

The whole chloroplast genome sequence of *Macadamia tetraphylla* (Proteaceae)

Jin Liu^{a*}, Ying-Feng Niu^{a*}, Shu-Bang Ni^a, Xi-Yong He^a, Cheng Zheng^a, Zi-Yan Liu^a, Hao-Hong Cai^b and Chao Shi^{a,b,c}

^aYunnan Institute of Tropical Crops, Jinghong, Xishuangbanna, PR China; ^bCollege of Marine Science and Biological Engineering, Qingdao University of Science and Technology, Qingdao, PR China; ^cKunming Institute of Botany, Chinese Academy of Sciences, Kunming, PR China

ABSTRACT

Macadamia tetraphylla (Proteaceae) is one of the two macadamia plants that are edible and of cultivated value. Only two chloroplast genomes were reported in Proteaceae so far. In this study, we report the complete chloroplast genome sequence of *M. tetraphylla*, which is the third reported chloroplast genome in Proteaceae. The chloroplast genome is 159,195 bp long and includes 113 genes. Its LSC, SSC and IR regions are 87,951, 18,748 and 26,248 bp long, respectively. Phylogenetic analysis indicates that *M. tetraphylla* was clustered with other two species of Proteaceae, the *M. integrifolia* and *M. ternifolia*.

ARTICLE HISTORY

Received 23 August 2018
Accepted 31 August 2018

KEYWORDS

Macadamia tetraphylla;
proteaceae; chloroplast genome

Macadamia (*Macadamia tetraphylla*; Proteaceae) is one of the two macadamia plants that are edible and of cultivated value. Its kernel rate and oil content are low. Because of the high sugar content (Fourie and Basson 1990), processed products are prone to browning (Wall and Gentry 2005). Whereas, macadamia tree is a gracefully plant with dense branches and leaves, beautiful and aromatic flowers, solid and fine wood, and it conquers most insects and disease. Thus, it is an excellent landscaping and timber species. However, *M. tetraphylla* is becoming extinct (Shapcott and Powell 2011). Genetic analysis can determine the potential long-term viability of this species. Only two chloroplast genomes were reported in Proteaceae, the *M. integrifolia* (Nock et al. 2014) and *M. ternifolia* (Liu et al. 2017). In the

study, we report the complete chloroplast genome of *M. tetraphylla*, and hope it will help restore the macadamia species to prevent the extinction of this single wild population.

The DNA material was isolated from mature leaves of a *M. tetraphylla* plant cultivated in the plant garden of Yunnan Institute of Tropical Crops (YITC), Jinghong, China and the specimen of this tree was conserved in YITC. The isolated DNA was sent to BGI Shenzhen for library construction and genome sequencing on the Illumina HiSeq 2000 Platform (Illumina, San Diego, CA) with subsequent of 3.5 Gbp reads in fastq format were obtained and used in chloroplast genome assembly. The complete chloroplast genome was annotated with Dual Organelle GenoMe Annotator (DOGMA; Wyman et al. 2004) and submitted to the Genbank

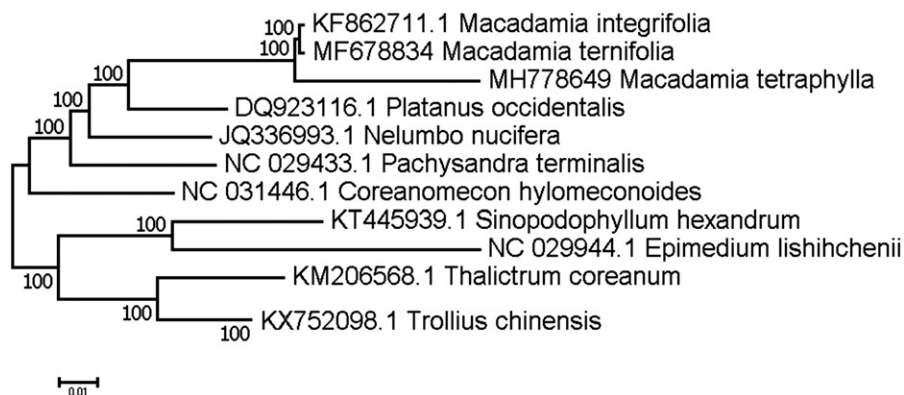


Figure 1. Maximum-likelihood (ML) phylogenetic tree of *M. tetraphylla* and the other 10 representative early diverging eudicotyledons species. Number above each node indicates the ML bootstrap support values.

(<http://www.ncbi.nlm.nih.gov/>) under the accession number of MH778649. A physical map of the chloroplast genome was generated by OGDRAW (<http://ogdraw.mpimp-golm.mpg.de/>) (Lohse et al. 2013). Our assembly of the *M. tetraphylla* resulted in a final sequence of 159,195 bp in length with no gap. The overall A-T content of the chloroplast genome was 61.2%. This chloroplast genome included a typical quadripartite structure with the Large Single Copy (LSC), Small Single Copy (SSC) and Inverted Repeat (IR) regions of 87,951, 18,748 and 26,248 long, respectively. Genome annotation showed 113 full length genes including 79 protein-coding genes, 30 tRNA genes and four rRNA genes. The genome organization, gene content and gene relative positions were almost identical to the previously reported *M. integrifolia* chloroplast genome (Nock et al. 2014). Eighteen genes were duplicated in the IR regions. Fifteen genes contained one intron, while three had two introns. Phylogenetic tree included *M. tetraphylla*, *M. integrifolia* and the other 10 representative early diverging eudicotyledons species was constructed. Maximum-likelihood (ML) analysis exhibited that *M. tetraphylla* clustered with *M. integrifolia* (Nock et al. 2014) and *M. ternifolia* (Liu et al. 2017), and other species were highly supported by a bootstrap value of 100 (Figure 1).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The project was funded by National Natural Science Foundation of China [No. 31760215], the Technology Innovation Talents Project of Yunnan Province [2018HB086], and Sci-Tech Innovation System Construction for Tropical Crops Grant of Yunnan Province [No. RF2018].

References

- Fourie PC, Basson DS. 1990. Sugar content of almond, pecan, and macadamia nuts. *J Agric Food Chem.* 38:101–104.
- Liu J, Niu Y-F, Ni S-B, He X-Y, Shi C. 2017. Complete chloroplast genome of a subtropical fruit tree *Macadamia ternifolia* (Proteaceae). *Mitochondrial Dna Part B: Resources.* 2:738–739.
- Lohse M, Drechsel O, Kahlau S, Bock R. 2013. Organellar GenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* 41:575–581.
- Nock CJ, Baten A, King GJ. 2014. Complete chloroplast genome of *Macadamia integrifolia* confirms the position of the Gondwanan early-diverging eudicot family Proteaceae. *BMC Genomics.* 15:13.
- Shapcott A, Powell M. 2011. Demographic structure, genetic diversity and habitat distribution of the endangered, Australian rainforest tree *Macadamia janseni* help facilitate an introduction program. *Austral J Botany.* 59:215–225.
- Wall MM, Gentry TS. 2005. Sugar composition and color development during drying and roasting of macadamia nuts. *Proceedings of the Hawaii Macadamia Nut Association 45th Annual Conference.* p. 11–19.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with dogma. *Bioinformatics.* 20:3252–3255.