Characterization of the complete chloroplast genome of traditional Chinese herb, *Solanum japonense* Nakai. (Solanaceae)

Xiaoxia Zhang\(^a\), Xiaofeng Zhang\(^a\), Huan Wang\(^b\), Shilong Chen\(^b\), Tingfeng Cheng\(^b\), Shengbo Shi\(^b\) and Dangwei Zhou\(^b\)

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

**ABSTRACT**

*Solanum japonense* Nakai. (Solanaceae) is a traditional Chinese medicine and animal fodder in Asian continent. Here, the complete chloroplast genome sequence of *S. japonense* was determined by the Illumina HiSeq technology. The complete chloroplast genome of *S. japonense* was 155,415 bp and the GC content was 37.81%. The typical circular quadripartite structure was composed with two inverted repeat (IR) regions with 25,588 bp, a large single-copy (LSC) region (85,931 bp) and a small single-copy (SSC) region (18,344 bp). The chloroplast genome of *S. japonense* contained 132 unigenes, which contained 86 protein-coding genes, 37 tRNA and 8 rRNA genes and one pseudo-gene. Moreover, 183 SSRs were identified and 65% (119) of them located at LSC region. A Maximum-Likelihood (ML) phylogenetic analysis based on chloroplast genomes indicated that *S. japonense* was closely related to *S. dulcamara*, *S. nigrum*. Our results would provide a valuable resource for resource utilization and the phylogenetic studies of species of Dulcamara sect. in Solanaceae.

*S. japonense* Nakai. (Yehaigie in Chinese) is a widely perennial herb, belongs to Solanaceae family, which distributes in wastelands, slopes or forests ranged 300 m to 2900 m in China (Flora in China 1979). This species and *Solanum dulcamara* both belong to Sect. Dulcamara (Moench), while the former has deltate, broadly lanceolate or ovate-lanceolate leaves, often with 2 basal lobes (Flora in China 1979). As its vine, flower and fruits were similar to *S. dulcamara*, it is often confused with *S. dulcamara* and to treat rheumatism and skin diseases in Asia and India (Li 2004; Kumar et al. 2009; Yuan et al. 2020). Solanaceae has various alkaloids and is an important medicine resource. Recently, the study also showed *S. dulcamara* and other related species were abundant in terpenoids and has various pharmacological activities (Yuan et al. 2020). As the important wild plant resource, *S. japonense* taxon and pharmacological traits have been paid attention (Flora in China 1979; Li 2004), however there still known limit on chloroplast genome of this species till now. In this paper, we sequenced and assembled the complete chloroplast of *S. japonense* using Illumina Hiseq platform. The cp genome was annotated and submitted to the Genbank (Accession number: MW077727).

We collected fresh leaves from a single individual of this species from Minhe County (102.733\(^\circ\), 36.143\(^\circ\), 2809 m), Qinghai, China, and dried leaves with silica gel. Voucher specimens were deposited in the herbarium of Northwest Institute of Plateau Biology, CAS (HNWP, Zhou2020024). 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. The average insertion of the library was 320 bp and whole genome sequencing with 150 bp pair-end reads using the Illumina Hiseq platform (San Diego CA, USA) at Genepioneer Biotechnologies Inc., Nanjing, China.

In total, we obtained about 17,793,213 high quality clean reads and QC30 value was 93.96%. The cp genome was assembly using NovoPlasty software (Dierckxsens et al. 2016) and the previously published cp genome of *S. dulcamara* (Amiryousefi et al. 2018) 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 *S. japonense* was done with CpgAVAS (Liu et al. 2012). The complete cp genome sequence and its annotations were submitted to Genbank (MW077727).

The complete chloroplast genome of *S. japonense* is 155,415 bp in size and the average GC content is 37.81%.
In the typical circular quadripartite structure, there was a pair of inverted repeat (IR) regions with 25,588 bp in length, which separated by a large single-copy (LSC) region (85,931 bp) and a small single-copy (SSC) region (18,344 bp). The chloroplast genome of *S. japonense* contained 132 unigenes, which was composed of 86 protein-coding genes, 37 tRNA and 8 rRNA genes and one psedo-gene. Moreover, 183 SSRs were identified and 65% (119) of them existed in LSC region. Compared the four cp genome of Solanaceae, most genes were conserved and the variation displayed in the intergeneric regions.

To identified the phylogenetic relationship of *S. japonense* and other 39 species of Solanaceae, the whole chloroplast genome of them were aligned with MAFFT software (Katoh and Standley 2013), and species of *Scrophularia dentata* was set as the outgroup species. Maximum -Likelihood (ML) phylogenetic analysis was conducted adopting HKY85 model by PhyML3.0 (Guindon et al. 2010) with 100 bootstrap replicates. The phylogenetic tree displayed that *S. japonense* was closer clustered with *S. dulcamara* while *S.nigrum* was out of them clearly (Figure 1). Our cp genome data of *S. japonense* would facilitate population, genetic identification and cp genic engineering research of this traditional Chinese 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-976Q], and the Construction Project for Innovation Platform of Qinghai Province [2017-ZJ-Y14].

**Data availability statement**
*Solanum japonense* N cp genome data was submitted to Genbank (accession number: MW077727).

**References**
Amiryousefi A, Hyvönen J, Poczai P. 2018. The chloroplast genome sequence of bittersweet (*Solanum dulcamara*): plastid genome structure evolution in Solanaceae. PLoS One. 13(4):e0196069.
Dierckxsens N, Mardulyn P, Smits G. 2016. Novoplasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45:e18. Flora in China. 1979. Delectis Florae Reipublicae Popularis Sinicae Agendae Academiae Sinicae Edita. Beijing: Science Press.
Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate Maximum-Likelihood
Phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3): 307–321.
Kumar P, Sharma B, Bakshi N. 2009. Biological activity of alkaloids from Solanum dulcamara L. Nat Prod Res. 23(8):719–723.
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.
Li SZ. 2004. Compendium of Materia Medica. Beijing: Foreign Languages Press.
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.
Wyman SK, Jansen RK, Boore J. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics. 20(17):3252–3255.
Yuan PL, Fan QL, Li YJ, Lv GY. 2020. Advance study of sesquiterpenes and their pharmacological activity on 13 species of Solanaceae. Chin Tradit Patent Med. 42(2):439–445.