Plant Breeding, **134**, 589–598 (2015)

doi:10.1111/pbr.12296

Development, characterization and mapping of microsatellite markers for lentil (*Lens culinaris* Medik.)

Enver Ersoy Andeden[1,3], Faheem S. Baloch[1,4], Esra Çakir[2], Faruk Toklu[1,2] and Hakan Özkan[1,2,5]

[1]Department of Biotechnology, Institute of Natural and Applied Sciences, University of Çukurova, Adana, 01330, Turkey; [2]Department of Field Crops, Faculty of Agriculture, University of Çukurova, Adana, 01330, Turkey; [3]Present address: Department of Molecular Biology and Genetics, Faculty of Science & Letters, University of Nevşehir Hacı Bektaş Veli, Nevşehir, Turkey; [4]Present address: Department of Field Crops, Faculty of Agricultural and Natural Sciences, Abant Izzet Baysal University, Bolu, Turkey; [5]Corresponding author, E-mail: hozkan@cu.edu.tr

With 2 figures and 2 tables

Received January 21, 2015 / Accepted May 14, 2015
Communicated by Jens Léon

Abstract

Lentil is the sixth most important pulse crop terms of production in the world, but the number of available and mapped SSR markers are limited. To develop SSR markers in lentil, four genomic libraries for (CA)n, (GA)n, (AAC)n and (ATG)n repeats were constructed. A total of 360 SSR primers were designed and validated using 15 Turkish lentil cultivars and genotypes. The most polymorphic repeat motifs were GA and CT, with a mean number of alleles per locus of 7.80 and 6.55, respectively. Seventy-eight SSR primers amplified a total of 400 polymorphic alleles, whereas 71 SSR primers produced markers within the expected size range. For 78 polymorphic SSR primers, the average number of alleles per locus was 5.1 and PIC value ranged from 0.07 to 0.89, with an average of 0.58. A linkage map was constructed using 92 individual F_2 plants derived from a cross between Karacadağ × Silvan, with 47 SSR markers. The SSR markers developed in this study could be used for germplasm classification and identification and mapping of QTL in lentil.

Key words: simple sequence repeats — lentil — enriched genomic libraries — linkage mapping — genetic diversity

Lentil (*Lens culinaris* Medik.), which has a large genome size of 4063 Mbp/1C, is a diploid (2n = 14) self-pollinating ancient crop of classical Mediterranean civilization and continues to play an important role in the global human diet and in modern agriculture. Lentil was domesticated about 9000 BC from its wild progenitor *Lens culinaris* spp. *orientalis*, in areas that comprise modern-day south-east Turkey and the northern part of Syria (Sonnante et al. 2009). It is a very important pulse crop worldwide due to its high seed protein content; furthermore, it is a good source of essential minerals (Karaköy et al. 2012) and is highly appreciated by consumers. Additionally, recent studies have demonstrated that lectins, which are found in some legumes species, can reduce the risk of developing some forms of cancer, can activate innate defence mechanisms and manage obesity (Roy et al. 2010). Moreover, lentils are able to fix atmospheric nitrogen via symbiotic association and can thereby assist in the management of soil fertility (Karaköy et al. 2012).

Among legumes, lentil ranks sixth in terms of global production (FAOSTAT 2013) and is mainly grown on the Indian subcontinent, the Mediterranean region, North America and Australia. Turkey is one of the most important lentil-producing countries worldwide (after Canada and India) and ranks first in the Mediterranean region in terms of the total lentil production (about 55% of total lentil production of all countries bordering the Mediterranean Sea). Historically, lentils have mostly been grown in areas of the developing world where funding for genetic research is scarce, and there is paucity in the development and implementation of molecular techniques into lentil breeding in comparison with that available for cereal and other crop species. Even though lentils have been an important food legume for centuries, very little effort has been made to perform genetic studies and gene mapping in this crop.

Microsatellites, also known as simple sequence repeats (SSRs), consist of tandem repeats of simple nucleotide units (1–6 bp), which are widely spread throughout the genomes of plants and animals (Jarne and Lagoda 1996). Following their description in plants by Condit and Hubbell in 1991, SSRs have become the marker system of choice by many researchers because of their reproducibility, transferability, multiallelic codominant inheritance, whole-genome coverage and high degree of polymorphism (Alsaleh et al. 2015). Thousands of SSR markers have been developed for economically important crops such as wheat (Song et al. 2005), barley (Zhang et al. 2014) and corn (Sharopova et al. 2002) and have been usefully applied for a variety of purposes.

The lack of available molecular markers limits genetic and molecular studies on lentils. Nonetheless, several genetic studies, including those on genetic diversity, have been carried out on lentils based on isoenzyme and DNA markers (Tahir and Muehlbauer 1994, 1995, Tahir et al. 1994, Sharma et al. 1996, Rodriguez et al. 1997, Eujayl et al. 1998, Duran et al. 2004, Kahraman et al. 2004, Toklu et al. 2009, Tanyolac et al. 2010, Baloch et al. 2015). Despite the existence of substantial diversity among lentil landraces and cultivars, at both the genotypic and phenotypic levels, no effective molecular breeding programme has been developed. The genetic basis of most of the traits in lentil remains unknown, and no comprehensive genetic map is available. Recently, some genetic linkage maps were developed with the PCR-based markers, and the number of available markers across the *Lens* genome increased dramatically (Kumar et al. 2015). First linkage map using DNA molecular markers was produced by Havey and Muehlbauer (1989). Subsequently, Hamwieh et al. (2005) added 39 SSR and 50 AFLP markers to the map constructed by Eujayl et al. (1998) to produce a comprehensive *Lens* map comprising 283 genetic markers covering 715 cM. In the last decades, some maps were produced using different DNA molecular markers (Perez de la Vega et al. 2011, Gupta et al. 2012a,b, Saha et al. 2013, Sharpe et al. 2013). Recently, EST-SSR markers were developed and used to study genetic diversity (Alo et al. 2011, Kaur et al. 2011). The first report describing the isolation of microsatellites in lentil was published by Hamwieh et al. (2005, 2009). These authors developed

44 genomic SSR markers for lentil and used them in genetic diversity studies and in linkage mapping analysis. Recently, 5673 EST-SSR primers and 122 genomic SSR primers were designed, and a subset of these was utilized for diversity analyses (Verma et al. 2013, 2014). Although these examples represent significant advances in the development of molecular tools for lentil characterization, they are still limited in their ability to comprehensively analyse the lentil genome, and to be applied to in-depth molecular breeding studies, when we compare it with wheat and barley, where thousands of the SSR markers are available for constructing the saturated linkage map and QTL analysis which in turn helped to identify linked markers for traits of interest. In conclusion, many functional markers are now available in wheat and related crops for genomic-assisted breeding due to the availability of genomic resources. Similarly, in comparison with major legume crops such as soybean, common, bean, pigeon pea and chickpea, the pace of development of genomic resources was slow in lentil (Kumar et al. 2014). One of the major concerns in the genetic characterization and breeding of lentil was the lack of informative SSR markers. The recent application of the next-generation sequencing and genotyping by sequencing technologies has accelerated the lentil genome sequencing project and large discovery of genomewide single nucleotide polymorphism (SNP) markers. Transcriptome analysis of the lentil genome showed that 10 341 ESTs available for lentil (NCBI 2015), which could be beneficial source of genomic studies in lentil. Despite these advances, the number of available genomic SSR markers for lentil is only 166 (Hamwieh et al. 2005, 2009, Verma et al. 2014). Thus, the goals of this study were (i) to develop a collection of reproducible SSRs using genomic libraries enriched in CA, GA, AAC and ATG repeats, (ii) to assess polymorphism of SSR markers in 15 Turkish lentil genotypes using capillary electrophoresis and (iii) to confirm segregation of new SSR markers in an F_2 population.

Materials and Methods

Plant material and DNA extraction: Fifteen lentil genotypes, consisting of nine lentil cultivars ('Emre', 'Seyran-96', 'Çifçi', 'Özbek', 'Kafkas', 'Şakar-91', 'Çağıl-2004', 'Fırat-87' and 'Altıntoprak') and six lentil landraces (Karacadağ, Silvan, Kumçatı, Kışlık kırmızı, Hacıbey and Yerli kırmızı), were used as plant material to evaluate the amplification and polymorphism of the developed microsatellites. For linkage analysis, an F_2 population was developed from a cross between Karacadağ and Silvan genotypes. Two local red lentil landraces from Anatolian Plateau having contrasting agronomic and morphological features were selected as parents for developing mapping population. Both of the parents were selected as single plants from local landraces and selfed for two generations to stabilize their genetic background. Karacadağ is local Turkish landrace collected from Diyarbakır Province located in south-east Turkey, core area of lentil domestication and diversity. This landrace exhibited pubescent leaves, erect growth habit and tendril leaves, owing to have a wide adaptation. These landraces have many contrasting and distinct features. Toklu et al. (2009) clearly mentioned that Karacadağ landrace was very diverse from the rest of the landraces based on AFLP and ISSR analyses. Karacadağ landrace is early-flowering and early-maturing and suitable for mechanical harvesting, and exhibited higher 100-grain weight, larger grain size, taller plant height and other agronomical characters when compared with Silvan. The F_2 populations consisted of 92 plants produced by four self-fertilizing F_1 plants. Genomic DNA extraction was carried out according to the CTAB protocol with minor modification, as described by Ozkan et al. (2005).

Construction of enriched genomic libraries: Genomic DNA isolated from the Karacadağ line was used to construct the genomic libraries. Four different repeats [(CA)n, (GA)n, (AAC)n and (ATG)n] were selected for the construction of four independent SSR-enriched genomic libraries. This selection was based on evidence that these SSRs are abundant in plant genomes (Gupta and Varshney 2000). Four microsatellite-enriched libraries were constructed by Genetic Identification Services (http://www.genetic-id-services.com). Enriched DNA was ligated into the pUC19 plasmid following digestion with *Hin*dIII (New England Biolabs), and the recombinant plasmids were electroporated into *Escherichia coli* strain DH5α. Colonies were grown overnight on LB agar plates containing ampicillin (100 mg/l), Xgal and Bluo-Gal. Twelve recombinant clones from each library were selected at random to undergo sequencing on an ABI 377, using Amersham's DYEnamic™ ET Terminator Cycle Sequencing Kit (Amersham Biosciences P/N US81050, Little Chalfont, Buckinghamshire, UK), according to the manufacturer's instructions. Based on these enrichment results, 108 recombinant clones from each enriched library were sequenced. Sequences were compared using ClustalW, and duplicated sequences were eliminated. Microsatellite repeats were classified as perfect, imperfect and compound according to Weber (1990). PCR primers were designed from flanking regions using DesignerPCR, v.1.03 (Research Genetics, Huntsville, AL, USA) with the following parameters: annealing temperature 55–60°C, GC content 35–60% and amplicon size 100–350 bp.

PCR conditions and SSR marker genotyping: PCRs were carried out in a total volume of 12 µl containing 50 ng of genomic DNA (Karacadağ line), 1X Dream *Taq* buffer (Fermentas, USA), 1.2 mM dNTP, 5 µM forward and reverse primer, and 0.60 U/µl Dream *Taq* DNA polymerase (Fermentas, USA). For each primer pair, a gradient PCR with temperatures ranging from 48°C to 65°C was carried out to determine the optimum annealing temperature using an Eppendorf Gradient Thermocycler. The PCR programme consisted of an initial period at 95°C for 5 min followed by 30 cycles of 1-min denaturation at 95°C, 1 min at the appropriate annealing temperature and 1-min extension at 72°C, and a final extension period of 10 min at 72°C. Amplification products were analysed by gel electrophoresis on 3% agarose gels in 0.5X TBE buffer stained with ethidium bromide and photographed under ultraviolet light. An additional lowercase letter (a or b) was added when two SSR markers were isolated from the same clone sequence.

SSR primers were screened for amplification and polymorphisms in 15 lentil genotypes. M13-tailed primer PCR amplification of SSRs was used and was performed according to the methods described by Schuelke (2000) in 12 µl PCR mixes containing 1X buffer, 0.125 mM dNTP, 0.4 pmol M13-sequence-tailed forward primer (TGTAAAACGACGGCC AGT), 0.3 pmol reverse primer, 3 pmol universal M13 primer (TGTAAAACGACGGCCAGT) labelled with one of four fluorescent dyes (6-FAM, VIC, NED and PET), 0.12 U/µl *Taq* DNA polymerase and ~50 ng genomic DNA. The PCR conditions consisted of an initial denaturation at 95°C for 5 min, 35 cycles of denaturing at 95°C for 1 min, 2 min at the appropriate annealing temperature and extension at 72°C for 2 min, followed by five cycles of denaturation at 94°C for 30 s, annealing at 53°C for 1 min and extension at 72°C for 1 min, with a final extension at 72°C for 10 min. For all SSR primers with an annealing temperature lower than 55°C, an alternative method was used, as reported by De Arruda et al. (2010) with reduced non-specific amplifications. The use of this protocol reduced the amount of unspecific binding by the M13 primers. The PCR products were then stored at 4°C until analysis. A set of four PCR products (0.75 µl of each), each labelled with a different dye, was combined with 0.14 µl GeneScan-500 LIZ® size standards (Applied Biosystems) and 6.86 µl Hi-Di™ formamide (Applied Biosystems), denatured at 94°C for 5 min, chilled on ice and separated using an ABI 3130xl Genetic Analyzer (Applied Biosystems). GENEMAPPER software v3.7 (Applied Biosystems) was used to determine fragment size, as described in the user manual.

Data analysis: The quality of amplification and the number of loci were recorded for each SSR marker. After determining the allelic profile at

each SSR locus, polymorphism information content (PIC), gene diversity (He) and the number of polymorphic alleles were calculated using PowerMarker v.3.25 (Liu and Muse 2005). Genetic similarities were calculated according to the method developed by Jaccard (1908). The Jaccard genetic similarity matrix was used to build an unweighted pair-group method with arithmetic means (UPGMA) tree. NTSYS-PC version 2.1 (Rohlf 2004) was used for genetic similarity computing and dendrogram construction.

SSR markers for F_2 populations were scored as codominant. Observed segregation ratios were compared with the expected Mendelian ratios using chi-square (χ^2) goodness-of-fit tests. A linkage map was constructed using the program 'JOINMAP' v.4.0 (Van Ooijen and Voorrips 2001). To determine and calculate the linkage groups, a log of odds (LOD) threshold was set at a minimum of six. Kosambi's mapping function was applied to estimate genetic distances in cM.

Results

To determine the success of the enrichment process for each of the four enriched libraries, 108 clones from each library were sequenced and scanned for microsatellite motifs. Nine of 432 clones were duplicated, and around 30% of clones were redundant. Thus, after accounting for redundancy and duplicated sequences, 360 unique microsatellite-containing sequences remained in 301 clones, including 70, 81, 67 and 83 from the CA, GA, AAC and ATG libraries, respectively (Table 1).

Of the 360 microsatellite loci identified, we successfully designed flanking primer pairs for all SSR loci including 84, 98, 83 and 95 from the CA, GA, AAC and ATG libraries, respectively. The remaining clones produced sequences that contained no discernible microsatellite repeats or that generated insufficient flanking DNA to construct PCR primers. A total of 360 SSR markers were developed from 301 clone sequences, which contained 133 perfect dinucleotide motifs, 169 perfect trinucleotide motifs, 14 perfect tetranucleotide repeats, 4 perfect pentanucleotide repeats, 9 compound repeats and 31 imperfect repeats (Table 1). The longest repeat motif was the GA with 35 uninterrupted repeats. Some information about 360 SSR markers developed in this study is presented in Table S1.

All primers were initially tested and optimized on the lentil genotypes Karacadağ and Silvan by gradient PCR. A total of 220 SSR markers produced PCR products of expected sizes, while 140 primers either failed to generate an amplification product or produced a complex pattern of bands that was difficult to evaluate. For 57, 57, 57 and 49 SSR markers from CA, GA, AAC and ATG libraries, respectively, an annealing temperature was successfully detected. To characterize how informative these markers were, 149 SSR markers were screened for polymorphisms among 15 lentil genotypes from Turkey. Seventy-eight SSR markers (52%) detected polymorphisms among 15 lentil genotypes (Table 2), while 71 SSR markers (48%) were found to be monomorphic. The 78 polymorphic SSRs included 31 that belonged to the CA-enriched library, 24 belonging to the GA-enriched library, 12 belonging to the AAC-enriched library and 11 belonging to the ATG-enriched library. Among 78 polymorphic SSR, 61 were perfect repeats (40 dinucleotide, 19 trinucleotide, 2 tetranucleotide), and the remaining 17 were six compound SSRs and 11 imperfect SSRs. The 78 SSR markers amplified 400 alleles on 15 lentil genotypes with a mean number of 5.1 alleles per locus, ranging from 2 to 11. The expected heterozygosity per locus was 0.12–0.90 with an average of 0.62. The average value of PIC for SSR marker sets was 0.58, ranging from 0.07 (CULD309) to 0.89 (CULA109) (Table 2). The allelic data obtained with 78 primer pairs across 15 lentil genotypes were scored and computed to obtain the neighbour-joining dendrogram, which was able to clearly distinguish all lentil genotypes (Fig. 1).

To evaluate the utility of these markers for future linkage map construction, inheritance of the microsatellite loci was investigated in a segregating F_2 population (Karacadağ × Silvan). In total, 47 SSR markers were polymorphic between the parents and 92 screened individuals of the F_2 population. Linkage analysis, performed using the program JOINMAP (Van Ooijen and Voorrips 2001), identified seven linkage groups consisting of 43 linked microsatellite loci spanning 303.9 cM with a marker density of 7.06 cM/marker (Fig. 2). The mean length for all linkage groups was 43.4 cM; LG2 was the longest with 68.9 cM, whereas LG7 was the shortest with 10.6 cM. The total number of mapped loci per linkage group ranged from 3 for LG6 to 10 for LG1. However, the distribution of SSR markers between linkage groups was unequal. Even though the average marker density was 7.06 cM/marker, there was a large gap between markers in LG3 and also very close markers in LG1.

Discussion

In the last few decades, breeding efforts have started to pay attention to lentil. However, molecular breeding efforts in lentil are falling behind those made for other crops. A weakness in genomic studies in lentil is the lack of available molecular markers. Until now, RAPD, ISSR and AFLP markers, which are

Table 1: Summary of SSR marker development and PCR analysis

Development stage	(CA)n	(GA)n	(AAC)n	(ATG)n	Total
Clones sequenced	108	108	108	108	432
Number of clones used for SSR markers design	70	81	67	83	301
Number of SSR markers designed	84	98	83	95	360
Number of perfect dinucleotide repeats	49	82	0	2	133
Number of imperfect dinucleotide repeats	9	7	0	0	16
Number of perfect trinucleotide repeats	8	2	74	85	169
Number of imperfect trinucleotide repeats	0	0	8	7	15
Number of perfect tetranucleotide repeats	5	7	1	1	14
Number of perfect pentanucleotide repeats	4	0	0	0	4
Number of compound repeats	9	0	0	0	9
Number of SSR markers detected AT[1]	57	57	57	49	220
Number of SSR markers analyzed by PCR	42	29	44	34	149
Number of polymorphic SSR markers	31	24	12	11	78
Number of monomorphic SSR markers	11	5	32	23	71

[1]Annealing temperature.

Table 2: Primer name, their sequence, repeat motif, annealing temperature and some diversity parameters

SSR markers	Forward primer	Reverse primer	Repeat motif	Ta (°C)	Observed allele size range (bp)	Na	He	PIC
CA-enriched library								
CULA3	ATCCTTCTTCGGCACTTG	AAACGATTGTGTGTAGTTGTTG	$(CA)_{12}$	65	249–257	4	0.68	0.61
CULA7	CACGCGATTAGAGGATCA	CTCACCTGGTTTATGAAAGAAT	$(CA)_{16}$	53	231–284	6	0.75	0.71
CULA9	TTCCTTTCCTTATTTCCTCTTG	AACGAATCTGAGCCACTTG	$(TG)_{12}$	68	241–253	5	0.68	0.64
CULA10	ATTCCTTGTGTCATTTACGTTC	CATAGGTTTTGGGAACAGATC	$(TA)_7(TG)_8$& $(TG)_{14}$	53	131–161	6	0.81	0.78
CULA103	TGTTCGATTTTTTAAGGTGCTG	GGAAGTTGGAAGTGGATTACGT	$(CA)_{11}$	63	284–316	6	0.76	0.72
CULA105	CGACAGATATGTCCACACTC	CCAAACTTTTGCTTTTGTC	$(TA)_7(TG)_{22}$	58	142–178	9	0.88	0.86
CULA107	TTGGTTGACAAGATCACAATC	CTCGTCACGGTAATCTATCATC	$(CA)_7$&$(CA)_7$	63	287–290	2	0.50	0.37
CULA109	CGAAGAGAGATAACAACAATG	TTTTTTGTCCCTATGATGG	$(TG)_{15}A(GA)_{29}$	50	355–417	11	0.90	0.89
CULA114	GCCACAGCCATGCTTTAC	TATCGTATGGGGTTGTGTAATC	$(AC)_{11}(AT)_9$	63	244–246	2	0.12	0.12
CULA116	ATGCAACAAATATAGCCACTGC	GAGGTTGATTCTGCCATGTTAC	$(CA)_{17}$	65	108–116	4	0.71	0.66
CULA119	AACAAGCTGCAACAAACTTG	TGCAACAAAGACCTTTTATCC	$(AC)_{12}$	63	98–110	6	0.78	0.75
CULA121	CGACAAAACTCCAAAGAACC	GAGGGCGAGGAAGAAGAG	$(AC)_8$	58	270–315	5	0.70	0.65
CULA123	TCGAGCTGAACACATCAAC	TAGCAGTGTATGTAGCCATGAG	$(AC)_{10}$	61	189–195	4	0.58	0.53
CULA211	AATTGGCTAGGTCTTGAAAAC	GAGGAAGTGAGAAAACTCGTC	$(GT)_{23}(GA)_{18}$	63	243–276	10	0.85	0.84
CULA216	GGAAGAAGAACCTGAAAATAC	ATGCAGAAACGCTCTCTT	$(AG)_9$	50	147–157	4	0.65	0.60
CULA219	AAATCCCTCAAGTGTTTATGTG	TAACCCTATCCCTTTTACAACC	$(AC)_{15}$	58	156–167	5	0.78	0.75
CULA301	AAATCCCTCCCTCACATTC	TTTCCAGTGGGTTTTATCG	$(AC)_{13}$	58	251–266	5	0.70	0.65
CULA305	CGTTGTTCCTCTTCTCATTTC	TGGTTGAATCAAGACGTAGTG	$(CA)_{18}(CT)_8$	58	230–239	4	0.73	0.68
CULA308	ATTTGGAGTGCAAGTAACCTAC	CCTGAACACACGAACATTG	$(TC)_{20}A(CA)_6$	58	237–273	8	0.82	0.80
CULA309	CACATTAGTGAAAGATCCTTGTGC	CGTCGGAGCTACTACCGAGT	$(AC)_{21}$	58	179–192	5	0.77	0.74
CULA311	AGTGCGGAACTTGTTCTTGA	TTGGTTTGGAACTACACAACG	$(GT)_{14}(GA)_{19}$	58	164–264	7	0.80	0.77
CULA312	TCCAACATCTTGTCCAACATC	CAGGACGTAACTCATGTGACC	$(CA)_7$&$(CA)_9$ &$(CA)_7$	65	261–287	5	0.70	0.65
CULA323	GTTCCTGACAATTCTTGAAGTC	AGGCTTGAAAACATGCTTT	$(CA)_9$	58	185–205	7	0.78	0.74
CULA405-b	CCCACGTGTGTTTAGAACCAT	TGGCATATTGGAACATTGACAT	$(ATTA)_3$	65	164–177	2	0.12	0.12
CULA408	CAACTTTCTCTCTCTCTCTCTCTC	GGTTTGTGGGGTCAGTCAGT	$(CA)_{11}$	65	134–157	8	0.83	0.81
CULA413-b	CGACACTCTGGGTGAGAAGAG	CATGCAAAAATCAAGCGAAA	$(AC)_{14}$	58	273–319	9	0.86	0.85
CULA413-d	CACACACACACACACACACA	CATGCAAAAATCAAGCGAAA	$(TAAC)_4$	63	190–194	3	0.56	0.48
CULA414	TCACTTGTTGTGTGTGTGTG	AAAAATTCACTGGCACCAAA	$(GT)_7$	63	123–151	5	0.70	0.65
CULA415	CATGCCAAATTTCAAAATGC	CAACACAAATGGCATGAAACA	$(TG)_{21}$	48	109–144	8	0.83	0.81
CULA421	ACCACGTAAAATATGTCTTTGG	CAGAGAACTTCGACCAACTTAG	$(AC)_{15}$	53	193–216	6	0.78	0.75
CULA422	TGTCACTAGTCTTATATGTGCCAAA	TGAAGGGGTGTGTTTACCAGA	$(TG)_{12}$	58	264–302	3	0.42	0.37
GA-enriched library								
CULB3	TCAAACTTTCGCACAGAATAAC	GTTCGCTGCATGAGGAAG	$(TC)_{19}$	55	80–109	6	0.76	0.73
CULB7	CCAAAGGAAAGGGATGAG	AGGGAGATGAGAGTGAGGTC	$(CT)_7$	63	192–237	8	0.85	0.83
CULB9	ACGTGGTGAAACTTTTGG	TGGGATTTGTTTTGAGAAG	$(CT)_{24}$	58	178–204	9	0.85	0.83
CULB107	GCCAAAATTGAATAAACCCTC	GATTTGAGTGGCGGATTTC	$(CT)_{28}$	58	163–207	9	0.83	0.81
CULB113	GTTTGGTTTGAGGAATAGGTC	AATTACACTAGGTCGCCATTAG	$(GA)_{23}$	63	108–130	6	0.80	0.77
CULB114	CAACTTTTGCATTGGAGATAAC	CGAGTGTCCTCATCAATTTAAC	$(TC)_{26}A(CA)_{13}$	58	102–152	8	0.81	0.79
CULB115	TGGAAGAGTCAAGGAGTGG	AGGTTCACACCCAGTTTCTC	$(GA)_{21}$	63	213–250	7	0.76	0.73
CULB118	CTCTCGTCGGGATTCATAC	GAAAAGGGGGATGTGTTTAG	$(CT)_{23}$	53	80–125	5	0.76	0.72
CULB205	AACCGATAGATTCGTTGGG	GTTTCAGCCTGGACATTGA	$(GA)_{24}$	63	256–272	6	0.78	0.75
CULB206	AACCCACGCAATCAGTTT	ACAACAACACCTCTTCTCAGTC	$(CA)_{17}$&$(CA)_6$	48	224–258	9	0.86	0.85
CULB217	TAGGGCTTTTCCTCCTTTCC	GGTGAAACATTCACGTAACAGAG	$(CT)_{31}$	58	144–178	10	0.86	0.85
CULB218	TTCCTACGTTTCCTCACATATT	AGCCAAACTAATAGCAGCATAC	$(AG)_{19}$	63	163–197	4	0.57	0.51
CULB222	ATATGGGTGCGTGTGAATATAC	ACCAAACTTTTCTCGTTTCTTC	$(GA)_{28}$	63	125–179	9	0.83	0.82
CULB305	CGTCAAAATCGTAAAGAAAGTG	GAGCGACAAGAATCAACATC	$(GA)_{15}$	48	205–235	7	0.83	0.80
CULB308	TCATGGACCTAACCTAGATGC	GGTTTGAGGGTTCTATGAGATC	$(TG)_6(AG)_{14}$	63	229–302	7	0.81	0.78
CULB310	AGACGCTGACATCCTGTATG	AAGAGAAGGGAGAAGGTGATT	$(TC)_{18}$	63	267–315	8	0.85	0.83
CULB311	TTTCTTTTGCCACCACACAA	AAGGGTGGAAGAGTCAAGGAG	$(CT)_{22}$	58	108–148	7	0.76	0.74
CULB402	TCTACGAACAAAGGGGTCTC	GGACTGAAATCCTCCATAGG	$(TC)_8$	63	232–236	3	0.57	0.50
CULB405	ATTGTCATCACACACCTACCC	TCCTGCACTTGGAAACATGA	$(TC)_{30}$	63	131–172	6	0.78	0.74
CULB414	TCACGAAAGAAGGCAACAA	GCAAGGAGAGCAAAGAAGC	$(CT)_{18}$	55	151–153	2	0.12	0.12
CULB416	TGACGACTCTGTTGATTTACTG	GTAATTGGGTCTAATGGAGTGA	$(TC)_{10}C(CT)_7$	58	288–315	5	0.78	0.74
CULB418	TAGGCAAGAGAGGAATAGGAG	GCATACAACCACCATCATAAC	$(GA)_{28}$	63	218–279	8	0.85	0.83
CULB419	GGAAAACAGAGCATGTGAAT	TCCCAATTCCATGATTCTC	$(GA)_{35}$	58	117–160	7	0.84	0.82
CULB423-a	ACACTCTCACACACGCACATA	ATTAGTCCGCAAGACAAGTGAC	$(TC)_6$	63	209–265	10	0.85	0.84
AAC-enriched library								
CULC109	TGGGGAATTCCTATGCATGT	AACCCAAACTTCCCAAACCT	$(CTG)_4$	65	207–210	2	0.46	0.36
CULC113-b	TGGGGTGTTTGTGTTTGTTG	CCAATCCCAATCCAATCAAG	$(TTG)_5$	58	110–116	3	0.66	0.58
CULC206	CCGGTTCAAGCTCAATTTTC	TCGGTATTGGGTCGAAACTC	$(CAA)_4$	65	150–153	2	0.06	0.06
CULC207	CCAGAAGGAATTACTGTGAGTT	CATGGCTTAATCCTAAATCATC	$(GTT)_4$&$(TGT)_4$	53	192–210	2	0.39	0.31
CULC208	TGTGGGTCGTATGACCATTG	CATGCGACTCATACGGACAA	$(TGT)_5$	65	172–194	5	0.67	0.62
CULC221	ATCTCAACATCGACTCCACTAG	GAGTTGTGACCTCACGTTCTAG	$(ACA)_4$&$(ACA)_2$	65	199–202	2	0.32	0.27
CULC302	CCCATTTACGCCAACTTAATC	GGGATTGTTGTTGTCAATGTG	$(ACA)_7$	63	225–231	3	0.63	0.56
CULC404	CTTGCCTAATCGTGACATGC	TCATGCAACAACACGTAACG	$(GTT)_4$	63	291–300	3	0.52	0.44

(continued)

Table 2. (continued)

SSR markers	Forward primer	Reverse primer	Repeat motif	Ta (°C)	Observed allele size range (bp)	Na	He	PIC
CULC409	CATCCGTGCCATAGACTTATC	TGAGTTATCCAGAGGGGATTAC	$(ACA)_5$	63	248–251	2	0.26	0.23
CULC410	AGAAGGAATTACGGTGAGTGG	GCTGTGTAACGCTCCATCTAA	$(TGT)_5$	65	248–260	3	0.27	0.26
CULC411-b	TGATGATGAGGTTAGGAACGAA	GGACCACCGTCCAAATATGA	$(TGT)_4$	65	195–203	2	0.12	0.12
CULC414	TTAGCTCCAACTCAAAACATG	AACTTGTTCCAACATTGTTACC	$(CAA)_7$	58	156–162	2	0.39	0.31
ATG-enriched library								
CULD12-a	TCTCAGGCTCAGCAAAATCC	AGGAGGGTGATGATGACGAT	$(TCA)_4$	63	151–166	2	0.32	0.27
CULD117-a	TCTCACAAACCACCTCTCTCAA	TCCCACGGTATGGACGTAGT	$(AT)_6$	58	144–165	5	0.61	0.58
CULD123	ATGGATGCGTGGACTCTC	TTTGCCTCGTTTGAGTAGC	$(CAT)_4$	63	264–288	2	0.12	0.12
CULD206	CGGATGGTAATTGATTTAGTG	CCACAAAACCTCTCATCG	$(GAT)_9$	58	215–224	4	0.68	0.62
CULD207	TCCGAAAGGAAACAAAACA	CCGAAGGTGGTGTCTCTAA	$(ATC)_9$	53	271–282	3	0.43	0.39
CULD222	CATCCACAACCACATCGAGA	ATGCGGATCGTGTGTGTTTA	$(CAA)_4$	63	189–192	2	0.49	0.37
CULD303	CTACCCCATTCACAGAAAATC	GGTTGAGCTGCTTAATAATACG	$(TCA)_5$	58	281–284	2	0.12	0.12
CULD309	GCCATGAATTTATGTTGAGTTG	ATACCCCTCTTAGGCAGGAG	$(GTT)_5$	65	234–237	2	0.07	0.07
CULD318-b	GCTGTTTTTGGAGTTGTTGTTG	CATCACTGGACCGAAGTCCT	$(TGA)_4$	65	158–160	2	0.12	0.12
CULD415	GCATGGACTCTCATACCACAC	TCCGACGTATAGGGATGAAAT	$(CAT)_3$&$(ATC)_3$	63	219–229	3	0.58	0.51
CULD416	CACTGGATCGAAGTCCTTGAC	ACGTTTAGGCAGCAATGTGTT	$(CAT)_3$	63	281–300	2	0.12	0.12

Ta, annealing temperature; Na, number of alleles; He, expected heterozygosity; PIC, polymorphism information contents; &, more than one nucleotide.

unspecific and dominant in nature, have been used for studies in lentil. Linkage mapping, QTL mapping and marker-assisted breeding studies in lentil have been fairly limited. Therefore, powerful and informative genetic markers are needed to serve different purposes. Genomic and genic SSR markers have been developed by lentil researchers and used in studies investigating linkage mapping and genetic diversity (Hamwieh et al. 2005, 2009, Gupta et al. 2012a, Kaur et al. 2014, Verma et al. 2014); however, the number of available genomic SSR markers in this legume is still limited in number compared with crops of economic importance. Therefore, the present study aimed to develop genomic SSR markers from four enriched genomic libraries.

The development of SSR markers is not an easy task due to the high cost, time and labour required to design the primers (Zane et al. 2002). Once developed, however, microsatellite marker approaches become cost-effective. Currently, many different protocols and methods can be used to develop SSR markers, which include microsatellite enrichment (Hamilton and Baulcombe 1999), 5′-anchor polymerase chain reaction (Fisher et al. 1996), sequenced-tagged microsatellite profiling (Hayden and Sharp 2001), database BLAST search (Altschul et al. 1990) and selectively amplified microsatellite and next-generation sequences (Wang et al. 2010). The enrichment technique is based on the principle of capturing microsatellites from genomic DNA by hybridization with synthetic oligonucleotides bound to nylon membranes or magnetic particles (Zane et al. 2002). This method has been widely used in many plant species such as safflower (Hamdan et al. 2011), coffee (Missio et al. 2010), chickpea (Sethy et al. 2006), rice (Brondani et al. 2002), hazelnut (Gürcan and Mehlenbacher 2010), wheat (Song et al. 2005) and bean (Benchimol et al. 2007).

It is apparent that not all SSR motifs are equally abundant in eukaryotic genomes and the relative abundance of different motifs varies among different species (Ferguson et al. 2004). In the target regions, TC, CT and GA dinucleotide repeats, and CAT, CAA and TGT trinucleotide repeats have the highest frequency among 133 dinucleotide perfect repeats (36.9%) and 169 trinucleotide perfect repeats (46.9%), respectively. The dinucleotide motif, TC, was the most abundant motif, followed closely by CAT, then CT and CAA. This is in contrast to previous surveys that have assessed microsatellite abundance in plant genomes, where AT repeats were found to be the most predominant (Mace and Godwin 2002, Odeny et al. 2007). On the other hand, Hamwieh et al. (2009) reported that CA and GT dinucleotide motifs might be the major microsatellites in the lentil genome. However, the finding that TC is the most abundant repeat motif in the lentil genome might be due to the microsatellite enrichment procedure used. The enrichment procedure used, involving several PCR steps, a selective hybridization step, and a mixture of different repeats motifs can affect the frequency of the repeat clones (Van Der Schoot et al. 2000). Meanwhile, CT and GA repeats have been reported to be highly polymorphic in other plant genomes such as rice, bean, tomato and peanut (Cho et al. 2000, Gaitan-Solis et al. 2002, He et al. 2003, Ferguson et al. 2004). Differences in the frequencies of dinucleotide and trinucleotide motifs have variously been ascribed to conformational properties of DNA (Ashworth et al. 2004). For genetic diversity studies, a subset of highly informative loci that are robust and well defined, and that give good coverage of the genome, would be very useful. Therefore, to confirm the functionality of these newly developed SSRs in our study, 149 primer pairs were randomly selected out of 220 with known annealing temperatures and were used to amplify the genomic DNA of 15 lentil cultivars and landraces. Of these 149 primers, 78 produced polymorphic fragments, whereas the remaining 71 primers exhibited a monomorphic banding pattern in 15 lentil cultivars and landraces. The number of alleles observed at microsatellite loci ranged from 2 to 11 with an average of 5.1 alleles per locus for the 78 polymorphic SSR loci. Hamwieh et al. (2009) reported that the average number of alleles per locus was 9.14 for 14 SSR loci in 30 *L. culinaris* Medik. genotypes. However, the numbers of genotypes and SSR markers were different in our study compared with the findings of Hamwieh et al. (2009). The average number of alleles per locus was also relatively lower than in other species such as chickpea (6.4 alleles per locus for 25 loci; Sethy et al. 2006) and buckwheat (12.2 alleles per locus for 54 SSR loci; Kanishi et al. 2006). A higher mean number of alleles per polymorphic SSR locus was observed for GA repeats (7.80), followed by CT (6.55) and CA (5.91), making them the most suitable and appropriate motifs to target in further SSR markers for lentil.

PIC provides a better estimate of diversity than the raw number of alleles, because it takes into account the relative frequen-

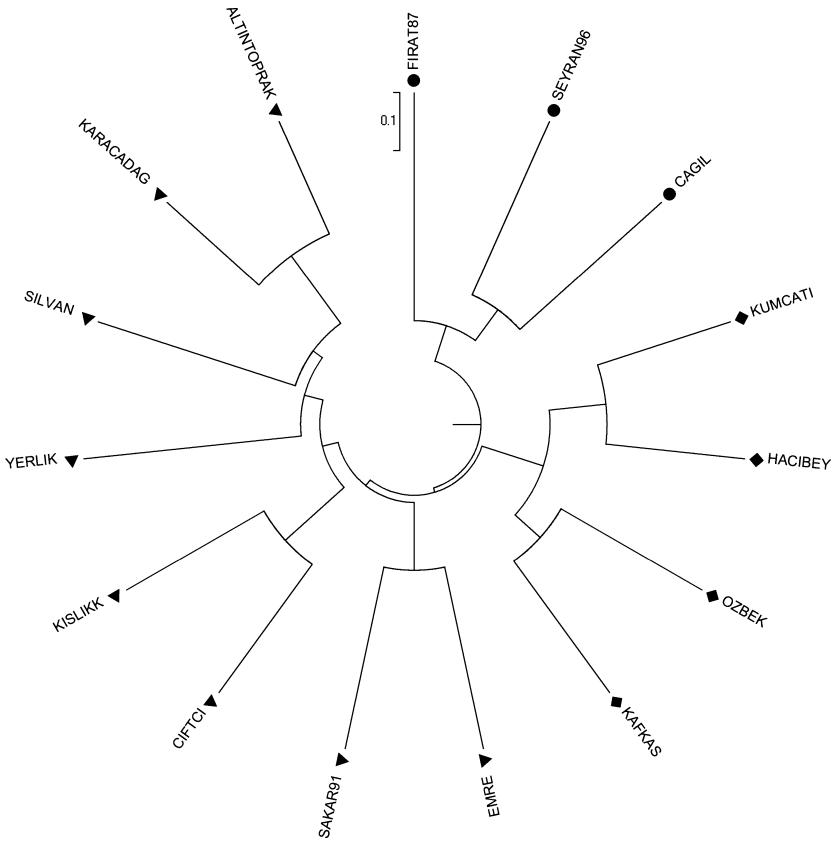

Fig. 1: Neighbour-joining analysis of eight Turkish lentil cultivars and seven Turkish landraces based on SSR markers developed in this study

cies of each allele present (Laborda et al. 2005). Moreover, the PIC value is also an indicator of marker ability to discriminate among genotypes. Here, the mean PIC value was 0.58, with 49 of 78 polymorphic SSR markers having the mean PIC value >0.58. A higher PIC value (0.89) was observed in dinucleotide repeats, whereas among trinucleotide repeats it was 0.62. Similar findings have been reported that show that the level of polymorphism in dinucleotide repeats was higher than the level of polymorphism in trinucleotide repeats, such as in tomato (He et al. 2003) and avocado (Ashworth et al. 2004). Additionally, the average PIC value among dinucleotides in the present study was much higher than that observed in other studies; for example, the average PIC value for dinucleotide repeats in pigeon pea was 0.48 (Odeny et al. 2007) and in flax was 0.41 (Soto-Cerda et al. 2011). Temnykh et al. (2001) reported that dinucleotide repeats typically reside outside the coding regions of genes and/or are characterized by having a higher number of repeats (Li et al. 2004), making them the best source of highly polymorphic SSR markers (Odeny et al. 2007).

In the present study, the longest repeat motifs were GA, CT and CA with 35, 31 and 27 uninterrupted repeats, respectively. Longer repeats have previously been associated with a higher percentage of polymorphism. We found that there was a positive correlation ($r = 0.611$) between repeat length and PIC value. Similar positive correlations have been described in maize (Sharopova et al. 2002) and rice (Temnykh et al. 2001, Singh et al. 2010), but this contrasts with the results obtained in soybean (Shultz et al. 2007) and sunflower (Yi et al. 2006). Odeny et al. (2007) noted that theoretically, the number of repeats is correlated with the mutation rate, meaning that more recently evolved microsatellites would have low polymorphism due to the lower chances of mutation, even if they have longer repeats.

The neighbour-joining graph (Fig. 1) obtained by analysing these SSR marker data suggests that they may be used in diversity analysis. Analysis of the neighbour-joining graph divided the small set of cultivars and landraces into two groups. The local landrace Karacadağ and one cultivar 'Altıntoprak' were grouped together under the same subcluster. Toklu et al. (2009) reported that the lentil landrace from Karacadağ was clearly different from the rest of the Turkish germplasm. Similarly, the lentil cultivars 'Kafkas' and 'Özbek' were also developed by single plant selection from landraces collected from south-east Anatolia, and in this study were clustered closely with local landraces. Similar results were also observed in our previous study, where 'Özbek' and 'Kafkas' were grouped with local landraces from south-eastern Turkey (Toklu et al. 2009). The neighbour-joining dendrogram clearly showed that these SSR markers were able to effectively discriminate all lentil genotypes.

In this study, we observed cultivar-specific alleles that were detected by 32 SSR markers. Cultivar-specific alleles (called rare alleles) were distributed over 12 genotypes with one or more SSR locus. Two local landraces, Karacadağ and Kışlık kırmızı, and one cultivar, 'Şakar-91', did not harbour any cultivar-specific bands (further information is provided in Table S2). If these unique alleles could be confirmed over broader range of accession, these cultivar-specific bands observed by different SSR markers could be successfully used for cultivar identification, and to protect breeder's rights through DNA fingerprinting of cultivars.

SSR markers developed in this study were used to confirm their use for the construction of genetic linkage maps in lentil. A population consisting of 92 F_2 segregating progeny from a cross between Karacadağ x Silvan was used to construct a linkage map based on these SSR markers. A total of 220 newly devel-

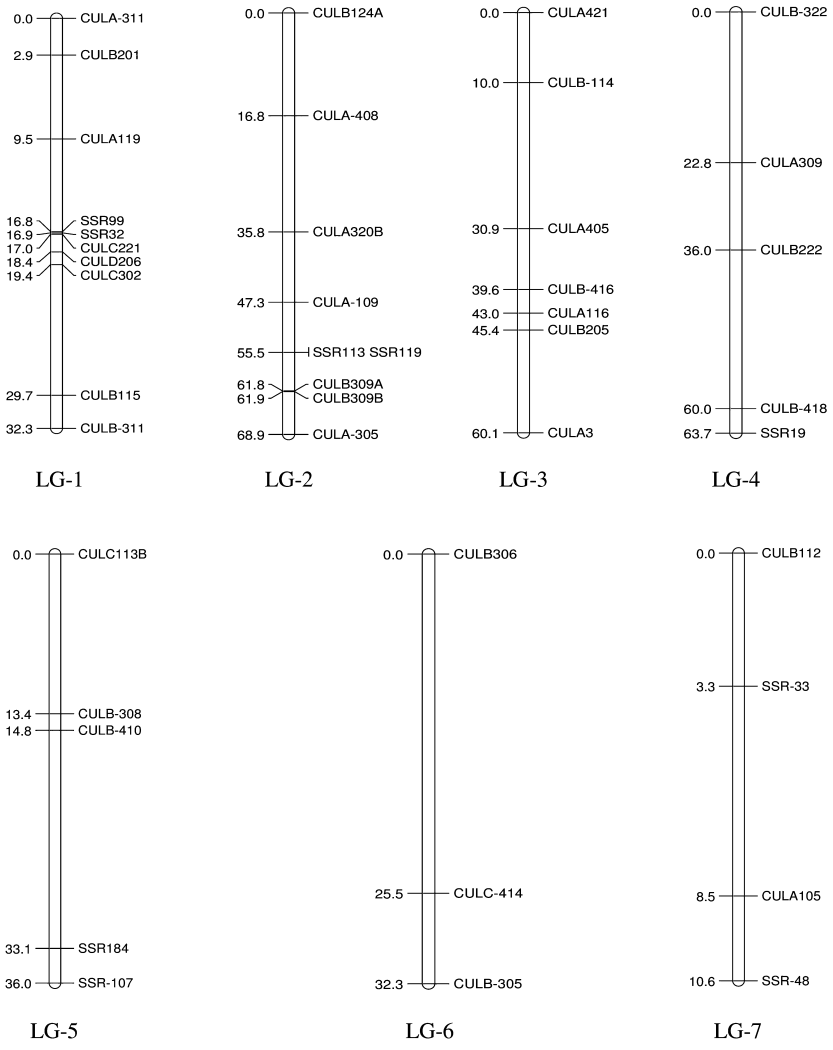

Fig. 2: Linkage map of F_2 population (Karacadağ x Silvan). Distances are indicated in centimorgan (cM) on the left side, and SSR markers are shown to the right of each linkage group

oped SSR markers in the present study, and 40 SSR markers developed by Hamwieh et al. (2005) were used to screen for polymorphisms between the parents. Of 260 SSR markers, only 47 showed polymorphism among the parents. The microsatellite polymorphism among parents was lower when compared with that observed by Hamwieh et al. (2005). Finally, 47 SSR loci were used to construct a genetic linkage map of the Karacadağ x Silvan parents. Four microsatellite markers were not linked. Of 43 SSR markers mapped in this study, nine SSR markers were developed by Hamwieh et al. (2005). The total number of markers used for constructing the linkage map in the present study was very low; however, the number of genomic SSR markers mapped in this study was good in comparison with previously constructed genetic linkage map (Phan et al. 2007, Gupta et al. 2012b, Saha et al. 2013, Kaur et al. 2014). Moreover, the population used for linkage map construction was F_2 mapping population. However, the objective of this work was just to check the utility of these genomic SSR markers in linkage analysis. Now, F_2 mapping population used in this study has been successfully used to develop recombinant inbred lines, and we are using different markers to develop saturated linkage map of lentil from Anatolia. The linkage groups were consecutively numbered as LG1–LG7 in descending order (Fig. 2). The map spanned a total length of 303.9 cM, with a marker density of 7.06 cM/marker, but the distribution of microsatellite markers between linkage groups was unequal. For instance, there was a large distance between markers (LG3) and also very close markers (LG1) in linkage groups (Fig. 2).

Chi-square test was conducted to check the segregation analysis of SSR markers used to construct the linkage map. The chi-square test showed that four SSR markers deviated from the expected Mendelian segregation ratio, whereas the remaining 43 markers were consistent with expected segregation pattern. A similarly distorted segregation percentage was observed in the genetic linkage map of *Lens* (Hamwieh et al. 2005), where the authors reported that 4 of 42 microsatellite loci deviated from normal segregation behaviour. Segregation distortion of 14% in the F_2 population of lentil has been reported by Rubeena and Taylor (2003).

Theoretically, the number of linkage groups should be equal to the number of haploid chromosomes; therefore, the genetic linkage map should include seven linkage groups of the lentil genome. The seven linkage groups observed in this map (Fig. 2) were equal the haploid chromosome number of lentils ($2n = 2x = 14$). However, large regions will need to be covered by more DNA markers in future work. Comparison based on the localization of SSR loci showed that the current map is consistent with a previous map (Hamwieh et al. 2005). We had also mapped nine SSR markers previously developed and described by Hamwieh et al. (2005), and these were mapped on the same genomic position as previously described. This suggests that the microsatellite markers developed in this study are locus specific and are highly useful for mapping studies.

Despite recent advances in the development of genetic markers, the development of microsatellite markers in lentil has been fairly slow. However, the recent application of the next-generation sequencing and genotyping by sequencing technologies has accelerated the discovery of genomewide single nucleotide polymorphism (SNP) markers. Last decade, several linkage maps in lentil have been developed using EST-SSR and SNP markers. Actually, the number of genomic SSR markers published so far is insufficient for the construction of a genetic linkage map that covers the whole lentil genome. The genomic SSR markers reported here are a valuable source and increase the number of markers available for genetic and genomic analyses and comprehensive genetic linkage maps, to identify novel traits linked with SSR markers, and will therefore contribute to genomic-assisted breeding in lentil and related species.

Acknowledgement

The authors express their gratitude to TÜBİTAK (The Scientific and Technological Research Council of Turkey, TOVAG-104O186) and the University of Çukurova, Scientific Research Projects Unit (ZF2004BAP17), for their financial support.

References

Alo, F., B. J. Furman, E. Akhunov, J. Dvorak, and P. Gepts, 2011: Leveraging genomic resources of model species for the assessment of diversity and phylogeny in wild and domesticated lentil. J. Hered. **102**, 315—329.

Alsaleh, A., F. S. Baloch, M. Derya, M. Azrak, B. Kilian, H. Özkan, and M. Nachit, 2015: Genetic linkage map of anatolian durum wheat derived from a cross of Kunduru-1149 × Cham1. Plant Mol. Biol. Rep. **33**, 209—220.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, 1990: Basic local alignment search tool. J. Mol. Biol. **215**, 403—410.

Ashworth, V. E. T. M., M. C. Kobayashi, M. De La Cruz, and M. T. Clegg, 2004: Microsatellite markers in avocado (*Persea americana* Mill.): development of dinucleotide and trinucleotide markers. Sci. Hortic. **101**, 255—267.

Baloch, F. S., M. Derya, E. E. Andeden, A. Alsaleh, G. Cömertpay, B. Kilian, and H. Özkan, 2015: Inter-primer binding site retrotransposon and inter-simple sequence repeat diversity among wild *Lens* species. Biochem. Syst. Ecol. **58**, 162—168.

Benchimol, L. L., T. De Campos, S. A. M. Carbonell, C. A. Colombo, A. F. Chioratto, E. F. Formighieri, L. R. L. Gouvea, and A. P. De Souza, 2007: Structure of genetic diversity among common bean (*Phaseolus vulgaris* L.) varieties of Mesoamerican and Andean origins using new developed microsatellite markers. Genet. Resour. Crop Evol. **54**, 1747—1762.

Brondani, C., P. Rangel, R. Brondani, and M. Ferreira, 2002: QTL mapping and introgression of yield-related traits from *Oryza glumaepatula* to cultivated rice (*Oryza sativa*) using microsatellite markers. Theor. Appl. Genet. **104**, 1192—1203.

Cho, Y. G., T. Ishii, S. Temnykh, X. Chen, L. Lipovich, S. R. Mccouch, W. D. Park, N. Ayres, and S. Cartinhour, 2000: Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor. Appl. Genet. **100**, 713—722.

Condit, R., and S. P. Hubbell, 1991: Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. Genome **34**, 66—71.

De Arruda, M. P., E. C. Gonçalves, M. P. C. Schneider, A. L. D. C. da Silva, and E. Morielle-Versute, 2010: An alternative genotyping method using dye-labeled universal primer to reduce unspecific amplifications. Mol. Biol. Rep. **37**, 2031—2036.

Duran, Y., R. Fratini, P. Garcia, and M. Perez de la Vega, 2004: An intersubspecific genetic map of *Lens*. Theor. Appl. Genet. **108**, 1265—1273.

Eujayl, I., M. Baum, W. Powell, W. Erskine, and E. Pehu, 1998: A genetic linkage map of lentil (*Lens* sp.) based on RAPD and AFLP markers using recombinant inbred lines. Theor. Appl. Genet. **97**, 83—89.

Faostat, 2013: Available at: http://faostat3.fao.org/download/Q/QC/E (last accessed on January 5, 2015).

Ferguson, M. E., M. D. Burow, S. R. Schulze, P. J. Bramel, A. H. Paterson, S. Kresovich, and S. Mitchell, 2004: Microsatellite identification and characterization in peanut (*A. hypogaea* L.). Theor. Appl. Genet. **108**, 1064—1070.

Fisher, P. J., R. C. Gardner, and T. E. Richardson, 1996: Single locus microsatellite isolated using 5' anchored PCR. Nucleic Acid Res. **24**, 4369—4371.

Gaitan-Solis, E., M. C. Duque, K. J. Edwards, and J. Tohme, 2002: Microsatellite repeats in common Bean. Crop Sci. **42**, 2128—2136.

Gupta, P. K., and R. K. Varshney, 2000: The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica **113**, 163—185.

Gupta, D., P. W. J. Taylor, P. Inder, H. T. T. Phan, S. R. Ellwood, P. N. Mathur, A. Sarker and R. Ford, 2012a: Integration of EST-SSR markers of *Medicago truncatula* into intraspecific linkage map of lentil and identification of QTL conferring resistance to ascochyta blight at seedling and pod stages. Mol. Breed. **30**, 429—439.

Gupta, M., B. Verma, N. Kumar, R. K. Chahota, R. Rathour, S. K. Sharma, S. Bhatia, T.R. Sharma. 2012b: Construction of intersubspecific molecular genetic map of lentil based on ISSR, RAPD and SSR markers. J. Genet. **91**, 279—287.

Gürcan, K., and S. A. Mehlenbacher, 2010: Development of microsatellite marker loci for European hazelnut (*Corylus avellana* L.) from ISSR fragments. Mol. Breeding **26**, 551—559.

Hamdan, Y. A. S., M. J. Garcia-Moreno, J. Redondo-Nevado, L. Velasco, and B. Perez-Vich, 2011: Development and characterization of genomic microsatellite markers in safflower (*Carthamus tinctorius* L.). Plant Breed. **130**, 237—241.

Hamilton, A. J., and D. C. Baulcombe, 1999: A species of small antisense RNA in posttranscriptional gene silencing in plants. Science **286**, 950—952.

Hamwieh, A., S. M. Udapa, W. Choumane, A. Sarker, F. Dreyer, C. Jung, and M. Baum, 2005: A genetic linkage map of *Lens* sp. Based on microsatellite and AFLP markers and the localization of fusarium vascular wilt resistance. Theor. Appl. Genet. **110**, 669—677.

Hamwieh, A., S. M. Udapa, F. Sarker, C. Jung, and M. Baum, 2009: Development of new microsatellite markers and their application in the analysis of genetic diversity in lentils. Breed. Sci. **59**, 77—86.

Havey, M. J., and F. J. Muehlbauer, 1989: Linkages between restriction fragment length, isozyme and morphological markers in lentil. Theor. Appl. Genet. **77**, 395—401.

Hayden, M. J., and P. J. Sharp, 2001: Sequence-tagged microsatellite profiling (STMP): a rapid technique for developing SSR markers. Nucleic Acids Res. **29**, e43.

He, C., V. Poysa, and K. Yu, 2003: Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. Theor. Appl. Genet. **106**, 363—373.

Jaccard, P., 1908: Nouvelles recherches sur la distribution florale. Bull. Soc. Vaud. Sci. Nat. **4**, 223—270.

Jarne, P., and P. J. Lagoda, 1996: Microsatellites, from molecules to populations and back. Trends Ecol. Evol. **11**, 424—429.

Kahraman, A., I. Kusmenoglu, N. Aydin, A. Aydogan, W. Erksine, and F. J. Muehlbauer, 2004: QTL mapping of winter hardiness genes in lentil. Crop Sci. **44**, 13—22.

Kanishi, T., H. Iwata, K. Yashiro, Y. Tsumura, R. Ohsawa, Y. Yasui, and O. Ohnishi, 2006: Development and characterization of microsatellite markers for common buckwheat. Breed. Sci. **56**, 277—285.

Karaköy, T., H. Erdem, F. S. Baloch, F. Toklu, S. Eker, B. Kilian, and H. Özkan, 2012: Diversity of macro-and micronutrients in the seeds of lentil landraces. Sci. World J. **2012**, 710412.

Kaur, S., N. O. Cogan, L. W. Pembleton, M. Shinozuka, K. W. Savin, M. Materne, and J. W. Forster, 2011: Transcriptome sequencing of

lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. BMC Genom. **12**, 265.

Kaur, S., N. I. Cogan, A. Stephens, D. Noy, M. Butsch, J. Forster, and M. Michael, 2014: EST-SNP discovery and dense genetic mapping in lentil (*Lens culinaris* Medik.) enable candidate gene selection for boron tolerance. Theor. Appl. Genet. **127**, 703—713.

Kumar, S., A. Hamweih, A. Manickavelu, J. Kumar, T. R. Sharma, and M. Baum, 2014: Advances in lentil genomics. In: S. Gupta, N. Nadarajan, and D. S. Gupta (eds), Legumes in Omics Era, 111—130. Springer Science+Business Media, New York.

Kumar, S., K. Rajendran, J. Kumar, A. Hamwieh, and M. Baum, 2015: Current knowledge in lentil genomics and its application for crop improvement. Front. Plant Sci. **6**, 78.

Laborda, P. R., K. M. Oliveira, A. A. F. Garcia, M. E. A. G. Z. Paterniani, and A. P. de Souza, 2005: Tropical maize germplasm: what can we say about its genetic diversity in the light of molecular markers? Theor. Appl. Genet. **111**, 1288—1299.

Li, Y. C., A. B. Korol, T. Fahima, and E. Nevo, 2004: Microsatellites within genes: structure, function, and evolution. Mol. Biol. Evol. **21**, 991—1007.

Liu, K., and S. V. Muse, 2005: PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics **21**, 2128—2129.

Mace, E. S., and I. D. Godwin, 2002: Development and characterization of polymorphic microsatellite markers in taro (*Colocasia esculenta*). Genome **45**, 823—832.

Missio, R. F., E. T. Caixeta, E. M. Zambolim, L. Zambolim, C. D. Cruz, and N. S. Sakiyama, 2010: Polymorphic information content of SSR markers for *Coffea* spp. Crop Breed. Appl. Biotechnol. **10**, 89—94.

NCBI, 2015: Available at: http://www.ncbi.nlm.nih.gov/nucest/?term=lentil [accessed on January 12, 2015].

Odeny, D. A., B. Jayashree, M. Ferguson, D. Hoisington, J. Crouch, and C. Gebhardt, 2007: Development, characterization and utilization of microsatellite markers in pigeonpea. Plant Breeding **126**, 130—136.

Ozkan, H., A. Brandolini, C. Pozzi, S. Effgen, J. Wunder, and F. Salamini, 2005: A reconsideration of the domestication geography of tetraploid wheat. Theor. Appl. Genet. **110**, 1052—1060.

Perez de la Vega, M., R. M. Fratini, and F. J. Muehlbauer, 2011: Lentil. In: Perez de la Vega M., A. M. Torres, J. I. Cubero, and C. Kole (eds), Genetics, Genomic and Breeding of Cool Season Grain Legumes (Genetics, Genomics and Breeding in Crop Plants), 98—150. Science Pubs, Boca Raton, FL.

Phan, H. T., S. R. Ellwood, J. K. Hane, R. Ford, M. Materne, and R. P. Oliver, 2007: Extensive microsynteny between *Medicago truncatula* and *Lens culinaris* ssp. *culinaris*. Theor. Appl. Genet. **114**, 549—558.

Rodriguez, M. M., O. M. Paredes, and V. L. Becerra, 1997: Isozyme diversity of Chilean lentil germplasm (*Lens culinaris* Medik.). In: Int Food Legume Res Conf III, Adelaide, Australia, 135.

Rohlf, F. J., 2004: TPS Software. State University of New York, Stony Brook, NY.

Roy, F., J. I. Boye, and B. K. Simpson, 2010: Bioactive proteins and peptides in pulse crops: pea, chickpea and lentil. Food Res. Int. **43**, 432—442.

Rubeena, F., and P. W. J. Taylor, 2003: Construction of an intraspecific linkage map of lentil (*Lens culinaris* ssp. *culinaris*). Theor. Appl. Genet. **107**, 910—916.

Saha, G. C., A. Sarker, W. Chen, G. J. Vandemark, and F. J. Muehlbauer, 2013: Inheritance and Linkage Map Positions of Genes Conferring Agromorphological Traits in *Lens culinaris* Medik. Int. J. Agron. 2013, 618926. http://dx.doi.org/10.1155/2013/618926

Schuelke, M., 2000: An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. **18**, 233—234.

Sethy, N. K., B. Shokeen, K. J. Edwards, and S. Bhatia, 2006: Development of microsatellite markers and analysis of intraspecific genetic variability in chickpea (*Cicer arietinum* L.). Theor. Appl. Genet. **112**, 1416—1428.

Sharma, S. K., M. R. Knox, and T. H. N. Ellis, 1996: AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD analysis. Theor. Appl. Genet. **93**, 751—758.

Sharopova, N., M. D. McMullen, L. Schultz, S. Schroeder, H. Sanchez-Villeda, J. Gardiner, D. Bergstrom, K. Houchins, S. Melia-Hancock, T. Musket, N. Duru, M. Polacco, K. Edwards, T. Ruff, J. C. Register, C. Brouwer, R. Thompson, R. Velasco, E. Chin, M. Lee, W. Woodman-Clikeman, M. J. Long, E. Liscum, K. Cone, G. Davis, and E. H. Coe, 2002: Development and mapping of SSR markers for maize. Plant Mol. Biol. **48**, 463—481.

Sharpe, A. G., L. Ramsay, L. A. Sanderson, M. J. Fedoruk, W. E. Clarke, L. Rong, et al., 2013: Ancient orphan crop joins modern era: gene-based SNP discovery and mapping in lentil. BMC Genom. **14**, 192.

Shultz, J. L., S. Kazi, R. Bashir, J. A. Afzal, and D. A. Lightfoot, 2007: The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean. Theor. Appl. Genet. **114**, 1081—1090.

Singh, H., R. K. Deshmukh, A. Singh, A. K. Singh, K. Gaikwad, T. R. Sharma, T. Mohapatra, and N. K. Singh, 2010: Highly variable SSR markers suitable for rice genotyping using agarose gels. Mol. Breed. **25**, 359—364.

Song, Q. J., J. R. Shi, S. Singh, E. W. Fickus, J. M. Costa, J. Lewis, B. S. Gill, R. Ward, and P. B. Cregan, 2005: Development and mapping of microsatellite (SSR) markers in wheat. Theor. Appl. Genet. **110**, 550—560.

Sonnante, G., K. Hammer, and D. Pignone, 2009: From the cradle of agriculture a handful of lentils: history of domestication. Rend. Lincei **20**, 2137.

Soto-Cerda, B. J., R. A. Carrasco, G. A. Aravena, H. A. Urbina, and C. S. Navarro, 2011: Identifying novel polymorphic microsatellites from cultivated flax (*Linum usitatissimum* L.) following data mining. Plant Mol. Biol. Rep. **29**, 753—759.

Tahir, M., and F. J. Muehlbauer, 1994: Gene mapping in lentil with recombinant inbred lines. J. Hered. **85**, 306—310.

Tahir, M., and F. J. Muehlbauer, 1995: Association of quantitative trait loci with isozyme markers in lentil (*Lens culinaris* L.). J. Genet. Breed. **49**, 145—150.

Tahir, M., F. J. Muehlbauer, and S. C. Spaeth, 1994: Association of isozyme markers with quantitative trait loci in random single seed descent derived lines of lentil (*Lens culinaris* Medik.). Euphytica **75**, 111—119.

Tanyolac, B., S. Ozatay, A. Kahraman, and F. Muehlbauer, 2010: Linkage mapping of lentil (*Lens culinaris* L.) genome using recombinant inbred lines revealed by AFLP, ISSR, RAPD and some morphologic markers. J. Agric. Biotechnol. Sustain. Dev. **2**, 001—006.

Temnykh, S., G. DeClerck, A. Lukashova, L. Lipovich, S. Cartinhour, and S. McCouch, 2001: Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res. **11**, 1441—1452.

Toklu, F., T. Karaköy, E. Haklı, T. Bicer, A. Brandolini, B. Kilian, and H. Özkan, 2009: Genetic variation among lentil (*Lens culinaris* Medik) landraces from Southeast Turkey. Plant Breeding **128**, 178—186.

Van Der Schoot, J., M. Pospiskova, B. Vosman, and M. J. M. Smulders, 2000: Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). Theor. Appl. Genet. **101**, 317—322.

Van Ooijen, J. W., and R. E. Voorrips, 2001: JoinMap 3.0 Software for The Calculation of Genetic Linkage Maps. Plant Research International, Wageningen, the Netherlands.

Verma, P., N. Shah, and S. Bhatia, 2013: Development of an expressed gene catalogue and molecular markers from the de novo assembly of short sequence reads of the lentil (*Lens culinaris* Medik.) transcriptome. Plant Biotechnol. J. **11**, 894—905.

Verma, P., T. R. Sharma, P. S. Srivastava, M. Z. Abdin, and S. Bhatia, 2014: Exploring genetic variability within lentil (*Lens culinaris* Medik.) and across related legumes using a newly developed set of microsatellite markers. Mol. Biol. Rep. **41**, 5607—5625.

Wang, J. Y., L. S. Zheng, B. Z. Huang, W. L. Liu, and Y. T. Wu, 2010: Development, characterization, and variability analysis of microsatellites from a commercial cultivar of *Musa acuminata*. Genet. Resour. Crop Evol. **57**, 553—563.

Weber, J. L., 1990: Informativeness of human (dC-dA)n (dG-dT)n polymorphisms. Genomics **7**, 524—530.

Yi, G., J. Lee, S. Lee, D. Choi, and B. D. Kim, 2006: Exploitation of pepper EST-SSRs and an SSR-based linkage map. Theor. Appl. Genet. **114**, 113—130.

Zane, L., L. Bargelloni, and T. Patarnello, 2002: Strategies for microsatellite isolation: a review. Mol. Ecol. **11**, 1—16.

Zhang, M., W. Mao, G. Zhang, and F. Wu, 2014: Development and characterization of polymorphic EST-SSR and genomic SSR markers for Tibetan annual wild barley. PLoS One **9**, e94881.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Repeat motif, annealing temperature and expected PCR product size of all SSR markers.

Table S2. Information about the rare alleles observed in lentil cultivars and landraces using SSR markers.