MITOCHONDRIAL DNA PART B
2019, VOL. 4, NO. 2, 4163–4165
https://doi.org/10.1080/23802359.2019.1693301

Taylor & Francis
Taylor & Francis Group

MITOGENOME ANNOUNCEMENT

Complete chloroplast genome sequence of *Pueraria thomsonii*, an important traditional Chinese medicine plant

Xiao-Rong Miao[a,b]*, Jun-Qi Niu[a,b]*, Ai-Qin Wang[a,b], Dao-Bo Wang[a,b] and Jing Fan[c]

[a]College of Biology and Pharmacy, Yulin Normal University, Yulin, China; [b]Guangxi Key Laboratory of Agricultural Resources Chemistry and Biotechnology, Yulin, China; [c]College of Life Sciences, Leshan Normal University, Leshan, China

ABSTRACT

Pueraria thomsonii is a leguminous plant with high root yield and starch content. It is also a medicinal material in the Chinese pharmacopeia. However, the raw materials of *P. thomsonii* are often confused with some non-medicinal Pueraria plants. To enrich the genetic resources of *P. thomsonii* and guide its molecular identification, the complete chloroplast genome was sequenced and reported. The total genome of *P. thomsonii* is 153,434 bp in length. consisting of two inverted repeat regions (IR_S, 25,640 bp each) separated by a large single-copy (LSC, 84,155 bp) and a small single-copy region (SSC, 17,999 bp). The overall GC content is 35.41%. It contains 130 genes, including 85 protein coding genes, 8 rRNA genes and 37 tRNA genes. Phylogenetic analysis showed that *P. thomsonii* could be distinguished from other plants and closely related to the legume *Pachyrhizus erosus*. This study enriches the genetic information of *P. thomsonii* and contributes to the screening of excellent germplasm.

ARTICLE HISTORY
Received 31 October 2019
Accepted 9 November 2019

KEYWORDS
Pueraria thomsonii; complete chloroplast genome; legume; illumina sequencing

Pueraria plant is a perennial vine of the papilionaceae family. It is reported that there are nearly 20 species in the world (Egan et al. 2016), among which there are 9 species and 2 variants of Pueraria in China, but only the *Pueraria lobata* (Willd.) Ohwi. and *Pueraria thomsonii* Benth were used as medicinal materials and archived in the Chinese pharmacopeia (Zhao et al. 2011; Wang et al. 2018). Due to the scarcity of wild *P. lobata* resources, *P. thomsonii* can also produce isoflavones and is used to prevent various chronic diseases, it is widely grown to supply medicinal materials (Wong et al. 2015; Liang et al. 2017). However, non-medicinal plants of the genus Pueraria often appear in the market, therefore, the accurate identification of *P. thomsonii* can guarantee the source and quality of pueraria. Chloroplast genome has been widely used in plant evolution and taxonomy due to its maternal inheritance and conserved structure (Fan and Huang 2019). However, the complete chloroplast genome of *P. thomsonii* has not been reported. In this work, the chloroplast genome sequence of *P. thomsonii* was decrypted for the first time, which can better understand the genetic background of Pueraria and provide a basis for species identification of *P. thomsonii*.

The leaves of *P. thomsonii* were collected from the medicinal botanical garden of Guangxi university in Nanning, Guangxi (108°33′45″E, 22°82′13″N), the specimen (YS20171014) was stored in the herbarium of Yulin normal university. Total DNA of *P. thomsonii* was extracted from fresh leaves by SDS method, and sequenced by Illumina HiSeqXten platform. The obtained genome sequences were de novo assembled using SPAdesv.3.11.0 (Bankevich et al. 2012), low-quality reads and adapters were removed by FastQC software (Andrews 2010), and finally annotated by Plann software (Huang and Cronk 2015). The chloroplast genome size of *P. thomsonii* (GenBank accession no. MN515038) is 153,434 bp and has four sub-regions: a large single copy (LSC) of 84,155 bp and a small single copy region (SSC) of 17,999 bp, which are separated by two inverted repeats (IRs) of 25,640 bp. It contains 130 genes, including 85 protein-coding genes, 37 tRNA genes and 8 rRNAs, 18 genes (7 Protein coding genes, 4 rRNAs, and 7 tRNAs) are duplicated in the inverted repeat regions, the total GC content is 35.41%.

To analyze the phylogeny of *P. thomsonii*, 39 complete chloroplast genomes were aligned using MAFFT7.037 software (Katoh and Standley 2013), then the Maximum Likelihood (ML) phylogenetic tree was constructed by Mega-X v10.0.5 software (Kumar et al. 2018), with the operating parameters of GTR + G model and 1000 bootstrap replicates. The result of ML phylogenetic tree showed that *P. thomsonii* is closely related to *Pachyrhizus erosus* (Figure 1). Our results

CONTACT Dao-Bo Wang d.b.wang@foxmail.com College of Biology and Pharmacy, Yulin Normal University, No. 1303, East Jiaoyu Road, Yuzhou District, Yulin 537000, Guangxi Province, P.R. China. Jing Fan fanjing972001@126.com College of Life Sciences, Leshan normal University, No. 778, Binhe Road, Shizhong District, Leshan 614000, Sichuan province, P.R. China
*Xiao-Rong Miao and Jun-Qi Niu contributed equally to this work.

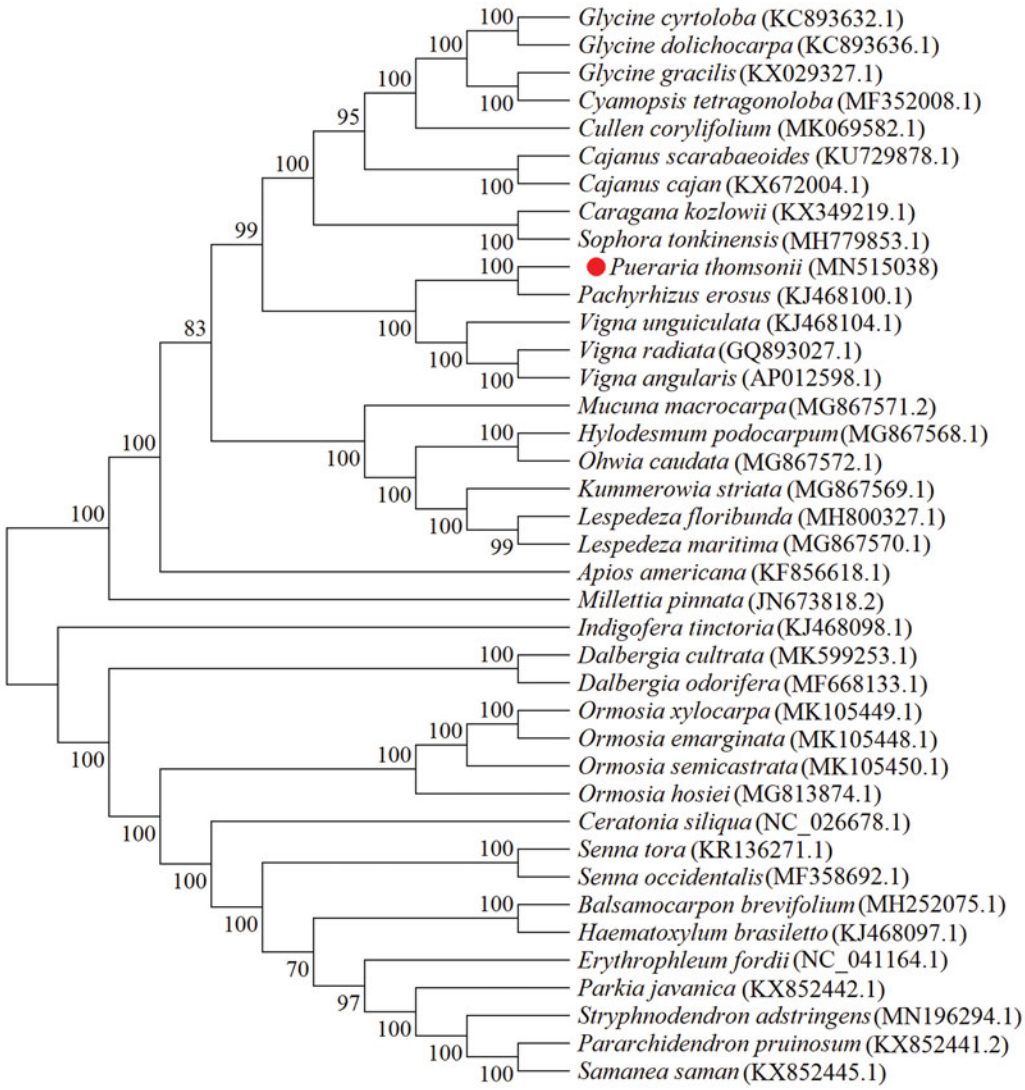

Figure 1. The Maximum Likelihood (ML) phylogenetic tree based on the 39 plant chloroplast genome. Note: The number near each node represents the support value of 1000 bootstrap replicates.

provide useful resources for molecular identification and phylogeny of *P. thomsonii*.

Disclosure statement

The authors confirm this article content has no conflict of interest, and all the authors are responsible for the content of this article.

Funding

This work was supported by the National Natural Science Foundation of China [31860403] and Guangxi Natural Science Foundation projects [2017GXNSFBA198023].

ORCID

Jing Fan http://orcid.org/0000-0001-8295-5772

References

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 19(5):455–477.

Egan AN, Vatanparast M, Cagle W. 2016. Parsing polyphyletic Pueraria: delimiting distinct evolutionary lineages through phylogeny. Mol Phylogenet Evol. 104:44–59.

Fan J, Huang MY. 2019. Chloroplast genome structure and phylogeny of *Spiranthes sinensis*, an endangered medicinal orchid plant. Mitochondrial DNA B. 4(2):2994–2996.

Huang DI, Cronk QC. 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.

Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 35(6):1547–1549.

Liang J, Maeda T, Tao XL, Wu YH, Tang HJ. 2017. Physicochemical properties of pueraria root starches and their effect on the improvement of buckwheat noodles quality. Cereal Chem. 94(3):554–559.

Wang Y, Yang Y, Jiao J, Wu Z, Yang M. 2018. Support vector regression approach to predict the design space for the extraction process of *Pueraria lobata*. Molecules. 23(10):2405.

Wong KH, Razmovski-Naumovski V, Li KM, Li GQ, Chan K. 2015. Comparing morphological, chemical and anti-diabetic characteristics of *Puerariae Lobatae* Radix and *Puerariae Thomsonii* Radix. J Ethnopharmacol. 164:53–63.

Zhao C, Chan HY, Yuan D, Liang Y, Lau TY, Chau FT. 2011. Rapid simultaneous determination of major isoflavones of *Pueraria lobata* and discriminative analysis of its geographical origins by principal component analysis. Phytochem Anal. 22(6):503–508.