

The complete chloroplast genome sequence of *Keteleeria fortunei* (Pinaceae)

Yuanyuan Li^{a,b*}, Cheng Zhang^{a,b*}, Min Zhang^{a,b*}, Yongfu Li^{a,b}, Xianrong Wang^{a,b} and Yifan Duan^{a,b}

^aInternational Cultivar Registration Center for Osmanthus, College of Biology and the Environment, Nanjing Forestry University, Nanjing, Jiangsu, China; ^bCo-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing, Jiangsu, China

ABSTRACT

Here, the complete chloroplast genome of *Keteleeria fortunei*, a vulnerable species in China, was sequenced by next-generation sequencing platform. Its circular genome was 117,183 bp in length and the GC content was 38.5%. A total of 101 genes were annotated, including 4 rRNA genes, 20 tRNA genes, and 71 protein coding genes. This study would further our understanding of the genomics and the conservation and utilization of *K. fortunei*.

ARTICLE HISTORY

Received 6 August 2019
Accepted 16 August 2019

KEYWORDS

Chloroplast genome;
SPAdes; Illumina
sequencing; phylogenetic
analysis

Keteleeria fortunei (Murr.) Carr., a vulnerable species belonging to Pinaceae in China (Qin et al. 2017), is widely cultivated for afforestation and as an ornamental species, and its timber is used for construction and furniture (Fu et al. 1999). However, the genomic background of this species is still lacking and no chloroplast genome of *K. fortunei* has been reported. Here, we assembled and characterized the complete chloroplast genome of *K. fortunei* using the Illumina paired-end sequencing data.

The total genomic DNA was extracted from the fresh leaves of *K. fortunei* which were collected at Kunming Institute of Botany (102°45'4.68"E, 25°8'30.12"N). The voucher specimen was kept in the herbarium of Nanjing Forestry University (accession number: NF0000005). Illumina paired-end (PE) DNA library was prepared and sequenced in Nanjing Genepioneer Biotechnologies Inc., Nanjing, China. The raw reads were filtered by CLC Genomics Workbench v9 (CLC Bio, Aarhus, Denmark), and the obtained clean reads were assembled into the chloroplast genome using SPAdes (Bankevich et al. 2012). Finally, gene structure annotation was carried out with GENEIOUS R11 (Kearse et al. 2012) and the physical map was generated with OGDRAW (Lohse et al. 2013). A phylogenetic tree was inferred based on the maximum-likelihood (ML) by using HomBlocks (Bi et al. 2018) and GTR + I + G were selected as the best substitution model for the ML analyses using PartitionFinder2 (Lanfear et al. 2012).

Like some other conifers (Xie et al. 2019; Zhang et al. 2019), the chloroplast genome of *K. fortunei* did not show a stable tetrad structure for having no typical inverted repeat (IR) regions. The circular genome of *K. fortunei* was 117,183 bp in size. It contained 101 genes including 71

protein-coding genes, 20 tRNA genes, and 4 rRNA genes. Except that there were four copies of *trnS*, three copies in *trnI*, *trnL*, and *trnR* and two copies in *trnT*, *trnP*, *trnV*, *psbI*, *trnG*, and *trnFM*, all the other genes were single-copy. MicroSATellite (MISA) software (<http://pgrc.ipk-gatersleben.de/misa/>) was used to search for simple sequence repeat (SSR), and it found that cpSSR distribution range on chloroplast genome was 530–117,182 bp, in which the number of repeats of single nucleotide, dinucleotide, and trinucleotide was 21, 7, and 1, respectively. The codons corresponding to protein-coding genes in the chloroplast genome of *K. fortunei* preferred to use A/T base, among which leucine had the largest number of codons (1995), isoleucine and serine were 1684 and 1445, ranking 2nd and 3rd respectively, while cysteine had the smallest number of 227.



The phylogenetic analysis of 24 cp genomes suggested a close relationship between *Keteleeria* (*K. fortunei*, *Keteleeria davidiana*) and *Abies* (*Abies alba*, *Abies koreana*) (Figure 1), agreeing with previous studies based on transcriptomic data (Ran et al. 2018). This study would provide a theoretical basis for further understanding the genomic information and the development, conservation, and utilization of *K. fortunei*.

Acknowledgements

We thank Dr. Zhang Liang (Kunming Institute of Botany, Chinese Academy of Sciences) for collecting samples.

Disclosure statement

No potential conflict of interest was reported by the authors.

CONTACT Yifan Duan  yifan419@hotmail.com  International Cultivar Registration Center for Osmanthus, College of Biology and the Environment, Nanjing Forestry University, 159# Longpan Road, Xuanwu District, Nanjing 210037, Jiangsu, China

*Co-first authors.

© 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.

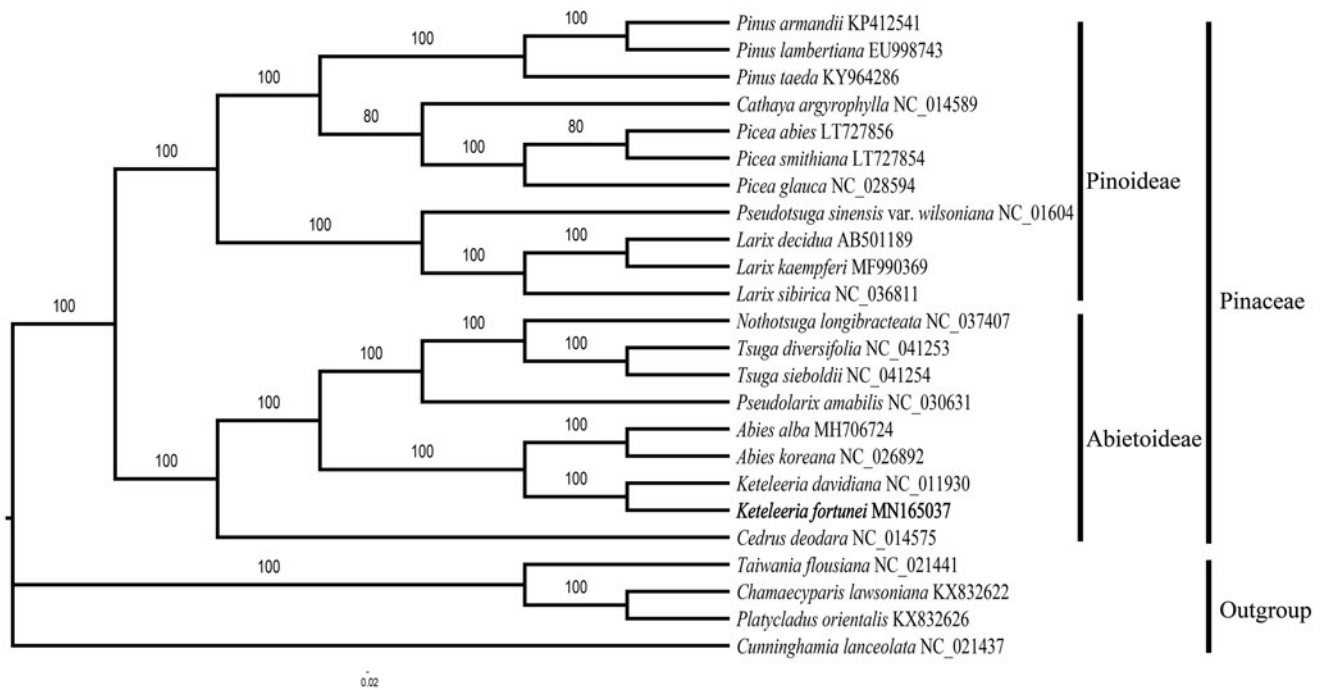


Figure 1. The phylogenetic tree based on 24 complete chloroplast genome sequences. Bootstrapping values were listed for each node.

Funding

This study was partially funded by the Priority Academic Program Development (PAPD) of Jiangsu High Education Institutions, Jiangsu Province, China.

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.
- Bi GQ, Mao YX, Xing QK, Cao M. 2018. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics.* 110:18–22.
- Fu LG, Li N, Mill RR. 1999. Pinaceae. In: Wu ZY, Raven PH, editors. *Flora of China*. 4th ed. Beijing (China): Science Press.
- 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: 1647–1649.
- Lanfear R, Calcott B, Ho SY, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol.* 29:1695–1701.
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013. Organellar Genome DRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41:W575–W581.
- Qin HN, Yang Y, Dong SY, He Q, Jia Y, Zhao LN, Yu SX, Liu HY, Liu B, Yan YH, et al. 2017. Threatened species list of China's higher plants. *Biodiv Sci.* 25:696–744.
- Ran JH, Shen TT, Wu H, Gong X, Wang XQ. 2018. Phylogeny and evolutionary history of Pinaceae updated by transcriptomic analysis. *Mol Phylogenet Evol.* 129:106–116.
- Xie SY, Li JL, Miao JB, Xu JJ, Mao KS. 2019. The complete chloroplast genome of *Juniperus squamata* (Cupressaceae), a shrubby conifer from Asian Mountains. *Mitochondrial DNA B.* 4:2137–2139.
- Zhang J, Cao ZP, Li S, Liu MM, Huo XW. 2019. The complete chloroplast genome sequence of *Amentotaxus yunnanensis* (Taxaceae). *Mitochondrial DNA B.* 4:2477–2478.