The complete chloroplast genome sequence of a critically endangered orchid *Paphiopedilum gratrixianum* (Orchidaceae)

Chao Hu, Kai Jiang, Xinhua Zeng and Weichang Huang

Eastern China Conservation Centre for Wild Endangered Plant Resources, Shanghai Chenshan Botanical Garden, Shanghai, China

ABSTRACT

*Paphiopedilum gratrixianum* is a critically endangered orchid with extremely small populations and of great horticulture value. We assembled and annotated the complete chloroplast genome of *P. gratrixianum* in this study. The genome is 157,292 bp in length and 35.8% of the GC content. It encodes a total of 102 unique genes, including 68 protein-coding genes, 30 tRNA genes, and four rRNA genes. Phylogenetic analysis results strongly supported that it was closely related to *P. tranlienianum*. The complete chloroplast genome of *P. gratrixianum* (GenBank accession number: MW284890) can be useful for the evolutionary study in Orchidaceae.

**Paphiopedilum** Pfitzer is an important genus of Orchidaceae which is famous for the showy flower. *Paphiopedilum gratrixianum* Rolfe is distributed in SE Yunnan, China, Laos and Vietnam (Wu et al. 2009). It is terrestrial or lithophytic and the habitat is restricted to karst area. It had been classified as a worldwide protected wild plant in CITES appendix I. The IUCN Red List of Threatened Species ([https://www.iucnredlist.org/species/191732/2000229](https://www.iucnredlist.org/species/191732/2000229)) defined *P. gratrixianum* as Critically Endangered (CR). It is also listed as one of the Plant Species with Extremely Small Populations (PSESP) in China. So it is urgent to protect the germplasm resource of *P. gratrixianum*.

The fresh leaves were collected from the greenhouse of Shanghai Chenshan Botanical Garden and dried in silica gel. The specimen and DNA were deposited in Shanghai Chenshan Herbarium (CSH). The plant total genomic DNA kit (Tiangen, Beijing, China) was used for total genomic DNA extraction. And the libraries was prepared for subsequent genome skimming sequencing. Paired-end reads of 2 × 150 bp were generated by Illumina HiSeq 4000 platform.
Finally, 4G raw data was used for complete chloroplast genome assembly. GetOrganelle software (Jin et al. 2020) was used for the assembly with *P. purpuratum* (Lindl.) Stein (GenBank accession number: MN535015) as a reference. Annotation was performed by DOGMA (Wyman et al. 2004) and modified manually on Geneious (Gibbs, XXXX). Phylogenetic tree based on whole chloroplast genome was performed by RAXML v7.0.4 (Stamatakis 2006) with 1000 bootstraps for the calculation of the branch supports (Figure 1).

Chloroplast genome which is derived from endosymbiont of cyanobacterium (Dyall et al. 2004) is conserved in gene content and quadripartite structure. The total length of the genome was 157,292 bp, and the length for large single-copy region (LSC), small single-copy region (SSC), and two inverted repeat regions (IRa and IRb) were 87,252 bp, 1828 bp and 34,106 bp, respectively. The genome encodes 79 (68 unique) protein-coding genes and 4 pseudogenes (3 unique), 38 (30 unique) tRNA genes, and 8 (4 unique) rRNA genes. The IR region contains 13 genes (including pseudogenes) with some genes from SSC shift to this region such as *ndhD* and *ccsA* compared with *Apostasia odorata* (KM244734). The GC contents for the whole genome and three different parts (LSC, SSC and IR) are 35.8%, 33.2%, 23.8%, 39.5%, respectively. The phylogenetic study indicated that *Paphiopedilum* formed a sister lineage to *Phagmipedium* Rolfe within Cypripedioideae. And *P. gratrixianum* was highly supported to be closely related to *P. tranlienianum* O.Gruss & Perner.

**Disclosure statement**

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

**Funding**

This work was supported by a grant from the Science and Technology Commission Shanghai Municipality to Weichang Huang [19390743600] and project of Shanghai Landscaping and City Appearance Administrative Bureau to Chao Hu [G192424, G202401].

**Data availability statement**

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov/] under the accession no. MW284890. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA694727, SRR13526459 and SAMN17572096, respectively.

**Acknowledgments**

The authors are grateful to Ziyi Ni, Li Shao for the cultivation of samples.

**References**

Dyall SD, Brown MT, Johnson PJ. 2004. Ancient invasions: from endosymbionts to organelles. Science. 304(5668):253–257.

Gibbs MD. De novo assembly and reconstruction of complete circular chloroplast genomes using Geneious Prime. [https://assets.geneious.com/documentation/geneious/App+Note+-+De+Novo+Assembly+of+Chloroplasts.pdf](https://assets.geneious.com/documentation/geneious/App+Note+-+De+Novo+Assembly+of+Chloroplasts.pdf)

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):1–31.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22(21):2688–2690.

Wu Z, Raven P, Hong D. 2009. Flora of China. Vol. 25 (Orchidaceae). Beijing: Science Press and St. Louis: Missouri Botanical Garden Press.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organelar genomes with DOGMA. Bioinformatics. 20(17):3252–3255.