View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

THE CROP JOURNAL 3 (2015) 335–342

HOSTED BY

Available online at www.sciencedirect.com

ScienceDirect

Identification of EST–SSRs and molecular diversity analysis in *Mentha piperita*

Birendra Kumar[a], *Umesh Kumar*[b], *Hemant Kumar Yadav*[b,c,*]

[a]*CSIR-Central Institute of Medicinal and Aromatic Plant Sciences, P.O. CIMAP, Near Kukrail Picnic Spot, Lucknow-226 015, India*
[b]*CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow-226 001, India*
[c]*Academy of Scientific and Innovative Research (AcSIR), New Delhi, India*

ARTICLE INFO

Article history:
Received 8 July 2014
Received in revised form
2 February 2015
Accepted 16 February 2015
Available online 28 February 2015

Keywords:
EST–SSRs
Genetic diversity
Mentha piperita
Polymorphic information content
Transferability

ABSTRACT

EST sequences of *Mentha piperita* available in the public domain (NCBI) were exploited to develop SSR markers. A total of 1316 ESTs were assembled into 155 contigs and 653 singletons and of these, 110 sequences were found to contain 130 SSRs, with a frequency of 1 SSR/3.4 kb. Dinucleotide repeat SSRs were most frequent (72.3%) with the AG/CT (43.8%) repeat motif followed by AT/AT (16.2%). Primers were successfully designed for 68 SSR-containing sequences (62.0%). The 68 primers amplified 13 accessions of *M. piperita* and 54 produced clear amplicons of the expected size. Of these 54, 33 (61%) were found to be polymorphic among *M. piperita* accessions, showing from 2 to 4 alleles with an average of 2.33 alleles/SSR, and the polymorphic information content (PIC) value varied between 0.13 and 0.51 (average 0.25). All the amplified SSRs showed transferability among four different species of *Mentha*, with a highest in *Mentha arvensis* (87.0%) and minimum in *Mentha citrata* (37.0%). The newly developed SSRs markers were found to be useful for diversity analysis, as they successfully differentiated among species and accessions of *Mentha*.

© 2015 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The genus *Mentha* (mint) is one of the most important taxa of the family Lamiaceae and comprises 25 to 30 species grown in different parts of the world. Only five of them, Japanese/menthol mint (*Mentha arvensis* L. var. *piperascens forma* Holmes), spearmint (*Mentha spicata* L.), peppermint (*Mentha piperita* L.), scotch spearmint (*Mentha cardiaca* Baker.) and Bergamot mint (*Mentha citrata* Ehrh.) are commercially grown in India and other countries [1,2]. These five species are the major natural source of aroma compounds of industrial importance; namely-menthol, menthofuran, carvone, linalool, and linalyl acetate. Because of their cooling, pleasant aroma and flavor, essential oils of mint are used in perfumery, cosmetics, confectionery, and the pharmaceutical industries. The oil of *M. piperita*, known as peppermint oil, is widely used for headache, nerve pain, toothache, oral inflammation, joint conditions, itchiness, allergic rash, repelling mosquitoes, rheumatism, muscular pains, etc. [3,4]. Menthol is the major constituent of the essential oil constituents of peppermint oil [5]. Peppermint oil of globally accepted quality contains high amounts of menthol, moderate amounts of menthone, and very low amounts of pulegone and menthofuran [6,7]. The presence and concentrations of certain chemical constituents of essential oils change

* *Corresponding author at:* CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow-226 001, India. Tel.:+91 522 2297938.
E-mail address: h.yadav@nbri.res.in (H.K. Yadav).
Peer review under responsibility of Crop Science Society of China and Institute of Crop Science, CAAS.

http://dx.doi.org/10.1016/j.cj.2015.02.002
2214-5141/© 2015 Crop Science Society of China and Institute of Crop Science, CAAS. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

according to the season, soil, climate, and site of plant growth. Peppermint is cultivated in several parts of India and has great economic value and a strong export potential for its volatile oil extracts.

DNA-based molecular markers have been shown to be an efficient tool to assist conventional plant breeding in various ways, such as by assessing the gene pool for diverse parental lines, hybridity testing, QTL mapping, gene tagging, and marker-assisted selection. Despite the importance of peppermint as an aromatic and medicinal plant, no comprehensive molecular marker systems are available. A few studies have assessed genetic diversity in species of *Mentha* based on RAPD [8–10] and AFLP fingerprinting [11]. There is a complete lack of *Mentha*-specific molecular markers for use in genetic studies and genetic improvement programs.

Simple sequence repeats (SSRs) also called microsatellites, are 1–6-base tandem repeats of DNA sequences, abundant in both prokaryotic and eukaryotic organisms in coding and noncoding regions [12]. SSRs are a preferred marker system owing to their codominant inheritance, multiallelic nature, abundance in the genome, high reproducibility, hyperpolymorphism, and high rate of transferability across genera and species [13–15]. Expressed sequence tags (ESTs) available in the public domain are the easiest and cheapest source for SSR development. EST–SSRs offer various advantages including ease of access, presence in gene-rich regions, and high transferability across species and genera [16], which enable them to serve as anchor markers for comparative mapping and evolutionary studies [15]. Given that no SSR markers are reported in *Mentha*, the present study was undertaken to exploit the EST database of *M. piperita* to (1) analyze the frequency and distribution of SSRs in ESTs, (2) develop and characterize EST–SSRs (3) test their transferability in related species and (4) detect polymorphism/diversity among accessions and species of *Mentha*.

2. Materials and methods

2.1. Plant materials and DNA isolation

The plant material included 13 accessions of *M. piperita*, 5 of *M. arvensis*, 4 of *M. spicata*, and 1 each of *Mentha longifolia* and *M. citrata*. These accessions were previously evaluated for essential oil content and other components. The details of these plant materials are given in Table 1. *M. piperita* is characterized by moderate oil content and high menthofuran. *M. arvensis*, also known as menthol mint, contains comparatively high oil content rich in menthol. The accessions of *M. spicata* and *M. longifolia* contain carvone-rich essential oils and *M. citrata* is rich in linalool and linalyl acetate. These accessions were grown during the 2012–2013 crop season in the experimental field of the Central Institute of Medicinal and Aromatic Plants (CIMAP) Lucknow, Uttar Pradesh, India. Fresh leaves combined from five randomly chosen plants of each accession were used to isolate total genomic DNA with a Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. The quality of DNA was assessed on 0.8% agarose gel and quantity was checked with a NanoDrop spectrophotometer ND 1000 (NanoDrop Products, USA). Finally, the DNA was normalized to 10 ng μL^{-1} for PCR amplification.

2.2. Data mining and EST–SSR identification

A total of 1316 raw EST sequences of *M. piperita* were downloaded from the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/dbest/) on January 14, 2013. The 5′ or 3′ end poly A or poly T stretches were removed from the raw EST sequences using EST-Trimmer software (http://pgrc.ipk-

Table 1 – Details of cultivars and species of *Mentha* used in the study.

Species	Cultivars	Type	Oil content (%)	Origin
M. piperita	CIM-Indus	Cultivar	0.25 (menthofuran 15–25%)	CSIR-CIMAP, Lucknow
M. piperita	CIMAP-Patra	Cultivar	0.35 (menthofuran 40–45%)	CSIR-CIMAP, Lucknow
M. piperita	Kukrail	Cultivar	0.50 (menthofuran 5–8%)	CSIR-CIMAP, Lucknow
M. piperita	Madhuras	Cultivar	0.45 (menthofuran 2–5%)	CSIR-CIMAP, Lucknow
M. piperita	MPS-5	Breeding line	0.40 (menthol 65%)	CSIR-CIMAP, Lucknow
M. piperita	MPS-6	Breeding line	0.45	CSIR-CIMAP, Lucknow
M. piperita	MPS-9	Breeding line	0.35	CSIR-CIMAP, Lucknow
M. piperita	MPS-16	Breeding line	0.45	CSIR-CIMAP, Lucknow
M. piperita	MPS-20	Breeding line	0.45	CSIR-CIMAP, Lucknow
M. piperita	MPS-25	Breeding line	0.30	CSIR-CIMAP, Lucknow
M. piperita	MPS-36	Breeding line	0.35 (menthofuran 30–35%)	CSIR-CIMAP, Lucknow
M. piperita	MPP	Landrace	0.40	Purara, Bageshwar
M. piperita	MPS-21	Breeding line	0.55	CSIR-CIMAP, Lucknow
M. arvensis	Kosi	Cultivar	0.85 (menthol rich)	CSIR-CIMAP, Lucknow
M. arvensis	MAS-1	Cultivar	0.85 (menthol rich)	CSIR-CIMAP, Lucknow
M. arvensis	MAS-49	Breeding line	0.90 (menthol rich)	CSIR-CIMAP, Lucknow
M. arvensis	MAS-10-11-45	Breeding line	0.80 (menthol rich)	CSIR-CIMAP, Lucknow
M. arvensis	MAS-13-2-125	Breeding line	0.80 (menthol rich)	CSIR-CIMAP, Lucknow
M. spicata	Neera	Cultivar	0.55 (marvone rich)	CSIR-CIMAP, Lucknow
M. spicata	Neerkalka	Cultivar	0.60 (marvone rich)	CSIR-CIMAP, Lucknow
M. spicata	Arka	Cultivar	0.75 (marvone rich)	CSIR-CIMAP, Lucknow
M. spicata	MSP	Landrace	0.45 (marvone rich)	Nagar, Kullu, H.P.
M. longifolia	MLP	Landrace	0.25 (carvone rich)	Palampur, H.P.
M. citrata	Kiran	Cultivar	0.35 (minalool and minalyl acetate rich)	CSIR-CIMAP, Lucknow

gatersleben.de/misa/download/est_trimmer.pl). EST sequences were then assembled using the CAP3 assembler [17] with criteria of 40 bp overlap and 90% identity. The assembled EST sequences were subjected to SSR search using MISA (http://pgrc.ipk-gatersleben.de/misa/) with criteria of minimum number of repeats of 5 for dinucleotide (DNR) and trinucleotide (TNR) and of 4 for tetra-, penta-, and hexanucleotide SSRs.

2.3. Primer design and PCR amplification

The SSR-containing sequences were used to design flanking primers with PRIMER3 software (http://frodo.wi.mit.edu/primer3) with major primer design parameters as follows: product length 100–300 bp, primer size 18–25 bp, and melting temperature 57–63 °C (optimum 60 °C). In some cases, where primers could not be designed, the criteria were relaxed. The primers were synthesized with an additional 18-base (5′-TGTAAAACGACGGCCAGT-3′) tag at the 5′ end to all the forward primers [18]. Additionally, four 18-base primer named as "M13 tag" was also synthesized labeled with either FAM, VIC, PET, or NED fluorescent dye. PCR amplification of genomic DNA was performed in a 10 μL reaction volume in an Veriti Thermal Cycler PCR (Applied Biosystems, Foster City, CA, USA) containing 10 ng genomic DNA, 1× PCR Master Mix (AmpliTaq Gold 360), 5 pmol forward primer (tailed with M 13 tag), 15 pmol reverse primer, and 15 pmol "M13 tag". The PCR programs employed initial denaturation for 5 min at 95 °C followed by 35 cycles of denaturation for 1 min at 94 °C, annealing for 45 s at 48–52 °C (primer-specific) and extension for 1 min at 72 °C. These were followed by 10 cycles of denaturation for 30 s at 95 °C, annealing for 45 s at 53 °C, and extension for 45 s at 72 °C followed by a final extension for 12 min at 72 °C. The PCR products were separated by capillary electrophoresis using the ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA). After PCR amplification confirmation on 1.5% agarose gel, post-PCR multiplex sets were constructed based on fluorescence-labeled primer dyes. For post-PCR multiplexing, 1 μL of FAM and 2 μL each of VIC, NED, and PET-labeled PCR products representing different SSRs were mixed with 60 μL water. The mixed product (2 μL) was then added to 8 μL of Hi-Di formamide containing 0.20 μL GeneScan 600 LIZ as internal size standard, denatured for 5 min at 95 °C, quick-chilled on ice for 5 min, and loaded on the ABI3730xl DNA Analyzer for electrophoresis. SSR amplicon size was determined with GeneMapper 4.0 software (Applied Biosystems, USA).

2.4. Data acquisition and statistical analyses

The PCR-amplified SSR markers were scored by allele size as well as presence (1) or absence (0). Statistical analysis for the calculation of observed heterozygosity (Ho), gene diversity or expected heterozygosity (He), major allele frequency, and polymorphic information content (PIC) of EST–SSR markers was performed with Power Marker 3.25 [19]. PIC was calculated following Botstein et al. [20]:

$$\mathrm{PIC} = 1-\left[\sum_{i=1}^{n} P_j^2\right]-\left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2P_i^2P_j^2\right],$$

where P_i and P_j are the frequencies of the ith and jth alleles. A pairwise similarity matrix among the accessions was calculated with the 0–1 data matrix using Jaccard's coefficient. This matrix was used to construct a dendrogram using the unweighted pair-group method with arithmetic mean (UPGMA) with NTSYS-pc 2.2 [21].

3. Results

3.1. Frequency and distribution of EST–SSRs

A total of 1316 raw EST sequences varying in length from 105 to 1147 bases (average 525) were downloaded, cleaned, and assembled (Table 2). The assembly resulted in 155 contigs and 653 singletons. The length of contigs varied between 409 and 2118 bp with an average of 743 bp. Of the 808 assembled sequences, 110 were found to contain 130 SSRs, with a frequency of 1 SSR/3.4 kb of the available ESTs. Of the 110 SSR-containing ESTs, 14 (13%) contained more than one SSR, and 15 SSRs were found in compound form. Among the types of SSRs, the highest proportion was represented by dinucleotide repeat (DNR) (72.3%), followed by trinucleotide (TNR) (21.5), tetranucleotide (3.8%), and hexanucleotide (2.3%) repeats (Table 3). No pentanucleotide SSR was identified under the criteria used for the SSR search. The majority of the SSR motifs were of smaller repeat length and only 8 were found to contain 10 or more repeats (Table 3). The most common type of SSR motif was AG/CT (43.8%), followed by AT/AT (16.2%), AC/GT (12.3%), AAG/CTT (10.8%), and AAT/ATT (3.8%) (Table 3).

3.2. Polymorphism analysis and cross-species transferability

All of the 110 SSR-containing sequences were used to design primers flanking the SSR motif. We successfully designed primers for 68 (62%) SSR-containing sequences (Table S1). The remaining 42 (38%) SSR containing sequences were not found suitable for designing primers, owing either to marginal SSRs or inappropriate flanking sequences. All the 68 primer pairs were synthesized and characterized for various marker attributes with 13 accessions of *M. piperita* and also used to assess cross-species transferability in four species of *Mentha*. Of the 68, 54 primer pairs produced clear amplicons of the expected sizes. The 54 EST–SSRs amplified with *M. piperita* accessions resulted in 33 polymorphic SSRs with a total of 77

Table 2 – Details of ESTs and SSRs identified in *M. piperatai*.

Parameters	Numbers
Total raw EST-sequences	1316
Contig	155
Singletons	653
Total number of sequences examined	808
Total size of examined sequences (bp)	454795
Total number of identified SSRs	130
Number of SSR containing sequences	110
Number of sequences containing more than 1 SSR	14
Number of SSRs present in compound formation	15
Number of primers designed	68

Table 3 – SSR frequencies by repeat type and motif in *M. piperita*.

SSR type	4	5	6	7	8	9	>10	Total	%
DNR	0	46	17	12	9	3	7	94	72.3
TNR	0	17	6	2	2	0	1	28	21.5
TtNR	4	1	0	0	0	0	0	5	3.8
PNR	0	0	0	0	0	0	0	0	0.0
HNR	2	1	0	0	0	0	0	3	2.3
Repeat motif	**4**	**5**	**6**	**7**	**8**	**9**	**>10**	**Total**	**%**
AC/GT	0	8	3	1	2	1	1	16	12.3
AG/CT	0	29	11	7	4	2	3	57	43.8
AT/AT	0	9	3	4	3	0	2	21	16.2
AAG/CTT	0	8	4	0	2	0	0	14	10.8
AAT/ATT	0	4	1	0	0	0	0	5	3.8
ACG/CGT	0	1	0	0	0	0	0	1	0.8
ACT/AGT	0	0	1	0	0	0	0	1	0.8
AGG/CCT	0	2	0	1	0	0	1	4	3.1
ATC/ATG	0	1	0	0	0	0	0	1	0.8
CCG/CGG	0	1	0	1	0	0	0	2	1.5
AAAG/CTTT	0	1	0	0	0	0	0	1	0.8
AAAT/ATTT	3	0	0	0	0	0	0	3	2.3
AATT/AATT	1	0	0	0	0	0	0	1	0.8
AAGGTC/ACCTTG	1	0	0	0	0	0	0	1	0.8
ACCGCC/CGGTGG	1	1	0	0	0	0	0	2	1.5

alleles and 21 monomorphic SSRs (Table 4, Fig. S1). The number of alleles varied from 2 to 4 with an average of 2.33 alleles/SSR. The PIC values of polymorphic SSRs ranged between 0.13 and 0.51 with an average of 0.325 ± 0.09. The primer pair EMM_003 showed the highest PIC (0.51), followed by EMM_049 (0.48), EMM_055 (0.36), and EMM_41 (0.36). The expected heterozygosity (He) varied from 0 to 0.69 (EMM_026) with an average of 0.21 (Table 4). The potential cross-species transferability of the 54 EST–SSRs was assessed in four species of *Mentha* including five accessions of *M. arvensis*, four of *M. spicata*, and one each of *M. citrata* and *M. longifolia*. Among these, *M. arvensis* showed the highest EST–SSR transferability (87.0%) followed by *M. spicata* (83.0%), *M. longifolia* (55.0%), and *M. citrata* (37.0%).

3.3. *Genetic diversity analysis*

Jaccard's similarity coefficients for 24 mint accessions were calculated based on the genotypic data for 54 EST–SSRs and used to prepare a dendrogram (Fig. 1). The genetic similarity coefficient varied from 0.32 to 0.87 with an average of 0.57 ±0.12. The maximum similarity was found between CIM-Indus and CIMAP-Patra (0.87) followed by Kosi and MAS-10-11-45 (0.84), MPS-16 and MPS-20 (0.83), and MPS-21 and MPS-20 (0.81). The minimum similarity or maximum dissimilarity was found between *M. citrata* and MPS-16 (0.32) and between *M. longifolia* and MPS-20 (0.32). UPGMA clustering classified all the accessions into three major clusters (Fig. 1). Cluster I was the largest, containing all the accessions of *M. piperita* and *M. arvensis*. This cluster could be further subdivided into two subclusters, one containing all the accessions of *M. piperita* and other all the accessions of *M. arvensis*. The main cluster II contained all four accessions of *M. spicata*. The species *M. citrata* and *M. longifolia* grouped together as an outgroup (cluster III) in the dendrogram.

4. Discussion

Among various molecular markers, SSR markers have been widely exploited for several plant genetics and breeding studies and applications, including evaluation of genetic relationship between individuals, tagging useful genes/alleles, linkage/QTL mapping, and phylogenetic analyses. Their features including hypervariability, multiallelic nature, wide genome coverage, relative abundance, and amenability to automation and high-throughput genotyping make SSRs preferential markers [22]. In the present investigation, we exploited the publically available EST database of *M. piperita* to develop EST–SSRs for various genetic studies. A total of 110 (8.4%) of the *M. piperita* EST sequences contained microsatellites, yielding 130 SSRs. This was a relatively high abundance of SSRs as compared to those reported earlier for maize (1.4%), barley (3.4%), sorghum (3.6%), and rice (4.7%) [23]. However, it was lower than those reported for tea (15.5%) [24], castor bean (28.4%) [25], and opium poppy (18.8%) [26]. The relative frequency of EST–SSRs was found to be 1/3.4 kb, comparable to those reported for pepper (1/3.8 kb) [27], and tea (1/3.5 kb) [28] and much higher than those reported for lotus (1/13.0 kb) [29], wheat (1/15.6 kb) [23], tomato (1/11.1 kb) [30], and lily (1/15.9 kb) [31]. However, comparing the frequency and abundance of SSRs reported in different plant species may not give conclusive information, as these values are dependent on SSR search criteria, size of data set, database mining tools, and EST sequence redundancy [15]. DNR and TNR have been reported to be the predominant types of repeat motif in EST–SSRs in several plants, but the most abundant motif varied with species [15]. In *M. piperita* EST–SSRs, DNR was the most dominant motif (72.3%) followed by TNR (21.5%), with both accounting for 93.8% of EST–SSRs. A high proportion of DNR was also reported in EST–SSRs in coffee [32], lotus [29], cassava [33], and *Jatropha* [34]. However, TNR was reported as the most common repeat motif in EST–SSRs in citrus [35], Hawaiian mint [36], peanut [37], and lily [31]. The EST dataset used in the present investigation was small, and accordingly various features of EST–SSRs such as frequency, abundance, and motif type might vary if a larger dataset were used. AG/CT (43.8%) and AAG/CTT (10.8%) were the most common DNR and TNR respectively. Similar findings have also been reported earlier by Cardle et al. [30], Kantety et al. [23], Raji et al. [33], Pan et al. [29], Zhang et al. [38], Akash and Myers [39]. In the present study no GC/GC repeat motif SSRs were identified. The GC/GC motif was also not reported in EST–SSRs of *Medicago truncatula* [40], cassava [33], coffee [32], peanut [37], sesame [37], or alfalfa [41].

The 68 SSR flanking primers developed in the present study showed a higher rate of successful primer design (62.0%) than reported for alfalfa (14.0%) [41], sesame (49%) [38], or faba bean (29.0%) [39], but lower than reported for opium poppy (86.4%) [26]. Of 54 primers, 33 amplified across 13 accessions of *M. piperita*, showing 61.0% of polymorphism. The percent polymorphism in the present investigation was

Table 4 – Details of 54 EST–SSRs of *M. piperita* and their cross-species transferability among four other species of *Mentha*.

SSR_ID	SSR motif	Expected size (bp)	Size range (allele no.)	Ho/He [1]	PIC [2]	Cross-species transferability [3]			
						M. arvensis	*M. citrata*	*M. longifolia*	*M. spicata*
EMM_001	(GAA)5	258	258–282 (2)	0.16/0.00	0.16	258, 282	NA	282	258, 282
EMM_003	(CCACCG)5	247	247–274 (3)	0.56/0.42	0.51	NA	NA	247	247, 265
EMM_006	(GAA)5	102	102–148 (3)	0.36/0.31	0.34	124	NA	NA	124
EMM_007	(TC)9	141	146–160 (3)	0.21/0.08	0.21	152	NA	160	146, 152, 160
EMM_010	(GAG)5	273	273–285 (3)	0.15/0.00	0.15	273, 277	273	273	273, 285
EMM_011	(TC)5	188	188–198 (2)	0.13/0.00	0.13	188, 198	NA	198	198
EMM_012	(AGA)5	345	342 (1)	0.00/0.00	0.00	342	342	342	342
EMM_014	(CA)5(CT)6	228	220–230 (2)	0.27/0.33	0.24	220, 230	230	230	220, 230
EMM_015	(GAA)6	261	264–290 (3)	0.25/0.15	0.23	264, 270	NA	NA	290, 270
EMM_016	(AC)6	325	311–315 (2)	0.25/0.15	0.23	NA	315	NA	311, 315
EMM_017	(TG)6	207	201–223 (3)	0.25/0.15	0.23	201, 207	NA	NA	207, 223
EMM_018	(GAA)5	298	298 (1)	0.00/0.00	0.00	298	NA	298	NA
EMM_020	(GAA)5	221	201 (1)	0.00/0.00	0.00	201	NA	NA	201
EMM_021	(AT)7	208	198–220 (4)	0.26/0.23	0.25	198, 210	NA	216	220
EMM_022	(AT)5	341	341 (1)	0.00/0.00	0.00	341	NA	341	341
EMM_023	(TC)5	177	157 (1)	0.00/0.00	0.00	NA	NA	NA	NA
EMM_024	(TA)5	190	200 (1)	0.00/0.00	0.00	200	200	NA	200
EMM_025	(CCG)5	250	265–280 (3)	0.30/0.23	0.26	265, 271	265	265, 280	NA
EMM_026	(CTT)6	173	187–191 (2)	0.44/0.69	0.35	187, 191	NA	187	187
EMM_027	(TC)6	229	224–230 (2)	0.14/0.15	0.13	NA	NA	NA	NA
EMM_028	(TAA)5	329	329 (1)	0.00/0.00	0.00	329	329	329	329
EMM_031	(CT)5	291	290 (1)	0.00/0.00	0.00	290	290	290	290
EMM_032	(AAG)5	201	198–244 (4)	0.37/0.23	0.32	198, 210	225	244	210, 244
EMM_033	(TC)5	117	117 (1)	0.00/0.00	0.00	117	NA	NA	117
EMM_034	(CT)5	198	190 (1)	0.00/0.00	0.00	190	190	190	190
EMM_035	(CCT)5	152	172 (1)	0.00/0.00	0.00	172	NA	NA	172
EMM_036	(AT)5	228	220–224 (2)	0.30/0.23	0.26	NA	NA	NA	224
EMM_037	(TA)6	298	260–280 (3)	0.25/0.15	0.23	260, 266	NA	NA	260, 280
EMM_039	(TAA)5	342	342–361 (3)	0.19/0.08	0.18	342, 348	NA	NA	361
EMM_040	(GAT)5	229	229 (1)	0.00/0.00	0.00	229	229	229	229
EMM_041	(GAC)5	186	178–186 (3)	0.40/0.36	0.36	186	NA	NA	NA
EMM_042	(CA)5	272	272 (1)	0.00/0.00	0.00	272	272	NA	272
EMM_043	(TA)8	206	206–220 (2)	0.14/0.15	0.13	220	NA	220	206
EMM_044	(CT)5	350	340–355 (3)	0.20/0.23	0.18	340, 349	355	349	340, 355
EMM_045	(GAA)5	300	300 (1)	0.00/0.00	0.00	300	300	300	300
EMM_046	(AT)5	306	306 (1)	0.00/0.00	0.00	306	NA	NA	306
EMM_047	(CT)5	221	229–250 (3)	0.30/0.38	0.26	229, 241	NA	NA	NA
EMM_048	(CT)6	211	206 (1)	0.00/0.00	0.00	NA	NA	NA	206
EMM_049	(CT)6	150	139–158 (4)	0.51/0.38	0.48	139, 156	144, 152	152, 156	144, 152
EMM_050	(CT)5	206	198–218 (3)	0.26/0.31	0.25	210, 218	NA	218	198, 210
EMM_051	(TTAA)4	184	180–184 (2)	0.20/0.23	0.18	NA	NA	NA	184
EMM_053	(CT)5	121	117 (1)	0.00/0.00	0.00	117	117	117	117
EMM_054	(GAG)11	157	143–154 (2)	0.29/0.08	0.26	143	NA	143	NA
EMM_055	(GAG)7	182	153–159 (2)	0.44/0.00	0.36	153, 159	159	159	153, 159
EMM_057	(AG)5	154	148–170 (4)	0.35/0.23	0.32	148, 152	148	148	NA
EMM_058	(TA)5	206	210–228 (3)	0.30/0.23	0.26	210, 218	NA	NA	224
EMM_059	(CT)5	258	270 (1)	0.00/0.00	0.00	270	NA	NA	270
EMM_062	(CTA)6	124	122–171 (4)	0.21/0.23	0.21	122, 124	NA	NA	122, 124
EMM_063	(GT)5	151	151 (1)	0.00/0.00	0.00	151	151	151	151
EMM_064	(TA)10	215	215–230 (2)	0.21/0.25	0.19	215, 217	NA	230	217
EMM_065	(TA)8	285	271 (1)	0.00/0.00	0.00	271	NA	NA	271
EMM_066	(GAA)6	132	131–135 (2)	0.19/0.08	0.18	135	NA	135	135
EMM_067	(GT)6	183	185 (1)	0.00/0.00	0.00	185	185	NA	185
EMM_068	(CT)5	250	250–265 (3)	0.25/0.31	0.23	250, 256	NA	265	NA

[1] Observed/expected heterozygosity.
[2] Polymorphism information content.
[3] Five genotypes of *M. arvensis*, four of *M. spicata*, and one each of *M. citrata* and *M. longifolia* were tested for EST–SSR transferability.

found to be higher than that observed across 24 accessions of sesame (11.6%) [38] but lower than that reported among 28 alfalfa accessions (97.0%) [41] and among 37 accessions of opium poppy [26]. The polymorphism found in the EST–SSRs of *M. piperita* suggested the potential of these markers for use in future genetic studies. However, most of the SSRs showed

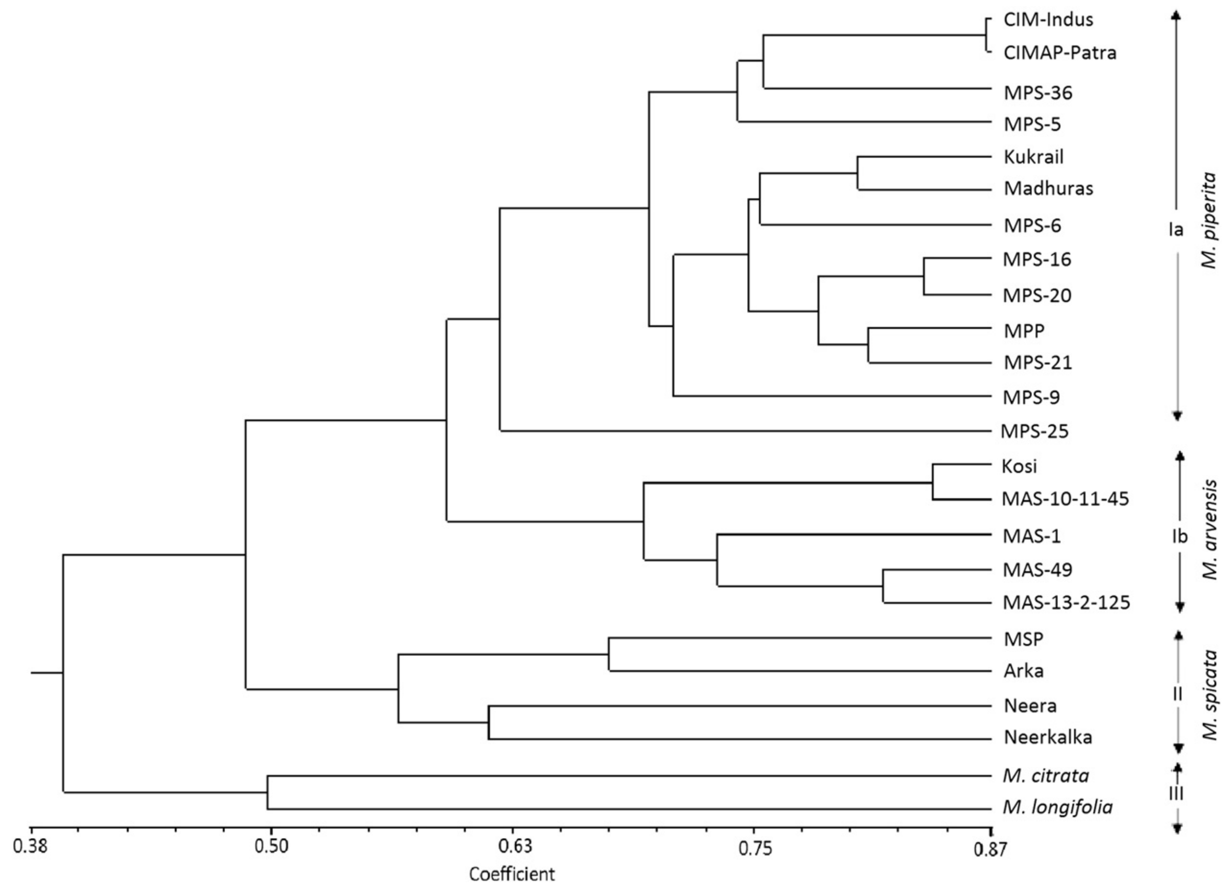

Fig. 1 – UPGMA dendrogram of *Mentha* species/accessions. Genetic distance was based on Jaccard's similarity coefficient calculated from EST–SSR data.

low PIC values, with an average of 0.25 ± 0.09. The low average PIC value may reflect the small number of accessions surveyed or low allelic polymorphism. The utility of the SSRs could be expanded by studying their transferability to closely associated genera/species. Transferability of EST–SSRs has been reported in many crop plants, including sugarcane [42], pear [43], *Prunus* [44], lettuce [45], alfalfa [41], faba bean [39], and opium poppy [26]. *M. arvensis* and *M spicata* showed the highest transferability, indicating the closeness of their relationship to *M. piperita*. This close relationship has also been previously established and reported by Khanuja et al. [8], Gobert et al. [11], and Shiran et al. [10] based on RAPD and AFLP analysis. Reports of the interspecific hybrids *M. arvensis* × *M. spicata* [2] and *M. arvensis* × *M. piperita* [46] also support these findings. The EST–SSRs developed in the present study were also used to evaluate genetic relatedness among accessions and species of *Mentha*. The EST–SSRs were able to distinguish different accessions within and among *Mentha* species. All 13 accessions of *M. piperita* were grouped in one cluster and accessions of *M. arvensis* clustered together. The adjacent clustering of *M. arvensis* with *M. piperita* further supports their close relatedness. Likewise, 4 accessions of *M. spicata* clustered together and *M. citrata* and *M. longifolia* did not group with any cluster. This result indicates that the EST–SSRs reported here have potential for various genetic studies in *Mentha*. The dendrogram clustering of *M. citrata* and *M. longifolia* close to *M. spicata* and *M. arvensis* indicates some common ancestry. Gobert et al. [11] showed high similarity between *M. spicata* and *M. longifolia* based on AFLP analysis.

In the present study, we have developed EST–SSRs using the *M. piperita* EST database. The EST–SSRs showed a high level of polymorphism and were transferable across several species of *Mentha*. Genetic relatedness among different accessions and species of *Mentha* was established using the EST–SSRs. The present study demonstrates the potential of EST–SSRs for genetic and phylogenetic studies in *Mentha*.

Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.cj.2015.0x.00x.

REFERENCES

[1] N.K. Patra, H. Tanveer, N.K. Tyagi, S. Kumar, Cyto-taxonomical status and genetical and breeding perspectives of *Mentha*, J. Med. Arom. Plant Sci. 22 (2000) 419–430.

[2] N.K. Patra, H. Tanveer, S.P.S. Khanuja, A.K. Shasany, H.P. Singh, V.R. Singh, S. Kumar, A unique interspecific hybrid spearmint clone with growth properties of *Mentha arvensis* L. and oil qualities of *Mentha spicata* L, Theor. Appl. Genet. 102 (2001) 471–476.

[3] S. Behnam, M. Farzaneh, M. Ahmadzadeh, A.S. Tehrani, Composition and antifungal activity of essential oils of *Mentha piperita* and *Lavendula angustifolia* on post-harvest phytopathogens, Commun. Agric. Appl. Biol. Sci. 71 (2006) 1321–1326.

[4] S. Kizil, N. Hasimi, V. Tolan, E. Kilinc, U. Yuksel, Mineral content, essential oil components and biological activity of two mentha species (*M. piperita* L., *M. spicata* L.), Turk. J. Field Crops. 15 (2010) 148–153.

[5] W.A. Court, R.C. Roy, R. Pocs, Effect of harvest date on the yield and quality of the essential oil of peppermint, Can. J. Plant Sci. 73 (1993) 815–824.

[6] S.S. Mahmoud, R.B. Croteau, Menthofuran regulates essential oil biosynthesis in peppermint by controlling a downstream monoterpene reductase, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 14481–14486.

[7] R.B. Croteau, E.M. Davis, K.L. Ringer, M.R. Wildung, Menthol biosynthesis and molecular genetics, Naturwissenschaften 92 (2005) 562–577.

[8] S.P.S. Khanuja, A.K. Shasany, A. Srivastava, S. Kumar, Assessment of genetic relationships in *Mentha* species, Euphytica 111 (2000) 121–125.

[9] A.L. Fenwick, S.M. Ward, Use of random amplified polymorphic DNA markers for cultivar identification in mint, Hort. Sci. 36 (2001) 761–764.

[10] B. Shiran, M. Soheila, R. Khorshid, Assessment of genetic diversity among Iranian mints using RAPD markers, in: J. Vollmann, H. Grausgruber, P. Ruckenbauer (Eds.),Genetic Variation for Plant Breeding. Proceedings of the 17th EUCARPIA General Congress, Tulln, Austria, 8–11 September, 2004, 2004, pp. 121–125.

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

[12] D. Field, C. Wills, Abundant microsatellite polymorphism in *Saccharomyces cerevisiae*, and the different distributions of microsatellites in eight prokaryotes and *S. cerevisiae*, result from strong mutation pressures and a variety of selective forces, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 1647–1652.

[13] W. Powell, G. Machray, J. Provan, Polymorphism revealed by simple sequence repeats, Trends Plant Sci. 1 (1996) 215–222.

[14] P.K. Gupta, S. Rustgi, S. Sharma, R. Singh, N. Kumar, H.S. Balyan, Transferable EST–SSR markers for the study of polymorphism and genetic diversity in bread wheat, Mol. Gen. Genomics. 270 (2003) 315–323.

[15] R.K. Varshney, A. Graner, M.E. Sorrells, Genic microsatellite markers: features and applications, Trends Biotechnol. 23 (2005) 48–55.

[16] T. Thiel, W. Michalek, R.K. Varshney, A. Graner, Exploiting EST database for the development and characterization of gen-derived SSR-markers in barley (*Hordeum vulgare* L.), Theor. Appl. Genet. 10 (2003) 6411–6422.

[17] X. Huang, A. Madan, CAP3: a DNA sequence assembly program, Genome Res. 9 (1999) 868–877.

[18] M. Schuelke, An economic method for the fluorescent labeling of PCR fragments, Nat. Biotechnol. 18 (2000) 233–234.

[19] K. Liu, S.V. Muse, Power marker: integrated analysis environment for genetic marker data, Bioinformatics 21 (2005) 2128–2129.

[20] D. Botstein, R.L. White, M. Skolnick, R.W. Davis, Construction of genetic linkage map in man using restriction fragment length polymorphisms, Am. J. Hum. Genet. 32 (1980) 314–331.

[21] F.J. Rohlf, NTSYS-pc Version 2.2: Numerical Taxonomy and Multivariate Analysis System, Exeter Software, New York, 2000.

[22] R.K. Kalia, M.K. Rai, S. Kalia, R. Singh, A.K. Dhawan, Microsatellite markers: an overview of the recent progress in plants, Euphytica 177 (2011) 309–334.

[23] R.V. Kantety, M.L. Rota, D.E. Mathews, M.E. Sorrells, Data mining for simple sequence repeats in expressed sequence tags from barely, maize, rice, sorghum and wheat, Plant Mol. Biol. 48 (2002) 501–510.

[24] J.Q. Ma, Y.H. Zhou, C.L. Ma, M.Z. Yao, J.Q. Jin, Identification and characterization of 74 novel polymorphic EST–SSR markers in the tea plant, *Camellia sinensis* (Theaceae), Am. J. Bot. 97 (2010) 153–156.

[25] L. Qui, C. Yang, B. Tian, J.B. Yang, A. Liu, Exploiting EST databases for the development and characterization of EST–SSR markers in castor bean (*Ricinus communis* L.), BMC Plant Biol. 10 (2010) 278.

[26] H. Selale, I. Celik, V. Gultekin, J. Allmer, S. Doganlar, A. Frary, Development of EST–SSR markers for diversity and breeding studies in opium poppy, Plant Breed. 132 (2013) 344–351.

[27] G. Yi, J.M. Lee, S. Lee, D. Choi, B.D. Kim, Exploitation of pepper EST–SSRs and an SSR-based linkage map, Theor. Appl. Genet. 114 (2006) 113–130.

[28] R.K. Sharma, P. Bhardwaj, R. Negi, T. Mohapatra, P.S. Ahuja, Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.), BMC Plant Biol. 9 (2009) 53.

[29] L. Pan, Q. Xia, Z. Quan, H. Liu, W. Ke, Y. Ding, Development of novel EST–SSRs from sacred lotus (*Nelumbo nucifera* Gaertn) and their utilization for the genetic diversity analysis of *N. nucifera*, J. Hered. 101 (2010) 71–82.

[30] L. Cardle, L. Ramsay, D. Milbourne, M. Macaulay, D. Marshall, R. Waugh, Computational and experimental characterization of physically clustered simple sequence repeats in plants, Genetics 156 (2000) 847–854.

[31] S. Yuan, G. Liang, C. Liu, J. Ming, The development of EST–SSR markers in *Lilium regale* and their cross-amplification in related species, Euphytica 189 (2013) 393–419.

[32] R.K. Aggarwal, P.S. Hendre, R.K. Varshney, P.K. Bhat, V. Krishnakumar, L. Singh, Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species, Theor. Appl. Genet. 114 (2007) 359–372.

[33] A.A. Raji, J.V. Anderson, O.A. Kolade, C.D. Ugwu, A.G. Dixon, I.L. Ingelbrecht, Gene-based microsatellites for cassava (*Manihot esculenta* Crantz): prevalence, polymorphisms, and cross-taxa utility, BMC Plant Biol. 9 (2009) 118.

[34] H.K. Yadav, A. Ranjan, M.H. Asif, S. Mantri, S.V. Sawant, R. Tuli, EST-derived SSR markers in *Jatropha curcas* development, characterization, polymorphism, and transferability across the species/genera, Tree Genet. Genomes 7 (2011) 207–219.

[35] C. Chen, P. Zhou, Y. Choi, S. Huang, F. Gmitter, Mining and characterizing microsatellites from citrus ESTs, Theor. Appl. Genet. 112 (2006) 1248–1257.

[36] C. Lindqvist, A.C. Scheen, M.J. Yoo, P. Grey, D.G. Oppenheimer, J.H.L. Mack, D.E. Soltis, P.S. Soltis, V.A. Albert, An expressed sequence tag (EST) library from developing fruits of an Hawaiian endemic mint (*Stenogyne rugosa*, Lamiaceae): characterization and microsatellite markers, BMC Plant Biol. 6 (2006) 16.

[37] X. Liang, X. Chen, Y. Hong, H. Liu, G. Zhou, S. Li, B. Gu, Utility of EST-derived SSR in cultivated peanut (*Arachis hypogaea* L.) and *Arachis* wild species, BMC Plant Biol. 9 (2009) 35.

[38] H. Zhang, W. Libin, M. Hongmei, Z. Tide, W. Cuiying, Development and validation of genic-SSR markers in sesame by RNA-seq, BMC Genomics 13 (2012) 316.

[39] M.W. Akash, G.O. Myers, The development of faba bean expressed sequence tag–simple sequence repeats (EST–SSRs) and their validity in diversity analysis, Plant Breed. 131 (2012) 522–530.

[40] I.M. Eujayl, K. Sledge, L. Wang, G.D. May, K. Chekhovskiy, *Medicago truncatula* EST–SSRs reveal cross-species genetic markers for *Medicago* spp, Theor. Appl. Genet. 108 (2004) 414–422.
[41] Z. Wang, Y. Hongwei, F. Xinnian, L. Xuehui, G. Hongwen, Development of simple sequence repeat markers and diversity analysis in alfalfa (*Medicago sativa* L.), Mol. Biol. Rep. 40 (2013) 3291–3298.
[42] G.M. Cordeiro, R. Casu, C.L. McIntyre, J.M. Manners, R.J. Henry, Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to *Erianthus* and *Sorghum*, Plant Sci. 160 (2001) 1115–1123.
[43] L. Fan, M.Y. Zhang, Q.Z. Liu, T.L. Li, Y. Song, L.F. Wang, S.L. Zhang, J. Wu, Transferability of newly developed pear ssr markers to other rosaceae species, Plant Mol. Biol. Report. 31 (2013) 1271–1282.
[44] M. Mnejja, M.J. Garcia, J.M. Audergon, P. Arús, *Prunus* microsatellite marker transferability across rosaceous crops, Tree Genet. Genomes 6 (2010) 689–700.
[45] I. Simko, Development of EST–SSR markers for the study of population structure in lettuce (*Lactuca sativa* L.), J. Hered. 100 (2009) 256–262.
[46] B. Kumar, A.K. Shukla, A. Samad, Development and characterization of menthofuran-rich inter-specific hybrid peppermint variety CIMAP-Patra, Mol. Breed. 34 (2014) 717–724.