

## The complete chloroplast genome sequence of a new variety of *Dendrobium officinale* 'zhong ke IV hao'

Zhimin Zhong<sup>a</sup>, Guifang Zhang<sup>a</sup>, Xiaoping Lai<sup>a</sup> and Song Huang<sup>b</sup>

<sup>a</sup>School of Chinese Materia Medica, Guangzhou University of Chinese Medicine, Guangzhou, China; <sup>b</sup>Dongguan Mathematical Engineering Academy of Chinese Medicine, Guangzhou University of Chinese Medicine, Guangdong Dongguan, China

### ABSTRACT

Here, we reported and characterized the complete chloroplast (cp) genome sequence of *Dendrobium officinale* 'zhong ke IV hao', a new variety from self-cross plants of imported Sichuan *D. officinale*, obtained exclusively using Illumina and PacBio sequencing technology. The genome size is 152,185 bp, containing a large single copy (LSC) region (85,094 bp) and a small single copy (SSC) region (14,521 bp) that were separated by two inverted repeat (IRs) regions (26,285 bp). The GC content was 37.46%. In total, the complete cp DNA contains 89 protein-coding genes, 30 tRNA genes, and 8 rRNA genes. Twelve genes contained one or two introns. Phylogenetic analyses showed that the chloroplast genome of *D. officinale* 'zhong ke IV hao' is related to that of the traditional *D. officinale*.

### ARTICLE HISTORY

Received 11 July 2016  
Revised 20 July 2016  
Accepted 29 July 2016

### KEYWORDS

Complete chloroplast genome; *Dendrobium officinale* 'zhong ke IV hao'; single-molecule real-time (SMRT); PacBio

The complete chloroplast (cp) genome is commonly a typical circular double-stranded DNA molecule which has been considered as a super-barcode (Li et al. 2014). Sequencing and analyzing the cp-genomes of plants can provide evolutionary information for plant species identification, phylogeny, taxonomy, and ecology. Recently, the Pac Bio platform offers unprecedented sufficient throughput and long reads was recognized as the third generation sequencer, which would provide a potential and reliable sequence. Insufficiently, the Pac Bio long reads are error-prone and cannot be analyzed by traditional assembly programs effectively. Thus, we sequenced complete chloroplast genome of *Dendrobium officinale* 'zhong ke IV hao' with both Illumina and Pac Bio next-generation technology, in order to contribute to the current understanding of phylogenetic relationships within *Dendrobium*.

*Dendrobium* is one of the largest genera of Orchidaceae, which is the largest family in angiosperms (Xiang et al. 2013). As *D. officinale*, Kimura et Migo were Endangered and often adulterated with other cheaper and more common orchids (Pei Yang et al. 2016) for its economic and medicinal value. In order to solve the problems of wild *D. officinale* in short supply, *D. officinale* 'zhong ke IV hao', which has the characteristics of high production, strong resistance, good quality, and great market potential was cultivated of wild Sichuan *D. officinale*.

Genomic DNA was extracted from fresh leaves of *D. officinale* 'zhong ke IV hao' that lives in greenhouse and cultivated by Professor Duan jun from South China Botanical Garden (113.36N, 23.18W), Chinese Academy of Science

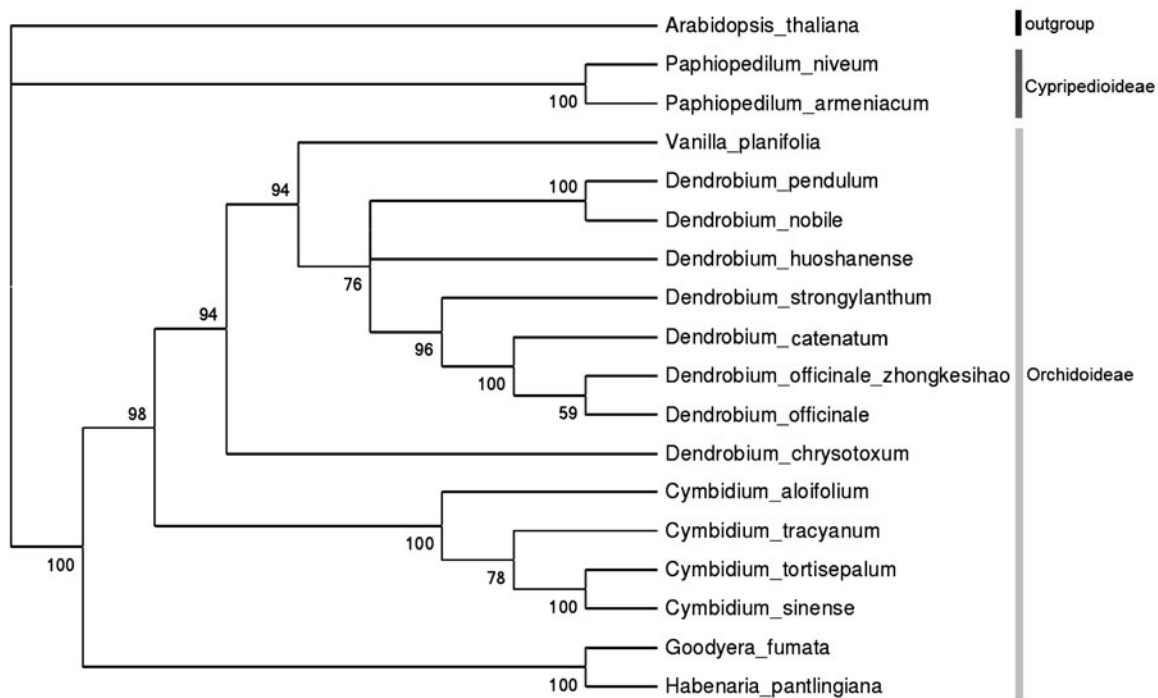
(Guangzhou, China). We sequenced its genome using Illumina HiSeq4000 (San Diego, CA) and Pac Bio next-generation technology. Genome sequencing was carried out using the Pacific Biosciences RSII (Menlo Park, CA) sequencing platform. Pac Bio long reads (one single-molecule real-time [SMRT] cells, 125 coverage) were assembled using the Hierarchical Genome Assembly Process 2 (HGAP2) protocol from SMRT Analysis version 2.0 package (18), resulting in the complete circular 152,185 bp cp-genome. The cp-genome was assembled using the SOAPdenovo (v2.04) and Celera Assembler (v8.0) and annotated with Dual Organellar GenoMe Annotator (DOGMA) (Wyman et al. 2004). The complete cp-genome sequence was submitted to GenBank under accession number of KX507360.

The chloroplast genome of *D. officinale* 'zhong ke IV hao' has a length of 152,185 bp and make up of a large single-copy region (LSC, 85,094 bp), a small single-copy region (SSC, 14,521 bp), and two inverted repeat regions (IRs, 26,285 bp). The total GC content is 37.46% while that of the IR regions is 43.0%. This is a little higher than that in the LSC region and the SSC region (35.0% and 30.0%). A total of 127 genes were successfully annotated, including 89 protein-coding genes, 30 tRNA genes, and 8 rRNA genes. The tRNA-coding genes are distributed throughout the genome, 17 in the LSC, one in the SSC, and 12 in the IR regions while rRNAs only situate in IR regions. The portions of protein-coding genes, tRNA genes and rRNA genes are 53.86%, 1.48 %, and 6.92 % in the 152,185 bp, respectively. Twelve genes contained one or two introns, containing the protein-coding genes, *atpF*, *clpP*, *ndhB*, *ndhB*, *ndhF*, *rpl2*, *rpl2*, *rpoC1*, *ycf1*, *ycf15*, *ycf15*, and *ycf3*.

**CONTACT** Xiao-ping Lai ✉ [lxp88@gzucm.edu.cn](mailto:lxp88@gzucm.edu.cn)  College of Chinese Materia Medica, Guangzhou University of Chinese Medicine, Guangzhou 510006, China

© 2016 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 based on neighbour-joining analysis of one outgroup plant and 17 chloroplast genome sequences belonging to Orchidaceae family, which Genbank accession numbers as follows: *A. thaliana* (NC\_000932), *Paphiopedilum niveum*(KJ524105), *P. armeniacum* (KJ566307), *Vanilla planifolia* (KJ566306), *D. pendulum* (KT695604), *D. nobile* (KT591465), *D. huoshanense* (KT630834), *D. strongylanthum* (KR673323), *D. catenatum* (KC771275), *D. officinale* 'zhong ke IV hao'(KX507360), *D. officinale* (KJ862886), *D. chrysotoxum* (KT633996), *Cymbidium aloifolium* (KC876122), *C. tracyanum* (KC876127), *C. tortisepalum* (KC876124), *C. sinense* (KC876123), *Goodyera fumata* (KJ501999), *Habenaria pantlingiana* (KJ524104).

The phylogenetic tree of *D.officinale* 'zhong ke IV hao' was produced using 17 chloroplast genome sequences of Orchidaceae family which were reported on Genbank (Figure 1). Neighbour-joining analysis was performed with Mega 6. (MEGA Inc., Englewood, NJ) (Tamura et al. 2013), setting model plant *Arabidopsis thaliana* as outgroup and using 500 bootstrap analysis. The bootstrap values were ranged from 59 to 100 in all nodes. As expected *D.officinale* 'zhong ke IV hao' is mostly related to traditional *D.officinale* that collected from Zhejiang Province.

### Disclosure statement

The authors report no conflicts of interest, and alone are responsible for the content and writing of the paper.

### Funding

This study was financially supported by the HongKong, Macao and Taiwan Science & Technology Cooperation Program of China [No. 2014DFH30010], Science and Technology Plan Projects of

Guangdong Province in China [2015B020234008], Natural Science Foundation of Guangdong Province China [2014A030304059] and National Natural Science Foundation of China [81001601] Excellent Young Teachers Program of Guangdong Province, China [yq2014042].

### References

- Li XW, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. 2014. Plant DNA barcoding: from gene to genome. *Biol Rev Camb Philos Soc.* 90:157–166.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol.* 30:2725–2729.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.
- Xiang XG, Schuiteman A, Li DZ, Huang WC, Chung SW, Li JW, Zhou HL, Jin WT, Lai YJ, Li ZY, Jin XH. 2013. Molecular systematics of *Dendrobium* (Orchidaceae, Dendrobieae) from mainland Asia based on plastid and nuclear sequences. *Mol Phylogenet Evol.* 69:950–960.
- Yang P, Zhou H, Qian J, Xu H, Shao Q, Li Y, Yao H. 2016. The complete chloroplast genome sequence of *Dendrobium officinale*. *Mitochondrial DNA A DNA MappSeq Anal.* 27:1262–1264.