The complete chloroplast genome of *Ottelia acuminate* var. *crispa*, an endangered aquatic herb with extremely narrow distribution

Chaoyong Tu, Hao Lin, Qiang Wang, Lijun Wu, Yulian Yang, Lijuan Zou and Qinggui Wu

Ecological Security and Protection Key Laboratory of Sichuan Province, College of Life Science and Biotechnology, Mianyang Normal University, Mianyang, China

**ABSTRACT**

*Ottelia acuminate* var. *crispa* is an endangered aquatic herb with extremely narrow distribution. In this study, we assembled the chloroplast genome of this species based on the second-generation high-throughput sequencing. The genome is 157,783 bp in length with a typical quadripartite structure including a large single-copy region (LSC) of 88,294 bp, a small single-copy region (SSC) of 49,379 bp, and a pair of inverted repeats (IRs) of 10,055 bp each. A total of 128 genes were annotated, including 83 protein-coding genes (PCG), 37 tRNA genes, and 8 rRNA genes. The phylogenetic tree shows that *O. acuminate* var. *crispa* has a close relationship with the genus *Elodea*. The chloroplast genome presented here provides a valuable resource to conserve this endangered species.

*Ottelia acuminate* var. *crispa* (Hand.-Mazz.) H. Li (*Ottelia*, Hydrocharitaceae) is an endangered aquatic herb that currently only exists in Lugu Lake in southwest China. As a submerged dioecious plant, water and light environment, climatic conditions, and nutrients are often a factor to it and now have been an endangered species. It was classified in the IUCN Red list as critically endangered species in 2004 (IUCN Nepal 2004). *Ottelia acuminate* var. *crispa* flowering throughout the year could be an important ornamental aquatic plant attract tourists, while its leaves are also of certain edible value. In addition, *O. acuminate* var. *crispa* could be used in monitoring water pollution and environmental protection due to its high requirements for water quality environment. However, little is known about the complete chloroplast genome (cp-genome) information of such an endangered species. In this study, we reported the cp-genome of *O. acuminate* var. *crispa*. The annotated genome has been submitted to GenBank under the accession number MT993696 (https://www.ncbi.nlm.nih.gov/nuccore/MT993696).

We collected the fresh leaves of an *O. acuminate* var. *crispa* individual from Lugu Lake in Yunnan Province, China (100.8207°E, 27.7002°N). Voucher specimen (MNU-PHO-0154) of the species was stored in the Ecological Security and Protection Key Laboratory of Sichuan Province, China. We used the DNA-secure Plant Kit (TIANGEN) to extract the total DNA from samples. The whole-genome sequencing was carried out with the HiSeq X Ten Platform (Illumina, USA). Finally, we generated about 10 G high-quality base pairs of raw data. The software fast QC (Gdula et al. 2019) and Trimmomatic (Bolger et al. 2014) were used to filter the raw data to obtain clean reads for the subsequent analysis. The software BWA v0.7.12 (Li and Durbin 2009) and SAMtools v1.3.1 (Li et al. 2009) were used to map the clean reads to the reference *Ottelia cordata* (MN056354.1) chloroplast genome. The mapped reads were extracted and then used to assemble the genome with the software NOVOPlasty v4.1 (Dierckxsens et al. 2017). The resulting contig was aligned to the reference genome by Bwa and Samtools again. We use Geneious version 8.1.4 (Kearse et al. 2012) to compare and adjust the assembled cp-genome sequence manually. Finally, we obtained a cp-genome of *O. acuminate* var. *crispa* with a size of 157,783 bp after filling the gaps with GapClosers (Luo et al. 2012). We annotated the cp-genome with Plann v1.0 (Huang and Cronk 2015) and checked the quality with Sequin v15.10 (Clark et al. 2016).

The cp-genome of *O. acuminate* var. *crispa* is with a typical quadripartite structure which consists of a large single-copy region of 88,294 bp, a small single-copy region of 49,379 bp, two reverse repeat regions of 10,055 bp each and a GC content of 36.58%. Genome annotation reveals a total of 127 genes, including 84 protein-coding genes (PCGs), 35 transfer RNA (tRNA) genes, and 8 ribosomal RNA (rRNA) genes. To infer the phylogenetic position of *O. acuminate* var. *crispa*, we reconstructed a phylogenetic tree with other 6 complete chloroplast genomes downloaded from NCBI (Figure 1) including *O. acuminate* var. *crispa*. The sequences were aligned using the software MAFFT (Katoh and Standley 2013) and the maximum likelihood analysis worked on
software RAxML v8.2.9 (Stamatakis 2014) setting GTRGAMMA as the best model and 1000 bootstrap tests. The phylogenetic tree demonstrates that *O. acuminate* var. *crispa* is closely related to genus *Elodea* with strong support (Figure 1). In summary, the complete chloroplast genomic of *O. acuminate* var. *crispa* will not only provide valuable informative resource to facilitate the identification, conservation, and utilization of this extremely endangered species but also provide insightful information to further research on phylogenetic study of Hydrocharitaceae.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was financially supported by the Sichuan Provincial Science and Technology Project (2018JY0244 and 2018JY0192), the Scientific research initiation project of Mianyang normal university (QD2019A12) and the Open Project from the Ecological Security and Protection Key Laboratory of Sichuan Province (ESP1903).

**Data availability statement**

The annotated genome has been submitted to GenBank under the accession number of MT993696 (https://www.ncbi.nlm.nih.gov/nuccore/MT993696). The voucher specimen of the species is free accessible at Ecological Security and Protection Key Laboratory of Sichuan Province, China with the No. MNU-PHO-0154.

**References**

Bolger A, Lohse M, Usadel B. 2014. Trimmmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30(15):2114–2120.

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. 2019. GenBank. Nucleic Acids Res. 44:D67–D72.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.

Gdula MR, Nesterova TB, Pintacuda G, Godwin J, Zhan Y, Ozadam H, McClellan M, Moralli D, Krueger F, Green CM, et al. 2019. The non-canonical SMC protein SmcHD1 antagonises TAD formation and compartmentalisation on the inactive X chromosome. Nat Commun. 10(1):30.

Huang D, Cronk Q. 2015. Plann: a command-line application for annotating plastome sequences. Appl Plant Sci. 3(8):1500026.

IUCN, A Review of the Status and Threats to Wetlands in Nepal, IUCN Nepal, Kathmandu, Nepal, 2004.

Katoth K, Standley D. 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.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 25(14):1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25(16):2078–2079.

Luo R, Liu B, Xie Y, Li Z, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Giga Sci. 1:1–6.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.