Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

🔓 OPEN ACCESS Check for updates

The complete chloroplast genome of *rhododendron williamsianum* (ericaceae)

Daoyuan Zhou[a], Juan Wang[a], Fengjuan Zhou[a], Zhongxiang Li[b], Yangyang Qi[a] and Tingting Hu[a]

[a]Forestry College, Xinyang Agriculture and Forestry University, Xinyang, China; [b]Shangcai Forestry Development Service Center, Zhumadian, China

ABSTRACT

Rhododendron williamsianum Rehder & E. H. Wilson 1913, is a plant with important horticultural value. Here we report its chloroplast genome. The total length of the chloroplast genome was 205,424 bp, with a GC content of 35.8%. It consisted of a 107,968 bp large single copy, a 2606 bp small single copy, and a pair of 47,425 bp inverted repeats separating them. Within the chloroplast genome, there were a total of 110 unique genes, which included 76 protein-coding genes, 4 rRNA genes, and 30 tRNA genes. Our phylogenetic analyses indicated that *R. williamsianum* was closely genetically related to *R. sutchuenense* and *R. jingangshanicum*. The findings from this study not only contribute to the genetic database of *Rhododendron* plants but also have implications for evolutionary research within the family Ericaceae.

Background

Rhododendron Linnaeus 1753 is one of the ten most famous flowers in China, not only has high ornamental value, but also can be used for ecological protection and medicine and other research (Li et al. 2018; Liang et al. 2016). *Rhododendron* belongs to the family Ericaceae, with about 1025 species in the world, of which about 580 species are distributed in China, and most of them are endemic to China. Although *Rhododendron* is rich in germplasm resources, with the development of society, some species have been seriously disturbed and even extinct or endangered (Ma et al. 2017; Ma et al. 2014). Among the vast and diverse family of *Rhododendron*, *Rhododendron williamsianum* Rehder & E. H. Wilson 1913 stands out for its unique beauty and intriguing biological characteristics. Also known as the round-leaved azalea, this species is admired for its distinctive circular leaves and bell-shaped flowers, which add a vibrant splash of color to the landscapes of Southeast Tibet and adjacent areas, its native habitat. Beyond its esthetic value, *R. williamsianum* holds significant potential for scientific research, particularly in the investigation of its chloroplast genome. Chloroplasts, the site of photosynthesis in plants, contain their own DNA, which can offer valuable insights into the evolutionary history, genetic diversity, and adaptive mechanisms of plants. This article aims to shed light on the complete chloroplast genome of *R. williamsianum*, discussing its structure, gene content, and the implications of these findings for conservation, taxonomy, and evolutionary studies within the *Rhododendron*.

Material and methods

DNA extraction of sample, genome sequencing

R. williamsianum was gathered from the Emei Mount, Sichuan, China (Figure 1, 103.3456680°E, 29.5671674°N). The specimen was stored in the herbarium of the Xinyang Agriculture and Forestry University (voucher number: YYDJN02, under the name of Tingting Hu, 19003766172@163.com). The process of genomic DNA extraction was carried out using the CTAB (Doyle and Doyle, 1987). The library for next-generation sequencing, possessing a 150 bp insert size, was prepared and sequenced utilizing the Illumina HiSeq 2000 platform developed by Origingene Company in Shanghai.

Assembly and annotation

Approximately 4.5 Gb of raw data was obtained through this process. GetOrganelle v.1.7.7 (Jin et al. 2020) and Geneious Prime v.2022 (Kearse et al. 2012) was utilized for *De novo* genome assembly. CPGAVAS2 (Shi et al. 2019) was employed for genome annotation. Visualization of the genome map was carried out using CPGView (Liu et al. 2023).

Repeat sequence and codon preference analysis

simple sequence repeats (SSRs) were detected using MISA (Beier et al. 2017) with parameters set to 10, 6, 5, 5, 5, and 5. Long repeats were evaluated using REPuter (Kurtz et al.
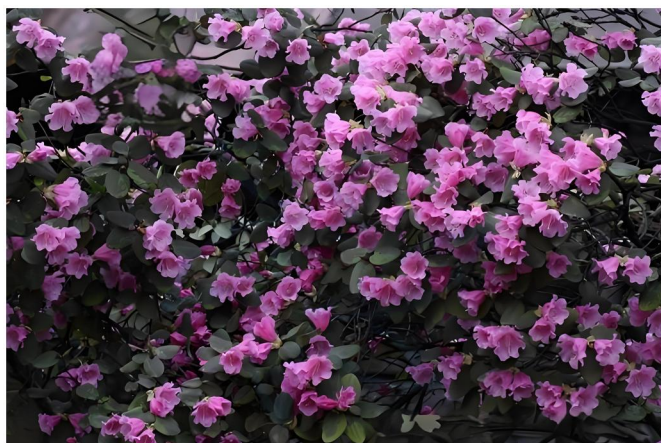
Figure 1. Species reference map of *R. williamsianum*. This picture was taken by Tingting Hu from the Emei Mount, Sichuan, China (voucher number: YYDJN02; 103.3456680°E, 29.5671674°N). Core features: Shrubs, 1-2 m tall; leaves leathery, broadly ovate or suborbicular, apex rounded, apiculate, base cordate or suborbicular, dark green above, glabrous, grayish-white below, finely veined in a raised reticulation; petiole 1-1.5 cm long, flat above, terete below, sparsely stipitate glandular. Racemose umbels with 2-6 flowers; pedicels 2-3 cm long, stout; calyx small, discoid, 6-lobed, lobes broadly triangular; corolla broadly campanulate, pink, lobes suborbicular; stamens 1.8-3 cm long, filaments glabrous, anthers ovate-orbicular, deep purplish-red; pistil subequal to the corolla; ovary ovate-orbicular, ca. 5 mm long; styles 3.5 cm long; stigmas dilated into heads. Capsule terete, 1.5-2.5 cm long, 6 mm in diam.

2001) with the following parameters: minimum length of 30 bp, hamming distance of 3. The comparison of IR boundaries was conducted using CPJSdraw (Li et al. 2023). Sequence hotspot analysis was conducted using mVISTA (Frazer et al. 2004). In addition, codon preference analysis was performed using PhyloSuite v. 1.2.3 (Zhang et al. 2020).

Phylogenetic analysis

We conducted a chloroplast phylogenetic analysis of the 29 *Rhododendron* species. These chloroplast protein-coding genes were extracted using PhyloSuite v. 1.2.3 (Zhang et al. 2020), then aligned using MAFFT v. 7.4 (Katoh and Standley, 2013). The chloroplast phylogenetic tree was generated using the maximum likelihood method implemented in IQ-TREE v. 2.1.2 (Nguyen et al. 2014) with 1000 ultrafast bootstrap replicates. To better analyze the phylogenetic position of *R. williamsianum*, we obtained the angiosperms353 gene set based on transcriptome and genome resequencing data from 232 *Rhododendron* species using Easy353 software (Zhang et al. 2022). These sequencing data were obtained from the NCBI SRA database.

Results

Genome feature analysis

We conducted an analysis on the depth of genome coverage and the structure of splicing genes. As shown in Figure S1 and Figure S2, the depth of the cp genome of *R. williamsianum* ranged from 196× to 15313×, with most of the loci concentrated at 4000× to 6000×, and the structure of these splicing genes was similar to that of most flowering plants. This circular quadripartite cp genome spans 205,424 bp

(Figure 2), comprising the LSC region (107,968 bp), the SSC region (2606 bp), and a pair of IR regions (47,425 bp). The overall GC content of this cp genome is 35.8%. Upon annotation, a total of 142 genes were identified, leading to 110 unique genes after eliminating 32 duplicate genes. These unique genes encompass 76 PCGs, 4 rRNA genes, and 30 tRNA genes (Table S1). Notably, of the 32 duplicated genes detected, there were 15 PCGs, 8 tRNAs, and 13 rRNAs. In addition, mVISTA analysis revealed that the five *Rhododendron* relatives had more regions of variation, also indicating greater chloroplast variation in *Rhododendron* species. However, compared with *R. delavayi* var. *delavayi*, the differences were relatively small, indicating that *R. williamsianum* is closely related to *R. delavayi* var. *delavayi* (Figure S3).

Repeat and IR boundaries analysis

A total of 79 simple sequence repeats (SSRs) were found in the cp genome of *R. williamsianum*. These SSRs were comprised of 73 mononucleotides, 5 dinucleotides and one trinucleotide. The predominant type of SSRs observed were mononucleotides, making up 92.41% of the total. The SSRs were most prevalent in the LSC region, where they were concentrated in the non-coding areas (Table S2). Furthermore, we detected 50 long repeats, encompassing 22 palindromic repeats and 28 forward repeats, which were primarily situated in the LSC and IR regions (Table S3). The IR boundary structures of the cp genomes of the five *Rhododendron* relatives differed significantly, but the IR/SSC boundaries differed less (Figure S4).

Relative synonymous codon usage analysis

By calculating the RSCU values of the protein-coding genes, we found that CUG encoded leucine had the lowest RSCU value of 0.34, while UUA encoded leucine had the highest RSCU value of 2.16 (Figure S5; Table S4). In addition, methionine (AUG) and tryptophan (UGG), both of which are encoded by single codons, had an RSCU value of 1, suggesting no preference.

Phylogenetic analysis

The chloroplast phylogenetic tree showed that our results were similar to those of previous studies, with most nodes having high bootstrap values (Zhou et al. 2023) and *R. williamsianum* was closely related to *R. delavayi* var. *delavayi* (Figure 3). In order to better analyze the phylogenetic position of *R. williamsianum*, we obtained and phylogenetically analyzed the angiosperms353 gene set based on the transcriptome and genome resequencing data of 232 *Rhododendron* species. The results showed that *R. williamsianum* was more closely related to *R. sutchuenense* and *R. jingangshanicum* (Figure S6).
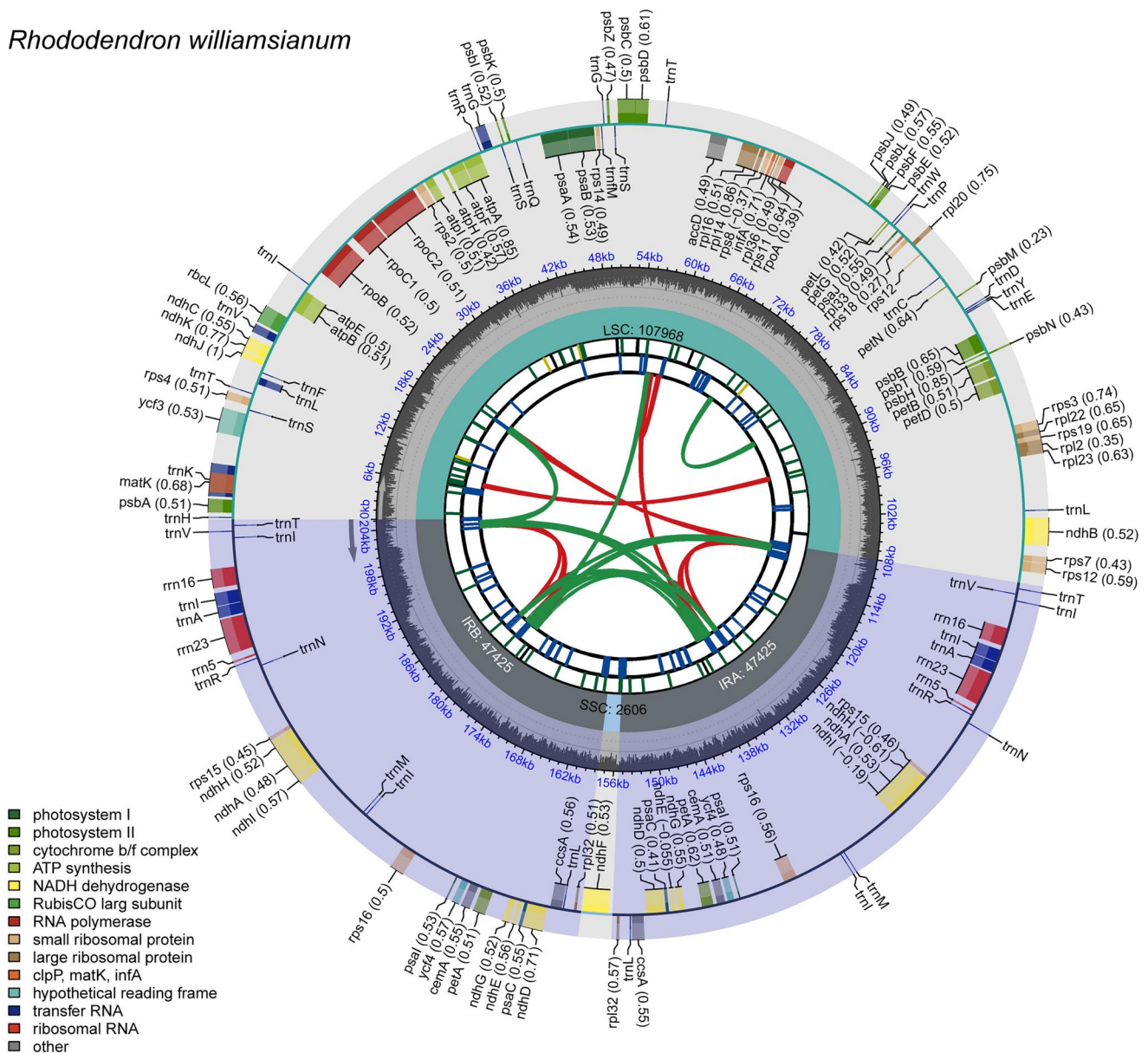
Rhododendron williamsianum

Figure 2. The chloroplast genome map of *R. williamsianum*. Genes on the inside of the circle are transcribed clockwise, while those on the outside are transcribed counter-clockwise. The gene name is followed by the optional codon usage bias in parentheses. The GC contents are displayed in the small grey bar graphs in the inner circle. Gene functional classification is color-coded and shown in the center.

Discussion

The length of *R. williamsianum* was 205,424 bp, and it encoded a total of 110 genes. It is similar to the published *Rhododendron* species (An et al. 2022; Li et al. 2020; Liu et al. 2019; Liu et al. 2021; Ma et al. 2021; Shen and Huang, 2024; Shen et al. 2022; Shen et al. 2019; Wang et al. 2023; Wang et al. 2021a; Wang et al. 2021b; Yu et al. 2022; Zhou et al. 2023; Zhou et al. 2021; Zhu et al. 2022). Our phylogenetic results support that *R. williamsianum* was closely related to *R. sutchuenense* and *R. jingangshanicum*. The chloroplast genomes of most *Rhododendron* species are still unknown, which has important implications for the phylogenetic position assessment of *Rhododendron* species, and there is a need to enrich the genome sequencing of *Rhododendron* species in the future. This research not only enriches the genetic data of *Rhododendron* but also forms a foundation for comprehending the evolutionary path of Ericaceae species.

Conclusion

The cp genome of *R. williamsianum* was assembled using short-read data in this study. It showed a typical tetrameric structure, similar to the cp genomes of other Ericaceae plants. The phylogenetic analyses indicated that *R. williamsianum* was closely genetically related to *R. sutchuenense* and *R. jingangshanicum*. Thus, the cp genome of *R. williamsianum* not only contributes to the genomic knowledge of *Rhododendron* but also provides insights into the evolution of the species.
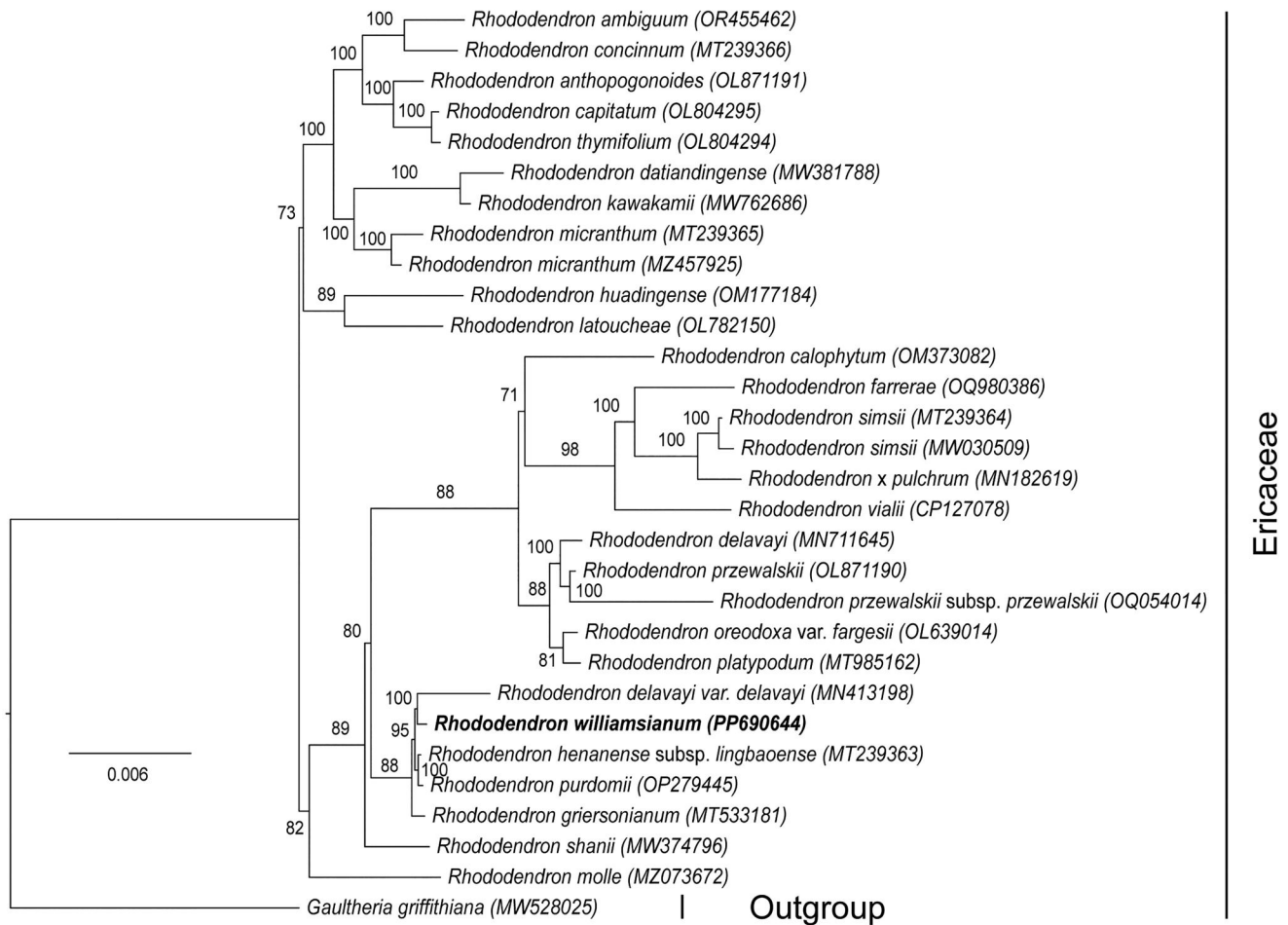
Figure 3. Phylogenetic tree based on the concatenated sequences of 74 protein-coding genes in 30 species by maximum-likelihood method. Values split by back-slashes above branches represent bootstraps. Branch supports were tested using ultrafast bootstrap with 1,000 replicates. *Rhododendron williamsianum* (PP690644) was marked in bold. The following sequences were used: *R. ambiguum* (OR455462), *R. concinnum* (MT239366) (Zhou et al. 2023), *R. anthopogonoides* (OL871191), *R. capitatum* (OL804295), *R. thymifolium* (OL804294), *R. datiandingense* (MW381788) (Wang et al. 2021b), *R. kawakamii* (MW762686) (Wang et al. 2021a), *R. micranthum* (MT239365), *R. micranthum* (MZ457925) (Zhou et al. 2023), *R. huadingense* (OM177184) (an et al. 2022), *R. latoucheae* (OL782150), *R. calophytum* (OM373082) (Zhou et al. 2023), *R. farrerae* (OQ980386) (Shen and Huang, 2024), *R. simsii* (MT239364) (Zhou et al. 2023), *R. simsii* (MW030509), *rhododendron x pulchrum* (MN182619) (Shen et al. 2022), *R. vialii* (CP127078), *R. delavayi* (MN711645) (Li et al. 2020), *R. przewalskii* (OL871190), *R. przewalskii* subsp. *przewalskii* (OQ054014) (Wang et al. 2023), *R. oreodoxa* var. *fargesii* (OL639014) (Zhu et al. 2022), *R. platypodum* (MT985162) (Ma et al. 2021), *R. delavayi* var. *delavayi* (MN413198) (Liu et al. 2019), *R. henanense* subsp. *lingbaoense* (MT239363) (Zhou et al. 2021), *R. purdomii* (OP279445) (Shen et al. 2019), *R. griersonianum* (MT533181), *R. shanii* (MW374796) (Yu et al. 2022), *R. molle* (MZ073672) (Liu et al. 2021), *gaultheria griffithiana* (MW528025).

Author contributions

DYZ and TTH conceived and designed the experiments; WJ and DYZ performed the experiments; DYZ, FGZ and ZXL analyzed the data and modified the article; DYZ, WJ and YYQ wrote the paper. All authors agree to be accountable for all aspects of the work.

Disclosure statement

No potential conflict of interest was reported by the authors.

Ethical approval

Leaf material of *R. williamsianum* was provided by the Xinyang Agriculture and Forestry University. Experimental researches do not involve the genetic transformation, preservation of the genetic background of the species used, and any other processes requiring ethical approval. Therefore, no special permission was required.

Data availability statement

The genome sequence data that support the findings of this study are available in GenBank of NCBI (http://www.ncbi.nlm.nih.gov/) under the accession no. PP690644. The associated BioProject, BioSample, and SRA numbers are PRJNA1108497, SAMN41244686, SRR28923684, respectively.

References

An R, Niu M, Lou X, Huang H, Lin E., 2022. The complete chloroplast genome of *Rhododendron huadingense* (Ericaceae). Mitochondrial DNA B Resour. 7(11):1910–1912. doi:10.1080/23802359.2022.2135403.

Beier S, Thiel T, Münch T, Scholz U, Mascher M., 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics. 33(16):2583–2585. doi:10.1093/bioinformatics/btx198.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissues. Phytochem Bull. 19:11–15.

Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32(Web Server issue):W273–279., doi:10.1093/nar/gkh458.

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):241. doi:10.1186/s13059-020-02154-5.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780. doi:10.1093/molbev/mst010.

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. doi:10.1093/bioinformatics/bts199.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R., 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 29(22):4633–4642. doi:10.1093/nar/29.22.4633.

Li H, Guo Q, Li Q, Yang L., 2020. Long-reads reveal that Rhododendron delavayi plastid genome contains extensive repeat sequences, and recombination exists among plastid genomes of photosynthetic Ericaceae. PeerJ. 8:e9048. doi:10.7717/peerj.9048.

Li H, Guo Q, Xu L, Gao H, Liu L, Zhou X., 2023. CPJSdraw: analysis and visualization of junction sites of chloroplast genomes. PeerJ. 11: e15326. doi:10.7717/peerj.15326.

Li T, Liu X, Li Z, Ma H, Wan Y, Liu X, Fu L., 2018. Study on reproductive biology of Rhododendron longipedicellatum: a newly discovered and special threatened plant surviving in limestone habitat in Southeast Yunnan, China. Front Plant Sci. 9:33. doi:10.3389/fpls.2018.00033.

Liang J-y, You C-x, Guo S-s, Zhang W-j, Li Y, Geng Z-f, Wang C-f, Du S-s, Deng Z-w, Zhang J, et al. 2016. Chemical constituents of the essential oil extracted from Rhododendron thymifolium and their insecticidal activities against Liposcelis bostrychophila or Tribolium castaneum. Ind Crops Prod. 79:267–273. doi:10.1016/j.indcrop.2015.11.002.

Liu J, Chen T, Zhang Y, Li Y, Gong J, Yi Y., 2019. The complete chloroplast genome of Rhododendron delavayi (Ericaceae). Mitochondrial DNA B Resour. 5(1):37–38. doi:10.1080/23802359.2019.1689860.

Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, Liu C., 2023. CPGView: a package for visualizing detailed chloroplast genome structures. Mol Ecol Resour. 23(3):694–704. doi:10.1111/1755-0998.13729.

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. doi:10.1080/23802359.2021.1959458.

Ma H, Li T, Liu X, et al. 2017. Research progress in conservation biology of Rhododendron. World Forestry Res. 30(4):13–17.

Ma L-H, Zhu H-X, Wang C-Y, Li M-Y, Wang H-Y., 2021. The complete chloroplast genome of Rhododendron platypodum (Ericaceae): an endemic and endangered species from China. Mitochondrial DNA B Resour. 6(1):196–197. doi:10.1080/23802359.2020.1860714.

Ma Y, Nielsen J, Chamberlain DF, Li X, Sun W., 2014. The conservation of Rhododendrons is of greater urgency than has been previously acknowledged in China. Biodivers Conserv. 23(12):3149–3154. doi:10.1007/s10531-014-0764-9.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ., 2014. IQ-TREE: a Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol Biol Evol. 32(1):268–274. doi:10.1093/molbev/msu300.

Shen J, Huang L. 2024. The complete chloroplast genome sequence of Rhododendron farrerae Tate ex Sweet (Ericaceae). Mitochondrial DNA B Resour. 9(1):1–4. doi:10.1080/23802359.2023.2294897.

Shen J, Li X, Li M, Cheng H, Huang X, Jin S., 2022. Characterization, comparative phylogenetic, and gene transfer analyses of organelle genomes of Rhododendron × pulchrum. Front Plant Sci. 13:969765. doi:10.3389/fpls.2022.969765.

Shen JS, Li XQ, Zhu XT, Huang XL, Jin SH., 2019. Complete chloroplast genome of Rhododendron pulchrum, an ornamental medicinal and food tree. Mitochondrial DNA B Resour. 4(2):3527–3528. doi:10.1080/23802359.2019.1676181.

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. doi:10.1093/nar/gkz345.

Wang T, Tang C, Chen J, Cao Z, He H, Li Y, Li X., 2023. Phylogenetic relationship and characterization of the complete chloroplast genome of Rhododendron przewalskii subsp. przewalskii Maximowicz 1877, an ornamental and medicinal plant in China. Mitochondrial DNA B Resour. 8(9):998–1002. doi:10.1080/23802359.2023.2259040.

Wang ZF, Chang LW, Cao HL. 2021a. The complete chloroplast genome of Rhododendron kawakamii (Ericaceae). Mitochondrial DNA B Resour. 6(9):2538–2540. doi:10.1080/23802359.2021.1959439.

Wang Z-F, Feng H-F, Li Y-Y, Wang H-F, Cao H-L., 2021b. The complete chloroplast genome of Rhododendron datiandingense (Ericaceae). Mitochondrial DNA B Resour. 6(6):1749–1751. doi:10.1080/23802359.2021.1931504.

Yu H-X, Xu W-G, Zhang J-L, Lv Y-Y, Chen P, Xu B-K, Qu C, Wang Z., 2022. The complete chloroplast genome of Rhododendron shanii W.P. Fang (Ericaceae), a endemic plant from the Southern Dabie Mountains of China. Mitochondrial DNA B Resour. 7(10):1787–1788. doi:10.1080/23802359.2022.2127336.

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT., 2020. PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 20(1):348–355. doi:10.1111/1755-0998.13096.

Zhang Z, Xie P, Guo Y, Zhou W, Liu E, Yu Y., 2022. Easy353: A Tool to Get Angiosperms353 Genes for Phylogenomic Research. Mol Biol Evol. 39(12):msac261. doi:10.1093/molbev/msac261.

Zhou X, Liu M, Song L. 2023. Structural characterization of four Rhododendron spp. chloroplast genomes and comparative analyses with other azaleas. Biocell. 47(3):657–668. doi:10.32604/biocell.2023.026781.

Zhou X-J, Wei M-J, Zhang K, Han J-W, Wang H-L, Dong S-W., 2021. Characterization of the complete chloroplast genome of Rhododendron henanense subsp. lingbaoense Fang. Mitochondrial DNA B Resour. 6(12):3325–3326. doi:10.1080/23802359.2021.1994898.

Zhu Q, Cao Y, Yao Y, Zhang Y., 2022. The complete chloroplast genome sequence of Rhododendron oreodoxa var. fargesii (Ericaceae), an ornamental plant. Mitochondrial DNA B Resour. 7(7):1234–1236. doi:10.1080/23802359.2022.2093667.