The complete chloroplast genome of *Craigia yunnanensis*, an endangered plant species with extremely small populations (PSESP) from South China

Hafiz Muhammad Warissa, Yaling Chen and Jing Yang

*a*Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China; *b*Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China; *c*University of Chinese Academy of Sciences, Beijing, China

**ABSTRACT**

*Craigia yunnanensis* is an Endangered tree species, enlisted in the national plant species with extremely small population (PSESP) of China for urgent protection. The chloroplast genome is characterized using Illumina pair-end sequencing data. The whole chloroplast genome is 163,166 bp in length and consists of a pair of inverted repeat (IR) regions of 25,573 bp each, which are separated by the large single copy (LSC: 91,580 bp) region and small single copy (SSC: 20,440 bp) regions. The chloroplast genome encodes 113 unique genes, including 79 protein-coding genes (PCGs), 30 tRNA, and 4 rRNA genes. The overall GC content of the whole chloroplast genome is 36.4%. The phylogenetic tree revealed a close relationship between *C. yunnanensis* and *Tilia* species of Malvaceae family.

*Craigia yunnanensis* W. W. Smith & W. E. Evans is a deciduous tree that can grow up to 30 m tall, having a straight trunk and fine wood structure, which produces top-quality timber for construction and furniture (Li 2005; Gao et al. 2010). However, because of deforestation and cash crop cultivation in southern China, its natural population size has been greatly reduced and is on the verge of extinction. As the only surviving species in relic genus *Craigia* (once widespread in the Tertiary era) (Yang et al. 2016), it has been proposed as a second-ranked plant for national protection in China and categorized as ‘Endangered’ by IUCN Red List (Sun 1998; Wang and Xie 2004; IUCN 2008). Moreover, *C. yunnanensis* is now recorded in plant species with extremely small population (PSESP) of China in the national-level Implementation Plan of Rescuing and Conserving China’s PSESP (2010e2015). A conservation strategy focusing on PSESP and aimed at rescuing the most endangered plants (Ma et al. 2013). In this study, we characterized the complete chloroplast genome of *C. yunnanensis* to investigate the phylogenetic relationship of *C. yunnanensis* and provide genomic resources for conservation.

Fresh leaves were collected from a wild individuals of *C. yunnanensis* from the Jiangdong town of Dehong prefecture (Yunnan, China: 98°38’E, 24°53’N) and dried using silica gel and vouchar specimen (XTBG-LW001-M01) was deposited in Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences. Total genomic DNA was isolated using Tiangen Plant Genomic DNA Kits (Tiangen Biotech, Beijing). Illumina Hi-Seq 2500 were used to generate the 150 bp pair-end reads. There were total 16,626,54 reads for both ends. GetOrganelle pipeline (https://github.com/Kinggerm/GetOrganelle) was used to de novo assembling the chloroplast genome, with SPAdes v.3.10.1 as assembler (Bankevich et al. 2012), then visualized by Bandage v.0.8.1 (Wick et al. 2015) to determine paths of the plastome. We used CpGAVAS (Liu et al. 2012) to annotate the chloroplast genome automatically, then manually adjusted and visualized in Geneious Prime v.2019.1.3 (Kearse et al. 2012). A final circular chloroplast genome map was drawn using OGDRAW (Lohse et al. 2013). The newly annotated complete chloroplast genome was submitted to GenBank with accession number MN088379.

The complete chloroplast genome of *C. yunnanensis* is 163,166 bp in length, contains a quadripartite structure that consists of a large single copy (LSC) region of 91,580 bp and a small single copy (SSC) region of 20,440 bp with two inverted repeat (IR) regions of 25,573 bp. It harbors 113 unique genes, including 79 protein-coding genes, 30 tRNA genes, and 4 rRNA genes. The base compositions of *C. yunnanensis* chloroplast genome are uneven (31.4% A, 18.4% C, 18.1%G, 32.2% T), with an overall GC content of 36.4%.

To further examine its phylogenetic position, the maximum likelihood (ML) tree with 1000 bootstrap replicates was inferred by RAxML-HPC2 Workflow on XSEDE (v.8.2.12) CIPRES online portal (Miller et al. 2010) from alignments created by MAFFT (Katoh and Standley 2013) in Geneious Prime.
v.2019.1.3 (Kearse et al. 2012). We selected 14 published chloroplast genomes including *Aquilaria sinensis* KT148967 and *Aquilaria yunnanensis* MG656407 from Thymelaeaceae and 11 species from Malvaceae, with three species from Brassicaceae as outgroup. The phylogeny reconstructed revealed that *C. yunnanensis* is closely related to *Tilla* species of Malvaceae (Figure 1). This study will provide a fundamental resource for studying the phylogenetic relationship and conservation of this Endangered plant species with extremely small populations.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by grant from the National Natural Science Foundation of China [grants no. 31600439], National Key R&D Program of China [grants no. 2017YFC0505200].

**References**

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prijibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comp Biol. 19:455–477.

Gao ZR, Zhang CQ, Milne RL 2010. Size-class structure and variation in seed and seedling traits in relation to population size of an endangered species *Craigia yunnanensis* (Tiliaceae). Aust J Bot. 58: 214–223.

International Union for Conservation of Nature and Natural Resources (IUCN). 2008. The IUCN red list of threatened species. http://www.iucnredlist.org.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30:772–780.

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: 1647–1649.

Li YY. 2005. National protected wild plants in Yunnan Province China. Yunnan, Kunming, China: Science and Technology Press.

Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC Genomics. 13:715.

Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW—a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. Nucleic Acids Res. 41:W575–W581.

Ma YP, Chen G, Grumbine RE, Diao ZL, Sun WB, Guo HJ. 2013. Conserving plant species with extremely small populations (PSESP) in China. Biodivers Conserv. 22:803–809.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, LA. 1-8

Sun, W. 1998. *Craigia yunnanensis*. The IUCN red list of threatened Species 1998: e.T32335A9699030 [Accessed on 2019 May 22]. http://dx.doi.org/10.2305/IUCN.UK.1998.RLTS.T32335A9699030.en.

Wang S, Xie Y. 2004. China species red list, vol 1. Beijing: Higher Education Press.

Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics. 31:3350–3352.

Yang J, Gao ZR, Sun WB, Zhang CQ. 2016. High regional genetic differentiation of an endangered relict plant *Craigia yunnanensis* and implications for its conservation. Plant Diversity. 38:221–226.

**Figure 1.** The phylogenetic tree based on 14 complete chloroplast genome sequences. GenBank accession numbers: *Abelmoschus esculentus* (KY635876), *Althaea officinalis* (KY085914), *Aquilaria sinensis* (KT148967), *Aquilaria yunnanensis* (MG656407), *Arabidopsis thaliana* (NC_000932), *Bombax ceiba* (NC_037494), *Brassica juncea* (KT381449), *Brassica nigra* (KT878383), *Craigia yunnanensis* (MN088379), *Hibiscus syriacus* (KP688069), *Talipariti hamabo* (KR259988), *Tilia amurensis* (KT894772), *Tilia mandshurica* (KT894773), *Tilia oliveri* (KT894774), *Tilia paucicostata* (KT894775).