Complete chloroplast genomes of two mint species (Lamiaceae) from Al-Madinah, Saudi Arabia: phylogenetic and genomic comparative analyses

Faten Zubair Filimban, Samaila Samaila Yaradua, Abubakar Bello and Hani Choudhry

CONTACT Faten Zubair Filimban faten_filimban@yahoo.com Department of Biology, Faculty of Sciences, Division of Botany, King Abdulaziz University, Jeddah, Saudi Arabia; Samaila Samaila Yaradua dryaradua@gmail.com Department of Biology, Center for Biodiversity and Conservation, Umaru Musa Yaradua University, Katsina, Nigeria

The genus Mentha (mint) is one of the most significant taxa of the Lamiaceae, and it comprises 24 species, among which 15 are hybrids (POWO). Species in the genus are almost exclusively perennial, herbaceous plants, which are cosmopolitan in distribution. The genus is represented by three species in Saudi Arabia: Mentha longifolia (L.) Huds 1762 (locally called Hassawi and Habaqайн دما نا), Mentha pulegium L. 1753 (Mugrabi) and M. × verticillata L. 1759 (Al-Madinah mintائن دما نا) with M. longifolia being the only known native (Collenette 1985; POWO 2022). All the species are used in herbal teas, alone or as spice mixtures for many foods to offer aroma and flavor. In addition, Mentha species have been used for the treatment of many diseases such as throat infection, bronchitis, nausea, and ulcerative colitis (Ahmed et al. 2015; Vining et al. 2017). This makes many Mentha species very valuable for industry, as mint oil is a supplement for wide spectrum of products. Other uses of mint essential oils include antipruritic, astringent, antiseptic, and antimicrobial activities, and for treating neuralgia, myalgia, headaches, and migraines (Ahmed et al. 2015; Jedrzejczyk and Rewers 2018).

Despite the significance of the species of the genus Mentha, identification and phylogenetic relationships among the species are very complicated and still uncertain. This is largely as a result of the variation in basic chromosome number, frequent interspecific hybridization, cytomixis, polymorphism in morphology and essential oil composition under different environmental conditions, colonial mutant propagation, as well as the occurrence of polyploidy, aneuploidy, and nothomorphs (Gobert et al. 2002; Tucker and Chambers 2002; Denslow and Poindexter 2009; Jabeen et al. 2012; Tucker 2012; Jedrzejczyk and Rewers 2018). This is the cause of the existence of numerous synonymous names (over 3000 published names in the genus), and still uncertain taxonomy of the genus. To solve this problem, there is need to sequence and characterize the plastome of the species.

Two sampled specimens were identified by Dr. Faten Z. Filimban and deposited in the Herbarium of King Abdulaziz University in Saudi Arabia (KAU) (contact info: malgandaby@yahoo.com) including M. longifolia (voucher no. FK1021) and M. × verticillata (voucher no. FK1022) all from Wadi Aqiq near Dhu Hulaifah area, Al-Madinah Province, Saudi Arabia (24.423372'N, 39.562133'E and 24.42398'N, 39.56723'E, respectively). The plants were collected in non-protected areas; hence no permission is required for the collection of the species. Total genomic DNA was extracted using plant...
Nodes represent the posterior probability (PP) values. (M/C2) of clean data were obtained for ing was performed on an Illumina HiSeq 2500 platform Figure 1. Bayesian NOVOPlasty (Dierckxsens et al. 2017). The complete chloroplast genome of Mentha longifolia (NC_032054.1) was used as seed and reference. The orientation of the assembled sequences was further corrected manually using Geneious prime 2020.0.5 (Biomatters Ltd., Auckland, New Zealand).

The complete cp genome was annotated with the online annotation tool GeSeq and annotations were corrected manually with the sequin v. 15. 50 (Biomatters Ltd., Auckland, New Zealand).

The circular chloroplast genome of M. longifolia was 152,078 bp (GenBank accession no. ON124917) in size, contained a large single-copy (LSC) region of 83,220 bp and a small single-copy (SSC) region of 17,652 bp, and separated by a pair of inverted repeats (IRs) regions of 25,603 bp. The cp genome has 133 annotated genes, including 88 protein coding genes, 37 tRNA genes, and eight rRNA genes. The base compositions of the chloroplast genome were uneven (31.5% A, 19.2% C, 18.6% G, and 31.5% T), with an overall GC content of 37.8%.

The complete chloroplast genome of M. × verticillata was 152,026 bp in length (GenBank accession no. ON124918) and composed of IRs of 25,605 bp which divide LSC of 83,165 bp and SSC of 17,655 bp, the average GC content was 37.8% (30.7% A, 19.2% C, 18.6% G, and 31.5% T). There are 133 genes annotated, including 87 protein-coding genes, 38 tRNA genes, and eight rRNA genes. To identify the phylogenetic positions of M. longifolia and M. × verticillata among the species Nepetoideae, we downloaded 13 published complete chloroplast genome sequences of Nepetoideae and chloroplast genome of Teucrium mascatense (Ajugoideae), Congea tomentosa (Symphorematoideae), and Dicrastylis parvifolia (Prostantheroideae) to be used as outgroup to construct the phylogenetic tree. The plastome sequences were aligned with MAFFT v.7 (Katoh and Standley 2013), and analyzed using Bayesian’s inference analysis. The result clearly showed that the tribe is monophyletic based on the samples of species and also supported the division of Salviniae, Nepetinae, and Menthiniae subtribes with strong support with posterior probability (PP) = 1.0 (Figure 1). The phylogenetic tree revealed that M. × verticillata is a sister to Mentha spicata whereas M. longifolia formed a sister relationship with a clade of M. canadensis and M. × piperita. All these relationships were strongly supported (PP = 1.0). Interestingly, M. longifolia sequenced in this study did not form a sister relationship with M. longifolia from South Africa, this might be as a result of phylogeographic factors. This opens a window for further investigations.

**Author contributions**

SSY, FZB, and AB were involved in the conception and design; HC and SSY performed the analysis; AB and HC drafted the paper; FZB and SSY revised it critically for intellectual content; all authors approved the final version to be published and agreed to be accountable for all aspects of the work. All the authors meet the criteria for authorship as per the ICMJE criteria.

**Disclosure statement**

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

**Funding**

The research was supported by The Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Grant No. [IFPRC-187-130-2020]. The authors are highly grateful for this support.

**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/ under the accession no. ON124917–ON124918. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA842226 for M. verticillata, PRJNA842225 for M. longifolia, SRR19913734 for M. verticillata, SRR20637368 for M. longifolia and SAMN28650027 for M. verticillata, SAMN28649942 for M. longifolia, respectively.

**References**

Ahmed AM, Ozbak HA, Hemeg HA. 2015. Effect of essential oil of traditional two Saudi mint types and its possible role in cardiovascular and throat health. Int J Clin Exp Med. 8(5):8060.

Colletette S. 1985. Illustrated guide to the flowers of Saudi Arabia. London: Scorpio Publishing Ltd.

Denslow MW, Poindexter DB. 2009. Mentha suaveolens and M. × rotundifolia in North Carolina: a clarification of distribution and taxonomic identity. J Bot Res Inst Texas. 3:389–398.

Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 45(4):e18.

Gobert V, Moja S, Colson M, Taberlet P. 2002. Hybridization in the section Mentha (Lamiaceae) inferred from AFLP markers. Am J Bot. 89(12): 2017–2023.

Jabeen A, Guo B, Abbasi BH, Shinwari ZK, Mahmood T. 2012. Phylogenetics of selected Mentha species on the basis of rps8, rps11 and rps14 chloroplast genes. J Med Plants Res. 6:30–36.

Figure 1. Bayesian’s inference phylogenetic tree of Nepetoideae species (n = 13) based on the complete chloroplast genome sequence. Numbers on the nodes represent the posterior probability (PP) values.

DNA (BioTeke Corporation, Beijing, China). Genome sequencing was performed on an Illumina HiSeq 2500 platform (illumina Inc., San Diego, CA). Approximately, 6.0 and 5.1 GB of clean data were obtained for M. longifolia and M. × verticillata, respectively, and assembled using NOVOPlasty (Dierckxsens et al. 2017). The complete chloroplast genome of Mentha longifolia (NC_032054.1) was used as seed and reference. The orientation of the assembled sequences was further corrected manually using Geneious prime 2020.0.5 (Biomatters Ltd., Auckland, New Zealand).
Jedrzejczyk I, Rewers M. 2018. Genome size and ISSR markers for Mentha L. (Lamiaceae) genetic diversity assessment and species identification. Ind Crops Prod. 120:171–179.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780.
POWO. 2022. Plants of the world online. Kew: Facilitated by the Royal Botanic Gardens; [accessed 2022 Apr 5]. http://www.plantsoftheworldonline.org/.

Tucker AO. 2012. Genetics and breeding of the genus Mentha: a model for other polyploid species with secondary constituents. J Med Active Plants. 1:19–29.

Tucker AO, Chambers HE. 2002. Mentha canadensis L. (Lamiaceae): a relict amphidiploids from the lower tertiary. Taxon. 51(4):703–718.

Vining KJ, Johnson SR, Ahkami A, Lange I, Parrish AN, Trapp SC, Croteau RB, Straub SCK, Pandelova I, Lange BM. 2017. Draft genome sequence of Mentha longifolia and development of resources for mint cultivar improvement. Mol Plant. 10(2):323–339.