Characterization of the complete chloroplast genome of *Cycas hongheensis* (Cycadaceae), an endemic species in the red river region of China

Li-Xin Wu,^a,b,c,d^ Yue-Hua Wang^a^ and Xun Gong^b,c^ 

^a^Plant Science Institute, School of Life Sciences, Yunnan University, Kunming, Yunnan, China; ^b^Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China; ^c^Key Laboratory of Economic Plants and Biotechnology, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan, China; ^d^University of Chinese Academy of Sciences, Beijing, China

**ABSTRACT**

In this study, we determined the complete chloroplast genome of *Cycas hongheensis* (Cycadaceae), one of the first-class protected plants in China. The chloroplast genome is 162,048 bp in length with 133 genes, including 87 protein-coding genes, eight ribosomal RNA genes, and 37 transfer RNA genes. The overall GC content is 39.4%. Phylogenomic analysis showed that *C. hongheensis* as sister to all other *Cycas* species that with reported chloroplast genomes. The chloroplast genome of *C. hongheensis* reported here will contribute to further comparative chloroplast genome of Cycads and helpful to study the phylogeography of Cycadaceae.

Cycad is an ancient plant group, flourished from Late Triassic to early Cretaceous of Mesozoic era, with a diversification shrink until Late Cretaceous (Gao and Thomas 1989). There are about 24 Cycas species from China, in which Yunnan and Guangxi have the most species diversity (Ma et al. 2005). Due to wild *Cycas* populations were seriously damaged, all of the *Cycas* species in China were given ‘First Grade’ conservation status.

*Cycas hongheensis* S. Y. Wang et S. L. Yang ex D. Y. Wang was first proposed without any type specimen and formal Latin description in 1994, which was considered as invalidly publication (Vorster 1997). Despite a formal description was given (Wang et al. 1996), their publication lacked any morphological descriptions from reproductive organs (megasporophylls, microsporophylls and seeds). It was until the year 2016 that the morphological description of reproductive organs of *C. hongheensis* was supplemented (Liu 2016), showing distinct differences of *C. hongheensis* to the other Chinenses Cycas. This species only distributes in Gejiu City on the North Bank of Red River and grows on the dry and hot limestone hillside at low altitude. Based on the assessment of IUCN red list of threatened species, this species is considered to be Critically Endangered (Bösenberg 2010).

The fresh leaf tissues were collected from Gejiu City in Yunnan Province (23°11′N, 102°57′E) and stored at Kunming Institute of Botany (http://www.kib.ac.cn/), Linxin Wu, wulixin@mail.kib.ac.cn) under the voucher number KIB194344. Total genomic DNA was extracted by modified CTAB method (Doyle 1991). A total of 2 G raw data from Illumina Hiseq platform (Novogene, Beijing, China) were screened and assembled into contigs by Get_Organelle pipeline (Jun et al. 2020) with *Cycas szechuanensis* (NC_042668) as reference. The resulted contigs were reordered and further trimmed according to obtain the complete chloroplast genome. We applied PGA (Qu et al. 2019) to annotate the *C. hongheensis* chloroplast genome and double-checked in Geneious v11.0.3 (Kearse et al. 2012). The complete chloroplast genome sequence was deposited in NCBI under the accession (MZ562967). To infer the phylogenetic position of *C. hongheensis*, we used MAFFT v7 (Katoh and Standley 2013) align *C. hongheensis* and other seven *Cycas* chloroplast genomes (Figure 1) by using all the rest nine cycad genera and *Ginkgo biloba* as outgroup, all the chloroplast genomes were downloaded from NCBI. The maximum likelihood (ML) tree with 1000 bootstrap replicates was implemented in IQTREE 2.1.3 (Minh et al. 2020).

The complete chloroplast genome of *C. hongheensis* showed typical quadripartite structure with a length of 162,048 bp, which contains 133 genes in total, including 87 protein-coding genes (83 PCG species), 37 transfer RNA genes (30 tRNA species) and 8 ribosomal RNA genes (4 rRNA species). Most of these genes occurred as a single copy, while 15 gene species occurred as double copies, including all 4 PCG species (*ndhB*, *rps12*, *rps7*, *ycf2*), 7 tRNA species (*trnA-UGC*, *tmH-GUG*, *tml-GAU*, *tmL-CAA*, *tmN-GUU*, *tmR-AGC*, *tvm-V-GAC*) and 4 rRNA species (*rrn4.5*, *rrn5*, *rrn16*, *rrn23*). According to the reference genome, there are eight genes (*atpF*, *rpoC1*, *petD*, *rpl16*, *rpl2*, *ndhA*, *ndhB*, *ndhH*) occupying a single intron.
and two genes (clpP and ycf3) that have two introns. The complete chloroplast genome of *C. hongheensis* contains a large single-copy (LSC) region with 88,815 bp in length, a pair of inverted repeats (IRA and IRB) regions of 25,088 bp each, and a small single-copy (SSC) region with a length of 23,057 bp. The overall GC content of *C. hongheensis* chloroplast genome is 39.4%, with the number 42.0% in IR, 38.7% and 36.6% in LSC and SSC regions, respectively.

Phylogenomic analysis generated a well-resolved tree for cycads (Figure 1), and two clades were revealed in Cycadales: clade I consisted of all the Zamiaceae, and clade II corresponded to Cycadaceae (*Cycas*). Within clade II, *C. hongheensis* was resolved and strongly supported as sister to all other *Cycas* species that with reported chloroplast genomes. In general, this study can contribute to further comparative studies on chloroplast genomes in Cycadaceae.

**Disclosure statement**

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

**Funding**

This work was supported by the West Light Foundation of the Chinese Academy of Sciences [Y8246811W1].

**Data availability statement**

The data that support the findings of this study are openly available in NCBI at [https://www.ncbi.nlm.nih.gov](https://www.ncbi.nlm.nih.gov) under the accession number MZ562967. The BioProject, SRA, and Bio-Sample numbers are PRJNA752001, SRR15342259, and SAMN20566043 respectively.

**References**

Bøsenberg JD. 2010. *Cycas hongheensis*. The IUCN Red List of Threatened Species; [https://www.iucnredlist.org/species/42040/10635591](https://www.iucnredlist.org/species/42040/10635591) [accessed 2021 July 12]. 2010: e.T42040A10635591.

Doyle J. 1991. DNA protocols for plants. Vol. 57. In: Hewitt GM, Johnston AWB, Young JPW, editors. Molecular techniques in taxonomy. Switzerland: Springer; p. 283–293.

Gao ZF, Thomas BA. 1989. A review of fossil cycad megasporophylls, with new evidence of *Crossozamia* Pomel and its associated leaves from the Lower Permian of Taiyuan, China. Rev Palaeobot Palynol. 60(3–4):205–223.

Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21(1):241.

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.

Liu J. 2016. Molecular phylogeny and biogeography of *Cycas* L. Beijing: University of Chinese Academy of Sciences.

Ma Y, Li N, Su JX, Lin JZ. 2005. Taxonomy of *Cycas* in China: current status and perspective. J Shaxi Normal University. 19(2):73–77.

Minh BQ, Schmidt HA, Chenomor Q, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 37(5):1530–1534.

Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes . Plant Methods. 15:50.

Vorster P. 1997. The correct author citation for *Cycas hongheensis* and *C. parvula* (Cycadaceae). Acta Phytotaxonomica Sinica. 35(6):550–551.

Wang FX, Liang HB, Chen TQ, Wang DY. 1996. Cycads in China. Guangzhou: Guangdong Science and Technology Press. p. 62.

**Figure 1.** Phylogenetic inference of cycads based on the maximum-likelihood (ML) analysis using complete chloroplast genome data. Numbers on nodes represent the bootstrap values based on 1000 replicates from IQTREE.