Characterization of the complete chloroplast genome sequence of medicinal plant: *Potentilla bifurca* (Rosaceae)

Yi Liu\(^a\), Shengxia Xiang\(^b\) and Xi Fu\(^a\)

\(^a\)Hospital of Chengdu, University of Traditional Chinese Medicine, Chengdu, Sichuan, P. R. China; \(^b\)Department of Oncology, Sichuan Integrative Medicine Hospital, Chengdu, Sichuan, P. R. China

**ABSTRACT**

*Potentilla bifurca* is a medicinal plant, the root extracts have been applied for the treatment of certain viral infections as folk medicinal herbs. The complete chloroplast genome sequence of *P. bifurca* has been assembled, containing 128 complete genes, including 84 protein-coding genes (84 PCGs), 8 ribosomal RNA genes (4 rRNAs), and 37 tRNA genes (37 tRNAs). The overall GC content of cp DNA is 37.1%, the corresponding values of the LSC, SSC, and IR regions are 35.0%, 31.0%, and 42.8%. Phylogenetic tree shows that *P. bifurca* was identified as the basal clade of *Potentilla*.

*Potentilla bifurca* Linnaeus belongs to genus *Potentilla* (Rosaceae), which is distributed in temperate and alpine zones of the Northern hemisphere. *Potentilla bifurca* has been known for its decorative value and curative properties (Eriksson et al. 1998). Extracts from the aerial and underground parts have been widely applied in Chinese traditional medicine. It is applied externally to treat antioxidant, antitumor and antiulcerogenic diseases (Gürbüz et al. 2005; Tomczyk et al. 2010; Tomczyk et al. 2013). In Tibet, *P. bifurca* root extracts are also used for the treatment of certain viral infections as folk medicinal herbs (Zhao et al. 2008). However, the chloroplast genome of *P. bifurca* has not been reported despite its important medical value. In this study, we assembled the complete chloroplast genome of *P. bifurca*, hoping to lay a foundation for further research.

Fresh leaves of *P. bifurca* were collected from Xiji (Guyuan, Ningxia, China; coordinates: 105°54'E, 35°50'N) and dried with silica gel. The voucher specimen was stored in Sichuan University Herbarium with the number is QTPLJQ14383036. Total genomic DNA was extracted with a modified CTAB method (Doyle and Doyle 1987) and a 350-bp library was constructed. This library was sequenced on the Illumina NovaSeq 6000 system with 150 bp paired-end reads. We obtained 10 million high quality pair-end reads for *P. bifurca*, and after removing the adapters, the remained reads were used to assemble the complete chloroplast genome by NOVOPlasty (Dierckxsens et al. 2017). The complete chloroplasts genome sequence of *D. longicalyx* was used as a reference. Plann v1.1 (Huang and CronK 2015) and Geneious v11.0.3 (Kearse et al. 2012) were used to annotate the chloroplasts genome and correct the annotation.

The total plastome length of *P. bifurca* (MT712015) is 155,841 bp, exhibits a typical quadripartite structural organization, consisting of a large single copy (LSC) region of 85,128 bp, two inverted repeat (IR) regions of 26,061 bp and a small single copy (SSC) region of 18,591 bp. The cp genome contains 129 complete genes, including 84 protein-coding genes, 8 ribosomal RNA genes, and 37 tRNA genes. Most genes occur in a single copy, while 16 genes occur in double, including all rRNAs (4.5S, 5S, 16S, and 23S rRNA), 7 tRNAs (trnA-UGC, trnI-GAU, trnL-CAA, trnN-GUU, trnR-ACG, and trnV-GAC), and 5 PCGs (psp7, ndhB, ycf2, rpl2 and rpl23). The overall GC content of cp DNA is 37.1%, the corresponding values of the LSC, SSC, and IR regions are 35.0%, 31.0%, and 42.8%.

In order to further clarify the phylogenetic position of *P. bifurca*, plastome of 8 representative *Potentilla* species were obtained from NCBI to reconstruct the plastome phylogeny, with *Fragaria viridis* as an outgroup. All the sequences were aligned using MAFFT v.7.313 (Katoh and Standley 2013) and maximum likelihood phylogenetic analyses were conducted using RAxML v.8.2.11 (Stamatakis 2014) under GTRCAT model with 500 bootstrap replicates. The phylogenetic tree shows that *P. bifurca* was the basal clade of *Potentilla* (Figure 1).

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Data availability statement**

The raw sequencing data were deposited in the NCBI Sequence Read Archive (SRA) database, the accession number is SAMN16562210. The
final genome assembly are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov, the reference number is MT712015.

References

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 19:11–15.

Eriksson T, Donoghue MJ, Hibbs MS. 1998. Phylogenetic analysis of Potentilla using DNA sequences of nuclear ribosomal internal transcribed spacers (ITS), and implications for the classification of Rosoideae (Rosaceae). Plant Syst Evol. 211(3–4):155–179.

Gürbüz I, Özkán AM, Yesilada E, Kutsal O. 2005. Anti-ulcerogenic activity of some plants used in folk medicine of Pinarbasi (Kayseri, Turkey). J Ethnopharmacol. 101(1-3):313–318.

Huang DI, Cronk QCB. 2015. Plann: a command-line application for annotating plastome sequences. Appl Plant Sci. 3(8):1500026.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.

Kearse M, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12):1647–1649.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.

Tomczyk M, Paduch R, Wiater A, Pleszczynska M, Kandefer-Szersze M, Szczerak J. 2013. The influence of aqueous extracts of selected Potentilla species on normal human colon cells. Acta Pol Pharm. 70(3):523–531.

Zhao YL, Cai GM, Hong X, Shan LM, Xiao XK. 2008. Anti-hepatitis B virus activities of triterpenoid saponin compound from Potentilla anserine L. Phytomedicine. 15(4):253–258.