

## The complete chloroplast genome sequence of the medicinal plant *Persicaria filiformis*

Kun Zhang, Zhihui Gao, Zhiping Han and Xiaofei Shan

College of Life Sciences, Shanxi Datong University, Datong, China

### ABSTRACT

*Persicaria filiformis*, known as a traditional Chinese herbal medicine, is a perennial plant of Polygonaceae widely distributed in China. The complete chloroplast (cp) genome of *P. filiformis* was assembled and analyzed in this study. The length of the circular genome is 159,741 bp, with a rich GC content of 41.3%. The cp genome structure consists of a large single-copy region (LSC 84,432 bp), a small single-copy region (SSC 13,073 bp) and a pair of inverted repeat regions (IR 31,118 bp). The complete genome encodes 130 genes, including 85 protein-coding genes, 37 tRNA genes and 8 rRNA genes. Phylogenetic analysis indicated that *P. filiformis* is most related to *P. japonica*.

### ARTICLE HISTORY

Received 25 May 2021  
Accepted 26 June 2021

### KEYWORDS



*Persicaria filiformis*;  
Polygonaceae; chloroplast  
genome; illumina  
sequencing

*Persicaria filiformis* (Thunb.) Nakai is a specie of genus *Persicaria* in the plant family Polygonaceae represented by 50 genera and 1150 species around the world (Wang et al. 2015). The genus *Persicaria* includes three species and distributes in the East Asia and North America. Only one species and two varieties are originated in China (Ma et al. 2017). *Persicaria filiformis* is endemic to China, and scattered in the mountain forest margin in Guizhou, Shandong, Henan, Shanxi, Sichuan and Guangxi. In Flora of China, this species has been recognized as an *Antenoron filiforme* (Thunb.) Rob. et Vaut., which has two variants, namely *Antenoron filiforme* (Thunb.) Rob. et Vaut. var. *neofiliforme* (Nakai) and *Antenoron filiforme* (Thunb.) Rob. et Vaut. var. *kachinum* (Nieuw.), respectively. As a Chinese herbal medicine, *P. filiformis* is widely used in anti-inflammatory, analgesic and anticoagulant treatments for its significant pharmaceutical value. At present, relatively few reports have been made on the study of *P. filiformis*. The reported researches on *P. filiformis* mainly focused on phytochemical constituents (Zhao et al. 2011), pharmacological action (Huang et al. 2004; Ma et al. 2017) and phylogenetic origin (Xu et al. 2011). In recent years, the complete chloroplast (cp) genomes have been sequenced in diverse species, and comparative analysis with different close species cp genomes has provided an efficient means for revealing the genetic characteristics and evolutionary relationship of valuable species (Cai et al. 2020). The cp genome plays a more crucial role in the study of cp molecular markers development, species identification and population dynamics (Shaw et al. 2014). In the present study, we established and characterized the complete *P. filiformis* cp genome to provide valuable information for further genetic and phylogenetical studies of this plant.

The fresh leaves of *P. filiformis* were collected from Hengshan mountain in Datong, Shanxi, China (39°42'N,

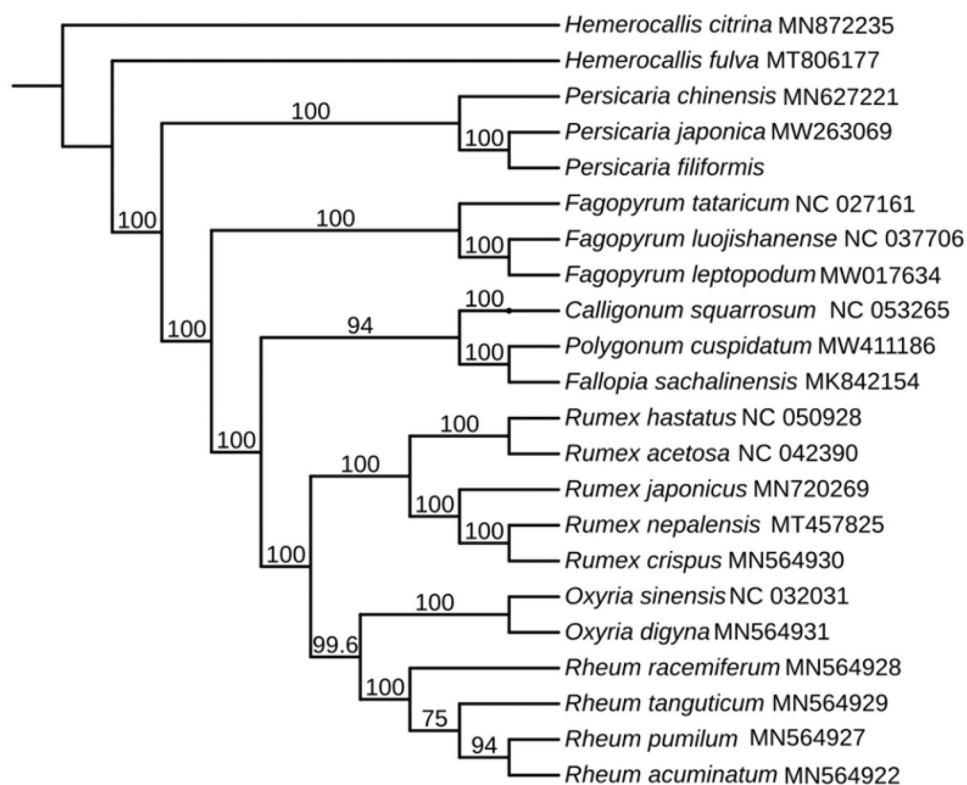
113°41'E), and the voucher specimen (Accession number: HSEW202103) was deposited in Shanxi Datong University (<http://www.sxdtu.edu.cn/>, Kun Zhang, 876828320@qq.com). The sample's total genomic DNA was extracted from about 100 mg fresh leaves using a modified CTAB method (Murray and Thompson 1980). The libraries with an average length of 350 bp were constructed using the NexteraXT DNA Library Preparation Kit (Illumina, San Diego, CA), and then sequenced on Illumina Novaseq 6000 platform. Raw sequence reads were edited using NGS QC Tool kit (Patel and Jain 2012). In total, 3.7 Gb clean data was *de novo* assembled by SPAdes v.3.11.0 software (Bankevich et al. 2012) using the cp genome of *Persicaria chinensis* (L.) H. Gross (Accession number: MW627221) as the reference. Finally, the assembled complete cp genome was annotated via PGA (Qu et al. 2019). The annotated sequence was submitted to GenBank and data were openly available at (<https://www.ncbi.nlm.nih.gov/nucleotide/MW915468.1/>) under the accession MW915468. The associated SRA number is SRR14160983.

The cp genome sequence of *P. filiformis* is 159,741 bp in length, with a quadripartite structure. Total GC content of this genome is 41.3%. The complete cp genome of *P. filiformis*, containing a large single-copy (LSC) region of 84,432 bp, a small single-copy (SSC) region of 13,073 bp and two inverted repeat (IR) regions of 31,118 bp. The cp genome encodes 130 genes, including 85 protein-coding genes, 37 tRNA genes and 8 rRNA genes. To further determine the phylogenetic position of *P. filiformis*, the complete cp genomes of other 19 species in Polygonaceae family and two species of Asphodelaceae as out-group from GenBank were aligned using MAFFT (Katoh and Standley 2013, available at <https://mafft.cbrc.jp/alignment/server/>). Subsequently, a maximum-likelihood phylogenetic tree

**CONTACT** Kun Zhang  876828320@qq.com  College of Life Sciences, Shanxi Datong University, Datong, China

© 2021 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.** Maximum-likelihood phylogenetic tree for *P. filiformis* based on 20 complete cp genome in Polygonaceae, with *Hemerocallis citrina* Baroni and *Hemerocallis fulva* (L.) L. as outgroup. The bootstrap values are located on each node and the Genbank accession numbers are shown beside the Latin name of the species.

was established by IQTREE v1.6 (Jana et al. 2016) with 1000 bootstrap replicates. The phylogenetic analysis indicated that *P. filiformis* was closed to *Persicaria japonica* (Meisn.) H. Gross ex Nakai (Figure 1). The information derived from this work may help facilitate the identification and utilization of *P. filiformis*, and further elucidate the evolution of this medicinal plant.

## Disclosure statement

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## Funding

This work was supported by Youth Science and Technology Research Fund of Applied Basic Research Program in Shanxi Province [Grant No. 201901D211426], Science and Technology Innovation Project of Higher Education in Shanxi Province [Grant No. 2019L0770], Datong Science and Technology Bureau Applied Basic Research Project [Grant No. 2020137].

## Data availability statement

The assembled complete cp genome sequence of *P. filiformis* has been submitted to GenBank of NCBI and are openly available under the accession number: MW915468 (<https://www.ncbi.nlm.nih.gov/nuccore/MW915468.1/>). The associated BioProject, SRA and Bio-Sample numbers are PRJNA720258, SRR14160983 and SAMN18644362 respectively.

## References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, et al. 2012.

- SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Cai ZP, Wang HX, Wang GX. 2020. Complete chloroplast genome sequence of *Pinellia ternata* (Thunb.) Breit, a medicinal plants to China. *Mitochondrial DNA B Resour.* 5(3):2107–2108.
- Huang YQ, Luo HM, Chen XF, Geng XZ, Mo YZ. 2004. Pharmacological effect of *Antenoron filiforme*. *Chin Tradit Patent Med.* 26(11):918–921.
- Jana T, Lam-Tung N, Arndt VH, Quang MB. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 44:W232–W235.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Ma SZ, Luan SH, Zhu LJ, Zhang X, Yao XS. 2017. Antiviral phenolics from *Antenoron filiforme* var. *neofiliforme*. *J Asian Nat Prod Res.* 20(8):1–7.
- Murray MG, Thompson WF. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8(19):4321–4326.
- Patel RK, Jain M. 2012. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. *PLOS One.* 7(2):e30619.
- Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods.* 15(1):50–61.
- Shaw J, Shafer HL, Leonard OR, Kovach MJ, Schorr M, Morris AB. 2014. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. *Am J Bot.* 101(11):1987–2004.
- Wang JR, Liu M, Ru J, Cao DL, Cheng XY, Zhang XX. 2015. The structures of tepals and fruits of Polygonaceae in China with a note on their systematic significance. *Acta Prataculturae Sin.* 24(2):116–129.
- Xu CM, Qu CY, Yu WG, Zhang XJ, Li FZ. 2011. Phylogenetic origin of *Antenoron filiforme* inferred from nuclear ITS sequences. *Guihaia.* 31(3): 300–303.
- Zhao YX, Li HF, Ma QY, Li N, Huang SZ. 2011. Studies on the chemical constituents of the herb of *Antenoron filiforme*. *Zhong Yao Cai.* 34(5): 704–707.