Characterization of the complete chloroplast genome of traditional Tibetan herb, *Rheum Pumilum* Maxim. (Polygonaceae)

Rui Li\(^a\), Xiaofeng Zhang\(^a\), Jiuli Wang\(^c\), Dangwei Zhou\(^b\), Huan Wang\(^b\), Shengbo Shi\(^b\) and Tingfeng Cheng\(^b\)

\(^a\)College of Medicine, Xi’an International University, Shaanxi, P. R. China; \(^b\)Key Laboratory of Adaptation and Evolution of Plateau Biota (AEPB), Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Qinghai, P. R. China; \(^c\)College of Ecological Environment and Resources, Qinghai Nationalities University, Xining, China

**ABSTRACT**

The complete chloroplast genome sequence of traditional Tibetan herb, *Rheum pumilum* Maxim. was analyzed. The complete chloroplast genome of *R. pumilum* is 162,213 bp in size and has a 27.27% GC content. In the typical circular quadripartite structure, there was a pair of inverted repeat (IR) regions with 31,023 bp in length, which separated by a large single-copy (LSC) region (87,424 bp) and a small single-copy (SSC) region (12,743 bp). The chloroplast genome of *R. pumilum* contained 131 unigenes, which was composed of 86 protein-coding genes, 37 tRNA and 8 rRNA genes. Moreover, 238 SSRs were identified and 58.8% of them existed in LSC region. A maximum likelihood (ML) phylogenetic analysis based on chloroplast genomes indicated that *R. pumilum* was closely related to *R. palmatum*, *R. tanguticum*, and *R. officinale*. Our results would provide a valuable resource for resource utilization and the phylogenetic studies of Rheum in Polygonaceae.

*Rheum pumilum* Maxim. (XiaoDahuang in Chinese) is an alpine perennial herb, belongs to Polygonaceae family. Unlike *Rheum palmatum*, *Rheum tanguticum*, and *Rheum officinale*, its leaves appear ovule or wide ovule and shoot is dwarf (5–25 cm). This species distributes in sand soil, alpine meadow or shrub from 3000 m to 4700 m on the Qinghai-Tibet Plateau (Liu 1996), and the stout roots also have great medicinal values (Wang et al. 1988; Zhang 2004; Dong et al. 2016). Although complete chloroplast (cp) genomes of *R. palmatum*, *R. tanguticum*, and *R. officinale* have been studied (Fan et al. 2016; Zhou et al. 2018), there is no record about this species till now. In this paper, we sequenced and assembled the complete chloroplast of *R. pumilum* using Illumina Hiseq platform. The cp genome was annotated and submitted to the Genbank with the accession number (MN652917).

We collected fresh leaves from a single individual of this species from Daban Mountain (N37.34°, E101.40°; Alt. 3940 m), Menyuan County, Qinghai, China, and dried leaves with silica gel. Voucher specimens were deposited in the herbarium of Northwest Institute of Plateau Biology, CAS (HNWP, Zhou2019021). Total DNA was extracted from the fresh leaves with the DNeasy Plant MiniKit (QIAGEN, CA, USA) according to the manufacturer’s instructions. DNA quality was assessed based on spectrophotometry and electrophoresis in 1% (w/v) agarose gel, and then the good integrity and purity DNA was used for library construction and sequencing with the Illumina Hiseq platform (San Diego CA, USA) at Genepioneer Biotechnologies Inc., Nanjing, China.

In total, we obtained about 21,341,719 high quality clean reads. The cp genome was assembly using NovoPlasty software (Dierckxsens et al. 2017) and the previously published cp genome of *R. palmatum* (Fan et al. 2016) was used as seed reference. We visualized the genome by Geneious version 8.05 (Kearse et al. 2012). Gene annotation firstly perform with DOGMA (Wyman et al. 2004) and CpGAVAS (Liu et al. 2012), then corrected manually with the Geneious (Kearse et al. 2012). Finally, the physical map of cp genome of *R. pumilum* was done with CpGAVAS (Liu et al. 2012). The complete cp genome sequence and its annotations were submitted to Genbank (MN652917).

The complete chloroplast genome of *R. pumilum* is 162,213 bp in size and has a 27.27% GC content. In the typical circular quadripartite structure, there was a pair of inverted repeat (IR) regions with 31,023 bp in length, which separated by a large single-copy (LSC) region (87,424 bp) and a small single-copy (SSC) region (12,743 bp). The chloroplast genome of *R. pumilum* contained 131 unigenes, which was composed of 86 protein-coding genes, 37 tRNA and 8 rRNA genes. Moreover, 238 SSRs were identified and 58.8% of them existed in LSC region. Compared the four cp gnome of *R. pumilum* was closely related to *R. palmatum*, *R. tanguticum*, and *R. officinale*. Our results would provide a valuable resource for resource utilization and the phylogenetic studies of Rheum in Polygonaceae.
To identified the phylogenetic relationship of *R. pumilum* and other 24 species, 65 protein coding genes extracted from the chloroplast genome of 16 taxa, which including 6 species of Polygnonaceae and species of *Amborella trichopoda* in Amborellacea was set as the outgroup. Sequences were aligned using the program MAFFT (Katoh and Standley 2013), and maximum likelihood (ML) analysis was performed by RaxML based on Kimura 2-parameter model with 1000 bootstrap replicates (Alexandros et al. 2017). The phylogenetic tree displayed that *R. pumilum* were closer clustered with *R. palmatum*, *R. tanguticum*, and *R. officinale* while *R. witrochii* is out of them clearly (Figure 1). Our cp genome data of *R. pumilum* would facilitate population, genetic identification and cp genetic engineering research of this traditional Tibetan herb in the future.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by grants from the Foundation of Science in Qinghai [2017-ZJ-702; 2019-ZJ-9760], and the Construction Project for Innovation Platform of Qinghai Province [2017-ZJ-Y14].

**ORCID**

Jiuli Wang  
http://orcid.org/0000-0002-7426-8101

Dangwei Zhou  
http://orcid.org/0000-0003-4126-614X

**References**

Alexandros S, Paul H, Jacques R. 2017. A rapid bootstrap algorithm for the RAxML Web servers. System Biol. 57:562.

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

Dong HJ, Chen XH, Zeng R. 2016. Rapid analysis on chemical constituents in roots of *Rheum pumilum* by UPLC coupled with hybrid quadrupole-orbit trap MS. Chin Tradit Herb Drugs. 47(4):2428–2435.

Fan K, Sun XJ, Huang M, Wang XM. 2016. The complete chloroplast genome sequence of the medicinal plant *Rheum palmatum* L. (Polygonaceae). Mitochondrial DNA Part A. 27:2935–2936.

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, Moir R, 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.
Liu C, Shi L, Zhu Y, Chen H, Zhang J, Lin X, Guan X. 2012. CpGAVAS, an integrated web server for the annotation, visualization, analysis, and GenBank submission of completely sequenced chloroplast genome sequences. BMC Genomics. 13(1):715.

Liu SW. 1996. Qinghai flora. Xi’ning (China): Qinghai People’s Publishing house.

Wang AJ, Zhu YQ, Guo JX. 1988. Pharmacognosy identification of the three Tibetan Rheum (Polygonaceae). Chin Tradit Herbal Drugs. 8:27–29.

Wyman SK, Jansen RK, Boore J. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics. 20(17):3252–3255.

Zhang XF. 2004. Xining Dahuang (history, resources and utilization). Xi’an (China): Shaanxi’Sci&Tech Publishing house.

Zhou YX, Nie J, Xiao L, Hu ZG, Wang B. 2018. Comparative chloroplast genome analysis of Rhubarb botanical origins and the development of specific identification markers. Molecules. 23(11): 2811.