Complete chloroplast genome of *Holboellia grandiflora* Réaulebour (Lardizabalaceae), a plant with an edible fruit

Yan Li¹, Qinxiang Chang¹, Caijuan Zhang², Huimin Li² and Pengguo Xia²

¹Department of Art Design, Taiyuan University, Taiyuan, PR China; ²Key Laboratory of Plant Secondary Metabolism and Regulation of Zhejiang Province, College of Life Sciences and Medicine, Zhejiang Sci-Tech University, Hangzhou, PR China

**ABSTRACT**

*Holboellia grandiflora* Réaulebour (Lardizabalaceae) is an evergreen twining perennial woody vine. To our knowledge, this is the first report on the complete chloroplast genome sequence of *H. grandiflora*. The complete chloroplast genome sequence was 157,811 bp in length and contained a large single-copy region of 86,554 bp and a small single-copy region of 18,975 bp. A pair of inverted repeats of 26,141 bp were included. It contained 130 genes, comprising 37 transfer RNA genes, and eight ribosomal RNA genes, as well as 85 coding sequences (CDSs). The GC content of the complete chloroplast genome sequence was 38.7%. The phylogenetic tree showed a close relationship among the three species of *Holboellia* (*H. grandiflora*, *H. angustifolia*, and *H. latifolia*). These findings provide a reference for phylogenetic relationships and assessment of the genetic structure of the Lardizabalaceae family.

*Holboellia grandiflora*, Réaulebour 1906, which belongs Lardizabalaceae, is an evergreen twining perennial woody vine mainly distributed at an altitude of 800–3000 m (Flora of China Editorial Committee 2001). *Holboellia grandiflora* does not currently exist on the IUCN Red List. Its pistillate and staminate flowers are purple and the staminate ones are green, respectively. The fruit of *H. grandiflora* can be consumed and also used for wine making. In addition, its seeds can be pressed into oil. Currently, this species is widely distributed in the wild and has not been developed and used. It is well adapted to environmental conditions and has strong resistance to stress (Tao et al. 2013).

Fresh leaf tissue of *H. grandiflora* was sampled from a mountain located in Longwang Town, Ningshan County, Ankang City, Shaanxi Province (33°10′39.88″N, 108°31′28.14″E, and altitude 853 m). The voucher specimen was preserved at the Herbarium of Xi’an Botanical Garden (voucher number: XunLulu et al. 00780, Lulu Xun, xunlulu20032006@126.com).

Total genomic DNA was extracted from fresh leaves using a modified CTAB method (Doyle 1987). The sequencing library was constructed using the Illumina Hiseq6000 platform with an insert size of approximately 490 bp. In total, raw PE reads of approximately 1.86 Gb were generated and approximately 1.85 Gb of clean PE reads were obtained after using fastp application (Chen et al. 2018) to trim and filter them. The complete chloroplast genome of *H. grandiflora* was assembled using the software NOVOPlast version 2.7.2 software (Dierckxsens et al. 2017) (Brussels, Belgium). The chloroplast genome was determined using Geneious Prime software, referring to the sequence of *H. angustifolia* (NC053741), and was submitted to GenBank (GenBank ID MW970138). The complete chloroplast genome sequence of *H. grandiflora* was 157,811 bp in length, with a large single-copy region of 86,554 bp and a small single-copy region of 18,975 bp. A pair of inverted repeats of 26,141 bp were included. It contained 130 genes, comprising 37 tRNA genes, 8 rRNA genes, and 85 coding sequences (CDSs).

Based on the complete chloroplast genomes of *H. grandiflora* and related species, we constructed a phylogenetic tree using maximum likelihood (ML) phylogenetic analyses (Figure 1); phylogenetic tree was constructed using IQTREE version 1.6.7 (Nguyen et al. 2015) (Vienna, Austria) based on 13 complete chloroplast genome sequences of Lardizabalaceae and *Euptelea pleiosperma* (NC029449) as the outgroup, under K3Pu + F + G4 model with 1000 bootstrap replicates. Phylogenetic analysis revealed that *H. grandiflora* has a strong sister relationship with *H. angustifolia* (NC053741) and *H. latifolia* (MH394375).

The gene content, GC content, and gene order of *H. grandiflora* were found to be similar to those of *H. angustifolia* (Wu et al. 2019). Additionally, a total of 49 microsatellites (SSRs) were identified in the *H. grandiflora* cp genome using MISA. Among them, 37 were located in the LSC regions, whereas 2 and 8 were found in the IR and SSC regions, respectively. The majority of the SSRs (44/49) were A or T repeats, which was consistent with the A/T-richness in the complete cp genome (Xuan et al. 2013). The complete chloroplast genome sequence provides the necessary data for the study of organ development genes and phylogenetic studies of the Lardizabalaceae family. The findings of this
study could be useful for further studies on *H. grandiflora* and its various applications.

**Ethical approval**

Research and collection of plant material were conducted according to the guidelines provided by Xi’an Botanical Garden. Permission was granted by Hangzhou Academy of Agricultural Sciences to carry out research on the species.

**Author contributions**

Y.L. and P.X. conceived and designed this study. C.Z. and H.L. conducted analysis. P.X. and C.Z. contributed the analytical methods. Y.L. wrote the manuscript. P.X. edited the manuscript. All authors have read and agreed to the published version of the manuscript.

**Disclosure statement**

No conflict of interest exits in the submission of this manuscript, and manuscript is approved by all authors for publication. And we declared comply with the International Union for Conservation of Nature (IUCN) policies research involving species at risk of extinction, the Convention on Biological Diversity and the Convention on the Trade in Endangered Species of Wild Fauna and Flora.

**Funding**

This work was supported by the Basic Public Welfare Research Project of Zhejiang Province [LGN19H280004].

**ORCID**

Pengguo Xia  http://orcid.org/0000-0003-3572-7616

**Data availability statement**

The data that support the findings of this study are openly available in NCBI (https://www.ncbi.nlm.nih.gov) GenBank with the accession number MW970138. The associated BioProject, SRA, and BioSample funding numbers are PRJNA723130, SRR14278902, and SAMN18805688, respectively.

**References**

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 34(17):i884–i890.

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

Doyle JJ. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.

Flora of China Editorial Committee. 2001. Flora of China. Chin Acad. 29:16–17.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.

Tao GJ, Yang TF, Tao GL. 2013. Investigation of *Holboellia grandiflora* Reaub Fruit. China Wild Plant Resources. 32(002):68–69.

Wu XZ, Li TJ, Wen F. 2019. The complete chloroplast genome of *Holboellia angustifolia* (Ranales: Lardizabalaceae), a traditional herbal species. Mitochondrial DNA Part B. 4(2):3358–3359.

Xuan Y, Lei G, Bo W, Su YJ, Wang T. 2013. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): evolutionary comparison of Cephalotaxus chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms. Genome Biol Evol. 5(4):688–698.