

The complete chloroplast genome of *Mentha spicata*, an endangered species native to South Europe

Kangyu Wang^{a*}, Li Li^{a*}, Ying Hua^b, Mingzhu Zhao^a, Shaokun Li^a, Honghua Sun^a, Yanxi Lv^a and Yi Wang^a

^aCollege of Life Science, Jilin Agricultural University, Changchun, China; ^bJilin ZhongKe Bio-engineering Joint Stock Co., Ltd, Changchun, China

ABSTRACT

Mentha spicata, also known as mint, is the best known source of aromatic essential oil. It was widespread in Europe and Asia. But due to human activity, it has been classified as Least Concern (LC) species in the IUCN Red List of Threatened Species. Here, we presented the complete chloroplast genome of *M. spicata*. The circular genome is 16,430 bp in length and contains 132 genes, including 87 protein-coding genes (PCG), 37 transfer RNA genes (tRNA) and eight ribosomal RNA genes (rRNA). The overall nucleotide composition is: 30.7% A, 19.2% C, 18.6% G, 31.5% T, with a total G + C content of 37.85%. The phylogenetic tree was constructed to explore the taxonomic status of *M. spicata*, which contributes to phylogenetic studies and further conservation strategies for this species.

ARTICLE HISTORY

Received 23 November 2017
Accepted 1 December 2017

KEYWORDS

Mentha spicata; chloroplast genome; HomBlocks; phylogenomics; phylogenetic relationship




Mentha spicata, also known as mint, belongs to a genus of plants in the family Lamiaceae. For decades, *M. spicata* is used to extract aromatic essential oil, which is the source of the best-known monoterpenes, menthol and carvone (Jin et al. 2014). They are extensively used in flavour and fragrance industries, pharmaceuticals and cosmetic products (Lange et al. 2011). *M. spicata* has a very widespread distribution, but due to its long-term and excessive utilization by man, the amount of this species has significantly declined. In 2014, they were classified as Least Concern(LC) species in the IUCN Red List of Threatened Species (Lansdown 2014). The availability of the complete chloroplast genome sequences is helpful to the conservation of endangered species. However, sequences of complete chloroplast genomes of the genus *Mentha* is limited to *M. longifolia* (Vining et al. 2017). In this study, we obtained and characterized the complete chloroplast genome sequence of *M. spicata* using the Illumina paired-end sequencing data, which will contribute to develop protection measures for this endangered species.

The specimen of *M. spicata* was isolated from Jilin Agricultural University test field in Changchun, Jilin, China (125.24E; 43.48N) and the DNA of *M. spicata* was stored in Jilin Agricultural University College of Life Science (No. JLAUCL53). The DNA sample was sequenced using the Illumina X-Ten Sequencing Platform (Illumina, CA). Before chloroplast genome assembling, adapters and low-quality sequences were removed using the FastQC software (Andrews 2010). The chloroplast genome was assembled with SPAdes v3.8 (<http://bioinf.spbau.ru/spades>) (Bankevich et al.

2012) and annotated by DOGMA (<http://dogma.cccb.utexas.edu/>) (Wyman et al. 2004). The tRNA genes were further identified using ARAGORN (Laslett and Canback 2004). The annotated chloroplast genome was submitted to GenBank database under accession No. MG256495.

The complete chloroplast genome of *M. spicata* is a circular molecule with a length of 152,132bp, comprising a pair of inverted repeat regions (IRs) of 25,625bp, a large single-copy region (LSC) of 83,218bp and a small single-copy region (SSC) of 17,664bp. It contained 132 genes, including 87 protein-coding genes (PCG), 37 transfer RNA genes (tRNA) and eight ribosomal RNA genes (rRNA). Among those genes, 17 genes duplicated in the IR region. Additionally, 20 genes were found containing a single intron and two genes (*ClpP* & *ycf3*) had two introns. The base compositions of *M. spicata* chloroplast genome were uneven (30.7% A, 19.2% C, 18.6% G, 31.5% T). The overall GC content of this chloroplast genome was 37.85%.

To ascertain phylogenetic position of *M. spicata* among other higher plants, we selected 79 published complete chloroplast genome sequences of higher plants to construct alignment using Homblocks (<https://github.com/fenghen360/HomBlocks>) (Bi et al. 2017). The phylogenetic trees were reconstructed using maximum-likelihood (ML) and neighbour-joining (NJ) methods. ML analysis were performed using RaxML-8.2.4 (Stamatakis 2014), of which the bootstrap values were calculated using 1000 replicates to assess node support. NJ phylogenetic tree was constructed using MEGA7 with 1000 bootstrap replicate (Kumar et al. 2016). All the nodes were inferred with strong support by the ML and NJ

CONTACT Kangyu Wang  wky427@sina.com; Yi Wang  wanglaoshi0606@163.com  College of Life Science, Jilin Agricultural University, Changchun, China
*Joint first author: These authors contributed equally to this study.

© 2017 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.

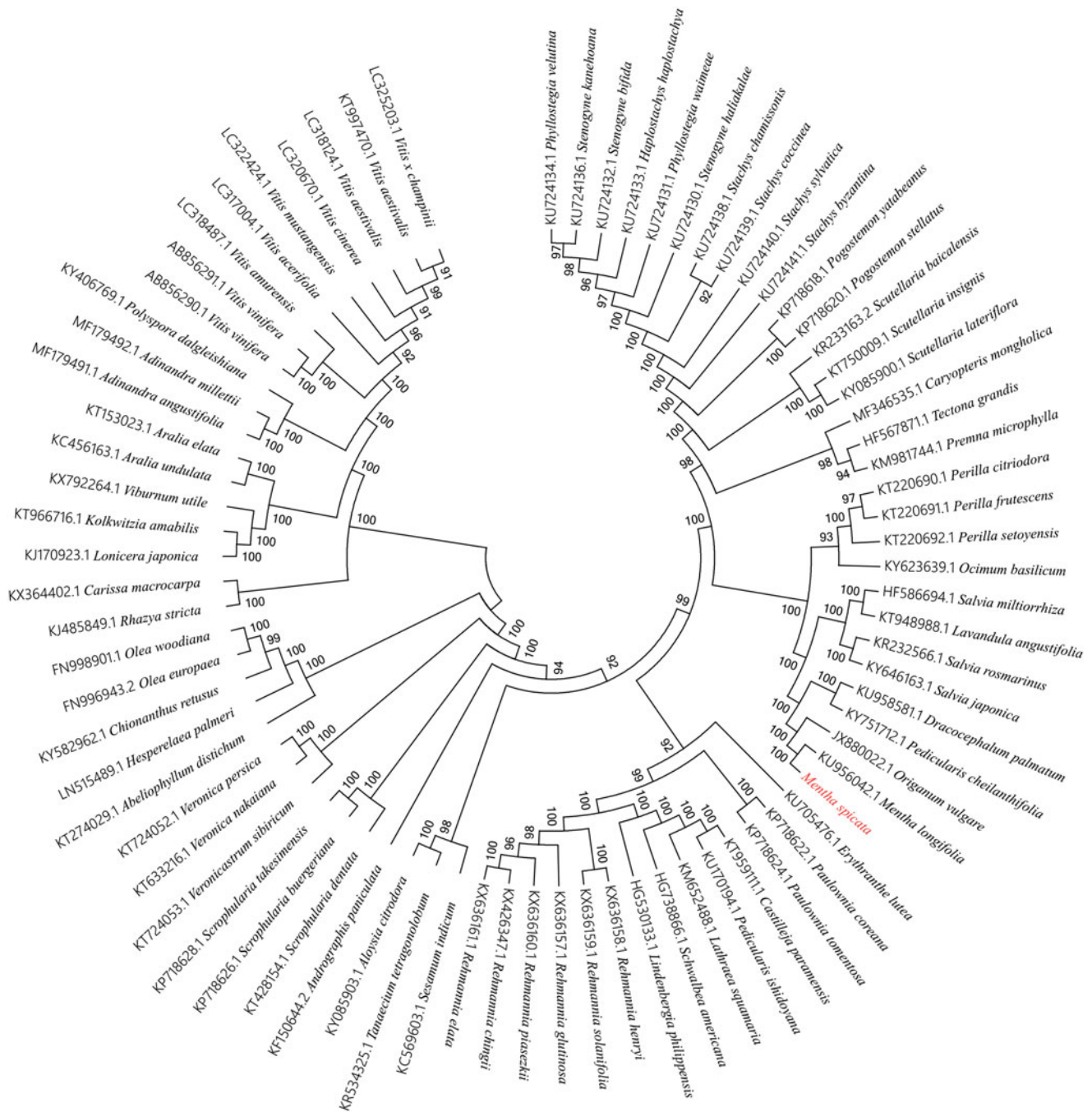


Figure 1. Phylogenetic relationships among 79 plant chloroplast genomes. This tree was drawn without setting of an outgroup. All nodes exhibit above 90% bootstraps. The length of branch represents the divergence distance.

methods. As shown in the phylogenetic tree (Figure 1), The chloroplast genome of *M. spicata* was clustered with *M. longifolia* and *Origanum vulgare*.

Disclosure statement

The authors have declared that no competing interests exist.

Funding

This work was supported by the Doctoral Scientific Research Foundation of Jilin Agricultural University [201618].

References

- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19:455–477.
- Bi G, Mao Y, Xing Q, Cao M. 2017. HomBlocks: a multiple-alignment construction pipeline for organelle phylogenomics based on locally collinear block searching. *Genomics.* DOI:10.1016/j.ygeno.2017.08.001.
- Jin J, Panicker D, Wang Q, Kim MJ, Liu J, Yin JL, Wong L, Jang IC, Chua NH, Sarojam R. 2014. Next generation sequencing unravels the biosynthetic ability of spearmint (*Mentha spicata*) peltate glandular trichomes through comparative transcriptomics. *BMC Plant Biol.* 14:292.

- Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33:1870–1874.
- Lange BM, Mahmoud SS, Wildung MR, Turner GW, Davis EM, Lange I, Baker RC, Boydston RA, Croteau RB. 2011. Improving peppermint essential oil yield and composition by metabolic engineering. *Proc Natl Acad Sci.* 108:16944–16949.
- Lansdown RV. 2014. *Mentha spicata*. The IUCN Red List of Threatened Species 2014: e.T164464A42395980. <http://dx.doi.org/10.2305/IUCN.UK.2014-1.RLTS.T164464A42395980.en>.
- Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32:11–16.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 30:1312–1313.
- Vining KJ, Johnson SR, Ahkami A, Lange I, Parrish AN, Trapp SC, Croteau RB, Straub SC, Pandelova I, Lange BM. 2017. Draft genome sequence of *Mentha longifolia* and development of resources for mint cultivar improvement. *Mol Plant.* 10:323–339.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.