


## The complete chloroplast genome sequence of the *Pueraria lobata* (Willd.) Ohwi (Leguminosae)

Tianxu Cao, Xiling Ma, Yu Zhang, Wenzheng Su, Bicong Li, Qinghong Zhou and Qianglong Zhu 

Department of Horticulture, College of Agronomy, Jiangxi Agricultural University, Nanchang, PR China

### ABSTRACT

*Pueraria lobata* (Willd.) Ohwi is an essential traditional oriental medicine with therapeutic effects. In this study, we assembled the complete chloroplast genome of *P. lobata*. The total genome size was 153,442 bp in length, containing a large single-copy (LSC) region of 84,162 bp, a small single-copy (SSC) of 17,998 bp, and a pair of inverted repeats (IRs) of 25,641 bp, and possessing 35.41% GC content. In addition, the whole chloroplast genome encodes a total of 129 genes, including 84 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Phylogenetic tree analysis of 48 species in the family Papilionoideae of Leguminosae indicated that *P. lobata* was belong to Papilionoideae and closely related to the genus, *Pachyrhizus*, *Vigna* and *Phaseolus*.

### ARTICLE HISTORY

Received 7 August 2020  
Accepted 7 October 2020

### KEYWORDS



*Pueraria lobata* (Willd.) Ohwi; kudzu; chloroplast genome; phylogenetic analysis

*Pueraria lobata* (Willd.) Ohwi is a wild leguminous creeper, commonly known as kudzu or Ge-gen, its swollen fleshy root possesses a high content of puerarin, daidzein, and rutin and is taken as a famous Chinese traditional medicine used to therapy muscle stiffness, common cold, influenza, cardiovascular disease, other ailments (Guerra et al. 2000; Chen et al. 2001; Guo et al. 2020), Kudzu starch has been used as an herb additive in food (Zhu et al. 2018). China is one of the main original and distribution centers of kudzu, but its genetic diversity has been seldom investigated, which seriously hinder germplasm resource conservation, new variety breeding and detection of species invasion (Sun et al. 2005; Yuan et al. 2017). Chloroplast genome was an important part of plant whole genome and play an important role in photosynthesis and other metabolic pathways in kudzu, and the genome sequence is a great molecular resource for genetic diversity, plant taxonomy, and phylogenetic relationships at different taxonomic levels (Ji et al. 2020). So far, about 296 complete chloroplast genomes of Leguminosae have been sequenced and published, but the Chloroplast genome of *P. lobata* has not been reported. Therefore, we reported the complete chloroplast genome sequence of *P. lobata* based on the high-throughput sequencing with hope to promote the studies based on the kudzu chloroplast genome.

The *P. lobata* was collected at Hengfeng, Shangrao, Jiangxi (117°6'E, 28°42'N) in May 2019, and preserved in the Yulan experimental station of Jiangxi Agricultural University, Nanchang, China (Voucher number: JAGE-01). Total genomic DNA was extracted from fresh leaves using modified CTAB (Allen et al. 2006), the high-quality DNA was sent to construct a genomic library and sequenced by BGISEQ-500

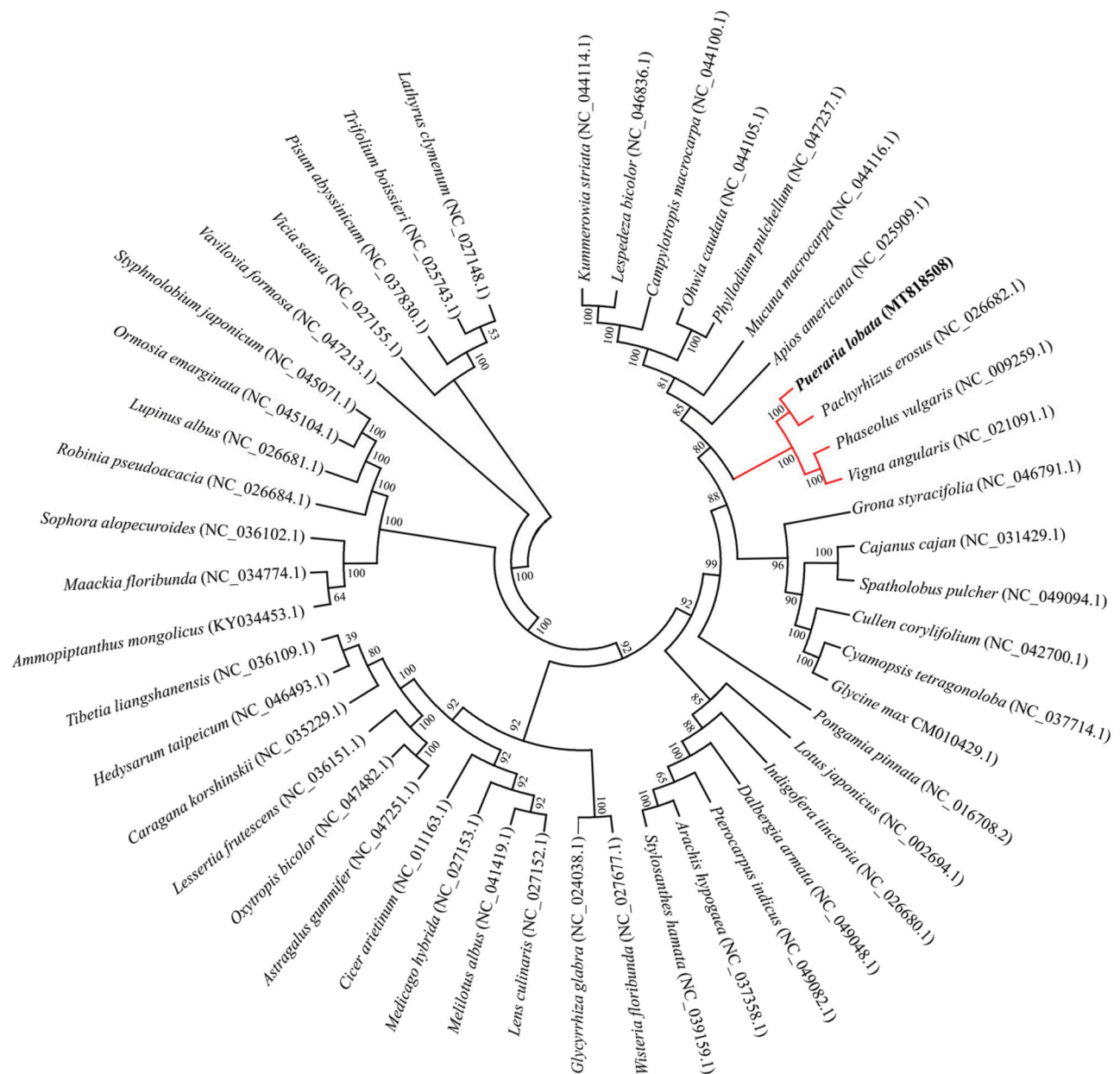
platform (BGI, Shenzhen, China). About 3.0 Gb of sequence data was obtained after sequencing and base quality control, 0.5 Gb of clean pair-end reads with length of 150 bp was randomly extracted using Seqtk. The kudzu chloroplast genome was assembled by using SPAdes (v 3.12.0) (Bankevich et al. 2012). Specific method is as follows: first, these reads were assembled with using the Plasmidspades.py in SPAdes, Contigs representing the chloroplast genome were then retrieved, ordered, and incorporated into a single draft sequence by BlastN against the chloroplast genome of *Glycine mix* (CM010429) (Shen et al. 2019), third, the gaps in the chloroplast single draft sequence of were closed by using GapCloser (v1.12-r6). Finally, the complete genome sequence was annotated using CPGAVAS2 (Shi et al. 2019), and manually checked and corrected by Sequin. The annotated genomic sequence was submitted to GenBank under accession number MT818508.

The complete chloroplast genome of *P. lobata* is 153,442 bp in length, and exhibits a typical quadripartite structure, consisting of a pair of IRs (25,641 bp) separated by the LSC (84,162 bp) and SSC (17,998 bp) regions. The LSC, SSC, IR regions and GC contents of the whole chloroplast genome are 34.7%, 28.90%, 41.88%, and 35.41%, respectively. A total of 129 genes are encoded, including 86 protein-coding genes, 8 rRNA genes, and 37 tRNA genes; 6 of the protein-coding genes, 6 of the tRNA genes, and 4 rRNA genes are duplicated within the IRs. Moreover, 11 protein-coding and 6 tRNA genes contained one intron, and 3 protein-coding genes contained two introns respectively. Noticeably, the start code of *ndhD* is by RNA editing, and an exon of *rps12* is by trans-splicing.

**CONTACT** Qianglong Zhu  [longzhu2011@126.com](mailto:longzhu2011@126.com)  Department of Horticulture, College of Agronomy, Jiangxi Agricultural University, No. 1101 Zhimin Street, Qingshanhu District, Nanchang 330045, PR China

© 2020 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.** Phylogenetic tree showing relationship between *P. lobata* and 48 species belonging to Papilionoideae in Leguminosae family. Phylogenetic tree was constructed based on 63 protein-coding genes of chloroplast genomes using maximum likelihood (ML) with 1000 bootstrap replicates. Numbers in each the node indicated the bootstrap support values.

To date, the complete chloroplast genomes of 47 genera of Papilionoideae have been reported at GenBank, to confirm the phylogenetic location of *P. lobata*, a phylogenetic analysis was conducted using common 63 chloroplast protein-coding sequences of *P. lobata* and other 47 published chloroplast genomes represented different genus. These chloroplast protein-coding sequences were aligned by the MAFFT v7.407 (Nakamura et al. 2018), the phylogenetic tree was constructed by maximum-likelihood method with 1000 bootstrap value using MEGA v10.0.4 (Kumar et al. 2018). The phylogenetic analysis results clearly showed that *P. lobata* was closer to *Pachyrhizus* (*P. erosus*), *Vigna* (*V. angularis*), and *Phaseolus* (*P. vulgaris*), but keep remote phylogenetic relationship with these genera, *Lathyrus*, *Trifolium*, and *prisum* in the family Leguminosae (Figure 1). The chloroplast genome of *P. lobata* will provide useful genetic information for

promoting the evolutionary studies and conservation of Leguminosae species.

### Disclosure statement

No potential conflict of interest was reported by the author(s).

### Funding

This work was funded by the Youth Science Foundation of Jiangxi Provincial Department of Education [GJJ180218], and the Construction Project of Root and Tuber Crops Industry and Technology System [JXARS-19-01].

### ORCID

Qianglong Zhu  <http://orcid.org/0000-0001-5355-8815>

## Data availability statement

The data that support the findings of this study are openly available in GenBank at <https://www.ncbi.nlm.nih.gov>, reference number MT818508.

## References

- Allen GC, Flores-Vergara MA, Krasynanski S, Kumar S, Thompson WF. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc.* 1(5):2320–2325.
- 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.
- Chen G, Zhang J, Jiannong Y. 2001. Determination of puerarin, daidzein and rutin in *Pueraria lobata* (Willd.) Ohwi by capillary electrophoresis with electrochemical detection. *J Chromatogr A.* 923(1–2):255–262.
- Guerra MC, Speroni E, Broccoli M, Cangini M, Pasini P, Minghetti A, Crespi-Perellino N, Mirasoli M, Cantelli-Forti G, Paolini M. 2000. Comparison between Chinese medical herb *Pueraria lobata* crude extract and its main isoflavone puerarin antioxidant properties and effects on rat liver CYP-catalysed drug metabolism. *Life Sci.* 67(24): 2997–3006.
- Guo K, Yao Y, Yang M, Li Y, Wu B, Lin X. 2020. Transcriptome sequencing and analysis reveals the molecular response to selenium stimuli in *Pueraria lobata* (Willd.) Ohwi. *PeerJ.* 8:e8768.
- Ji Y, Zhang W, Li D, Shen L. 2020. The complete chloroplast genome sequence of *Annamocarya sinensis* (Juglandaceae), an endangered species endemic to Yunnan Province, China. *Mitochondrial DNA Part B.* 5(3):2021–2023.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol E.* 35(6):1547–1549.
- Nakamura T, Yamada KD, Tomii K, Katoh K. 2018. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics.* 34(14):2490–2492.
- Shen Y, Du H, Liu Y, Ni L, Wang Z, Liang C, Tian Z. 2019. Update soybean *Zhonghuang 13* genome to a golden reference. *Sci China Life Sci.* 62(9):1257–1260.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73.
- Sun JH, Li ZC, Jewett DK, Britton KO, Ye WH, Ge XJ. 2005. Genetic diversity of *Pueraria lobata* (kudzu) and closely related taxa as revealed by inter-simple sequence repeat analysis. *Weed Res.* 45(4):255–260.
- Yuan C, Zhong WJ, Gong YY, De-Qiang PU, Pei-Cheng JI, Huang HY, Yang ZH, Zhang C. 2017. Genetic diversity and trait association analysis of *Pueraria lobata* resources. *J Plant Genet Resour.* 18(2):233–241.
- Zhu X, Du X, Chen X, Hu J, Zhou X, Guo L. 2018. Determining the effects of annealing time on the glass transition temperature of *Pueraria lobata* (Willd.) Ohwi starch. *Int J Food Sci Technol.* 53(1):43–49.