

MITOGENOME ANNOUNCEMENT



The complete chloroplast genome and phylogenetic analysis of *Salix triandra* from China

Dongyang Wu^a , Yupeng Wang^a, Li Zhang^a, Lijun Dou^b and Linming Gao^c

^aCollege of Information Science and Technology, Nanjing Forestry University, Nanjing, Jiangsu, China; ^bInformation Center of Nanjing Forestry University, Nanjing, Jiangsu, China; ^cComputer Center of Information Science and Technology, Nanjing Forestry University, Nanjing, Jiangsu, China

ABSTRACT

Salix triandra is a great willow for bees and an excellent choice for living willow structures. In this study, we assembled and annotated the complete chloroplast (cp) genome sequence of *S. triandra*. The whole cp genome is 155,821 base-pairs (bp) in size, which comprises one small single copy (SSC) region of 16,223 bp and one large single copy (LSC) region of 84,532 bp separated by a pair of inverted repeats (IRs) of 27,533 bp. There are 131 genes, including 86 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Phylogenetic analysis with the Neighbour-joining method indicates that *S. triandra* is closely related to *S. tetrasperma*. The complete cp genome will facilitate the biological studies in the order Malpighiales in future.

ARTICLE HISTORY

Received 21 September 2019
Accepted 25 September 2019

KEYWORDS

Salix triandra; chloroplast genome; phylogeny

Salix triandra (common names: almond willow, almond-leaved willow) is a species of willow native to Western and Central Asia and Europe. *S. triandra* has immense economic value because it provides bees with abundant pollen and nectar early in the year (Farkas and Zajácz 2007). *Salix triandra* belongs to flowering plants of Salicaceae family, which contains 56 genera and about 1220 species summarized by the Angiosperm Phylogeny Group (APG) (Christenhusz and Byng 2016). In this paper, the complete cp genome sequence of *S. triandra* is characterized for further phylogenetic studies of the family Salicaceae.

The total genomic DNA was extracted from fresh leaves of *Salix triandra* collected in Maoer Mountain (China; N45°19'40.60", E127°37'9.03") using TakaRa MiniBEST Plant Genomic DNA Extraction Kit (Tokyo, Japan). Voucher specimen was deposited in the Key Laboratory of Forest Genetics and Biotechnology, Ministry of Education, Nanjing Forestry University (DB447). Then, the whole-genome sequences were conducted on PacBio RS II Sequencing Instrument (Pacific Bioscience, USA) by Nextomics Biosciences (Wuhan, China). After controlling quality and correcting of 17.6 G raw data, the remaining high-quality reads were assembled with Falcon version 3.0 (Chin et al. 2016). There are many similar cp genomes of homologous species available, which usually as reference genomes to acquire the order of contigs (Wang et al. 2018). So we used BLASTN (Camacho et al. 2009) to isolate cp contigs of *S. triandra* based on *S. suchowensis*

(NC_026462.1). The annotation of *S. triandra* cp genome was conducted with DOGMA (Wyman et al. 2004) with manual check and the physical map was drawn by OGDRAW (Lohse et al. 2007).

The complete cp genome of *S. triandra* (GenBank accession number: MK722343) is 155,821 bp, and shares the common feature of comprising an LSC region of 84,532 bp, an SSC region of 16,223 bp, and two IRs region of 27,533 bp. The total length of protein-coding genes is 91,877 bp, accounting for 58.96% of the complete cp genome and encoding 26,646 amino acids. The overall GC content is 36.65%, higher than those of LSC (34.39%) and SSC (30.94%) regions, but lower than those of IRs regions 41.79%). The cp genome possesses 131 functional genes, including 86 protein-coding genes, 8 rRNA genes, and 37 tRNA genes. Among these genes, three genes (*rps12*, *clpP*, *ycf3*) have two introns, and other 17 genes have one intron each. There are 18 genes (7 protein-coding genes, 7 tRNA and 4 rRNA genes) duplicated in the IR regions.

To identify the phylogenetic position of *S. triandra*, 24 cp genomes (excluding *S. triandra*) of the family Salicaceae were downloaded from National Centre for Biotechnology Information (NCBI). These 25 cp genomes sequences were aligned with ClustalW (Thompson et al. 2002). The phylogenetic tree (Figure 1) based on 70 protein-coding genes using the neighbour-joining method was reconstructed by MEGA7 (Kumar et al. 2016). The phylogenetic analysis showed that *S.*

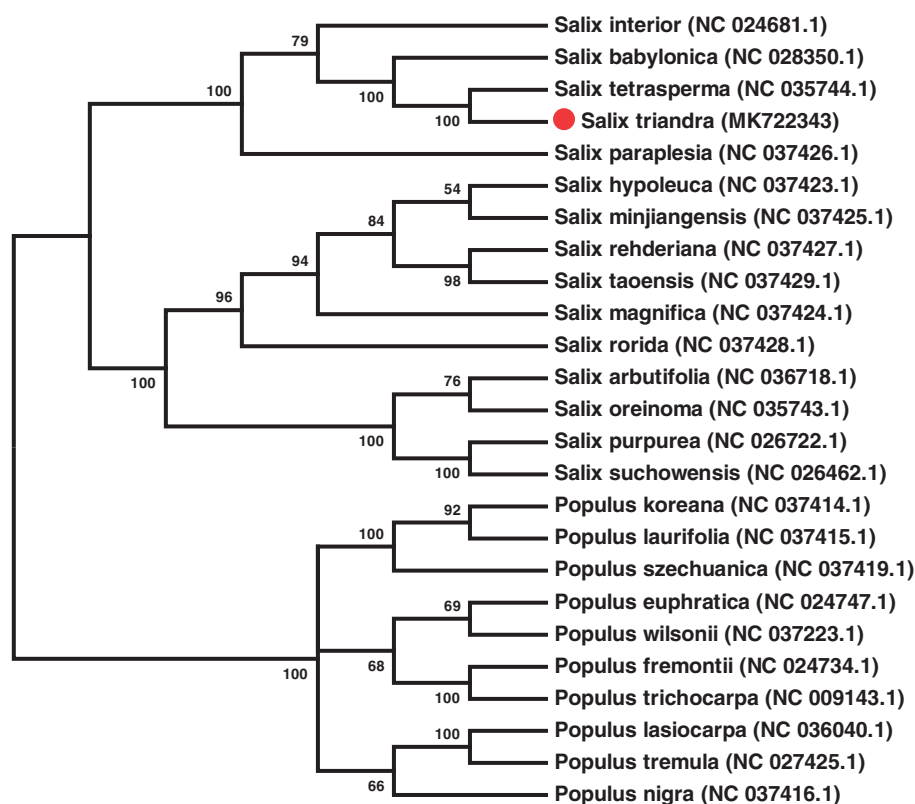


Figure 1. Phylogenetic tree with neighbour-joining method was constructed by MEGA7. The 70 protein-coding genes of 25 cp genomes from the Salicaceae family were aligned by ClustalW. The bootstrap values (indicating on the branches) were based on 1000 replicates and the accession number is list after each species name.

triandra is evolutionarily closest to *S. tetrasperma* with 100% bootstrap value.

Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Funding

This study was supported by the National Key Research and Development Plan of China [2016YFD0600101], the National Natural Science Foundation of China [31570662, 31500533 and 61401214], Jiangsu Provincial Department of Housing and Urban-Rural Development [2016ZD44], and the PAPD (Priority Academic Programme Development) programme at Nanjing Forestry University, Natural Science Foundation of Jiangsu Province [BK20170934], National Natural Science Foundation of China [61802193].

ORCID

Dongyang Wu  <http://orcid.org/0000-0002-2570-6494>

References

- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLASTp: architecture and applications. *BMC Bioinformatics*. 10(1):421.
- Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, Dunn C, O'Malley R, Figueroa-Balderas R, Morales-Cruz A, et al. 2016. Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods*. 13(12):1050.
- Christenhusz MJ, Byng JW. 2016. The number of known plants species in the world and its annual increase. *Phytotaxa*. 261(3):201–217.
- Farkas Á, Zajtác E. 2007. Nectar production for the Hungarian honey industry. *Eur J Plant Sci Biotechnol*. 1(2):125–151.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 33(7):1870.
- Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet*. 52(5–6):267–274.
- Thompson JD, Gibson TJ, Higgins DG. 2002. Multiple sequence alignment using clustalW and clustalX. *Current protocols in bioinformatics/editorial board, Andreas D. Baxevanis. [et al.], Chapter 2(Unit 2), Unit 2.3.*
- Wang X, Cheng F, Rohlsen D, Bi C, Wang C, Xu Y, Wei S, Ye Q, Yin T, Ye N, et al. 2018. Organellar genome assembly methods and comparative analysis of horticultural plants. *Hortic Res*. 5(1):3.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics*. 20(17):3252–3255.