Chloroplast Genome Characteristics and Phylogenetic Analysis of the Alpine Plant *Rhodiola tangutica* (Postprint)

Authors: Qu Rongju, Tsering Tashi, Mao Xuanrui, Liu Yuping, Su Xu, Hu Xiayu, Yang Ping, Li Xiaoli, Sun Chenglin, Jin Jiarui

Date: 2024-12-02T00:00:00+00:00

Abstract

To elucidate the structural characteristics, gene information, and phylogenetic relationships of the chloroplast genome of Rhodiola tangutica, an important medicinal plant in Northwestern China, this study performed genome sequencing using Illumina NovaSeq6000 and analyzed its structure, gene function, and phylogenetic relationships using various bioinformatics analysis software including GeSeq, PGA, NOVOPlasty, IRscope, and MISA. The results showed that: (1) The chloroplast genome of R. tangutica exhibited a typical quadripartite circular structure, consisting of an 82,121 bp LSC region, a 16,996 bp SSC region, and a 25,873 bp IR region, with a total length of 150,863 bp and an overall GC content of 37.8%; the IR region had the highest GC content (42.9%). It encoded 131 genes, including 85 protein-coding genes (PCGs), 38 tRNA genes, and 8 rRNA genes. (2) A total of 32,471 codons were detected, among which cysteine (Cys) codons were the least abundant (1.18%) and isoleucine (Ile) codons were the most abundant (8.24%), with RSCU values greater than 1 for 29 codon types. (3) Analysis of the IR region revealed that both rps19 and ndhF showed expansion into the IRB region. (4) Phylogenetic analysis indicated that R. tangutica is most closely related to Rhodiola quadrifida, with divergence time estimation revealing that the origin of Rhodiola species occurred at an average of 15.50 Mya (95% HPD: 6.0-21.0 Mya). This study clarified the chloroplast genome characteristics of R. tangutica, obtained a relatively reasonable phylogenetic relationship among Rhodiola species, and provides a theoretical basis for research on genetic diversity, discussion of adaptive evolution mechanisms, and germplasm resource conservation in the genus Rhodiola.

Full Text

Characteristics of the Chloroplast Genome and Phylogenetic Analysis of the Alpine Plant *Rhodiola tangutica* (Crassulaceae)

Qu Rongju[1], Cai Rangzhaxi[1], Mao Xuanrui[1], Liu Yuping[1,2,3], Su Xu[1,2,3*], Hu Xiayu[1], Yang Ping[1], Li Xiaoli[1], Sun Chenglin[1], Jin Jiarui[1]

[1]School of Life Sciences, Qinghai Normal University, Xining 810008, China
[2]Key Laboratory of Biodiversity Formation Mechanism and Comprehensive Utilization of the Qinghai-Tibet Plateau in Qinghai Province, Qinghai Normal University, Xining 810008, China
[3]Academy of Plateau Science and Sustainability, Qinghai Normal University, Xining 810016, China

Abstract

To elucidate the structural characteristics, gene content, and phylogenetic relationships of the chloroplast genome in *Rhodiola tangutica*, an important medicinal plant from northwestern China, we sequenced its genome using the Illumina NovaSeq6000 platform and conducted comprehensive bioinformatic analyses using GeSeq, PGA, NOVOPlasty, IRscope, and MISA. The results revealed: (1) The chloroplast genome of *R. tangutica* exhibits a typical quadripartite circular structure spanning 150,863 bp, comprising an 82,121 bp LSC region, a 16,996 bp SSC region, and 25,873 bp IR regions. The overall GC content is 37.8%, with the highest GC content (42.9%) observed in the IR regions. The genome encodes 131 genes, including 85 protein-coding genes (PCGs), 38 tRNAs, and 8 rRNAs. (2) A total of 32,471 codons were detected, with cysteine (Cys) showing the lowest usage (1.18%) and isoleucine (Ile) the highest (8.24%). Twenty-nine codons displayed RSCU values greater than 1. (3) IR boundary analysis revealed that both *rps19* and *ndhF* genes have expanded into the IRB region. (4) Phylogenetic analysis indicated that *R. tangutica* is most closely related to *R. quadrifida*, with divergence time estimates suggesting that *Rhodiola* species originated approximately 15.50 Mya (95% HPD: 6.0–21.0 Mya). This study clarifies the chloroplast genome characteristics of *R. tangutica* and establishes a robust phylogenetic framework for the genus, providing a theoretical foundation for future research on genetic diversity, adaptive evolution, and germplasm resource conservation in *Rhodiola*.

Keywords: *Rhodiola*; *Rhodiola tangutica*; chloroplast genome; phylogeny; codon usage bias

---

Chloroplasts are essential semi-autonomous organelles in green plants that perform photosynthesis and energy conversion (Corriveau et al., 1988; Sudan et al., 2022). They convert solar energy into chemical energy and play crucial

roles in nutrient accumulation, growth and development, and stress tolerance (Suzuki et al., 2012; Song et al., 2017). In angiosperms, chloroplast genomes typically range from 120–180 kb, exhibit predominantly maternal inheritance (though some species show biparental or paternal inheritance), maintain highly conserved structures, and evolve at a relatively slow rate (Yang et al., 2019; Zheng et al., 2022). These genomes feature a characteristic circular quadripartite structure consisting of a large single copy (LSC) region, a small single copy (SSC) region, and two inverted repeat regions (IRA and IRB). They generally encode 110–130 genes (Daniell et al., 2016; Raman et al., 2020; Xu et al., 2020) and have been widely applied in species classification and identification (Jansen et al., 2006; Liang et al., 2018; Jiang et al., 2020), phylogenetic reconstruction (Jansen et al., 2006; Zhang et al., 2021), population genetic structure analysis (Twyford et al., 2017; Yu et al., 2022), and speciation mechanism studies (Liu et al., 2023). For instance, Gao et al. (2023) analyzed the chloroplast genome of the endangered orchid *Paphiopedilum parishii* and found high intraspecific diversity in SSRs, long repeats, SNPs, InDels, and nucleotide sequences, with all six individuals forming a monophyletic group sister to *Paphiopedilum dianthum*. Jiang et al. (2023) compared chloroplast genomes across *Trichosanthes* species, revealing sequence rearrangements and two major clades, with *Trichosanthes* being most closely related to *Indofevillea*, suggesting that positive selection drove population differentiation.

Rhodiola tangutica is a perennial medicinal herb belonging to Crassulaceae, genus *Rhodiola*, primarily distributed in northwestern China, typically inhabiting alpine rock crevices and riparian zones at elevations of 2,000–4,700 m (Editorial Committee of Flora of China, Chinese Academy of Sciences, 1984). As a traditional Tibetan and Chinese medicinal material, *R. tangutica* is used to treat physical weakness, shortness of breath, fatigue, and lung heat cough, with both roots and stems possessing medicinal properties including fever reduction, qi replenishment, heart nourishment, lung clearing, and detumescence (He and Zhang, 2012). Modern pharmacological research has demonstrated anticancer, immunomodulatory, antimicrobial, anti-inflammatory, and anti-fatigue effects (He et al., 2013). Previous studies have investigated seed germination characteristics (He et al., 2013), chemical composition (Li et al., 2014), pharmacological effects (Wang et al., 1992; Jia and Han, 2005), and germplasm resources (Shi et al., 2021) of *R. tangutica*. For example, He et al. (2013) found that 0.3 $g \cdot L^{-1}$ urea treatment significantly improved seed germination rate, vigor, and index, while higher concentrations inhibited germination. Li et al. (2014) extracted chemical components using solvents of different polarities, identifying alkanes, alcohols, ketones, and esters, with salidroside content significantly higher than in *R. crenulata*. Shi et al. (2021) determined that leaf explants on MS medium supplemented with 6-BA (0.5 $mg \cdot L^{-1}$) and NAA (0.5 $mg \cdot L^{-1}$) were optimal for callus induction, providing foundational data for germplasm conservation and artificial cultivation. However, the chloroplast genome structure, characteristics, and phylogenetic relationships of *R. tangutica* remain unreported. Therefore, this study sequenced the chloroplast genome of *R. tangutica* using high-

throughput sequencing technology and employed NOVOPlasty, GeSeq, PGA, IRscope, and MISA to address: (1) structural characteristics of the chloroplast genome; (2) gene number, function, and distribution; (3) SSR loci number and characteristics; (4) gene selective pressure; (5) gene contraction and expansion; and (6) phylogenetic relationships and systematic position of *R. tangutica* within *Rhodiola*.

1. Materials and Methods

1.1 Experimental Materials

Fresh leaves of *R. tangutica* were collected from the southern bank of Hala Lake, Delingha City, Haixi Prefecture, Qinghai Province (38°11 44.8 N, 97°36 1.33 E, elevation 4,056 m). Leaves were placed in sample bags and dried in sealed containers with silica gel. Voucher specimens are deposited at the Qinghai-Tibet Plateau Biological Specimen Museum, Northwest Institute of Plateau Biology, Chinese Academy of Sciences (HNWP).

1.2.1 Genome DNA Extraction and Sequencing

Total genomic DNA was extracted using a modified CTAB method (Doyle and Doyle, 1987; Tai and Tanksley, 1990). DNA integrity, purity ($OD_{260}/_{280}$), and concentration were assessed via 1% agarose gel electrophoresis, Nanodrop, and Qubit 2.0 fluorometry. Qualified DNA samples were randomly fragmented using a Covaris ultrasonic disruptor, followed by end repair, A-tailing, adapter ligation, purification, and PCR amplification for library construction and quality control. The prepared libraries were sequenced on the Illumina NovaSeq6000 platform.

1.2.2 Chloroplast Genome Assembly and Annotation

Using the *R. rosea* chloroplast genome (GenBank: MH410216) as a reference, the *R. tangutica* chloroplast genome was assembled with NOVOPlasty (Dierckxsens et al., 2017; Ding et al., 2020) using default parameters. Gene annotation was performed using GeSeq (Tillich et al., 2017) and PGA (Qu et al., 2019) with default settings, followed by manual correction to remove errors and redundancies. The circular map was visualized using OGDRAW (Lohse et al., 2013). The complete sequence and annotation information were submitted to NCBI GenBank (OR120372).

1.2.3 Codon Usage Bias and SSR Analysis

Simple sequence repeats (SSRs) in the *R. tangutica* chloroplast genome were detected using the online tool MISA (Beier et al., 2017) with minimum repeat thresholds of 10 for mononucleotides, 5 for dinucleotides, 4 for trinucleotides, and 3 for tetra- to hexanucleotides. Codon usage bias was analyzed using CodonW (Sharp and Li, 1987; Shields and Sharp, 1987).

1.2.4 IR Boundary Contraction/Expansion and Selective Pressure Analysis

The chloroplast genome sequences of *R. tangutica* and six closely related species (*R. fastigiata*, *R. crenulata*, *R. sacra*, *R. kirilowii*, *R. quadrifida*, and *R. himalensis*) were compared to identify variable hotspots. IRscope (Amiryousefi et al., 2018) was used to visualize differences in LSC/IRB/SSC/IRA boundaries across the seven *Rhodiola* species. Using *R. sacra* as a reference, DnaSP v5 (Librado and Rozas, 2009) was employed to calculate Ka/Ks ratios for all protein-coding genes in *R. tangutica*, *R. quadrifida*, *R. kirilowii*, and *R. himalensis*.

1.2.5 Phylogenetic Analysis and Divergence Time Estimation

Chloroplast genome sequences of 23 *Rhodiola* species were downloaded from NCBI, with *Prunus humilis* (Rosaceae) as the outgroup. Multiple sequence alignment was performed using MAFFT (Katoh et al., 2002) with default settings. Aligned sequences were trimmed with Gblocks, and the optimal nucleotide substitution model was selected using ModelFinder. A maximum likelihood (ML) tree was constructed using IQ-TREE in the PhyloSuite package (Guindon et al., 2010; Nguyen et al., 2015) with 5,000 bootstrap replicates.

Fossil records for *Rhodiola* were obtained from literature (Zhang et al., 2014) and the TimeTree Database (http://www.timetree.org/). Divergence times between *R. quadrifida* and *R. kirilowii* (0.93–3.33 Mya) and between *R. hobsonii* and *R. macrocarpa* (6.32–20.23 Mya) were used as calibration points. Divergence times were estimated using MCMCTree in the PAML package, employing the approximate method, a molecular clock correlated with rate, and the JC69 substitution model. Two independent runs were performed to assess convergence (correlation = 1), confirming result consistency.

2. Results

2.1 Basic Characteristics of the Chloroplast Genome

High-throughput sequencing yielded 8,704,701 raw reads (2.6 Gb), which were filtered to 8,489,447 clean reads (2.55 Gb). The assembled chloroplast genome of *R. tangutica* is 150,863 bp in length, exhibiting a typical quadripartite circular structure composed of a pair of inverted repeats (IRA and IRB, each 25,873 bp), a large single copy region (LSC, 82,121 bp), and a small single copy region (SSC, 16,996 bp) [FIGURE:1, TABLE:1]. The overall GC content is 37.8%, with the IR regions showing the highest GC content (42.9%), followed by LSC (35.8%) and SSC (31.8%). The genome encodes 131 genes, including 8 rRNAs, 38 tRNAs, and 85 protein-coding genes (PCGs). Among these, 86 genes are located in the LSC region, 13 in the SSC region, and 18 in the IR regions. IR region genes consist of 4 rRNAs (*rrn23*, *rrn4.5*, *rrn16*, *rrn5*), 7 tRNAs (*trnI-GAU*, *trnL-CAA*, *trnI-CAU*, *trnR-ACG*, *trnV-GAC*, *trnA-UGC*, *trnN-UGG*), 5 PCGs (*rpl23*, *rps12*, *rpl2*, *ndhB*, *rps7*), and 2 genes of unknown function (*ycf2* and

ycf1) [TABLE:1, TABLE:2]. Additionally, *clpP1* and *rps12* contain two introns, while *petB*, *petD*, *atpF*, *ndhA*, *ndhB*, *rpoC1*, *rpl2*, *trnK-UUU*, *trnL-UAA*, *trnA-UGC*, *trnG-UCC*, *trnI-GAU*, and *trnV-UAC* each contain one intron, totaling 9 PCGs and 6 tRNAs with introns; all other genes lack introns .

2.2 Codon Usage Bias

Analysis of codon usage bias in the *R. tangutica* chloroplast genome revealed that 104 CDS sequences contain 32,471 codons representing 23 amino acids [FIGURE:2, TABLE:3]. Leucine (Leu) codons were most abundant (10.79%), followed by serine (Ser) (8.39%) and isoleucine (Ile) (8.24%), while cysteine (Cys) was the least frequent (1.18%) . Twenty-nine synonymous codons showed RSCU values greater than 1, indicating codon usage bias in the *R. tangutica* chloroplast genome. The codon AGA (encoding arginine) exhibited the strongest bias (RSCU = 1.803). Nearly all codons with RSCU $>$ 1 end with A or U (except UUG), and tryptophan (Trp) and methionine (Met), each encoded by a single codon, showed no usage bias [Figure 2: see original paper].

2.3 Simple Sequence Repeat (SSR) Analysis

A total of 38 SSRs were identified in the *R. tangutica* chloroplast genome, comprising two repeat unit types: mononucleotides (35) and dinucleotides (3). The specific counts were (A) (13), (T) (22), (AT) (1), and (TA) (2). SSR lengths ranged from 10–14 bp, with 24 SSRs located in the LSC region, 9 in the SSC region, and 1 each in IRA and IRB . Most SSR loci were found in intergenic spacer (IGS) regions, with only a few present in coding sequences (CDS) .

2.4 IR Region Contraction/Expansion and Selective Pressure Analysis

Comparative analysis of chloroplast genome boundaries (LSC, IRB, SSC, IRA) across seven *Rhodiola* species revealed similar gene composition, position, and length [Figure 3: see original paper]. The JLB (LSC/IRB) boundary gene *rps19* consistently expanded 109–110 bp into the IRB region. The JSB (IRB/SSC) boundary comprised *ycf1* and *ndhF* genes; except for *R. fastigiata*, *R. crenulata*, and *R. sacra* where *ycf1* contracted toward IRB, all other species showed *ycf1* expansion (1–27 bp) into SSC, while *ndhF* consistently expanded 15–61 bp into IRB. All seven species showed *ycf1* spanning the JSA (SSC/IRA) boundary with similar total lengths (16,993–17,059 bp), including 4,089–4,104 bp in SSC and 5,165–5,183 bp in IRA. The JLA (IRA/LSC) boundary was consistently located between *trnH* and *rps19*. Thus, no gene loss was observed across *Rhodiola* species, with highly conserved gene distribution, quantity, type, and expansion/contraction patterns, indicating relatively low structural variation.

The Ka/Ks ratio, representing the ratio of nonsynonymous to synonymous substitution rates, indicates selective pressure: Ka/Ks $>$ 1 suggests positive selection, while Ka/Ks $<$ 1 indicates purifying selection. Genes with Ka/Ks $>$ 0 are

shown in [Figure 4: see original paper], revealing variable nucleotide substitution rates among protein-coding genes. Among genes with Ka/Ks > 1, *psaB* in *R. tangutica* was under positive selection, whereas *psaB* in other species experienced purifying selection. The *accD* gene in *R. tangutica* and *R. quadrifida* showed purifying selection, while *accD* in the other two species was under positive selection.

2.5 Phylogenetic Analysis and Divergence Time Estimation

Using *Prunus humilis* as the outgroup, we constructed a maximum likelihood (ML) tree based on chloroplast genome sequences from 24 *Rhodiola* species. The results revealed a strongly supported monophyletic group (bootstrap support = 100%) comprising all *Rhodiola* species [Figure 5: see original paper]. This clade divided into two major branches: Clade1 included *R. dumulosa*, *R. rhodantha*, *R. wallichiana*, *R. stapfii*, *R. sexifolia*, *R. hobsonii*, *R. sinuata*, *R. humilis*, *R. prainii*, and *R. ovatisepala*; Clade2 comprised *R. macrocarpa*, *R. gelida*, *R. rosea*, *R. calliantha*, *R. bupleuroides*, *R. yunnanensis*, *R. fastigiata*, *R. crenulata*, *R. sacra*, *R. tangutica*, *R. quadrifida*, *R. kirilowii*, and *R. himalensis*. *Rhodiola tangutica* was positioned within Clade2, showing the closest relationship with *R. quadrifida* [Figure 5: see original paper].

Divergence time estimation based on the *Rhodiola* phylogeny [Figure 6: see original paper] indicated that the genus originated approximately 15.50 Mya (95% HPD: 6.0–21.0 Mya), with the split between Clade1 and Clade2 occurring in the mid-Miocene. Further diversification of major lineages occurred from 1.97 Mya (95% HPD: 0.77–2.63 Mya) to 10.64 Mya (95% HPD: 4.16–14.18 Mya).

3. Discussion and Conclusion

Chloroplasts are organelles with independent, complete genomes that serve as critical sites for photosynthesis in higher plants. Their genomes are relatively conserved, typically ranging from 120–180 kb and encoding 100–130 genes (Xu et al., 2020; Yu et al., 2022). Our results demonstrate that the *R. tangutica* chloroplast genome possesses a typical quadripartite circular structure of 150,863 bp, encoding 131 genes with an overall GC content of 37.8%. The IR region exhibits higher GC content than LSC and SSC regions. Genes can be categorized into photosynthesis-related, self-replication, other functional, and unknown-function genes, consistent with previously reported chloroplast genomes of other *Rhodiola* species (Zhao, 2020; Zhang et al., 2022; Zhao et al., 2022). For example, Zhao et al. (2022) assembled chloroplast genomes of six *Rhodiola* species, finding genome sizes of 150,771–151,891 bp, GC contents of 37.7%–37.8%, and similar lengths of SSC, LSC, and IR regions, confirming the high quality of our assembly and annotation and indicating high similarity among *Rhodiola* chloroplast genomes. We attribute the elevated GC content in IR regions to the presence of rRNA genes with high GC content (Zhang et al., 2012).

Codon usage bias, where certain codons are used more frequently than synony-

mous alternatives, significantly influences gene function and protein expression, providing valuable insights into gene evolution and species origin (Ye et al., 2018; Sudan et al., 2022). In *R. tangutica*, 29 codons showed RSCU > 1, displaying clear AU bias (except UUG, all ended with A/U). We speculate this may result from high A/U content and nucleotide mutation patterns (Du et al., 2020; Zhang et al., 2022). Simple sequence repeats are 1–6 bp repeats in higher eukaryote chloroplast genomes, offering advantages including high polymorphism, abundance, and low recombination rates (Yuan et al., 2021; Mao et al., 2022), making them valuable for species delimitation, genetic diversity assessment, and fingerprinting (Kane and Cronk, 2008; Asaf et al., 2017). The *R. tangutica* chloroplast genome SSRs consist solely of mononucleotides (35) and dinucleotides (3), all being AT/TA repeats, consistent with most reported angiosperms where chloroplast SSRs primarily comprise short poly-A and poly-T rather than C or G repeats (Pugh et al., 2004).

Previous studies have shown that length variation and structural changes in chloroplast genomes often result from expansion and contraction of IR, LSC, and SSC regions (Huang et al., 2014; Jiang et al., 2021). Our comparison of boundary regions across seven *Rhodiola* species revealed that the JLB boundary consistently resides within the *rps19* gene interval. Except for *R. fastigiata*, *R. crenulata*, and *R. sacra* where *ycf1* contracts toward IRB, the IRB/SSC boundary in other species lies within the *ycf1–ndhF* overlap region. JSA and JLA boundaries show similar gene types, lengths, and positions, particularly between *R. tangutica* and its closest relative *R. quadrifida*, which exhibit high consistency in gene size, cross-region distribution, and location. We conclude that *Rhodiola* chloroplast genomes are highly conserved, with boundary variations at JLB, JSB, JSA, and JLA likely representing the primary mechanism of IR expansion/contraction. To further assess evolutionary divergence, we conducted Ka/Ks analysis of protein-coding genes. Ka/Ks ratios evaluate gene differentiation rates, with values >1 indicating positive selection, 1 neutral selection, and <1 purifying selection. Our analysis revealed that *psaB* in *R. tangutica* experienced strong positive selection, while *accD* showed significant interspecific variation.

Chloroplast genome sequences have been widely applied in plant phylogenetic and relationship studies (Nock et al., 2011). *Rhodiola* species are traditional Tibetan-Chinese medicinal plants with complex morphological variation, controversial systematic positions, and challenging species delimitation (He and Zhang, 2012). To clarify phylogenetic relationships, we analyzed 24 *Rhodiola* species, revealing a strongly supported monophyletic group (bootstrap = 100%) divided into two major clades. *Rhodiola tangutica* and *R. quadrifida* formed sister taxa with close affinity. Our phylogenetic topology is consistent with Yu et al. (2023) using chloroplast genomes and Zhang et al. (2014) based on five DNA barcodes (*rbcL*, *matK*, *trnH-psbA*, *trnL-F*, and ITS), confirming the accuracy of our phylogeny. Additionally, Hu et al. (2004) using RAPD markers and Ni and Wang (2004) using UPGMA clustering both found close relationships between *R. crenulata* and *R. fastigiata*, which our results further support, validating our

phylogenetic reconstruction. Divergence time analysis suggests that rapid diversification of *Rhodiola* occurred approximately 15.50 Mya, with major lineages further diversifying to generate extant species. Due to incomplete fossil records, future studies should incorporate more comprehensive data to obtain reliable calibration points (Zhang et al., 2014). Our results provide not only a robust phylogenetic framework for *Rhodiola* but also a theoretical basis for future research on genetic diversity, adaptive evolution, and germplasm conservation.

References

Amiryousefi A, Hyvönen J, Poczai P. 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics*, 34:

Asaf S, Waqas M, Khan AL, et al. 2017. The complete chloroplast genome of wild rice (*Oryza minuta*) and its comparison to related species. *Frontiers in Plant Science*, 8: 304.

Beier S, Thiel T, Münch T, et al. 2017. MISA-web: a web server for microsatellite prediction. *Bioinformatics*, 33: 2583-2585.

Corriveau JL, Coleman AW. 1988. Rapid screening method to detect potential biparental inheritance of plastid DNA and results for over 200 angiosperm species. *American Journal of Botany*, 75(10): 1443-1458.

Daniell H, Lin CS, Yu M, et al. 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology*, 17(1): 134.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19: 11-15.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, 45: e18.

Ding BC, Sun Y, Rong FX, et al. 2020. The complete mitochondrial genome of *Holothuria spinifera* (Théel, 1866). *Mitochondrial DNA Part B*, 5:

Du XY, Zeng T, Feng Q, et al. 2020. The complete chloroplast genome sequence of yellow mustard (*Sinapis alba* L.) and its phylogenetic relationship to other Brassicaceae species. *Gene*, 731: 144340.

Editorial Committee of Flora of China, Chinese Academy of Sciences. 1984. *Flora Reipublicae Popularis Sinicae*. Beijing: Science Press.

Gao XZ, Tang L, Wang Y, et al. 2024. Intraspecific genetic variation within chloroplast genome of a rare and endangered species *Paphiopedilum parishii* (Orchidaceae). *Guihaia*, 44(1): 1-14.

Guindon S, Dufayard JF, Lefort V, et al. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59: 307-321.

He MY, Zhang XL. 2012. Study on the optimized extraction technology of *Rhodiola algida*. *Chinese Journal of Modern Applied Pharmacy*, 29(6): 495-497.

He SL, Ma LF, Yang JJ, et al. 2013. Effects of different concentrations of urea solution on seed germination of *Rhodiola algida* var. *tangutica*. *Northern Horticulture*, 3: 173-175.

Huang H, Shi C, Liu Y, et al. 2014. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships. *BMC Evolutionary Biology*, 14: 151.

Hu TS, Lu YM, Ma LQ, et al. 2004. Identification of four *Rhodiola* L. plants by RAPD. *Chinese Traditional and Herbal Drugs*, 11: 90-93.

Jansen RK, Kaittanis C, Sasaki C, et al. 2006. Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. *BMC Evolutionary Biology*, 6: 32.

Jiang WJ, Guo MY, Pang XH, et al. 2020. Application of chloroplast genome in identification and phylogenetic analysis of medicinal plants. *World Chinese Medicine*, 15(5): 702-708.

Jiang ZZ, Hu SB, Yang HD, et al. 2023. Comparative evolution of *Trichosanthes* based on chloroplast genomes. *Journal of Anqing Normal University (Natural Science Edition)*, 29(2): 87-95.

Jia JC, Han T. 2005. Mechanism of *Rhodiola rosea*. *Western Journal of Traditional Chinese Medicine*, 11: 47-49.

Jiang M, Wang JF, Zhu Y, et al. 2021. Structure, sequence characteristics, and phylogenetic evolution analysis of *Belamcanda chinensis* chloroplast genome. *Chinese Traditional and Herbal Drugs*, 52(13): 4039-4046.

Kane NC, Cronk Q. 2008. Botany without borders: barcoding in focus. *Molecular Ecology*, 17(24): 5175-5176.

Katoh K, Misawa K, Kuma K, et al. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30: 3059-3066.

Liang FP, Wen XN, Gao HY, et al. 2018. Analysis of chloroplast genome features of Asteraceae species. *Genomics and Applied Biology*, 37(12): 5437-5447.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25: 1451-1452.

Liu C, Lü YQ, Shi ZL, et al. 2023. Characteristics of chloroplast genomes and phylogenetic analysis of *Artemisia* species. *Journal of Northwest Forestry University*, 38(3): 78-86.

Li Y, Wang XF, Li J, et al. 2014. Extraction of chemical components from *Rhodiola tangutica* by different solvents and its salidroside content determination.

Journal of Gansu Agricultural University, 49(5): 63-69.

Lohse M, Drechsel O, Kahlaou S, et al. 2013. OrganellarGenomeDRAW: a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research*, 41: W575-W581.

Mao XR, Liu YP, Su X, et al. 2022. Characteristics analysis of simple sequence repeat (SSR) loci in *Psammochloa villosa* (Poaceae) based on transcriptome data. *Acta Agrestia Sinica*, 30(8): 1990-2001.

Nguyen LT, Schmidt HA, von Haeseler A, et al. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32: 268-274.

Nock CJ, Waters DLE, Edwards MA, et al. 2011. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal*, 9(3):

Ni NC, Wang WH. 2004. Molecular identification of common medicinal *Rhodiola rosea* in Yunnan. *Journal of Shanxi University of Chinese Medicine*, 5(3): 48-52.

Powell W, Morgante M, McDevitt R, et al. 1995. Polymorphic simple sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. *Proceedings of the National Academy of Sciences of the United States of America*, 92: 7759-7763.

Pugh T, Fouet O, Risterucci AM, et al. 2004. A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. *Theoretical and Applied Genetics*, 108: 1151-1161.

Qu XJ, Moore M, Li DZ, et al. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods*, 15: 1-12.

Raman G, Park KT, Kim JH, et al. 2020. Characteristics of the completed chloroplast genome sequence of *Xanthium spinosum*: comparative analyses, identification of mutational hotspots and phylogenetic implications. *BMC Genomics*, 22(2): 855.

Shi GM, Ju XT, Yang LN, et al. 2021. Callus induction and plant regeneration of *Rhodiola tangutica*. *Molecular Plant Breeding*, 21(16): 5411-5418.

Sharp PM, Li WH. 1987. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Research*, 15: 1281-1295.

Shields DC, Sharp PM. 1987. Synonymous codon usage in *Bacillus subtilis* reflects both translational selection and mutational biases. *Nucleic Acids Research*, 15: 8023-8040.

Su DD, Liu YP, Liu T, et al. 2022. Structure of chloroplast genome and its characteristics of *Sphaerophysa salsula*. *Bulletin of Botanical Research*, 42(3): 446-454.

Suzuki N, Koussevitzky S, Mittler R, et al. 2012. ROS and redox signaling in the response of plants to abiotic stress. *Plant, Cell & Environment*, 35(2): 259-270.

Song J, Long YH, Lin LM, et al. 2017. Analysis on structure and phylogeny of chloroplast genomes in Araliaceae species. *Chinese Traditional and Herbal Drugs*, 48(24): 5070-5075.

Tai TH, Tanksley SD. 1990. A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. *Plant Molecular Biology Reporter*, 8(4): 297-303.

Tillich M, Lehwark P, Pellizzer T, et al. 2017. GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Research*, W1(3): W6-11.

Twyford AD, Ness RW. 2017. Strategies for complete plastid genome sequencing. *Molecular Ecology Resources*, 17(5): 858-868.

Wang XQ, Li JB, Zhang HY, et al. 1992. Experimental study on the anti-tumor effect of rhodosin. *Jilin Journal of Chinese Medicine*, 1992(3): 40.

Xu WB, Xia BS, Li XW. 2020. The complete chloroplast genome sequences of five pinnate-leaved *Primula* species and phylogenetic analyses. *Scientific Reports*, 10:

Yu XL, Wei P, Chen ZYF, et al. 2023. Comparative analysis of the organelle genomes of three *Rhodiola* species provide insights into their structural dynamics and sequence divergences. *BMC Plant Biology*, 23: 156.

Yu T, Pu F, Guan Q, et al. 2022. Characteristics of chloroplast genome and phylogenetic analysis of *Euphorbia peplus*. *Molecular Plant Breeding*, 20(6): 1828-1837.

Ye YJ, Ni ZX, Bai TD, et al. 2018. The analysis of chloroplast genome codon usage bias in *Pinus massoniana*. *Genomics and Applied Biology*, 37(10): 4464-4471.

Yang YM, Jiao J, Fan XC, et al. 2019. Complete chloroplast genome sequence and characteristics analysis of *Vitis ficifolia*. *Acta Horticulturae Sinica*, 46(4): 635-648.

Yuan JB, Zhang XJ, Wang M, et al. 2021. Simple sequence repeats drive genome plasticity and promote adaptive evolution in penaeid shrimp. *Communications Biology*, 4: 186.

Zhao KH, Li LQ, Quan H, et al. 2022. Comparative analyses of chloroplast genomes from six *Rhodiola* species: variable DNA markers identification and phylogenetic relationships within the genus. *BMC Genomics*, 23: 577.

Zheng CY, Su X, Liu YP, et al. 2022. Characteristics of complete chloroplast genome and phylogenetic analysis of *Cirsium souliei* (Asteraceae). *Acta Pharmaceutica Sinica*, 57(7): 2224-2233.

Zhang JQ, Meng SY, Allen GA, et al. 2014. Rapid radiation and dispersal out of the Qinghai-Tibetan Plateau of an alpine plant lineage *Rhodiola* (Crassulaceae). *Molecular Phylogenetics and Evolution*, 77: 147-158.

Zhang TW, Fang YJ, Wang XM, et al. 2012. The complete chloroplast and mitochondrial genome sequences of *Boea hygrometrica*: insights into the evolution of plant organellar genomes. *PLoS One*, 7(1): e30531.

Zhang Y, Su X, Liu YP, et al. 2022. Characteristics of complete chloroplast genome and phylogenetic analysis of *Rhodiola himalensis* (Crassulaceae). *Bulletin of Botanical Research*, 42(4): 602-612.

Zhao DN. 2020. Phylogeny of *Rhodiola* based on plastomes. Xi'an: Shaanxi Normal University.

Zhang DX, Kong LL. 2023. Genome sequencing and characteristic analysis of *Cercidiphyllum japonicum*. *Molecular Plant Breeding*, 21(17): 5654-5662.

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