



The complete chloroplast genome of *Mentha suaveolens* Ehrh. and comparative analysis of six *Mentha* taxa (Lamiaceae)

Mahmoud Magdy · Khairiah Mubarak Alwutayd · Hassan Mansour · Amnah M. Alamri · Salha Mesfer Alshamrani · Nora M. Al Aboud · Daa Abdel Moneim

Received: 30 December 2024 / Accepted: 22 April 2025 / Published online: 24 May 2025
© The Author(s) 2025

Abstract Apple mint (*Mentha suaveolens* Ehrh.) is a plant of the genus *Mentha* in the family Lamiaceae, commonly known by pineapple mint, woolly mint, and/or round leaf mint. Similar to many aromatic plants, it is valued for its aroma, essential oil content, flavor, and/or other secondary metabolites associated with pharmaceutical properties. *M. suaveolens*,

a sweet-scented aromatic herb, is well-known for flavoring candies, medicines, teas, and toothpaste, as well as being used in many other products. It usually found along streams and humid places in western and southern Europe, as well as the Mediterranean region. In the current study, the chloroplast genome of *M. suaveolens* was assembled characterized, and compared with other seven *Mentha* spp. chloroplast genomes, including three natural hybrids. The *M. suaveolens* circular genome length is 152,053

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10722-025-02456-2>.

M. Magdy (✉)
Department of Plant Biology, Faculty of Biology, Murcia University, 30100 Murcia, Spain
e-mail: mahmoud.magdy@um.es

M. Magdy
Department of Genetics, Faculty of Agriculture, Ain Shams University, Cairo 11241, Egypt

K. M. Alwutayd
Department of Biology, College of Science, Princess Nourah bint Abdulrahman University, 11671 Riyadh, Saudi Arabia
e-mail: kmalwateed@pnu.edu.sa

H. Mansour
Department of Biological Sciences, College of Science and Arts, King Abdulaziz University, 21911 Rabigh, Saudi Arabia
e-mail: hmansor@kau.edu.sa

H. Mansour
Department of Botany and Microbiology, Faculty of Science, Suez Canal University, Ismailia 41522, Egypt

A. M. Alamri
Department of Biological Sciences, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia
e-mail: aalomri@Kau.edu.sa

S. M. Alshamrani
Department of Biological Sciences, College of Science, University of Jeddah, 21959 Jeddah, Saudi Arabia
e-mail: smalshmrane@uj.edu.sa

N. M. Al Aboud
Department of Biology, Faculty of Sciences, Umm Al-Qura University, Makkah, Saudi Arabia
e-mail: nmaboud@uqu.edu.sa

D. Abdel Moneim
Department of Plant Production (Genetic Branch), Faculty of Environmental Agricultural Sciences, Arish University, El-Arish, Egypt
e-mail: dabdelmoniem@aru.edu.eg

bp, and there are 134 genes, including 88 protein-coding genes, 37 transfer RNA genes, and 8 ribosomal RNA genes (rRNA), with a total GC content of 37.9%. A chloroplast genome feature comparison between *M. suaveolens* and six *Mentha* species was conducted, revealing genomic variability among the studied accessions. Variation among the LSC-IR-SSC boundaries was observed, as well as the RSCU varied according to the region, while the protein-coding regions were more conservative than the intergenic spacers in number of single nucleotide polymorphisms (SNPs). Based on the phyloplastomic, two clades were identified beside *M. canadensis*. In addition to being informative for any plastomic and/or evolutionary studies of the genus *Mentha* and family Lamiaceae, the provided information report the maternal source of the *Mentha* hybrid species.

Keywords *Mentha suaveolens* · Apple mint · Chloroplast genome comparison · Phylogenetic analysis

Introduction

About 18 species and at least 11 hybrids were reported for the genus *Mentha*, all are found on the five continents, mainly between temperate and subtemperate climate zones (Šarić-Kundalić 2009). The majority of *Mentha* varieties grow in Europe, Australia, Central Asia, and North Africa (El-Ghorab 2006). There are many species belonging to the genus *Mentha*, which grow in wet and moist environments and grow between 10 and 120 cm tall (Šarić-Kundalić 2009). Among the Lamiaceae family, *Mentha* contains several taxa with highly valuable essential oils, making it the most important genus. Over 23,000 metric tons of oil are produced annually, valued at over \$400 million. Due to this, they are considered the most economically valuable essential oils (El-Ghorab 2006).

The genus *Mentha* has complex systematics due to its polyploid nature, morphology, chromosome number, frequent interspecific hybridizations, and vegetative reproduction (Attiya Jabeen 2012). Natural interspecific hybridization occurs with high frequency both in the cultivation and wild population of *Mentha* (Gobert et al. 2002), for example, *Mentha* × *piperita* L., *Mentha* × *verticillata* L., *Mentha* × *villosa* Huds.,

are naturally occurring hybrids with established lineages, among others. There have been several methods used to develop relationships among *Mentha* species, including cytological, morphological, chemical, and molecular markers; diversity in pollen morphology (Celenk et al. 2008), flavonoids variations (Voinin et al. 1999), isozyme polymorphism (Fadhel and Boussaïd 2004), inter simple sequence repeats–polymerase chain reaction (ISSR-PCR), random amplified polymorphic DNA (RAPD) markers (Khanuja et al. 2000; Shinwari et al. 2011), amplified fragment length polymorphism (AFLP) markers (Gobert et al. 2002), and chloroplast DNA variations (Attiya Jabeen 2012) have been used in past to study genus *Mentha* systematics. Having knowledge of a plant's genetic diversity allows us to have a superior parental combination selected (Schlötterer 2004).

Originally known as *Mentha rotundifolia* (L.) Huds. (Nagell et al. 1974), apple mint, *Mentha suaveolens* Ehrh., an aromatic herb with a sweet scent, apple mint grows along streams as well as in humid environments, being native to the Mediterranean region and southern and western Europe (Oumzil et al. 2002). As a result of its diverse effects, it has been used in traditional medicine as a carminative, antispasmodic, diaphoretic, and analgesic (El-Kashoury et al. 2013). Furthermore, it is known to have biological effects, such as anti-inflammatory, antioxidant, and antimicrobial (Oumzil et al. 2002; López et al. 2010). Apple mint is an herbaceous, upright perennial plant that is mostly grown as a culinary herb or for ground cover. Most commonly, it grows between 40 and 100 cm (16–39 inches) and forms clonal colonies by stolon. In contrast to its light green leaves, the sessile, opposite leaves are oblong to nearly ovate, measuring 4–6 cm (1.2 inches) long and 2–4 cm (0.8 inches) wide, with serrated edges and somewhat hairy tops, they have a downy underlayer (Vining et al. 2019). However, apple mint's secondary metabolites have been shown to possess numerous biological activities and pharmaceutical properties (Son et al. 2018). Among the many uses of mint's essential oils, it has been used as a flavoring agent for candies, confections, medicines, teas, and toothpaste, but it is also used in a variety of other products (Moreno et al. 2002; Oumzil et al. 2002; Shin et al. 2008; Chen et al. 2012). In addition to their psychoactive properties, mint essential oils are also known to have anti-inflammatory effects (Umezue et al. 2001). Among its

uses are treating nausea, bronchitis, flatulence, anorexia, ulcerative colitis, and liver disease (Moreno et al. 2002). In addition, it has antioxidant properties, which makes it useful for treating chronic and degenerative conditions (Lachance et al. 2001). It has been proven that peppermint oil is effective in reducing hyperperistaltic symptoms during endoscopy (Asao et al. 2003).

Inferring the maternal lineage and determining the origin of polyploids can be done using the chloroplast genome, which is normally non-recombining and uniparentally inherited (Chen et al. 2012). In higher plants, chloroplast genomes are double-stranded circular DNA with a highly conserved structure and gene content. Sizes range from 120 to 170 kb and carry 120–130 genes, depending on the plant (Ruhlman and Jansen 2014). The chloroplast (cp) genomes typically consist of two inverted repeat regions (IR) that separate the small single-copy region (SSC) from the large single-copy region (LSC) (Huaizhu et al. 2020). The genome is relatively small and has a slow mutation rate, as well as parthenogenetic inheritance (Wang et al. 2017). In addition to developing molecular markers to classify medicinal plants, identify plant species, study population genetics, and study genome evolution, chloroplast genomic information has been widely used (Yao et al. 2020). Since the advent of high-throughput sequencing methods, scientists have been able to sequence the genomes of more than 6000 plant species and deposit them at the National Center for Biotechnology Information (NCBI) (Zhu et al. 2016).

The main objective of this study is to undertake a comprehensive exploration of the chloroplast genome of *M. suaveolens*, encompassing its sequencing, annotation and shedding light on its genetic composition and structural characteristics. Furthermore, this study extends its scope to encompass a comparative analysis involving the chloroplast genomes of an additional six *Mentha* species including three established hybrids. Additionally, we aimed to analyze codon usage bias, long repeat sequences, and simple sequence repeats (SSRs) to gain deeper insights into the evolutionary and structural features of the chloroplast genomes. To uncover and decipher the intricate phylogenetic relationships and to identify unique features and genetic variations specific to each species, enhancing our understanding of the evolutionary

dynamics and genetic diversity existing within the *Mentha* genus.

Materials and methods

Sample collection and DNA extraction

The *M. suaveolens* plant materials were collected from the growing plants in the El-Orman botanical garden in Giza, Egypt, in May 2023. Using the WizPrep™ gDNA Mini Kit (Cell/Tissue, WIZ-BIO, Seoul, South Korea), total genomic DNA was extracted. The DNA quality was evaluated using 1% TBE agarose gels and measured using a Quantus™ Fluorometer (Promega, USA). For further processing, the extracted gDNA was stored at a constant temperature of −20 °C. Voucher specimens are kept in the germplasm collection of the Genetics Department, Faculty of Agriculture, Ain Shams University (coded: MS101).

Library construction, chloroplast genome assembly, and annotation

The TruSeq library preparation kit (Illumina, San Diego, California, USA) was used to prepare DNA libraries using 350 bp fragmented short inserts. Sequencing of the library was performed on an Illumina HiSeq 4000 platform with 150 bp lengths/reads (Novogene, China). In this study, single-contig assembly was performed by using clean, high-quality reads (Magdy et al. 2019, 2022; Magdy and Ouyang 2020). The genome assembly of *M. suaveolens* cp and coding sequences confirmation and correction were performed using Geneious Prime (Kearse et al. 2012). In order to annotate the chloroplast genome, GeSeq was used with the assembled chloroplast genome (Tillich et al. 2017). Additionally, the tRNA scan-SE 2.0 search server, used to identify anticodon sequences and typical cloverleaf secondary structures of all tRNA, was verified (Lowe and Chan 2016).

Tandem repeats and codon usage bias analysis

Based on the minimum repeat size set at 30 and the Hamming distance of 50, REPuter predicted long repeat sequences (Darling et al. 2004). By using Microsatellite Identification Software (MISA), simple

sequence repeats (SSRs) were detected (Beier et al. 2017). In this study, the codon preference (i.e., the relative synonymous codon usage RSCU) of the *Mentha* species protein encoded sequences was analyzed by MEGA× (Kumar et al. 2018).

Genetic polymorphism and evolutionary pressures

After removing stop codons and gaps between the comparative sequences, the chloroplast genomes were aligned using MAFFT v7.3 (Kato and Standley 2013). The single nucleotide polymorphism (SNP) and their impact on protein translation of the protein-coding sequences of seven *Mentha* species was determined using Geneious Prime. Ultimately, we used PAML's CODEML program (London, UK) to calculate Synonymous (dS) and Nonsynonymous (dN) Substitution Rates (dS/dN ratios) (Yang and Rannala 2012). The results of the experiment indicated that positive, neutral, or purifying selections would occur if $dS/dN > 1$, $dS/dN = 1$, and $dS/dN < 1$, respectively (Nielsen 2005).

Chloroplast genomes comparative analysis and phylogenetics

A comparison of the IR/SSC and IR/LSC boundaries and junctions among the seven *Mentha* species including three hybrids (*Mentha canadensis* L., *Mentha longifolia* (L.) L., *Mentha spicata* L., *M. × piperita*, *M. × verticillata*, and *M. × villosa*) was conducted using the online tool IRscope (<https://iscope.shinyapps.io/irapp/>, accessed on the 12 th of November 2022). Using the Mauve genome aligner, the total chloroplast genomes were aligned, as well as each chloroplast region (IR, SSC, and LSC) (Darling et al. 2004). These alignments were used for Maximum Likelihood (ML) phylogenetic reconstruction supported with minimum-evolution NNIs and SPRs methods using FastTree v2.1.11 (Price et al. 2010) in accordance with the software's default parameters for bootstrap support (1000 runs), switching between Jukes-Cantor and generalized time-reversible (GTR) models. The *Ocimum basilicum* cp genome (NC_035143) was used as an outgroup to root the trees.

In cases where the ML trees exhibited chaining effects—particularly with IR-based reconstructions—hierarchical clustering was applied using Ward's

minimum variance method to improve tree topology and reveal more distinct clustering. Pairwise genetic distances were computed under the Kimura 2-parameter model (K80) using the `dist.dna()` function in the `ape` package. WARD clustering was performed using the `hclust()` function with `method = "ward.D2"`, with bootstrap of 1000 run, and resulting trees were visualized and compared to ML trees using "dendextend".

Results

Characteristics of *M. suaveolens* chloroplast genome

To explore the genetic intricacies of the *Mentha* genus, a comparative plastomic analysis was undertaken among seven *Mentha* species. Among these, the newly sequenced chloroplast, identified as *M. suaveolens* (MS101), exhibited a genome size of 152,053 bp and a GC content of 37.9% (Fig. 1). Notably, the plastome sizes for the other species ranged narrowly, from 150,048 bp for *M. × villosa* to 152,154 bp for *M. canadensis*, suggesting a high level of conservation within this genus. The GC content, a fundamental parameter in genome stability and structure, was also remarkably consistent across the board, with most of the compared species of 37.8%, excluding *M. suaveolens* and *M. × verticillata*, registering a GC content of 37.9%. Deeper into the genomic regions, the LSC region in *M. suaveolens* (MS101) spanned 83,168 bp. The LSC regions in the other *Mentha* species exhibited minimal variation, ranging from 83,165 bp in *M. × verticillata* to 83,278 bp in *M. canadensis*. The SSC region, another pivotal segment in plastomic studies, was 17,657 bp in MS101. The other species showcased only minor variations in this region, oscillating between 17,652 bp and 17,676 bp. Furthermore, the IR regions recorded 25,614 bp, the largest across all species, which remained within 25,600 and 25,608 bp. In terms of gene content, there was a unanimous consensus among the *Mentha* species. Each species possessed a total of 134 genes. This tally comprises 88 Protein-Coding Genes (PCGs), 37 tRNA genes essential for protein synthesis, and 8 rRNA genes critical for ribosomal structure and function (Table 1).

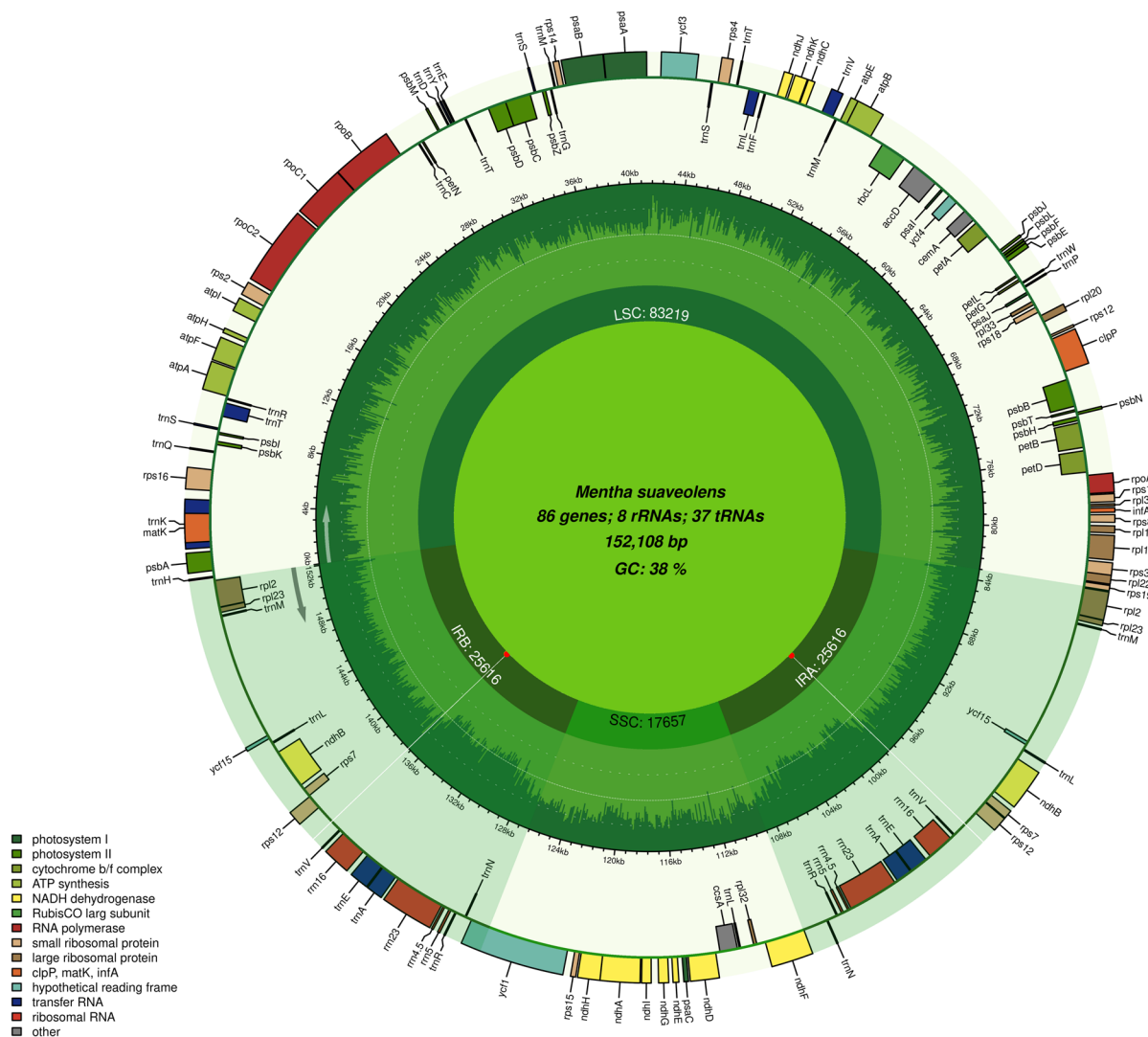


Fig. 1 The complete chloroplast genome map of *Mentha suaveolens* with a total length of 152,108 bp. The inner circle represents the GC content and is defined by cp regions (LSC, SSC, IRA, and IRB). The outer circle represents the complete cp sequence, with the genes annotated outside the circle to rep-

resent forward genes and the genes annotated inside the circle to represent reversed genes. All genes are colored according to their functional groups. The GC content % = 37.9% and is shown approximated to 38%

SSR and long repeat sequence analysis

From the repeat finder analysis, the *M. suaveolens* chloroplast genome showcased a total of 420 SSRs. These repeats were categorized based on the repeat unit size, from mononucleotide to decanucleotide repeats. The most abundant SSRs were the pentanucleotide repeats, amounting to 125 (or 30% of the total SSRs). Following closely were the tetranucleotide repeats, with a count of 92, contributing to 22%

of the total. Trinucleotide and dinucleotide repeats also held significant numbers, with 73 (17%) and 52 (12%) SSRs, respectively. It's worth noting that as the length of the repeat units increased, the frequency generally decreased, with exceptions being the decanucleotide repeats, which held 30 (7%) SSRs. When we break down the distribution of these SSRs across different genomic regions, it's evident that the Large Single Copy (LSC) region is a hotspot, holding 296 SSRs, which translates to a whopping 70.4% of the

Table 1 The basic characteristics of *Mentha suaveolens* cp genome and the other available *Mentha* species

Species	GC%	Length bp	LSC bp	SSC bp	IR bp	Genes	PCGs	tRNA	rRNA
<i>Mentha suaveolens</i> (MS101)	37.9	152,053	83,168	17,657	25,614	134	88	37	8
<i>Mentha canadensis</i> (NC_044082)	37.8	152,154	83,278	17,676	25,600	134	88	37	8
<i>Mentha longifolia</i> (ON124917)	37.8	152,078	83,220	17,652	25,603	134	88	37	8
<i>Mentha spicata</i> (OM617844)	37.8	152,048	83,179	17,653	25,608	134	88	37	8
<i>Mentha</i> × <i>piperita</i> (NC_047475)	37.8	152,048	83,179	17,653	25,608	134	88	37	8
<i>Mentha</i> × <i>verticillata</i> (ON124918)	37.9	152,026	83,165	17,655	25,603	134	88	37	8
<i>Mentha</i> × <i>villosa</i> (OP404081)	37.8	150,048	83,179	17,653	25,608	134	88	37	8

total SSRs. The IR and SSC regions held lesser proportions, with 15% (63 SSRs) and 14.6% (61 SSRs) respectively (Table 2).

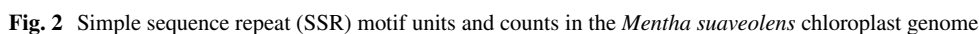
In an in-depth analysis of SSRs within the *M. suaveolens* chloroplast genome, distinct patterns emerge in relation to motif prevalence and their localization within specific genomic regions. In summation, the genomic landscape of *M. suaveolens* is intricately woven with a diverse array of SSR motifs, amounting to an aggregate of 299. Notably, the motifs "AT", "AAT", "AAAT", "AAAAT", "AAAG", and "AAAAAG", are predominant. The LSC region distinctly emerges as a pivotal domain for these prevalent motifs, constituting 207 of the 299 total SSRs, while the SSC and IR regions contribute lesser frequencies, tallying at 48 and 44, respectively (Fig. 2).

Based on each category, within the dinucleotide SSRs, collectively found in 30 occurrences, the motif "AT" conspicuously dominates, registering 20 instances. A significant proportion of these, precisely 16, are localized within the Large Single

Copy (LSC) region, underscoring the region's predisposition towards this specific motif. Transitioning to hexanucleotide SSRs, which account for a total of 78 occurrences, the motif "AAAAAG" emerges preeminent with a total of 10 instances. A meticulous dissection reveals the LSC region as a significant reservoir for this motif, harboring 5 out of the total 10 instances. Delving into the pentanucleotide SSRs, which cumulatively mark 70 occurrences, the "AAAAT" sequence emerges superior, manifesting 15 times. Notably, the LSC region substantiates its proclivity for harboring dominant motifs, encapsulating 12 out of the 15 instances of the "AAAAT" motif. In the context of tetranucleotide SSRs, which cumulate at 66, the "AAAG" motif is most pronounced, with 17 instances. Detailed stratification reveals its pronounced concentration within the LSC region, amounting to 12 out of its 17 total occurrences. Concluding with the trinucleotide SSRs, which enumerate at 49, the "AAT" motif is unequivocally the most recurrent, observed 19 times. The LSC region,

Table 2 Repeated sequences in the *Mentha suaveolens* chloroplast genome, including repeat Class, repeat abundances, and percentage abundance

Row labels	IR	LSC	SSC	Grand Total	~ Percentage
Mononucleotide repeat	5	24	2	31	7
Dinucleotide repeat	9	34	9	52	12
Trinucleotide repeat	11	50	12	73	17
Tetranucleotide repeat	15	61	16	92	22
Pentanucleotide repeat	17	92	16	125	30
Hexanucleotide repeat	4	18	4	26	6
Septanucleotide repeat	–	11	–	11	3
Octanucleotide repeat	2	2	1	5	1
Novanucleotide repeat	1	4	1	6	1
Decanucleotide repeat	4	24	2	30	7
Grand total	63	296	61	420	100
Percentage	15%	70.4%	14.6%	100.00%	–



The SSR motifs across different *Mentha* species exhibit similarities and distinctions in distribution and frequencies. When evaluating dinucleotide repeats, there's a tight clustering in frequency among the various *Mentha* species, ranging from 36 to 40 repeats. The cumulative count for dinucleotide repeats stands at 261, accounting for approximately 8% of the total SSRs. Trinucleotide repeats present a slightly more varied distribution but remain consistent, with values oscillating between 61 and 63 repeats across the species. They accumulate to a significant 435, forming 13% of the total repeats. The tetranucleotide repeats, which contribute to 18% of the total, manifest between 84 and 89 occurrences across the diverse species, with a cumulative count of 602. Pentanucleotide repeats are particularly noteworthy, with frequencies ranging narrowly from 104 to 107. These motifs represent a substantial portion, with a sum of 740, translating to 22% of the entire SSR landscape. Hexanucleotide repeats emerge as the most prevalent motif, with frequencies spanning between 131 and 143. The aggregated count is the highest among all categories at 977, constituting a dominant 29% of total SSRs. Septanucleotide repeats showcase uniformity, each with frequencies hovering around 29 to 30. With an aggregate of 204, they form 6% of the total. Octanucleotide repeats, though fewer in comparison, display a modest variability, ranging from 9 to 14. Their combined frequency is 73, accounting

Across various *Mentha* species, *M. suaveolens* presents intriguing comparative data where exists a combined total of 3386 SSR motifs. The hexanucleotide repeats conspicuously dominate the genomic landscape, whereas motifs like decanucleotide remain considerably sparse. The SSR frequency in *M. suaveolens* stands at 484, which is closely aligned with the ranges observed in the other six *Mentha* species. Specifically, while *M. canadensis* (NC_044082) and *M. × verticillata* (ON124918) have slightly higher frequencies at 487 and 485 respectively, the others, including *M. longifolia* (ON124917), *M. spicata* (OM617844), *M. × piperita* (NC_047475), and *M. × villosa* (OP404081), all match or slightly deviate from the count in *M. suaveolens*, all having 482 or 484 repeats. This highlights a genomic consistency in SSR motif distribution among these species, indicating potential shared evolutionary or genomic structures. However, it's the nuanced differences in specific repeat types, such as tetranucleotide or hexanucleotide motifs, that may offer insights into the unique genetic characteristics or evolutionary trajectories of *M. suaveolens* in comparison to its counterparts (Table 3).

Table 3 Comparative analysis of simple sequence repeat (SSR) motif frequencies across seven *Mentha* species chloroplast genomes, including repeat Class, repeat abundances, and percentage abundance

Repeat type	<i>M. suaveolens</i> (MS101)	<i>M. canadensis</i> (NC_044082)	<i>M. longifolia</i> (ONI24917)	<i>M. spicata</i> (OM617844)	<i>M. × piperita</i> (NC_047475)	<i>M. × verticillata</i> (ONI24918)	<i>M. × villosa</i> (OP404081)	Grand total	Percentage
Dinucleotide repeat	36	40	37	37	37	37	37	261	8
Trinucleotide repeat	61	62	63	62	62	63	62	435	13
Tetranucleotide repeat	84	89	86	86	86	85	86	602	18
Pentanucleotide repeat	107	106	105	106	106	104	106	740	22
Hexanucleotide repeat	142	131	141	140	140	143	140	977	28
Septanucleotide repeat	29	30	29	29	29	29	29	204	6
Octanucleotide repeat	11	14	11	9	9	10	9	73	2
Nonanucleotide repeat	7	11	7	7	7	8	7	54	2
Decanucleotide repeat	7	4	5	6	6	6	6	40	1
Grand total	484	487	484	482	482	485	482	3386	100

Codon usage bias of *Mentha* species

The RSCU values for *M. suaveolens* MS101 are compared to those of *M. canadensis*, *M. longifolia*, *M. spicata*, *M. × piperita*, *M. × verticillata*, and *M. × villosa*. The values represent the frequency of usage of each codon for a specific amino acid relative to other synonymous codons coding for the same amino acid. Generally, a value of 1 indicates no bias, greater than 1 indicates higher usage, and lower than 1 indicates lower usage. Utilizing the sequences of protein-coding genes, the frequency of codon usage for various *Mentha* species was discerned. With the standard repertoire of 64 codons, numerous codons across these species coded for the standard 20 types of amino acids. Barring methionine (AUG) and tryptophan (UGG), which possess an unbiased usage (RSCU = 1) across all species, other amino acids showcased varying codon preferences. For example, leucine, serine, and arginine are represented by six codons each, with other amino acids having representation by two or four synonymous codons.

Within the data set, slight variations in RSCU values are noted, as shown in Fig. 3. Of the codons presented, several exhibit an RSCU value greater than 1, indicating a preference for their usage. Notably, a majority of these preferred codons end in A or U, echoing the observation of an A/U content bias in the third codon position. However, certain codons like AGU and UGG, encoding serine and tryptophan, respectively, maintain an unbiased usage across the *Mentha* species (RSCU = 1), emphasizing a conserved codon preference for these amino acids.

Under more specific comparison, *M. suaveolens* largely mirrors the RSCU values observed in the other *Mentha* species. The slight variations are minute, underscoring a high level of conservation in codon usage preferences across these species. For instance, the codon 'GCA' has an RSCU value of 1.12 in *M. suaveolens*, which is identical in four other species but slightly decreases to 1.11 in *M. × verticillata*. Similarly, for the codon 'CUG', *M. suaveolens* showcases an RSCU value of 0.42, while *M. × piperita*, *M. longifolia*, and *M. × villosa* have a marginally reduced value of 0.41. Such minuscule disparities are observed across multiple codons (Fig. 3).

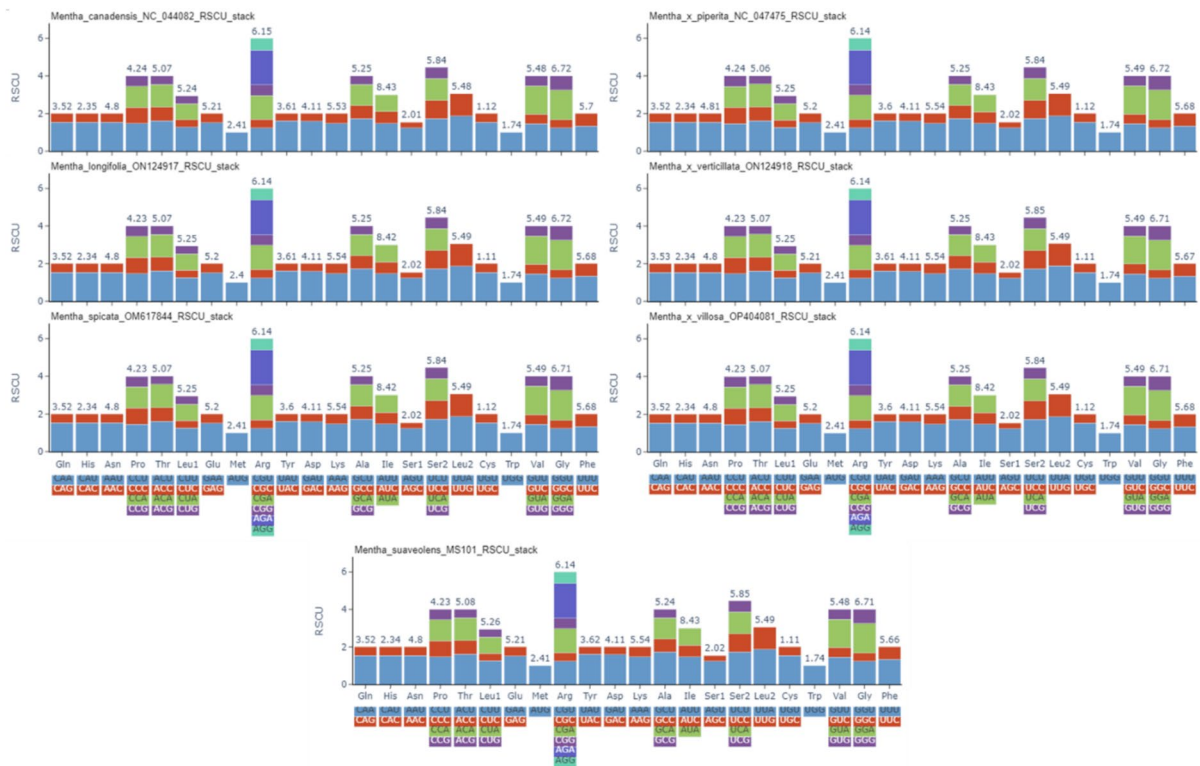


Fig. 3 Relative Synonymous Codon Usage (RSCU) values occurred in the chloroplast genomes of seven *Mentha* species. Values presented in the figure are provided in the Supplementary Data

IR expansion and contraction

The IR region of the chloroplast genome stands out as the most conserved section. Yet, it's the contractions and expansions at its border regions that largely dictate the changes in the chloroplast genome length throughout evolutionary processes. The pivotal junction sites demarcating the transitions between these regions are labeled as JLB (joining the IRb and LSC regions), JSA (connecting the SSC region to IRa), JSB (linking the IRb to SSC), and JLA (fusing the IRa and LSC). Upon the examination of these four junctions across the seven *Mentha* species, it was discerned that subtle size variations in their plastomes were largely inconsequential to any significant shifts in the IR boundaries. However, an exception emerged in *M. canadensis*. In this species, the gene *rps19* was located 87 base pairs away from the JLB boundary, which subsequently resulted in a JSB boundary shift. This shift incorporated an additional 5 base pairs of the *ycf1* gene while simultaneously omitting 5 base

pairs of the *ndhF* from the IRb region. Another notable variance was observed in *M. longifolia*. Here, the annotated *ycf1* gene was extended by 61 base pairs relative to its counterparts in other *Mentha* species. This elongation led to its presence in the SSC region, yet it did so without causing any shift in the JSB boundary (Fig. 4).

Genetic polymorphism and substitution rates (dS/dN)

Single Nucleotide Polymorphisms (SNPs) were identified in various *Mentha* species when compared against the *M. suaveolens* that was used as a reference genome. In the Coding Sequence (CDS) region, the total number of detected SNPs was 190 (73 transversions and 117 transitions), of which 97 were causing a nonsynonymous mutation in the protein sequence. Across the six *Mentha* species, *M. canadensis* exhibits the highest SNP count with 140, followed by *M. longifolia* at 52. *M. x piperita*, *M. spicata*, and *M. x villosa* each display 48 SNPs, while

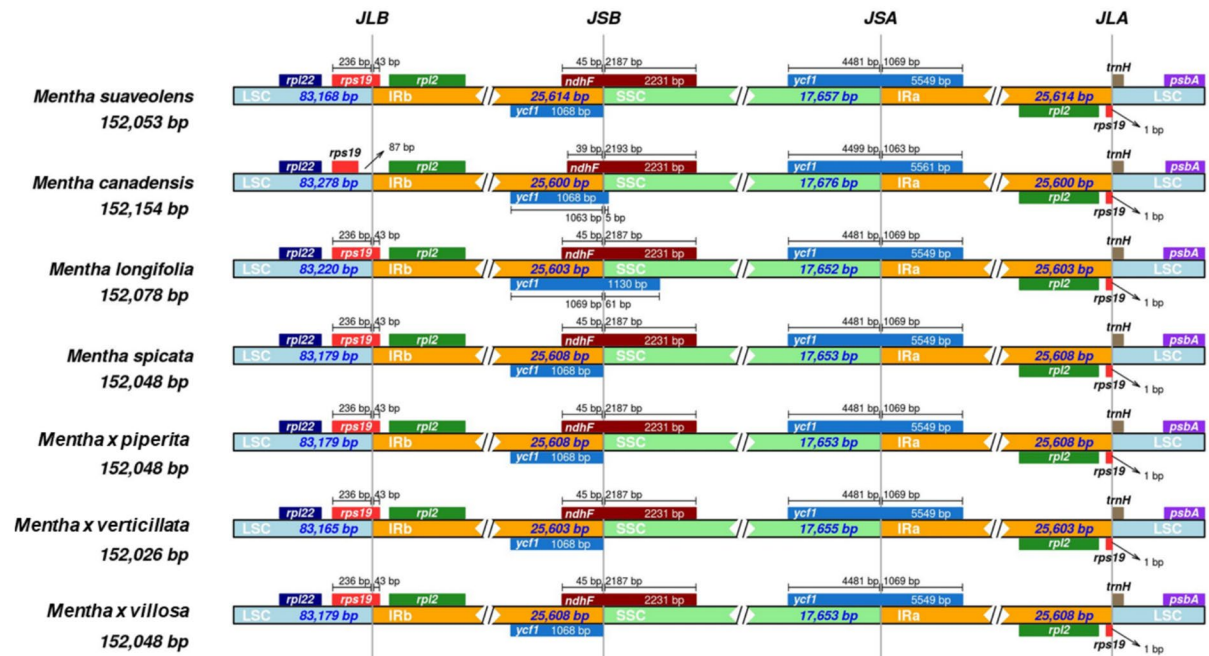


Fig. 4 Comparisons of LSC, SSC, and IRs junctions among the *Mentha* species. Colored boxes represent genes, while arrows show the coordinate positions of each gene near the

junctions. Abbreviations denote the junction site of the plastid genome JLA (IRa/LSC), JLB (IRb/LSC), JSA (SSC/IRa), and JSB (IRb/SSC)

M. × verticillata reveals the least with 25 SNPs. Delving into the specifics of SNP types within the CDS, transversion SNPs for *M. canadensis* are reported at 53, whereas *M. × piperita*, *M. spicata*, and *M. × villosa* each have 14. The *M. longifolia* manifests 18, and *M. × verticillata* presents 10. In terms of transition SNPs within the CDS, *M. canadensis* once again leads with 87, with *M. × piperita*, *M. spicata*, *M. longifolia*, and *M. × villosa* each recording 34, and *M. × verticillata* at 15. Non-synonymous SNPs within the CDS for *M. canadensis* tally at 76, followed closely by *M. × piperita*, *M. spicata*, and *M. × villosa*, each at 22, with *M. longifolia* slightly lower at 21, and *M. × verticillata* at 15.

In the Intergenic Spacer (IGS) region, the total number of detected SNPs was 277 (129 transversions and 148 transitions). Across the compared *Mentha* species, *M. canadensis* possesses the highest SNP count with 215, in contrast to *M. × piperita*, *M. spicata*, and *M. × villosa* each having 78, *M. longifolia* with 78, and *M. × verticillata* at 39. Furthermore, in the IGS region, transversion SNPs for *M. canadensis* are 93, with *M. × piperita*, *M. spicata*, and *M. × villosa* each having 36, *M. longifolia* with 35, and

M. × verticillata with 21. Transition SNPs in the IGS region for *M. canadensis* count 122, *M. × piperita*, *M. spicata*, and *M. × villosa* each have 42, *M. longifolia* slightly higher at 43, and *M. × verticillata* presents 18. It is noteworthy to mention that SNPs, which were found common across all six *Mentha* species in comparison to *M. suaveolens*, were 14 in the CDS and 24 in the IGS regions (Fig. 5).

At the genes level, the hypothetical protein gene *ycf1* exhibited the highest number of SNPs with a total of 50, followed by the RNA polymerase gene *rpoB* and *rpoC2*, each having 11 SNPs. Several genes, such as the Cytochrome b6-f complex gene *petA*, NADH dehydrogenase genes *ndhA*, *ndhD*, and *ndhF*, and Photosystem I genes *psaA* and *psaB*, had SNPs ranging between 3 to 8. Numerous genes like *atpE*, *infA*, *ycf3*, *rpl14*, *rpl16*, *rpl32*, *ndhJ*, *cemA*, *psbA*, *psbH*, *psbK*, *psbZ*, *rps14*, *rps15*, *rps4*, and *rps8* reported a singular SNP (Fig. 6-SNPs).

Among the *Mentha* species analyzed in relation to *M. suaveolens*, the recorded total of 190 SNPs (93 synonymous (S) and 97 nonsynonymous mutations (N)) were used to calculate the dS/dN ratio for each gene. This data is crucial as the dS/dN ratio,

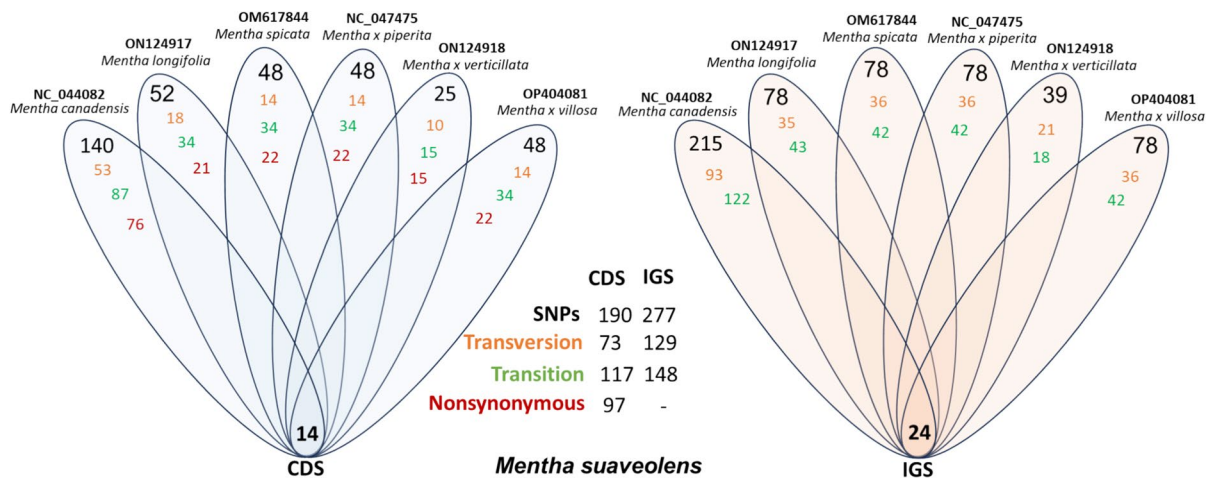


Fig. 5 Genetic diversity of six *Mentha* species compared to *M. suaveolens* chloroplast genome

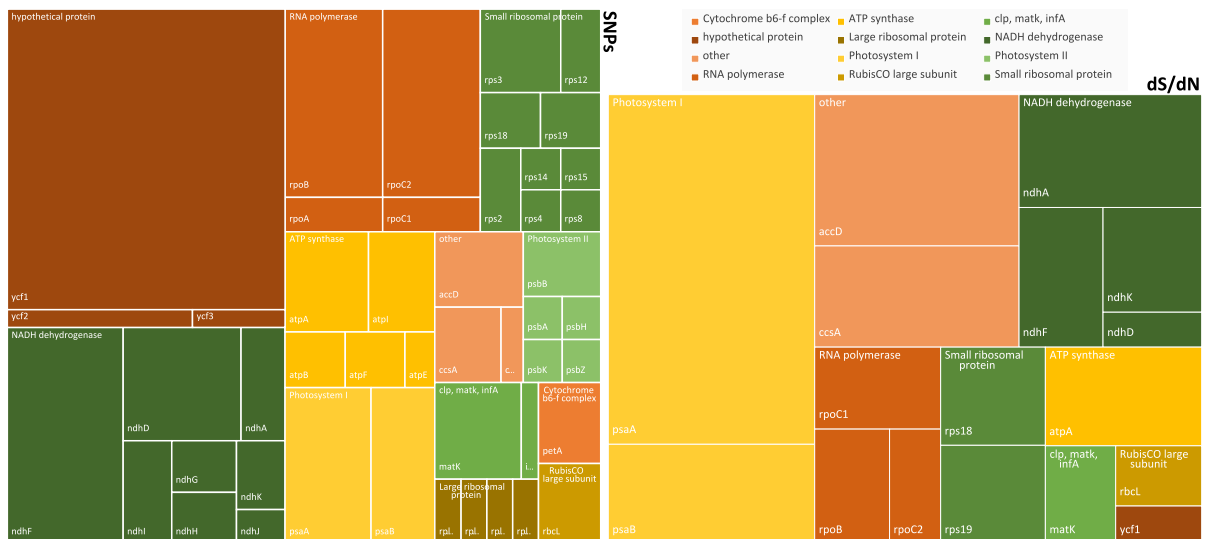


Fig. 6 Genetic polymorphism (SNPs) and genes under selection (dN/dS) ratio of chloroplast protein-coding genes of six *Mentha* species compared to *M. suaveolens* chloroplast

genome. Square size is directly proportional to the number of SNPs and dN/dS ratio. Values presented in the figure are provided in the Supplementary Data

representing the ratio of synonymous to nonsynonymous mutations, is indicative of the selection pressure acting on a gene. The assorted dS/dN ratios spread across the genes signal diverse selection pressures ranging from positive and negative to potentially neutral influences.

The Photosystem I group, containing genes *psaA* and *psaB*, demonstrates dS/dN values of 7 and 2, respectively, both indicative of positive

selection. Conversely, the genes in Photosystem II, including *psbA*, *psbB*, *psbH*, *psbK*, and *psbZ*, all have no nonsynonymous mutations. Within the "other" group, the *accD* gene revealed three synonymous and one nonsynonymous mutation, resulting in a dS/dN ratio of 3. This indicates a pronounced positive selection. Similarly, the *ccsA* gene displayed two synonymous and one nonsynonymous, leading to a dS/dN ratio of 2, which

again implies a positive selection pressure. On the contrary, the *cemA* gene showed only one synonymous SNP with no nonsynonymous mutations, rendering its dS/dN ratio undefined. Meanwhile, in the NADH dehydrogenase group, genes *ndhA* exhibit positive selection with dS/dN values of 2. The gene *ndhF*, however, portrays a slight positive selection with a dS/dN ratio of 1.1429. The RNA polymerase group reveals varied selection pressures across its genes. Notably, *rpoB* has a dS/dN ratio of 0.8333, suggesting slight purifying selection, while *rpoC2* exhibits a ratio of 0.5714, also implying purifying selection.

Genes in the Small ribosomal protein group, such as *rps12* and *rps14*, show dS/dN ratios of zero, indicating purifying selection. In contrast, *rps18*, *rps19*, and *rps8*, similar to multiple previously mentioned genes, lack a dS/dN value due to the absence of nonsynonymous mutations. The ATP synthase group, encompassing genes like *atpA*, *atpB*, *atpE*, *atpF*, and *atpI*, displays varying selection pressures. Specifically, *atpA* has a ratio of 1.5, signifying positive selection, while genes *atpB*, *atpE*, *atpF*, and *atpI*, have no nonsynonymous mutations. The genes within the *clp*, *matK*, and *infA* groups, including *infA* and *matK*, exhibit dS/dN ratios of 0.6667 and zero, respectively, implying purifying selection. In the RubisCO large subunit group, the gene *rbcL* demonstrates negative selection with a dS/dN ratio of 0.5. Among hypothetical proteins, *ycf1* shows a dS/dN ratio of 0.3158, *ycf2*, and *ycf3*, similar to some prior genes having no nonsynonymous mutations.

In the Cytochrome b6-f complex group, the gene *petA* shows three nonsynonymous mutations without any synonymous mutation, resulting in a dS/dN ratio of 0, suggesting purifying selection. Lastly, Large ribosomal proteins, such as *rpl14*, *rpl16*, *rpl32*, and *rpl36*, all lack a dS/dN value due to the absence of either synonymous or nonsynonymous mutations (Fig. 6 dS/dN).

Phylogenetic relationships of *Mentha* species

The Maximum Likelihood (ML) phylogenetic tree, constructed from the complete chloroplast genomes of seven *Mentha* species and rooted with *Ocimum basilicum* from the Lamiaceae family, elucidates the evolutionary relationships among the

Mentha species. In the tree, three species, *M. spicata* (OM617844), *M. × piperita* (NC_047475), and *M. × villosa* (OP404081), form a monophyletic clade with no discernible differences in their chloroplast genome distances. Adjacent to this triad is *M. longifolia* (ON124917), which positions itself as a sister lineage. This placement signifies a near, but slightly more distant, evolutionary kinship with the previously mentioned trio. A distinct branch emerging close to *M. longifolia* is shared by *M. suaveolens* and *M. × verticillata* (ON124918). Further separated from the rest is *M. canadensis* (NC_044082).

Upon comparing the phylogenetic trees derived from various chloroplast regions to the one constructed using the complete chloroplast genome, distinct insights are discerned concerning the evolutionary relationships among the *Mentha* species. The tree derived from the LSC region aligns more closely with the complete chloroplast genome tree, especially regarding the relationship between *M. suaveolens* and *M. × verticillata*. In this case, the congruence between the two trees underscores the significant influence of the LSC region on the overall phylogenetic structure. The SSC region-based phylogenetic tree echoes the pattern seen in the full chloroplast genome tree, with minor deviations. The close association of *M. suaveolens* with *M. × verticillata* is again evident. However, the branch lengths, which are indicative of the evolutionary distance, differ slightly, suggesting the SSC region might encompass unique evolutionary signals distinct from the entire genome. The IR region-based phylogenetic tree showed a chaining effect in the ML based method, which was reconstructed with ward method, and demonstrated a very similar pattern to the full chloroplast genome tree, except of the *M. longifolia* which was not disguisable from *M. spicata*, *M. × piperita*, and *M. × villosa* cluster. This deviation suggests that the IR region offers a contrasting evolutionary perspective for *M. longifolia* compared to the entire chloroplast genome (Fig. 7).

Discussion

Taxonomic and evolutionary studies have used chloroplast genomes to determine genome structure and determine evolutionary relationships between closely related species (Mehmood et al. 2020; Henriquez et al. 2020). The *Mentha* genus, well-known

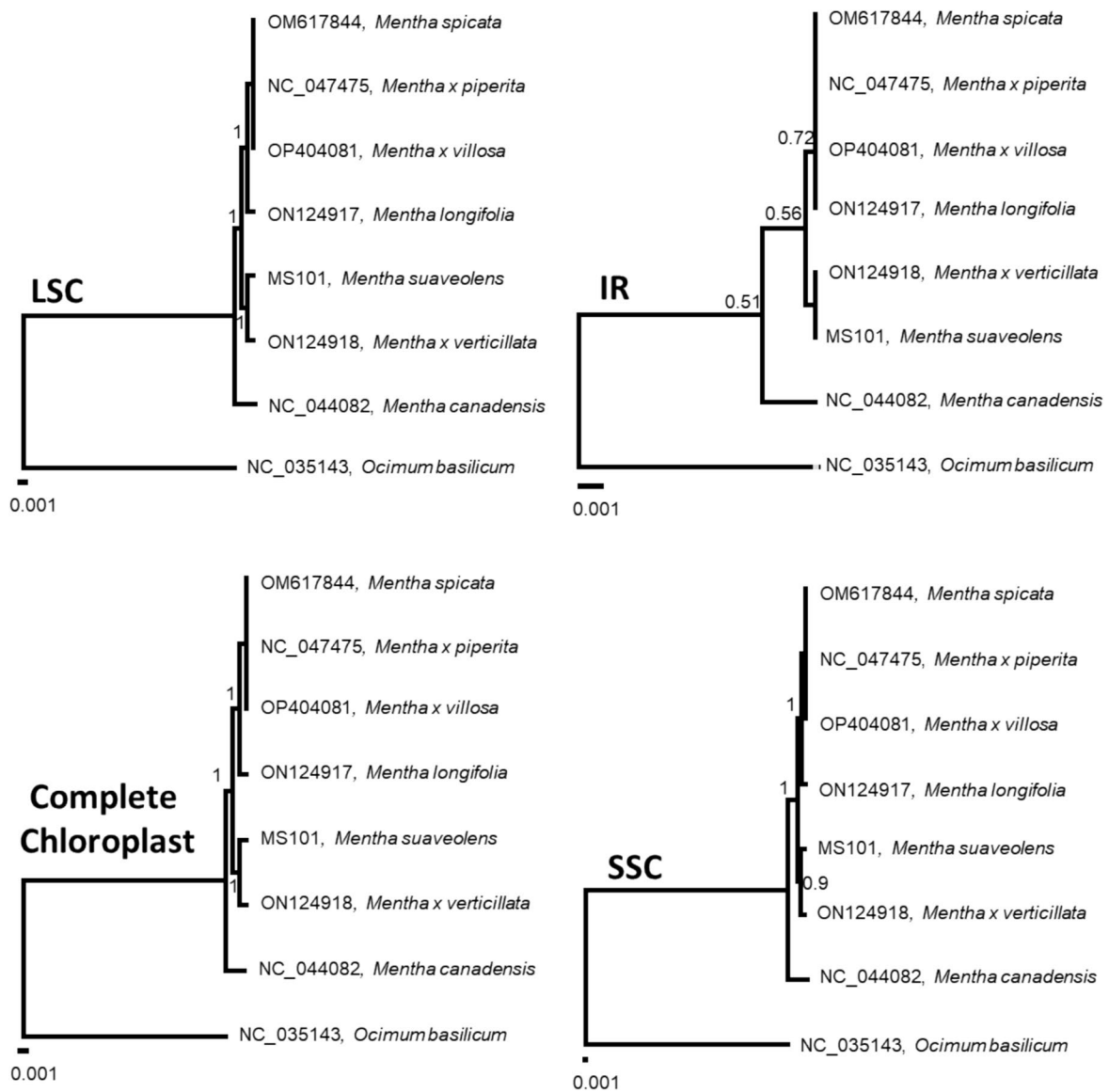


Fig. 7 Rooted molecular phylogenetic tree of *Mentha* species based on ML methods (LSC, SSC and complete chloroplast) and Ward method (IR) to overcome chaining effect showed

by ML for the IR region. Bootstrap values are shown on each node, and the trees were scaled at 0.001

for its economic and therapeutic significance, has been a subject of fascination for biologists due to its diverse array of species (Salehi et al. 2018; Tafrihi et al. 2021). In our pursuit to unravel the genetic foundation behind this diversity, we conducted a comprehensive plastome analysis of seven *Mentha* species, including the newly sequenced *M. suaveolens*. This study provides valuable insights

into genetic conservation, evolution, and diversity within the genus.

One of the most striking observations in our analysis was the remarkable conservation in genome size across the analyzed species. The genome sizes hovered around 152k base pairs, suggesting that significant genomic contractions or expansions have been rare in the evolutionary history of the *Mentha* genus (Soltis et al. 2003; Jin et al. 2014; Vining et al. 2019).

This stability in genome size implies the presence of a fundamental genomic backbone that has remained largely undisturbed over evolutionary timescales. The GC content of a genome is an important metric that influences DNA stability, molecular structure, and vulnerability to mutations (Bunsawat et al. 2004; Sharp et al. 2010). Our study found a narrow range of GC content across the studied species, primarily around 37.8%. However, *M. suaveolens* and *M. × verticillata* displayed a slightly higher GC content of 37.9%, hinting at nuanced genomic alterations or unique evolutionary pressures in these species (Huaizhu et al. 2020; Zubair Filimban et al. 2022). The LSC and SSC in chloroplast genomes typically house essential genes, and their conservation is of paramount importance (Ruhlman and Jansen 2014; Rogalski et al. 2015; Henriquez et al. 2020; Sibbald and Archibald 2020). Our analysis showed high uniformity across these regions in the *Mentha* genus, with minor differences indicating specific adaptations. The IR regions varied slightly, with *M. suaveolens* having the largest segment. These regions play key roles in genome stability and plastid evolution (Vining et al. 2019; Balaji et al. 2021).

Despite the species diversity within the *Mentha* genus, our analysis revealed consistent gene content across all analyzed species, including protein-coding genes, tRNAs, and rRNAs (Wu et al. 2022). This highlights the strong genetic foundation of the *Mentha* genus and the essential role of these genes in its metabolism and life cycle. Simple Sequence Repeats (SSRs) serve as key markers for genetic diversity, evolution, and phylogenetic studies (Vining et al. 2019; Huang et al. 2020; Liang et al. 2021). Within the chloroplast genome of *M. suaveolens*, we identified 420 SSRs, shedding light on the genome's intricate and dynamic nature. These SSRs exhibited various repeat units, ranging from mono to decanucleotide, reflecting genomic variety and functional diversity. Penta-nucleotide repeats were particularly abundant, accounting for nearly a third of the total identified SSRs (George et al. 2015; Maheswari et al. 2021). Tetranucleotide repeats also represented a substantial fraction. Interestingly, with increasing repeat unit length, there was a decline in frequency, except for decanucleotide repeats, which marked their presence as outliers (Balaji et al. 2021). SSRs were predominantly found in the LSC region, which accounted for over 70% of the total, suggesting it

as a hotspot for genetic variation and evolution. In contrast, the IR and SSC regions showed fewer repeats and appeared more stable. Common motifs like "AT," "AAT," and "AAAT" suggest conserved or favored genomic patterns through evolutionary time (Coyne et al. 2011; Lichman et al. 2020). The LSC region consistently harbored a majority of these dominant motifs, suggesting its role as a genetic reservoir (Vining et al. 2019; Cui et al. 2021). Comparing SSR distributions across different *Mentha* species revealed intriguing insights. The uniformity in the frequency distribution of dinucleotide and trinucleotide repeats suggested shared evolutionary pressures or genetic conservation mechanisms among these species (Liang et al. 2019; Niu et al. 2023). In contrast, variations in other repeat units, such as tetra, penta, and hexanucleotide SSRs, highlighted unique genomic peculiarities specific to each species. The prevalence of hexanucleotide repeats, both in frequency and distribution, merits further exploration for their potential role in species-specific functionalities or adaptations (Dong et al. 2020; Fu et al. 2022).

Although chloroplast genomes are often considered stable and conserved, our analysis revealed intriguing dynamism at the boundaries of the IR regions. These boundaries play a vital role in determining plastome length and have undergone subtle changes over evolutionary time (Moore et al. 2007; Turmel et al. 2017; Wang et al. 2017). The junctions (JLB, JSA, JSB, and JLA) demarcating transitions between different chloroplast genome regions were analyzed across *Mentha* species. Most species exhibited consistency in these junction locations, but *M. canadensis* stood out as an anomaly, with repositioning of the gene *rps19* leading to intricate rearrangements in the genome (Huaizhu et al. 2020). The *M. longifolia* displayed unique elongation of the *ycf1* gene into the SSC region, showcasing the genome's ability to accommodate genetic novelties (Bunsawat et al. 2004).

The Relative Synonymous Codon Usage (RSCU) values provided insights into the genetic nuances and preferences of *Mentha* species (Khan et al. 2019). Consistent codon usage for methionine and tryptophan across species suggests conserved evolutionary traits. Amino acids with multiple codons—like leucine, serine, and arginine—showed patterns favoring specific third-position nucleotides, likely driven by structural, translational, or evolutionary factors.

While overall codon usage trends were shared, subtle species-specific differences, particularly in *M. suaveolens*, reflect the genome's fine-tuned codon optimization (Sharp et al. 2010). These nuanced differences could be indicative of differential, adaptive pressures, or genetic drifts influencing codon preferences.

Exploring SNPs within *Mentha* species, when compared to the reference genome of *M. suaveolens*, shed light on genetic diversity and evolution within the genus. The CDS region contained a total of 190 SNPs, with *M. canadensis* having the highest count of 140, reflecting its unique evolutionary journey. Transition and transversion SNPs distribution across species indicated varied genomic alterations, with *M. canadensis* consistently leading in counts. Non-synonymous SNPs within the CDS region were of particular interest, as they can alter protein structures and functions (Huaizhu et al. 2020). The *M. canadensis* exhibited a significant count of 76 non-synonymous SNPs, suggesting potential differences in protein functionalities compared to other *Mentha* species. In the IGS region, SNP counts soared to 277, potentially influencing gene regulation and genome stability. Common SNPs shared across all six *Mentha* species suggested shared evolutionary experiences or conserved genetic alterations with functional significance.

The hypothetical protein gene *ycf1* exhibited the highest number of SNPs followed by the RNA polymerase gene *rpoB* and *rpoC2*, which can be used as molecular markers for population analysis, molecular identification, and species delimitation in the genus *Mentha* (Rogalski et al. 2015; Welch et al. 2016). The dS/dN ratios provided insights into the selective pressures on genes. According to Nelson (2005), positive selection was indicated by dS/dN values greater than one, while values less than one implied purifying or negative selection. Genes like *psaA* and *psaB* in Photosystem I exhibited positive selection, while other genes in Photosystem II and other categories displayed purifying selection, emphasizing their vital roles (Xu et al. 2015; Jiang et al. 2017).

Evolution's tale, encoded in DNA, unfolds through ML phylogenetic trees from chloroplast genomes (George et al. 2015). These trees illuminate *Mentha* species' connections and deviations. In the ML tree, *M. spicata*, *M. × piperita*, and *M. × villosa* form a closely-knit group with near-identical chloroplast genomes, as well as the *M. suaveolens* and

M. × verticillata shared a distinct branch, suggesting a strong genetic bond and likely share maternal ancestry for the hybrid species (Jin et al. 2014), a new information to clarify based on the this study. The *M. longifolia*, positioned as a sister lineage to the *M. spicata* clade, indicates a slightly more distant genetic relationship. In contrast, *M. canadensis* appears as a distinct lineage, suggesting early divergence and a separate evolutionary path, possibly due to limited maternal gene flow from other *Mentha* taxa. The plastid phylogeny helps resolve incongruences observed in AFLP-based studies (Gobert et al. 2002), underscoring the value of chloroplast genomes in tracing maternal lineages—unlike AFLP, which reflects nuclear variation. Similarly, recent ITS-based phylogenies failed to accurately assign hybrid species to their maternal origins, even when supported by morphological data (Naseem et al. 2025). Comparative phylogenetics provides nuanced insights, emphasizing the LSC region's significant evolutionary narratives. Contrasts in the SSC region tree, with minor branch length variations, hint at unique evolutionary events specific to this region. Regardless of the detected changing effect in ML tree, that was resolved by the ward method, the IR region showed a deviation from the full chloroplast genome tree for *M. longifolia*, prompting consideration of selective pressures or historical events shaping the IR region (Chen et al. 2022).

Conclusion

In summary, the comparative study of *Mentha* genus chloroplast genomes reveals a lineage marked by genomic conservation with occasional variations. This conservation, especially in essential genomic regions, highlights their evolutionary importance, laying the groundwork for future studies and biotechnological applications. Exploring plastomic intricacies, even when observations suggest uniformity, may reveal keys to understanding evolution and adaptation. *Mentha suaveolens*' chloroplast genome shows diverse SSR motifs, mainly in the LSC region. Trends in SSR distribution balance conservation and flexibility, aiding adaptation. Nuanced SSR differences among *Mentha* species enrich our understanding. Comparative analysis of RSCU values shows predominantly conserved patterns, with occasional

variation. These deviations hint at adaptive strategies and evolution. Exploring the chloroplast genome's IR region reveals conservation and change. Genetic differences in *M. canadensis* and *M. longifolia* prompt inquiries into functional implications. SNPs unveil the complexity of *Mentha* genus evolution, with some genes conserved and others rapidly evolving to adapt to environmental challenges. Phylogenetic exploration highlights intricate relationships among *Mentha* species and the potential “maternal” origin of the analyzed chloroplast genome of three established hybrids *M. × piperita*, and *M. × villosa*, *M. × verticillata*, emphasizing genetic authenticity and the multifaceted nature of evolution. In essence, *Mentha* species epitomize the dynamic dance of evolution.

Acknowledgements We extend our deepest gratitude to Prof. Dr. Rosa Maria Ros of Murcia University for her meticulous revision of the manuscript. Her expertise and constructive criticism significantly improved the quality of this work.

Author contribution Conceptualization, MM; methodology, KMA, MM; software, MM; validation, MM, HM, DA; formal analysis, KMA, AA, SMA; investigation, KMA, MM; resources, MM, DA; data curation, MM; writing—original draft preparation, MM, HM; writing—review and editing, KMA, AA, SMA, DA; visu-alization, MM; supervision, MM, DA; project administration, KMA; funding acquisition, KMA, HM, AA, SMA. All authors has read and agreed to the published version of the manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This research was funded by Fundación Séneca – Región de Murcia through the Saavedra Fajardo program, supporting the research contract of Mahmoud Magdy (grant number 22416/SF/23).

Data availability The assembled chloroplast genome is freely available at: <https://doi.org/10.6084/m9.figshare.24999416>.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds

the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Asao T, Kuwano H, Ide M et al (2003) Spasmolytic effect of peppermint oil in barium during double-contrast barium enema compared with Buscopan. Clin Radiol 58:301–305. [https://doi.org/10.1016/S0009-9260\(02\)00532-9](https://doi.org/10.1016/S0009-9260(02)00532-9)
- Attiya J (2012) Phylogenetics of selected *Mentha* species on the basis of rps8, rps11 and rps14 chloroplast genes. J Med Plants Res. <https://doi.org/10.5897/JMPR11.658>
- Balaji R, Ravichandiran K, Tanuja PM (2021) The complete chloroplast genome of *Ocimum gratissimum* from India—a medicinal plant in the Lamiaceae. Mitochondr DNA Part B 6:948–950. <https://doi.org/10.1080/23802359.2021.1889413>
- Beier S, Thiel T, Münch T et al (2017) MISA-web: a web server for microsatellite prediction. Bioinformatics 33:2583–2585. <https://doi.org/10.1093/bioinformatics/btx198>
- Bunsawat J, Elliott NE, Hertweck KL et al (2004) Phylogenetics of *Mentha* (Lamiaceae): evidence from chloroplast DNA sequences. Syst Bot 29:959–964. <https://doi.org/10.1600/0363644042450973>
- Celenk S, Tarimcilar G, Bicakci A et al (2008) A palynological study of the genus *Mentha* L. (Lamiaceae). Botan J Linn Soc 157:141–154. <https://doi.org/10.1111/j.1095-8339.2008.00789.x>
- Chen X, Zhang F, Yao L (2012) Chloroplast DNA molecular characterization and leaf volatiles analysis of mint (*Mentha*; Lamiaceae) populations in China. Ind Crops Prod 37:270–274. <https://doi.org/10.1016/j.indcrop.2011.11.011>
- Chen J, Xie D, He X et al (2022) Comparative analysis of the complete chloroplast genomes in allium section bromatorrhiza species (Amaryllidaceae): phylogenetic relationship and adaptive evolution. Genes 13:1279. <https://doi.org/10.3390/genes13071279>
- Coyne RS, Hannick L, Shanmugam D et al (2011) Comparative genomics of the pathogenic ciliate *Ichthyophthirius multifiliis*, its free-living relatives and a host species provide insights into adoption of a parasitic lifestyle and prospects for disease control. Genome Biol 12:R100. <https://doi.org/10.1186/gb-2011-12-10-r100>
- Cui G, Wang C, Wei X et al (2021) Complete chloroplast genome of *Hordeum brevisubulatum*: genome organization, synonymous codon usage, phylogenetic relationships, and comparative structure analysis. PLoS ONE 16:e0261196. <https://doi.org/10.1371/journal.pone.0261196>
- Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394–1403. <https://doi.org/10.1101/gr.2289704>
- Dong W, Xu C, Wen J, Zhou S (2020) Evolutionary directions of single nucleotide substitutions and structural mutations in the chloroplast genomes of the family

- Calycanthaceae. BMC Evol Biol 20:96. <https://doi.org/10.1186/s12862-020-01661-0>
- El-Ghorab AH (2006) The chemical composition of the *Mentha pulegium* L. Essential oil from Egypt and its antioxidant activity. J Essent Oil Bear Plants 9:183–195. <https://doi.org/10.1080/0972060X.2006.10643491>
- El-Kashoury E-SA, El-Askary HI, Kandil ZA, Salem MA (2013) Botanical and genetic characterization of *Mentha suaveolens* Ehrh. cultivated in Egypt. Pharmacogn J 5:228–237. <https://doi.org/10.1016/j.phcgj.2013.10.002>
- Fadhel NB, Boussaïd M (2004) Genetic diversity in wild Tunisian populations of *Mentha pulegium* L. (Lamiaceae). Genet Resour Crop Evol 51:309–321. <https://doi.org/10.1023/B:GRES.0000024016.13743.8b>
- Fu G, Liu Y, Caraballo-Ortiz MA et al (2022) Characterization of the complete chloroplast genome of the dragon-head herb, *Dracocephalum heterophyllum* (Lamiaceae), and comparative analyses with related species. Diversity 14:110. <https://doi.org/10.3390/d14020110>
- George B, Bhatt BS, Awasthi M et al (2015) Comparative analysis of microsatellites in chloroplast genomes of lower and higher plants. Curr Genet 61:665–677. <https://doi.org/10.1007/s00294-015-0495-9>
- Gobert V, Moja S, Colson M, Taberlet P (2002) Hybridization in the section *Mentha* (Lamiaceae) inferred from AFLP markers. Am J Bot 89:2017–2023. <https://doi.org/10.3732/ajb.89.12.2017>
- Henriquez CL, Abdullah AI et al (2020) Evolutionary dynamics of chloroplast genomes in subfamily Aroideae (Araceae). Genomics 112:2349–2360. <https://doi.org/10.1016/j.ygeno.2020.01.006>
- Huaizhu L, Lin Y, Bai J et al (2020) The complete chloroplast genome sequence of *Mentha canadensis* (Labiatae), a traditional Chinese herbal medicine. Mitochondrial DNA Part B 5:55–56. <https://doi.org/10.1080/23802359.2019.1687031>
- Huang S, Ge X, Cano A et al (2020) Comparative analysis of chloroplast genomes for five *Dicliptera* species (Acanthaceae): molecular structure, phylogenetic relationships, and adaptive evolution. PeerJ 8:e8450. <https://doi.org/10.7717/peerj.8450>
- Jiang D, Zhao Z, Zhang T et al (2017) The Chloroplast genome sequence of *Scutellaria baicalensis* provides insight into intraspecific and interspecific chloroplast genome diversity in Scutellaria. Genes 8:227. <https://doi.org/10.3390/genes8090227>
- Jin J, Panicker D, Wang Q et al (2014) Next generation sequencing unravels the biosynthetic ability of Spearmint (*Mentha spicata*) peltate glandular trichomes through comparative transcriptomics. BMC Plant Biol 14:292. <https://doi.org/10.1186/s12870-014-0292-5>
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. <https://doi.org/10.1093/molbev/mst010>
- Kearse M, Moir R, Wilson A et al (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
- Khan A, Asaf S, Khan AL et al (2019) Complete chloroplast genomes of medicinally important *Teucrium* species and comparative analyses with related species from Lamiaceae. PeerJ 7:e7260. <https://doi.org/10.7717/peerj.7260>
- Khanuja SPS, Shasany AK, Srivastava A, Kumar S (2000) No title found. Euphytica 111:121–125. <https://doi.org/10.1023/A:1003829512956>
- Kumar S, Stecher G, Li M et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Lachance PA, Nakat Z, Jeong W-S (2001) Antioxidants: an integrative approach. Nutrition 17:835–838. [https://doi.org/10.1016/S0899-9007\(01\)00636-0](https://doi.org/10.1016/S0899-9007(01)00636-0)
- Liang C, Wang L, Lei J et al (2019) A comparative analysis of the chloroplast genomes of four salvia medicinal plants. Engineering 5:907–915. <https://doi.org/10.1016/j.eng.2019.01.017>
- Liang C, Wang L, Ma W, Xu J (2021) A comparative study of complete chloroplast genome for the genus salvia. J Plant Biochem Biotechnol 30:117–125. <https://doi.org/10.1007/s13562-020-00575-8>
- Lichman BR, Godden GT, Buell CR (2020) Gene and genome duplications in the evolution of chemodiversity: perspectives from studies of Lamiaceae. Curr Opin Plant Biol 55:74–83. <https://doi.org/10.1016/j.pbi.2020.03.005>
- López V, Martín S, Gómez-Serranillos MP et al (2010) Neuroprotective and neurochemical properties of mint extracts: neurochemical properties of mints. Phytother Res 24:869–874. <https://doi.org/10.1002/ptr.3037>
- Lowe TM, Chan PP (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucl Acids Res 44:W54–W57. <https://doi.org/10.1093/nar/gkw413>
- Magdy M, Ouyang B (2020) The complete mitochondrial genome of the chiltepin pepper (*Capsicum annuum* var. *glabriusculum*), the wild progenitor of *Capsicum annuum* L. Mitochondrial DNA Part B 5:683–684
- Magdy M, Ou L, Yu H et al (2019) Pan-plastome approach empowers the assessment of genetic variation in cultivated *Capsicum* species. Hortic Res 6:108. <https://doi.org/10.1038/s41438-019-0191-x>
- Magdy M, El-Sherbeny EA, Ramirez Sanchez A (2022) The complete chloroplast genome of the Egyptian henbane (*Hyoscyamus muticus* L., Solanaceae). Mitochondrial DNA Part B 7:1109–1111. <https://doi.org/10.1080/23802359.2022.2086493>
- Maheswari P, Kunhikannan C, Yasodha R (2021) Chloroplast genome analysis of Angiosperms and phylogenetic relationships among Lamiaceae members with particular reference to teak (*Tectona grandis* L.f.). J Biosci 46:43. <https://doi.org/10.1007/s12038-021-00166-2>
- Mehmood F, Abdullah SI et al (2020) Characterization of *Withania somnifera* chloroplast genome and its comparison with other selected species of Solanaceae. Genomics 112:1522–1530. <https://doi.org/10.1016/j.ygeno.2019.08.024>
- Moore MJ, Bell CD, Soltis PS, Soltis DE (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proc Natl Acad Sci USA

- 104:19363–19368. <https://doi.org/10.1073/pnas.0708072104>
- Moreno L, Bello R, Primo-Yfera E, Esplugues J (2002) Pharmacological properties of the methanol extract from *Mentha suaveolens* Ehrh. *Phytother Res* 16:10–13. <https://doi.org/10.1002/ptr.744>
- Nagell A, Hefendehl FW, Hoyer J (1974) Zwei stereoisomere 1,2-Epoxymenthylacetate aus einem Öl von *Mentha rotundifolia* × *Mentha longifolia* Two Stereoisomeric 1,2-Epoxymenthylacetates from an Oil of *Mentha rotundifolia* × *Mentha longifolia*. *Z Naturforschung C* 29:294–295. <https://doi.org/10.1515/znc-1974-5-620>
- Naseem I, Khan MA, Habib U et al (2025) Morphological profiling and DNA barcoding revealed genetic diversity and phylogeny of *Mentha* species cultivated in Pakistan. *Genet Resour Crop Evol* 72:2977–2995. <https://doi.org/10.1007/s10722-024-02140-x>
- Nielsen R (2005) Molecular signatures of natural selection. *Annu Rev Genet* 39:197–218. <https://doi.org/10.1146/annurev.genet.39.073003.112420>
- Niu Y, Qin Q, Dong Y et al (2023) Chloroplast genome structure and phylogenetic analysis of 13 lamiaceae plants in tibet. *Front Biosci Landmark* 28:110. <https://doi.org/10.31083/j.fbl2806110>
- Oumzil H, Ghoulami S, Rhajaoui M et al (2002) Antibacterial and antifungal activity of essential oils of *Mentha suaveolens*. *Phytother Res* 16:727–731. <https://doi.org/10.1002/ptr.1045>
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>
- Rogalski M, Do Nascimento Vieira L, Fraga HP, Guerra MP (2015) Plastid genomics in horticultural species: importance and applications for plant population genetics, evolution, and biotechnology. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2015.00586>
- Ruhlman TA, Jansen RK (2014) The plastid genomes of flowering plants. In: Maliga P (ed) *Chloroplast biotechnology*. Humana Press, Totowa, pp 3–38
- Salehi B, Stojanović-Radić Z, Matejić J et al (2018) Plants of genus *Mentha*: from farm to food factory. *Plants* 7:70. <https://doi.org/10.3390/plants7030070>
- Šarić-Kundalić B (2009) Multivariate numerical taxonomy of *Mentha* species, hybrids, varieties and cultivars. *Sci Pharm* 77:851–876. <https://doi.org/10.3797/scipharm.0905-10>
- Schlötterer C (2004) The evolution of molecular markers—Just a matter of fashion? *Nat Rev Genet* 5:63–69. <https://doi.org/10.1038/nrg1249>
- Sharp PM, Emery LR, Zeng K (2010) Forces that influence the evolution of codon bias. *Philos Trans R Soc B* 365:1203–1212. <https://doi.org/10.1098/rstb.2009.0305>
- Shin MH, Moon YJ, Seo J-E et al (2008) Reactive oxygen species produced by NADPH oxidase, xanthine oxidase, and mitochondrial electron transport system mediate heat shock-induced MMP-1 and MMP-9 expression. *Free Radical Biol Med* 44:635–645. <https://doi.org/10.1016/j.freeradbiomed.2007.10.053>
- Shinwari ZK, Sultan S, Mehmood T (2011) Molecular and morphological characterization of selected *Mentha* species. *Pak J Bot* 43:1433–1436
- Sibbald SJ, Archibald JM (2020) Genomic insights into plastid evolution. *Genome Biol Evol* 12:978–990. <https://doi.org/10.1093/gbe/evaa096>
- Soltis DE, Soltis PS, Bennett MD, Leitch IJ (2003) Evolution of genome size in the angiosperms. *Am J Bot* 90:1596–1603. <https://doi.org/10.3732/ajb.90.11.1596>
- Son D, Kim M, Woo H et al (2018) Anti-thermal skin aging activity of aqueous extracts derived from Apple Mint (*Mentha suaveolens* Ehrh.) in human dermal fibroblasts. *Evid-Based Complement Altern Med* 2018:1–7. <https://doi.org/10.1155/2018/4595982>
- Tafrihi M, Imran M, Tufail T et al (2021) The wonderful activities of the genus *Mentha*: not only antioxidant properties. *Molecules* 26:1118. <https://doi.org/10.3390/molecules26041118>
- Tillich M, Lehwark P, Pellizzer T et al (2017) GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Res* 45:W6–W11. <https://doi.org/10.1093/nar/gkx391>
- Turmel M, Otis C, Lemieux C (2017) Divergent copies of the large inverted repeat in the chloroplast genomes of ulvophyceae green algae. *Sci Rep* 7:994. <https://doi.org/10.1038/s41598-017-01144-1>
- Umez T, Sakata A, Ito H (2001) Ambulation-promoting effect of peppermint oil and identification of its active constituents. *Pharmacol Biochem Behav* 69:383–390. [https://doi.org/10.1016/S0091-3057\(01\)00543-3](https://doi.org/10.1016/S0091-3057(01)00543-3)
- Vining KJ, Pandelova I, Hummer K et al (2019) Genetic diversity survey of *Mentha aquatica* L. and *Mentha suaveolens* Ehrh., mint crop ancestors. *Genet Resour Crop Evol* 66:825–845. <https://doi.org/10.1007/s10722-019-00750-4>
- Voirin B, Bayet C, Faure O, Jullien F (1999) Free flavonoid aglycones as markers of parentage in *Mentha aquatica*, *M. citrata*, *M. spicata* and *M. × piperita*. *Phytochemistry* 50:1189–1193. [https://doi.org/10.1016/S0031-9422\(98\)00672-4](https://doi.org/10.1016/S0031-9422(98)00672-4)
- Wang K, Li L, Hua Y et al (2017) The complete chloroplast genome of *Mentha spicata*, an endangered species native to South Europe. *Mitochondrial DNA Part B* 2:907–909. <https://doi.org/10.1080/23802359.2017.1413311>
- Welch AJ, Collins K, Ratan A et al (2016) Data characterizing the chloroplast genomes of extinct and endangered Hawaiian endemic mints (Lamiaceae) and their close relatives. *Data Brief* 7:900–922. <https://doi.org/10.1016/j.dib.2016.03.037>
- Wu J, Li H, Lei J, Liang Z (2022) The complete chloroplast genome sequence of *Trigonotis peduncularis* (Boraginaceae). *Mitochondrial DNA Part B* 7:456–457. <https://doi.org/10.1080/23802359.2022.2048212>
- Xu J-H, Liu Q, Hu W et al (2015) Dynamics of chloroplast genomes in green plants. *Genomics* 106:221–231. <https://doi.org/10.1016/j.ygeno.2015.07.004>
- Yang Z, Rannala B (2012) Molecular phylogenetics: principles and practice. *Nat Rev Genet* 13:303–314. <https://doi.org/10.1038/nrg3186>
- Yao J, Zhao F, Xu Y et al (2020) Complete chloroplast genome sequencing and phylogenetic analysis of two

- dracocephalum plants. Biomed Res Int 2020:1–9. <https://doi.org/10.1155/2020/4374801>
- Zhu A, Guo W, Gupta S et al (2016) Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. New Phytol 209:1747–1756. <https://doi.org/10.1111/nph.13743>
- Zubair Filimban F, Yarádua SS, Bello A, Choudhry H (2022) Complete chloroplast genomes of two mint species (Lamiaceae) from Al-Madinah, Saudi Arabia: phylogenetic and genomic comparative analyses. Mitochondrial DNA Part B 7:1797–1799. <https://doi.org/10.1080/23802359.2022.2130713>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.