

MITOGENOME ANNOUNCEMENT



## Characterization of the complete chloroplast genome of *Phtheirospermum japonicum* (Rhinanthaeae)

Yong-Chao Li, Xin-Liang Zhao, Jing Yang and Qing-Yun Fu

School of Life Science and Technology, Henan Institute of Science and Technology, Xinxiang, Henan Province, China

### ABSTRACT

In this study, we report the complete chloroplast (cp) genome of *Phtheirospermum japonicum* was determined through Illumina sequencing method. The complete chloroplast genome of *Ph. japonicum* was 153,397 bp in length and contained a pair of IR regions (25,601 bp) separated by a small single copy region (17,728 bp) and a large single copy region (84,467 bp). The cp genome of *Ph. japonicum* encoded 125 genes including 86 protein-coding genes, 31 tRNA genes and eight ribosomal RNA genes. The overall GC content of *Ph. japonicum* cp genome is 38.3%. By phylogenetic analysis using ML method, *Ph. japonicum* was placed in family Orobanchaceae and showed the closest relationship with *Castilleja paramensis*.

### ARTICLE HISTORY

Received 23 August 2019  
Accepted 23 September 2019

### KEYWORDS

*Phtheirospermum japonicum*;  
chloroplast genome;  
phylogenetic analysis

*Phtheirospermum japonicum*, belonging to Rhinanthaeae (Orobanchaceae) is a hemiparasitic plant that widely distributed in Eastern Asia. In this study, we report the chloroplast genome of *Ph. japonicum* using next-generation sequencing, aiming to figure out its phylogenetic position precisely through molecular phylogenetic analysis.

Plant materials of *Phtheirospermum japonicum* sequenced in this study were collected from hengduan mountains of Southwest China (27°30'N, 99°61'E). Both plant materials and total genomic DNA that extracted from fresh young leaves using cetyltrimethylammonium bromide (CTAB) method was stored in Henan Institute of Science and Technology. The corresponding specimen was also stored in herbarium of Henan Institute of Science and Technology under accession NO. IST-20190611-Phjap.

For high-throughout sequencing (NGS), paired-end library from DNA extracts were prepared with a NEBNext Library building kits, following manufacturer's protocol. Then, the library was sequenced on an Illumina HiSeq2500 platform. After reads quality filtration, the clean reads were assembled by SPAdes 3.11.0 (Bankevich et al. 2012). We used chloroplast genome of *Schwalbea americana* (accession NO.: HG738866) as a reference sequence to align the contigs and identify gaps. To fill the gap, Price (Ruby et al. 2013) and MITObim v1.8 (Hahn et al. 2013) were applied and Bandage (Wick Ryan et al. 2015) was used to identify the borders of the IR, LSC, and SSC regions. The complete sequence was primarily annotated by Plann (Huang et al. 2015) combined with manual correction. All tRNAs were confirmed using the tRNAscan-SE

search server (Lowe and Eddy 1997). Other protein-coding genes were verified by BLAST search on the NCBI website (<http://blast.ncbi.nlm.nih.gov/>), and manual correction for start and stop codons were conducted. This complete chloroplast genome sequence together with gene annotations was submitted to GenBank under the accession numbers of MN075943.

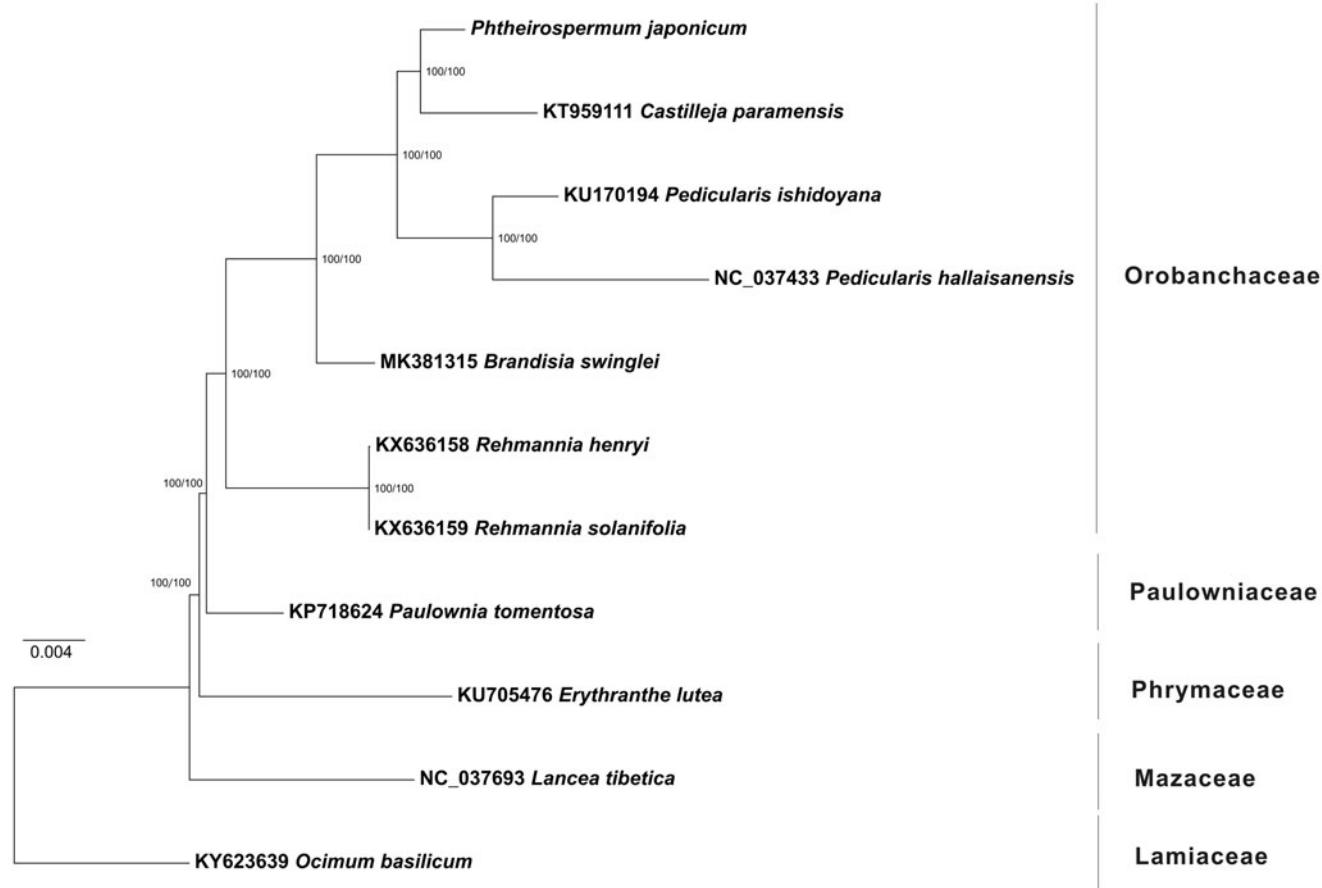
The chloroplast genome of *Ph. japonicum* is a typical quadripartite structure with a length of 153,397 bp. The whole cp genome contains a large single-copy (LSC) region of 84,467 bp, a small single-copy (SSC) region of 17,728 bp, and two inverted repeat (IRs) regions of 25,601 bp. The cp genome possesses 125 genes, including 86 protein-coding genes, 8 ribosomal RNA genes (4 rRNA species) and 31 tRNA genes. The overall GC content of the cp genome is 38.1%.

To infer the phylogenetic relationships between *Ph. japonicum* and the related species, the whole cp genome sequences of 11 species of Lamiales were aligned using HomBlocks (Guigui et al. 2017), resulting in 101,544 positions in total. The whole genome alignment was analyzed by IQ-TREE version 1.6.6 (Nguyen et al. 2015) under the GTR + F + R2 model. The tree topology was verified under both 1000 bootstrap and 1000 replicates of SH-aLRT test. As shown in Figure 1, the phylogenetic positions of these 11 cp genomes were successfully resolved with full bootstrap supports across almost all nodes. *Phtheirospermum japonicum* was placed in family Orobanchaceae and showed the closest relationship with *Castilleja paramensis*.

**CONTACT** Yong-Chao Li ✉ [histlyc@126.com](mailto:histlyc@126.com) School of Life Science and Technology, Henan Institute of Science and Technology, Xinxiang, Henan Province, China

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** Phylogenetic tree yielded by ML analysis of 11 lamiids cp genomes. Nodes harvesting both full bootstrap and SH-like aLRT values are indicated by pink points. Scale indicates substitutions per site.

**Disclosure statement**

The authors have declared that no competing interests exist.

**Funding**

This study was financially supported by the Major Science and Technology Projects of Henan Province Advanced Education [17A530002].

**References**

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

Guiqi B, Mao Y, Xing Q, Cao M. 2017. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics.* 110:18–22.

Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* 41: e129–e129.

Huang Daisie I., Quentin CB Cronk. 2015. Plann: a command-line application for annotating plastome sequences. *Applications in plant sciences.* 3:1500026.

Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964.

Nguyen L-T, Schmidt HA, Haeseler A. V, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32:268–274.

Ruby JG, Priya B, Joseph LD. 2013. PRICE: software for the targeted assembly of components of (meta) genomic sequence data. *G3: Genes| Genomes. G3 (Bethesda).* 3:865–880.

Wick Ryan R, Schultz Mark B, Schultz J, Holt Kathryn E. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics.* 31:3350–3352.