The complete chloroplast genome sequence of *Styrax serrulatus* Roxburgh (Styracaceae)

Lu Tian\textsuperscript{a,b}, Xiaogang Xu\textsuperscript{a,b}, Lili Tong\textsuperscript{c}, Chongli Xia\textsuperscript{c} and Yao Cheng\textsuperscript{a,b}

\textsuperscript{a}Co-Innovation Center for Sustainable Forestry in Southern China, College of Biology and the Environment, Key Laboratory of State Forestry and Grassland Administration on Subtropical Forest Biodiversity Conservation, Nanjing Forestry University, Nanjing, China; \textsuperscript{b}State Environmental Protection Scientific Observation and Research Station for Ecology and Environment of Wuyi Mountains, Nanjing, China; \textsuperscript{c}School of Horticulture & Landscape Architecture, Jinling Institute of Technology, Nanjing, China

**ABSTRACT**

*Styrax serrulatus* Roxburgh (William Roxburgh 1832), which plays an important role in ecology and economy, is a deciduous species of Styracaceae. In this paper, we sequenced, assembled, and annotated the chloroplast (cp) genome of *S. serrulatus* by using the sequencing data from Illumina Novaseq platform (Illumina, San Diego, CA). The complete cp genome of *S. serrulatus* is 157,929 base pairs (bp) in length, containing a pair of inverted repeat regions (IRs) of 26,048 bp each, a large single-copy (LSC) region of 87,552 bp, and a small single-copy (SSC) region of 18,281 bp. It contains 133 genes, including 8 rRNA genes, 37 tRNA genes, 87 protein-coding genes, and 1 pseudo gene. The GC content of *S. serrulatus* cp genome is 36.96%. The phylogenetic analysis suggests that *S. serrulatus* is a sister species to *Styrax agrestis* in Styracaceae.

**ACKNOWLEDGMENTS**

This work was supported by the National Key Research and Development Program of China (2017YFD0600900). Xiaogang Xu and Lili Tong were involved in sample collection and data analysis. Chongli Xia and Yao Cheng were involved in the seedling collection and sample preparation. Xiaogang Xu was responsible for the manuscript writing. All authors have read and approved the final manuscript.

**CONTACT**

Xiaogang Xu \quad 1208657126@qq.com

Co-Innovation Center for Sustainable Forestry in Southern China, College of Biology and the Environment, Key Laboratory of State Forestry and Grassland Administration on Subtropical Forest Biodiversity Conservation, Nanjing Forestry University, Nanjing, China

\textcopyright{} 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 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.
S. serrulatus genome was 36.96%, and the corresponding values in LSC, SSC, and IR regions were 34.80%, 30.28%, and 42.92%, respectively.

To reveal the phylogenetic evolution of S. serrulatus, we constructed a ML phylogenetic tree based on 40 cp genomes from Styracaceae and 5 cp genomes as outgroups from 3 taxa (Ebenaceae, Symplocaceae, Theaceae). We found that S. serrulatus was clustered with other families of Styracaceae with 100% bootstrap values (Figure 1). In addition, S. serrulatus was highly supported to be a sister species to Styrax agrestis in Styracaceae.

**Disclosure statement**

The authors report no conflict of interest.

**Funding**

This work was supported by The Biodiversity Investigation, Observation and Assessment Program of Ministry of Ecology and Environment of China #1 under Grant [number 20191226]; The Special Fiscal Funds for Repair and Purchase in National Public Institutions #2 under Grant [number 2010002002]; Jiangsu Forest Science and Technology Innovation and Extension Program #3 under Grant [number LYKJ[2018]13]; and The Priority Academic Program Development of Jiangsu Higher Education Institutions #4 under Grant [number PAPD].

**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. MZ152917. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA739084, SRR14861495, and SAMN19771193, respectively.

**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.

Gao F-L, Shen J-G, Liao F-R, Cai W, Lin S-Q, Yang H-K, Chen S-L. 2018.
The first complete genome sequence of Narcissus latent virus from Narcissus. Arch Virol. 163(5):1383–1386.

Huang S-M, Grimes JW. 2003. Styracaceae. In: Wu Z-Y, Raven PH, Hong D-Y, editors. Flora of China, vols. 15 (Styracaceae). Beijing: Science Press; St. Louis: Missouri Botanic Garden Press; p. 260.

Liu C, Shi L-C, Zhu Y-J, Chen H-M, Zhang J-H, Lin X-H, Guan X-J. 2012.
CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC Genomics. 13(715):715.

Rozewicki J, Li S, Amada KM, Standley DM, Katoh K. 2019. MAFFT-DASH: integrated protein sequence and structural alignment. Nucleic Acids Res. 47(W1):W5–W10.

William R. 1832. Descriptions of Indian Plants (Flora Indica). 2:415–416.