Comparative Analysis of Chloroplast Genomes of *Rhododendron capitatum*, *Rhododendron przewalskii* and Rhododendron Species (Postprint)

Authors: Jia Shouning, Zhang Yingmin, Zhao Guofu, Chen Wenjuan, Li Guodong

Date: 2022-07-05T00:00:00+00:00

Abstract

Rhododendron capitatum and R. przewalskii are wild flowers and medicinal plants with high ornamental value. To investigate the genetic structure and evolutionary characteristics of the chloroplast genomes of these two species, this study sequenced their complete chloroplast genomes using the Illumina HiSeq 4000 platform. Following assembly and annotation, comparative genomics and phylogenetic analyses were conducted in conjunction with seven previously published chloroplast genomes of Rhododendron species. The results demonstrated that: (1) The complete chloroplast genomes of R. capitatum and R. przewalskii exhibited a typical circular quadripartite structure, each comprising a large single-copy region (105,990 bp and 109,191 bp), a small single-copy region (2,617 bp and 2,606 bp), and a pair of inverted repeat regions (45,825 bp and 47,516 bp), with total lengths of 200,257 bp and 206,829 bp, respectively. (2) A total of 263 SSR loci were identified in the chloroplast genomes, with the majority showing a preference for A/T bases; codons also exhibited a preference for ending with A/U. (3) Structural variations, including gene loss and genome rearrangement, were prevalent across the complete chloroplast genomes of Rhododendron species. This study enriches the genomic resources for Rhododendron and provides a theoretical foundation for research on breeding, genetic evolution, and phylogeny of R. capitatum and R. przewalskii.

Full Text

Preamble

Comparative Analysis of Chloroplast Genomes of *Rhododendron capitatum*, *R. przewalskii*, and Related *Rhododendron* Species

JIA Shouning[1], ZHANG Yingmin[2], ZHAO Guofu[1], CHEN Wenjuan[1], LI Guodong[3,4*]

[1] Qinghai Hospital of Traditional Chinese Medicine, Xining 810000, China
[2] First College for Clinical Medicine, Yunnan University of Chinese Medicine, Kunming 650500, China
[3] College of Traditional Chinese Medicine, Yunnan University of Chinese Medicine, Kunming 650500, China
[4] Yunnan Key Laboratory of Dai and Yi Medicines, Yunnan University of Chinese Medicine, Kunming 650500, China

Abstract

Rhododendron capitatum and *R. przewalskii* are wild ornamental flowers and medicinal plants with high economic value. To investigate the genetic structure and evolutionary characteristics of their chloroplast genomes, this study sequenced the complete chloroplast genomes of both species using the Illumina HiSeq 4000 platform. Following assembly and annotation, comparative genomic and phylogenetic analyses were performed incorporating seven previously published chloroplast genomes from other *Rhododendron* species. The results revealed: (1) Both *R. capitatum* and *R. przewalskii* possess typical circular quadripartite chloroplast genomes, each comprising a large single-copy region (105,990 bp and 109,191 bp, respectively), a small single-copy region (2,617 bp and 2,606 bp), and a pair of inverted repeat regions (45,825 bp and 47,516 bp), with total lengths of 200,257 bp and 206,829 bp, respectively. (2) A total of 263 simple sequence repeat (SSR) loci were identified in the two chloroplast genomes, with most SSRs showing A/T base preference and codons preferentially ending with A/U. (3) Structural variations, including gene loss and genome rearrangement, were common across *Rhododendron* chloroplast genomes. This study enriches the genomic resources for *Rhododendron* and provides a theoretical foundation for breeding, genetic evolution, and phylogenetic research on *R. capitatum* and *R. przewalskii*.

Keywords: *Rhododendron capitatum*, *Rhododendron przewalskii*, *Rhododendron* L., chloroplast genome, sequence characteristics, structural mutation

Introduction

The genus *Rhododendron* L. (Ericaceae) represents the largest and most diverse genus in the heath family, comprising over 1,000 species worldwide, with approximately 571 species distributed primarily in southwestern and southern China. *Rhododendron* species are renowned internationally for their beautiful floral forms and vibrant colors, ranking among China' s traditional top ten ornamental flowers and earning titles such as "King of Woody Flowers" and "Beauty among Flowers." China possesses the richest wild rhododendron resources glob-

ally, providing favorable conditions for developing the rhododendron horticultural industry and cultivar research. However, research on introduction, domestication, breeding, and exploitation of *Rhododendron* resources lags significantly behind other horticulturally advanced countries and regions. This discrepancy arises from two main factors: first, *Rhododendron* exhibits rapid radiation divergence and frequent interspecific hybridization, creating numerous variants and hybrid types that complicate germplasm identification; second, rhododendron breeding in China started relatively late with a weak foundation, substantially constraining new cultivar development and resource mining.

Rhododendron capitatum and *R. przewalskii* are evergreen shrubs within *Rhododendron* that possess both high medicinal value and ornamental appeal. Both species are used medicinally for their leaves and flowers in Tibetan medicine to treat "Bacon disease" and "Long disease," representing characteristic ethnic medicinal materials in Tibetan regions. Additionally, their attractive tree forms, brilliant flower colors, extended flowering periods, and strong adaptability to cold, drought, and harsh conditions make them excellent parental materials for shrub ornamental breeding, offering promising prospects for horticultural application and industrialization. Current research on these two species has primarily focused on functional traits, chemical constituents, and ecological adaptability, while genetic background studies remain scarce. Clarifying their genetic background is crucial for understanding and effectively utilizing superior trait gene resources and for conducting germplasm innovation and cultivar improvement through molecular genetics-based breeding approaches.

Chloroplasts, derived from endosymbiotic cyanobacteria, are photosynthetic organelles in plant cells that play vital roles in plant growth, development, and long-term evolution. The chloroplast genome in higher plants typically exists as a double-stranded circular molecule of 120–160 kb, consisting of a large single-copy region (LSC, 81–90 kb), a small single-copy region (SSC, 18–20 kb), and a pair of identical inverted repeat regions (IRs, 20–30 kb). Most chloroplast genomes are maternally inherited and structurally conserved, containing abundant genetic information that serves as a routine source for plant genetic transformation, genetic diversity analysis, adaptive evolution, and molecular breeding research, providing a molecular basis for economic crop improvement, horticultural cultivar breeding, and conservation of endangered species. Previous studies have employed DNA barcoding, reduced-representation genome sequencing, and genome skimming to explore phylogenetic relationships among *Rhododendron* species, achieving basic resolution of their genetic affinities. However, most taxonomic studies have been limited to genus, subgenus, or section levels, with weak discriminatory power for species, subsections, or even sections. Chloroplast genome sequences, with their moderate evolutionary rate and length several hundred times greater than conventional barcode sequences, contain more genetic information and offer stronger resolution, enabling accurate species identification as a "super DNA barcode."

This study utilized *R. capitatum* and *R. przewalskii* as materials, comprehen-

sively analyzing their sequence and structural characteristics based on de novo sequencing, assembly, and annotation, and conducting comparative analyses with seven published *Rhododendron* chloroplast genomes. We addressed three scientific questions: (1) What are the sequence characteristics of the *R. capitatum* and *R. przewalskii* chloroplast genomes? (2) What are the evolutionary features of *Rhododendron* chloroplast genomes? (3) What are the variation patterns in *Rhododendron* chloroplast genomes and their underlying causes? This research provides genetic resources for studies on breeding, species identification, and resource exploitation of *R. capitatum*, *R. przewalskii*, and the broader *Rhododendron* genus.

Materials and Methods

Plant Materials

Rhododendron capitatum samples were collected on September 6, 2021, from Huzhu County, Haidong City, Qinghai Province (37°00 16.11 N, 102°15 90.11 E). *R. przewalskii* samples were collected on August 11, 2021, from Menyuan County, Haibei Tibetan Autonomous Prefecture, Qinghai Province (37°25 19.25 N, 101°80 12.39 E). Healthy young leaves were selected, dried in silica gel for DNA extraction, and plants with flowers and fruits were collected for subsequent morphological identification. All voucher specimens were identified by Associate Chief Pharmacist Zhao Guofu from Qinghai Hospital of Traditional Chinese Medicine and deposited in the hospital' s herbarium (collection numbers: *R. capitatum*: 632126LY0192; *R. przewalskii*: 632126LY0128).

DNA Extraction, Sequencing, and Assembly

Total genomic DNA was extracted using the Plant Genomic DNA Kit (Beijing Biotech). DNA purity was assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA), concentration was measured with a Quantus Fluorometer (Picogreen), and integrity was verified via agarose gel electrophoresis. DNA libraries were constructed and sequenced on the Illumina HiSeq 4000 platform. Raw reads were filtered using NGS QC Toolkit to remove low-quality regions, yielding clean reads for assembly.

Chloroplast Genome Assembly, Annotation, and Physical Mapping

Clean reads were assembled using GetOrganelle software with the chloroplast genome of *R. platypodum* (NC_{053746}) as a reference. The assembled chloroplast genomes were annotated using CPGAVAS2 and the results were corrected in Geneious Prime by referencing *R. platypodum* (NC_{053746}) and *R. concinnum* (MT239366). Physical maps were generated using OrganellarGenomeDRAW. The validated chloroplast genome sequences of *R. capitatum* and *R.*

przewalskii were submitted to GenBank.

Basic Feature Analysis

SSRs were detected in the chloroplast genomes of *R. capitatum* and *R. przewalskii* using the MISA web tool. Minimum repeat thresholds were set at 10, 5, 4, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide SSRs, respectively. For codon usage bias analysis, protein-coding genes from nine *Rhododendron* chloroplast genomes were screened, excluding duplicate genes and those shorter than 300 bp. The remaining sequences were analyzed for codon usage preferences using MEGA X software. Comparative analysis of nine *Rhododendron* chloroplast genomes was performed using mVISTA with the shuffle-LAGAN mode, using *R. platypodum* (NC_{053746}) as the reference.

Phylogenetic Analysis

To determine the phylogenetic positions of *R. capitatum* and *R. przewalskii* within *Rhododendron*, 17 chloroplast genome sequences were selected for phylogenetic tree construction. In addition to the two newly sequenced genomes, 15 sequences were downloaded from NCBI, including seven *Rhododendron* species, three *Gaultheria* species, three *Vaccinium* species, and two outgroup species (*Pyrola* and *Chimaphila*). The optimal nucleotide substitution model and parameters were determined using ModelFinder. Maximum likelihood (ML) and Bayesian inference (BI) analyses were conducted. The ML tree was constructed in RAxML v8.2.4 with the GTR+GAMMA model and 1,000 bootstrap replicates. The BI tree was generated in MrBayes v3.2.6 with the GTR+GAMMA model, 2,000,000 generations, sampling frequency of 1,000, and burn-in fraction of 0.25.

Results

Basic Features of Chloroplast Genomes

The chloroplast genomes of both *R. capitatum* and *R. przewalskii* are circular double-stranded molecules with classic quadripartite structure, consisting of a large single-copy region (LSC), a small single-copy region (SSC), and two inverted repeat regions (IR) [Figure 1: see original paper]. The *R. capitatum* chloroplast genome (GenBank accession: OL804295) spans 200,257 bp, with LSC, SSC, and IR regions of 105,990 bp, 2,617 bp, and 45,825 bp, respectively. The *R. przewalskii* chloroplast genome (GenBank accession: OL871190) spans 206,829 bp, with LSC, SSC, and IR regions of 109,191 bp, 2,606 bp, and 47,516 bp, respectively. The overall GC content is 35.8% in *R. capitatum* and 35.7% in *R. przewalskii*, with IR regions (36.6% and 36.5%) showing higher GC content than LSC (35.4% and 35.1%) and SSC (29.8% and 30.0%) regions. Across nine *Rhododendron* species, chloroplast genome length ranges from 193,798 bp

to 208,015 bp, with *R. henanense* subsp. *lingbaoense* being the largest and *R. delavayi* var. *delavayi* the smallest. LSC regions range from 105,990 bp to 110,593 bp, SSC regions from 26 bp to 2,621 bp, and IR regions from 40,583 bp to 47,516 bp. Total GC content varies modestly from 35.7% to 36.0% .

The *R. capitatum* chloroplast genome encodes 138 genes, including 86 protein-coding genes, 44 tRNA genes, and 8 rRNA genes, while *R. przewalskii* encodes 143 genes (89 protein-coding, 46 tRNA, and 8 rRNA). Among the nine published *Rhododendron* chloroplast genomes, total gene number ranges from 123 (*R. delavayi* var. *delavayi*) to 150 (*R. griersonianum*), protein-coding genes from 80 to 95, tRNA genes from 35 to 47, and rRNA genes consistently number 8 . These differences primarily result from gene loss events involving *ycf1*, *ycf15*, *trnR-UCU*, *trnM-CAU*, *trnH-GUG*, *accD*, and *infA* in some species .

Simple Sequence Repeat Analysis

MISA analysis identified 263 SSRs in the chloroplast genomes of *R. capitatum* and *R. przewalskii*, comprising six types (mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide SSRs). In *R. capitatum*, 130 SSRs were detected: 74 in the LSC region, 2 in the SSC region, and 54 distributed across the two IR regions. Mononucleotide SSRs were most abundant (86), followed by dinucleotide (19), trinucleotide (12), tetranucleotide (12), and hexanucleotide (1) types; no pentanucleotide SSRs were found. In *R. przewalskii*, 133 SSRs were identified: 87 in LSC, 3 in SSC, and 43 in IR regions, including 71 mononucleotide, 19 dinucleotide, 25 trinucleotide, 16 tetranucleotide, and one each of penta- and hexanucleotide SSRs .

Codon Usage Bias Analysis

For codon usage bias analysis across nine *Rhododendron* species, 436 protein-coding gene sequences were selected (47 from *R. capitatum*, 49 from *R. przewalskii*, 49 from *R. concinnum*, 44 from *R. delavayi* var. *delavayi*, 54 from *R. griersonianum*, 49 from *R. henanense* subsp. *lingbaoense*, 50 from *R. micranthum*, 49 from *R. molle*, and 45 from *R. platypodum*). Analysis revealed 30 high-frequency codons with RSCU > 1 (e.g., AGA, UAA, GCU), of which 29 end with A/U and only one ends with G/C, indicating a strong preference for A/U-ending codons in *Rhododendron* chloroplast genomes. Thirty-two codons showed RSCU < 1 (e.g., AGC, CGC, UGA), with 29 ending in G/C and three ending in A/U. In *R. capitatum*, *R. micranthum*, and *R. molle*, all codons except UGA (which is preferentially used with RSCU > 1) showed consistent preferences, demonstrating high conservation of codon usage bias across the nine species [Figure 2: see original paper].

Comparative Analysis of *Rhododendron* Chloroplast Genomes

Whole-genome alignment of nine *Rhododendron* chloroplast genomes using mVISTA revealed high overall conservation, with coding regions more con-

served than non-coding regions and IR regions more conserved than single-copy regions [Figure 3: see original paper]. Seven highly divergent intergenic spacers were identified: *matK-ycf3*, *trnM-CAU-rpoB*, *trnT-GGU-accD*, *rpoA-psbJ*, *trnM-CAU-rrn16*, *trnI-CAU-rps16*, and *rps16-psaI*.

Phylogenetic Analysis

Phylogenetic analysis based on 17 chloroplast genome sequences showed identical topologies between ML and BI methods, with strong support for all branches (BS $\geq$ 63%, PP = 1). The monophyly of *Rhododendron* was strongly supported (BS = 100%, PP = 1). *Gaultheria* and *Vaccinium* formed a sister clade (BS = 100%, PP = 1) that was sister to *Rhododendron* (BS = 100%, PP = 1). Within *Rhododendron*, nine species formed a monophyletic group. Five species from subgenus *Hymenanthes* (*R. przewalskii*, *R. platypodum*, *R. griersonianum*, *R. delavayi* var. *delavayi*, and *R. henanense* subsp. *lingbaoense*) clustered together (BS = 100%, PP = 1), while three species from subgenus *Rhododendron* (*R. concinnum*, *R. micranthum*, and *R. capitatum*) formed a group but were not monophyletic. *R. przewalskii* and *R. platypodum* were sister taxa with moderate support (BS = 67%, PP = 1), and *R. przewalskii* showed a relatively close relationship with *R. micranthum* (BS = 100%, PP = 1) [Figure 4: see original paper].

Discussion and Conclusion

Comparative analysis of the chloroplast genomes of *R. capitatum* and *R. przewalskii* with seven previously reported *Rhododendron* species revealed significant structural changes in chloroplast genome organization and gene order. Compared with most angiosperm chloroplast genomes, *Rhododendron* species exhibit a dramatically reduced SSC region and expanded LSC and IR regions, with the IR region nearly doubled in size. This IR expansion appears to be the primary cause of increased genome size in *Rhododendron*. Similar extensive structural rearrangements have been observed in other species, such as *Pelargonium* × *hortorum*, whose IR region extends to 75 kb with a total genome size of approximately 200 kb, and *Cryptomeria japonica*, which completely lacks IR regions and has a genome of only ~130 kb. Additionally, the GC content of *Rhododendron* chloroplast genomes (35.7%-36.0%) is lower than that of many angiosperms such as *Aconitum* (~38.0%) and Adoxaceae (~38.3%). This low GC content correlates with increased structural rearrangements, though the underlying mechanisms remain unclear. Some studies suggest that chloroplast genome rearrangements may be related to inheritance patterns, and the frequent natural hybridization in *Rhododendron*—where genetically divergent parents produce fertile offspring—may contribute to genomic recombination and structural variation.

Structural variations in *Rhododendron* chloroplast genomes have led to gene duplication and loss events. In most angiosperms, *ycf1* has two copies, with

a complete gene at the IRa-SSC boundary and a truncated pseudogene at the IRb-SSC boundary. However, eight of the nine *Rhododendron* species examined completely lack the *ycf1* gene, suggesting that gene loss or duplication patterns are associated with differential IR expansion or contraction. For tRNA genes, loss of one copy can be compensated by other copies without functional impairment, whereas protein-coding gene loss may require nuclear-encoded products to be imported into chloroplasts. For example, the absence of *infA* in *R. concinnum* and *R. delavayi* var. *delavayi* may reflect transfer of this gene to the nucleus, followed by transcription and protein import into chloroplasts. Conversely, genes such as *ndhF* show duplication when transferred from the SSC to IR region. The chloroplast NADH dehydrogenase complex is involved in photosynthetic electron transport and chloroplast respiration, with high mutation rates and sensitivity to environmental stress. Given that most *Rhododendron* species inhabit cold, hypoxic plateau regions, duplication of *ndhF* and related genes may be associated with environmental adaptation. Overall, chloroplast genome rearrangements commonly accompany gene loss or duplication, and the substantial interspecific variation in gene loss patterns indicates high mutability in *Rhododendron* chloroplast genomes.

Codon usage bias refers to the non-uniform usage of synonymous codons during DNA encoding, representing the outcome of long-term evolution and natural selection. This bias manifests across different species, genomes, and genes, and constitutes an important evolutionary feature with significant implications for gene expression and organismal evolution. Our bioinformatic analysis of codon usage frequencies in *Rhododendron* chloroplast genomes identified preferentially used codons, providing valuable references for improving exogenous gene expression efficiency, transgenic research, and breeding improvement.

Chloroplast SSRs offer abundant polymorphic loci with advantages of easy replication and high information content, making them widely applicable in genetic diversity assessment, molecular-assisted breeding, and phylogenetic studies. In *R. capitatum* and *R. przewalskii*, mononucleotide SSRs predominate, followed by di- and trinucleotide SSRs, with longer motifs being less common. Nearly all SSR repeat units consist primarily of A and T bases, indicating that poly-A/T motifs dominate chloroplast SSRs—a pattern also observed in other angiosperms. These SSRs provide candidate molecular markers for cultivar breeding and genetic diversity research in *Rhododendron*.

Rhododendron represents a taxonomically challenging group. Our phylogenetic tree based on chloroplast genomes showed good resolution, strongly supporting the monophyly of *Rhododendron* and clear separation of subgenus *Hymenanthes* from other subgenera. This demonstrates that chloroplast genomes are highly effective for resolving species identification and phylogenetic relationships within *Rhododendron*. However, given the vast number of *Rhododendron* species, more chloroplast genome sequences are needed to generate a more realistic evolutionary lineage for the genus.

References

BAI L, JIAO ML, ZANG HY, et al., 2019. Chemical composition of essential oils from four *Rhododendron* species and their repellent activity against three stored-product insects[J]. *Environ Sci Pollut Res*, 26(22): 23198-23205.

BAI XX, SHEN FJ, LI ZB, et al., 2019. DNA barcode identification of genus *Rhododendron* L.: reference library, molecular markers and cultivar wild relatives tracing[J]. *Tianjin Agric Sci*, 25(10): 1-7.

BEIER S, THIEL T, MÜNCH T, et al., 2017. MISA-web: a web server for microsatellite prediction[J]. *Bioinformatics*, 33(16): 2583-2585.

BRUDNO M, MALDE S, POLIAKOV A, et al., 2003. Glocal alignment: Finding rearrangements during alignment[J]. *Bioinformatics*, 19: i54-i62.

CAO WX, 2001. Ecological adaptive strategy of genus *Rhododendron* L. in eastern Qilian mountain alpine ecosystem[D]. Gansu: Northwest Normal University: 35-37.

CAO XJ, LIU JJ, YANG M, 2009. Photosynthetic characteristics and anatomical structure of five species of *Rhododendron* in the Taibai mountain[J]. *Acta Bot Boreal-Occident Sin*, 29(12): 2483-2491.

CHENG SY, 2017. Study on resources and classification of azalea cultivars in China[D]. Ganzhou: Gannan Normal University: 11-12.

CHUMLEY TW, PALMER JD, MOWER JP, et al., 2006. The complete chloroplast genome sequence of *Pelargonium* × *hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants[J]. *Mol Biol Evol*, 23(11): 2175-2190.

DANIELL H, LIN CS, YU M, et al., 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering[J]. *Genome Biol*, 17(1): 134.

DRUMMOND A, 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data[J]. *Bioinformatics*, 28(12): 1647-1649.

FAN WB, WU Y, YANG J, et al., 2018. Comparative chloroplast genomics of Dipsacales species: insights into sequence variation, adaptive evolution, and phylogenetic relationships[J]. *Front Plant Sci*, 9: 689.

FU CN, MO ZQ, YANG JB, et al., 2022. Testing genome skimming for species discrimination in the large and taxonomically difficult genus *Rhododendron*[J]. *Mol Ecol Resour*, 22(1): 404-414.

GENG XM, HUAN ZQ, SU JL, et al., 2021. Researches advances in germplasm innovation of *Rhododendrons*[J]. *Mol Plant Breed*, 19(2): 604-613.

GREINER S, LEHWARK P, BOCK R, 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of or-

ganellar genomes[J]. *Nucl Acid Res*, 47(W1): W59-W64.

GU CH, MA L, WU ZQ, et al., 2019. Comparative analyses of chloroplast genomes from 22 Lythraceae species: inferences for phylogenetic relationships and genome evolution within Myrtales[J]. *BMC Plant Biol*, 19(1): 1-19.

HE JL, ZHOU TY, SONG YK, et al., 2021. Variations in floral traits of *Rhododendron przewalskii* with slope aspect in the southeastern Tibetan Plateau[J]. *Chin J Appl Environ Biol*, 27(4): 860-868.

HIRAO T, WATANABE A, KURITA M, et al., 2008. Complete nucleotide sequence of the *Cryptomeria japonica* D. Don. chloroplast genome and comparative chloroplast genomics: diversified genomic structure of coniferous species[J]. *BMC Plant Biol*, 8(1): 1-20.

HUANG CL, YAO G, TIAN XL, et al., 2021. Phylogenomic analysis of *Rhododendron* species in Guizhou Baili Rhododendron reserve based on RAD sequencing[M]. *Sci Silv Sin*, 57(2): 72-81.

HUELSENBECK JP, RONQUIST F, 2001. MRBAYES: Bayesian inference of phylogenetic trees[J]. *Bioinformatics*, 17: 754-755.

JANSEN RK, RAUBESON LA, BOORE JL, et al., 2005. Methods for obtaining and analyzing whole chloroplast genome sequences[J]. *Meth Enzym*, 395: 348-384.

JIN JJ, YU WB, YANG JB, et al., 2020. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes[J]. *Genome Biol*, 21(1): 1-31.

KANE N, SVEINSSON S, DEMPEWOLF H, et al., 2012. Ultra-barcoding in cacao (*Theobroma* spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA[J]. *Am J Bot*, 99(2): 320-329.

KATOH K, STANDLEY DM, 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability[J]. *Mol Biol Evol*, 30(4): 772-780.

KUMAR S, STECHER G, LI M, et al., 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms[J]. *Mol Biol Evol*, 35(6): 1547.

LEE SB, KAITTANIS C, JANSEN RK, et al. 2006. The complete chloroplast genome sequence of *Gossypium hirsutum*: organization and phylogenetic relationships to other angiosperms[J]. *BMC Genomic*, 7(1): 1-12.

LI Y, LÜ GH, ZHANG XN, et al., 2017. Chloroplast genome structure and variation analysis of *Brasicaceae* species[J]. *Acta Bot Boreal-Occident Sin*, 37(6): 1090-1101.

LI Y, SYLVESTER SP, LI M, et al., 2019. The complete plastid genome of *Magnolia zenii* and genetic comparison to Magnoliaceae species[J]. *Molecules*, 24(2): 261.

LI YF, LI SM, JIN X, et al., 2019. Classification of 85 species of *Rhododendron* in China based on RAD high throughput sequencing[J]. *For Res*, 32(3): 1-8.

LI YG, 2011. Current situation and garden use of wild *Rhododendron* resources in Qinghai province[D]. Xi' an: Northwest A & F University: 25-26.

LIAO HB, HUANG GH, YU MH, et al. 2017. Five pairs of meroterpenoid enantiomers from *Rhododendron capitatum*[J]. *J Org Chem*, 82(3): 1632-1637.

LUO DS, 2018. The original interpretation of Jingzhu Materia Medica[M]. Chengdu: Sichuan Science and Technology Press: 485.

MILNE RI, DAVIES C, PRICKETT R, et al., 2010. Phylogeny of *Rhododendron* subgenus *Hymenanthes* based on chloroplast DNA markers: between-lineage hybridisation during adaptive radiation?[J] *Plant Syst Evol*, 285: 233-244.

MA YP, XIE WJ, SUN WB, et al., 2016. Strong reproductive isolation despite occasional hybridization between a widely distributed and a narrow endemic *Rhododendron* species[J]. *Sci Rep*, 6(1): 1-11.

MAYOR C, BRUDNO M, SCHWARTZ JR, et al., 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length[J]. *Bioinformatics*, 16(11): 1046-1047.

MENG J, LI XP, LI HT, et al., 2018. Comparative analysis of the complete chloroplast genomes of four *Aconitum* medicinal species[J]. *Molecules*, 23(5): 1015.

PATEL RK, JAIN M, 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data[J]. *PLoS ONE*, 7(2): e30619.

PENG L, YAMAMOTO H, SHIKANAI T, 2011. Structure and biogenesis of the chloroplast NAD (P) H dehydrogenase complex[J]. *Biochim Biophys Acta-Biomembr*, 1807(8): 945-953.

PING JY, FENG PP, LI JY, et al., 2021. Molecular evolution and SSRs analysis based on the chloroplast genome of *Callitropsis funebris*[J]. *Ecol Evol*, 11(9): 4786-4802.

Qinghai Provincial Drug Inspection Institute, Qinghai Tibetan Medicinal Institute, 1996. Chinese Tibetan Medicine[M]. Shanghai: Shanghai Science & Technology Press: 210-211.

ROZAS J, FERRER-MATA A, SÁNCHEZ-DELBARRIO JC, et al., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets[J]. *Mol Biol Evol*, 34(12): 3299-3302.

SHARP PM, COWE E, 1991. Synonymous codon usage in *Saccharomyces cerevisiae*[J]. *Yeast*, 7(7): 657-678.

SHI LC, CHEN HM, JIANG M, et al., 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer[J]. *Nucl Acids Res*, 47(W1): W65-W73.

SILVA SR, DIAZ YC, PENHA HA, et al., 2016. The chloroplast genome of *Utricularia reniformis* sheds light on the evolution of the *ndh* gene complex of terrestrial carnivorous plants from the Lentibulariaceae family[J]. *PLoS ONE*, 11: e0165176.

STAMATAKIS A, 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies[J]. *Bioinformatics*, 30: 1312-1313.

SUGIURA C, KOBAYASHI Y, AOKI S, et al., 2003. Complete chloroplast DNA sequence of the moss *Physcomitrella patens*: evidence for the loss and relocation of *rpoA* from the chloroplast to the nucleus[J]. *Nucl Acids Res*, 31(18): 5324-5331.

TANG HQ, TANG L, SHAO SC, et al., 2021. Chloroplast genomic diversity in *Bulbophyllum* section *Macrocaulia* (Bl.) Aver. (Orchidaceae, Epidendroideae, Malaxideae): insights into species divergence and adaptive evolution[J]. *Plant Divers*, 43(5): 350-361.

WANG L, DONG WP, ZHOU SL, 2012. Structural mutations and reorganizations in chloroplast genomes of flowering plants[J]. *Acta Bot Boreal-Occident Sin*, 32(6): 1282-1288.

WANG XQ, SONG WW, XIAO JJ, et al., 2021. Phylogeny of Myrtales and related groups based on chloroplast genome[J]. *Guihaia*, 41(1): 68-80.

WICKE S, MÜLLER KF, DEPAMPHILIS CW, et al., 2016. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants[J]. *Proc Natl Acad Sci*, 113(32): 8846-8851.

XIE DF, YU HX, PRICE M, et al., 2019. Phylogeny of Chinese *Allium* species in section *Daghestanica* and adaptive evolution of *Allium* (Amaryllidaceae, Allioideae) species revealed by the chloroplast complete genome[J]. *Front Plant Sci*, 10: 460.

YAN LJ, LIU J, MICHAE M, et al., 2015. DNA barcoding of *Rhododendron* (Ericaceae), the largest Chinese plant genus in biodiversity hotspots of the Himalaya-Hengduan Mountain[J]. *Mol Ecol Resour*, 15(4): 932-944.

YANG HB, FANG RZ, JIN CL, et al., 1999. *Flora of China* (Vol 57, 1)[M]. Beijing: Science Press: 13-213.

YANG KT, CHEN GP, XIAN JR, et al., 2021. Scaling relationship between leaf mass and leaf area: A case study using six alpine *Rhododendron* species in the Eastern Tibetan Plateau[J]. *Global Ecol Conserv*, 30: e01754.

YI X, GAO L, WANG B, et al., 2013. The complete chloroplast genome sequence of *Cephalotaxus oliveri* (Cephalotaxaceae): evolutionary comparison of *Cephalotaxus* chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms[J]. *Genome Biol Evol*, 5(4): 688-698.

ZHANG DG, CAO WX, PU XP, et al., 2003. Botanical morpha and ecological adaptation of *Rhododendron* in Eastern Qilian Mountains[J]. *Grassl Turf*, (1):

27-30.

ZHANG Q, LIU Y, SODMERGEN, 2003. Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species[J]. *Plant Cell Physiol*, 44(9): 941-951.

ZHANG YM, HAN LJ, YANG CW, et al., 2022. Comparative chloroplast genome analysis of medicinally important *Veratrum* (Melanthiaceae) in China: insights genomic characterization and phylogenetic relationships[J]. *Plant Divers*, 44: 70-82.

ZHUANG P, 2019. Progress on the fertility of *Rhododendron*[J]. *Biodivers Sci*, 27(3): 327-338.

Note: Figure translations are in progress. See original paper for figures.