.* 9:357.
- Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32:11–16.
- Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 22:2688–2690.
- Whiting MD, Lang G, Ophardt D. 2005. Rootstock and training system affect sweet cherry growth, yield, and fruit quality. *HortScience.* 40:582–586.
- Xu X, Wen J, Wang W, Zheng W. 2018. The complete chloroplast genome of the threatened *Prunus cerasoides*, a rare winter blooming cherry in the Himalayan region. *Conserv Genet Resour.* 10:499–502.