

## The complete chloroplast genome of *Isodon rubescens*, a traditional Chinese herb

Zhiyong Yue<sup>a</sup>, Yan Wang<sup>b</sup>, Baozhen Zhou<sup>a</sup> and Hanping Wang<sup>a</sup>

<sup>a</sup>College of Medicine, Xi'an International University, Xi'an, Shaanxi, China; <sup>b</sup>Microbiology Institute of Shaanxi, Xi'an, Shaanxi, China

### ABSTRACT

*Isodon rubescens* (Hemsley) H. Hara is a kind of traditional medical herb which can be used for cancer treatment. In this study, the complete chloroplast genome sequence of *I. rubescens* was assembled. Its complete circular chloroplast DNA length was 152,761 bp. The genome was made up of a large single-copy region of 83,655 bp, a small single-copy region of 17,660 bp, and a pair of inverted repeat regions of 25,723 bp. The genome totally encoded 129 genes, containing 85 protein-coding genes, 36 tRNA genes, and eight rRNA genes. Phylogenetic analysis indicated that *I. rubescens* had a close relationship with basil (*Ocimum basilicum*).

### ARTICLE HISTORY

Received 22 September 2020  
Accepted 27 November 2020

### KEYWORDS

*Isodon rubescens*;  
chloroplast genome;  
phylogenetic tree



*Isodon rubescens* (Hemsley) H. Hara, also known as *Rabdosia rubescens*, is one of the important medical herbs in the Ocimeae tribe of Lamiaceae family (Neelamkavil and Thoppil 2017). Current phytochemical study had isolated several diterpenoids from *I. rubescens*, which showed positive functions to treat various diseases (Zhang et al. 2017). Oridonin as one of the identified diterpenoids was possible for the treatment of cancer, autoimmune, and inflammatory diseases through the inhibition of NF- $\kappa$ B (Leung et al. 2005). The key enzymes involved in terpenoid biosynthesis in *I. rubescens* were predicated through transcriptome analysis (Su et al. 2016). Lamiaceae is one of the largest family with more than 7000 species, which are classified into 12 primary clades (Li et al. 2016). The complete chloroplast genome sequences of *I. rubescens* can provide more useful information for phylogeny analysis and taxonomy.

The plant sample of *I. rubescens* was cultivated in the herb nursery of Xianyang (108.69E, 34.35 N), Shaanxi Province, China. The voucher specimen was deposited in the Herbarium of the Microbiology Institute of Shaanxi, Microbiology Institute of Shaanxi, China (zw2020003). The DNA from fresh leaves was extracted by the CTAB method (Porebski et al. 1997). The DNA insert fragments about 400 bp in length were used for library construction. Based on the Illumina Novaseq Platform at Personal Biotechnology Co. Ltd (Shanghai, China) and 2  $\times$  150 bp pair-end sequencing mode, total 1.7 Gbp of sequencing data was generated (NCBI accession number: SRR13002570). The cleaned reads were assembled to form contig and scaffold via software of A5-MiSeq (v 20150522, Coil et al. 2015) and SPAdes (v 3.9.0, Bankevich et al. 2012). The locations of assembled contig sequences were determined using MUMmer (v 3.1, Kurtz

et al. 2004) by *Mentha  $\times$  piperita* as the reference genome (NCBI accession number: NC\_047475.1). The assembled result was corrected via pilon (v1.18, Walker et al. 2014). The final chloroplast genome was annotated by CPGAVAS2 (Shi et al. 2019).

Using the next generation sequencing technology, we assembled a circular complete chloroplast genome of *I. rubescens*. The genome sequence and all gene annotations were submitted to the NCBI database with an accession number of MW018469. The complete chloroplast genome sequence was 152,761 bp in length. The genome was consisted of a large single-copy region (LSC, 83,655 bp), a small single-copy region (SSC, 17,660 bp), and two inverted repeat regions (IR, 25,723 bp). The whole genome encoded 129 genes, including 85 protein-coding genes, 36 tRNA genes, and 8 rRNA genes.

To determine the evolutionary relationship between *I. rubescens* and other Lamiaceae species, a maximum-likelihood phylogenetic tree based on total 21 complete chloroplast sequences from Lamiaceae family was constructed (Figure 1). The species of *Teucrium mascatense* and *Ajuga reptans* from Ajugoideae subfamily were used as the outgroups. The GTRGAMMA + F + R3 model was chosen according to BIC and 1000 bootstrap replicates were used. All genome sequences were retrieved from the GenBank database. These sequences were aligned using MAFFT (v 7.407, Katoh and Standley 2013). Then, trimAl (v 1.4.1, Capella-Gutiérrez et al. 2009) was applied to remove poorly-aligned and divergent regions with algorithm automated1. The remaining sequences were used to infer the phylogenetic position using IQtree (v 1.6.12, Nguyen et al. 2015) under parameters '-nt AUTO -m MFP -bb 1000 -bnni'. The result

**CONTACT** Hanping Wang  [hanpingwang@126.com](mailto:hanpingwang@126.com)  College of Medicine, Xi'an International University, Xi'an, Shaanxi, China

© 2021 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-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

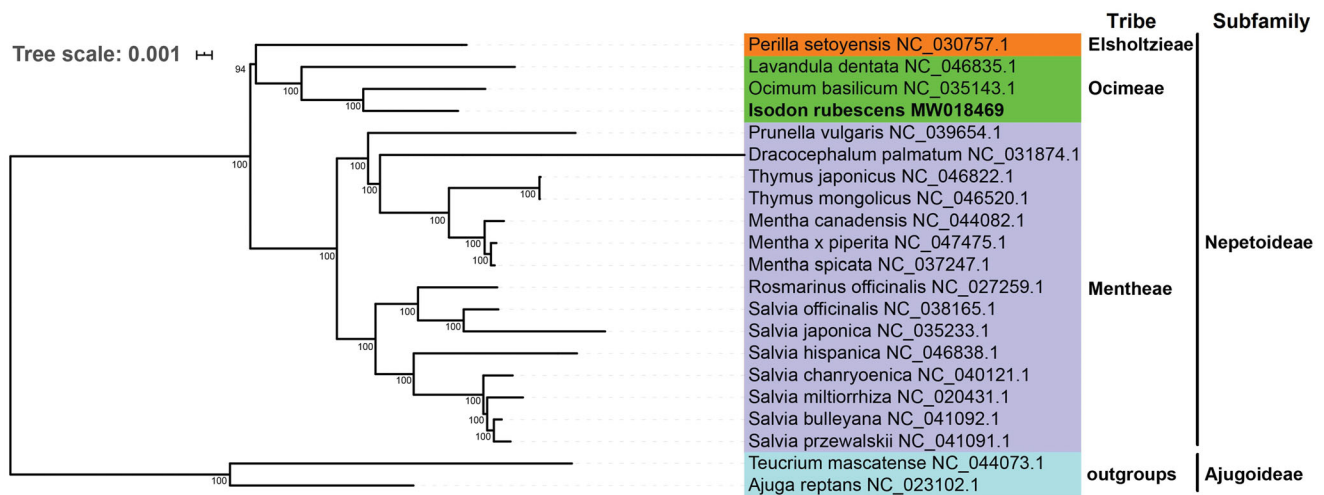


Figure 1. Phylogenetic analysis of 21 complete chloroplast genomes from Lamiaceae family. The bootstrap support values are marked at the nodes.

showed that *I. rubescens* had a close relationship with basil (*Ocimum basilicum*).

## Disclosure statement

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

## Funding

This research was supported by the Science and Research special Project of Education Department of Shaanxi Provincial Project [20JK0864], the Doctoral Scientific Research Foundation of Xi'an International University [XAIU2019014], and the Key Research and Development Program of Shaanxi Province [2020NY-197].

## Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank at <https://www.ncbi.nlm.nih.gov/> under the accession no. MW018469. The associated BioProject, BioSample, and SRA numbers are PRJNA674459, SAMN16657551, and SRR13002570 respectively.

## References

- 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.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 25(15):1972–1973.
- Coil D, Jospin G, Darling AE. 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics.* 31(4):587–589.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* 5(2):R12.
- Leung C-H, Grill SP, Lam W, Han Q-B, Sun H-D, Cheng Y-C. 2005. Novel mechanism of inhibition of nuclear factor-kappa B DNA-binding activity by diterpenoids isolated from *Isodon rubescens*. *Mol Pharmacol.* 68(2):286–297.
- Li B, Cantino PD, Olmstead RG, Bramley GLC, Xiang C-L, Ma Z-H, Tan Y-H, Zhang D-X. 2016. A large-scale chloroplast phylogeny of the Lamiaceae sheds new light on its subfamilial classification. *Sci Rep.* 6(1):34318–34343.
- Neelamkavil SV, Thoppil JE. 2017. Chromosome complements of *Isodon* (Family: Lamiaceae-Subfamily: Nepetoideae): trends in karyotype evolution. *Proc Natl Acad Sci, India, Sect B Biol Sci.* 87(2): 267–272.
- Nguyen L-T, 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.
- Porebski S, Bailey LG, Baum BR. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol Biol Rep.* 15(1):8–15.
- 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.
- Su X, Li Q, Chen S, Dong C, Hu Y, Yin L, Yang J. 2016. Analysis of the transcriptome of *Isodon rubescens* and key enzymes involved in terpenoid biosynthesis. *Biotechnol Biotechnological Equipment.* 30(3): 592–601.
- Walker BJ, Abeeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLOS One.* 9(11):e112963.
- Zhang Y-Y, Jiang H-Y, Liu M, Hu K, Wang W-G, Du X, Li X-N, Pu J-X, Sun H-D. 2017. Bioactive *ent*-kaurane diterpenoids from *Isodon rubescens*. *Phytochemistry.* 143:199–207.