

Characterization of the complete chloroplast genome of the medicinal herb *Veronica polita* Fr. 1819 (Lamiales: Plantaginaceae)

Shouning Jia^a, Wenjuan Chen^a, Guofu Zhao^a, Shuangxi Wang^a, Zhiwei Xu^a and Qien Li^b

^aQinghai Provincial Hospital of Traditional Chinese Medicine, Xining, People's Republic of China; ^bTibetan Medicine Research Center, Tibetan Medical College, Qinghai University, Xining, People's Republic of China

ABSTRACT

Veronica polita Fr. 1819 (synonym: *Veronica didyma* Ten. 1981) is a species of annual herb with high medicinal value. It is originally from Southwest Asia, but has been naturalized widely in many regions of the world. In this study, the complete chloroplast genome of *V. polita* was determined to be 150,191 bp long with a typical quadripartite structure, comprising two inverted repeat regions (IRa and IRb, 25,465 bp each), a large single-copy (LSC) region (81,847 bp) and a small single-copy (SSC) region (17,414 bp). It encodes a panel of 114 genes (including 79 protein-coding, 31 tRNA, and four rRNA genes) with 18 of them being completely or partially duplicated and 19 of them possessing one or two introns. Phylogenetic analysis supported the tribal-level taxonomy of the family Plantaginaceae, and revealed that *V. polita* was most closely related to the congener *Veronica persica* Poir. 1808.

ARTICLE HISTORY

Received 10 March 2022
Accepted 26 May 2022

KEYWORDS

Chloroplast genome; high-throughput sequencing; medicinal herb; phylogenetic analysis; *Veronica polita* Fr. 1819

Veronica polita Fr. 1819 (synonym: *Veronica didyma* Ten. 1981) is a species of annual herb in the family Plantaginaceae (order Lamiales) (Schoch et al. 2020). It is originally from Southwest Asia, but has been naturalized widely in the temperate and subtropical regions of the world (Hong and Fischer 1998; Shanghai Academy of Science & Technology 1999). This herb has long been used in traditional Chinese medicine for treating such illnesses as dysmenorrhea, metrorrhagia, and malaria (Nanjing University of Chinese Medicine 2005), and has also recently been reported for anti-inflammatory and antioxidant activities (Akanda et al. 2018). To date, previous studies of *V. polita* have mainly focused upon its chemical composition (Wang et al. 1995), agricultural implications (Li 2003; Zhang J et al. 2006; Zhong and Wang 2010), economic values (Yang et al. 2006), ecology (Tang et al. 2006), physiology (Zhang Y et al. 2010; Li and Xu 2015), and pharmacological effects (Akanda et al. 2018). Its genetics and genomics have yet to be investigated. In this study, the complete chloroplast (cp) genome of *V. polita* was retrieved from Illumina sequencing reads, and its phylogenetic placement was investigated.

Leaf tissues were collected from an individual of *V. polita* in Haihu New District, Xining, Qinghai Province, China (101.73°E, 36.64°N), and were used for the genomic DNA isolation with the DNeasy Plant Mini Kit (Qiagen, Valencia, CA). A specimen was deposited at the Herbarium of Qinghai Provincial Hospital of Traditional Chinese Medicine (<http://www.qhzyy.com.cn/>; Shouning Jia, Email: iashouning@163.com).

under the voucher number VDIDY20210930. High-throughput DNA sequencing was conducted on the Illumina HiSeq X Ten Sequencing System (Illumina, San Diego, CA), yielding 17.88 M of 150-bp paired-end reads in all. The cp genome was assembled using the software MITObim v1.9 (Hahn et al. 2013) with that of *Veronica persica* Poir. 1808 (GenBank accession: KT724052) (Choi et al. 2016) as the initial reference. Annotation of the cp genome was done in Geneious Prime (Biomatters Ltd., Auckland, New Zealand) by aligning with those of its congeners.

The cp genome of *V. polita* is 150,191 bp long with a typical quadripartite structure, comprising two inverted repeat regions (IRa and IRb, 25,465 bp each), a large single-copy (LSC) region (81,847 bp) and a small single-copy (SSC) region (17,414 bp). The base composition is asymmetric (30.7%A, 19.2%C, 18.7%G, and 31.4%T) with an overall A + T content of 62.1% ('light strand'). The A + T contents of LSC (64.1%) and SSC (68.4%) are obviously higher than that of IR regions (56.8%). The cp genome encodes a total of 114 genes, including 79 protein-coding gene (PCG), 31 tRNA, and four rRNA genes. In all, 18 genes are completely or partially duplicated, including seven PCGs (*ndhB*, *rpl2*, *rpl23*, *rps7*, *rps12*, *ycf1*, and *ycf2*), seven tRNA (*trnA*-UGC, *trnI*-CAU, *trnI*-GAU, *trnL*-CAA, *trnN*-GUU, *trnR*-ACG, and *trnV*-GAC), and all four rRNA genes (*rrn4.5*, *rrn5*, *rrn16*, and *rrn23*). Ten PCGs (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl2*, *rpl16*, *rpoC1*, *rps12*, and *rps16*) and seven tRNA genes (*trnA*-UGC, *trnG*-UCC, *trnI*-GAU, *trnK*-UUU, *trnL*-UAA, *trnL*-UAG, and *trnV*-UAC) harbor a single intron, and

CONTACT Qien Li  qienli@outlook.com  Tibetan Medical College, Qinghai University, 251 Ningda Road, Xining, Qinghai Province 810016, People's Republic of China; Zhiwei Xu  hackbowl@163.com  Qinghai Provincial Hospital of Traditional Chinese Medicine, 338 Qiyi Road, Chengzhong District, Xining 810000, Qinghai Province, People's Republic of China

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

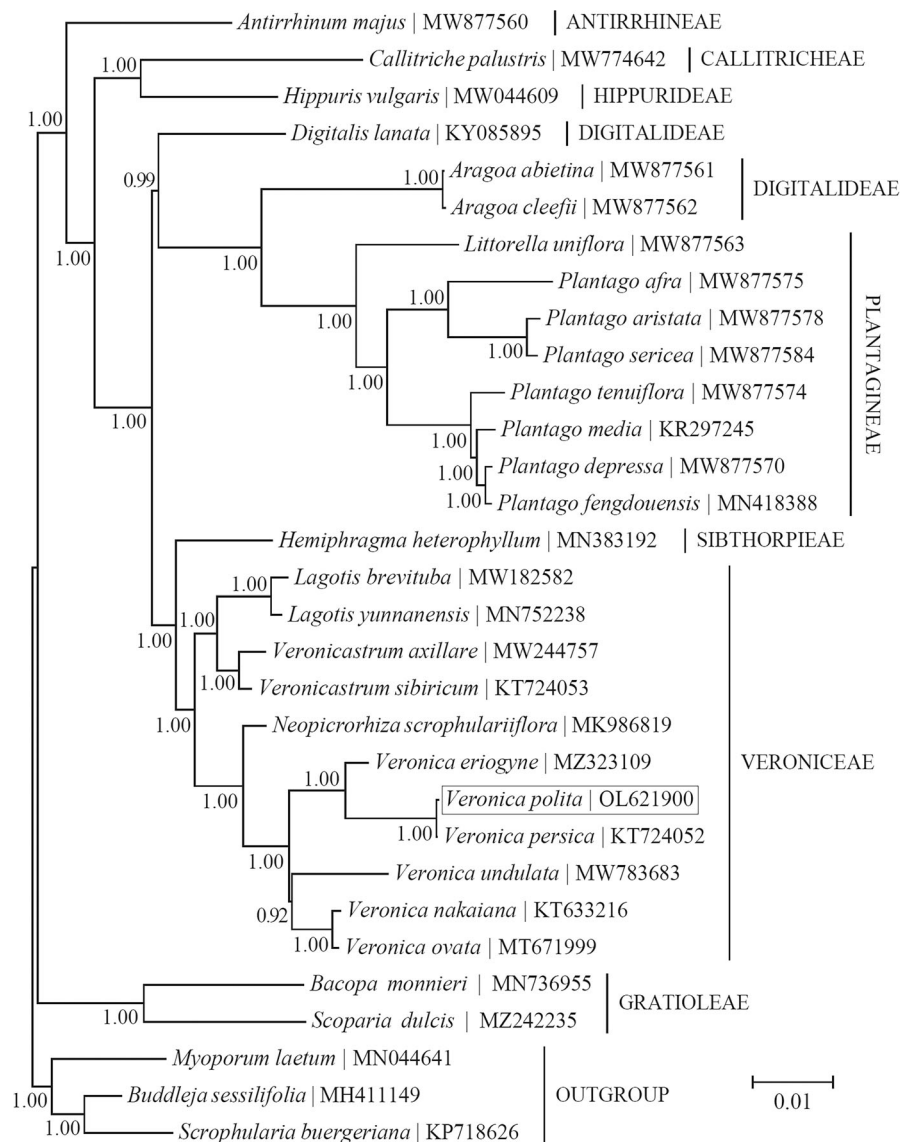


Figure 1. Phylogeny of the family Plantaginaceae based on the Bayesian analysis of the concatenated coding sequences of chloroplast PCGs. The best-fit nucleotide substitution model is 'GTR + G+I'. The support values next to the nodes were Bayesian posterior probabilities according to the Bayesian analysis. The outgroup taxa included were three species from the family Scrophulariaceae (order Lamiales). Tribal-level taxonomy was presented for each taxon in capital letters.

double introns occur in a couple of PCGs (*clpP* and *ycf3*). In addition, four microsatellite loci ((A)₁₆, (A)₁₇, (AT)₇, and (TTTTG)₃) were detected in the cp genome using SciRoKo v3.4 (Kofler et al. 2007) with default parameters.

Phylogenetic analysis was conducted based on the Bayesian analysis of cp PCGs for a panel of 28 taxa within the family Plantaginaceae using MrBayes v3.1.1 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003) (Figure 1). As suggested by the 'Model Selection' function of TOPALI v2.5 (Milne et al. 2009), <GTR + G+I> was implemented as the best-fit nucleotide substitution. The outgroup taxa used in this study were three species from the family Scrophulariaceae (order Lamiales), i.e. *Buddleja sessilifolia* B.S. Sun ex S.Y. Bao 1983 (MH411149) (Ge et al. 2018), *Myoporum laetum* G. Forst. 1786 (MN044641) (Fowler et al. 2020), and *Scrophularia buergeriana* Miq. 1865 (KP718626) (Yi and Kim 2016). The phylogenetic analysis appears to support the tribal-level taxonomy of the family Plantaginaceae. The 11 taxa belonging to the tribe Veroniceae were clustered together,

and *V. polita* was found to be most closely related to the congener *V. persica*.

Ethical approval

The materials used in this study are not included in the IUCN Red List of Threatened Species or the List of State-protected Plant Species, and the sampling site is not located in any protected area. The field study and laboratory study were conducted in accordance with guidelines provided by Qinghai Provincial Hospital of Traditional Chinese Medicine and Qinghai University.

Author contributions

Zhiwei Xu and Qien Li were involved in the conception and design; Shouning Jia, Wenjuan Chen, Guofu Zhao, and Wangshuang Xi analyzed and interpreted the data; Shouning

Jia drafted the paper; Shouning Jia, Zhiwei Xu, and Qien Li revised it critically for intellectual content. All authors approved the final version to be published, and agreed to be accountable for all aspects of the work.

Disclosure statement

The authors report no conflict of interest.

Funding

This research was supported by the Qinghai Province Chinese-Tibetan Medicine Innovation Research Project [J2020005].

Data availability statement

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 number OL621900. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA689716, SRR17062764, and SAMN23489229, respectively. The annotated chloroplast genome is also available in Zenodo at <https://doi.org/10.5281/zenodo.5746098>.

References

- Akanda MR, Nam H-H, Tian W, Islam A, Choo B-K, Park B-Y. 2018. Regulation of JAK2/STAT3 and NF- κ B signal transduction pathways; *Veronica polita* alleviates dextran sulfate sodium-induced murine colitis. *Biomed Pharmacother.* 100:296–303.
- Choi KS, Chung MG, Park S. 2016. The complete chloroplast genome sequences of three Veroniceae species (Plantaginaceae): comparative analysis and highly divergent regions. *Front Plant Sci.* 7:355.
- Fowler RM, McLay TGB, Schuster TM, Buirchell BJ, Murphy DJ, Bayly MJ. 2020. Plastid phylogenomic analysis of tribe Myoporeae (Scrophulariaceae). *Plant Syst Evol.* 306(3):52.
- Ge J, Cai L, Bi G-Q, Chen G, Sun W. 2018. Characterization of the complete chloroplast genomes of *Buddleja colvilei* and *B. sessilifolia*: implications for the taxonomy of *Buddleja* L. *Molecules.* 23(6):1248.
- Hahn C, Bachmann L, Chevreaux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* 41(13):e129.
- Hong D, Fischer MA. 1998. *Veronica* Linnaeus. *Flora Chin.* 18:65–80.
- Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics.* 17(8):754–755.
- Kofler R, Schlötterer C, Lelley T. 2007. SciRoKo: a new tool for whole genome microsatellite search and investigation. *Bioinformatics.* 23(13):1683–1685.
- Li G-F, Xu C. 2015. Study on the tolerance of *Veronica didyma* Tenore to lead. *Nor Hort.* 2015(8):70–73.
- Li J. 2003. Studies on eco-economic thresholds of *Veronica didyma* in wheat field. *J Anhui Agric Sci.* 31(6):1062–1064.
- Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF, Wright F. 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics.* 25(1):126–127.
- Nanjing University of Chinese Medicine. 2005. Dictionary of traditional Chinese medicine. 2nd ed. Shanghai: Shanghai Scientific and Technical Publishers.
- Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics.* 19(12):1572–1574.
- Schoch CL, Ciufu S, Domrachev M, Hottot CL, Kannan S, Khovanskaya R, Leipe D, McVeigh R, O'Neill K, Robbette B, et al. 2020. NCBI taxonomy: a comprehensive update on curation, resources and tools. Database. 2020:baaa062.
- Shanghai Academy of Science & Technology. 1999. Flora of Shanghai. Shanghai: Shanghai Scientific and Technical Publishers.
- Tang J, Chen J, Chen X. 2006. Response of 12 weedy species to elevated CO₂ in low-phosphorus-availability soil. *Ecol Res.* 21(5):664–670.
- Wang CZ, Jia ZJ, Liao JC. 1995. Flavonoid and iridoid glycosides from *Veronica didyma* Tenore. *Indian J Chem Sect B Org Chem Incl Med Chem.* 34(10):914–916.
- Yang W-Q, Kou J-C, Liu B. 2006. Studies on the turf feature of *Veronica didyma*. *Grassl Turf.* 2006(1):54–59.
- Yi D-K, Kim K-J. 2016. The two complete plastomes from *Scrophularia* (Scrophulariaceae): *Scrophularia buergeriana* and *S. takesimensis*. *Mitochondrial DNA B Resour.* 1(1):710–712.
- Zhang J, Yue J, Mu X, Yuan L, Zhang R, Xu M. 2006. Preliminary study of allelopathy mechanism of *Veronica didyma* Tenore. *Chin Agric Sci Bull.* 22(11):151–153.
- Zhang Y, Li B, An Y, Li G. 2010. Establishment of regeneration system for *Veronica didyma*. *J Northeast Univ.* 38(12):46–48.
- Zhong L-G, Wang C-H. 2010. Research on the biological characteristics and hazards of the exotic species *Veronica didyma*. *J Anhui Agric Sci.* 38(19):10113–10115.