

The complete chloroplast genome sequence of *Salix kochiana* Trautv. and its phylogenetic analysis

Jing Wang^{a,b} , Zicheng Yu^{a,b}, Xu Yao^{a,b}, Jie Wan^{a,b}, Zhengxuan Wang^{a,b} and Xiaoping Li^{a,b,c}

^aCollaborative Innovation Center of Southern Modern Forestry, Nanjing Forestry University, Nanjing, China; ^bCollege of Forestry, Nanjing Forestry University, Nanjing, China; ^cJiangsu Key Laboratory for Poplar Germplasm Innovation and Variety Improvement, Nanjing Forestry University, Nanjing, China

ABSTRACT

Salix kochiana Trautvetter 1837 is one of the highest value shrubs present in northern China with important economic and ecological benefits. This study revealed the structural characteristics and phylogenetic relationships of chloroplast genes in *S. kochiana* Trautv. The results showed that the length of the complete chloroplast genome was 155,657 bp, which was a typical circular double-stranded structure, including an 84,458 bp large single-copy region (LSC), a 16,221 bp small single-copy region (SSC) and a 27,489 bp pair of inverted repeat regions (IRA and IRB). The chloroplast genome contains 48,757 A bases, 28,017 G bases, 49,843 T bases, and 29,040 C bases, with a GC content of 36.66%. Through bioinformatics annotation, a total of 126 genes were found in the chloroplast genome, including 81 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Phylogenetic analysis showed that *S. kochiana* Trautv. was closely related to *S. triandroides*.

ARTICLE HISTORY

Received 12 January 2022
Accepted 3 June 2022

KEYWORDS


Chloroplast genome;
phylogenetic analysis; *Salix kochiana* Trautv.; Salicaceae

The *Salix* genus is composed of ~ 350–520 species that are distributed worldwide (Zhou et al. 2021). In China, the *Salix* genus contains 257 species, including 122 varieties and 33 variants, and they are distributed across provinces (Wang et al. 2012). *Salix kochiana* Trautv., an important species of the *Salix* genus, is mainly distributed in the Heilongjiang and Jilin provinces of China and is also found in Mongolia and Russia. *S. kochiana* Trautv. can grow in sand dunes and river wetlands, so it can serve as an important tree species for windbreaks, sand fixation, and soil conservation in northern China. With its fast growth, luxuriant foliage, and slender and soft branches, *S. kochiana* Trautv. can be made into a variety of woven goods. Its twigs and leaves can also be used as animal feed. Therefore, *S. kochiana* Trautv. has high economic value and provides high ecological benefits. However, the available genetic and genomic resources for *S. kochiana* Trautv. are very limited. In addition, due to the highly efficient crossing rate among *Salix* species, the classification of the *Salix* genus is still unclear (Chen et al. 2010). Here, we revealed the structural characteristics and phylogenetic relationships of chloroplast genes in *S. kochiana* Trautv. These results are of great significance for the taxonomic study of *Salix* and even Salicaceae and can also be used to evaluate the breeding availability of *Salix* at the genomic level (Chen and Liu 2008; Ren and Li 2021).

The test materials were collected from Mao'er Mountain of Heilongjiang Province, China (44°29'29"N, 127°17'41"E). The specimens were stored in Room 60708 Biotechnology

Building (University sample room), Nanjing Forestry University, China (Xiaoping Li, xpli@njfu.edu.cn, voucher number: MRSSQL2017_1451). The modified CTAB method (cetyl trimethylammonium bromide) (Doyle and Doyle 1987) was used to extract whole-genome DNA from the leaves of *S. kochiana* Trautv., and DNA concentration and purity were detected by 1% agarose gel electrophoresis and a NanoDrop spectrophotometer (Niu et al. 2020). Qualified DNA fragment libraries were constructed with an average insert size of 300 bp using the Illumina Nextera kit and sequenced by the Illumina NovaSeq 6000 platform (Illumina, San Diego, CA). A total of 3.5 G raw data were obtained and filtered to eliminate low-quality reads by Fastp (Chen et al. 2018). Finally, approximately 3.2 G of high-quality clean reads were harvested and used to assemble the chloroplast genome with NOVOPlasty (version 4.1) (<https://github.com/ndierckx/NOVOPlasty/>) (Dierckxsens et al. 2017). The chloroplast genes of *S. kochiana* Trautv. were aligned and annotated with the chloroplast genome of the reference species *S. triandroides* (MW929215.1) by PGA (Plastid Genome Annotator) (Qu et al. 2019). After manual correction by Geneious v11.0.3 (Kearse et al. 2012), the resulting annotation of the complete chloroplast DNA was obtained and submitted to NCBI online by BankIt (GenBank accession number: OL339478.1).

The annotated results showed that the complete chloroplast genome length was 155,657 bp, which was a typical circular double-stranded structure, including an 84,458 bp large single-copy region (LSC), a 16,221 bp small single-copy region

CONTACT Xiaoping Li  xpli@njfu.edu.cn  College of Forestry, Nanjing Forestry University, Nanjing, 210037, China

© 2022 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.

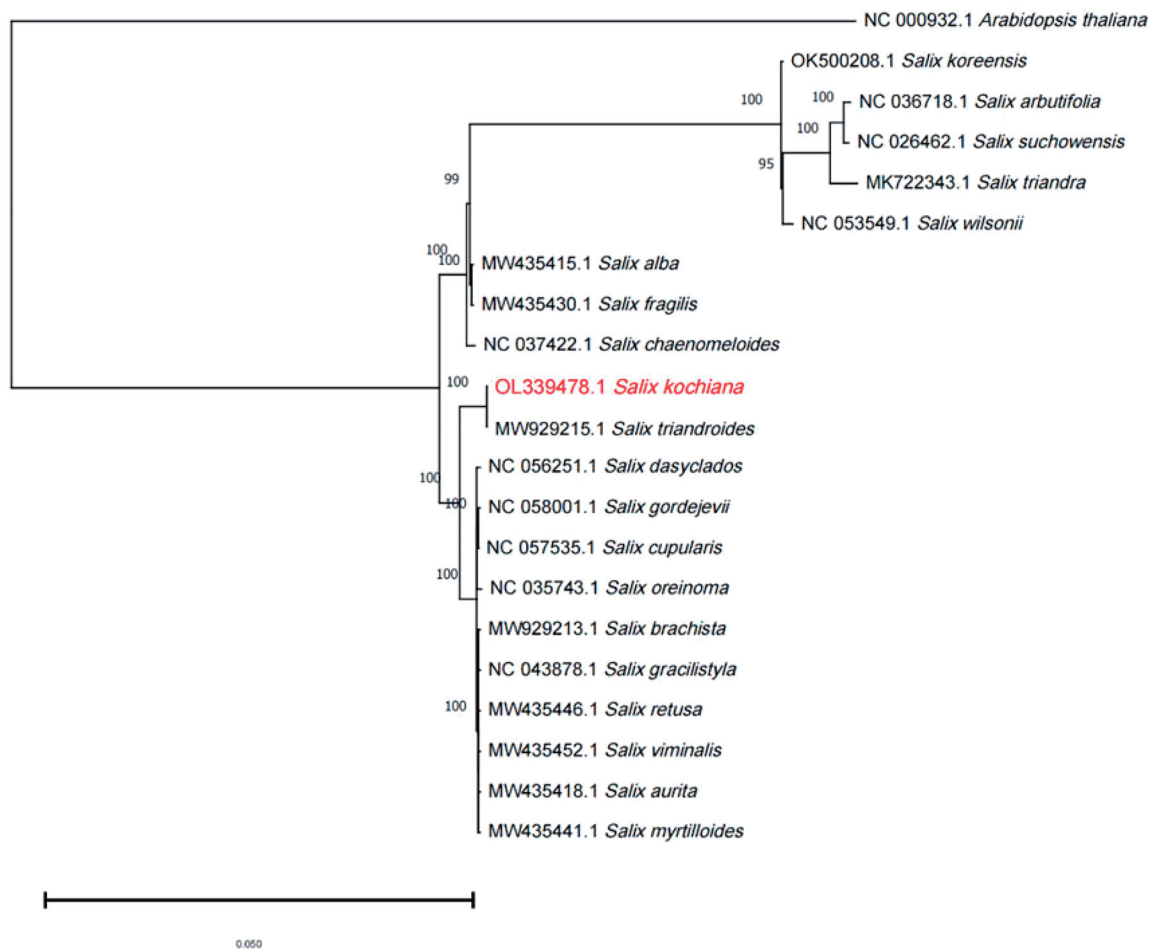


Figure 1. Maximum likelihood phylogenetic tree based on 21 chloroplast genome sequences of selected species. The numbers near branches indicate bootstrap support values.

(SSC), and a 27,489 bp pair of inverted repeat regions (IRA and IRB). The chloroplast genome contains 48,757 A bases, 28,017 G bases, 49,843 T bases, and 29,040 C bases, with a GC content of 36.66%. A total of 126 genes were identified, including 81 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Among them, 7 protein-coding genes, 7 tRNA genes and 4 rRNA genes were repeated in the inverted repeat region.

To reveal the relationship between *S. kochiana* Trautv. and other species in the *Salix* genus, we downloaded the chloroplast genome sequences of 19 published *Salix* species from the NCBI database and performed phylogenetic analysis. Comparison of the *S. kochiana* Trautv. plastome to previously published data shows a high level of gene synteny with some publicly available *Salix* sequences (Zhou et al. 2021). *Arabidopsis thaliana* (NC_000932.1) was used as an outgroup. MAFFT v7 software (<https://mafft.cbrc.jp/alignment/software/>) (Kato and Standley 2013) was used to perform multiple weight ratio pairs for the common sequences of chloroplast genomes of these 21 species. A phylogenetic tree was constructed based on the maximum likelihood (ML) method with the general time reversible model and 1000 bootstrap replicates using MEGA11 (Koichiro et al. 2021). The results showed that *S. kochiana* Trautv. was most closely related to *S. triandroides* (Figure 1). This finding was similar to those of a previous *Salix* whole-plastome phylogenetic study (Wang

and Li 2021; Chen et al. 2019). These results explain the systematic origin and molecular phylogenetic location of *S. kochiana* Trautv., which can be used for molecular identification and resource exploitation. Overall, this study lays a foundation for *Salix* phylogenetics in the future.

Disclosure statement

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

Author contributions

Xiaoping Li contributed to the conception of the study; Jing Wang, Zicheng Yu, Xu Yao, Jie Wan and Zhengxuan Wang performed the experiment; Jing Wang contributed significantly to analysis and manuscript preparation; Jing Wang performed the data analyses and wrote the manuscript; Xiaoping Li helped perform the analysis with constructive discussions and revised it critically for intellectual content. All authors approve the version to be published and agree to be accountable for all aspects of the work.

Funding

This work was supported by the National Key Research and Development Plan of China (2021YFD2200202), Natural Science Foundation of Jiangsu Province [BK20161525] and National Natural Science Foundation of China [30970245].

ORCID

Jing Wang  <http://orcid.org/0000-0002-8897-4226>

Data availability statement

The complete chloroplast genome sequence data of *S. kochiana* Trautv. supporting the findings of this study are openly available in Genebank of NCBI at <http://www.ncbi.nlm.nih.gov/> under the accession number OL339478.1. The associated BioProject, SRA, and BioSample numbers are PRJNA778909, SRR16924947, and SAMN23003006 respectively.

Regulation statement

This study strictly follows the wild plant protection regulations of China, which were firstly promulgated by Decree No. 204 of the state council of China on Sept. 30, 1996 and amended on Oct.7, 2017 under Decision of The State Council on Revising Some Administrative Regulations.

References

- Chen J-H, Sun H, Wen J, Yang Y-P. 2010. Molecular phylogeny of *Salix L.* (Salicaceae) inferred from three chloroplast datasets and its systematic implications. *Taxon*. 59(1):29–37.
- Chen Y, Hu N, Wu H. 2019. Analyzing and characterizing the chloroplast genome of *Salix wilsonii*. *Biomed Res Int*. 2019:5190425.
- Chen X, Liu C. 2008. Progress in chloroplast genome analysis. *Prog Biochem Biophys*. 35:21–28.
- Chen S, Zhou Y, Chen Y, Jia G. 2018. FASTP: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 34(17):i884–i890.
- Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOplasty: *de novo* assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 45(4):e18.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull*. 19:11–15.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12):1647–1649.
- Koichiro T, Glen S, Sudhir K. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol*. 38(7):3022–3027.
- Niu YF, Li GH, Liu ZY, et al. 2020. Assembly and annotation of chloroplast genome sequencing of *Hevea glabra*. *Study Econ Forest*. 38(04):62–71.
- Qu X-J, Moore MJ, Li D-Z, Yi T-S. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods*. 15:50.
- Ren RR, Li XP. 2021. Characterization of the complete chloroplast genome of *Salix linearistipularis* (Franch.) Hao 1936. *Mitochondrial DNA B Resour*. 6(9):2764–2766.
- Wang W, Li X. 2021. Characterization of the complete chloroplast genome of *Salix chienii* W.C. Cheng (Salicaceae) from Southern China. *Mitochondrial DNA B Resour*. 6(10):3087–3088.
- Wang DC, Yang YP, Chen JH, et al. 2012. Leaf epidermal micromorphological characteristics of 28 *Salix* species and their taxonomic significance. *J Plant Taxon Res*. 34(005):430–442.
- Zhou J, Jiao Z, Guo J, Wang BS, Zheng J. 2021. Complete chloroplast genome sequencing of five *Salix* species and its application in the phylogeny and taxonomy of the genus. *Mitochondrial DNA B Resour*. 6(8):2348–2352.