

MITOGENOME ANNOUNCEMENT



Characterization of the complete chloroplast genome sequence of *Daphne retusa* Hemsl. (Thymelaeaceae), a rare alpine plant species in northwestern China

Fang Yan^{a,d}, Chun-Yun Zhang^b, Qin-Li Wang^{a,c}, Jian-Dong Wang^b, Hai-Ping Wang^b, Tao Xu^d, En-Jun Wang^d and Liang-Yu Hou^d

^aKey Laboratory of Hexi Corridor Resources Utilization of Gansu, Hexi University, Zhangye, China; ^bTaikang Pharmaceutical Limited Liability Company of Gansu Province, Wuwei, China; ^cGansu Engineering Research Center for Resource Utilization of SMS, Hexi University, Zhangye, China; ^dApplied Technology Research Institute on original plants for Zushima of Hexi University, Zhangye, China

ABSTRACT

Daphne retusa Hemsl. (Thymelaeaceae) is an evergreen shrub plant. First, we characterized the complete nucleotide sequence of chloroplast (cp) genome of *D. retusa*. The total length of cp genome was found to be 170,553 bp, including a large single copy (LSC) region of 84,886 bp, a small single copy (SSC) region of 2,437 bp, and a pair of inverted repeats (IRs) of 41,617 bp. The cp genome of *Daphne retusa* Hemsl. contains 134 genes, including 90 protein-coding genes (75 PCG species), 37 transfer RNA genes (29 tRNA species), and 6 rRNA genes (3 RNA species). A total of 13 genes (*trnK-UUU*, *trnS-CGA*, *atpF*, *rpoC1*, *trnL-UAA*, *trnC-ACA*, *petD*, *rpl16*, *rpl2*, *ndhB*, *trnE-UUC*, *ndhA*, and *trnA-UGC*) contain a single intron, and one gene (*ycf3*) contains two introns. The GC content in whole cp genome, LSC region, SSC region, and IR region was 36.75%, 34.83%, 28.19%, and 38.96% respectively, like other Thymelaeaceae plants. Phylogenetic analysis suggested that *D. retusa* has a close relationship with congeneric *Daphne tangutica*.

ARTICLE HISTORY

Received 25 February 2021
Accepted 12 June 2021

KEYWORDS

Daphne retusa;
Thymelaeaceae; complete
chloroplast genome;
phylogenetic relationship

Daphne retusa Hemsl. (Thymelaeaceae) is a slow-growing evergreen shrub with a maximum height of 1 m, and it has flexible, off-white stems. This plant often grows on virgin land on hillsides at elevations of 2500–3000 m in the mountains in northwestern China. The dried stem and root bark of *D. retusa* are known as ‘Zushima’ which is a common traditional Chinese medicine (Jamila et al. 2018). However, the wild sources of this plant have been nearly used up due to over-exploitation (Geng et al. 2016). In this study, we report the chloroplast genome sequence of *D. retusa*, and it would be essential to the conservation of this endangered species. The annotated genome sequence has been submitted to GenBank under accession number MW245832.

Fresh leaves of *D. retusa* were collected from Longnan Mountains (33°38′23″N, 104°50′30″E, 2910 m) in Gansu, China. The total DNA was extracted using the modified CTAB method (Allen et al. 2006) and used for the high-throughput Illumina HiSeq sequencing by Biomarker Technologies, Inc. (Beijing, China). A voucher specimen (ayrx_a) is deposited at the Building B2 of Kechuang 6 Road, Yard 88, Beijing Economic-Technological Development Area (E-mail: lingfeng_zeng@163.com). In total, 47,102,044 raw PE150 reads were

obtained and quality-trimmed using the software in the NGSQC Toolkit v2.3.3 (Patel and Jain 2012) with 2% allowed ambiguous bases. The Geneious Prime (Kearse et al. 2012) was used to assemble the complete chloroplast genome. SPAdes v3.11.1 (<http://cab.spbu.ru/software/SPAdes>) was employed to obtain the optimal assembly results using multiple Kmer parameters (107, 117, 127), with the closely related species *Daphne tangutica* (GenBank: MK455800) serving as the reference. The assembled genome sequence was annotated using CPGAVAS2 (Biomatters Ltd., Auckland, New Zealand) by comparing it to the cpDNA of *Daphne tangutica* (GenBank: MK455800). Subsequently, a maximum likelihood (ML) tree with 1000 bootstrap replicates was inferred by MEGA7 with default parameters (Kumar et al. 2016). The iTOL (<https://itol.embl.de/>) was used for tree visualization.

The total length of nucleotide sequence of plastid genome is 170,553 bp, including a large single copy (LSC) region of 84,886 bp, a small single copy (SSC) region of 2,437 bp, and a pair of inverted repeats (IRs) of 41,617 bp (Figure 1). The cp genome of *D. retusa* contains 134 genes, including 90 protein-coding genes (75 PCG species), 43 transfer RNA genes (37 tRNA species), and 6 rRNA genes (4 RNA species). A total of 13

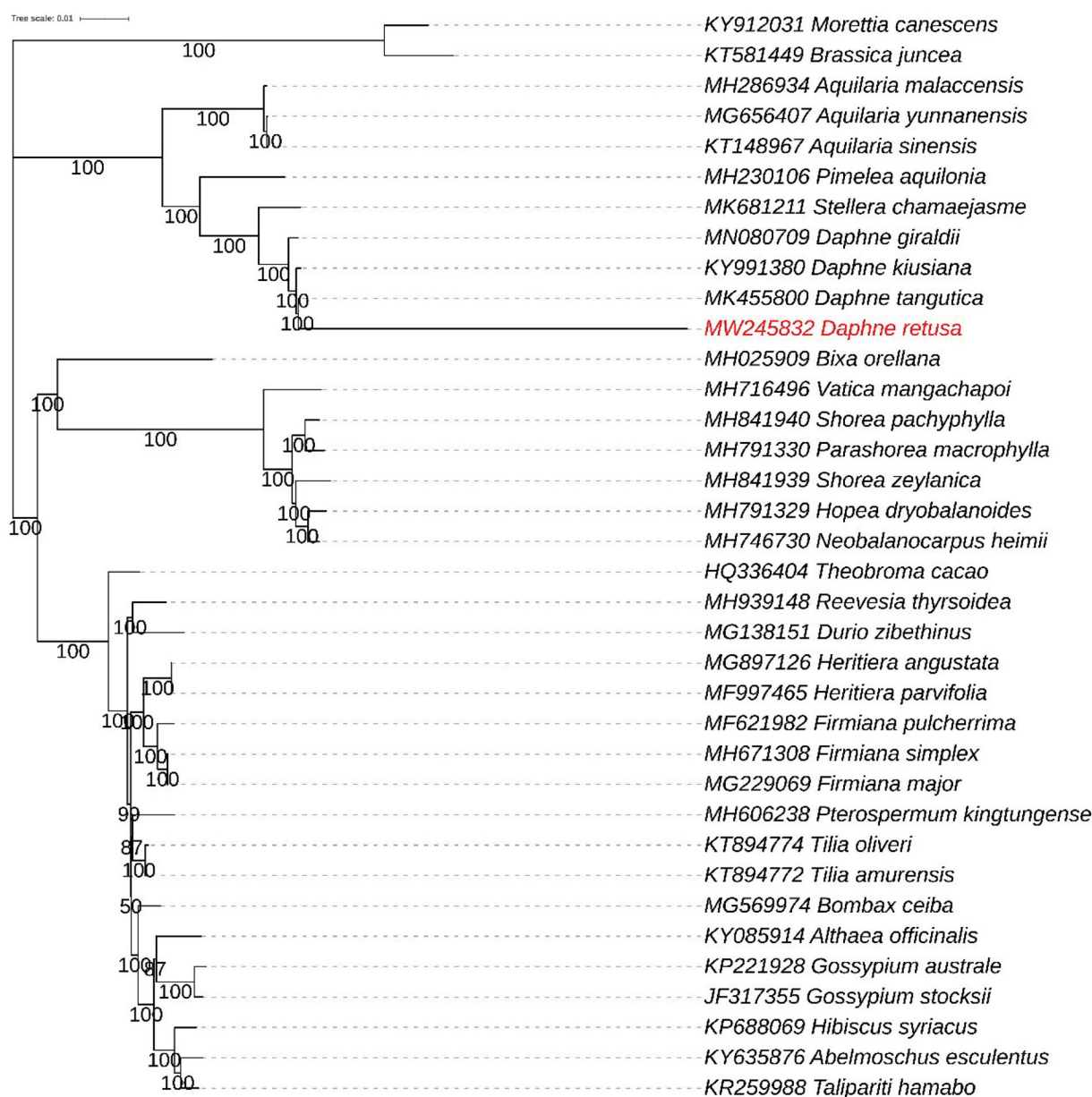


Figure 1. Maximum likelihood (ML) tree of *D. retusa* and its related relatives based on the complete chloroplast genome sequences.

genes (*trnK-UUU*, *trnS-CGA*, *atpF*, *rpoC1*, *trnL-UAA*, *trnC-ACA*, *petD*, *rpl16*, *rpl2*, *ndhB*, *trnE-UUC*, *ndhA*, and *trnA-UGC*) contain a single intron, and one gene (*ycf3*) contains two introns. The GC content in whole cp genome, LSC region, SSC region, and IR region was 36.75%, 34.83%, 28.19%, and 38.96%, respectively, like other Thymelaeaceae plants.

To investigate the phylogenetic position of *D. retusa*, an alignment of Thymelaeaceae chloroplast genome sequences was processed by MAFFT (Katoh and Standley 2013), and a neighbor-joining (NJ) phylogenetic tree (Figure 1) was completed with MAFFT (Katoh and Standley 2013). The results showed close phylogenetic relationships among species in genus *Daphne*, especially among species *Daphne kiusiana*, *Daphne tangutica*, and *Daphne giraldii*. The relationships within *Daphne* can be well distinguished using whole chloroplast genome data. *D. retusa* has a close relationship with congeneric *Daphne tangutica*. The complete chloroplast

genome sequence of *D. retusa* facilitates the phylogenetic studies of Thymelaeaceae.

Acknowledgements

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

Disclosure statement

The authors have no potential or actual conflicts of interest to report.

Funding

The authors acknowledge the financial assistance from the Long-yuan Youth Innovation and Entrepreneurship Talent Team Project of Gansu Province [2020RCXM130] and the Gansu Province Higher Education Industry Support and Guidance [2021C-07; 2019C-01].

Data availability statement

The genome sequence data that support the findings of this study are freely available in GenBank at (<https://www.ncbi.nlm.nih.gov/>) under accession number MW245832. The associated BioProject, SRA, and BioSample numbers are PRJNA693110, SRX9895515, and SAMN17377581, respectively.

References

- Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc.* 1(5):2320–2325.
- Geng L, Li ST, Yang XH, Chen MY, Ren GX, Li YF, Zhang X, Zhang ZY, Liu CS. 2016. [Resource situation investigation and analysis about endangered Chinese herbal medicine *Daphnes Cortex*]. *Zhongguo Zhong Yao Za Zhi.* 41(18):3323–3328.
- Jamila N, Shafiq AT, Naeem K, Sadia N, Muhammad RS, Syed Muhammad AH, Haroon K. 2018. Evaluation of *In-vitro* antimicrobial potential of *Daphne retusa* Hemsl. Against human pathogenic bacteria and fungi. *Curr Top Med Chem.* 18:779–786.
- Katoh 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.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33(7):1870–1874.
- Patel RK, Jain M. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One.* 7(2):e30619.