

The complete chloroplast genome of *Araucaria cunninghamii* (Araucariaceae)

Jingyao Ping^a, Xin Luo^b, Ming Zhu^a, Rongjing Zhang^a, Chunmei Qian^a, Yingjuan Su^{c,d} and Ting Wang^{a,e}

^aCollege of Life Sciences, South China Agricultural University, Guangzhou, China; ^bInstitute of Tropical and Subtropical Cash Crops, Yunnan Academy of Agricultural Sciences, Baoshan, China; ^cSchool of Life Sciences, Sun Yat-sen University, Guangzhou, China; ^dResearch Institute of Sun Yat-sen University in Shenzhen, Shenzhen, China; ^eGuangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

ABSTRACT

The chloroplast genome of the *Araucaria cunninghamii* has been completely sequenced. The genome size is 146,337 bp, and the overall GC content is 36.7%. This cp genome does not contain canonical IRs, it encodes a total 122 genes including 82 protein-coding genes, 36 tRNA genes and four rRNA genes. Among them, eight protein-coding genes (*rpl2*, *rpl16*, *petB*, *petD*, *rpoC1*, *atpF*, *ndhA* and *ndhB*) have two exon, and two genes (*rps12* and *ycf3*) have three exons. Also, *trnI*-CAU and *rrn5* has two copies. The Maximum-likelihood tree shows that *A. cunninghamii* is sister to *A. heterophylla*.

ARTICLE HISTORY

Received 5 June 2020
Accepted 20 June 2020

KEYWORDS

Araucaria cunninghamii;
chloroplast genome;
phylogenetic analysis

Araucaria cunninghamii (Araucariaceae) is a perennial evergreen tree. It is native to the southeastern coastal area of Oceania and is the main afforestation tree species in tropical and subtropical regions. Its tall tree shape and graceful posture are known as one of the five major park tree species in the world and are important ornamental plants in the garden, and cultivated in many places in China (Zheng and Fu 1978). *A. cunninghamii* is a common group of *Araucaria*, and its chloroplast genome characteristics have not been

published. Obtaining its complete chloroplast genome can provide molecular resources for the study of phylogenetic genomics and population variation of *A. cunninghamii*.

Fresh leaves were sampled from the campus of South China Agricultural University (SCAU), location (E113°20', N23°9'). The specimens were stored in Herbarium of the College of Life Sciences, SCAU (specimen no.: PJY-NYS1910). Genomic DNA was extracted using Tiangen Plant Genomic DNA Kit (Tiangen Biotech Co., Beijing, China). Illumina HiSeq

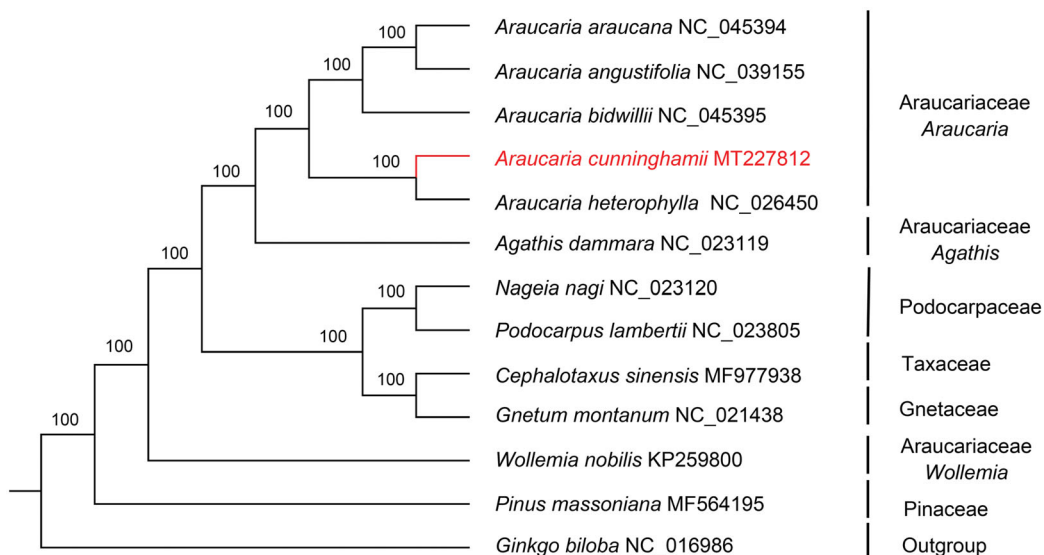




Figure 1. The ML phylogenetic tree constructed based on the complete chloroplast genome sequences of 13 species. Outgroup is *Ginkgo biloba*. Bootstrap support values are marked on nodes.

CONTACT Ting Wang  tingwang@scau.edu.cn  College of Life Sciences, South China Agricultural University, Guangzhou, China; Guangdong Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

© 2020 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.

2500 platform was used for sequencing. A total of 12,206,029 raw reads were generated. Clean reads were obtained after filtering with Trimmomatic v0.32 (Bolger et al. 2014), which were *de novo* assembled into contigs by Velvet (Zerbino and Birney 2008). DOGMA (Wyman et al. 2004) was used for gene prediction. The protein sequence obtained by gene prediction was compared with the NCBI CDD (Marchler-Bauer et al. 2015, Conserved Domain Database) database, and COG function information of genes was predicted by rpsblast v2.2.30+ (Altschul et al. 1997). The ClustalW model in MEGA X (Kumar et al. 2018) was used to create a multiple sequence alignment of the complete cp genome of *A. cunninghamii* with those of other 12 plants downloaded from GenBank. Based on maximum-likelihood (ML), the phylogenetic tree was constructed using RaxMLGUI2 (Stamatakis 2014) with 1000 bootstrap replicates.

The complete cp genome of *A. cunninghamii* is 146,337 bp in size with circular DNA molecular structure (GenBank accession number: MT227812). Like other conifers, it does not contain inverted repeat (IR) region. The cpDNA contains 122 genes, including 82 protein-coding genes, 36 tRNA genes and four rRNA genes. Among these genes, 14 genes (eight protein-coding genes and six tRNA genes) contain one intron, two genes (*rps12*, *ycf3*) have two introns. Like other species of *Araucaria* (such as *A. angustifolia*) (Brandão et al. 2019), it also has two duplicated genes (*trnI*-CAU and *rnn5*). The ML-tree highly supports *A. Cunninghamii* and *A. heterophylla* are sister groups, and the monophyletic clade formed by them are sister groups of remaining *Araucaria* (Figure 1). The cp genome of *A. cunninghamii* provides a reliable molecular resource for phylogenetic studies and chloroplast genomics of gymnosperms.

Disclosure statement

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

Funding

This work was supported by the National Natural Science Foundation of China [31670200, 31872670].

Data availability statement

The data that support the findings of this study are available in [NCBI] at [<https://www.ncbi.nlm.nih.gov/>]. These data were derived from the following resources available in the public domain: [https://www.ncbi.nlm.nih.gov/nucleotide/NC_021438,KP259800,MF977938,MT227812,NC_039155,NC_045394,NC_045395,NC_026450,NC_023805,MF564195,NC_023120,NC_016986]

References

- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389–3402.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics.* 30(15):2114–2120.
- Brandão JHSG, Rodrigues NF, Eguiluz M, Guzman F, Margis R. 2019. *Araucaria angustifolia* chloroplast genome sequence and its relation to other Araucariaceae. *Genet Mol Biol.* 42(3):671–676.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 35(6):1547–1549.
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, et al. 2015. CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43(Database issue): D222–D226.
- Stamatakis A. 2014. RaxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30(9):1312–1313.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20(17):3252–3255.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18(5):821–829.
- Zheng WJ, Fu LG. 1978. Cupressaceae. *Flora of China*. Volume 7. Beijing: Science Press (in Chinese).