

## The complete chloroplast genomes of *Lycopus lucidus* and *Agastache rugosa*, two herbal species in tribe Mentheae of Lamiaceae family

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

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

### ABSTRACT

Lamiaceae is one of the largest families in the kingdom Plantae, including lots of traditional Chinese herbs. *Lycopus lucidus* and *Agastache rugosa* are two Lamiaceae species, which are most frequently used in Chinese traditional medicine. In the current study, the complete chloroplast genome sequences of two species were assembled. Their circular DNA lengths were 152,096 and 151,922 bp respectively. Both genomes were made up of a large single-copy region, a small single-copy region, and a pair of inverted repeat regions. Each genome totally encoded 133 genes, containing 88 protein-coding genes, 37 tRNA genes, and eight rRNA genes. Phylogenetic analysis indicated that both species belong to the Mentheae tribe of the Lamiaceae family.

### ARTICLE HISTORY

Received 13 September 2020  
Accepted 2 November 2020

### KEYWORDS

*Lycopus lucidus*; *Agastache rugosa*; chloroplast genome; phylogeny



Lamiaceae is one of the largest families in the plant kingdom, with more than 7000 species from over 200 genera (Karaca et al. 2013). A variety of plants in the Lamiaceae family are important sources as traditional Chinese medical herbs, which are used for pain treatment (Uritu et al. 2018), anti-fungal (Waller et al. 2017), anti-cancer (Mesquita et al. 2019) and some other diseases. *Lycopus lucidus* and *Agastache rugosa* are two widely used medicinal Lamiaceae plants. In this study, the chloroplast genomes of both species were assembled and their phylogenetic relationship within the Lamiaceae family was analyzed.

Chloroplast genomic sequences are most conserved during plants' evolution and are often used for the evolutionary analysis. Through the phylogenetic analysis of some chloroplast genes from 288 species of Lamiaceae, the family was divided into 12 large branches (Li et al. 2016). On this basis, using the next-generation sequencing technology, we assembled the complete chloroplast genome of *Lycopus lucidus* and *Agastache rugosa*, and obtained their circular chloroplast genome sequences, which were submitted to the NCBI database (accession numbers: MT980792 and MT980793)

Two plant samples in the current experiment were collected from the herb nursery of Xianyang (108.69E, 34.35N), Shaanxi, China. The voucher specimens were deposited in the Herbarium of the Microbiology Institute of Shaanxi, Microbiology Institute of Shaanxi, China (zw2020001 and zw2020002). Total genomic DNA was extracted from the fresh leaves of plants by the CTAB method (Porebski et al. 1997). The DNA library was constructed and the inserted fragment was about 400 bp. The Illumina Novaseq sequencing system was used, and paired-end 150 bp mode was selected. After data cleaning, 1.79 and 1.71 Gb high-quality

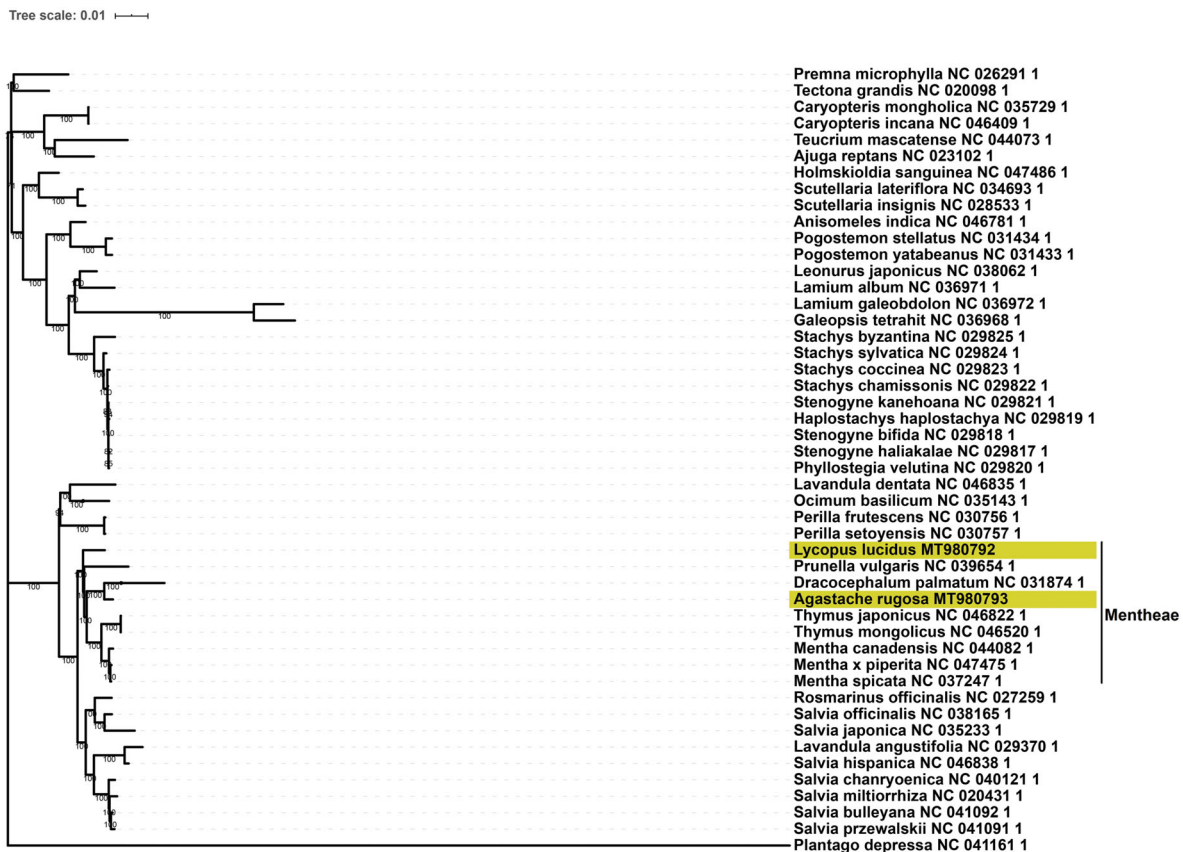
reads data were obtained from two species. These reads were assembled *de novo* by software of A5-MiSeq (v 20150522, Coil et al. 2015) and SPAdes (v 3.9.0, Bankevich et al. 2012). The locations of assembled contig sequences were determined using MUMmer (v 3.1, Kurtz et al. 2004) software with *Mentha × piperita* as the reference (NCBI accession number: NC\_047475.1). The length of two circle chloroplast genome sequences were 152,096 (*Lycopus lucidus*) and 151,922 bp (*Agastache rugosa*) respectively. Both genomes consisted of a large single-copy region, a small single-copy region and two inverted repeat regions. The web server CPGAVAS2 (Shi et al. 2019) was used to annotate the assembled genomes. Both genomes encoded 133 genes, including 88 coding genes, 37 tRNA genes and 8 rRNA genes.

To determine the phylogenetic relationship of both species in the Lamiaceae family, Total 48 complete sequences of the chloroplast genome were retrieved from the GenBank to construct a maximum-likelihood tree. The GTRGAMMA + F+R3 model was chosen according to BIC and 1000 bootstrap replicates were used. *Plantago depressa* in the family of Plantaginaceae was as an out-group. 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 showed both *Lycopus lucidus* and *Agastache rugosa* were in the Mentheae clade (Figure 1), and had a closed relationship with mint.

**CONTACT** Zhiyong Yue  [yuezhiyong@xaiu.edu.cn](mailto:yuezhiyong@xaiu.edu.cn)  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 48 complete chloroplast genomes. Bootstrap support values are marked at the nodes.

## 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 [20JK0865], the Science and Technology Research Project of Shaanxi Province Academy of Sciences [2018nk-01], and the Key Research and Development Program of Shaanxi Province [2019SF-291].

## Data availability statement

The data that support the findings of this study are openly available in NCBI at <https://www.ncbi.nlm.nih.gov/>, with accession number [MT980792; MT980793].

## References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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.
- Karaca M, Ince AG, Aydin A, Ay ST. 2013. Cross-genera transferable e-microsatellite markers for 12 genera of the Lamiaceae family. *J Sci Food Agric.* 93(8):1869–1879.
- 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.
- 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):34343–34318.
- Mesquita L. S S d, Luz TRSA, Mesquita J. W C d, Coutinho DF, Amaral F. M M d, Ribeiro M. N d S, Malik S. 2019. Exploring the anticancer properties of essential oils from family Lamiaceae. *Food Rev Int.* 35(2): 105–131.
- 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.
- 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.
- Uritu CM, Mihai CT, Stanciu G-D, Dodi G, Alexa-Stratulat T, Luca A, Leon-Constantin M-M, Stefanescu R, Bild V, Melnic S, et al. 2018. Medicinal plants of the family Lamiaceae in pain therapy: a review. *Pain Res Manag.* 2018:7801543. 2018.
- Waller SB, Cleff MB, Serra EF, Silva AL, Gomes ADR, de Mello JRB, de Faria RO, Meireles MCA. 2017. Plants from Lamiaceae family as source of antifungal molecules in humane and veterinary medicine. *Microb Pathog.* 104:232–237.