The complete chloroplast genome of *Indigofera stachyodes* (Fabaceae), a traditional Chinese medicinal plant

Na Zhang, Jin-lan Long, Yuan Wu, Yong-pin Zhang and Zhi-kun Wu

Department of Pharmacy, Guizhou University of Traditional Chinese Medicine, Guiyang, China

**ABSTRACT**

*Indigofera stachyodes* Lindl. is a traditional medicinal plant in southwestern China. In this study, we report the complete chloroplast genome sequence of *I. stachyodes*, using next-generation sequencing technology. The complete chloroplast genome of *I. stachyodes* was 158,039 bp in length with an overall GC content of 35.80%, containing a large single-copy (LSC) region of 88,772 bp, a small single-copy (SSC) region of 18,733 bp, and a pair of inverted repeats (IRs) regions of 25,267 bp. In total, there are 128 genes (83 protein-coding genes (PCGs), eight ribosomal RNA (rRNA) genes, and 37 tRNA genes) in the whole chloroplast genome, including 113 unique genes (78 unique PCGs, 31 unique tRNAs, and four unique rRNAs). The phylogenetic analysis indicated that *I. stachyodes* formed a monophyletic clade with *I. tinctoria* and *I. linifolia*, showing that they have close relationship. The complete chloroplast genome of *I. stachyodes* provides valuable genomic information for the phylogeny, molecular identification and sustainable utilization of this species.

The plant *Indigofera stachyodes* Lindley 1843, which belongs to Papilionoideae of Fabaceae, is a perennial shrub, 1–3 m tall, mainly distributed in Guizhou province, Yunnan province, and Guangxi Zhuang National Autonomous Region of southwestern China, as well as Bhutan, Cambodia, India, Indonesia, Laos, Myanmar, Nepal, Thailand, and Vietnam (Gao and Brian 2010). The roots of *I. stachyodes* are well known as Xue-ren-shen in Chinese, have been traditionally used as a traditional medicine by the Miao people to treat a wide array of human ailments, such as wounds, dysentery, cirrhosis, and rheumatism (Editorial Committee of Chinese Ben-cao 1999). A variety of biological properties including antioxidant, α-glucosidase inhibitory, and anti-inflammatory activities have been reported for crude extracts from its roots (Li et al. 2011; Qiu et al. 2013). In this study, we report the first chloroplast genome of *I. stachyodes*, which will provide valuable genomic information for the study of phylogeny, molecular identification and sustainable utilization of this species.

Fresh leaves of *I. stachyodes* were collected from a wild population (106°40′29″ E, 26°26′31″ N) in Huaxi district, Guizhou Province, China. The voucher specimen was deposited in the herbarium of Guizhou University of Traditional Chinese Medicine (Cheng-gang Hu, 2274547063@qq.com) under the voucher number ZN20201026. We isolated the DNA from the plants and sequenced the chloroplast genome using Illumina HiSeq X Ten sequencer to construct the Illumina paired-end (PE) sequencing. NOVOplasty v2.7.2 (Dierckxsens et al. 2017) was then used to assemble the complete chloroplast genome of *I. stachyodes*. We also used Geneious v 8.0.2 software to annotate the assembled chloroplast genome (Kearse et al. 2012). The annotated chloroplast genome of *I. stachyodes* was deposited into GenBank with the accession number MZ768851.

The complete chloroplast genome of *I. stachyodes* was 158,039 bp in length with a typical quadripartite structure, and an overall GC content of 35.8%. The assembled genome contained a large single-copy (LSC) region of 88,772 bp, a small single-copy (SSC) region of 18,733 bp, and a pair of inverted repeats (IRs) regions of 25,267 bp. In total, there are 128 genes (83 protein-coding genes (PCGs), eight ribosomal RNA (rRNA) genes, and 37 tRNA genes) in the whole chloroplast genome, including 113 unique genes (78 unique PCGs, 31 unique tRNAs, and four unique rRNAs, respectively).

The maximum-likelihood (ML) phylogenetic tree was constructed based on 17 complete chloroplast genomes (all coding and noncoding sequences) of Papilionoideae species and two complete chloroplast genomes of Caesalpinioideae species as outgroups, using IQ-TREE v1.6.10 (Nguyen et al. 2015) and performed based on a TVM + F+I+R2 model according to Bayesian’s information criteria using ModelFinder (Kalyaanamoorthy et al. 2017). Ultrafast bootstrap (UFBoot) was used to test branch supports (Hoang et al. 2018) and an SH-like approximate likelihood ratio with 10,000 bootstrap replicates (Figure 1). The phylogenetic analysis indicated that the plants in the same genus, *I. stachyodes*, *I. linifolia*, and *I. tinctoria*, formed a monophyletic clade with 100% bootstrap support; *I. tinctoria* formed a monophyletic clade with 99% bootstrap support.
value, showing that they have a close relationship. Within this monophyletic clade, *I. stachyodes* was most closely related to *I. tinctoria* with 99.3% protein coding sequences identical, compared 98.7% of protein coding sequences and 91.5% of non-coding sequences identical to those of *I. linifolia*. This reported *I. stachyodes* chloroplast genome will provide useful information for molecular identification of close species of *Indigofera*, and also for phylogenetic and evolutionary studies in Fabaceae.

**Acknowledgements**

Ethical approval: Research and collection of plant material were conducted according to the guidelines provided by GZY (Guizhou University of Traditional Chinese Medicine). Permission was granted by the National Natural Science Foundation of China.

**Authors contributions**

Z-KW and Y-PZ planned and designed the research. NZ, J-LL, and YW collected the plant materials, NZ performed experiments, and Z-KW analyzed the data. Z-KW and NZ wrote the manuscript.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by the National Natural Science Foundation of China [No. U1812403-2].

**Data availability statement**

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov/](https://www.ncbi.nlm.nih.gov/) under the accession no MZ768851. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA752685, SRR15367629, and SAMN20606238, respectively.

**References**

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):18.

Doyle J. 1991. DNA protocols for plants: CTAB total DNA isolation. In: Hewitt GM, Johnston A, editors. Molecular techniques in taxonomy. Berlin: Springer-Verlag Press; p. 283–293.

Editorial Committee of Chinese Ben cao. 1999. The State Administration of Traditional Chinese Medicine of People’s Republic of China. Zhong Hua Ben Cao. Vol. 4. Shanghai: Shanghai Science and Technology Press; p. 533.

Gao XF, Brian DS. 2010. *Indigofera*. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China. Vol. 10. Beijing; St. Louis: Science Press; Missouri Botanical Garden Press; p. 154.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultratfast bootstrap approximation. Mol Biol Evol. 35(2):518–522.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 14(6):587–589.

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 YY, Li CQ, Xu QT, Kang WY. 2011. Antioxidant, α-glucosidase inhibitory activities in vitro and alloxan-induced diabetic rats’ protective effect of *Indigofera stachyodes* Lindl. root. J Med Plants Res. 5:3321–3328.

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274.

Qiu L, Liang Y, Tang GH, Yuan CM, Zhang Y, Hao XY, Hao X-J, He H-P. 2013. Two new flavonols, including one flavan dimer, from the roots of *Indigofera stachyodes*. Phytochem Lett. 6(3):368–371.