TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2015) 39: 69-77
© TÜBİTAK
doi:10.3906/biy-1404-29

E-microsatellite markers for some naturally occurring *Salvia* species in the Mediterranean region

Ayşe Gül İNCE, Mehmet KARACA*
Department of Field Crops, Faculty of Agriculture, Akdeniz University, Antalya, Turkey

Received: 08.04.2014 • **Accepted:** 01.07.2014 • **Published Online:** 02.01.2015 • **Printed:** 30.01.2015

Abstract: The genus *Salvia* L. contains about 1000 species, more than 40 of which are endemic to the Mediterranean region. Species boundaries and classifications of this genus have been mainly conducted based on morphological and anatomical characteristics. Although genomic and genic microsatellites are commonly used in a variety of genetic studies in many plant and animal species, their use has been limited in *Salvia* species. This research was undertaken to study microsatellite contents of expressed sequence tags (ESTs) and develop a new set of expressed-microsatellite (e-microsatellite) markers using the GenBank databases. In in silico studies of the research, microsatellites were identified and e-microsatellite primer pairs flanking the repeats were designed. In in vitro studies of the present work, e-microsatellite primer pairs were used to amplify genomic DNA samples of different *Salvia* species naturally occurring in the Mediterranean region of Turkey. The results indicated that ESTs of related species deposited in GenBank were reliable sources for the development of e-microsatellite markers. The results of this study also indicated that cross-transferability of e-microsatellite markers decreased as the evolutionary distance among the tested *Salvia* species increased. E-microsatellite loci presented in this work could be very useful for measuring genetic diversity, gene flow, and species identification studies in the genus *Salvia*.

Key words: Cross-transferability, DNA markers, expressed sequence tags, Labiatae, sage, touch-down-PCR

1. Introduction

The genus *Salvia* L., commonly known as sage, is one of the largest genera in the family Lamiaceae (Labiatae). It is known that this genus contains about 1000 plant species, more than 40 of which are endemic to the Mediterranean area. *Salvia* consists of annual, biennial, perennial, and woody shrubs species, many of which are ethnomedicinally significant in regions where they naturally grow. *Salvia* has been placed in the subfamily Nepetoideae and is commonly considered as monophyletic. However, the high level of cross-hybridization in this genus makes the study of its chemical features, morphology, and anatomy difficult. Phylogenetic and genetic studies based on morphology, anatomy, and chemical features have many potential disadvantages. For instance, there are many unanswered questions concerning the biogeography, character evolution, and relationships within the genus *Salvia* (Walker and Sytsma, 2007; Karaca et al., 2008; Radosavljevic et al., 2011; Xu et al., 2013).

Microsatellite markers, also known as simple sequence repeats (SSRs) or short tandem repeats (STRs), have become one of the most important molecular genetic markers in modern research. Microsatellites are employed in a number of fields, including population genetics, molecular ecology, systematic biodiversity, and conservation genetics. Although microsatellite markers have been successfully used in many plant and animal species for different purposes (Karaca et al., 2004; Wang et al., 2008; Teacher et al., 2012), their use has been limited in *Salvia* species (Deng et al., 2009; Radosavljevic et al., 2011; Ince and Karaca, 2012; Xu et al., 2013). For instance, a search for '*Salvia*' in titles using Web of Science (Thomson Reuters) returned 2616 articles, but when this search was refined with microsatellites, SSRs, or STRs, the number of articles decreased to 3 as of 6 October 2013. This clearly shows the need for more microsatellite research studies in *Salvia*.

Microsatellite markers have many advantages, including their high abundance, high information content, codominant inheritance, and reproducibility. Genomic (conventional) microsatellites are usually developed from a variety of DNA libraries such as genomic, bacterial artificial chromosome, and public genomic databases (Karaca et al., 2004; Wang et al., 2008; Teacher et al., 2012). Microsatellite primer pairs are usually obtained using a microsatellite enrichment procedure that includes

* Correspondence: mkaraca@akdeniz.edu.tr

genomic DNA extraction, digestion studies of the genomic DNA, and multiple purification steps to remove small and degraded DNA fragments. Target products are hybridized using microsatellite domains (repeat-containing region of sequences), and probe-target fragments are isolated and ligated into vectors and sequenced. Finally, microsatellite primer pairs flanking the microsatellite domains (repeat regions) are designed using appropriate software programs. Development of genomic microsatellite markers is usually expensive and it is time-consuming to isolate and characterize them. Genomic microsatellites are usually species-specific markers that must be developed de novo for each species under study. Furthermore, genomic microsatellites commonly occur in the noncoding regions of the genome (Ince et al., 2008; Deng et al., 2009; Karaca et al., 2013; Xu et al., 2013).

Recently developed next-generation sequencing platforms provide novel avenues for isolation of genomic and genic (transcribed region of a genome) microsatellites. For instance, a single 454 run is capable of generating about 400 Mb of sequence data to capture individual microsatellites along with enough flanking sequence to design polymerase chain reaction (PCR) primer pairs (Karan et al., 2012). However, next-generation sequencing platforms are expensive and not easily available in many laboratories. Furthermore, genomic microsatellite markers that are developed from genomic DNA sequences using conventional or next-generation sequencing platforms have low levels of cross-transferability among closely related species (Ince et al., 2010a; Polat et al., 2010). It has been commonly recognized that sequences of expressed sequence tags (ESTs) can be successfully used in the development of genic microsatellite markers (Karaca et al., 2013; Xu et al., 2013).

In comparison to genomic microsatellite markers, genic microsatellite markers have several disadvantages. Genomic microsatellites usually do not have genetic functions and are not closely linked to transcriptional regions, while genic (expressed) ones are potentially tightly linked with functional genes that perhaps control certain important genetic characters (Gao et al., 2004; Ince et al., 2008; Teacher et al., 2012). In addition, genic microsatellites contain high transferability between unrelated species and even different genera in a family due to the fact that expressed genes are highly conserved, thus increasing the likelihood of cross transferability (Ince et al., 2008; Deng et al., 2009; Teacher et al., 2012; Karaca et al., 2013).

The National Center for Biotechnology Information (NCBI) EST databases contain an ever-increasing number of in silico sequences. Due to the current lapse in government funding, however, the information in the NCBI has not been up to date. However, some other resources are easily available for the efficient development of large numbers of e-microsatellite markers. E-microsatellite markers developed from ESTs may be directly related with a coding gene; therefore, they could be used as target sequences for developing functional microsatellite markers. These functional markers are good resources for breeding studies since they may be directly associated with the gene or genes affecting a particular important trait (Ince et al., 2008; Deng et al., 2009; Teacher et al., 2012). In this study, expressed sequence tag-based microsatellite primer pairs (e-microsatellites) were developed using a contigs-based approach and characterized in 6 different *Salvia* species naturally occurring in the Mediterranean region of Turkey. To the best of our knowledge, this is the first report on cross-species transferable e-microsatellite markers.

2. Materials and methods

2.1. Plant materials and DNA extraction

Genomic DNA samples of 6 different species of *Salvia* (*S. sclarea* L., *S. fruticosa* Mill., *S. tomentosa* Mill., *S. dichroantha* Stapf., *S. virgata* Jacq., and *S. pisidica* Boiss. & Heldr) were extracted from frozen fresh leaves of single individual and bulked samples consisting of several individual plant leaves using an established protocol (Karaca et al., 2005b; Ince and Karaca, 2009). Polysaccharides of some samples that coisolated with genomic DNA samples were further purified using a method described by Ince et al. (2011). Further information about some of the plant materials used in this study could be found in a previous study (Ince and Karaca, 2012). Briefly, 2 g of leaf samples from an individual plant of each species and 10 g of leaf samples from several individuals from a species were ground in a small, wide-sized mortar using liquid nitrogen. DNA samples extracted from those bulked samples were used to assess the amplification patterns of primer pairs developed in this study. These DNA samples were not utilized in cloning, sequencing, and data analyses studies.

Powdered plant samples were homogenized in the extraction buffer; following centrifugation, cells were lysed using lysis buffer. Genomic DNA samples were purified from proteins, polysaccharides, and RNA contaminants according to protocol described by Karaca et al. (2005b). DNA amount, purity, and integrity of the samples were determined using a spectrophotometer and agarose gel electrophoresis technique (Karaca and İnce, 2011). Prior to PCR amplification studies, concentrations of DNA samples were readjusted so that each PCR sample contained 0.065 µg in 8.5 µL volume for PCR amplification. All chemicals used in this study were of molecular biology grade and purchased from Amresco Inc. (Solon, OH, USA) or Invitrogen Corp. (Carlsbad, CA, USA).

2.2. Data mining

ESTs were downloaded from GenBank (http://www.ncbi.nlm.nih.gov/) of the NCBI databases and these sequences were used as targets for microsatellite mining studies. Sequences of ESTs were analyzed using the Tandem Repeats Analyzer 1.5 software (Bilgen et al., 2004). In order to classify and detect microsatellites in ESTs, minimum motif length criteria were defined as being 10 repeats (20 bp) for dinucleotide, 7 repeats (21 bp) for trinucleotide, and 4 repeats for tetra- (16 bp), penta- (20 bp), and hexanucleotide (24 bp) motif lengths (Karaca et al., 2005a). After identification of ESTs containing microsatellites, they were assembled into contiguous sequences (contigs) using the Sequencher software (Gene Codes, Ann Arbor, MI, USA). Contigs assembly parameters were set to a minimum overlap of 50 bases, 90% identity match, and the large gap option was implemented (Ince, 2012). In order to utilize all the EST sequences of *Salvia* species, the contigs-based approach (Ince, 2012) was utilized. This approach does not remove redundancy in EST sequences and therefore does not omit valuable information in sequences of different alleles at the same locus. After identification of nonredundant ESTs, they were further analyzed to design microsatellite primer pairs using the Primer 3 program (Rozen and Skaletsky, 2000). Briefly, selected ESTs were used to design primer pairs based on the following main parameters: GC content was set between 40% and 80%, annealing temperature (Tm) was set between 58 and 62 °C, and expected amplified products size was defined as 150–500 bp.

2.3. Touch-down PCR

Amplification reactions of the targeted region of the genomic DNA samples were carried out using a touch-down style polymerase chain reaction (Td-PCR) to avoid amplifying the nonspecific sequences. The first step of the touch-down polymerase chain reaction cycle was done at 60 °C, which is considered a high annealing temperature. The annealing temperature was decreased in increments for every subsequent set of 10 cycles. At the higher annealing temperature primer pairs annealed with the least-permissive of nonspecific binding.

The following Td-PCRs were carried out in 25 µL reaction volume containing 60 ± 5 ng of genomic DNA samples as templates, 0.5 µM of each e-microsatellite primer pair (Table 1), 80 mM Tris-HCl (pH 8.8), 19 mM $(NH_4)_2SO_4$, 0.009% Tween-20 (w/v), 0.28 mM of each dNTP, 3 mM $MgCl_2$, and 1 U of Taq DNA polymerase (Invitrogen). The Td-PCR amplification profile was as follows: initial denaturation at 94 °C for 180 s, 10 cycles with denaturation at 94 °C for 20 s, annealing at 60 °C for 30 s in the first cycle diminishing by 0.5 °C in each subsequent cycle, and extension reactions at 72 °C for 1 min using a Veriti 96-well thermal cycler (Applied Biosystems, Foster City, CA, USA). Thirty additional PCR cycles were run using the same cycling parameters mentioned above with constant annealing at 55 °C. Denaturation and extension conditions were the same as indicated above. The amplification reactions were completed with a final extension step at 72 °C for 10 min (Ince et al., 2010c).

2.4. Electrophoresis

Five microliters of DNA loading buffer consisting of 0.25 g L^{-1} bromophenol blue, 0.25 L^{-1} xylene cyanol FF, and 400 g L^{-1} sucrose prepared in water was added to each 25 µL of amplified reaction and 8–12 µL of these mixtures was loaded in 3% high-resolution agarose gels (Serva Electrophoresis GmbH, Heidelberg, Germany). In addition to the 0.5 µg/mL ethidium bromide in agarose gels added during the gel preparation, 0.5 µg/mL ethidium bromide solution was also added to 1 L of running buffer. Amplicons (PCR products) were electrophoresed at 5 V/cm at constant voltage for 8–12 h in the presence of 1X Tris-borate ethylenediamine tetraacetic acid (EDTA) buffer (89 mM Tris-borate, 2 mM EDTA, pH 8.3) and photographed on an ultraviolet (UV) transilluminator for analysis (Ince et al., 2010c).

2.5. Cloning and sequencing

For confirmation of microsatellites in PCR amplified products, several single amplicons (alleles) were recovered from agarose gels (Karaca et al., 2013). Purified amplicons were transferred to vectors using the InsTAclone cloning kit/TransformAid kit and protocol (Fermentas Life Sciences, Hanover, MD, USA). The sequence map and the MCS region of the vector (pTZ57R/T) are available for download at www.thermoscientificbio.com/fermentas/.

Plasmids were extracted from bacterial strains (JM107) using a plasmid isolation kit (Fermentas). Target regions in plasmids were commercially sequenced in duplicates using M13F and M13R primer pairs (Macrogen Inc., Amsterdam, the Netherlands). DNA sequences were manually analyzed for detecting the microsatellite repeats. Sequencing analyses confirmed that all 5 tested alleles contained microsatellite repeats as expected from the EST sequences. This also indicated that the primer pairs developed in this study successfully and specifically amplified targeted regions in *Salvia* genomic DNA samples.

2.6. Data analysis

Microsatellite markers were photographed on a UV transilluminator and scored as presence (1) and absence (0). Scores were then used in determination of cross-species transferability ratio and polymorphism information content (PIC) values (also called heterozygosity, H). Polymorphism information content (PIC or H) of each microsatellite locus was calculated according to the formula:

Table 1. Characteristics of 22 e-microsatellite primer pairs and markers in *Salvia* species.[a]

ID	Sequence	Motif	Size (bp)
S01	F: ACCGTGATTCGGCGTTTA R: GCCTCTTCCATTCCTTTTCC	GA_{17}	190–210
S02	F: TCATCTTTCCTCAGACAAACA R: AGACCAGCAAGCATCTCAACT	GA_{10}	195–220
S03	F: GCTACCACCCGATTTACGAA R: GCCAGAAAACTGCCATTCC	GAT_{7}	180–200
S04	F: CAAATCCAGCCCCAAATCA R: TTCCTCTTTCAGGTTCCATCAG	CT_{13}	550–650
S05	F: GGGGGAGAAAGCACAGAAC R: GTCTTGGGAGGCTGGTTG	CA_{11}	680–750
S06	F: CGCAATCCTCCCTCATAAAT R: GACCTTCTTCACGCTGGTG	TC_{19}	250–260
S07	F: TTCTGCTGGAAGGGATGG R: AAGGGGAGAGACCACAAGTC	GA_{12}	190–205
S08	F: GCCGAGTTTCGTCTTCAACC R: AGAAAAGGGACTGGGGCTG	TCA_{7}	190–200
S09	F: TGACCCTGTGGTGTTTGATG R: TCTTGCTCTCGTGCTCTTGA	GAG_{8}	180–200
S10	F: TTATCCGTGGGCTTGGCTA R: TCTTGACAGGCTTCTACTCA	TC_{18}	500–560
S11	F: GCCTTTCCGTTTTGGTTACT R: GCTCGTTCTCACCTTCCAC	GA_{12}	600
S12	F: GCATCAAGCCAGGTATTTTGG R: GACGAAGGACAACAGCATAA	AG_{11}	330–355
S13	F: CCTCGACTACTATTGAGCGTA R: TGTCTTATTTCTTCGTGGGTCA	CA_{10}	200–210
S14	F: GGAGCTGGAGAAAGAGAACA R: TGCAAGAAAAGCAAGCTACA	TA_{14}	180–190
S15	F: TCGATTTCCTTCGATTGCTC R: GCCTCCAAACACTACCTTGC	$ATTT_{5}$	180–190
S16	F: AGGCTGTTGTAGGTTCTCATCT R: GCGTGTAGTTTAGTAGGAGGA	$GTGAT_{5}$	300–310
S17	F: CTGTGAAGCAAGCAGGTTCT R: ACGCCATAACTGAGACACG	AG_{13}	390–420
S18	F: TGATAATGAGATTGGCAGCAG R: GAGAACTTTTGACACCCGTTT	TAT_{8}	180–195
S19	F: CGCTCTACCACTTCTCGGATTC R: CTACACCAACGCAGCCACTT	TAT_{8}	500
S20	F: AGACTGAGGGCTTTCGTT R: GGAGATGATTCTTGTTGAAG	TCA_{8}	200
S21	F: GAATCCGAGTAGAGGGTGTTG R: TGAAGCACATACCAAAATGG	$TACA_{5}$	200–230
S22	F: CGAGATGGTGAACGAGCA R: ATTTGCCCCAGTGCCTCT	CAG_{7}	180–200

[a]: Shown for each primer pair are the forward (F) and reverse (R) sequence, repeat type (motif), repeat numbers, and size of the markers (bp).

$PIC = 1 - \sum fi^2$,

where fi is the frequency of the ith allele in the species examined. Cross-species transferability of e-microsatellite loci was determined using the following basic formula:

TC: (SMM/TST) × 100,

where TC is the cross-species transferability (%), SMM is the number of species with at least one allele (marker) transferred, and TST is the total number of species tested.

3. Results

3.1. In silico analysis

The GenBank (http://www.ncbi.nlm.nih.gov/) databases contain ESTs and other sequences for many organisms. The dbEST release of 130101 reported that there were 1459 ESTs for *Salvia fruticosa*, 19 ESTs for *S. sclarea*, and 10288 ESTs for *S. miltiorrhiza* on 1 January 2013. In this study a total of 11766 ESTs were utilized and mined for microsatellites. A total of 199 ESTs with microsatellite domains were identified, varying from di- to hexanucleotide repeats from *S. fruticosa* and *S. miltiorrhiza*. Analysis revealed that none of the ESTs of *S. sclarea* contained microsatellite domains. Analysis showed that the highest number of microsatellite repeat type was di- (90), followed by tri- (51), tetra- (17), penta- (17), and hexa-nucleotides (24).

A total of 71 ESTs containing microsatellites were assembled into 23 different contigs, while a total of 125 ESTs could not be assembled into contigs. They were considered singletons since they were nonhomologous as per the assembly criteria. The number of ESTs in contigs varied from 2 to 10. Analysis of 23 contigs and 125 singletons indicated that 120 primer pairs could be designed since sequences contained enough primer flanking regions outside the microsatellite domains and some sequences contained polymorphism within the same contigs. In the present study e-microsatellite primer pairs for hexanucleotide repeats could not be designed since ESTs with hexanucleotides did not follow the criteria mentioned above.

3.2. In vitro analysis

A total of 30 candidate primer pairs were selected from contigs and singletons and screened across 6 *Salvia* species that naturally grow along the Mediterranean coast of Turkey. However, 8 primer pairs from the 30 candidate microsatellite markers were not suitable for genetic studies of the 6 *Salvia* species used in this study because some primer pairs produced amplified products larger than 1 kb, some primer pairs produced multiple products or no amplification, and some primer pairs produced very weak amplifications (data not shown). Primer dimer and hairpinning were not possible because during the primer designing these 2 possible causes was omitted. Amplification failure may have resulted from annealing primer pairs onto neighboring exonic region(s) separated by intron(s). Failure may also have resulted from the primer synthesis errors and the incomplete PCR optimization for different annealing temperatures and $MgCl_2$ concentrations.

A total of 22 e-microsatellite primer pairs (Table 1) were successfully amplified across the plant samples used in this study and produced expressed microsatellite (e-microsatellite) markers. The higher amplification ratio indicates that the use of the contigs-based approach could not only reduce the analysis time but also increase the success rate of cross-transferable microsatellites (Ince, 2012). These primer pairs were named S primer pairs, standing for *Salvia*. Sequences, motif types, and repeat numbers of these microsatellite primer pairs are listed in Table 1.

Microsatellite repeat types and motif lengths of the amplified markers varied depending on the primer pairs. The most common microsatellite repeat type was dinucleotide repeats. Among the 22 e-microsatellite primer pairs, 10 produced dinucleotide repeats. The number of repeats in dinucleotide type (motif) microsatellites varied from 10 to 19 depending on the primer pairs. The second most common microsatellite repeat type was trinucleotides, varying from 7 to 8 repeats depending on the primer pairs. Within the 22 e-microsatellite primer pairs, 2 amplified tetra- and 1 amplified pentanucleotides. Figures 1a–1i show amplification patterns of some e-microsatellite markers in different *Salvia* species. The size of the markers varied from 180 bp to 750 bp in 3% agarose gels. It was noted that some of the primer pairs produced larger amplified products than their expected size in ESTs (Figures 1a–1i).

Based on the high-resolution agarose gel electrophoretic studies, a total of 89 alleles were observed across 22 loci; the number of alleles per locus ranged from 1 to 8 alleles (present or absent) depending on the species and primer pairs (Table 2). The mean genetic diversity index value (H or PIC) was 0.42 and the average number of alleles per locus was 4.05. The highest PIC value was obtained with the primer pairs S03 and S05, while lower PIC values were obtained with S11, S19, and S20. Relationships between repeat types, repeat numbers, and the PIC values of microsatellite loci were also studied and the results indicated that there were no relationships among these characteristics.

Transferability of e-microsatellite markers developed in the present study varied depending on the primer pair used (Table 2). All the 22 microsatellite markers transferred to 2 *Salvia* species, *S. tomentosa* and *S. dichroantha*, indicating that these 2 species were genetically related. Among the 22 primer pairs, 21 transferred to *S. fruticosa* and 19 transferred to *S. pisidica*, while 13 transferred to *S. sclarea* and *S. virgata* (Table 2).

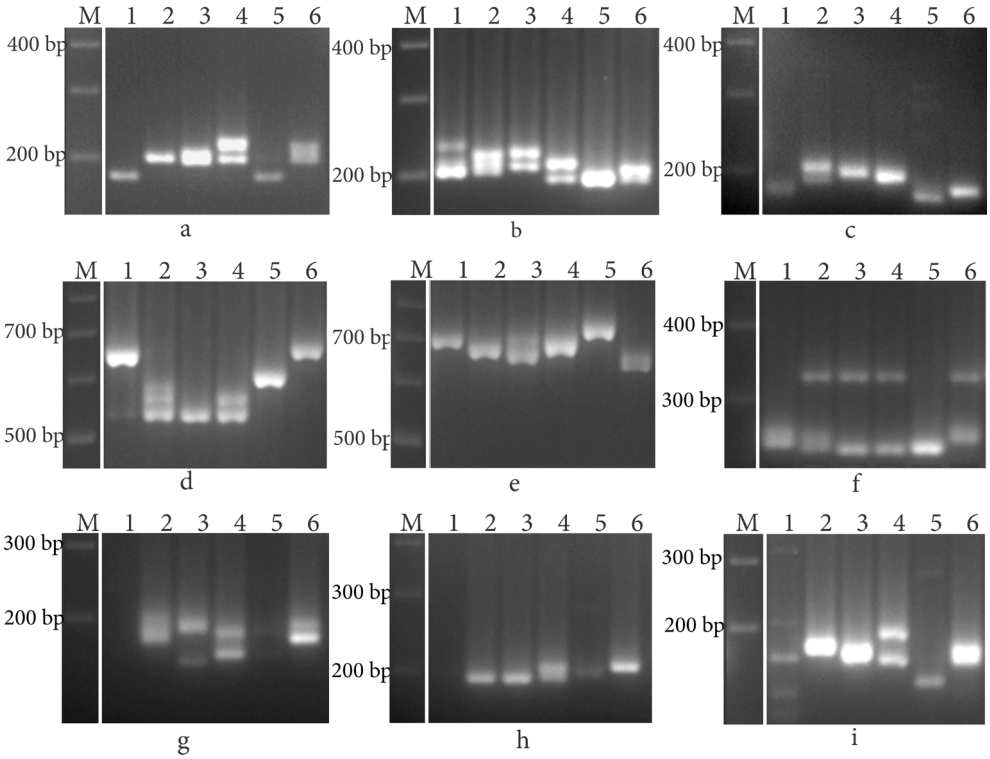


Figure 1. High resolution agarose gel images of some e-microsatellite markers in 6 different *Salvia* species. Numbers above each panel are genomic DNA templates from *Salvia* species used in this study. 1: European sage (*Salvia sclarea* L.), 2: Anatolian sage (*Salvia fruticosa* Mill.), 3: balsamic sage (*Salvia tomentosa* Mill.), 4: Turkish sage (*Salvia dichroantha* Stapf.), 5: meadow sage (*Salvia virgata* Jacq.), and 6: yurt sage (*Salvia pisidica* Boiss. & Heldr.). M: DNA markers in 3% high resolution agarose gels. Each panel shows e-microsatellite markers amplified: **a)** S01, **b)** S02, **c)** S03, **d)** S04, **e)** S05, **f)** S06, **g)** S07, **h)** S08 and **i)** S09.

4. Discussion

The majority of microsatellites of *Salvia* ESTs were mononucleotide type, but ESTs containing mononucleotide microsatellites were not used in primer designing studies. A total of 1524 ESTs contained the mononucleotide repeat type, while 199 ESTs contained di- to hexanucleotide microsatellites. This indicated that the occurrence and intensities of mononucleotide repeats were dominant in *Salvia* ESTs, as previously reported in *Arabidopsis*, human, and other plant species including *Salvia* (Bilgen et al., 2004; Mishra et al., 2011; Duran et al., 2013; Xu et al., 2013).

In silico analyses showed that CT/TC was the most frequent motif in the dinucleotide motif type, and GAA/AAG and AGA were the most frequent motifs in trinucleotide repeats. The highest number of repeats in dinucleotides was CT type, which had 38 repeats, and the highest number of repeats in trinucleotides was CTT, which had 26 repeats. These results were generally consistent with previous reports in barley, maize, cotton, sorghum, sage, and wheat (Ince et al., 2010c; Mishra et al., 2011; Radosavljevic et al., 2012; Duran et al., 2013; Xu et al., 2013).

The present study showed that some of the primer pairs produced larger amplified products than their expected size in ESTs. This was probably due to the presence of intron or introns as seen in one of our previous studies in which larger products were obtained using EST primer pairs (Ince et al., 2010c). Those primer pairs producing larger amplified products are also useful in cleavage amplified polymorphic sequence (CAPS)-microsatellite analysis, which is a relatively new approach as a modest modification of the CAPS technique (Ince et al., 2010b).

Studies of in vitro analyses indicated that there were no relationships among repeat types, repeat numbers, and the PIC values of microsatellite loci. Further studies are required to test whether there are any relationships among these characteristics since the number of samples and loci used in this study was limited. Development of primer pairs with higher ratio of cross-species and genera transferability is critical in *Salvia* species since the application of microsatellite markers in *Salvia* has been limited. Although random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism

Table 2. Cross-species transferability of 22 primer pairs in 6 different *Salvia* species.[a]

ID	*Salvia* species used						nA	H	T_C
	SS	SF	ST	SD	SV	SP			
S01	●	●	●	●	●	●	6	0.53	100
S02	●	●	●	●	●	●	8	0.61	100
S03	●	●	●	●	●	●	5	0.78	100
S04	●	●	●	●	●	●	6	0.53	100
S05	●	●	●	●	●	●	5	0.78	100
S06	●	●	●	●	●	●	4	0.61	100
S07		●	●	●		●	6	0.25	66.6
S08		●	●	●		●	4	0.19	66.6
S09	●	●	●	●	●	●	5	0.68	100
S10		●	●	●	●	●	2	0.36	83.3
S11		●	●	●		●	1	0.0	66.6
S12	●	●	●	●		●	5	0.69	83.3
S13	●	●	●	●	●	●	3	0.25	100
S14		●	●	●			3	0.25	66.6
S15			●	●		●	3	0.33	50
S16	●	●	●	●	●	●	3	0.41	100
S17	●	●	●	●	●	●	6	0.67	100
S18		●	●	●		●	5	0.31	66.6
S19		●	●	●			1	0.0	50
S20		●	●	●			1	0.0	50
S21	●	●	●	●	●	●	5	0.50	100
S22	●	●	●	●	●	●	2	0.44	100
Total	13	21	22	22	13	19	89		

[a]: Shown for each primer pair are the number of alleles (present or absent) (nA) and heterozygosity values (H). T_C is the cross-species transferability (%). SS: *Salvia sclarea* L., SF: *Salvia fruticosa* Mill., ST: *Salvia tomentosa* Mill., SD: *Salvia dichroantha* Stapf., SV: *Salvia virgata* Jacq., SP: *Salvia pisidica* Boiss. & Heldr. ● indicates successful transferability of primer pair.

(AFLP), inter-simple sequence repeat (ISSR), target region amplified polymorphism (TRAP), and sequence-related amplified polymorphism (SRAP) have been used to assess the genetic diversity and population structure of *Salvia* (Xu et al., 2013) microsatellite markers have many advantages in comparison to these techniques. Therefore, continued efforts should be made in order to increase the number of microsatellite loci available for genetic and conservation studies of *Salvia*.

Microsatellites reported in this study are e-microsatellite markers. Higher cross-species transferability ratios of these microsatellites were expected since the coding regions of genes are more conserved than other genomic regions. The results of the present study clearly indicate that cross-species transferability of e-microsatellite markers are related to the genetic relationship between the source and the target species. For instance, *S. sclarea* and *S. virgata* are genetically less similar to other species based on the chloroplast and mitochondrial fragment analyses (Karaca et al., 2008). A study by Walker and Sytsma in 2007 also revealed that *S. sclarea* and *S. miltiorrhiza* were in different clades of *Salvia*. This indicates that the low number of transferability between *S. sclarea* and *S. miltiorrhiz*a was due to their larger genetic distance. A recent study dealing with the development of cross-species transferable microsatellite markers clearly showed that cross-species transferability of microsatellite markers is closely related with the genetic relationship between the source and the

target (Karaca et al., 2013). Thus, EST sources of the genus *Salvia* or some other genera in the Lamiaceae family could be successfully used to identify cross-genera transferable e-microsatellite markers for uncharacterized genomes of the genera in the Lamiaceae family.

Comparison studies between individual and bulked samples collected from different species revealed that larger numbers of alleles were detected in the bulked samples than in the individual samples. This was expected since bulked samples consisted of different individuals. It was also noted that band density differences existed in some markers obtained from bulked samples and individual samples. This was probably due to the low amount of genomic DNA of individuals in bulked samples. E-microsatellite markers developed in this study do not require a capillary electrophoresis platform, but the resolution of the agarose gel system may be very low in comparison to capillary electrophoresis. In order to get high resolution of alleles produced with e-microsatellite markers developed in this study, the use of capillary electrophoresis system is recommended. With the use of a capillarity system most of the primer pairs could be used in multiplex PCR. The use of multiplex PCRs will definitely reduce the reaction costs and labor time for population genetic studies of *Salvia* species.

The use of e-microsatellite markers reported in this paper could be very useful in genetic and phylogenetic studies of nonmonophyletic genus of *Salvia* (Walker et al., 2004). The high level of species transferability of e-microsatellite markers developed in the present study could be used in comparative mapping studies in *Salvia*. For instance, Sim et al. (2009) reported that 57% of cereal e-microsatellite markers were transferable to ryegrass (*Lolium* spp.), and 37 markers were mapped on an existing ryegrass genetic map. Radosavljevic et al. (2012) developed 9 e-microsatellite markers and used them in a natural population of *S. officinalis* L. collected from the eastern Adriatic coast. It was reported that the number of e-microsatellite alleles was 8 to 21 depending on the primer, and PIC ranged from 0.46 to 0.93 (Radosavljevic et al., 2012). The number of alleles and the PIC values of this study are generally consistent with those results.

E-microsatellite markers developed from EST sequences are relatively inexpensive, require minimal labor to produce, and can be frequently associated with functionally annotated genes. For instance, Gao et al. (2004) reported that e-microsatellite markers showed significant similarities to some genes, including genes involved in metabolism, cellular structures, stress resistance, transcription, development, transporters, and storage proteins.

Despite many pharmacological and pharmaceutical properties of *Salvia* species, gene sequencing data are limited in comparison to other plant species. This study utilized EST sequences and developed 22 cross-species transferable expressed microsatellite markers. E-microsatellite markers developed in this study will be very useful for further investigation of the population genetic structure, gene flow, genetic diversity, and species identification studies. Recent studies revealed many pharmacological and pharmaceutical compounds. The use of e-microsatellite DNA markers could be a cost-effective alternative method to identify species with medicinal value, authenticate some *Salvia* species, and quantitatively detect the purity of their seeds or dry herbal products. To the best of our knowledge, this is the first report on cross-species transferable e-microsatellite markers for *Salvia* species, most of which are endemic to Turkey.

Acknowledgments

This work was partially supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) and the Scientific Research Projects Coordination Unit of Akdeniz University.

References

Bilgen M, Karaca M, Onus AN, Ince AG (2004). A software program combining sequence motif searches with keywords for finding repeats containing DNA sequences. Bioinformatics 20: 3379–3386.

Deng KJ, Zhang Y, Ren YL (2009). Identification, characteristics, and utilization of simple sequence repeat markers derived from *Salvia miltiorrhiza*. Acta Pharm Sinic 44: 1165–1172.

Duran C, Singhania R, Raman H, Batley J, Edwards D (2013). Predicting polymorphic EST-SSRs in silico. Mol Ecol Resour 13: 538–545.

Gao LF, Jing RL, Huo NX, Li Y, Li XP, Zhou RH, Chang XP, Tang JF, Ma ZY, Jia JZ (2004). One hundred and one new microsatellite loci derived from ESTs (EST-SSRs) in bread wheat. Theor Appl Genet 108: 1392–400.

Ince AG (2012). Contig-based microsatellite marker approach and its application in *Cichorium* ESTs. Rom Biotech Lett 17: 7177–7186.

Ince AG, Karaca M (2009). The MAGi RNA extraction method: a highly efficient and simple procedure for fresh and dry plant tissue. J Sci Food Agr 89: 168–176.

Ince AG, Karaca M (2012). Species-specific touch-down DAMD-PCR markers for *Salvia* species. J Med Plants Res 6: 1590–1595.

Ince AG, Karaca M, Bilgen M, Onus AN (2008). Digital differential display tools for mining microsatellite containing organism, organ, and tissue. Plant Cell Tiss Org 94: 281–290.

Ince AG, Karaca M, Onus AN (2010a). Polymorphic microsatellite markers transferable across *Capsicum* species. Plant Mol Biol Rep 28: 285–291.

Ince AG, Karaca M, Onus AN (2010b). CAPS-microsatellites: use of CAPS method to convert non-polymorphic microsatellites into useful markers. Mol Breeding 25: 491–499.

Ince AG, Karaca M, Onus AN (2010c). Differential expression patterns of genes containing microsatellites in *Capsicum annuum* L. Mol Breeding 25: 645–658.

Ince AG, Yildiz F, Karaca M (2011). The MAGi DNA extraction method for fresh tissues and dry seeds. J Med Plants Res 5: 5458–5464.

Karaca M, Bilgen M, Onus AN, Ince AG, Elmasulu SY (2005a). Exact tandem repeats analyzer (E-TRA) for DNA sequence mining. J Genet 84: 49–54.

Karaca M, İnce AG (2011). New non-redundant microsatellite and CAPS-microsatellite markers for cotton (*Gossypium* L.). Turk J Field Crops 16: 172–178.

Karaca M, Ince AG, Ay ST, Turgut K, Onus AN (2008). PCR-RFLP and DAMD-PCR genotyping for *Salvia* species. J Sci Food Agr 88: 2508–2516.

Karaca M, Ince AG, Aydin A, Ay ST (2013). Cross-genera transferable e-microsatellite markers for 12 genera of the Lamiaceae family. J Sci Food Agr 93: 1869–1879.

Karaca M, Ince AG, Elmasulu SY, Onus AN, Turgut K (2005b). Coisolation of genomic and organelle DNAs from 15 genera and 31 species of plants. Anal Biochem 343: 353–355.

Karaca M, Saha S, Callahan FE, Jenkins JN, Read JJ, Percy RG (2004). Molecular and cytological characterization of a cytoplasmic-specific mutant in pima cotton (*Gossypium barbadense* L.). Euphytica 139: 187–197.

Karan M, Evans DS, Reilly D, Schulte K, Wright C, Innes D, Holton TA, Nikles DG, Dickinson GR (2012). Rapid microsatellite marker development for African mahogany (*Khaya senegalensis*, Meliaceae) using next-generation sequencing and assessment of its intra-specific genetic diversity. Mol Ecol Resour 12: 344–353.

Mishra RK, Gangadhar BH, Yu JW, Kim DH, Park SW (2011). Development and characteristics of EST based SSR markers in Madagascar periwinkle (*Catharanthus roseus*) and their transferability in other medicinal plants. Plant Omics 4: 154–162.

Polat E, Ince AG, Karaca M, Onus AN (2010). Mining and utilization of mushroom ESTs for microsatellites. Conserv Genet 11: 1123–1126.

Radosavljevic I, Jakse J, Javornik B, Satovic Z, Liber Z (2011). New microsatellite markers for *Salvia officinalis* (Lamiaceae) and cross-amplification in closely related species. Am J Bot 98: e316–e318.

Radosavljevic I, Satovic Z, Jakse J, Javornik B, Greguras D, Jug-Dujakovic M, Liber Z (2012). Development of new microsatellite markers for *Salvia officinalis* L. and its potential use in conservation-genetic studies of narrow endemic *Salvia brachyodon* Vandas. Int J Mol Sci 13: 12082–12093.

Rozen S, Skaletsky H (2000). PRIMER 3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, editors. Bioinformatics Methods and Protocols: Methods in Molecular Biology. Totowa, NJ, USA: Humana Press, pp. 365–386.

Sim SC, Yu JK, Jo YK, Sorrells ME, Jung G (2009). Transferability of cereal EST-SSR markers to ryegrass. Genome 52: 431–437.

Teacher AGF, Kahkonen K., Merila J (2012). Development of 61 new transcriptome-derived microsatellites for the Atlantic herring (*Clupea harengus*). Conserv Genet Res 4: 71–74.

Walker JB, Sytsma KJ (2007). Staminal evolution in the genus *Salvia* (Lamiaceae): molecular phylogenetic evidence for multiple origins of the staminal lever. Ann Bot 100: 375–391.

Walker JB, Sytsma KJ, Treutlein J, Wink M (2004). *Salvia* (Lamiaceae) is not monophyletic: implications for the systematic, radiation, and ecological specializations of *Salvia* and tribe *Mentheae*. Am J Bot 91: 1115–1125.

Wang Y, Ren R, Yu Z (2008). Bioinformatics mining of EST-SSR loci in the Pacific oyster, *Crassostrea gigas*. Anim Genet 39: 287–289.

Xu G, Liu C, Huang L, Wang X, Zhang Y, Liu S, Liao C, Yuan Q, Zhou X (2013). Development of new EST-derived SSRs in *Salvia miltiorrhiza* (Labiatae) in China and preliminary analysis of genetic diversity and population structure. Biochem Syst Ecol 51: 308–313.