The complete chloroplast genome of Meconopsis simplicifolia and its genetic comparison to other Meconopsis species

Yixi Yang (yangyixi1011@cdu.edu.cn)
Chengdu University
Rui Li
Chengdu University
Zhidan Zhu
Chengdu University
Peng Huang
Tibet Rhodiola Pharmaceutical Holding Company

Research Article

Keywords: Chloroplast genome, Papaveraceae, Meconopsis simplicifolia, Phylogenetic analysis, Comparative genomics

Posted Date: August 23rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3242403/v1

Abstract

M. simplicifolia, an endangered Chinese herb, possesses medicinal properties used in the treatment of various disorders. Despite its importance, there is a lack of genomic information available for *M. simplicifolia*, hindering our understanding of its molecular biology. Therefore, this study aimed to assemble and compare the chloroplast (cp) genome of *M. simplicifolia* with other reported *Meconopsis* cp genomes. The assembled *M. simplicifolia*plastome spanned 152,772 bp and exhibited the typical quadripartite structure comprising large (LSC, 83,824 bp) and small (SSC, 17,646 bp) single-copy regions, separated by a pair of inverted repeats (IRs, 25,651 bp). Overall, 131 genes were predicted, including 84 protein-coding genes, 37 tRNAs, and 8 rRNAs. Additionally, 33 SSRs and 27 long repeat sequences were identified. The seven *Meconopsis* species shared conserved genomic features in terms of gene structure and gene order. However, a comparison of the IR boundaries of the seven *Meconopsis* cp genomes revealed minor differences at the IR/SC boundary regions. Interestingly, *M. simplicifolia* exhibited the loss of the *rpl2* gene in the IRb region, which could be valuable for future plant classification research. A phylogenetic analysis including *M. simplicifolia* and eight other *Meconopsis* species indicated that *M. simplicifolia* clustered together with *M. betonicifolia*. The availability of the cp genome sequence of*M. simplicifolia* is of utmost importance for understanding phylogenetic and evolutionary aspects at or above the *Meconopsis*level. Comparative analysis of the cp genomes of *Meconopsis* species will facilitate species identification and enable selective breeding studies of these medicinal plants.

Introduction

Meconopsis is a herb with ornamental and medicinal properties (Grierson and Long 1983), existing as both annual and perennial species. Globally, there are 49 known *Meconopsis* species, several of which have been reported to possess therapeutic efficacy (Shi et al. 2022). In China, *Meconopsis* plants hold significant importance as herbal medicines and are utilized in Tibetan pharmacopeias such as “The Four Medical Tantras”, “Yue Wang Yao Zhen”, and “Jing Zhu Materia Medica” for the treatment of various ailments (Liu et al. 2016). *Meconopsis simplicifolia* (D. Don) Walp. 1842 is a species that thrives in high-altitude habitats across the Himalayas and adjacent plateau and mountain areas. It typically reaches heights of 30–70 cm and features narrow oblong seed capsules, bluish flowers, and hairy stems and leaves (Wangchuk et al. 2013). In Tibetan folk medicine, the entire plant of *M. simplicifolia* is used in the form of decoctions and poultices to address inflammatory issues, liver and lung infections, blood infections, as well as for its anticancer and antibacterial properties (Guo et al. 2016). *M. simplicifolia* is also highly valued for its ornamental value due to its attractive foliage and colorful flowers, earning it names such as “fairy grass” and “Himalayan poppy” (Sulaiman et al. 1991). It is considered an endangered species in China.

Traditional Chinese Medicine (TCM) has documented the use of 18 *Meconopsis* species in the formulation of more than 47 distinct formulas (Wangchuk et al. 2023). However, TCM prescriptions rarely rely on individual plants, and the identification of closely related *Meconopsis* species based on morphological characteristics alone poses challenges. Furthermore, significant differences in chemical composition exist among the species, increasing the risk to patient’s health and safety when raw materials are misidentified during the application of medicine (Tétényi 2003; Wei et al. 2019). For instance, *M. simplicifolia* bears morphological resemblance to *M. horridula*, and their distribution ranges overlap (Grierson and Long 1983). Hence, there is a need to integrate morphological classification with molecular analyses to understand the relationships within the genus *Meconopsis.*

Chloroplasts are essential organelles involved in photosynthesis, providing vital energy for plant growth and development (Daniell et al. 2016). Plant chloroplast genomes have been extensively utilized for studying genetic diversity, molecular identification, and taxonomic characterization (Bi et al. 2018; Hu et al. 2016; Wang et al. 2021). The majority of cp genomes exhibit a characteristic quadripartite structure, consisting of a small single copy region (SSC), a large single copy region (LSC), and a pair of inverted repeats (IRs) (Kwon et al. 2020). Chloroplast DNA (cpDNA) is

abundant and highly conserved among green plants, making it suitable for genome-wide evolutionary studies (Wang et al. 2023). Previous research has indicated that genes associated with photosynthesis may undergo adaptive evolution during species diversification (Hao et al. 2010; Kapralov and Filatov 2007; Zhang et al. 2018b). Molecular markers based on cp genomes, such as *matK*, *rpoC2*, *petA*, *ndhF*, and *ycf1* genes, have proven useful for identifying *Meconopsis* species and elucidating their evolutionary relationships (Li et al. 2019). Thus, obtaining a draft assembly of the cp genome for *M. simplicifolia* holds significant implications for the study of molecular identification, phylogenetic relationships, genetic structure, and functional variations among *Meconopsis* species.

The main objective of this study was to characterize the assembly, annotation, and analysis of the complete cp genome of *M. simplicifolia*. A secondary objective was to investigate the detailed cp genomic feature of *M. simplicifolia* and its phylogenetic relationships with the cp genomes of other *Meconopsis* species. The characterization of highly variable regions will aid in the development of potential DNA barcodes for future research. The results contribute to advancing our understanding of the fundamental characteristics of *Meconopsis* species and their evolutionary patterns within the Papaveraceae family.

Materials and methods

Plant material

M. simplicifolia plants (Fig. 1) were collected from the wild in Lingzhi, Tibet, China (38.9784°N, 105.9035°E) in July 2022. A voucher specimen (voucher number: NMU00912) was deposited at the Herbarium of North Minzu University (contact person: Lei Zhang, zhangsanshi-0319@163.com). The collection of plant material adhered to relevant institutional, national, and international guidelines and legislation, and we obtained permission to collect *M. simplicifolia* materials.

Sequencing, assembly, and annotation of the M. simplicifolia cp genome

Genomic DNA was extracted from the leaves using the DNeasy Plant Mini Kit (Qiagen, CA, USA) following the manufacturer's protocol. For Illumina sequencing, paired-end libraries, with an insertion size of 350 bp, were constructed according to the manufacturer's instructions and sequenced using the Illumina HiSeq 2500 platform. Approximately 5.57 Gb of raw reads were generated and assembled into non-redundant contigs using NOVOPlasty (Nicolas et al. 2017), a *de novo* sequence assembly software package, with k = 39 and a genome range of 120,000-200,000. Initial gene annotation of the *M. simplicifolia* cp genome was performed using Plann (Huang and Cronk 2015) with the annotation of *M. racemosa* as the reference (GenBank accession number: MK533649; Li et al. 2019), and the annotation was refined using Geneious software (https://www.geneious.com/) (Kearse et al. 2012).

SSRs and repeated sequences

The Perl script MISA (http://pgrc.ipk-gatersleben.de/misa/misa.html) was employed to detect simple sequence repeats (SSRs) within the cp genome, with the following parameters: 10, 5, 4, 3, 3, and 3 for mono-, di-, tri-, tetra-, penta-, and hexanucleotide repeats, respectively. Repeats (forward, palindrome, complement, and reverse sequences) were identified using the online REPuter software (https://bibiserv.cebitec.uni-bielefeld.de/reputer?id=reputer_view_submission), with default settings.

Comparative analysis of cp genome structure

Comparative genomic analysis of *M. simplicifolia* and six other *Meconopsis* species was performed using the BLAST Ring Image Generator (BRIG) software (Alikhan et al. 2011). The comparison and variation of the entire genome architecture were visualized using the Shuffle-LAGAN mode of the mVISTA software (Frazer et al. 2004). The whole-

genome alignments of the cp genomes from the following seven *Meconopsis* species were performed: *M. simplicifolia* (this study), *M. horridula* (MK533646; Li et al. 2019), M. *integrifolia* (MK533647; Li et al. 2019), M. *punicea* (MK533648; Li et al. 2019), M. *racemosa* (MK533649; Li et al. 2019), M. *henrici* (MN488591; Zhu and Zhang 2020), and *M. quintuplinervia* (MK801686; Xu et al. 2019). The IRscope tool was used to compare and illustrate the IR border regions of the seven *Meconopsis* species (Amiryousefi et al. 2018).

Phylogenetic analysis

For phylogenetic analysis, the cp genomes of the reported *Meconopsis* species were utilized. A species from the Papaveraceae family, *Papaver orientale* (NC_037832; Zhou et al. 2018) was used as the 'outgroup'. The coding sequences of the protein-coding genes present in all the cp genomes of the *Papaveraceae* species were aligned using MAFFT-LINSI v7.313 (Katoh and Standley 2013). The optimal trees were inferred by maximum likelihood phylogenetic analysis using RAxML v8.2.11 (Stamatakis 2006), with the GTRGAMMA model and 500 bootstrap replicates.

Results

Features of the cp genome of M. simplicifolia

The assembled cp genome of *M. simplicifolia* has been deposited in GenBank under the accession number NC_070211. The cp DNA of *M. simplicifolia* measured 152,772 bp in length and exhibited the typical quadripartite structure observed in other complete cp genomes of *Meconopsis* species (Table 1 and Fig. 2). It consisted of a pair of inverted repeat regions (IRa and IRb) spanning 25,651 bp each, along with separate single-copy regions, including the small single-copy region (SSC) of 17,646 bp and the large single-copy region (LSC) of 83,824 bp. The *M. simplicifolia* cp genome contained 131 predicted functional genes, comprising 84 protein-coding genes, 37 tRNA genes, 8 rRNA genes, and two pseudogenes (*rps15* and *rps19*) (Table 2). Among these genes, 21 were intron-containing genes, with 13 being protein-coding genes and 8 being tRNA genes. Out of the intron-containing genes, 18 contained a single intron, while three (*clpP*, *rps12*, and *ycf*3) contained two introns.

Table 1
Summary of complete chloroplast genomes for nine *Meconopsis* species.

Feature	*M. simplicifolia*	*M. horridula*	*M. integrifolia*	*M. punicea*	*M. racemosa*	*M. henrici*	*M. quintuplinervia*
Accession number	NC_070211	MK533646	MK533647	MK533648	MK533649	MN488591	MK801686
Genome size (bp)	152772	153785	151864	153259	153816	153388	154997
IR length (bp)	25651	51988	51306	51548	51988	26107	25984
SSC length (bp)	17646	17898	17749	17729	17898	17822	17876
LSC length (bp)	83824	83899	82809	83982	83930	83698	85153
No. of total genes	131	127	127	127	127	112	129
No. of protein coding genes	84	90	90	90	90	78	84
No. of tRNA genes	39	37	37	37	37	30	37
No. of rRNA genes	8	8	8	8	8	3	8
Overall GC content (%)	38.7	38.8	38.8	38.5	38.7	38.5	38.5

Table 2
List of genes in the *M. simplicifolia* chloroplast genome.

Category	Gene Group	Gene Name
Self-replication	Ribosomal protein (large subunit) (9)	rpl14, rpl16[a], rpl20, rpl22, rpl23[b], rpl32, rpl33, rpl36
	Ribosomal protein (Small subunit) (16)	rps2, rps3, rps4, rps7[b], rps8, rps11, rps12[a,b], rps14, rps15, rps16[a], rps18, rps19
	DNA-dependent RNA polymerase (4)	rpoA, rpoB, rpoC1[a], rpoC2
	rRNA genes (8)	rrn16[b], rrn23[b], rrn4.5[b], rrn5[b],
	tRNA genes (37)	trnH-GUG, trnK-UUU[a], trnQ-UUG, trnS-GCU, trnG-UCC[a], trnR-UCU, trnC-GCA, trnD-GUC, trnY-GUA, trnE-UUC, trnT-GGU, trnS-UGA, trnG-GCC, trnS-GGA, trnT-UGU, trnL-UAA[a], trnF-GAA, trnV-UAC[a], trnfM-CAU[b], trnW-CCA, trnP-UGG, trnI-CAU[b], trnL-CAA[b], trnV-GAC[b], trnI-GAU[a,b], trnA-UGC[a,b], trnR-ACG[b], trnN-GUU[b], trnL-UAG
Photosynthesis	Photosystem I (5)	psaA, psaB, psaC, psaI, psaJ
	Photosystem II (15)	psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ
	NADH dehydrogenase (11)	ndhA[a], ndhB[a,b], ndhC, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK
	Cytochrome b/f complex (6)	petA, petB[a], petD[a], petG, petL, petN
	ATP synthase (6)	atpA, atpB, atpE, atpF[a], atpH, atpI
	Large subunit of rubisco (1)	rbcL
Other genes	Translational initiation factor (1)	infA
	ATP-dependent protease subunit p gene (1)	clpP[a]
	Maturase (1)	matK
	Envelop membrane protein (1)	cemA
Unknowns	Subunit of acetyl-CoA-carboxylase (1)	accD

Notes: [a] Genes containing introns; [b] Two gene copies in IR.

Category	Gene Group	Gene Name
	C-type cytochrome synthesis gene (1)	ccsA
	Conserved hypothetical chloroplast ORF (7)	ycf1[b], ycf2[b], ycf3[a], ycf4, ycf15,
	Pseudogene (2)	rps15, rps19

Notes: [a] Genes containing introns; [b] Two gene copies in IR.

The overall GC content of the *M. simplicifolia* cp genome was 38.74%. Regarding the protein-coding regions, the GC content of the first, second, and third codons was 45.91%, 38.17%, and 31.15%, respectively. The 84 protein-coding genes encoded 25,454 codons, with leucine (L) being the most frequently used amino acid (10.36%) and cysteine (C) being the least frequently used (1.19%).

SSR and Long-Repeat Analysis

Simple sequence repeats (SSRs), also known as microsatellites, are tandem repeat sequences consisting of 1–6 nucleotide repeat units. They are widely distributed in cp genomes and often used as genetic markers in population genetics and evolutionary studies due to their high intraspecific variability (Ellis and Burke 2007). In this study, we analyzed the SSRs in the cp genomes of *M. simplicifolia* and six other *Meconopsis* species. A total of 33 SSRs were identified in the *M. simplicifolia* cp genome. Similarly, *M. horridula*, *M. integrifolia*, *M. punicea*, *M. racemose*, *M. henrici*, and *M. quintuplinervia* contained 38, 33, 34, 40, 23, and 35 SSRs, respectively (Fig. 3A). Mononucleotide repeats were the most abundant in all seven species. Dinucleotide repeats were the second most prevalent, with *M. simplicifolia* having four and *M. racemose*, *M. horridula*, and *M. punicea* having eight dinucleotide repeats. In *M. simplicifolia*, all mononucleotide repeats (100%) and the majority of dinucleotide repeats (75%) consisted of A/T nucleotides (Fig. 3B).

Repeat sequences play a crucial role in phylogenetic research and genome reorganization. In the cp genome of *M. simplicifolia*, 27 dispersed repeats were identified, including 10 forward repeats and 17 palindromic repeats (Fig. 4A). This pattern is consistent with the six other *Meconopsis* cp genomes, with the number of repeats ranging from 29 in *M. punicea* to 50 in *M. integrifolia*. Palindromic repeats were the most prevalent repeat type among the seven *Meconopsis* species (Fig. 4B, C). Most of these repeats ranged from 30 bp to 44 bp in length.

Comparative analysis of cp genomes of Meconopsis Species

Comparative analysis of cp genomes provided valuable insights into intricate evolutionary relationships. In this study, we compared the cp genomes of *M. simplicifolia* and six other *Meconopsis* species. The size of the seven *Meconopsis* cp genomes ranged from 151,864 (*M. integrifolia*) to 154,997 bp (*M. quintuplinervia*). Notably, within the genus *Meconopsis*, the genome of *M. simplicifolia* exhibited a higher degree of conservation and could be accurately mapped (Fig. 5). To assess the sequence identity of the *Meconopsis* cp genomes, we used the mVISTA software. The results revealed that the IR regions exhibited fewer differences compared to the LSC and SSC regions (Fig. 6). Non-coding regions displayed more variability than coding regions, with significant changes observed in intergenic spacers among the seven cp genomes. These highly divergent regions included *trnH-psbA*, matK, *rps16-psbK*, *atpH-atpI*, *rpoC2*, *psbM-petN*, *trnE-trnT*, *trnT-psbD*, *psaA-ycf3*, *trnF-ndhJ*, *ndhK*, *ndhC-trnV*, *atpB-rbcL*, *accD*, *ycf4-cemA*, *petA-psbL*, *psbE-petL*, *clpP-psbB*, *rpl16*, *ndhF-rpl32-ccsA*, *ycf1*, among others.

A detailed comparison of the binding regions between the inverted repeats (IR/LSC and IR/SSC) was performed among the seven species (Fig. 7). In all species, the *rpl22* gene was located within the LSC region. Variations in gene content and order were observed, such as the presence of the *ycf1* gene at the SSC/IRa junction in *M. simplicifolia, M. horridula, M. punicea, M. racemose, M. henrici*, and *M. quintuplinervia*, while *M. integrifolia* had a missing *ycf1* gene in the SSC/IRa junction. Expansion and contraction of the inverted repeat region were observed. For example, the *rps19* gene was found within the LSC region in *M. racemose*, while in the other six *Meconopsis* species, it was located 67–158 bp away, spanning the LSC and IRb binding regions. Except for *M. simplicifolia*, the *rpl2* gene did not extend into the LSC region in the other species. Overall, there were only minor variations in the IR boundary regions among the cp genomes of these seven *Meconopsis* species.

Phylogenetic analysis of M. simplicifolia and related Meconopsis species cp genomes

To determine the phylogenetic position of *M. simplicifolia* within the Papaveraceae family, we utilized the cp genomes of ten Papaveraceae members, including *M. simplicifolia*. A species tree was constructed based on the alignment of 75 shared protein-coding genes from the cp genomes. The analysis revealed that *M. simplicifolia* clustered together with *M. betonicifolia* (Fig. 8). All *Meconopsis* plants formed a well-supported branch, indicating the high potential of cp genomes for species differentiation within the order Papaveraceae.

Discussion

The cp genome serves as a valuable resource for interspecific differentiation and various biotechnological applications (Bock 2015). In this study, we successfully sequenced and assembled the complete cp genome of *M. simplicifolia*. Furthermore, a comparative analysis of the cp genomes of *M. simplicifolia* and six other *Meconopsis* species was conducted. Similar to other seed plants (Daniell et al. 2016), the cp genome of *M. simplicifolia* exhibited a classic quadripartite structure and gene content. The size of the *M. simplicifolia* cp genome was 152,772 bp, consistent with other *Meconopsis* species ranging from 151–154 kb (Li et al. 2019; Xu et al. 2019; Zhu and Zhang 2020). All these complete cp genomes displayed a GC content of 38%, which is in line with the low GC content observed in the cp genomes of other angiosperms (Wicke et al. 2011).

Simple sequence repeats (SSRs) have been extensively used for phylogenetic relationships, genetic diversity studies, and species identification. The majority of SSRs found in the seven *Meconopsis* cp genomes consisted of mononucleotide repeats predominantly composed of A and T, consistent with previous research on *Meconopsis* species (Li et al. 2019). The availability of genomic resources in *Meconopsis* plants can enhance the understanding of population patterns and the identification of gene regions associated with important medicinal and environmental adaptive traits (Li et al. 2019). SSRs can serve as useful tools for species identification and genetic diversity studies of *Meconopsis* species and their relatives. Additionally, repeat sequences can help identify mutational hotspots and establish phylogenetic relationships. In this study, we identified 27 repeated sequences in *M. simplicifolia*, which is fewer compared to other *Meconopsis* species. Most of these repeats were located in genes, indicating that the cp genome of *M. simplicifolia* retains a significant amount of genetic material.

Analysis of genes with known functions revealed that *M. simplicifolia* shares 75 protein-coding genes with the other six *Meconopsis* species. Similar to other *Meconopsis* species, *M. simplicifolia* lacks the *rpl2* gene, which plays a crucial role in chloroplast development during early leaf development (Lan et al. 2023). It is hypothesized that the *Msrpl2* gene has either been functionally transferred to the nucleus or replaced by an eukaryotic gene. Comparative analysis of cp genomes using BRIG and mVISTA showed high sequence identity among all *Meconopsis* species. However, *Meconopsis* cp genomes have also undergone gene duplication, gene/intron loss, insertion/deletion, pseudogenization, and varying expansion/contraction of the inverted repeat region. These genomic events are consistent with observations in other

angiosperms (Jansen and Ruhlman 2012; Wicke et al. 2011). In *M. simplicifolia*, the pseudogenization of *rps19* in the IR regions is consistent with a previous study (Li et al. 2019). The sequence and content of the SC regions show less similarity compared to the IR regions in *Meconopsis* species. The most highly divergent non-coding regions were identified in the intergenic regions of *trnT-psbD*, *ndhC-trnV*, and *ndhF-rpl32-ccsA*, which have also been recognized as molecular markers in many land plants (Chiapella et al. 2019; Li et al. 2019; Zhang et al. 2018a).

Taxonomy and phylogeny of *Meconopsis* have been extensively studied at the genus level (Xiao 2013; Xiao and Simpson 2017). Previous studies on the evolutionary relationships among different *Meconopsis* species utilized internal transcribed spacer and cpDNA sequences, as well as amplified fragment length polymorphisms (AFLPs) (Valtuea et al. 2011; Xiao 2013; Xiao and Simpson 2017). However, complete genome sequencing offers a more comprehensive perspective (Li et al. 2019). In the case of *M. simplicifolia*, limited information was available. The phylogenetic position of *M. simplicifolia* within *Meconopsis* was determined using cp genomes and 75 protein-coding genes among nine *Meconopsis* species. Phylogenetic analysis revealed that all *Meconopsis* species formed a monophyletic clade with 100% bootstrap support. *M. simplicifolia* was found to be closely related to *M. betonicifolia*, supporting previous morphological and molecular data (Xiao 2013).

Conclusion

This study successfully characterized the complete cp genome of *M. simplicifolia* and conducted a comparative analysis of *Meconopsis* cp genomes, revealing the conserved genome structure and organization across seven *Meconopsis* species. The most divergent regions among these Meconopsis cp genomes were identified in three non-coding intergenic spacer (IGS) regions (*trnT-psbD*, *ndhC-trnV*, and *ndhF-rpl32-ccsA*) and three genic regions (*matK*, *rpoC2*, and *ycf1*). Moreover, the genetic resources such as SSRs and repetitive sequences discovered in the cp genomes can serve as valuable molecular markers for the identification of *Meconopsis* species. Phylogenetic analysis demonstrated that *M. simplicifolia* is closely related to *M. betonicifolia*. These comprehensive cp genomes provide essential resources for genetic and biological studies on *Meconopsis* and other species within the Papaveraceae family.

Declarations

Acknowledgements

This study was supported by Tibet Antonomous Region Science and Technology Plan Project (Key Research and Development Program) (No. XZ202201ZY0031G). The authors thank Dr. Lei Zhang from North Minzu University for helping with plant material collection, and thank Dr. Mingcheng Wang from Chengdu University for his useful suggestions in data analysis. We would like to thank MogoEdit (https://www.mogoedit.com) for its English editing during the preparation of this manuscript.

Disclosure statement

None of the co-authors have any conflict of interest to declare. The authors alone are responsible for the content and composition of the paper.

Author contribution

All authors contributed to the study conception and design. Rui Li and Yixi Yang conceived and supervised the project. Rui Li, Zhidan Zhu and Peng Huang analyzed the data and wrote the manuscript. All of the authors have read and approved the final manuscript and have agreed to be accountable for all aspects of the work.

Data availability

The genome sequence 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 no. NC_070211. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA935484, SRR23491443, and SAMN33317127, respectively.

References

1. Ali A, Jaakko H, Peter P (2018) IRscope: an online program to visualize the junction sites of chloroplast genomes. Bioinformatics 34(17):3030-3031. https://doi.org/10.1093/bioinformatics/bty220
2. Alikhan NF, Petty NK, Zakour NL, Beatson SA (2011) BLAST ring image generator (BRIG): simple prokaryote genome comparisons. BMC Genomics 12:402-411. https://doi.org/10.1186/1471-2164-12-402
3. Bi Y, Zhang M, Xue J, Dong R, Du Y, Zhang X (2018) Chloroplast genomic resources for phylogeny and DNA barcoding: a case study on *Fritillaria*. Sci Rep 8(1):1184-1195. https://doi.org/10.1038/s41598-018-19591-9
4. Bock R (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. Annu Rev Plant Biolo 66:211-241. https://doi.org/10.1146/annurev-arplant-050213-040212
5. Chiapella JO, Barfuss MHJ, Xue Z, Greimler J (2019) The plastid genome of *Deschampsia cespitosa* (Poaceae). Molecules 24(2):216-227. https://doi.org/10.3390/molecules24020216
6. Daniell H, Lin C, Yu M, Chang W (2016) Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol 17(1):134-162. https://doi.org/10.1186/s13059-016-1004-2
7. Dierckxsens N, Mardulyn P, Smits G (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res 45(4):e18-e26. https://doi.org/10.1093/nar/gkw955
8. Ellis JR, Burke JM (2007) EST-SSRs as a resource for population genetic analyses. Heredity(Edilib) 99(2):125-132. https://doi.org/10.1038/sj.hdy.6801001
9. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for comparative genomics. Nucleic Acids Res 32(1):273-279. https://doi.org/10.1093/nar/gkh458
10. Grierson AJC, Long DG (1983) Flora of Bhutan, including a record of plants from Sikkim (Volume 1 parts 1 and 2). Royal Botanic Garden, Edinburgh
11. Guo Q, Bai R, Zhao B, Feng X, Zhao Y, Tu P, Chai X (2016) An ethnopharmacological, phytochemical and pharmacological review of the genus *Meconopsis*. Am J Chin Med 44(3):439-462. https://doi.org/10.1142/S0192415X16500257
12. Hao D, Chen S, Xiao P (2010) Molecular evolution and positive Darwinian selection of the chloroplast maturase matK. J Plant Res 123(2):241-247. https://doi.org/10.1007/s10265-009-0261-5
13. Hu H, Hu Q, Shehbaz IA, Luo X, Zeng T, Guo X, Liu J (2016) Species delimitation and interspecific relationships of the genus *Orychophragmus* (Brassicaceae) inferred from whole chloroplast genomes. Front Plant Sci 7:1826-1835. https://doi.org/10.3389/fpls.2016.01826
14. Huang DI, Cronk QCB (2015) Plann: a command-line application for annotating plastome sequences. *Appl in Plant Sci* 3(8):1500026-1500028. https://doi.org/10.3732/apps.1500026
15. Jansen RK, Ruhlman TA (2012) Plastid genomes of seed plants. In: Knoop V (ed) Genomics of chloroplasts and mitochondria. Springer, Netherlands, pp 103-126
16. Grey-Wilson C (2016) The genus *Meconopsis*: blue poppies and their relatives. Royal Botanic Gardens, Kew
17. Kapralov MV, Filatov DA (2007) Widespread positive selection in the photosynthetic Rubisco enzyme. BMC Evol Biol 7:73-82. https://doi.org/10.1186/1471-2148-7-73
18. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30(4):772-780. https://doi.org/10.1093/molbev/mst010

19. Kearse M, Moir R, Wilson A, Havas SS, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12):1647-1649. https://doi.org/10.1093/bioinformatics/bts199
20. Kwon EC, Kim JH, Kim NH (2020) Comprehensive genomic analyses with 115 plastomes from algae to seed plants: structure, gene contents, GC contents, and introns. Genes Genomics 42(6): 553-570. https://doi.org/10.1007/s13258-020-00923-x
21. Lan J, Lin Q, Zhou C, Liu X, Miao R, Ma T, Chen Y, Mou C, Jing R, Feng M, Nguyen T, Ren Y, Cheng Z, Zhang X, Liu S, Jiang L, Wan J (2023) Young Leaf White Stripe encodes a P-type PPR protein required for chloroplast development. J Integr Plant Biol 65(7):1687-1720. https://doi.org/10.1111/jipb.13477
22. Li D, Zhao C, Liu X (2019) Complete chloroplast genome sequences of *Kaempferia galanga* and *Kaempferia elegans*: molecular structures and comparative analysis. Molecules 24(3):474-491. https://doi.org/10.3390/molecules24030474
23. Li X, Tan W, Sun J, Du J, Zheng C, Tian X, Zheng M, Xiang B, Wang Y (2019) Comparison of four complete chloroplast genomes of medicinal and ornamental *Meconopsis* species: genome organization and species discrimination. Sci Rep 9(1):10567-10578. https://doi.org/10.1038/s41598-019-47008-8
24. Liu F, Rong H, Li W (2016) Analysis of effective components of Tibetan medicine *meconopsis* antioxidant five veins. Medicine and Biopharmaceutical: Proceedings of 2015 International Conference 972-978
25. Nicolas D, Patrick M, Guillaume S (2017) NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res 45(4):e18-e26. https://doi.org/10.1093/nar/gkw955
26. Shi N, Naudiyal N, Wang J, Gaire NP, Wu Y, Wei Y, He J, Wang CY (2022) Assessing the impact of climate change on potential distribution of *Meconopsis punicea* and its influence on ecosystem services supply in the southeastern margin of qinghai-tibet plateau. Front Plant Sci 12:830119-830135. https://doi.org/10.3389/fpls.2021.830119
27. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22(21):2688-2690. https://doi.org/10.1093/bioinformatics/btl446
28. Sulaiman IM, Rangaswamy NS, Babu CR (1991) Formation of plantlets through somatic embryogeny in the Himalayan blue poppy, *Meconopsis simplicifolia* (Papaveraceae). Plant Cell Rep 9(10):582-585. https://doi.org/10.1007/BF00232337
29. Tétényi P (2003) Chemo-Dahlgrenogram of the Tribe Papavereae. III WOCMAP Congress on Medicinal and Aromatic Plants-Volume 1: Bioprospecting and Ethnopharmacology 675:49-57
30. Valtuea FJ, Preston CD, Kadereit JW (2011) Evolutionary significance of the invasion of introduced populations into the native range of *Meconopsis cambrica*. Mol Ecol 20(20):4318-4331. https://doi.org/10.1111/j.1365-294X.2011.05273.x
31. Wang M, Li R, Zhao Q (2023) Multi-omics techniques in genetic studies and breeding of forest plants. Forests 14(6):1196. https://doi.org/10.3390/f14061196
32. Wangchuk P, Keller PA, Pyne SG, Lie W, Willis AC, Rattanajak R, Kamchonwongpaisan S (2013) A new protoberberine alkaloid from *Meconopsis simplicifolia* (D. Don) Walpers with potent antimalarial activity against a multidrug resistant Plasmodium falciparum strain. J Ethnopharmacol 150(3):953-959.https://doi.org/10.1016/j.jep.2013.09.052
33. Wang Y, Wang S, Liu Y, Yuan Q, Sun J, Guo L (2021) Chloroplast genome variation and phylogenetic relationships of *Atractylodes* species. BMC Genomics 22(1):1-12. https://doi.org/10.1186/s12864-021-07394-8
34. Wangchuk P, Jamtsho T (2023) Phytopharmaceutical properties and quality assessment of two Himalayan medicinal plants, *Meconopsis horridula* and *Meconopsis simplicifolia*. J Herb Med 38:100628.

https://doi.org/10.1016/j.hermed.2023.100628

35. Wei L, Zhou Y, Liu Q (2019) Molecular and morphological evaluation of *Meconopsis georgei* and *M. castanea* in southwest China. Nordic J Bot 37(9):2369. https://doi.org/10.1111/njb.02369
36. Wicke S, Schneeweiss GM, Depamphilis CW, Müller KF, Quandt D (2011) The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol 76(3-5):273-297. https://doi.org/10.1007/s11103-011-9762-4
37. Xiao W (2013) Molecular systematics of *Meconopsis Vig*. (Papaveraceae): taxonomy, polyploidy evolution, and historical biogeography from a phylogenetic insight. Dissertation, University of Texas
38. Xiao W, Simpson BB (2017) A New infrageneric classification of *Meconopsis* (Papaveraceae) based on a well-supported molecular phylogeny. Syst Bot 42(2):226-233. https://doi.org/ 10.1600/036364417X695466
39. Xu B, Kang H, Luo S, Yang Q, Zhao Z, Peng F, Liu Z (2019) Complete chloroplast genome of the *Meconopsis quintuplinervia* (Papaveraceae), a traditional medicine of Tibetan. Mitochondrial DNA B Resour 4(2):2335-2336. https://doi.org/10.1080/23802359.2019.1629343
40. Zhang Z, An M, Miao J, Gu Z, Liu C, Zhong B (2018) The Antarctic sea ice alga Chlamydomonas sp. ICE-L provides insights into adaptive patterns of chloroplast evolution. BMC Plant Biol 18(1):53-64. https://doi.org/10.1186/s12870-018-1273-x
41. Zhang X, Rong C, Qin L, Mo C, Fan L, Yan J, Zhang M (2018) Complete chloroplast genome sequence of Malus hupehensis: genome structure, comparative analysis, and phylogenetic relationships. Molecules 23(11):2917-2933. https://doi.org/10.3390/molecules23112917
42. Zhou J, Cui Y, Chen X, Li Y, Xu Z, Duan B, Yao H (2018) Complete chloroplast genomes of *Papaver rhoeas* and *Papaver orientale*: molecular structures, comparative analysis, and phylogenetic analysis. Molecules 23(2):437-451. https://doi.org/10.3390/molecules23020437
43. Zhu Y, Zhang D (2020) Complete chloroplast genome sequences of two species used for Tibetan medicines, *Meconopsis punicea* vig. and *M. henrici* vig. (Papaveraceae). Mitochondrial DNA B Resour 5(1):48-50. https://doi.org/10.1080/23802359.2019.1693918

Figures

Figure 1

Photograph of an adult *M. simplicifolia* plant taken by Peng Huang.

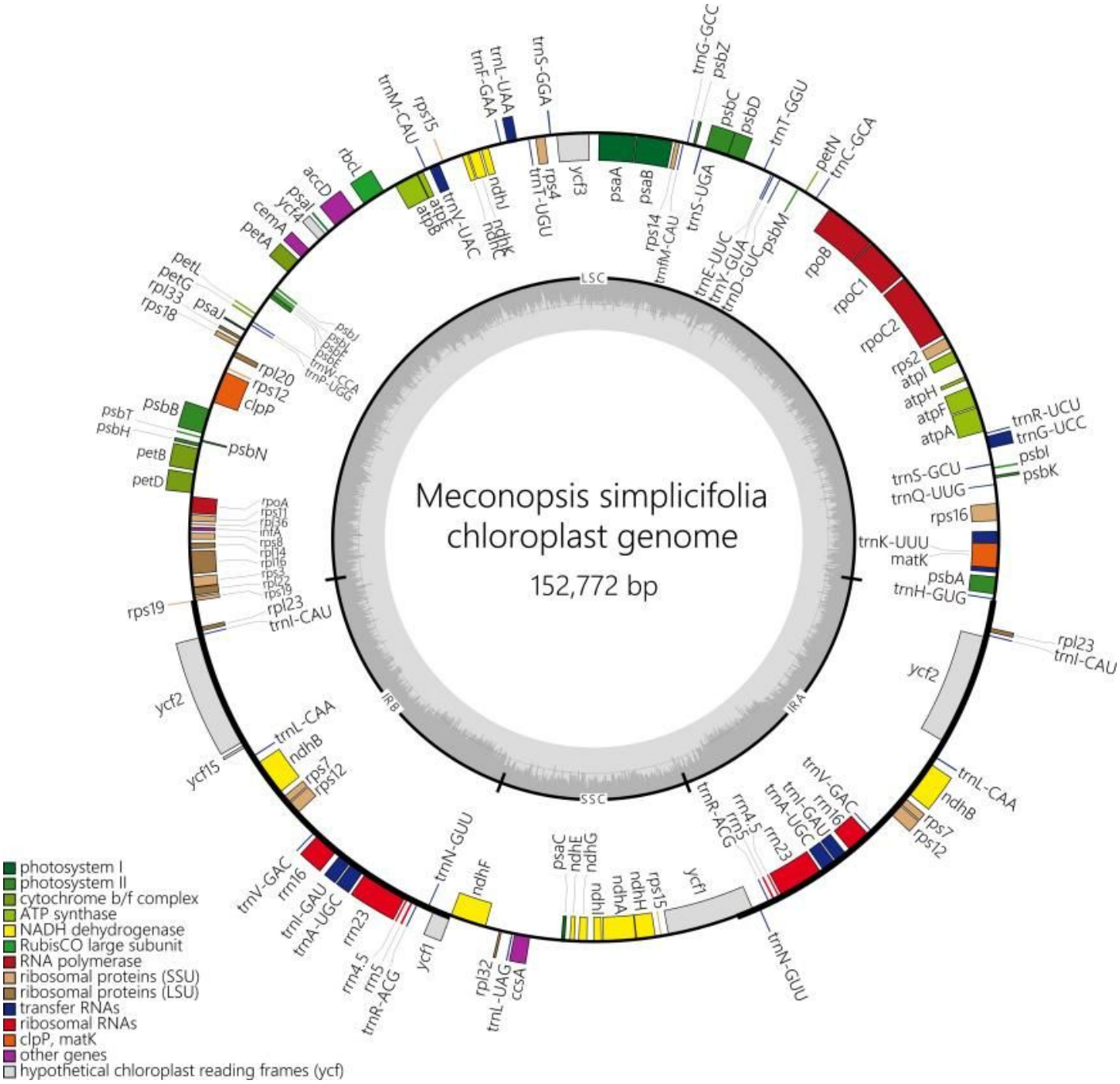

Figure 2

Circos map illustrating the cp genome of *M. simplicifolia.* The inner track displays the quadripartite structure and GC content, while the outer track shows the location and type of genes.

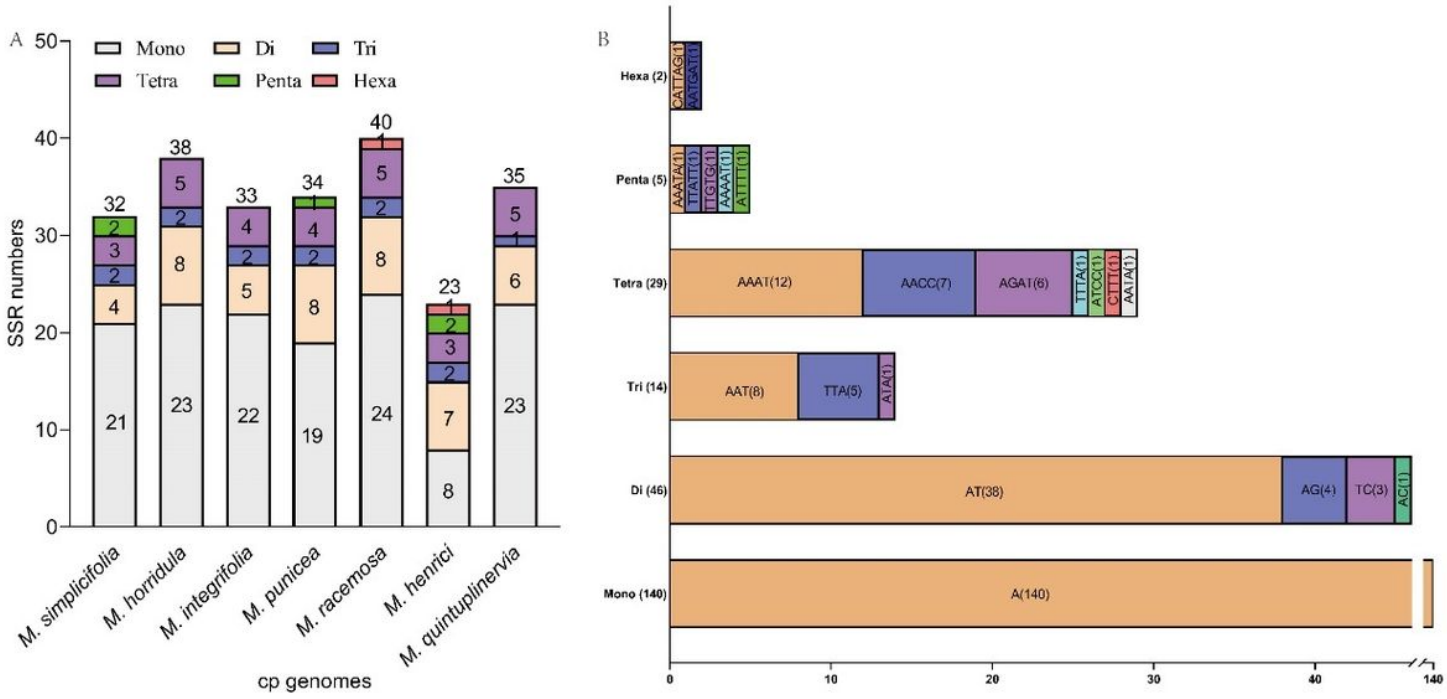

Figure 3

Analysis of simple sequence repeats (SSRs) in the cp genomes of the seven *Meconopsis* species. (A) Number of different types of SSRs; (B) Frequency of identified SSR motifs in different repeat classes.

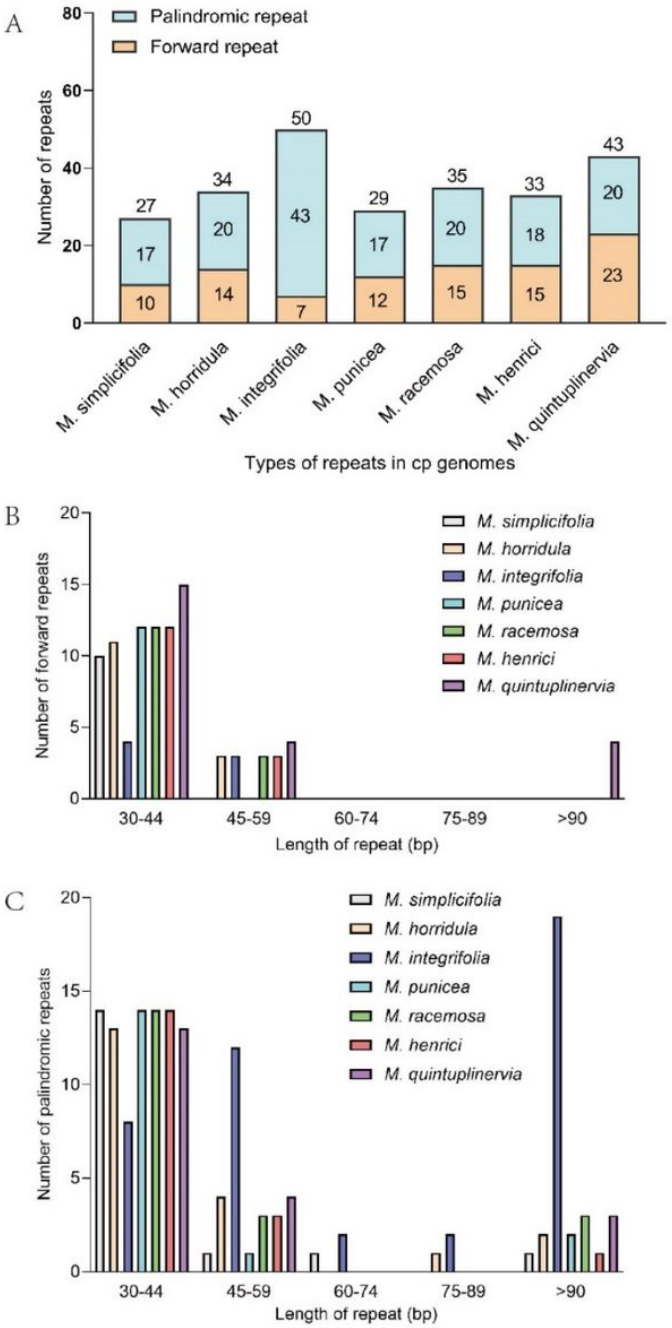

Figure 4

Analysis of repeated sequences in the cp genomes of the seven *Meconopsis* species. (A) Frequency of palindromic and forward repeats; (B) Frequency of forward repeats by length; (C) Frequency of palindromic repeats by length.

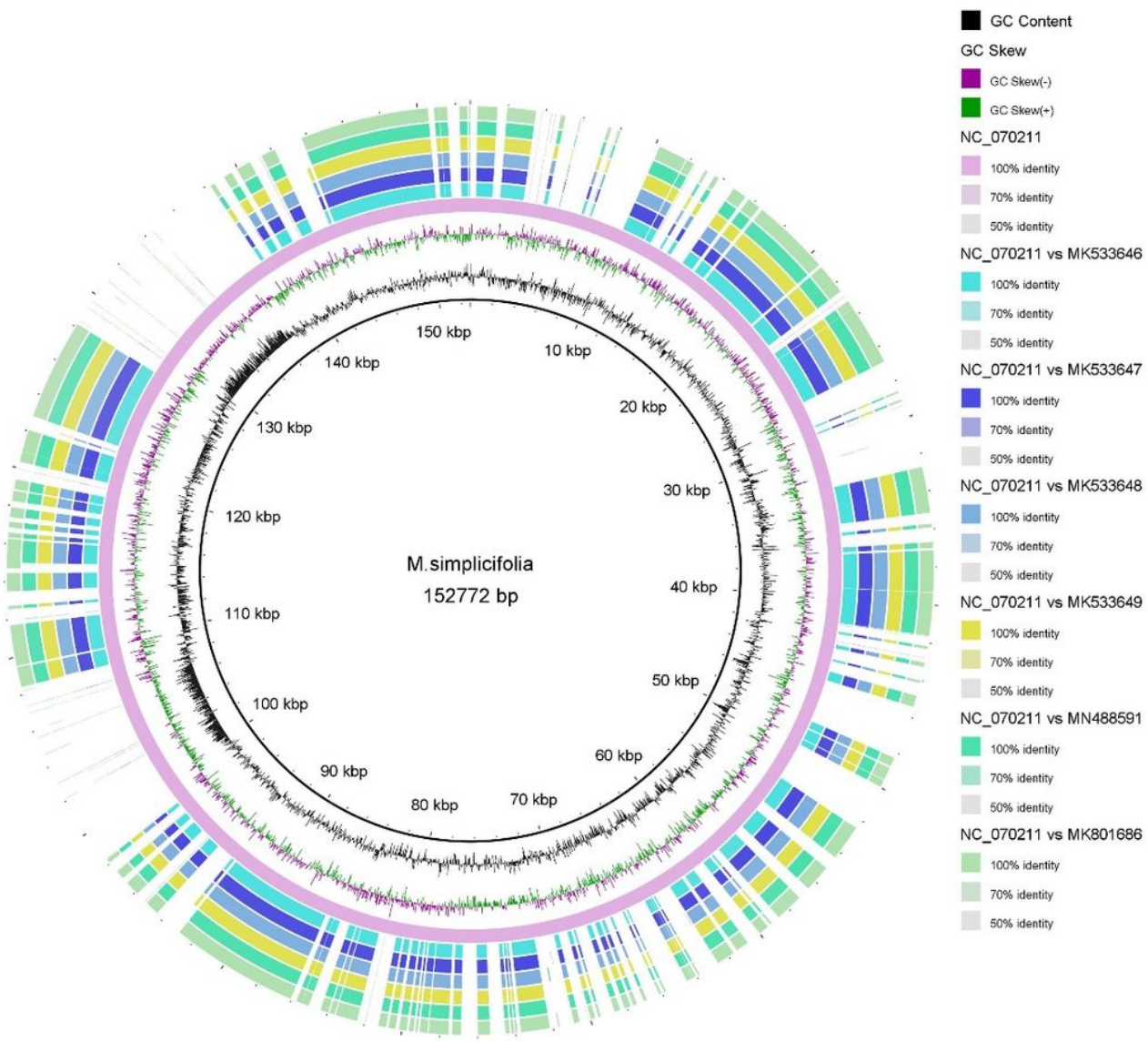

Figure 5

BRIG ring comparisons of the cp genomes of the seven *Meconopsis* species. The innermost rings represent the GC content (black) and GC skew (purple/green). The seven outer circles depict the similarity results compared to the reference genome (*M. simplicifolia*. NC_070211).

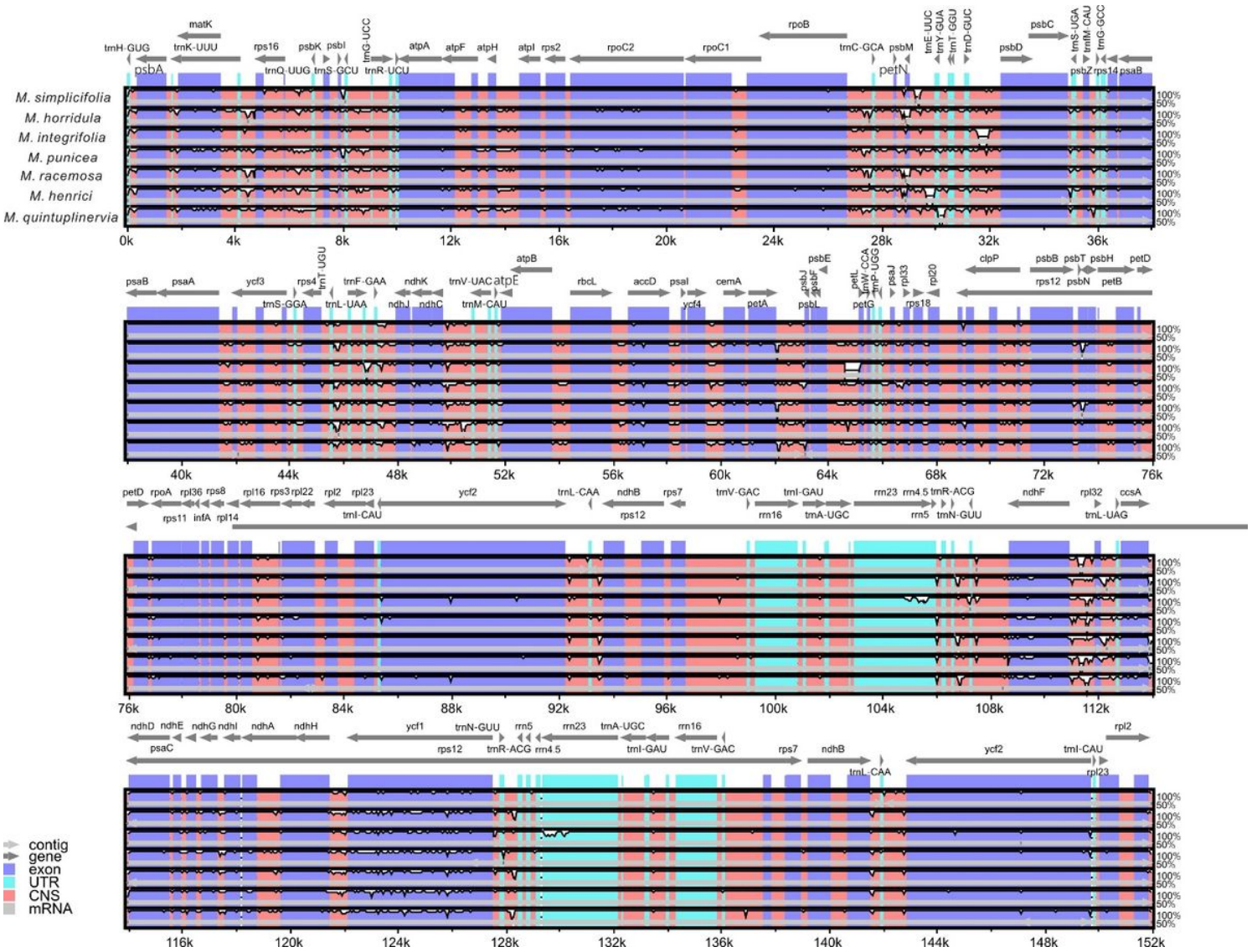

Figure 6

Alignment of the cp genomes of the seven *Meconopsis* species.

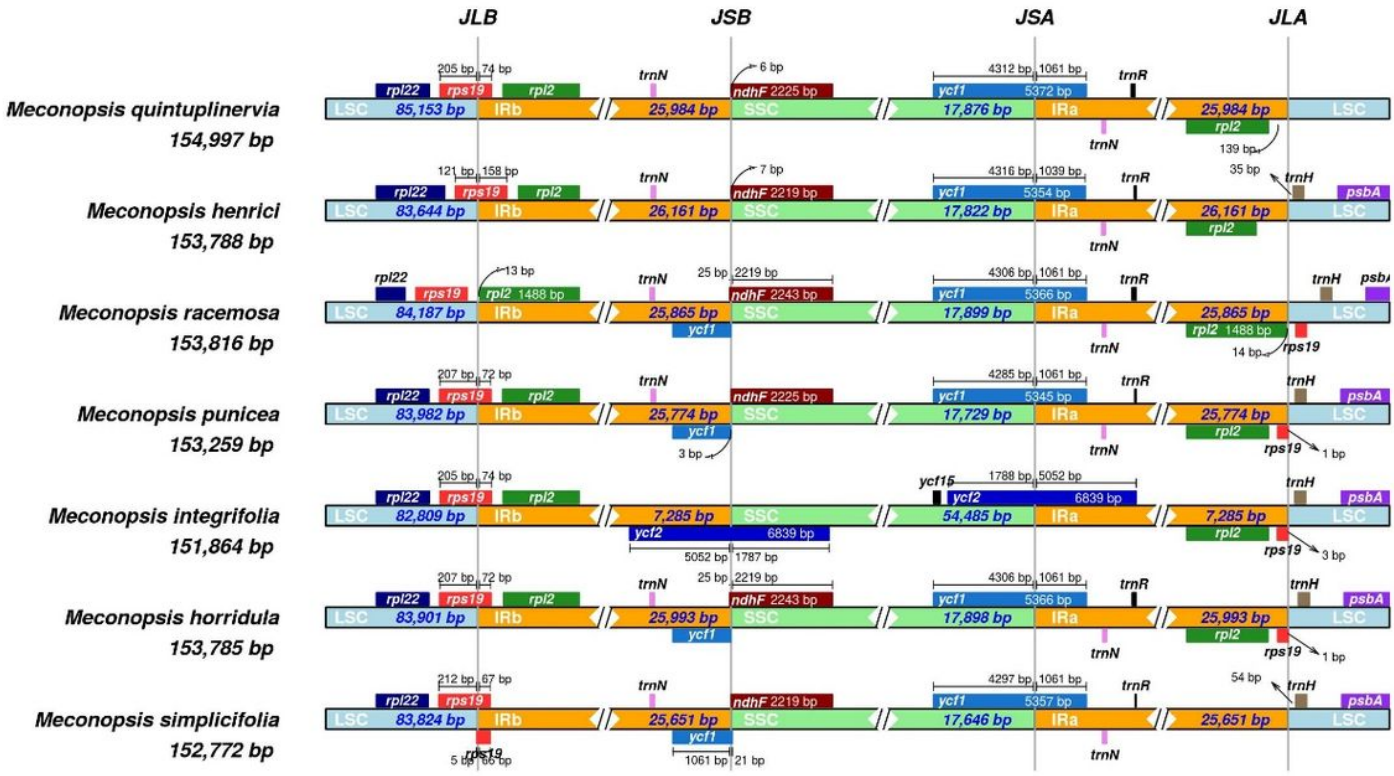

Figure 7

Comparison of the border regions of LSC, SSC, and IR regions among the cp genomes of the seven *Meconopsis* species.

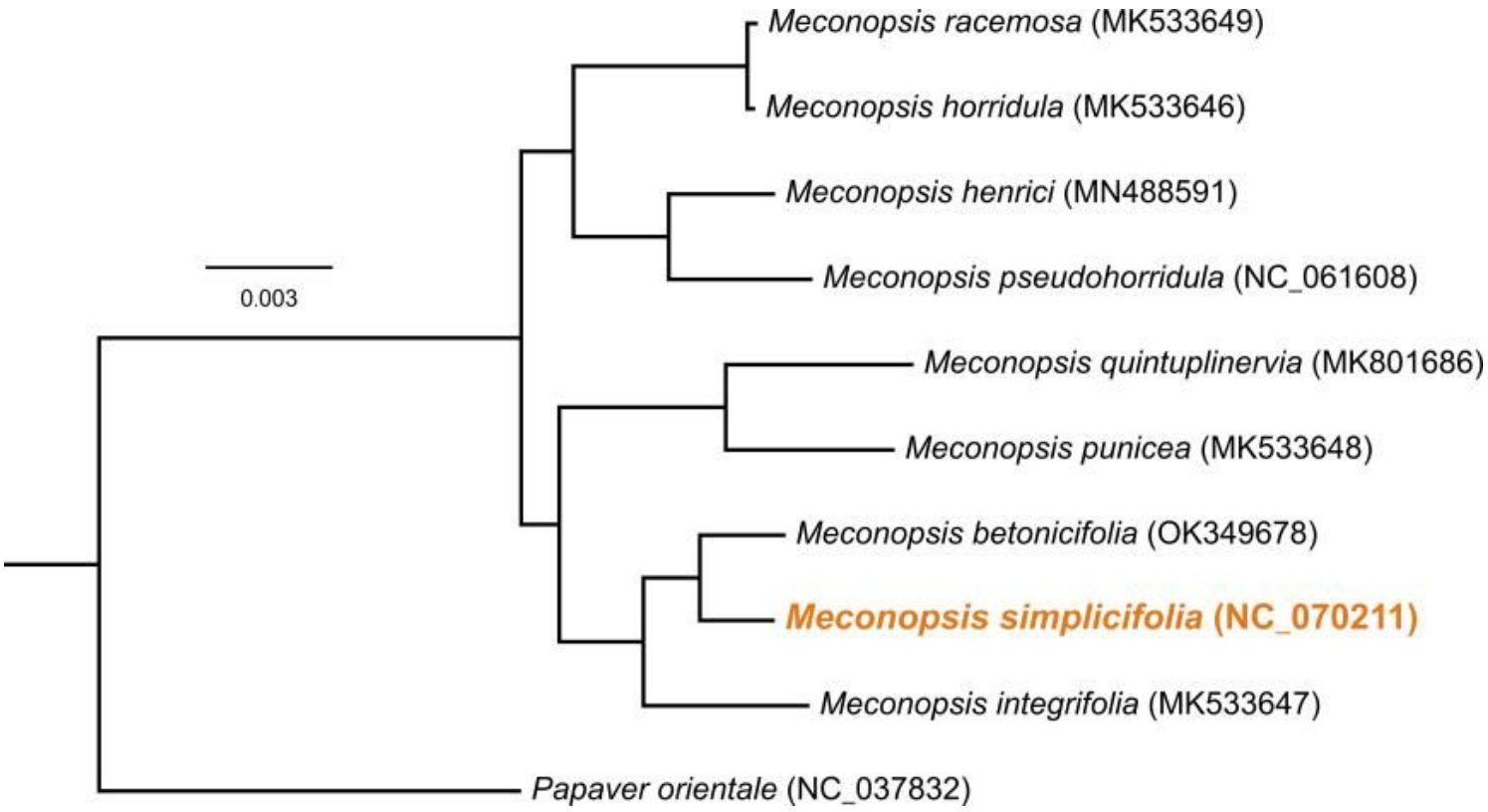

Figure 8

Species tree based on the coding sequences of 75 protein-coding genes shared by the cp genomes of 10 Papaveraceae members, including *M. simplicifolia* (NC_070211; this study), *M. horridula* (MK533646; Li et al., 2019), *M. integrifolia* (MK533647; Li et al., 2019), *M. punicea* (MK533648; Li et al., 2019), *M. racemosa* (MK533649; Li et al., 2019), *M. henrici* (MN488591; Zhu and Zhang, 2020), *M. quintuplinervia* (MK801686; Xu et al., 2019), *M. pseudohorridula*(NC_061608; unpublished), *M. betonicifolia* (OK349678; unpublished), and the outgroup species *Papaver orientale* (NC_037832; Zhou et al., 2018). All nodes received 100% bootstrap support.