

The complete chloroplast genome sequence of *Styrax obassia* (Styracaceae)

Xiaogang Xu^{a,b}, Lili Tong^c, Yaoqin Zhang^{a,b}, Chongli Xia^c and Yabo Wang^{a,b}

^aCollege of Biology and Environment, Nanjing Forestry University, Nanjing, China; ^bCo-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, China; ^cSchool of Horticulture and Landscape Architecture, Jinling Institute of Technology, Nanjing, China

ABSTRACT

Styrax obassia Siebold & Zucc. is a distinct member of the family Styracaceae with fragrant and beautiful flowers. In this study, we determined the complete chloroplast (cp) genome sequence of *S. obassia* in an effort to provide genomic resources useful for promoting its conservation. The circular genome of *S. obassia* was 157,910 bp in size and contained two inverted repeat (IRa and IRb) regions of 26,051 bp, which were separated by a large single copy (LSC) region of 87,528 bp, and a small single copy (SSC) region of 18,280 bp. A total of 133 genes are encoded, including 88 protein-coding genes, 37 tRNA genes, and eight rRNA genes. The overall GC content of *S. obassia* genome is 36.97%. The phylogenetic analysis suggests that *S. obassia* is a sister species to *Styrax suberifolius* and *Styrax zhejiangensis* in Styracaceae.

ARTICLE HISTORY

Received 8 December 2019
Accepted 23 December 2019

KEYWORDS

Styrax obassia; Styracaceae;
complete chloroplast
genome; phylogenomics

Styrax obassia Siebold & Zucc., is a broad-leaved tree species with fragrant and beautiful flowers, which is mainly distributed northern China which has high medicinal value and can be as applied for ornamental purposes. Despite this progress in understanding Styracaceae systematics, many of the relationships among species or genera in the family remain poorly resolved. Comparison of complete plastome sequences further provides the opportunity to explore sequence variation and molecular evolutionary patterns associated with gene loss, rearrangements, duplication, and transfer events. Here, we characterized the complete cp genome sequence of *S. obassia* (GeneBank accession number: MN560143) based on Illumina pair-end sequencing to provide a valuable complete cp genomic resource.

The total genomic DNA was extracted from the fresh leaves of *S. obassia* grown in Zhongshan Botanical Garden (N 32.1566, E 118.9690) in Nanjing, Jiangsu, China. The voucher specimen was kept in the herbarium of Nanjing Forestry University (accession number: NF2019061). The whole genome sequencing was conducted by Nanjing Genepioneer Biotechnologies Inc. (Nanjing, China) on the Illumina HiSeq platform. The raw reads were filtered by CLC Genomics Workbench v9, and the obtained clean reads were assembled into chloroplast genome using SPAdes assembler v3.10.1 (Bankevich et al. 2012). Finally, gene structure annotation was carried out with CpGAVAS (Liu et al. 2012), and sequences were aligned by MAFFT v7.307 (Katoh and Standley 2014).

A Neighbor-Joining (NJ) tree with 100 bootstrap replicates was inferred using MEGA version 7 (Kumar et al. 2016).

The circular genome of *S. obassia* was 157,910 bp in size and contained two inverted repeat (IRa and IRb) regions of 26,051 bp, which were separated by a large single copy (LSC) region of 87,528 bp, and a small single copy (SSC) region of 18,280 bp. A total of 133 genes are encoded, including 88 protein-coding genes (81 PCG species), 37 tRNA genes (30 tRNA species), and eight rRNA genes (four rRNA species). Most of genes occurred in a single copy, however, seven protein-coding genes (*ndhB*, *rpl2*, *rpl23*, *rps12*, *rps7*, *ycf15*, and *ycf2*), seven tRNA genes (*trnA-UGC*, *trnI-CAU*, *trnI-GAU*, *trnL-CAA*, *trnN-GUU*, *trnR-ACG*, and *trnV-GAC*), and four rRNA genes (*16S*, *23S*, *4.5S*, and *5S*) are totally duplicated. A total of nine protein-coding genes (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl16*, *rpl2*, *rpoC1*, and *rps16*) contained one intron while the other three genes (*clpP*, *rps12*, and *ycf3*) had two intron each. The overall GC content of *S. obassia* genome is 36.97%, and the corresponding values in LSC, SSC and IR regions are 34.81%, 30.31% and 42.92%, respectively.

To ascertain the phylogenetic evolution of *S. obassia*, the fasta format file containing all the chloroplast genome sequences of 33 species (28 Styracaceae chloroplast genomes, 2 Symplocaceae chloroplast genomes, 1 Clethraceae chloroplast genomes, and 2 Ericaceae chloroplast genomes). The phylogenetic analysis suggests that *S. obassia* is a sister species to *Styrax suberifolius* and *Styrax*

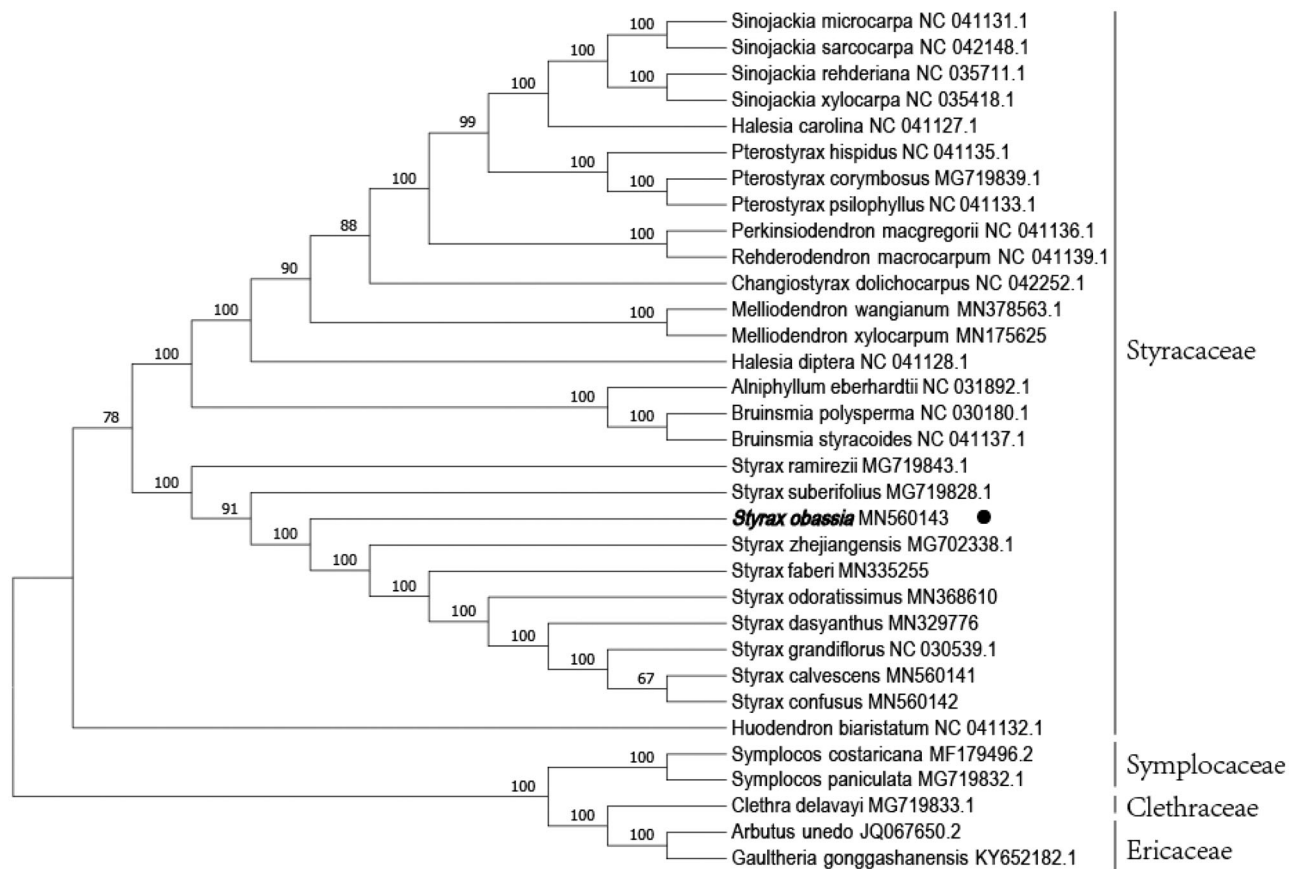


Figure 1. Phylogenetic tree inferred by Neighbor-Joining (NJ) method based on the complete chloroplast genome of 33 representative species. The bootstrap support values are shown at the branches.

zhejiangensis in Styracaceae, with bootstrap support values of 100% (Figure 1).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research work was supported by Jiangsu Forestry Science and Technology Innovation and Extension Program (No. LYKJ[2018]13), The Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Katoh K, Standley DM. 2014. MAFFT: iterative refinement and additional methods. *Methods Mol Biol.* 1079:131–146.
- Kumar S, Stecher G, Tamura K. 2016. Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33(7): 1870–1874.
- Liu C, Shi LC, Zhu YJ, Chen HM, Zhang JH, Lin XH, Guan XJ. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. *BMC Genomics.* 13(1):715.