ABSTRACT

Rhododendron huadingense is an important horticultural plant that belongs to the Ericaceae family. In this study, the chloroplast genome sequence of R. huadingense is reported. The chloroplast genome of R. huadingense was 198,952 bp in length and had an angiosperm-typical quadruplicate structure with a large single-copy (LSC) region of 108,557 bp, a small single-copy (SSC) region of 53 bp, and two inverted repeat regions (IRs) of 45,171 bp. One hundred and thirteen unique genes including 79 protein-coding genes, 30 transfer RNA (tRNA) genes, and four ribosomal RNA (rRNA) genes were identified in the chloroplast genome. Further phylogenetic analysis revealed a close relationship between R. huadingense and R. molle. The complete chloroplast genome of R. huadingense provides valuable genetic information for the phylogeny, varieties breeding and sustainable utilization of this species.

ARTICLE HISTORY

Received 18 May 2022
Accepted 9 October 2022

KEYWORDS

Rhododendron huadingense; chloroplast genome; phylogenetic analysis; Ericaceae

Rhododendron huadingense B.Y. Ding & Y.Y. Fang 1990 was first discovered and named by Ding Bingyang etc in Huading Mountain, Tiantai County, Zhejiang Province, China in 1990 (Ding and Fang 1990). It is an endangered species and has been designated as a priority protected wild plant in Zhejiang Province due to its limited habitats and tiny population (Li et al. 2019). R. huadingense is unique in terms of hairiness, inflorescence type, fruit form, and seed size, all of which are important in phylogenetic investigations. Besides, it also has a high ornamental value due to its beautiful purple flowers (Ding et al. 1995). On the other hand, with the development of high throughput sequencing technology, it became much easier to sequence and obtain the chloroplast genomic sequences (Guo et al. 2018). To date, the chloroplast genomes of many species of Rhododendron have been reported, including R. delavayi (Liu et al. 2020), R. platypodium (Ma et al. 2021), and R. molle (Liu et al. 2021). In this study, the chloroplast genome sequencing of R. huadingense will be a great help to molecular identification as well as conservation strategy developing for this species.

Fresh and clean leaves of R. huadingense were collected from Taizhou, Zhejiang, China (29°15′27.49″N, 121°5′38.47″E, 850 m). A voucher specimen was deposited at the Herbarium of Hangzhou Botanical Garden with no. HZ041835 (contact person: Yahong Gao; email: 164314353@qq.com). According to the International Union for Conservation of Nature (IUCN) policy on endangered species research, the sample collection and the study were conducted with permission from Huading Forestry Farm, Tiantai, China. The total genomic DNA was extracted by using the modified CTAB protocol. DNA libraries were prepared with an insert size of 350 bp using NEBNext Ultra DNA Library Prep Kit. Pair-end sequencing (150 bp reads) was performed on the Illumina Hiseq 2500 Platform (Tianjin, China). In total, 46.1 GB clean data were generated and utilized for the chloroplast genome assembly using NOVOPlasty 4.3.1 (Dierckxsens et al. 2020). The entire genome sequence was annotated by CPGAVAS2 (Shi et al. 2020) and GeSeq (Tillich et al. 2017) with manual adjustments, using chloroplast genome of R. griersonianum (MT533181) as a reference. The genome and its annotated genes were submitted to NCBI under the accession number OM177184.

The chloroplast genome of R. huadingense was 198,952 bp in length and presented an intact circular structure with two inverted repeat regions (IRs), a small single-copy (SSC), and a large single-copy (LSC). The SSC and LSC regions were divided by two IR regions. The SSC was only composed of 53 base pairs, the LSC was 108,557 bp in length, and the IR region was 45,171 bp in length. The overall GC content of the chloroplast genome was 35.89%. The GC content of the two IR regions (36.63%) was higher than that of the LSC (35.29%) and SSC (5.66%). The relative high GC content was caused by rRNA and tRNA genes, which was also observed in other plants (He et al. 2016). The chloroplast genome encoded 113 unique genes, including 79 protein-coding genes, 30 tRNAs, and four rRNAs.

Microsatellite of the chloroplast genome was analyzed using the MicroSAtellite (MISA) program (Beier et al. 2017). The parameters were set as follows: 1–10, 2–6, 3–5, 4–4, 5–3, and 6–3; the minimum distance between two SSRs was set to 100 bp. In this study, we detected 75 SSRs, which included 67 mononucleotides (A/T), four dinucleotides (AT/AT), two...
pentanucleotides (AAAAT/ATTTT, ACATG/ATGTC), and two hexanucleotide (AAGGGT/ACCCTT, AATATT/AATATT). Mononucleotides are the most frequently repeated type, accounting for 89.3% of the total SSR markers.

According to the previous studies (Liu et al. 2021; Zhou et al. 2021), the chloroplast genomes of 15 related species were chosen for the phylogenetic analysis. Among these species, four Rhododendron species and three Vaccinium species of Ericaceae were used to interpret the phylogenetic relationships between R. huadingense and other species of Ericaceae, and the other eight species were used as outgroups. The phylogenetic analysis was carried out based on the coding sequences (CDS) of 59 shared genes from chloroplast genomes of R. huadingense and 15 related species. The chloroplast genome sequences of these 15 species were obtained from GenBank. All sequences were aligned using MAFFT v7.490 (Nakamura et al. 2018). Then, based on alignment results, the best tree model was determined by using jModeltest (Posada 2008). The maximum-likelihood (ML) analyses were conducted using RAxML 8.0 (Stamatakis 2006). For ML analyses, the best-fit model (general time reversible) with gamma distribution (GTR + G) was used with 1000 bootstrap replicates. According to the phylogenetic analysis, the Rhododendron plants formed a monophyletic clade with 100% bootstrap value, and R. huadingense had the closest relationship with R. molle (Figure 1).

**Ethical approval**

Research and collection of plant material was conducted with permission of related institution, and complied with national or international guidelines and legislation.

**Author contributions**

R A, M-Y N, and E-P L planned and designed the research. X-Z L, R A, M-Y N, and H-H H collected the plant materials. R A, M-Y N, and X-Z L performed experiments, and R A, E-P L, and H-H H analyzed the data. R A, H-H H, and E-P L wrote the manuscript.

**Disclosure statement**

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

**Funding**

The research was supported by Zhejiang Science and Technology Major Program on Agricultural New Variety Breeding [2021C02071-2].

**Data availability statement**

The 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 number OM177184. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA838786, SRR19256810, and SAMN28463168, respectively.

**References**

Beier S, Thiel T, Münch T, Scholz U, Mascher M. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics. 33(16):2583–2585.
Dierckxsens N, Mardulyn P, Smits G. 2020. Unraveling heteroplasmy patterns with NOVOPlasty. NAR Genom Bioinform. 2(1):lqz011.
Ding B, Fang Y. 1990. A new species of Rhododendron from Zhejiang, China. Bull Bot Res. 10(1):31–33.
Ding B, Wu H, Zhang H, Fang Y. 1995. Seed morphology of Rhododendron L. (Ericaceae) from Zhejiang and its taxonomic significance. Acta Bot Bor Occid Sin. 15(6):36–42.
Guo S, Guo L, Zhao W, Xu J, Li Y, Zhang X, Shen X, Wu M, Hou X. 2018. Complete chloroplast genome sequence and phylogenetic analysis of Paeonia ostii. Molecules. 23(2):246.
He Y, Xiao H, Deng C, Xiong L, Yang J, Peng C. 2016. The complete chloroplast genome sequences of the medicinal plant Pogostemon cablin. Int J Mol Sci. 17(6):820.
Li X, Shen D, Xu Q, Zhang J, He L, Li D, Yv M, Shen B. 2019. Effects on pollen viabilities of Rhododendron huadingense with different storage methods. J Zhejiang Sci Technol. 39(5):72–76.
Liu J, Chen T, Zhang Y, Li Y, Gong J, Yi Y. 2020. The complete chloroplast genome of *Rhododendron delavayi* (Ericaceae). Mitochondrial DNA Part B Resour. 5(1):37–38.

Liu Y, Li Q, Wang L, Wu L, Huang Y, Zhang J, Song Y, Liao J. 2021. The complete chloroplast genome of *Rhododendron molle* and its phylogenetic position within Ericaceae. Mitochondrial DNA B Resour. 6(9):2587–2588.

Ma L, Zhu H, Wang C, Li M, Wang H. 2021. The complete chloroplast genome of *Rhododendron platypodum* (Ericaceae): an endemic and endangered species from China. Mitochondrial DNA B Resour. 6(1):196–197.

Nakamura T, Yamada KD, Tomii K, Katoh K. 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. Bioinformatics. 34(14):2490–2492.

Posada D. 2008. jModelTest: phylogenetic model averaging. Mol Biol Evol. 25(7):1253–1256.

Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 47(W1):W65–W73.

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

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11.

Zhou X, Wei M, Zhang K, Han J, Wang H, Dong S. 2021. Characterization of the complete chloroplast genome of *Rhododendron henanense* subsp. *lingbaense* Fang. Mitochondrial DNA B Resour. 6(12):3325–3326.